

PAPER • OPEN ACCESS

A non-intrusive parametric reduced order model for urban wind flow using deep learning and Grassmann manifold.

To cite this article: Mandar V Tabib *et al* 2021 *J. Phys.: Conf. Ser.* **2018** 012038

View the [article online](#) for updates and enhancements.

You may also like

- [Performance Simulation of Tubular Segmented-in-Series SOFC Using Simplified Equivalent Circuit](#)
Shun Yoshida, Tadashi Tanaka and Yoshitaka Inui
- [\(Invited\) Potential of Double Network Gel as a Tribological Material Realizing Low Friction in Water](#)
Koki Kanda and Koshi Adachi
- [Analysis of the Effect of Surface Diffusion on Effective Diffusivity of Oxygen in Catalyst Layer By Direct Simulation Monte Carlo](#)
Tomoki Hori, Takuya Mabuchi, Ikuya Kinefuchi et al.



The Electrochemical Society
Advancing solid state & electrochemical science & technology

242nd ECS Meeting

Oct 9 – 13, 2022 • Atlanta, GA, US

Abstract submission deadline: **April 8, 2022**

Connect. Engage. Champion. Empower. Accelerate.

MOVE SCIENCE FORWARD



Submit your abstract



A non-intrusive parametric reduced order model for urban wind flow using deep learning and Grassmann manifold.

Mandar V Tabib¹, Suraj Pawar², Shady E. Ahmed², Adil Rasheed¹, Omer San²

¹SINTEF DIGITAL, Trondheim, Norway

²School of Mechanical & Aerospace Engineering, Oklahoma State University, Stillwater, OK 74078, USA.

E-mail: mandar.tabib@sintef.no

Abstract. In this study, we present a parametric non-intrusive reduced order modeling framework as a potential digital twin enabler for fluid flow related applications. The case study considered here involves building-induced flows and turbulence with inlet turbulence value as a parameter. The framework proposed employs a Grassmann manifold interpolation approach (GI) for obtaining basis functions corresponding to new values of parameter, and exploits the time series prediction capability of long short-term memory (LSTM) recurrent neural network for obtaining temporal coefficients associated with the new basis functions. The methodology works in the following way: (i) in the training phase, the LSTM model is trained on the modal coefficients extracted from the high-resolution data using proper orthogonal decomposition (POD) transform for the known values of parameter, and (ii) in the testing phase, the trained model predicts the modal coefficients for the total time recursively based on the initial time history for the new value of parameter. Then, we reconstruct the flow fields for the new value of parameter (new inlet turbulent value) using the GI modulated basis functions and LSTM predicted associated temporal coefficients. To assess the performance of the proposed model, the ROM-LG predictions are compared with the high-dimensional full-order model solutions using L1 and L2 error analyses as well as with the conventional POD based ROM (ROM-POD) solutions. The results indicate that the non-intrusive ROM (ROM-LG) framework yields a stable solution for the velocity fields and for short-term prediction of dynamic turbulent kinetic energy fields. This work has scope for further development and will be useful for building-integrated wind energy and urban drone operation in a smart-city digital twin platform.

1. Introduction

Turbulent flows in fluid dynamics are non-linear dynamical systems involving a wide range of spatio-temporal scales. Numerical simulation of the such flows involving using Direct Numerical Simulation will be computationally intractable even for the computational infrastructure foreseen in the near future. Even with relatively computationally lighter models like Large Eddy Simulation (LES) and Reynolds Averaged Navier Stokes (RANS) models, such simulations are not sufficient to fulfill the realtime simulation capability in the context of emerging technologies like Digital Twins [1]. As an alternative to the existing techniques, reduced order modeling (ROM) is emerging as a promising approach to reduce the computational burden of the existing



high-fidelity simulators. In general, ROM works in such a way that the high-dimensional complex dynamical systems are represented with much lower-dimensional (but dense) systems while keeping the solution quality within the acceptable range [2].

Among the different variants of ROM strategies, the Galerkin Projection (GP) based ROMs have been utilized extensively in several domains. One of the commonly used methods for constructing the reduced basis which are used in the projection step is the Proper Orthogonal Decomposition (POD). It is one of the most popular model reduction methods which decomposes the flow field into a set of basis functions that optimally describes the system and selects only the most energy-conserving bases to represent the system [3]. However, it has been observed that discarding the modes results in instabilities and modeling errors in the solution [4] as these modes contribute to the dynamics and dissipation of flow structures. Furthermore, the GP-ROM approach requires the governing equation of the high fidelity simulators before the model can be developed. This is a major hurdle in building digital twins where different vendors and software developers contribute with their own proprietary software. To address these two major problems, we propose a fully non-intrusive ROM (NIROM) which exploits the strengths of Long Short Term Memory (LSTM) Recurrent Neural Networks (RNN) in accurately predicting timeseries. Such NIROM has been previously shown to outperform the state-of-the-art GP-ROMs ([5, 6, 7]). The current work differs from all the previous works in two regards: (a) It involves application consists of highly turbulent wind flows involving real complex geometries encountered in cities, and (b) utilization of Grassmann manifold interpolation [8] to account for the changes in the bases when the flow dynamics changes drastically.

The layout of the paper is as follows: Section 2 provides an overview of the methodology and its implementation. In Section 3, we evaluate the predictive performance of the proposed ROM framework with respect to the standard POD-ROM and full order model solutions. Finally, Section 4 provides summary and conclusions drawn from the study.

2. Non-intrusive ROM-LG Methodology

Figure 1 shows the methodology of the ROM-LG framework to obtain field for new values of the parameter. Intrusive ROMs needs exact equations driving the phenomena, while non-intrusive ROMs as the ones developed here can work without knowing them. Here, we bypass the physics-based Galerkin projection part with completely data-driven neural network approach to predict the modal coefficients, and exploit the Grassman manifold interpolation for basis function. The methodology comprises of a computationally heavy offline training phase to develop the NIROM and an efficient online phase for reconstructing flow fields for a new value of parameter from newly computed basis function and temporal coefficients. In online testing phase, we recursively predict the modal coefficients for the total time using the trained model \mathcal{M} . The input of the trained LSTM model \mathcal{M} will be the states (i.e. the R modal coefficients corresponding to each of selected R spatial modes) for the pre-selected time value of σ (which is look-back in time i.e. number of previous time steps), so the input is denoted as: $\{a_1^{(1)}, \dots, a_R^{(1)}; \dots; a_1^{(\sigma)}, \dots, a_R^{(\sigma)}\}$ and the output will be recursive prediction of corresponding future time states. The key steps of the ROM-LG framework are outlined below in detail (see also Figure 1).

Step 1. Database generation for known values of parameter: We choose the 3D building-induced turbulent flows as our test bed to evaluate the performance of parametric non-intrusive ROMs where the parameter is the inlet turbulence profile. Database generation involves collection of 900 2D snapshot data for the variable fields Velocity and turbulent kinetic energy (TKE) for each unique value of the parameter from the full-order model (FOM) simulation for different values of parameter. The FOM model comprises of 3D transient Reynolds averaged Navier stokes (RANS) and the OpenFOAM software has been used. The details of the 3D model used for building-induced flows are described in this work [9]. The database is generated by

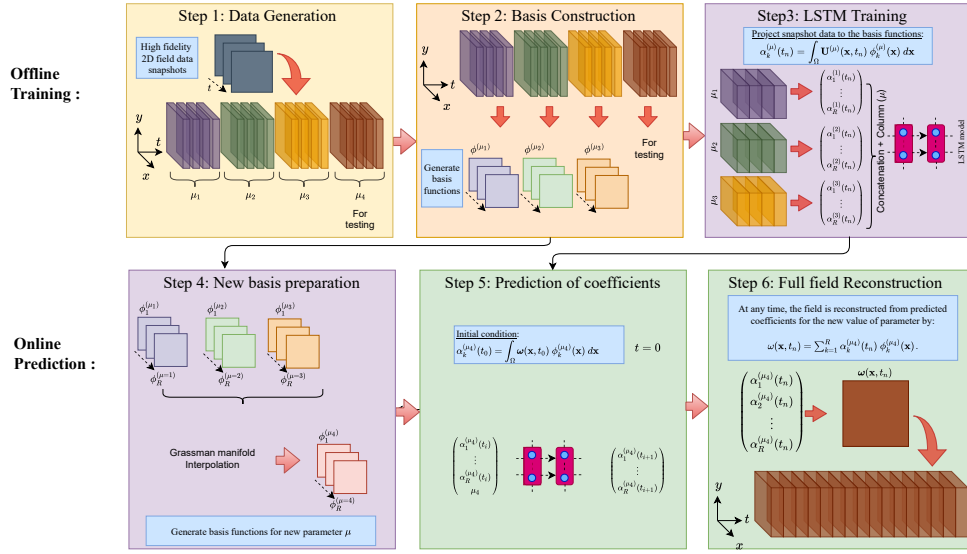


Figure 1: Workflow of the ROM-LG framework.

conducting four unsteady simulations, each corresponding to a particular value of the parameter (inlet turbulent level characterized by inlet effective turbulent viscosity). This parameter is chosen to test the methodology for the turbulent flows. The unsteady flow dynamics is seen to become more turbulent with each simulation with the increase in the effective viscosity (figure 2). Figure 2 shows the four values of the parameter chosen for this work (three of those are used for training purposes in the offline stage and the fourth one is to test the framework performance in online phase in next stage). Each of the four simulation are conducted for a total of 210 s. The flow-field from 30 s to 210 s is used for training and testing. The flow-field is saved with a time-step of 0.2 s, thus generating a database of around 900 snapshots per value of the parameter.

Step 2: Basis function (mode) generation for each field for a given value of parameter: Using POD snapshot method, we compute R POD modes $\Phi = [\phi_1, \phi_2, \dots, \phi_R]$ for velocity and TKE for each value of parameter, respectively. Figure 3 shows the first 3 modes (basis functions) of turbulent kinetic energy obtained for each of the three different training values of the parameter. The basis functions (modes) show the influence of the parameter (change in values of inlet turbulence level) on the fields.

Temporal coefficient associated with the basis functions: Construct temporal modal coefficients by a forward transform through projection of the variable field (here $\omega(x, y, t_n)$ denotes variables, like velocity or turbulent kinetic energy) onto the the basis functions (ϕ_k) as shown in equation below.

$$a_k(t_n) = \langle \omega(x, y, t_n); \phi_k \rangle, \quad (1)$$

where $a_k(t_n) = \{a_k^{(1)}, a_k^{(2)}, \dots, a_k^{(N)}\}$.

Step 3: Data-driven model: Train the LSTM model on reduced order snapshots for selected lookback time-window σ ,

$$\mathcal{M} : \{a_1^{(n)}, \dots, a_R^{(n)}; \dots; a_1^{(n-\sigma)}, \dots, a_R^{(n-\sigma)}\} \Rightarrow \{a_1^{(n+1)}, \dots, a_R^{(n+1)}\}. \quad (2)$$

The parameters used for the training are described in Table 1.

Figure 4 shows the loss performance of the LSTM model while training to capture the dynamics of temporal coefficients for turbulent kinetic energy. Similar behaviour is observed for the velocity database as well (not shown).

Parameter	Inlet effective viscosity
Test parameter value	6.35 m2/s
Training parameter value 1	1.58 m2/s
Training parameter value 2	25.4 m2/s
Training parameter value 3	63.5 m2/s

Figure 2: Parameter values used in our case.

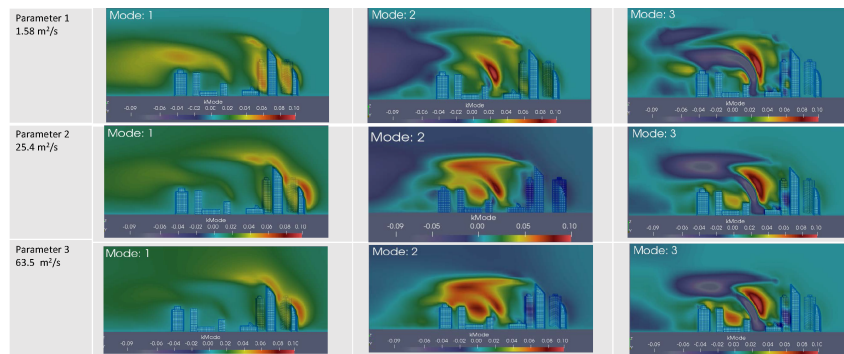


Figure 3: First 3 modes (basis functions) of turbulent kinetic energy.

Table 1: Hyper-parameters used to train the LSTM network

Parameters	Values
Number of hidden layers	3
Number of neurons in each hidden layer	80
Batch size	50
Epochs	350
Activation functions in the LSTM layers	tanh
Validation data set	20%
Loss function	MSE
Optimizer	ADAM

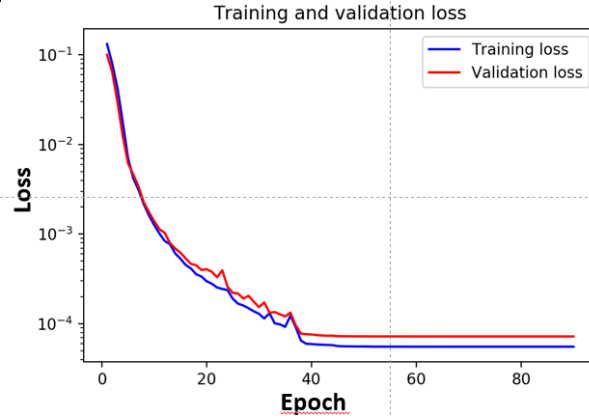


Figure 4: LSTM training performance: loss while training.

Step 4: New basis functions corresponding to new unseen value of a parameter: During online deployment, we utilize a Grassmann manifold interpolation approach to approximate the POD basis set for the unseen test parameter value using the known available POD basis sets. All the known computed sets of basis functions $\{\Phi_i, i = 1, 2, \dots, p\}$ corresponding to different parameter values i lies as points on a non-flat Grassmann manifold. We conduct the interpolation on the flat tangent space of a selected point on the Grassmann manifold. Hence, as a first step, the Grassmann manifold interpolation consists of first choosing a reference point S_0 corresponding to a parameter value and its computed set of basis functions Φ_0 , and then finding the tangent space at this point. Then the neighbouring points (denoted by S_i) on the manifold corresponding to the sub-spaces spanned by basis functions $\{\Phi_i\}$ are mapped onto this tangent space using logarithmic mapping. Here, matrix Γ_i represents the tangent space where each point S_i is mapped (see equations 3-4). In this work, the Grassmannian manifold interpolation utilizes the basis set corresponding to the training parameter value of 1.58 as the reference point and it obtains a new set of basis functions for the unseen test parameter value training parameter value of 6.35.

$$(\Phi_i - \Phi_0 \Phi_0^T \Phi_i)(\Phi_0^T \Phi_i)^{-1} = \mathbf{U}_i \Sigma_i \mathbf{V}_i^T, \quad (3)$$

$$\Gamma_i = \mathbf{U}_i \tan^{-1}(\Sigma_i) \mathbf{V}_i^T. \quad (4)$$

Then, the matrix Γ_t corresponding to the test parameter ν_t is obtained by the Lagrange interpolation of the matrices Γ_i corresponding to known parameter values $\nu_i, i = 1, 2, \dots, P$ and corresponding basis set, as shown in equation 5.

$$\Gamma_t = \sum_{i=1}^P \left(\prod_{\substack{j=1 \\ j \neq i}}^P \frac{\nu_t - \nu_j}{\nu_i - \nu_j} \right) \Gamma_i, \quad (5)$$

where P refers to the number of the control parameters for offline simulations (e.g., $P = 3$ in our numerical examples). Finally, the POD basis Φ_t corresponding to the test parameter ν_t is computed using the exponential mapping as follows:

$$\Gamma_t = \mathbf{U}_t \Sigma_t \mathbf{V}_t^T, \quad (6)$$

$$\Phi_t = [\Phi_0 \mathbf{V}_t \cos(\Sigma_t) + \mathbf{U}_t \sin(\Sigma_t)] \mathbf{V}_t^T. \quad (7)$$

Readers can refer to [8] for more in-depth information on Grassmann interpolation. Figure 5 compares the GI obtained basis functions with the basis functions obtained by POD based on high-fidelity test dataset. Note that the basis functions seem quite similar to each other. The GI obtained basis functions are orthogonal, and they are generated without knowing the database of fields for new parameter or without knowing the equations.

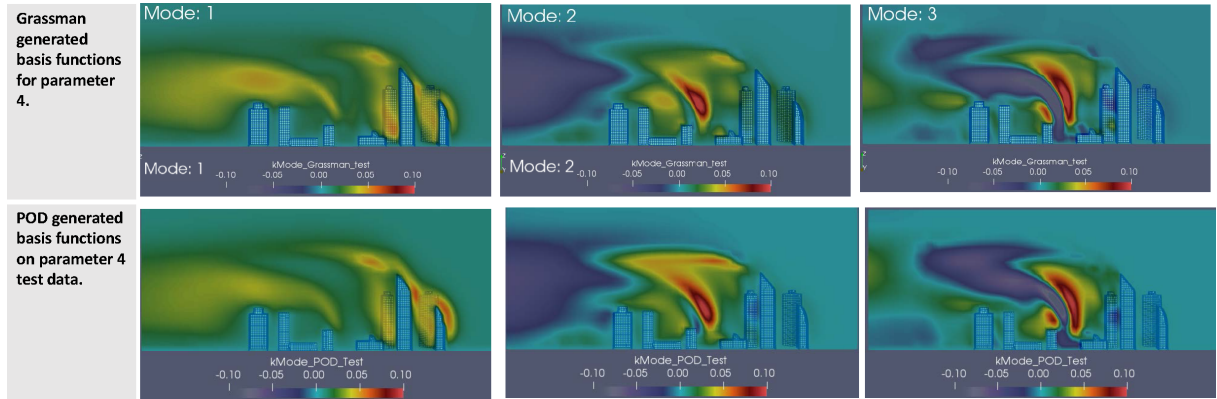


Figure 5: A comparison of the GI obtained basis functions with the basis functions obtained by POD based on high-fidelity test dataset for turbulent kinetic energy.

Step 5: LSTM generated temporal coefficients: Use the trained LSTM model \mathcal{M} to recursively predict temporal coefficients $a_k(t)$ until final time reached, given initial values $\{a_k^{(1)}, a_k^{(2)}, \dots, a_k^{(\sigma)}\}$ based on σ and GI obtained basis functions. Figure 6 shows the comparison between LSTM predicted temporal coefficient for the new parameter value for the four new basis modes versus the true coefficient obtained from the POD based algorithm on the actual data. The results indicate that for most basis modes, the LSTM is able to capture the dynamics over 50 s duration well. For longer-terms, some deviation has been observed, especially in turbulent kinetic energy. In Section 3, we check the results of this on the actual flow reconstruction.

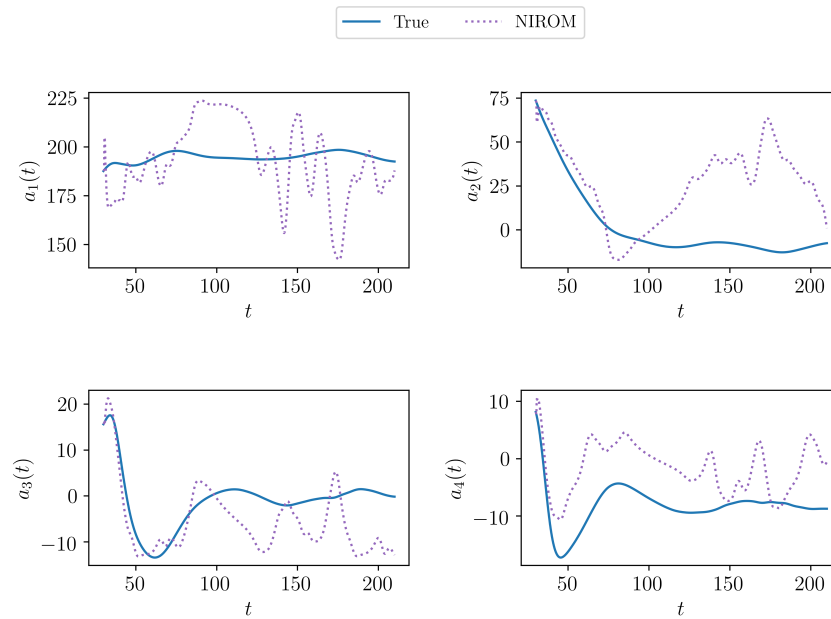
Step 6: Reconstruct the flow fields for the new parameter value by inverse transform.

3. Numerical Results

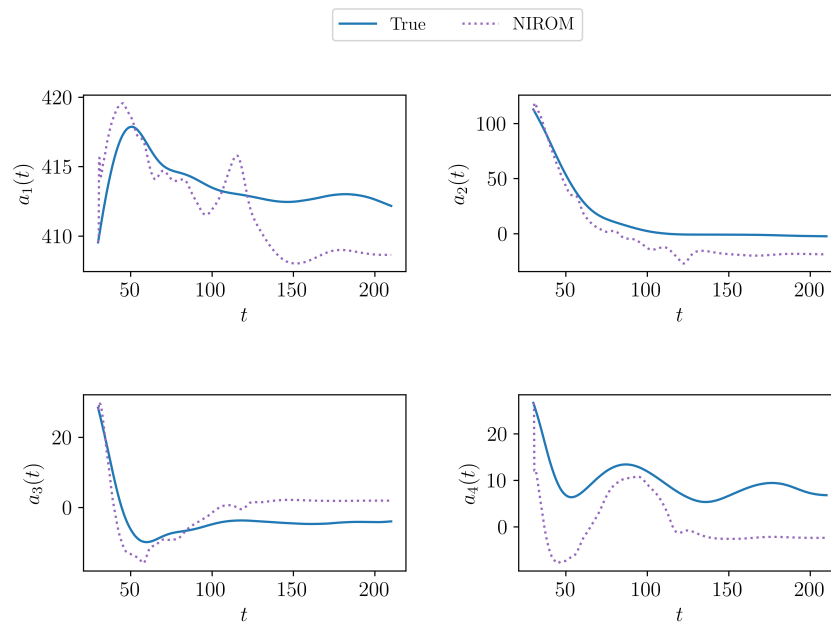
The transient flow fields that are obtained for the test data (i.e. for a new value of the parameter) using the ROM-LG NIROM methodology have been compared with the (a) actual flow fields and, (b) with the POD-ROM flow fields for three representative times. Figure 7 and Figure 8 show this comparison for velocity fields and turbulent kinetic energy fields, respectively with the error in the last row. This error for the ROM-LG methodology is computed related to the FOM as shown in equation 8, where ω represent the fields (velocity or turbulent kinetic energy):

$$error_i = \frac{\omega_{ROM-LG,i} - \omega_{FOM,i}}{\omega_{FOM,i}} \quad (8)$$

for all i 's from $i=1..N$, where i is the computational cell where flow fields are computed. The results at the three representative time values (36 s, 66 s, 174 s) for the velocity field (Figure 7), show that the majority of region shows lower values of error (nearer to the value of zero) except few regions of the flow. The non-intrusive methodology is showing decent comparisons with the FOM results for velocity. For the turbulent kinetic energy (TKE) field (Figure 7), the error is higher than the velocity field error at longer-time predictions of 174 s as the TKE fields are more dynamic. Further, the L2 norm of normalized error of NIROM with respect to FOM is shown in Figure 9 over all the times. Here, it confirms the lower values of errors for the velocity field than



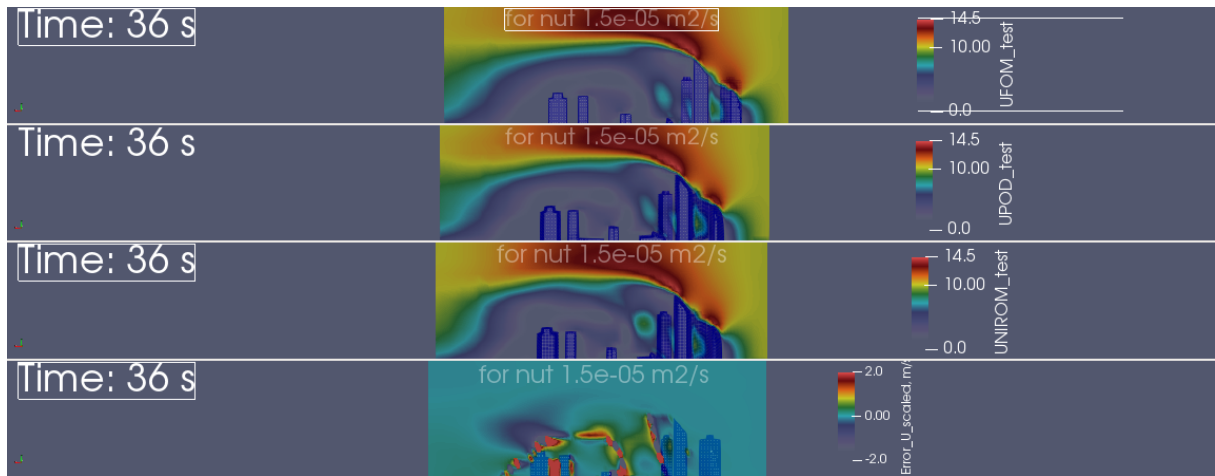
(a) Turbulent kinetic energy: LSTM vs True



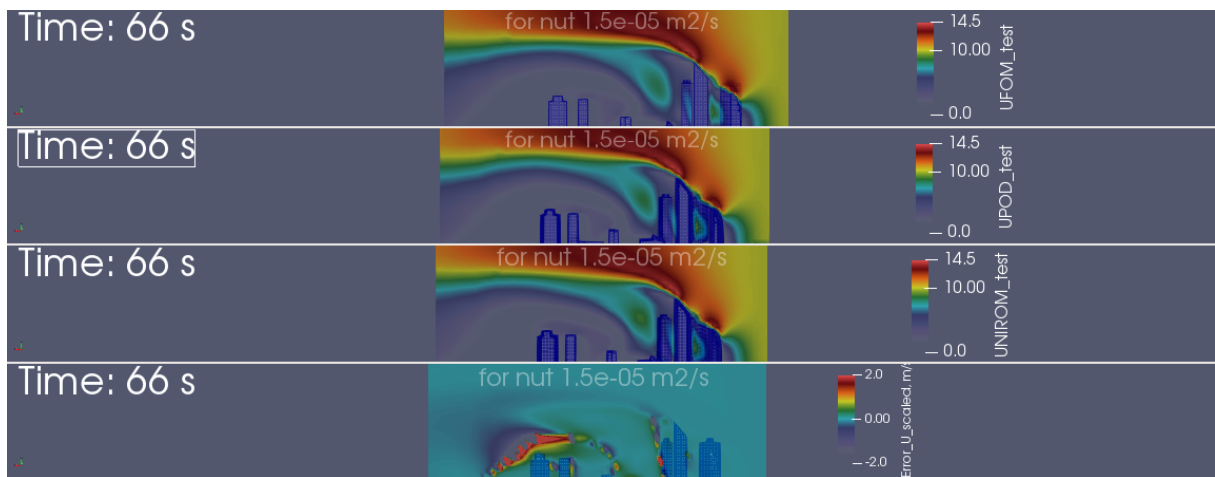
(b) Velocity: LSTM vs True

Figure 6: Comparison: The LSTM predicted dynamics of temporal coefficient corresponding to four basis modes of a new unseen parameter value, Vs the true temporal dynamics of the coefficient . The X-axis is time and has units of seconds. The out-of-sample temporal coefficients for the four basis modes are evaluated from time ($t = 30$ to $t = 200$ s).

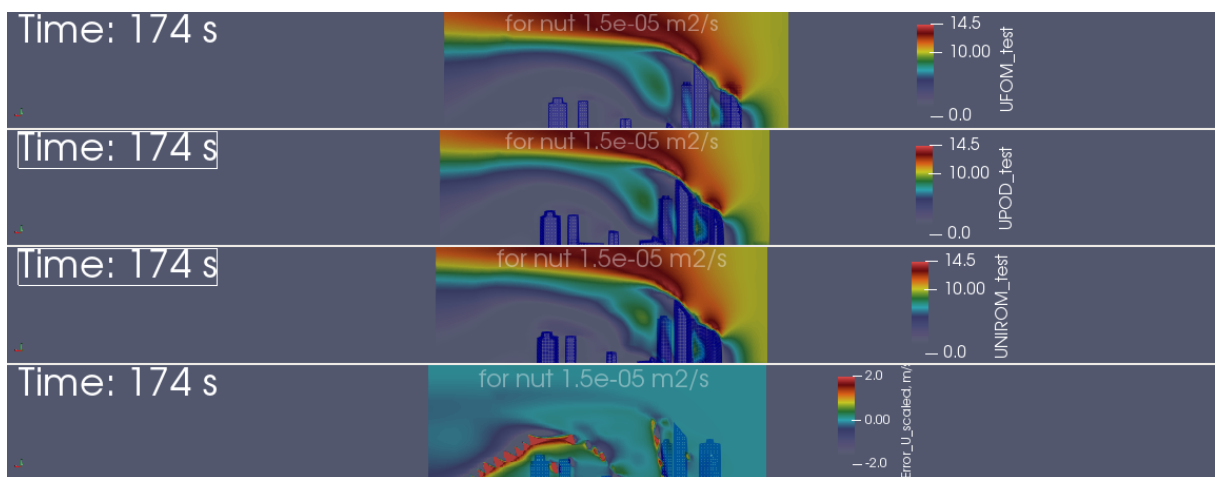
the TKE field, and the error increases with time for the TKE field. This could be attributed to the difficulties faced by the current LSTM architecture in capturing long-term dynamics in TKE beyond 50 s of operation. The non-intrusive methodology is showing decent comparisons with



(a) 36 s

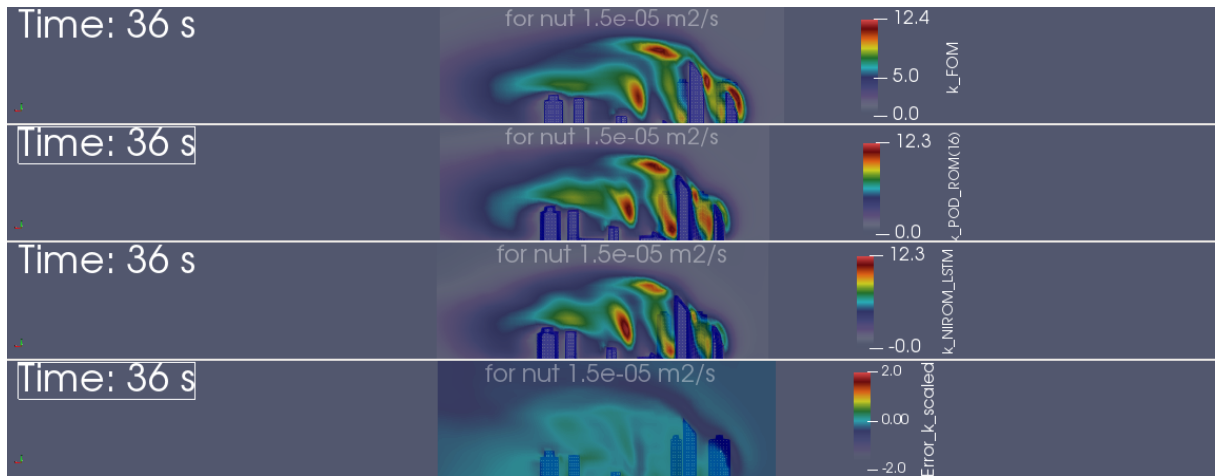


(b) 66 s

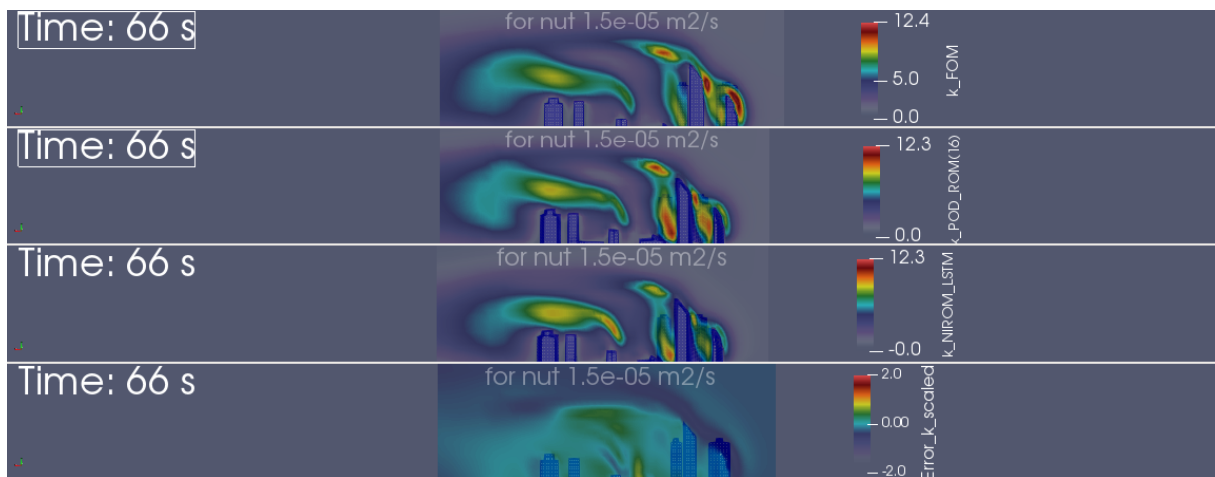


(c) 174 s

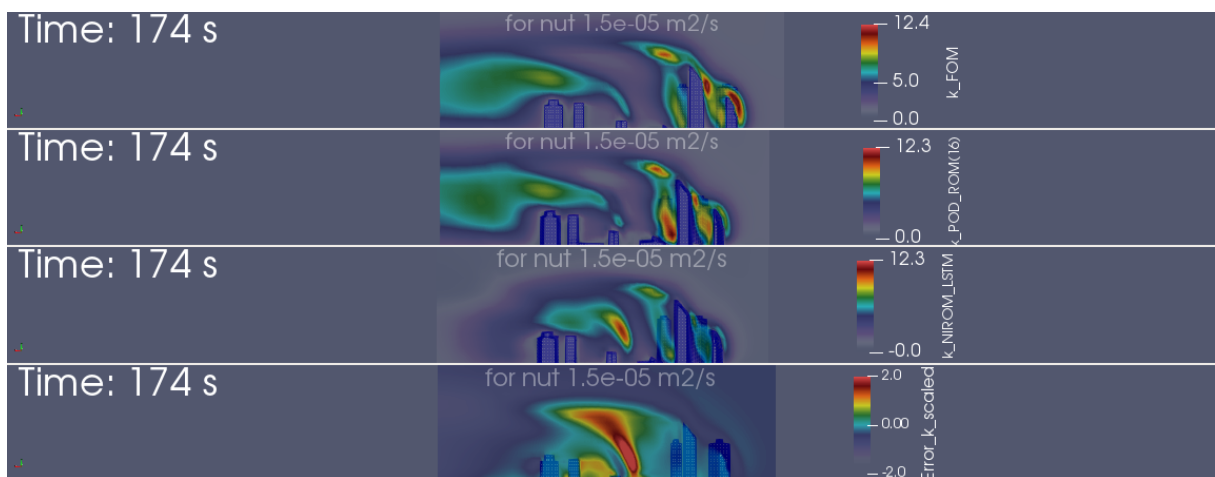
Figure 7: Magnitude of Velocity fields at different times: Comparison of FOM Vs POD-ROM Vs ROM-LG prediction for new unseen value of parameter. Last row shows the L2 norm of normalized relative error between FOM and ROM-LG predictions (normalized by FOM value).



(a) 36 s



(b) 66 s



(c) 174 s

Figure 8: Turbulent kinetic energy fields at different times: Comparison of FOM Vs POD-ROM Vs ROM-LG prediction for new unseen value of parameter. Last row shows the L2 norm of normalized relative error between FOM and ROM-LG predictions (normalized by FOM value).

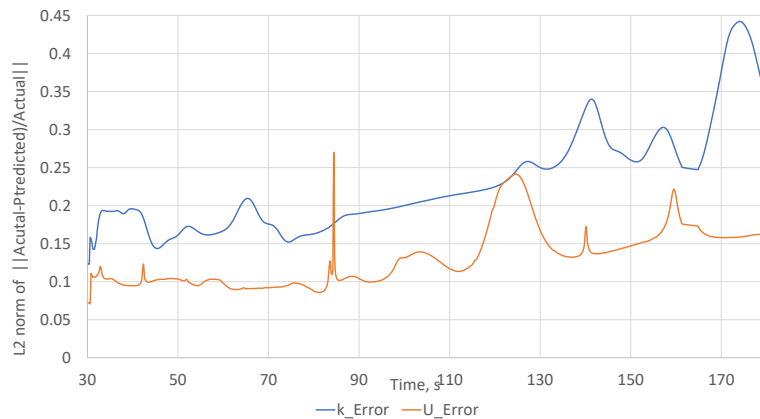


Figure 9: Temporal variation of L2 error between FOM and ROM-LG predictions.

the FOM results for short-term turbulent kinetic energy fields. Overall, considering the values of errors, the performance of the present non-intrusive methodology can be deemed decent with scope for further improvements by testing with different LSTM architectures for fields that are generally highly dynamic in nature (like TKE). Comparing the computational efficiency, the FOM takes a day to simulate the 200 s flow on 8 processor CPU set-up, whereas the current non-intrusive methodology provides the flow simulation results in an hour. Overall, the present methodology demonstrates good potential for real-life urban flows with satisfactory error range.

4. Conclusions

In this paper, a fully non-intrusive parametric ROM framework has been developed and used to capture the building-induced flows for different levels of turbulence. Based on our findings for the present case, we conclude that the ROM-LG involving LSTM and Grassmann interpolation has shown promising potential for obtaining unsteady velocity field predictions and unsteady short-term turbulent kinetic energy prediction for new values of chosen parameter. The current methodology does not rely on the governing equations to obtain the solution, which means that there are no numerical constraints while predicting the solutions. Additionally, it is computationally more efficient to predict the solution using a trained model rather than the physics-based approach of solving ODEs/PDEs. The future work can focus on improving the ROM-LG framework for long-term predictions of fields that are highly-dynamic in nature and show its utility for predicting forces on the buildings.

5. Acknowledgment

The authors acknowledge the financial support from the Norwegian Research Council for the project “ATB-Helsevel project call - Project No.: 282207” under which this work is carried out.

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] Akhtar I, Nayfeh A H and Ribbens C J 2009 *Theoretical and Computational Fluid Dynamics* **23** 213–237
- [5] Mohan A T and Gaitonde D V 2018 *arXiv preprint arXiv:1804.09269*
- [6] Pawar S, Ahmed S E, San O and Rasheed A 2020 *Physics of Fluids* **32** 036602
- [7] Ahmed S E, Pawar S, San O, Rasheed A and Tabib M 2021 *Computers Fluids* **221** 104895 ISSN 0045-7930
- [8] Amsallem D and Farhat C 2008 *AIAA Journal* **46** 1803–1813
- [9] Tabib M, Helge K, Tor S, Rasheed A and Kvamsdal T 2020 (SINTEF Academic Press) ISBN 978-82-536-1684-1