Paper:

Xiao, D., Fang, F., Zheng, J., Pain, C. & Navon, I. (2019). Machine learning-based rapid response tools for regional air pollution modelling. *Atmospheric Environment*

http://dx.doi.org/10.1016/j.atmosenv.2018.11.051

Machine learning-based rapid response tools for regional air pollution modelling

D. Xiao[a,b,e], F. Fang[a,e,*], J. Zheng[c], C.C.Pain[a,e], I.M.Navon[d]

[a]*Applied Modelling and Computation Group,*
Department of Earth Science and Engineering,
Imperial College London,
Prince Consort Road, London, SW7 2BP, UK.
URL: `http://amcg.ese.imperial.ac.uk`
[b]*ZCCE College of Engineering, Swansea University,*
Bay Campus, Fabian Way, Swansea SA1 8EN, UK
[c]*Institute of Urban Environment,Chinese Academy of Sciences.*
[d]*Department of Scientific Computing,*
Florida State University,
Tallahassee, FL, 32306-4120, USA
[e]*Data Assimilation Laboratory, Data Science Institute, Imperial College London*

Abstract

A parameterised non-intrusive reduced order model (P-NIROM) based on proper orthogonal decomposition (POD) and machine learning methods has been firstly developed for model reduction of pollutant transport equations. Our motivation is to provide rapid response urban air pollution predictions and controls. The varying parameters in the P-NIROM are pollutant sources. The training data sets are obtained from the high fidelity modelling solutions (called snapshots) for selected parameters (pollutant sources, here) over the parameter space $\mathcal{R}^P$. From these training data sets, the machine learning method is used to generate the relationship between the reduced solutions and inputs (pollutant sources) over $\mathcal{R}^P$. Furthermore a set of of hyper-surface functions associated with each POD basis function is constructed for representing the fluid dynamics over the reduced space. The accuracy of the P-NIROM is highly dependent on the quality of the training set, here obtained from the high fidelity model. Over existing machine learning methods, the P-NIROM algorithm proposed here has the advantages that (1) it is combined with NIROM, thus providing rapid and reasonably accurate solutions; and (2) it is a robust and efficient approach for representation of any parametrised partial differential equations as the model parameters/inputs

*Corresponding author
Email address: `f.fang@imperial.ac.uk` (F. Fang)

vary. In this study, we demonstrate the way how to implement the P-NIROM for the pollutant transport equation (but not limited to due to its robustness). Its predictive capability is illustrated in a three-dimensional (3-D) simulation of power plant plumes over a large region in China, where the varying parameters are the emission intensity at three locations. Results indicate that in comparison to the high fidelity model, the CPU cost is reduced by factor up to five orders of magnitude while reasonable accuracy remains.

Keywords: Machine learning, Finite Element, Proper orthogonal decomposition, Reduced order modelling, air pollution.

1. Introduction

Pollution in cities has a strong impact on the health of communities and affects global warming with dire consequences to humanity. The dynamic and pollutant transport processes involve a wide range of scales. The highly disparate scale poses a formidable challenge for atmospheric and air pollution modelling. In recent years, the spatial resolution in operation air pollution models has been increased significantly, thus improving predictive capability. However, this unavoidably leads to an increase in computational cost [1]. Our motivation is to develop numerical tools for rapid responses/predictions of pollutants without sacrificing solution accuracy, especially in emergency situations.

Reduced-order models (ROMs) have become important to many fields as they offer the potential to simulate dynamical systems with considerably reduced computational cost in comparison to high fidelity models [2, 3, 4, 5]. Recently, reduced order methods have been applied to studies of air pollution [6, 7, 8, 9]. Existing ROMs can be classified into two categories: intrusive and non-intrusive approaches in the sense that whether the implementation of ROMs requires knowledge of the details of original numerical source codes [10]. The intrusive reduced order methods have been widely used in many fields [11, 12, 13, 14, 15, 16, 17, 18]. More recently, the non-intrusive methods have became popular since they are less dependent on complex dynamic systems and are therefore easy to implement even when the numerical source code is not available. Existing non-intrusive methods used for generating ROMs are POD in combination with radial basis function (RBF), Smolyak sparse grid and artificial neural networks etc.[19, 20, 21, 22, 23, 24]. The applications of NIROMs can be found in the work of [25, 26, 27, 28].

More recently, Wang *et al.* introduced a deep learning technique to NIROMs and applied it to fluid problems [29]. Deep learning technologies represent the most recent progress in artificial neural networks [30], and have been applied to a number of areas such as speech recognition [31], image recognition [32], medical science [33], self-driving cars [34], language understanding [35] etc.

In this work, we have developed a Parameterised NIROM (P-NIROM) based on machine learning techniques for parameterised pollutant transport problems. The input parameters are the emission intensity of pollutants released at different source locations. The P-NIROM enables rapid simulations and controls of the impact of pollutant sources without excessive computational costs. Given a set of selected pollutant sources $\mathbf{Q}_{tr}$ over the parameterised space R^P, the training data sets (also called solution snapshots) can be obtained by running the high fidelity model. From the snapshot solutions, the corresponding reduced basis functions are calculated using singular value decomposition (SVD)/POD. The reduced basis functions are used for constructing the reduced space. The original high fidelity model can be projected onto the reduced space, which is several orders of magnitude smaller than the dimensional size of the high fidelity full model, thus significantly reducing the computational cost. For any unseen emission intensity of pollutant sources $\mathbf{Q} \in \mathcal{R}^P$, the P-NIROM is constructed using the machine learning methods. From the training solution snapshots, a Gaussian process is used for generating the snapshots and POD basis functions for the unseen pollutant sources $\mathbf{Q}$. Furthermore, the relationship (P-NIROM) between the reduced solutions and the inputs (the pollutant emission intensities) can be obtained using the machine learning techniques. Finally, the solutions from the P-NIROM are projected back the full space.

The P-NIROM is a robust and efficient numerical tool for rapid prediction of pollutants released from different sources and assessment of their impact on specified cities/locations. In this work, we have been successfully applied the P-NIROM to air pollution simulations over a large region in China which covers 55 cities including Beijing. The efficiency and accuracy of the P-NIROM have been evaluated by comparing the results with those from the high fidelity full model.

The remainder of this article is arranged as follows. The pollutant transport equation and its discretisation are described in section 2. In section 3, the details of forming the P-NIROM using POD and machine learning methods are provided. Section 4 presents a numerical experiment of simulating the spatial and temporary distribution of pollutants released from 100 power plants in China. Conclusions are drawn in section 5.

2. Pollutant transport equation and its discretisation

The dispersion of the tracer concentration (c) is modelled by:

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c + \nabla \cdot (\kappa \nabla c) - Q = 0, \tag{1}$$

where $\mathbf{u}$ is the velocity vector, Q is a source term and κ the diffusivity. In general, the discretised form of (1) at each time level n (where a time interval of Δt is set during the simulation period) can be written:

$$\mathcal{M}(\mu)\mathbf{c}^n = \mathbf{s}^n(\mu, \mathbf{c}^{n-1}), \tag{2}$$

where $\mathcal{M}$ is the full numerical operator with varying input parameter μ, $\mathbf{c}^n = (c_1^n \ldots, c_j^n, \ldots, c_{\mathcal{N}}^n)^T$ $(1 \leq j \leq \mathcal{N}$, $\mathcal{N}$ is the number of nodes in the computational domain), $\mathbf{s}^s$ includes the source term, boundary conditions and the variable solutions from the previous time level. In this study, the varying input parameters in air pollutant problems are set to be the pollutant sources, $\mu = \mathbf{Q} = (Q_1, \ldots, Q_s, \ldots, Q_S)$ (here, S is the number of pollutant sources).

2.1. Parameterised reduced order transport equation

In this work, the POD approach in combination with machine learning techniques is used for model reduction. POD has proven to be a powerful tool for circumventing the intensive computational burden in large complex numerical simulations. POD is capable of representing large complex dynamical systems using a few number of optimal basis functions. In POD reduced order modelling, the tracer concentration in (2) can be expressed as an expansion of the POD basis functions $\Phi = (\Phi_1, \ldots, \Phi_m, \ldots, \Phi_M)$:

$$\mathbf{c}^n = \Phi \mathbf{c}^{r,n}, \tag{3}$$

where $\mathbf{c}^{r,n} = (c_1^{r,n} \ldots, c_m^{r,n}, \ldots, c_M^{r,n})^T$ $(1 \leq m \leq M) \in \mathcal{R}^M$ is the reduced state variable vector (the superscript r indicates the variable associated with the reduced order model) to be determined over the reduced space. The POD basis functions are constructed from a collection of snapshots that are taken from the high fidelity model solution (2) for the selected training pollutant sources. Using SVD, a set of orthogonal basis functions $\{\Phi_m\}$ can be obtained in an optimal way. The POD basis functions can represent the dynamics of snapshot solutions. The loss of information due to the truncation of the POD expansion set to M vectors can be quantified by the following ratio,

$$E = \frac{\sum_{j=1}^{M} \lambda_j^2}{\sum_{j=1}^{I} \lambda_j^2}, \tag{4}$$

where λ denotes eigenvalues, and I is the total number of eigenvectors (here equivalent to the number of solution snapshots used for generating the POD basis functions). The value of E will tend to 1 as M is increased to the value I, this would imply no loss of information. A few number of leading eigenvectors can represent most of dynamical energy within the solution snapshots.

Projecting (2) from the $\mathcal{N}$ dimensional space onto the M dimensional reduced space ($M << \mathcal{N}$), yields:

$$\Phi^T \mathcal{M}(\mu)\mathbf{c}^n = \Phi^T \mathbf{s}^n(\mu, \mathbf{c}^{n-1}). \tag{5}$$

The parameterised reduced order model can thus be written as:

$$\mathcal{M}^r(\mu)\mathbf{c}^{r,n} = \mathbf{s}^{r,n}(\mu, \mathbf{c}^{r,n-1}), \tag{6}$$

where $\mathcal{M}^r_{t_0,t_j} = \Phi^T \mathcal{M}(\mu)$ is the model operator over the reduced space, $\mathbf{c}^{r,n} = \Phi^T \mathbf{c}^n$ and $\mathbf{s}^{r,n} = \Phi^T \mathbf{s}^n(\mu, \mathbf{c}^{r,n-1})$.

Equations (3) and (6) can be used for efficient air pollution operational prediction where the CPU time can be reduced by several orders of magnitude. In this work, the parameter set μ in (6) consists of the pollutant source inputs. A recently developed NIROM [29] is extended to construct the parameterised reduced order model in (6). The P-NIROM based on the machine learning techniques described below is capable of predicting problems with unseen or different parameters (for example, unseen pollutant sources). It is also non-intrusive and independent of the original source code.

3. Construction of P-NIROM based on POD and machine learning methods

The parameterised reduced order model (6) is re-written for the variable c^r_m associated with each POD basis function Φ_m over the reduced space in a general form:

$$c^{r,n}_m = \mathcal{F}_m(\mu, \mathbf{c}^{r,n-1}), \quad m \in (1, \ldots, M). \tag{7}$$

In non-intrusive reduced order modelling, one searches a set of functions $\mathcal{F}_m$ to represent the dynamics in (7). In this work, we introduce the Gaussian process regression (GPR) [36] and deep learning learning methods [29] to construct the relationship functions $\mathcal{F}_m$ to represent the fluid dynamics of system (6) for any unseen input parameter $\mu = \mathbf{Q} \in \mathcal{R}^P$.

3.1. Gaussian process regression for calculation of POD coefficient and snapshot solutions for any input over the parameter space

In GPR, the relationship between the input $\mu = \mathbf{Q}$ (here, the pollutant source) and the corresponding output $\mathbf{c}^n$ at each time level $n \in (1, 2, \ldots, T_{tr})$ can be expressed as follows [36]:

$$\mathbf{c}^n(\mathbf{Q}) = g^n(\mathbf{Q}) + \epsilon(\mathbf{Q}), \tag{8}$$

where, $\epsilon = \mathcal{G}(0, \sigma_n)$ is the Gaussian distribution with zero mean and variance σ_n.

In GPR, it is assumed that the function $g^n(\mathbf{Q})$ has a Gaussian distribution (with zero mean, here):

$$g^n(\mathbf{Q}) \sim \mathcal{G}\left(\mathbf{0}, k^n(\mathbf{Q}, \mathbf{Q}')\right), \tag{9}$$

where the covariance function $k^n(\mathbf{Q}, \mathbf{Q}')$ represents the dependency between the function values at two different input points $\mathbf{Q}$ and $\mathbf{Q}'$, that is,

$$cov\left(g^n(\mathbf{Q}), g^n(\mathbf{Q}')\right) = k^n(\mathbf{Q}, \mathbf{Q}') = \sigma_{w^n} exp(-\frac{1}{2l}|\mathbf{Q} - \mathbf{Q}'|), \tag{10}$$

where, l is the length scale and σ_{w^n} is the variance. The correlation between the functions $g^n(\mathbf{Q})$ and $g^n(\mathbf{Q}')$ is dependent on the distance between the two input points. Given a set of training input-output pairs $\{\mathbf{Q}_{tr,i}, \mathbf{c}^n_{tr,i}\}, i \in (1, \ldots, N_{tr})$ (where, N_{tr} is the number of training points), one aims to predict the pollutant concentration $\mathbf{c}^n$ in (8) for any new input $\mathbf{Q}$. The joint Gaussian distribution of the training and predicted outputs ($\mathbf{c}^n_{tr}$ and $\mathbf{c}^n$) for the training and new inputs ($\mathbf{Q}_{tr}$ and $\mathbf{Q}$) respectively can be written:

$$\begin{bmatrix} \mathbf{c}^n_{tr} \\ \mathbf{c}^n \end{bmatrix} = \begin{bmatrix} g^n(\mathbf{Q}_{tr}) \\ g^n(\mathbf{Q}) \end{bmatrix} \sim \mathcal{N}\left(0 \quad \begin{bmatrix} K^n(\mathbf{Q}_{tr}, \mathbf{Q}_{tr}) & K^{nT}(\mathbf{Q}, \mathbf{Q}_{tr}) \\ K^n(\mathbf{Q}, \mathbf{Q}_{tr}) & K^n(\mathbf{Q}, \mathbf{Q}) \end{bmatrix}\right), \tag{11}$$

where, $K^n(\mathbf{Q}_{tr}, \mathbf{Q}_{tr})$ is the covariance matrix between all training points and is written below:

$$K^n(\mathbf{Q}_{tr}, \mathbf{Q}_{tr}) = \begin{bmatrix} k^n(\mathbf{Q}_{tr,1}, \mathbf{Q}_{tr,1}) & k^n(\mathbf{Q}_{tr,1}, \mathbf{Q}_{tr,2}) & \dots & k^n(\mathbf{Q}_{tr,1}, \mathbf{Q}_{tr,N_{tr}}) \\ k^n(\mathbf{Q}_{tr,2}, \mathbf{Q}_{tr,1}) & k^n(\mathbf{Q}_{tr,2}, \mathbf{Q}_{tr,2}) & \dots & k^n(\mathbf{Q}_{tr,2}, \mathbf{Q}_{tr,N_{tr}}) \\ \vdots & \vdots & \ddots & \vdots \\ k^n(\mathbf{Q}_{tr,N_{tr}}, \mathbf{Q}_{tr,1}) & k^n(\mathbf{Q}_{tr,N_{tr}}, \mathbf{Q}_{tr,N_{tr}}) & \dots & k^n(\mathbf{Q}_{tr,N_{tr}}, \mathbf{Q}_{tr,N_{tr}}) \end{bmatrix}, \tag{12}$$

and the matrices

$$K^n(\mathbf{Q}_{tr}, \mathbf{Q}) = \begin{bmatrix} k^n(\mathbf{Q}_{tr,1}, \mathbf{Q}) & k^n(\mathbf{Q}_{tr,2}, \mathbf{Q}) & \dots & k^n(\mathbf{Q}_{tr,N_{tr}}, \mathbf{Q}) \end{bmatrix}, \tag{13}$$

$$K^n(\mathbf{Q}, \mathbf{Q}) = k^n(\mathbf{Q}, \mathbf{Q}). \tag{14}$$

Given a set of the training inputs (here, the pollutant sources) $\mu_{tr} = \mathbf{Q}_{tr} = (Q_{tr,1}, \dots, Q_{tr,S})$ over the parameter space R^P, the snapshot solutions $\mathbf{c}_{tr} = (\mathbf{c}^1_{tr}, \dots, \mathbf{c}^n_{tr}, \dots, \mathbf{c}^{N_t}_{tr})$ can be obtained by running the high fidelity model (2) during the training simulation period $[0, T_{tr}]$.

For efficient calculations, one can project $\mathbf{c}^n_{tr}$ from the high dimensional full space onto the reduced space:

$$\mathbf{c}^{r,n}_{tr} = \Phi^T \mathbf{c}^n_{tr}, \tag{15}$$

For any given input parameter (pollution source $\mathbf{Q}$), the probability of the prediction of the reduced variable $\mathbf{c}^r$ is:

$$\mathbf{c}^{r,n} | \mathbf{c}^{r,n}_{tr} \sim \mathcal{N}(K^n_* K^{n\,-1}_{tr} \mathbf{c}^{r,n}_{tr}, K^n_{**} - K^n_* K^{n\,-1}_{tr} K^{n\,T}_*), \tag{16}$$

where, $K^n_* = K^n(\mathbf{Q}, \mathbf{Q}_{tr})$, $K^n_{**} = K^n(\mathbf{Q}, \mathbf{Q})$ and $K^n_{tr} = K^n(\mathbf{Q}_{tr}, \mathbf{Q}_{tr})$. The best estimate of $\mathbf{c}^{r,n}$ is the mean of the Gaussian distribution:

$$\bar{\mathbf{c}}^{r,n} = K^n_* K^{n\,-1}_{tr} \mathbf{c}^{n,r}_{tr}. \tag{17}$$

3.2. Deep learning method for construction of P-NIROM and calculation of reduced solutions for any input over the parameter space

In this section, an alternative method for calculation of reduced solutions for any given input is introduced. A Recurrent Neural Network (RNN) using the Long Short Term Memory (LSTM) architecture is used to construct the P-NIROM (7). Compared to traditional RNNs, the LSTM has a special memory block in the hidden layer of the recurrent neural network, allowing information to persist. This type of network has cyclic connections, which makes the network a powerful method to model temporal data since it has an internal memory system to deal with temporal sequence inputs. A memory cell is composed of four main elements: an input gate, a neuron with a self-recurrent connection (a connection to itself), a forget gate and an output gate.

The input gate of each memory block controls the information transmitting from the input activations into the memory cell and the output gate controls the information transmitting from the memory cell activations into other nodes. The forget gate decides what information is to be deleted from the memory cell state [29].

The LSTM technique is utilised to construct the set of functions (hyper-surfaces) F_m in (7). In the LSTM network, the input is the reduced solution $\mathbf{c}^{r,n-1} = (c_1^{r,n-1}, ..., c_M^{r,n-1})$ at the previous time level $n-1$ while the output is the reduced solution $c_m^{r,n}$ associated with the m^{th} POD basis function Φ_m ($m \in (1, \dots, M)$). The relationship function (hyper-surface F_m) between the input $\mathbf{c}^{r,n-1}$ and output $c_m^{r,n}$ can be obtained using the following equations:

$$
\begin{aligned}
I^n &= \varrho(W_{ic}\mathbf{c}^{r,n-1} + W_{ih}h^{n-1} + W_{iCe}Ce^{n-1} + b_i),\\
f^n &= \varrho(W_{fc}\mathbf{c}^{r,n-1} + W_{fh}h^{n-1} + W_{fCe}Ce^{n-1} + b_r),\\
o^n &= \varrho(W_{oc}\mathbf{c}^{r,n-1} + W_{oh}h^{n-1} + W_{oCe}Ce^{n} + b_o),\\
Ce^n = r^n \odot Ce^{n-1} + i^n \odot Ce_i(W_{Cec}&\mathbf{c}^{r,n-1} + W_{Ceh}h^{n-1} + b_{Ce}),\\
h^n &= o^n \odot Ce_o(Ce^n),\\
c_m^{r,n} &= \zeta(W_{rh}h^n + b_r),
\end{aligned}
\tag{18}
$$

where I, f and o denote the input, forget and output gate vectors respectively, Ce is the cell activation vector, b is the bias vector, ϱ is the activation function, W denotes the weight matrices (*e.g.* W_{ic} is the weight matrix from the input gate to the input), $\odot$ is the element wise product of the vectors, Ce_o and Ce_i are the cell output and cell input activation functions respectively and ζ is the network output activation function.

After obtaining the function F_m, it can then be used to predict the POD coefficients at current time level n. The offline calculation of snapshots at the training stage and the online procedure for constructing and resolving the P-NIROM can be algorithmically summarised in Figure 1. The details of the offline and on-line calculations are further given in Algorithm 1 and 2 respectively.

Algorithm 1: Offline Calculations

---

(1) Select a set of training inputs (here the emission intensity) $\mathbf{Q}_{tr,i} \in R^P$, where, $i \in (1, \ldots, N_{tr})$;

(2) Given the input $\mathbf{Q}_{tr,i}$, generate the solution snapshots by running the high fidelity full model during the training period $[0, T_{tr}]$

for $i = 1$ *to* N_{tr} **do**

  for $n = 1$ *to* T_{tr} **do**

    Solving Equation (2):

    $\mathcal{M}(\mathbf{Q}_{tr,i})\mathbf{c}^n = \mathbf{s}^n(\mathbf{Q}_{tr,i}, \mathbf{c}^{n-1})$

  endfor

endfor

(3) Calculate the POD basis functions using SVD.

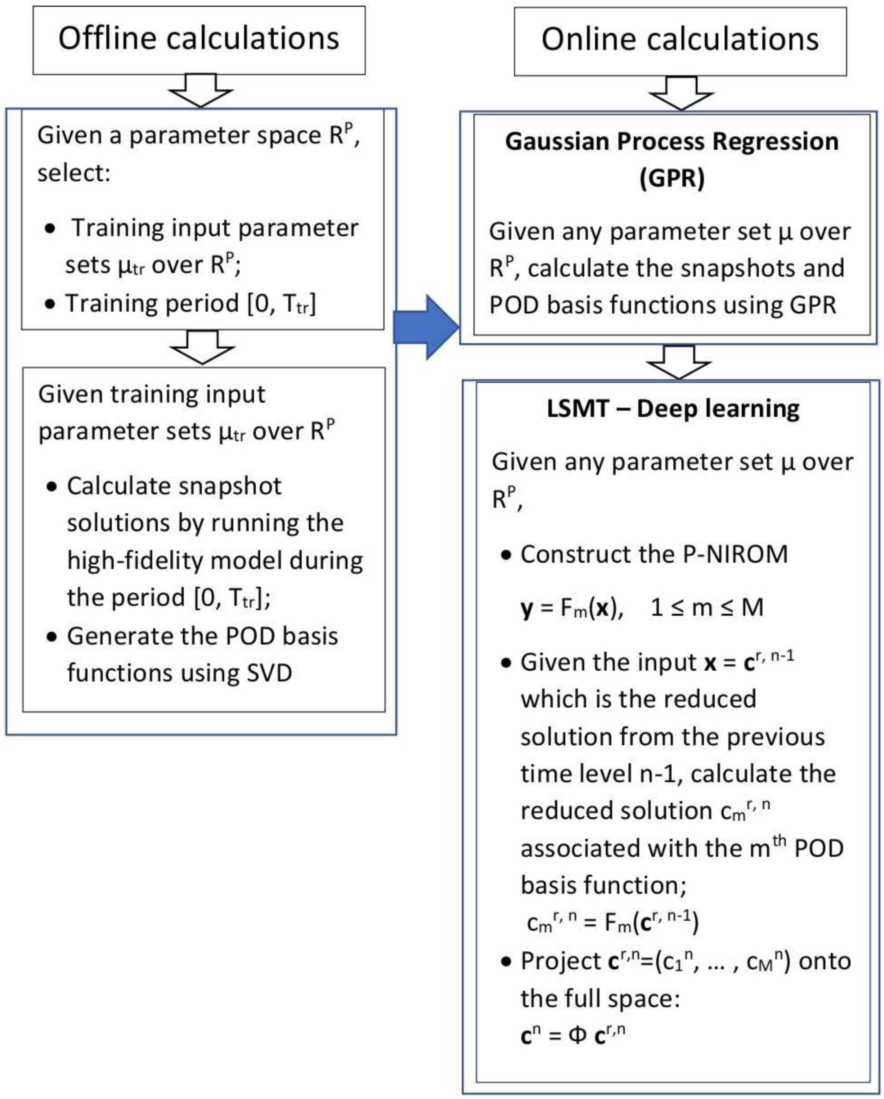


Figure 1: The figure displayed above shows the online and offline procedures of constructing and resolving the P-NIROM for any given parameter $\mu \in \mathcal{R}^{\mathcal{P}}$.

Algorithm 2: Online Calculation

(1) Using GPR, for any given unseen input $\mu = \mathbf{Q}$ (here the pollutant emission intensity), calculate the snapshot solutions

  (i) Calculate the covariance matrices between the given unseen ($\mathbf{Q}$) and training ($\mathbf{Q}_{tr}$) points in (12) - (14);

  (ii) For the given unseen emission intensity $\mathbf{Q}$, calculate the probability of the prediction of the tracer variable $\mathbf{c}^r$ in (16) over the reduced space;

  (iii) Calculate the best estimate of solutions over the reduced space using (17), then project back the full space.

(2) Using LSTM, construct the set of P-NIROMs and calculate reduced solutions for any input over the parameter space;

  (i) Using (18), construct the set of P-NIROMs F_m $(m \in (1, \ldots, M))$ for the associated POD basis function Φ_m;

  (ii) Calculate the solutions at time level n using the P-NIROM.

    (a) Give the reduced solution $\mathbf{c}^{r,n-1}$ at the previous time level $n-1$;

    (b) Calculate the solution $\mathbf{c}^{r,n} = (c_1^{r,n}, \ldots, c_m^{r,n}, \ldots, c_M^{r,n})$ at time level n over the reduced space
    for $m = 1$ *to* M **do**

$$c_m^{r,n} = F_m(\mu, \mathbf{c}^{r,n-1})$$

    endfor

    (c) Obtain the solution $\mathbf{c}^n(\mu)$ at the current time level n by projecting $\mathbf{c}^{r,n}(\mu)$ onto the full space via:

$$\mathbf{c}^n(\mu) = \sum_{m=1}^{M} c_m^{r,n} \Phi_m$$

4. Regional pollutant dispersion in China

To demonstrate the capability of the new P-NIROM based on machine learning techniques, it has been applied to a realistic case in China where the SO_2 emissions from power plants disperse through the atmosphere in time. The SO_2 emission intensity at the power plant locations was obtained from the Regional Emission inventory in ASia (REAS 2.1) data developed by National Institute of Environmental Sciences of Japan. The simulated domain covers the whole Shanxi-Hebei-Shandong-Henan region of China with an area encompassing $1090km \times 1060km$, and there are about 100 power plants in this area [37].

Using adaptive mesh techniques, the 2D top adaptive mesh ($20\,km$ above the sea level) is first constructed to ensure a high resolution of $2.5\,km$ around the power plan points within a radius of 6 km. The 3D unstructured mesh with 61479 nodes is then obtained by extending the 2D top mesh onto the terrain surface, with 11 terrain-following layers, where 7 vertical layers are within $1\,km$ above the terrain. The pollutant SO_2 sources around the power plants are released into the atmosphere at the hight of $200\,m$ above the terrain.

In the study, the simulation started at 00:00 UTC on the 10 January 2013 and ran through to the 15 January 2013. A time interval of $\Delta t = 0.5\,hr$ was used. Assuming that the mixing layer height is $600\,m$ and the turbulent horizontal diffusivity is $100\,m^2/s$ while the vertical eddy diffusivity is parameterised based on a scheme by Byun and Dennis [38]. The meteorological fields are provided by the mesoscale meteorological model WRF (v3.5) [39].

In this case, the varying input parameter, $\mu = \mathbf{Q}$, is the emission intensity of pollutant sources at locations ξ_1, ξ_2 and ξ_3 (see Table 1). The emission intensity of pollutant sources is ranged from 0 to 5000 $g\,s^{-1}$. A set of training pollutant sources $\mu_{tr} = \mathbf{Q}_{tr}$ at three locations is listed in Table 1. The solution snapshots $\mathbf{c}_{tr}$ with the training parameters were obtained by running the high fidelity model (Fluidity [40]) and stored at equally spaced time intervals ($3\,hrs$) during the simulation period (5 days).

To illustrate the capability of the P-NIROM based on machine learning techniques, an unseen test case, the emission intensity of pollutant sources $\mu = \mathbf{Q} = (2400, 2400, 5000)\,g\,s^{-1}$, was given at locations ξ_1, ξ_2 and ξ_3 respectively (T_1 in Table 1). Following the online procedure shown in Figure 1, using the GPR, the solution snapshots (the distribution of pollutants at every $3\,hrs$) for the given unseen pollutant sources were calculated from the

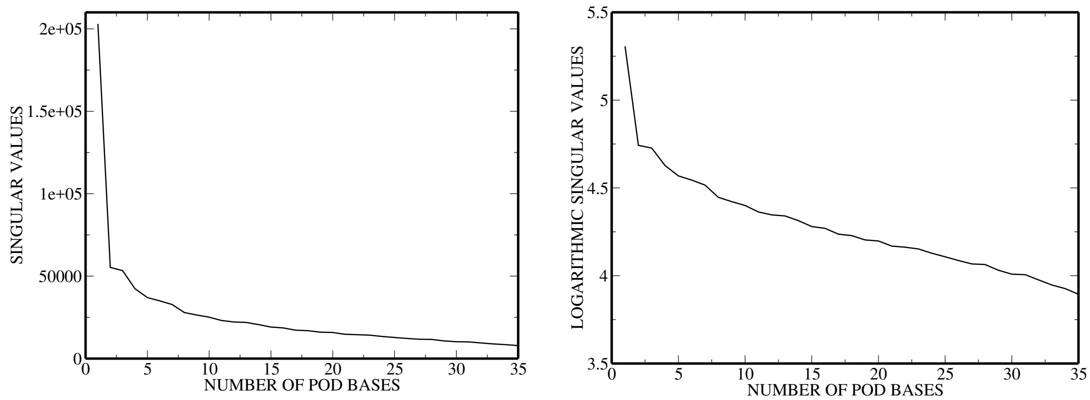
Table 1: The emission intensity ($g\,s^{-1}$) of SO_2 at locations ξ_1 (x=540,y=752) km, ξ_2 (x=603, y=670) km and ξ_3 (x=753, y=679) km. $A_1 - A_{28}$ are the training cases while $T_1 - T_2$ are the unseen cases used for evaluating the predictive capability of the new P-NIROM.

Cases	ξ_1	ξ_2	ξ_3	cases	ξ_1	ξ_2	ξ_3	cases	ξ_1	ξ_2	ξ_3
A1	1047	1678	1160	A11	4267	2500	2500	A21	1440	1140	3570
A2	0	0	0	A12	2500	732	2500	A22	1250	1250	1250
A3	5000	5000	5000	A13	2500	4267	2500	A23	3750	3750	3750
A4	2500	2500	2500	A14	2500	2500	732	A24	0	5000	5000
A5	0	2500	2500	A15	2500	2500	4267	A25	1250	5000	5000
A6	5000	2500	2500	A16	534	589	910	A26	5000	5000	0
A7	2500	5000	2500	A17	600	639	1580	A27	5000	0	0
A8	2500	2500	0	A18	181	1061	1356	A28	2500	0	2500
A9	2500	2500	5000	A19	428	1881	329				
A10	732	2500	2500	A20	1300	1380	3000				
T1	2400	2400	5000	T2	5500	6000	6000				

training solutions for the selected training parameters (28 training parameter sets in Table 1).

Figure 2 shows the singular values and a logarithmic scale of singular values. From the calculation in (4), the sharp decrease of singular values suggests that the first 36 leading POD basis functions can capture 99% of dynamical energy within the solution snapshots. In this study, two cases of 6 and 36 POD basis functions were chosen to construct the P-NIROM. The larger the number of POD basis functions chosen, the higher the accuracy of P-NIROM. Figure 4 provides some of the first 36 leading basis functions. It can be seen that the first leading basis function captures a large part of the spatial distribution of pollutant concentration solutions, while the remaining basis functions represent the details of pollutant distributions of different regions.

A comparison of coefficients for the POD basis functions between using the standard ROM and machine learning ROM (based on LSTM and GPR) is provided in Figure 3. It is clearly seen that the POD coefficients are in very close agreement with each other. Compared to the standard ROM, the machine learning ROM has a wider range of application areas, especially where observational data is concerned, for example, data assimilation, data reduction by condensing the information into the required dynamical



(a) Singular values (b) Logarithmic scale of singular values

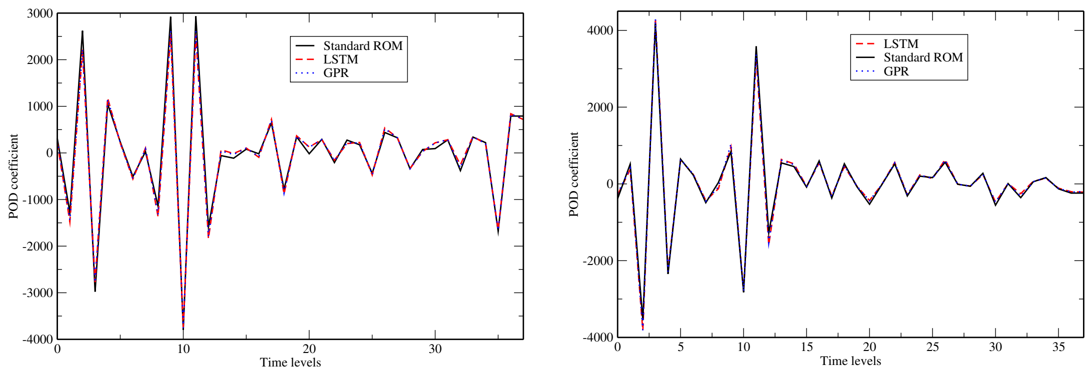
Figure 2: The singular values and logarithmic scale of singular values.

features.

Figure 5 presents the spatial distribution of pollutant solutions at time levels $t = 30\,hrs$ and $t = 105\,hrs$, as calculated by the fidelity model and P-NIROM with 6 and 36 POD basis functions. It is illustrated that P-NIROM with 6 POD basis captures most of the details of pollutant distribution at time level $t = 105\,hrs$, but fails at time level $t = 30\,hrs$. With an increased number of 36 POD basis functions, the P-NIROM has performed well at resolving the flow dynamics and evolution of power plant plumes (see Figure 5(e) and (d)). This is further highlighted in Figure 6 which shows the solutions from different angles. Further comparison is provided in Figure 7 which illustrates the evolution of pollutant concentrations predicted by the fidelity model and P-NIROM at the location $(x = 379, y = 786)\,km$. We can see that the P-NIROM with 6 and 36 POD basis function is in close agreement with the high fidelity model at this location.

An error analysis of P-NIROM has been carried out. Visual inspection of Figure 8 shows the spatial distribution of absolution errors of pollutant solutions between the high fidelity model and P-NIROM. It is visually evident that the accuracy of P-NIROM solutions is improved by increasing the number of retained POD basis functions from 6 to 36. Figure 9 illustrates the RMSE and correlation coefficients of pollutant solutions between the high fidelity model and P-NIROM with 36 POD basis functions. The correlation coefficients achieve results above $80\% - 90\%$. This again demonstrates that the P-NIROM is in good agreement with the high fidelity full model.

To further investigate the predictive ability of the P-NIROM, another

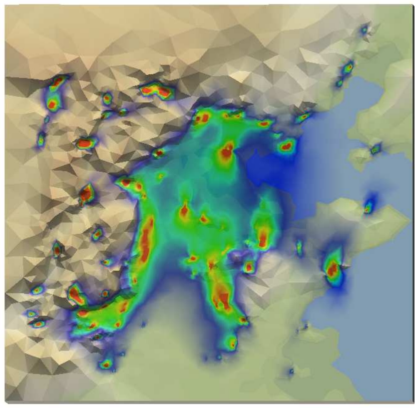


(a) First POD coefficient

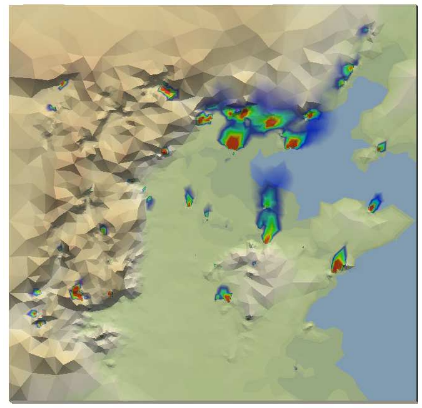
(b) Second POD coefficient

Figure 3: The first and second POD coefficients obtained from the standard ROM and machine learning ROM (the black solid line: standard ROM, the red dash line: LSTM-ROM, and the blue dot line: GPR-ROM.

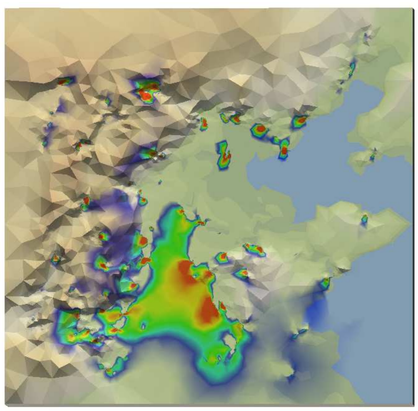
unseen cases (T2) was set up, where the emission intensities of pollutants at three source locations ($\mu = \mathbf{Q} = (5500, 6000, 6000)\, g\, s^{-1}$, see Table1) were given beyond the range of the training data $(0, 5000)\, g\, s^{-1}$. The pollutant solutions (at time level $t = 24\, hrs$) from the high fidelity full model and P-NIROM are shown in Figures 10 (a) and (b) respectively while the corresponding absolute error is illustrated in Figure 10 (d). A comparison of results between the high fidelity full model and P-NIROMs at a particular location $(x = 599, y = 569)\, km$ is provided in Figures 10 (c). As shown in the figures, the predictive ability of the P-NIROM in cases T2 is acceptable although the given test data goes beyond the range of the training data.



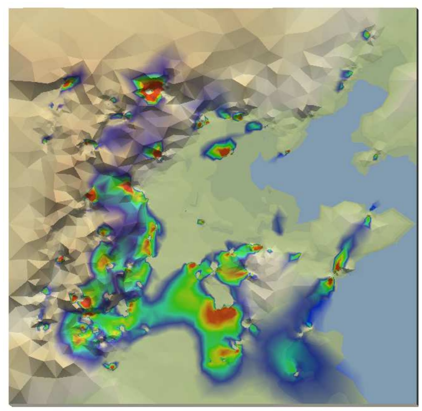
(a) 1st basis function



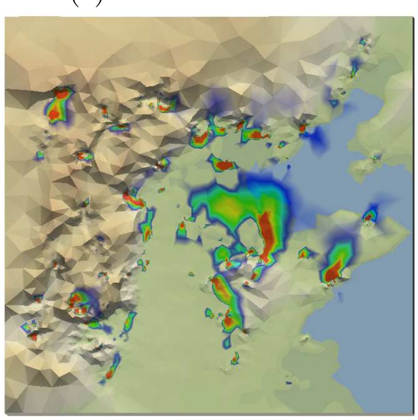
(b) 2nd basis function



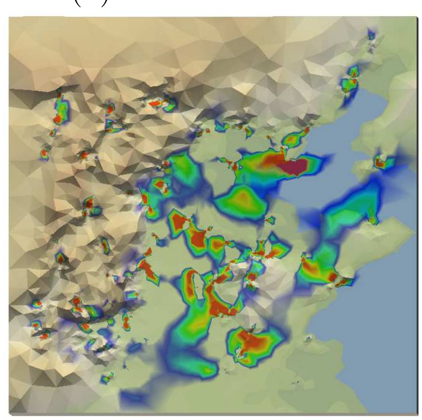
(c) 3rd basis functions



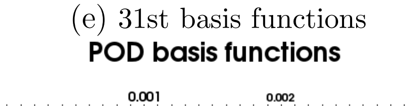
(d) 4th basis functions



(e) 31st basis functions



(f) 36th basis functions



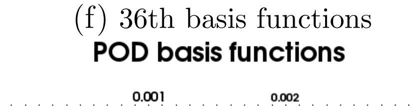
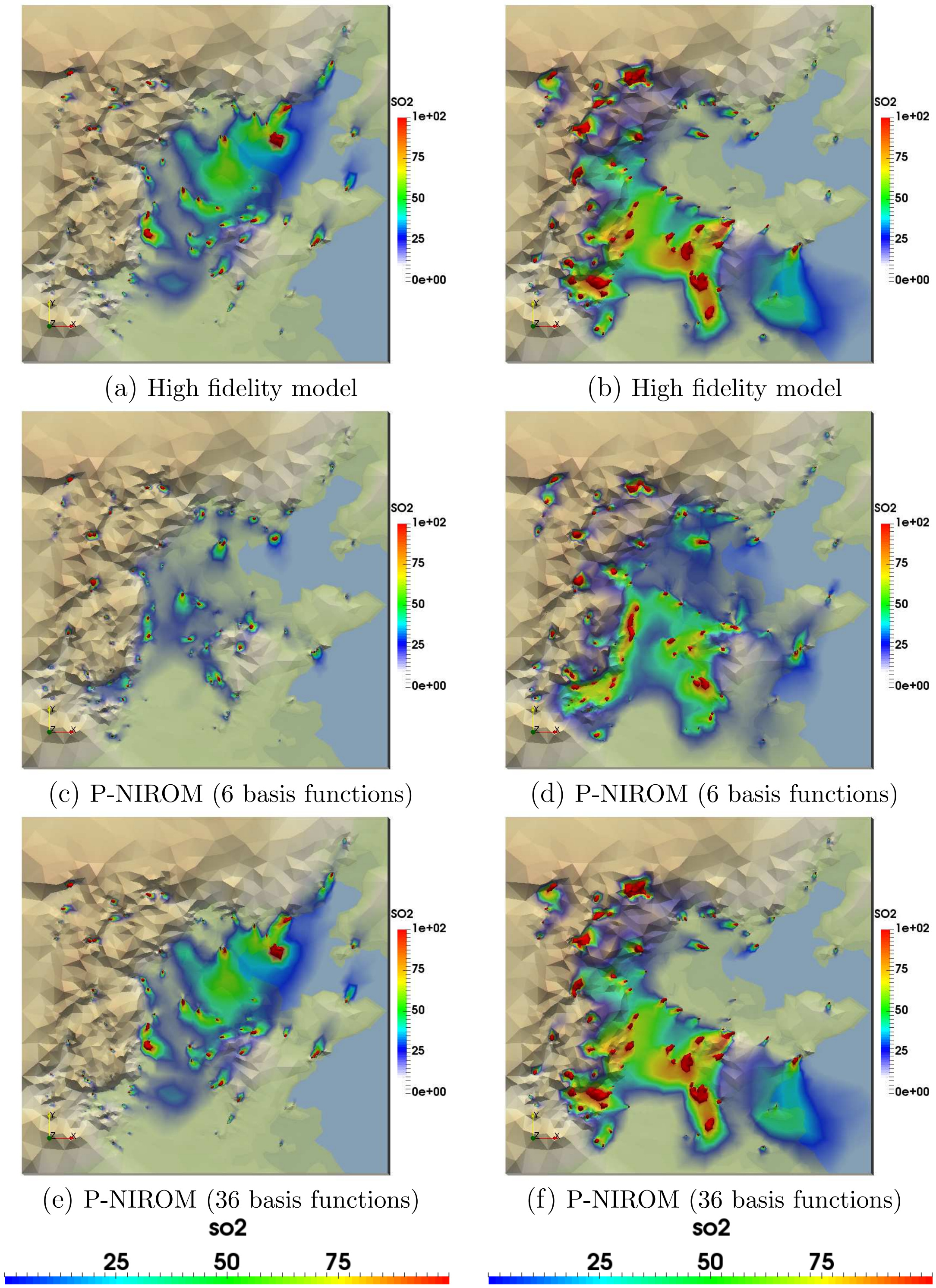


Figure 4: Some of the first 36 leading POD basis functions



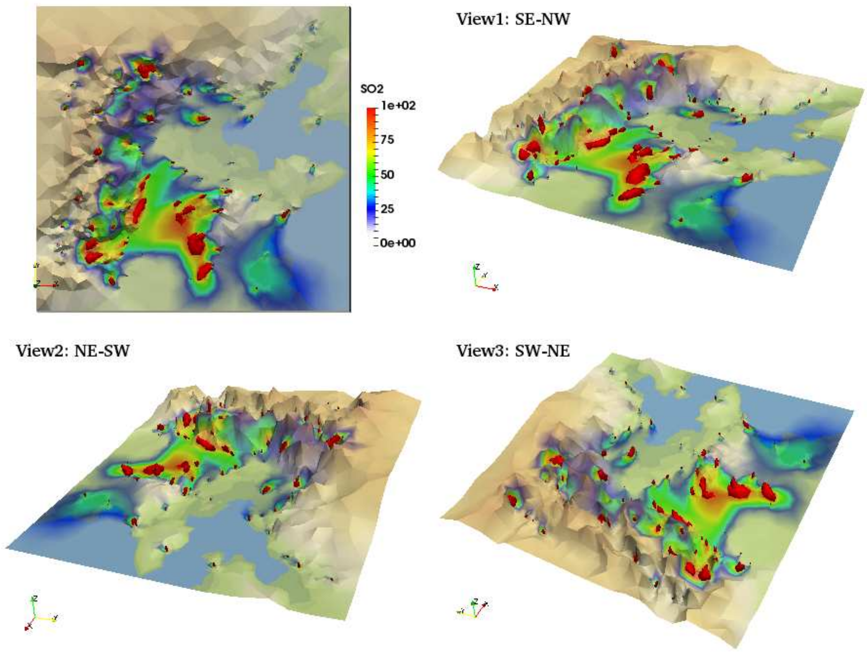
(a) High fidelity model (b) High fidelity model

(c) P-NIROM (6 basis functions) (d) P-NIROM (6 basis functions)

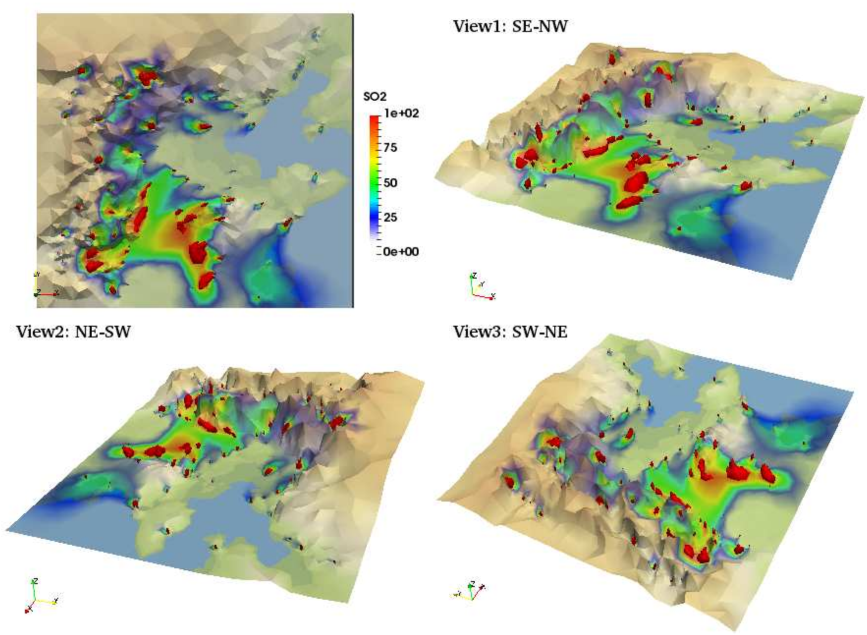
(e) P-NIROM (36 basis functions) (f) P-NIROM (36 basis functions)

$t = 30\,hrs$ $t = 105\,hrs$

Figure 5: Case T1 ($\xi_1 = 2400, \xi_2 = 2400, \xi_3 = 5000$) $g\,s^{-1}$: the comparison of pollutant concentration solutions at time levels $t = 30\,hrs$ (left panel) and $t = 105\,hrs$ (right panel) between the high fidelity model and P-NIROM with 6 and 36 POD basis functions.



(a) Full model



(b) P-NIROM with 36 basis functions

Figure 6: Case T1 ($\xi_1 = 2400, \xi_2 = 2400, \xi_3 = 5000$) $g\,s^{-1}$: the comparison of pollutant results at time levels $t = 30\,hrs$ between the high fidelity full model and P-NIROM with 36 basis functions from different angles.

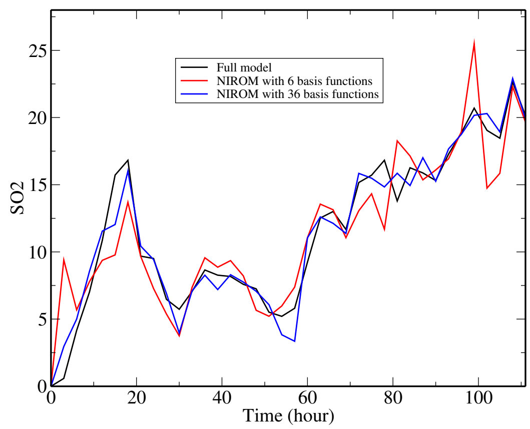


Figure 7: Case T1 $(\xi_1 = 2400, \xi_2 = 2400, \xi_3 = 5000)\, g\, s^{-1}$: the evolution of pollutant concentration solutions predicted by the high fidelity model and P-NIROM at a specified location $x = 378\, km, y = 786\, km$.

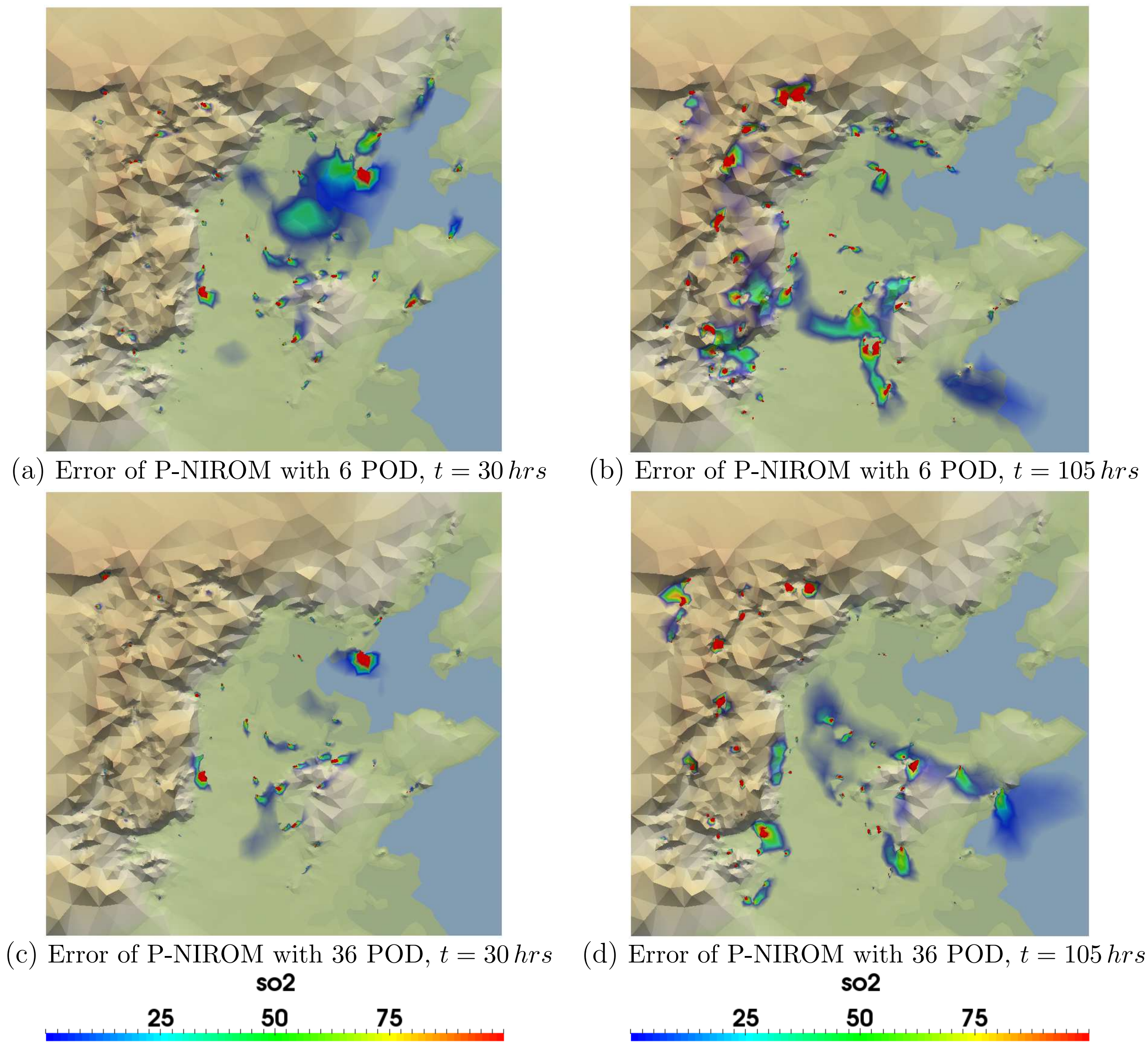
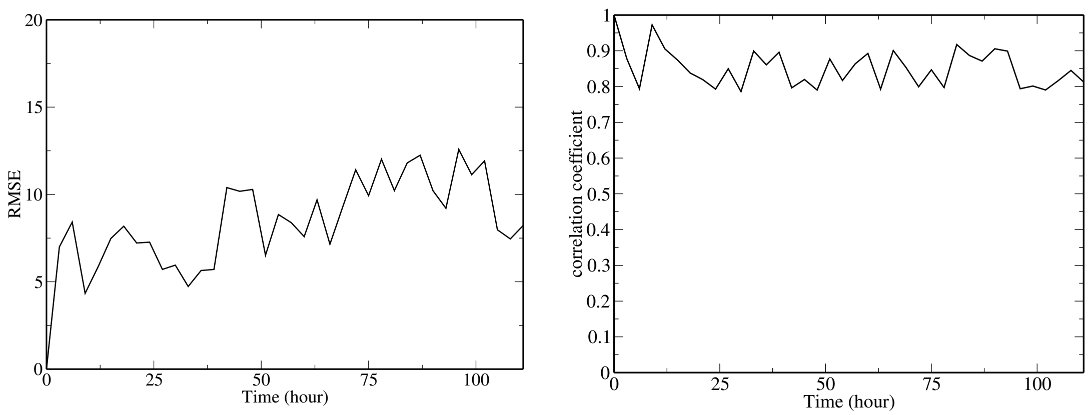
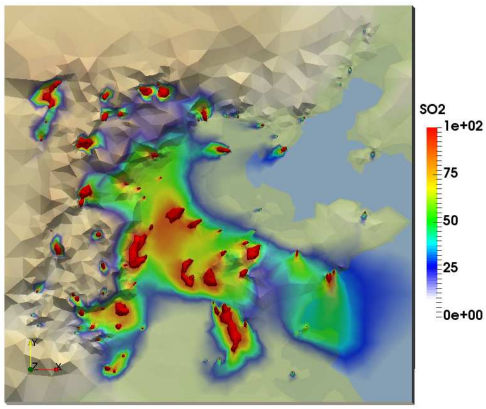


Figure 8: Case T1 ($\xi_1 = 2400, \xi_2 = 2400, \xi_3 = 5000$) $g\,s^{-1}$: the spatial distribution of absolute errors between the high fidelity model and P-NIROM which is constructed with 6 and 36 POD basis functions.

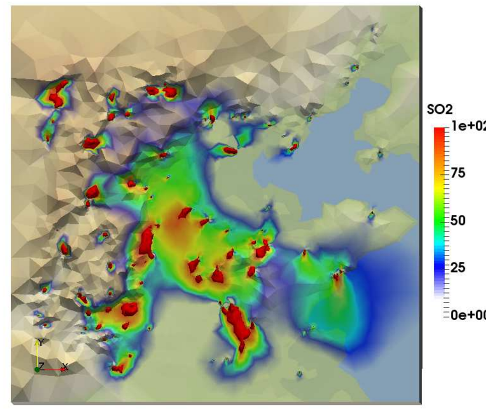


(a) RMSE (b) correlation coefficients

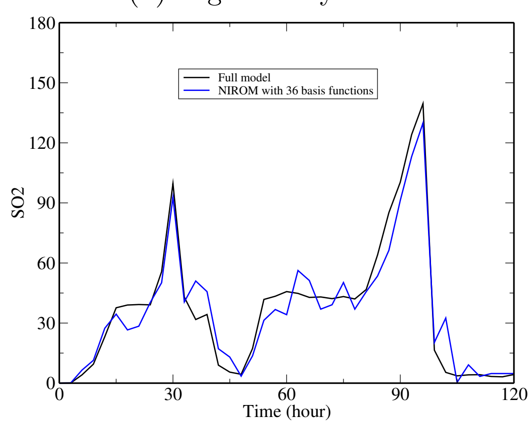
Figure 9: Case T1 $(\xi_1 = 2400, \xi_2 = 2400, \xi_3 = 5000)\, g\, s^{-1}$: the RMSE and correlation coefficients of pollutant concentration solutions between the high fidelity model and P-NIROM with 36 POD basis functions.



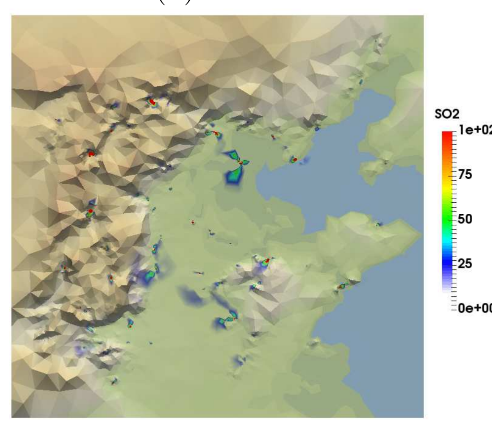
(a) High fidelity model



(b) P-NIROM



(c) Solutions at a location: $x = 599\,km$, $y = 569\,km$



(d) Error

Figure 10: Case T2 ($\xi_1 = 5500, \xi_2 = 6000, \xi_3 = 6000)\,g\,s^{-1}$: comparison of pollutant concentration solutions between the high fidelity full model and P-NIROM with 36 basis functions: (a) and (b) the spatial solution at time level $t = 102\,hrs$ from the high fidelity model and P-NIROM respectively; (c) the evolution of pollutant concentration solutions at a location: $x = 599\,km$, $y = 569\,km$; and (d) the spatial error at time level $t = 102\,hrs$.

4.1. Computational efficiency

This section provides a comparison of the online computational CPU cost required by the high fidelity full model and P-NIROM. The specifications of the machine for simulations were: 12 cores with a frequency of 3.33GHz ($Intel^{\textregistered}$ Xeon(R) CPU X5680 @3.33GHz × 12) and a 62.9GB memory. One core was used for the simulations since the cases were simulated in serial. Table 2 lists the online CPU cost required for running the high fidelity model and P-NIROM. The offline cost (see Figure 1) at the training stage is not listed in this table. It can be seen that using the P-NIROM, the CPU time is reduced by five order of magnitude in comparison to the high fidelity model.

Table 2: Online CPU cost required for running the high fidelity model and P-NIROM during one time step.

Cases	Model	assembling and solving	projection	interpolation	total
Test case	Full model	616.9	0	0	696.59
	NIROM	0	0.003	0.001	0.004

5. Conclusions

This article has presented a new P-NIROM for predictive modelling of pollutant transport phenomena. The machine learning techniques in combination with POD are used for constructing the P-NIROM. First, at the training stage, for the selected input parameters $\mu_{tr} \in R^P$, the solution snapshots (serving as training datasets) and POD basis functions are obtained by running the high fidelity model. From the training data sets, for any given input parameters $\mu \in R^P$, using the machine learning technique a set of hyper-surface functions (P-NIROMs) is constructed to represent the dynamics of pollutant transport over the reduced space. The P-NIROM is then used for calculating the reduced solutions (POD coefficients) for the given μ (the emission intensity). The unique combination of the P-NIROM and machine learning techniques enables rapid and reasonably accurate simulations. The P-NIROM techniques developed here are robust and can be used for a large number of disciplines not least of pollutant flow based disciplines.

The P-NIROM has been applied to a realistic case in China involving plumes released from over 100 power plants. The varying input parameter is the emission intensity of pollutant sources. A comparison of pollutant solutions between the high fidelity model and P-NIROM has been undertaken. The P-NIROM with 36 POD basis functions exhibits an overall good agreement with the high fidelity model. The online computation cost required by the P-NIROM is reduced by several orders of magnitude in comparison to the high fidelity model.

Compared to existing P-NIROM techniques (for example, based on radial basis functions), the P-NIROM based on machine learning methods provides a wider range of application areas, for example, uncertainty analysis in both data and modelling results, real-time interactive use, data management (real-time data monitoring/analysis), data assimilation and better-informed decision making. In particular, the machine learning techniques with ROM can be used for data selection and data reduction by condensing the information into the desired number of features and recovering the original data from the reduced feature set.

Acknowledgements

The authors are grateful for the support of the EPSRC grant: Managing Air for Green Inner Cities (MAGIC)(EP/N010221/1).

References

[1] Kristen M Foley, Sergey L Napelenok, Carey Jang, Sharon Phillips, Bryan J Hubbell, and Charles M Fulcher. Two reduced form air quality modeling techniques for rapidly calculating pollutant mitigation potential across many sources, locations and precursor emission types. *Atmospheric environment*, 98:283–289, 2014.

[2] Laurent Cordier, Bernd R Noack, Gilles Tissot, Guillaume Lehnasch, Joël Delville, Maciej Balajewicz, Guillaume Daviller, and Robert K Niven. Identification strategies for model-based control. *Experiments in fluids*, 54(8):1580, 2013.

[3] Bernard Haasdonk. Reduced basis methods for parametrized pdesa tutorial introduction for stationary and instationary problems. *Model reduction and approximation: theory and algorithms*, 15:65, 2017.

[4] Peter Benner, Serkan Gugercin, and Karen Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.

[5] Michael Hinze and Stefan Volkwein. Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control. In *Dimension reduction of large-scale systems*, pages 261–306. Springer, 2005.

[6] Rafik Djouad and Bruno Sportisse. Solving reduced chemical models in air pollution modelling. *Applied Numerical Mathematics*, 44(1-2):49–61, 2003.

[7] Janelle K HAMMOND, Rachida CHAKIR, Frédéric Bourquin, and Yvon Maday. PBDW: a non-intrusive Reduced Basis Data Assimilation Method and its application to outdoor Air Quality Models. working paper or preprint, January 2018.

[8] Farha A Alkuwari, Serge Guillas, and Yuhang Wang. Statistical downscaling of an air quality model using fitted empirical orthogonal functions. *Atmospheric environment*, 81:1–10, 2013.

[9] F. Fang, T. Zhang, D. Pavlidis, C.C. Pain, A.G. Buchan, and I.M. Navon. Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons. *Atmospheric Environment*, 96:96–106, 2014.

[10] H. Chen. Blackbox stencil interpolation method for model reduction. Master's thesis, Massachusetts Institute of Technology, 2012.

[11] Michael Schlegel and Bernd R. Noack. On long-term boundedness of Galerkin models. *Journal of Fluid Mechanics*, 765:325–352, 2 2015.

[12] Jan Osth, Bernd R. Noack, Sinia Krajnovi, Diogo Barros, and Jacques Bore. On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body. *Journal of Fluid Mechanics*, 747:518–544, 5 2014.

[13] David Amsallem and Charbel Farhat. Stabilization of projection-based reduced-order models. *International Journal for Numerical Methods in Engineering*, 91(4):358–377, 2012.

[14] Leopoldo P. Franca and Sergio L. Frey. Stabilized finite element methods: II. The incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 99(2-3):209–233, 1992.

[15] S. Chaturantabut and D.C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput*, 32:2737–2764, 2010.

[16] Feriedoun S. and Alireza J. α Regularization of the POD-Galerkin dynamical systems of the Kuramoto-Sivashinsky equation. *Applied Mathematics and Computation*, 218(10):6012 – 6026, 2012.

[17] D. Xiao, F. Fang, A.G. Buchan, C.C. Pain, I.M. Navon, J. Du, , and G. Hu. Non-linear model reduction for the Navier-Stokes equations using Residual DEIM method. *Journal of Computational Physics*, 263:1–18, 2014.

[18] D. Xiao, F. Fang, J. Du, C.C. Pain, I.M. Navon, A.G. Buchan, A.H. ElSheikh, and G. Hu. Non-linear Petrov-Galerkin methods for reduced order modelling of the Navier-Stokes equations using a mixed finite element pair. *Computer Methods In Applied Mechanics and Engineering*, 255:147–157, 2013.

[19] C. Audouze, F.D. Vuyst, and P.B. Nair. Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations*, 29(5):1587–1628, 2013.

[20] J.S. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363:55 – 78, 2018.

[21] D. Xiao, F. Fang, A.G. Buchan, C.C. Pain, I.M. Navon, and A. Muggeridge. Non-intrusive reduced order modelling of the Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 293:552–541, 2015.

[22] Wang Chen, Jan S Hesthaven, Bai Junqiang, Zhang Yang, and Yang Tihao. A greedy non-intrusive reduced order model for fluid dynamics. *Journal of Northwestern Polytechnical University*, 2017.

[23] Diana Alina Bistrian and Ionel Michael Navon. Randomized dynamic mode decomposition for nonintrusive reduced order modelling. *International Journal for Numerical Methods in Engineering*, 112(1):3–25, 2017.

[24] D. Xiao, F. Fang, C. Pain, and G. Hu. Non-intrusive reduced order modelling of the Navier-Stokes equations based on RBF interpolation. *International Journal for Numerical Methods in Fluids*, 79(11):580–595, 2015.

[25] D Xiao, F Fang, C Pain, IM Navon, and A Muggeridge. Non-intrusive Reduced Order Modelling of Waterflooding in Geologically Heterogeneous Reservoirs. In *ECMOR XV-15th European Conference on the Mathematics of Oil Recovery*, 2016.

[26] Toomaj Foroud and Abbas Seifi. A guided pattern search with a non-intrusive reduced order modeling for oil production optimization: Brugge field case study. *Journal of Petroleum Science and Engineering*, 147:570–584, 2016.

[27] D Xiao, Z Lin, F Fang, C C Pain, Ionel M Navon, P Salinas, and A Muggeridge. Non-intrusive reduced-order modeling for multiphase porous media flows using smolyak sparse grids. *International Journal for Numerical Methods in Fluids*, 83(2):205–219, 2017.

[28] D. Xiao, P. Yang, F. Fang, J. Xiang, C.C. Pain, and I.M. Navon. Non-intrusive reduced order modeling of fluid-structure interactions. *Computer Methods in Applied Mechanics and Engineering*, 303:35–54, 2016.

[29] Z. Wang, D. Xiao, F. Fang, R. Govindan, C.C. Pain, and Y. Guo. Model identification of reduced order fluid dynamics systems using deep learning. *International Journal for Numerical Methods in Fluids*, 86(4):255–268, 2018.

[30] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[31] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick

Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[32] Jonathan J Tompson, Arjun Jain, Yann LeCun, and Christoph Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. In *28th Annual Conference on Neural Information Processing Systems 2014, NIPS 2014 - Montreal, Canada, Advances in neural information processing systems*, pages 1799–1807, 2014.

[33] Michael KK Leung, Hui Yuan Xiong, Leo J Lee, and Brendan J Frey. Deep learning of the tissue-regulated splicing code. *Bioinformatics*, 30(12):i121–i129, 2014.

[34] Raia Hadsell, Pierre Sermanet, Jan Ben, Ayse Erkan, Marco Scoffier, Koray Kavukcuoglu, Urs Muller, and Yann LeCun. Learning long-range vision for autonomous off-road driving. *Journal of Field Robotics*, 26(2):120–144, 2009.

[35] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug):2493–2537, 2011.

[36] Carl Edward Rasmussen. *Gaussian Processes in Machine Learning*, pages 63–71. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[37] J Zheng, J Zhu, Z Wang, F Fang, CC Pain, and J Xiang. Towards a new multiscale air quality transport model using the fully unstructured anisotropic adaptive mesh technology of fluidity (version 4.1. 9). *Geoscientific Model Development*, 8(10):3421, 2015.

[38] Daewon W Byun and Robin Dennis. Design artifacts in eulerian air quality models: Evaluation of the effects of layer thickness and vertical profile correction on surface ozone concentrations. *Atmospheric Environment*, 29(1):105–126, 1995.

[39] William C Skamarock, Joseph B Klemp, Jimy Dudhia, David O Gill, Dale M Barker, Wei Wang, and Jordan G Powers. A description of the advanced research wrf version 3. Technical report, NCAR Tech, June 2008.

[40] AMCG and Imperial. College. London. Fluidity manual v4.1.12. figshare. *https://dx.doi.org/10.6084/m9.figshare.1387713.v2*, pages 1–329, 2015.