13th ANKARA INTERNATIONAL AEROSPACE CONFERENCE **AIAC-2025-097**
17-19 September 2025 – METU, Ankara, TURKIYE

DATA-DRIVEN OPTIMIZATION OF TRANSITION MODELING FOR IMPROVED BOUNDARY LAYER PREDICTIONS

Baris Berkay Peksag[1]
Middle East Technical University (METU)
Ankara, Türkiye

Harika S. Kahveci[2]
Middle East Technical University (METU)
Ankara, Türkiye

ABSTRACT

Turbulence models are widely used in modeling of fluid flow problems to capture the phenomenon of boundary layer transition which may occur in a variety of flow conditions such as flow separation and high levels of turbulence or due to flow instabilities. With the use of turbulence models, a proper set of coefficients needs to be selected for modeling flow physics accurately. To accomplish this task, the general approach has been to utilize experimental data and data-driven methods. In this study, a data-driven model is trained to select proper coefficients for the Transition Shear Stress Transport (SST) model also known as the γ-Re_θ SST model. The data used in the analysis is generated via Computational Fluid Dynamics (CFD) runs that are performed with different combinations of the destruction/relaminarization source coefficients used in the transport equation for the intermittency γ. For the analysis, a flat plate test case with zero-pressure gradient is simulated. First, initial predictions are obtained using a set of selected coefficients, and results are compared with experimental data and predictions from literature. Then, several data sets are generated via computations where the coefficients are incrementally varied within their respective intervals in the turbulence model, which is followed by the optimization of those coefficients for different cases. For optimization, the Bayesian Optimization method is used along with the Gaussian Process as the probabilistic model. With the use of optimization, observable improvement is demonstrated in prediction of transition for the test case studied. Results also show the sensitivity of the predictions to the turbulence intensity level as well as to the effective Prandtl number for the transport of the specific dissipation rate.

INTRODUCTION

Turbulence has been one of the most studied aspects of flow physics through decades. The chaotic yet deterministic nature of this fluid phenomenon made it the center of attention among researchers. Solving the Navier-Stokes equations which are governing the fluid motion is rather challenging due to the complexity of these equations and high-resolution demand of turbulence computations. This has led to the widespread use of turbulence models implemented in the computational tools, although these models are bounded by assumptions simplifying the nature of the flow analyzed. Some of those models use the Reynolds Averaged Navier-Stokes (RANS) equations as framework. The Reynolds stresses in the RANS equations arising due to turbulent fluctuations cause a closure problem due to the addition of these extra unknowns excessing the number of available equations. The k-ω SST model is one of the turbulence models introduced to solve this problem using two transport equations

[1] M.Sc. Student in Department of Aerospace Engineering, Email: baris.peksag@metu.edu.tr
[2] Assoc. Prof. in Department of Aerospace Engineering, Email: kahveci@metu.edu.tr

for the turbulent kinetic energy, k, and turbulent frequency, ω [Menter, 1994]. However, this model assumes the whole boundary layer as turbulent which leads to prediction inaccuracy, whereas in nature, the early phase of a boundary layer presents laminar flow which eventually is disturbed and transitioned into turbulent flow. Boundary layer transition may be triggered by different mechanisms such as natural transition, bypass transition, or could even be induced by flow separation. Hence, to capture this complex phenomenon accurately, two new transport equations were introduced, defined as the γ-Re_θ *transition* model [Menter, Langtry, Likki, Suzen, Huang and Völker, 2006]. The first of those equations is for the intermittency function which is used to trigger transition locally. The second equation is for the transition momentum-thickness Reynolds number, $Re_{\theta t}$. While intermittency plays a switch role between laminar and turbulent regions of boundary layers, transition momentum-thickness Reynolds number is an empirical parameter used to capture the influence of turbulence intensity. The intermittency function is coupled with the k-ω SST model. Addition of these two equations to k and ω equations makes this a four-equation model. A modification to the blending function used to switch between k-ε and k-ω models was also included in the model. Later, the full model was made available to the research community including some improvements for prediction of natural transition [Langtry and Menter, 2009]. This turbulence model is known as γ-Re_θ *SST* or *Transition SST* model and is the subject of the current study.

In literature, various studies have been performed using the Transition SST model. A parametric uncertainty study was done on this model to determine which coefficients of the model played a significant role on predictions of flow field over flat plates and airfoils [Song, Liu, Lu and Yan, 2022]. It was concluded that the most critical coefficients contributing to the uncertainty for both types of geometries were the coefficients in the destruction equation of intermittency, namely Ca_2 and Ce_2. There are validation studies that have implemented this turbulence model into different solvers. In one of these studies, the transport equations of the k-ω SST and the γ-Re_θ transition model were implemented into the commercial software STAR-CCM+ [Malan, Suluksna and Juntasaro, 2009]. In another study, the γ-Re_θ transition model was coupled with the k-ω SST and Spalart-Allmaras turbulence models and was implemented within the open-source software SU2 [Rausa, Guardone and Auteri, 2023]. Predictions for 2D and 3D test cases were performed and compared with experimental results, reporting good agreement on transition onset location for the flat plate cases tested. In one study, the empirical correlation of critical Reynolds number, $Re_{\theta c}$, in the Transition SST model was modified and used with a high-order numerical method [Yuntao, Yulun, Song and Dehong, 2014]. Doing so resulted in improved flow predictions of typical low-speed flows and improvement in transition location prediction for turbulent flat plates. All these studies mentioned above used the flat-plate experiments of the European Research Community on Flow Turbulence and Combustion (ERCOFTAC) [Savill, 1993, Savill, 1996].

Boundary-layer transition has also pulled strong attention in the research of high-speed flows, as the transition from laminar to turbulent regime may result in significant increases in the wall heat flux. Since the γ-Re_θ transition model in its current form is suitable for the analysis of low-speed flows, there have been attempts by researchers to modify this model to obtain accurate predictions in hypersonic flow conditions as well. In one study, the model was extended to include correlations for the calculation of the transition onset end length [Krause, Behr and Ballmann, 2008]. For validation, both low-speed flows over flat plates and hypersonic flows over double-ramp intake configurations were studied. Later, two key empirical correlations of the model were recalibrated and the pressure gradient parameter and turbulent Prandtl number were corrected to simulate hypersonic boundary layer transition [Zhang, Zhang, Chen, Mao and Jiang, 2017]. In a more recent study, the γ-Re_θ transition model was modified by using the compressible boundary-layer self-similar solutions, compressibility correction and crossflow correlation [Fan, Liu, Zhao, Zhang, Yuan and Liu, 2024]. Based on this work, the effect of surface irregularities such as local cavities on transition in hypersonic flows was analyzed by further modifying the γ-Re_θ transition model in a follow-up study [Zhao, Zhang, Shen, Fan and Liu, 2025]. Among the tested cases were a flat plat with installed cavities and the windward surface of a shuttle orbiter. It was concluded that the main driver of the cavity-

induced transition was the adverse pressure gradient occurring near the trailing edge and the instability caused by the flow separation in the cavity.

Since transition is a complex phenomenon, empirical constants play an important role in its modeling. This is where the experimental data come into play. Although not data-driven, the Transition SST model uses experimental data in its correlations. On the other hand, data-driven studies have been a trending topic for the last decades. With progress in statistical studies enhancing machine learning algorithms, the implementation of CFD in data-driven approaches has also become one of the widely studied areas by data scientists.

METHOD

Transition SST Model

The *Shear Stress Transport (SST)* k-ω model [Menter, 1994] uses Reynolds Averaged Navier Stokes (RANS) equations, as in all eddy viscosity models, and assumes a linear stress-strain relation. The problem with these models is that they assume the boundary layer to be fully turbulent in all locations. This has brought the need to model the transition behavior in boundary layer flows. In the Transition SST or γ-Re_θ SST model [Langtry and Menter, 2009], two new transport equations are introduced. As explained by Langtry and Menter [Langtry and Menter, 2009], Eq. (1) is the transport equation for the intermittency, γ:

$$\frac{\partial(\rho\gamma)}{\partial t} + \frac{\partial(\rho u_j \gamma)}{\partial x_j} = P_\gamma - E_\gamma + \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_f}\right)\frac{\partial \gamma}{\partial x_j}\right] \tag{1}$$

where μ_t is the eddy viscosity and μ is the molecular viscosity. P_γ and E_γ are the production and destruction/relaminarization terms as given in Eqs. (2) and (3), respectively:

$$P_\gamma = F_{length} C_{a1} \rho S [\gamma F_{onset}]^{0.5} (1 - C_{e1}\gamma) \tag{2}$$

$$E_\gamma = C_{a2} \rho \Omega \gamma F_{turb} (C_{e2}\gamma - 1) \tag{3}$$

F_{length} and F_{onset} control the length of the transition region and its onset location, respectively. In the destruction term, F_{turb} parameter controls the relaminarization process. F_{onset} is calculated by the set of Eqs. (4-9) given below:

$$F_{onset} = \max(F_{onset2} - F_{onset3}, 0) \tag{4}$$

$$F_{onset2} = \min(\max(F_{onset1}, F_{onset1}^4), 2) \tag{5}$$

$$F_{onset1} = \frac{Re_v}{2.193\, Re_{\theta c}} \tag{6}$$

$$Re_v = \frac{\rho y^2 S}{\mu} \tag{7}$$

$$F_{onset3} = max\left(1 - \left(\frac{R_T}{2.5}\right)^3, 0\right) \tag{8}$$

$$R_T = \frac{\rho k}{\mu \omega} \tag{9}$$

Here, Re_v is the vorticity Reynolds number and $Re_{\theta c}$ is the critical Reynolds number where the intermittency's value first starts to increase inside the boundary layer. R_T is the viscosity ratio, and k is the turbulence kinetic energy. The constants that appear in the intermittency equation are C_{a1}= 2.0, C_{e1}= 1.0, C_{a2}= 0.06, C_{e2}= 50, and σ_f= 1.0.
The second transport equation is for the scalar transition momentum-thickness Reynolds number, $\tilde{R}e_{\theta t}$, as shown in Eq. (10):

$$\frac{\partial(\rho\tilde{R}e_{\theta t})}{\partial t}+\frac{\partial(\rho u_j\tilde{R}e_{\theta t})}{\partial x_j}=P_{\theta t}+\frac{\partial}{\partial x_j}\left[\sigma_{\theta t}(\mu+\mu_t)\frac{\partial\tilde{R}e_{\theta t}}{\partial x_j}\right] \tag{10}$$

The production term is described via Eqs. (11-12), where $Re_{\theta t}$ represents the local value:

$$P_{\theta t}=c_{\theta t}\frac{\rho}{t}\left(Re_{\theta t}-\tilde{R}e_{\theta t}\right)(1-F_{\theta t}) \tag{11}$$

$$t=\frac{500\mu}{\rho U^2} \tag{12}$$

U is the local velocity of the flow. $F_{\theta t}$ in Eq. (11) is a blending function that turns off production and enables to diffuse the momentum-thickness Reynolds number into the boundary layer from the free stream as described in Eq. (13):

$$F_{\theta t}=min\left(max\left(F_{wake}e^{-\left(\frac{y}{\delta}\right)^4},1-\left(\frac{\gamma-1/C_{e2}}{1-1/C_{e2}}\right)^2\right),1\right) \tag{13}$$

F_{wake} given in Eq. (14) is used to deactivate this blending function in the downstream wake regions:

$$F_{wake}=e^{\left[-\left(\frac{Re_\omega}{10^5}\right)^2\right]} \tag{14}$$

The constants used in the second transport equation are $c_{\theta t}$= 0.03 and $\sigma_{\theta t}$= 2.0. F_{length} and $Re_{\theta c}$ are correlated with $\tilde{R}e_{\theta t}$ as in Eqs. (15) and (16):

$$F_{length}=\begin{cases}[39.8189\cdot10^{-1}+(-119.270\cdot10^{-4})\tilde{R}e_{\theta t}+(-132.567\cdot10^{-6})\tilde{R}e_{\theta t}^2], & \tilde{R}e_{\theta t}<400\\ [263.404+(-123.939\cdot10^{-2})\tilde{R}e_{\theta t}+(194.548\cdot10^{-5})\tilde{R}e_{\theta t}^2+(-101.695\cdot10^{-8})\tilde{R}e_{\theta t}^3], & 400\le\tilde{R}e_{\theta t}<596\\ [0.5-(\tilde{R}e_{\theta t}-596)\cdot3\cdot10^{-4}], & 596\le\tilde{R}e_{\theta t}<1200\\ [0.3188], & 1200\le\tilde{R}e_{\theta t}\end{cases} \tag{15}$$

$$Re_{\theta c}=\begin{cases}\left[\tilde{R}e_{\theta t}-(396.035\cdot10^{-2}+(-120.656\cdot10^{-4})\tilde{R}e_{\theta t}+(868.230\cdot10^{-6})\tilde{R}e_{\theta t}^2\right. & \tilde{R}e_{\theta t}\le1870\\ \left.+(-696.506\cdot10^{-9})\tilde{R}e_{\theta t}^3+(174.105\cdot10^{-12})\tilde{R}e_{\theta t}^4\right] & \\ [\tilde{R}e_{\theta t}-(593.11+(\tilde{R}e_{\theta t}-1870)\cdot0.482)], & \tilde{R}e_{\theta t}>1870\end{cases} \tag{16}$$

The empirical correlation for $Re_{\theta t}$ based on turbulence intensity and other details of the numerical scheme are provided by Langtry and Menter [Langtry and Menter, 2009].
Additionally, according to the Baseline (k-ω) model, the effective diffusivity for the specific dissipation rate ω, is given by Eq. (17):

$$\Gamma_\omega = \mu + \frac{\mu_t}{\sigma_\omega} \tag{17}$$

which shows that the turbulent Prandtl number for ω, that is, the σ_ω term, is a multiplier of the eddy viscosity term μ_t. σ_ω is defined by Eq. (18):

$$\sigma_\omega = \frac{1}{F_1/\sigma_{\omega,1}+(1-F_1)/\sigma_{\omega,2}} \tag{18}$$

Here, F_1 is yet another blending function applied on the standard k-ω model in the near-wall region and on a transformed k-ε model in the far region. $\sigma_{\omega,1}$ is defined as the SDR-inner Prandtl number. It is the effective Prandtl number for the transport of the specific dissipation rate inside the near-wall region. In a similar fashion, $\sigma_{\omega,2}$ is defined as the SDR-outer Prandtl number and applies to the outside of the near-wall region. The Transition SST model uses the definitions for $\sigma_{\omega,1}$, and $\sigma_{\omega,2}$ based on the Baseline (k-ω) model [ANSYS, Inc., 2024].The constants in Eq. (18) are $\sigma_{\omega,1}$= 2.0 and $\sigma_{\omega,2} = 1.168$.

In the current study, ANSYS Fluent Release 2024 R2 is used as the solver which enables the use of this transition model in the CFD analyses performed.

Strategy

This study focuses on transition predictions for flow over flat plates. One of the ERCOFTAC T3 series of flat-plate experiments [Savill, 1993] is considered in the current study. The T3A case with turbulence intensity 3.5% is the subject case utilized for the validation effort. This was also one of the test cases investigated by Langtry and Menter [Langtry and Menter, 2009]. As mentioned before, the uncertainty analysis presented by Song et al. [Song, Liu, Lu and Yan, 2022] on the same T3A case, in addition to some other test cases, revealed that the coefficients in the destruction equation of intermittency (Ca_2, Ce_2) are the most important contributors to the uncertainty. Therefore, this study specifically focuses on the effect of these two coefficients on prediction accuracy. Initial CFD runs are performed with combinations of lower and upper bounds and default values. Later, multiple CFD runs scanning those two coefficients from lower to upper bound will be executed. The work will progress with optimization of the results to determine the combination producing the best agreement with the data.

Optimization Method

Unlike gradient-based methods which use gradients to get the optimum result, the Bayesian optimization is a tool taking the probabilistic model that quantifies the uncertainty of a solution field as its main mechanism. This enables the model not to do the gradient descent blindly but also to evaluate the points in new areas where there is uncertainty to explore. To do that, the algorithm needs a probabilistic model to provide a continuous distribution between points and an acquisition function to evaluate the point's state to provide a new point with more potential.

As a probabilistic model, the Gaussian Process is often recommended [Snoek, Larochelle and Adams, 2012] due to its adaptability into many kinds of functions and its well assessment of uncertainty. The model basically sheds continuous Gaussian distributions along the data field and enables itself to judge the reliability of the guess [Rasmussen and Williams, 2006].

As the acquisition function, Snoek et al. [Snoek, Larochelle and Adams, 2012] recommended to use the Expected Improvement (EI). In a simplified form, this function can be interpreted as in Eq. (19):

$$EI(x) = \sigma(x)\big[\gamma(x)\phi\big(\gamma(x)\big) + N\big(\gamma(x)\big)\big] \tag{19}$$

where σ is the standard deviation of the predicted data. Multiplying it with the rest of the equation increases the value of EI in the areas where uncertainty is present. γ is the Z-score of the data point interpreting how far it is from the best point (i.e. the point having the best EI value), in terms of standard deviation. The Z-score is defined by Eq. (20):

$$Z = \frac{X-\mu}{\sigma} \tag{20}$$

Here, μ is the mean value. The parameter Φ in Eq. (19) is the Cumulative Distribution Function (CDF) which is denoted by $F(x)$ in Eq. (21). For a continuous random variable X:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \tag{21}$$

for $-\infty < x < \infty$. The CDF of the Z-score gives the probability of the lower Z-score, or in other words the probability of improvement. A higher Z-score results in a higher CDF value. This increases EI and encourages the model to dive into lower uncertainty or higher probability areas of the probability field. Finally, the parameter N in Eq. (19) is the Probability Density Function (PDF) of the Z-score of the data point. Details are explained by Montgomery and Runger [Montgomery and Runger, 2014].

In the current study, one of the inputs to the optimization loop is the relative error of the variation of the skin friction coefficient with local Reynolds number from the CFD result with respect to the ERCOFTAC T3A experiment. The error is calculated pointwise using the experimental data, and its overall mean for N points is defined as in Eq. (22):

$$e = \frac{\left|\sum_{n=1}^{\mathrm{N}} \frac{C_{f,CFD,n} - C_{f,ERCOFTAC,n}}{C_{f,ERCOFTAC,n}}\right|}{\mathrm{N}} \tag{22}$$

In the optimization, Python's Bayesian Optimization module from its bayes-opt library is used. This module is designed to discover the values of parameters resulting in a maximum value. Therefore, the error values are multiplied by -1 to be suitable for the minimization task.

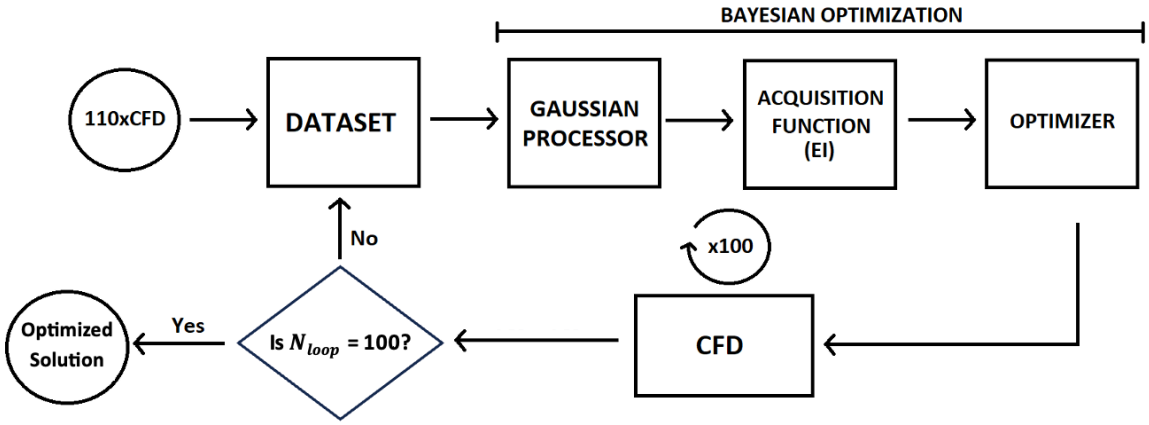


Figure 1: Optimization loop used in the study

The optimization loop used in the study is given in Figure 1. The model reads all the generated data from the CFD results. Within the data, there are the local Reynolds numbers and the corresponding skin friction coefficients calculated over the flat plate. Inside the dataset, C_{a2} and C_{e2} coefficients are defined for the relevant cases. Then, the absolute relative error of skin

friction coefficients with respect to the experimental data is calculated. Before optimization, the coefficients C_{a2} and C_{e2} are scaled to values between 0 and 1, with 0 being the minimum and 1 being the maximum values, to make them of the same weight. Then, the data points having C_{a2}, C_{e2} and relative error values are fitted into the Gaussian Processor (GP), which is a surrogate model used for machine learning. For this, Python's Sklearn.Gaussian process module is utilized. The model fits the data with Gaussian distributions and whenever a random point (C_{a2}, C_{e2}) is asked, it returns its fitted error value and the uncertainty of the model. The data is then registered into the acquisition function, which predicts the data from the GP fit as an objective function and calculates their EI values of error. Then, with the maximize function embedded in the library, 50 iterations are performed trying new points with the EI values based on the fitted data. The function selects the maximum value, and then with the corresponding C_{a2} and C_{e2} coefficients of the minimum-error case, the optimization returns the new coefficients in a scaled format between 0 and 1. Thus, the coefficients are re-scaled into their original interval and are used in the next CFD run. After the run is completed, the new CFD output is added into the whole dataset that includes 110 CFD results. Hence, in each round, a new CFD result is added into the dataset, and the GP is fitted again with the new data. The loop ends once the whole process is repeated for 100 times, returning the optimal solution.

Model and Test Conditions

The model is prepared with the dimensions used in the study of Langtry and Menter [Langtry and Menter, 2009], where the length of the flat plate is taken as 1.5 m. The domain has an upstream extension of 0.04 m and a height of 0.22 m. In contrast to the rounded leading edge with a radius of 0.75 m used by Langtry and Menter [Langtry and Menter, 2009], a sharp leading edge was modeled for simplicity. The type of boundary applied along the flat plate is defined as wall, whereas the upstream portion of the plate as well as the upper boundary of the domain are defined as symmetry (or slip wall). Dimensions and boundary types of the prepared grid can be seen in Figure 2. In the computations, the velocity inlet and pressure outlet type of boundaries are used.

Table 1 gives the inlet conditions for the flat plate test case used in the study. These conditions are defined to be at 0.04 m upstream of the leading edge of the plate [Langtry and Menter, 2009]. On the other hand, the turbulence intensity level quoted in this table corresponds to the leading edge of the plate [Langtry, 2006]. These values were defined at these locations in the current model. The air properties were assumed to be constant in the simulations. The viscosity ratio is defined as the ratio of the eddy viscosity, μ_t, to the molecular viscosity, μ. According to the procedure followed by Langtry and Menter [Langtry and Menter, 2009], in the current study, a viscosity ratio of 12.5 was found to be better matching the measured turbulence decay along the plate instead of the ratio of 12. Hence, that value was used in current computations. Details of this procedure are explained in the next section. As a result, the study is performed with the initial conditions of free stream turbulence intensity 3.3% at the plate leading edge (x=0 m) and the viscosity ratio of 12.5.

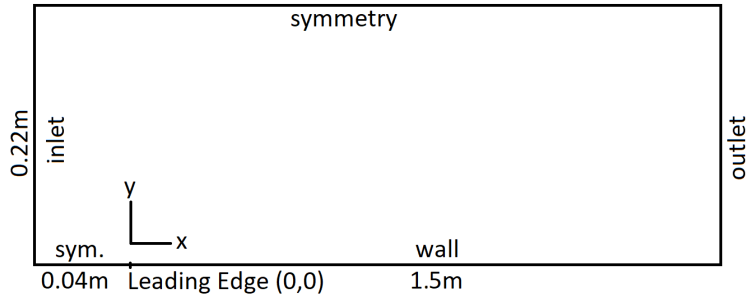


Figure 2: Dimensions and boundary types of the domain

Table 1: Test case details

Test Case	Inlet Velocity (m/s)	Turbulence Intensity (%)	Viscosity Ratio, μ_t/μ	Density (kg/m^3)	Dynamic Viscosity (kg/ms)
T3A	5.4	3.3	12	1.2	1.8x10^{-5}

Grid Sensitivity and Validation

The generated structured grid is presented in Figure 3. The leading edge of the flat plate is marked with a red dot, and the plate continues in streamwise direction towards right until the end of the domain. The dark section around the leading-edge region is due to the further refined grid that was applied to capture the development of the upstream boundary layer and transition behavior correctly.

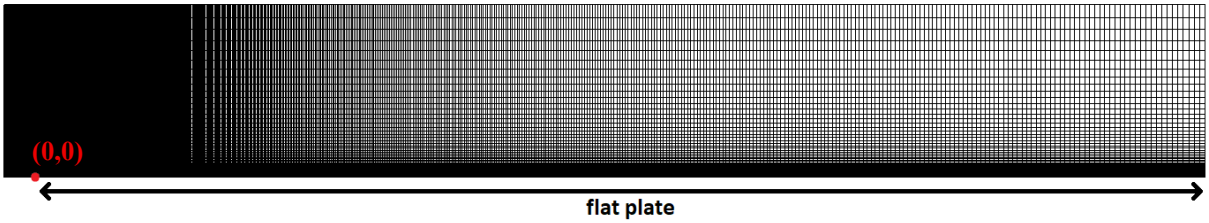


Figure 3: Grid used in the study

In the sensitivity study, four different structured grids were used and named as the coarse (352x70), medium (495x100), fine (704x140) and finest (990x200). These grid sizes were chosen with the aim of doubling the element number as the grid became finer. First layer thickness, growth rate and near-leading edge spacing values at the first cell are also adjusted to keep the aspect ratio of the cells similar between the grids. Details can be found in Table 2. y+ values showed a varying trend having the maximum value at the leading edge and the minimum in the laminar region. In transition, the values increased and stayed level for the rest of the plate. All y+ values were below 0.1.

Table 2: Grid details

Grid	Wall Nodes	Upstream Nodes	Vertical Nodes	First Layer Thickness	Vertical Growth Rate	Near-Leading Edge Spacing	Horizontal Growth Rate	Max y^+
Coarse	320	32	70	1.4x10^{-6}	1.16	1.4 x10^{-3}	1.008	0.051
Medium	450	45	100	1 x10^{-6}	1.11	1 x10^{-3}	1.006	0.04
Fine	640	64	140	7 x10^{-7}	1.08	7 x10^{-4}	1.004	0.031
Finest	900	90	200	5 x10^{-7}	1.05	5 x10^{-4}	1.003	0.025

The measured turbulence intensity decay along the flat plate was used in order to match the inlet boundary condition in the computations. For this, the turbulence intensity level at the inlet of the domain was defined such that the intensity value matches 3.3% at the leading edge of the plate. Once matched, this value was fixed and the inlet viscosity ratio was adjusted via trial and error so that the measured turbulence levels along the flat plate were also matched, which is the procedure described by Langtry and Menter [Langtry and Menter, 2009].

A very good agreement is obtained between the data and the prediction of the turbulence decay with a viscosity ratio of 12.5 as shown in Figure 4. This agreement level is also very

similar to that previously demonstrated [Menter, Smirnov, Liu and Avancha, 2015]. Hence, in all cases presented in this paper, this value was used.

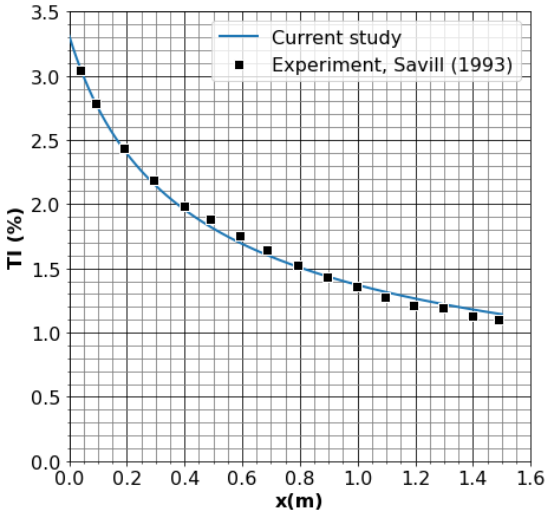
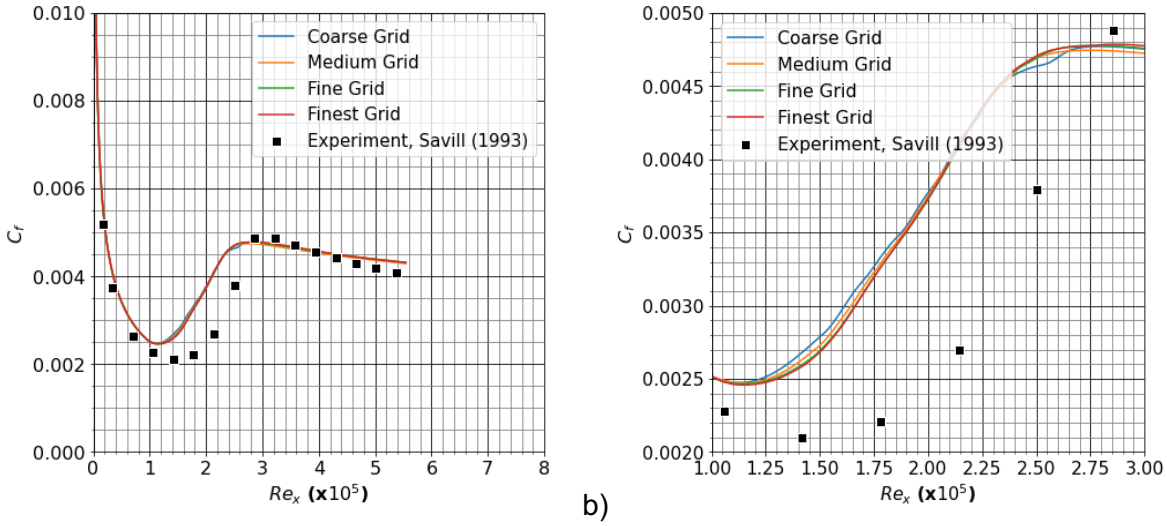


Figure 4: Turbulence decay along the plate



a) b)

Figure 5: Grid sensitivity study: a) variation of skin friction coefficient with Reynolds number, b) zoomed-in version

All CFD runs in the sensitivity analysis utilized the default values of the model coefficients. Figure 5 shows the comparison of the results with each other. In the figure, the y axis represents the skin friction coefficient, C_f, which is the shear stress on the wall divided by the dynamic pressure of free stream flow as given in Eq. (23):

$$C_f = \frac{\tau_w}{\frac{1}{2}\rho U^2} \tag{23}$$

where τ_w is the shear stress acting on the wall, and ρ and U are the density and velocity of the free stream, respectively. The x axis is the local Reynolds number, Re_x, with x being the distance from the leading edge of the plate, which is given by Eq. (24):

$$Re_x = \frac{\rho U x}{\mu} \tag{24}$$

According to the experiment, the transition starts at around 25% of the plate length and ends around the midway. With the start of the transition, the high-momentum fluid replaces the low-momentum fluid by the wall due to promoted mixing. This results in an increase in the shear stress, and therefore in the skin friction coefficient. According to Figure 5(b), there is no significant difference between the fine and the finest grid. The turbulence decay shown in Figure 4 was not affected by grid resolution. To save computational effort, the fine grid is selected for use in the remaining of the study.

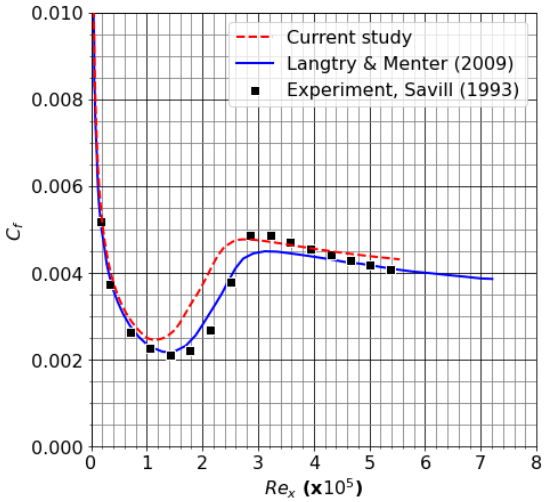


Figure 6: Validation - variation of skin friction coefficient with Reynolds number

Next, the result obtained with the fine grid is compared with both the experiment result [Savill, 1993] and the prediction [Langtry and Menter, 2009] as demonstrated in Figure 6. The trend captured by the current prediction is very similar to that of Langtry and Menter [Langtry and Menter, 2009]. However, there is a mismatch between the two predictions to some degree. The current prediction suggests an earlier transition and overpredicts the value of the skin friction coefficient at the start of transition compared to both the existing prediction and the experiment. On the other hand, the existing prediction underpredicts the value of the skin friction coefficient at the location where the transition ends compared to the experiment and the current prediction. For the two solutions, in addition to the difference in the leading-edge geometries, the solver used in the two studies was also different. Langtry and Menter [Langtry and Menter, 2009] used CFX-5 in their analysis, whereas Fluent Release 2024 R2 was used in the current study. It should be noted that the default values of the coefficients in the destruction equation of intermittency (C_{a2}, C_{e2}) were used for the current prediction given in Figure 6. In this study, it is aimed to obtain the best agreement with the experimental data via optimization of these coefficients.

Preparing the Training Data

After selection of the grid, initially five different cases are prepared for a quick understanding of the effect of the model coefficients C_{a2} and C_{e2}. An initial set of CFD runs is performed with combinations of lower and upper bounds of these coefficients and one with the default values, which are shown in Table 3. Comparison of these five cases with the experiment can be observed in Figure 7(a). The solution with the default values of the coefficients is designated with a red dashed line. The region that is spanned by the bounds, in other words the design space, demonstrate all possible solutions that could be obtained by solely changing C_{a2} and C_{e2}. All these predictions correspond to the test conditions with TI = 3.3% at the plate leading edge and VR = 12.5.

Table 3: Coefficient values used in optimization

Coefficient	Min Bound	Max Bound	Default	Value Array
C_{a2}	0.042	0.078	0.06	0.042, 0.046, … , 0.074, 0.078 (10 values)
C_{e2}	35	65	50	35, 38, … , 62, 65 (11 values)

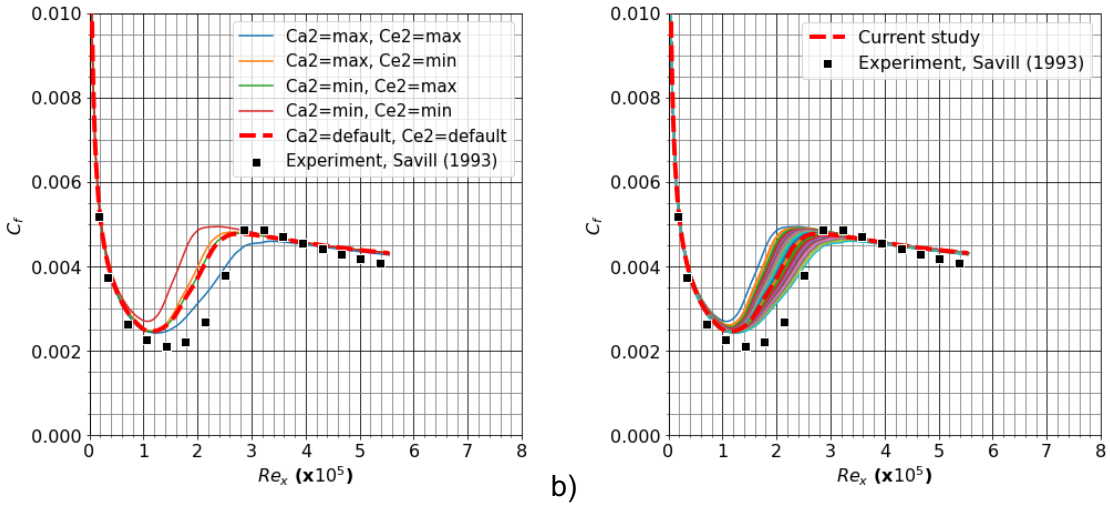


Figure 7: a) Bound and default values of coefficients, b) entire set used for the optimization

After observing these boundaries, this design space is incrementally spanned to create a training set. Several data sets are generated via computations where the coefficients are incrementally varied within their respective intervals in the turbulence model. For the Bayesian optimization, 110 CFD runs are used as the initial dataset. These runs are prepared to have 10 different values for C_{a2} and 11 different values for C_{e2}, as given in Table 3. The bounded region in Figure 7(b) is occupied by all of those runs to serve as a visual aid, including the upper and lower bounds of the coefficients. In Figure 7, it is observed that the solutions obtained using the default values of C_{a2} and C_{e2} fall somewhere inside the design space generated by the extreme values of these coefficients, as expected. The goal of the optimization is to find the closest curve to the experimental data set. However, it is also obvious that the bounding case shown in blue curve in Figure 7(a) is the closest prediction to the data, since according to Figure 7, the design space to be used for optimization mostly misses the experimental data points in the transition region.

Hence, in addition to the current test case, it was decided to explore some other cases as well with some deviation from the target test conditions in order to obtain a design space that covers the experimental data, while keeping this deviation at a minimum. For this, a parametric approach is employed, which also makes it possible to observe the sensitivity of the results to the inlet conditions as well as to the tunable parameters used in the CFD tool. As a result, two parametric studies are planned for the optimization of the coefficients:

Parametric Study 1 - varying the turbulence intensity with respect to the target value, and

Parametric Study 2 - varying the parameter ‘SDR-inner Prandtl number’, $\sigma_{\omega,1}$, while keeping the free stream turbulence intensity 3.3% at the leading edge of the plate.

Both these sets will have a viscosity ratio of 12.5. Since $\sigma_{\omega,1}$ is not varied in Study 1, it is kept at its default value for those cases ($\sigma_{\omega,1}$= 2.0). The effect of the SDR-outer Prandtl number, $\sigma_{\omega,2}$, described in the Method section was also investigated, but it was observed that varying this parameter in the vicinity of its default value was not resulting in an obvious change in the predicted solution, unless it was changed significantly. This is expected since $\sigma_{\omega,2}$ applies to the flow region away from the wall. Hence, the variation of this parameter is not included in the parametric study, and it is used with its default value ($\sigma_{\omega,2}$= 1.168) for all cases studied.

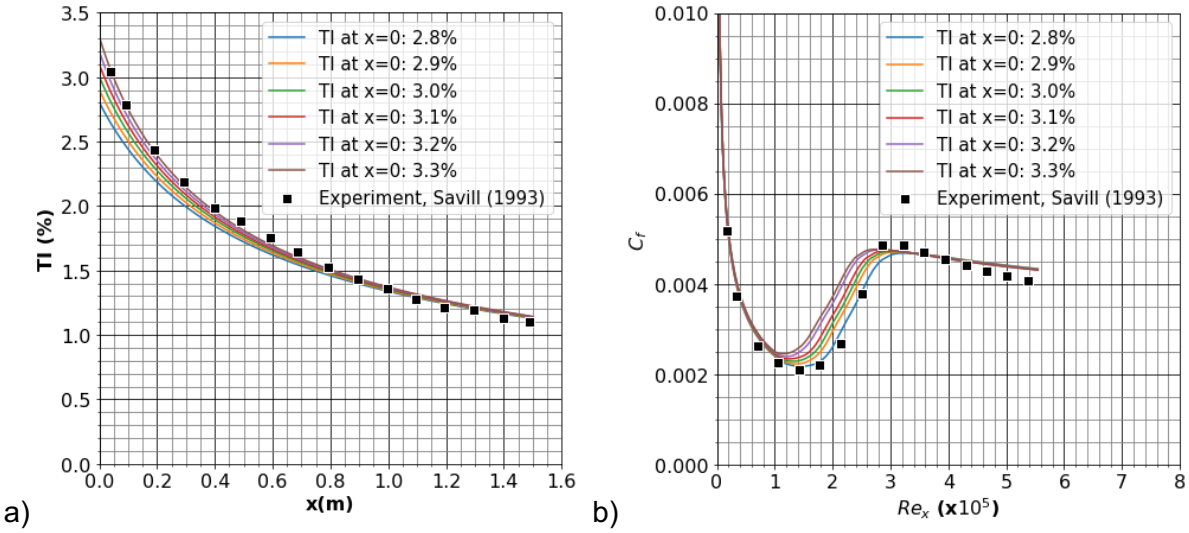


Figure 8: Parametric study 1: effect of turbulence intensity (TI), a) TI decay, b) variation of skin friction coefficient

Since the boundary layer stability can be affected by the free stream turbulence, the transition behavior is expected to differ across different turbulence intensity (TI) values. The resulting curves for Study 1 can be seen in Figure 8. It can be observed that decreasing the turbulence intensity at the leading edge given in Figure 8(a) results in a later transition as in the experiment according to Figure 8(b), hence the agreement with the data improves.

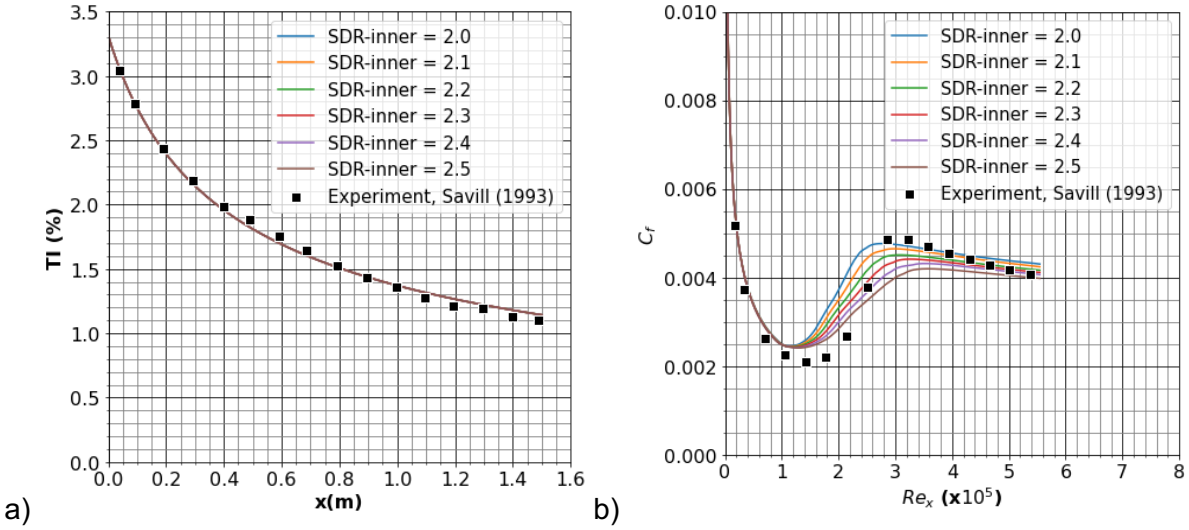


Figure 9: Parametric study 2: effect of SDR-inner Prandtl number with TI = 3.3%, a) TI decay, b) variation of skin friction coefficient

The turbulence decay obtained with the default value of SDR-inner parameter ($\sigma_{\omega,1}$= 2.0) in Figure 9(a) is the same with that shown in Figure 4. Figure 9(a) shows that changing this parameter does not affect the turbulence decay rate acting along the plate. This is not surprising as this parameter controls the dissipation rate in the near-wall region of the plate. On the other hand, as this parameter is increased, the boundary layer transition is predicted to be occurring over a longer distance along the plate, while the magnitude of the skin friction coefficient is mostly underpredicted at the location where transition ends. This is contrary to

the trends observed in Figure 8(b) with varying the turbulence intensity level. There, the turbulence intensity was shown not to have an obvious effect on the transition length.

RESULTS AND DISCUSSION

In this section, first, the design spaces obtained in each parametric study are demonstrated. Next, optimization is performed for all the cases included in each study to present how the optimization results are affected by the selection of the considered parameters. In all presented figures, the CFD solution that was obtained using the default values of the coefficients C_{a2} and C_{e2} is shown with the red dashed line.

In Figure 10, the turbulence intensity at the leading edge of the plate is incrementally decreased starting from the test condition of 3.3%. The increments are kept considerably small so that the condition to be studied does not deviate too much from the actual test condition, therefore still providing a reasonable suggestion for the coefficient set that should be used for this flow problem. It can be observed that the design space in each case is different, which means that it is affected by the turbulence intensity level used. As was observed earlier with Figure 7, the default solution falls inside the design space for each case. As the turbulence intensity is reduced, regardless of C_{a2} and C_{e2} values, transition happens later, and also the predicted lowest value of the skin friction coefficient decreases. This is quite expected since lower turbulence intensity levels result in lower turbulent kinetic energy effective inside the boundary layer which delays transition. Comparison of the default solutions between Figure 10(a) and Figure 10(f) shows that with TI=2.8%, transition starts at a local Reynolds number value of around 1.4×10^5 and the lowest skin friction coefficient observed is close to 0.002. However, with TI=3.3%, transition starting point is closer to a local Reynolds number of around 1.2×10^5 and the predicted lowest value of the skin friction coefficient is closer to 0.0025.

In Figure 11, the value of the SDR-inner parameter, $\sigma_{\omega,1}$, is incrementally increased starting from its default value ($\sigma_{\omega,1}$= 2.0). As this value is increased, the effective diffusivity for the specific dissipation rate decreases as given in Eq. (17). Although varying this parameter does not affect the predicted lowest value of the skin friction coefficient, it affects the length of transition. This can be observed by comparing Figure 11(a) where $\sigma_{\omega,1}$ = 2.0 and Figure 11(f) where $\sigma_{\omega,1}$ = 2.5 via the red dashed lines, which use the default values of the coefficients. Figure 11(a) and Figure 10(f) correspond to the same case where TI = 3.3% and $\sigma_{\omega,1}$ = 2.0, hence transition starts at around a local Reynolds number of 1.2×10^5 and the predicted lowest value of the skin friction coefficient is around 0.0025. What's more visible is the location where the predicted largest value of the skin friction coefficient is observed, which is where the transition ends. In default solution of Figure 11(a), this point corresponds to a local Reynolds number of 2.6×10^5, whereas in the default solution of Figure 11(f), it is shifted to a later point closer to 3.4×10^5, which is a clear increase in the transition length. The corresponding value of the skin friction coefficient also clearly drops and underpredicts the experimental data.

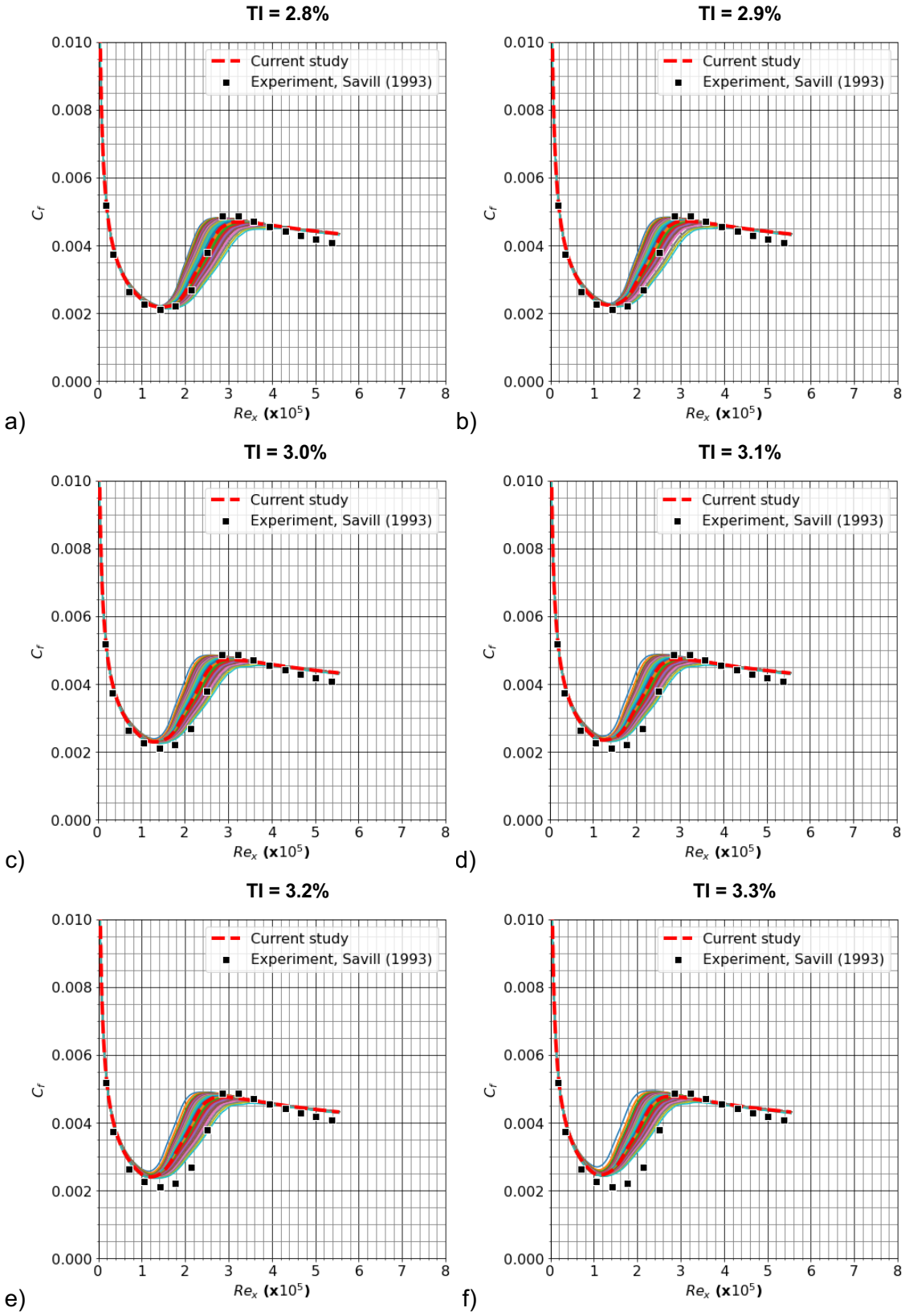


Figure 10: Design space for Study 1 for different turbulence intensity levels

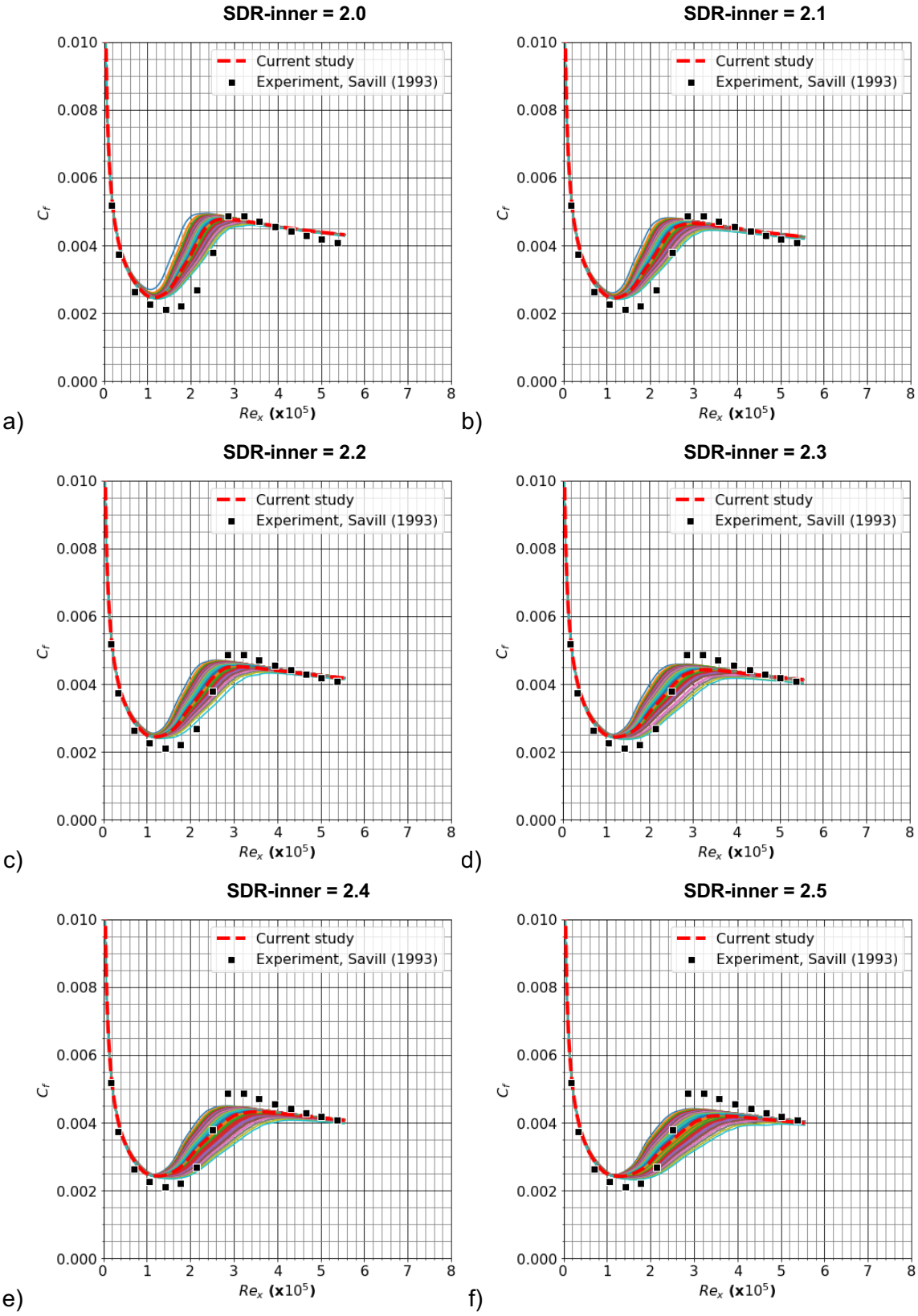


Figure 11: Design space for Study 2 for different SDR-inner parameter values

In both Figure 10 and Figure 11, for the cases where the design space mostly misses the experimental data points in the transition region, which are TI = 3.3%, 3.2%, 3.1%, and $\sigma_{\omega,1}$ = 2.0, 2.1, 2.2, the optimization is expected to come up with the lower bound of the design space which corresponds to the maximum values of C_{a2} and C_{e2}.

Results of Optimization

Using the methodology described in the Optimization Method section, the results of all runs performed for each case are used as an input to the optimization loop in terms of the parameters C_f, Re_x, C_{a2}, and C_{e2}. As a result of this process, the optimized coefficients of C_{a2} and C_{e2} for each case are obtained. The resulting values are presented in Table 4.

Table 4: Coefficients used for the optimized solutions

Case	C_{a2}	C_{e2}
TI = 3.3%	0.078	65
TI = 3.2%	0.078	65
TI = 3.1%	0.078	65
TI = 3.0%	0.074	65
TI = 2.9%	0.066	64.16
TI = 2.8%	0.049	65
SDR-inner = 2.0	0.078	65
SDR-inner = 2.1	0.078	65
SDR-inner = 2.2	0.078	65
SDR-inner = 2.3	0.066	65
SDR-inner = 2.4	0.056	63.93
SDR-inner = 2.5	0.048	65

These results show that the optimization loop selected the maximum values of both C_{a2} and C_{e2} for the cases with TI = 3.1%, 3.2% 3.3% and $\sigma_{\omega,1}$ = 2.0, 2.1, 2.2, which was already expected as discussed previously. On the other hand, for the remaining cases with TI=2.8%, 2.9%, 3.0% and $\sigma_{\omega,1}$ = 2.3, 2.4, 2.5, the optimization still kept C_{e2} close or equal to its maximum value but decreased C_{a2} from its maximum value to produce a closer result to the experiment. Apparently, as it becomes even a later or a longer transition case, C_{a2} must be further decreased to match the experiment. According to these observations, it can be said that tuning C_{a2} is more effective to optimize the prediction.

Set 1 – Variation of Turbulence Intensity

In Figure 12, the curves of skin friction coefficient predicted by both the default and optimized values of C_{a2} and C_{e2} under the effect of different free stream turbulence intensity values are presented in comparison with the experimental data and the prediction by Langtry and Menter [Langtry and Menter, 2009]. As the TI value at the plate leading edge decreases, the optimized prediction is observed to get closer to the prediction with the default values of coefficients. Hence, Figure 12(a) is the case where the best agreement with the experiment is observed despite the under-predicted highest value of the skin friction coefficient; however, the turbulence intensity value used in this case is lower by 0.5% than that used in the experiment. In Table 4, it was observed that as the TI value decreases, the optimizer suggests a lower value for C_{a2}. According to Eq. (3), a decrease in C_{a2} decreases the destruction of intermittency inside the boundary layer, which results in improved agreement with the data.

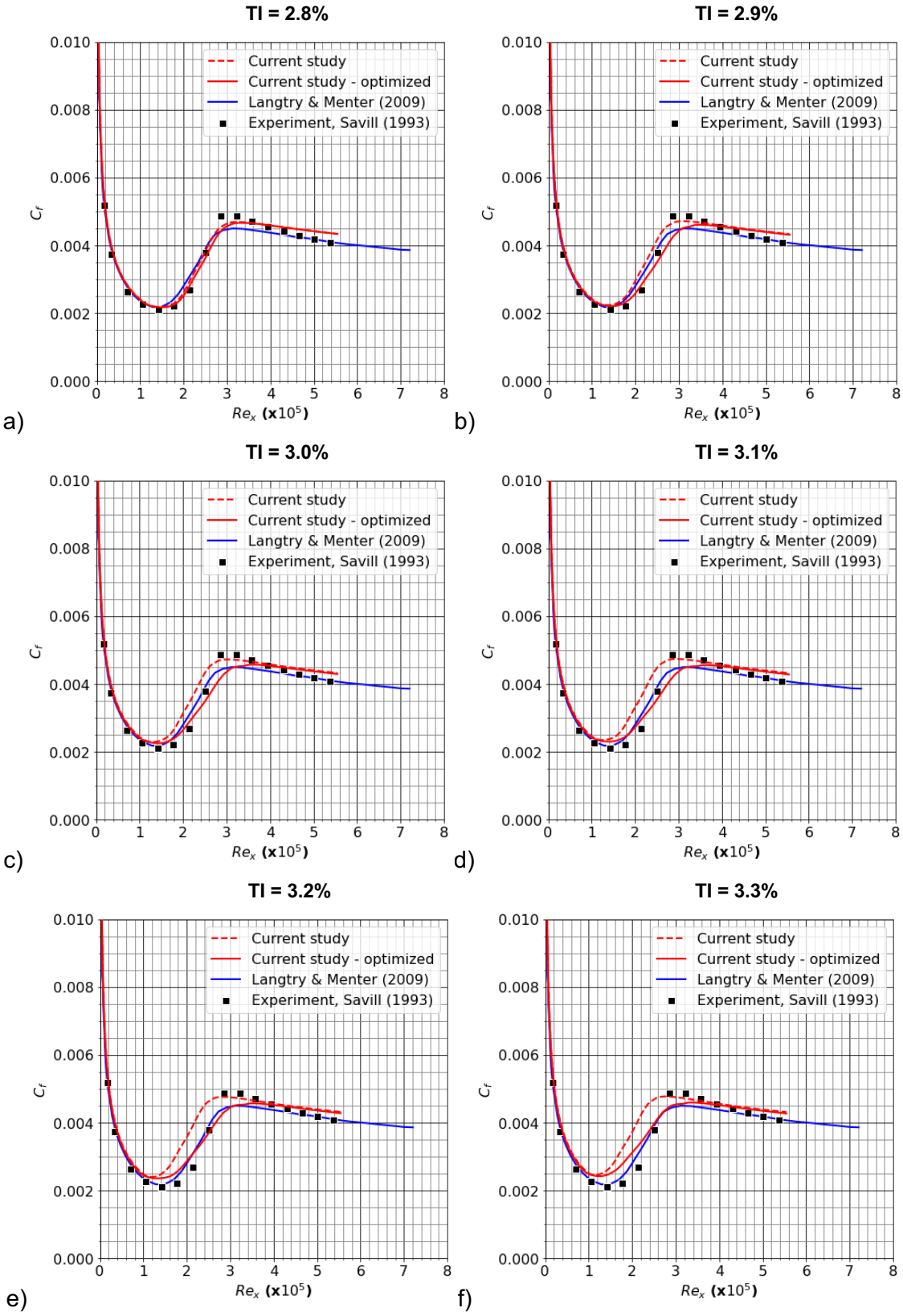


Figure 12: Optimization in Study 1 for different turbulence intensity levels

Set 2 Variation of SDR-inner Prandtl Number

Finally, in Figure 13, the curves of skin friction predicted by both the default and optimized values of C_{a2} and C_{e2} using different values of SDR-inner parameter $\sigma_{\omega,1}$ are presented.

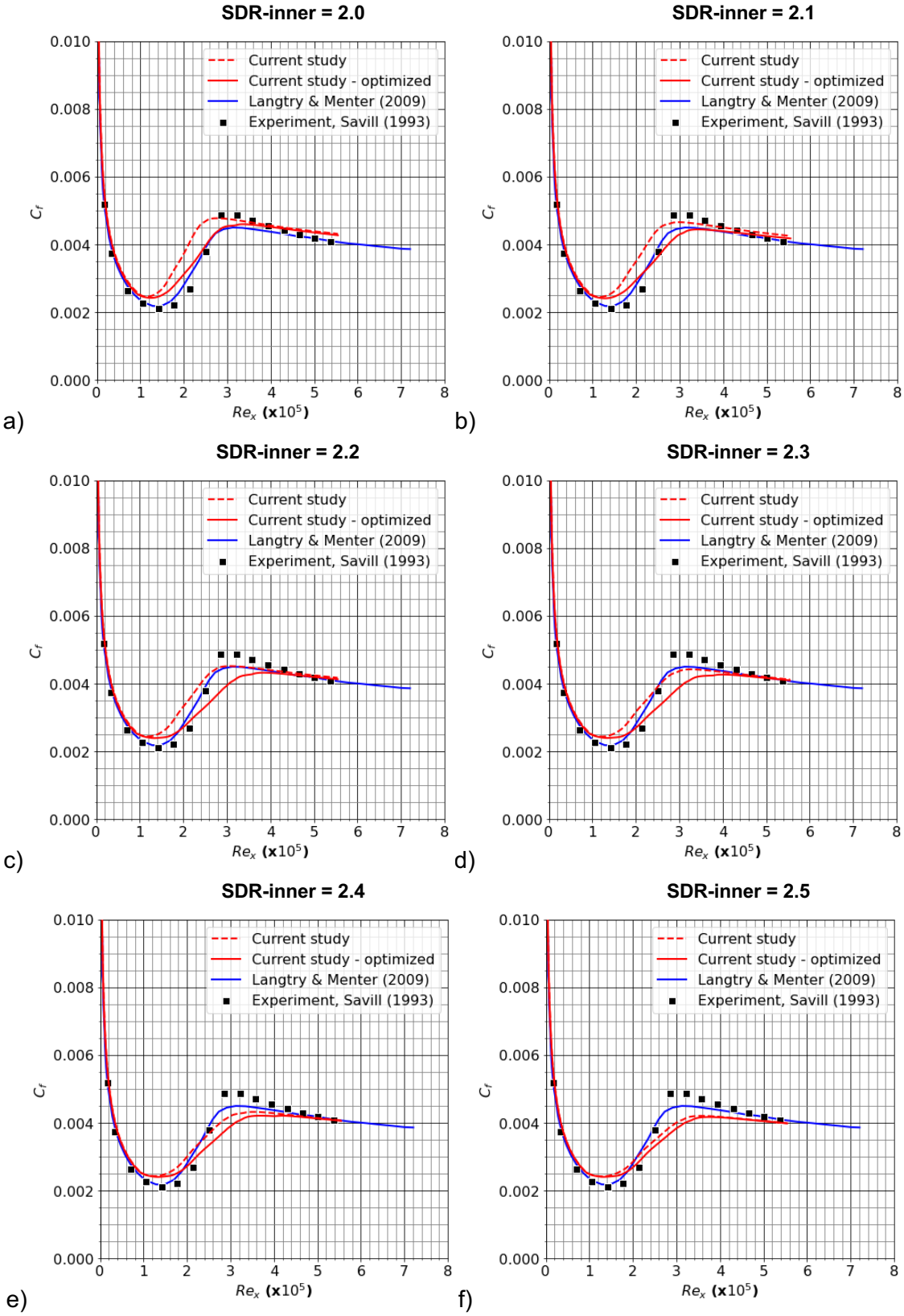


Figure 13: Optimization in Study 2 for different SDR-inner parameter values

Results are again compared with the experimental data and the prediction by Langtry and Menter [Langtry and Menter, 2009]. The value of the SDR-inner parameter, $\sigma_{\omega,1}$, is incrementally increased starting from its default value ($\sigma_{\omega,1}$= 2.0) as was done in Figure 10. For the higher values of this parameter in the studied range, the optimized prediction is observed to be closer to the prediction with the default values of C_{a2} and C_{e2}. Another observation is that increasing $\sigma_{\omega,1}$ results in an increased mismatch with the experimental data. The predicted transition behavior considerably changes since a longer transition is predicted. Among all these cases in Figure 13, the case in Figure 13(a) that corresponds to the experimental conditions for the flat plate examined gives the best optimized solution.

Overall, these results show that predictions can be improved by tuning the coefficients in the Transition SST Model. In the framework of this study, the coefficients in the intermittency destruction term, C_{a2} and C_{e2}, were studied. The optimizer suggested the maximum value of C_{e2} for most of the cases for the optimized solution. Hence, replacing it with one of the coefficients in the intermittency production term, C_{a1} or C_{e1}, might enhance the performance of the optimization. On another note, the optimizer tries to minimize the overall mean error based on the experimental data, which might cause to overlook local errors. Hence, implementing different error definitions could further improve the optimized predictions.

CONCLUSIONS

In this study, a data-driven model is trained to select proper values of coefficients in the intermittency destruction term of the Transition Shear Stress Transport (SST) model, also known as the γ-Re_θ SST model. The test case used in the computations is the ERCOFTAC T3 series of flat-plate experiments. In the initial study presented in this paper, predictions are obtained with combinations of lower and upper bounds of these coefficients and one with default values. Comparisons of these initial results with literature provide satisfactory agreement. Later, two parametric studies are performed, one with varying the free stream turbulence intensity boundary condition and the other with varying the specific dissipation rate inside the near-wall region. With the default values of the coefficients under investigation, it is observed that transition happens later as the free stream turbulence is reduced, which is expected. On the other hand, increasing the specific dissipation rate inside the near-wall region is found to be increasing the transition length. The Bayesian Optimization method is used to optimize the coefficients that were used for all these cases studied to obtain the best matching prediction with the experimental data. The optimization loop uses the Gaussian Processor as the probabilistic model, and the Expected Improvement as the acquisition function. Together, they produce a model which does not work through gradients but gives a chance to the lower uncertainty regions to explore a given case. After the optimization, the optimized results are compared with the experimental data and an existing prediction from the literature. With the demonstrated results, the sensitivity of the predictions to the inlet conditions as well as to the tunable parameters used in the turbulence model could be analyzed.

References

ANSYS, Inc., (2024) *ANSYS Fluent User's Guide*, Release 2024 R2, Canonsburg, PA, USA.

Fan, Y.X., Liu, X., Zhao, R., Zhang, X., Yuan, W., Liu, X., (2024) *Local Correlation, Compressibility and Crossflow Corrections of γ-Re$_\theta$ Transition Model for High-Speed Flow,* Physics of Fluids, 36, 014103.

Krause, M., Behr, M., Ballmann, J., (2008) *Modeling of Transition Effects in Hypersonic Intake Flows Using a Correlation-Based Intermittency Model,* 15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference, AIAA Paper 2008-2598.

Langtry, R. B., (2006) *A Correlation-Based Transition Model Using Local Variables for Unstructured Parallelized CFD Codes,* Ph.D. Thesis, Univ. of Stuttgart, Stuttgart, Germany.

Langtry, R. B. and Menter, F. R., (2009) *Correlation-Based Transition Modelling for Unstructured Parallelized Computational Fluid Dynamics Codes,* AIAA Journal Vol.47, No 12, pp.2894-2906.

Malan, P., Suluksna, K., and Juntasaro, E., (2009) *Calibrating the γ-Re$_\theta$ Transition Model for Commercial CFD, 47th* AIAA Aerospace Sciences Meeting, AIAA Paper 2009-1142.

Menter, F. R., (1994) *Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications,* AIAA Journal Vol.32, No 8, pp.1598-1605.

Menter, F. R., Langtry, R. B., Likki, S. R., Suzen, Y. B., Huang, P. G., and Völker, S., (2006) *A Correlation Based Transition Model Using Local Variables Part 1: Model Formulation,* Journal of Turbomachinery, Vol. 128, No. 3, pp.413–422.

Menter, F. R., Smirnov, P. E., Liu, T., and Avancha, R., (2015) *A One-Eqution Local Correlation-Based Transition Model,* Flow, Turbulence and Combustion Vol. 95, pp.583-619.

Montgomery, D. C., and Runger, G. C. (2014). *Applied Statistics and Probability for Engineers*. John Wiley & Sons.

Rasmussen, C., E., and Williams, C., (2006) *Gaussian Processes for Machine Learning,* MIT Press.

Rausa, A., Guardone, A., and Auteri, F., (2023) *Implementation of γ-Re$_\theta$ Transition Model within SU2: Model Validation,* AIAA SciTech Forum, AIAA Paper 2023-1570.

Savill, A. M., (1993) *Further Progress in the Turbulence Modeling of By-Pass Transition*, edited by Rodi W. and Martelli F., Engineering Turbulence Modelling and Experiments, Elsevier, New York, pp.583-592.

Savill, A. M., (1996) *One-Point Closures Applied to Transition,* Turbulence and Transition Modelling*,* ERCOFTAC Series, Lecture Notes from the ERCOFTAC/IUTAM Summerschool held in Stockholm, 12-20 June, 1995, edited by M. Hallbäck, D.S. Henningson, A.V. Johansson, P.H. Alfredsson, Vol. 2, Springer Netherlands, pp. 233–268.

Snoek, J., Larochelle, H., and Adams, R. P., (2012) *Practical Bayesian Optimization of Machine Learning Algorithms*, Advances in Neural Information Processing Systems 25, pp.2951-2959.

Song, Z., Liu, Z., Lu, J., and Yan, C., (2022) *Quantification of Parametric Uncertainty in γ-Re$_\theta$ Model for Typical Flat Plate and Airfoil Transitional Flows,* Chinese Journal of Aeronautics, Vol.34, No 4, pp.237-251.

Yuntao, W., Yulun, Z., Song, L., and Dehong, M., (2014) *Calibration of a γ-Re$_\theta$ Transition Model and Its Validation in Low-Speed Flows with High-Order Numerical Method,* Chinese Journal of Aeronautics, Vol. 28, No 3, pp.704-711.

Zhao, R., Zhang, X., Shen, L., Fan, Y., and Liu, F., (2025) *Improved γ-Re$_\theta$ Transition Model for Hypersonic Cavity-Induced Transition Predictions,* Acta Mechanica, 236, 2529-2549.

Zhang, Y. F., Zhang, Y. R., Chen, J. Q., Mao, M. L., Jiang, Y., (2017) *Numerical Simulations of Hypersonic Boundary Layer Transition Based on the Flow Solver Chant 2.0*, 21st AIAA International Space Planes and Hypersonic Technologies Conference, AIAA Paper 2017-2409.