21st Australasian Fluid Mechanics Conference
Adelaide, Australia
10-13 December 2018

Magnetic Resonance Imaging, Optimization, and Machine Learning to Understand and Model Turbulent Mixing

Pedro M. Milani and John K. Eaton

Department of Mechanical Engineering
Stanford University, California, United States

Abstract

We summarize an evolving set of "big data" approaches to study and model turbulent mixing with specific application to gas turbine film cooling. New magnetic resonance imaging (MRI) diagnostics allow rapid acquisition of 3D mean velocity and scalar concentration data at millions of spatial points. Optimization techniques were used with the large data sets to tune the constants of existing scalar transport models and to show the required characteristics of new models. Substantial improvements in predictive accuracy can be achieved by introducing an anistropic scalar diffusivity formulation and correcting an error in the standard near-wall formulation of the scalar transport model. The numerical experiments showed that the turbulence model in the scalar transport equation is the main culprit in the poor accuracy of modern Reynolds-Averaged Navier Stokes (RANS) codes. Our latest modeling effort uses experimentally validated, wall-resolved Large Eddy Simulation (LES) results as training data for a machine learning algorithm that replaces the scalar transport model in an existing computational fluid dynamics code. Major reductions of the error in scalar concentration have been documented using this approach.

Introduction

Turbulent mixing of scalar contaminants is a critical process in both nature and technology. Scalar contaminants are quantities that are transported by the flow without affecting the velocity field. Examples include trace gases in airflows, dissolved chemicals that do not affect the fluid properties of the carrier liquid, heat, and dispersed particles or droplets that are small enough to follow the turbulent motions. In many cases such as fuel mixing, volcanic ash dispersion, and thermal mixing in enclosures turbulent scalar mixing is the dominant mechanism controlling the behavior of the system.

Accurate prediction of turbulent scalar mixing is needed for several problems ranging from engineering design to estimating the impact of ongoing natural disasters. Modern eddy-resolving simulations including both Direct Numerical Simulation (DNS) and high-resolution Large Eddy Simulation (LES) can accurately predict scalar transport and dispersion. Generally such simulations are limited to relatively simple flow configurations at moderate Reynolds number and require long run times on large supercomputers. Therefore, a less computationally intensive approach is required for highly complex configurations, rapid response predictions such as needed during a natural disaster, or for engineering optimization where many variations of a design must be analyzed numerically. Coupled solutions of the Reynolds-Averaged Navier Stokes (RANS) equations for the velocity and pressure fields and the Reynolds-Averaged Advection-Diffusion (RAAD) equation for a transported scalar are used in such situations, and Reynolds-averaged codes are expected to be the industrial workhorse for the foreseeable future. The equation that is solved for the mean scalar field is shown below in Eq. 1, where $\overline{u_i'c'}$ is the unclosed term that demands modeling and α is the molecular diffusivity.

$$\bar{u}_i \frac{\partial \bar{c}}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\alpha \frac{\partial \bar{c}}{\partial x_i}\right) - \frac{\partial}{\partial x_i}\left(\overline{u_i'c'}\right) \tag{1}$$

RANS/RAAD codes use semi-empirical models to represent the effects of turbulence on the mean velocity, pressure, and scalar concentration distributions. The turbulent scalar transport models are a particular weakness of this approach, and there has been little effort to improve the models over the past 20 years [25]. Most practical models are based on a simple gradient diffusion hypothesis with a scalar turbulent diffusivity thus representing the turbulent scalar transport as Eq.2 below:

$$\overline{u_i'c'} = -\alpha_t(x,y,z)\frac{\partial \bar{c}}{\partial x_i} \tag{2}$$

Here α_t is the turbulent scalar diffusivity, which is a function of position in the flow field. Usually, α_t is assumed to be directly proportional to the eddy viscosity (ν_t) calculated as part of the RANS momentum solver. This is called the Reynolds analogy and represented as Eq. 3:

$$\alpha_t = \frac{\nu_t}{Sc_t} \tag{3}$$

The turbulent Schmidt number (Sc_t), or equivalently the turbulent Prandtl number for heat transport (Pr_t), is taken as a constant that may be tuned to match experimental data. Most often a value of 0.85 is used because it best matches experimental profiles in the log-layer of turbulent boundary layers over a flat plate [11].

Our group and many others have shown that conventional scalar transport models are not predictive. We have spent the past eight years compiling extensive data sets for a wide range of jet-in-crossflow configurations relevant to gas turbine film cooling. We have found that well resolved RANS/RAAD simulations using conventional gradient diffusion models can produce very large errors in the mean scalar concentration and even the relative performance between different jet configurations is not consistently predicted. This is true even when the velocity field computations are reasonably accurate, pointing the finger directly at the turbulent scalar transport models. More advanced generalized gradient diffusion hypothesis models produce at most minor improvements [22].

About ten years ago, our group began developing "big data" approaches to improve scalar transport modeling with a specific focus on gas turbine cooling technology. Film cooling systems consist of arrays of angled jets in crossflow and/or 3D slot jets with the goal of forming a layer of relatively cool air over the turbine blade surface to protect it from the extremely hot mainstream gases. The effectiveness of this cooling method is controlled by the rate of turbulent mixing of the coolant jets with the mainstream.

We developed new Magnetic Resonance Imaging (MRI) protocols to provide detailed mean velocity measurements of the three-component mean velocity and scalar concentration on regular Cartesian grids throughout complex flow fields. These techniques were applied to approximately 15 different geometries each for trailing edge slot cooling and jet-in-crossflow configurations. Each of these experiments included measurements at 0.5 to 4 million spatial points, providing a very large database to develop models. Our first big data approach to model development involved using the MRI data in optimization procedures to find the best constants to use for various types of models. One curious finding was that accurate prediction of jet-in-crossflow configurations required a region of negative α_t immediately downstream of the jet exit, indicating that gradient diffusion is an insufficient model for these flows. This led us to conduct wall-resolved LES computations for the exact configuration of several of our experiments. The experimentally-validated simulations provide turbulence data that enriches the experimental results. In particular, the turbulent scalar transport is resolved throughout the flow field. We are using machine learning techniques to develop new models that have demonstrated substantial improvements in predictive accuracy relative to conventional models and new methods to interpret machine learning models are helping us to understand why existing models fail.

The goal of this paper is to show the evolution of our big data approaches to scalar transport modeling with particular emphasis on film cooling configurations. Much of this work has been reported in previous works by members of our lab, but new results are also reported. We make no attempt to review other methods for 3D measurement of turbulent mixing or advanced scalar transport model development in this short paper.

Magnetic Resonance Experiments

The research reported here relies on two MRI-based techniques to measure the mean velocity and scalar concentration in turbulent flows. Both techniques use water as the working fluid and 3D printed plastic flow models. Experiments are conducted in conventional 3.0 Tesla GE medical MRI scanners. Signal is generated in an MRI scanner by magnetizing the protons in hydrogen atoms bound in water. Thus, there is no question that flow tracers follow the flow. Copper sulfate dissolved in the water is used to enhance the signal strength for the velocity measurement and as the passive scalar for the scalar concentration measurements.

Magnetic Resonance Velocimetry (MRV) uses a phase contrast measurement protocol [21, 8] to measure three components of mean velocity. Typical flow velocities are around 0.5 m/s giving Reynolds numbers of the order of 25,000 in a standard 50 mm channel. A single acquisition of a full velocity field requires around 5 minutes of scanning. This acquisition sequence is repeated 15 to 25 times to ensure well-converged statistics for high turbulence mixing experiments. Velocity scans are interspersed with reference "flow-off" scans bringing the total time for a single full field experiment to approximately two to three hours depending on the size of the test section. The velocity measurement uncertainty is around 5% at 95% confidence, and it is significantly lower over most of the measurement domain. Uncertainties are highest near the walls of the channel and in the neighborhood of features of the apparatus such as flanges that perturb the magnetic fields.

Magnetic Resonance Concentration measurement (MRC) is based on the linear relationship between copper sulfate concentration and MRI signal strength [1]. It produces a 3D map of scalar concentration on the same measurement grid as the MRV data. The signal strength is also a function of the flow velocity, so the output must be calibrated in situ for each experiment. Typically setup, calibration, and 20 to 30 repeated scans takes around 12 hours in the scanner. Accurate measurements down to about 2% of the inlet copper sulfate concentration are routine. Higher resolution measurements are achieved by feeding high concentration copper sulfate at the jet inlet, in which case the measurements near the jet inlet are not useful due to signal non-linearity at high concentration. Our latest measurements using this technique have achieved resolutions below 1% concentration.

A typical experiment studying a single pitched jet in crossflow uses a 6 mm hydraulic diameter jet feeding into the crossflow in a 50 mm square test section. Measurements cover a 240 mm length of the test section with spatial resolution of approximately 0.6 mm in all three directions. The measurements also include the flow in the jet supply plenum and inside the jet feed hole. Thus a typical data set includes three components of velocity and the scalar concentration at approximately 2.5 million spatial points. Figure 1 shows 5 of the 356 measurement planes of data from a simple case of a round hole pitched at 30° from horizontal with a blowing ratio ($BR \equiv V_{jet}/V_{main}$) of 2.0. Within the data planes, the velocity and concentration data are only shown for points where the concentration is greater than 20% of the inlet concentration. The figure shows the strong coupling between the mean velocity and concentration fields. The jet separates from the wall and a counter-rotating vortex pair sweeps mainstream fluid under the jet reducing the cooling effectiveness. Turbulent mixing acts to spread the jet laterally and upward, but in this near jet region, the turbulence is unable to counteract the upward flow between the vortices. The comprehensive and detailed nature of these data sets drove us to consider big data approaches to analyze, interpret, and model the results.

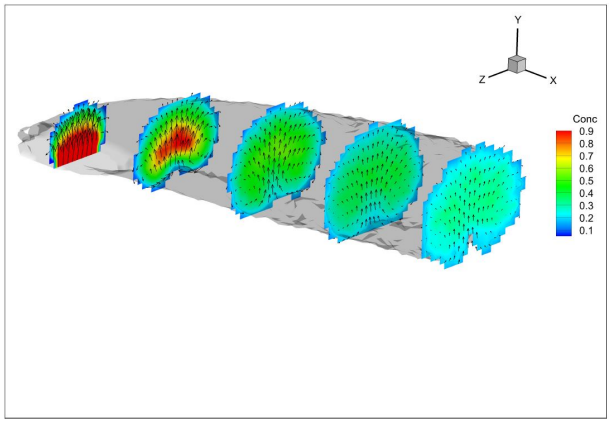

Figure 1: Slices showing mean secondary flows (arrows) and concentration (contour) data from an MRI experiment. Data are plotted only for locations where $\bar{c} > 0.2$ (delineated by the gray isosurface). Only $1/4$ of velocity vectors shown for clarity.

Optimization Studies

A typical MRI dataset contains mean velocity and concentration data on millions of points of a 3D Cartesian grid. Such data can be analyzed directly to understand physical phenomena in flows where turbulent mixing is important. For example, Coletti et al. [5] used MRI data to study the effect of mainstream pressure gradients on a circular jet-in-crossflow geometry. With full field measurements, the authors were able to analyze, at the same time, the incoming boundary layer, the in-hole velocities, the

structure of the counter-rotating vortex pair, and the resulting passive scalar concentration, which is nearly impossible to do with other experimental techniques.

Another important objective of MRI studies is to obtain the means to improve turbulence modeling, towards the goal of making RANS simulations predictive. A problem is that magnetic resonance techniques cannot directly measure the quantities that turbulence models aim to predict. For example, we are able to measure the mean concentration field $\bar{c}(x,y,z)$, but not the turbulent scalar flux $\overline{u_i'c'}(x,y,z)$, which is the term that RANS models need to close in the Reynolds-averaged scalar transport equation. To resolve this, our first approach involved an optimization procedure.

Formally, in the case of turbulent scalar mixing, we postulate that the turbulent scalar flux is a known function f of the mean quantities available to RANS solvers, such as the mean velocity, mean concentration, and other modeled quantities such as eddy viscosity ν_t, as shown in Eq. 4. This function depends on one or more parameters, denoted by θ. Then, we define an error metric E, which is a non-negative number that represents how far off the mean scalar concentration field predicted by the RANS solver, $\bar{c}_{RANS}(x,y,z)$, is from the MRC field data, $\bar{c}_{MRC}(x,y,z)$, shown in Eq. 5. Since the RANS solver uses f as a closure for the turbulent scalar flux, the error E will be a function of the parameters θ used. Finally, an optimization algorithm is used to set the parameters θ such that the error E is minimized, shown in Eq. 6. With this, we hope to produce a better closure f than the standard model for turbulent scalar flux based on $Sc_t = 0.85$.

$$\overline{u_i'c'} = f(\bar{u}_j, \bar{c}, ...; \theta) \tag{4}$$

$$E(\theta) = \text{discrepancy between } \bar{c}_{RANS} \text{ and } \bar{c}_{MRC} \tag{5}$$

$$\theta = \arg\min(E) \tag{6}$$

Trailing Edge Slot Film Cooling Studies

In this subsection, we discuss the results of applying the general optimization framework described above to the problem of predicting scalar concentration in trailing edge slot film cooling. Figure 2(a) shows a gas turbine blade, with a typical slot film cooling geometry presented in Fig. 2(b). Cooler fluid is pumped out through the slots to protect this critical region; when designing the cooling system, it is desirable to minimize the coolant flow while providing sufficient protection against thermal stresses.

Ling et al. [13] applied the optimization framework coupled with MRC data in a simple way. As the function f, they picked the typical gradient diffusion hypothesis with an isotropic diffusivity and a uniform Sc_t, as shown in Eq. 7. The turbulent Schmidt number was allowed to vary as the only free parameter θ in this case. For the error metric, they interpolated simulation results and the experimental data onto the same mesh and averaged the absolute value of their difference over all the computational cells of a region of interest (ROI) as shown in Eq. 8 (N is the number of cells in the region). To minimize E, they used exhaustive search: different RANS simulations were performed using different values of Sc_t, and the optimal value was chosen as the one that minimized the error E. Note that this approach is feasible only when there is a small number of parameters (1 or 2), because its computational cost grows exponentially with the number of parameters that need to be set.

$$\overline{u_i'c'} = -\frac{\nu_t}{Sc_t}\frac{\partial \bar{c}}{\partial x_i} \tag{7}$$

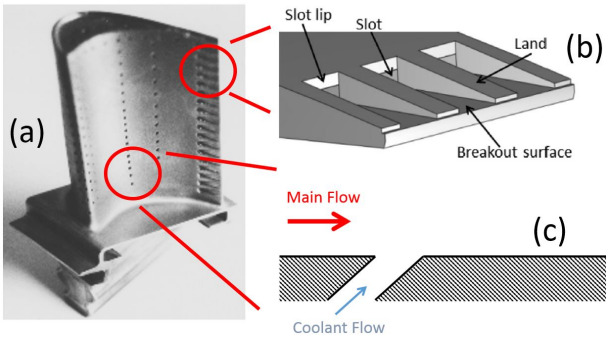

Figure 2: Schematic showing film cooling schemes in a gas turbine blade. (a) shows a picture of a turbine blade, taken from `http://www.me.umn.edu/labs/tcht/measurements/what.html`. (b) zooms onto the trailing edge, where slot film cooling is used, adapted from [16]. (c) shows the center plane of a single inclined jet-in-crossflow, which is used to model the film cooling holes in the middle of the blade.

$$E(Sc_t) = \frac{1}{N}\sum_{ROI} |\bar{c}_{RANS} - \bar{c}_{MRC}| \tag{8}$$

They used 4 different flow configurations (on which 4 separate MRV/MRC experiments were run). They also explored two different regions of interest (and thus two different versions of the error metric E): one that encompassed the full mixing region of a single slot, and a second, smaller region, that only encompassed cells near the wall. The latter was chosen because it is critical for film cooling applications that RANS simulations be able to predict correct mean scalar distributions close to the blade surface. In total, 8 different optimal values of Sc_t were found. In general, optimal values of 0.85 (or a little lower in some cases) were found when the region of interest consisted of the full mixing region; but they were noticeably lower (as low as 0.25) when the near-wall ROI was used. In all cases, the improvement over the baseline model ($Sc_t = 0.85$) was modest. This suggests that a fixed Sc_t model for the whole domain is inappropriate, since the only way to improve results near the wall is to worsen the results in the overall domain. Besides, it suggests that different physics are governing the mixing near the wall compared to the rest of the domain, in a way that the Reynolds analogy is not able to capture.

These results then led Ling et al. [16] to propose a more general improvement to the turbulent scalar mixing model. They realized that, due to the different boundary conditions at the wall for velocity u and scalar concentration c (no-slip versus adiabatic), the turbulent transport of the two quantities ($\overline{u'v'}$ and $\overline{v'c'}$) should behave differently as one asymptotically approaches the wall [26]. This means that the Reynolds analogy breaks down: α_t, which governs $\overline{v'c'}$, cannot be proportional to ν_t, which governs $\overline{u'v'}$, close to the wall. Thus, they proposed removing the Van Driest correction from the turbulent diffusivity to obtain the correct asymptotic scaling, which is equivalent to setting a corrected Schmidt number shown in Eq. 9.

$$Sc_t^{corrected} = Sc_t(1 - e^{-y^+/70}) \tag{9}$$

y^+ is the distance from a cell to the nearest solid surface in wall coordinates, and 70 is a dimensionless constant chosen to ensure that the proposed correction would affect not only the viscous sublayer, but also the buffer layer and parts of the log layer for these film cooling flows. In this formulation, Sc_t is still a single constant, but the solver uses $\alpha_t = \nu_t / Sc_t^{corrected}$ for

the turbulent diffusivity field. Ling et al. [16] used the correction of Eq. 9 together with $Sc_t = 0.45$ across 15 different slot film cooling cases for which they had MRI data. They obtained significant improvement in most cases when predicting the near wall concentration (see Fig. 3), and also obtained improvement in overall error metric in 13 out of the 15 cases.

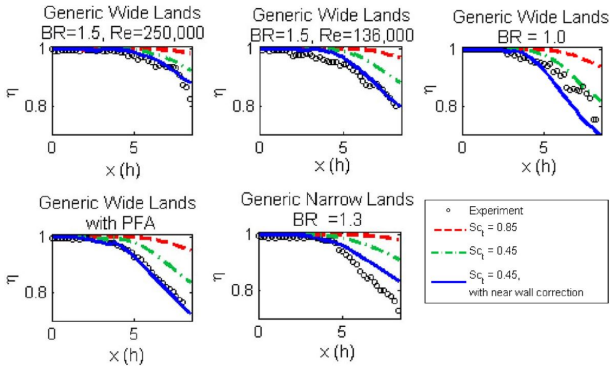

Figure 3: Plots adapted from Ling et al. [16] showing spanwise-averaged adiabatic effectiveness (i.e., scalar concentration at the wall) η versus streamwise coordinate x. Only 5 out of the 15 cases are shown for brevity.

Inclined Jet-in-Crossflow Uniform Models

Another important flow configuration for turbine blades is discrete hole film cooling (as shown in Fig. 2(c)). The design requirements are identical to the slot film cooling case (minimize coolant flow while providing enough thermal protection). Ideally, the flow coming out of the hole would remain close to the bottom wall at all times, but due to mean secondary flows and turbulent mixing, hot fluid from the mainstream is entrained and touches the bottom wall. A passive scalar is used to model the transport of heat because at such high speeds, buoyancy effects and molecular transport are negligible. It is a well known fact in the turbomachinery community that RANS simulations do a poor job of predicting the temperature distribution, particularly close to the blade surface (e.g. as shown, for example, by Nikparto et al. [20], Coletti et al. [4], and Hoda and Acharya [10]).

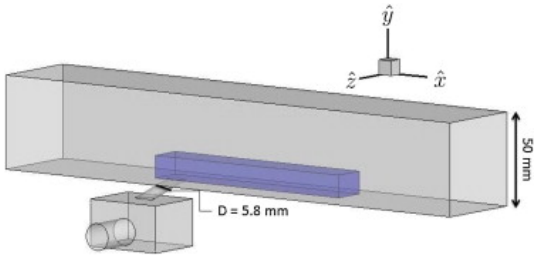

Figure 4: Baseline jet-in-crossflow geometry, taken from Ling et al. [12]. The region of interest for evaluation of the error metric is shown in blue, and D is the diameter of the circular hole.

Ling et al. [12] introduced the optimization framework for discrete hole film cooling using data from a single MRV/MRC dataset in a baseline geometry, shown in Fig. 4. In this case, a single cylindrical hole is inclined 30° with respect to the horizontal, the blowing ratio is 1.0, and the Reynolds number based on hole diameter is $Re_D = 2900$. Their function f was again based on the gradient diffusion hypothesis; instead of using the Reynolds analogy with Sc_t as a parameter, they used a uniform turbulent diffusivity α_t to parameterize f. They calculated their error metric E over the whole region where the jet interacts with the mainstream, shown in blue in Fig. 4. They explored different options for the error metric E, including the L1 norm (average of $|\bar{c}_{RANS} - \bar{c}_{MRC}|$), the L2 norm (average of $(\bar{c}_{RANS} - \bar{c}_{MRC})^2$), and the Huber norm (a mixture between the L1 and L2 norms). They concluded that the approach was robust to the specific form of E and decided to utilize the L1 norm (shown in Eq. 8). Another interesting result from the same paper [12] was obtained using an anisotropic diffusivity coupled with the gradient diffusion hypothesis. They used two free parameters θ to describe f, a diffusivity in the wall-normal direction ($\alpha_{t,y}$), and a diffusivity in the spanwise direction ($\alpha_{t,z}$), as shown in Eq. 10. Note that the diffusivity prescribed along the streamwise direction barely affects the results, because transport of $\bar{c}$ in that direction is dominated by mean advection.

$$\overline{u'c'} = -\frac{\alpha_{t,y}+\alpha_{t,z}}{2}\frac{\partial \bar{c}}{\partial x};\ \overline{v'c'} = -\alpha_{t,y}\frac{\partial \bar{c}}{\partial y};\ \overline{w'c'} = -\alpha_{t,z}\frac{\partial \bar{c}}{\partial z} \quad (10)$$

The optimization, again based on exhaustive search, showed that allowing for two parameters instead of one caused a decrease in the resultant error function; and, we usually found that the optimal wall-normal diffusivity is smaller than the optimal spanwise diffusivity. This is probably due to the presence of the wall, which suppresses eddies (and thus mixing) in the y direction. It is important to note that, to obtain $\bar{c}_{RANS}$, only the RAAD equation, Eq. 1, is solved; the mean velocity field from the MRV experiment is used, so that any error will come from the mixing model.

Inclined Jet-in-Crossflow Non-Uniform Models

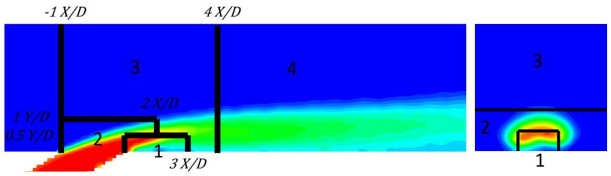

Figure 5: Contours of the MRC data for the baseline jet-in-crossflow case showing the four distinct regions. For the anisotropic and non-uniform optimization, we set different values of diffusivity, independently, for each region. The left panel shows a center plane and the right panel shows a streamwise plane $2D$ downstream of the center of the hole.

The optimized anisotropic model for discrete hole film cooling was clearly limited by the use of constant scalar diffusivity values throughout the field. We extended this approach by allowing the diffusivities to vary over space, and this paper is the first report of that work. The full domain was split into four distinct partitions as shown in Fig. 5, and then we used optimization to choose a different value of $\alpha_{t,y}$ and $\alpha_{t,z}$ for each partition. Partitions 1-3 cover the near injection region (where we expect most of the interesting physics to happen), and partition 4 encompasses the rest of the domain. Under this formulation, function f would be similar to the one shown in Eq. 10, except that the diffusivities are now piece-wise constant. This function f contains 8 optimizable parameters, so it is virtually impossible to perform optimization using exhaustive search, especially because in the present case one evaluation of the error metric requires solving a 3D partial differential equation. Instead, we

deployed a genetic optimization algorithm. The flowchart in Fig. 6 contains brief descriptions of the different steps. The algorithm attempts to mimic the biological concept of evolution. We start with a random population of individuals (a set of possible parameters). In each iteration, the individuals reproduce (two "parent" solutions create an "offspring" solution based on their parameter values), random mutations are introduced, and less fit individuals (defined by their error metric) are eliminated. After enough iterations (called generations), the fittest individual emerges as the algorithm's final answer. More details can be found in Elbeltagi et al. [7].

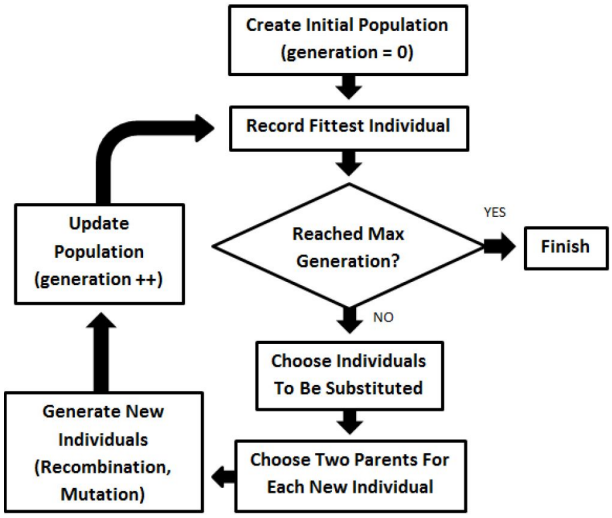

Figure 6: Schematics describing our genetic algorithm.

We performed this non-uniform optimization across 5 different jet-in-crossflow cases where MRI data were available: the baseline case shown in Fig. 4, two cases where the the mainstream pressure gradient is different (favorable and adverse pressure gradients), and two where the incoming boundary layer thickness is different (thin and thick boundary layers). For more information on these datasets, consult Ryan et al. [23]. As before, the concentration $\bar{c}_{RANS}$ was calculated using the experimental mean velocity field and compared to $\bar{c}_{MRC}$ through an error metric E; in this case, the Huber norm was chosen due to it being insensitive to small deviations (a good property due to experimental uncertainty), but not dominated by outliers. However, as shown by Ling et al. [12], the optimization results are robust to different types of metrics, so we would have gotten very similar results had the L1 norm of Eq. 8 been used instead.

Allowing for spatial variation of the diffusivities improved the calculated mean concentration results significantly: it yielded an average 18% reduction in overall error metric compared to the best anisotropic uniform solution of Ling et al. [12]. The improvement in the critical regions (close to the wall and close to injection) was even more striking: when evaluated only on partition 1, the error metric reduced by an average of 32% across the 5 cases. Fig. 7 shows mean concentration results from the baseline jet-in-crossflow case; it indeed shows that the reduction in error metric translates to qualitative improvements in the simulation results. An important characteristic of the mean concentration field for this case is the low values of concentration right under the jet, past injection (where partition 1 is). This happens because the coolant jet separates from the bottom wall before reattaching, and the counter-rotating vortex pair acts to bring mainstream flow under the jet. The experiments show that the turbulent mixing in that region is not enough to counteract those two effects, but in typical RANS simulations this mixing is significantly overestimated in the region. As is clear from the second row in Fig. 7, this causes an over-prediction of the mean concentration at the wall, and also causes the concentration contour in the streamwise plane to lose its characteristic kidney shape. The optimal non-uniform diffusivity ameliorates these issues, while maintaining the good results far downstream of injection. Especially due to the improvement near the wall, such results are of great interest in predicting film cooling effectiveness.

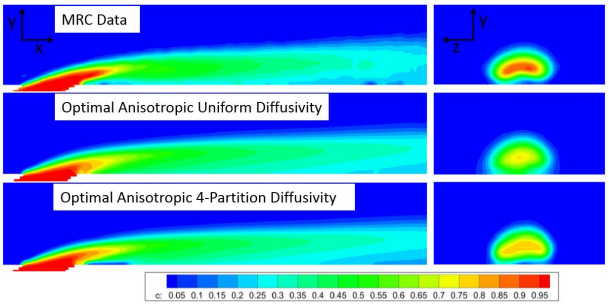

Figure 7: Contour plots of mean concentration in the baseline geometry. The left panels show center planes, and the right panels show streamwise planes $2D$ downstream of the center of the hole. Note that the optimal 4-partition results matches the MRC data significantly better.

Across the 5 datasets in which this optimization was performed, some patterns became evident in the optimal diffusivities yielded by the optimization algorithm. Most noteworthy is the diffusivity in partition 1: in all cases, the optimal values of $\alpha_{t,y}$ were slightly negative, while $\alpha_{t,z}$ was a moderately high positive value. This shows that anisotropy is extremely important in that region, and the only way to obtain the correct mean concentration is to shut down any turbulent transport in the wall-normal direction (or even add some counter-gradient transport if numerically stable). Almost all other diffusivities were positive, and no patterns were apparent in diffusities in partitions 2 and 3. In partition 4, which represents most of the flow, the optimal diffusivities were similar across the 5 cases, and significantly anisotropic (with wall-normal diffusivities being about half of the spanwise diffusivities in all cases). Table 1 contains the optimal diffusivities in all the 5 experimental cases.

Machine Learning

The 4-partition optimization work showed us that default mixing models give poor mean concentration predictions even when the correct mean velocity field is used. Furthermore, enhancing these mixing models can generate noticeable qualitative and quantitative improvements on the resulting scalar concentration for a jet-in-crossflow, as demonstrated by Fig. 7. However, if the objective is to propose an alternative model that can improve results on a general film cooling geometry, this approach is insufficient. Ideally, we would desire a model that does not require the user to select specific partitions and that contains a set of parameters that can be tuned once and then applied to different geometries (rather than optimized for each specific case given experimental data). That is where a machine learning approach becomes useful.

Machine learning (ML) consists of a broad class of algorithms to process large amounts of data and extract patterns from the data. In particular, we are interested in supervised learning algorithms for regression, which try to build a complex and generalizable mapping which is then used to predict a real variable

Table 1: Optimal diffusivities in each of the 5 cases under the three optimization regimes. All reported diffusivities are non-dimensional, normalized by $V_{jet}D$. APG stands for "Adverse Pressure Gradient", FPG stands for "Favorable Pressure Gradient", and BL stands for the "Boundary Layer" cases.

		Baseline		APG		FPG		ThickBL		ThinBL	
		$\alpha_{t,y}$	$\alpha_{t,z}$	$\alpha_{t,y}$	$\alpha_{t,z}$	$\alpha_{t,y}$	$\alpha_{t,z}$	$\alpha_{t,y}$	$\alpha_{t,z}$	$\alpha_{t,y}$	$\alpha_{t,z}$
Uniform, isotropic		0.01		0.012		0.006		0.0125		0.009	
Uniform, anisotropic		0.009	0.0105	0.014	0.011	0.0045	0.008	0.0185	0.01	0.0095	0.0085
4-partition, anistropic	**Partition 1**	-0.003	0.0127	-0.003	0.0167	-0.003	0.0124	-0.003	0.0183	-0.003	0.0198
	Partition 2	0.0015	-0.003	0.0108	0.0021	-0.003	0.0222	0.0322	0.0009	0.0055	0.0039
	Partition 3	0.0161	0.0107	0.0255	0.001	0.0076	-0.0009	0.0256	0.0045	0.0216	0.0021
	Partition 4	0.0073	0.0136	0.0067	0.0169	0.006	0.0088	0.0079	0.0154	0.0063	0.0116

of interest (or label) given other quantities that it should depend on (or features). To build this mapping, one needs training data, with several examples where the label and the features are known. Linear regression is the simplest example of these types of algorithms, and random forests (RF) and neural networks are more complex, state-of-the-art ones. Bishop [2] presents a good summary of the subject.

Machine learning applied for turbulence modeling is a new, but growing field. Outside of our research group, active authors in the field include Dr. Ling (e.g. [15]), Prof. Duraisamy at University of Michigan (e.g. [6]), Prof. Sandberg at University of Melbourne (e.g. [24]), and Prof. Xiao at Virginia Tech (e.g.[27]). All have slightly different approaches. The current section is not meant to be a comprehensive review, but instead introduce our group's approaches, and summarize past results and future directions.

The work of Milani et al. [19] introduced our machine learning approach to improve turbulent scalar modeling and showed great promise. We fixed the isotropic gradient diffusion hypothesis of Eq. 2 as the model form, but did not use the Reynolds analogy. Instead, we used a machine learning algorithm to determine a model for α_t as a function of local variables available to a RANS solver. These local variables (the features) were based on the mean velocity gradient $\partial\bar{u}_j/\partial x_i$, the mean scalar gradient $\partial\bar{c}/\partial x_i$, the eddy viscosity ν_t, and the distance to the nearest wall d. They were properly non-dimensionalized based on local turbulent quantities (k and ε, calculated by the $k-\varepsilon$ model), and the tensors were converted to Galilean invariant bases according to the formulation proposed by Ling et al. [14]. A full list of the 19 features used can be found in Milani et al. [19].

A random forest machine learning algorithm was used due to its ease of implementation, insensitivity to tunable hyperparemeters, and robustness to noise and outliers. For more details on this specific algorithm, the reader is directed to Louppe's comprehensive guide [17].

To train the random forest, we used high-fidelity simulation data (either highly-resolved LES or DNS) of simple turbulent flows. We did so because the training step of a supervised learning model requires the "true" field for α_t, and thus requires turbulent statistics (which our MRI data cannot provide). However, we consider our simulations as means to enrich the MRI datasets with turbulence information, because we go to great lengths to ensure that the same problem is being simulated. We tune the inlet conditions to produce correct velocity and turbulence profiles and thoroughly validate the mean field data against MRI results. Two highly-resolved LES simulations of distinct jet-in-crossflow configurations are further described in Bodart et al. [3] and Folkersma et al. [9].

When the machine learning model was trained on two of the datasets (a skewed jet-in-crossflow and a wall-mounted cube in crossflow) and applied to a third one, the baseline jet-in-crossflow, it was able to significantly improve the scalar concentration results, particularly close to the wall. Figure 8 is taken from Milani et al. [19] and it shows the level of improvement obtained for near-wall concentration.

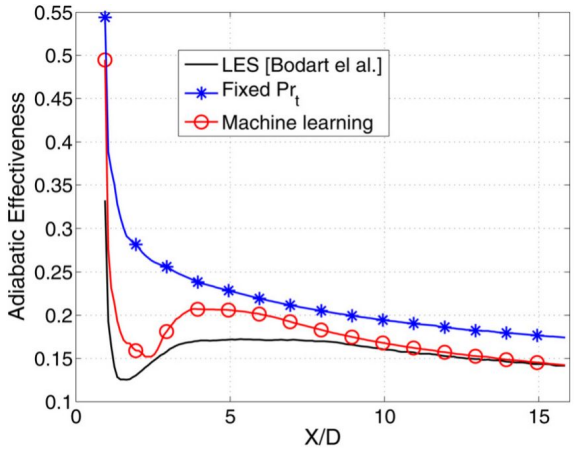

Figure 8: Plot taken from Milani et al. [19] showing spanwise-averaged adiabatic effectiveness (i.e., mean scalar concentration at the wall) in the baseline jet-in-crossflow geometry. The LES curve matches the experimental data for this case.

Another question that we sought to answer regards the interpretability of physics-based ML models. High-performing machine learning algorithms, such as neural networks and random forests, are notorious for making "black-box" decisions: they predict accurately when everything goes well, but it is difficult to explain why it made a prediction. This might be acceptable in image classification or ad targeting applications, but in physical applications (where fundamental laws exist and experts have physical intuition) it is not. So, Milani et al. [18] sought to explain random forest predictions by proposing a spatially-resolved metric, called the Pointwise Feature Usage (PFU), to determine how important each feature is in predicting the turbulent diffusivity. Their main conclusions is that most of the decisions can be explained by the eddy viscosity ν_t and dis-

tance from the wall d; but, critical regions of the flow require non-linear closures.

Our most recent work on scalar flux modeling showed that the dramatic improvement provided by ML models cannot always be repeated across different datasets and different strategies. The results shown in Fig. 8 give us the confidence that ML methods are promising, but that remains an open research path for us. Next steps to improve models include understanding when the amount of training data is enough (and adding more when it is not), and also deploying more complex models in conjunction with the machine learning approach (e.g. an anisotropic diffusivity model). Furthermore, research into improving models for the turbulent momentum flux is also on the agenda for the lab. We have discovered in soon-to-be-published work that simple modifications to the turbulent $k-\varepsilon$ equations coupled with ML techniques yield improved mean velocity predictions in separated turbulent flows. Using such techniques in film cooling simulations together with the model of Milani et al. [19] could further improve mean scalar predictions, since it would add a more accurate velocity field to the more accurate turbulent mixing model.

Conclusions

In the current paper, we presented an overview of the research conducted in our lab in the past 10 years in the area of turbulent mixing models. Thanks to the MRI experimental techniques, we have datasets that are not easily available in the community: 3D mean data on complex internal geometries with reasonably fast turnaround. The optimization and machine learning paths show how we leveraged that experimental capability (which, at this point, is as close to "big data" as experimental fluid mechanics gets) to study turbulent mixing models specifically for RANS simulations.

Most of the results presented here have already been presented in previous publications. However, the application of genetic algorithms to optimize the 8 parameters of the anisotropic, 4-partition model are being presented here for the first time. That was an important stepping stone in the progression of this research because it clearly showed that improving the mixing model is key to obtaining better results in RANS simulations, while at the same time pointing at glaring weaknesses of the Reynolds analogy, the typical closure in commercial RANS solvers today. The ongoing machine learning research, combined with high-fidelity simulations to enrich the MRI data, allows us to generalize those ideas and apply them to arbitrary geometries with promising results.

Acknowledgments

Work on the development and application of MRI measurement techniques to study film cooling was supported by Honeywell Aerospace and GE Aircraft Engines. Research in machine learning for scalar transport modeling has been supported by Sandia National Labs and Honeywell Aerospace. We have collaborated on the works described in this paper with Dr. Chris Elkins, Col. Michael Benson (USMA), Dr. Julia Ling (Citrine Informatics), Dr. Andrew Banko, Dr. Emin Issakhanian, Dr. Kevin Ryan, Dr. Daniel Borup, Dr. Julien Bodart (ISAE-Supaero), Mr. Ian Gunady, Mr. David Ching, Dr. Filippo Coletti, and Dr. Sayuri Yapa.

References

[1] Benson, M. J., Elkins, C. J., Mobley, P. D., Alley, M. T. and Eaton, J. K., Three-Dimensional Concentration Field Measurements in a Mixing Layer Using Magnetic Resonance Imaging, *Experiments in Fluids*, **49**, 2010, 43–55.

[2] Bishop, C., *Pattern Recognition and Machine Learning*, Springer, New York, NY, 2006.

[3] Bodart, J., Coletti, F., Bermejo-Moreno, I. and Eaton, J., High-Fidelity Simulation of a Turbulent Inclined Jet in a Crossflow, *Center for Turbulence Research Annual Research Briefs*, 263–275.

[4] Coletti, F., Benson, M., Ling, J., Elkins, C. and Eaton, J., Turbulent Transport in an Inclined Jet in Crossflow, *International Journal of Heat and Fluid Flow*, **43**, 2013, 149–160.

[5] Coletti, F., Elkins, C. J. and Eaton, J. K., An Inclined Jet in Crossflow Under the Effect of Streamwise Pressure Gradients, *Experiments in Fluids*, **54**, 2013, 1589.

[6] Duraisamy, K., Zhang, Z. J. and Singh, A. P., New Approaches in Turbulence and Transition Modeling Using Data-Driven Techniques, 1284.

[7] Elbeltagi, E., Hegazy, T. and Grierson, D., Comparison Among Five Evolutionary-based Optimization Algorithms, *Advanced engineering informatics*, **19**, 2005, 43–53.

[8] Elkins, C., Markl, M., Pelc, N. and Eaton, J., 4D Magnetic Resonance Velocimetry for Mean Velocity Measurements in Complex Turbulent Flows, *Experiments in Fluids*, **34**, 2003, 494–503.

[9] Folkersma, M. and Bodart, J., Large Eddy Simulation of an Asymmetric Jet in Crossflow, in *Direct and Large-Eddy Simulation X*, Springer International Publishing, Basel, Switzerland, 2018, 85–91, 85–91.

[10] Hoda, A. and Acharya, S., Predictions of a Film Coolant Jet in Crossflow With Different Turbulence Models, *Journal of Turbomachinery*, **122**, 1999, 558–569.

[11] Kays, W. M., Turbulent Prandtl Number - Where Are We?, *Journal of Heat Transfer*, **116**, 1994, 284–295.

[12] Ling, J., Coletti, F., Yapa, S. D. and Eaton, J. K., Experimentally Informed Optimization of Turbulent Diffusivity for a Discrete Hole Film Cooling Geometry, *International Journal of Heat and Fluid Flow*, **44**, 2013, 348–357.

[13] Ling, J., Elkins, C. J. and Eaton, J. K., Optimal Turbulent Schmidt Number for RANS Modeling of Trailing Edge Slot Film Cooling, *Journal of Engineering for Gas Turbines and Power*, **137**, 2015, 072605.

[14] Ling, J., Jones, R. and Templeton, J., Machine Learning Strategies for Systems with Invariance Properties, *Journal of Computational Physics*, **318**, 2016, 22–35.

[15] Ling, J., Kurzawski, A. and Templeton, J., Reynolds Averaged Turbulence Modelling Using Deep Neural Networks with Embedded Invariance, *Journal of Fluid Mechanics*, **807**, 2016, 155–166.

[16] Ling, J., Rossi, R. and Eaton, J. K., Near Wall Modeling for Trailing Edge Slot Film Cooling, *Journal of Fluids Engineering*, **137**, 2015, 021103.

[17] Louppe, G., Understanding Random Forests: From Theory to Practice, *arXiv preprint*.

[18] Milani, P. M., Ling, J. and Eaton, J. K., Physical Interpretation of Machine Learning Models Applied to Film Cooling Flows, in *ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition*, American Society of Mechanical Engineers, 2018, V05AT12A019–V05AT12A019, V05AT12A019–V05AT12A019.

[19] Milani, P. M., Ling, J., Saez-Mischlich, G., Bodart, J. and Eaton, J. K., A Machine Learning Approach for Determining the Turbulent Diffusivity in Film Cooling Flows, *Journal of Turbomachinery*, **140**, 2018, 021006.

[20] Nikparto, A., Rice, T. and Schobeiri, M. T., Experimental And Numerical Investigation of Heat Transfer and Film Cooling Effectiveness of a Highly Loaded Turbine Blade Under Steady and Unsteady Wake Flow Condition, in *ASME Turbo Expo 2017: Turbomachinery technical conference and exposition*, American Society of Mechanical Engineers, 2017, V05CT19A029–V05CT19A029, V05CT19A029–V05CT19A029.

[21] Pelc, N. J., Sommer, F., Li, K., Brosnan, T. J., Herfkens, R. J. and Enzmann, D. R., Quantitative Magnetic Resonance Flow Imaging, *Magnetic resonance quarterly*, **10**, 1994, 125–147.

[22] Ryan, K. J., Bodart, J., Folkersma, M., Elkins, C. J. and Eaton, J. K., Turbulent Scalar Mixing in a Skewed Jet in Crossflow: Experiments and Modeling, *Flow, Turbulence and Combustion*, **98**, 2017, 781–801.

[23] Ryan, K. J., Coletti, F., Elkins, C. J. and Eaton, J. K., Building Block Experiments in Discrete Hole Film Cooling, in *ASME Turbo Expo 2015: Turbine Technical Conference and Exposition*, American Society of Mechanical Engineers, 2015, V05BT12A042–V05BT12A042, V05BT12A042–V05BT12A042.

[24] Sandberg, R., Tan, R., Weatheritt, J., Ooi, A., Haghiri, A., Michelassi, V. and Laskowski, G., Applying Machine Learnt Explicit Algebraic Stress and Scalar Flux Models to a fundamental Trailing Edge Slot, *Journal of Turbomachinery*, **140**, 2018, 101008.

[25] Slotnick, J., Khodadoust, A., Alonso, J., Darmofal, D., Gropp, W., Lurie, E. and Mavriplis, D., CFD Vision 2030 Study: a Path to Revolutionary Computational Aerosciences, *NASA Technical Report*.

[26] So, R. M. and Speziale, C., A Review of Turbulent Heat Transfer Modeling, *Annual Review of Heat Transfer*, **10**.

[27] Wang, Jian-Xun and Wu, Jinlong and Ling, Julia and Iaccarino, Gianluca and Xiao, Heng, A Comprehensive Physics-Informed Machine Learning Framework for Predictive Turbulence Modeling, *arXiv preprint arXiv:1701.07102*.