Imperial College London

Controlling ventilation in buildings using Adversarial Neural Networks to produce comfortable healthy environments that are energy efficient

Imperial College London

Project Plan

First Supervisor
Christopher Pain

Second Supervisor
Boyang Chen

Other Supervisors
Claire Heaney
Rossella Arcucci

Author:
Donghu Guo

CID:
02140865

Email:
dg321@ic.ac.uk

MSc EDSML
Department of Earth Science and Engineering

July 2022

1 Introduction & Literature review

Energy efficiency on the ventilation side is important to claim, especially in the winter, which usually takes up a large proportion of energy we use in cities [1]. The challenge is not only about to over ventilate but also to ventilate adequately. At the same time, we should consider the air quality. Sometimes we do not want to bring polluted air into the indoor space as well. Therefore, it is essential to make quantitative analyses of indoor air, such as EnergyPlus modelling and Fluidity modelling [2].

In view of the previous monitoring and simulation of human exhaled gas flow, its propagation is often very irregular and difficult to explain intuitively, which will even complicate the existing indoor air changes [3]. At present, Computational Fluid Dynamics (CFD) is widely used to simulate such complex physical processes through numerical calculations [4], which is one of the main tools allowing us to understand air flows [5]. However, its computational cost and time spent are both very high. So that is why here we want to produce AI models, which have great potential to enhance CFD [6].

Generative adversarial network (GAN) [7] uses an adversarial training strategy to directly shape the output distribution through back propagation. Similarly, Adversarial Autoencoder (AAE) was proposed [8] to perform variational inference by matching the aggregate posterior value of the hidden code vector of the autoencoder with any prior distribution. These adversarial networks have rich applications in Super Resolution, Image-to-Image Translation, etc. After some proper adjustment, they can be applied to predict data in time series. With adversarial training, the prediction results will follow the expected distribution [9]. For instance, GAN was used to quantify the uncertainty of forward simulations in the presence of observed data [10]. To distinguish, we refer to this modified GAN as PredGAN. Later on, PredAAE was proposed by modifying the classic AAE mentioned above. The corresponding prediction algorithm can be seen in Figure 1b. In addition, consisting of a normal neural network and a discriminator, the predictive adversarial neural network (PredANN) [11] also has excellent prediction performance. The corresponding loss function could be written as follows:

$$\min_{G,H} E(\|\alpha_{g,C}^k - \tilde{\alpha}_{g,C}^k\|^2) + \min_{G} \max_{D} E_{z\sim p_{prior}}[logD(z)] + E_{\alpha\sim p_{data}}[log(1 - D(G(\alpha))] \tag{1}$$

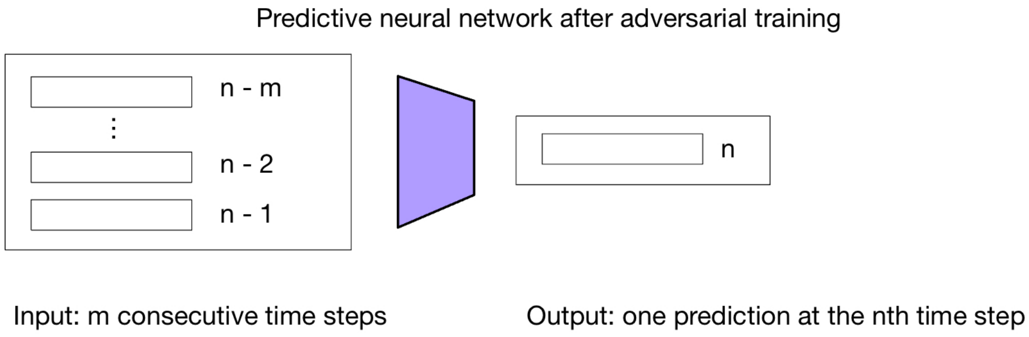
Data Assimilation (DA) can be used to improve the accuracy of dynamic model, which aims to combine the prediction with the observation to reduce the uncertainty in the model [12]. DA techniques are generally divided into stochastic data assimilation (best linear unbiased estimator (BLUE), kalman filter algorithm) and variational data assimilation (3D variational assimilation (3D-Var), 4D variational assimilation (4D-Var)). The cost functions of 3D-Var and 4D-Var could be written as follows respectively:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H[\mathbf{x}])^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]) \tag{2}$$

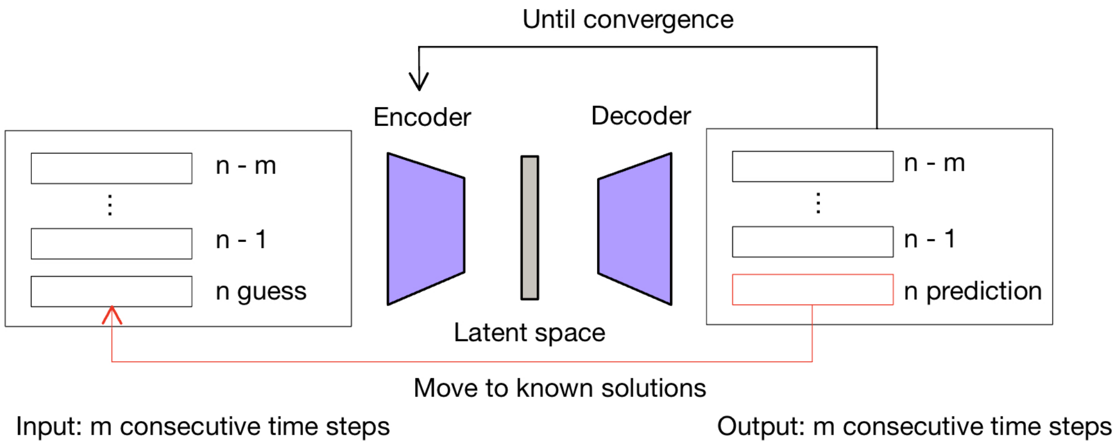
$$= 2\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - 2H^T\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]) \tag{3}$$

where $\mathbf{B}$ denotes the background error covariance, $\mathbf{R}$ the observational error covariance, H is a linear or nonlinear observation operator.

In this application field, the original data is mainly 3D space, which is high-dimensional samples. If they are directly used in model training, a very large amount of computational cost will be produced. Therefore, it is particularly important to reduce the dimension of data in advance, which is usually referred to as Reduced-Order Modelling (ROM) [13]. By implementing Principal component analysis (PCA), proper orthogonal decomposition (POD) [14] is a common way to achieve ROM and is suitable for 3D data.



(a) Direct prediction algorithm: PredANN.



(b) Indirect prediction algorithm: PredAAE.

Figure 1: Two prediction strategies: direct and indirect.

2 Problem Description and Objectives

The project mainly focuses on energy efficiency and control of the ventilation based on AI and a purely data-driven approach. We want to combine the ANNs and DA [10] to simulate and predict the relevant air indicators of the enclosed space according to different ventilation settings. Based on this prediction result, the air condition in the future can be quickly judged, and corresponding countermeasures can be put forward efficiently. The ultimate goal in this field can be making buildings produce at least the amount of site energy they consume (Net-zero buildings) [15].

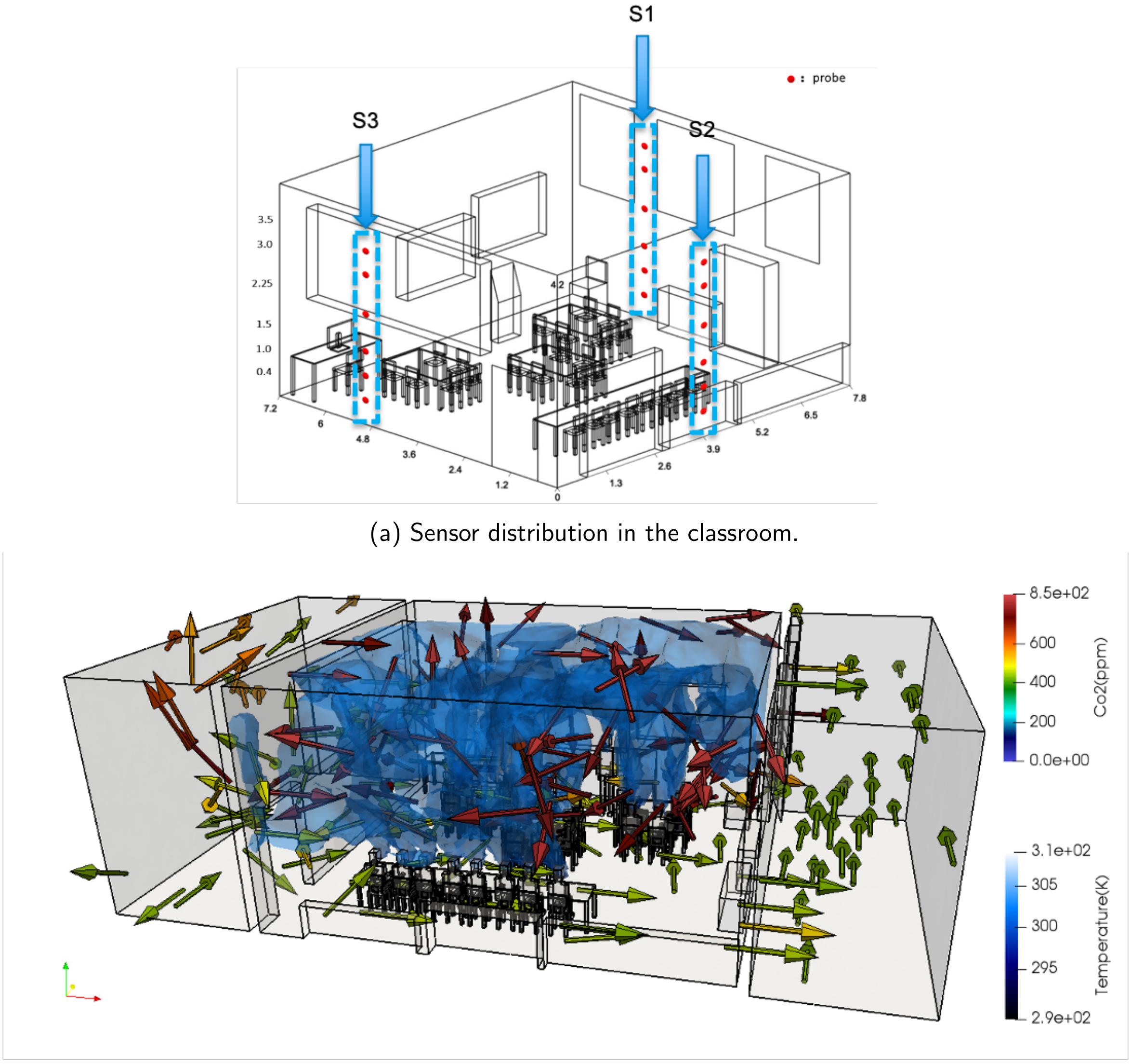
Model selection, construction and training are the key parts of the project. Here, we will focus on comparing the prediction performance of different adversarial network frameworks and will select models with more stable and accurate output, and carry out the corresponding hyperparameter optimisation [16], where we want to introduce new hyperparameters related to different ventilation controls such as ventilation rate into our model. In addition, the design of the whole prediction system also needs in-depth research, such as what size of sequence data $(\mathbf{m})$ are selected as the input of prediction, or whether to select the strategy of direct prediction or indirect iterative prediction, shown in Figure 1, and so on. At this stage, we hope to implement PredANN based on PredAAE, so as to achieve direct prediction.

After training an ideal model, our objective next is trying to quickly predict the future change trend of indoor air indicators such as CO_2 and temperature, timely find out possible abnormal conditions, and then research how to take corresponding ventilation control.

3 Progress to Date

- Completed mostly the collection, collation and review of relevant literature.
- Built up the whole workflow for simulation and prediction of the enclosed space.
  At this stage, the data we use is a subset obtained by putting 18 sensors into the 3D simulation data, involving CO_2, temperature and humidity. The location of sensor distribution is shown in Figure 2a. In the next phase, we will access the full 3D CFD data to train our model formally (Figure 2b shows CO_2 plume on the 800ppm Iso-surface when doing CFD.).

  Now the PredAAE model has been built and trained. The DA can be implemented subsequently. As we currently are not using the formal dataset, all these models will be retrained afterwards. Thus, only show the autoencoder reconstruction performance of PredAAE in Figure 3.
- Set up software development life cycle.
  At present, I have established a basic framework on GitHub, integrated the model-relevant implementations into a tool package. Jupyter notebooks can carry out relevant experiments by directly importing the package.



(a) Sensor distribution in the classroom.

(b) CO_2 plume in a ventilated classroom: 800ppm Iso-surface.

Figure 2: Description of CFD model.

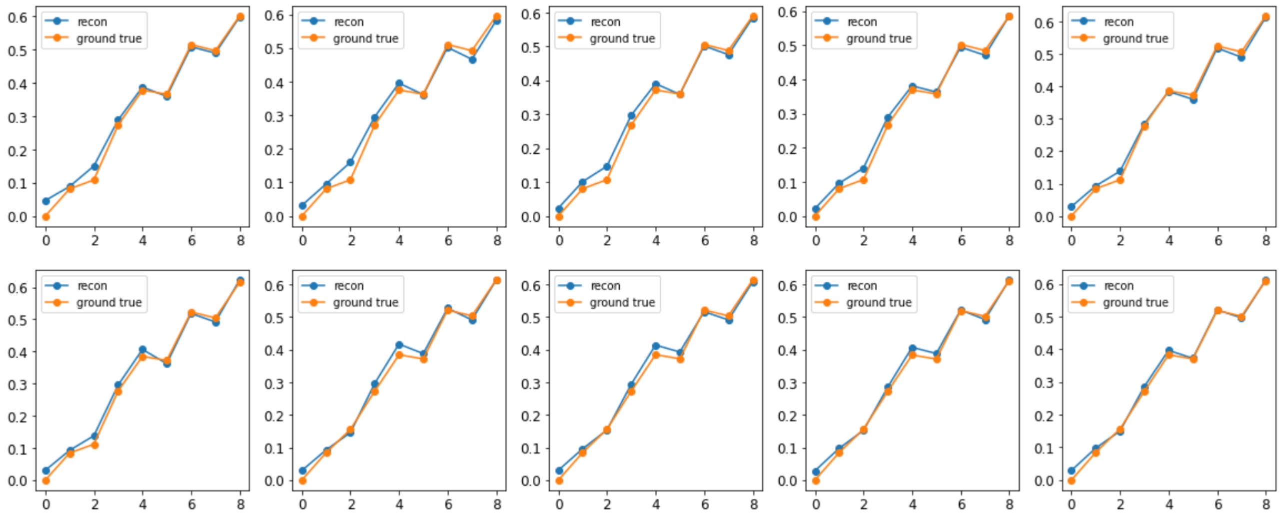


Figure 3: Autoencoder reconstruction performance of PredAAE. (Different subplot represents different sensor, here we show 10 of them. In each subplot, the 'ground true' line represents 9 timesteps' CO_2 data, whereas the 'recon' represents the autoencoder output.)

4 Future Plan

According to the review of the literature and the implementation of the experiment so far, I put my idea about the overall workflow of the whole project into the Figure 4.

And corresponding to this workflow, I list the main to-do tasks in the future as the following points. See Figure 5 for more details.

- Continue improving the model and workflow.
  Try to make predictions in time, by building ANNs and comparing the performance between PredAAE and PredANN to choose the better one. Try to work out 4D-Var to better combine the model and observation data.
- Train predictive model with the full 3D simulation data, where I would try to use the POD coefficients training the selected adversarial network to learn the evolution of the system in time.
- Implement different ventilation setting experiments to generate relevant hyperparemeters such as ventilation rate for the model, which would be key innovation of the project.
- Collect sensor data in a real meeting room in the next few weeks and try to implement experiments on it. Compared with the simulated data, the measured data has more noise and higher uncertainty, which can make our model more robust.
- Continuously improve the GitHub repository, optimise the code, and write the final report.

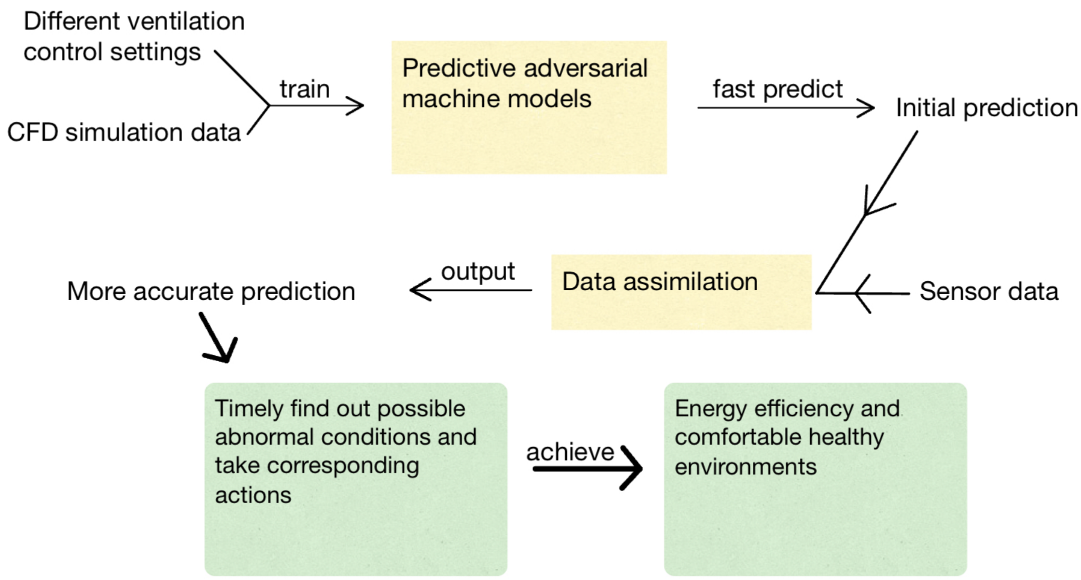


Figure 4: Project workflow.

Name	Start Date	End Date
Continuously improve the GitHub repository, optimise the code, and write the final report.	Jul 01, 2022	Sep 02, 2022
ROM and Train different ANNs, compare the performance and choose the best one	Jul 01, 2022	Jul 12, 2022
Train the chosen predictive model with the full 3D simulation data.	Jul 05, 2022	Jul 27, 2022
Train and test the chosen model on the collected sensor data (Data assimilation)	Jul 12, 2022	Aug 10, 2022
Implement different ventialtion setting experiments to generate hyperparameter for the model	Jul 15, 2022	Aug 24, 2022
Presentation	Sep 14, 2022	Sep 16, 2022
Submit final report, codes, and al relevant files	Sep 02, 2022	Sep 02, 2022

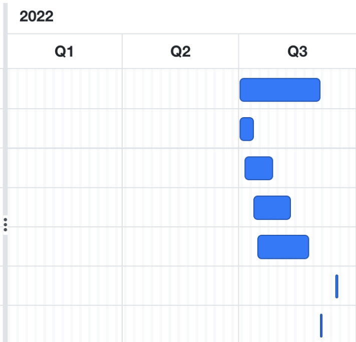


Figure 5: Future plan. (Q represents quarter, where Q3 includes July, August, and September.)

References

[1] EnergyPlus Version. 8.7 documentation—input output reference. *US Department of Energy: Washington, DC, USA*, 2017.

[2] Mottet Laetitia, Song Jiyun, Short C Alan, Chen Shuqin, Wu Jindong, Yu Wei, Xiong Jie, Zhang Qiulei, Ge Jian, Liu Meng, et al. The hot summer-cold winter region in china: Challenges in the low carbon adaptation of residential slab buildings to enhance comfort. *Energy and Buildings*, 223:110181, 2020.

[3] Marija Trčka and Jan LM Hensen. Overview of hvac system simulation. *Automation in Construction*, 19(2):93–99, 2010.

[4] TJ Chung et al. *Computational fluid dynamics*. Cambridge university press, 2002.

[5] Dunhui Xiao, CE Heaney, L Mottet, F Fang, W Lin, IM Navon, Y Guo, OK Matar, AG Robins, and CC Pain. A reduced order model for turbulent flows in the urban environment using machine learning. *Building and Environment*, 148:323–337, 2019.

[6] Ricardo Vinuesa and Steven L Brunton. The potential of machine learning to enhance computational fluid dynamics. *arXiv preprint arXiv:2110.02085*, 2021.

[7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[8] Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey. Adversarial autoencoders. *arXiv preprint arXiv:1511.05644*, 2015.

[9] Vinicius LS Silva, Claire E Heaney, and Christopher C Pain. Gan for time series prediction, data assimilation and uncertainty quantification. *arXiv preprint arXiv:2105.13859*, 2021.

[10] Vinicius LS Silva, Claire E Heaney, Yaqi Li, and Christopher C Pain. Data assimilation predictive gan (da-predgan): applied to determine the spread of covid-19. *arXiv preprint arXiv:2105.07729*, 2021.

[11] Claire Heaney, Xiangqi Liu, Hanna Go, Zef Wolffs, Pablo Salinas, Ionel M Navon, and Christopher C Pain. Extending the capabilities of data-driven reduced-order models to make predictions for unseen scenarios: applied to flow around buildings. *Frontiers in Physics*, page 563.

[12] Rossella Arcucci, Jiangcheng Zhu, Shuang Hu, and Yi-Ke Guo. Deep data assimilation: integrating deep learning with data assimilation. *Applied Sciences*, 11(3):1114, 2021.

[13] Bernd R Noack, Marek Morzynski, and Gilead Tadmor. *Reduced-order modelling for flow control*, volume 528. Springer Science & Business Media, 2011.

[14] Anindya Chatterjee. An introduction to the proper orthogonal decomposition. *Current science*, pages 808–817, 2000.

[15] Igor Sartori, Assunta Napolitano, and Karsten Voss. Net zero energy buildings: A consistent definition framework. *Energy and buildings*, 48:220–232, 2012.

[16] César Quilodrán-Casas, Vinicius LS Silva, Rossella Arcucci, Claire E Heaney, YiKe Guo, and Christopher C Pain. Digital twins based on bidirectional lstm and gan for modelling the covid-19 pandemic. *Neurocomputing*, 470:11–28, 2022.