

**Analysis and Reduced-Order Modeling of Urban Airflow and Pollutant
Dispersion under Thermal Stratification Conditions**

Shahin Masoumi-Verki

A Thesis

in the Department

of

Building, Civil, and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Building Engineering) at

Concordia University

Montreal, Quebec, Canada

July 2022

© Shahin Masoumi-Verki, 2022

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Shahin Masoumi-Verki

Entitled: Analysis and Reduced-Order Modeling of Urban Airflow and Pollutant Dispersion under
Thermal Stratification Conditions

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Building Engineering)

complies with the regulations of the University and meets the accepted standards with respect to
originality and quality.

Signed by the final examining committee:

_____ (Chair)
Dr. Andrea Schiffauerova

_____ (Thesis Supervisor)
Dr. Fariborz Haghighat

_____ (Thesis Supervisor)
Dr. Ursula Eicker

_____ (Examiner)
Dr. Zhi Chen

_____ (Examiner)
Dr. Mohamed Ouf

_____ (Examiner)
Dr. Marius Paraschivoiu

_____ (External Examiner)
Dr. John Z. Zhai

Approved by _____
Dr. Mazdak Nik-Bakht, Graduate Program Director

August 22, 2022 _____
Dr. Mourad Debbabi, Dean,
Gina Cody School of Engineering and Computer Science

Abstract

Analysis and Reduced-Order Modeling of Urban Airflow and Pollutant Dispersion under Thermal Stratification Conditions

Shahin Masoumi-Verki, Ph.D.

Concordia University, 2022

Different thermal stratification conditions, namely, stable, isothermal (or neutral), and unstable, can locally occur in urban areas. Alteration in the thermal condition of an urban area may significantly change the airflow pattern and pollutant dispersion process by affecting both the mean and fluctuating components of the variables. The unstable effects can increase the vertical flow movement, while the stable ones can suppress it. Furthermore, unstable conditions increase turbulence kinetic energy (TKE), which increases the fluctuations in concentration. On the other hand, stable conditions lead to buoyant destruction. Due to frequent changes in the boundary conditions, a model is required for monitoring these situations, which can be used as a fast-response (near real-time) model. This thesis aims to propose a systematic approach for analysis and reduced-order modeling of the airflow and concentration fields under non-isothermal conditions.

The present study uses a high-fidelity computational fluid dynamics approach, i.e., embedded large eddy simulation (ELES), to simulate the impact of the aforementioned thermal conditions on the airflow and concentration fields. The model considers the pros of both the Reynolds-averaged Navier-Stokes, RANS, (i.e., high speed), and large eddy simulation, LES, (i.e., high accuracy) approaches. After thoroughly analyzing the results, the proper orthogonal decomposition (POD) and frequency analyses are performed to investigate the impact of thermal conditions on the turbulence structure of the flow field. Considering the most energetic POD modes can lead to a good approximation of the whole airflow field, which is an important finding in developing a reduced-order model (ROM). Due to the limitations arising from the linear nature of POD, convolutional autoencoder (CAE)-based methods are used for model order reduction, using the unstable dataset generated by ELES. In addition to the conventional CAE, multiscale CAE (MS-CAE) and self-attention CAE (SA-CAE) are developed to capture multiscale and long-range dependencies among the datapoints, respectively. Afterwards, a parallel long short-term memory (LSTM) network is used to compute the temporal dynamics of the low-dimensional subspaces. ROMs maintain prediction accuracy at an acceptable level compared to ELES, while reducing the data reconstruction time to the order of seconds.

Acknowledgment

First and foremost, I would sincerely like to express my gratitude to my supervisors, Professor Fariborz Haghghat and Professor Ursula Eicker, for their support and guidance during my study. I am grateful for the invaluable advice they gave me over these years. For me, it was a great honor to meet, know, and work with them.

Also, I would like to acknowledge other members of my thesis committee, Dr. Marius Paraschivoiu, Dr. Zhi Chen, and Dr. Mohamed Ouf, for their valuable advice.

I would like to name a group of extraordinary people I am lucky to have as friends: Amirhossein, Majid, Shahrooz, Alireza, Payam, Hesam, and Mohammadreza. Also, I would like to thank my friends and colleagues in the Energy and Environment group and the Next-Generation Cities institute for their help and stimulating exchange of ideas.

Last but not least, I wish to express my appreciation and love to my mother, Nasrin, and my father, Shahrooz, for their continuous inspiration and support. I dedicate this work to my loving parents, whom I dearly miss.

Contributions of the Author

Journal Papers:

Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses

Authors	Shahin Masoumi-Verki, Payam Gholamalipour, Fariborz Haghighat, Ursula Eicker
Journal, Date, and DOI	Building and Environment (2021), vol. 206, 108388, https://doi.org/10.1016/j.buildenv.2021.108388
Abstract	<p>Non-isothermal stratification conditions can alter the airflow pattern and pollutant dispersion process within urban areas. The present study is focused on the impact of various stratification conditions, namely, stable, isothermal (neutral), and unstable, on the airflow and concentration fields around an isolated high-rise building. Zonal Reynolds-averaged Navier-Stokes (RANS)-large eddy simulation (LES), also known as embedded large eddy simulation (ELES), is employed for simulating the airflow and concentration fields under non-isothermal boundary layers in order to make a balance between computational costs and accuracy. Comparing the results predicted by the present ELES with an available LES study, with almost similar computational settings (i.e., the inflow turbulence generation method, grid resolution, etc.) shows better performance of ELES in predicting the concentration field. The findings also illustrate that the impact of the unstable stratification condition on turbulence statistics is more pronounced than that of the stable stratification condition. The present article also investigates the effect of thermal stratification conditions on the mechanisms of pollutant dispersion, namely, convective and turbulent diffusion fluxes. The findings reveal that an increase in turbulence kinetic energy (TKE), caused by the unstable thermal stratification condition, increases the concentration fluctuations, which causes the pollutant concentration to be decreased. Furthermore, spectral and proper orthogonal decomposition (POD) analyses are performed for all stratification scenarios. The results show that by altering the thermal condition from isothermal to non-isothermal, either stable or unstable conditions, the contribution of the primary dominant modes to total TKE increases.</p>
Chapters	Chapter 3

A review of advances towards efficient reduced-order models (ROM) for predicting urban airflow and pollutant dispersion

Authors	Shahin Masoumi-Verki, Fariborz Haghighat, Ursula Eicker
Journal, Date, and DOI	Building and Environment (2022), vol. 216, 108966, https://doi.org/10.1016/j.buildenv.2022.108966
Abstract	Computational fluid dynamics (CFD) models have been used for the simulation of urban airflow and pollutant dispersion, due to their capability to capture different length scales and turbulence nature of the flow field. However, their high computational costs prevent them from being used for (near) real-time simulations, long-term predictions, and simulations with dynamic boundary conditions. Reduced-order models (ROMs) are proposed as reliable alternatives to CFD approaches to solve the mentioned issues. This article aims to comprehensively review the state-of-the-art application of different methodologies to develop a non-intrusive ROM (NIROM) for predicting urban airflow and pollutant dispersion. Developing such models comprises two steps: dimensionality reduction and computing the feature dynamics of the reduced space. Various methodologies, with the focus on machine learning algorithms, are proposed for the mentioned stages, while their capabilities and limitations are discussed. Furthermore, different approaches are introduced to overcome the issue of the physical interpretation of these models. Also, several methods are proposed to make the models suitable for being used in long-term predictions and multi-query problems (i.e., considering changes in boundary conditions).
Chapters	Chapter 4

Performance analysis of different reduced-order models for predicting urban turbulent flow field

Authors	Shahin Masoumi-Verki, Fariborz Haghighat, Ursula Eicker
---------	---

Journal, Date, and DOI	Submitted (Under Review)
Abstract	<p>Computational fluid dynamics (CFD) has been considered as a promising numerical approach in fluid dynamics problems, such as urban airflow prediction. However, flow field prediction using CFD models is time-consuming. Thus, they cannot be used for (near) real-time and long-term simulations. Reduced-order models (ROMs) are emerged to obviate this limitation. Deep learning (DL) algorithms have been used for developing non-intrusive ROMs (NIROMs) in fluid dynamics applications. In the present study, three different approaches, namely, convolutional autoencoder (CAE), multi-scale CAE (MS-CAE), and self-attention CAE (SA-CAE), are developed for dimensionality reduction, which is considered the first step of the development of a NIROM. The developed models are then used to find a low-dimensional representation of the original data. Afterward, a parallel long short-term memory (LSTM) network is employed for computing the temporal dynamics of the obtained low-dimensional space. The models are trained to reconstruct a turbulent airflow field in the wake region of an isolated high-rise building, located in an unstable thermal stratification condition, using validated CFD data. The models show promising performance in reconstructing the flow field. However, discrepancies can be observed in the regions with intense gradients. Also, power spectral density functions (PSD) obtained from the reconstructed data are in good agreement with those obtained from the CFD results. On the whole, SA-CAE performs better in reconstructing the flow field than the other models, followed by MS-CAE and CAE.</p>
Chapters	Chapter 5

Conference Paper:

Data-Driven Reduced-Order Model for Urban Airflow Prediction

Authors	Shahin Masoumi-Verki, Fariborz Haghighat, Ursula Eicker
Publisher	COBEE 2022, 5th International Conference on Building Energy and Environment (Montreal, Canada)

Abstract

Due to computational barriers of computational fluid dynamics (CFD) models, they cannot be used for tasks such as (near) real-time simulations. Reduced-order model (ROM) can be used as an alternative to CFD since it can approximate the results in a fraction of the CFD simulation time. The present article generates a data-driven ROM, using convolutional autoencoders (CAEs) and long short-term memory (LSTM) networks, to reconstruct the turbulent flow field within a simplified urban area. Furthermore, the effect of the kernel size on capturing spatial information is investigated. The results indicate that, although the model has some deficiencies in the flow field reconstruction in high-gradients regions, the model's overall performance is acceptable. Moreover, it is shown that the kernel size has a negligible impact on the model performance for the present model and dataset.

Table of Contents

List of Figures.....	xiii
List of Tables.....	xvii
Abbreviations.....	xviii
Chapter 1. Introduction.....	1
1.1 Motivations.....	1
1.2 Impact of thermal stratification condition on airflow pattern and pollutant dispersion...	1
1.3 Problem definition.....	6
1.4 Objectives and thesis organization.....	7
Chapter 2. Literature review on the impact of non-isothermal conditions on flow and concentration fields.....	10
2.1 Introduction	10
2.2 Experimental studies	10
2.2.1 Open-environment studies	10
2.2.2 Wind tunnel measurements.....	12
2.3 Numerical (CFD) studies	16
2.3.1 Dimensionality and geometry.....	19
2.3.2 CFD approach.....	20
2.4 Conclusions for computational parameter setting.....	24
Chapter 3. Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses	26
3.1 Introduction	26
3.2 Simulation cases, boundary conditions and computational settings	26
3.2.1 Computational domain and grid resolution.....	28

3.2.2	Solution method.....	34
3.3	Results validation.....	35
3.3.1	Time-averaged flow field and TKE.....	35
3.3.2	Time-averaged temperature.....	39
3.3.3	Time-averaged and fluctuating concentration.....	41
3.3.4	ELES vs. LES: computational costs.....	43
3.4	Impact of thermal stratification on flow field, turbulence statistics, and pollutant dispersion.....	44
3.4.1	Time-averaged flow field.....	44
3.4.2	Turbulence kinetic energy.....	46
3.4.3	Frequency analysis.....	47
3.4.4	POD of the flow field.....	48
3.4.5	Concentration field.....	53
3.5	Conclusions.....	58
Chapter 4. A review of advances towards efficient reduced-order models (ROM) for predicting urban airflow and pollutant dispersion.....		60
4.1	Introduction.....	60
4.1.1	Motivation.....	60
4.1.2	Pros and cons of CFD application in urban-related problems.....	63
4.1.3	Objective and contribution.....	64
4.2	Reduced-order modeling.....	65
4.2.1	Coherent structures.....	65
4.2.2	Reduced-order models (ROMs).....	65
4.3	Dimensionality reduction.....	68

4.3.1	POD: a linear dimensionality reduction technique	68
4.3.2	The relationship between POD and autoencoder	70
4.3.3	Nonlinear dimensionality reduction.....	72
4.4	Feature dynamics.....	75
4.5	Challenges and recent advances	77
4.6	Conclusion.....	80
Chapter 5. Performance analysis of different reduced-order models for predicting urban turbulent flow field		81
5.1	Introduction	81
5.2	Methods used in NIROM development	81
5.2.1	Convolutional autoencoder (CAE)	81
5.2.2	Multiscale CAE (MS-CAE).....	83
5.2.3	Self-attention CAE (SA-CAE).....	84
5.2.4	LSTM.....	86
5.3	Dataset.....	88
5.4	NIROM training.....	89
5.5	Results and discussions	90
5.5.1	Dimensionality reduction part.....	90
5.5.2	NIROM performance evaluation	94
5.6	Conclusions and future works	109
Chapter 6. Conclusions and outlook.....		111
6.1	Conclusions	111
6.2	Contributions.....	114
6.3	Limitations and future remarks	114

Bibliography.. ..	117
Appendix A: Equations of the employed CFD models	147
Appendix B: POD mathematics.....	153

List of Figures

Figure 1.1: Schematic view of the impact of the induced buoyant force, due to surface heating, on the flow field by considering the mutual directions of the wind and heated surface, in the condition with perpendicular wind with respect to the canyon axis: (a) leeward wall is heated (unidirectional act of the background wind and the induced buoyant flow), and (b) windward wall is heated (bi-directional act of the background wind and the induced buoyant flow) (obtained from [17]).....	2
Figure 1.2: Comparison between the streamlines and the normalized Reynolds stresses ($u'w'/U^2$) under different stratification conditions: (a) $Ri = 0$, (b) $Ri = -0.1$, (c) $Ri = 0.1$, and (d) $Ri = 0.188$ (obtained from [38])	5
Figure 2.1: Comparison between the results of simulations of [134] (with the RANS, URANS and LES approaches) with the wind tunnel measurements of [64] with $Re=9000$ and $Ri=1.56$, for a 2D street canyon with the aspect ratio of 1 (obtained from [134])	22
Figure 3.1: The schematic view of the model building, located in different stratification conditions (obtained from [163]). z is the coordinate in the vertical direction. u , θ and c_{gas} represent the streamwise velocity, temperature, and gaseous pollutant concentration, respectively. <i>build</i> , <i>gas</i> and <i>f</i> are the abbreviations for building, gaseous pollutant and floor, consecutively.....	27
Figure 3.2: First row: computational domain for ELES: a) top view in $x - y$ plane, and b) side view in $x - z$ plane (orange: the model building; chromatic blue: LES zone; pale blue: RANS zone),.....	30
Figure 3.3: Ratio of time-averaged SGS kinematic viscosity to the physical one, $\langle v_{SGS} \rangle / \nu$ obtained from isothermal ELES simulation over the sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, $x/H = 1.5$, along $y/H = 0$	31
Figure 3.4: Dimensionless form of the Inflow boundary conditions, based on the experiment [162], ELES (conducted in the empty domain) at the inlet boundary (i.e., RANS zone), denoted by ELES_R, and at the RANS-LES interface (i.e., beginning of the LES zone), represented by ELES_L:	33
Figure 3.5: Comparison between profiles of dimensionless time-averaged streamwise velocity, u/U_H , obtained by present simulation (ELES), LES of [101] and experiment [162] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (“S”, “I”, and “U” represent stable, isothermal and unstable conditions, respectively)	36
Figure 3.6: Comparison between contours of dimensionless TKE, k/U_H^2 , obtained by ELES and experiment of [162] (experimental contours of TKE are obtained from [162])	38

_Toc112612578

Figure 3.7: Comparison between profiles of dimensionless time-averaged temperature, $(\langle \theta \rangle - \theta_f) / \Delta \theta$, obtained by present simulation (ELES), LES of [101] and experiment of [162] over sampling lines, $x/H = -1$, $x/H = -0.625$, $x/H = 0.0625$, and $x/H = 0.625$, along $y/H = 0$ (“S” and “U” denote stable and unstable conditions, respectively)..... 40

Figure 3.8: Comparison between profiles of dimensionless time-averaged pollutant concentration, c/C_0 , obtained by present simulation (ELES), LES of [101] and experiment of [162] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (“S”, “I” and “U” denote stable, isothermal and unstable conditions, respectively)..... 41

Figure 3.9: Comparison between profiles of dimensionless pollutant concentration fluctuations, c_{rms}/C_0 , obtained by present simulation (ELES), LES of [101] and experiment of [162] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (“S”, “I” and “U” denote stable, isothermal and unstable conditions, respectively)..... 43

Figure 3.10: Time-averaged streamlines and dimensionless time-averaged vertical velocity, w/U_H , under different thermal stratification conditions..... 45

Figure 3.11: Comparison between profiles of dimensionless TKE profiles, k/U_H^2 under different stability conditions over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ 46

Figure 3.12: First row: time histories of (a) streamwise and (b) vertical velocity components at point $x/H = 2$ and $z/H = 0.5$, along $y/H = 0$, for different thermal stratification conditions, 49

Figure 3.13: Contribution of each POD mode to total TKE for different stratification conditions 50

Figure 3.14: Iso-lines of the first four velocity POD modes in the streamwise direction along $y/H = 0$ for different thermal stratification conditions (positive values: solid orange lines; negative values: dashed blue lines)..... 51

Figure 3.15: Iso-lines of the first four velocity POD modes in the vertical direction along $y/H = 0$ for different thermal stratification conditions (positive values: solid orange lines; negative values: dashed blue lines) 52

Figure 3.16: Power spectral density (PSD) functions of the coefficients of the first four POD modes 53

Figure 3.17: Contours of time-averaged pollutant concentration, $\langle c \rangle / C_0$, under different thermal stratification conditions along $y/H = 0$ 55

Figure 3.18: Comparison between three-dimensional time-averaged and instantaneous plume shapes, obtained by iso-surfaces of $\langle c \rangle / C_0 = 1$ and $c / C_0 = 1$, respectively, colored by dimensionless turbulence kinetic energy, k/U_H^2 56

Figure 3.19: Convective flux, $\langle u \rangle \langle c \rangle$, and turbulent diffusion flux, $\langle u' c' \rangle$, in streamwise direction along $y/H = 0$ under different thermal stratification conditions	57
Figure 3.20: Convective flux, $\langle w \rangle \langle c \rangle$, and turbulent diffusion flux, $\langle w' c' \rangle$, in vertical direction along $y/H = 0$ under different thermal stratification conditions	58
Figure 4.1: Difference between linear and nonlinear dimensionality reduction methods (obtained from [268]).....	71
Figure 4.2: Difference between linear and nonlinear dimensionality reduction methods (obtained from [226]).....	73
Figure 4.3: An example of the convolutional operation with a 3×3 convolutional kernel, with no padding and strides (obtained from [275]).....	73
Figure 4.4: The schematic view of the structure of a convolutional autoencoder (modified from [266]).....	74
Figure 4.5: The schematic view of the multiscale CAE used in [311], using three different kernel sizes of $3 \times 3 \times 3$, $5 \times 5 \times 5$, and $7 \times 7 \times 7$ for dimensionality reduction of DNS data (modified from [311]).....	79
Figure 5.1: The schematic view of a 3×3 kernel, with a stride value of 2, acting on a 7×7 input data	82
Figure 5.2: The encoder part structure of the CAE. $k = 3 \times 3 \times 32$ means that the size of the convolution kernel is 3×3 , and 32 convolution kernels are used in this layer. s indicates the stride of the convolutional layer.	83
Figure 5.3: The structure of the encoder part of MS-CAE. Three different encoders with the same structure as the one used in CAE, but with different kernel sizes, i.e., 3×3 , 5×5 , and 7×7 ..	84
Figure 5.4: The structure of the self-attention module. \otimes and \oplus represent matrix addition and multiplication, respectively. C' and $c.r.$ denote the intermediate channel and the compression ratio, consecutively.	86
Figure 5.5: The structure of parallel LSTM.....	87
Figure 5.6: The data sampling region, located in the LES zone	89
Figure 5.7: Contours of the dimensionless time-averaged velocity components obtained by CFD and the models used in the present study.....	92
Figure 5.8: Comparison between the profiles of dimensionless time-averaged longitudinal velocity, u/u_H , over the training dataset, along the sampling lines, $x/H = 0.375$, $x/H = 0.625$,	

$x/H = 1$, and $x/H = 1.5$, reconstructed by different models against those obtained by the CFD study of [208].....	93
Figure 5.9: The contours of the dimensionless instantaneous longitudinal velocity component, u/u_H , reconstructed by NIROMs, and the available CFD data at different time steps [208]	96
Figure 5.10: The profiles of the dimensionless longitudinal velocity fluctuations reconstructed by the models and those obtained by the CFD simulation over $x/H = 0.375$ and $x/H = 0.625$ at three different time steps.....	97
Figure 5.11: The contours of the absolute deviation of the instantaneous longitudinal velocity components reconstructed by NIROMs, with respect to the CFD data at different time steps [208].....	98
Figure 5.12: Time history of u'/u_H obtained from the CFD simulation and the NIROMs, at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$, with related time-averaged and standard deviation values	100
Figure 5.13: Power spectrum density (PSD) function of streamwise velocity fluctuations at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$	102
Figure 5.14: The contours of the dimensionless instantaneous vertical velocity component, w/u_H , reconstructed by NIROMs, and the available CFD data [208]	104
Figure 5.15: The contours of the absolute deviation of the instantaneous vertical velocity components reconstructed by NIROMs, with respect to the CFD data at different time steps [208].....	105
Figure 5.16: Time history of w'/u_H , obtained from the CFD simulation and the NIROMs, at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$, with related time-averaged and standard deviation values	107
Figure 5.17: PSD function of vertical velocity fluctuations at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$	108
Figure 5.18: Time trace of mse of the reconstructed flow field (a) longitudinal velocity component, and (b) vertical velocity component.....	109
Figure A. 1: The energy spectrum with the test and grid filters [35].....	149

List of Tables

Table 2.1: An overview of wind tunnel studies on the impact of non-isothermal boundary layers on the flow field and the pollutant dispersion process (Ref.: reference number; Ri/Fr number: Richardson/Froude number; Heated (cooled) surf.: heated (cooled) surface; AR: canyon aspect ratio; Dimen.: study dimensionality)	13
Table 2.2: An overview of CFD studies on the impact of non-isothermal boundary layers on non-reactive pollutant dispersion (Ref.: reference number; Turb. model: turbulence model; Dimen.: dimensionality of studies).....	16
Table 3.1: Conditions for the wind tunnel experiment of [162]. (Ri , $\Delta\theta$ and θ_0 represent bulk Richardson number, the absolute value of the temperature difference ($ \theta_H - \theta_f $), and the space-averaged airflow temperature at the inflow boundary layer, respectively.....	28
Table 3.2: Boundary conditions of the computational domain.....	34
Table 3.3: Hit rate (q) and $FAC2$ of u/U_H profiles obtained by the present study (ELES) and LES of [101] with respect to the experimental data of [162], over the sampling lines along the $y/H = 0$ plane (i.e., 44 experimental data), for different stratification conditions.....	37
Table 3.4: Mean relative error percentage of $(\langle\theta\rangle - \theta_f)/\Delta\theta$ profiles obtained by the present study (ELES) and LES of [101] with respect to the experimental data of [162], over the sampling lines along $y/H = 0$, for different stratification conditions	40
Table 3.5: Mean relative error (RE) and $FAC2$ of c/C_0 profiles obtained by the present study (ELES) and LES [101] with respect to the experimental data of [162], over the sampling lines along the $y/H = 0$ plane (i.e., 44 experimental data), for different stratification conditions	42
Table 4.1: An overview of previous review articles on pollutant dispersion and urban airflow ..	60
Table 4.2: An overview of previous studies on the development of NIROM for predicting urban airflow field and pollutant dispersion	67
Table 5.1: The mse of the reconstruction of the velocity field components for different models	93
Table 5.2: The average of the directional turbulence kinetic energy percentage, E_k , of the velocity field components for different models	94
Table 5.3: The mean value of mse of the reconstruction of the velocity field components for different NIROMs.....	109

Abbreviations

ABL	Atmospheric boundary layer
AE	Autoencoder
AE-FCNN	Nonlinear fully-connected neural network
AE-linear-FCNN	Linear fully-connected neural network
CAE	Convolutional autoencoder
CCNN	Causal convolutional neural network
CFD	Computational fluid dynamics
CFL	Courant-Friedrichs-Lewy number
CNN	Convolutional neural network
CPU	Central processing unit
DDES	Delayed detached eddy simulation
DES	Detached eddy simulation
DL	Deep learning
DMD	Dynamic mode decomposition
DNN	Deep neural network
DNS	Direct numerical simulation
dof	Degrees of freedom
DSLMM	Dynamic Smagorinsky-Lilly model
ELES	Embedded large eddy simulation
GPR	Gaussian process regression
GRU	Gated recurrent unit
HVAC	Heating, ventilation, air conditioning
IROM	Intrusive reduced-order model
LES	Large eddy simulation
LSTM	Long short-term memory
MS-CAE	Multi-scale convolutional autoencoder
<i>mse</i>	Mean squared error
NIROM	Non-intrusive reduced-order model
PALM	Parallelized large eddy simulation model
PC	Principal component
PCA	Principal component analysis
POD	Proper orthogonal decomposition
PSD	Power spectral density
RANS	Reynolds-averaged Navier-Stokes
RBF	Radial basis function
ReLU	Rectified linear unit
ResNet	Residual neural network
RLZ	Realizable $k - \varepsilon$
RNG	Re-normalization group $k - \varepsilon$
RNN	Recurrent neural network
ROM	Reduced-order model

SA-CAE	Self-attention convolutional autoencoder
SGS	Subgrid-scale
SKE	Standard $k - \varepsilon$
SSLM	Standard Smagorinsky-Lilly model
SST	Shear stress transport $k - \omega$
SVD	Singular value decomposition
TKE	Turbulence kinetic energy
UHI	Urban heat island
URANS	Unsteady Reynolds-averaged Navier-Stokes
VAEDC	Variational autoencoder with deep convolutional layers
WALE	Wall-adapting local eddy-viscosity
WHO	World Health Organization
XGBoost	eXtreme Gradient Boost

Chapter 1. Introduction

1.1 Motivations

Changes in urban temperature distribution (e.g., caused by the urban heat island (UHI) formation or diurnal and seasonal changes) may significantly affect pollutant dispersion [1–4]. Alteration of the urban temperature distribution leads to the formation of different thermal stratification conditions in the atmospheric boundary layers (ABL): (1) isothermal, in which the temperature difference between the flow field and surfaces is zero, (2) stable condition occurs when the airflow temperature is higher than that of surfaces, and (3) unstable condition, which can occur when the temperature of surfaces is higher than that of the airflow. An increase in surface temperature can induce buoyant forces, which leads to local instabilities, and thus, turbulent mixing is increased [1,2]. Furthermore, in wintertime, or at night [5], when the dominant thermal stratification may become stable, considering local temperature rises can induce local instabilities.

Surfaces in urban areas are usually warmer than the surrounding airflow due to the absorption of solar radiation. The increase in the surface temperature leads to the induction of buoyant forces in the proximity of surfaces. A field measurement experiment conducted by Niachou et al. [6] during hot summer weather conditions showed that unstable stratification conditions were present in 85% of cases during the daytime and 64% of cases at night. Santamouris et al. [7] measured up to 19 °C of difference between the temperatures of the canyon's façades, which affected the airflow pattern. Such a large temperature difference can change the airflow pattern and pollutant dispersion process within an urban area. It is noteworthy to mention that the buoyant force can become a dominant mechanism in the flow field in urban airflow when the background wind velocity is less than 2.5 m/s; a situation that can occur 18% of the time in a year in low-latitude cities [8,9]. From the wind rose presented in the study of Mirzaie et al. [10], it can be found out that a significant percentage of the annual wind speed is less than 2-3 m/s in Zurich. A remarkable contribution of the induced buoyant force to pollution removal in a highly-packed high-rise city, Hong Kong, in no wind conditions, has been reported [11]. Thus, non-isothermal conditions can play a significant role in defining airflow pattern, and thus, pollutant dispersion in urban areas.

1.2 Impact of thermal stratification condition on airflow pattern and pollutant dispersion

In the present section, different articles are reviewed to fundamentally tackle the impact of non-isothermal conditions on the airflow pattern and pollutant dispersion process.

Bulk Richardson number, Ri_b , is used in atmospheric studies in order to quantify the significance of thermal effects in flow movement. The definition of Ri_b is as follows [12]:

Eq. 1-1

$$Ri_b = \left(\frac{g}{T}\right) \left(\frac{\Delta T}{\Delta Z}\right) \left(\frac{\Delta \bar{u}}{\Delta Z}\right)^{-2}$$

where, g is the gravitational acceleration, and T and ΔT represent the mean canyon temperature and mean vertical difference of air temperature, respectively. Also, $\Delta \bar{u}$ is the mean vertical difference of horizontal wind speed in the street canyon. The bulk Richardson number can be defined as the ratio between the buoyant convective flow and the shear flow. By substituting Δz with the building (or the canyon) height, H , the bulk Richardson number can be re-written as follows [13]:

Eq. 1-2

$$Ri_b = \frac{gH(T_H - T_0)}{T U_H^2}$$

where, T_H and T_0 denote the temperature at the top and bottom of the canyon, respectively. Moreover, U_H is the mean wind speed at the canyon height.

As can be seen in Eq. 1-2, the impact of the induced buoyant force on the flow field is correlated with various factors. Other than the background airflow velocity and the strength of surface heating, which can be expressed by the bulk Richardson number, the mutual directions of the wind and heated surface (see Figure 1.1), the canyon's aspect ratio, etc. [2,14–16] are also important. Furthermore, UHI mitigation strategies, such as vegetation planting and surface modification, can change the (surface) UHI intensity within an urban area, which influences the temperature distribution of an urban area.

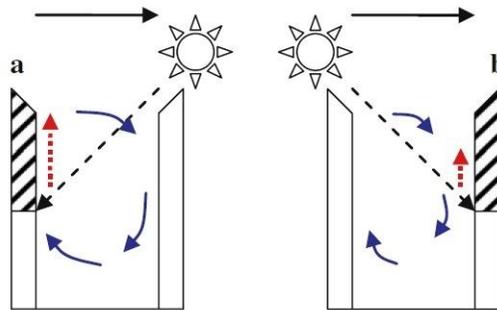


Figure 1.1: Schematic view of the impact of the induced buoyant force, due to surface heating, on the flow field by considering the mutual directions of the wind and heated surface, in the condition with perpendicular wind with respect to the canyon axis: (a) leeward wall is heated (unidirectional act of the background wind and the induced buoyant flow), and (b) windward wall is heated (bi-directional act of the background wind and the induced buoyant flow) (obtained from [17])

The dominant mechanism, among the shear flow and buoyancy-driven one, in airflow movement within a canyon can be found by using a non-dimensional parameter, called buoyancy parameter, B , defined as below [18]:

Eq. 1-3

$$B = \left(\frac{g \alpha (T_{hi} - T_{lo}) H}{u_0^2 [1 + (H/W)^2]} \right)$$

where, α is the thermal expansion of air. Furthermore, T_{hi} and T_{lo} denote the temperatures of the leeward wall (i.e., heated wall) and the windward façade (i.e., unheated wall, or the wall with lower temperature), respectively. u_0 is the characteristic background velocity, and H and W are the height and width of the canyon, consecutively. Dallman et al. [18] stated that as B exceeds a critical value, $B_c = 0.05$, the buoyancy-driven flow becomes dominant.

From Figure 1.1 (a), for the heated leeward wall, the induced upward flow is prohibited by the background flow from completely being mixed with the airflow aloft, and part of it tends to circulate within the canyon [13]. Furthermore, due to the downwash effect of the windward wall, a portion of the background airflow may also circulate within the canyon. Wang et al. [19] reported that when the leeward wall or the ground was heated, the flow structure and the dispersion pattern were similar to those in the isothermal case. However, it has been shown that by altering the surface temperature, the location of the vortex core in the wake region can be changed [20]. The mentioned change obviously affects the dispersion process. Based on the results, even weakly unstable conditions should not be neglected in dispersion studies. Also, Three-dimensional simulations have shown that by inducing thin thermal boundary layers in the vicinity of the heated surfaces, the air's entrainment into the canyon increases, and thus, the pollutant concentration decreases [21]. A similar result was reported by [22,23], which showed the entrainment of fresh air from the leeward corner at the roof level and decrement of the amount of the pollutant concentration when the ground was heated. Wang et al. [19] also stated that an opposing force is induced when the windward wall is heated, and thus, two counter-rotating forces, one caused by advection and the other caused by the buoyant force, change the flow and pollutant dispersion patterns (see Figure 1.1 (b)).

As reported by [24], under unstable conditions, due to ground heating, the recirculation region was destroyed due to the formation of an upward buoyancy-driven flow. The mentioned upward flow increased the interaction of the canyon flow with the airflow aloft. Mixing the canyon flow with the flow above it has also been reported by the experiment of [17].

Two fluxes are involved in the pollutant dispersion process: (1) convective flux, which is defined as the product of time-averaged velocity and time-averaged concentration values ($\langle u_i \rangle \langle c \rangle$), and (2) turbulent diffusion flux, defined as the average of the product of the fluctuations of velocity and concentration ($\langle u_i' c' \rangle$) over time [25]. In dense urban areas, inside street canyons, where the airflow velocity is small, the role of the induced buoyant force in flow movement increases. Li et al. [22] reported that increasing the instability level by ground heating enhances the flow advection within

a canyon, which results in an increase in the contribution of the convective flux to the dispersion process.

Uehara et al. [26] reported a positive relationship between the turbulence intensity and the level of flow instability. The results of [27] also show a significant increase in the turbulence kinetic energy (TKE) value near the heated wall. Generally, as mentioned by [28], in the case of unstable stratified conditions, buoyancy becomes important in the TKE production, and thus, the transfer of momentum. An increase in the TKE level leads to an increase in the concentration fluctuations, which results in a shorter plume shape in the wake region of a building [29]. Furthermore, based on the results of [22], the roof-level turbulent diffusion, which is a result of the mixing of the shear flow and vertically-induced buoyant one, increases as the canyon ground is heated. Therefore, due to the increase in the contribution of turbulent diffusion flux to the dispersion process with changing the stratification condition from stable to unstable, the pollutant dispersion process becomes stronger [30].

Pulvirenti and Di Sabatino [31] showed that the effect of turbulent diffusion fluxes in pollutant transportation was significant for almost all the Richardson numbers investigated. However, it has been reported that as the stratification condition became more unstable, the ratio of the amount of pollutant transported by the mean flow to the turbulent flux became larger [32]. Changes in the share of each flux in pollutant transportation can be considered as a result of the generation of the secondary vortex in the canyon due to the buoyancy effect. It is noteworthy to mention that the role of the studied geometry is of utmost importance in defining the dominant dispersion mechanism. Thus, generalizing the results to other geometries should be done carefully.

The stable stratification condition has always attracted a lot of attention (e.g., [33,34]) since it can cause pollution to be accumulated in the affected regions. The reason for the mentioned phenomenon is that the stable stratification condition causes the turbulence generation to be suppressed [35–37]. Stable stratification condition causes a fluid parcel to get back to its equilibrium state after a vertical displacement, which results in the energy extraction from $\overline{w'^2}$ (buoyant destruction) [37], where w' is the velocity fluctuation in the vertical direction. Furthermore, as mentioned by [37], the energy is extracted from the turbulence generated from shear stress ($\overline{u'w'}$) due to the excessive energy needed to overcome buoyancy (u' is the velocity fluctuation in the longitudinal direction). Therefore, the contribution of the turbulent diffusion flux to the dispersion process decreases. It should be noted that, as observed experimentally [36,37], weakly stable conditions do not significantly change the turbulence structure and statistics of the flow. Figure 1.2 shows that although the Reynolds stress, $\overline{u'w'}$, for the weakly stable conditions is weaker than that for the isothermal and unstable conditions, its value has not changed significantly compared to the isothermal condition [38].

It has been reported that the impact of unstable thermal stratification on the concentration field was more pronounced than that of the stable one with its counterpart Richardson number [39]. However, stable stratification conditions can also change the flow and concentration fields within a street canyon. Figure 1.2 shows a comparison between the streamlines and normalized Reynolds stress ($\langle u'w' \rangle / U^2$) under the isothermal, unstable ($Ri = -0.1$), and two stable ($Ri = 0.1$ and 0.188) conditions within a canyon [38]. In Figure 1.2 (c) and (d), a stagnant region is visible in

the proximity of the ground, which can trap pollution. A decrease in the Reynolds stress, which can be seen in Figure 1.2 (c) and (d) in comparison with the other two cases, can weaken the pollutant removal. Also, Duan and Ngan [24] compared the Reynolds stress in an urban area for different stratification conditions, and the results were in-line with those obtained by [38].

It has been reported that coherent structures and airflow patterns around an isolated building are quite similar to each other under the isothermal and stably stratified boundary layers, with some differences in the magnitude of the streamwise velocity [40]. Therefore, it can be concluded that the overall pattern of pollutant dispersion can be quite similar for the different cases, as is the case with [40]. However, small differences in the width of the polluted area in the wake region and also the concentration magnitude for each case at different locations are observable in their reported results.

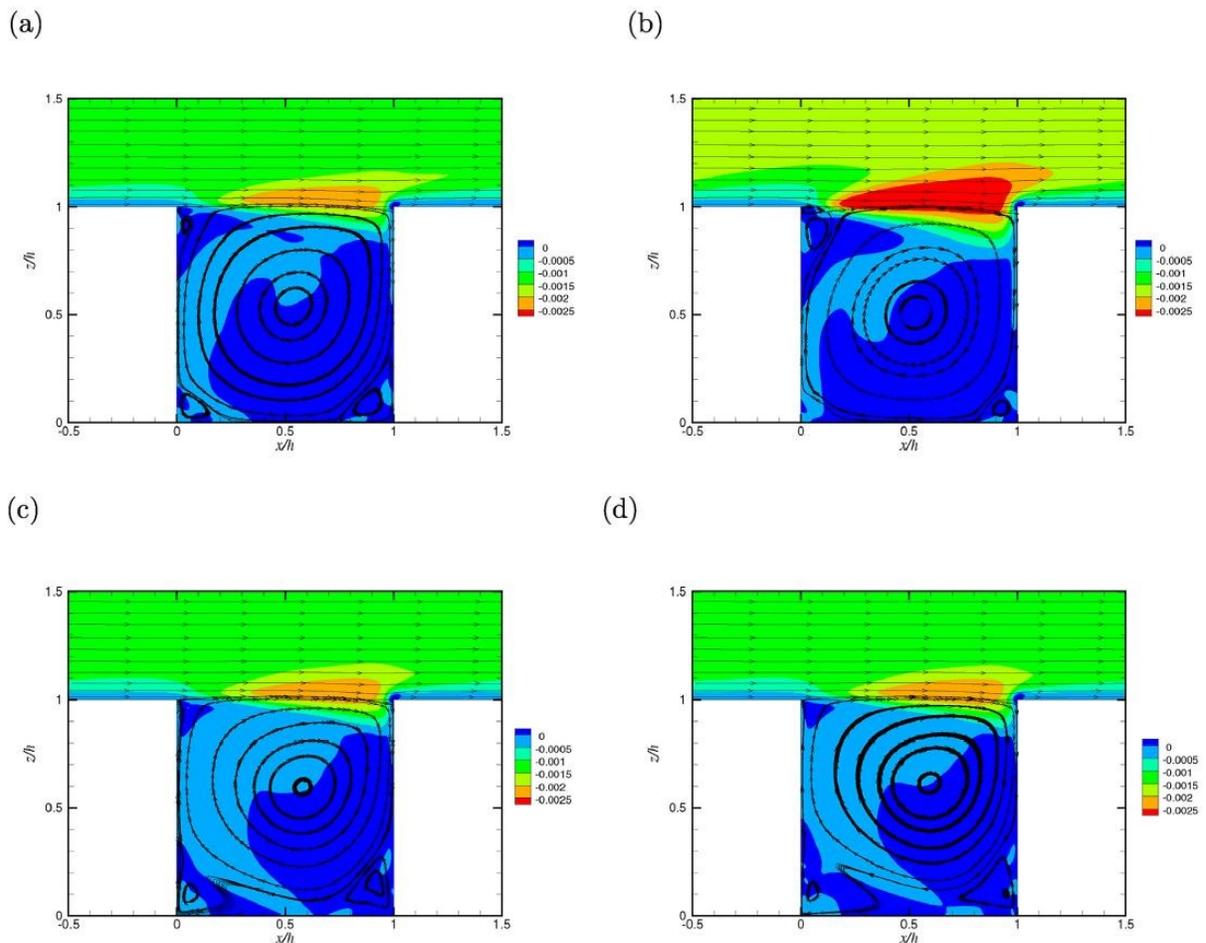


Figure 1.2: Comparison between the streamlines and the normalized Reynolds stresses ($\langle u'w' \rangle / U^2$) under different stratification conditions: (a) $Ri = 0$, (b) $Ri = -0.1$, (c) $Ri = 0.1$, and (d) $Ri = 0.188$ (obtained from [38])

The growth of boundary layers has been reported to be slower in the stably stratified condition compared to the isothermal case [41]. The mentioned phenomenon caused the airflow to be entrained farther downstream in the stably stratified condition in comparison with the other case. The influence of the stably stratified condition on increasing the pollutant concentration in the downstream canyons was reported by [42] and also compared with the isothermal and unstable conditions.

In addition to the effect of the temperature of surfaces on generating different stratification conditions (i.e., local stratification), considering the thermal stratification of the approaching flow (i.e., boundary layer) is also of utmost importance. By using wind tunnel measurements, Marucci and Carpentieri [43] reported that the stable condition caused the pollutant concentration to be increased by up to two times compared to the isothermal case, while the unstable one reduced the pollution concentration by up to three times. In-line with the findings of the mentioned study, Sessa et al. [44] stated that the stable condition increased the pollutant concentration in the tested urban area.

1.3 Problem definition

Based on the above explanations, it can be concluded that the occurrence of different thermal conditions is quite frequent within urban areas, which can be influenced by other factors, such as wind speed and direction. These meteorological factors frequently change through time, and the airflow and concentration fields can constantly be under the influence of the mentioned factors. Therefore, it would be reasonable to account for these alterations in simulating the field variables. Furthermore, when dealing with the concentration field, monitoring the situation becomes important. For instance, in case of sudden release of pollutants upstream of an urban area, it would be essential to monitor the concentration field. Also, long-term simulations are required to assess the ventilation performance of an urban area [45].

Experimental and numerical (e.g., computational fluid dynamics, CFD) approaches are among the most common ones for studying the flow and concentration fields within urban areas. Characteristics of these approaches make them suitable for being used to achieve various goals. For instance, experimental methods can be used to discover the underlying physics of the problem. Also, they can be used to validate the results of CFD simulations. On the other hand, CFD simulations can generate high-resolution data, which can be used for a comprehensive analysis of the problem. However, conducting experimental and CFD studies is time-consuming, which makes them unsuitable for the tasks mentioned in the previous paragraph. Data-driven methods can obviate the limitations regarding the simulation time.

Data-driven algorithms have emerged in fluid dynamics applications for different purposes, including turbulence closure, super-resolution, and reduced-order modeling, which are described in detail in [46,47]. The idea behind reduced-order modeling is to reduce the degree of freedom of a system to achieve significant speed-up. This order reduction in fluid dynamics problems can be performed by considering the coherent structures of field variables (e.g., the flow field), which is thoroughly discussed in section 4.2 of the present dissertation. Therefore, by providing datasets

using CFD simulations, data-driven models can be trained to reduce computational costs, which makes them suitable for the aforementioned tasks. It should be noted that some efforts have recently been made to use sparse experimental data for training a data-driven model [48–50].

1.4 Objectives and thesis organization

The present dissertation employs numerical and data-driven approaches to:

1. Fundamentally analyze the impact of thermal stratification conditions on the urban airflow field, pollutant dispersion process, and turbulence characteristics of the airflow field, and,
2. Develop a fast-response model, i.e., reduced-order model (ROM), for prediction tasks using datasets generated by the CFD simulations.

It should be noted that, as mentioned in section 1.3, having access to fast-response models for urban areas is of utmost importance since:

1. The monitoring task is essential due to the occurrence of emergency situations, such as sudden pollution release from an industrial plant located upstream of an urban area,
2. Long-term simulations may be required (for instance, to assess the ventilation performance within an urban area),
3. Capturing frequent changes in the boundary conditions (i.e., meteorological factors, such as alteration of the thermal stratification condition, inflow wind speed, and wind direction), may be needed.

Thus, in the present dissertation, after validating the CFD results with available experimental data, various non-intrusive ROMs (NIROMs) are developed for predicting the turbulent urban airflow. The outline of each chapter of this dissertation is mentioned below.

Chapter 2 of the present thesis comprehensively reviews experimental and numerical studies considering the impact of non-isothermal conditions on the airflow field and pollutant dispersion. Although this study is conducted numerically, reviewing experimental research works is of utmost importance since they shed light on the fundamentals of the process under investigation and can be used for numerical models, which are crucial in the present study. Therefore, being aware of the conducted experimental studies and their reported results is beneficial for finding an appropriate test case for numerical ones. Thus, these studies provide opportunities for researchers to build up a roadmap for their simulations' goals.

Chapter 3 introduces the simulation test cases: stable, isothermal, and unstable conditions. In this chapter, for the first time, the applicability of embedded large eddy simulation (ELES), also known as zonal large eddy simulation, in studying the airflow field and pollutant dispersion under non-isothermal conditions is shown. After validating the simulations' results with available experimental data, the generated datasets are used for proper orthogonal decomposition (POD) and spectral analyses to study the impact of thermal conditions on the turbulence statistics of the urban airflow field. The POD technique is considered the most-common dimensionality reduction

technique in fluid dynamics problems. Therefore, the results obtained from the POD analysis show the feasibility of generating a ROM for the present test cases.

Chapter 4 introduces different techniques for developing NIROMs in fluid dynamics applications by thoroughly reviewing the existing studies. The NIROM development procedure is divided into two parts, and the limitations of the frequently-used approaches are mentioned. The scarcity of studies on the development of NIROMs for highly turbulent flow fields leads to the development of different NIROMs in Chapter 5, in order to try to obviate the limitations of the most-common approaches by modifying them based on the physics of the fluid dynamics problems, i.e., multiscale nature of turbulence and long-range dependencies that might be present among datapoints. The employed models are introduced in this chapter, and the results are comprehensively analyzed to prove that the developed NIROMs can be considered a reliable approach for CFD simulations when fast-response models are required.

The present dissertation is a manuscript-based thesis, in which the contents of chapters 3 to 5 are part of the published and under review manuscripts. Furthermore, chapter 2 is the longer and more comprehensive version of the introduction part of the article used in chapter 3. Also, the introduction and the simulation case sections of the article used in chapter 5 are already addressed in chapters 4 and 3, respectively. Therefore, the mentioned sections are modified for the sake of the dissertation flow.

Chapter 2:

The chapter describes the influence of non-isothermal thermal conditions on the urban airflow and pollutant dispersion process by reviewing experimental and numerical articles. Furthermore, the limitations of previous works are mentioned.

- Chapter 3:

Masoumi-Verki, Shahin, Gholamalipour, Payam, Haghigat, Fariborz, Eicker, Ursula. “Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses”, *Building and Environment* 206 (2021), 108388.

- Chapter 4:

Masoumi-Verki, Shahin, Haghigat, Fariborz, Eicker, Ursula. “A review of advances towards efficient reduced-order models (ROM) for predicting urban airflow and pollutant dispersion”, *Building and Environment* 216 (2022), 108966.

- Chapter 5:

Masoumi-Verki, Shahin, Haghigat, Fariborz, Eicker, Ursula. “Performance analysis of different reduced-order models for predicting urban turbulent flow field”, *under review*.

The present dissertation also includes two appendices. Appendix A explains the governing equations of the CFD models used for the simulations, while Appendix B mentions the mathematics behind the POD technique. It should be noted that Appendix B is similar to the appendix section of the article used in Chapter 3.

Chapter 2. Literature review on the impact of non-isothermal conditions on flow and concentration fields

2.1 Introduction

In this chapter, the experimental and numerical studies investigating the impact of non-isothermal thermal conditions on the urban airflow and pollutant dispersion process are reviewed. The results are analyzed, and the limitations of the numerical studies are addressed.

2.2 Experimental studies

Experimental methods in urban studies consist of two major approaches: scaled laboratory measurements (i.e., wind tunnel and water tunnel measurements) and field measurements, which, itself, can be categorized into two classes, namely, reduced-scale and full-scale outdoor measurements. Torres et al. [51] also classified experimental studies into two categories: open-environment, consisting of reduced-scale and full-scale measurements, and closed-environment, which is similar to scaled laboratory measurements. In the present section, outdoor field measurements and wind tunnel measurements are discussed.

2.2.1 Open-environment studies

Full-scale field measurements are not usually suitable for investigating the impact of a particular factor (i.e., parametric study) on a process, e.g., dispersion process, since the boundary conditions are neither controllable nor repeatable [52]. Furthermore, field measurement studies are usually time-consuming and costly. However, field measurements provide valuable information on processes within an urban area, considering its complexities and influential factors, as is the case with BUBBLE [53], which was conducted in the city of Basel, Switzerland, to study boundary layer characteristics. Contrary to full-scale outdoor measurements, scaled outdoor measurements (e.g., [18,54]) provide researchers with the opportunity of parametric studies, without having trouble with the law of similitude, which is hard to maintain in other scaled studies, such as wind tunnel, when thermal effects are considered [54]. In the following paragraphs, some outdoor field measurement studies are introduced to find out the role of thermal stratification conditions in urban areas.

Niachou et al. [55] reported that the temperature difference between the opposite sides of the investigated canyons (i.e., the two façades) was not significant. However, the temperature difference between the ground (particularly near the windward wall) and the airflow was remarkable (up to 30°C), which resulted in the induction of strong buoyant forces. The Nantes'99 experiment [56] was conducted in a three-lane one-way traffic street, with an aspect ratio of 1.4. It should be noted that the height of the leeward wall has mentioned to be 3.4 m taller than that of

the windward façade. The pollution concentration (carbon monoxide, CO), temperature, flow and turbulence conditions were measured during hot seasons with low wind speed. Furthermore, vehicle-induced turbulence was also measured. It should be noted that the pollutant concentration was measured at the pedestrian level, 1.5 m.

Using the data of the Nantes'99 experiment, Louka et al. [57] reported that thin thermal boundary layers were formed in the proximity of the windward and leeward walls in the morning and the afternoon, respectively. These thermal boundary layers were responsible for generating a strong buoyant force close to the wall, which may affect the flow field and pollutant dispersion. Due to the absence of wind velocity measurements at the location of temperature measurements in the Nantes'99 experiment, studying the induced thermal effects may become difficult and conclusions may be erroneous.

High-resolution wind field measurements in a street canyon, with high geometrical complexity, called Joint Urban 2003, were conducted in Oklahoma City, and the effect of stability condition on the flow and turbulence statistics was investigated [58]. They reported that the impact of stratification conditions on the flow and turbulence statistics was highly dependent upon the variables used for normalization. For instance, they stated that by choosing the wind speed at the mean height of the buildings as the reference velocity, only minor influences of stability conditions could be drawn. On the other hand, when the wind velocity at a higher elevation was selected as the reference one, the impact of thermal stratification became remarkable. It should be noted that due to the high variability in the heights of the buildings, the influence of aspect ratio cannot be interpreted from the results. Also, the sensor was placed in just two different positions at different sides of the canyon. Using the data of Joint Urban 2003, Zajic et al. [59] stated that dynamically stable conditions did not happen in the canyon, and the thermal characteristics in cities could be highly affected by the built environment.

In another study, using data from the project “Innovative Laboratory for Research and Education in Urban Meteorology (ILREUM)”, conducted on the campus of the University of Oklahoma and two urban canyons in Oklahoma City, Klein and Galvez [60] reported that the upwind stability condition significantly affected TKE in the shear-layer region at the roof level. The results show TKE decrement with the increase in the stability level. The same trend was also reported for TKE values inside the canyon. However, they reported that the impact of upwind stability conditions on the mean flow pattern inside the canyon was quite small. It should be noted that TKE and mean flow velocity were normalized with the friction velocity. Also, the aspect ratio of the target canyon was reported to be 3.5, in which the flow can be considered as the isolated roughness one [61], in which thermal effects are less evident than the other types of flows.

Using a quasi-2D canyon with an aspect ratio of unity, Aliabadi et al. [5] conducted experiments to investigate the impact of several factors on flow and temperature dynamics, including thermal stability, wind direction and wind speed, in Guelph, Ontario, Canada. They observed stable stratification conditions at night and unstable conditions during the mid-afternoon. The results show a gradual decline of TKE with an increase in the value of Ri_b .

As mentioned earlier, reduced-scale outdoor measurements provide researchers with the opportunity of parametric studies. Dallman et al. [18] used ship containers to create a simplified street canyon for scaled outdoor measurements, considering the leeward wall as the heated one. By defining a dimensionless number, called buoyancy parameter (see Eq. 1-3), they stated that in very high ranges of B compared to B_c , the flow becomes independent of the background airflow velocity. Furthermore, the fluctuating component of velocity was reported to be dependent upon B when the buoyant force became more significant. It has been reported that turbulent mixing increased with the increase in the value of B [54].

The canyon aspect ratio plays an important role in defining the airflow pattern [62,63]. By using scaled outdoor measurements, Chen et al. [54] reported that the critical value of the buoyancy parameter varies with the canyon aspect ratio. It should be noted that they neglected the case with the bi-directional act of the background wind and the induced buoyant force. The higher the aspect ratio, the lower the critical buoyancy parameter becomes [54]. Thus, in deep canyons, even a small temperature difference between the airflow and surfaces may strongly affect the urban airflow and pollutant dispersion.

Based on the reviewed articles, it can be concluded that in most of the studies, the pollutant dispersion process has been neglected. Therefore, there is a lack of study on outdoor measurements on the impact of thermal stratification on pollutant dispersion. Furthermore, in order to have an accurate investigation on the effect of local instabilities on the airflow pattern and dispersion process, surface temperature should be measured in different locations since, as reported by [64], 81% and 94% of the variation in the daytime and nighttime air temperature can be explained by the surface temperature at the time, respectively.

2.2.2 Wind tunnel measurements

Contrary to full-scale field measurement studies, closed-environment ones, such as those using wind tunnels, are useful for parametric studies [65]. However, one of the most critical issues of using wind tunnels in urban studies is the complexity of holding the law of similitude, especially in the non-isothermal condition [66]. Detailed information on the Reynolds and Richardson similarity criteria can be found in [66–72].

An overview of the wind tunnel studies on the impact of non-isothermal boundary layers on the airflow field and pollutant dispersion in urban areas is given in Table 2.1.

Table 2.1: An overview of wind tunnel studies on the impact of non-isothermal boundary layers on the flow field and the pollutant dispersion process (Ref.: reference number; Ri/Fr number: Richardson/Froude number; Heated (cooled) surf.: heated (cooled) surface; AR: canyon aspect ratio; Dimen.: study dimensionality)

Ref.	Ri/Fr number	Heated (cooled) surf.	Geometry (canyon AR)	Dim.	Findings/goals
<i>Studies without pollutant source</i>					
[26]	Bulk Ri ranging from 0.79 (stable) to -0.21 (unstable)	Ground	Building arrays - AR=1	3D	Unstable (stable) stratification led to stronger (weaker) cavity eddy compared to neutral case - high stability led to zero wind speed in canyon
[27]	Fr ranging from 0.27 to 2.03	WW	Isolated canyon – AR=1	2D	Heating windward wall led to generation of a weak secondary flow in the proximity of ground
[69]	Bulk Ri=0.9 and 1.6 (unstable)	LW	Isolated building block	2D	Impact of Ri on recirculation region in wake region
[66]	Fr ranging from 0.65 to 17.3	WW, LW, ground	Isolated canyon - AR=1	2D	Effect of heated surface on the main recirculation region in the canyon
[73]	Fr ranging from 0.3 to 21.9	Ground	Street canyon placed in an urban area - AR=1 (considered non-uniform buildings heights and non-equal buildings lengths)	3D	Non-uniform buildings heights led to improvement of canyon ventilation since lateral airflow became more significant
[2]	Bulk Ri ranging from 0 to 1.25 (unstable)	WW, LW, ground	Street canyon placed in a realistic city geometry - AR=0.68	3D	As Ri increased, TKE increased - significant effect of building height on buoyancy
[16]	Bulk Ri=0.131 and 1.176 (unstable)	WW, LW, ground, roof	Street canyon placed within a number of blocks - AR=0.67, 1, 2	2D	Impacts of heated surface and canyon AR on flow and temperature fields

Studies with pollutant source

[74]	Bulk $Ri = -0.25$ (unstable)	Ground	Isolated high-rise building – pollutant was injected from a point source located on the ground in wake region	3D	Experimental data used for validating different CFD approaches (LES and RANS)
[75]	Bulk Ri ranging from 0 to 4.77 (unstable)	Ground between two buildings	One high-rise building and one low-rise in its wake, with a stack on the roof of the low-rise one	3D	Experimental data used for validating the simulations with two different turbulence models and two different wall treatment models
[30]	Bulk $Ri = 2.3$ (stable) and bulk $Ri = -3$ (unstable)	Ground	Building arrays with $AR = 0.96$ or 1.91 – pollutant was injected from a point source located on the ground	3D	Significant effect of stratification on pollutant diffusion even in the proximity of the source
[76]	Fr ranging from 0.31 to 0.62	WW, LW	Arrays of street canyons with $AR = 1$ and 1.5 – pollutant was injected from a line source located on the ground	2D	Impact of WW heating was reported to be more significant than LW heating in pollutant removal from square cavities – WW heating resulted in lowering air quality in narrow canyons
[67]	Bulk Ri ranging between 0 and 1.15 (unstable)	Roof of the low-rise building	One high-rise building and one low-rise in the wake of that, with a stack on the roof of the low-rise one	3D	Increasing Ri more than 0.78 did not affect pollutant concentration significantly
[77]	Bulk Ri ranging from -0.23 (unstable) to 0.29 (stable)	Ground	Building arrays with $AR = 1$ – pollutant was injected from a line source located on the ground	3D	Pollutant concentration was increased with the increase in bulk Ri – the stability effect ratio was independent of the measurement location
[33]	Bulk Ri ranging from -1.5 (unstable) to 0.29 (stable)	Ground	Array of buildings with $AR = 1$ – pollutant was injected from a circular point source located on the ground – wind direction was 45 degree deviated from the canyon axis	3D	Significant impact of stratification on vertical dispersion – unstable (stable) stratification decreased (increased) pollutant concentration remarkably

Abbreviations: WW: windward wall; LW: leeward wall

From Table 2.1, it can be seen that plenty of the experiments have been conducted in 2D format. However, Allegrini [73] reported that buoyant force can induce 3D flow structure, which prevents the formation of standing vortices within canyons. Therefore, in addition to the flow in the vertical plane, that in the horizontal plane is also important, particularly under non-isothermal flows.

Furthermore, in most of the studies, urban geometry has been simplified to a number of buildings (or even an isolated building) or an isolated canyon. However, in some studies building arrays have been used to model complexities of an urban area, such as the sheltering and channeling effects. Uehara et al. [26] used several cubes to model a city. They used a laser Doppler anemometer (LDA) and a cold wire for measuring the wind velocity components and temperature within and above the canyon. The results show an increase in the value of the vertical velocity component in the proximity of both the leeward and windward walls. They reported that, under the unstable case, due to the increase in the mixing process, the gradients of the vertical temperature became smaller, and thus, the instability decreased.

The impact of uniform wall heating on the airflow field was investigated by Kovar-Panskus et al. [27]. They reported that with the decrease in the Froude number's value (i.e., the inverse of the Richardson number), the tendency to generate the secondary vortex became stronger. In their results, there is no evidence of having the upward dominating buoyant force which may be due to the generation of a very thin boundary layer in the proximity of the windward wall and data collection far from the wall, outside of the thin boundary layer.

Using LDA and a cold wire thermometer, Kanda and Yamao [30] studied the impact of thermal stratification conditions by controlling the temperature of the ground and airflow on the dispersion process. The results illustrate that, inside the canyon, the stratification condition did not change the Reynolds stress, $\overline{u'w'}$, significantly, and its value is very small for all the conditions. However, above the canopy, the magnitude of $\overline{u'w'}$ became very larger compared to its magnitude inside the canyon, and it had a positive relationship with the increase in the instability level. The results of the concentration field also show that as the stability level increased, the concentration level also increased. Furthermore, in agreement with [36], pollution distributed more horizontally under the stable condition compared to the other two situations. Ogawa et al. [36] stated that for strong stable conditions, instead of vertically-wise dispersion, the plume spread in a very thin horizontal layer as its vertical movement was inhibited.

Based on the wind tunnel experiment of Ogawa et al. [36], increasing the stability level increases the generated shear in the vicinity of the ground. However, the velocity gradients for the strong stable condition were reported to be smaller than those for the weak stable condition above a specific height from the ground. Also, the results show that as the stable stratification condition became stronger, the mean airflow velocity in the longitudinal direction became smaller, particularly with the increase in height from the ground.

Marucci and Carpentieri [43] investigated the combined effects of the stably stratified approaching flow profile and local stratification caused by surface heating on pollutant dispersion from a point source. In terms of TKE, when the windward wall was heated, the generated TKE within the canyon was reported to be the most significant. In contrast, in the case of the heated leeward wall,

there was no remarkable increase in the generated TKE. Furthermore, it has been reported that the stably stratified approaching flow did not significantly affect the turbulent diffusion flux. However, it influenced the convective flux, mainly by increasing the magnitude of the negative flux close to the windward wall [43].

In most wind tunnel studies, the impact of vegetation coverage on the flow field and dispersion process has been neglected. Some wind tunnel studies focus on the aerodynamic effects of trees on the flow field and pollutant dispersion [78–80]. However, their thermal effects have been overlooked. One of the main barriers to conducting such studies is the lack of ability to model differential wall heating, similar to solar radiation, considering the shading effect of trees and buildings. As can be seen in Table 2.1, all the studies have been done by controlling the temperature of surfaces and airflow, and also the surface heating is uniform. Lin et al. [81] conducted a wind tunnel experiment to investigate the possibility of applying artificial light as radiation. They stated that this method can provide a better representation of the situation regarding the solar angle and shading effects. However, there exist some difficulties, such as air temperature measurements, in this situation. Researchers can focus on the possibility of using such an approach to model differential wall heating for future studies.

2.3 Numerical (CFD) studies

Table 2.2 shows an overview of CFD studies on the effect of non-isothermal boundary layers on pollutant dispersion. In the current section, the limitations of the numerical studies are addressed based on the indicators mentioned in Table 2.2.

Table 2.2: An overview of CFD studies on the impact of non-isothermal boundary layers on non-reactive pollutant dispersion (Ref.: reference number; Turb. model: turbulence model; Dimen.: dimensionality of studies)

Ref.	CFD approach	Turb. model	Geometry	Dimen.
[82]	RANS	SKE [83]	Isolated canyon	2D
[84]	RANS	SKE	Isolated canyon	2D
[21]	RANS	RNG	Isolated canyon	3D
[14]	RANS	SKE	Isolated canyon	2D
[85]	RANS	SKE	Isolated canyon	2D

[86]	RANS	RNG [87]	Isolated canyon	2D
[88]	RANS	SKE	Realistic urban area	3D
[32]	RANS	RNG	Isolated canyon	2D
[23]	LES	1-eq. SGS model	Isolated canyon	3D
[22]	LES	1-eq. SGS model	Isolated canyon	3D
[19]	RANS	RNG	Street canyons	2D
[74]	RANS, LES	SKE, SSLM [89]	Isolated building	3D
[90]	LES	SSLM	Isolated canyon	3D
[91]	LES	1-eq. SGS model	Isolated canyon	3D
[39]	LES	SSLM	Arrays of buildings - realistic	3D
[92]	RANS	RLZ [93]	Four buildings	3D
[94]	RANS	RNG	Street canyons	2D
[75]	RANS	SKE, RLZ	Cavity	3D
[38]	LES	1-eq. SGS model	Isolated canyon	3D
[95]	RANS	RNG	Street canyons	2D
[41]	LES	Dutch atmospheric large eddy simulation (DALES) [96]	Arrays of buildings	3D

[97]	RANS	SKE	Canyons with different aspect ratios and roof shapes	2D
[31]	LES	1-eq. SGS model	Isolated canyon	3D
[40]	LES	Meneveau's Lagrangian-dynamic Smagorinsky Model [98]	Different generic configurations - Arrays of buildings	3D
[99]	LES	SSLM	Arrays of buildings	3D
[100]	LES	PALM model [101]	Arrays of buildings	3D
[102]	LES	PALM model	Arrays of buildings	3D
[103]	LES	DSLML [104,105]	Isolated building	3D
[106]	RANS	RNG	Street canyons	2D
[107]	RANS	SKE	Arrays of buildings	3D
[108]	RANS	SKE	Street canyons	2D
[109]	RANS	SST	Street canyons	2D
[42]	RANS	Modified SKE	Arrays of buildings	3D
[110]	RANS	RNG	Street canyons	2D
[111]	RANS	SKE	Isolated building	3D
[112]	RANS	RNG	Arrays of buildings	3D

[113]	RANS	RNG	Isolated street canyon with different heights of buildings	3D
[20]	RANS	RNG	Isolated building	3D
[114]	RANS	SKE, RNG	Isolated building	3D

Abbreviations: RANS: Reynolds-averaged Navier-Stokes; LES: large eddy simulation; DES: detached eddy simulation; SKE: standard k- ϵ ; RNG: renormalization group k- ϵ ; RLZ: realizable k- ϵ ; SSLM: standard Smagorinsky-Lilly model; DSLM: dynamic Smagorinsky-Lilly model; 1-eq. SGS model: one-equation subgrid-scale model; WMLES: wall-modeled large eddy simulation; DDES: delayed detached eddy simulation; SA: Spalart-Allmaras; PALM: parallelized large eddy simulation model; SST: shear stress transport (SST) k- ω

2.3.1 Dimensionality and geometry

As can be seen in Table 2.2, several studies have used 2D street canyons [94,95,106]. Chen et al. [115] stated that the airflow structure in 2D and 3D studies, especially under an oblique wind direction, is notably different from each other. Therefore, in order to have a reliable simulation of microscale pollutant dispersion in urban areas under thermal effects, it is necessary to conduct a 3D simulation.

Plenty of previous studies have performed simulations using an isolated street canyon or building [38,116], as reported in Table 2.2. This can be due to the lack of experimental data in neighborhood-like geometries. However, considering an isolated canyon or building may neglect some of the urban airflow features, such as the sheltering effect. Therefore, the airflow pattern in an isolated canyon may be remarkably different from that in a more realistic area, e.g., building arrays [115]. Also, in cases with an isolated building, the strong downwash effect caused by the downstream building is neglected. Furthermore, in the majority of the tabulated studies, researchers used a generic geometry without considering the complexities of a real urban area. Allegrini and Carmeliet [117] showed that the UHI is a local phenomenon that varies in an urban area due to building geometry, surface materials, and wind direction and speed. Thus, the dispersion process, particularly under non-isothermal boundary layers, can be influenced by the canyon aspect ratio as well. On the other hand, the use of an isolated building or canyon in studies is common since it provides researchers with the opportunity of focusing on the fundamentals of a particular process, a specific feature of a simulation approach, etc.

The impact of buildings' heights on the airflow temperature and air quality has been investigated using two step-up and two step-down canyons as the test cases [113]. The results show a significant effect of the height of the buildings on both airflow temperature and air quality. They stated that when the incoming wind speed is low and when the windward wall was heated, an increase in the height of the upstream building led to the strengthening of the natural convection. Therefore, due to the formation of convergent flows, the averaged wind velocity at the pedestrian level increased

significantly. Several other studies considered the impact of variations in the building's height on the ventilation performance and the dispersion process within an urban area [118–121].

Yang et al. [122] claimed that urban density has a remarkable influence on the airflow and concentration fields. These factors also affect the buoyant force, and thus, the airflow pattern in urban areas. Xie et al. [39] simulated pollutant dispersion in a realistic site in non-isothermal conditions, considering two different wind directions. The results show that the concentration field is sensitive to the stratification condition, even a weak one, under oblique wind directions. The results show that the value of mean concentration under different stratification conditions can be different by up to an order of magnitude. Therefore, it can be concluded that the airflow and dispersion patterns in a realistic urban area can differ significantly from those in a generic one, especially when thermal effects are considered.

2.3.2 CFD approach

In order to use CFD for simulating the impact of non-isothermal boundary layers on pollutant dispersion, two different aspects should be considered: (1) the ability of the model in capturing the concentration field, and (2) its viability under non-isothermal conditions. This means that the model should be able to capture the mean and fluctuating variables under thermal effects.

Many efforts have been made to determine the capability of CFD models in predicting the airflow and pollutants concentration fields under the isothermal condition [123–128]. The RANS approach has extensively been used due to its affordable computational costs. However, with the advent of high-performance computers, the use of LES has become more common. A comprehensive comparison between RANS and LES in urban-/building-related fields can be found in [129].

Usually, RANS models over-predict the concentration field [124]. On the other hand, fluctuations due to large-scale eddy motions can be captured by LES [124,130]. Capturing fluctuations of the airflow and concentration fields is essential in dispersion studies since it is related to pollutant dispersion mechanisms (i.e., turbulent diffusion flux) [107,116,131]. It has been mentioned that LES can generate acceptable results, regarding the mean flow field and TKE, compared to the experimental data even with a simplified subgrid-scale (SGS) model, since the vortex shedding around the building can be captured by LES, while RANS cannot reproduce it [124].

SSLM-LES showed a better agreement (up to 4 times better) with experimental data than RNG, especially in the proximity of the canyon floor, in simulating the airflow and concentration fields within a 3D generic street canyon, under the isothermal condition [132]. Using a part of the downtown of Montreal as the test case, Gousseau et al. [133] reported that DSLM-LES performed better than the SKE in predicting the concentration field.

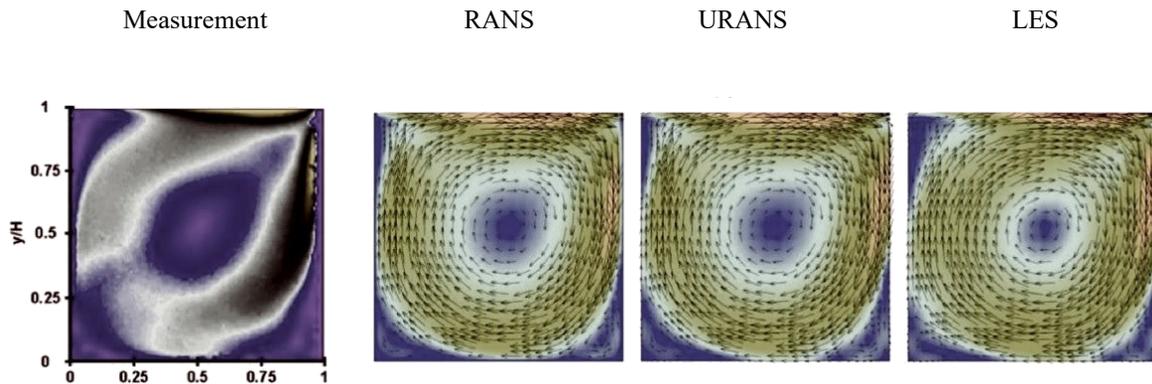
Jadidi et al. [128] compared the capability of different unsteady methods, namely, unsteady RANS (URANS), using SST, scale-adaptive simulation (SAS) [134], and SSLM-LES, in predicting the airflow and pollutant concentration fields around an isolated building in the isothermal condition. It has been reported that SST could not predict the time-dependent concentration distribution [128]. Furthermore, the results showed the incapability of SST in capturing the turbulence coherent

structures in the wake region of the building, while the other two models performed better. The authors also compared the three models regarding their computational costs. SAS showed 67% increase in the computational costs compared to LES, and LES had an additional computational cost of 88% with respect to the URANS.

As reviewed above, LES provides more accurate results in predicting the concentration field compared to RANS-based models. However, large discrepancies have been reported between the simulated concentration fields and the experimental data in the vicinity of the pollutant source, located on the ground, for both indoor and outdoor regions [25,135], which may be due to lack of accuracy in predicting TKE in that region.

In order to account for thermal effects, the turbulence model should be capable of capturing the field variables under non-isothermal boundary layers. RANS, URANS, and LES were employed to predict the airflow field within street canyons under non-isothermal conditions [136]. The results show that, when the windward wall was heated, RANS and URANS could not predict the stagnant area in the canyon. However, LES was able to capture the mentioned region, as can be seen in Figure 2.1 (obtained from [136]). Nazarian and Kleissl [137] also reported the incapability of URANS in predicting the turbulence nature of the flow, particularly in the canyon region.

Isothermal condition



Windward wall heating condition

Measurement	RANS	URANS	LES
-------------	------	-------	-----

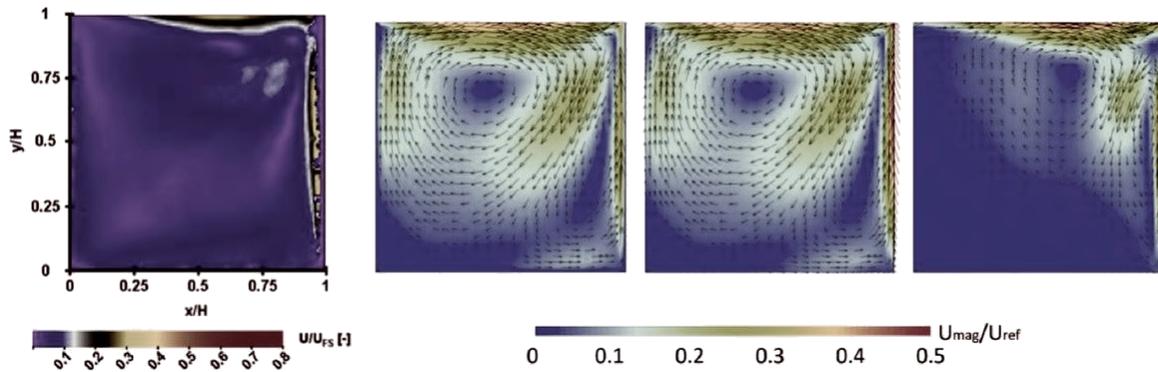


Figure 2.1: Comparison between the results of simulations of [136] (with the RANS, URANS and LES approaches) with the wind tunnel measurements of [66] with $Re=9000$ and $Ri=1.56$, for a 2D street canyon with the aspect ratio of 1 (obtained from [136])

Yoshie et al. [74] showed that the RANS approach could not capture the vortex shedding phenomena and vortical regions behind the tested building. However, LES showed a good performance in predicting the airflow and concentration fields. Regarding the temperature field, the vertical temperature distribution in the proximity of the leeward wall, yielded from RANS, was more vertical than that obtained by the experiment, which can be attributed to the strong vertical flow predicted by the RANS approach in that region. However, based on the experiment [27] and the simulation [21], the thermal boundary layers generated in the vicinity of the walls are very thin, and their effects on the overall flow field are not significant. Therefore, it can be concluded that the RANS approach overestimated the results in [74].

Since RANS is computationally efficient, its usage is beneficial in urban studies, particularly in those covering large areas. The computational resources required for LES is much higher than that for RANS [126]. It has been reported that the CPU time required for the statistical convergence using LES is 25 times more than that for RANS in studying the dispersion process around an isolated building. The situation can even be worse for a large number of buildings [138]. RANS models can be revised to overcome their limitations. Several research items have been performed to investigate the impact of modeling/resolving the boundary layer on heat transfer from surfaces using the RANS approach [139–142]. It has been reported that the impact of choosing a proper wall function on the results is more remarkable in strong stratification conditions than weak ones [75]. The results of [143] showed an overestimation of the velocity profiles in the vicinity of the walls when the low-Reynolds number modeling was used. Furthermore, the performance of the non-equilibrium wall functions in predicting TKE was better than low-Reynolds number modeling.

In order to clearly illustrate the mechanisms of the dispersion process, especially under the non-isothermal boundary layer, a more complex approach than RANS, such as LES, is required [129,136]. Since most of the structures in urban areas are sharp-edged bodies, the flow around these bodies is extremely unsteady and turbulent [29]. In order to reduce the computational costs

while retaining the pros of LES, a combination of URANS and LES, such as detached eddy simulation (DES) [144], delayed detached eddy simulation (DDES) [145], and also multi-domain (multi-zonal) models (e.g., zonal or embedded large eddy simulation, ELES) [29,146] can be used [145–150].

Using DES, the flow in the boundary layer is modeled using URANS, while LES resolves the outer regions. Modeling the near-wall flow employing the URANS approach allows users to coarsen the computational grid. The grid size is an essential factor in switching to the LES mode from the URANS mode and vice versa. Therefore, in DES, computational grids should be generated carefully to determine the proper place for the switching process. In some cases, the grid resolution in the boundary layer is not suitable for performing DES, and the switching process takes place in the wrong location. Therefore, LES is being used on a relatively coarse mesh, which yields inaccurate results [35]. Thus, it is reasonable to find a way to delay the process of switching to LES from URANS in the boundary layer, using the modified version of DES, called DDES [145] (also see [145,150] for the modifications).

The main idea behind the ELES method is to divide the computational domain into two different regions with different concepts of turbulence modeling [29,146]. ELES computational domain is composed of two zones: (1) the inner zone encompassing the target region, in which the LES approach is used, and (2) the outer zone in which the URANS approach is used. At the URANS-LES interface, TKE, which is modeled in the URANS zone, should be explicitly converted to the resolved scales [29]. In other words, since the URANS approach does not generate turbulent fluctuations, an action should be performed to produce the fluctuations at the interface [29,151].

Researchers have employed the above-mentioned approaches in dispersion and other urban-related studies [29,135,151,152]. Liu and Niu [153] showed that DDES can reduce computational costs by 40% compared to LES, while retaining accuracy at an acceptable level in simulating the airflow field around an isolated high-rise building under the isothermal condition. However, the results showed some discrepancies between the mean flow patterns of DDES and LES in the wake region. Paik et al. [154] reported that both URANS and DES (on coarse grids) could not capture the coherent structures around cubes. On a finer grid resolution, the results of DES with low-Reynolds modification were reported to be in good agreement with the measurements.

DES was successfully used in predicting the airflow and pollutant concentration fields around building arrays at the isothermal condition [155]. In the study of Lateb et al. [156] under the isothermal condition, the flow field obtained by DES agreed well with those from the wind tunnel measurements. However, DES yielded almost the same average error regarding the concentration field compared to RANS.

In the study of Liu and Niu [157] under the isothermal condition, DDES reduced computational costs by approximately 20% while retaining model accuracy at the LES level. Dai et al. [135] reported that the performance of DES in predicting the concentration field was better than RANS, while it was outperformed by LES. They stated that DES underestimated the concentration field in most parts of the simulated case.

Although ELES has been used in problems such as channel flows, flow over bluff bodies, etc. [149,150,158–160], it has rarely been employed in predicting urban airflow and concentration fields [29,151,152,161,162]. The model has been used to enhance the performance of a Gaussian model, called AERMOD, by gaining knowledge about the building downwash and sidewash effects [152,161,162]. In [152,161,162], similar to [151], the SST model was selected for the URANS zone, while the wall-adapting local eddy-viscosity (WALE) subgrid scale (SGS) model was employed in the LES part of the domain. The constant turbulent Schmidt number (Sc_t) approach with $Sc_t = 0.7$ was used in [151], as is the case with [29]. Although using dynamic method for turbulent Schmidt number calculation results in higher accuracy in concentration field prediction, it increases computational costs [163]. The results of ELES performed in [151] were promising under perpendicular and oblique wind directions. Jadidi et al. [29] also reported that, with an increase in computational costs by 49% compared to LES, ELES was able to improve the accuracy in concentration field prediction.

It is noteworthy that none of the above-mentioned studies were performed under non-isothermal conditions. Thus, it is reasonable to investigate the performance of hybrid/zonal models in predicting turbulent airflow and concentration fields under non-isothermal conditions.

2.4 Conclusions for computational parameter setting

In this chapter, experimental and numerical studies were reviewed in order to realize the impact of non-isothermal conditions on the flow and concentration fields. The reviewed full-scale and reduced-scale field measurements have usually neglected the pollutant concentration. On the other hand, wind tunnel measurements have been conducted considering pollutant vents and different geometries representing urban areas. In addition to providing knowledge on the fundamentals of the problem, experimental studies can be used to verify numerical models. In order to have a reliable simulation, it would be beneficial to validate the results of various parameters with experimental data, including the time-averaged flow and concentration fields, and turbulence statistics. Thus, having access to the mentioned experimental data is a vital factor in conducting a CFD simulation.

Numerical studies have been reviewed and their limitations were categorized into various parameters, including the dimensionality, the geometry, and the CFD approach. Due to significant differences between the airflow pattern in 2D and 3D studies, particularly under thermal effects, conducting 3D studies seem to be reasonable. Regarding the geometry used in simulations, structural complexities of a realistic urban areas, such as changes in the urban area density and buildings' heights, have usually been overlooked. This neglect can be caused by the unavailability of experimental data. Furthermore, conducting simulations employing simplified geometries can provide opportunities to fundamentally investigate a process, which can even be expandable to more realistic geometries.

Regarding the CFD approaches used in the reviewed studies, in order to properly simulate the airflow and concentration fields in an urban area, the model should be able to capture the unsteady and turbulence nature of the flow field due to high turbulence level in urban airflows, arising from

the presence of sharp-edged geometries and high Reynolds number flows. Furthermore, non-isothermal conditions can even increase the turbulence intensity within urban areas. Therefore, more complex models than RANS ones, such as LES, can be a reasonable choice. However, high computational costs of LES are a barrier for the mentioned task, particularly in the regions covering vast areas. Thus, hybrid/zonal URANS-LES approaches were introduced to make a balance between computational costs and accuracy. It was reported that thermal effects have usually been neglected in studies, particularly those using hybrid/zonal URANS-LES approaches. Since the inclusion of thermal effects can increase the turbulence intensity of the flow field, and the use of URANS approach in the model may decrease the models' capability in capturing the turbulence nature of the flow field, the viability of these models should be assessed under non-isothermal conditions.

Therefore, in the following chapters, a simplified geometry (an isolated high-rise building) is used to simulate the flow and concentration fields around it. Furthermore, ELES is used for the first time under thermal effects in simulating highly-turbulent airflow and concentration fields. ELES is considered a complex model since different parameters, including the size of the computational domain, computational grids, simulation time step, etc., can significantly affect the results. Since there is no comprehensive guideline for conducting ELES, the author of the present thesis had to experiment with different parameters to obtain reasonable results. Thus, it is noteworthy to mention that the settings used for the simulation in the next chapter may not be expandable to other test cases.

Chapter 3. Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses

3.1 Introduction

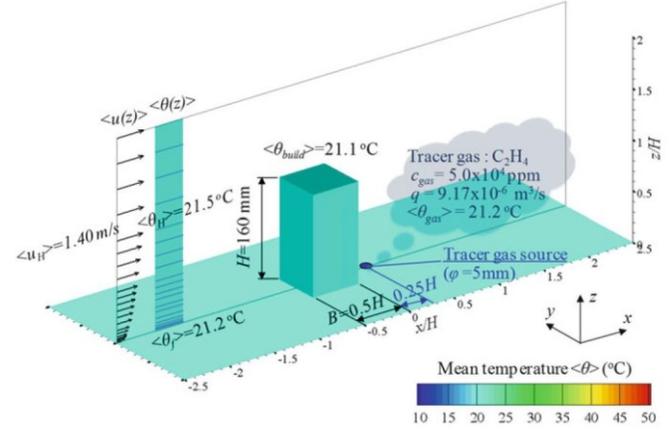
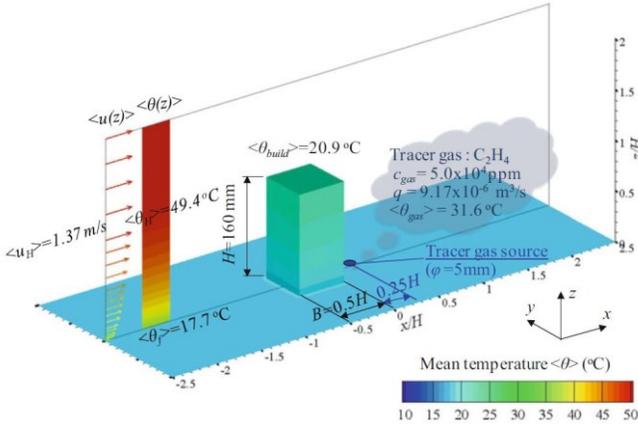
Non-isothermal stratification conditions can alter the airflow pattern and pollutant dispersion process within urban areas. The present chapter is focused on the impact of various stratification conditions, namely, stable, isothermal (neutral), and unstable, on the airflow and concentration fields around an isolated high-rise building. Zonal Reynolds-averaged Navier-Stokes (RANS)-large eddy simulation (LES), also known as embedded large eddy simulation (ELES), is employed for simulating the airflow and concentration fields under non-isothermal boundary layers in order to make a balance between computational costs and accuracy. The present chapter, at first, qualitatively and quantitatively compares the results obtained by ELES with the available experimental data [164] and the prediction made by LES model [103]. The current chapter evaluates the performance of ELES for being used in urban studies under thermal effects, with the emphasize on the results accuracy and computational costs. Furthermore, after validating the results, a comprehensive analysis of the impact of thermal stratification conditions on the airflow and concentration fields is carried out. Also, the proper orthogonal decomposition (POD) technique and time-frequency analysis are employed to understand the dominant structures of the airflow field and assess its transient behavior.

3.2 Simulation cases, boundary conditions and computational settings

In the present chapter, an isolated high-rise building, under different stratification conditions, is used as the test case. The geometry and conditions are the same as the ones used in the wind tunnel experiment [164]. Figure 3.1, obtained from [165], depicts the schematic view of the model building under three different stratification conditions, namely, stable, isothermal, and unstable. The building height is $H = 2W = 2B = 0.16m$, where W and B are the width and depth of the building, respectively. The background wind velocity at the building height, u_H , is set to be almost equal for all cases; $1.37 m/s$ for both the stable and unstable conditions, and $1.4 m/s$ for the isothermal case. Therefore, the Reynolds number at the building height is approximately 15,000.

a) Stable

b) Isothermal



c) Unstable

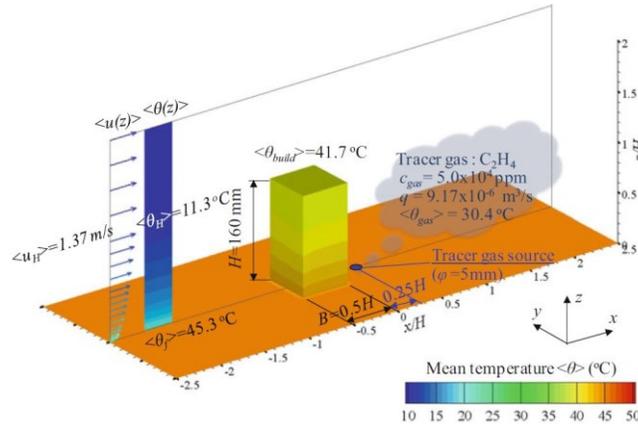


Figure 3.1: The schematic view of the model building, located in different stratification conditions (obtained from [165]). z is the coordinate in the vertical direction. u , θ and c_{gas} represent the streamwise velocity, temperature, and gaseous pollutant concentration, respectively. *build*, *gas* and *f* are the abbreviations for building, gaseous pollutant and floor, consecutively.

The volume flow rate discharging from a circular point source located in the building's wake region is set to be $q = 9.17 \times 10^{-6}\text{ m}^3/\text{s}$ with the ethylene (C_2H_4) concentration of 5% [165]. The diameter of the pollutant source is 5mm. It should be noted that the surface temperature is constant. In the stable and unstable conditions, the temperature of façades and that of the floor are different from each other. The experimental conditions of [164] are summarized in Table 3.1.

Table 3.1: Conditions for the wind tunnel experiment of [164]. (Ri , $\Delta\theta$ and θ_0 represent bulk Richardson number, the absolute value of the temperature difference ($|\theta_H - \theta_f|$), and the space-averaged airflow temperature at the inflow boundary layer, respectively).

Parameter	Thermal stratification conditions		
	Stable	Isothermal	Unstable
Ri	0.08	0.00	-0.10
H [m]	0.16	0.16	0.16
u_H [m/s]	1.37	1.40	1.37
$\langle\theta_f\rangle$ [°C]	17.7	21.2	45.3
$\langle\theta_H\rangle$ [°C]	49.4	21.5	11.3
$\langle\Delta\theta\rangle$ [°C]	31.6	0.4	33.9
$\langle\theta_0\rangle$ [°C]	41.9	21.5	16.6
$\langle\theta_{build}\rangle$ [°C]	20.9	21.1	41.7
$\langle\theta_{gas}\rangle$ [°C]	31.6	21.2	30.4

3.2.1 Computational domain and grid resolution

The computational domain is generated based on the recommendations of [103,166]. The inlet boundary is placed $2H$ upstream of the model building, while the outlet one is located $10H$ downstream of the model building. Furthermore, the computational domain is extended in the lateral and vertical directions by $7.5H$ and $6.25H$, respectively. Since the present computational domain is aimed to be used with the ELES approach, based on the guidelines of [29], it is divided

into two different zones, namely, the URANS zone and the LES zone. The computational domain for the present simulation is shown in Figure 3.2 (a) and (b).

A grid sensitivity analysis is performed on the computational grid to ensure its usability. It should be mentioned that this analysis is based on the simulation of the isothermal condition. Three different grids, namely, coarse, medium, and fine, with 541,590, 1,457,560, and 2,148,192 cells, respectively, were employed to assess the independence of the results from the grid resolution. For the concentration field, the results of the medium and fine grids were almost similar to each other. Furthermore, two different metrics, namely, hit rate and the factor of 2 of the observations $FAC2$ (see Eqs. 3-1 and 3-2) have shown no remarkable results improvements for the medium and fine grids, compared to the coarse one. For instance, $FAC2$ increased from 0.88, for the coarse grid, to 0.90, for the fine one (for the isothermal case). Moreover, regarding the concentration field, the mentioned metric remained the same (0.97) for all the mentioned grids. Also, the fine grid improved the time-averaged concentration results by 5% of mean relative error compared to the coarse one. Findings of [25,103,167,168] also show that refining the computational grid does not necessarily guarantee the provision of higher concentration accuracy in the LES approach. Furthermore, the time required for the convergence of the simulations with the medium and fine grids was much higher than that for the coarse grid (more than three times for the medium one). Therefore, due to the high computational costs of the medium and fine grids, the coarse grid is selected for further simulations in the present chapter. Moreover, the primary goal of the present chapter is to compare the capability of LES and ELES in simulating the airflow and concentration fields under non-isothermal boundary layers. Therefore, the coarse grid, which is almost equal in cell number with the one employed in the LES study of [103], is chosen. The coarse grid is shown in Figure 3.2 (c) and (d). The performance evaluation of ELES with respect to the LES [103] and the experiment [164] are then carried out in more detail.

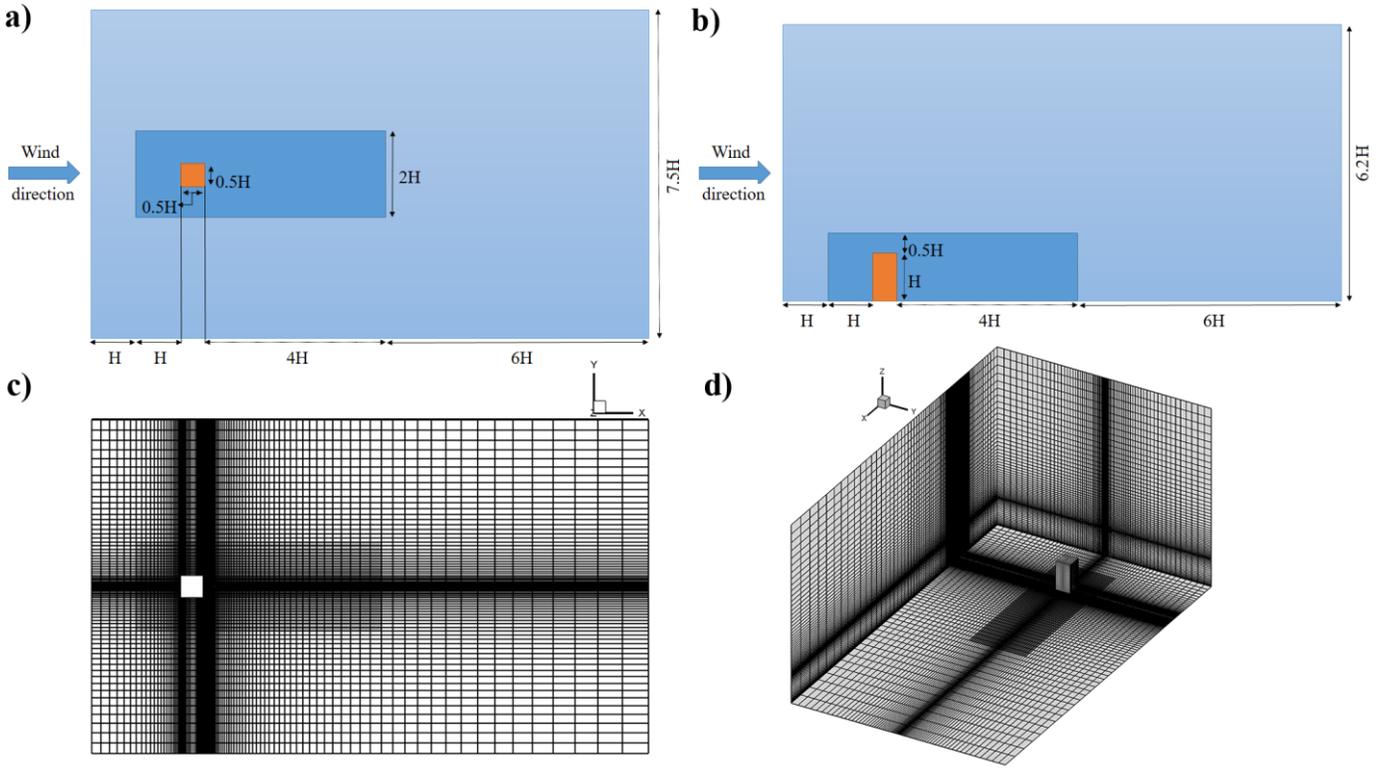


Figure 3.2: First row: computational domain for ELES: a) top view in $x - y$ plane, and b) side view in $x - z$ plane (orange: the model building; chromatic blue: LES zone; pale blue: RANS zone),

Second row: coarse computational grids: c) top view in the $x - y$ plane (white square represents the model building), and d) 3D view

The ratio of the number of the grid cells in the LES zone to that in the RANS zone is approximately $1/3$. Hexahedral cells with the maximum aspect ratio of 1.06 are employed for generating the computational grids. The maximum value of the dimensionless wall distance, y^+ , over all the building façades and the domain ground for the coarse grid in the LES zone is less than 26, which is within the acceptable range of using the Werner and Wengle wall functions [169]. Furthermore, for grid resolution assessment in the LES zone, the $\langle u_{SGS} \rangle / v$ approach, which has been shown to be useful [131,170–172], is applied. It should be noted that this describes the ratio of the SGS kinematic viscosity to the physical one. As the grid becomes finer, the mentioned ratio approaches smaller values [172]. Figure 3.3 shows the profiles of $\langle u_{SGS} \rangle / v$ over the sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, $x/H = 1.5$, along the vertical center-plane, $y/H = 0$, for the coarse grid. Figure 3.3 shows the mentioned ratio, at its maximum value, is in order of one in the wake region. Thus, it can be concluded that the coarse grid is suitable for the present simulations.

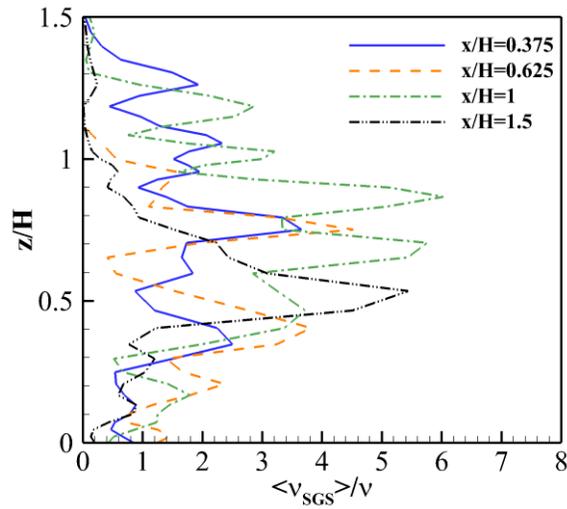
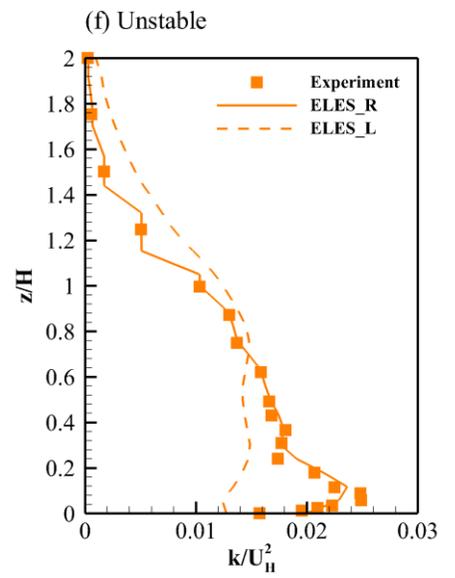
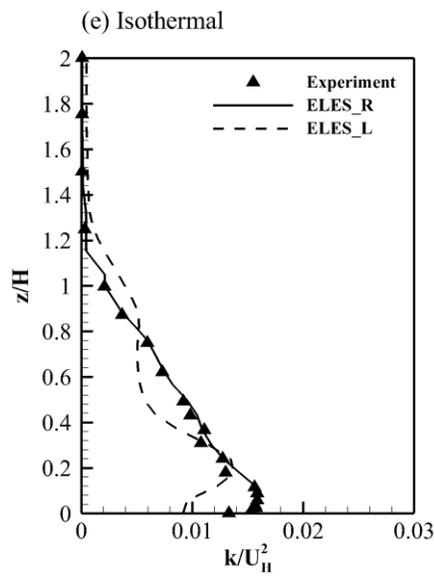
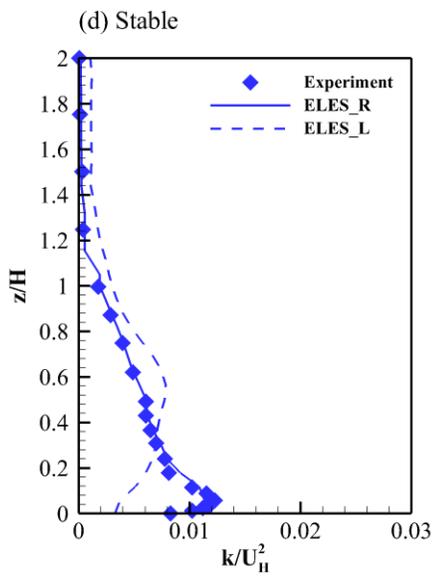
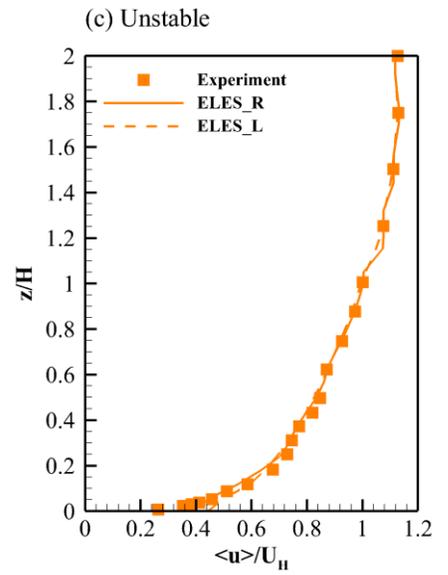
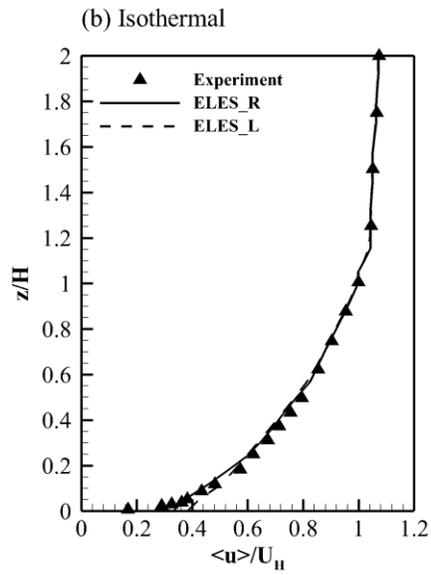
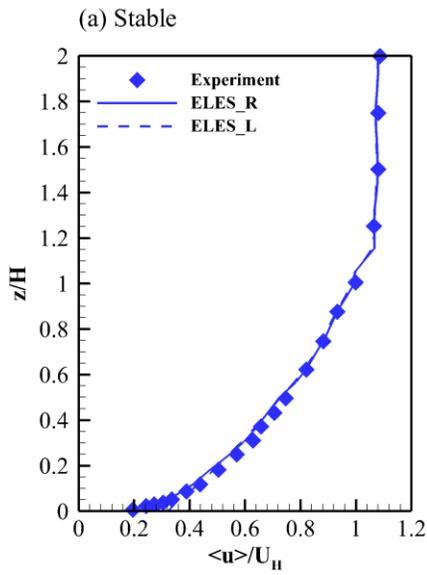


Figure 3.3: Ratio of time-averaged SGS kinematic viscosity to the physical one, $\langle v_{SGS} \rangle / \nu$, obtained from isothermal ELES simulation over the sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, $x/H = 1.5$, along $y/H = 0$

Figure 3.4 shows the profiles of the non-dimensional form of the streamwise velocity, TKE and temperature at the inlet boundary of the wind tunnel experiment of [164]. As mentioned [166], the values of turbulence dissipation rate, ε , at the inlet boundary can be derived using $\varepsilon = P_k + G_k$, where, P_k and G_k are the TKE production and buoyancy production terms, respectively.

Table 3.2 shows the boundary condition type at each surface of the domain. It should be noted that at the URANS-LES interface, the vortex method [173] is used to explicitly convert the modeled TKE into the resolved scales [29]. Since there is no study regarding the usage of different turbulence generation methods at the URANS-LES zone for wind engineering problems, in the present study, the authors have relied on previous studies for LES computations.



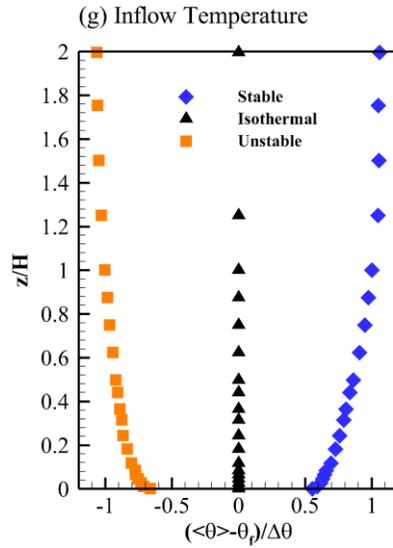


Figure 3.4: Dimensionless form of the Inflow boundary conditions, based on the experiment [164], ELES (conducted in the empty domain) at the inlet boundary (i.e., RANS zone), denoted by ELES_R, and at the RANS-LES interface (i.e., beginning of the LES zone), represented by ELES_L:

(a), (b), and (c): streamwise velocity, (d), (e) and (f) TKE, and (g) inflow temperature profiles [103]

It should be noted that the authors are aware of the difference in the location of the LES inflow boundary between ELES and LES (which, obviously, affects the employment of different turbulence generation methods regarding their overprediction/under-estimation of the results). Studies regarding the usage of these methods in the field of wind engineering using LES are also rare [174–176]. Based on the results of the mentioned studies, no inflow turbulence generation method can model real turbulence. Therefore, in order to have a reliable simulation, several inflow turbulence generation methods should be assessed. However, since the goal of the present study is not to evaluate the performance of these methods for being used in ELES computations, the results of the mentioned comparative studies have been considered. Since the performance of the vortex method in predicting the flow field structures [174] and turbulence statistics and concentration field [176] have been reported to be reasonable, the authors of the present study used the mentioned method for the simulations.

Table 3.2: Boundary conditions of the computational domain

Inflow boundary	The profiles which are shown in Figure 3.3, and computing ε based on $\varepsilon = P_k + G_k$, mentioned by [166]
URANS-LES interface	Vortex method [173] with number of vortices of 399, based on the recommendation of [177]
Upper and side boundaries	Symmetry boundary conditions which imply zero normal velocity components and zero normal gradients of all the flow variables
Floor and the building façades	Werner and Wengle [169] wall treatment
Pollutant vent	Uniform and constant velocity, zero turbulence intensity ($U_{vent} = 0.467 \text{ m/s}$)
Outlet boundary	Zero static gauge pressure and zero gradients for all the variables

In order to assess the homogeneity of the approaching flow profiles, ELES simulations, for the same domain size and grid resolution, are conducted (see Figure 3.4). The alterations in the time-averaged velocity profiles, for all stratification conditions, remain very limited, indicating that horizontal homogeneity is achieved. However, some changes in the TKE profiles have been observed. It should be noted that since in ELES, the inflow boundary is located in the RANS zone, which does not need any inflow turbulence generation method, the inflow TKE profiles are coincident with the experimental ones. However, due to the employment of the vortex method at the RANS-LES interface, which produces low velocity fluctuations in the streamwise direction [178], TKE underestimation has been observed. A similar finding has also been reported by Refs. [29,178]. However, this underestimation has low impact on the dispersion process since the turbulence in the wake region is mainly produced due to the presence of the building [178].

3.2.2 Solution method

In this study, ELES is used for the simulations; RLZ [93] in the URANS zone and the dynamic Smagorinsky-Lilly model (DSLM) [104,105] in the LES zone. Regarding the usage of RLZ in the URANS zone, its applicability in predicting the flow and concentration fields in the ELES under the isothermal condition has been considered [29]. Furthermore, the performance of three different turbulence models, standard $k - \varepsilon$ (SKE), re-normalization group $k - \varepsilon$ (RNG) [87], and RLZ in predicting the flow field, turbulence statistics, and the concentration field were evaluated. Based

on the results, RLZ, which performed relatively good (but not the best in predicting every variable), has been selected for being used in the present study. The governing equations of the mentioned models can be found in Appendix A.

In order to capture the flow characteristics' evolutions, a fixed time step of $\Delta t = 1 \times 10^{-4}$ s is selected to maintain the Courant-Friedrichs-Lewy (CFL) number less than 1. After initializing the simulations with the RLZ model, the effect of initial conditions was disappeared after approximately 7 s regarding the flow time to reach a semi-steady state condition. Then, the time-averaging process was started to achieve a satisfactory statistical convergence, which was more than 30 s regarding the flow time. Furthermore, the constant turbulent Schmidt number approach, with the value of 0.7, is selected for the present simulations [29]. Also, the turbulent Prandtl number has been set to 0.9, as is the case with [166].

The unsteady SIMPLE [179] algorithm is employed for the pressure-velocity coupling procedure. The second-order scheme is used for the pressure interpolation, and the second-order upwind schemes are employed for turbulence quantities, energy, and concentration. Furthermore, the second-order bounded central differencing schemes are used for the momentum spatial discretization.

3.3 Results validation

3.3.1 Time-averaged flow field and TKE

Comparison between the non-dimensional time-averaged streamwise velocity profiles, $\langle u \rangle / U_H$, obtained by the ELES approach with those from the LES [103] and the wind tunnel experiment [164,165], along $y/H = 0$, for the stable, isothermal and unstable conditions is shown in Figure 3.5. The comparison is made over the measuring points on four sampling lines, namely, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, in the wake of the model building. It should be noted that the experiment was performed with an uncertainty of 5% regarding the time-averaged velocity [164,165].

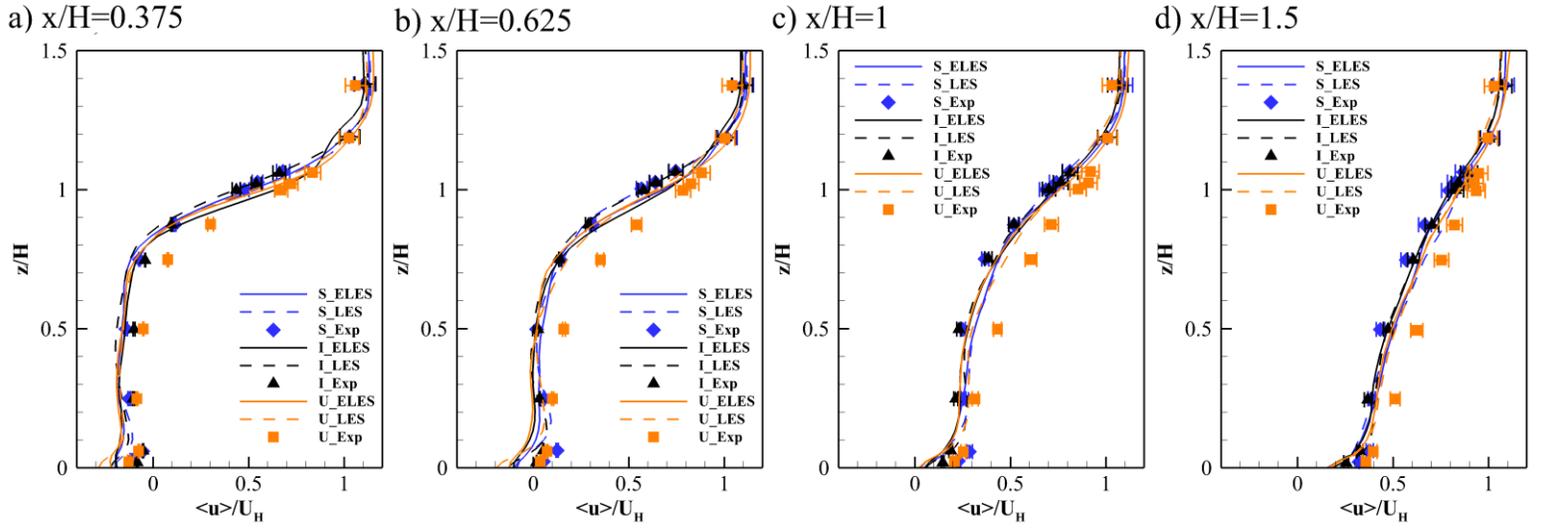


Figure 3.5: Comparison between profiles of dimensionless time-averaged streamwise velocity, $\langle u \rangle / U_H$, obtained by present simulation (ELES), LES of [103] and experiment [164] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (“S”, “I”, and “U” represent stable, isothermal and unstable conditions, respectively)

Figure 3.5 shows that ELES and the LES of [103] perform almost similarly in predicting $\langle u \rangle / U_H$. Large deviations between the results of the mentioned models can be found in the proximity of the walls, particularly for $x/H \leq 0.625$, at $z/H < 1$. As the flow moves toward the downstream (i.e., getting farther from the recirculating region in the wake of the building), the mentioned deviation decreases. In the vicinity of the ground, almost for all datapoints, both ELES and LES under-predicted the velocity field. Moreover, LES [103] performs slightly better than the present ELES near the ground at $x/H = 0.375$ and 0.625 . This can be due to different mesh resolutions in the proximity of the ground and the difference between the nature of the models (i.e., most of the domain is modeled using URANS in ELES. Also, perturbing the URANS results at the URANS-LES interface plays an important role in the accuracy). However, the model’s tendency in overprediction is observable at $z/H > 0.5$ as the flow progresses toward the downstream for the stable and isothermal cases. On the other hand, at almost all data points in the $z/H < 1$ zone, both ELES and the LES of [103] under-predicted the results under the unstable condition.

Two different evaluation metrics, namely hit rate (q) and the factor of 2 of the observations ($FAC2$), are employed to quantitatively evaluate the performance of the modeling approaches (i.e., ELES of the current study and LES of [103]) with respect to the available experimental data [164]. These metrics can be defined as follows [180–184]:

Hit rate (q):

$$q = \frac{1}{N} \sum_{i=1}^N n_i \quad \text{with} \quad n_i = \begin{cases} 1 & \text{for } \left| \frac{P_i - O_i}{O_i} \right| \leq D_q \\ 0 & \text{else} \end{cases} \quad \text{Eq. 3-1}$$

where, N is the number of data points, O_i and P_i are the observed (measured) and predicted values of a particular variable. Also, D_q corresponds to the allowed relative deviations, which has been set to $D_q = 0.25$ as suggested by [180,181,183].

$FAC2$:

$$FAC2 = \frac{1}{N} \sum_{i=1}^N n_i \quad \text{with} \quad n_i = \begin{cases} 1 & \text{for } 0.5 < \frac{P_i}{O_i} \leq 2 \\ 0 & \text{else} \end{cases} \quad \text{Eq. 3-2}$$

The hit rate and $FAC2$ values of $\langle u \rangle / U_H$, obtained by each of the mentioned models, with respect to the experiment of [164] is summarized in Table 3.3. It should be noted that the number of data points is 44 for each case (i.e., 11 points on each of the four vertical lines).

Table 3.3: Hit rate (q) and $FAC2$ of $\langle u \rangle / U_H$ profiles obtained by the present study (ELES) and LES of [103] with respect to the experimental data of [164], over the sampling lines along the $y/H = 0$ plane (i.e., 44 experimental data), for different stratification conditions

Modeling approach	Thermal stratification condition		
	Stable	Isothermal	Unstable
ELES	$q = 0.76$ $FAC2 = 0.88$	$q = 0.80$ $FAC2 = 0.88$	$q = 0.71$ $FAC2 = 0.80$
LES [103]	$q = 0.79$ $FAC2 = 0.88$	$q = 0.77$ $FAC2 = 0.86$	$q = 0.71$ $FAC2 = 0.81$

Table 3.3 shows that the performance of both ELES and LES deteriorates under the unstable condition compared to the other two cases, which can be attributed to the presence of stronger shear layers under the unstable condition. The overall performance of both the present ELES and LES of [103] are almost similar to each other, with some minor differences mostly in the proximity of the ground. Furthermore, the results of the present ELES are in good agreement (fairly) with experimental data. As presented in Table 3.3, the values of both metrics, for all the simulated conditions, are reported to be higher than the quality acceptance criteria (i.e., $q \geq 0.66$ and $FAC2 > 0.5$) [182,185].

Figure 3.6 illustrates the contours of dimensionless TKE, k/U_H^2 , obtained from ELES and the experiment [164]. Since the TKE contours of the LES of [103] have not been reported, there is no comparison between the performance of ELES and LES in predicting TKE.

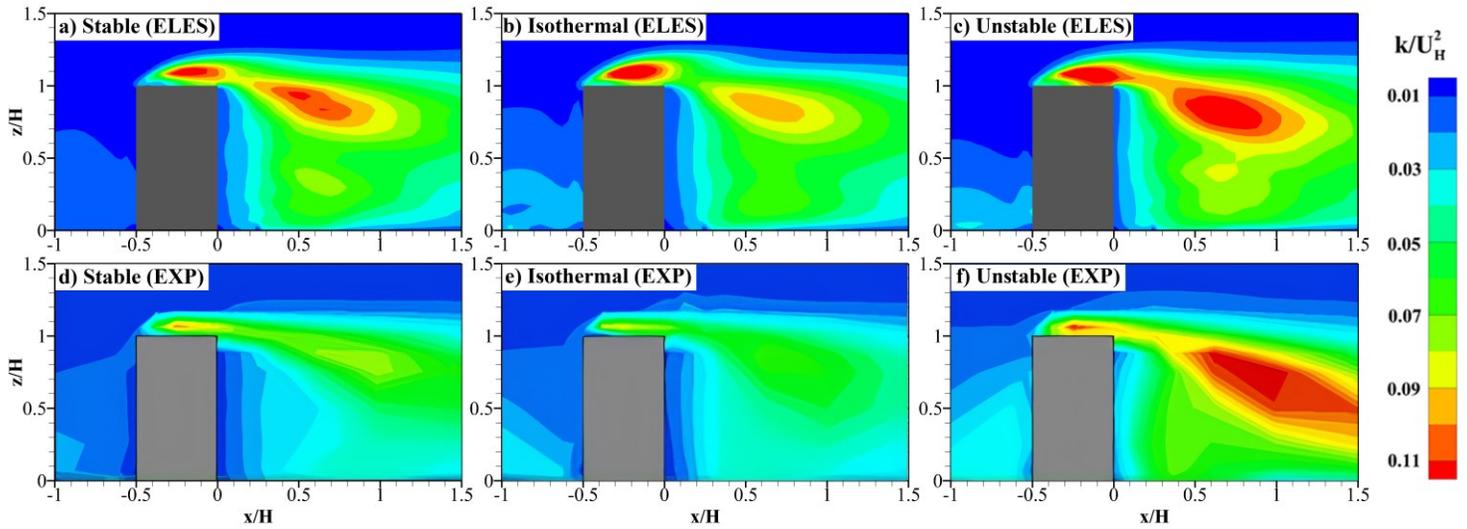


Figure 3.6: Comparison between contours of dimensionless TKE, k/U_H^2 , obtained by ELES and experiment of [164] (experimental contours of TKE are obtained from [164])

Figure 3.6 illustrates that below the high-TKE region, where the shear layers weaken, the ELES performance is better. Furthermore, in the vicinity of the leeward wall, the predicted values are in good agreement with the experimental data. The location of the maximum TKE values in shear layers, predicted by ELES, is in good agreement with the experimental data. However, in general, ELES over-predicts the values of TKE, which can be attributed to the fact that the flow is perturbed at the URANS-LES interface, near the target region. Thus, the TKE level at the roof and in the wake region increases. Also, it should be noted that the q and $FAC2$ values for TKE, obtained by ELES, are in the acceptable range mentioned above (q values for the stable, isothermal, and unstable conditions are about 0.71, 0.74, and 0.75, respectively, and $FAC2$ values are approximately 0.88, 0.92, and 0.95 for the stable, isothermal, and unstable cases, consecutively).

Contrary to LES, for ELES, the inlet TKE profile is almost similar to that of the wind tunnel since it is located in the RANS zone [29]. Due to the flow perturbation at the URANS-LES interface, the incident TKE profile of ELES may slightly be different from that of the experiment. TKE over-prediction at the building roof and in the wake region has also been mentioned by [29]. Moreover, [151] noted that, downstream of the URANS-LES interface, the TKE values increase for ELES with respect to the experimental data. Also, different inflow turbulence generation methods yield different TKE results in the wake region [176]. Using the vortex method led to higher fluctuations in the wake region compared to other inflow turbulence generation approaches studied in [176].

3.3.2 Time-averaged temperature

Comparison between the non-dimensional time-averaged temperature profiles, $(\langle\theta\rangle - \theta_f)/\Delta\theta$, obtained by the ELES with those from LES [103] and the experimental data [164], along $y/H = 0$, for the stable, isothermal and unstable conditions is shown in Figure 3.7. The comparison is made over four sampling lines, $x/H = -1$, $x/H = -0.625$ (upstream of the model building), $x/H = 0.0625$, and $x/H = 0.625$.

Figure 3.7 shows that the results obtained from ELES are in good agreement with those from the LES of [103] and the wind tunnel experiment [164]. At all the vertical lines, the results of ELES are almost similar to those obtained by the LES of [103], with some negligible differences in the vicinity of the floor. Over $x/H = 0.0625$, particularly for $z/H < 1$, the performance of ELES and LES for both the stable and unstable conditions are slightly different. This shows that the LES performance in predicting the temperature field, in the very proximity of the wall, is better than that of ELES.

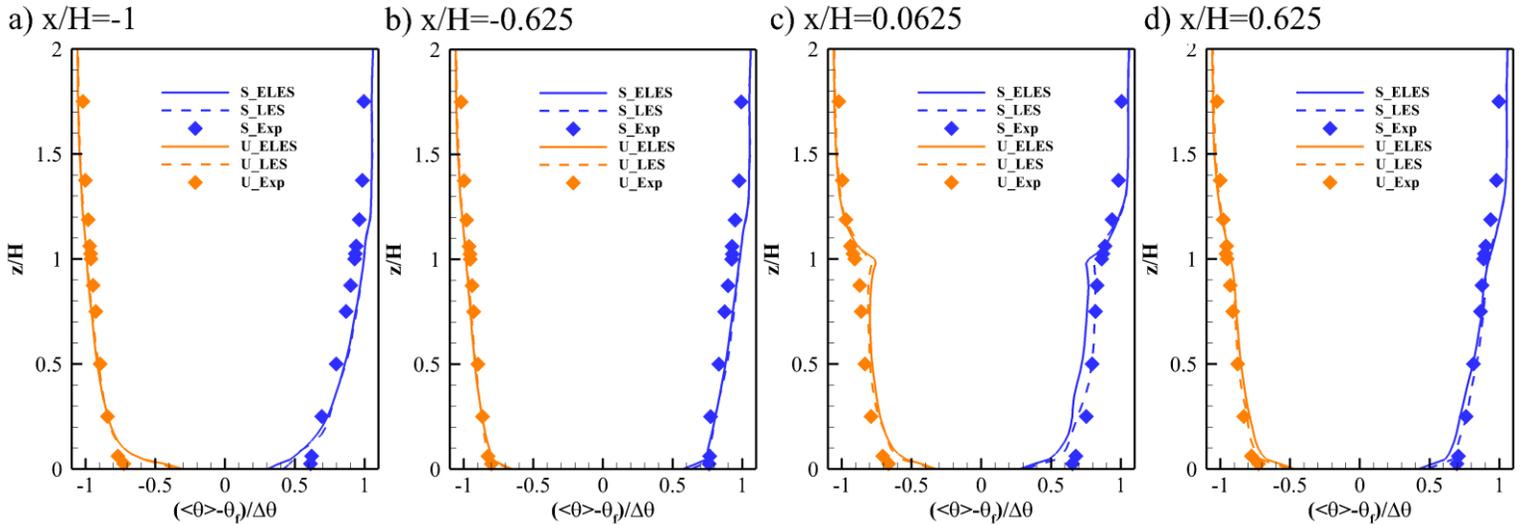


Figure 3.7: Comparison between profiles of dimensionless time-averaged temperature, $(\langle\theta\rangle - \theta_f)/\Delta\theta$, obtained by present simulation (ELES), LES of [103] and experiment of [164] over sampling lines, $x/H = -1$, $x/H = -0.625$, $x/H = 0.0625$, and $x/H = 0.625$, along $y/H = 0$ (“S” and “U” denote stable and unstable conditions, respectively)

Table 3.4 gives the mean relative error percentage, $(1/M \sum_{i=1}^M (P_i - O_i)/O_i) \times 100$ (where, M is the number of datapoints) of $(\langle\theta\rangle - \theta_f)/\Delta\theta$ obtained by each of the mentioned models with respect to the experimental data [164].

Table 3.4: Mean relative error percentage of $(\langle\theta\rangle - \theta_f)/\Delta\theta$ profiles obtained by the present study (ELES) and LES of [103] with respect to the experimental data of [164], over the sampling lines along $y/H = 0$, for different stratification conditions

Sampling line	Thermal stratification condition	
	Stable	Unstable
$x/H = -1$	ELES = 9.0%	ELES = 5.1%
	LES = 9.6%	LES = 4.8%
$x/H = -0.625$	ELES = 5.5%	ELES = 3.1%
	LES = 5.5%	LES = 2.6%
$x/H = 0.0625$	ELES = 11.2%	ELES = 10.3%
	LES = 7.5%	LES = 8.6%
$x/H = 0.625$	ELES = 5.1%	ELES = 3.5%
	LES = 4.9%	LES = 3.0%
Average over sampling lines	ELES = 7.7%	ELES = 5.5%
	LES = 6.9%	LES = 4.8%

3.3.3 Time-averaged and fluctuating concentration

Figure 3.8 compares the profiles of dimensionless time-averaged pollutant concentration, $\langle c \rangle / C_0$, obtained by ELES with those from the LES of [103] and the wind tunnel experiment [164]. The experiment was conducted with an uncertainty of 15% regarding the time-averaged concentration. The comparison is made along $y/H = 0$, over four sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, for the stable, isothermal, and unstable conditions. It should be noted that the reference concentration, C_0 , is defined as follows:

$$C_0 = \frac{q}{U_H H^2}$$

Eq. 3-3

Figure 3.8 shows that the ELES and LES perform almost similar in predicting the concentration field under the isothermal and stable conditions, even in the proximity of the ground. However, the difference between the capabilities of these approaches in predicting the concentration field under the unstable condition is more pronounced. This difference may be due to the changes in the role of each mechanism of pollutant dispersion, particularly the turbulent diffusion flux, in the whole process under different stratification conditions. Thus, the capability of the model in capturing the variables fluctuations may be a reason for the above-mentioned differences.

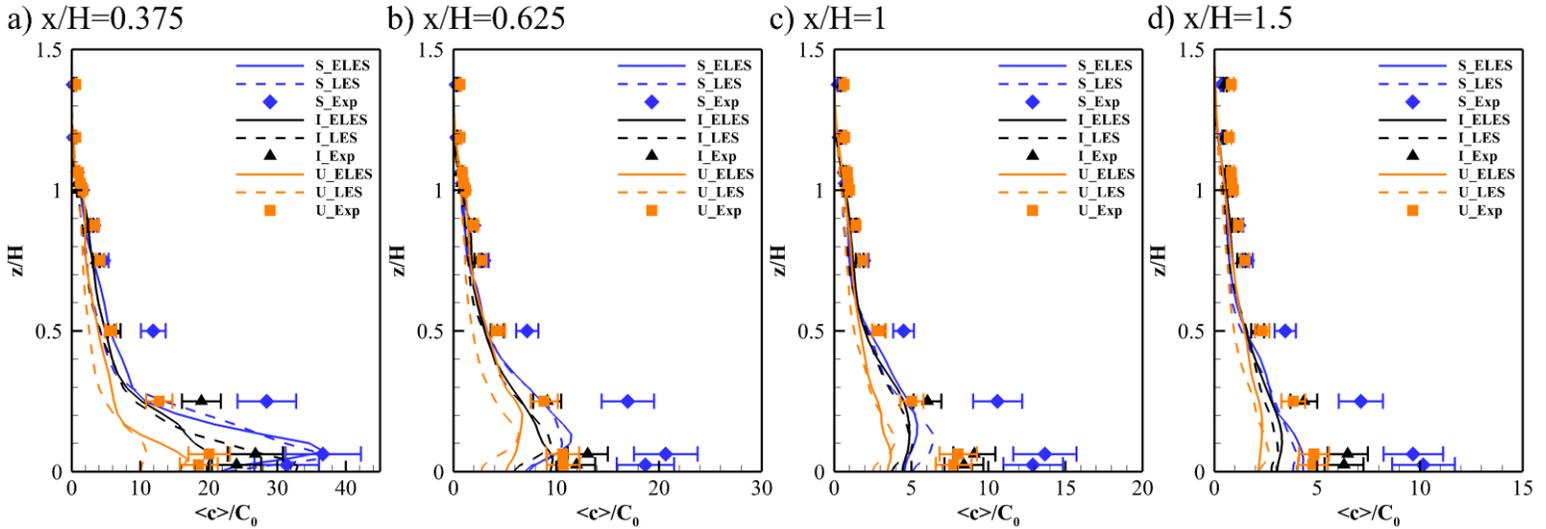


Figure 3.8: Comparison between profiles of dimensionless time-averaged pollutant concentration, $\langle c \rangle / C_0$, obtained by present simulation (ELES), LES of [103] and experiment of [164] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (‘‘S’’, ‘‘I’’ and ‘‘U’’ denote stable, isothermal and unstable conditions, respectively)

The mean relative error and $FAC2$ of $\langle c \rangle / C_0$, obtained by ELES and the LES of [103], with respect to the experiment [164] is listed in Table 3.5.

Table 3.5: Mean relative error (RE) and $FAC2$ of $\langle c \rangle / C_0$ profiles obtained by the present study (ELES) and LES [103] with respect to the experimental data of [164], over the sampling lines along the $y/H = 0$ plane (i.e., 44 experimental data), for different stratification conditions

Modeling approach	Thermal stratification condition		
	Stable	Isothermal	Unstable
ELES	$RE = 28.82\%$	$RE = 17.42\%$	$RE = 24.60\%$
	$FAC2 = 0.77$	$FAC2 = 0.97$	$FAC2 = 0.84$
LES [103]	$RE = 31.61\%$	$RE = 24.23\%$	$RE = 35.01\%$
	$FAC2 = 0.70$	$FAC2 = 0.89$	$FAC2 = 0.62$

According to Table 3.5, ELES performs better than LES in predicting the concentration field. As the instability increases, the difference between the models' accuracy becomes larger. In other words, under the stable condition, ELES and LES of [103] yielded almost similar results, while under the unstable condition, ELES outperformed the LES simulation. The local intensity, defined as the ratio of the root mean square of the concentration fluctuations to the time-averaged one at the same point, plays an important role in the model performance in predicting the time-averaged concentration field [29]. Furthermore, TKE and pollutant concentration are related to each other. As mentioned by [29], higher TKE values lead to higher concentration fluctuations. Therefore, the better performance of ELES in predicting the time-averaged concentration field can be attributed to its superior performance in predicting TKE, and thus, concentration fluctuations, compared to LES.

The profiles of dimensionless pollutant concentration fluctuations, c_{rms}/C_0 , obtained by ELES, LES [103], and the experimental data [164] are illustrated in Figure 3.9. It should be noted that c_{rms} is the root mean square of the concentration fluctuations. These profiles are depicted along $y/H = 0$, over four sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$ for the stable, isothermal, and unstable conditions.

Accurate prediction of concentration fluctuations is of the utmost importance since for the purpose of assessing pollutant exposure, the role of concentrations measured in the span of a few seconds, or even less, is significant, as opposed to those measured over minutes or hours as well as time-averaged ones [29]. As shown in Figure 3.9, as the flow progresses toward the downstream, the values of c_{rms}/C_0 decreases. The location of the peak value, predicted by ELES, is in good

agreement with that reported by the experimental data [164]. Furthermore, the trends of the ELES results are similar to those from [164]. As Figure 3.9 shows, the ELES has significantly outperformed the LES in predicting pollutant concentration fluctuations, which, as mentioned earlier, can be attributed to the better performance of ELES in predicting TKE compared to LES. Considering the concentration field, better performance of the ELES compared to the LES has also been reported by [29].

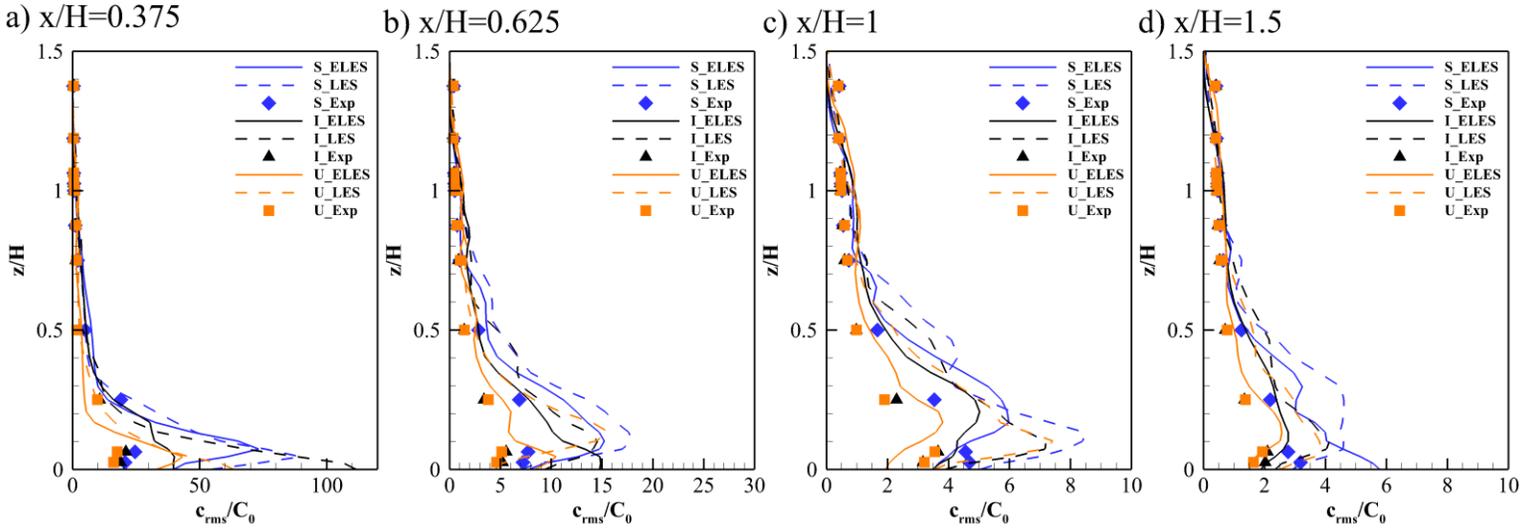


Figure 3.9: Comparison between profiles of dimensionless pollutant concentration fluctuations, c_{rms}/C_0 , obtained by present simulation (ELES), LES of [103] and experiment of [164] over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ (“S”, “I” and “U” denote stable, isothermal and unstable conditions, respectively)

3.3.4 ELES vs. LES: computational costs

The sampling time of the present simulation is different from the LES of [103], which was sampled for 21 s regarding the flow time. However, the results of the present ELES have once been validated with the experimental data and those obtained from the LES of [103] after about 21 s. After 21 s of sampling, in order to achieve better results, the sampling process was continued for another 9 s. However, the difference between the results obtained after 21 s and 30 s was negligible. It should be noted that the figures shown in the present study have been depicted after 30 s of sampling. Therefore, in order to compare the required CPU time for the simulations, the sampling time (regarding the flow time), per day have been compared for the present ELES and the LES of [103]. For the isothermal case, the LES of [103] sampled 0.7 s per day, while the present ELES sampled 2.75 s each day. For the stable and unstable conditions, the LES of [103] sampled 0.65 s each day, while ELES sampled the results for 1.5 s per day. Therefore, the speed-up obtained from using ELES against LES is about 292% for the isothermal condition, and 130% for the other two cases.

3.4 Impact of thermal stratification on flow field, turbulence statistics, and pollutant dispersion

3.4.1 Time-averaged flow field

Considering Figure 3.5, and by averaging the values of $\langle u \rangle / U_H$ over the sampling lines for each case, it can be found out that the stratification condition slightly changes the streamwise velocity. For example, the average of $\langle u \rangle / U_H$ over $x/H = 0.375$ is about 5% higher for the unstable case compared to the isothermal one. A small influence of thermal stratification conditions on the streamwise velocity has been reported by [24]. However, considering $z/H < 0.5$, over $x/H = 0.375$, the average of the absolute value of $\langle u \rangle / U_H$ under the unstable condition is significantly higher than that under the stable and isothermal ones by about 18.1% and 7.9%, respectively. Furthermore, by analyzing the results, it can be noted that the flow instability increases the strength of the recirculating region in the wake region.

Figure 3.10 shows the impact of thermal stratification conditions on time-averaged streamlines and dimensionless time-averaged vertical velocity, $\langle w \rangle / U_H$, along $y/H = 0$. This figure shows that in the proximity of the leeward wall, for $z/H < 0.5$, streamlines are steeper in the isothermal and unstable conditions compared to the stable case. The same pattern, with a smaller difference between the cases, can be observed above the mentioned height. For $0.3 < z/H < 0.5$, the flow is directed downward for the stable and isothermal conditions. However, in the same region, the flow is more horizontally directed in the unstable case, which can be attributed to the buoyant flow induction. Furthermore, a strong recirculating region is formed under the unstable condition in the wake region, approximately at the building height level, which can be related to the generation of strong shear layers due to the opposition of the downward background wind velocity and the upward buoyant force.

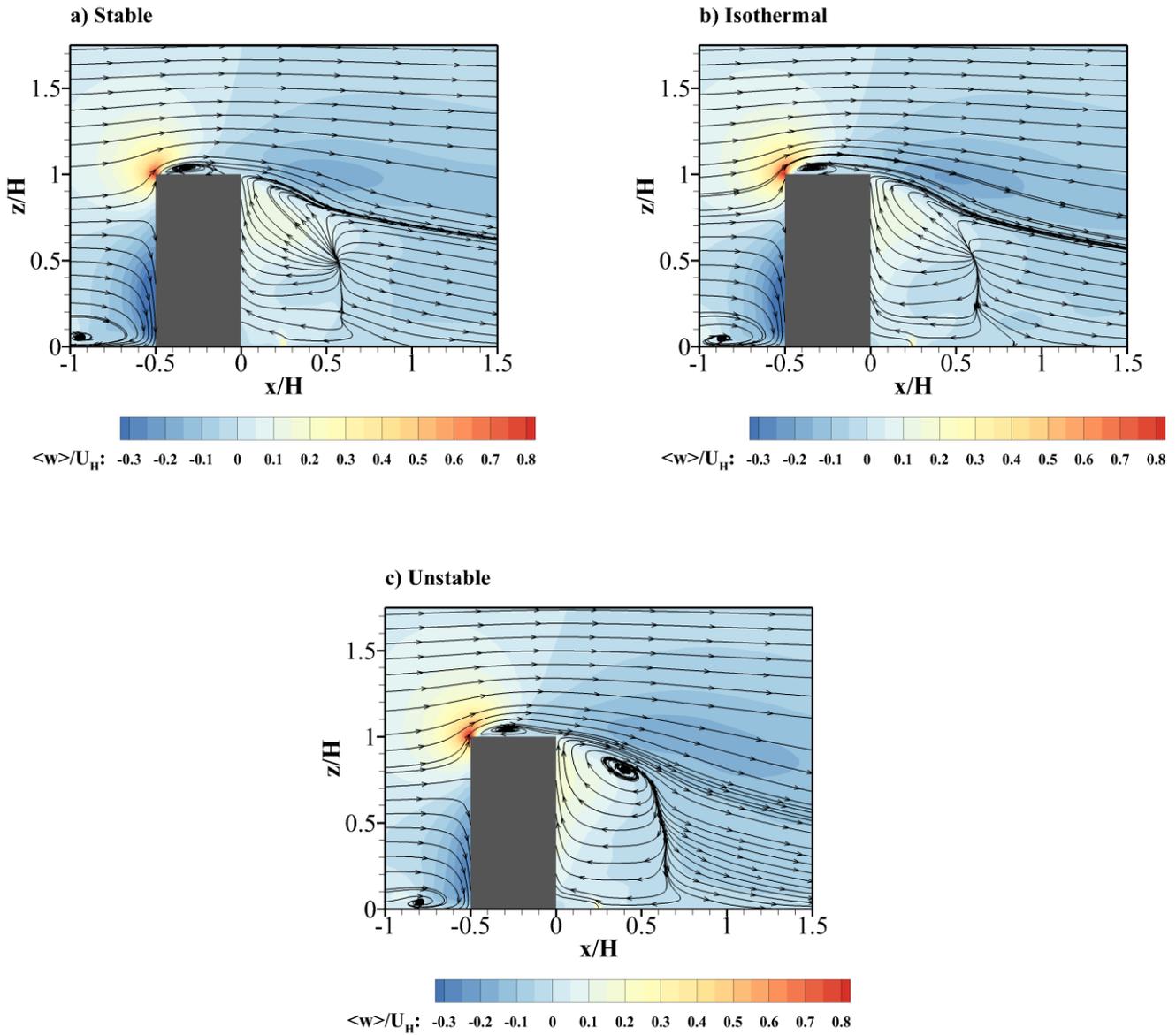


Figure 3.10: Time-averaged streamlines and dimensionless time-averaged vertical velocity, $\langle w \rangle / U_H$, under different thermal stratification conditions

Figure 3.10 shows the overall pattern of $\langle w \rangle / U_H$ is similar for the stable and isothermal cases. However, the region with positive values of $\langle w \rangle / U_H$, i.e., upward direction, becomes smaller for the stable condition compared to the isothermal one. In contrast, due to the formation of a thin thermal boundary layer in the vicinity of the leeward façade, a larger region with positive $\langle w \rangle / U_H$ is observable under the unstable condition in comparison with the other two cases.

3.4.2 Turbulence kinetic energy

Changes in the turbulence level of the airflow considerably influence the dispersion process through the turbulent diffusion flux. It has been reported that the turbulent diffusion flux significantly affects the dispersion process [31]. Therefore, it is reasonable to study the impact of thermal stratification conditions on the TKE generation/suppression. Figure 3.6 shows that as the instability increases, in the proximity of the leeward wall, TKE values increase. Furthermore, the overall increase in the TKE values under the unstable condition compared to the other two cases is observable. Therefore, changes in TKE values due to wall heating are not only restricted to the very proximity of the façade.

Figure 3.11 illustrates the dimensionless TKE profiles, k/U_H^2 , over four sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$ under different stability conditions. The figure shows the stable condition has a very small impact on the TKE value compared to the isothermal case. However, a slight reduction in the TKE value can be observed, particularly in the vicinity of the leeward façade and the ground. Similar results have been reported by Duan and Ngan [24]. They stated that the average of TKE in a street canyon is almost independent of the Richardson number for weak stable conditions. The trend of the TKE profile under the unstable condition is almost similar to those for the other cases, which may be attributed to the uniform surface heating. However, as mentioned earlier, an overall increase in the TKE values under the unstable case can be observed. Furthermore, thermal effects have not altered the vertical location of the maximum TKE value.

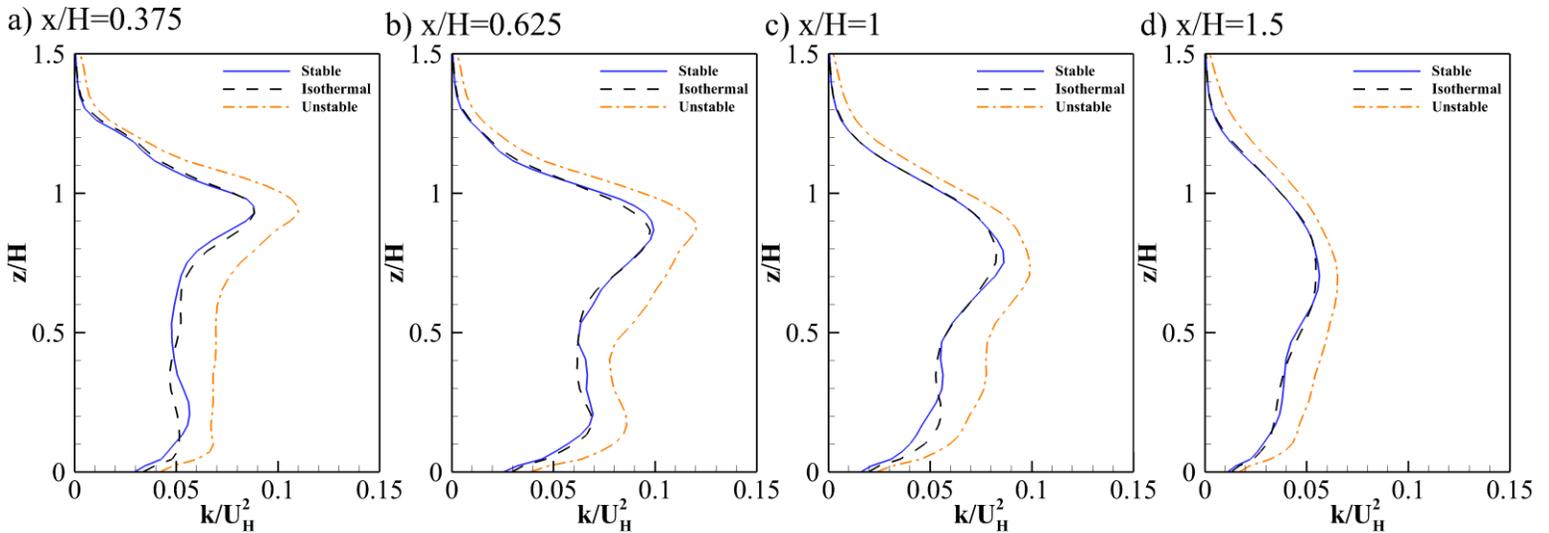


Figure 3.11: Comparison between profiles of dimensionless TKE profiles, k/U_H^2 , under different stability conditions over sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$

3.4.3 Frequency analysis

Time histories of the streamwise and vertical velocity components, at the point $x/H = 2$ and $z/H = 0.5$, along $y/H = 0$, for different thermal stratification conditions, are shown in Figure 3.12 (a) and (b). It should be noted that the mentioned time histories were depicted for various points in the LES zone, and the findings were almost similar. However, due to higher differences between the results at the mentioned point amongst the tested ones, this point has been chosen for further analyses. It should be noted that the dimensionless time unit, t^* , is defined as $t^* = t \times U_H/H$, where t is time (here, the time range in which the required data are sampled).

For both the streamwise and vertical directions, and for all the stratification conditions, non-sinusoidal and fast-changing signals are observable in Figure 3.12 (a) and (b). The time series obtained under different thermal effects show, to some extent, similarity to each other, in which high-frequency structures can be seen. Based on the results, for the streamwise direction, the standard deviation, σ , for the stable, isothermal and unstable conditions are $\sigma_{u,stable} = 0.15 \text{ m/s}$, $\sigma_{u,isothermal} = 0.18 \text{ m/s}$ and $\sigma_{u,unstable} = 0.20 \text{ m/s}$, respectively. The mentioned standard deviation values for the vertical velocity component are $\sigma_{w,stable} = 0.17 \text{ m/s}$, $\sigma_{w,isothermal} = 0.20 \text{ m/s}$ and $\sigma_{w,unstable} = 0.23 \text{ m/s}$. Therefore, it can be concluded that, on the whole, with an increase in the flow instability, the deviation of velocity signals from the mean value becomes larger. In other words, as the flow becomes unstable, the intensity of fluctuations gets larger.

Figure 3.12 (c) and (d) compares the power spectrum density (PSD) functions of the streamwise and vertical velocity fluctuations, normalized using the velocity variance (σ^2) along each direction, under different thermal stratification conditions at $x/H = 2$ and $z/H = 0.5$. It should be noted that PSD has been sketched for several points; however, for the sake of brevity, just the one at $x/H = 2$ and $z/H = 0.5$ is provided in the present study.

Considering Figure 3.12 (c) and (d), at low frequencies, a large eddy behavior can be seen. Furthermore, $-5/3$ power law inertial sub-range, and decay at high frequencies are observable, as is the case with a general spectrum shape. This is attributed to the cascade of energy from low-frequency eddies to the part of the spectrum with higher frequencies. All cases show several high-energy frequencies (i.e., dominant frequencies). At low frequencies, for each of the studied conditions, the turbulence energy is almost maintained at a specific level. It is noteworthy that higher energy levels for the isothermal and unstable cases, compared to the stable condition, indicate a larger mixing effect caused by larger velocity fluctuations. Therefore, smaller recirculating regions form downstream of the building under the isothermal and unstable conditions compared to the stable case. As the frequency increases, the energy decays at different rates. In both directions, it can be seen that the frequency content of the velocity fluctuations for the stable case is less than that under the other two stratification conditions, which agrees with the time history of fluctuations. This can be due to lower vortex shedding under the stable condition compared to the isothermal and unstable cases [186]. It can be observed that the spectra generated for the stable case decay more rapidly than the other two cases, followed by the isothermal and unstable conditions.

3.4.4 POD of the flow field

POD is a powerful tool for analyzing turbulent flow since, contrary to the conventional CFD approaches (which use time-averaged and pointwise statistic), it can capture the combined spatio-temporal nature of the flow dynamics [128]. In other words, the chronological vortex development is not obtainable using the time-averaged flow field [187]. In the present study, the snapshot POD technique is used to quantitatively investigate the impact of thermal stratification conditions on the shape and amplitude of the dynamic modes. POD is applied to the flow field behind the building, where $0.025 \leq x/H \leq 4$ and $0.025 \leq x/H \leq 1.5$, along $y/H = 0$. In the mentioned zone, 3500 snapshots are used for the analysis. It should be noted that POD analysis has been performed over 3000, 3500 and 4500 snapshots, and the results were almost similar. However, in order to reduce minor changes in high POD modes, 3500 snapshots are considered to be adequate for the present analysis. Figure 3.13 illustrates the contribution of each POD mode to the total TKE for different stratification conditions.

As can be seen in Figure 3.13, the contribution of the first thirty POD modes to the total TKE is increased under both the stable and unstable conditions compared to the isothermal case. For the stable and unstable conditions, the first thirty POD modes contain about 80.89% and 81.39% of the total TKE, respectively, while those contain about 65.58% of the total TKE for the isothermal case. Furthermore, the first four POD modes contain about 31.09%, 20.94%, and 34.23% for the stable, isothermal, and unstable conditions, consecutively. From Figure 3.13, it could be interpreted that the impact of thermal stratification conditions on the contribution of each mode to the total TKE is more dependent upon the magnitude of Ri rather than its sign. However, this claim should be further investigated for different values of Ri since a general assertion cannot be announced with a limited number of simulations.

Figures 3.14 and 3.15 illustrate the first four POD modes for the velocity field along the streamwise and vertical directions, respectively, in the $y/H = 0$ plane. As outlined in Appendix B, each velocity component (i.e., in the streamwise, lateral and vertical directions) can be attributed to the spatial eigenvector along the specified direction, $\varphi_i^n(\underline{x})$, which separately represents the contribution of a particular mode to the spatial structure [128].

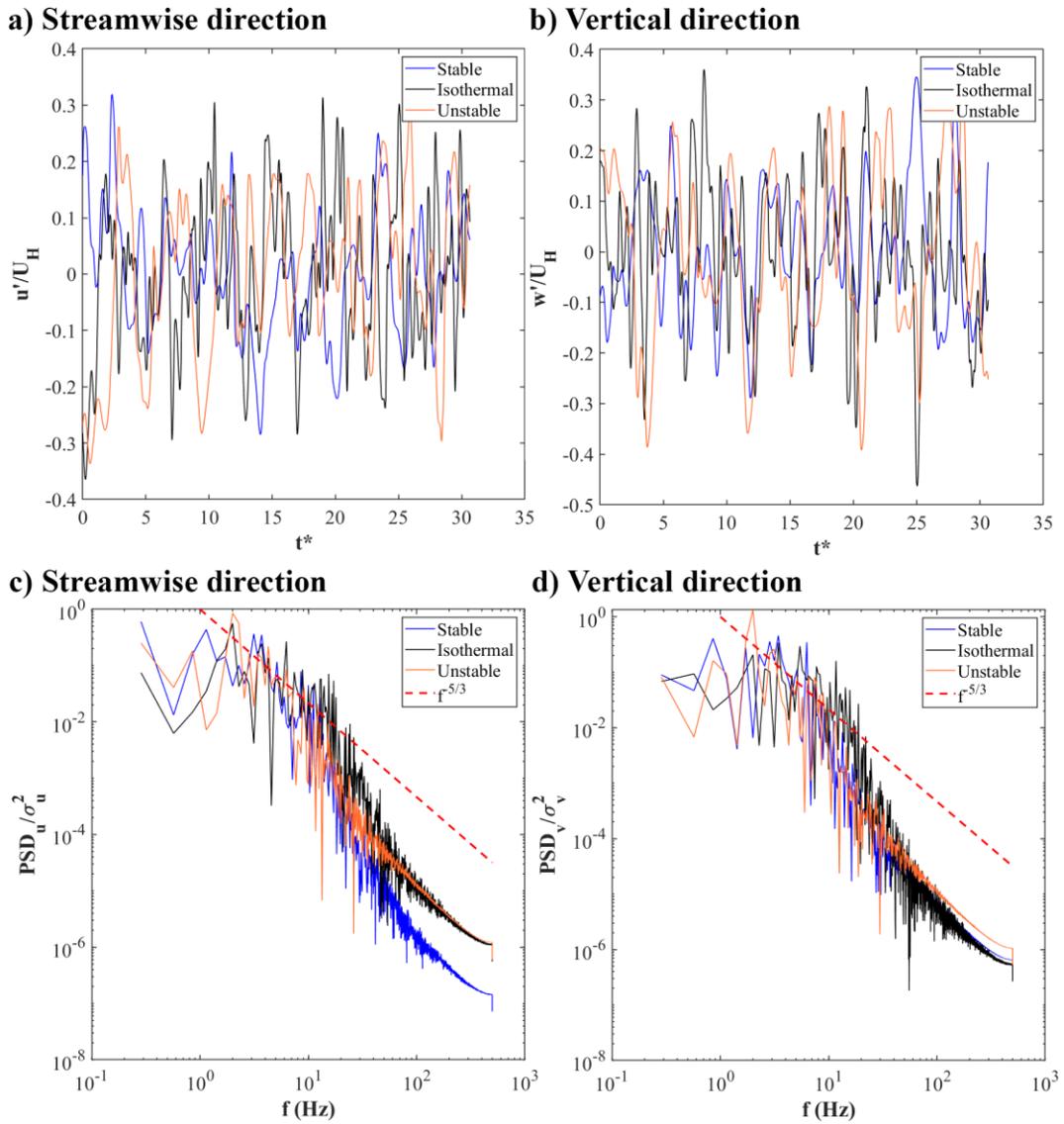


Figure 3.12: First row: time histories of (a) streamwise and (b) vertical velocity components at point $x/H = 2$ and $z/H = 0.5$, along $y/H = 0$, for different thermal stratification conditions,

Second row: power spectrum density (PSD) functions of (c) streamwise and (d) vertical velocity fluctuations, normalized by directional velocity variance (σ^2), under different thermal stratification conditions at $x/H = 2$ and $z/H = 0.5$

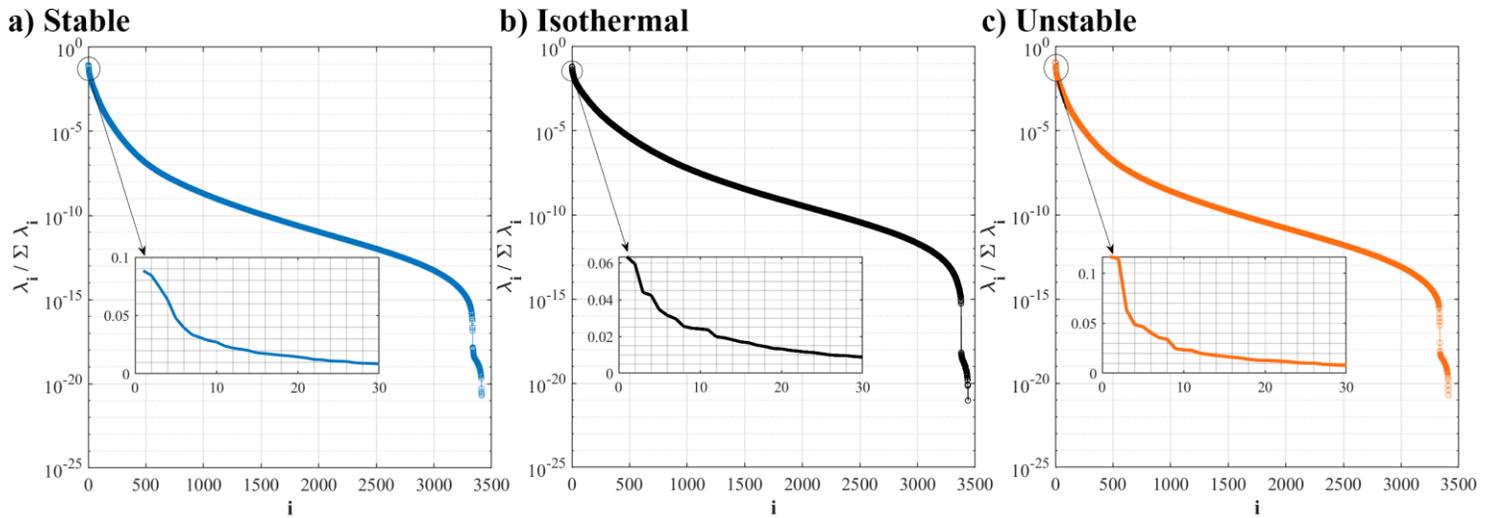


Figure 3.13: Contribution of each POD mode to total TKE for different stratification conditions

Changes in the sign of the iso-lines can be interpreted as the presence of large vortical regions in the region [128]. As shown in Figure 3.14, the stratification condition has significantly changed the flow field structure in the streamwise direction. However, the vortical regions under the stable and isothermal conditions are somewhat identical to each other. For example, a vortical region can be seen in the first mode of the two mentioned cases at approximately $z/H = 0.5$. On the other hand, the instability condition has completely changed the turbulence structure, especially in the proximity of the wall. Furthermore, zone alterations in the unstable case usually occur in the vicinity of the leeward wall. This can be due to the fact that the mixing process is affected by the leeward wall temperature. Therefore, mixing occurs in the vicinity of the wall. However, for the stable and isothermal conditions, the mentioned changes can be seen farther from the leeward wall compared to the unstable case.

a) Stable

b) Isothermal

c) Unstable

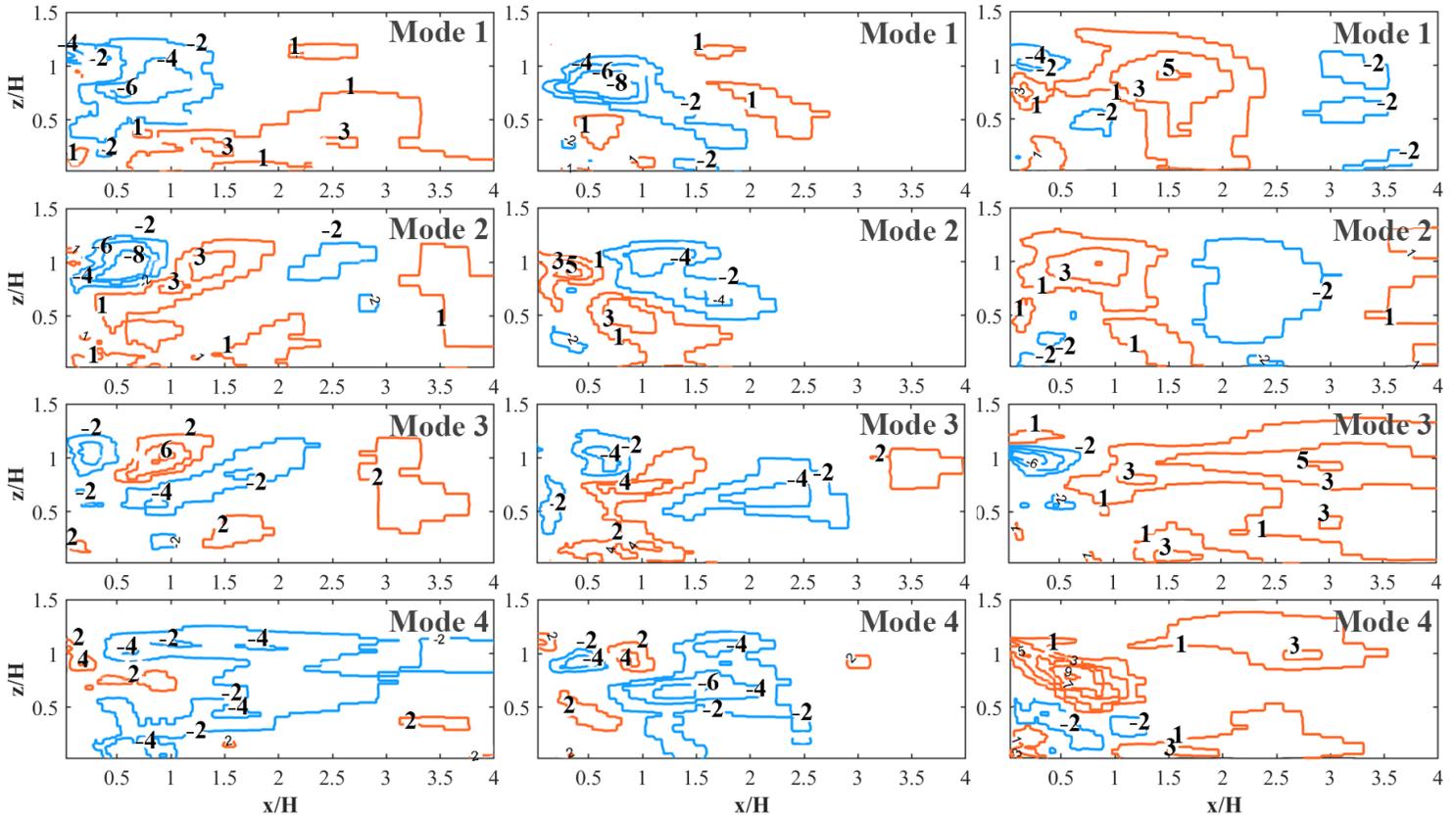


Figure 3.14: Iso-lines of the first four velocity POD modes in the streamwise direction along $y/H = 0$ for different thermal stratification conditions (positive values: solid orange lines; negative values: dashed blue lines)

The positive and negative values of the POD modes of the vertical velocity components represent the upwash and downwash flows [128]. The mentioned feature is of utmost importance in studying pollutant dispersion since these structures are responsible for pollutant transportation in a transient manner. As can be seen in Figure 3.15, the stratification condition has significantly altered the velocity POD modes in the vertical direction with respect to the isothermal case. The positive iso-lines (i.e., upwash flow) in the proximity of the leeward wall are observable in all the modes shown for the unstable case. However, since the thermal effects are weak, it can be seen in the figure that the generated upwash flow is not very strong compared to the other cases. On the other hand, in the mentioned region, in other conditions, particularly the stable case, the downwash flow is observed.

a) Stable

b) Isothermal

c) Unstable

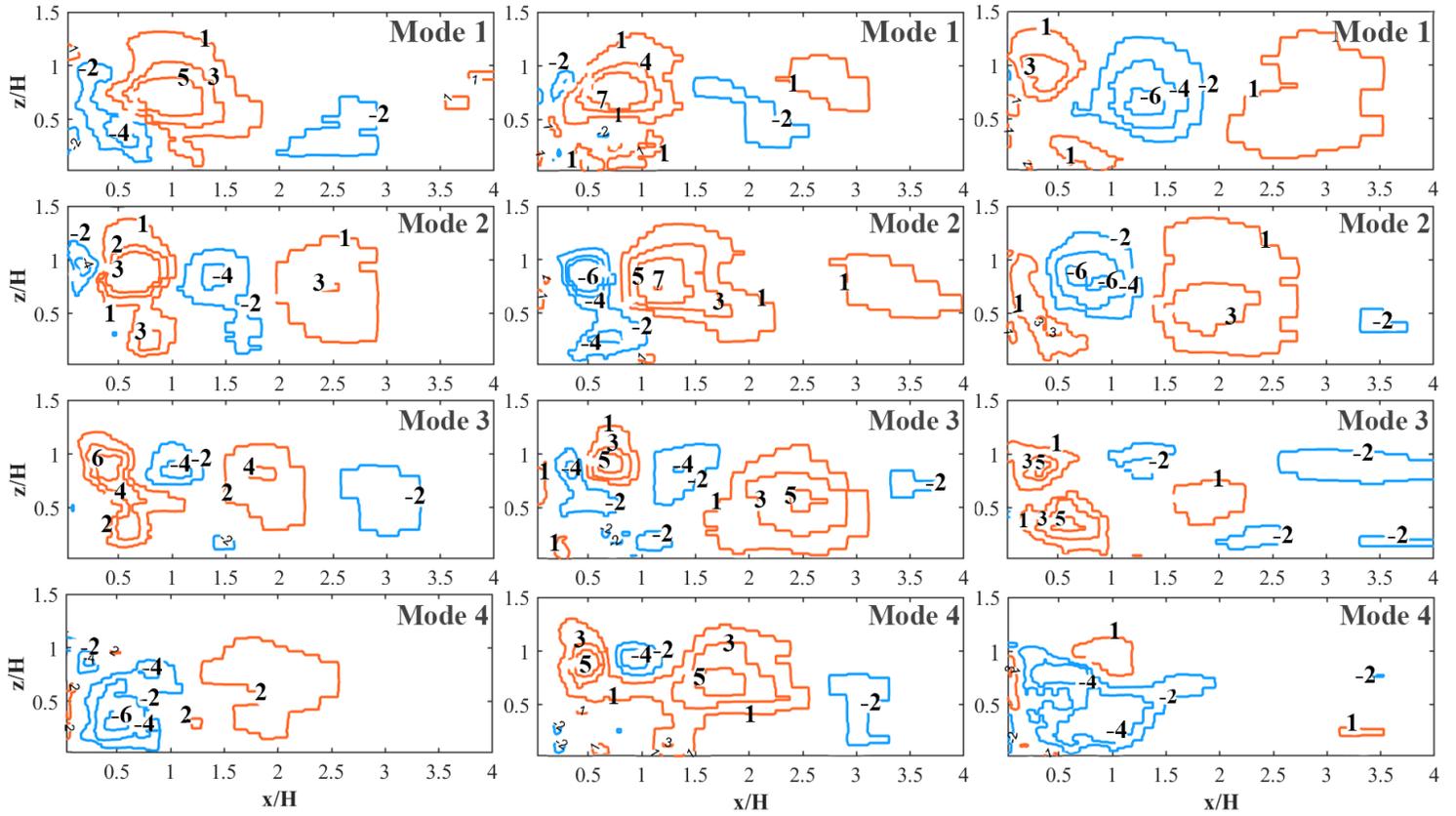


Figure 3.15: Iso-lines of the first four velocity POD modes in the vertical direction along $y/H = 0$ for different thermal stratification conditions (positive values: solid orange lines; negative values: dashed blue lines)

The PSD functions of the coefficients of the first four POD modes for different stratification conditions are provided in Figure 3.16. As can be observed, the stability/instability condition increases the energy content at high frequencies. The trends of the profiles for the isothermal case are almost similar for the first four modes, with some alteration in the range of the energy content at higher frequencies, particularly for the fourth mode. For the unstable condition, for all the modes shown, $f = 1.9989 \text{ Hz}$, corresponding to the Strouhal number of $St \approx 0.23$, is one of the frequencies at which a peak occurs. The mentioned frequency is the same as the one at which the peak occurs in the spectral analysis of the streamwise and vertical velocities (see Figure 3.12 (c) and (d)). This mode plays an essential role in vortex shedding in the wake region. For the isothermal case, $f = 3.4266 \text{ Hz}$ ($St \approx 0.39$) is the one at which the peak occurs in the third and fourth mode, which is exactly the same as the one at which the spectral analysis of the vertical velocity shows its peak in Figure 3.12 (d). For the stable condition, although the ultimate peaks in Figures 3.12 and 3.16 do not exactly occur at the same frequency, their correspondence frequencies are close to each other. By matching the mentioned frequencies, the contribution of the dominant modes to the turbulent flow structure and pollutant dispersion process can be addressed. In future

works, researchers can focus on relating the dominant modes and the frequencies at which the peak occur to the turbulence structures to reveal the impact of different stratification conditions on turbulence features.

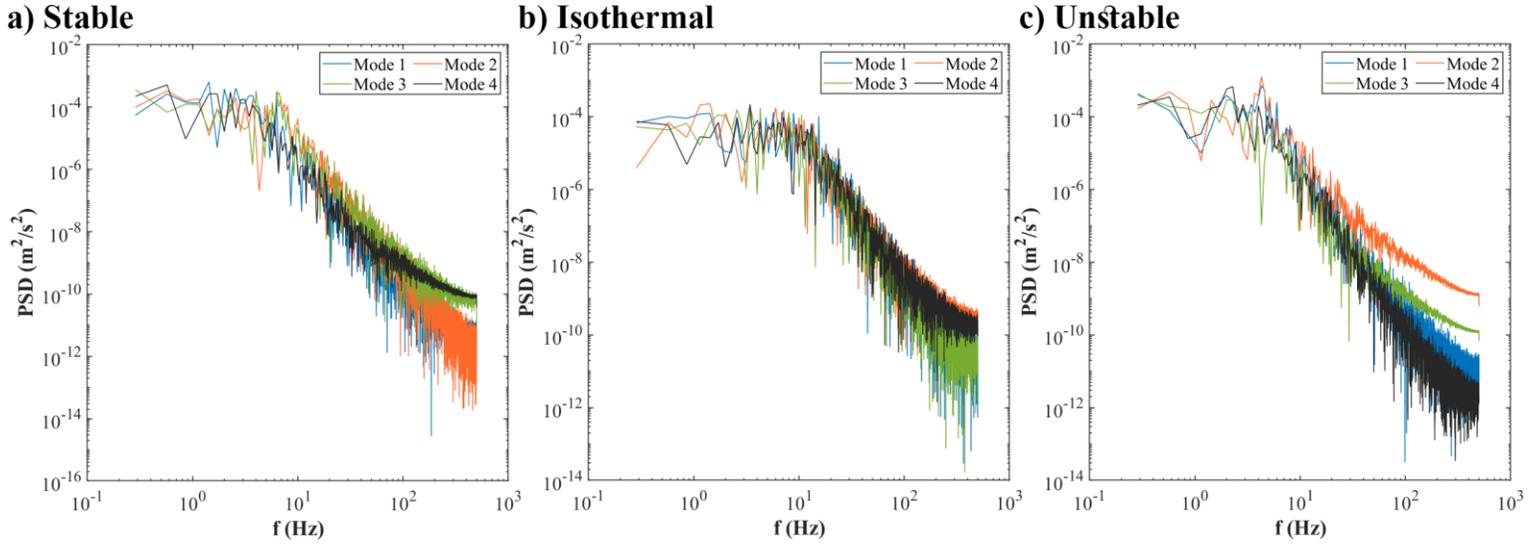


Figure 3.16: Power spectral density (PSD) functions of the coefficients of the first four POD modes

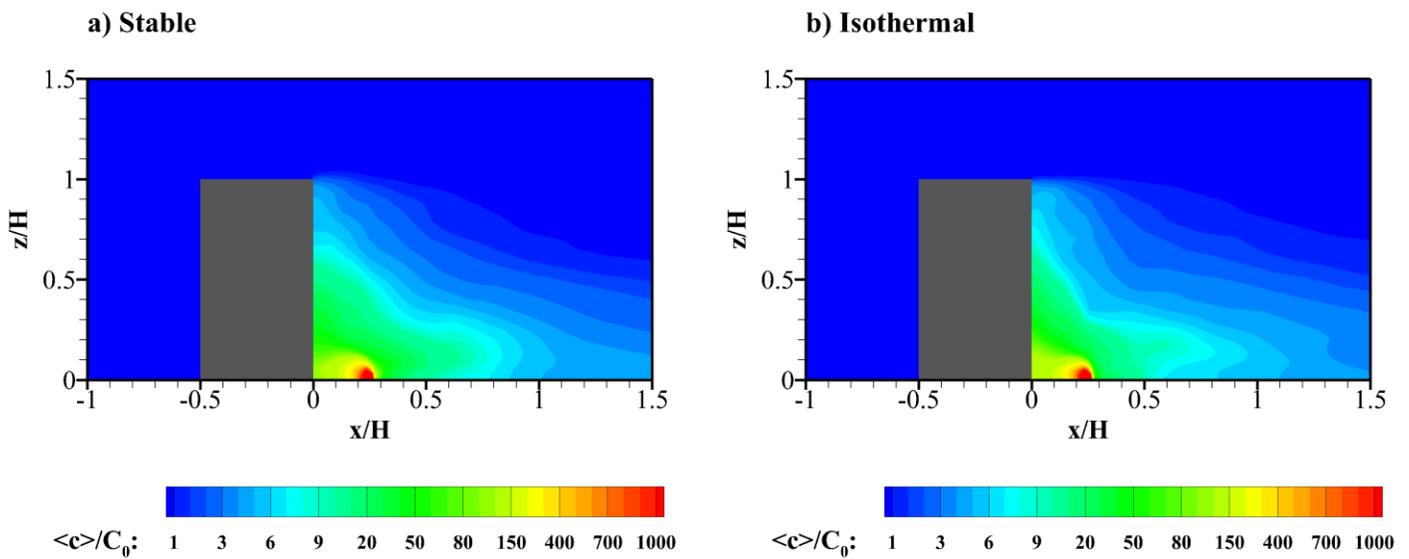
3.4.5 Concentration field

By considering the ELES results depicted in Figure 3.8, it can be found out that the impact of thermal stratification condition on the time-averaged concentration field is more pronounced below $z/H = 0.5$. On the other hand, as the height increases, the role of thermal effects in altering the time-averaged concentration field diminishes. Therefore, it can be concluded that even a weakly stable/unstable stratification condition can significantly affect the pollutant concentration at the pedestrian level. Over $x/H = 0.375$, for $z/H \leq 0.5$, the stable condition increases the concentration level by about 33.9%, with respect to the isothermal case, while the unstable condition reduces the concentration level by approximately 28.2%. The mentioned increase (reduction) in the concentration level due to the stable (unstable) condition with respect to the isothermal case, at $z/H \leq 0.5$, over $x/H = 0.625$, $x/H = 1$ and $x/H = 1.5$ is about 15.0% (17.3%), 8.2% (28.1%), and 17.3% (22.6%), respectively. Thus, it can be interpreted that the influence of thermal stratification conditions (even weak ones) on the concentration field is not only restricted to the region near the leeward wall.

Figure 3.17 illustrates the contours of dimensionless time-averaged pollutant concentration, $\langle c \rangle / C_0$, under different stratification conditions along $y/H = 0$. Pollution is more vertically dispersed along the leeward façade under the isothermal and unstable conditions compared to the stable one. Furthermore, the highly polluted region is more compact under the unstable condition in comparison with the other two cases. This is in line with the finding of [36], in which they

reported that the plume disperses more horizontally as the stability level increases since the vertical movement is suppressed.

Figure 3.18 illustrates the time-averaged and instantaneous plume shapes, obtained by the iso-surfaces of $\langle c \rangle / C_0 = 1$ and $c / C_0 = 1$, respectively, colored by k / U_H^2 . It can be observed that TKE increases at the building height and also in the wake region. An increase in TKE values leads to an increase in the concentration fluctuations [29]. Thus, the plume shape is shorter under the unstable condition compared to the other two cases. Furthermore, it can be seen that, for the unstable case, the width of the plume shape is less than that under the isothermal and stable conditions. This is in-line with the findings from the frequency analysis part regarding the length of the recirculating regions under different thermal conditions.



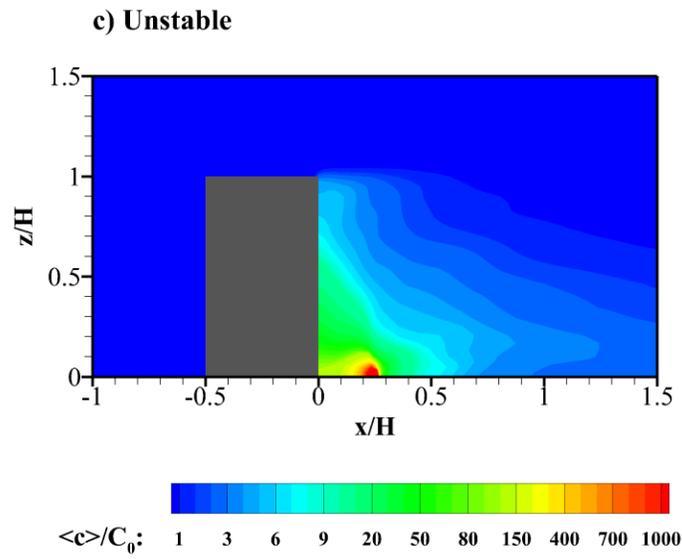


Figure 3.17: Contours of time-averaged pollutant concentration, $\langle c \rangle / C_0$, under different thermal stratification conditions along $y/H = 0$

Two different mechanisms, namely, convective flux (i.e., $\langle u_i \rangle \langle c \rangle$) and turbulent diffusion flux (i.e., $\langle u_i' c' \rangle$) are responsible for the pollutant transportation process. Figure 3.19 shows the contours of the mentioned fluxes in the streamwise direction, normalized by the product of the streamwise velocity at the building height and the reference concentration, along $y/H = 0$. Based on the results, thermal stratification condition has a negligible effect on the convective flux in the wake region, in the proximity of the pollutant source. On the other hand, the impact of the stratification condition on the turbulent diffusion flux is more remarkable. As the instability level increases, the streamwise flux in the vicinity of the leeward wall also increases. Furthermore, the iso-lines of the turbulent diffusion flux under the unstable condition tend toward the leeward wall, while those under the stable condition are extended in the streamwise direction. This can explain the reason for the more elongated plume shape under the stable condition compared to the other two cases.

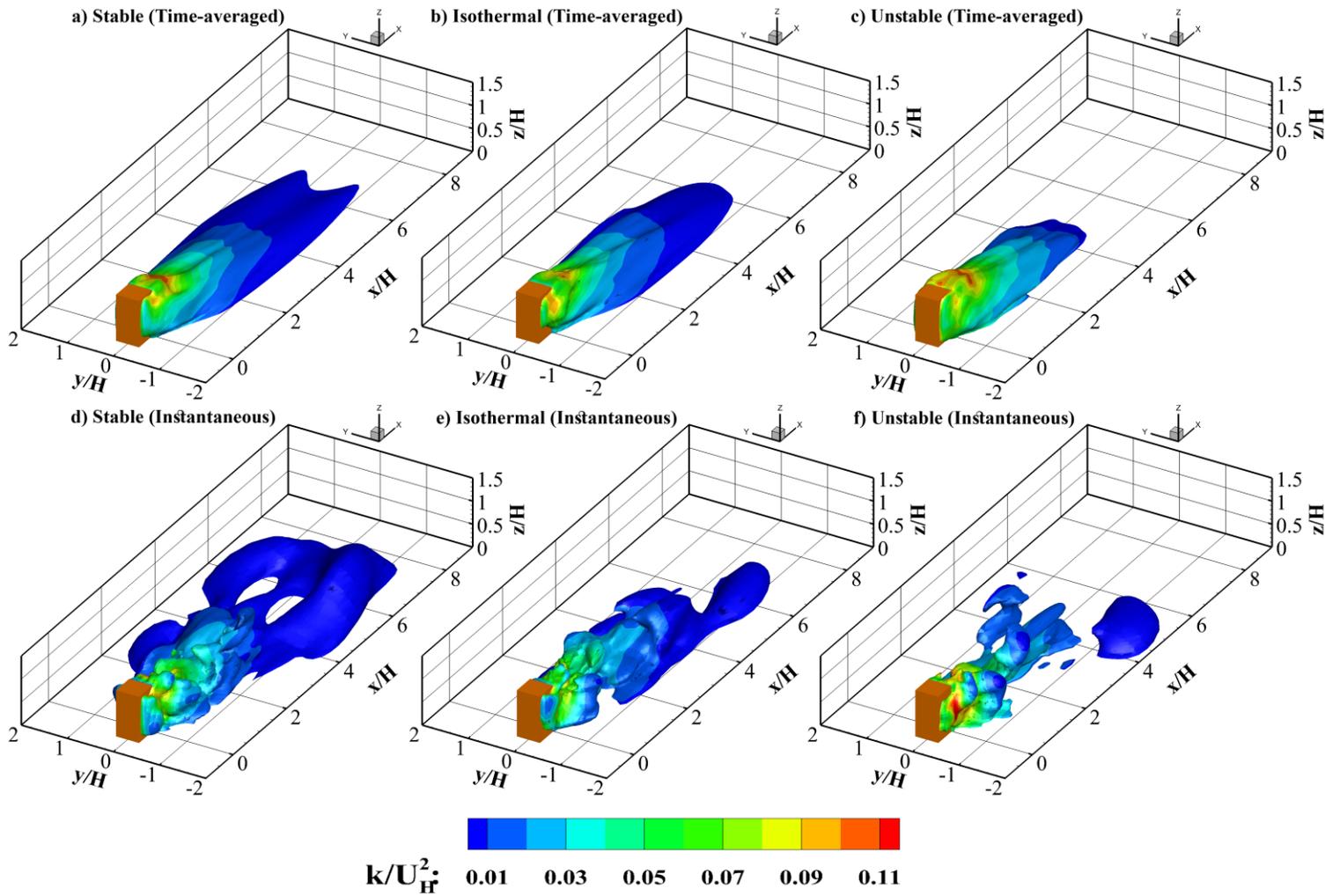


Figure 3.18: Comparison between three-dimensional time-averaged and instantaneous plume shapes, obtained by iso-surfaces of $\langle c \rangle / C_0 = 1$ and $c / C_0 = 1$, respectively, colored by dimensionless turbulence kinetic energy, k / U_H^2

The contours of non-dimensional convective and turbulent diffusion fluxes in the vertical direction, along $y/H = 0$, are depicted in Figure 20. It can be observed that, near the leeward façade, the convective flux under the stable condition is weaker than that under the other two cases. Moreover, below $z/H < 0.3$, the contours of convective flux under the unstable and isothermal conditions are more elongated in the streamwise direction compared to the stable one. This can be considered as an explanation for more dispersion along the leeward wall under the unstable and isothermal conditions compared to the stable case. It can be seen that the unstable condition enhances the strength of the turbulent diffusion flux in the vertical direction, in the proximity of the leeward wall.

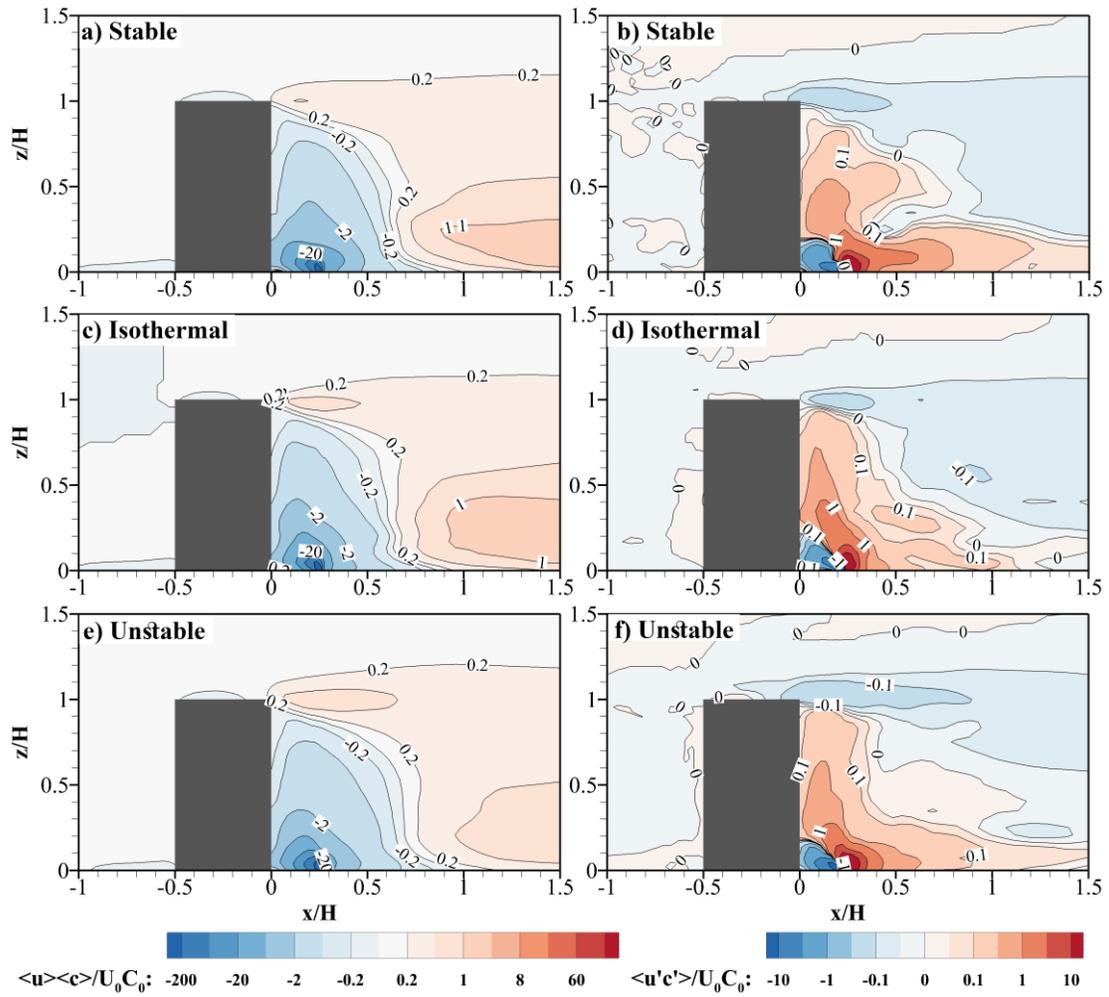


Figure 3.19: Convective flux, $\langle \mathbf{u} \rangle \langle c \rangle$, and turbulent diffusion flux, $\langle \mathbf{u}'c' \rangle$, in streamwise direction along $y/H = 0$ under different thermal stratification conditions

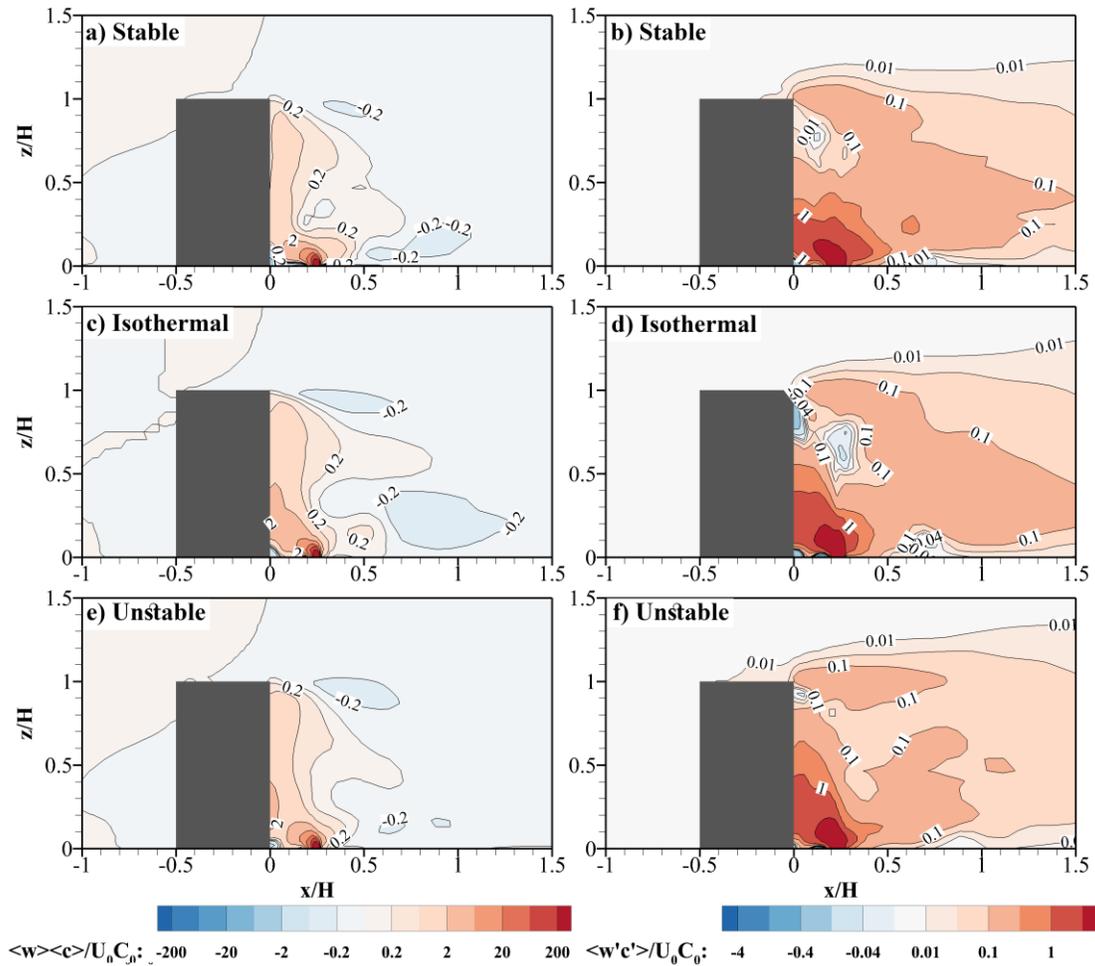


Figure 3.20: Convective flux, $\langle w \rangle \langle c \rangle$, and turbulent diffusion flux, $\langle w' c' \rangle$, in vertical direction along $y/H = 0$ under different thermal stratification conditions

3.5 Conclusions

Thermal stratification conditions significantly affect the airflow pattern and pollutant dispersion process within an urban area. In order to have a reliable simulation, particularly under non-isothermal conditions, in which the role of turbulence is of utmost importance, the use of a high-fidelity CFD model seems to be necessary. So as to make a balance between computational costs and accuracy, the zonal URANS-LES approach, also known as the ELES, was used in the present study. ELES was employed for three different stratification conditions, namely, stable, isothermal, and unstable.

In the present study, the ELES results, including the time-averaged flow, temperature and concentration fields, TKE, and concentration fluctuations, were quantitatively compared with a wind tunnel experimental data [164] and an LES study [103]. Regarding the time-averaged

velocity field, the ELES results were in good agreement with both the wind tunnel data and the prediction made by LES. ELES over-predicted the TKE values in locations with strong shear layers. The findings showed that ELES has the capability to predict the time-averaged temperature field similar to LES, if it is set up properly. Furthermore, regarding the time-averaged and fluctuating concentration fields, ELES outperformed LES, particularly under the unstable condition. Therefore, in addition to lower computational costs of ELES compared to LES, the results showed that, under proper settings, its performance in predicting the concentration field can be better than LES.

Regarding the impact of thermal stratification conditions on the flow field and TKE, the unstable condition had a significant effect on the mentioned fields, while the stable condition slightly altered the results compared to the isothermal case. Considering the time-averaged concentration field, the stable condition increased the pollutant concentration by 33.9%, while the unstable one lowered that by 28.2%, both compared to the isothermal case. Regarding the pollutant dispersion mechanisms, an increase in the flow instability led to higher convective and turbulent diffusion fluxes, particularly in the vicinity of the leeward wall.

The PSD functions of the streamwise and vertical velocity components showed lower energy content of the stable condition compared to the isothermal and unstable cases for almost all the frequency range. This can be attributed to the lower vortex shedding under the stable case. POD analysis revealed that altering the stratification condition to non-isothermal, either stable or unstable, increased the contribution of the primary dominant POD modes to the total TKE. This shows that the mentioned contribution may be mainly dependent upon the magnitude of the Richardson number. Furthermore, the results showed that the unstable condition significantly changed the turbulence structures of the POD modes in both streamwise and vertical directions.

Chapter 4. A review of advances towards efficient reduced-order models (ROM) for predicting urban airflow and pollutant dispersion

4.1 Introduction

4.1.1 Motivation

Due to emissions from power plants located upstream of an urban area, from transportation or other anthropogenic activities, pollution often accumulates in urban areas, with higher levels than in the outskirts of those regions. Not only does pollution accumulation influence the outdoor air quality, but it can also affect the indoor air quality through the windows (i.e., air infiltration) or even the intake of heating, ventilation, air conditioning (HVAC) systems [188]. As reported by the World Health Organization (WHO), in 2019, 99% of the world population lived in areas where the WHO air quality guidelines were not met [189]. It has also been mentioned that, in 2016, about 4.2 million premature deaths occurred worldwide due to outdoor air pollution. These issues have led researchers to focus on outdoor air quality and related studies like urban airflow prediction.

Factors influencing the airflow structure and pollutant dispersion process within urban areas can be categorized into three groups [190,191]:

1. Meteorological factors such as wind speed and direction, and thermal stratification conditions, etc.,
2. Different urban configurations (i.e., urban morphology), such as the canyon's aspect ratio, non-uniformity of buildings' heights, the density of the urban area, surface coverage, etc.,
3. Moving vehicles.

Table 4.1 lists the research items reviewing the impact of the aforementioned factors on pollutant dispersion and urban airflow.

Table 4.1: An overview of previous review articles on pollutant dispersion and urban airflow

Reference	Year	Research focus
-----------	------	----------------

[192]	2005	Impact of urban configuration, variable wind direction, and traffic-induced turbulence on the flow and dispersion patterns
[193]	2012	Studying the role of urban physics in various fields such as pollutant dispersion and thermal comfort
[194]	2013	Reviewing techniques of using computational fluid dynamics (CFD) in near-field dispersion studies (reviewing a limited number of studies to address the impact of the inflow stability on pollutant dispersion)
[195]	2014	State-of-the-art review of different modeling and measurement techniques in studying flow and concentration fields, considering several influential factors, such as thermal effects, building geometries, etc.
[196]	2016	State-of-the-art review of experimental and numerical studies focusing on near-field pollutant dispersion
[197]	2017	Reviewing the impact of green infrastructure on pollutant dispersion in open roads and built-up canyons, mainly considering its aerodynamic and barrier effects
[198]	2017	Reviewing microclimate CFD studies, categorizing based on the urban morphology and the methodology
[199]	2018	Reviewing urban tree modeling using CFD, considering aerodynamic, deposition, and thermal effects
[200]	2018	Review on isothermal (neutral) urban airflow and pollutant dispersion, considering time-varying inflow conditions
[191]	2020	Reviewing the impact of different factors, such as urban configuration, time-varying inflow conditions, thermal effects, tree planting, etc., on urban airflow and pollutant dispersion
[13]	2020	Reviewing theoretical, experimental, and numerical research items on isothermal and non-isothermal flows in street canyons

[190]	2021	Review on the impact of various factors, such as inflow conditions, thermal effects, geometrical factors, etc., on pollutant dispersion
-------	------	---

[51]	2021	Focusing on urban airflow using experimental and numerical simulations, and shedding light on the usage of data-driven methods to dynamics of urban flows
------	------	---

[201]	2021	Reviewing challenges in using CFD in urban studies, such as simplifications of boundary conditions, data acquisition, and computational barriers
-------	------	--

[46]	2021	Reviewing applications of data-driven methods in built environment problems, such as turbulence closure, super-resolution, reduced-order modeling, etc.
------	------	---

Factors related to urban morphology are usually constant throughout a long period of time for a particular area. However, considering changes in urban morphology can be useful for an efficient design of urban environments regarding pollutant dispersion, pedestrian's health and comfort, etc. Detailed information regarding the impact of urban morphology on urban airflow and pollutant dispersion can be found in [190,191].

Contrary to the urban configuration, the other two factors, namely metrological and pollution sources, frequently change. Moving vehicles can be considered as moving pollution sources. Furthermore, traffic-related factors and shapes of vehicles can alter the airflow and turbulence fields in urban areas. Earlier research work reported the impact of the motion of vehicles on the airflow and contaminants concentration fields [202–206].

Changes in the wind speed and direction significantly influence the dispersion process [190,191]. For instance, in simulating the airflow and contaminants concentration fields in a street canyon, changes in the inflow wind conditions caused alteration in the vortical region formed in the canyon [207]. Also, considering the real-time incoming wind speed improves prediction of contaminant dispersion [191,207,208]. Furthermore, in low wind speed conditions, it is required to consider the thermal condition of the region [190,191]. Changes in the thermal condition of an urban area can result in alterations of the turbulent diffusion and convective fluxes, which lead to changes in the pollutant dispersion pattern [94,209]. Therefore, considering the frequent changes in the meteorological conditions is of utmost importance in urban-related problems. Furthermore, (near) real-time simulation of the airflow and contaminant concentration fields is required in built environment applications, particularly for pollutant dispersion in urban areas in emergency situations [46,210]. Also, long-term simulations are necessary for investigating the ventilation and pollution characteristic of an urban area [45].

A comprehensive model must account for the aforementioned meteorological complexities. CFD is considered the most common and promising simulation approach in fluid dynamics problems,

such as urban-related ones. According to Table 4.1, the review articles are mainly focused on CFD methods. However, due to the limitations of CFD, the use of data-driven approaches in fluid dynamics applications is becoming more common in different aspects, such as turbulence closure, super-resolution, reduced-order modeling, etc. [46,51,211]. In order to understand the necessity of employing data-driven methods in simulating fluid dynamics applications, the following subsection is devoted to the application of CFD in urban-related problems and its limitations.

4.1.2 Pros and cons of CFD application in urban-related problems

CFD methods have extensively been used in urban-related problems [129] due to their capability of capturing different length scales and the turbulence nature of the flow field. Also, their grid-based approach provides the opportunity of controlling simulation accuracy. CFD approaches can be categorized into three main groups, namely, Reynolds-averaged Navier-Stokes (RANS), large eddy simulation (LES), and direct numerical simulation (DNS). Among these methods, the RANS and LES approaches have commonly been used in urban studies such as airflow prediction and pollutant dispersion [15,24,184,186,209,212,213] due to their more reasonable computational costs compared to DNS. A complete comparison between the performance of RANS and LES in built environments can be found in [129].

While RANS provides the average quantities, based on time-averaging, LES uses a spatial filtering method to separate large eddies from small ones [179]. By selecting a filter width (i.e., cutoff width), LES resolves the eddies larger than the cutoff width while modeling those with smaller sizes than the mentioned width using a subgrid-scale (SGS) model. Thus, LES can capture fluctuations due to large-scale eddy motions [124,130]. Capturing fluctuations of the airflow and concentration fields is essential in dispersion studies since it is related to pollutant dispersion mechanisms (i.e., turbulent diffusion flux) [107,116,131]. LES can generate acceptable results regarding the mean flow field and turbulence kinetic energy (TKE) even with a simplified SGS model since the vortex shedding around buildings can be captured by LES, while RANS cannot reproduce it [124].

RANS models tend to over-predict pollutant dispersion [124]. Several research items have shown the superiority of LES compared to RANS in predicting the airflow and concentration fields [128,132,133]. However, it should be mentioned that significant discrepancies have been reported between the simulated concentration fields by LES and the experimental data in the vicinity of the pollutant source, located on the ground, for both indoor and outdoor regions [25,135].

It is noteworthy to mention that non-isothermal flows can be formed in urban areas, particularly dense ones, due to the urban heat island (UHI) effect. Since increasing the flow instability increases the TKE level, which increases the concentration fluctuations [209], the capability of the model in predicting fluctuations becomes crucial. LES has shown a better performance than RANS and unsteady RANS (URANS) in predicting the airflow field within street canyons under non-isothermal conditions [136]. Also, the incapability of URANS in predicting the turbulence nature of the flow under unstable thermal conditions has also been reported by Nazarian and Kleissl [137]. It should be noted that although URANS is an unsteady approach, it is not able to capture the internally-induced fluctuations of flow [214]. URANS models can be used for non-stationary

flows, and they can somehow predict vortex shedding, which is related to the largest unsteady scales. However, they fail to capture the remaining eddy scales. This can be due to the fact that they rely on the mean airflow equations but perform ensemble averaging. Furthermore, LES has performed better than RANS in capturing vortex shedding and vortical regions in the wake of an isolated building under an unstable non-isothermal condition [74].

However, since RANS is computationally efficient, its usage is beneficial in urban studies, particularly those covering large areas. The computational resources required for LES are much higher than those for RANS. The CPU time needed for the statistical convergence of LES was reported to be 25 times more than that for RANS in studying the dispersion process around an isolated building [126]. The situation can even be exacerbated for a large number of buildings [138]. In order to reduce computational costs while retaining the pros of LES, a combination of URANS and LES, such as detached eddy simulation (DES) [144], delayed detached eddy simulation (DDES) [145], and also multi-domain (multi-zonal) models (e.g., zonal or embedded large eddy simulation, ELES) [29,146] can be used [145–150].

Masoumi-Verki et al. [209] reported that using ELES resulted in the speed-up of about 292% and 130% compared to LES for the isothermal and non-isothermal cases, respectively. Furthermore, Liu and Niu [153] showed that DDES reduced computational costs by 40% compared to LES while retaining accuracy at an acceptable level in simulating the airflow field around an isolated high-rise building under the isothermal condition. In another study, the use of DDES resulted in reducing computational costs by approximately 20% compared to LES [157]. However, even with the mentioned reductions in computational costs, using high-fidelity CFD models in predicting the airflow and concentration fields is still costly. For instance, in order to use ELES for simulating the airflow and contaminant concentration field around an isolated high-rise building, about 7.6 and 14 days of data sampling were required to reach statistical convergence under the isothermal and non-isothermal conditions, consecutively [209].

Therefore, it can be concluded that, even with the aforementioned modifications, conducting long-term CFD simulations with dynamic boundary conditions is still time-consuming [45]. Also, (near) real-time simulations are practically infeasible using CFD approaches. In order to overcome the mentioned limitations, using data-driven methods, in tandem with CFD, or in some cases that the required data is available from other sources than numerical simulations, as an alternative for CFD, is becoming more popular [46,51,211].

4.1.3 Objective and contribution

The present chapter aims at reviewing the advances and challenges of data-driven approaches, particularly machine learning algorithms, in the development of reduced-order models (ROMs) for predicting the airflow and contaminant concentration fields. It should be noted that using the mentioned algorithms in fluid dynamics is not restricted to reduced-order modeling. They can also be used in dimensionality reduction, feature extraction, flow control, turbulence closure, shape optimization, etc. [46,47]. However, the present chapter is focused on reduced-order modeling.

Using ROMs can lower the CPU time for predicting the airflow and contaminant concentration fields compared to conventional CFD simulations, enabling them to be used for (near) real-time simulations. Also, the mentioned models can provide the opportunity of conducting long-term simulations with dynamic boundary conditions. However, due to the black-box nature of machine learning methods, they are not usually physically interpretable (i.e., it is difficult to analyze the obtained results based on the physics of the problem). Increasing the physical interpretability of such models and embedding physics of the problem into them are challenging tasks. After providing an exhaustive review of the existing models, the present chapter tackles the mentioned challenges for the development of an efficient while physically interpretable ROM. In addition to addressing different data-driven methods for the development of ROMs, the present review tries to relate them to the physical aspects of the problems as much as possible.

4.2 Reduced-order modeling

4.2.1 Coherent structures

Coherent structures in turbulent flows are considered as regions that contain the most significant characteristics of flows [215]. Generally, turbulence is composed of eddies with different length scales [216]. Within the energy spectrum, these ranges are the energy-containing (permanent) sub-range, inertial sub-range, and dissipation sub-range. Each of these ranges contains a part of the energy of the flow. Therefore, in order to reduce computational costs, considering the ranges of eddies that cover the most important parts of the energy of the flow results in a good approximation of the whole ranges. These low dimensional structures (i.e., coherent structures) can be described by dynamical systems [217].

Proper orthogonal decomposition (POD) [218,219] is one of the most common approaches for analyzing turbulent flows since it can capture the combined spatio-temporal nature of the flow dynamics [128]. POD was first introduced to the field of fluid dynamics by Lumley [218], and it can capture the most dominant features of turbulent flows. In other words, the most energetic modes in the system can be extracted using the singular value decomposition [217]. Thus, instead of using a high-dimensional system, a limited number of deterministic functions can be used for analysis and representing the turbulent flow [220]. For instance, Masoumi-Verki et al. [209] reported that within the wake region of an isolated high-rise building located in stable and unstable thermal stratification conditions, the contribution of the first 30 POD modes to the total TKE is about 80.89% and 81.39%, respectively. This value has been reported to be 65.58% for the isothermal case. Thus, a few characteristic modes can capture a large portion of the total TKE, which is an important finding in the development of ROMs.

4.2.2 Reduced-order models (ROMs)

The application of ROMs has gained a lot of attention in the past few years in fluid dynamics problems, such as urban-related ones [221,222]. Complex systems can be simulated using ROMs with reduced computational costs [221]. In other words, it reduces the degrees of freedom (dof) of the systems, while the key features will be maintained [223]. Furthermore, multi-query problems

(here, alterations in the system's boundary conditions) can be solved using reduced-order modeling approaches [224,225].

Generally, ROMs can be classified into two main categories: intrusive and non-intrusive [226,227]. The intrusive ROMs (IROMs) are projection-based methods in which the governing equations describing the problem are projected onto a low-dimensional space, so-called the reduced space [228]. Therefore, a reduced-size, physics-based model is generated. The most common approach for deriving the reduced space is the POD technique. Other popular approaches are dynamic mode decomposition (DMD) [229,230] and greedy algorithms [231], and among nonlinear ones, the gappy POD method [232] and discrete empirical interpolation [233], which are modifications of POD to reduce the complexity related to the evaluation of the nonlinear terms of the ROM. IROMs have been developed for built environment (i.e., both indoor and outdoor) studies, such as deriving low-dimensional ventilation model [234], simulating indoor pollutant dispersion [235], simulating time-dependent pollution release from a power station [236], simulating pollutant dispersion within street canyons [237], etc.

The most common projection-based approach is the POD-Galerkin method. It is noteworthy to mention that the projection-based methods have usually been combined with solvers based on the finite element method [237–240]. However, since the majority of CFD solvers are based on the finite volume method, applying the POD-Galerkin method to the finite volume approximation has recently become popular [228,241,242].

Stability and nonlinear inefficiency can be named as the main issues of the POD-Galerkin method [225,243,244]. For instance, the truncational approach of the POD technique may result in less energy dissipation in IROMs compared to high-fidelity solvers. Thus, IROMs can blow up [244,245]. A complete discussion on the stability issues of IROMs can be found in [244]. It should be noted that several methods, such as Petrov-Galerkin projection [246–249], discrete empirical interpolation [250,251], etc., have been proposed to overcome the previously-mentioned issues of the POD-Galerkin method. However, developing IROMs requires source code manipulation and knowledge about the discretization schemes, which can be cumbersome, and in some cases, impractical since there is no access to the source code of the majority of software packages [226].

On the other hand, non-intrusive ROMs (NIROMs), which do not require any source code manipulations, can overcome the IROMs' limitations. Contrary to IROMs, researchers can use NIROMs to avoid instability and overcome the issues regarding the nonlinear inefficiency [249,251,252]. However, they suffer from high computational costs during the training stage. There exist a direct relationship between computational costs and the dimensional size of the system [253]. In other words, since each variable obtained from snapshots should be treated separately, an increase in the number of the variables exponentially increases the computational costs. However, their reconstruction time is much faster than conventional CFD models. For instance, Xiang et al. [225] reported that using NIROM reduced the simulation time by the factor of 226 compared to conventional CFD approaches. In another study, this factor has been reported to be 800 [45]. Xiao et al. [224] also reported that using NIROMs in predicting pollutant dispersion resulted in five orders of magnitude reduction in the CPU time compared to their high-fidelity model. However, NIROMs neglect to incorporate the problems' physics since they only rely on

the obtained data, which may be problematic in engineering applications, thoroughly discussed in section 4.5.

Machine learning (ML) algorithms (and particularly deep learning ones) have become popular in developing NIROMs, due to their capability of extracting correlations among data, dealing with high dimensional data, and low computational costs of data reconstruction. Table 4.2 lists the studies on the development of NIROMs in urban areas to predict airflow field and pollutant dispersion.

Generating a NIROM comprises two different steps: dimensionality reduction and computing feature dynamics. Dimensionality reduction is the process of finding a low-dimensional representation of the original data, while the feature dynamics step is computing the temporal dynamics of the mentioned space. Afterward, the low-dimensional space is reconstructed back to the original space. The following sections of the present chapter discuss these two steps of NIROMs' training.

Table 4.2: An overview of previous studies on the development of NIROM for predicting urban airflow field and pollutant dispersion

Reference	Year	Method(s)	Research focus
<i>Linear dimensionality reduction</i>			
[221]	2019	POD - GPR	Predicting the turbulent flow field in a realistic urban area
[254]	2019	POD - GPR	Predicting the turbulent flow field in a realistic urban area using the domain decomposition method
[255]	2019	POD - Gaussian RBF	Predicting the airflow field within 2D and 3D street canyons using the domain decomposition method
[224]	2019	POD - GPR, LSTM	Predicting urban air pollution with different parameters, i.e., pollutant sources
[256]	2021	POD - RBF	Predicting urban air pollution within a large urban area
[257]	2021	PCA – LSTM and adversarial LSTM	Predicting the airflow and concentration fields – comparing the performance of classic LSTM with the adversarial one

[258]	2021	PC-based adversarial AE – adversarial LSTM	Evaluating the performance of the methodology in predicting the airflow field
<i>Nonlinear dimensionality reduction</i>			
[210]	2018	VAEDC - DNN	Real-time analysis of the probability of death in a realistic urban area with varying wind speed, direction, and pollution discharge rate
[225]	2021	POD, AE-linear-FCNN, AE-FCNN, CAE - XGBoost	Developing a NIROM for predicting urban airflow with transient boundary conditions – comparing the performance of different dimensionality reduction techniques
[45]	2021	CAE - XGBoost	Long-term simulation of urban airflow at city scale

GPR: Gaussian process regression; RBF: radial basis function; LSTM: long short-term memory; PCA: principal component analysis; PC: principal component; AE: autoencoder; VAEDC: variational autoencoder with deep convolutional layers; DNN: deep neural network; AE-linear-FCNN: linear fully-connected neural network; AE-FCNN: nonlinear fully-connected neural network; CAE: convolutional neural network-autoencoder; CNN: convolutional neural network; XGBoost: eXtreme Gradient Boost

4.3 Dimensionality reduction

The idea of dimensionality reduction is to embed the solution manifold onto a lower-dimensional latent space (i.e., reduced space). In simpler words, as can be seen in the previous sections, a few numbers of POD modes can represent the behavior of the flow field [209]. Therefore, one can use the idea of finding a low-dimensional representation of the original data to perform simulations with much lower computational costs. The process can be categorized into two main approaches: the linear and nonlinear dimensionality reduction methods.

4.3.1 POD: a linear dimensionality reduction technique

The POD technique, also known as principal component analysis (PCA), has traditionally been used for deriving reduced basis functions of physical systems [253,259]. Although other linear dimensionality reduction techniques, such as DMD, have been used in fluid dynamics problems for analysis or ROM development [260,261], according to Table 4.2, the linear technique, which has dominated urban-related problems, is still POD. Thus, the focus of the current section is on the POD technique.

In the first step, snapshots' matrix (consider the velocity field, \mathbf{u} , at this step) should be generated from the data obtained by high-fidelity simulation (or from an experiment) [221]:

Eq. 4-1

$$S = [\mathbf{u}^1 \ \mathbf{u}^2 \ \mathbf{u}^3 \ \dots \ \mathbf{u}^{N_s}]$$

where, N_s denotes the total number of snapshots. At the n^{th} time step, velocity can be in the following form:

Eq. 4-2

$$\mathbf{u}^n = \{u_1^n, u_2^n, \dots, u_N^n, v_1^n, v_2^n, \dots, v_N^n, w_1^n, w_2^n, \dots, w_N^n\}^T$$

where, N represents the number of nodes. Furthermore, u_i^n , v_i^n , and w_i^n show the horizontal, lateral, and vertical velocity components at node i and time n . In order to find out the POD basis functions, singular value decomposition (SVD) can be applied to the snapshots' matrix:

Eq. 4-3

$$S = U\Sigma V^T$$

where, $U \in \mathbb{R}^{3N \times 3N}$ states the spatial correlations in data, and its columns are orthogonal. $V \in \mathbb{R}^{N_s \times N_s}$ gives the corresponding time dynamics for each of these vectors in U . Also, $\Sigma \in \mathbb{R}^{3N \times N_s}$, which is a diagonal matrix, provides an orthogonal set of vectors in which data is embedded, and indicates the significance of each direction. The first N_s columns of U are considered as the POD basis functions. The number of basis functions can be reduced to r by considering a specific tolerance, e.g., $\eta \leq 1$, [221] (i.e., low-rank truncation of data):

Eq. 4-4

$$\frac{\sum_{j=1}^r \sigma_j^2}{\sum_{j=1}^{N_s} \sigma_j^2} \geq \eta$$

Eq. 4-4 shows the ratio between the amount of energy captured by the first r modes to the total energy. The first r modes, then, can be stacked in a matrix, Φ_r , showing the low-dimensional subspace on which, the data is embedded. Afterward, a variable, \mathbf{u} (here, velocity or pollutant concentration), can be expressed using the following equation:

Eq. 4-5

$$\mathbf{u} = \bar{\mathbf{u}} + \sum_{j=1}^r a_j \Phi_j$$

where, a is a time-dependent coefficient of the POD expansion. In other words, a shows the time dynamics of this subspace. Furthermore, $\bar{\mathbf{u}}$ shows the mean value of \mathbf{u} . The next step is to find a

by, for example, training a time-stepping neural network to reconstruct the original data using the reduced space.

In [254,255], a domain decomposition method has been applied to finite element nodes to divide the domain into a number of subdomains in order to compute a set of POD modes for each subdomain. Thus, a set of hypersurfaces can be constructed to explain the dynamics of the system. Dividing the computational domain into multiple subdomains can be performed using a weighting constraint on vertices (e.g., uniform weight, the maximum value of the nodal Reynolds stresses, etc.) so as to reach an equal accuracy for each of them while minimizing the dependence of each subdomain regarding the system's dynamic [254]. In order to achieve this goal, different numbers of POD modes may be used in each subdomain to represent the low-dimensional system of a specific subdomain. This method can improve the model accuracy, particularly in large domains with variations in the geometric specifications.

Quilodrán-Casas et al. [258] combined PCA with adversarial autoencoders to develop a ROM. In order to reduce the problem's dimensionality, they used PCA to derive the dominant modes of a 3D velocity field obtained by CFD. Afterward, they applied an adversarial autoencoder on the principal components obtained by PCA to obtain the latent space, on which computations can be done. This method of dimensionality reduction showed better performance than the traditional truncation of principal components.

It is worth noting that since POD is, in nature, an eigenvalue problem, it is physically interpretable. For instance, as mentioned before, each POD mode contributes to the total TKE [209]. Although this technique has extensively been used in fluid dynamics problems (e.g., see Table 4.2 and [227,253]), it has a major drawback, i.e., its linear nature. In other words, it computes an optimal linear subspace on which data can be projected. However, real-world data, such as those related to fluid dynamics, are strongly nonlinear, particularly in advection-dominated problems. Previous studies have compared the performance of POD with some nonlinear methods in generating ROMs for different applications [225,262,263]. These studies support the fact that the POD-based ROMs are not as efficient and accurate as those generated by nonlinear dimensionality reduction methods. Eivazi et al. [264] reported that in order to capture 99% of the energy from the flow reconstruction, 247 POD modes were required for a simplified urban configuration (i.e., two cubes in a row). Thus, a linear approximation may result in insignificant speed-ups for heavily nonlinear and advection-dominated problems [265]. Furthermore, the performance of POD-based ROMs may be deteriorated under dynamic boundary conditions compared to cases with steady ones [225]. As previously explained, boundary conditions in urban areas constantly change. Therefore, using linear dimensionality reduction techniques, such as POD, might not be the most suitable choice for urban-related applications.

4.3.2 The relationship between POD and autoencoder

Under-complete autoencoder can be considered as the generalized form of POD [266]. As mentioned before, POD (or PCA) finds a linear low-dimensional hyperplane to describe the original data. On the other hand, in autoencoders, nonlinearities can be introduced to the network

using nonlinear activation functions. Thus, autoencoders can learn complex data representation, which can be employed to describe the original data in a lower dimensionality [211]. Figure 4.1, obtained from [267], shows the difference between linear and nonlinear dimensionality reduction methods.

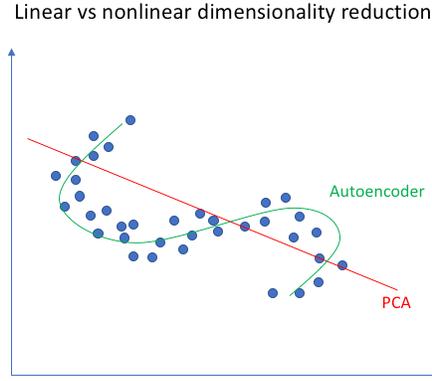


Figure 4.1: Difference between linear and nonlinear dimensionality reduction methods (obtained from [267])

An autoencoder consists of two parts, namely encoder, and decoder, which are designed to act as the inverse of each other. The encoder takes in a high-fidelity state (i.e., the original data) and encodes it into a latent space (i.e., low-dimensional representation of the original data), in which computations can be done [268]:

$$\mathbf{h} = f_E(\mathbf{x}; \theta_E)$$

Eq. 4-6

where, $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{h} \in \mathbb{R}^{N_h}$ represent the input state and the representation vector (i.e., latent space), respectively. N and N_h denote the dimensionality of the original space and latent space, consecutively. It should be noted that $N_h < N$. Also, θ_E represents the parameters of the encoder part. Afterward, a decoder network is employed to take back the feature state to reconstruct $\hat{\mathbf{x}}$, which is an approximation of \mathbf{x} [268]:

$$\hat{\mathbf{x}} = f_D(\mathbf{h}; \theta_D)$$

Eq. 4-7

where, θ_D denotes the parameters of the decoder part. In order to identify the latent space, snapshots obtained from the high-fidelity solution are used. The procedure consists of finding the parameters so as to minimize the reconstruction error over the training examples [268]:

$$\theta_E^*, \theta_D^* = \arg \min_{\theta_E, \theta_D} \mathbb{E}_{\mathbf{x} \sim P_{data}} [\mathcal{L}(\hat{\mathbf{x}}, \mathbf{x})]$$

Eq. 4-8

where, $\mathcal{L}(\hat{\mathbf{x}}, \mathbf{x}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$ is a measure for calculating the deviation of the reconstructed $\hat{\mathbf{x}}$ from \mathbf{x} [211]. Also, \mathbb{E} in statistics is used to denote the “expected value”. In order to prevent the autoencoder network from learning the identity function, $N_h < N$ can be applied. In this way, hidden features of P_{data} (i.e., the data generating distribution) can be obtained [268]. It is noteworthy to mention that f_E , f_D , and $\mathcal{L}(\hat{\mathbf{x}}, \mathbf{x})$ are usually selected based on the application.

Single-layer encoder and decoder, with linear activation function, result in a solution similar to that obtained by POD [211,269]. Thus, Eqs. 4-6 and 4-7 can be reformed as follows [268]:

$$\mathbf{h} = \mathbf{W}_E \mathbf{x}$$

Eq. 4-9

$$\hat{\mathbf{x}} = \mathbf{W}_D \mathbf{h}$$

Eq. 4-10

where, $\mathbf{W}_E \in \mathbb{R}^{N_h \times N}$ and $\mathbf{W}_D \in \mathbb{R}^{N \times N_h}$. The autoencoder can learn the same subspace as the one captured by the first N_h POD modes if $\mathbf{W}_E = \mathbf{W}_D = \mathbf{W}$ and $\mathbf{W}^T \mathbf{W} = \mathbf{I}_{N_h \times N_h}$ and with the following equation for $\mathcal{L}(\hat{\mathbf{x}}, \mathbf{x})$ [268]:

$$\mathcal{L}(\hat{\mathbf{x}}, \mathbf{x}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 = \|\mathbf{x} - \mathbf{W} \mathbf{W}^T \mathbf{x}\|_2^2$$

Eq. 4-11

4.3.3 Nonlinear dimensionality reduction

In nonlinear dimensionality reduction, the assumption made for the latent space as a linear solution manifold should be removed. Therefore, it is not possible to write the projection operator as $\mathbf{W} \mathbf{W}^T$ anymore, and its general form should be used since any nonlinear function may be employed [265]. It is noteworthy to mention that one of the most common nonlinear dimensionality reduction methods is the kernel PCA. However, using this approach, in some cases, may be problematic since there exist no general guidelines for choosing the nonlinear mapping or the kernel [265]. For more information regarding the kernel PCA method, one can refer to [270].

As can be seen in Table 4.2, the use of nonlinear dimensionality reduction techniques becomes more common in urban-related problems. Xiang et al. [225] compared the performance of different autoencoder-based dimensionality reduction methods, namely, AE-linear-FCNN, AE-FCNN, and CAE, with the POD technique. They applied these methods on 2D velocity fields on the pedestrian level. For AE-FCNN, they used SELU activation function [271] to consider the impact of nonlinear activation function on dimensionality reduction. The results show that nonlinear approaches performed better in finding a low-dimensional representation for the original data than POD. A comparison between the structures of the AE-linear-FCNN and AE-FCNN can be observed in Figure 4.2, obtained from [225], in which the blue neuron illustrates the nonlinear layer.

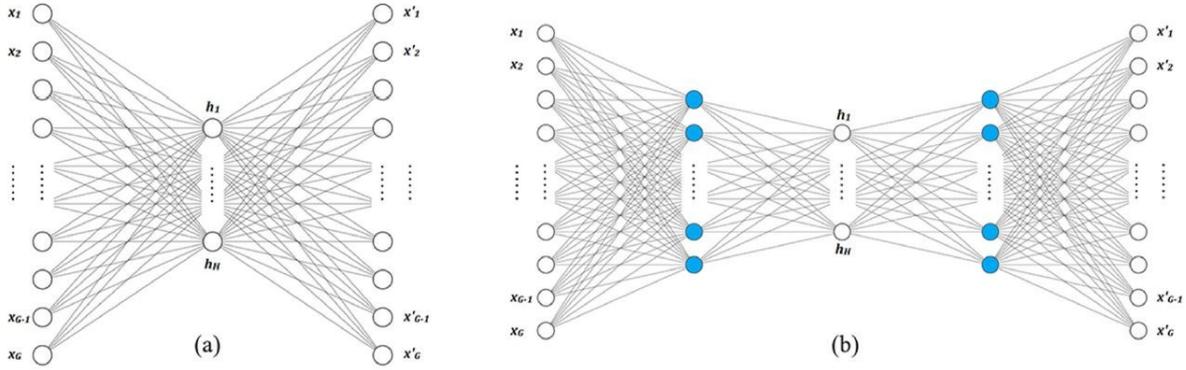


Figure 4.2: Difference between linear and nonlinear dimensionality reduction methods (obtained from [225])

Although deep fully-connected autoencoders performed well in previous studies [225,272], the approach alone may not be scalable as the dof of the system increases since the number of the trainable parameters increases significantly [268]. Also, an increase in the depth of the network can even exacerbate the situation (see Figure 4.2). Besides, using fully-connected autoencoders for very high-dimensional data may result in over-fitting. Using convolutional layers can solve the problem due to their features, namely, local connection and shared weight across the input domain [266,273], making them an appropriate feature detection choice. Figure 4.3 shows a schematic view of how the input data (i.e., the blue matrix) can be mapped to the output (i.e., the green matrix) by sliding the 3×3 kernel through the input data [274].

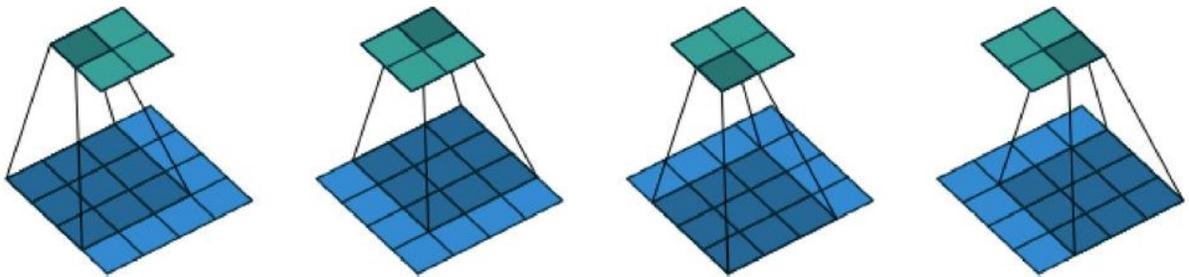


Figure 4.3: An example of the convolutional operation with a 3×3 convolutional kernel, with no padding and strides (obtained from [274])

Downsampling in convolutional layers is usually done by employing pooling layers (i.e., maximum or average pooling layers) or strides larger than one, which is related to the filtering process [266]. It is noteworthy to mention that if pattern recognition is important in a specific application and data is obtained from a grided topology, CNNs can be helpful [46,266]. Using CAE can facilitate the dimensionality reduction procedure. Instead of applying a fully-connected autoencoder to the original data (high-dimensional), in CAE, the network can be applied to a vectorized feature map with lower dimensionality obtained from a deep CNN [268]. In this way,

nonlinear autoencoders can be applied to a large dataset in order to extract hidden features of the system. Figure 4.4, showing a schematic view of a CAE, illustrates that in the encoding part, a series of convolutional layers with an increasing number of filters leads to a decrease in the dimension [265]. Therefore, the dof of the system decreases. After obtaining the hidden layer (from CNN-encoder), a similar approach, but in the reverse form, is designed to reconstruct the input from the hidden layer. At this point, the loss function should be minimized for training the model.

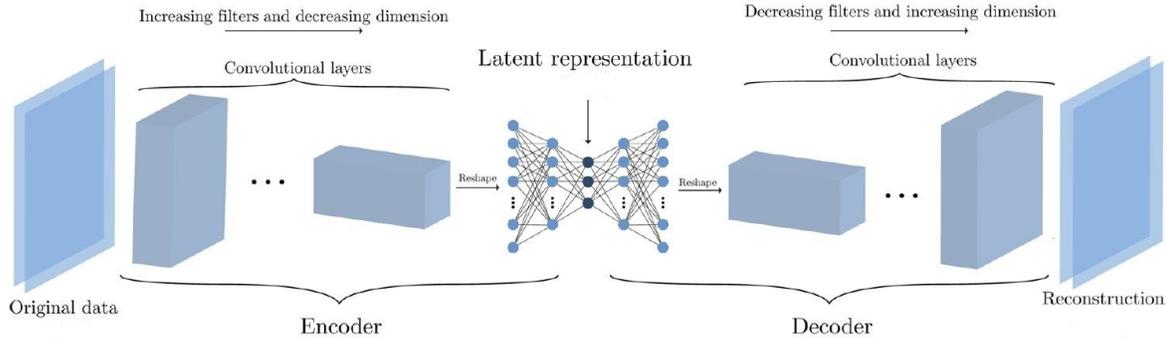


Figure 4.4: The schematic view of the structure of a convolutional autoencoder (modified from [265])

In order to use CAE for dimensionality reduction, Xiang et al. [225] employed five convolutional layers with 3×3 kernels and ReLU activation function [275], using maximum pooling layers after each of the first four ones. Then the reduced, vector-shaped data were fed into fully-connected layers. The predictive performance of NIROMs, based on different dimensionality reduction techniques, was sorted as $CAE > AE-FCNN > AE-linear-FCNN > POD$. However, they did not compare these methods' training and reconstruction times. Furthermore, Mücke et al. [265] compared the performance of CAE and POD in an advection-dominated problem. They reported that considering a latent space with the dimension of 2 for CAE led to the flow reconstruction with a mean squared error in the range of $10^{-4} - 10^{-3}$. However, in order to achieve a similar accuracy using the POD technique, at least 17 POD modes were required. This finding is considered as a supporting fact on the claim of the inefficacy of the POD technique for being used in advection-dominated problems.

In 3D convolutional layers, instead of just sliding in a 2D space, the kernel also goes through the depth, which is helpful to extract dependencies in the third dimension. This method has been used in several applications, particularly in studies with video data as the input (e.g., action recognition and video segmentation [276,277]) since it can exploit spatio-temporal dependencies in the system. Since the nature of fluid dynamics problems is also spatio-temporal, 3D convolutional layers can logically be employed. The method has successfully been applied to a synthetic jet problem in a transitional regime for prediction tasks [278].

Based on the above review, it can be concluded that the most popular approach of nonlinear dimensionality reduction in fluid dynamics problems, such as urban-related ones, is CAE due to its capability of handling systems with high dof. Using adversarial autoencoders may also increase the ability of the method to find a better low-dimensional representation of the original data, as is

used in [258] for the POD technique. However, some limitations exist regarding the physical interpretation of nonlinear methods, which is of utmost importance in developing NIROMs for engineering problems. These issues and recommendations to overcome them are discussed in section 4.5.

4.4 Feature dynamics

The next step of generating a ROM is to compute the temporal evolution of the low-dimensional features obtained from the previous step. Based on Table 4.2, several approaches, including neural networks and regression techniques, have been used for the mentioned task. The RBF interpolation method, which can be used for the interpolation of multi-dimensional scattered data [279] (using a weighted combination of RBFs), has successfully been applied to fluid dynamics problems [255,280]. The approach is suitable for handling a large amount of data. The method has been used along with the POD technique to generate NIROMs for different fluid dynamics applications such as a simplified street canyon, flow past a cylinder, and a lock exchange problem, using unstructured finite element grids [255,280]. The results obtained from NIROM showed a generally good agreement with those from the high-fidelity solver.

GPR [281] has also been employed to study the prediction of urban airflow and pollutant dispersion [221,224,254]. This linear method is trained to find a combination of Gaussian-shaped basis functions to provide hypersurfaces for the NIROM development. The method is suitable for small datasets since it does not require intensive training procedures due to its linear nature. However, its performance may deteriorate with large datasets [221,282]. It should be noted that some efforts have been made to use GPR with large datasets [282,283].

Time series data can efficiently be handled using recurrent neural networks (RNNs), such as LSTM [284], gated recurrent unit (GRU) [285], and echo-state networks [286]. LSTM, which is a modified version of RNN, can solve the issue of vanishing or exploding gradients in the backpropagation procedure in classic RNNs. The LSTM layer consists of four parts: cell state, input gate, output gate, and forget gate. Using these gates, LSTM has the ability to control the flow of information. The gates can learn which information is important and which is not. More details on the functionality of each gate can be found in [284]. In this method, values from the reduced space are fed into the LSTM model as the initial encoded field [287]. Afterward, the predicted field is recursively used as the model's cell state input. After feeding the input data, the LSTM model is trained to minimize a loss function in the latent space. LSTM has successfully been employed in modeling the temporal evolution of low-dimensional representations obtained by POD and autoencoders [227,265,268]. For instance, the time-averaged error for CAE-LSTM in predicting the flow past a cylinder, for Reynolds number ranging between 120 and 200, was reported to be 3.92×10^{-2} [265]. The mentioned value for the POD-LSTM method was stated to be of the order of magnitude of 10. Furthermore, It has been reported that LSTM has performed reasonably well in generating a NIROM for a large urban area with varying parameters, i.e., pollution emission intensity [224]. In [224], the root-mean-square error of pollutant concentration results, in a realistic urban area, between the developed parametrized NIROM and high-fidelity

model was mainly of the order of magnitude of 1. As Mücke et al. [265] reported GRU performed similarly to LSTM in problems such as linear advection and flow past a cylinder.

Adversarial training of LSTM has also been employed in urban-related studies to improve the performance of the conventional LSTM [257,258]. In these studies, a discriminator was used to differentiate between the original principal components and those predicted by LSTM. Afterward, the discriminator training has been performed by minimizing the binary cross-entropy loss. The comparison between the performance of the conventional LSTM and the adversarial one exhibits that after 80 time steps of forecasting, the former diverged quickly from the underlying state. On the other hand, the latter could preserve the underlying physics during the prediction phase [258]. Furthermore, the adversarial LSTM was able to create eddies better than the conventional one. It should be noted that using the adversarial LSTM resulted in the reduction of, in some locations, one order of magnitude in the absolute error with respect to the original data, compared to the conventional LSTM [258].

Echo-state networks, a form of reservoir computing, have been used to learn the dynamics of chaotic systems, including those related to fluid dynamics, such as the Kuramoto-Sivashinsky equation, Charney-DeVore system, and turbulent moist Rayleigh-Bénard convection flow [288–290]. In echo-state networks, the weights of the input and reservoir are initialized and kept constant during the training phase, while those related to the output layer are trained. Therefore, the issues related to the backpropagation procedure are eliminated. Furthermore, training these networks is usually faster than LSTM. Although echo-state networks have shown promising performance in learning the feature dynamics of nonlinear systems [288–290], they have not been comprehensively compared to other RNNs, such as LSTM, in predicting turbulent flows to investigate their functionality in complex problems. Moreover, echo-state networks are known to face stability issues, particularly when using noisy datasets. [291]. Also, their performance in long-term prediction attenuates. These limitations may be due to the dependence of such networks on hyper-parameters. It is known that echo-state networks are able to perform well in a narrow region of hyper-parameter space [292]. Although there exist several research items on overcoming the issues arising from the usage of these networks [292–294], their functionality in learning the feature dynamics of complex turbulent flows (e.g., urban airflows) is in doubt.

Xiang et al. [225] used XGBoost [295], a scalable decision tree boosted system, as the regression model on the reduced space obtained by different dimensionality reduction methods (see Table 4.2). The “boosting” term refers to a set of algorithms that can convert weak learners to strong ones [296]. The weak model was fitted to the whole space of the input data, while the second model was fitted to residuals of the weak model [225]. This process eventually stops by reaching a certain criterion. The predictive model is the sum of the prediction of each model. It is noteworthy to mention that Xiang et al. [225] used the model for developing a NIROM with transient boundary conditions. In other words, the XGBoost algorithm was used to map the dynamics of boundary conditions to each low-dimensional space obtained by a particular dimensionality reduction technique. The results show that their proposed model was able to capture the dominant features of the urban airflow under both low and high wind speed conditions.

Causal CNN (CCNN), also known as temporal CNN [297], combined with feed-forward neural networks, has been applied to fluid dynamics problems, such as linear advection, discontinuous compressible flow, and flow past a cylinder, to capture the time evolution of the latent space [265,298]. Using CCNN, the information leakage from the future time steps into the previous ones is prevented since it operates only on the data prior to the current time step in the previous layer. Comparing CCNN with LSTM for sequence modeling across a diverse range of tasks and datasets exhibited that the former has a longer effective memory length than the latter [299]. However, a comparison between the results of CAE-LSTM and CAE-CCNN in reconstructing the flow field past a cylinder shows that the relative error of the latter approach is about one order of magnitude larger than that for the former one beyond the training horizon [265]. On the other hand, CAE-CCNN slightly performed better than CAE-LSTM for the linear advection problem.

Self-attention deep learning [300] has recently been applied to fluid dynamics problems, such as flow past a cylinder, with $Re = 3200$, to deal with long-term prediction issues arising from the use of autoencoder networks [263]. The self-attention mechanism is the one in which the inputs interact with each other to realize which one requires more attention. Furthermore, the input data sequence is not fed to the model in its sequential order in the self-attention mechanism. After feeding the low-dimensional space to the self-attention network, the network predicted the space for the next time step [263]. The results showed that using the self-attention method reduced the root-mean-square error to approximately 1/3 of that for the conventional POD-based ROM [263].

Recent advances in calculating feature dynamics for the NIROM development in the fluid dynamics applications are discussed so far in section 4.4. One of the challenges in generating NIROMs for urban-related problems is to parameterize it, since, for instance, the weather conditions in an urban area constantly change. There exist some studies on the development of NIROMs with varying boundary conditions in different applications, such as flow past a cylinder [265,287] and urban airflow and concentration fields prediction [224,225]. It should be noted that reviewing the advances and challenges of the development of such parametric models is not among the topics of the present dissertation. However, noises and small changes in the input data may result in significant changes in the models' outputs since these data-driven models are known to be domain-agnostic [301]. Thus, some modifications may be required for generating a robust and stable model, which are addressed in section 4.5.

4.5 Challenges and recent advances

As discussed in the previous sections, autoencoder-based methods are the most common nonlinear dimensionality reduction techniques in fluid dynamics problems (see Table 4.2 and [263,265,268,287,302]). However, nonlinear low-dimensional mapping strategies are not physically interpretable. In other words, contrary to the modes obtained from POD, which can be related to the physical structures present in the system, it is difficult to interpret the physical meaning of the latent vectors exploited by autoencoders with nonlinear activation functions [303]. This is due to the fact that no concept like eigenvalues can be interpreted from these nonlinear modes since they are not orthogonal. Furthermore, unlike POD, the latent modes obtained from

conventional autoencoder-based strategies are not sorted in the order of their contribution to the total TKE [303].

In order to overcome the aforementioned issues, Fukami et al. [303] used CNN combined with hierarchical autoencoder [304] to rank the latent modes based on their energy. The method is based on dividing a network into different subnetworks and then feeding the latent vector of each subnetwork to the next one. In this way, subnetworks already consider the features obtained by previous subnetworks in their latent space. Thus, they will try to find the remaining modes that have not been observed in the previous subnetworks. The method was, at first, applied to laminar and transient wake regions behind a 2D cylinder and then to a turbulent cross-sectional flow in a channel with $Re_\tau = 180$, where Re_τ represents the Reynolds number based on the friction velocity and channel half height. Although the method solved the problem of ranked modes regarding their contribution to reconstructing the flow field, the uniqueness problem remained an issue, which can be attributed to the probabilistic process of neural networks during the training phase. Also, since multiple neural networks should be trained, the method may be cumbersome, particularly for exploiting higher-order modes.

In order to solve the interpretability issue of nonlinear mode extraction and their rank regarding the modes' energy content, Eivazi et al. [264] used a probabilistic deep neural network algorithm based on the combination of CNN and a modified version of variational autoencoder [305,306] to exploit near-orthogonal nonlinear modes for turbulent flows by applying the method on a simplified urban configuration. For doing so, the correlation between the latent variables was minimized, while the size of the latent vector was penalized. Furthermore, in order to rank the obtained modes based on their energy content, they used an expression to find out the relative TKE percentage captured by the reconstructed model with respect to the original values of fluctuations. Although the proposed model outperformed the POD technique, it slightly decreased the reconstruction accuracy compared to the conventional CAE and hierarchical CAE.

Since turbulence is composed of eddies with different scales, several studies have used multiscale CNN for the dimensionality reduction process to accommodate the multiscale nature of a complex turbulent flow [307–310]. Thus, different kernel sizes can be considered for the mentioned purpose. Nakamura et al. [310] used 3D CNN with three different kernel sizes, $3 \times 3 \times 3$, $5 \times 5 \times 5$, and $7 \times 7 \times 7$, in their model and obtained the weights of their multiscale CAE using an iterative optimization. The method was applied to a turbulent channel flow with $Re_\tau = 110$ in order to develop a NIROM, which resulted in a good statistical agreement with the DNS data. It should be noted that since more complex flows, such as those with higher Reynolds number, may require a larger number of latent modes than the case conducted in [310], the combination of this method with the hierarchical one introduced in [303] may solve the issue to a great extent. The schematic view of the multiscale CAE model used in [310] is shown in Figure 4.5, obtained from [310].

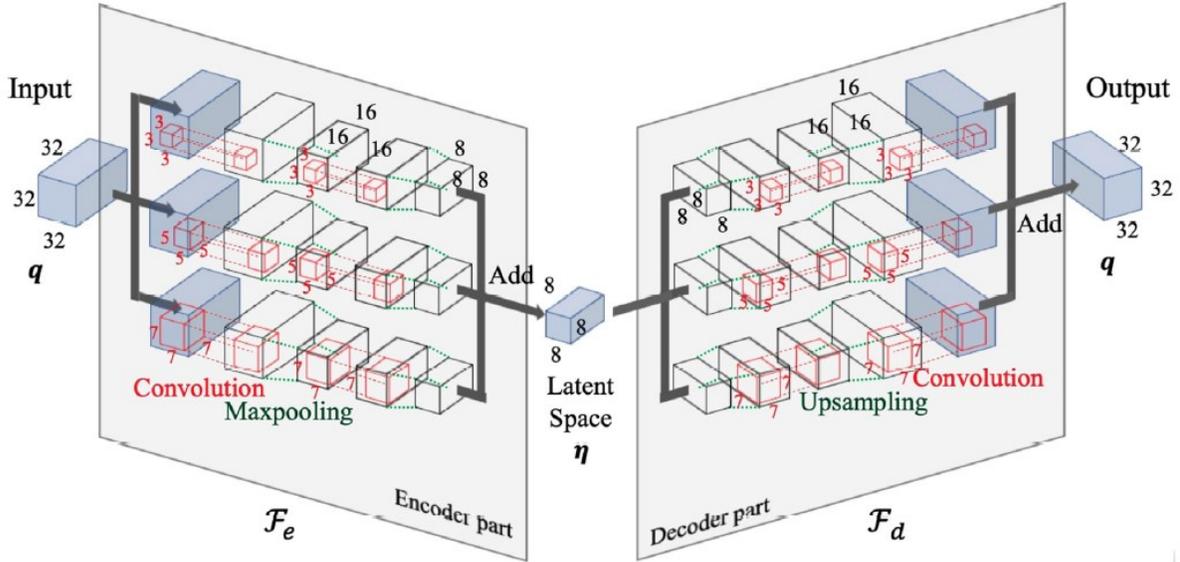


Figure 4.5: The schematic view of the multiscale CAE used in [310], using three different kernel sizes of $3 \times 3 \times 3$, $5 \times 5 \times 5$, and $7 \times 7 \times 7$ for dimensionality reduction of DNS data (modified from [310])

As mentioned previously, since neural network models are black-box models, they tend to neglect the physics and characteristics related to the domain of the problem [301], which may lead to erroneous results in the presence of noisy data, small training data, or many hyperparameters. In order to overcome this issue to some extent, it is beneficial to embed information related to the physics and domain of the problem into the model. In other words, customizing the loss function to achieve physics-informed models, such as those pioneered by Raissi et al. [311,312], can improve the model regarding its stability and robustness.

By applying automatic differentiation on the output variables, Raissi et al. [313] embedded the coupled dynamics of the passive scalar and the Navier-Stokes equations into the outputs to achieve a physics-informed learning algorithm, which is agnostic to the geometry of the problem and its initial and boundary conditions. They customized the loss function during the training stage by adding a term related to the residuals of the momentum equations, the transport equation of a passive scalar, and the continuity equation. The proposed model has been reported to be robust to remarkable noises in the input data. Cheng and Zhang [314] also applied automatic differentiation with respect to time and space to the output of a residual neural network (ResNet) in order to develop a physics-informed network for applications governed by the Navier-Stokes equations. They reported that including ResNet can lead to the improvement of the model performance.

Erichson et al. [301] used the empirical risk minimization to enhance the generalization of a data-driven model for fluid flow prediction by proposing a method for autoencoders to preserve the Lyapunov stability. Using Jacobian regularization in the loss function of a simple feed-forward neural network, Pan and Duraisamy [315] suppressed the sensitivity of the predictive model to the local error. In this way, they improved the accuracy and robustness of the model in long-term

predictions using limited available data. These modifications seem to be of utmost importance in developing a NIROM for urban-related applications (e.g., predicting the turbulent airflow field and pollutant dispersion), particularly when multi-query and long-term predictions are required.

4.6 Conclusion

Due to their high computational costs, CFD solvers cannot usually be used for multi-query problems and (near) real-time predictions of the flow and concentration fields. Using ROMs as an alternative to high-fidelity models can overcome these issues. In the present chapter, two different types of ROMs, namely IROM and NIROM, were introduced, and also their capabilities and limitations were mentioned. It was explained that using NIROMs can solve the issues arising from the use of IROMs, regarding source code manipulation and nonlinear inefficiency. Thanks to the progress of machine learning algorithms, the use of deep learning in generating NIROMs becomes popular. Developing a NIROM comprises two stages: dimensionality reduction and calculating feature dynamics. Different methodologies, with their advantages and disadvantages, were introduced for each of these two stages.

It was explained that using linear dimensionality reduction techniques, such as POD, results in inefficient models, particularly when dealing with advection-dominated problems. Furthermore, using a nonlinear autoencoder as a standalone approach is not suitable for complex systems such as fluid dynamics problems. Therefore, the combination of CNN-autoencoder can be used for reducing the dimensionality of such complex systems. Moreover, using more advanced methods, such as adversarial autoencoder, can increase the model's accuracy. Afterward, several approaches for computing the feature dynamics of the derived reduced space were introduced. Among these methods, LSTM has successfully been used in fluid dynamics problems. It was stated that using adversarial LSTM instead of conventional one can result in a longer prediction horizon over the test dataset without a significant decrease in the model accuracy. CCNN was introduced as another approach for calculating the feature dynamics of the low-dimensional space due to its long effective memory length. However, more investigations are required to find out its effectiveness in fluid dynamics problems.

Due to the black-box nature of machine learning algorithms, there still exist several challenges in developing these models, such as the physical interpretation of the models, their stability, and robustness under dynamic boundary conditions. In the present chapter, different approaches were introduced to make the dimensionality reduction procedure physically interpretable by sorting them based on their energy content and making them near-orthogonal. Furthermore, several efforts were reviewed to embed physics into the model by customizing loss functions, which is of utmost importance in developing NIROMs for the long-term prediction of urban airflow and pollutant dispersion under dynamic boundary conditions. The present chapter is expected to shed light on the current stage of the development of NIROMs for urban-related applications and pave the way for future works by mentioning the existing challenges.

Chapter 5. Performance analysis of different reduced-order models for predicting urban turbulent flow field

5.1 Introduction

Computational fluid dynamics (CFD) has been considered as a promising numerical approach in fluid dynamics problems, such as urban airflow prediction. However, flow field prediction using CFD models is time-consuming. Thus, they cannot be used for (near) real-time and long-term simulations. Reduced-order models (ROMs) are emerged to obviate this limitation. Deep learning (DL) algorithms have been used for developing non-intrusive ROMs (NIROMs) in fluid dynamics applications. In the present study, three different approaches, namely, convolutional autoencoder (CAE), multiscale CAE (MS-CAE), and self-attention CAE (SA-CAE) are developed for dimensionality reduction, which is considered as the first step of the development of a NIROM. The developed models are then used for finding a low-dimensional representation of the original data. Afterward, a parallel long short-term memory (LSTM) network is employed for computing the temporal dynamics of the obtained low-dimensional space. The models are trained to reconstruct a turbulent airflow field in the wake region of an isolated high-rise building, located in an unstable thermal stratification condition, using validated CFD data.

The methods used in this chapter are chosen to overcome the limitations of conventional CAEs arising from using a fixed kernel size in convolutional layers. In other words, efforts are made to capture multiscale and long-range dependencies among datapoints. Multiscale CAE (MS-CAE) and self-attention CAE (SA-CAE) are, for the first time, being used for reconstructing highly turbulent flow fields. In addition to qualitative and quantitative comparison of the results of developed models with the CFD data using conventional approaches, the spectral analysis is performed on the constructed dataset to analyze the large-scale unsteadiness of the reconstructed flow field.

5.2 Methods used in NIROM development

The current section explains the deep learning (DL) methods used in the present study for developing NIROMs. These methods are CAE, MS-CAE, SA-CAE, and LSTM.

5.2.1 Convolutional autoencoder (CAE)

Convolutional layers, pooling layers, and fully-connected layers can constitute a CAE. Using the encoder part of a trained CAE, a high-dimensional system (here, an airflow field) is mapped onto

a low-dimensional latent space. Afterward, the decoder part is designed to reconstruct the encoded datapoints to the original space.

Considering a 2D flow field dataset (i.e., longitudinal and vertical velocity components), 2D convolutional layers can be used in CAE. Thus, a 2D kernel with a specified size can be used to capture local correlations between the data points by sliding through the input data. An input dataset, $X \in \mathbb{R}^{H' \times W'}$, is convolved using a 2D convolutional kernel, $K \in \mathbb{R}^{M \times N}$, to obtain a feature map, $Y \in \mathbb{R}^{H'' \times W''}$. It should be noted that H' (H'') and W' (W'') are, respectively, the height and width of the input data (output feature map). For any datapoint $x \in X$, the datapoint $y \in Y$ can be calculated as follows [316]:

$$y_{i,j} = \sum_{m=1}^M \sum_{n=1}^N k_{m,n} \cdot x_{i-m+1,j-n+1}$$

Eq. 5-1

where, $1 \leq i \leq H''$ and $1 \leq j \leq W''$. In the present study, the padding method is selected as “same”; thus, $H'' = H'$ and $W'' = W'$.

Downsampling can be performed using pooling layers, usually maximum or average pooling operations [266]. A 2D pooling operation, maximum or average, reports the values inside a rectangle with its maximum or average value, respectively. Dimensionality reduction can also be performed using strides larger than one. In this case, the stride value defines the movement of the kernel of the convolutional layer. Figure 5.1 schematically illustrates how a 3×3 kernel acts on a 7×7 input. Strides larger than one was used in the present study since it can keep spatial information more effectively than pooling layers [316].

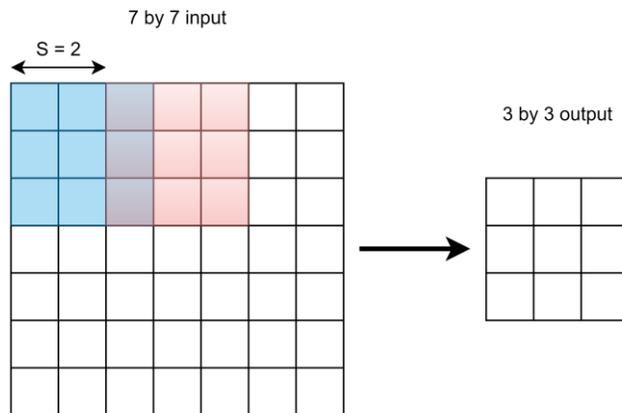


Figure 5.1: The schematic view of a 3×3 kernel, with a stride value of 2, acting on a 7×7 input data

The structure of the encoder part of the CAE used in the present study is depicted in Figure 5.2. It should be noted that the decoder part tries to reconstruct the encoded data back to the original space using combined upsampling and convolutional layers.

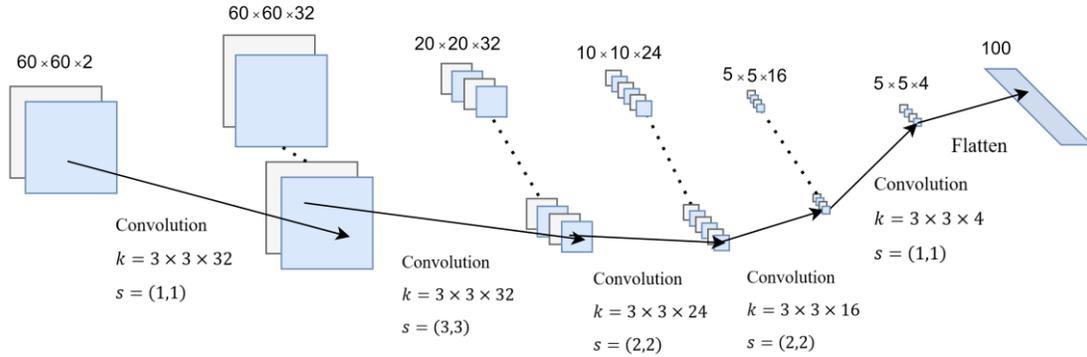


Figure 5.2: The encoder part structure of the CAE. $k = 3 \times 3 \times 32$ means that the size of the convolution kernel is 3×3 , and **32** convolution kernels are used in this layer. s indicates the stride of the convolutional layer.

5.2.2 Multiscale CAE (MS-CAE)

To capture the multiscale sense of images, Du et al. [317] proposed a multiscale convolutional neural network (CNN) structure. In order to account for the multiscale nature of turbulent flow fields, dimensionality reduction can be performed using MS-CAE. MS-CAE has previously been applied to fluid dynamics applications, such as turbulent channel flow [309], laminar and transitional flow (with the Reynolds number ranging between 20 and 160) past a cylinder [287], and unsteady flows (with the Reynolds number of 100) around bluff bodies of various shapes [307]. However, as previously mentioned, MS-CAE is being used for the first time in this study for highly turbulent urban airflow.

In the present study, after experimenting with various kernel sizes, 3×3 , 5×5 , and 7×7 kernels are selected. Employing various kernel sizes helps the model capture multiscale spatial correlations among data, which can be beneficial in more effective dimensionality reduction compared to the conventional CAE. The structure of the encoder part of the MS-CAE used in the present study is shown in Figure 5.3.

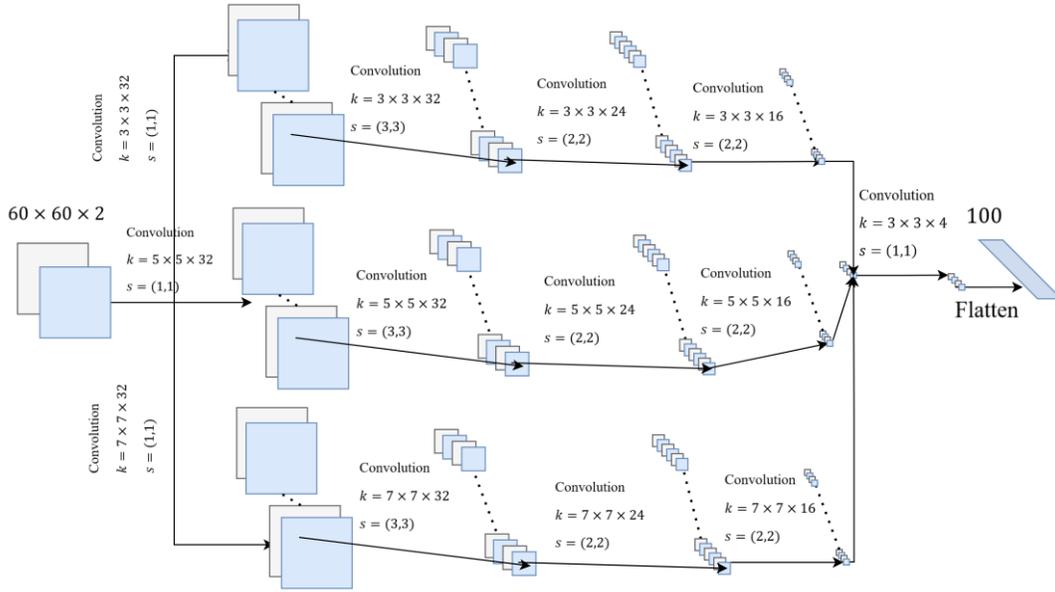


Figure 5.3: The structure of the encoder part of MS-CAE. Three different encoders with the same structure as the one used in CAE, but with different kernel sizes, i.e., 3×3 , 5×5 , and 7×7 .

5.2.3 Self-attention CAE (SA-CAE)

As explained previously, kernels in convolutional layers are responsible for capturing local correlations among data. Therefore, the kernel size might be a limiting factor in finding non-local correlations. In sub-section 5.2.2, MS-CAE was introduced, which enables the model to consider the multiscale nature of the flow field. The goal of developing SA-CAE can be considered similar to MS-CAE. In both models, the aim is to overcome the limitations caused by the fixed size of convolutional kernels. In the present study, the methodology proposed by Wang et al. [318], called non-local neural networks, was used to capture long-range dependency among data.

In the self-attention module employed in the present study (see Figure 5.4), the output is the sum of the input data and the attention module output multiplies by a factor controlling the attention intensity [316]:

$$Y = X + X'$$

Eq. 5-2

where, $X \in \mathbb{R}^{H' \times W' \times C}$ is the input to the self-attention module. C denotes the number of channels. $X' = \alpha Y'$, where α is the aforementioned trainable controlling factor and $Y' \in \mathbb{R}^{H' \times W' \times C}$ is the output feature map of attention part. Also, $Y \in \mathbb{R}^{H' \times W' \times C}$ is the output feature map. In order to compute the attention weight, the similarity of any $x_j \in X$ is compared to a specific $x_i \in X$, using the following equation [316,318]:

Eq. 5-3

$$\Phi(x_i, x_j) = \exp(\phi(x_i)^T \theta(x_j))$$

where, T refers to “transposed”, $1 \leq i, j \leq H'W'$. In order to compress the size of channels, 1×1 convolutions are used in each path, shown in Figure 5.4, i.e., ϕ , θ , and g paths. The compression value to decrease the number of channels plays an important role in reducing computational costs. Then, the attention weight of a specific x_i to any x_j in the input domain can be calculated as the ratio of $\Phi(x_i, x_j)$ to the sum of $\Phi(x_i, x_k)$, where $k \in [1, H'W']$:

Eq. 5-4

$$\psi(i, j) = \text{softmax}(\phi(x_i)^T \cdot \theta(x_j))$$

where, *softmax* represents the normalized exponential function. It should be noted that another compression ratio (*c.r.*) is used to reduce the dimensionality in the ϕ and g paths, as can be observed in Figure 5.4. Afterward, the output of *softmax* is multiplied by the output of the g path to obtain a feature map with the dimensionality of $H'W' \times C'$. Then, the output is reshaped to $H' \times W' \times C'$ and it goes through a convolutional layer to have a dimensionality similar to that of the input feature map. Finally, the output of the module is added to the input feature map to obtain Y . It should be noted that in the present study, the compression ratio for reducing the number of channels (i.e., to obtain the intermediate channel) and *c.r.* are both selected to be 2. The attention module is used with the second, third, and fourth convolutional blocks of the CAE shown in Figure 5.2, before reducing the dimensionality.

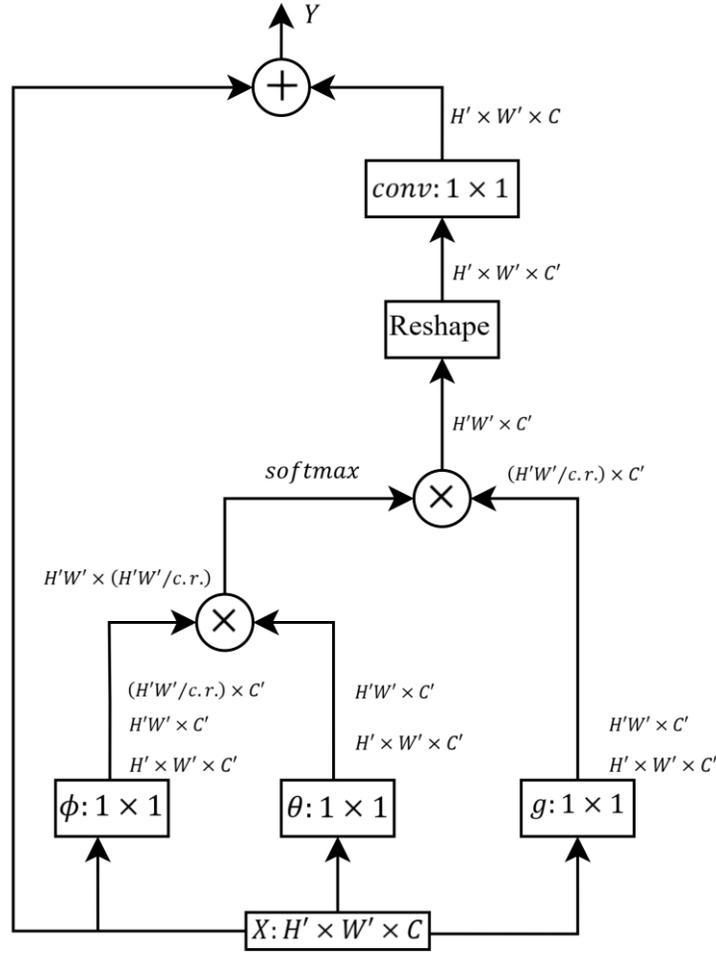


Figure 5.4: The structure of the self-attention module. \otimes and \oplus represent matrix addition and multiplication, respectively. C' and $c.r.$ denote the intermediate channel and the compression ratio, consecutively.

5.2.4 LSTM

LSTM is a recurrent neural network (RNN) in which the gating mechanism controls the information flow [284]. Using the mentioned characteristic, LSTM can solve the vanishing and exploding gradient issues faced in standard RNNs. These gates are the input gate (i), the output gate (o), and the forget gate (f). The mathematics behind the LSTM procedure is as follows:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f),$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i),$$

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c), \quad \text{Eq. 5-5}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t,$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o),$$

$$h_t = o_t \times \tanh(C_t)$$

where, t denotes the time step, w and b are the weight and bias of each gate, respectively. Furthermore, C and \tilde{C}_t represent the cell state and a vector of new candidate values created by a \tanh layer, consecutively. Also, h denotes the cell output. σ and \tanh are the sigmoid and hyperbolic tangent activation functions.

In order to handle time-series data related to complex nature, such as turbulence, the parallel form of LSTM [310] is used in this investigation. In other words, the summation of different LSTM networks yields the final output. The results of different parallel LSTM configurations with each other and with conventional LSTM was compared to find the best configuration for this study. The results indicated that the parallel LSTM in the current configuration (see Figure 5.5) outperforms the other tested forms. Furthermore, various values were used as the input sequence length to find the best option. The outcome showed that the mentioned value does not significantly impact the results in this application. Therefore, the length of the input sequence is chosen to be five, which means that the future state output contains the information from the five previous time steps.

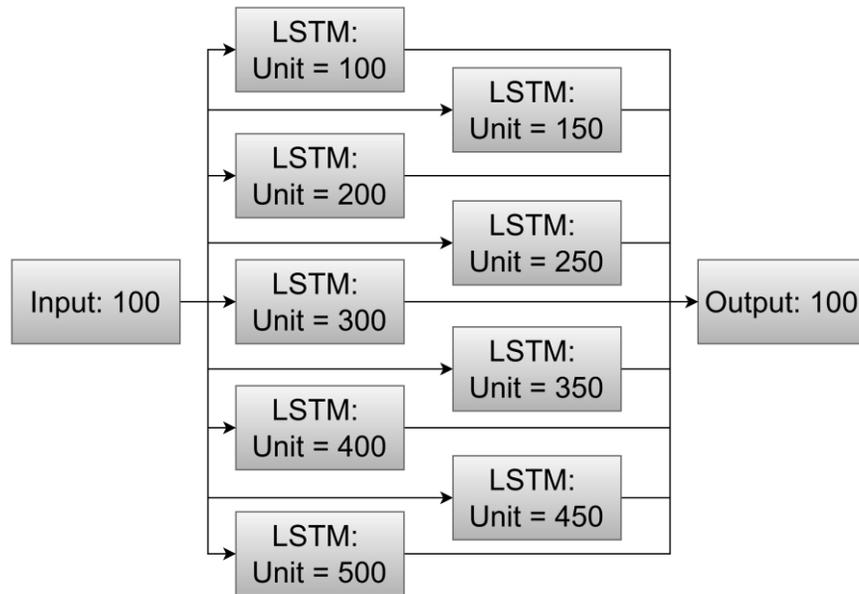


Figure 5.5: The structure of parallel LSTM

5.3 Dataset

The test case used in the present study is the unstable case used in chapter 2. The dataset used for this study is composed of the longitudinal and vertical velocity components, u and w , respectively, obtained from the above-mentioned CFD simulation, collected in the wake region of the building, located in the LES zone, shown in Figure 3.1 (c). The unstable dataset is chosen for this study because, based on the CFD simulations and the obtained time histories of the velocity components performed in [209], velocity fluctuations under the unstable condition are greater than those in the isothermal and stable cases. Furthermore, the number of available snapshots for the unstable case is more than those available for the other two cases. Therefore, the NIROM training can be performed using different numbers of snapshots to minimize the impact of the training dataset size on the results. The snapshots of the velocity components are gathered in the $x - z$ plane, with 160 points in the longitudinal direction and 60 points in the vertical direction. However, in order to focus on the regions with intense gradients, only the first 60 points in the longitudinal direction are selected. Furthermore, the results obtained from the NIROMs trained on the whole datapoints were in-line with the models' findings. It is noteworthy that $\delta t = 10\Delta t = 1 \times 10^{-3}$ s, where δt is the data sampling time step.

In order to train the models, after experimenting with different numbers of snapshots, 7605 snapshots are used. 80% of the dataset is used for training, while the rest is employed as the test dataset. The dimension of each image is $60 \times 60 \times 2$, where the last dimension indicates the number of channels. The first and second channels represent u and w , consecutively. Originally, data were sampled in the LES zone, where $\delta x \leq x \leq 4H$ and $\delta x \leq z \leq 1.5H$, where $\delta x = 0.025H = 0.004$ m. However, the datasets used for training and testing the models are collected in the hatched region shown in Figure 5.6.

The dataset was preprocessed by subtracting the mean value from it in order to work with fluctuations. After this step, the velocity values are scaled to be in the range of 0 and 1. The whole procedure can be found below:

$$v'_i = v_i - \bar{v}$$

Eq. 5-6

$$v''_i = (v'_i - \min(v')) / (\max(v') - \min(v'))$$

where, v_i indicates the either u or w at the i^{th} location. v' and v'' also represent the velocity fluctuation and the scaled velocity fluctuation, respectively, for either u or w .

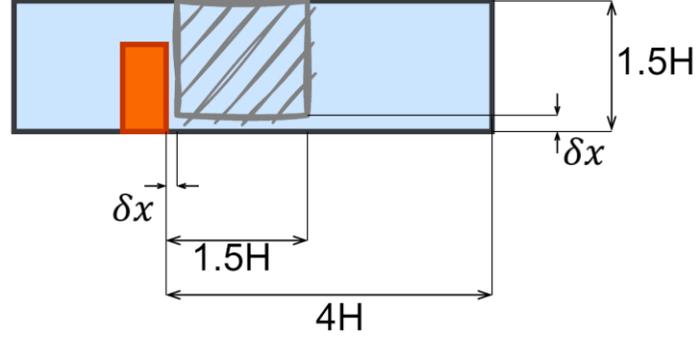


Figure 5.6: The data sampling region, located in the LES zone

5.4 NIROM training

Different NIROMs were developed by combining the aforementioned dimensionality reduction techniques, CAE, MS-CAE, and SA-CAE, with the parallel LSTM. In order to train the model, instantaneous velocity field data is used. As previously mentioned, the image size is $(N_z, N_x, N_c) = (60, 60, 2)$, where N_z and N_x are the number of datapoints in the z and x directions, respectively, and N_c represents the number of channels. After experimenting with different activation functions, rectified linear unit (ReLU) [319] is used for training CAE and SA-CAE, and sigmoid [320–322] is employed to train MS-CAE due to their better performance compared to other ones.

In the dimensionality reduction part, in order to train the model, a loss function (ε) is minimized. A combination of two functions, namely mean squared error (mse), ε_{mse} , and gradient difference loss [323], ε_{gdl} , is employed. The mentioned loss functions are defined as follows:

$$\varepsilon_{mse} = (1/N_x N_z N_c) \sum_{i=1}^{N_x} \sum_{j=1}^{N_z} \sum_{k=1}^{N_c} (X_{ijk} - X'_{ijk})^2$$

$$\varepsilon_{gdl} = (1/N_x N_z N_c) \sum_{i=1}^{N_x} \sum_{j=1}^{N_z} \sum_{k=1}^{N_c} [|(X_{ijk} - X_{(i-1)jk}) - (X'_{ijk} - X'_{(i-1)jk})| + |(X_{ijk} - X_{i(j-1)k}) - (X'_{ijk} - X'_{i(j-1)k})|]$$

Eq. 5-7

$$\varepsilon = \varepsilon_{mse} + \varepsilon_{gdl}$$

where, X and X' denote the original data and the reconstructed one, consecutively. The gradient difference loss function was used in other studies, such as [287], since it is beneficial for the model to avoid blurry predictions [324]. Adam gradient descent method [325], with the learning rate of 0.0001, is used to train CAE and SA-CAE, while the stochastic gradient descent (SGD) algorithm

with the learning rate of 0.001 is employed to train MS-CAE since it provides better generalization over the test dataset. The training process stopped by reaching convergence after about 220 and 400 epochs for CAE and MS-CAE, respectively, while it took about 700 epochs for SA-CAE. It is noteworthy that the training process was performed using different batch sizes, and the results showed the negligible influence of this factor on the training phase. In order to avoid overfitting during the training stage, an earlystopping criterion is used.

Afterward, the encoded field, \hat{X} , obtained from the encoder part of the dimensionality reduction methods, is used as the input of the LSTM model. In order to train the LSTM model, the weights are optimized to minimize a loss function, $\hat{\varepsilon}$, using the Adam algorithm, in the latent space:

Eq. 5-8

$$\hat{\varepsilon} = \left\| \hat{X}^{(n+1)\delta t} - \mathcal{F}_{LSTM}(\hat{X}^{n\delta t}, \hat{X}^{(n-1)\delta t}, \dots, \hat{X}^{(n-4)\delta t}) \right\|_2^2$$

where, \mathcal{F}_{LSTM} denotes the LSTM network. The Adam optimizer, with the learning rate of 0.001, is employed for the training loss minimization for all models. It took almost 200 epochs for the models to reach convergence. Again, an early stopping criterion is used in order to avoid overfitting during the training process. Similar to the previous step, the batch size had a negligible impact on the training phase.

5.5 Results and discussions

In the current section, the results obtained from the developed NIROMs are compared to each other and verified by the CFD results [209].

5.5.1 Dimensionality reduction part

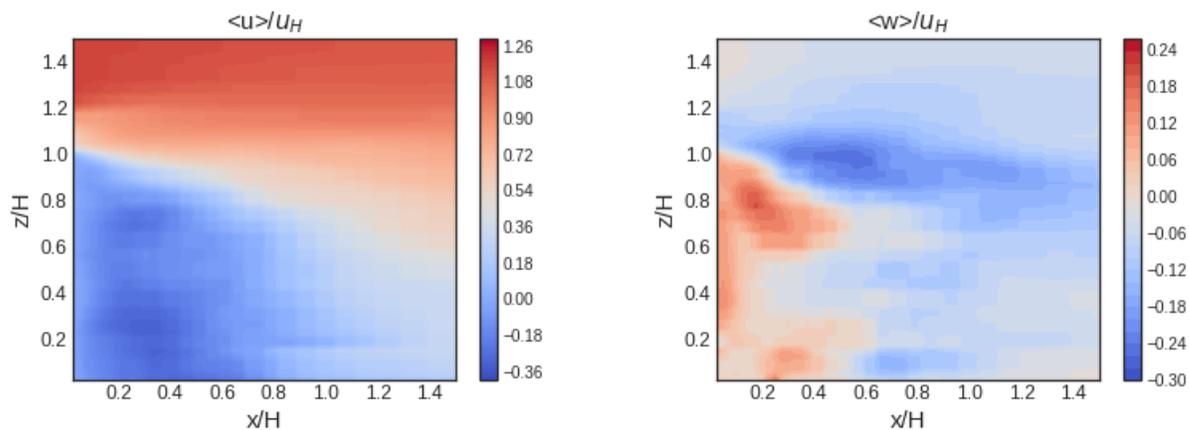
In the first step of evaluating the models' performance, the mapping capability of each model is assessed. As previously mentioned, the input of the models is a series of $60 \times 60 \times 2$ images. Also, the latent vector of all the methods is set to be $n_l = 100$. It should be noted that the models have been trained with different latent vector dimensions, i.e., $n_l = 50, 100, 150,$ and 200 .

The dimensionless time-averaged contours of the velocity components, reconstructed by each model over the test dataset, are shown in Figure 5.7 against the results of CFD.

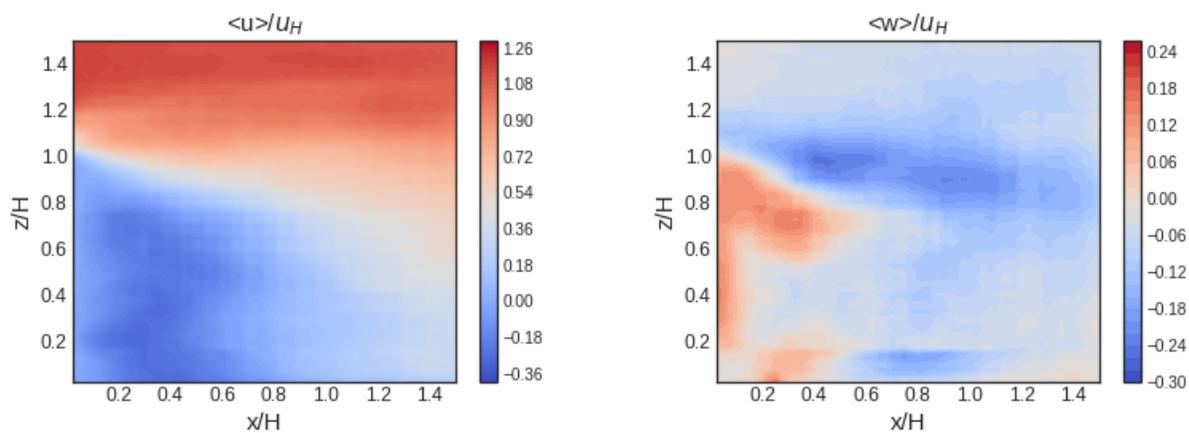
Longitudinal velocity component

Vertical velocity component

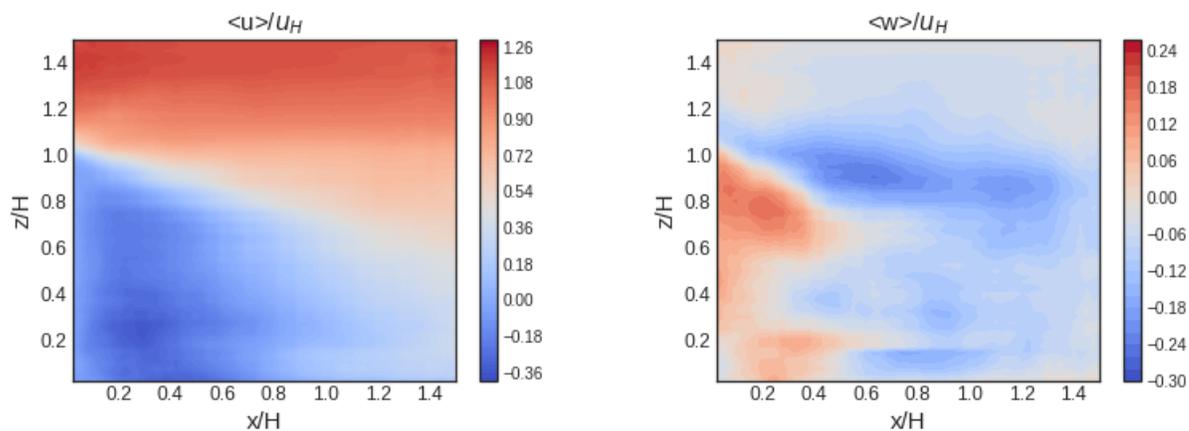
CFD



CAE



MS-CAE



SA-CAE

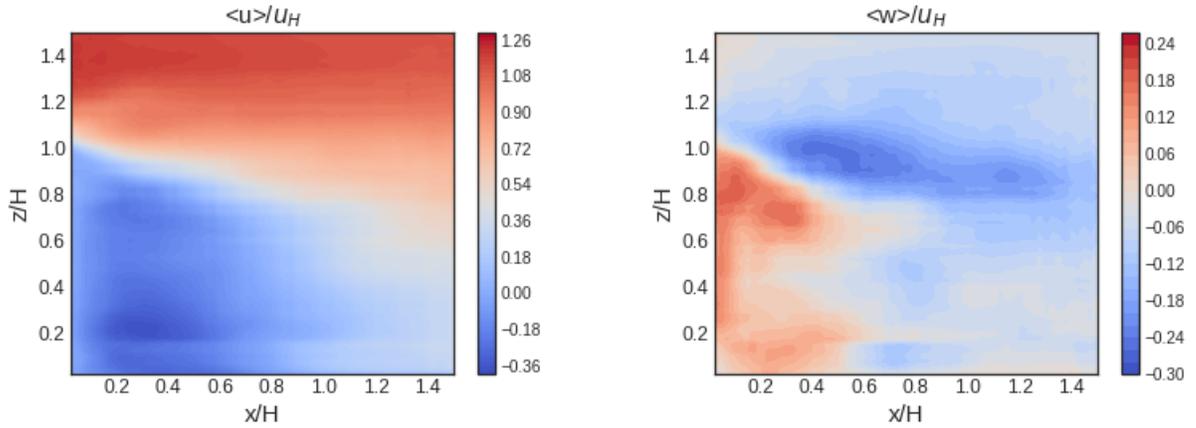
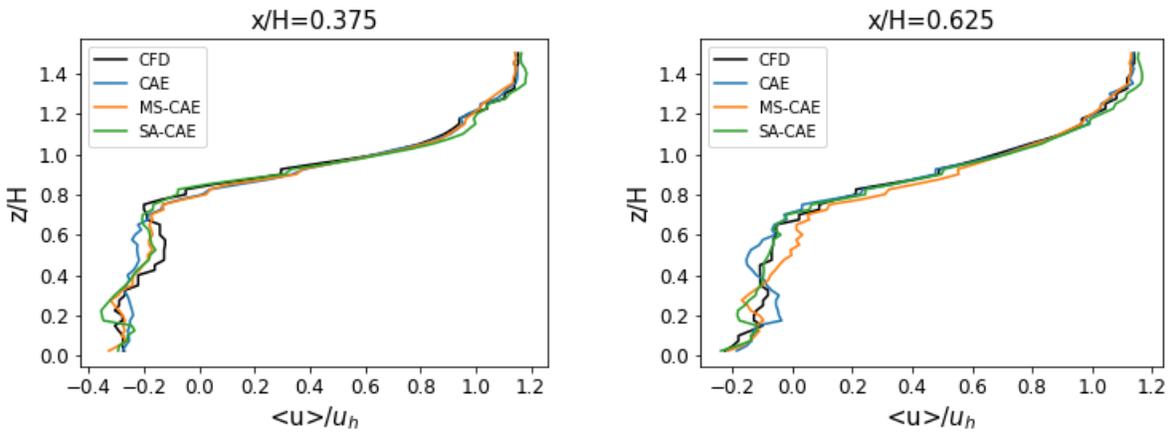


Figure 5.7: Contours of the dimensionless time-averaged velocity components obtained by CFD and the models used in the present study

Qualitatively, as can be observed in Figure 5.7, the performance of all the models in reconstructing u is promising. The location of the core of the wake region is reconstructed well by the models. MS-CAE and SA-CAE performed better than CAE in reconstructing w , particularly in regions near the building height. However, all models reconstructed the positive vertical velocity in the proximity of the heated ground and wall.

The profiles of dimensionless time-averaged longitudinal velocity, $\langle u \rangle / u_H$, reconstructed by each model, are depicted in Figure 5.8 over four sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, along $y/H = 0$. It should be noted that the center of the coordinate system is located on the ground, right behind the building. These sampling lines are the ones used for validating the CFD results [209]. Thus, for the sake of consistency, the same lines are used in this study.



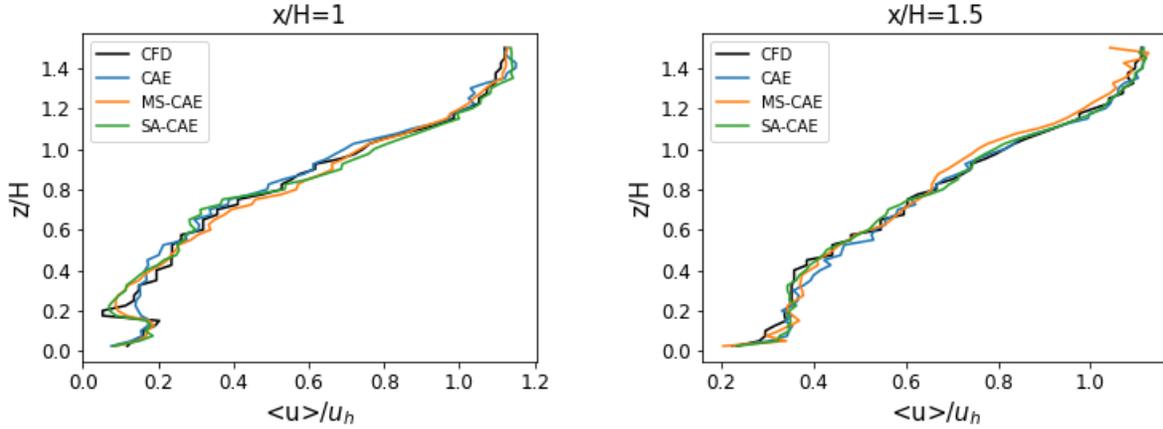


Figure 5.8: Comparison between the profiles of dimensionless time-averaged longitudinal velocity, $\langle u \rangle / u_h$, over the training dataset, along the sampling lines, $x/H = 0.375$, $x/H = 0.625$, $x/H = 1$, and $x/H = 1.5$, reconstructed by different models against those obtained by the CFD study of [209]

As shown in Figure 5.8, the models perform almost similarly in reconstructing the longitudinal velocity, particularly for $z/H > 0.8$. Furthermore, as the flow progresses downstream, the reconstructed velocity becomes more similar to the CFD results. The reason for the mentioned findings is that, with the increase in x and z , the fluctuations in the flow field dampen. In $z/H < 0.8$, the deviation of the reconstructed data from the CFD results becomes more significant compared to other regions since a strong recirculating region is formed in the mentioned area. Therefore, the flow field fluctuations are remarkably larger than in other areas. It can be observed that the reconstructed longitudinal velocity by SA-CAE and MS-CAE are in better agreement with the CFD data compared to CAE, in the regions close to the ground.

In order to quantify the reconstruction error, the average of the reconstruction mse , over time, for each velocity component is reported in Table 5.1.

Table 5.1: The mse of the reconstruction of the velocity field components for different models

Model	CAE	MS-CAE	SA-CAE
mse_u	0.0187	0.0164	0.0153
mse_w	0.0141	0.0100	0.0092

As reported in Table 5.1, mse_u is larger than mse_w for all the models. The results show that considering multi-scale and long-range dependencies among datapoints can improve the models' performance, which is supported by the results illustrated in Figures 5.7 and 5.8. In order to study the models' reconstruction quality, the turbulence kinetic energy percentage, E_k , is used for each direction to see what percentage of the fluctuations is captured by the models' reconstruction [264]:

$$E_k = \left(1 - \left\langle \frac{\sum_{i=1}^n (v' - \tilde{v}')^2}{\sum_{i=1}^n v'} \right\rangle \right) \times 100$$

where, where, $\langle . \rangle$ is the ensemble averaging in time, and v' and \tilde{v}' represent the reference and the reconstructed values of a velocity component. Also, the number of grid points is denoted by n . The average of E_k values, through time, for different models are reported in Table 5.2.

Table 5.2: The average of the directional turbulence kinetic energy percentage, E_k , of the velocity field components for different models

Model	CAE	MS-CAE	SA-CAE
$E_{k,u}$	94.48%	95.10%	95.16%
$E_{k,w}$	82.35%	83.60%	86.15%

Excellent reconstruction is performed by all the models, particularly for the longitudinal velocity component. It can be observed that the inclusion of the self-attention module can enhance the model's performance, particularly for the reconstruction of the vertical velocity component. On the whole, the results show excellent reconstructions by all the models.

5.5.2 NIROM performance evaluation

The next step is to feed the latent space, obtained by the encoder part of the autoencoder, into the LSTM model to predict its temporal evolution. Afterward, the decoder part of the autoencoder is used to take back the temporally evolved data to the original space. Figure 5.9 illustrates the contours of instantaneous u/u_H , obtained from the NIROMs, at different time steps, against the available CFD results [209]. It should be noted that the time steps for which the figures are depicted represent the snapshot numbers in the dataset, and they are not related to the CFD simulation time step. Hereinafter, for the sake of brevity, and in order for the figures' legends to fit in them, the DL-CAE, DL-MS-CAE, and DL-SA-CAE are briefed as CAE, MS-CAE, and SA-CAE, respectively. Also, "the dimensionality reduction part" of these models is pointed out whenever it is required. Furthermore, Figure 5.10 depicts the profiles of the dimensionless longitudinal velocity fluctuations reconstructed by the models and those obtained by the CFD simulation over the first two sampling lines (located in the recirculating regions) at different time steps. Also, in order to quantitatively compare the models' performance, the contours of the absolute deviation of the reconstructed longitudinal velocity component from the CFD data at different time steps are shown in Figure 5.11. These three figures are simultaneously used for the following analysis since they can provide more details on the models' performance.

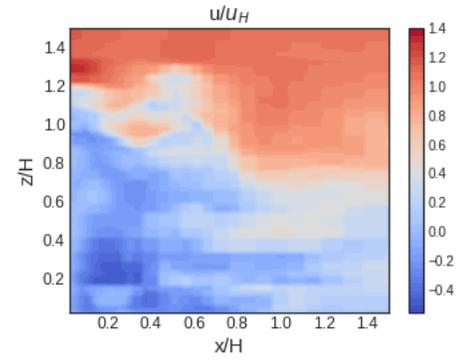
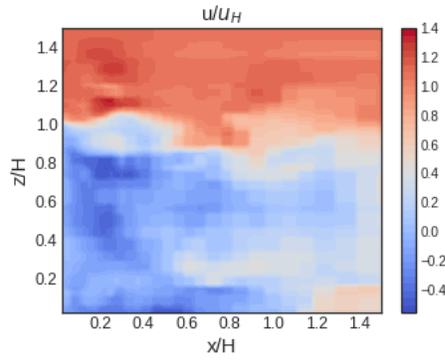
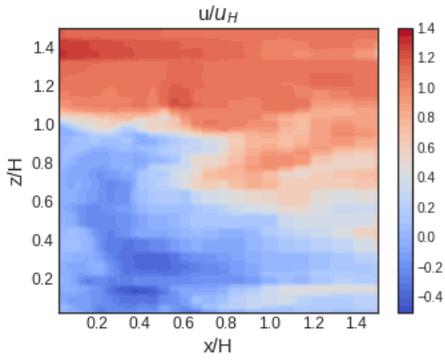
Time step

6500

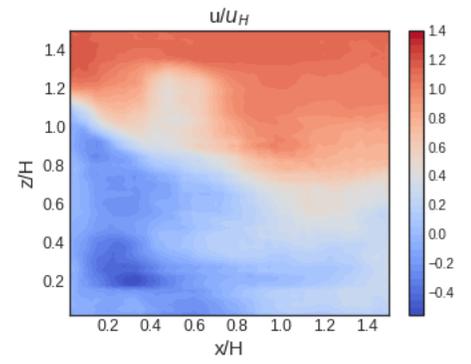
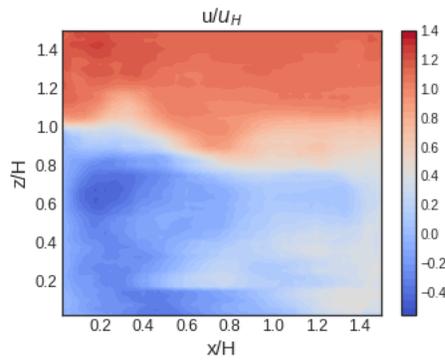
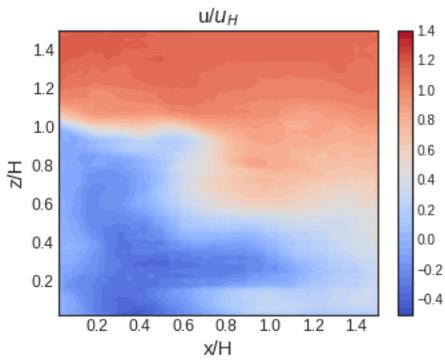
7000

7500

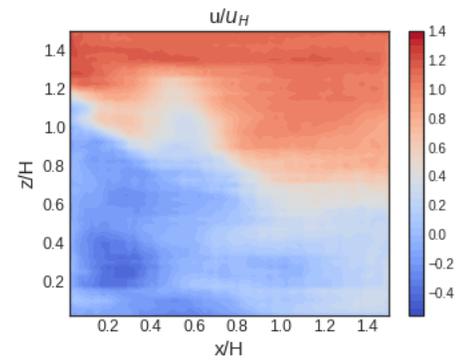
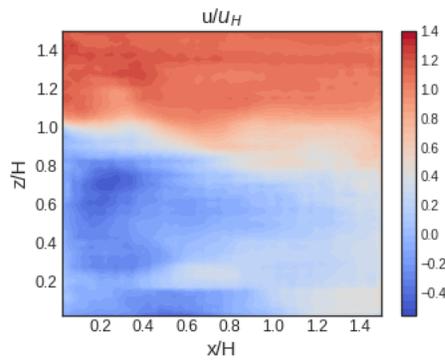
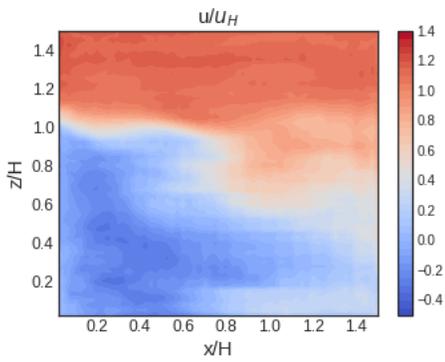
CFD



CAE



MS-CAE



SA-CAE

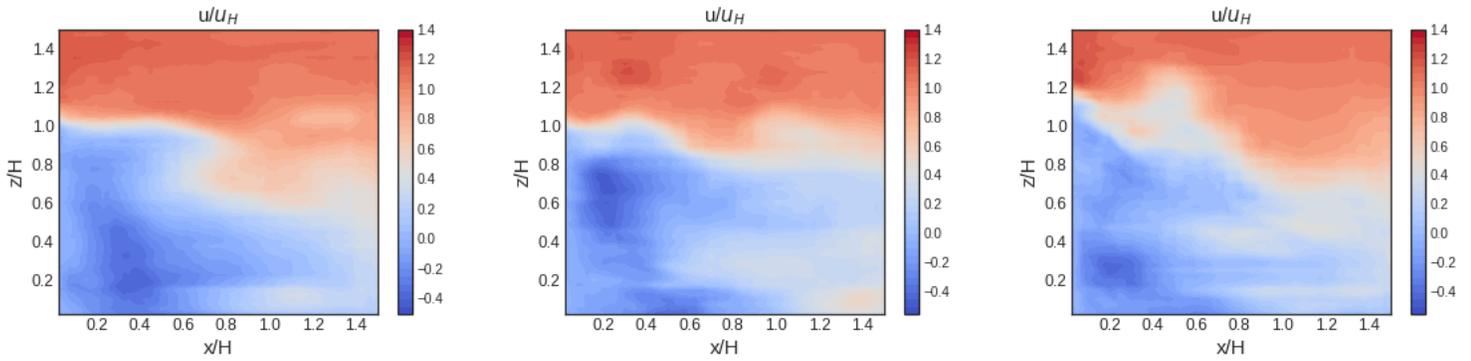
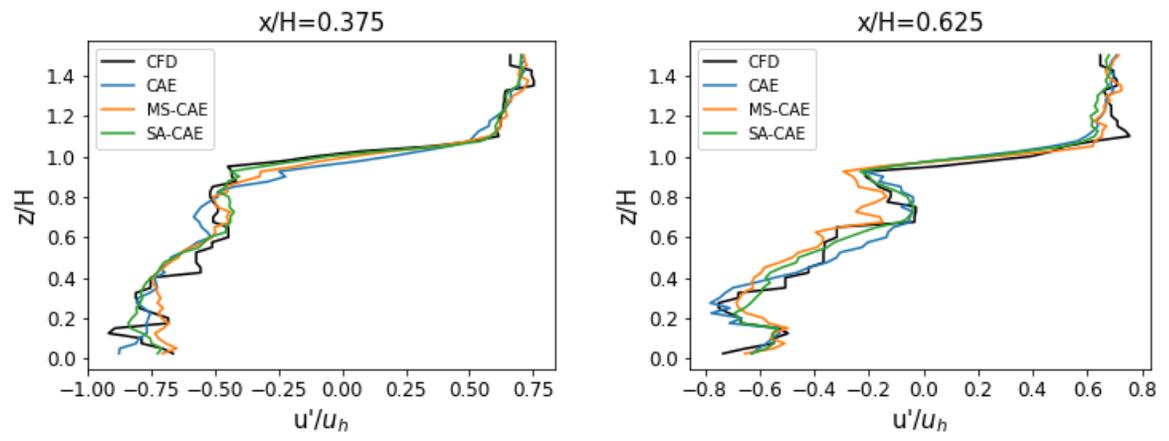


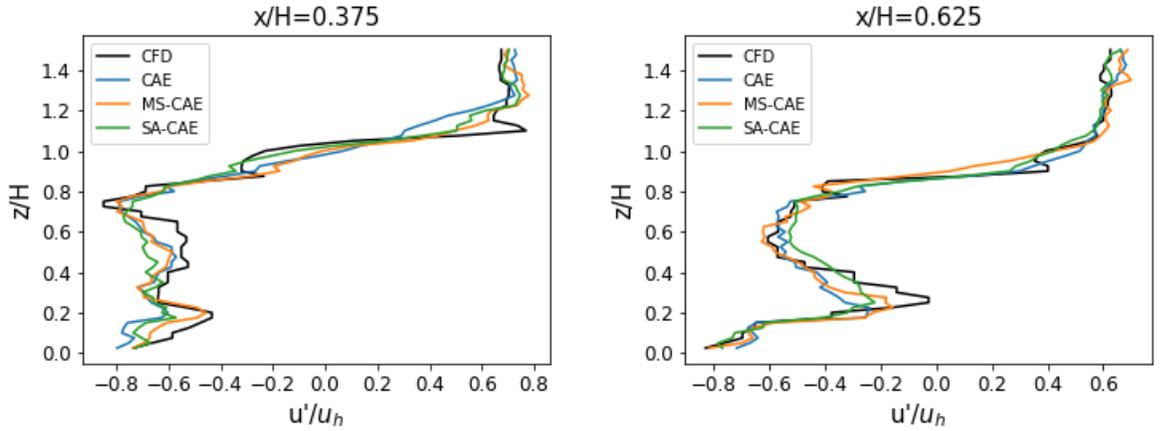
Figure 5.9: The contours of the dimensionless instantaneous longitudinal velocity component, u/u_H , reconstructed by NIROMs, and the available CFD data at different time steps [209]

Figure 5.9 shows that the longitudinal velocity component, reconstructed by the models, is in good agreement with the CFD results and the reconstructed results show similar characteristics to the reference data. However, the models are not able to capture all the fluctuations. This incapability is obviously more pronounced in the recirculating region, where significant fluctuations are observable in the results due to the effect of the walls on the flow field (see Figure 5.10).

Time step = 6500



Time step = 7000



Time step = 7500

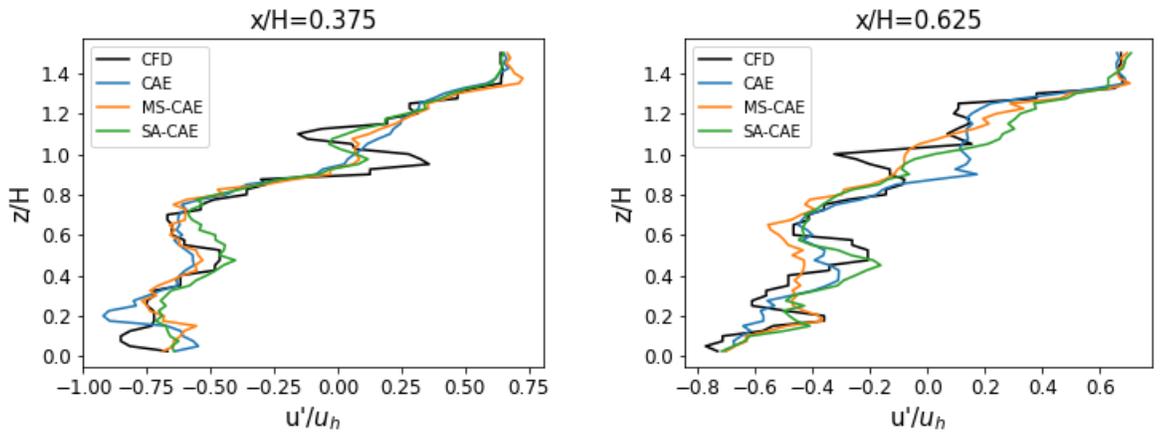


Figure 5.10: The profiles of the dimensionless longitudinal velocity fluctuations reconstructed by the models and those obtained by the CFD simulation over $x/H = 0.375$ and $x/H = 0.625$ at three different time steps

Furthermore, Figures 5.10 and 5.11 show remarkable deviations between the fluctuations reconstructed by the models and those from CFD at/above building height at *time step* = 7500, where velocity gradients are high. However, this limitation may become less significant if a larger dataset can be used for the model training since the model can see more vortex shedding cycles. At *time step* = 7000, the shear layer at the building height is better reconstructed by SA-CAE, followed by MS-CAE and CAE (see Figure 5.9), which is supported by Figures 5.10 and 5.11. At *time step* = 7500, the low-velocity region above the building height is better captured by SA-CAE and MS-CAE, compared to CAE. As Figure 5.10 illustrates, the modified models try to predict the sudden changes at $z/H > 1$ (particularly at $x/H = 0.375$). However, a remarkable underestimation in the values is visible. Furthermore, SA-CAE and MS-CAE show better performance in reconstructing the longitudinal velocity component in the vicinity of the ground compared to CAE.

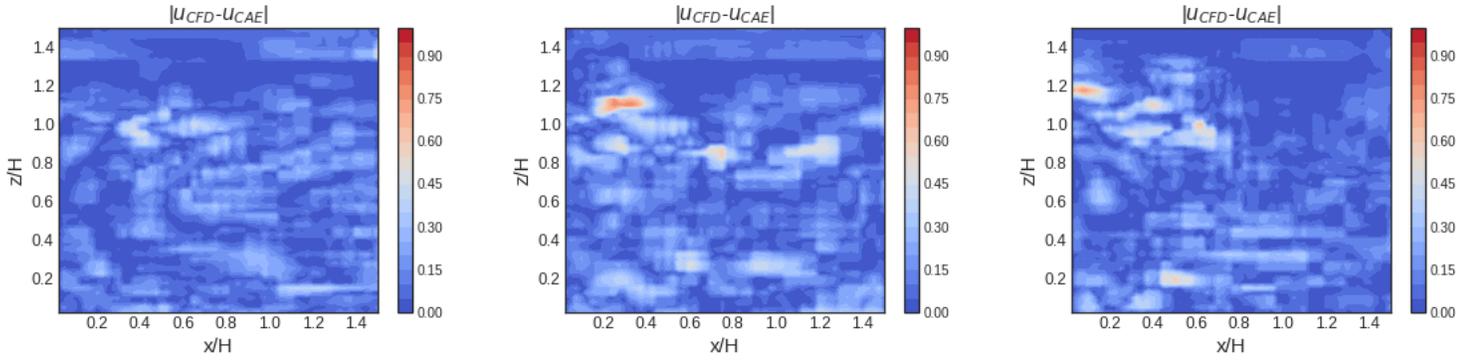
Time
step

6500

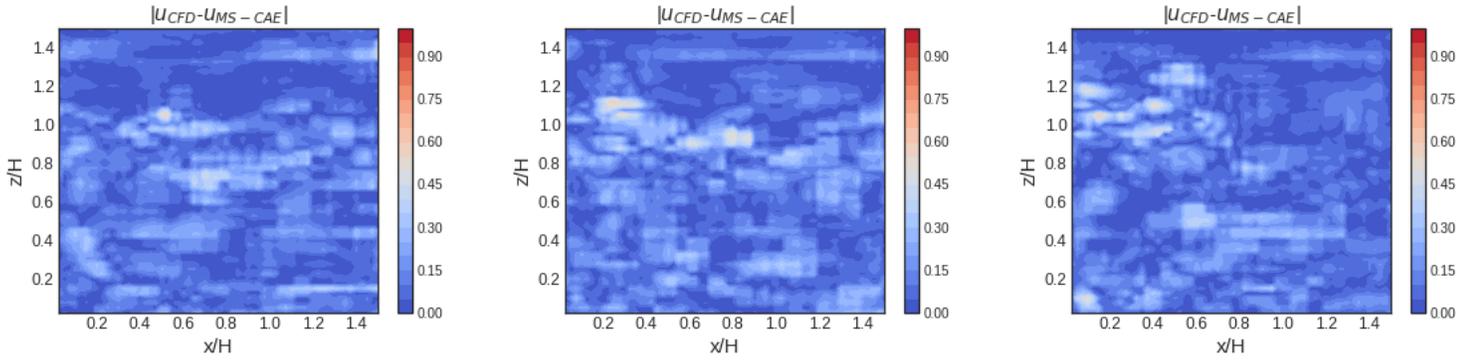
7000

7500

CAE



MS-
CAE



SA-
CAE

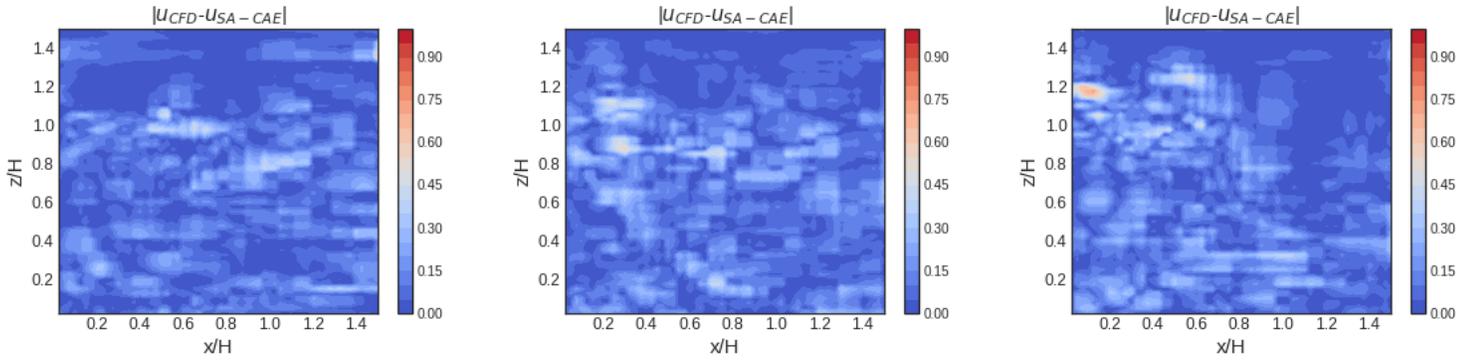
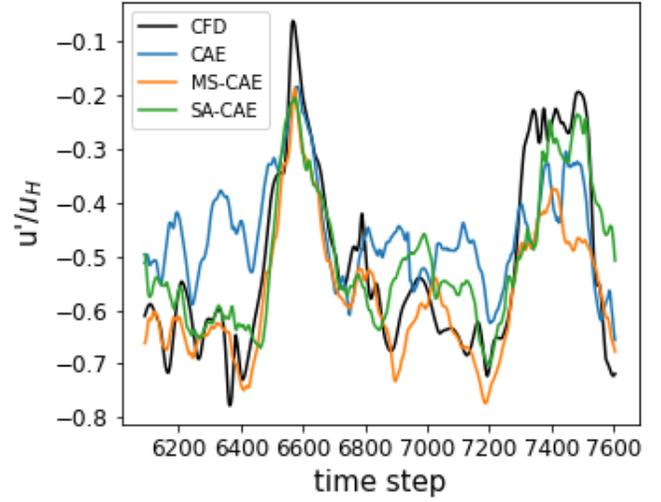
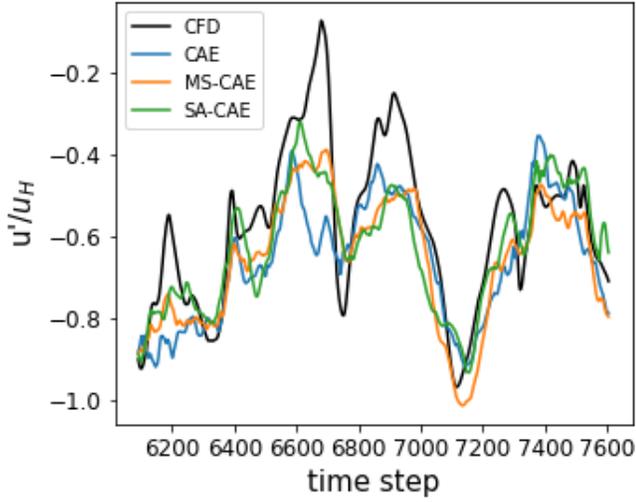


Figure 5.11: The contours of the absolute deviation of the instantaneous longitudinal velocity components reconstructed by NIROMs, with respect to the CFD data at different time steps [209]

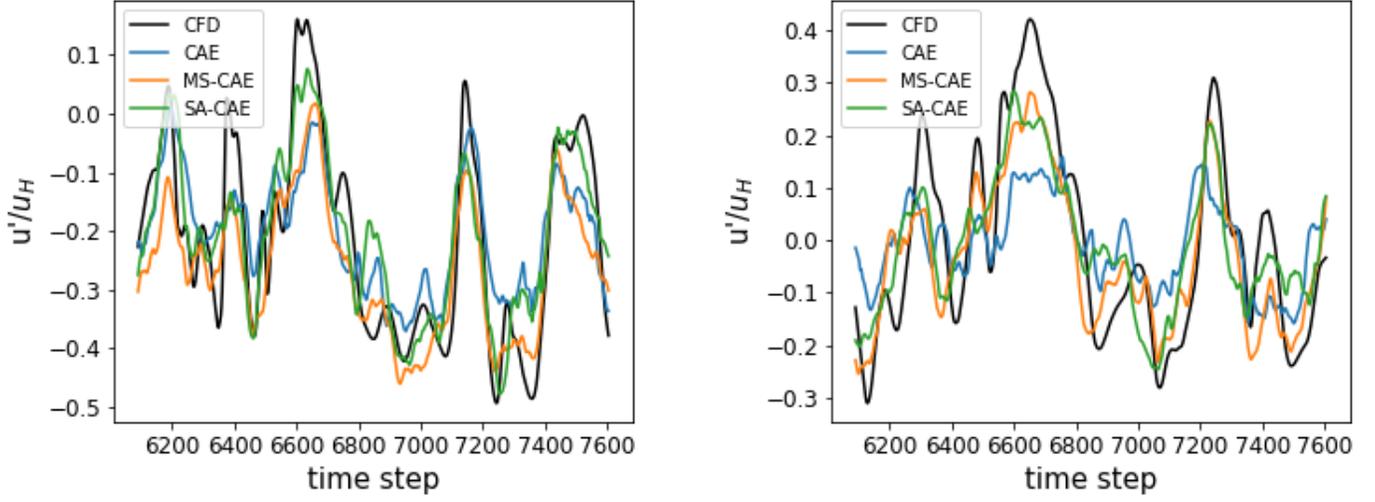
Figure 5.12 illustrates the time history of u'/u_H , obtained from the NIROMs and the CFD simulation, at four different points, $(x/H, z/H) = (0.375, 0.5)$, $(0.625, 0.5)$, $(1, 0.5)$ and $(1.5, 0.5)$. Furthermore, the time-averaged value of u'/u_H and its standard deviation, σ , in the time horizon shown in the graphs, are also included below them.

$(x/H, z/H) = (0.375, 0.5)$

$(x/H, z/H) = (0.625, 0.5)$



$\langle u'/u_H \rangle_{CFD} = -0.5725$	$(\sigma)_{\langle u'/u_H \rangle_{CFD}} = 0.1955$	$\langle u'/u_H \rangle_{CFD} = -0.5135$	$(\sigma)_{\langle u'/u_H \rangle_{CFD}} = 0.1705$
$\langle u'/u_H \rangle_{CAE} = -0.6581$	$(\sigma)_{\langle u'/u_H \rangle_{CAE}} = 0.1481$	$\langle u'/u_H \rangle_{CAE} = -0.4586$	$(\sigma)_{\langle u'/u_H \rangle_{CAE}} = 0.0951$
$\langle u'/u_H \rangle_{MS-CAE} = -0.6578$	$(\sigma)_{\langle u'/u_H \rangle_{MS-CAE}} = 0.1558$	$\langle u'/u_H \rangle_{MS-CAE} = -0.5696$	$(\sigma)_{\langle u'/u_H \rangle_{MS-CAE}} = 0.1198$
$\langle u'/u_H \rangle_{SA-CAE} = -0.6251$	$(\sigma)_{\langle u'/u_H \rangle_{SA-CAE}} = 0.1457$	$\langle u'/u_H \rangle_{SA-CAE} = -0.5046$	$(\sigma)_{\langle u'/u_H \rangle_{SA-CAE}} = 0.1265$
$(x/H, z/H) = (1, 0.5)$		$(x/H, z/H) = (1.5, 0.5)$	



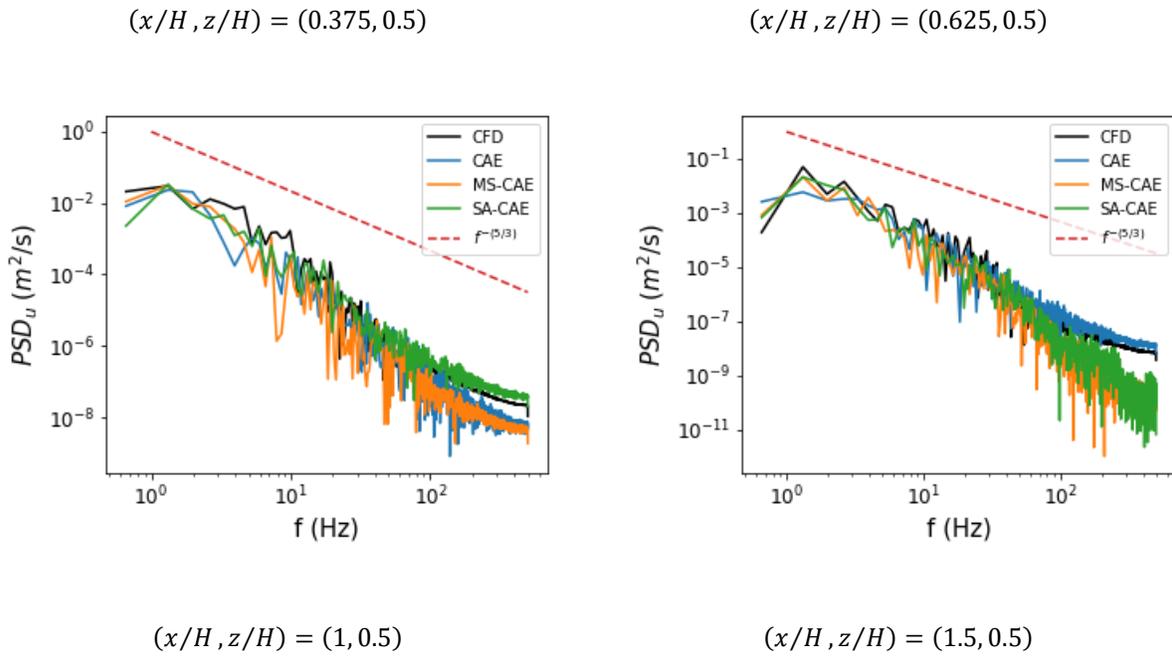
$\langle u'/u_H \rangle_{CFD} = -0.2088$	$(\sigma)_{\langle u'/u_H \rangle_{CFD}} = 0.1653$	$\langle u'/u_H \rangle_{CFD} = -0.0051$	$(\sigma)_{\langle u'/u_H \rangle_{CFD}} = 0.1843$
$\langle u'/u_H \rangle_{CAE} = -0.2001$	$(\sigma)_{\langle u'/u_H \rangle_{CAE}} = 0.0981$	$\langle u'/u_H \rangle_{CAE} = -0.0082$	$(\sigma)_{\langle u'/u_H \rangle_{CAE}} = 0.0827$
$\langle u'/u_H \rangle_{MS-CAE} = -0.2590$	$(\sigma)_{\langle u'/u_H \rangle_{MS-CAE}} = 0.1155$	$\langle u'/u_H \rangle_{MS-CAE} = -0.0297$	$(\sigma)_{\langle u'/u_H \rangle_{MS-CAE}} = 0.1362$
$\langle u'/u_H \rangle_{SA-CAE} = -0.2033$	$(\sigma)_{\langle u'/u_H \rangle_{SA-CAE}} = 0.1342$	$\langle u'/u_H \rangle_{SA-CAE} = -0.0080$	$(\sigma)_{\langle u'/u_H \rangle_{SA-CAE}} = 0.1246$

Figure 5.12: Time history of u'/u_H , obtained from the CFD simulation and the NIROMs, at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$, with related time-averaged and standard deviation values

Non-sinusoidal and fast-changing signals can be seen in the results of the NIROMs in Figure 5.12. The results obtained by SA-CAE and MS-CAE are in better agreement with the CFD results, compared to those reconstructed by CAE. In other words, the tendency of the modified models to capture the jumps in the CFD results is higher than the baseline model, i.e., CAE. However, the reconstructed results by the models have a lower standard deviation, and thus, a lower variance, than the CFD results. This means that, on the whole, the models have the tendency to underestimate TKE related to the longitudinal velocity component. Although the modified models can improve the results, some coherent structures are still missing. It should be noted that errors at $(x/H, z/H) = (1.5, 0.5)$ may also be related to the location of the point, which is located at the very end of the domain (i.e., weaker performance of convolutional layers compared to the locations in the middle of the domain).

The power spectrum density (PSD) function of the longitudinal velocity fluctuations at the aforementioned four sampling points is depicted in Figure 5.13. The general spectrum shape (i.e., large eddy behavior at low frequencies, $-5/3$ power law inertial sub-range, and energy decay at high frequencies) is captured by all the models. At low frequencies, the energy content of the

reconstructed velocity field of SA-CAE and MS-CAE is closer to that of the CFD. However, as mentioned in the explanation of Figure 5.12, the models tend to underestimate the TKE, related to the longitudinal velocity component, which can also be observed in Figure 5.13. In most parts of the low-frequency region, SA-CAE and MS-CAE perform better than CAE, which may be due to considering the long-range and multi-scale dependencies among the datapoints in the decoder part. At high frequencies, on the whole, the energy decays at almost similar rates for the results obtained by SA-CAE and MS-CAE. This decay rate for CAE is more similar to that of the CFD data, compared to the other two models. Furthermore, as can be seen in Figure 5.13, most of the frequencies at which peaks occur are captured by the models. For instance, at $(x/H, z/H) = (0.375, 0.5)$, the most distinct frequency obtained by the models is similar to that yielded by the CFD simulation ($= 1.4\text{Hz}$), which corresponds to the Strouhal number (St) of 0.16. Also, at $(x/H, z/H) = (1, 0.5)$, the most dominant frequency resulting from SA-CAE is similar to that of the CFD results, corresponding to $St = 0.23$.



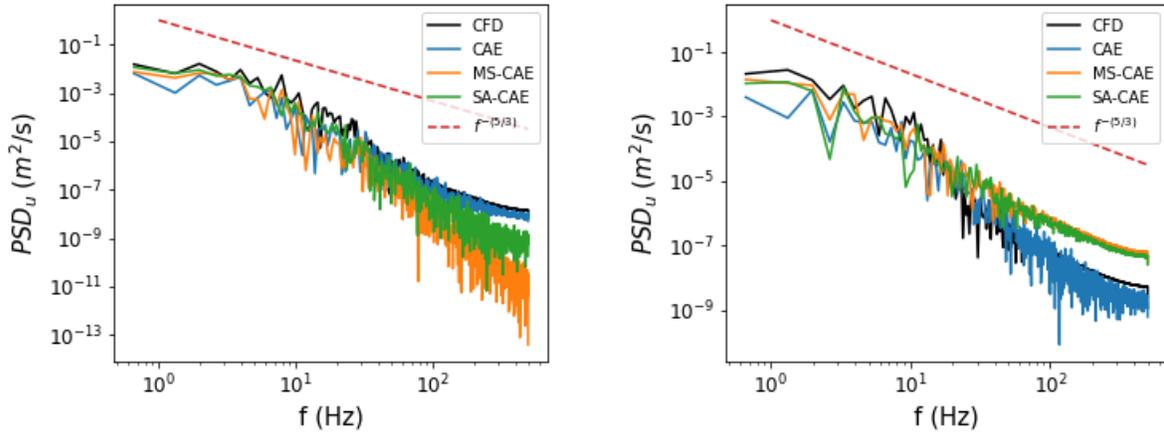


Figure 5.13: Power spectrum density (PSD) function of streamwise velocity fluctuations at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$

Figure 5.14 shows the contours of instantaneous w/u_H , obtained from the NIROMs, at different time steps, against the available CFD results [209]. Furthermore, the contours of the absolute deviation of the reconstructed w from the CFD data, at different time steps, are illustrated in Figure 5.15. Although the models are able to reconstruct the overall pattern of the vertical velocity component, they, on the whole, tend to underestimate the results. This tendency may lead to erroneous results even in predicting the velocity direction as reported for CAE, close to the building height, at *time step* = 6500. Similar to the reconstruction of the longitudinal velocity component, the models are not able to capture all the fluctuations, and the models show smoother contours compared to those obtained from CFD. It should be noted that smaller deviations between the reconstructed vertical velocity component and the CFD results, compared to the similar deviations for the longitudinal component, may be rooted in the difference between their datasets; the larger standard deviation of the longitudinal dataset compared to that of the vertical velocity component.

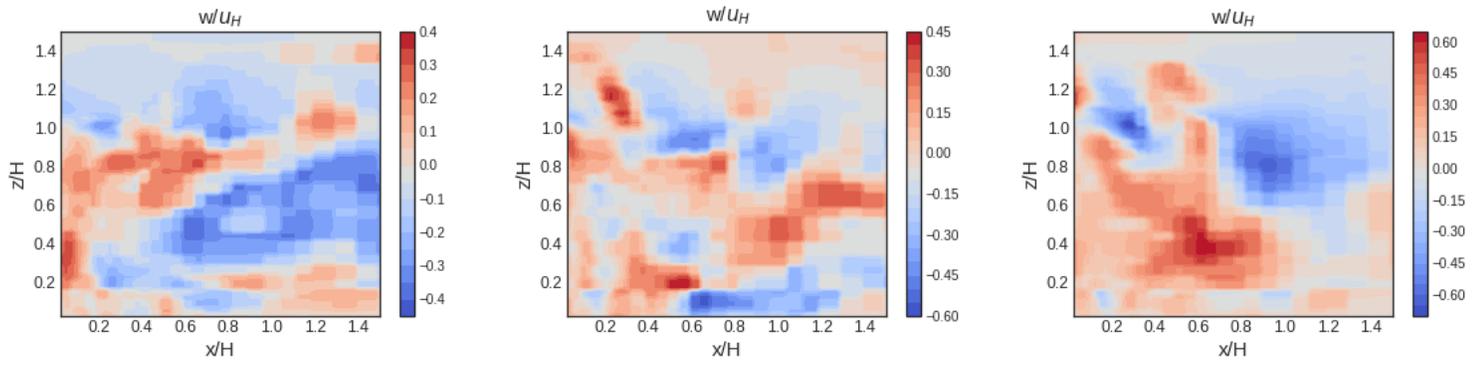
*Time
step*

6500

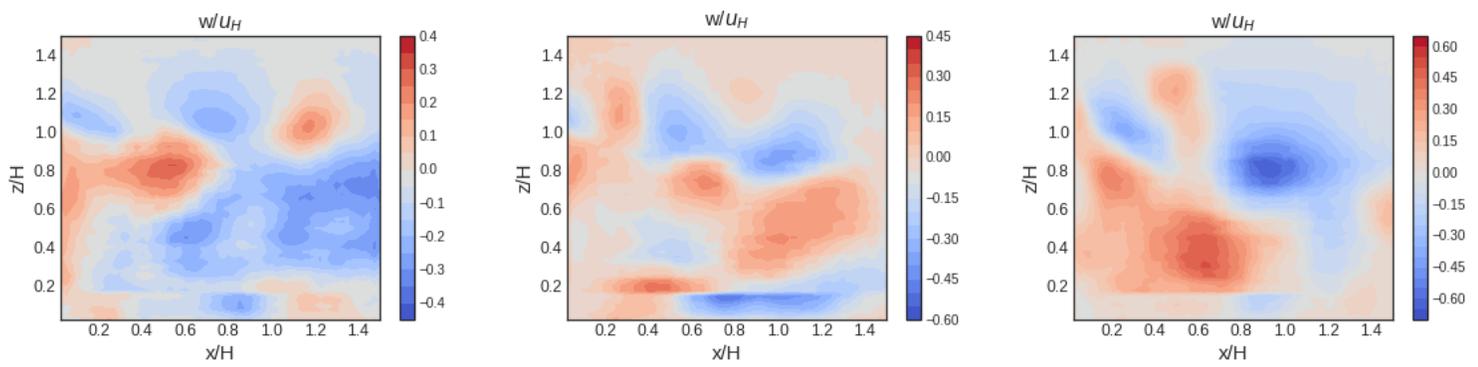
7000

7500

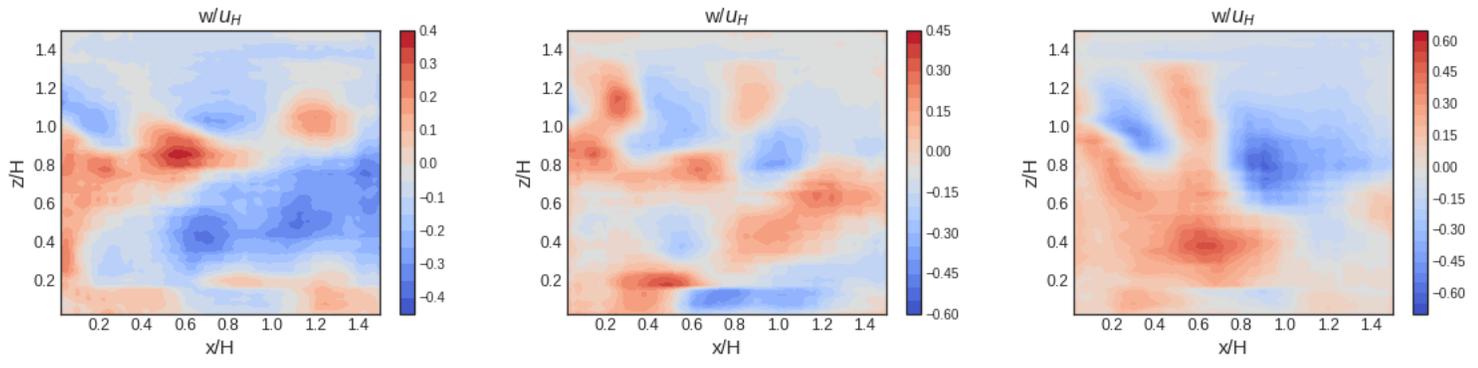
CFD



CAE



MS-CAE



SA-CAE

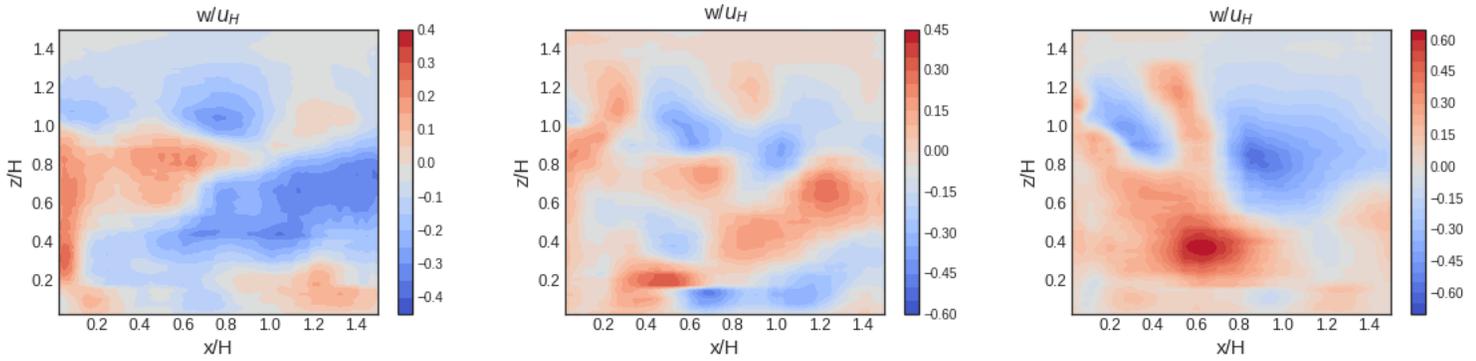


Figure 5.14: The contours of the dimensionless instantaneous vertical velocity component, w/u_H , reconstructed by NIROMs, and the available CFD data [209]

The reconstructed vertical velocity by SA-CAE and MS-CAE in the proximity of the ground are in better agreement with the CFD results compared to that reproduced by CAE. Furthermore, CAE reconstructs a smooth and underestimated velocity field in the vicinity of the leeward wall. For instance, at *time step* = 7000 and 7500, almost stagnant or, in some cases, downward velocity is predicted by CAE, which is not in line with the physics of the problem (i.e., heated wall). On the other hand, the modified models are capable of reconstructing the thin upward layer, close to the leeward wall, in better agreement with the CFD results. On the whole, the performance of SA-CAE and MS-CAE in reconstructing the vertical velocity component is better than CAE, particularly in the regions with intense gradients, such as at the building height, in the proximity of the building wall, and close to the ground.

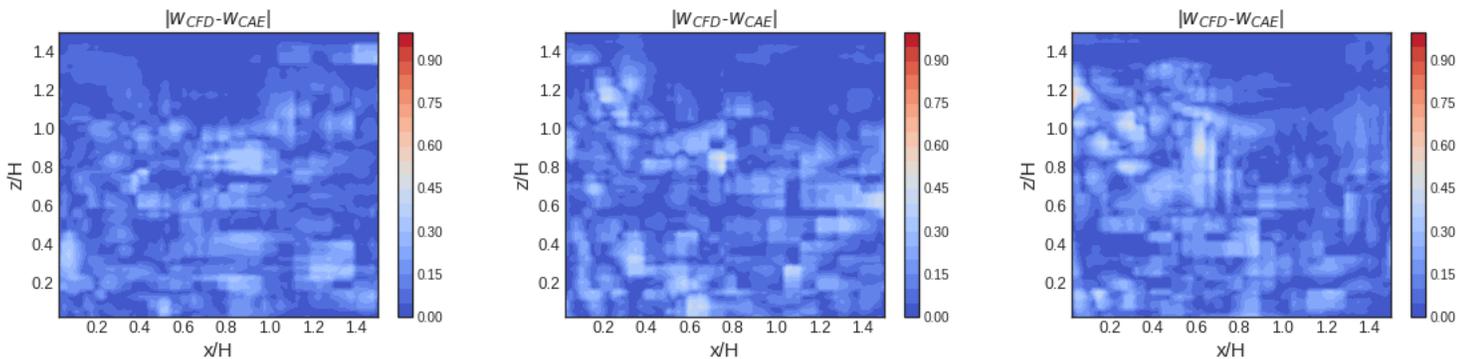
Time step

6500

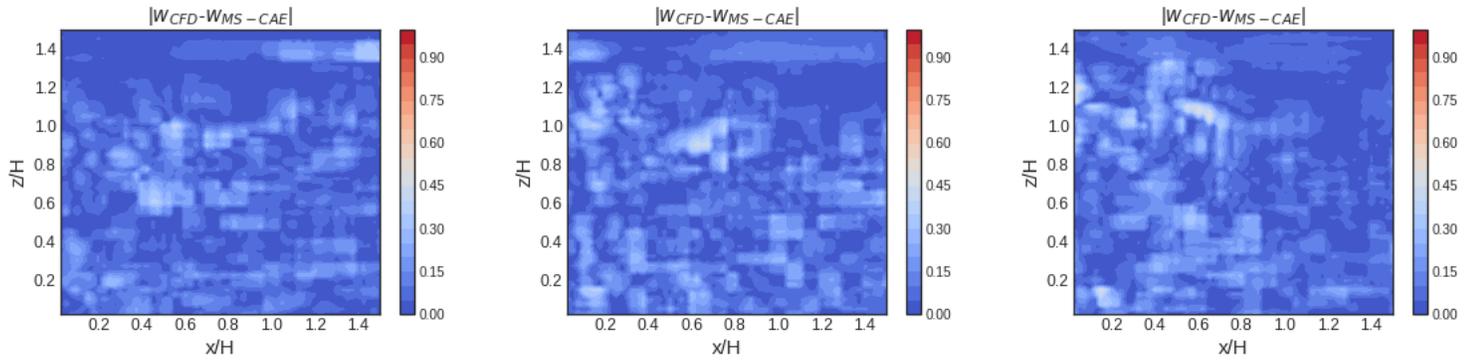
7000

7500

CAE



MS-CAE



SA-CAE

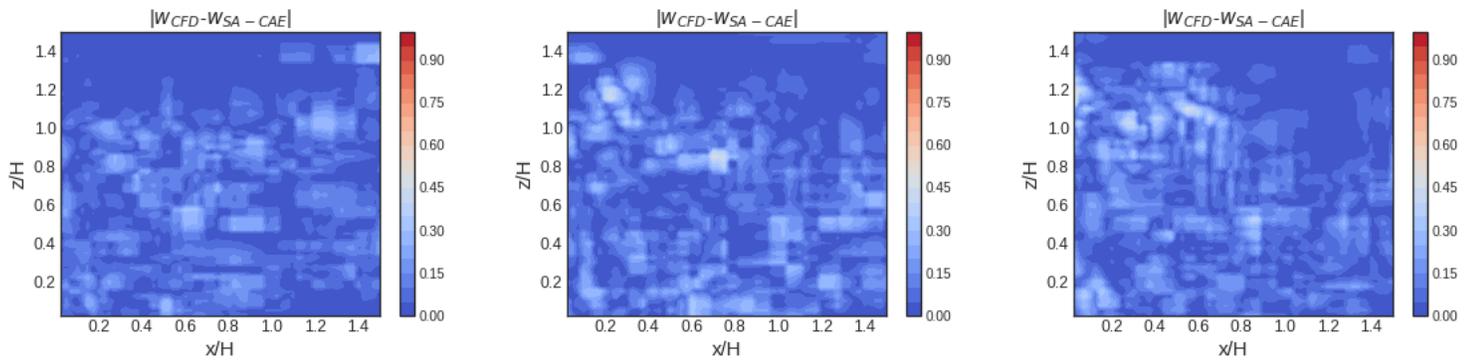
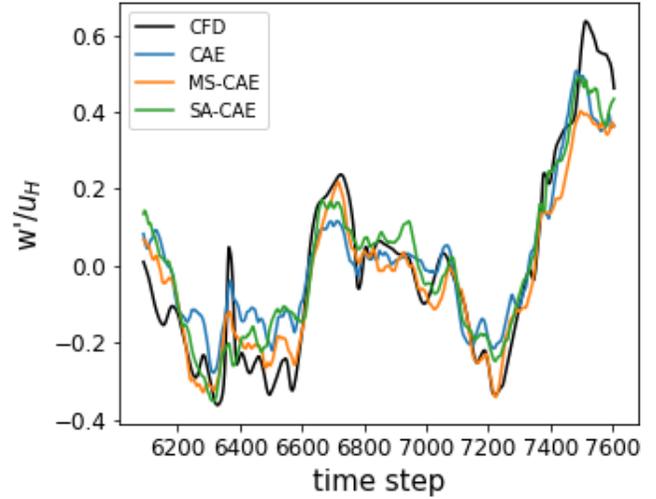
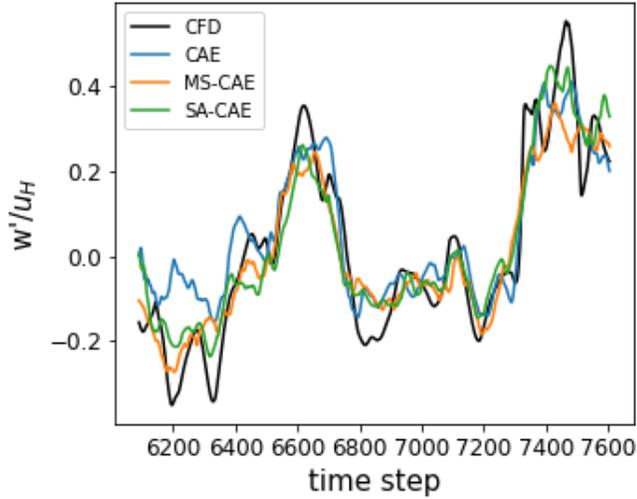


Figure 5.15: The contours of the absolute deviation of the instantaneous vertical velocity components reconstructed by NIROMs, with respect to the CFD data at different time steps [209]

The time history of the vertical velocity fluctuations, at the previously-mentioned sampling points, along with the related time-averaged and standard deviation obtained from CFD and the developed NIROMs, are shown in Figure 5.16. Similar to the longitudinal velocity component, the modified models perform better than the baseline one, i.e., CAE, in capturing fast-changing signals. Furthermore, on the whole, the mean values and standard deviations reported for SA-CAE are in better agreement with those of the CFD simulation, compared to the other two models. Higher standard deviation, and thus, variance, predicted by SA-CAE compared to the other two models results in higher TKE related to the vertical velocity component, which is in better agreement with the CFD results.

$$(x/H, z/H) = (0.375, 0.5)$$

$$(x/H, z/H) = (0.625, 0.5)$$



$$\langle (w')/u_H \rangle_{CFD} = 0.0138$$

$$(\sigma)_{\langle (w')/u_H \rangle_{CFD}} = 0.2135$$

$$\langle (w')/u_H \rangle_{CFD} = -0.0186$$

$$(\sigma)_{\langle (w')/u_H \rangle_{CFD}} = 0.2502$$

$$\langle (w')/u_H \rangle_{CAE} = 0.0490$$

$$(\sigma)_{\langle (w')/u_H \rangle_{CAE}} = 0.1618$$

$$\langle (w')/u_H \rangle_{CAE} = 0.0064$$

$$(\sigma)_{\langle (w')/u_H \rangle_{CAE}} = 0.1793$$

$$\langle (w')/u_H \rangle_{MS-CAE} = 0.0054$$

$$(\sigma)_{\langle (w')/u_H \rangle_{MS-CAE}} = 0.1698$$

$$\langle (w')/u_H \rangle_{MS-CAE} = -0.0384$$

$$(\sigma)_{\langle (w')/u_H \rangle_{MS-CAE}} = 0.1923$$

$$\langle (w')/u_H \rangle_{SA-CAE} = 0.0196$$

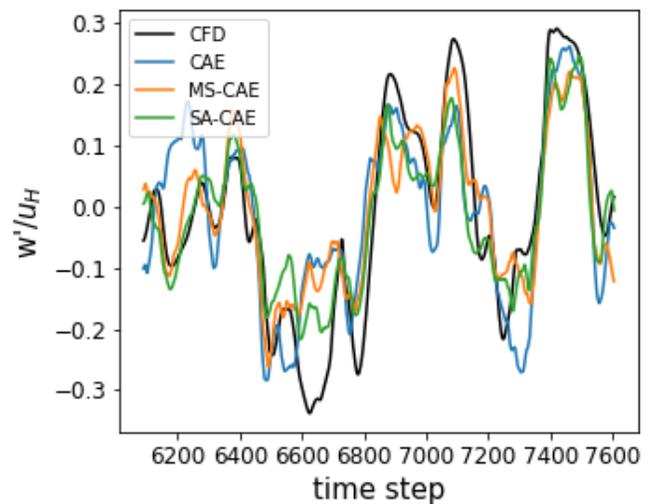
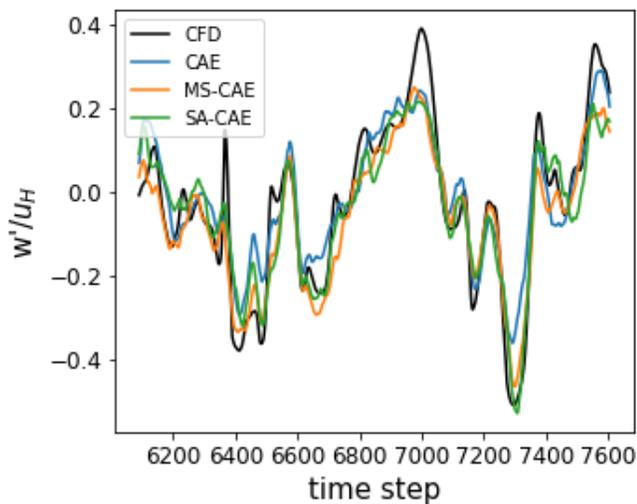
$$(\sigma)_{\langle (w')/u_H \rangle_{SA-CAE}} = 0.1828$$

$$\langle (w')/u_H \rangle_{SA-CAE} = 0.0008$$

$$(\sigma)_{\langle (w')/u_H \rangle_{SA-CAE}} = 0.2037$$

$$(x/H, z/H) = (1, 0.5)$$

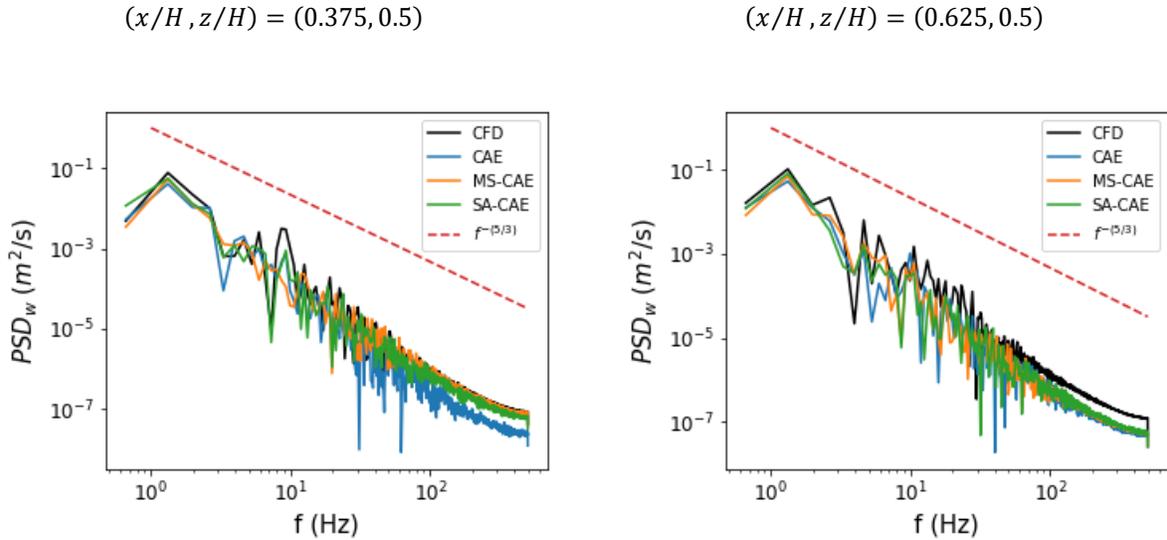
$$(x/H, z/H) = (1.5, 0.5)$$



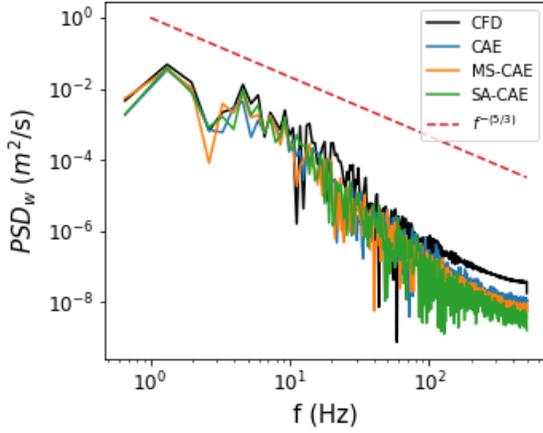
$\langle w' \rangle / u_H)_{CFD} = -0.0299$	$(\sigma)_{\langle w' \rangle / U_H, CFD} = 0.1978$	$\langle w' \rangle / u_H)_{CFD} = -0.0102$	$(\sigma)_{\langle w' \rangle / U_H, CFD} = 0.1405$
$\langle w' \rangle / u_H)_{CAE} = -0.0138$	$(\sigma)_{\langle w' \rangle / U_H, CAE} = 0.1479$	$\langle w' \rangle / u_H)_{CAE} = -0.0084$	$(\sigma)_{\langle w' \rangle / U_H, CAE} = 0.1186$
$\langle w' \rangle / u_H)_{MS-CAE} = -0.0590$	$(\sigma)_{\langle w' \rangle / U_H, MS-CAE} = 0.1602$	$\langle w' \rangle / u_H)_{MS-CAE} = -0.0030$	$(\sigma)_{\langle w' \rangle / U_H, MS-CAE} = 0.1847$
$\langle w' \rangle / u_H)_{SA-CAE} = -0.0412$	$(\sigma)_{\langle w' \rangle / U_H, SA-CAE} = 0.1612$	$\langle w' \rangle / u_H)_{SA-CAE} = -0.0135$	$(\sigma)_{\langle w' \rangle / U_H, SA-CAE} = 0.1166$

Figure 5.16: Time history of w'/u_H , obtained from the CFD simulation and the NIROMs, at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$, with related time-averaged and standard deviation values

The PSD of the vertical velocity fluctuations at the sampling points is illustrated in Figure 5.17. Similar to the PSD of the longitudinal velocity components, the general spectrum shape is captured by all NIROMs. Although there is good agreement between the results of the models and those obtained by CFD at low-frequency ranges, there still exist some discrepancies between the reconstructed and reference data, which is supported by the results provided in Figure 5.16. In low-frequency ranges, the frequencies at which peaks occur are captured by the models, particularly by SA-CAE and MS-CAE. However, some errors in the PSD levels are visible at these dominant frequencies. The dominant frequencies captured by SA-CAE are similar to those of the CFD simulation. At $(x/H, z/H) = (0.375, 0.5)$, the dominant frequency is 1.4 Hz , corresponding to $St = 0.16$. Also, the dominant frequencies captured by CAE and MS-CAE agree with those predicted by the CFD simulation, except at $(x/H, z/H) = (1.5, 0.5)$, which may be due to the location of the point, as mentioned before in the discussion of Figure 5.12. At higher frequencies, the energy level is predicted well by SA-CAE compared to CFD, particularly at $(x/H, z/H) = (0.375, 0.5)$ and $(0.625, 0.5)$, which are located in the recirculating region.



$$(x/H, z/H) = (1, 0.5)$$



$$(x/H, z/H) = (1.5, 0.5)$$

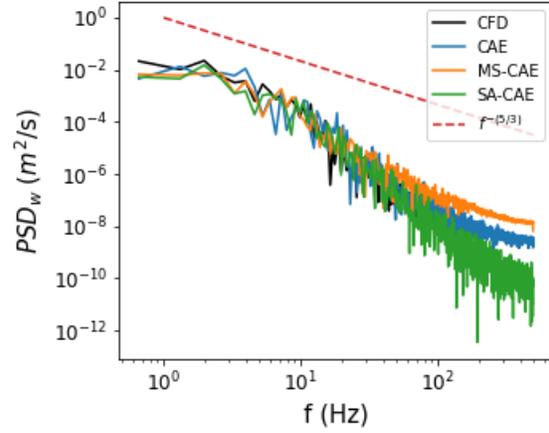


Figure 5.17: PSD function of vertical velocity fluctuations at four sampling points, $(x/H, z/H) = (0.375, 0.5), (0.625, 0.5), (1, 0.5)$ and $(1.5, 0.5)$

The time trace of the mean squared error of the longitudinal velocity component (mse_u) and that of the vertical velocity component (mse_w) are depicted in Figure 5.18. As previously mentioned, the time step on the x -axis represents the number of snapshots in the dataset. The order of mse for all the models remains the same over the whole test dataset. Although there exist some oscillations in the error values, they do not grow. As can be observed in Figure 5.18, on the whole, NIROMs based on SA-CAE and MS-CAE perform better than that based on CAE in reconstructing both the longitudinal and vertical velocity components. Regarding the reconstruction of the vertical velocity component, the SA-CAE has the upper hand in almost all time steps. The mean value of mse for each NIROM is listed in Table 5.3. Furthermore, the time required for the model training and velocity field reconstruction is mentioned.

(a)

(b)

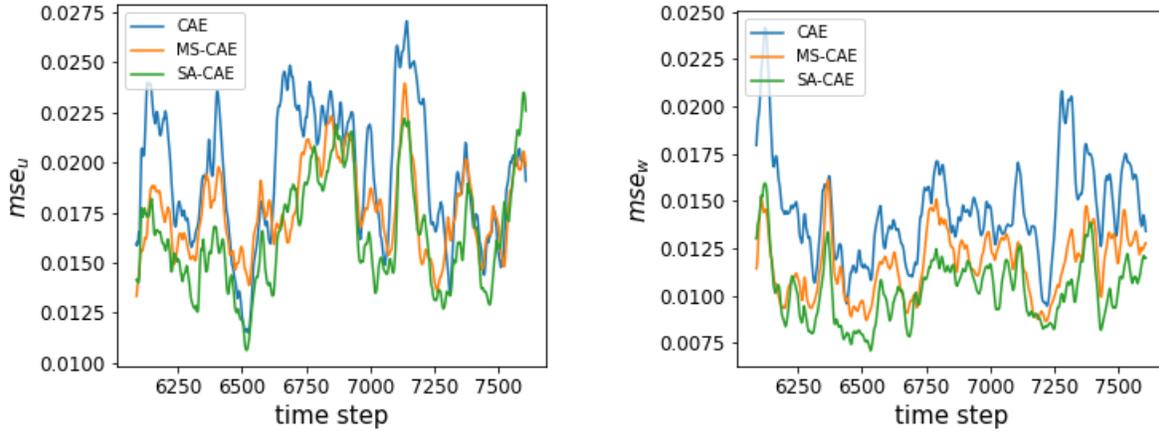


Figure 5.18: Time trace of mse of the reconstructed flow field (a) longitudinal velocity component, and (b) vertical velocity component

It can be concluded that considering long-range and multi-scale dependencies among datapoints, can improve the reconstruction ability of the model over the present dataset, without significantly increasing the reconstruction times. On the other hand, the training time may increase significantly. More experiments with more complicated datasets are required to understand the models' capabilities.

Table 5.3: The mean value of mse of the reconstruction of the velocity field components for different NIROMs

Model	CAE	MS-CAE	SA-CAE
mse_u	0.0194	0.0177	0.0164
mse_w	0.0145	0.0116	0.0103
$mse_{avg} = (mse_u + mse_w)/2$	0.0170	0.0147	0.0134
Training time (min)	80	150	165
Reconstruction time (s)	22	33	26

5.6 Conclusions and future works

Developing a NIROM consists of two steps: dimensionality reduction and feature dynamics computation. The former step seeks a low-dimensional representation of the original data, in which the dominant features of the high-fidelity data are present, while the latter computes the temporal dynamics of the obtained low-dimensional space. Different NIROMs were used to approximate the airflow field in the wake region of an isolated high-rise building under an unstable thermal

stratification condition. The developed NIROMs differ from each other in the dimensionality reduction stage. Three different techniques, namely, CAE, MS-CAE, and SA-CAE, were employed for the mentioned stage. The last two models were developed to overcome the limitation of convolutional layers arising from a fixed kernel size by expanding the receptive field. The temporal dynamics computation was also performed using parallel LSTM networks, which outperforms the conventional LSTM.

The capability of the dimensionality reduction parts and the whole models were assessed by comparing the results with available CFD data. It was shown that the reconstructed velocity components obtained from the dimensionality reduction parts are in good agreement with CFD results. The models' performance was almost similar in the longitudinal velocity component reconstruction. However, SA-CAE performed better in reconstructing the vertical velocity component, followed by MS-CAE and CAE.

Considering the whole NIROM model, the instantaneous velocity contours obtained from the models showed that the overall trends match the CFD results. However, there exist discrepancies between the reconstructed fields and the CFD results in the regions with intense gradients. Furthermore, information loss, regarding small-scale coherent structures, is visible in the models' predictions. The PSD function profiles of the reconstructed data showed that the models could capture the general spectrum shape. Although the models were able to predict the dominant frequencies at which peaks occur, discrepancies were reported in the PSD levels, particularly in the longitudinal velocity component. On the whole, SA-CAE performed better than the other two models in reconstructing the velocity field, followed by MS-CAE and CAE.

The present study developed the models over a dataset obtained from a CFD simulation in a simplified urban geometry. However, more experiments over more complicated datasets are required to reach a consistent conclusion on the models' performance. Furthermore, larger datasets may improve the model performance, particularly in calculating temporal dynamics, since the model can be trained over more vortex shedding cycles. Moreover, using physics-informed ROMs may improve the models' performance.

Chapter 6. Conclusions and outlook

6.1 Conclusions

Studying the impact of thermal stratification conditions on the flow and concentration fields within urban areas is of utmost importance. Due to frequent changes in meteorological conditions, it would be beneficial to capture these alterations to investigate their impact on the flow and concentration fields within an urban area. Furthermore, monitoring tasks may become important in studying the concentration field since a sudden release of pollutants upstream of an urban area may remarkably affect the residents' health. Moreover, having access to long-term simulations is of utmost importance in studying the ventilation performance of an urban area. Therefore, a model with low computational costs is required to be used for the aforementioned tasks. On the other hand, due to the significant importance of variables' fluctuations in the dispersion process (i.e., the impact of turbulent diffusion flux in pollutant transportation), the model should be able to capture the fluctuations of the flow and concentration fields. Thus, a comprehensive model must be fast while accounting for the complexities arising from the physics of the problem. CFD approaches have shown promising performance in investigating the impact of thermal stratification conditions on the urban airflow field and pollutant dispersion. Although CFD approaches, such as LES, can capture the fluctuations of the flow and concentration fields, their computational cost is a barrier to being used for the tasks mentioned earlier. In order to obviate the mentioned limitations, the present dissertation proposed a framework to generate fast models that are able to capture the complexities arising from the physics of the problem, i.e., turbulence. In order to do so, this thesis comprises two steps:

- (1) The CFD simulation step: in this step, after conducting CFD simulations, the importance of studying the impact of thermal stratification conditions on the flow and concentration fields within urban areas was addressed using different analyses. Afterward, the results provided by the CFD simulations were used to generate the aforementioned fast model.
- (2) The model order reduction step: in this step, model order reduction techniques were applied to the CFD results to generate a fast model. Afterward, the developed model was modified in order to account for the complexity of the problem (i.e., turbulence).

To capture the variable fluctuations, particularly under unstable conditions, complex, high-fidelity CFD approaches, such as LES, are more suitable than RANS-based ones for simulations. Using high-fidelity and unsteady approaches like LES provides the opportunity to capture the unsteady and turbulent nature of the flow field. However, conducting LES in urban areas, particularly in studies covering large areas, is complex and, in some cases, impractical due to its high computational costs. Thus, to make a balance between computational costs and accuracy, ELES was selected to conduct the CFD simulation in the present dissertation. ELES was used to simulate the flow and concentration fields under three different stratification conditions, namely, stable, isothermal, and unstable. The results of simulations, including the time-averaged flow, temperature and concentration fields, TKE, and concentration fluctuations, were compared with wind tunnel experimental data [164] and an LES study [103]. The findings showed that the time-averaged velocity field yielded by ELES was in good agreement with the wind tunnel and LES data.

However, overprediction was observed in the TKE results obtained by ELES in locations with strong shear layers. Regarding the time-averaged temperature field, findings showed that ELES has the capability to predict the results similar to LES, if it is set up properly. Also, ELES performance in predicting the time-averaged and fluctuating concentration fields was more promising than LES, particularly under the unstable condition, which reduced the relative error against the wind tunnel data by about 10.41%. It should be noted that the results showed that as TKE increased, the ELES performance became better than LES, which can be due to the difference between the locations of the turbulence generator in the models. Furthermore, the speed-up resulting from the use of ELES instead of LES was reported to be significant: 292% for the isothermal condition and 130% for the non-isothermal cases. These findings showed that ELES has the potential to be used in urban-related studies, since it can produce results similar to LES with less computational costs.

After validating the results with available experimental [164] data, the impact of thermal stratification conditions on the flow and concentration fields was thoroughly investigated, considering time-averaged and fluctuating variables, spectral analysis, and POD analysis. The results obtained from the simulations showed that the impact of the unstable condition was significant on the airflow field and TKE compared to the isothermal case. On the other hand, the mentioned variables were slightly changed under the stable condition in comparison with the isothermal case. Furthermore, it was reported that the unstable condition could effectively lower the pollutant concentration in the wake region of the building. The pollutant concentration under the unstable condition was 28.2% lower than that under the isothermal case. On the other hand, the pollutant concentration increased by about 33.9%, compared to the isothermal case, due to the occurrence of the stable condition. The mentioned effects of the non-isothermal conditions on the concentration field were due to the alteration of the convective and turbulent diffusion fluxes under the mentioned conditions, which resulted in changes in the pollutant dispersion pattern.

The spectral and POD analyses were conducted to investigate the impact of thermal stratification conditions on the turbulence structure of the flow field. The time history of the velocity fluctuations, in both the longitudinal and vertical directions, showed that the impact of flow instability on increasing TKE was significant. This finding was supported by the PSD profiles of the velocity fluctuations. Furthermore, the results of the spectral analysis showed that the recirculating region in the wake of the building becomes smaller with an increase in flow instability. Moreover, the PSD functions of the streamwise and vertical velocity components illustrated that an increase in the flow instability resulted in an increase in the vortex shedding around the building.

Furthermore, the turbulence structures of the POD modes in both the streamwise and vertical directions were remarkably changed under the unstable condition compared to the isothermal case. Based on the POD analysis of the flow field, it was mentioned that the first thirty POD modes contain 65.58% of the total TKE for the isothermal case. This value was changed for the stable and unstable cases to 80.89% and 81.39%, respectively. From this analysis, it can be interpreted that considering regions with the most dominant characteristics of the flow can largely represent the flow field's behavior. It can be concluded that considering the ranges of eddies with the most significant share of the energy of the flow results in a good approximation of the whole ranges.

Thus, low-dimensional structures (i.e., coherent structures) can be obtained, which is beneficial for describing a dynamical system.

In order to reduce the dimensionality of the dataset obtained from the CFD simulation, two different approaches can be performed: linear and non-linear dimensionality reduction techniques. Using a linear dimensionality reduction technique, such as POD, for finding a low-dimensional subspace, describing the whole system was reported to be ineffective for advection-dominated problems since they are strongly nonlinear. Therefore, this study employed nonlinear dimensionality reduction techniques to compute a low-dimensional representation of the original data. It should be noted that the results of the unstable case in the gradients-intense regions in the wake of the model building were used as the dataset for generating NIROMs. In order to reduce the dimensionality of the dataset, CAE-based methods were used as the baseline model to avoid over-fitting by taking advantage of the features of convolutional layers, namely, local connection and shared weights across the input domain. In addition to the baseline model, two other models, namely, MS-CAE and SA-CAE, were employed to account for the physics of the problem by capturing multi-scale and large-range dependencies among the datapoints. Low dimensional representations of the original dataset were obtained using the mentioned three methods.

After obtaining low-dimensional representations of the original dataset, the temporal dynamics of these subspaces should be computed. In order to capture the temporal evolution of the low-dimensional subspaces, a parallel LSTM network was used since it outperformed conventional LSTM networks in the present study.

The models' capabilities were assessed by comparing the results of both the dimensionality reduction part and the whole model to the previously validated unstable dataset. Regarding the dimensionality reduction part, the models' performance in reconstructing the longitudinal velocity component was slightly different from each other. Also, the results were in good agreement with the prediction made by a validated CFD model. However, the difference between the performance of the models in reconstructing the vertical velocity component was more pronounced. For both directions, SA-CAE performed better than the other models, followed by MS-CAE and CAE.

Considering the whole NIROMs, the instantaneous velocity contours captured the trends of the CFD results. However, in regions with high gradients, there exist discrepancies between the models' results and those from ELES. Furthermore, the time history of the velocity fluctuations showed significant errors when there are sudden changes through time. Therefore, it can be interpreted that some small-scale structures are missing. Also, the models underestimated the TKE predicted by the CFD simulation. However, this underestimation was less significant for the modified models compared to the baseline one. Moreover, based on the PSD functions, the models, particularly the modified ones, were able to capture the dominant frequencies at which peaks occur. The average mse , defined as $(mse_u + mse_w)/2$, showed that the overall performance of SA-CAE, with $mse_{avg} = 0.0134$, was better than the other two models, followed by MS-CAE ($mse_{avg} = 0.0147$) and CAE ($mse_{avg} = 0.0170$). Also, it should be noted that the mse for the models through time was reported to have an oscillatory behavior, but without any signs of growing. Thus, the models are useful for further predictions. Although the training time for the modified models was reported to be almost twice that of the baseline one, the reconstruction time

for the models was mentioned to be in the same order. On the whole, it seems that trying to consider the problem's physics can improve models' performance.

6.2 Contributions

The present dissertation first used a CFD model, i.e., ELES, to thoroughly analyze the impact of thermal stratification conditions on the airflow field, pollutant dispersion process, and turbulence statistics. It should be noted that ELES was employed for the first time to predict the mentioned field variables under thermal effects. This study is considered an incentive for researchers to take advantage of the characteristics of this model to conduct their research, particularly for cases with high turbulence intensity (e.g., unstable thermal conditions), since the model can provide results comparable to LES with less computational costs.

Afterward, the generated results were used to develop NIROMs for a highly-turbulent flow field. This study went beyond linear dimensionality reduction techniques, e.g., POD, which is the most common technique for this purpose in fluid dynamics problems such as urban-related ones. Furthermore, efforts were made to overcome the limitations of a conventional CAE, arising from the fixed kernel size, considering the underlying physics of the problem, i.e., multiscale and long-range dependencies among datapoints. Moreover, in order to improve the accuracy of the temporal dynamics computation, a parallel LSTM network was used to show its applicability in capturing the dynamics of highly-turbulent flow fields. Also, the speed-up achieved from using the models was reported to be approximately four orders of magnitude compared to ELES, making the models suitable for near real-time predictions.

The results were analyzed using both statistical and fluid dynamics viewpoints to shed light on the strengths and weaknesses of the models. Using this study, researchers can figure out and tackle the existing limitations of developing a NIROM for turbulent flow fields.

6.3 Limitations and future remarks

In order to investigate the impact of non-isothermal conditions on the airflow and concentration fields, and turbulence characteristics of the flow field, more realistic experiments are required. Although some experiments on the mentioned topic exist, most neglected critical aspects, such as realistic geometry, differential surface heating, pollutant dispersion, etc. Furthermore, data availability is a barrier to conducting comprehensive simulations in some cases. Due to the mentioned limitations, the present study used an isolated high-rise building as a representative of a structure within urban areas. Therefore, more experiments with other geometries, considering the complexities of a realistic urban area, are required to establish the present study's findings.

ELES has shown promising performance in simulating flow and concentration fields within urban areas under non-isothermal conditions in a fraction of the time required for conducting LES. However, as previously stated in section 3.2, no guidelines are available for conducting ELES in urban areas regarding the domain generation, the use of proper turbulence generation approach at

the URANS-LES interface, etc. Therefore, guidelines are needed for conducting ELES within urban areas.

The present study used the POD technique to analyze the turbulent flow structures by decomposing the flow field into a set of deterministic spatial functions and temporal coefficients. However, since the POD technique ranks its modes based on their kinetic energy, it contains different spectral information [326]. Thus, a continuous frequency spectrum is embedded within a POD mode. This characteristic of the POD technique may lead to the ignorance of flow structures with relatively small energy contents but with strong connections to other structures sharing the same frequency. The DMD method can solve this problem since, contrary to the POD modes, the DMD modes are based on their dynamics. Thus, each DMD mode contains a single frequency. The DMD technique, which decomposes the system into a series of approximated Koopman modes, was reported to be capable of capturing the spatial structures and dominant frequencies [327]. Another approach that can be used to decompose the flow field into fluctuation patterns is the spectral POD (SPOD) technique. SPOD is an advanced form of POD, considering spatio-temporal modes for statistically stationary flows [328]. Therefore, each mode oscillates at a single frequency. In other words, SPOD modes can be regarded as optimally averaged DMD modes [328]. However, they can consider the statistical variability of turbulent flows. It should be noted that each of the mentioned techniques is designed to account for a particular goal. Therefore, in order to fully analyze the urban airflow field and the pollutant dispersion mechanisms, these techniques can be used in tandem with each other.

Data was collected from structured computational grids for developing NIROMs in the present dissertation. However, using complex geometries for CFD simulations may restrict the use of structured grids. In order to properly use CAE-based methods for dimensionality reduction over unstructured grids, modifications should be made. Recently, some studies have focused on the mentioned topic [329,330]. Also, using graph convolution instead of the one used here, i.e., the Euclidean one, can be beneficial for performing the operations over unstructured grids.

Furthermore, NIROMs have usually been developed for small-scale fluid dynamics applications, as is the case with the present thesis. Using the existing approaches for more realistic geometries, such as a neighborhood or a city, may lead to the deterioration of the results. In order to use NIROMs for large-scale domains, it would be beneficial to divide them into multiple subdomains, using domain decomposition methods, and treat them separately. Afterward, an iteration-by-subdomain technique can be used to converge the solution over the whole domain [331].

Plenty of previous studies have used simple problems with low Reynolds number flows, which produce almost sinusoidal and slow-changing signals. Therefore, by using a sufficient amount of data, RNN methods can capture the dynamics of the systems. However, in a more complex problem, such as the one studied here, the temporal variations of the signals at each point are non-sinusoidal and fast-changing. In order to calculate the temporal dynamics of systems, several approaches were introduced in section 4.4. However, based on the literature review, LSTM is considered one of the most promising approaches for calculating the temporal dynamics of complicated systems. This study tried to improve the performance of the mentioned network by using a parallel form of it. Although the results were improved compared to the model using a

conventional LSTM, using more robust approaches are required. Using 3D CAE can result in the extraction of spatio-temporal characteristics embedded in the dataset [332]. Using discriminators in tandem with LSTM may also reduce the error arising from the temporal dynamics calculation. Also, probabilistic neural networks (PNNs) can be useful since they provide the model with a probabilistic posterior [48]. Using physics-informed learning algorithms may also improve the model's performance since it accounts for the residuals of the governing equations, which make the model robust to noisy input data.

Thus, possible future works based on the limitations of the present thesis can be summarized as follows:

- Preparing guidelines for using ELES in urban areas,
- Using decomposition methods other than POD to extract more detailed information about the turbulence structure of the flow field,
- Expanding the use of CAE-based dimensionality reduction techniques to unstructured grids,
- Applying the proposed methods to more complex domains using domain decomposition methods,
- Improving accuracy of temporal dynamics calculation by using discriminators in neural networks, PNNs, physics-informed learning algorithms, etc.

Bibliography

- [1] C. Sarrat, A. Lemonsu, V. Masson, D. Guedalia, Impact of urban heat island on regional atmospheric pollution, *Atmos. Environ.* 40 (2006) 1743–1758. <https://doi.org/10.1016/j.atmosenv.2005.11.037>.
- [2] C. Tsalicoglou, J. Allegrini, J. Carmeliet, Wind tunnel measurements of buoyancy-driven urban flows, in: *Proc. 18th Int. Symp. Flow Vis.*, 2018. <https://doi.org/10.3929/ETHZ-B-000279213>.
- [3] Q. Mu, M. Shiraiwa, M. Octaviani, N. Ma, A. Ding, H. Su, G. Lammel, U. Pöschl, Y. Cheng, Temperature effect on phase state and reactivity controls atmospheric multiphase chemistry and transport of PAHs, *Sci. Adv.* 4 (2018). <https://doi.org/10.1126/sciadv.aap7314>.
- [4] X. Xie, Z. Zhu, Effects of Heat Intensity and Inflow Wind on the Reactive Pollution Dispersion in Urban Street Canyon, *J. Shanghai Jiaotong Univ.* 23 (2018) 109–116. <https://doi.org/10.1007/s12204-018-2030-x>.
- [5] A.A. Aliabadi, M. Moradi, D. Clement, W.D. Lubitz, B. Gharabaghi, Flow and temperature dynamics in an urban canyon under a comprehensive set of wind directions, wind speeds, and thermal stability conditions, *Environ. Fluid Mech.* 19 (2019) 81–109. <https://doi.org/10.1007/s10652-018-9606-8>.
- [6] K. Niachou, I. Livada, M. Santamouris, Experimental study of temperature and airflow distribution inside an urban street canyon during hot summer weather conditions-Part I: Air and surface temperatures, *Build. Environ.* 43 (2008) 1383–1392. <https://doi.org/10.1016/j.buildenv.2007.01.039>.
- [7] M. Santamouris, N. Papanikolaou, I. Koronakis, I. Livada, D. Asimakopoulos, Thermal and air flow characteristics in a deep pedestrian canyon under hot weather conditions, *Atmos. Environ.* 33 (1999) 4503–4521. [https://doi.org/10.1016/S1352-2310\(99\)00187-9](https://doi.org/10.1016/S1352-2310(99)00187-9).
- [8] A.B. Barlag, W. Kuttler, The significance of country breezes for urban planning, *Energy Build.* 15 (1990) 291–297. [https://doi.org/10.1016/0378-7788\(90\)90001-Y](https://doi.org/10.1016/0378-7788(90)90001-Y).
- [9] S.J. Mei, J.T. Hu, D. Liu, F.Y. Zhao, Y. Li, H.Q. Wang, Thermal buoyancy driven canyon airflows inside the compact urban blocks saturated with very weak synoptic wind: Plume merging mechanism, *Build. Environ.* 131 (2018) 32–43. <https://doi.org/10.1016/j.buildenv.2017.12.035>.
- [10] P.A. Mirzaei, J. Carmeliet, Dynamical computational fluid dynamics modeling of the stochastic wind for application of urban studies, *Build. Environ.* 70 (2013) 161–170.

- <https://doi.org/10.1016/j.buildenv.2013.08.014>.
- [11] L. Yang, Y. Li, City ventilation of Hong Kong at no-wind conditions, *Atmos. Environ.* 43 (2009) 3111–3121. <https://doi.org/10.1016/j.atmosenv.2009.02.062>.
- [12] Y. Nakamura, T.R. Oke, Wind, temperature and stability conditions in an east-west oriented urban canyon, *Atmos. Environ.* 22 (1988) 2691–2700. [https://doi.org/10.1016/0004-6981\(88\)90437-4](https://doi.org/10.1016/0004-6981(88)90437-4).
- [13] Y. Zhao, L. Wai, A. Kubilay, J. Carmeliet, Isothermal and non-isothermal flow in street canyons: A review from theoretical, experimental and numerical perspectives, *Build. Environ.* 184 (2020) 107163. <https://doi.org/10.1016/j.buildenv.2020.107163>.
- [14] X. Xie, Z. Huang, J. Wang, Z. Xie, The impact of solar radiation and street layout on pollutant dispersion in street canyon, *Build. Environ.* 40 (2005) 201–212. <https://doi.org/10.1016/j.buildenv.2004.07.013>.
- [15] P.A. Mirzaei, F. Haghighat, Pollution removal effectiveness of the pedestrian ventilation system, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 46–58. <https://doi.org/10.1016/j.jweia.2010.10.007>.
- [16] Y. Lin, T. Ichinose, Y. Yamao, H. Mouri, Wind velocity and temperature fields under different surface heating conditions in a street canyon in wind tunnel experiments, *Build. Environ.* 168 (2020) 106500. <https://doi.org/10.1016/j.buildenv.2019.106500>.
- [17] B. Offerle, I. Eliasson, C.S.B. Grimmond, B. Holmer, Surface heating in relation to air temperature, wind and turbulence in an urban street canyon, *Boundary-Layer Meteorol.* 122 (2007) 273–292. <https://doi.org/10.1007/s10546-006-9099-8>.
- [18] A. Dallman, S. Magnusson, R. Britter, L. Norford, D. Entekhabi, H.J.S. Fernando, Conditions for thermal circulation in urban street canyons, *Build. Environ.* 80 (2014) 184–191. <https://doi.org/10.1016/j.buildenv.2014.05.014>.
- [19] P. Wang, D. Zhao, W. Wang, H. Mu, G. Cai, C. Liao, Thermal effect on pollutant dispersion in an urban street canyon, *Int. J. Environ. Res.* 5 (2011) 813–820. <https://doi.org/10.22059/ijer.2011.388>.
- [20] X. Liu, X. Wu, M. Wu, C. Shi, The impact of building surface temperature rise on airflow and cross-contamination around high-rise building, *Environ. Sci. Pollut. Res.* 27 (2020) 11855–11869. <https://doi.org/10.1007/s11356-020-07671-1>.
- [21] M.Y. Tsai, K.S. Chen, C.H. Wu, Three-dimensional modeling of air flow and pollutant dispersion in an urban street canyon with thermal effects, *J. Air Waste Manag. Assoc.* 55 (2005) 1178–1189. <https://doi.org/10.1080/10473289.2005.10464716>.
- [22] X.X. Li, R.E. Britter, T.Y. Koh, L.K. Norford, C.H. Liu, D. Entekhabi, D.Y.C. Leung,

- Large-Eddy Simulation of Flow and Pollutant Transport in Urban Street Canyons with Ground Heating, *Boundary-Layer Meteorol.* 137 (2010) 187–204. <https://doi.org/10.1007/s10546-010-9534-8>.
- [23] X.-X. Li, T.-Y. Koh, R. Britter, C.-H. Liu, L.K. Norford, D. Entekhabi, D.Y.C. Leung, Large-Eddy Simulation of Flow Field and Pollutant Dispersion in Urban Street Canyons Under Unstable Stratifications, in: *Seventh Int. Conf. Urban Clim.*, 2009.
- [24] G. Duan, K. Ngan, Sensitivity of turbulent flow around a 3-D building array to urban boundary-layer stability, *J. Wind Eng. Ind. Aerodyn.* 193 (2019) 103958. <https://doi.org/10.1016/j.jweia.2019.103958>.
- [25] F. Bazdidi-Tehrani, S. Masoumi-Verki, P. Gholamalipour, M. Kiamansouri, Large eddy simulation of pollutant dispersion in a naturally cross-ventilated model building: Comparison between sub-grid scale models, *Build. Simul.* 12 (2019) 921–941. <https://doi.org/10.1007/s12273-019-0525-5>.
- [26] K. Uehara, S. Murakami, S. Oikawa, S. Wakamatsu, Wind tunnel experiments on how thermal stratification affects flow in and above urban street canyons, *Atmos. Environ.* 34 (2000) 1553–1562. [https://doi.org/10.1016/S1352-2310\(99\)00410-0](https://doi.org/10.1016/S1352-2310(99)00410-0).
- [27] A. Kovar-Panskus, L. Moulinneuf, E. Savory, A. Abdelqari, J.-F. Sini, J.-M. Rosant, A. Robins, N. Toy, A Wind Tunnel Investigation of the Influence of Solar-Induced Wall-Heating on the Flow Regime within a Simulated Urban Street Canyon, *Water, Air, Soil Pollut. Focus.* 2 (2002) 555–571. https://doi.org/10.1007/978-94-010-0312-4_40.
- [28] M.A.J. de Nijs, J.D. Pietrzak, On total turbulent energy and the passive and active role of buoyancy in turbulent momentum and mass transfer, *Ocean Dyn.* 62 (2012) 849–865. <https://doi.org/10.1007/s10236-012-0536-6>.
- [29] M. Jadidi, F. Bazdidi-Tehrani, M. Kiamansouri, Embedded large eddy simulation approach for pollutant dispersion around a model building in atmospheric boundary layer, *Environ. Fluid Mech.* 16 (2016) 575–601. <https://doi.org/10.1007/s10652-016-9444-5>.
- [30] I. Kanda, Y. Yamao, Passive scalar diffusion in and above urban-like roughness under weakly stable and unstable thermal stratification conditions, *J. Wind Eng. Ind. Aerodyn.* 148 (2016) 18–33. <https://doi.org/10.1016/j.jweia.2015.11.002>.
- [31] B. Pulvirenti, S. Di Sabatino, CFD characterization of street canyon heating by solar radiation on building walls, in: *18th Int. Conf. Harmon. within Atmos. Dispers. Model. Regul. Purp.*, 2017: pp. 902–906.
- [32] W.C. Cheng, C.H. Liu, D.Y.C. Leung, On the correlation of air and pollutant exchange for street canyons in combined wind-buoyancy-driven flow, *Atmos. Environ.* 43 (2009) 3682–3690. <https://doi.org/10.1016/j.atmosenv.2009.04.054>.

- [33] D. Marucci, M. Carpentieri, Dispersion in an array of buildings in stable and convective atmospheric conditions, *Atmos. Environ.* 222 (2020) 117100. <https://doi.org/10.1016/j.atmosenv.2019.117100>.
- [34] E. Fedorovich, J. Thäter, A wind tunnel study of gaseous tracer dispersion in the convective boundary layer capped by a temperature inversion, *Atmos. Environ.* 36 (2002) 2245–2255. [https://doi.org/10.1016/S1352-2310\(02\)00185-1](https://doi.org/10.1016/S1352-2310(02)00185-1).
- [35] L. Davidson, *Fluid mechanics, turbulent flow and turbulence modeling*, 2017. <https://doi.org/10.1111/j.1365-3032.2005.00450.x>.
- [36] Y. Ogawa, P.G. Diosey, K. Uehara, H. Ueda, Wind tunnel observation of flow and diffusion under stable stratification, *Atmos. Environ.* 19 (1985) 65–74. [https://doi.org/10.1016/0004-6981\(85\)90136-2](https://doi.org/10.1016/0004-6981(85)90136-2).
- [37] O. Williams, T. Hohman, T. Van Buren, E. Bou-Zeid, A.J. Smits, The effect of stable thermal stratification on turbulent boundary layer statistics, *J. Fluid Mech.* 812 (2017) 1039–1075. <https://doi.org/10.1017/jfm.2016.781>.
- [38] X.X. Li, R. Britter, L.K. Norford, Effect of stable stratification on dispersion within urban street canyons: A large-eddy simulation, *Atmos. Environ.* 144 (2016) 47–59. <https://doi.org/10.1016/j.atmosenv.2016.08.069>.
- [39] Z.T. Xie, P. Hayden, C.R. Wood, Large-eddy simulation of approaching-flow stratification on dispersion over arrays of buildings, *Atmos. Environ.* 71 (2013) 64–74. <https://doi.org/10.1016/j.atmosenv.2013.01.054>.
- [40] Z. Shen, G. Cui, Z. Zhang, Turbulent dispersion of pollutants in urban-type canopies under stable stratification conditions, *Atmos. Environ.* 156 (2017) 1–14. <https://doi.org/10.1016/j.atmosenv.2017.02.017>.
- [41] J.M. Tomas, M.J.B.M. Pourquie, H.J.J. Jonker, Stable Stratification Effects on Flow and Pollutant Dispersion in Boundary Layers Entering a Generic Urban Environment, *Boundary-Layer Meteorol.* 159 (2016) 221–239. <https://doi.org/10.1007/s10546-015-0124-7>.
- [42] D. Guo, P. Zhao, R. Wang, R. Yao, J. Hu, Numerical simulations of the flow field and pollutant dispersion in an idealized urban area under different atmospheric stability conditions, *Process Saf. Environ. Prot.* 136 (2020) 310–323. <https://doi.org/10.1016/j.psep.2020.01.031>.
- [43] D. Marucci, M. Carpentieri, Effect of local and upwind stratification on flow and dispersion inside and above a bi-dimensional street canyon, *Build. Environ.* 156 (2019) 74–88. <https://doi.org/10.1016/j.buildenv.2019.04.013>.
- [44] V. Sessa, Z.X. Steven, Thermal Stratification Effects on Turbulence and Dispersion in

- Internal and External Boundary Layers, *Boundary-Layer Meteorol.* 176 (2020) 61–83. <https://doi.org/10.1007/s10546-020-00524-x>.
- [45] S. Xiang, J. Zhou, X. Fu, L. Zheng, Y. Wang, Y. Zhang, K. Yi, J. Liu, J. Ma, S. Tao, Fast simulation of high resolution urban wind fields at city scale, *Urban Clim.* 39 (2021) 100941. <https://doi.org/10.1016/j.uclim.2021.100941>.
- [46] G. Calzolari, W. Liu, Deep learning to replace, improve, or aid CFD analysis in built environment applications: A review, *Build. Environ.* 206 (2021) 108315. <https://doi.org/10.1016/j.buildenv.2021.108315>.
- [47] S.L. Brunton, B.R. Noack, P. Koumoutsakos, Machine Learning for Fluid Mechanics, *Annu. Rev. Fluid Mech.* 52 (2020) 477–508. <https://doi.org/10.1146/annurev-fluid-010719-060214>.
- [48] R. Maulik, K. Fukami, N. Ramachandra, K. Fukagata, K. Taira, Probabilistic neural networks for fluid flow surrogate modeling and data recovery, *Phys. Rev. Fluids.* 5 (2020) 104401. <https://doi.org/10.1103/PhysRevFluids.5.104401>.
- [49] M. Morimoto, K. Fukami, K. Fukagata, Experimental velocity data estimation for imperfect particle images using machine learning, *Phys. Fluids.* 33 (2021) 087121. <https://doi.org/10.1063/5.0060760>.
- [50] X. Zhang, T. Ji, F. Xie, H. Zheng, Y. Zheng, Unsteady flow prediction from sparse measurements by compressed sensing reduced order modeling, *Comput. Methods Appl. Mech. Eng.* 393 (2022) 114800. <https://doi.org/10.1016/j.cma.2022.114800>.
- [51] P. Torres, S. Le Clainche, R. Vinuesa, On the experimental, numerical and data-driven methods to study urban flows, *Energies.* 14 (2021) 1310–1348. <https://doi.org/10.3390/en14051310>.
- [52] M. Schatzmann, B. Leidl, Issues with validation of urban flow and dispersion CFD models, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 169–186. <https://doi.org/10.1016/j.jweia.2011.01.005>.
- [53] M.W. Rotach, R. Vogt, C. Bernhofer, E. Batchvarova, A. Christen, A. Clappier, B. Feddersen, S.E. Gryning, G. Martucci, H. Mayer, V. Mitev, T.R. Oke, E. Parlow, H. Richner, M. Roth, Y.A. Roulet, D. Ruffieux, J.A. Salmond, M. Schatzmann, J.A. Voogt, BUBBLE - An urban boundary layer meteorology project, *Theor. Appl. Climatol.* 81 (2005) 231–261. <https://doi.org/10.1007/s00704-004-0117-9>.
- [54] G. Chen, X. Yang, H. Yang, J. Hang, Y. Lin, X. Wang, Q. Wang, Y. Liu, The influence of aspect ratios and solar heating on flow and ventilation in 2D street canyons by scaled outdoor experiments, *Build. Environ.* 185 (2020) 107159. <https://doi.org/10.1016/j.buildenv.2020.107159>.

- [55] K. Niachou, I. Livada, M. Santamouris, Experimental study of temperature and airflow distribution inside an urban street canyon during hot summer weather conditions. Part II: Airflow analysis, *Build. Environ.* 43 (2008) 1393–1403. <https://doi.org/10.1016/j.buildenv.2007.01.040>.
- [56] G. Vachon, J.M. Rosant, P. Mestayer, P. Louka, J.F. Sini, D. Delaunay, M.-J. Antoine, F. Ducroz, J. Garreau, R. Griffiths, C. Jones, Y. Loren, F. Molle, J.-P. Peneau, Y. Tetard, M. Violleau, Experimental investigation of pollutant dispersion within a street in low wind conditions, the experiment Nantes '99, in: 9th Int. Sci. Symp. Transp. Air Pollut., 2000.
- [57] P. Louka, G. Vachon, J. Sini, P.G. Mestayer, J. Rosant, Thermal effects on the airflow in a street canyon–Nantes'99 experimental results and model simulations, *Water, Air, Soil Pollut.* 2 (2002) 351–364.
- [58] P. Klein, J. V. Clark, Flow variability in a North American downtown street Canyon, *J. Appl. Meteorol. Climatol.* 46 (2007) 851–877. <https://doi.org/10.1175/JAM2494.1>.
- [59] D. Zajic, H.J.S. Fernando, R. Calhoun, M. Princevac, M.J. Brown, E.R. Pardyjak, Flow and turbulence in an urban canyon, *J. Appl. Meteorol. Climatol.* 50 (2011) 203–223. <https://doi.org/10.1175/2010JAMC2525.1>.
- [60] P.M. Klein, J.M. Galvez, Flow and turbulence characteristics in a suburban street canyon, *Environ. Fluid Mech.* 15 (2015) 419–438. <https://doi.org/10.1007/s10652-014-9352-5>.
- [61] T.R. Oke, Street design and urban canopy layer climate, *Energy Build.* 11 (1988) 103–113. [https://doi.org/10.1016/0378-7788\(88\)90026-6](https://doi.org/10.1016/0378-7788(88)90026-6).
- [62] T.R. Oke, *Boundary layer climates*, Second Edi, Routledge, 1987.
- [63] Z. Shen, B. Wang, G. Cui, Z. Zhang, Flow pattern and pollutant dispersion over three dimensional building arrays, *Atmos. Environ.* 116 (2015) 202–215. <https://doi.org/10.1016/j.atmosenv.2015.06.022>.
- [64] J. Nichol, Remote Sensing of Urban Heat Islands by Day and Night, *Photogramm. Eng. Remote Sens.* 71 (2005) 613–621.
- [65] A. Gupta, *Physical Modeling of the Downwash Effect of Rooftop Structures on Plume Dispersion*, Concordia University, 2009.
- [66] J. Allegrini, V. Dorer, J. Carmeliet, Wind tunnel measurements of buoyant flows in street canyons, *Build. Environ.* 59 (2013) 315–326. <https://doi.org/10.1016/j.buildenv.2012.08.029>.
- [67] X. Huang, Y. Huang, N. Xu, Y. Luo, P. Cui, Thermal effects on the dispersion of rooftop stack emission in the wake of a tall building within suburban areas by wind-tunnel experiments, *J. Wind Eng. Ind. Aerodyn.* 205 (2020) 104295.

<https://doi.org/10.1016/j.jweia.2020.104295>.

- [68] C. Shu, L. (Leon) Wang, M. Mortezaadeh, Dimensional analysis of Reynolds independence and regional critical Reynolds numbers for urban aerodynamics, *J. Wind Eng. Ind. Aerodyn.* 203 (2020) 104232. <https://doi.org/10.1016/j.jweia.2020.104232>.
- [69] K. Richards, M. Schatzmann, B. Leidl, Wind tunnel experiments modelling the thermal effects within the vicinity of a single block building with leeward wall heating, *J. Wind Eng. Ind. Aerodyn.* 94 (2006) 621–636. <https://doi.org/10.1016/j.jweia.2006.02.003>.
- [70] W.H. Snyder, Similarity criteria for the application of fluid models to the study of air pollution meteorology, *Boundary-Layer Meteorol.* 3 (1972) 113–134. <https://doi.org/10.1007/BF00769111>.
- [71] P.Y. Cui, Z. Li, W.Q. Tao, Numerical investigations on Re-independence for the turbulent flow and pollutant dispersion under the urban boundary layer with some experimental validations, *Int. J. Heat Mass Transf.* 106 (2017) 422–436. <https://doi.org/10.1016/j.ijheatmasstransfer.2016.08.038>.
- [72] B. Ruck, Wind-Tunnel Measurements of Flow Field Characteristics Around a Heated Model Building, *J. Wind Eng. Ind. Aerodyn.* 50 (1993) 139–152.
- [73] J. Allegrini, A wind tunnel study on three-dimensional buoyant flows in street canyons with different roof shapes and building lengths, *Build. Environ.* 143 (2018) 71–88. <https://doi.org/10.1016/j.buildenv.2018.06.056>.
- [74] R. Yoshie, G. Jiang, T. Shirasawa, J. Chung, CFD simulations of gas dispersion around high-rise building in non-isothermal boundary layer, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 279–288. <https://doi.org/10.1016/j.jweia.2011.01.006>.
- [75] P.Y. Cui, Z. Li, W.Q. Tao, Buoyancy flows and pollutant dispersion through different scale urban areas: CFD simulations and wind-tunnel measurements, *Build. Environ.* 104 (2016) 76–91. <https://doi.org/10.1016/j.buildenv.2016.04.028>.
- [76] S. Fellini, L. Ridolfi, P. Salizzoni, Street canyon ventilation: Combined effect of cross-section geometry and wall heating, *Q. J. R. Meteorol. Soc.* 146 (2020) 2347–2367. <https://doi.org/10.1002/qj.3795>.
- [77] T. Hu, R. Yoshie, Effect of atmospheric stability on air pollutant concentration and its generalization for real and idealized urban block models based on field observation data and wind tunnel experiments, *J. Wind Eng. Ind. Aerodyn.* 207 (2020) 104380. <https://doi.org/10.1016/j.jweia.2020.104380>.
- [78] C. Gromke, B. Ruck, Influence of trees on the dispersion of pollutants in an urban street canyon-Experimental investigation of the flow and concentration field, *Atmos. Environ.* 41 (2007) 3287–3302. <https://doi.org/10.1016/j.atmosenv.2006.12.043>.

- [79] C. Gromke, R. Buccolieri, S. Di Sabatino, B. Ruck, Dispersion study in a street canyon with tree planting by means of wind tunnel and numerical investigations - Evaluation of CFD data with experimental data, *Atmos. Environ.* 42 (2008) 8640–8650. <https://doi.org/10.1016/j.atmosenv.2008.08.019>.
- [80] C. Gromke, A vegetation modeling concept for building and environmental aerodynamics wind tunnel tests and its application in pollutant dispersion studies, *Environ. Pollut.* 159 (2011) 2094–2099. <https://doi.org/10.1016/j.envpol.2010.11.012>.
- [81] Y. Lin, T. Ichinose, R.T. Wu, Y. Yamao, H. Mouri, R. V Rivera-Virtudazo, An Experimental Study on Exploring the Possibility of Applying Artificial Light as Radiation in Wind Tunnel, *J. Heat Isl. Inst. Int.* 9 (2014) 108–112.
- [82] J.-F. Sini, S. Anquetint, P.G. Mestayer, Pollutant dispersion and thermal effects in urban street canyons, *Atmos. Environ.* 30 (1996) 2659–2677.
- [83] B.E. Launder, D.B. Spalding, The numerical computation of turbulent flows, *Comput. Methods Appl. Mech. Eng.* 3 (1974) 269–289. [https://doi.org/10.1016/0045-7825\(74\)90029-2](https://doi.org/10.1016/0045-7825(74)90029-2).
- [84] J.-J. Kim, J.-J. Baik, A Numerical Study of Thermal Effects on Flow and Pollutant Dispersion in Urban Street Canyons, *J. Appl. Meteorol.* 38 (1999) 1249–1261.
- [85] X. Xie, Z. Huang, J. Wang, Z. Xie, Thermal effects on vehicle emission dispersion in an urban street canyon, *Transp. Res. Part D.* 10 (2005) 197–212. <https://doi.org/10.1016/j.trd.2005.01.002>.
- [86] X. Xie, C.H. Liu, D.Y.C. Leung, Impact of building facades and ground heating on wind flow and pollutant transport in street canyons, *Atmos. Environ.* 41 (2007) 9030–9049. <https://doi.org/10.1016/j.atmosenv.2007.08.027>.
- [87] V. Yakhot, S.A. Orszag, S. Thangam, T.B. Gatski, C.G. Speziale, Development of turbulence models for shear flows by a double expansion technique, *Phys. Fluids A.* 4 (1992) 1510–1520. <https://doi.org/10.1063/1.858424>.
- [88] H. Huang, R. Ooka, H. Chen, S. Kato, T. Takahashi, T. Watanabe, CFD analysis on traffic-induced air pollutant dispersion under non-isothermal condition in a complex urban area in winter, *J. Wind Eng. Ind. Aerodyn.* 96 (2008) 1774–1788. <https://doi.org/10.1016/j.jweia.2008.02.010>.
- [89] J. Smagorinsky, General Circulation Experiments With the Primitive Equations I. The Basic Experiment, *Mon. Weather Rev.* 91 (1963).
- [90] X. Cai, Effects of differential wall heating in street canyons on dispersion and ventilation characteristics of a passive scalar, *Atmos. Environ.* 51 (2012) 268–277. <https://doi.org/10.1016/j.atmosenv.2012.01.010>.

- [91] X.-X. Li, R.E. Britter, L.K. Norford, T.-Y. Koh, D. Entekhabi, Flow and Pollutant Transport in Urban Street Canyons of Different Aspect Ratios with Ground Heating: Large-Eddy Simulation, *Boundary-Layer Meteorol.* 142 (2012) 289–304. <https://doi.org/10.1007/s10546-011-9670-9>.
- [92] N. Zhang, Y. Kang, K. Zhong, J. Liu, Numerical modeling of the flow and pollutant dispersion in street canyons with ground thermal effect, *Appl. Mech. Mater.* 548–549 (2014) 601–606. <https://doi.org/10.4028/www.scientific.net/AMM.548-549.601>.
- [93] T.-H. Shih, W.W. Liou, A. Shabbir, Z. Yang, J. Zhu, A new $k-\epsilon$ eddy viscosity model for high reynolds number turbulent flows, *Comput. Fluids.* 24 (1995) 227–238. [https://doi.org/10.1016/0045-7930\(94\)00032-T](https://doi.org/10.1016/0045-7930(94)00032-T).
- [94] Z. Tan, J. Dong, Y. Xiao, J. Tu, A numerical study of diurnally varying surface temperature on flow patterns and pollutant dispersion in street canyons, *Atmos. Environ.* 104 (2015) 217–227. <https://doi.org/10.1016/j.atmosenv.2015.01.027>.
- [95] L. Lin, J. Hang, X. Wang, X. Wang, S. Fan, Q. Fan, Y. Liu, Integrated Effects of Street Layouts and Wall Heating on Vehicular Pollutant Dispersion and their Reentry Toward Downstream Canyons, *Aerosol Air Qual. Res.* 16 (2016) 3142–3163. <https://doi.org/10.4209/aaqr.2016.04.0148>.
- [96] T. Heus, C.C. Van Heerwaarden, H.J.J. Jonker, A. Pier Siebesma, S. Axelsen, K. Van Den Dries, O. Geoffroy, A.F. Moene, D. Pino, S.R. De Roode, J.V.G. De Arellano, Formulation of the Dutch Atmospheric Large-Eddy Simulation (DALES) and overview of its applications, *Geosci. Model Dev.* 3 (2010) 415–444. <https://doi.org/10.5194/gmd-3-415-2010>.
- [97] S.H. Hosseini, P. Ghobadi, T. Ahmadi, J.K. Calautit, Numerical investigation of roof heating impacts on thermal comfort and air quality in urban canyons, *Appl. Therm. Eng.* 123 (2017) 310–326. <https://doi.org/10.1016/j.applthermaleng.2017.05.095>.
- [98] C. Meneveau, T.S. Lund, W.H. Cabot, A Lagrangian dynamic subgrid-scale model of turbulence, *J. Fluid Mech.* 319 (1996) 353–385. <https://doi.org/10.1017/S0022112096007379>.
- [99] G. Jiang, R. Yoshie, Large-eddy simulation of flow and pollutant dispersion in a 3D urban street model located in an unstable boundary layer, *Build. Environ.* 142 (2018) 47–57. <https://doi.org/10.1016/j.buildenv.2018.06.015>.
- [100] N. Nazarian, A. Martilli, J. Kleissl, Impacts of Realistic Urban Heating, Part I: Spatial Variability of Mean Flow, Turbulent Exchange and Pollutant Dispersion, *Boundary-Layer Meteorol.* 166 (2018) 367–393. <https://doi.org/10.1007/s10546-017-0311-9>.
- [101] B. Maronga, M. Gryschka, R. Heinze, F. Hoffmann, F. Kanani-Sühring, M. Keck, K. Ketelsen, M.O. Letzel, M. Sühring, S. Raasch, The Parallelized Large-Eddy Simulation

- Model (PALM) version 4.0 for atmospheric and oceanic flows: Model formulation, recent developments, and future perspectives, *Geosci. Model Dev.* 8 (2015) 2515–2551. <https://doi.org/10.5194/gmd-8-2515-2015>.
- [102] N. Nazarian, A. Martilli, L. Norford, J. Kleissl, Impacts of Realistic Urban Heating. Part II: Air Quality and City Breathability, *Boundary-Layer Meteorol.* 168 (2018) 321–341. <https://doi.org/10.1007/s10546-018-0346-6>.
- [103] F. Bazdidi-Tehrani, P. Gholamalipour, M. Kiamansouri, M. Jadidi, Large eddy simulation of thermal stratification effect on convective and turbulent diffusion fluxes concerning gaseous pollutant dispersion around a high-rise model building, *J. Build. Perform. Simul.* 12 (2019) 97–116. <https://doi.org/10.1080/19401493.2018.1486886>.
- [104] M. Germano, U. Piomelli, P. Moin, W.H. Cabot, A dynamic subgrid-scale eddy viscosity model, *Phys. Fluids A.* 3 (1991) 1760–1765. <https://doi.org/10.1063/1.857955>.
- [105] D.K. Lilly, A proposed modification of the Germano subgrid-scale closure method, *Phys. Fluids A.* 4 (1992) 633–635. <https://doi.org/10.1063/1.858280>.
- [106] J. Hang, R. Buccolieri, X. Yang, H. Yang, F. Quarta, B. Wang, Impact of indoor – outdoor temperature differences on dispersion of gaseous pollutant and particles in idealized street canyons with and without viaduct settings, *Build. Simul.* 12 (2019) 285–297.
- [107] G. Jiang, T. Hu, H. Yang, Effects of ground heating on ventilation and pollutant transport in three-dimensional urban street canyons with unit aspect ratio, *Atmosphere (Basel)*. 10 (2019). <https://doi.org/10.3390/atmos10050286>.
- [108] C.-W. Liu, S.-J. Mei, D. Liu, F.-Y. Zhao, Indoor and Built Convective dispersion of heat and airborne pollutants inside street canyons under the influence of urban ground heat flows, *Indoor Built Environ.* 28 (2019) 619–635. <https://doi.org/10.1177/1420326X17706186>.
- [109] S. Mei, J. Hu, D. Liu, F. Zhao, Y. Li, H. Wang, Airborne pollutant dilution inside the deep street canyons subjecting to thermal buoyancy driven flows : Effects of representative urban skylines, *Build. Environ.* 149 (2019) 592–606. <https://doi.org/10.1016/j.buildenv.2018.12.050>.
- [110] J. Hang, X. Chen, G. Chen, T. Chen, Y. Lin, Z. Luo, X. Zhang, Q. Wang, The influence of aspect ratios and wall heating conditions on flow and passive pollutant exposure in 2D typical street canyons, *Build. Environ.* 168 (2020) 106536. <https://doi.org/10.1016/j.buildenv.2019.106536>.
- [111] Y.D. Huang, N. Xu, S.Q. Ren, L.B. Qian, P.Y. Cui, Numerical investigation of the thermal effect on flow and dispersion of rooftop stack emissions with wind tunnel experimental validations, *Environ. Sci. Pollut. Res.* 28 (2021) 11618–11636. <https://doi.org/10.1007/s11356-020-11304-y>.

- [112] Z. Li, H. Zhang, C. Wen, A. Yang, Y. Juan, Effects of frontal area density on outdoor thermal comfort and air quality, *Build. Environ.* 180 (2020) 107028.
- [113] Z. Li, H. Zhang, C.Y. Wen, A.S. Yang, Y.H. Juan, Effects of height-asymmetric street canyon configurations on outdoor air temperature and air quality, *Build. Environ.* 183 (2020) 107195. <https://doi.org/10.1016/j.buildenv.2020.107195>.
- [114] M.F. Yassin, N.S. Alhajeri, A.A. Elmi, M.J. Malek, M. Shalash, Numerical simulation of gas dispersion from rooftop stacks on buildings in urban environments under changes in atmospheric thermal stability, *Environ. Monit. Assess.* 193 (2021). <https://doi.org/10.1007/s10661-020-08798-x>.
- [115] G. Chen, L. Rong, G. Zhang, Comparison of urban airflow between solar-induced thermal wall and uniform wall temperature boundary conditions by coupling CitySim and CFD, *Build. Environ.* 172 (2020) 106732. <https://doi.org/10.1016/j.buildenv.2020.106732>.
- [116] F. Bazdidi-Tehrani, P. Gholamalipour, M. Kiamansouri, M. Jadidi, Large eddy simulation of thermal stratification effect on convective and turbulent diffusion fluxes concerning gaseous pollutant dispersion around a high-rise model building, *J. Build. Perform. Simul.* 12 (2019) 97–116. <https://doi.org/10.1080/19401493.2018.1486886>.
- [117] J. Allegrini, J. Carmeliet, Simulations of local heat islands in Zürich with coupled CFD and building energy models, *Urban Clim.* 24 (2018) 340–359. <https://doi.org/10.1016/j.uclim.2017.02.003>.
- [118] J. Hang, Y. Li, M. Sandberg, R. Buccolieri, S. Di Sabatino, The influence of building height variability on pollutant dispersion and pedestrian ventilation in idealized high-rise urban areas, *Build. Environ.* 56 (2012) 346–360. <https://doi.org/10.1016/j.buildenv.2012.03.023>.
- [119] V.B.L. Boppana, Z.-T. Xie, I.P. Castro, Large-Eddy Simulation of Dispersion from Surface Sources in Arrays of Obstacles, *Boundary-Layer Meteorol.* 135 (2010) 433–454. <https://doi.org/10.1007/s10546-010-9489-9>.
- [120] Y. Toparlar, B. Blocken, P. Vos, G.J.F. Van Heijst, W.D. Janssen, T. Van Hooff, H. Montazeri, H.J.P. Timmermans, CFD simulation and validation of urban microclimate: A case study for Bergpolder Zuid, Rotterdam, *Build. Environ.* 83 (2015) 79–90. <https://doi.org/10.1016/j.buildenv.2014.08.004>.
- [121] Y. Tominaga, Visualization of city breathability based on CFD technique: case study for urban blocks in Niigata City, *J. Vis.* 15 (2012) 269–276. <https://doi.org/10.1007/s12650-012-0128-z>.
- [122] J. Yang, B. Shi, Y. Shi, S. Marvin, Y. Zheng, G. Xia, Air pollution dispersal in high density urban areas: Research on the triadic relation of wind, air pollution, and urban form, *Sustain. Cities Soc.* 54 (2020) 101941. <https://doi.org/10.1016/j.scs.2019.101941>.

- [123] Y. Tominaga, T. Stathopoulos, Numerical simulation of dispersion around an isolated cubic building: Comparison of various types of k- ϵ models, *Atmos. Environ.* 43 (2009) 3200–3210. <https://doi.org/10.1016/j.atmosenv.2009.03.038>.
- [124] Y. Tominaga, T. Stathopoulos, Numerical simulation of dispersion around an isolated cubic building: Model evaluation of RANS and LES, *Build. Environ.* 45 (2010) 2231–2239. <https://doi.org/10.1016/j.buildenv.2010.04.004>.
- [125] A. Dejoan, J.L. Santiago, A. Martilli, F. Martin, A. Pinelli, Comparison Between Large-Eddy Simulation and Reynolds-Averaged Navier – Stokes Computations for the MUST Field Experiment. Part I: Study of the Flow for an Incident Wind Directed Perpendicularly to the Front Array of Containers, *Boundary-Layer Meteorol.* 135 (2010) 133–150. <https://doi.org/10.1007/s10546-010-9466-3>.
- [126] S.M. Salim, R. Buccolieri, A. Chan, S. Di Sabatino, Numerical simulation of atmospheric pollutant dispersion in an urban street canyon: Comparison between RANS and LES, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 103–113. <https://doi.org/10.1016/j.jweia.2010.12.002>.
- [127] F. Bazdidi-Tehrani, M. Jadidi, Large eddy simulation of dispersion around an isolated cubic building: Evaluation of localized dynamic kSGS-equation sub-grid scale model, *Environ. Fluid Mech.* 14 (2014) 565–589. <https://doi.org/10.1007/s10652-013-9316-1>.
- [128] M. Jadidi, F. Bazdidi-Tehrani, M. Kiamansouri, Scale-adaptive simulation of unsteady flow and dispersion around a model building: spectral and POD analyses, *J. Build. Perform. Simul.* 11 (2018) 241–260. <https://doi.org/10.1080/19401493.2017.1326165>.
- [129] B. Blocken, LES over RANS in building simulation for outdoor and indoor applications: A foregone conclusion?, 2018. <https://doi.org/10.1007/s12273-018-0459-3>.
- [130] K.B. Shah, J.H. Ferziger, A fluid mechanics view of wind engineering: Large eddy simulation of flow past a cubic obstacle, *J. Wind Eng. Ind. Aerodyn.* 67–68 (1997) 211–224. [https://doi.org/10.1016/S0167-6105\(97\)00074-3](https://doi.org/10.1016/S0167-6105(97)00074-3).
- [131] F. Bazdidi-tehrani, S. Masoumi-verki, P. Gholamalipour, Impact of opening shape on airflow and pollutant dispersion in a wind-driven cross-ventilated model building: Large eddy simulation, *Sustain. Cities Soc.* 61 (2020) 102196. <https://doi.org/10.1016/j.scs.2020.102196>.
- [132] Y. Tominaga, T. Stathopoulos, CFD modeling of pollution dispersion in a street canyon: Comparison between LES and RANS, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 340–348. <https://doi.org/10.1016/j.jweia.2010.12.005>.
- [133] P. Gousseau, B. Blocken, T. Stathopoulos, G.J.F. Van Heijst, CFD simulation of near-field pollutant dispersion on a high-resolution grid: A case study by LES and RANS for a building group in downtown Montreal, *Atmos. Environ.* 45 (2011) 428–438. <https://doi.org/10.1016/j.atmosenv.2010.09.065>.

- [134] F.R. Menter, Y. Egorov, The scale-adaptive simulation method for unsteady turbulent flow predictions. part 1: Theory and model description, *Flow, Turbul. Combust.* 85 (2010) 113–138. <https://doi.org/10.1007/s10494-010-9264-5>.
- [135] Y. Dai, C.M. Mak, Z. Ai, J. Hang, Evaluation of computational and physical parameters influencing CFD simulations of pollutant dispersion in building arrays, *Build. Environ.* 137 (2018) 90–107. <https://doi.org/10.1016/j.buildenv.2018.04.005>.
- [136] L.W. Chew, L.R. Glicksman, L.K. Norford, Buoyant flows in street canyons: Comparison of RANS and LES at reduced and full scales, *Build. Environ.* 146 (2018) 77–87. <https://doi.org/10.1016/j.buildenv.2018.09.026>.
- [137] N. Nazarian, J. Kleissl, Realistic solar heating in urban areas: Air exchange and street-canyon ventilation, *Build. Environ.* 95 (2016) 75–93. <https://doi.org/10.1016/j.buildenv.2015.08.021>.
- [138] Y. Cheng, F.S. Lien, E. Yee, R. Sinclair, A comparison of large Eddy simulations with a standard $k - \epsilon$ Reynolds-averaged Navier – Stokes model for the prediction of a fully developed turbulent flow over a matrix of cubes, *J. Wind Eng. Ind. Aerodyn.* 91 (2003) 1301–1328. <https://doi.org/10.1016/j.jweia.2003.08.001>.
- [139] J. Allegrini, V. Dorer, T. Defraeye, J. Carmeliet, An adaptive temperature wall function for mixed convective flows at exterior surfaces of buildings in street canyons, *Build. Environ.* 49 (2012) 55–66. <https://doi.org/10.1016/j.buildenv.2011.09.025>.
- [140] T. Defraeye, B. Blocken, J. Carmeliet, CFD analysis of convective heat transfer at the surfaces of a cube immersed in a turbulent boundary layer, *Int. J. Heat Mass Transf.* 53 (2010) 297–308. <https://doi.org/10.1016/j.ijheatmasstransfer.2009.09.029>.
- [141] T. Defraeye, B. Blocken, J. Carmeliet, CFD simulation of heat transfer at surfaces of bluff bodies in turbulent boundary layers: Evaluation of a forced-convective temperature wall function for mixed convection, *J. Wind Eng. Ind. Aerodyn.* 104–106 (2012) 439–446. <https://doi.org/10.1016/j.jweia.2012.02.001>.
- [142] T. Defraeye, B. Blocken, J. Carmeliet, An adjusted temperature wall function for turbulent forced convective heat transfer for bluff bodies in the atmospheric boundary layer, *Build. Environ.* 46 (2011) 2130–2141. <https://doi.org/10.1016/j.buildenv.2011.04.013>.
- [143] J. Allegrini, V. Dorer, J. Carmeliet, Buoyant flows in street canyons: Validation of CFD simulations with wind tunnel measurements, *Build. Environ.* 72 (2014) 63–74. <https://doi.org/10.1016/j.buildenv.2013.10.021>.
- [144] P.R. Spalart, S.R. Allmaras, A One-Equation Turbulence Model for Aerodynamic Flows, in: *AIAA-92-0439*, 1992: p. 23.
- [145] P.R. Spalart, S. Deck, M.L. Shur, K.D. Squires, M.K. Strelets, A. Travin, A new version

- of detached-eddy simulation, resistant to ambiguous grid densities, *Theor. Comput. Fluid Dyn.* 20 (2006) 181–195. <https://doi.org/10.1007/s00162-006-0015-0>.
- [146] E. Labourasse, P. Sagaut, Reconstruction of turbulent fluctuations using a hybrid RANS/LES approach, *J. Comput. Phys.* 182 (2002) 301–336. <https://doi.org/10.1006/jcph.2002.7169>.
- [147] P.R. Spalart, W.-H. Jou, M. Strelets, S.R. Allmaras, Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach, in: *Proc. First AFOSR Int. Conf. DNS/LES*, Louisiana. Greyden Press. 4–8 Aug, 1997: pp. 137–147.
- [148] M. Shur, P.R. Spalart, M. Strelets, A. Travin, Detached-eddy simulation of an airfoil at high angle of attack, in: *Proc. 4th Int. Symp. Eng. Turbul. Model. Meas.* Ajaccio, Corsica, Fr. 24–26 May, Woodhead Publishing Limited, 1999: pp. 669–678. <https://doi.org/10.1016/b978-008043328-8/50064-3>.
- [149] P. Quéméré, P. Sagaut, V. Couailler, A new multi-domain/multi-resolution method for large-eddy simulation, *Int. J. Numer. Methods Fluids.* 36 (2001) 391–416. <https://doi.org/10.1002/flid.134>.
- [150] J. Fröhlich, D. von Terzi, Hybrid LES/RANS methods for the simulation of turbulent flows, *Prog. Aerosp. Sci.* 44 (2008) 349–377. <https://doi.org/10.1016/j.paerosci.2008.05.001>.
- [151] H. Foroutan, W. Tang, D.K. Heist, S.G. Perry, L.H. Brouwer, E.M. Monbureau, Numerical analysis of pollutant dispersion around elongated buildings: An embedded large eddy simulation approach, *Atmos. Environ.* 187 (2018) 117–130. <https://doi.org/10.1016/j.atmosenv.2018.05.053>.
- [152] B. Yang, J. Gu, K.M. Zhang, Parameterization of the building downwash and sidewash effect using a mixture model, *Build. Environ.* 172 (2020) 106694. <https://doi.org/10.1016/j.buildenv.2020.106694>.
- [153] J. Liu, J. Niu, CFD simulation of the wind environment around an isolated high-rise building: An evaluation of SRANS, LES and DES models, *Build. Environ.* 96 (2016) 91–106. <https://doi.org/10.1016/j.buildenv.2015.11.007>.
- [154] J. Paik, F. Sotiropoulos, F. Porté-agel, Detached eddy simulation of flow around two wall-mounted cubes in tandem, *Int. J. Heat Fluid Flow.* 30 (2009) 286–305. <https://doi.org/10.1016/j.ijheatfluidflow.2009.01.006>.
- [155] K.E. Kakosimos, M.J. Assael, Application of Detached Eddy Simulation to neighbourhood scale gases atmospheric dispersion modelling, *J. Hazard. Mater.* 261 (2013) 653–668. <https://doi.org/10.1016/j.jhazmat.2013.08.018>.
- [156] M. Lateb, C. Masson, T. Stathopoulos, C. Bedard, Simulation of near-field dispersion of

- pollutants using detached-eddy simulation, *Comput. Fluids*. 100 (2014) 308–320. <https://doi.org/10.1016/j.compfluid.2014.05.024>.
- [157] J. Liu, J. Niu, Delayed detached eddy simulation of pedestrian-level wind around a building array – The potential to save computing resources, *Build. Environ.* 152 (2019) 28–38. <https://doi.org/10.1016/j.buildenv.2019.02.011>.
- [158] B. Chaouat, R. Schiestel, Hybrid RANS/LES simulations of the turbulent flow over periodic hills at high Reynolds number using the PITM method, *Comput. Fluids*. 84 (2013) 279–300. <https://doi.org/10.1016/j.compfluid.2013.06.012>.
- [159] F.R. Menter, J. Schütze, M. Gritskevich, Global vs. Zonal Approaches in Hybrid RANS-LES Turbulence Modelling. In: Fu S., Haase W., Peng SH., Schwamborn D. (eds) *Progress in Hybrid RANS-LES Modelling, Notes Numer. Fluid Mech. Multidiscip. Des.* 117 (2012) 15–28. https://doi.org/10.1007/978-3-642-31818-4_2.
- [160] F. Mathey, D. Cokljat, Zonal Multi-Domain RANS/LES Simulation of Airflow Over the Ahmed Body, *Eng. Turbul. Model. Exp.* 6. (2005) 647–656. <https://doi.org/10.1016/B978-008044544-1/50062-5>.
- [161] E.M. Monbureau, D.K. Heist, S.G. Perry, L.H. Brouwer, H. Foroutan, W. Tang, Enhancements to AERMOD’s building downwash algorithms based on wind-tunnel and Embedded-LES modeling, *Atmos. Environ.* 179 (2018) 321–330. <https://doi.org/10.1016/j.atmosenv.2018.02.022>.
- [162] E.M. Monbureau, D.K. Heist, S.G. Perry, W. Tang, Modeling lateral plume deflection in the wake of an elongated building, *Atmos. Environ.* 234 (2020) 117608. <https://doi.org/10.1016/j.atmosenv.2020.117608>.
- [163] M. Jadidi, F. Bazdidi-Tehrani, M. Kiamansouri, Dynamic sub-grid scale turbulent Schmidt number approach in large eddy simulation of dispersion around an isolated cubical building, *Build. Simul.* 9 (2016) 183–200. <https://doi.org/10.1007/s12273-015-0257-0>.
- [164] http://www.wind.arch.t-kougei.ac.jp/info_center/pollution/Non-Isothermal_Flow.html, (2021).
- [165] R. Yoshie, Wind Tunnel Experiment and Large Eddy Simulation of Pollutant/Thermal Dispersion in Non-isothermal Turbulent Boundary Layer, in: *Adv. Environ. Wind Eng.*, 2016: pp. 167–196. https://doi.org/10.1007/978-4-431-55912-2_9.
- [166] R. Yoshie, G. Jiang, T. Shirasawa, J. Chung, CFD simulations of gas dispersion around high-rise building in non-isothermal boundary layer, *J. Wind Eng. Ind. Aerodyn.* 99 (2011) 279–288. <https://doi.org/10.1016/j.jweia.2011.01.006>.
- [167] Z.T. Ai, C.M. Mak, Large-eddy Simulation of flow and dispersion around an isolated

- building: Analysis of influencing factors, *Comput. Fluids*. 118 (2015) 89–100.
<https://doi.org/10.1016/j.compfluid.2015.06.006>.
- [168] Y. Dai, C.M. Mak, Z. Ai, J. Hang, Evaluation of computational and physical parameters influencing CFD simulations of pollutant dispersion in building arrays, *Build. Environ.* 137 (2018) 90–107. <https://doi.org/10.1016/j.buildenv.2018.04.005>.
- [169] H. Werner, H. Wengle, Large-Eddy Simulation of Turbulent Flow Over and Around a Cube in a Plate Channel. In: Durst F., Friedrich R., Launder B.E., Schmidt F.W., Schumann U., Whitelaw J.H. (eds), *Turbul. Shear Flows* 8. (1993) 155–168.
https://doi.org/10.1007/978-3-642-77674-8_12.
- [170] L. Davidson, How to estimate the resolution of an LES of recirculating flow. In: Salvetti M., Geurts B., Meyers J., Sagaut P. (eds) *Quality and Reliability of Large-Eddy Simulations II*, in: ERCOFTAC Ser., Springer, Dordrecht, 2011: pp. 269–286.
https://doi.org/10.1007/978-94-007-0231-8_25.
- [171] H. Kikumoto, R. Ooka, Large-eddy simulation of pollutant dispersion in a cavity at fine grid resolutions, *Build. Environ.* 127 (2018) 127–137.
<https://doi.org/10.1016/j.buildenv.2017.11.005>.
- [172] F. Bazdidi-Tehrani, A. Ghafouri, M. Jadidi, Grid resolution assessment in large eddy simulation of dispersion around an isolated cubic building, *J. Wind Eng. Ind. Aerodyn.* 121 (2013) 1–15. <https://doi.org/10.1016/j.jweia.2013.07.003>.
- [173] F. Mathey, D. Cokljat, J.P. Bertoglio, E. Sergent, Assessment of the vortex method for Large Eddy Simulation inlet conditions, *Prog. Comput. Fluid Dyn. An Int. J.* 6 (2006).
<https://doi.org/10.1504/PCFD.2006.009483>.
- [174] B.W. Yan, Q.S. Li, Inflow turbulence generation methods with large eddy simulation for wind effects on tall buildings, *Comput. Fluids*. 116 (2015) 158–175.
<https://doi.org/10.1016/j.compfluid.2015.04.020>.
- [175] G. Jiang, R. Yoshie, T. Shirasawa, X. Jin, Inflow turbulence generation for large eddy simulation in non-isothermal boundary layers, *J. Wind Eng. Ind. Aerodyn.* 104–106 (2012) 369–378. <https://doi.org/10.1016/j.jweia.2012.02.030>.
- [176] F. Bazdidi-Tehrani, M. Kiamansouri, M. Jadidi, Inflow turbulence generation techniques for large eddy simulation of flow and dispersion around a model building in a turbulent atmospheric boundary layer, *J. Build. Perform. Simul.* 9 (2016) 680–698.
<https://doi.org/10.1080/19401493.2016.1196729>.
- [177] X. Zheng, H. Montazeri, B. Blocken, CFD simulations of wind flow and mean surface pressure for buildings with balconies: Comparison of RANS and LES, *Build. Environ.* 173 (2020) 106747. <https://doi.org/10.1016/j.buildenv.2020.106747>.

- [178] P. Gousseau, B. Blocken, T. Stathopoulos, G.J.F. van Heijst, Near-field pollutant dispersion in an actual urban area: Analysis of the mass transport mechanism by high-resolution Large Eddy Simulations, *Comput. Fluids*. 114 (2015) 151–162. <https://doi.org/10.1016/j.compfluid.2015.02.018>.
- [179] H.K. Versteeg, W. Malalasekera, *An Introduction to Computational Fluid Dynamics. The Finite Volume Method*, Second Edi, 2007. <https://doi.org/10.2514/1.22547>.
- [180] P. Gousseau, B. Blocken, G.J.F. Van Heijst, Quality assessment of Large-Eddy Simulation of wind flow around a high-rise building: Validation and solution verification, *Comput. Fluids*. 79 (2013) 120–133. <https://doi.org/10.1016/j.compfluid.2013.03.006>.
- [181] N. Ikegaya, T. Okaze, H. Kikumoto, M. Imano, H. Ono, Y. Tominaga, Effect of the numerical viscosity on reproduction of mean and turbulent flow fields in the case of a 1:1:2 single block model, *J. Wind Eng. Ind. Aerodyn.* 191 (2019) 279–296. <https://doi.org/10.1016/j.jweia.2019.06.013>.
- [182] T. Okaze, H. Kikumoto, H. Ono, M. Imano, N. Ikegaya, T. Hasama, K. Nakao, T. Kishida, Y. Tabata, K. Nakajima, R. Yoshie, Y. Tominaga, Large-eddy simulation of flow around an isolated building: A step-by-step analysis of influencing factors on turbulent statistics, *Build. Environ.* (2021) 108021. <https://doi.org/10.1016/j.buildenv.2021.108021>.
- [183] M. Schatzmann, H. Olesen, J. Franke, *Cost 732 Model Evaluation Case Studies : Approach and Results*, Brussels, Belgium, 2010.
- [184] Y. Tominaga, T. Stathopoulos, CFD simulations of near-field pollutant dispersion with different plume buoyancies, *Build. Environ.* 131 (2018) 128–139. <https://doi.org/10.1016/j.buildenv.2018.01.008>.
- [185] VDI Guideline 3783 Part 9, *Environmental Meteorology Prognostic Microscale Wind Field Model Evaluation for Flow Around Buildings and Obstacles*, 2005.
- [186] X. Zhou, A. Ying, B. Cong, H. Kikumoto, R. Ooka, L. Kang, H. Hu, Large eddy simulation of the effect of unstable thermal stratification on airflow and pollutant dispersion around a rectangular building, *J. Wind Eng. Ind. Aerodyn.* 211 (2021) 104526. <https://doi.org/10.1016/j.jweia.2021.104526>.
- [187] B. Zhang, R. Ooka, H. Kikumoto, Identification of three-dimensional flow features around a square-section building model via spectral proper orthogonal decomposition, *Phys. Fluids*. 33 (2021). <https://doi.org/10.1063/5.0041395>.
- [188] D.Y.C. Leung, Outdoor-indoor air pollution in urban environment: Challenges and opportunity, *Front. Environ. Sci.* 2 (2015) 1–7. <https://doi.org/10.3389/fenvs.2014.00069>.
- [189] W.H.O. (WHO), [www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](http://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health), (2021).

- [190] Z. Li, T. Ming, S. Liu, C. Peng, R. de Richter, W. Li, H. Zhang, C.Y. Wen, Review on pollutant dispersion in urban areas-part A: Effects of mechanical factors and urban morphology, *Build. Environ.* 190 (2021) 107534. <https://doi.org/10.1016/j.buildenv.2020.107534>.
- [191] Y. Zhang, Z. Gu, C.W. Yu, Impact Factors on Airflow and Pollutant Dispersion in Urban Street Canyons and Comprehensive Simulations: a Review, *Curr. Pollut. Reports.* 6 (2020) 425–439. <https://doi.org/10.1007/s40726-020-00166-0>.
- [192] K. Ahmad, M. Khare, K.K. Chaudhry, Wind tunnel simulation studies on dispersion at urban street canyons and intersections - A review, *J. Wind Eng. Ind. Aerodyn.* 93 (2005) 697–717. <https://doi.org/10.1016/j.jweia.2005.04.002>.
- [193] P. Moonen, T. Defraeye, V. Dorer, B. Blocken, J. Carmeliet, Urban Physics: Effect of the micro-climate on comfort, health and energy demand, *Front. Archit. Res.* 1 (2012) 197–228. <https://doi.org/10.1016/j.foar.2012.05.002>.
- [194] Y. Tominaga, T. Stathopoulos, CFD simulation of near-field pollutant dispersion in the urban environment: A review of current modeling techniques, *Atmos. Environ.* 79 (2013) 716–730. <https://doi.org/10.1016/j.atmosenv.2013.07.028>.
- [195] A.W.M. Yazid, nor A.C. Sidik, S.M. Salim, K.M. Saqr, A review on the flow structure and pollutant dispersion in urban street canyons for urban planning strategies, *Simulation.* 90 (2014) 892–916. <https://doi.org/10.1177/0037549714528046>.
- [196] M. Lateb, R.N. Meroney, M. Yataghene, H. Fellouah, F. Saleh, M.C. Boufadel, On the use of numerical modelling for near-field pollutant dispersion in urban environments - A review, *Environ. Pollut.* 208 (2016) 271–283. <https://doi.org/10.1016/j.envpol.2015.07.039>.
- [197] K. V. Abhijith, P. Kumar, J. Gallagher, A. McNabola, R. Baldauf, F. Pilla, B. Broderick, S. Di Sabatino, B. Pulvirenti, Air pollution abatement performances of green infrastructure in open road and built-up street canyon environments – A review, *Atmos. Environ.* 162 (2017) 71–86. <https://doi.org/10.1016/j.atmosenv.2017.05.014>.
- [198] Y. Toparlar, B. Blocken, B. Maiheu, G.J.F. van Heijst, A review on the CFD analysis of urban microclimate, *Renew. Sustain. Energy Rev.* 80 (2017) 1613–1640. <https://doi.org/10.1016/j.rser.2017.05.248>.
- [199] R. Buccolieri, J. Santiago, E. Rivas, B. Sanchez, Review on urban tree modelling in CFD simulations: Aerodynamic, deposition and thermal effects, *Urban For. Urban Green.* 31 (2018) 212–220. <https://doi.org/10.1016/j.ufug.2018.03.003>.
- [200] Y. Zhang, Z. Gu, C.W. Yu, Review on numerical simulation of airflow and pollutant dispersion in urban street canyons under natural background wind condition, *Aerosol Air Qual. Res.* 18 (2018) 780–789. <https://doi.org/10.4209/aaqr.2017.09.0303>.

- [201] P.A. Mirzaei, CFD modeling of micro and urban climates: Problems to be solved in the new decade, *Sustain. Cities Soc.* 69 (2021). <https://doi.org/10.1016/j.scs.2021.102839>.
- [202] Z. Li, J. Xu, T. Ming, C. Peng, J. Huang, T. Gong, Numerical Simulation on the Effect of Vehicle Movement on Pollutant Dispersion in Urban Street, *Procedia Eng.* 205 (2017) 2303–2310. <https://doi.org/10.1016/j.proeng.2017.10.104>.
- [203] Q. Wang, W. Fang, R. de Richter, C. Peng, T. Ming, Effect of moving vehicles on pollutant dispersion in street canyon by using dynamic mesh updating method, *J. Wind Eng. Ind. Aerodyn.* 187 (2019) 15–25. <https://doi.org/10.1016/j.jweia.2019.01.014>.
- [204] C. Cai, T. Ming, W. Fang, R. de Richter, C. Peng, The effect of turbulence induced by different kinds of moving vehicles in street canyons, *Sustain. Cities Soc.* 54 (2020) 102015. <https://doi.org/10.1016/j.scs.2020.102015>.
- [205] T. Shi, T. Ming, Y. Wu, C. Peng, Y. Fang, R. De_Richter, The effect of exhaust emissions from a group of moving vehicles on pollutant dispersion in the street canyons, *Build. Environ.* 181 (2020) 107120. <https://doi.org/10.1016/j.buildenv.2020.107120>.
- [206] H. Woodward, M. Stettler, D. Pavlidis, E. Aristodemou, H. ApSimon, C. Pain, A large eddy simulation of the dispersion of traffic emissions by moving vehicles at an intersection, *Atmos. Environ.* 215 (2019) 116891. <https://doi.org/10.1016/j.atmosenv.2019.116891>.
- [207] W. Li, Y. He, Y. Zhang, J. Su, C. Chen, C.W. Yu, R. Zhang, Z. Gu, LES simulation of flow field and pollutant dispersion in a street canyon under time-varying inflows with TimeVarying-SIMPLE approach, *Build. Environ.* 157 (2019) 185–196. <https://doi.org/10.1016/j.buildenv.2019.04.049>.
- [208] Y.W. Zhang, Z.L. Gu, Y. Cheng, S.C. Lee, Effect of real-time boundary wind conditions on the air flow and pollutant dispersion in an urban street canyon-Large eddy simulations, *Atmos. Environ.* 45 (2011) 3352–3359. <https://doi.org/10.1016/j.atmosenv.2011.03.055>.
- [209] S. Masoumi-Verki, P. Gholamalipour, F. Haghighat, U. Eicker, Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses, *Build. Environ.* 206 (2021) 108388. <https://doi.org/10.1016/j.buildenv.2021.108388>.
- [210] J. Na, K. Jeon, W.B. Lee, Toxic gas release modeling for real-time analysis using variational autoencoder with convolutional neural networks, *Chem. Eng. Sci.* 181 (2018) 68–78. <https://doi.org/10.1016/j.ces.2018.02.008>.
- [211] R. Vinuesa, S.L. Brunton, The Potential of Machine Learning to Enhance Computational Fluid Dynamics, *ArXiv:2110.02085 [Physics.Flu-Dyn]*. (2021). <http://arxiv.org/abs/2110.02085>.

- [212] P. Gousseau, B. Blocken, G.J.F. Van Heijst, CFD simulation of pollutant dispersion around isolated buildings : On the role of convective and turbulent mass fluxes in the prediction accuracy, *J. Hazard. Mater.* 194 (2011) 422–434. <https://doi.org/10.1016/j.jhazmat.2011.08.008>.
- [213] Y. Tominaga, T. Stathopoulos, CFD Modeling of Pollution Dispersion in Building Array: Evaluation of turbulent scalar flux modeling in RANS model using LES results, *J. Wind Eng. Ind. Aerodyn.* 104–106 (2012) 484–491. <https://doi.org/10.1016/j.jweia.2012.02.004>.
- [214] S.M. Salim, K.C. Ong, Performance of RANS, URANS and LES in the Prediction of Airflow and Pollutant Dispersion., in: Kim H., Ao SI., Rieger B. IAENG Trans. Eng. Technol. Lect. Notes Electr. Eng. Vol 170. Springer, Dordr., 2013. https://doi.org/10.1007/978-94-007-4786-9_21.
- [215] H.E. Fiedler, Coherent structures in turbulent flows, *Prog. Aerosp. Sci.* 25 (1988) 231–269. [https://doi.org/10.1016/0376-0421\(88\)90001-2](https://doi.org/10.1016/0376-0421(88)90001-2).
- [216] S.B. Pope, *Turbulent flows*, Cambridge University Press, New York, 2013. [https://doi.org/10.1016/S0169-8095\(01\)00076-X](https://doi.org/10.1016/S0169-8095(01)00076-X).
- [217] J. Borggaard, A. Duggleby, A. Hay, T. Iliescu, Z. Wang, Reduced-order modeling of turbulent flows, in: 18th Int. Symp. Math. Theory Networks Syst., Blacksburg, Virginia, 2008.
- [218] J.L. Lumley, The structure of inhomogeneous turbulent flows, *Atmos. Turbul. Radio Wave Propag.* (1967) 166–178.
- [219] L. Sirovich, Turbulence and dynamics of coherent structures. I. Coherent structures, *Q. Appl. Math.* 45 (1987) 561–571. <https://doi.org/10.1090/qam/910462>.
- [220] P. Holmes, J.L. Lumley, G. Berkooz, C.W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Second, Cambridge University Press, Cambridge, 2012. <https://doi.org/10.1017/cbo9780511622700>.
- [221] D. Xiao, C.E. Heaney, L. Mottet, F. Fang, W. Lin, I.M. Navon, Y. Guo, O.K. Matar, A.G. Robins, C.C. Pain, A reduced order model for turbulent flows in the urban environment using machine learning, *Build. Environ.* 148 (2019) 323–337. <https://doi.org/10.1016/j.buildenv.2018.10.035>.
- [222] Y. Du, B. Blocken, S. Pirker, A novel approach to simulate pollutant dispersion in the built environment: Transport-based recurrence CFD, *Build. Environ.* 170 (2020) 106604. <https://doi.org/10.1016/j.buildenv.2019.106604>.
- [223] U. Baur, P. Benner, L. Feng, Model Order Reduction for Linear and Nonlinear Systems: A System-Theoretic Perspective, *Arch. Comput. Methods Eng.* 21 (2014) 331–358. <https://doi.org/10.1007/s11831-014-9111-2>.

- [224] D. Xiao, F. Fang, J. Zheng, C.C. Pain, I.M. Navon, Machine learning-based rapid response tools for regional air pollution modelling, *Atmos. Environ.* 199 (2019) 463–473.
- [225] S. Xiang, X. Fu, J. Zhou, Y. Wang, Y. Zhang, X. Hu, J. Xu, H. Liu, J. Liu, J. Ma, S. Tao, Non-intrusive reduced order model of urban airflow with dynamic boundary conditions, *Build. Environ.* 187 (2021) 107397. <https://doi.org/10.1016/j.buildenv.2020.107397>.
- [226] D. Xiao, Non-intrusive reduced order models and their applications, Imperial College, 2016.
- [227] Z. Wang, D. Xiao, F. Fang, R. Govindan, C.C. Pain, Y. Guo, Model identification of reduced order fluid dynamics systems using deep learning, *Int. J. Numer. Methods Fluids.* 86 (2018) 255–268. <https://doi.org/10.1002/fld.4416>.
- [228] S.K. Star, B. Sanderse, G. Stabile, G. Rozza, J. Degroote, Reduced order models for the incompressible Navier-Stokes equations on collocated grids using a ‘discretize-then-project’ approach, *Int. J. Numer. Methods Fluids.* 93 (2021) 2694–2722. <https://doi.org/10.1002/fld.4994>.
- [229] J.N. Kutz, S.L. Brunton, B.W. Brunton, J.L. Proctor, Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems, SIAM, 2016. <https://doi.org/10.1137/1.9781611974508>.
- [230] P.J. Schmid, Dynamic mode decomposition of numerical and experimental data, *J. Fluid Mech.* 656 (2010) 5–28. <https://doi.org/10.1017/S0022112010001217>.
- [231] C. Prud’homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods, *J. Fluids Eng.* 124 (2002) 70–80. <https://doi.org/10.1115/1.1448332>.
- [232] D. Galbally, K. Fidkowski, K. Willcox, O. Ghattas, Non-linear model reduction for uncertainty quantification in large-scale inverse problems, *Int. J. Numer. Methods Eng.* 81 (2010) 1581–1608. <https://doi.org/10.1002/nme.2746>.
- [233] S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM J. Sci. Comput.* 32 (2010) 2737–2764. <https://doi.org/10.1137/090766498>.
- [234] S.J. Cao, J. Meyers, On the construction and use of linear low-dimensional ventilation models, *Indoor Air.* 22 (2012) 427–441. <https://doi.org/10.1111/j.1600-0668.2012.00771.x>.
- [235] S.J. Cao, J. Meyers, Fast prediction of indoor pollutant dispersion based on reduced-order ventilation models, *Build. Simul.* 8 (2015) 415–420. <https://doi.org/10.1007/s12273-015-0240-9>.

- [236] L. Vervecken, J. Camps, J. Meyers, Stable reduced-order models for pollutant dispersion in the built environment, *Build. Environ.* 92 (2015) 360–367. <https://doi.org/10.1016/j.buildenv.2015.05.008>.
- [237] F. Fang, T. Zhang, D. Pavlidis, C.C. Pain, A.G. Buchan, I.M. Navon, Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons, *Atmos. Environ.* 96 (2014) 96–106. <https://doi.org/10.1016/j.atmosenv.2014.07.021>.
- [238] N.H. Nguyen, A rapid simulation of nano-particle transport in a two-dimensional human airway using POD/Galerkin reduced-order models, *Int. J. Numer. Methods Eng.* 105 (2016) 514–531. <https://doi.org/10.1002/nme.4986>.
- [239] R. Chakir, Y. Maday, A two-grid nite-element/reduced basis scheme for the approximation of the solution of parameter dependent PDE, in: *9e Colloq. Natl. En Calc. Des Struct. CSMA*, 2009.
- [240] S. Ullmann, M. Rotkvic, J. Lang, POD-Galerkin reduced-order modeling with adaptive finite element snapshots, *J. Comput. Phys.* 325 (2016) 244–258. <https://doi.org/10.1016/j.jcp.2016.08.018>.
- [241] S. Star, G. Stabile, G. Rozza, J. Degroote, A POD-Galerkin reduced order model of a turbulent convective buoyant flow of sodium over a backward-facing step, *Appl. Math. Model.* 89 (2021) 486–503. <https://doi.org/10.1016/j.apm.2020.07.029>.
- [242] S. Hijazi, S. Ali, G. Stabile, F. Ballarin, G. Rozza, The Effort of Increasing Reynolds Number in Projection-Based Reduced Order Methods: From Laminar to Turbulent Flows. In: van Brummelen H., Corsini A., Perotto S., Rozza G. (eds) *Numerical Methods for Flows.*, *Lect. Notes Comput. Sci. Eng.* 132 (2020) 245–264. https://doi.org/10.1007/978-3-030-30705-9_22.
- [243] L.P. Franca, S.L. Frey, Stabilized finite element methods: II. The incompressible Navier-Stokes equations, *Comput. Methods Appl. Mech. Eng.* 99 (1992) 209–233. [https://doi.org/10.1016/0045-7825\(92\)90041-H](https://doi.org/10.1016/0045-7825(92)90041-H).
- [244] S.K. Star, *Development of Reduced Order Modeling Methods for Incompressible Flows with Heat Transfer and Parametric Boundary Conditions*, Universiteit Gent, 2021.
- [245] S. Lorenzi, A. Cammi, L. Luzzi, G. Rozza, POD-Galerkin method for finite volume approximation of Navier–Stokes and RANS equations, *Comput. Methods Appl. Mech. Eng.* 311 (2016) 151–179. <https://doi.org/10.1016/j.cma.2016.08.006>.
- [246] E.J. Parish, C.R. Wentland, K. Duraisamy, The Adjoint Petrov–Galerkin method for non-linear model reduction, *Comput. Methods Appl. Mech. Eng.* 365 (2020) 112991. <https://doi.org/10.1016/j.cma.2020.112991>.

- [247] F. Fang, C.C. Pain, I.M. Navon, A.H. Elsheikh, J. Du, D. Xiao, Non-linear petrov-galerkin methods for reduced order hyperbolic equations and discontinuous finite element methods, *J. Comput. Phys.* 234 (2013) 540–559. <https://doi.org/10.1016/j.jcp.2012.10.011>.
- [248] R. Reyes, R. Codina, Projection-based reduced order models for flow problems: A variational multiscale approach, *Comput. Methods Appl. Mech. Eng.* 363 (2020) 112844. <https://doi.org/10.1016/j.cma.2020.112844>.
- [249] D. Xiao, F. Fang, J. Du, C.C. Pain, I.M. Navon, A.G. Buchan, A.H. ElSheikh, G. Hu, Non-linear Petrov-Galerkin methods for reduced order modelling of the Navier-Stokes equations using a mixed finite element pair, *Comput. Methods Appl. Mech. Eng.* 255 (2013) 147–157. <https://doi.org/10.1016/j.cma.2012.11.002>.
- [250] D. Sipp, M. Fosas de Pando, P.J. Schmid, Nonlinear model reduction: A comparison between POD-Galerkin and POD-DEIM methods, *Comput. Fluids.* 208 (2020) 104628. <https://doi.org/10.1016/j.compfluid.2020.104628>.
- [251] D. Xiao, F. Fang, A.G. Buchan, C.C. Pain, I.M. Navon, J. Du, G. Hu, Non-linear model reduction for the Navier-Stokes equations using residual DEIM method, *J. Comput. Phys.* 263 (2014) 1–18. <https://doi.org/10.1016/j.jcp.2014.01.011>.
- [252] J. Östh, B.R. Noack, S. Krajnović, D. Barros, J. Borée, On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body, *J. Fluid Mech.* 747 (2014) 518–544. <https://doi.org/10.1017/jfm.2014.168>.
- [253] D. Xiao, F. Fang, A.G. Buchan, C.C. Pain, I.M. Navon, A. Muggeridge, Non-intrusive reduced order modelling of the Navier – Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 293 (2015) 522–541. <https://doi.org/10.1016/j.cma.2015.05.015>.
- [254] D. Xiao, C.E. Heaney, F. Fang, L. Mottet, R. Hu, D.A. Bistran, E. Aristodemou, A domain decomposition non-intrusive reduced order model for turbulent flows, *Comput. Fluids.* 182 (2019) 15–27. <https://doi.org/10.1016/j.compfluid.2019.02.012>.
- [255] D. Xiao, F. Fang, C.E. Heaney, I.M. Navon, C.C. Pain, A domain decomposition method for the non-intrusive reduced order modelling of fluid flow, *Comput. Methods Appl. Mech. Eng.* 354 (2019) 307–330. <https://doi.org/10.1016/j.cma.2019.05.039>.
- [256] S. Ding, R. Yang, Reduced-order modelling of urban wind environment and gaseous pollutants dispersion in an urban-scale street canyon, *J. Saf. Sci. Resil.* 2 (2021) 238–245. <https://doi.org/10.1016/j.jnlssr.2021.09.001>.
- [257] C. Quilodrán-Casas, R. Arcucci, C. Pain, Y. Guo, Adversarially trained LSTMs on reduced order models of urban air pollution simulations, *ArXiv:2101.01568 [Cs.LG]*. (2021). <http://arxiv.org/abs/2101.01568>.

- [258] C. Quilodrán-Casas, R. Arcucci, L. Mottet, Y. Guo, C. Pain, Adversarial autoencoders and adversarial LSTM for improved forecasts of urban air pollution simulations, *ArXiv:2104.06297 [Cs.LG]*. (2021). <http://arxiv.org/abs/2104.06297>.
- [259] A. Tallet, C. Allery, F. Allard, POD approach to determine in real-time the temperature distribution in a cavity, *Build. Environ.* 93 (2015) 34–49. <https://doi.org/10.1016/j.buildenv.2015.07.007>.
- [260] J. Ding, S.J. Cao, Identification of zonal pollutant diffusion characteristics using dynamic mode decomposition: Towards the deployment of sensors, *Build. Environ.* 206 (2021) 108379 Contents. <https://doi.org/10.1016/j.buildenv.2021.108379>.
- [261] Z. Wu, D. Laurence, S. Utyuzhnikov, I. Afgan, Proper orthogonal decomposition and dynamic mode decomposition of jet in channel crossflow, *Nucl. Eng. Des.* 344 (2019) 54–68. <https://doi.org/10.1016/j.nucengdes.2019.01.015>.
- [262] S. Fresca, L. Dede', A. Manzoni, A Comprehensive Deep Learning-Based Approach to Reduced Order Modeling of Nonlinear Time-Dependent Parametrized PDEs, Springer US, 2021. <https://doi.org/10.1007/s10915-021-01462-7>.
- [263] R. Fu, D. Xiao, I.M. Navon, C. Wang, A data driven reduced order model of fluid flow by auto-Encoder and self-attention deep learning methods, *ArXiv:2109.02126 [Physics.Comp-Ph]*. (2021).
- [264] H. Eivazi, S. Le Clainche, S. Hoyas, R. Vinuesa, Towards extraction of orthogonal and parsimonious non-linear modes from turbulent flows, *ArXiv:2109.01514 [Physics.Fluid-Dyn]*. (2021). <http://arxiv.org/abs/2109.01514>.
- [265] N.T. Mücke, S.M. Boht, C.W. Oosterlee, Reduced order modeling for parameterized time-dependent PDEs using spatially and memory aware deep learning, *J. Comput. Sci.* 53 (2021) 101408.
- [266] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
- [267] H. Nugroho, M. Susanty, A. Irawan, M. Koyimatu, A. Yunita, Fully Convolutional Variational Autoencoder For Feature Extraction Of Fire Detection System, *J. Ilmu Komput. Dan Inf. (Journal Comput. Sci. Information)*. 13 (2020) 9–15. <https://doi.org/10.21609/jiki.v13i1.761>.
- [268] F.J. Gonzalez, M. Balajewicz, Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems, *ArXiv:1808.01346 [Math.DS]*. (2018).
- [269] P. Baldi, K. Hornik, Neural networks and principal component analysis: Learning from examples without local minima, *Neural Networks*. 2 (1989) 53–58. [https://doi.org/10.1016/0893-6080\(89\)90014-2](https://doi.org/10.1016/0893-6080(89)90014-2).

- [270] R. Vidal, Y. Ma, S. Sastry, Generalized Principal Component Analysis (GPCA), in: 2003 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2003. <https://doi.org/10.1109/CVPR.2003.1211411>.
- [271] https://tensorflow.google.cn/api_docs/python/tf/nn/relu, (2021).
- [272] D. Hartman, L.K. Mestha, A Deep Learning Framework for Model Reduction of Dynamical Systems, in: 2017 IEEE Conf. Control Technol. Appl., 2017: pp. 1917–1922. <https://doi.org/10.1109/CCTA.2017.8062736>.
- [273] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, *Nature*. 521 (2015) 436–444. <https://doi.org/10.1038/nature14539>.
- [274] V. Dumoulin, F. Visin, A guide to convolution arithmetic for deep learning, *ArXiv:1603.07285 [Stat.ML]*. (2018).
- [275] https://tensorflow.google.cn/api_docs/python/tf/nn/relu, (2021).
- [276] R. Hou, C. Chen, R. Sukthankar, M. Shah, An Efficient 3D CNN for Action/Object Segmentation in Video, *ArXiv:1907.08895 [Cs.CV]*. (2019).
- [277] J. Arunehru, G. Chamundeeswari, S.P. Bharathi, Human Action Recognition using 3D Convolutional Neural Networks with 3D Motion Cuboids in Surveillance Videos, *Procedia Comput. Sci.* 133 (2018) 471–477. <https://doi.org/10.1016/j.procs.2018.07.059>.
- [278] M. Lopez-Martin, S. Le, B. Carro, Model-free short-term fluid dynamics estimator with a deep 3D-convolutional neural network, *Expert Syst. Appl.* 177 (2021) 114924. <https://doi.org/10.1016/j.eswa.2021.114924>.
- [279] F.J. Narcowich, J.D. Ward, Scattered-data interpolation on \mathbb{R}^n : Error estimates for radial basis and band-limited functions, *SIAM J. Math. Anal.* 36 (2004) 284–300.
- [280] D. Xiao, F. Fang, C. Pain, G. Hu, Non-intrusive reduced-order modelling of the Navier – Stokes equations based on RBF interpolation, *Int. J. Numer. Methods Fluids.* 79 (2015) 580–595. <https://doi.org/10.1002/fld.4066>.
- [281] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, 2006.
- [282] C.J. Moore, A.J.K. Chua, C.P.L. Berry, J.R. Gair, Fast methods for training Gaussian processes on large datasets, *R. Soc. Open Sci.* 3 (2016) 160125. <https://doi.org/10.1098/rsos.160125>.
- [283] A. Banerjee, D.B. Dunson, S. Tokdar, Efficient Gaussian process regression for large datasets, *Biometrika.* 100 (2013) 75–89. <https://doi.org/10.1093/biomet/ass068>.

- [284] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural Comput.* 9 (1997) 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [285] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, in: *Conf. Empir. Methods Nat. Lang. Process. (EMNLP 2014)*, 2014: pp. 1724–1734. <https://doi.org/10.3115/v1/d14-1179>.
- [286] H. Jaeger, Echo state network, *Scholarpedia.* 2 (2007) 2330. <https://doi.org/10.4249/scholarpedia.2330>.
- [287] K. Hasegawa, K. Fukami, T. Murata, K. Fukagata, CNN-LSTM based reduced order modeling of two- dimensional unsteady flows around a circular cylinder at different Reynolds numbers, *Fluid Dyn. Res.* 52 (2020) 1–22.
- [288] J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach, *Phys. Rev. Lett.* 120 (2018) 024102. <https://doi.org/10.1103/PhysRevLett.120.024102>.
- [289] N.A.K. Doan, W. Polifke, L. Magri, Physics-informed echo state networks, *J. Comput. Sci.* 47 (2020) 101237. <https://doi.org/10.1016/j.jocs.2020.101237>.
- [290] F. Heyder, J. Schumacher, Echo state network for two-dimensional turbulent moist Rayleigh-Bénard convection, *Phys. Rev. E.* 103 (2021) 053107. <https://doi.org/10.1103/PhysRevE.103.053107>.
- [291] Q. Wu, E. Fokoue, D. Kudithipudi, On the Statistical Challenges of Echo State Networks and Some Potential Remedies, *ArXiv:1802.07369 [Stat.ML]*. (2018).
- [292] P. Verzelli, C. Alippi, L. Livi, Echo State Networks with Self-Normalizing Activations on the Hyper-Sphere, *Sci. Rep.* 9 (2019) 13887. <https://doi.org/10.1038/s41598-019-50158-4>.
- [293] D.J. Gauthier, E. Bollt, A. Griffit, W.A.S. Barbosa, Next generation reservoir computing, *Nat. Commun.* 12 (2021) 5564. <https://doi.org/10.1038/s41467-021-25801-2>.
- [294] K. Zheng, B. Qian, S. Li, Y. Xiao, W. Zhuang, Q. Ma, Long-short Term Echo State Network for Time Series Prediction, *IEEE Access.* 8 (2020) 91961–91974. <https://doi.org/10.1109/ACCESS.2020.2994773>.
- [295] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: *Proc. 22nd Acm Sigkdd Int. Conf. Knowl. Discov. Data Mining, ACM*, 2016: pp. 785–794.
- [296] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, Chapman and Hall/CRC, 2012.
- [297] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N.

- Kalchbrenner, A. Senior, K. Kavukcuoglu, WaveNet: A Generative Model for Raw Audio, ArXiv:1609.03499 [Cs.SD]. (2016).
- [298] J. Xu, K. Duraisamy, Multi-level convolutional autoencoder networks for parametric prediction of spatio-temporal dynamics, *Comput. Methods Appl. Mech. Eng.* 372 (2020) 113379. <https://doi.org/10.1016/j.cma.2020.113379>.
- [299] S. Bai, J.Z. Kolter, V. Koltun, An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling, ArXiv:1803.01271 [Cs.LG]. (2018).
- [300] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention Is All You Need, in: 31st Conf. Neural Inf. Process. Syst. (NIPS 2017), 2017.
- [301] N.B. Erichson, M. Muehlebach, M.W. Mahoney, Physics-informed Autoencoders for Lyapunov-stable Fluid Flow Prediction, ArXiv:1905.10866 [Physics.Comp-Ph]. (2019).
- [302] H. Eivazi, H. Veisi, M.H. Naderi, V. Esfahanian, Deep neural networks for nonlinear model order reduction of unsteady flows, *Phys. Fluids.* 32 (2020) 105104. <https://doi.org/10.1063/5.0020526>.
- [303] K. Fukami, T. Nakamura, K. Fukagata, Convolutional neural network based hierarchical autoencoder for nonlinear mode decomposition of fluid field data Convolutional neural network based hierarchical autoencoder for nonlinear mode decomposition of fluid field data, *Phys. Fluids.* 32 (2020) 095110. <https://doi.org/10.1063/5.0020721>.
- [304] R. Saegusa, H. Sakano, S. Hashimoto, Nonlinear principal component analysis to preserve the order of principal components, *Neurocomputing.* 61 (2004) 57–70. <https://doi.org/10.1016/j.neucom.2004.03.004>.
- [305] D.P. Kingma, M. Welling, Auto-encoding variational bayes, ArXiv:1312.6114 [Stat.ML]. (2014).
- [306] D.J. Rezende, S. Mohamed, D. Wierstra, Stochastic backpropagation and approximate inference in deep generative models, ArXiv:1401.4082 [Stat.ML]. (2014).
- [307] K. Hasegawa, K. Fukami, T. Murata, K. Fukagata, Machine-learning-based reduced-order modeling for unsteady flows around bluff bodies of various shapes, *Theor. Comput. Fluid Dyn.* 34 (2020) 367–383. <https://doi.org/10.1007/s00162-020-00528-w>.
- [308] K. Fukami, K. Fukagata, K. Taira, Super-resolution reconstruction of turbulent flows with machine learning, *J. Fluid Mech.* 870 (2019) 106–120. <https://doi.org/10.1017/jfm.2019.238>.
- [309] K. Fukami, K. Fukagata, K. Taira, Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows, *J. Fluid Mech.* 909 (2021).

<https://doi.org/10.1017/jfm.2020.948>.

- [310] T. Nakamura, K. Fukami, K. Hasegawa, Y. Nabaie, K. Fukagata, Convolutional neural network and long short-term memory based reduced order surrogate for minimal turbulent channel flow, *Phys. Fluids*. 33 (2021) 025116. <https://doi.org/10.1063/5.0039845>.
- [311] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations, *ArXiv:1711.10566 [Cs.AI]*. (2017). <http://arxiv.org/abs/1711.10566>.
- [312] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations, *ArXiv:1711.10566 [Cs.AI]*. (2017). <http://arxiv.org/abs/1711.10566>.
- [313] M. Raissi, A. Yazdani, G. Em Karniadakis, Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations, *Science (80-.)*. 367 (2020) 1026–1030. <https://doi.org/10.1126/science.aaw4741>.
- [314] C. Cheng, G.T. Zhang, Deep learning method based on physics informed neural network with Resnet block for solving fluid flow problems, *Water*. 13 (2021) 423. <https://doi.org/10.3390/w13040423>.
- [315] S. Pan, K. Duraisamy, Long-time predictive modeling of nonlinear dynamical systems using neural networks, *Complexity*. 2018 (2018). <https://doi.org/10.1155/2018/4801012>.
- [316] P. Wu, S. Gong, K. Pan, F. Qiu, W. Feng, C. Pain, Reduced order model using convolutional auto-encoder with self-attention, *Phys. Fluids*. 33 (2021). <https://doi.org/10.1063/5.0051155>.
- [317] X. Du, X. Qu, Y. He, D. Guo, Single image super-resolution based on multi-scale competitive convolutional neural network, *Sensors (Switzerland)*. 18 (2018) 789. <https://doi.org/10.3390/s18030789>.
- [318] X. Wang, R. Girshick, A. Gupta, K. He, Non-local Neural Networks, in: *IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Non-Local*, IEEE, 2018: pp. 7794–7803. <https://doi.org/10.1109/CVPR.2018.00813>.
- [319] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: *27th Int. Conf. Mach. Learn.*, 2010: pp. 807–814.
- [320] W.A. Little, The existence of persistent states in the brain, *Math. Biosci.* 19 (1974) 101–120. [https://doi.org/10.1016/0025-5564\(74\)90031-5](https://doi.org/10.1016/0025-5564(74)90031-5).
- [321] W.A. Little, G.L. Shaw, Analytic study of the memory storage capacity of a neural network, *Math. Biosci.* 39 (1978) 281–290. [https://doi.org/10.1016/0025-5564\(78\)90058-5](https://doi.org/10.1016/0025-5564(78)90058-5).

- [322] J. Han, C. Moraga, Han, J., Moraga, C. (1995). The influence of the sigmoid function parameters on the speed of backpropagation learning. In: Mira, J., Sandoval, F. (eds) From Natural to Artificial Neural Computation. IWANN 1995., Lect. Notes Comput. Sci. 930 (1995) 195–201. https://doi.org/10.1007/3-540-59497-3_175.
- [323] M. Mathieu, C. Couprie, Y. LeCun, Deep multi-scale video prediction beyond mean square error, ArXiv:1511.05440v6 [Cs.LG]. (2016).
- [324] S. Lee, D. You, Data-driven prediction of unsteady flow over a circular cylinder using deep learning, *J. Fluid Mech.* 879 (2019) 217–254.
- [325] D.P. Kingma, J.L. Ba, Adam: A method for stochastic optimization, 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. (2015) 1–15.
- [326] Z. Wu, D. Laurence, S. Utyuzhnikov, I. Afgan, Proper orthogonal decomposition and dynamic mode decomposition of jet in channel crossflow, *Nucl. Eng. Des.* 344 (2019) 54–68. <https://doi.org/10.1016/j.nucengdes.2019.01.015>.
- [327] C.W. Rowley, I. Mezi, S. Bagheri, P. Schlatter, D.S. Henningson, Spectral analysis of nonlinear flows, *J. Fluid Mech.* 641 (2009) 115–127. <https://doi.org/10.1017/S0022112009992059>.
- [328] A. Towne, O.T. Schmidt, T. Colonius, Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis, *J. Fluid Mech.* 847 (2018) 821–867. <https://doi.org/10.1017/jfm.2018.283>.
- [329] C.E. Heaney, Y. Li, O.K. Matar, C.C. Pain, Applying Convolutional Neural Networks to Data on Unstructured Meshes with Space-Filling Curves, ArXiv:2011.14820v2 [Math.NA]. (2021). <https://doi.org/10.48550/arXiv.2011.14820>.
- [330] J. Tencer, K. Potter, A Tailored Convolutional Neural Network for Nonlinear Manifold Learning of Computational Physics Data Using Unstructured Spatial Discretizations, *SIAM J. Sci. Comput.* 43 (2021) 2581–2613. <https://doi.org/10.1137/20M1344263>.
- [331] C.E. Heaney, Z. Wolffs, J.A. Tómasson, L. Kahouadji, P. Salinas, A. Nicolle, I.M. Navon, O.K. Matar, N. Srinil, C.C. Pain, An AI-based non-intrusive reduced-order model for extended domains applied to multiphase flow in pipes, *Phys. Fluids.* 34 (2022) 055111. <https://doi.org/10.1063/5.0088070>.
- [332] P. Pant, R. Doshi, P. Bahl, A. Barati Farimani, Deep learning for reduced order modelling and efficient temporal evolution of fluid simulations, *Phys. Fluids.* 33 (2021) 107101. <https://doi.org/10.1063/5.0062546>.
- [333] J.O. Hinze, *Turbulence*, McGraw-Hill, 1975. <https://books.google.com/books/about/Turbulence.html?id=xfRQAAAAMAAJ> (accessed November 13, 2018).

[334] Ansys fluent 12.0, 2009.

[335] J. Weiss, A tutorial on the proper orthogonal decomposition, AIAA Aviat. 2019 Forum. (2019) 1–21. <https://doi.org/10.2514/6.2019-3333>.

Appendix A: Equations of the employed CFD models

The governing equations are the continuity, Navier-Stokes, energy, and the scalar transport (pollutant concentration conservations) equations. In the present dissertation, the ELES approach is used to conduct the simulations. This approach comprises of two different models, namely, DSLM-LES and RLZ.

LES part:

For the LES part, the governing equations must go under a filtering operation, resulting in the filtered equations that can be used along with the LES approach [35]:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0$$

Eq. A- 1

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{u}_j) = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + 2\nu \frac{\partial \bar{S}_{ij}}{\partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} + g_i$$

Eq. A- 2

$$\frac{\partial \bar{T}}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{T}) = \frac{\partial}{\partial x_j} \left(\alpha \frac{\partial \bar{T}}{\partial x_j} - H_j^{\text{SGS}} \right)$$

Eq. A- 3

$$\frac{\partial \bar{c}}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{c}) = \frac{\partial}{\partial x_j} \left(D \frac{\partial \bar{c}}{\partial x_j} - J_j^{\text{SGS}} \right)$$

Eq. A- 4

where, \bar{u}_i , \bar{p} , \bar{T} and \bar{c} are the resolved filtered field variables denoting the velocity components, pressure, temperature, and concentration, respectively. ρ , ν , α , and D represent the density, molecular kinematic viscosity, thermal diffusivity, and molecular diffusivity, consecutively. \bar{S}_{ij} , τ_{ij} , H_j^{SGS} , and J_j^{SGS} show the symmetric stress tensor, the SGS stress tensor, the SGS heat flux, and the SGS concentration flux, respectively, which are defined by Eq. A-5 to Eq. A-8. Furthermore, g_i is the downward gravitational force; thus, $g_1 = g_2 = 0$.

$$\bar{S}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)$$

Eq. A- 5

$$\tau_{ij} = \overline{u_i u_j} - \bar{u}_i \bar{u}_j$$

Eq. A- 6

$$H_j^{\text{SGS}} = \overline{u_j T} - \bar{u}_j \bar{T}$$

Eq. A- 7

$$J_j^{\text{SGS}} = \overline{u_j c} - \bar{u}_j \bar{c}$$

Eq. A- 8

The simple box filter is selected as the filter function. The filter width, Δ , equals to $(\Delta_x \Delta_y \Delta_z)^{\frac{1}{3}}$. Afterward, the SGS stress tensor, also known as the residual stress tensor, can be modeled based on the Boussinesq hypothesis [216,333]:

$$\tau_{ij} - \frac{1}{3} \tau_{kk} \delta_{ij} = -2\nu_{\text{SGS}} \bar{S}_{ij}$$

Eq. A- 9

where, δ_{ij} is the Kronecker delta. Also, ν_{SGS} , known as the SGS viscosity, is being modeled by SGS models. It should be noted that after calculating ν_{SGS} , the heat and mass fluxes can be obtained using the SGS Prandtl number (Pr_{SGS}) and the SGS Schmidt number (Sc_{SGS}). The SGS model used in the present study is DSLM. However, SSLM, which is regarded as the simplest model, should be explained first since DSLM is achieved by modifying the calculation process in the model achieved by SSLM. In SSLM:

$$\nu_{\text{SGS}} = (C_s \Delta)^2 |\bar{S}|$$

Eq. A- 10

where, C_s and $|\bar{S}| = \sqrt{(2\bar{S}_{ij}\bar{S}_{ij})}$ denote the Smagorinsky constant and the magnitude of the strain rate, respectively. It is noteworthy that C_s is flow-dependent. However, SSLM assumes it to be constant everywhere in the model. $C_s = 0.1$ is an accepted value for a wide range of flows [35]. Eq. A-10 shows that by increasing the velocity gradients, the value of ν_{SGS} becomes excessively large. However, SGS turbulent fluctuations should approach to zero in the vicinity of a wall. Therefore, in order to modify the calculation of ν_{SGS} , a damping function is used in that region. A convenient way to dampen ν_{SGS} in the proximity of a wall is the upper limit length scale modification [35]:

$$\Delta = \min \left\{ (\Delta_x \Delta_y \Delta_z)^{\frac{1}{3}}, \kappa n \right\}$$

Eq. A- 11

where, κ and n represent the von-Karman constant and the distance to the nearest wall, consecutively.

As previously mentioned, DSLM is a modified version of SSLM. Instead of assuming a constant value for C_s , as is the case with SSLM, its value is dynamically computed based on the information from the resolved scales of motions in DSLM. In order to eliminate the smallest scales of the resolved field, a test filter, $\hat{\Delta} = 2\Delta$ is applied to Eq. A-2. It should be noted that the last term in Eq. A-2 is omitted since it makes no difference in the procedure. Furthermore, it is beneficial to pay attention to Figure A.1, obtained from [35], which shows the energy spectrum with the grid and test filters, to understand better how these filters work.

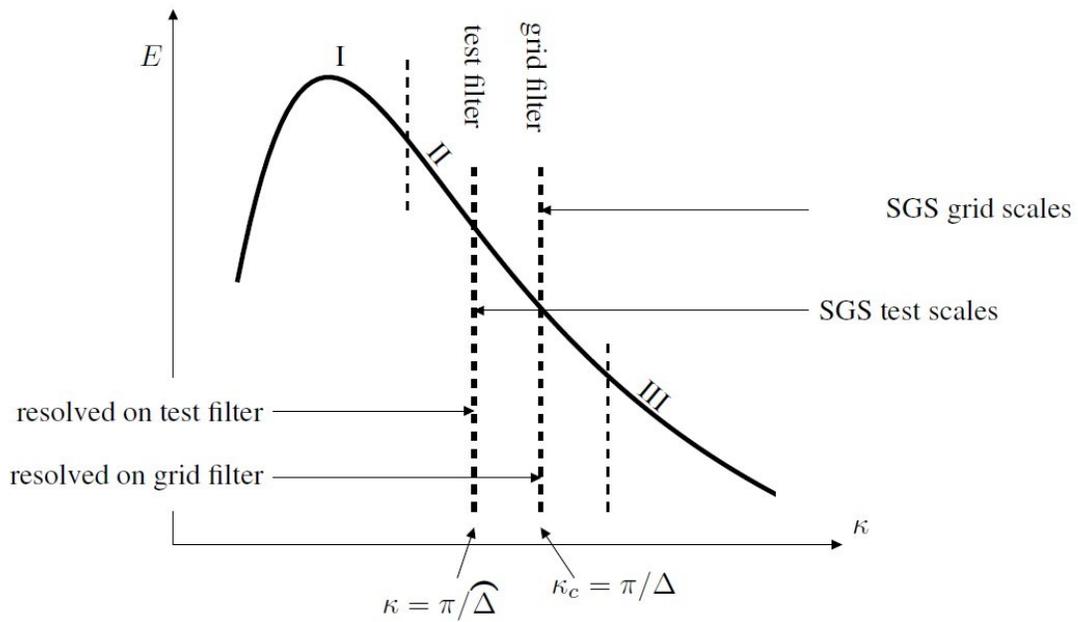


Figure A. 2: The energy spectrum with the test and grid filters [35]

Eq. A- 12

$$\frac{\partial \hat{u}_i}{\partial t} + \frac{\partial (\hat{u}_i \hat{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \hat{p}}{\partial x_i} + \nu \frac{\partial^2 \hat{u}_i}{\partial x_j \partial x_j} - \frac{\partial T_{ij}}{\partial x_j}$$

where, T_{ij} is the sub-test scale stress tensor, defined as:

$$T_{ij} = \widehat{u_i u_j} - \widehat{u_i} \widehat{u_j}$$

Eq. A-13

$$\frac{\partial \widehat{u_i}}{\partial t} + \frac{\partial (\widehat{u_i} \widehat{u_j})}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \widehat{p}}{\partial x_i} + \nu \frac{\partial^2 \widehat{u_i}}{\partial x_j \partial x_j} - \frac{\partial \widehat{\tau_{ij}}}{\partial x_j} - \frac{\partial}{\partial x_j} (\widehat{u_i} \widehat{u_j} - \widehat{u_i} \widehat{u_j})$$

Eq. A-14

Thus, by comparing Eq. A-12 to Eq. A-14:

$$T_{ij} = \widehat{u_i} \widehat{u_j} - \widehat{u_i} \widehat{u_j} + \widehat{\tau_{ij}}$$

Eq. A-15

Thus, the dynamic Leonard stresses can be defined as:

$$\mathcal{L}_{ij} = \widehat{u_i} \widehat{u_j} - \widehat{u_i} \widehat{u_j} = T_{ij} - \widehat{\tau_{ij}}$$

Eq. A-16

$$\mathcal{L}_{ii} = T_{ii} - \widehat{\tau_{ii}}$$

Eq. A-17

Thus, Eq. A-16 turns to:

$$\mathcal{L}_{ij} - \frac{1}{3} \mathcal{L}_{kk} \delta_{ij} = T_{ij} - \frac{1}{3} T_{kk} \delta_{ij} - \left(\widehat{\tau_{ij}} - \frac{1}{3} \widehat{\tau_{kk}} \delta_{ij} \right)$$

Eq. A-18

It is noteworthy to mention that the Leonard stresses represent those with the length scale ranging between Δ and $\widehat{\Delta}$ (intermediate level stresses). Therefore, the ranges of the length scale of τ_{ij} (grid level stresses) and T_{ij} (test level stresses) are also $\ell < \Delta$ and $\ell < \widehat{\Delta}$, respectively.

Considering similarity between T_{ij} and τ_{ij} , T_{ij} can be modeled as follows:

$$T_{ij} - \frac{1}{3} T_{kk} \delta_{ij} = -2\nu_{SGS} \widehat{S}_{ij}$$

Eq. A-19

where, $\nu_{SGS} = C_D \widehat{\Delta}^2 |\widehat{S}|$. Thus, C_D should be compared to C_S^2 . Also, $\widehat{S}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)$ and $|\widehat{S}| = \left(2\widehat{S}_{ij}\widehat{S}_{ij} \right)^{1/2}$.

By applying a test filter to Eq. A-9 (and substituting C_S^2 with C_D), and substituting it along with Eq. A-19 into Eq. A-18, the Leonard stresses, with the assumption of $C_D \widehat{\Delta}^2 |\widehat{S}| \bar{S}_{ij} = C_D \Delta^2 |\widehat{S}| \widehat{S}_{ij}$, can be rewritten as:

Eq. A-20

$$L_{ij} - \frac{1}{3} \mathcal{L}_{kk} \delta_{ij} = -2C_D \left(\widehat{\Delta}^2 |\widehat{S}| \widehat{S}_{ij} - \Delta^2 |\widehat{S}| \bar{S}_{ij} \right) = -2C_D M_{ij}$$

where, the $\frac{1}{3} \mathcal{L}_{kk} \delta_{ij} = \frac{1}{3} T_{kk} \delta_{ij} - \frac{1}{3} \hat{t}_{kk} \delta_{ij}$ is used. It should be noted that $C_D = C_D(x_i, t)$. The model factor, C_D , can be calculated by minimizing the difference between the left- and right-hand sides of Eq. A-20, in a least-square sense, as suggested by Lilly [105]:

Eq. A-21

$$Q = \left(L_{ij} - \frac{1}{3} \mathcal{L}_{kk} \delta_{ij} + 2C_D M_{ij} \right)^2$$

Eq. A-22

$$\frac{\partial Q}{\partial C_D} = 0 \Rightarrow C_D = -\frac{L_{ij} M_{ij}}{2M_{ij} M_{ij}}$$

In order to avoid numerical instability, C_D is clipped at 0.0, and in the positive side to 0.23 [35].

URANS part:

The governing equations of URANS is similar to the RANS equations, but with the unsteady term [35]:

Eq. A-23

$$\frac{\partial \langle u_i \rangle}{\partial x_i} = 0$$

Eq. A-24

$$\frac{\partial \langle u_i \rangle}{\partial t} + \frac{\partial}{\partial x_j} (\langle u_i \rangle \langle u_j \rangle) = -\frac{1}{\rho} \frac{\partial \langle p \rangle}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial \langle u_i \rangle}{\partial x_j} - \langle u_i' u_j' \rangle \right) + g_i$$

Eq. A- 25

$$\frac{\partial \langle T \rangle}{\partial t} + \frac{\partial}{\partial x_j} (\langle u_i \rangle \langle T \rangle) = \frac{\partial}{\partial x_j} \left(\alpha \frac{\partial \langle T \rangle}{\partial x_j} - H_j^{RANS} \right)$$

Eq. A- 26

$$\frac{\partial \langle c \rangle}{\partial t} + \frac{\partial}{\partial x_j} (\langle u_i \rangle \langle c \rangle) = \frac{\partial}{\partial x_j} \left(D \frac{\partial \langle c \rangle}{\partial x_j} - J_j^{RANS} \right)$$

where, $\langle u_i \rangle$, $\langle p \rangle$, $\langle T \rangle$, and $\langle c \rangle$ represent time averaged velocity, pressure, temperature, and concentration, respectively. Also, H_j^{RANS} and J_j^{RANS} are turbulent heat and turbulent concentration fluxes, consecutively. The stress term in the parentheses, located in the right-hand side of Eq, A-24 comprises of two stresses: the viscous stresses, and the Reynolds stresses (if they are multiplied by ρ). Considering eddy viscosity, the stress term in the parentheses can be written as:

Eq. A- 27

$$\mu \frac{\partial \langle u_i \rangle}{\partial x_j} - \rho \langle u'_i u'_j \rangle = (\mu + \mu_t) \frac{\partial \langle u_i \rangle}{\partial x_j}$$

where, μ_t is turbulent viscosity or eddy viscosity. The role of a $k - \varepsilon$ turbulence model is to model μ_t using the following equation:

Eq. A- 28

$$\mu_t = \rho C_\mu \frac{k^2}{\varepsilon}$$

where, k and ε denote the TKE and turbulence dissipation rate, respectively, which are calculated by the RLZ model k and ε transport equations. In the RLZ model, C_μ is variable [334]:

Eq. A- 29

$$C_\mu = \frac{1}{A_0 + A_s \frac{kU^*}{\varepsilon}}$$

where, $U^* = \sqrt{S_{ij}S_{ij} + \tilde{\Omega}_{ij}\tilde{\Omega}_{ij}}$. Also, $\tilde{\Omega}_{ij} = \Omega_{ij} - 2\varepsilon_{ijk}\omega_k$ and $\Omega_{ij} = \bar{\Omega}_{ij} - \varepsilon_{ijk}\omega_k$. It should be noted that $\bar{\Omega}_{ij}$ is the mean rate-of-rotation tensor located in a rotating reference frame with the angular velocity of ω_k [334]. Furthermore, $A_0 = 4.04$ and $A_s = \sqrt{6}\cos\phi$, where $\phi = \frac{1}{3}\cos^{-1}(\sqrt{6}W)$. Also, $W = \frac{S_{ij}S_{jk}S_{ki}}{\tilde{s}^3}$, where S_{ij} is the strain rate tensor, $S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$ and $\tilde{S} = \sqrt{S_{ij}S_{ij}}$. More detailed information on the RLZ model can be found in [334].

Appendix B: POD mathematics

Introduced to the fluid dynamics fields by Lumley [218], POD is a technique that can be used to capture dominant features of the turbulent flow. By decomposing a random vector field, representing the turbulent flow motion, into a set of deterministic functions, POD can capture the most energetic portion of the flow regarding its turbulence kinetic energy (TKE) [335]. Therefore, instead of analyzing a high-dimensional system, a limited number of deterministic functions can represent the most dominant features of the turbulent flow [220]. The technique can also be employed for model reduction by projecting the governing equations on a low-dimensional subspace [220].

The fluctuating velocity, $u'(\underline{x}, t)$, can be written as a set of deterministic spatial functions, $\varphi^n(\underline{x})$, and random time coefficients, $a_n(t)$, as follows [128,335]:

$$u'_i(\underline{x}, t) = \sum_{n=1}^N a_n(t) \varphi_i^n(\underline{x})$$

Eq. B- 1

where, t represents the time and \underline{x} denotes the position vector (x, y, z) , and N is the number of spatiotemporal modes, which form a complete orthogonal basis for Eq. B-1. The temporal correlation tensor, $C(t, t')$, is defined as Eq. B-2, and $a_n(t)$ are its eigenvectors [128]:

$$C(t, t') = \frac{1}{T} \int_V u'_i(\underline{x}, t) u'_i(\underline{x}, t') d\underline{x}$$

Eq. B- 2

$$\int_T C(t, t') a_n(t') dt' = \lambda^n a_n(t)$$

Eq. B- 3

where, λ^n is the n^{th} eigenvalue related to the n^{th} mode. Also, T is the upper limit of the time domain. By projecting the temporal velocity fields onto the temporal modes, the spatial modes can be obtained [128]:

$$\varphi_i^n(\underline{x}) = \frac{1}{T \lambda^n} \int_T a_n(t) u'_i(\underline{x}, t) dt$$

Eq. B- 4

It should be noted that the obtained eigenvalues contribute to the total TKE in the way that the first one has a more significant share in TKE than the rest. Thus, they can be re-ordered as $\lambda^n > \lambda^{n+1}$ [220]. More information on the mathematics of POD can be found in [220].