Hybrid deep-learning POD-based parametric reduced order model for flow around wind-turbine blade

Hybrid deep-learning POD-based parametric reduced order model for flow around wind-turbine blade

Mandar V Tabib[1], Vasileios Tsiolakis[2], Suraj Pawar[3], Shady E. Ahmed[3], Adil Rasheed [4], Trond Kvamsdal[2] and Omer San[3]

[1]SINTEF Digital, Trondheim, Norway
[2]Department of Mathematical Sciences, NTNU, Trondheim, Norway
[3]School of Mechanical & Aerospace Engineering, Oklahoma State University, Stillwater, USA
[4]Department of Engineering Cybernetics, NTNU, Trondheim, Norway

E-mail: mandar.tabib@sintef.no

Abstract. In this study, we present a parametric, non-intrusive reduced order modeling (NIROM) framework as a potential digital-twin enabler for fluid flow around an aerofoil. A wind turbine blade has its basic foundation in the aerofoil shape. A faster way of understanding dynamic flow changes around the aerofoil-shaped blade can help make quick decisions related to wind-turbine operations and lead to optimal aerodynamic performance and power production. In this direction, a case study involving the application of the NIROM methodology for flow prediction around a NACA 0015 aerofoil is considered. The Reynolds number (Re) is the varying parameter, ranging from 320 000 to 1.12 million and high-fidelity CFD simulations are performed to generate the database for developing the NIROM. The aforementioned NIROM framework employs a Grassmann manifold interpolation approach (GI) for obtaining basis functions corresponding to new values of the parameter (Reynolds number), and exploits the time series prediction capabilities of the long short-term memory (LSTM) recurrent neural network for obtaining temporal coefficients associated with the new basis functions. The methodology involves: (a) an offline training phase, where the LSTM model is trained on the modal coefficients extracted from the sampled high-resolution data using the proper orthogonal decomposition (POD), and (b) an online testing phase, where for the new parameter value, the corresponding flow field is obtained using the GI-modulated basis functions for new parameter and the LSTM-predicted temporal coefficients. The NIROM-approximated flow predictions at new parameters have been compared to the high-dimensional full-order model (FOM) solutions for the high-Re aerofoil case and for a low-Re number wake vortex merger case in order to put the performance of NIROM in perspective. The results indicate that the NIROM framework can qualitatively predict the complex flow scenario around the aerofoil for new values of Reynolds number, while it has quantitatively shown that the LSTM predictions improve with the enrichment of the training space. For the low-Re vortex merger case, NIROM works very well. Thus, it can be deduced that there is scope and potential for continued research in NIROMs as digital twin enablers in wind energy applications.

1. Introduction

Simulations of the flow field around wind turbines can help assess their power production and wake dynamics performance. The wakes emanating from turbine blades can adversely impact downstream turbines unless proper care is taken. Hence, such simulations along with experiments provide useful information for (a) the design of turbine blades, (b) the optimal wind turbine siting and relative placement of wind turbines, and (c) predicting the resultant power production in wind farms. However, such high-resolution numerical simulations are computationally intractable, particularly in the context of digital twins [1]. Moreover, the flow around wind turbine blades resides in the turbulent regime, which involves non-linear dynamics with a wide range of spatio-temporal scales, further increasing the complexity and cost of the simulation process.

The Reynolds-Averaged Navier-Stokes (RANS) strategy has become popular due to its ability to account for turbulence in a computationally efficient way. Nevertheless, such RANS simulations are still infeasible in real-time applications in the context of emerging technologies like Digital Twins. An alternative to the existing techniques, reduced order modeling (ROM) is increasing in popularity due to its ability to reduce the computational burden of the existing high-fidelity simulators. In general, ROM methods attempt to accurately approximate the high-dimensional, complex dynamical models by constructing a system of significantly lower degrees of freedom [2].

Proper orthogonal decomposition (POD) decomposes the flow field into a set of basis functions that optimally describes the original system, i.e. it captures the highest-energy modes to represent the system [3, 4]. However, it has been observed that discarding low-energy modes, viz. modes that correspond to smaller scales, results in instabilities and modeling errors in the ROM approximation [5, 6]. In complex systems, such as turbulent flow, the truncated modes contribute to the dynamics of the flow structures and the energy dissipation. Moreover, the Galerkin-projection based reduced order models (GP-ROMs) require knowledge of the governing equations of the full order model (FOM) solvers. This can prove to be a significant hurdle in developing digital twins, especially when different proprietary software is involved with no access to those governing equations.

To address these two hurdles, we propose a non-intrusive ROM (NIROM), which exploits the strengths of Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNN) in accurately predicting time-series. The current work differs from the previous NIROM related works (e.g., [7, 8, 9, 10]) in two regards. Firstly, it involves using a test case (NACA 0015 aerofoil) that consists of highly turbulent flow structures at high Reynolds number with complex physics. The previous works considered cases such as forced isotropic turbulence [7], magneto-hydrodynamic turbulence [7], Burgers equation [8, 9] or building-induced turbulence [10] at much lower Reynolds number values than the ones in the current work. The present study on high-Reynolds flow over an aerofoil offers a challenging and industrially-relevant test case for assessing the approximative power of the presented NIROM methodology. Secondly, our work considers the Grassmann manifold interpolation [11] to account for the changes in the bases when the flow dynamics change drastically with the parameter (Reynolds number).

The structure of the paper is as follows: Section 2 provides an overview of the methodology and its implementation. Then, in Section 3, we evaluate the predictive performance of the proposed ROM framework with respect to the FOM solutions. Finally, Section 4 provides a summary and conclusions drawn from the study.

2. Non-intrusive reduced order methodology (NIROM)

This section describes the NIROM methodology for predicting the flow fields for new parameter values. Given the non-intrusive nature of the presented framework, no knowledge of the exact FOM is required. Instead, a solely data-driven neural network approach will be used to predict

the time coefficients, coupled with the Grassmann manifold interpolation for the basis functions. As is generally the case with ROMs, this methodology comprises of a computationally heavy offline training phase to build the NIROM and an efficient online phase for reconstructing flow fields for any valid particularisation of the parameters.

The following section outlines the key steps of the NIROM framework (see Fig. 1). More precisely, the offline phase is detailed in sub-sections (2.1-2.3) where NIROM is trained with flow field database from known Reynolds numbers, and the online phase is described in sub-sections (2.4-2.6) where flow-fields for new Reynolds numbers are reconstructed.

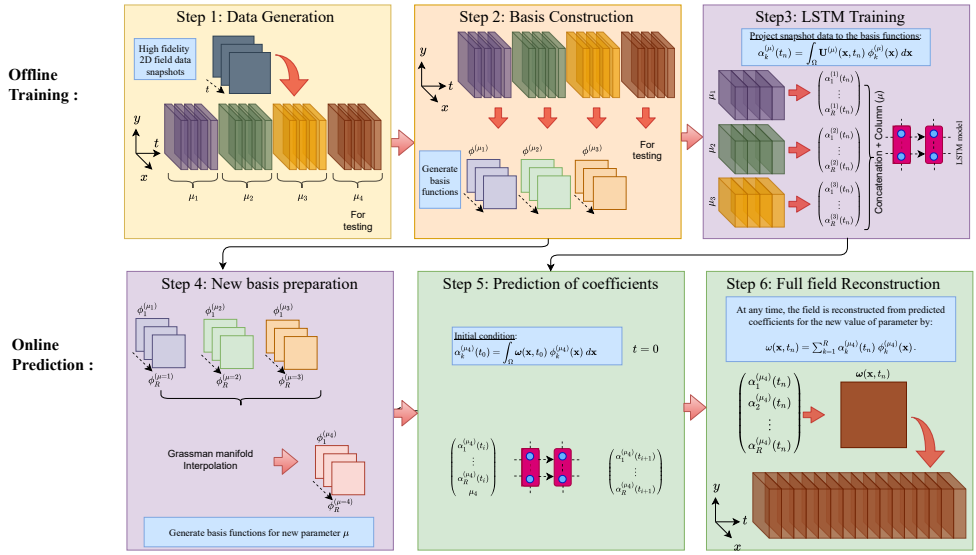


Figure 1: Workflow of the NIROM framework.

2.1. Database generation as the first step in offline phase

The first step in the offline phase consists of generating the database for known parameter values. In this work, the test case consists of the flow around a 2-dimensional NACA 0015 aerofoil with parameterized Reynolds number. The sampling procedure solves the FOM and collects 100 snapshots of the fully developed transient velocity fields for each Reynolds number value in the training set at regular time-intervals. Table 1 shows the values of Reynolds numbers used for training the NIROM in the offline stage and testing its performance in the online phase.

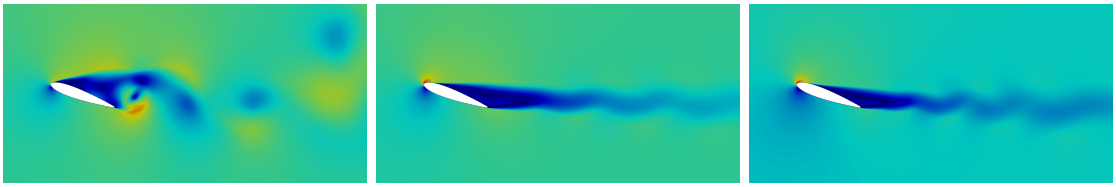
Reynolds number	
Training	Test
	160 000
320 000	
480 000	
550 000	
	640 000
750 000	
820 000	
960 000	
1 120 000	

Table 1: Training and test parameter values.

The Reynolds number range was chosen such that a significant change in the flow characteristics can be observed. Initially, a constant interval of Re = 160 000 was chosen with intermediate values added to enrich the training space and better identify alterations in the flow regime.

The FOM required for generating the snapshots solves the RANS equations using OpenFOAM's [12] finite-volume solvers. The exact transient solver throughout this work is the pressure-implicit with splitting of operators (PISO) algorithm [13]. The aerofoil floats in an O-domain with radius $20c$, where $c = 1\,\mathrm{m}$ is the chord length of the aerofoil, and its centre is the quarter-chord of the aerofoil ($c/4$). The domain is discretised using 130 000 hexahedral cells. On the inlet a

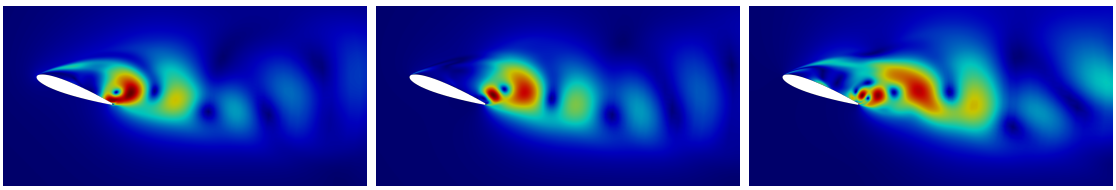
parameterized freestream velocity datum is imposed, while a free traction is imposed on the outlet. Finally, a no-slip condition is imposed on the aerofoil walls. Finally, the kinematic viscosity of the fluid is set to $\nu = 1.56 \cdot 10^{-5}\,\mathrm{m}^2/\mathrm{s}$. The time step is set to $1 \cdot 10^{-5}$ s, resulting in a maximum CFL number of approximately 0.5. It should also be noted that the temporal term is discretised using the Crank-Nicolson scheme and all spatial differential operators are discretised with second order schemes. Each simulation is run for a total of 4 s, while the flow-field from 2 s to 4 s, is used for training and testing. The flow-field is saved with a time-step of 0.02 s, thus generating a database of around 100 snapshots per value of the parameter. Figure 2 clearly depicts the unsteady flow dynamics and flow pattern. Moreover, changes to the Reynolds number introduce significant alterations to the boundary layer and wake structures.



(a) Re = 320 000 (b) Re = 640 000 (c) Re = 1 000 000

Figure 2: Velocity magnitudes depicting the flow structure around a NACA 0015 aerofoil for various Reynolds numbers at $t = 3$ s.

2.2. Construction of the reduced bases in the offline phase.
The second step in the offline phase consists of generating the basis functions (spatial modes) for the fields of interest for each Reynolds number in the training set and their corresponding temporal coefficients. Without loss of generality, let us examine the case for velocity. Using the POD, we compute R spatial modes $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_R]$ for velocity at each parameter value. Figure 3 shows the first three velocity basis functions associated to the Re = 320 000.



(a) $\boldsymbol{\phi}_1$ (b) $\boldsymbol{\phi}_2$ (c) $\boldsymbol{\phi}_3$

Figure 3: POD-constructed velocity spatial modes associated to Re = 640 000.

The time-coefficients are evaluated via projection of the solution matrix on the generated reduced basis

$$\boldsymbol{\alpha}_k = \langle \boldsymbol{U}(\boldsymbol{x}, \boldsymbol{t}), \boldsymbol{\phi}_k(\boldsymbol{x}) \rangle \,, \tag{1}$$

where $k \in [1, R]$ denotes the mode and $\boldsymbol{U}$ is the ensemble of velocity snapshots.

2.3. Training of the LSTM network in the offline phase
The third step in the offline phase involves training an LSTM model to predict the temporal coefficients on reduced order snapshots for any valid lookback time-window σ,

$$\mathcal{M} : \left\{ \alpha_1^{(n)}, \ldots, \alpha_R^{(n)}; \; \ldots \; ; \alpha_1^{(n-\sigma)}, \ldots, \alpha_R^{(n-\sigma)} \right\} \quad \Rightarrow \quad \left\{ \alpha_1^{(n+1)}, \ldots, \alpha_R^{(n+1)} \right\}. \tag{2}$$

The training dataset comprises of an input 3D matrix to LSTM containing the POD-evaluated time-coefficients and it has dimensions of $N \times \sigma \times R+1$, corresponding to the number of samples (N), the lookback time steps (σ) and the number of features ($R+1$). The number of features is equivalent to $R+1$ as the input data comprises of temporal coefficients values associated with R spatial basis function (modes) plus the value of the relevant Reynolds number (parameter). Moreover, an output (target) database of 2D array of the temporal coefficients for time t is provided with dimensions $N \times R$ for training the LSTM. Using the aforementioned data, the LSTM aims to map the inputs (σ previous timesteps) to the output, i.e. the time-coefficients for time t. Table 2 describes the parameters used for training the LSTM model, as obtained by the hyperparameter optimization routine optuna[14]. Figure 4 plots the loss performance of the LSTM model while training for the dynamics of the temporal coefficients of the velocity field. The loss stabilizes after about 300 epochs and training is completed.

Parameters	Values
Number of hidden layers	3
Number of neurons in each hidden layer	85
Batch size	60
Epochs	350
Activation functions in the LSTM layers	tanh
Validation data set	20%
Loss function	MSE
Optimizer	ADAM

Table 2: Hyper-parameters

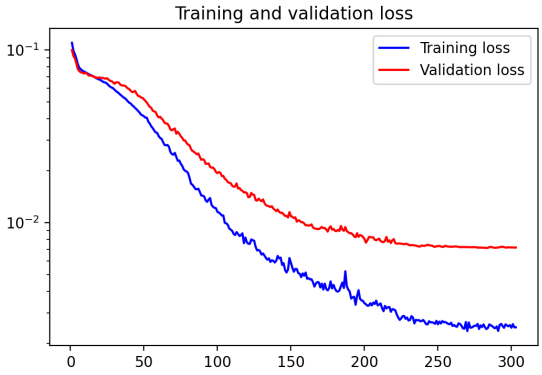


Figure 4: LSTM performance.

2.4. Online phase: Generating spatial modes for new parameter values
In this step, the basis functions for the new parameter value are approximated. This is the first step in the online phase. Starting from the existing POD-calculated reduced bases, the Grassmann manifold interpolation approach is employed to approximate new reduced bases for any unseen parameter value. Each of the R computed reduced bases $\mathbf{\Phi}_i, i = 1, 2..., R$ resides on a non-flat Grassmann manifold. The interpolation is carried out on the flat tangent space at a selected point on the Grassmann manifold. Hence, we start by selecting a reference point S_0 corresponding to a parameter value and its computed set of basis functions $\mathbf{\Phi}_0$, followed by finding the tangent space at this point. Then, all neighboring points, S_i, on the manifold corresponding to the sub-spaces spanned by basis functions $\{\mathbf{\Phi}_i\}$ are mapped onto this tangent space using logarithmic mapping. For a new test parameter, μ_k, the corresponding point S_k is obtained in the tangent space using Lagrange interpolation of known points in tangent space. Finally, the POD basis $\mathbf{\Phi}_k$ corresponding to the test parameter μ_k is computed using the exponential mapping. Readers can refer to [11] for more in-depth information on Grassmann interpolation (GI) and equations.

2.5. Online phase: Generating temporal coefficients
This is the second step in the online phase and consists of generating the new time-coefficients. Using the trained LSTM model $\mathcal{M}$, the time-coefficients of each basis function, i, for the test parameter μ_k, $\alpha_k^i(t)$, are predicted recursively, given σ initial values

$\{a_k^i(t_n), a_k^i(t_{n-1}), \ldots, a_k^i(t_{n-\sigma})\}$.

2.6. Online phase: Approximation of the new flow fields
The final step in the online phase consists of reconstructing the flow fields for the new parameter value by inverse transformation.

$$\boldsymbol{u}(\boldsymbol{x}, t; \mu_k) = \sum_{i=0}^{n} \alpha_k^i(t) \boldsymbol{\phi}_k^i \,, \tag{3}$$

where $n \leq N$ is any valid subset of basis functions.

3. Numerical results
In this section, the proposed methodology is applied to a two-dimensional test case, namely the transient flow around a NACA 0015. The varying feature/parameter of this work is the Reynolds number, which ranges from Re = 160 000 to Re = 1 200 000. It is significant to state that all numbers reside in the turbulent regime and the changes in the flow structure w.r.t. the parameter are extensive, increasing the complexity of the problem. Table 1 shows both the training and test Reynolds number values for this specific case. It should also be noted that Re = 640 000 is used to test the methodology's interpolation ability, while Re = 160 000 for its extrapolation ability.

The following sections will display the results of this methodology. Significant focus will be on the effect of enriching the training space, while also putting the complexity of the case into perspective by comparing the accuracy of the methodology when applied to low-Reynolds problems.

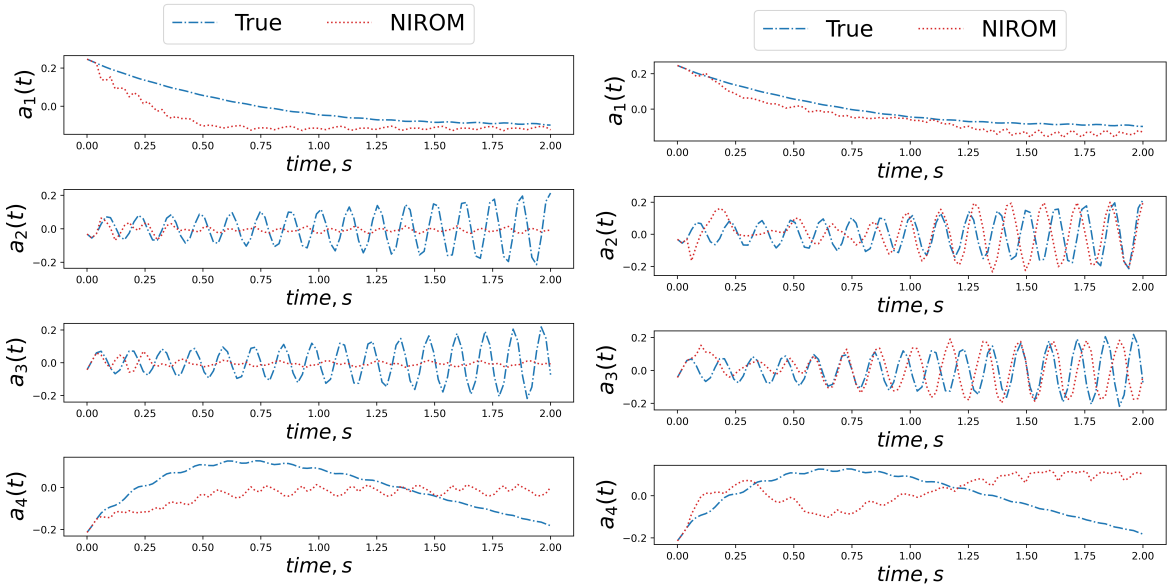
3.1. Performance of the LSTM model and Grassmann-Interpolation method
Following the training part of methodology above, i.e steps 2.1 - 2.3, we assemble the full-order snapshots for all Reynolds number cases in the training set, construct the respective reduced bases and time-coefficients, and use them to train the LSTM model.

First point of interest is the LSTM performance in predicting temporal coefficients for new Reynolds numbers. Figure 5(b) plots the time-coefficients for the first 4 most significant modes for Re = 640 000. It is easy to discern that even though the trained LSTM model can qualitatively identify patterns, it still contains significant errors. However, the influence of further enriching the training space becomes apparent when plotting the same coefficients as predicted from an LSTM model that was trained with fewer samples, namely 2 instead of 7 (Fig. 5(a)). On the other hand, the prediction for Re = 160 000 which resides outside the training space remains poor regardless of the training dataset (Fig. 6). But, this is along the expected lines as extrapolation has been an issue with machine learning applications.

Contrastingly, the influence of the training database on the Grassmann interpolation (GI) of the spatial basis behaves better with fewer training samples, but concentrated around the test point, e.g. two samples Re = 550 000 and Re = 750 000, respectively, when the test point is Re = 640 000 (as seen in figure 7). Due to the significant change of the flow structures with Reynolds number, interpolating becomes challenging for a wider parametric space.

3.2. Approximation of flow field
With GI producing a new reduced basis and LSTM predicting the time-coefficients, the flow fields can be approximated. This section considers the velocity field, however the results hold true for pressure and eddy viscosity.

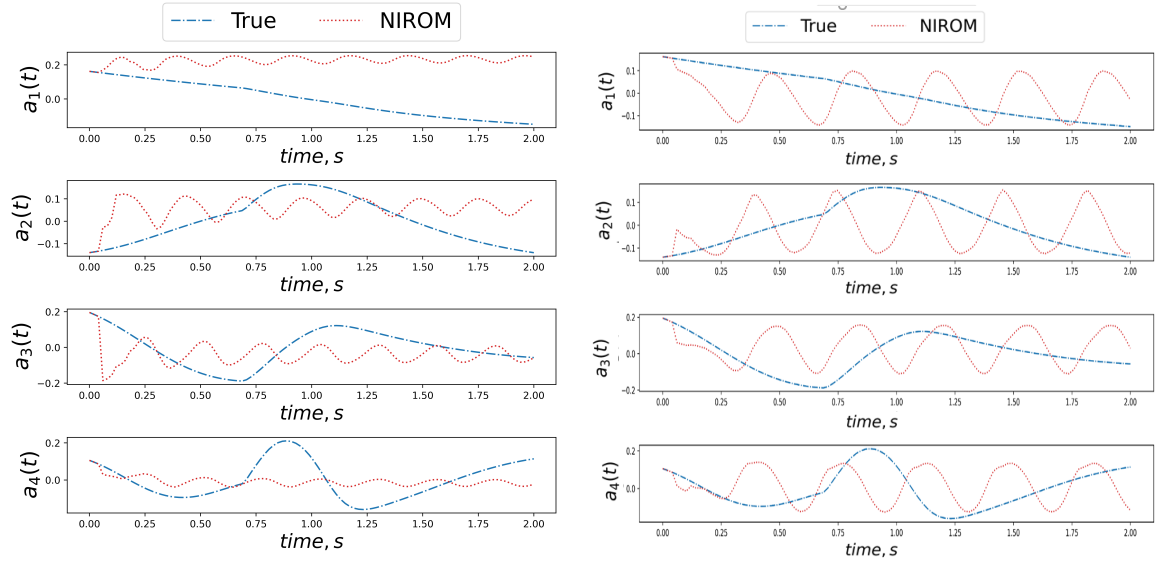
Figures 8-10 show the approximation/reconstruction of the velocity field for the two test points Re = 640 000 and Re = 160 000, respectively. They also show the error between NIROM



(a) LSTM trained with database from two Reynolds numbers

(b) LSTM trained with database from seven Reynolds numbers

Figure 5: Interpolation : Comparison of LSTM-predicted (NIROM) and POD-calculated (true) time-coefficients for Re = 640 000. The first 4 significant modes are plotted. As seen in the figure on right (b), the influence of richer training data-set is leading to better LSTM performance.



(a) LSTM trained with database from two Reynolds numbers

(b) LSTM trained with database from seven Reynolds numbers

Figure 6: Extrapolation: Comparison of LSTM-predicted (NIROM) and POD-calculated (true) time-coefficients for Re = 160 000. The first 4 significant modes are plotted. The influence of a richer training data-set is clear especially for mode 1, but the predictions are not so good for extrapolation case.

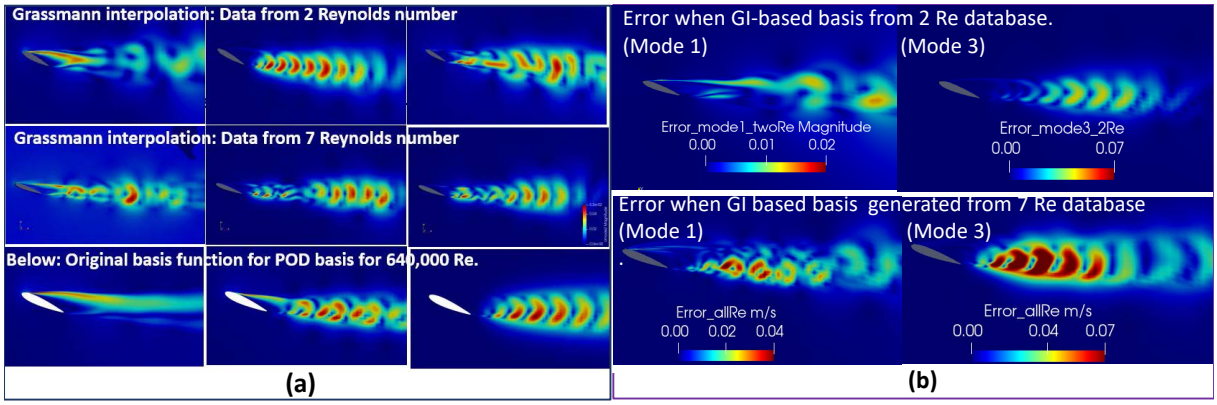


Figure 7: (a) A comparison of first 3 modes of velocity for Reynolds number of 640 000 as obtained by (1.) GI using two-Reynolds numbers database (550k, and 750k), (2.) GI using seven-Reynolds database and (3.) the POD generated basis functions using actual FOM CFD data, and (b) the corresponding error between POD-generated basis on actual FOM data and the GI basis (obtained with two-Reynolds database and with seven-Reynolds database) as shown for two mode 1 and mode 3. The error is less with two Reynolds database than for 7 Reynolds database because these two Reynolds numbers lie near to 640 000.

and FOM in figures 8-10. It can be seen that for the interpolation case, i.e. for Re = 640 000, the NIROM methodology is capturing the qualitative trend, viz. the wake structure and flow pattern are similar to that observed in the FOM. Quantitative errors can be observed, mainly in the wake, concentrated approximately 1 chord length downstream from the trailing edge (as seen in figures 8-9). The relative error reaches a maximum of nearly 100% in regions further than 1 chord length downstream as seen in figure 9. For the extrapolation case, NIROM struggles to capture the wake structure both near the wall and farther downstream in the wake region (figure 10). Such results are expected, based on the observations of Section 3.1 and the complexity of the case, with such extensive range of alterations in the flow structure as the Reynolds number changes.

3.3. Putting LSTM performance in perspective

In this section the performance of the proposed NIROM model as applied to the high-Reynolds case in Section 3 is put in perspective by comparing it with a low-Reynolds vortex merger application. The complexity associated with the high-Reynolds case ($\mathrm{Re} \in [320\,000, 1\,120\,000]$) becomes apparent when we consider the variation of the time-coefficients associated with the first spatial mode w.r.t. the Reynolds number (figure 11(a). In comparison, the same plot for the low-Reynolds case ($\mathrm{Re} \in [200, 800]$) (figure 11(b)) shows an easily discernible pattern.

It is therefore easy to comprehend the requirement for more rigorous training in the first case. Moreover, figure 12 shows how in the vortex merger case, the LSTM is able to learn the pattern and predict it more accurately for the new unseen Reynolds number of 500 and the NIROM can accurately approximate the dynamics of the problem. As such, it is expected that with a richer training dataset, or an adaptive parametric mode, both the LSTM and the GI procedures will improve in accuracy. So, the NIROM method has shown potential and scope for continued research as a good alternative method for ROMs at high Reynolds number applications.

4. Conclusions

In this paper, a fully non-intrusive parametric ROM framework has been developed and used to capture the aerofoil-induced flows for high-values of Reynolds numbers from 320 000 to 1 120 000.

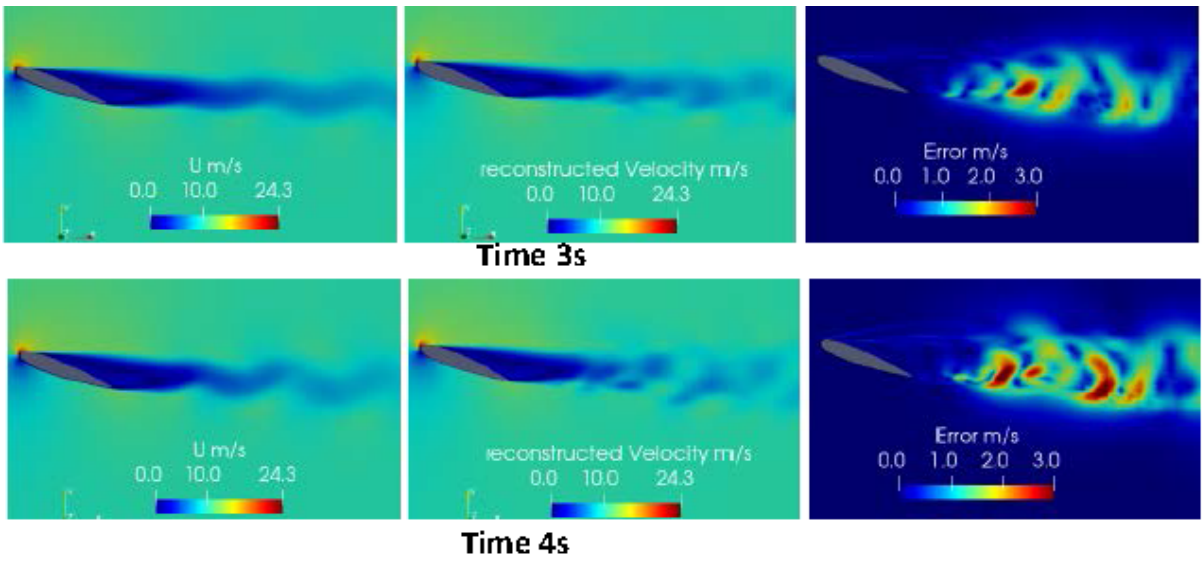


Figure 8: Re = 640 000 Case: Velocity comparison between FOM (left-most) and NIROM approximation (middle) at times t=3s and t=4s. The error $(U_{FOM} - U_{NIROM})$ is plotted in right-most figure.

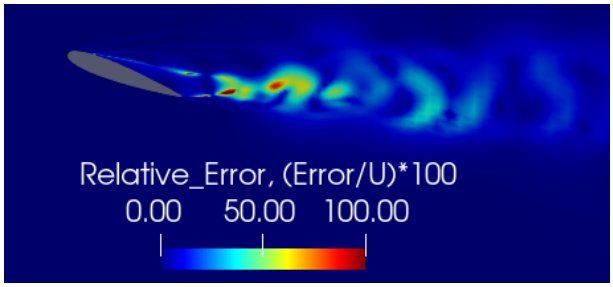


Figure 9: Re = 640 000: Relative error at $t = 4$s in percentage, $(error/U_{\text{FOM}})$*100.

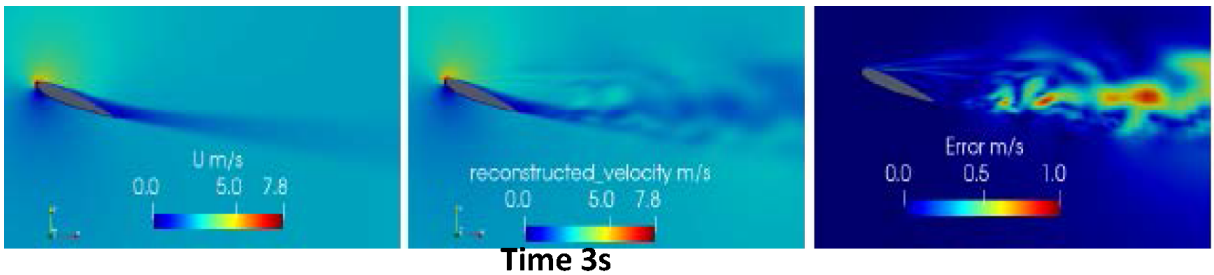
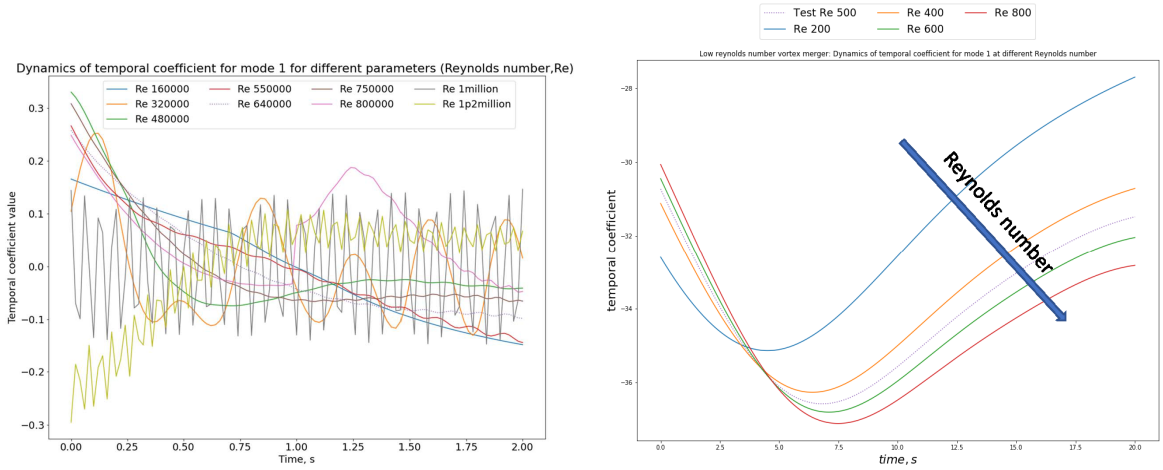
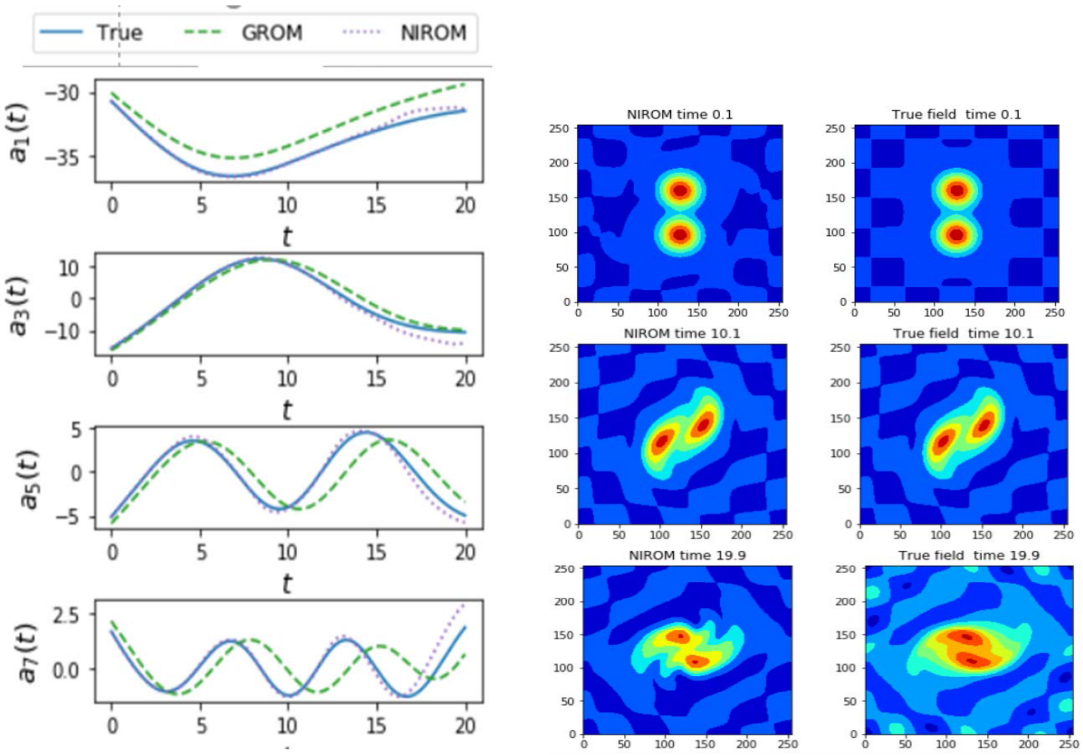


Figure 10: Re = 160 000 Case: Velocity comparison between FOM (left-most) and NIROM approximation (middle) at $t = 3$ s. The error $(U_{FOM} - U_{NIROM})$ is plotted in right-most figure.



(a) High Reynolds Airfoil, Re 320,000-1120,000 (b) Low Reynolds vortex merger, Re 200 to 800

Figure 11: Variation of the time-coefficient associated with the first spatial mode over time with Reynolds number. A comparison between the high-Re aerofoil case and the low-Re vortex-merger case.



(a) LSTM prediction of time-coefficients (b) Comparison of flow field between NIROM approximation true flow field for Re = 500

Figure 12: The proposed NIROM strategy applied to a low-Reynolds number vortex-merger case.

The case study is characterised by high complexity with major changes in the flow pattern (boundary layer separation and wake) with the Reynolds number variation.

Based on our findings we can conclude that the proposed methodology has shown potential by obtaining qualitatively similar wake structure and velocity field predictions for new values of Reynolds number as compared to the full order CFD model. We also experienced significant improvement in the approximative power of the NIROM as we refined the training space. Thus there is much room for improvement. The quantitative errors indicate that the methodology works better for interpolation applications, whereas extrapolation results were characterised by relatively high errors.

Finally, the performance of the presented NIROM methodology when applied to the complex high-Reynolds aerofoil case is put into perspective by comparing it with a low-Reynolds number vortex merger case. Hence, in this case the pattern of variation in the temporal coefficient profile is easily discernible. For this low-Reynolds number vortex-merger case, the NIROM managed to identify the pattern and accurately approximate the dynamics. Thus, the study reveals that NIROMs have good potential and there is scope for improvement of NIROMs for complex high Reynolds number flows with more training data.

5. Acknowledgment

The authors acknowledge the financial support from the Research Council of Norway and the industrial partners of the two projects *NORTHWIND:Norwegian Research Centre on Wind Energy* (project no. 321954) and *CONWIND: Research on smart operation control technologies for offshore wind farms* (project no. 304229). O.S. gratefully acknowledges the financial support of the National Science Foundation under Award Number DMS-2012255 and the Early Career Research Program (ECRP) support of the U.S. Department of Energy under Award Number DE-SC0019290.

6. References

[1] Rasheed A, San O and Kvamsdal T 2020 *IEEE Access* **8** 21980–22012
[2] Quarteroni A and Rozza G 2013 (Springer Publishing Company) ISBN 3319020897
[3] Sirovich L 1987 *Quarterly of Applied Mathematics* **45** 561–571
[4] Tabib M V and Joshi J B 2008 *Chemical Engineering Science* **63** 3695–3715
[5] Akhtar I, Nayfeh A H and Ribbens C J 2009 *Theoretical and Computational Fluid Dynamics* **23** 213–237
[6] Ahmed S E, Pawar S, San O, Rasheed A, Iliescu T and Noack B R 2021 *Physics of Fluids* **33** 091301
[7] Mohan A T and Gaitonde D V 2018 *arXiv preprint arXiv:1804.09269*
[8] Pawar S, Ahmed S E, San O and Rasheed A 2020 *Physics of Fluids* **32** 036602
[9] Ahmed S E, Pawar S, San O, Rasheed A and Tabib M 2021 *Computers & Fluids* **221** 104895 ISSN 0045-7930
[10] Tabib M V, Pawar S, Ahmed S E, Rasheed A and San O 2021 **2018** 012038
[11] Amsallem D and Farhat C 2008 *AIAA Journal* **46** 1803–1813
[12] OpenCFD Limited 2021 OpenFOAM: The Open Source CFD Toolbox. Programmer's Guide
[13] Issa R 1986 *Journal of Computational Physics* **62** 40–65 ISSN 0021-9991
[14] Akiba T, Sano S, Yanase T, Ohta T and Koyama M 2019 *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*