DOI: https://doi.org/10.1007/978-3-030-80716-0_19

Data-driven dynamics description of a transitional boundary layer

F. Foroozan[1,*], V. Guerrero[2], A. Ianiro[1] and S. Discetti[1]

Abstract Cluster analysis is applied to a DNS dataset of a transitional boundary layer developing over a flat plate. The stream-wise-span-wise plane at a wall normal distance close to the wall is sampled at several time instants and discretized into small sub-regions, which are the observations analysed in this work. Using K-medoids clustering algorithm, a partition of the observations is sought such that the medoids in each cluster represent the main local states. The clustering has been carried out on a two-dimensional reduced-order feature space, constructed with the multi-dimensional scaling technique. The clustered feature space provides a partitioning which consists of five different regions. The observations are automatically classified as laminar, turbulent spots, amplification of disturbances, or fully-developed turbulence. The Lagrangian evolution of the regions and the state transitions are described as a Markov process in terms of transition probability matrix and transition trajectory graph to determine the transition dynamics between different states.

1 Introduction

The dynamics of turbulent flows are non-linear and characterized by high dimensionality. The capability of machine-learning (ML) tools to deal with such kind of systems is now paving the way to promising research lines. The purpose of ML fluid dynamics is to provide accurate and efficient reduced-order models that capture the essential dynamical features of fluid flows at a reasonable cost [1]. A comprehensive overview of ML methods used for the turbulence modelling and control is reported in [2]. Dimensionality reduction techniques have received particular attention for flow modelling and control purposes in an automated manner.

In this work we propose an automated flow-modelling method to describe the transition to turbulence of a boundary layer (BL), which has important practical

[1] Aerospace Engineering Research Group, Universidad Carlos III de Madrid, Leganés, Spain
[2] Dept. of Statistics, Universidad Carlos III de Madrid, Getafe, Spain. firoozeh.foroozan@uc3m.es

implications due to the enhanced mixing of momentum, higher skin-friction drag, and heat transfer rates. Here we focus on zero-pressure-gradient BL in the presence of an external flow with free-stream turbulence, with the transition to turbulence being classified as bypass transition [3]. The bypass transition mechanism is described as follows. Firstly, there is a region of streaks i.e. the elongated regions of the perturbation of the stream-wise velocity component. Then these elongated distortions reach large amplitude, which can be larger than 10% of the mean flow speed when the free-stream turbulence intensity is only 3%. The streaks break down locally, to form turbulent spots, precursors of localized regions of turbulence. Once fully formed, these spots continue to grow and spread laterally until they merge with the downstream, fully-turbulent region.

In this work, we aim to obtain an automatic input-free domain partitioning technique, to represent the stages of development of a transitional bypass BL. The availability of well-understood theories for the boundary-layer transition makes it an excellent test case to explore the capabilities of ML to discover a flow theory, and also helps us to evaluate our approach. The use of unsupervised ML techniques to identify regions has already been explored on this dataset. An algorithm has been successfully implemented in the work of [4] to demonstrate the ability of an unsupervised self-organizing map as an automatic tool to identify the turbulent-boundary-layer interface in a transitional flow. In contrast to this method, which separates the points into TBL and non-TBL regions, our focus here is on detecting different stages of the transition based on feature similarity.

The overall approach of this work is presented in §2, to describe the dataset and also the implemented unsupervised-learning techniques, and the results are presented in two main parts as *kinematic* (in §3) and *dynamical* analysis (in §4).

2 Methodology

The dataset used in this work was downloaded from the John Hopkins Turbulence Databases, namely it is the Direct Numerical Simulation (DNS) of the transitional BL of the incompressible flow over a flat plate with an elliptical leading edge. The half-thickness of the plate (L) and the free-stream velocity (U_∞) are used as the reference length-scale and reference velocity respectively. The Reynolds number based on these references quantities is 800. The free-stream turbulence intensity is approximately 3% at the leading edge and slightly less than 0.5% at the outlet of the simulation domain. Stream-wise velocity distribution on a wall-parallel (x-z) plane is sampled. This plane is placed at $y/L = 0.25$ (with y being the wall-normal coordinate), sufficiently close to the wall to be representative of the wall-shear distribution. This selected domain is then discretized into small-sized square cells ($20L \times 20L$) to classify regions by inspecting sub-regions of the domain. The cell size here is of the same order of magnitude as the BL thickness, in order to being large enough to capture significant flow structures, and small enough to guarantee a good mapping of the state of the flow, i.e. a sufficient number of cells. The data in each cell is captured with the spacing of $0.1L$. Finally, the data is captured at several time instants with the time spacing equal to the convective time to cross a cell with a convection velocity equal to U_∞. This choice simplifies the analysis since a pseudo-Lagrangian dynamics of observations can be observed by simply stream-wise shift of one cell for each snapshots in time.

Unsupervised learning methods such as Clustering and Multidimensional Scaling (MDS) are applied and evaluated in this work. In order to get a more tractable dataset and remove noisy and redundant features, we propose to reduce the dimensionality before clustering. MDS is a dimensionality reduction technique that maps a set of N points in an original p-dimensional space to a q-dimensional space, where $q << p$, given only a proximity matrix. MDS does not require any prior knowledge of the system behavior. It only expects a time-resolved sequence of observations and a problem-dependent definition of a distance measure. Since we work with high-dimensional datasets, taking advantage of MDS we would be able to reduce the dimensionality of the problem while preserving the data structure in the state space.

Clustering partitions the data by introducing few representatives of the system as basis of the reduced-order models. Here we chose an algorithm called K-medoids, which follows the main procedure of K-means but with a different prototyping; K-medoids selects the most centered observation belonging to the cluster as its prototype instead of the average of the observations. This feature is advantageous since we are seeking to find a specific pattern as the representative of each cluster.

After clustering, the dynamical behaviour between different coherent states (or clusters) can be investigated using the Cluster Transition Matrix (CTM) P. It provides the transition of the observations from one cluster to another when advected downstream, that identifies how the groups of observations in the domain belonging to the same cluster change their grouping in one time step. The elements of this matrix P_{jk}, describing the probability of transition from a cluster C_k to C_j in a given forward step, are defined as the number of observations that move from C_k to C_j in Δt divided by the total number of observations (according to [5]).

In addition, to track the transition of cluster states in continuous time-steps, we defined the pseudo-Lagrangian trajectory of observations in the cluster space. Cluster Transition Trajectory (CTT) model the transition between these regions as a Markov process. Having the trajectory of the most probable transitions between clusters, discovers the prevailing sequence of the flow stages. This definition also helps us to track the most upstream observations while travelling downstream. We combine these ideas to develop our low-order model. Since this is a data-driven approach, the model can be as informative as provided by the data. In the present problem the temporal and spatial evolution are considered simultaneously. Here, the aim is not to predict the future states, but rather their interpretability, which means that we need a customized definition with links to the physics of the flow and not a hypothetical transition time-step. Accordingly, we have defined a physics-oriented transition time, which assumes compatible spatial and temporal steps: the convective time to cross a specified space-resolved cell with a convection velocity equal to U_∞. Since here we have defined a homogeneous coarse-grained spatial domain, this time-step will be fixed throughout the entire domain.

3 Kinematic analysis

Reducing the dimensions of the dataset enables us to plot all the data in a 2D MDS map with the coordinates of γ_1 and γ_2 that provide a useful interpretability tool to identify the characteristics of the observations which are the stream-wise velocity of the domain captured in small cells. To investigate these characteristics, we examined some of the points along γ_1 and γ_2 to illustrate them in the physical space and observe the flow structure inside each sample cell. Accordingly, we can find the parameters that correlate well with γ_1 and γ_2. This investigation showed that along γ_1, the stream-wise variance of the sampled data in the physical space, which is an indicator of the turbulence intensity, is increasing. Along γ_2, however, the high-variance regions of the sampled data move in span-wise direction. To quantify these characteristics of the observations the span-wise profile of the stream-wise velocity variance inside each cell are evaluated with γ_1 and γ_2. The mean value of this profile ($\overline{\sigma}$) and the span-wise position of the center of area of this profile (Z^*) are examined and we found that these two parameters have a linear behavior with respect to γ_1 and γ_2, respectively. This makes a valuable correlation between the axes and these parameters. Thus, it can be stated that the computer has found two consistent metrics to reveal important features. Accordingly, we can confirm that γ_2 is not related to the stream-wise evolution of the structures, thus the transition process.

After reducing the dimensionality of the problem, data clustering is performed on the low-dimensional space (Figure1a). In this work the elbow method [6] is used to choose the number of clusters. The result yield to $K = 6$. In addition to what the process found about γ_1 and γ_2 from correlation parameters, the symmetrical configuration of the clusters further confirms that γ_2 relates directly to the asymmetry of the cells rather than to region recognition. Therefore, it is relevant to shrink the 2D map by reflecting all the points to the upper half part of the map. Thus, clusters numbers 2 and 3 are merged into one cluster and five medoids are shown in Figure1b, in which we can detect different structures inside them. This difference between their internal structures reveals different flow regimes. We can clearly distinguish different stages of the transitional BL flow from streaks, formation of turbulent spots, high velocity fluctuations turbulent regions and fully turbulent region. There is an overshoot of turbulent activity in the transition region, then the turbulence activity decreases while moving towards fully-turbulent state and is in accordance to [7].

4 Dynamical analysis

The results shown in Figure2 report the CTM containing the probability of transition from one cluster (in columns) to another (in rows) within a forward time-step. It represents the dynamical change of state in the entire flow field. It can be seen that the second cluster play a role as transition phase between the first one and the last three which are coupled and reveal the regions of high turbulence. For cluster 3, it has a large probability to remain in the same state, and thus, it reveals the state of fully turbulent. Distance Matrix also discovers the similarity between clusters in terms of the distance parameter in the low-dimensional space (Figure2b). This distance represents the dissimilarity of the primitive cells, the more is the distance, the less similar are the clusters. The same configuration of clusters as the probability matrix can be seen here. It shows that the group of 3 clusters that have the most probability of transition in between are more similar to each other than to other clusters. To track the transition of cluster states in continuous time-steps, we need to model the transition between these regions as a process, which displays the graph of the trajectories shown in Figure2c. Here we assumed that after one time step, each cell travels one-step forward in space. With this assumption, the pseudo-Lagrangian tracking of the cells is captured from the leading edge to the end of the plate and the most probable trajectories are determined. In contrast to the cluster transition matrix that captures all the possible movements in time, this graph captures just the ones that happen between two different clusters. However, the typical path is consistent with the cluster transition matrix, and the cells typically move from laminar to turbulent spots, and finally to regions in the center which is the most spatially homogeneous configuration, thus fully-developed turbulence.

5 Conclusions

With the feature-space discretization, we obtained the regions of development of a transitional BL automatically identified as region containing streaks, turbulent spots, amplification of disturbances, and fully developed turbulent flow. The pseudo-Lagrangian evolution of the regions and the state transitions are employed in terms of transition probability matrix and transition trajectory graph to determine flow dynamics and transition mechanisms between the different states. The development of regions in transitional BL flow presents intermediate stages. The present methodology correctly identifies the bypass transition mechanism. This work thus shows that unsupervised algorithms can identify complex flow dynamics and extract theoretical information.

Acknowledgements This work was partially supported by the research projects:

1. PITUFLOW-CM-UC3M, funded by the call “Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020” under the frame of the Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid.
2. COTURB, funded by the European Research Council, under grant $ERC-2014.AdG-669505$.

References

1. SL Brunton, BR Noack, and P Koumoutsakos (2020) Machine Learning for Fluid Mechanics, J Fluid Mech 52:477–508

2. SL Brunton, and BR Noack (2015) Closed-Loop Turbulence Control: Progress and Challenges, Appl. Mech. Rev. 67(5): 050801
3. TA Zaki (2013) From Streaks to Spots and on to Turbulence: Exploring the Dynamics of Boundary Layer Transition, Flow, Turbulence and Combustion 91(3):451–473
4. Z Wu, J Lee, C Meneveau, and TA Zaki (2019) Application of a self-organizing map to identify the turbulent-boundary-layer interface in a transitional flow, Physical Review Fluid 4:023902
5. E Kaiser, BR Noack, L Cordier, A Spohn, M Segond, M Abel, G Daviller, J Östh, S Krajnović, and R Niven (2014) Cluster-based reduced-order modelling of a mixing layer, J Fluid Mech 754:365—414
6. RL Thorndike (1953) Who belongs in the family?, Psychometrika 18:267–276
7. X Wu, RG Jacobs, JCR Hunt, and PA Durbin (1999) Simulation of boundary layer transition induced by periodically passing wakes, J Fluid Mech 398:109–153

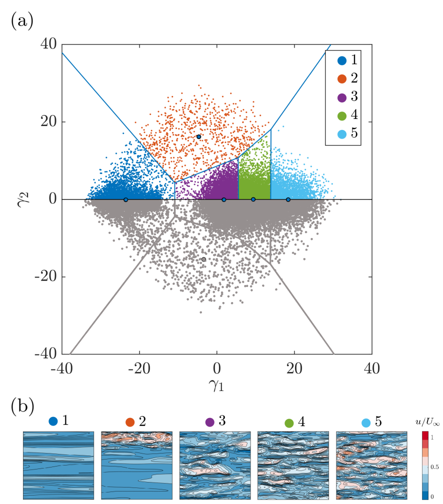

Fig. 1 Kinematic analysis: (a) Clustered two-dimensional MDS map after applying the symmetry. Clusters are specified by colors and the gray points are the merged points. (b) Original contour illustration of the cluster medoids; showing the streaks in medoid 1, turbulent spots in medoid 2, and turbulent stages in medoids 3 to 5 with increasing in intensity.

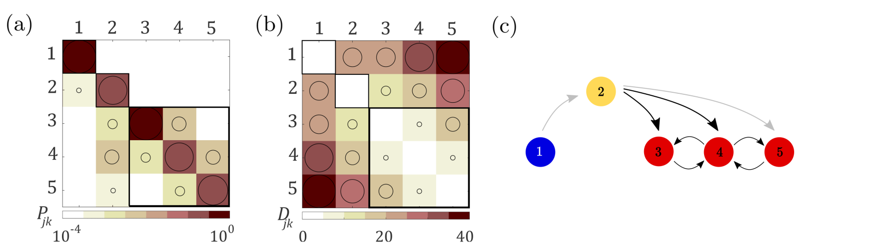

Fig. 2 Dynamical analysis: (a) Cluster Transition Matrix. (b) Cluster Distance Matrix. For the transition matrix the scale is logarithmic, while it is linear for the distance matrix. Cluster subsets are shown in black squares.(c) Graph of Cluster Transition Trajectory. The group of the three final clusters are depicted in one specific color to show their belonging to the turbulent region