11th International Symposium on Turbulence and Shear Flow Phenomena (TSFP11)
Southampton, UK, July 30 to August 2, 2019

VELOCITY ESTIMATION IN THE MIXING LAYER OF A SUBSONIC JET USING ARTIFICIAL NEURAL NETWORKS

Andrew S. Tenney
Department of Mechanical
and Aerospace Engineering
Syracuse University
Syracuse, NY 13210
astenney@syr.edu

Mark N. Glauser
Department of Mechanical
and Aerospace Engineering
Syracuse University
Syracuse, NY 13210
mglauser@syr.edu

Zachary P. Berger
Applied Research Laboratory
The Pennsylvania State University
State College, PA 16804
zpb1@psu.edu

ABSTRACT

Time-resolved Particle Image Velocimetry (PIV, 10 kHz) was used to measure the flow-field of a Mach 0.6 axisymmetric jet. The field was decomposed into its requisite spatial POD eigenfunctions, and the time dependent coefficients were recovered. An Artificial Neural Network (ANN) and a Linear Stochastic Estimation (LSE) model were then trained to estimate the first five time-dependent POD coefficients from five point velocity measurements made by "virtual crosswires" in the mixing layer. We show that the prediction accuracy is strongly dependent on the POD mode number for both models. On average, the ANN-based model is able to predict the velocity fluctuations more accurately than the LSE-based model. Finally, we examine the estimated reduced-order velocity fields and their correlation to analytically reconstructed reduced-order velocity fields. Possible extensions of this method are also discussed.

INTRODUCTION

LSE has been used as a tool for estimating conditional velocity fields in turbulent flows since its introduction by Adrian (1977). This method has proven its usefulness in modeling relationships between measured quantities in turbulent flows, both in its simplicity and overall accuracy. Several powerful extensions of this method have also been proposed and implemented. Bonnet *et al.* (1994) detailed one such extension that takes advantage of the low-dimensional capabilities of Proper Orthogonal Decomposition (*POD*) and the estimation abilities of LSE to identify structure in turbulent shear flows. This method, called the POD/LSE Complementary Technique, was later modified to be able to predict the time-dependent behavior of the POD modes given spatially sparse measurements. This modified version was shown by Pinier *et al.* (2007) and Picard & Delville (2000) to be useful in informing flow control schemes and modeling structure in turbulent flows, respectively.

In this document, we re-examine Stochastic Estimation, and formulate it to take advantage of the predictive power of Artificial Neural Networks (*ANNs*). This formulation is then integrated into the Modified POD/LSE Complementary Technique, where it is used to estimate the fluctuating velocity field in the mixing layer of a Mach 0.6 axisymmetric jet. We also offer a quantitative comparison to the traditional LSE-based variation to benchmark its performance.

EXPERIMENTS

The data used in this study was collected in the anechoic chamber at the Skytop Turbulence Lab at Syracuse University by Berger *et al.* (2015). The nozzle used in the experiments is converging only with an exit diameter of 50.8 mm. Exact specifications can be found in Tinney *et al.* (2004). The baseline operating condition of $M_j = 0.6$ corresponds to a dimensional velocity of 206 m/s, and a Reynolds number based on the nozzle diameter, Re_j of approximately 693,000.

In the near-field, *Kulite* XCE-093-5G series transducers sampled pressure at 14 locations, at a rate of 40.96 kHz. Two-component velocity was captured using a time-resolved 10 kHz PIV (*TRPIV*) setup. The laser sheet was positioned to illuminate the flow in a stream-wise oriented plane centered around a location approximately 4.75 D_h (11.5 in.) downstream of the nozzle exit plane. A single *Photron* FASTCAM camera was used to image the flow field and the image processing and vector calculation was performed in *LaVision's* DaVis software package. A photograph of the near-field can be seen in Figure 1. The TRPIV database consists of three sets of 45,000 two-component PIV snapshots collected over three 15 second experimental runs. Only the first 15,000 snapshots were used in this study. This corresponds to Case 6 in Berger's database, and will be referred to as such in future sections.

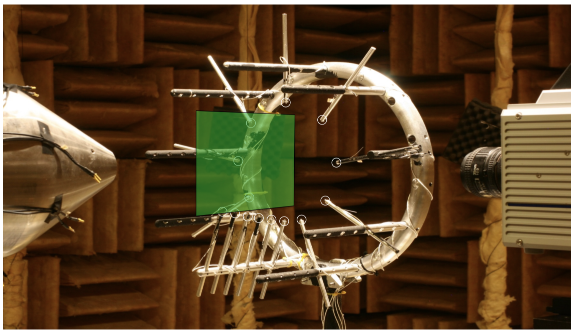

Figure 1. Experimental setup for 2013 TRPIV experiments. Location of *Kulite* transducers circled in white. Approximate location of TRPIV window indicated by green box. *Adapted from Berger (2014).*

METHODS

Stochastic Estimation

Stochastic estimation as proposed by Adrian (1977, 1979) relies on the idea of conditional averages. He hypothesized that conditional averaging can be utilized to make the best mean-square estimate the flow state at a time t and location $\mathbf{x}'$ given the flow state at a separate location, $x = x' + \delta x$. More succinctly, the estimate of conditional events at $\mathbf{x}'$ are a function of the unconditional events at location x, and the distance between the two locations (Equation 1).

$$\tilde{u}(\mathbf{x}',t) = f\{u(\mathbf{x},t), \delta\mathbf{x}\}. \tag{1}$$

This representation of the conditional events $\tilde{u}$ can be expanded using a Taylor series, as in Equation 2. Terms of order two and greater are neglected for simplicity, thus ensuring a linear estimate.

$$\tilde{u}_i(\mathbf{x}',t) = \sum_{j=1}^{N_c}\sum_{k=1}^{N_s} b_{ij}(\mathbf{x}',\mathbf{x}_k)u_j(\mathbf{x}_k,t) + O[u_j^2]. \tag{2}$$

The index i represents the i^{th} component of a vector valued flow state variable, N_s is the number of spatial points x where the unconditional events are measured, and N_c is the number of components of a vector valued unconditional event. This equation can be further simplified into a more generic form to collapse the double summation, as in Delville (1994).

$$\tilde{v}(\mathbf{x}',t) = \sum_{j=1}^{N_m} h_j(\mathbf{x}')g_j(t) + O[g_j^2]. \tag{3}$$

In Equation 3, the summation is now over the product of the total number of components and spatial measurements. This equation is the final, most simplified form of the conditional approximation, $\tilde{v}(\mathbf{x},t)$, based on unconditional events, $g_j(t)$. To produce an optimal estimate, the error between this approximation and the actual measured value of the conditional events must be minimized. This results in an over-determined system of equations for which an optimal least squares solution can be found using Equation 4.

$$\mathbf{H} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{V} \tag{4}$$

The resulting coefficient matrix $\mathbf{H}$ can be used to estimate conditional events $\tilde{v}(\mathbf{x},t)$ given measured unconditional events $g_j(t)$. Interesting features of the resulting estimate and some consequences of this formulation have been discussed by Adrian & Moin (1988), Tinney *et al.* (2006), and Delville (1994).

Modified POD/LSE Complementary Technique

The POD/LSE Complementary Technique was proposed by Bonnet *et al.* (1994) to take advantage of the strengths of POD and LSE. This method has been shown to be extremely useful in studying the low-dimensional behavior of turbulent mixing layers (Bonnet *et al.* (1998)). In the Modified Complementary Technique, LSE is used to estimate the time dependent POD coefficients ($a_i(t)$ in Equation 5) directly. These coefficients can then be combined with the spatial eigenfunctions ($\phi_i(\vec{x}$ in Equation 5) to reconstruct a reduced order representation of the flow field. In this formulation, some observable in the flow, such as pressure or velocity at one or several points, is used to define the unconditional events in Equation 6. The time-dependent POD coefficients serve as the conditional events. The LSE model is then trained to estimate a subset of the coefficients directly from these measurements.

$$\tilde{u}(\vec{x},t) = \sum_{i=1}^{N} a_i(t)\phi_i(\vec{x}) \tag{5}$$

$$\tilde{a}_n(t) = f\{u(\mathbf{x},t), \delta\mathbf{x}\}. \tag{6}$$

The time-dependent POD coefficients $a_i(t)$ are made up of contributions from each individual velocity component, u_i. When applying the Modified Complementary Technique, the LSE model can be trained to either estimate the overall time-dependent coefficients, or each one of these individual contributions. In the current study, we present results corresponding to estimation of the full coefficients only. This is called *Overall* estimation in future sections.

Artificial Neural Networks

A neural network consists of artificial neurons connected to one another with corresponding weights and biases and are commonly referred to as perceptrons. Perceptrons with two or more hidden layers between the input and output nodes are often classified as DNNs. The power of these networks to approximate complex non-linear functions lies in the Universal Approximation Theorem (*UAT*). It can be shown that on compact subsets of $\mathbb{R}^n$, a perceptron with only a single hidden layer and finite number of parameters can be used to approximate any continuous function to arbitrary precision as shown by Cybenko (1989). For very

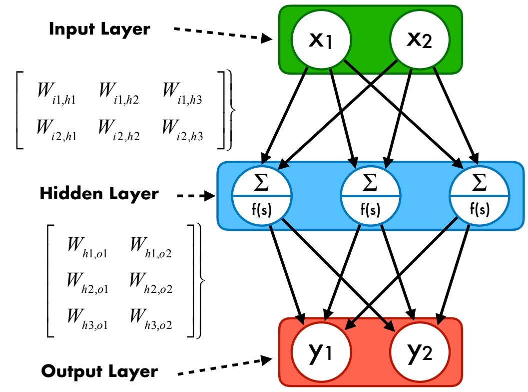

Figure 2. Schematic of artificial neural network with a single hidden layer. Associated symbols and matrices given in Equations 7 - 9.

complex problems, however, single layer networks quickly become unwieldy. Deepening the network increases its efficiency as compared to a single layer perceptron, and often allows for fewer overall nodes.

A fully-connected neural network is one in which each node in each respective layer is connected to every node in the layers both before and after it (sketched in Figure 2). A weight is associated with each of these connections, and a bias node can be included to account for any constant offsets in the data. At each node in a given layer, two mathematical operations occur: a linear combination of the outputs of the previous layer weighted by the corresponding inter-nodal weight matrix, and a non-linear mapping of the result of this combination onto a subset of $\mathbb{R}$. A generalized input vector and weight matrix are given by Equations 7 and 8, the linear combination equation is given by Equation 9, and an example nonlinear activation function is presented in Equation 10.

$$X = [x_1, x_2, \dots, x_m]_{1\times m} \tag{7}$$

$$W = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{bmatrix}_{m\times n} \tag{8}$$

$$\Sigma : s = XW + b \tag{9}$$

$$f(s) = \frac{1}{1+e^{-s}} \tag{10}$$

To train a network, a backpropagation algorithm is used. The inter-nodal weights and biases are randomly initialized, and the algorithm feeds the input data into the network. The output of the network is compared to the desired output, and the error (typically mean squared error in regression problems) is calculated. The gradient of the cost function with respect to each individual inter-nodal weight and bias can be calculated, and in training, the weights and biases can be nudged towards an optimum where the error is minimized, as noted by Géron (2017). Computation time can be reduced greatly if the gradients can be calculated explicitly. For this reason, the activation functions are easily differentiable.

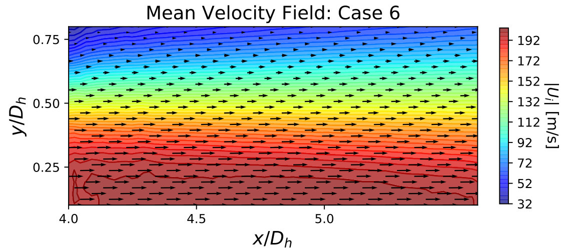

Figure 3. Mean velocity field in the upper mixing layer of Mach 0.6 axisymmetric jet.

RESULTS

Velocity Field

The PIV window spans the full width of the jet in a plane of symmetry between 4 and 5.75 nozzle diameters from the nozzle exit plane. This window is upstream of the point of potential core collapse. To simplify the estimation problem and make comparisons to other work, we focused only on the mixing layer on the upper side of the jet from y/D = 0.1 - 0.8. The mean velocity field in this region is illustrated in Figure 3. The contour plot is colored by the magnitude of the measured two-component velocity vector and overlaid with a subset of the average vectors.

The turbulent velocity fluctuations in this region are decomposed using POD. Only the first five modes were utilized in this study, accounting for approximately 32% of the total turbulent kinetic energy in the flow. The POD reconstruction was truncated at five modes due to the diminished accuracy of the ANN and LSE models in predicting the time-dependent POD coefficients associated with higher order modes (illustrated in the following section). Additionally, since mode pairing is observed in these leading modes, they can be used to qualitatively reconstruct large scale structures convecting within the mixing layer as noted by Taira *et al.* (2017).

"Virtual crosswire anemometers" measure the two-component velocity vectors at five stream-wise locations. These measurements were used as the unconditional events. The virtual anemometers were spaced evenly between $x/D = 4.25$ and 5.25, and are placed at the mean cross-stream location corresponding to peak TKE production due to cross-stream velocity gradients (Equation 11). This average cross stream location corresponds to $y/D = 0.575$. Finally, the signals at these locations were were band-pass filtered to improve model accuracy and consistency. They

Table 1. Velocity Prediction Network Hyperparameters

Nodes	512
Train/Test Split	33/67
Batch Size	200
Training Epochs	150
Hidden Activations	ELU
Kernel Initializer	Random Normal
Objective	Mean Squared Error
Optimizer	Adam
Learning Rate	0.001

were filtered between 100 and 4000 Hz, effectively removing the lowest and highest frequency components. This was an empirically driven decision, as it produced the best estimation results. The final ANN architecture is detailed in Table .

$$S = -\overline{u_i'v_j'}\frac{\partial \overline{U_i}}{\partial x_j} \tag{11}$$

Field Estimation

Figure 4 illustrates the predictions made by the LSE and ANN based formulations compared to the expected values. Only modes one and five are presented here, as they are illustrative of the upper and lower limits of this method's accuracy. The left column of Figure 4 compares the time-domain predictions, while the right column compares their power spectra. The predictions made by both the LSE and ANN models are very similar when compared in the time domain. In all cases, the predicted time series lie nearly on top of one another. In modes one through three, the large- and small-scale oscillations of the target signal were estimated accurately from a qualitative perspective. The amplitude of the predicted time-dependent coefficients decreases greatly with increasing mode number. This is observed most prominently in mode five, but is observed clearly after mode three. Some oscillations in the time dependent coefficient of mode five were predicted reasonably well (small peak near $t = 101ms$). In modes four and five, both methods under-predict the time-dependent coefficient by nearly one standard deviation across the entirety of the signal.

In the frequency domain, the discrepancies between the measured and predicted signals are also clear. The spectral trends were correctly predicted in general, however, there are clear amplitude discrepancies across all modes and all frequencies. In mode one (Figure 4, top right) the low frequencies of the estimated coefficient were under-predicted, while the mid-frequencies were slightly over-predicted. Mode five coefficients were severely under-predicted over the entire spectrum. The ANN based model provides a slightly more accurate estimate than the LSE model, however, the amplitude at each frequency is under-predicted by approximately one order of magnitude. The trend of deceasing accuracy with increasing mode number is also observed in the modes not shown in this figure.

Figure 5 illustrates the ratio of mean square error (MSE) between the estimated and expected time dependent coefficients to quantify their relative performance. The symbols are scaled by the inverse of the multiple correlation value between the model inputs and outputs to illustrate the degree of non-linearity in their relationship. For all modes estimated by both methods, the ratio of MSE_{ANN} to MSE_{LSE} is below one. This indicates that the ANN model outperforms the LSE model. The ANN model exhibits a 2% decrease in error when estimating the time-dependent coefficients of the first mode when compared to the LSE model. On average, the error reduction is less than 1.5% over all modes. In addition, there is not a clear relationship in the degree of linearity between the variables and the relative performance of the LSE and ANN models.

Figure 6 illustrates the reduced-order approximation of the velocity field using the first five modes and the coefficients analytically calculated time-dependent POD coefficients. This is used as a point of comparison. This reconstructed field is referred to as the baseline field going forward. The results from the ANN and LSE models are illustrated in Figure 7. All reconstructed snapshots are spaced 0.1 ms apart in time to illustrate the temporal evolution of the flow field. The baseline field exhibits alternating patches of positive and negative velocity that convect downstream. The regions of positive velocity in both the stream-wise and cross-stream components appear to stretch as they convect downstream. By identifying the center of these large positive and negative regions and tracking their motion in the stream-wise direction, we observed that they travel at approximately 85 m/s on average. This corresponds to the average velocity in the mixing layer. This convective velocity is also observed in the estimated fields.

The reconstructed fields in Figure 7 exhibit structure that is qualitatively similar to that observed in Figure 6. The alternating regions of positive and negative velocity are of similar size shape to the baseline reconstruction. They were also observed in nearly identical locations at each instant in time. The convection and stretching of the positive velocity region in this sequence was faithfully reproduced as well. While the spatial structure of the reconstructions is similar, there was a clear magnitude discrepancy between the baseline and estimated fields. Both reconstructions using the estimated coefficients have similar magnitude, however. The amplitude discrepancy between the estimated and calculated reconstructions follows directly from the difference observed in the time-dependent coefficients, particularly in the higher mode numbers.

Spatial correlations between the estimated and analytically reconstructed fields are shown in Figure 8. The ANN and LSE-based methods correlate best to the analytically reconstructed field between $y/D = 0.5$ and 0.75 for both the stream wise and cross-stream components. The largest discrepancies are observed closer to the position of the core of the jet, where the uncertainty in the data is highest. While the ANN-based method performs slightly better than the LSE-based method as evidenced by Figure 5, the differences between the estimated fields do not seem to be spatially coherent.

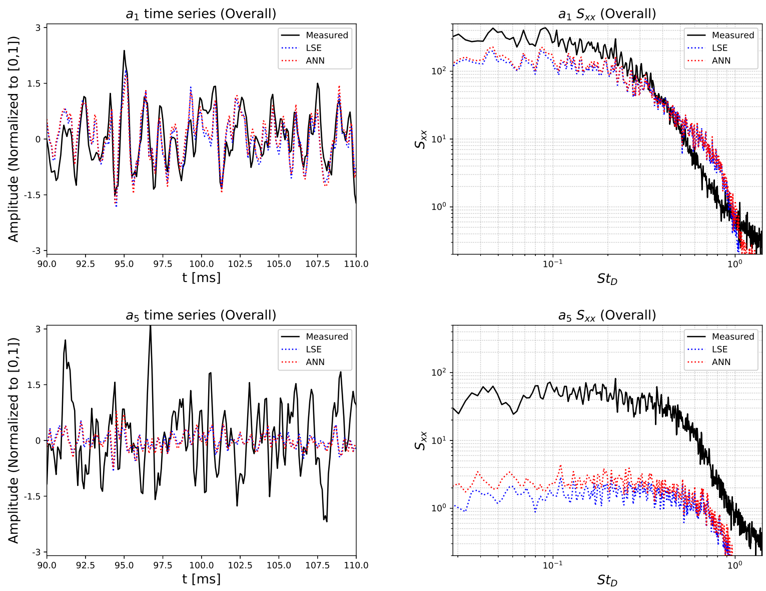

Figure 4. Comparison of LSE and ANN model predictions

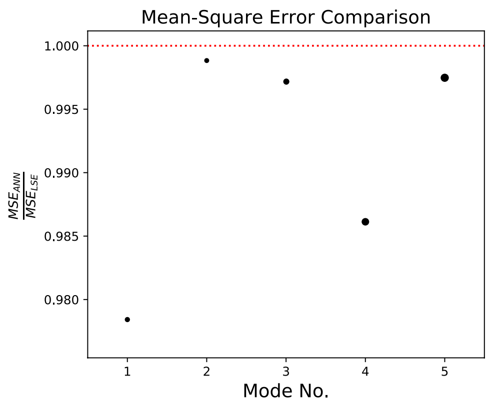

Figure 5. Comparison of mean squared error between predictions from LSE and ANN models.

DISCUSSION

Stochastic estimation can be reformulated to take advantage of machine learning. In this analysis, we demonstrated that a velocity field can be estimated using velocity measured at spatially sparse locations using a neural network-based version of the Modified Complementary Technique. This method was also compared to the Modified Complementary Technique implemented using a traditional LSE model. The ANN-based method exhibits a slight improvement over the traditionally formulated LSE-

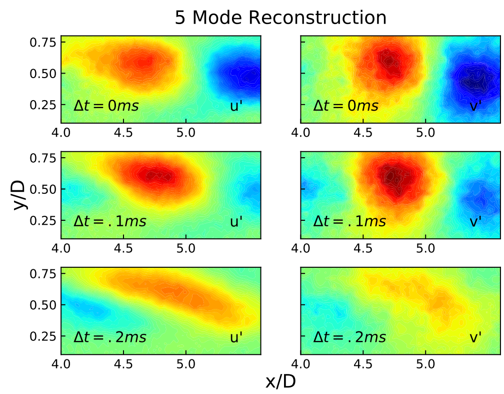

Figure 6. Five mode reconstruction of the fluctuating velocity field. Time increases from top to bottom. *Left:* Stream-wise velocity component. *Right:* Cross-stream velocity component.

based method. The improvement is on the order of 1.5% and is consistent with previous findings of Tenney *et al.* (2019). We also observed that the accuracy of the estimates produced by both formulations declines with increasing POD mode number. This suggests that more sophisticated methods may be needed to estimate the higher-order coefficients, or that the time-dependent coefficients associated with these modes are simply non-deterministic in nature. Finally, the reconstructed low-dimensional velocity fields produced by both estimation methods are qualitatively similar to the analytically calculated reduced-order fields. The neural network-based method estimates the magnitude of the fields slightly better than the LSE based method.

Several other aspects of this method should be studied in the future. It has been shown that by including higher order terms in the Stochastic Estimation formulation (such as using a quadratic model), slight improvement in prediction accuracy are shown. A comparison between the ANN-

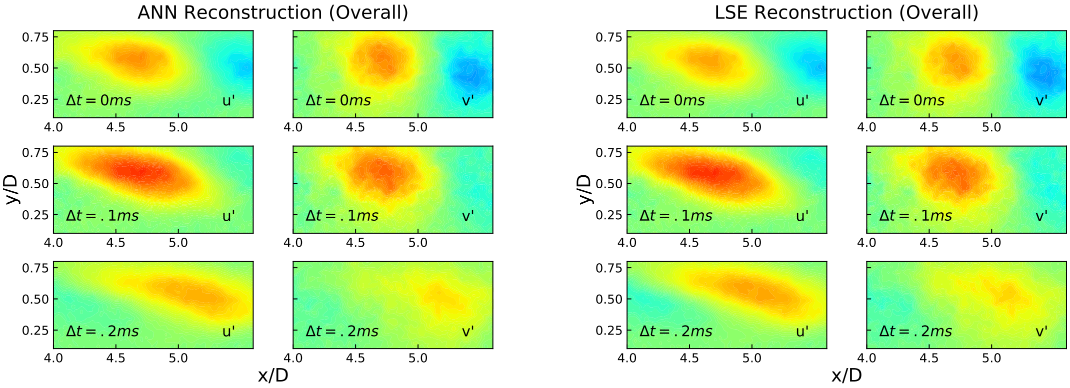

Figure 7. Five mode reconstruction of the fluctuating velocity field using the time-dependent coefficients estimated with the ANN and LSE-based models. Time increases from top to bottom. *Left subfigure:* ANN-based model. *Right subfigure:* ANN-based model.

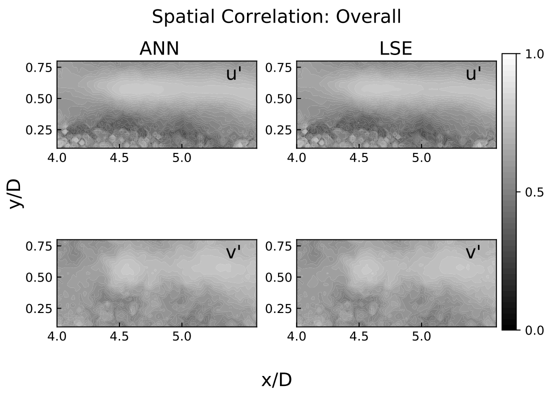

Figure 8. Spatial correlation between the estimated and analytically reconstructed fields. *Left:* ANN, *Right:* LSE, *Top:* stream-wise component, *Bottom:* Cross-stream component.

based method and the higher-order Stochastic Estimation model should be performed. Practically, the ANN based methods can also be applied to control problems, such as those studied by Pinier *et al.* (2007).

REFERENCES

Adrian, R.J. 1977 On the role of conditional averages in turbulence theory. *Symposium on Turbulence in Liquids* .

Adrian, R.J. 1979 Conditional eddies in isotropic turbulence. *Physics of Fluids* **22** (11), 2065–2070.

Adrian, R.J. & Moin, P. 1988 Stochastic estimation of organized turbulent structure: Homogeneous shear flow. *J. Fluid Mechanics* **190**, 531–559.

Berger, Z.P. 2014 The effects of active flow control on high-speed jet flow physics and noise. PhD thesis, Syracuse University.

Berger, Z.P., Shea, P.R., Berry, M.G., Noack, B.R., Gogineni, S. & Glauser, M.N. 2015 Active flow control for high speed jets with large window piv. *Flow, Turbulence, and Combustion* **94**, 97–123.

Bonnet, J.P., Cole, D.R., Delville, J., Glauser, M.N. & Ukeiley, L.S. 1994 Stochastic estimation and proper orthogonal decomposition: Complimentary techniques for identifying structure. *Experiments in Fluids* **14**, 307–314.

Bonnet, J.P., Delville, J., Glauser, M. N., Antonia, R. A., Bisset, D. K., Cole, D. R., Fiedler, H. E., Garem, J. H., Hilberg, D., Jeong, J., Kevlahan, N. K. R., Ukeiley, L. S. & Vincendeau, E. 1998 Collaborative testing of eddy structure identification methods in free turbulent shear flows. *Experiments in Fluids* **25**, 197–225.

Cybenko, G. 1989 Approximations by superpositions of sigmoidal functions. *Mathematics of Control, Signals, and Systems* **2**, 303–314.

Delville, J. 1994 La decomposition orthogonale aux valeurs propres et l'analyse de l'organisation tridimensionelle des ecoulements turbulents cisailles libres. PhD thesis, l'Universite de Poitiers.

Géron, A. 2017 *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media.

Picard, C. & Delville, J. 2000 Pressure velocity coupling in a round jet. *International Journal of Heat and Fluid Flow* **21**, 359–364.

Pinier, J., Ausseur, J., Glauser, M.N. & Higuchi, H. 2007 Proportial closed-loop feedback control of flow separation. *AIAAJ* **45**.

Taira, Kunihiko, Brunton, Steven L., Dawson, Scott T. M., Rowley, Clarence W., Colonius, Tim, McKeon, Beverley J., Schmidt, Oliver T., Gordeyev, Stanislav, Theofilis, Vassilios & Ukeiley, Lawrence S. 2017 Modal analysis of fluid flows: An overview. *AIAA Journal* **55**.

Tenney, Andrew S., Glauser, Mark N. & Berger, Zachary P. 2019 Stochastic estimation using modern methods in machine learning. AIAA, aIAA-2019-0324.

Tinney, C.E., Coiffet, F., Delville, J., Hall, A.M., Jordan, P. & Glauser, M.N. 2006 On spectral linear stochastic estimation. *Experiments in Fluids* **41**, 763–775.

Tinney, C.E., Hall, A.M., Glauser, M.N., Ukeiley, L.S. & Coughlin, T. 2004 Designing an anechoic chamber for the experimental study of high speed heated jets. In *42nd AIAA Aerospace Sciences Meeting and Exhibit*. (0010).