Reduced-Order Model Approaches for Predicting Airfoil Performance †

Antonio Colanera *, Eduardo Di Costanzo , Matteo Chiatto and Luigi de Luca

Department of Industrial Engineering, University of Naples Federico II, 80138 Naples, Italy
*
Correspondence: antonio.colanera@unina.it

Abstract: This study delves into the construction of reduced-order models (ROMs) of a flow field over a NACA 0012 airfoil at a moderate Reynolds number and an angle of attack of 8°. Numerical simulations were computed through the finite-volume solver OpenFOAM. The analysis considers two different reduction techniques: the standard Galerkin projection method, which involves projecting the governing equations onto proper orthogonal decomposition modes (POD−ROMs), and the cluster-based network model (CNM), a fully data-driven nonlinear approach. An analysis of the topology of the dominant POD modes was conducted, uncovering a traveling wave pattern in the wake dynamics. We compared the performances of both ROM techniques regarding their prediction of flow field behavior and integral quantities. The ROM framework facilitates the practical actuation of control strategies with significantly reduced computational demands compared to the full-order approach.

Keywords: reduced-order modeling; clustering; airfoils

1. Introduction

The quest to enhance aircraft performance, specifically in terms of drag reduction and improving the lift–drag ratio, is a paramount objective in aircraft design (Couto and Bergada [1]; Chiatto et al. [2]). To achieve these goals, the incorporation of flow control mechanisms, such as synthetic jets (Ceglia et al. [3]; Girfoglio et al. [4]; de Luca et al. [5]), has proven beneficial for delaying separation and managing turbulence transition (Tousi et al. [6,7]). Through careful optimization of airfoil shapes and control parameters, lift, drag, and virtual mass can be tailored to meet specific flight envelope requirements (Lahey et al. [8]). However, this optimization process often necessitates many computational simulations, which can be both time-consuming and computationally intensive (Li et al. [9]; Reumshüssel et al. [10]; Wang et al. [11]).

Modal decomposition techniques have emerged as valuable tools for gaining a deeper understanding of the key coherent flow structures (Taira et al. [12]) and for designing effective control strategies (Rowley and Dawson [13]). The recent increase in computational capabilities has paved the way for the development of various data-driven analytical tools for different aspects of fluid dynamics, including spectral characterization (Schmid [14]), system frequency response (Herrmann et al. [15]), and even the study of two-phase flows (Schmidt and Oberleithner [16]).

Recent research, as highlighted by Stabile et al. [17], emphasizes the relevance of formulating reduced-order models (ROMs) that retain the key features of full-order models (FOMs) while significantly reducing computational demands. This step is crucial for simplifying the governing equations’ complexity and enhancing the efficiency of the optimization process (Noack et al. [18]).
As a matter of fact, when employing reduced-order models one encounters several drawbacks, for instance, the introduction of unphysical artifacts and instabilities (Noack et al. [19]), the loss of fidelity at fine temporal and spatial scales (Taira et al. [20]), and the inability to accurately predict rare or extreme events due to nonlinear interactions (Racca and Magri [21]). Cluster-based network modeling (CNM) stands as a robust approach for investigating complex nonlinear dynamics using data (Fernex et al. [22]). CNM relies on an unsupervised machine learning technique to process a dataset of snapshots, reducing them to a small set of representative flow states. Subsequently, statistical methods are employed to construct a reduced-order model that presents a probabilistic representation of the system dynamics. CNM offers a fully automated framework for understanding system dynamics (Hou et al. [23], Colanera et al. [24]), estimating parameters, and implementing model-based control (Wang et al. [25]).

Furthermore, recent advancements in machine learning techniques (Brunton et al. [26]) have demonstrated their effectiveness, even in the prediction of stall behavior (Saetta et al. [27]). This study aims to develop reduced-order models (ROMs) for analyzing the flow around a specific airfoil configuration. Two distinct types of low-order models were constructed: proper orthogonal decomposition (POD) Galerkin ROMs (Stabile and Rozza [28]), which involve projecting the governing equations onto selected POD modes and cluster-based network models (CNMs) (Kaiser et al. [29] and Li et al. [30]). To the authors’ best knowledge, this paper represents an original application by comparing these two model approaches for airfoil configurations. This paper constitutes an expanded version of the conference paper by Di Costanzo et al. [31].

The paper is structured as follows: Section 2 concisely introduces the methodologies used, including both the full-order model (FOM) and ROMs. In Section 3, the results obtained are presented, along with a comparative analysis of the different low-order models.

2. Methodology

This section provides an overview of full-order modeling for unsteady turbulent flow analysis. Additionally, it describes two different approaches to reduced-order modeling. Full-order modeling offers a detailed representation of the flow field but can be computationally expensive. Reduced-order modeling techniques offer a more efficient alternative by capturing essential flow features while reducing computational costs.

2.1. Full Order Model

The present study investigates the behavior of NACA0012 airfoil under two-dimensional, unsteady, and incompressible flow conditions. To tackle this problem, Unsteady Reynolds-averaged Navier–Stokes (URANS) equations were employed. The numerical simulations were conducted using the pimpleFoam solver within the open-source software OpenFOAM. The pimpleFoam employs the PIMPLE algorithm for pressure–velocity coupling, enabling the modeling of a transient flow within a discretized domain, achieved by solving the Reynolds-averaged Navier–Stokes equations.

For the representation of turbulence in this analysis, the well-established $\kappa-\omega$ turbulence model was implemented. In this context, the equations system is denoted as a full-order model (FOM), as detailed in Equations (1) and (2).

$$\nabla \cdot \mathbf{u} = 0,$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p - \nabla \cdot \left( \nu_{\text{eff}} (\nabla \mathbf{u} + [\nabla \mathbf{u}]^T) \right) = 0,$$

where the time variable, $t$, kinematic pressure, $p$, and the velocity vector, $\mathbf{u}$, represented as $\mathbf{u} = [u, v, w]$, are considered. The effective kinematic viscosity is denoted as $\nu_{\text{eff}} = \nu + \nu_t$, where $\nu$ denotes the kinematic viscosity and $\nu_t$ denotes turbulent eddy viscosity. It is worth noting that in this work any additional effect related to the virtual mass has not been considered. The Reynolds number under free-stream conditions is defined as

$$Re = \frac{\rho_{\text{ref}} U_{\text{fre}} L}{\nu},$$

where $\rho_{\text{ref}}$ is the free-stream density, $U_{\text{fre}}$ is the free-stream velocity, $L$ is the characteristic length, and $\nu$ is the kinematic viscosity.
Re_{\infty} = U_{\infty} c / \nu = 7000, where $U_{\infty}$ represents the free-stream velocity, and $c$ is the airfoil chord. The angle of attack, denoted as $\alpha$, is fixed at $\alpha = 8^\circ$ for this investigation.

The simulations based on the FOM are executed until the values of time-averaged flow variables (such as velocity, drag, lift, and virtual mass when applicable) reach the periodic behavior. Subsequently, snapshots of flow field variables are stored for later use in constructing low-order models.

Utilizing the URANS approach and the $\kappa - \omega$ turbulence model, the FOM describes the unsteady turbulent flow surrounding the NACA0012 airfoil. This forms the basis for subsequent analyses and the creation of reduced-order models.

2.2. POD–ROM

Adopting the methodology outlined in Hijazi et al. [32], one can employ a systematic process that combines the POD technique with Galerkin projections to derive reduced-order models (ROMs) from simulations or experimental data.

Specifically, one can express the truncation of the POD expansions for the discretized fields of velocity $\mathbf{u}$, pressure $p$, and eddy viscosity $\nu$, as follows:

\[
\begin{bmatrix}
u(x_i, y_i, t) \\
u(x_i, y_i, t)
\end{bmatrix} = \Phi_u \mathbf{a}_r, \quad
\begin{bmatrix} p(x_i, y_i, t) \\
u(x_i, y_i, t)
\end{bmatrix} = \Phi_p \mathbf{b}_r, \quad
\begin{bmatrix} \phi(x_i, y_i, t) \\
u(x_i, y_i, t)
\end{bmatrix} = \Phi_\nu \mathbf{c}_r,
\]

where $\Phi_u$, $\Phi_p$, and $\Phi_\nu$ are, respectively, the POD mode matrices of velocity, pressure, and eddy viscosity. The computation of these mode matrices follows the procedure described by Stabile and Rozza [28], utilizing numerical data derived from the FOM.

The momentum equation is subsequently projected onto the POD modes by $\Phi_u$, resulting in the following reduced-order system:

\[
\frac{d\mathbf{a}_r}{dt} = f_u(\mathbf{a}_r, \mathbf{c}_r, \mathbf{b}_r).
\]

Here, $\mathbf{a}_r$, $\mathbf{c}_r$, and $\mathbf{b}_r$ denote the coefficients associated with the velocity, eddy viscosity, and pressure fields, respectively. Additional equations to determine coefficients $\mathbf{b}_r$ and $\mathbf{c}_r$ are derived from the pressure Poisson equation and the eddy viscosity model projections. Initial conditions for $\mathbf{a}_r$ and $\mathbf{c}_r$ are set by projecting $u(0)$ and $\nu(0)$ onto the POD modes. To stabilize the pressure, a velocity space supremizer enrichment strategy is employed, as detailed by Ballarin et al. [33]. The entire procedure is implemented using ITHACA-FV, a C++ library developed in the OpenFOAM framework [28].

The combination of proper orthogonal decomposition (POD) and Galerkin projections presents an effective method for constructing reduced-order models. This approach enables efficient analysis and prediction of flow patterns, substantially lowering computational costs.

2.3. CNM–ROM

The cluster-based network model (CNM) approach depicts the dynamics of a system through a directed network, wherein the clusters (i.e. the nodes) symbolize coarse-grained states of the system. The original snapshots are clustered into a few centroids representing the whole ensemble.

The clusters $C_k$ are computed using the standard k-means++ algorithm (MacQueen [34]; Lloyd [35]), which minimizes the inner-cluster variance. Consequently, the cost function $J$ is defined as follows:

\[
J(C_1, \ldots, C_k) = \frac{1}{M} \sum_{m=1}^{M} \|q^m - C_{\hat{k}(m)}\|_2^2,
\]

where $M$ represents the number of snapshots used for the analysis, $q^m$ denotes the $m$-th snapshot of velocity components, and $C_{\hat{k}(m)}$ denotes the cluster to which $q^m$ belongs. The cluster-affiliation function $k(m)$ associates a snapshot $q^m$ with the index of its nearest centroid:

\[
k(m) = \arg \min_i \|q^m - C_i\|_2.
\]
where \( \| \cdot \|_\Omega \) denotes a spatial norm. The optimal centroids, \( C^*_k \), are those that minimize \( J \):

\[
(C^*_1, \ldots, C^*_K) = \arg \min J(C_1, \ldots, C_K).
\]

After the snapshots are coarse-grained into \( K \) clusters, each cluster can be seen as a representative state of the system dynamics. The network nodes are identified as centroids, \( C_k \), of these clusters, calculated as the state averages within each cluster. The transition characteristics between such clusters rely on high-order direct transition probabilities identified from the data [22].

As sketched in Figure 1, CNM utilizes the direct transition matrix \( Q \), which considers only intercluster transitions and ignores the inner-cluster residence probability.

![Figure 1. Sketch of a directed network of clusters. The nodes denote cluster centroids, while arrows represent transition directions. Each transition includes a transition probability \( Q_{ij} \) and time \( T_{ij} \).](image)

The direct transition probability is inferred from the data as follows:

\[
Q_{k,j} = \frac{n_{k,j}}{n_j},
\]

where \( Q_{k,j} \) represents the probability of transitioning from cluster \( C_j \) to \( C_k \), and \( n_{k,j} \) denotes the number of such transitions, with \( n_j \) being the total number of transitions departing from cluster \( j \). Figure 1 shows an example of a directed network of clusters. For a generalized L-order model, the direct transition probability is expressed as \( \Pr(K_i = j|K_{i-1}, \ldots, K_{i-L}) \) (\( L \) being the model order), which is equivalent to using time-delay coordinates. Note that in the CNM framework, \( n_{i,i} = Q_{i,i} = 0 \).

The residence time \( \tau^n = t^{n+1} - t^n \) corresponds to the duration of the transit in a cluster \( C_k \), where \( t^n \) and \( t^{n+1} \) denote the time of the first and last snapshots to enter and leave \( C_k \). The individual transition time from cluster \( j \) to cluster \( k \) is defined as half the residence time of both clusters:

\[
\tau^{n}_{k,j} = \frac{\tau^{n-1} + \tau^n}{2} = \frac{t^{n-1} + t^n}{2}.
\]

By averaging all \( n_{k,j} \) individual transition times from cluster \( C_j \) to \( C_k \), the transition time \( T_{k,j} = 1/n_{k,j} \sum_{n=1}^{n_{k,j}} \tau^{n}_{k,j} \) is estimated.

Once the probabilistic model is determined through matrices \( Q \) and \( T \), a sequence of centroids, denoted as \( K_{0}, K_{1}, K_{2}, \ldots \), can be determined together with the following time instance:

\[
t_0 = 0, \quad t_1 = T_{k_1,k_0}, \quad t_2 = t_1 + T_{k_2,k_1}, \ldots
\]
In this framework, a smooth motion between these discrete jumps is obtained by a linear interpolation.

Within the CNM framework, the $N$-dimensional state of the system can be represented by various options, including the flow field variables themselves or any other observables. In this study, we consider the POD temporal coefficients as state vector components. Transition probabilities from one cluster centroid to another are estimated based on the training dataset. This estimation enables the computation and reconstruction of the temporal evolution of the reduced state. For further details, the reader is referred to Fernex et al. [22] and Kaiser et al. [29]. In this work, MATLAB (R2020b) was employed for the CNM implementation and test.

3. Results

The methodologies outlined earlier were employed in analyzing 2D numerical data of an unsteady, incompressible flow around an NACA0012 airfoil. These datasets have yielded valuable insights into the flow dynamics and effectiveness of the ROMs. In both the POD-based and CNM–ROM approaches, 400 snapshots were utilized during the training phase, with a sampling time interval of $\Delta t = 0.025$ s.

Figure 2 showcases a time-delay representation of the lift ($C_l$) and drag ($C_d$) coefficients with a time-delay parameter of $\tau = 6\Delta t$.

![Figure 2](image_url)

Figure 2. Time-delay representation of the lift/drag coefficients with a $\tau = 6\Delta t$.

It is possible to determine the lift and drag coefficients from the axial ($C_x$) and the transversal ($C_y$) force coefficients using the following relationship:

\[
C_d = C_x \cos(\alpha) + C_y \sin(\alpha),
\]

\[
C_l = C_x \sin(\alpha) - C_y \cos(\alpha),
\]

where $\alpha$ is the angle of attack. From the inspection of Figure 2, it is possible to recognize the limit cycle behavior of the flow dynamics.

In Figure 3, panels (a) and (b) depict the contours of the time-averaged streamwise velocity component $u$ and the transversal velocity component $v$. 
These contour plots offer a visual representation of the mean flow characteristics. Panels (c) to (f) showcase the leading velocity modes within the POD framework. The first two modes, in particular, capture the most significant fluctuations in the flow field. It is important to note that the spatial shift in the direction of advection of the POD modes corresponds to the traveling pattern of the emitted eddies.

In Figure 4, velocity contour plots from three methods are compared: the FOM, POD-ROM with 10 velocity modes and 12 pressure and viscosity modes, and CNM-ROM using 10 clusters with a 4th-order model. This comparison is made at a specific time instant, $t = 1$ second. This analysis allows us to assess the accuracy and fidelity of both the reduced-order models in capturing the flow features. The constructed flow fields obtained from the POD-ROM and CNM-ROM agree with the original FOM results, indicating the effectiveness of the reduced-order models in capturing the essential flow dynamics.

Figure 5 provides a comprehensive comparison of the axial and transversal force coefficients ($C_x$ and $C_y$) obtained from the FOM, the POD-ROM, and the CNM-ROM at various ROM parameters. This comparison allows us to evaluate the accuracy and predictive capability of the reduced-order models in capturing the aerodynamic forces acting on the airfoil. The black lines represent the $C_x$ and $C_y$ coefficients obtained from the FOM simulations. In panels (a,b), the blue lines correspond to the $C_x$ and $C_y$ coefficients obtained from the POD-ROM, utilizing 10 velocity modes and 12 pressure and viscosity modes. In panels (c,d), the red lines correspond to the coefficients obtained from a first-order CNM-ROM, while the blue lines correspond to a fourth-order CNM-ROM. The figure demonstrates the capabilities of the CNM-ROM technique in accurately predicting the lift and drag coefficients, surpassing the performance of the POD-ROM. Therefore, the CNM-ROM approach (with a higher order) captures the complex flow phenomena more effectively, resulting in improved predictive accuracy of the aerodynamic forces.
Figure 4. Snapshots computed by FOM (first line), POD–ROM (second line), and CNM–ROM (third line). All field variables were normalized to their maximum.

Figure 5. Force coefficients computed with FOM, POD–ROM, and CNM–ROM. In all the panels, the black lines refer to the FOM force coefficients. In panels (a,b), the blue lines correspond to the POD–ROM with 10 velocity modes and 12 pressure and viscosity modes; in panels (c,d), the red lines correspond to a first-order CNM–ROM, while the blue ones correspond to CNM–ROM with $L = 4$.

To further analyze the statistical properties of the flow, the unbiased autocorrelation function $R(\tau)$ between the FOM, POD–ROM, and CNM–ROM approaches is shown in Figure 6. The unbiased autocorrelation function (Protas et al. [36], Deng et al. [37]) is defined as follows:

$$R(\tau) = \frac{1}{T-\tau} \int_{\tau}^{T} u(t-\tau)^T u(t) dt, \quad \tau \in [0, T),$$

where $T$ is the simulation time and $\tau$ denotes the time delay, providing insights into the temporal correlation and persistence of flow features.

Figure 6 demonstrates that the ROM approaches perform quite well in capturing the flow statistical properties and temporal correlations. However, the CNM–ROM exhibits a
closer match to the autocorrelation function obtained from the FOM data, indicating its superior capability in reproducing the flow statistics.

Figure 6. Comparison of the autocorrelation function $R$ between the FOM, POD–ROM, and CNM–ROM. The black line represents the FOM, the red line represents the POD–ROM, and the blue dashed line corresponds to the CNM–ROM.

When considering the first-order CNM, the direct transition probability matrix $Q_{ij}$ and the direct transition times matrix $T_{ij}$ can be examined to describe the dynamics patterns. These matrices provide insights into how frequently and swiftly the system transitions between its various states or clusters. Figure 7 visually represents these matrices in panels (a) and (b), showcasing the direct transition probabilities ($Q_{ij}$) and times ($T_{ij}$) of the CNM.

Notably, the most likely transitions align with the subsequent clusters in the sequence, confirming the presence of a limit cycle in the system dynamics. Furthermore, by examining the transition times matrix, it is possible to estimate the dominant frequency by the following:

$$f_{\text{lead}} \approx \frac{1}{T_{2,1} + T_{3,2} + \ldots + T_{1,10}} \approx 1.4 \text{ Hz.}$$ (13)

The versatility of the CNM–ROM approach becomes evident as it enables predictions of system behavior under various parameter and control configurations. In Figure 8, panel (a) illustrates the predictions for $C_y$ when $\alpha = 6^\circ$. These predictions are generated by training the CNM–ROM with snapshots from different $\alpha$ values, excluding $\alpha = 6^\circ$. This demonstrates the CNM–ROM’s ability to generalize the learned dynamics to unobserved parameter settings.
Within Figure 8, the blue and red lines correspond to $\alpha = 5^\circ$ and $\alpha = 7^\circ$, respectively. The black dashed lines represent the $C_x$ and $C_y$ coefficients obtained from the FOM simulations, and the black solid lines represent the predictions made by the CNM–ROM. Remarkably, the CNM–ROM accurately captures the amplitude and frequency of the force coefficients for varying $\alpha$ values, including the unseen $\alpha = 6^\circ$. This showcases the robustness and generalization capability of the CNM–ROM, as it effectively predicts the aerodynamic forces even for parameter values that were not included in its training data.

![Figure 8](image-url)

**Figure 8.** Prediction of the effect of $\alpha$ using CNM–ROM. The blue and red lines correspond to $\alpha = 5^\circ$ and $\alpha = 7^\circ$, respectively. The black dashed lines represent the FOM simulations, while the black solid lines represent the CNM–ROM predictions.

Overall, the results presented in this section validate the effectiveness of the POD–ROM and CNM–ROM approaches in accurately capturing the flow dynamics, reproducing the flow fields, predicting force coefficients, and capturing the statistical properties of the flow. These findings highlight the potential of reduced-order modeling techniques in enhancing computational efficiency and enabling rapid design iterations for airfoil configurations.

4. Conclusions

This study employs two distinct methodologies for analyzing the flow field over a NACA0012 airfoil at a moderate Reynolds number. The first methodology, the POD–ROM, involves truncating POD expansions for the velocity, pressure, and eddy viscosity fields to a reduced number of modes. The second methodology, the CNM–ROM, utilizes a directed network model, where coarse-grained states are represented as nodes. These reduced-order modeling frameworks offer the potential to design control strategies with significantly reduced computational demands compared to the full-order approach.

Numerical simulations for the full-order model were carried out using the OpenFOAM finite-volume solver. The proper orthogonal decomposition reduced-order model was implemented through the ITHACA-FV library within the OpenFOAM framework.

The comparative analysis of POD–ROM and CNM–ROM methodologies has revealed the superior effectiveness of CNM–ROM in accurately capturing flow dynamics and statistical characteristics while demanding fewer computational resources. The CNM–ROM approach, relying on direct transition matrices and cluster analysis, demonstrated its ability to provide precise flow field predictions and aerodynamic forces.

Furthermore, the CNM–ROM model showcased its versatility by successfully predicting the effect of varying flow parameters, such as the angle of attack. Training the CNM–ROM with snapshots from different angles of attack settings accurately captured the amplitude and frequency of the lift and drag coefficients for unseen angles of attack values. This demonstrates the robustness and generalization ability of the CNM–ROM approach.

The promising results obtained from both the POD–ROM and CNM–ROM methodologies underscore their potential applicability in determining optimal flow control configurations within the reduced-order modeling framework. These methodologies provide
efficient alternatives for analyzing complex flow fields and have the potential to significantly reduce computational costs in aerodynamic design and optimization processes.

**Author Contributions:** Conceptualization, A.C. and E.D.C.; methodology, A.C.; software, E.D.C.; validation, A.C., M.C. and L.d.L.; formal analysis, A.C., M.C. and L.d.L.; investigation, A.C., M.C. and L.d.L.; resources, M.C. and L.d.L.; data curation, A.C., E.D.C., M.C. and L.d.L.; writing—original draft preparation, A.C.; writing—review and editing, A.C., E.D.C., M.C. and L.d.L.; visualization, A.C., E.D.C., M.C. and L.d.L.; supervision, M.C. and L.d.L.; project administration, M.C. and L.d.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**Abbreviations**

The following abbreviations are used in this manuscript:

- URANS: unsteady Reynolds-averaged Navier–Stokes
- FOM: full-order model
- ROM: reduced-order model
- POD: proper orthogonal decomposition
- CNM: cluster-based network model

**References**

34. MacQueen, J.B. Some methods for classification and analysis of multivariate observations 1967, 1, 281–297.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.