UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería del Diseño

Final Degree Project

Aerospace Engineering Degree

Spatio-temporal structures in a simplified urban flow

Author:
Álvaro Martínez Sánchez

Supervisor:
Dr. Ricardo Vinuesa Motilva

Advisor:
Dr. Sergio Hoyas Calvo

Universitat Politècnica de València

KTH Royal Institute of Technology

July, 2021

Acknowledgement

First and foremost, I would like to sincerely express my appreciation to my advisers, Dr. Ricardo Vinuesa Motilva and Dr. Sergio Hoyas Calvo, for their unwavering support during the completion of my Bachelor's thesis and research, as well as their patience, inspiration, excitement, and vast expertise. Their advice was invaluable at every stage of my research and thesis completion; I could not have wished for a finer opportunity for my final degree project.

Aside from my advisers, I would want to express my gratitude to Dr. Soledad Le Clainche Martínez for her continuous support, advice and encouragement, as well as for providing and explaining all of the tools required to complete the current project.

Special gratitude is also deserved by Pablo Torres, Eneko Lazpita, Adrián Corrochano and Hamidreza Eivazi for their encouragement, perceptive remarks and inspirational discussions. Without a shadow of a doubt, the research experience of this project has been considerably more profitable thanks to all of them.

I also appreciate my fellow classmates in Valencia for their inspiring conversations, the sleepless hours we spent working together before deadlines and the wonderful four years we shared together. The experience of studying at university will be unforgettable thanks to them.

Last but not least, I would want to express my gratitude to my family: my parents and sister, for spiritually supporting me and believing in me throughout my entire life.

Abstract

Understanding flow structures in urban areas is being widely recognised as an important concern due to its effect on urban development, air quality and pollutant dispersion. In this study, state-of-the-art data-driven methods for modal analysis of urban flows are used to achieve a better understanding of the dominant flow processes that occur in this dynamic phenomenon. Higher-order dynamic mode decomposition (HODMD), a highly efficient method in the analysis of turbulent flows, is used in conjunction with traditional techniques such as singular-value decomposition (SVD) and proper orthogonal decomposition (POD) over high-fidelity simulation data of a simplified urban environment. The flow simulation is carried out using the spectral-element code Nek5000 through a well-resolved large-eddy simulation (LES). The resulting flow regime will depend on the height-to-separation ratio h/ℓ varying from 1 to 0.25 for two wall-mounted obstacles with a constant width-to-height ratio $b/h = 0.5$. The interaction of the flow within the canopy influences the coherent structures, with the objective of the present project being to characterise the flow mechanisms driving the dynamics in urban areas with different separation ratios. Particularly, from all the vortical structures identified by the HODMD algorithm, low- and high-frequency modes are classified depending on their relation to the arch vortex and are referred to as arch generators and breakers, respectively. This classification implies that one of the processes driving the formation and destruction of major vortical structures in between the buildings is the interaction between low- and high-frequency structures. Additionally, when the separation increases, the flow inside the canopy becomes more correlated in the spanwise direction due to the increased interaction of the flow in this region. Therefore, HODMD is shown to be a powerful tool for describing the physics of very complicated turbulent flows.

Key words: direct numerical simulations (DNS), turbulent boundary layers (TBL), urban flows, spatio-temporal structures, data-driven methods

Resumen

La comprensión de las estructuras del flujo en áreas urbanas está adquiriendo una importancia cada vez mayor debido a su influencia sobre la planificación urbana, la calidad del aire y la dispersión de contaminantes. En este estudio, se utilizan métodos de vanguardia para el análisis modal de flujos urbanos con el objetivo de lograr una mejor comprensión de los patrones dominantes que tienen lugar en este fenómeno dinámico. *Dynamic mode decomposition* de orden superior (HODMD), un método sumamente efectivo para el análisis de flujos turbulentos, se utiliza junto con técnicas tradicionales como *singular-value decomposition* (SVD) y *proper orthogonal decomposition* (POD) sobre datos de simulación de alta fidelidad en un entorno urbano simplificado. La simulación del flujo se lleva a cabo utilizando el código de elementos espectrales Nek5000 a través de una *large-eddy simulation* (LES) bien resuelta. El régimen del flujo resultante dependerá en gran medida de la relación entre la altura y la separación de los obstáculos h/ℓ que variará entre 1 y 0.25 para dos obstáculos de pared con una relación entre ancho y altura constante $b/h = 0.5$. La interacción del flujo en la sección intermedia de los obstáculos influye en las estructuras coherentes, siendo el principal objetivo del presente estudio caracterizar los mecanismos del flujo que promueven esta dinámica en áreas urbanas con diferentes relaciones de separación. En particular, de todas las estructuras vorticales identificadas por el algoritmo HODMD, los modos de baja y alta frecuencia se clasifican según su relación con el vórtice de arco y se denominan generadores y rompedores de arco, respectivamente. Esta clasificación implica que la interacción entre estructuras de baja y alta frecuencia es uno de los mecanismos responsables de la generación y destrucción de estructuras vorticales relevantes entre los edificios. Asimismo, al aumentar la separación, el flujo se vuelve más correlacionado en la dirección transversal al flujo debido a la mayor interacción en la región comprendida entre los obstáculos. Por lo tanto, se ha demostrado que HODMD es una herramienta robusta para describir la física de flujos turbulentos de gran complejidad.

Palabras clave: simulación numérica directa (DNS), capas límites turbulentas (TBL), flujos urbanos, estructuras espacio-temporales

Resum

La comprensió de les estructures del flux en àrees urbanes està adquirint una importància cada vegada major degut a la seua influència sobre la planificació urbana, la qualitat de l'aire i la dispersió de contaminants. En aquest estudi, s'utilitzen mètodes d'avantguarda per a l'anàlisi modal de fluxos urbans amb l'objectiu d'aconseguir una millor comprensió dels patrons dominants que tenen lloc en aquest fenomen dinàmic. *Dynamic mode decomposition* d'ordre superior (HODMD), un mètode summament efectiu per a l'anàlisi de fluxos turbulents, s'utilitza junt a tècniques tradicionals com *singular-value decomposition* (SVD) i *proper ortogonal decomposition* (POD) sobre dades de simulació d'alta fidelitat en un entorn urbà simplificat. La simulació del flux es realitza utilitzant un codi d'elements espectrals Nek5000 a través d'una *large-eddy simulation* (LES) ben resolta. El règim del flux resultant dependrà en gran mesura de la relació entre l'altura i la separació dels obstacles h/ℓ que variarà entre 1 i 0.25 per a dos obstacles de paret amb una relació d'ample i altura constant $b/h = 0.5$. La interacció del flux en la secció intermèdia dels obstacles influeix en les estructures coherents, sent el principal objectiu del present estudi caracteritzar els mecanismes de flux que promouen aquesta dinàmica en àrees urbanes amb diferents relacions de separació. En particular, de totes les estructures vorticals trobades per l'algoritme HODMD, els modes de baixa i alta freqüència es classificaran segons la relació entre el vòrtex d'arc i es denominen generadors i destrossadors d'arc, respectivament. Aquesta classificació implica que la interacció entre estructures de baixa i alta freqüència es un dels mecanismes responsables de la generació i destrucció d'estructures vorticals rellevants entre els edificis. Així mateix, a l'augmentar la separació, el flux es torna més correlacionat en la direcció transversal al flux degut a la major interacció en la regió compresa entre els obstacles. Per tant, s'ha demostrat que HODMD és una ferramenta robusta per a descriure la física de fluxos turbulents de gran complexitat.

Paraules clau: simulació numèrica directa (DNS), capes límits turbulentes (TBL), fluxos urbans, estructures espacio-temporals

Contents

Acknowledgement i

Abstract ii

Resumen iii

Resum iv

I Report xii

1 Introduction 1

1.1 General concepts . . . 1

1.2 Motivation and objectives . . . 2

1.3 Historical perspective . . . 2

1.3.1 Fundamentals of urban flows . . . 2

1.3.1.1 Atmospheric boundary layer . . . 8

1.3.2 Coherent flow structures . . . 9

1.3.2.1 Experimental studies of coherent turbulent structures . . . 9

1.3.2.2 Methods for vortex identification . . . 11

1.3.3 Final comments on historical perspective . . . 14

2 Theoretical background 15

2.1 Theory in fluid mechanics . . . 15

2.1.1 The nature of turbulent flows . . . 15

2.1.2 The equations of fluid motion . . . 16

2.1.2.1 Continuum fluid properties . . . 17

2.1.2.2 Continuity equation . . . 17

2.1.2.3 Momentum equation . . . 18

2.1.2.4 Passive scalar equation . . . 18

2.1.2.5 Vorticity equation . . . 19

2.1.3 Definition of streamlines . . . 19

2.2 Computational theory . . . 20
2.3 Modal-decomposition theory . . . 21
2.3.1 Singular-value decomposition . . . 21
2.3.2 Fourier transforms . . . 23
2.3.2.1 Fourier series . . . 24
2.3.2.2 Fourier transform . . . 24
2.3.2.3 Discrete Fourier transform . . . 25
2.3.2.4 Fast Fourier transform . . . 26
2.3.3 Proper orthogonal decomposition . . . 27
2.3.4 Higher-order dynamic mode decomposition . . . 28
2.3.4.1 Step 1: dimension reduction . . . 29
2.3.4.2 Step 2: the DMD-d algorithm . . . 30
2.3.4.3 General remarks on the HODMD algorithm . . . 32
2.3.5 HODMD for spatially multidimensional data . . . 33
2.3.6 Iterative HODMD . . . 35
2.3.7 Summary on modal-decomposition analyses . . . 35

3 Problem implementation 36
3.1 Pre-processing tools . . . 36
3.1.1 Numerical domain . . . 36
3.1.2 Mesh design . . . 38
3.1.3 Boundary conditions . . . 39
3.2 Solution process . . . 39
3.3 Post-processing tools . . . 40
3.3.1 Interpolation mesh . . . 40
3.3.2 Modal decompositions . . . 41
3.3.2.1 Snapshot matrix . . . 42
3.3.2.2 Singular value decomposition . . . 42
3.3.2.3 Dynamic mode decomposition . . . 45
3.3.2.4 General conclusions on modal decompositions . . . 49

4 Results and discussion 50
4.1 Time-averaged velocity fields . . . 50
4.2 Proper orthogonal decomposition . . . 55
4.2.1 Temporal analysis using FFT . . . 60
4.3 Higher-order dynamic mode decomposition . . . 62
4.3.1 Initial calibration and parameter selection . . . 62
4.3.2 Global temporal modes . . . 63

4.3.2.1 Arch-generator modes 65
4.3.2.2 Arch-breaker modes 66
4.3.3 Streamline flow patterns 69

5 Conclusions **71**

II Blueprints, solicitation document and budget **77**

6 Plans and blueprints **78**

7 Solicitation document **79**
7.1 Functions of the involved parties 79
7.1.1 Functions of the student 79
7.1.2 Functions of the supervisor 79
7.1.3 Functions of the advisor 80
7.2 Working environment conditions 80

8 Budget **82**

List of Figures

1.1 Definition of sublayers involved in the study of the flow at the city scale . . . 3

1.2 Geometrical parameters and angle of incidence (AOI) of the incoming flow definition in a simplified urban model . . . 4

1.3 Threshold lines dividing the flow into three different regimes depending on the building (b/h) and the separation (h/ℓ) . . . 5

1.4 Flow regimes based on the obstacle separation, using the criterion of Oke . . . 7

1.5 Schematic representation of the flow around a surface-mounted rectangular obstacle . . . 10

1.6 Oil-film visualisation of the surface flow pattern of an in-line tandem arrangement of cubes . . . 11

1.7 Arch vortex schematic representation around a three-dimensional rectangular block with $b/h = 2$ and $w_b/h = 0.29$ depending on the angle of attack . . . 12

1.8 Representation of the iso-surfaces of $\Gamma_1 = 0.4$ for AOI $= 0°$ and $\Gamma_1 = 0.35$ for AOI $= 30°$, streamwise and spanwise rms velocity fluctuations, representing the 75% of the maximum value reported in each panel . . . 13

2.1 Time history of the axial component of velocity $U_1(t)$ on the centre line of a turbulent jet . . . 16

2.2 Schematic representation of the matrices involved in the full and economy singular value decomposition . . . 23

3.1 Generic geometrical scheme of the numerical domain . . . 37

3.2 Two-dimensional representation of the interpolation mesh at $z/h = 0$, proving the resolution in $y-$direction. . . . 41

3.3 Results of SVD applied to a 100×100 matrix for the toy model function in (3.2) 44

3.4 Results of DMD-d applied to 5000 snapshots with constant time interval $\Delta t = 8 \times 10^{-3}$ for the toy model in (3.3). . . . 47

3.5 Results of DMD-d applied to 5000 snapshots with constant time interval $\Delta t = 8 \times 10^{-3}$ for the toy model in (3.3) with a 10% of random noise. (a) Original toy model signal (black dashed line), original signal with noise (red) and DMD-d reconstruction (blue) using ω_0 and $\omega_2 \pm \omega_1$. (b) Zoomed-in version of panel (a). 48

4.1 Time-averaged streamwise velocity field U normalised with U_∞ at plane $z/h = 0$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes. . . . 51

4.2 Time-averaged wall-normal velocity field V normalised with U_∞ at plane $z/h = 0$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes. . . . 52

4.3 Time-averaged spanwise velocity field W normalised with U_∞ at plane $y/h = 0.25$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes. . . . 53

4.4 Main vortical structures formed around two wall-mounted obstacles with different separation ratios: (**a**) $h/\ell = 1$, (**b**) 0.5 and (**c**) 0.25, by means of the streamlines flow patterns . . . 54

4.5 Singular-value distribution normalised with its maximum value σ_1 of the POD modes corresponding to the streamwise velocity of the skimming-flow regime and symmetry imposed. . . . 55

4.6 POD analysis of the streamwise velocity at $y/h = 0.25$ for the skimming-flow regime ($h/\ell = 1$). Shown are the instantaneous and time-averaged fields and the associated four most dominant POD modes. . . . 56

4.7 Reconstructed streamwise velocity at $y/h = 0.25$ from POD at different latent space dimensions. The corresponding relative ℓ_2-norm of errors (ϵ) are listed in brackets. . . . 57

4.8 Singular-value distribution normalised with its maximum value σ_1 of the POD modes corresponding to the whole set of velocity components (streamwise, wall-normal and spanwise) of the (black) skimming-flow, (blue) wake-interference and (red) isolated-roughness regime. . . . 58

4.9 POD orthogonal basis of the streamwise velocity fields at $y/h = 0.25$ for the different flow regimes. . . . 59

4.10 POD orthogonal basis of the spanwise velocity fields at $y/h = 0.25$ for the different flow regimes. . . . 60

4.11 Temporal coefficients a_1 and a_2 of the first two POD modes for the (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness regime. 61

4.12 Power spectrum of FFT scaled with the Strouhal number St and applied to the temporal coefficients of the POD modes: (red) skimming-flow, (blue) wake-interference and (black) isolated-roughness regime. Peaks show the dominant frequencies associated with each POD mode. . . . 61

4.13 DMD-d calibration. Amplitude scaled with its maximum value ($\hat{a}_m = a_m/a_0$) versus frequency ω_m computed with different tolerances for (top) skimming flow, (middle) wake interference and (bottom) isolated roughness. . . . 63

4.14 Two-dimensional representation at $y/h = 0.25$ of the arch-breaker mode ($\omega_m = 1.22$) for the skimming-flow regime. . . . 64

4.15 Three-dimensional iso-surface of the (left) streamwise, (middle) wall-normal and (right) spanwise velocities of the arch-generator mode shown in Table 4.1. 66

4.16 Three-dimensional iso-surface of the (left) streamwise, (middle) wall-normal and (right) spanwise velocities of the arch-breaker mode shown in Table 4.1. . 67

4.17 DMD modes of the streamwise velocity fields at $y/h = 0.25$ for the different flow regimes. . . . 68

4.18 DMD modes of the spanwise velocity fields at $y/h = 0.25$ for the different flow regimes. . . . 68

4.19 Main flow patterns of the arch-generator mode ($\omega = 0.13$) visualised by means of streamlines for the wake-interference regime 69

4.20 Main flow patterns of the arch-breaker mode ($\omega = 1.05$) visualised by means of streamlines for the wake-interference regime 70

List of Tables

2.1 Tunable parameters of the HODMD algorithm 33

3.1 Geometrical parameters defining the numerical domain of the three flow regimes identified by Oke 38

3.2 Boundary conditions representing the physical behaviour of the flow simulation 39

3.3 Interpolation mesh parameters for the different flow regimes 41

3.4 Temporal parameters of the datasets used for the modal decomposition analysis 42

3.5 Results of SVD applied to the toy model function in (3.2) 45

3.6 Summary of the workflow stages for the DMD-d algorithm. 46

3.7 Results of the DMD-d used to approximate the toy model 47

3.8 Results of the DMD-d used to approximate the toy model with a 10% of random noise 48

4.1 Summary of the frequencies and amplitudes of the DMD modes shown in Figure 4.13. From top to bottom, the modes are ordered from lower to higher frequency. 64

8.1 Breakdown of the main costs before and after taxes of the present project . . 82

Part I

Report

1

Introduction

1.1 General concepts

The world we live in today is increasingly characterised by urban areas: around 55% of the global population has been reported to live in urban environments and, based on formerly consistent definitions of urbanised areas, this ratio is expected to increase up to 68% by 2050 [1]. In such a fashion, cities worldwide, which are the primary source of climate change, are required to play a decisive role in the accomplishment of the widely-known Paris Agreement targets [2], according to UN-Habitat Executive Director [3]. Moreover, the great levels of air pollution, to which the vast majority of the urban population is exposed, is undoubtedly related to a myriad of health issues; some recent research has estimated that climate change and, precisely, the global warming suffered by cities with respect to rural areas (UHI phenomenon), was strictly associated to $70,000$ deaths [4]. This problematic situation boosts the search for physical models capable of reproducing the pollutant and thermal distributions within cities.

Although some predictive models have already been introduced by the EU [5], their lack of providing the spatio-temporal required accuracy to model pollutant dispersion through urban environments forces researchers to improve those methods in order to ensure the must-have urban sustainability in the hereafter. Particularly, with the aim of establishing a proper action plan to alleviate the associated adverse consequences, several studies have focused their efforts on analysing the coherent structures of the flow: those three-dimensional flow regions responsible for the pollutant dispersion within a given urban geometry.

Furthermore, the plethora of data to which society has access nowadays is pushing big data techniques, including Artificial Intelligence (AI), forward. Even though a failure to properly manage this information overexposure might result in gaps in transparency or ethical standards, several studies are being currently developed in order to analyse the implications derived from applying those techniques to a wide range of disciplines of this data-driven world. The thorough research performed by Vinuesa et al. [6] analyses the role that AI is progressively adopting and its wider impact on various sectors and, notably, on achieving the Sustainable Development Goals (SDG) proposed by United Nations [7]. Focusing on the aforementioned urban sustainability goals, the analysis executed over SDG 11 (Sustainable Cities and Communities) showed clear evidence that AI could behave as an enabler towards the achievement of such targets. Therefore, using state-of-the-art data-driven techniques, general descriptions of the main patterns and instabilities driving the flow dynamics in urban areas can be provided.

Closely related to the point mentioned above, thanks to recent developments in computational techniques, direct numerical or large-eddy simulations can now be solved in a feasible time range and thus enabling researchers to have available more detailed information regarding the pollutant dispersion and its interaction within urban environments. This emerging sustainable discipline focuses on the analysis of the flow in order to properly characterise the urban air quality. It is, therefore, in this section where modern fluid simulations play a prominent role towards the understanding of the physics behind the flow in urban areas and, together with AI, the main patterns driving the flow dynamics can be identified. These coherent structures can be connected to the presence of polluting processes in such a way as to allow for a series of alternatives to minimise their impact when designing the cities of the time to come.

1.2 Motivation and objectives

The main motivation upon which this study relies is to gain further understanding into the physics of the flow within urban environments. This will enable to obtain new physical models of the dynamical processes responsible for pollutant dispersion and thermal effects in this type of environments. Particularly, in contrast to experimental studies, thanks to the increased spatial and temporal resolution provided by numerical simulations, the purpose is to characterise the spatio-temporal features of the flow, using data-driven methods that will shed light on the obtaining of generalisable and interpretable models of urban flows. Furthermore, the obtained results will be contrasted with the extensive work in both experimental and numerical analyses performed within the field.

Using the results of a large-eddy simulation (LES) over a simplified urban environment consisting of two building blocks, the objective is to use cutting-edge techniques to conduct modal decompositions over the three-dimensional instantaneous fields of the flow. Higher-order dynamic mode decomposition (HODMD), a highly efficient method in the study of turbulent flows, will be used together with common techniques such as singular-value decomposition (SVD) and proper orthogonal decomposition (POD), aiming at defining the key flow patterns with high precision and designing high-fidelity reduced-order models (ROMs). The resulting models will serve as a foundation for the creation of more complex algorithms, including machine learning, with reliable predictive capabilities.

1.3 Historical perspective

1.3.1 Fundamentals of urban flows

The issues associated with the dispersion of material through and above the urban area can be addressed by analysing the problem at the different length scales in which those environments can be described: regional and city scales (from 10 to 200 km), neighbourhood scale (up to 1 or 2 km) and street scale (between 100 and 200 m). Britter and Hanna [8] discussed how the physical processes at each of the different scales may vary and how they can be combined into mathematical models but conducted in a way that each model, reconstructed within its own regime, can serve as a reference for further interpretations at larger scales.

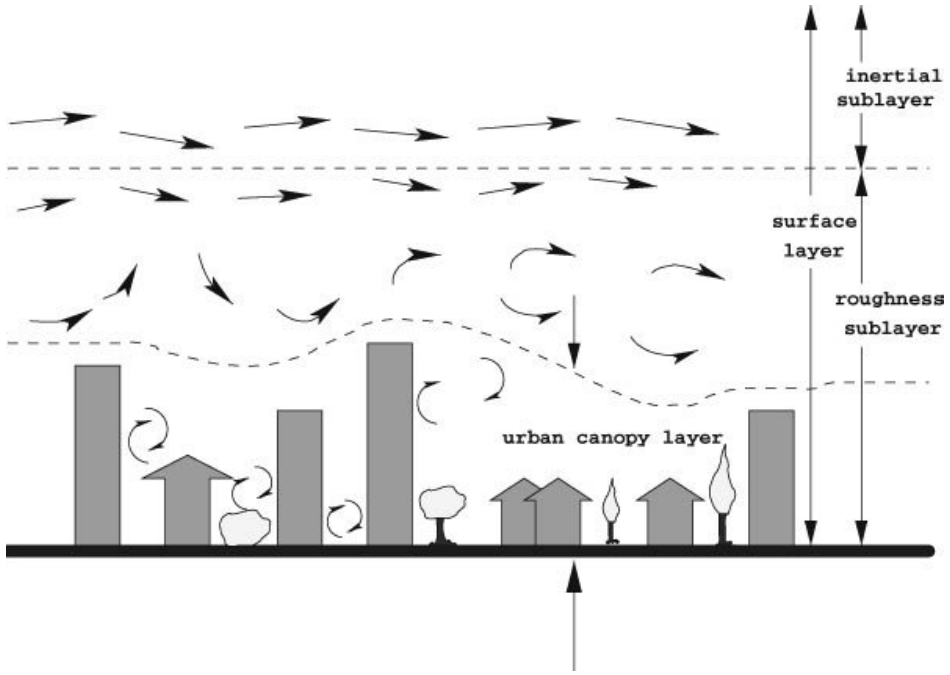


Figure 1.1: Definition of sublayers involved in the study of the flow at the city scale (urban area). Extracted from Britter and Hanna [8]

Regional and city scales. Britter and Hanna [8] describe the city scale as an urban environment that is clearly distinguished from its surroundings by its comparatively large obstacles that impose a significant drag force on the atmosphere. This scale is used to represent the diameter of a typical metropolitan region, allowing individual variations in flow and dispersion to be averaged out. In other words, at this size, additional information in the urban canopy will not be examined, and the emphasis will be on pollutants above the height of the buildings. Additionally, the authors contend that the regional scale refers to the surrounding zone that is mostly influenced by the metropolitan centre, i.e. the previously mentioned city scale. As a result, whilst the interaction between those zones can be of relevant importance, the flow understanding will be focused on the city scale, where Britter and Hanna [8], based on the work of Grimmond and Oke [9], identified the three main sublayers present in urban areas. In Figure 1.1, these urban layers are illustrated.

- **Inertial sublayer**: It is the flow region where the boundary layer has absorbed the obstacles perturbations (top sublayer). As a result, the layer can be thought of as a free stream layer, allowing normal atmospheric models to be applied. In truth, in theoretical aerodynamics, this is a well-known theory in which the effects of flow perturbations are ignored.

- **Urban-canopy sublayer**: In contrast to the inertial sublayer, the urban canopy sublayer is the one with the utmost interaction with urban obstacles. Therefore, those local obstacles will have a strong impact on the flow at a given location of this layer.

- **Roughness sublayer**: Having characterised the farthest and closest sublayers with respect to the local obstacles, i.e. the buildings, the roughness sublayer will refer to a transient band in which the flow adapts to the perturbations imposed in the urban canopy sublayer over time. It is worth saying that the urban sublayer is included within this layer and its limits expand up to the inertial sublayer.

Neighborhood scale. The neighbourhood scale is a 1–2 km spatial scale at which a gross parameterisation of the flow can be attempted as well as a scale at which extensive quantitative analysis is possible, though at a cost, in clear contrast with the previous scales. It is also a scale on which some statistical homogeneity can be expected; the city is then seen as a collection of these neighbourhoods. On this scale, dispersion research would almost certainly necessitate a more detailed understanding of the flow inside and directly above the urban canopy [8]. This is especially important when contemplating the effects and dangers of releasing dangerous chemicals into cities, whether by mistake or on purpose.

Street scale. The street (canyon) scale is analysed in particular in the sense of urban air quality, where the most significant source of pollutants, i.e. automobile emissions, is located near the pollutant receptors of concern, i.e. citizens [8]. These minor local vector factors are influenced and influence the flow at such a small scale. As the variability of the events has increased, characterising the flow under certain conditions continues to be considerably more difficult.

Once the length scales at which the urban problem can be assessed have been introduced, it is worth recalling that the majority of factors impacting pedestrian comfort and air quality, among others, must be evaluated at the street level. The urban-canopy sublayer is the area of the flow that is directly influenced by the obstacles. Oke [12] presented an analytical overview

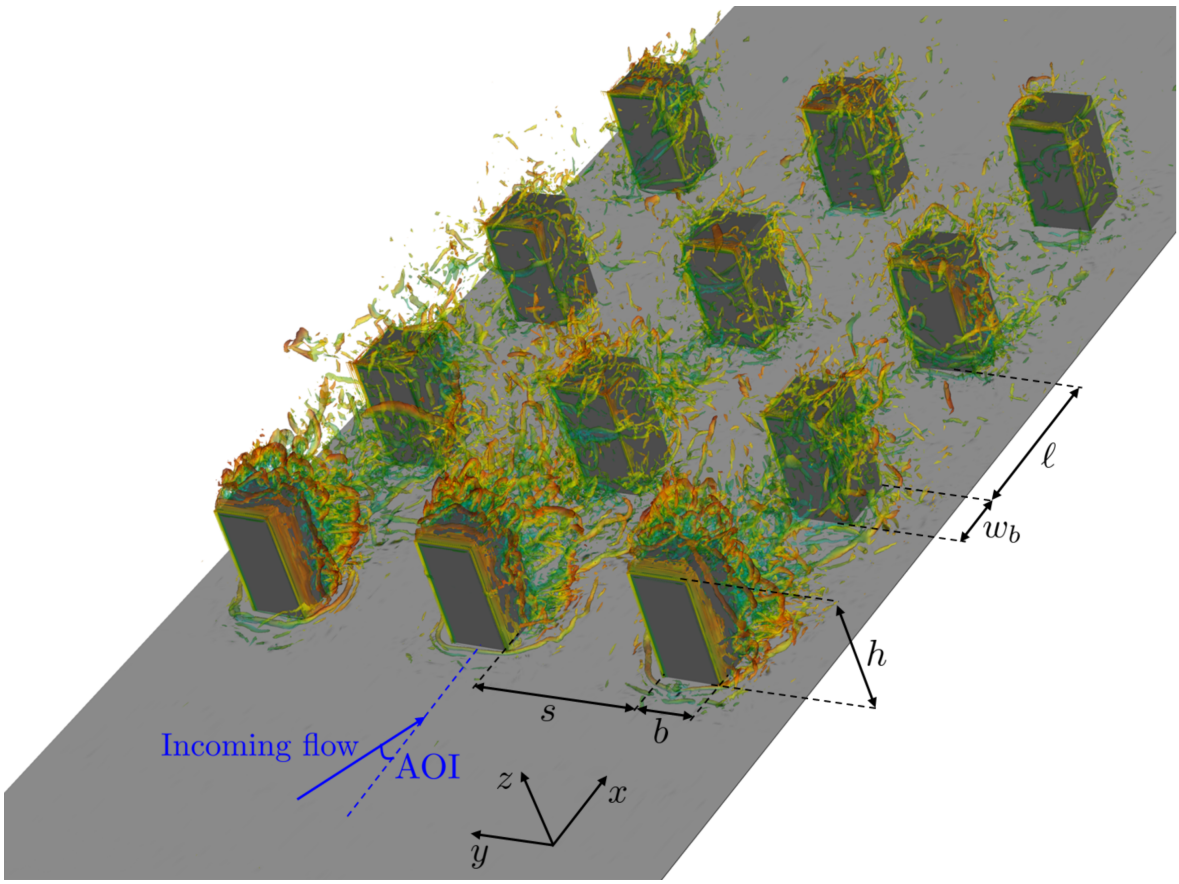


Figure 1.2: Geometrical parameters and angle of incidence (AOI) of the incoming flow definition in a simplified urban model. λ_2 method [10] is employed to characterise the vortical structures, which are coloured by streamwise velocity, from (dark blue) -1.2 to (dark red) $+1.8$. Extracted from Torres et al. [11].

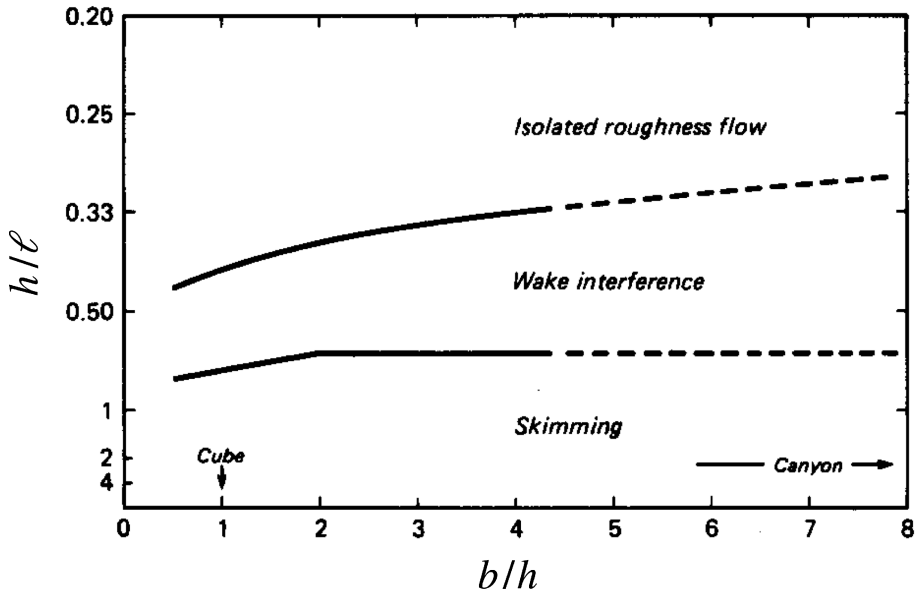


Figure 1.3: Threshold lines dividing the flow into three different regimes depending on the building (b/h) and the separation (h/ℓ). Extracted from Oke [12].

of the flow regimes apparent in this layer, identifying three flow regimes based on the ratio of street width to building height: in the case of narrow streets, the flow above the canopy can barely reach down to the street (*skimming flow*), and only one vortex can be seen between the obstacles; gradually broader streets lead to the *wake interference regime* first and then to the *isolated roughness flow* which shows much more contact with the flow above the roofline and dramatically more complicated physics than the previous cases. Figure 1.2 depicts the geometrical parameters defining an urban environment, together with the reference frame that will be used throughout the present. In particular, the width-to-height b/h and height-to-separation h/ℓ ratios need to be highlighted, since they will be relevant in determining the flow regimes identified by Oke. In Figure 1.3, the resulting flow regimes depending on the geometrical parameters that define an urban model are depicted. According to Oke [12], the height-to-separation ratio, h/ℓ, stands for the most critical parameter to define the resulting flow regime, as opposed to the width-to-height ratio, b/h, whose influence is not remarkable for considerably large ratios.

The study of Torres et al. [11] aims to evaluate Oke flow regimes [12] using a high-resolution large-eddy simulation (LES) with the high-order spectral code Nek5000 [13] over a simplified urban environment consisting of two wall-mounted rectangular obstacles. These simulations employ a spectral-element mesh with elements refined in the near-obstacle area, as well as a Gauss-Lobatto-Legendre (GLL) quadrature of eight points within each element to ensure the proper resolution. They were carried out on the Cray XC40 machine "Beskow" at KTH Royal Institute of Technology PDC Centre for High-Performance Computing. Further details regarding the simulation procedure can be found in Reference [14]. However, the geometrical parameters defining an urban environment shown in Figure 1.2 and the conclusions of Figure 1.3 highlight the importance of the width-to-height b/h and height-to-separation h/ℓ ratios, which determine the resulting flow regime. Being the h/ℓ ratio the most crucial one (recall Figure 1.3), the analysis of Torres et al. focuses on the variation of the h/ℓ ratio, keeping b/h constant.

In Figure 1.4, the discrepancies that occur in the flow regimes when the obstacle separation is changed are shown. The skimming-flow regime, obtained with $h/\ell = 1$, can be seen in Figure 1.4 (top), in which the flow above the obstacles can scarcely penetrate in the area between the two blocks due to the strong impact of the wake of the first obstacle. Figure 1.4 (middle), which depicts the wake-interference regime, shows a greater separation, resulting in $h/\ell = 0.5$. Whilst there is still some combined effects between the wake of the first building and the second block, it is apparent that some energy can be transferred between both obstacles from the flow above the canopy. Lastly, with $h/\ell = 0.25$, Figure 1.4 (bottom) shows the isolated-roughness regime. Owing to the high separation between blocks, a considerable momentum exchange will appear between this area and the flow above the canopy. These interactions lead, however, to more disorganised flow patterns, which hinder the obtaining of high-fidelity physical models.

Please note that these simulations are the ones that will be analysed throughout the present project in order to examine the changes appearing in the coherent structures of the flow when modifying the separation ratio between obstacles.

The study on flow and contaminant dispersion through building clusters was further investigated by Zajic et al. [15]. Hinged upon the classification of Oke [12] and multiple experiments on flow patterns across wall-mounted obstacles [16, 17, 18], the authors aimed to compare predicted flow patterns in well-known simplified configurations, e.g. those seen in Figure 1.4, with real atmospheric data obtained within a portion of Park Avenue in the Oklahoma City CBD[1]. They demonstrated that using canonical flow configurations as building blocks to simulate more complex multi-building configurations, e.g. an array of equispaced rectangular wall-mounted obstacles, performed well, at least in that situation. This conclusion makes first-order predictions of flow through canopies possible, such that field studies can be planned and interpreted and non-linear phenomena, which occur due to the strong interactions of canonical building processes, delineated.

They also contemplated the changes introduced when considering different building heights. This geometry variation, together with the previously discussed separation ratio h/ℓ, provokes a considerable alteration in the number of vortices appearing within the urban canopy [15]. In fact, the larger the ratios, the more vortices will appear. However, Zajic et al. [15] accentuated the importance of accountability for three-dimensional effects. Indeed, as a result of this 3D approach, the corner vortices appearing on either side of the obstacle modify the flow distribution inside the canopy, which is particularly noticeable for taller buildings, i.e. with a length-to-height ratio $w_b/h < 1$, since the side separation layer alters the recirculation bubble on the leeward face of the upwind building, resulting in intense turbulent fluctuations which may destroy the separation bubble [15]. Therefore, due to the three-dimensional approach, the flow analysis should be conducted in both the streamwise and spanwise coordinates.

Until now, the gaze has been kept on the classification proposed by Oke [12]. Nonetheless, several studies have focused their attention on the search for alternative categorisations of flow regimes in urban environments. One of those projects relies on Britter and Hunt [19], which seeks to obtain a general description of the flow, based on the interaction between

[1]The term CBD refers to the central business district, which is the central sector of urban centres and is characterised by a wide range in constructions of varying forms and sizes in (1–10) km^2. The regular disposition of buildings in new planned communities, such as Phoenix or Arizona, has steadily increased the interest in flow patterns in street canyons of CBDs over the past decade due to issues about pedestrian safety, internal and outside air quality.

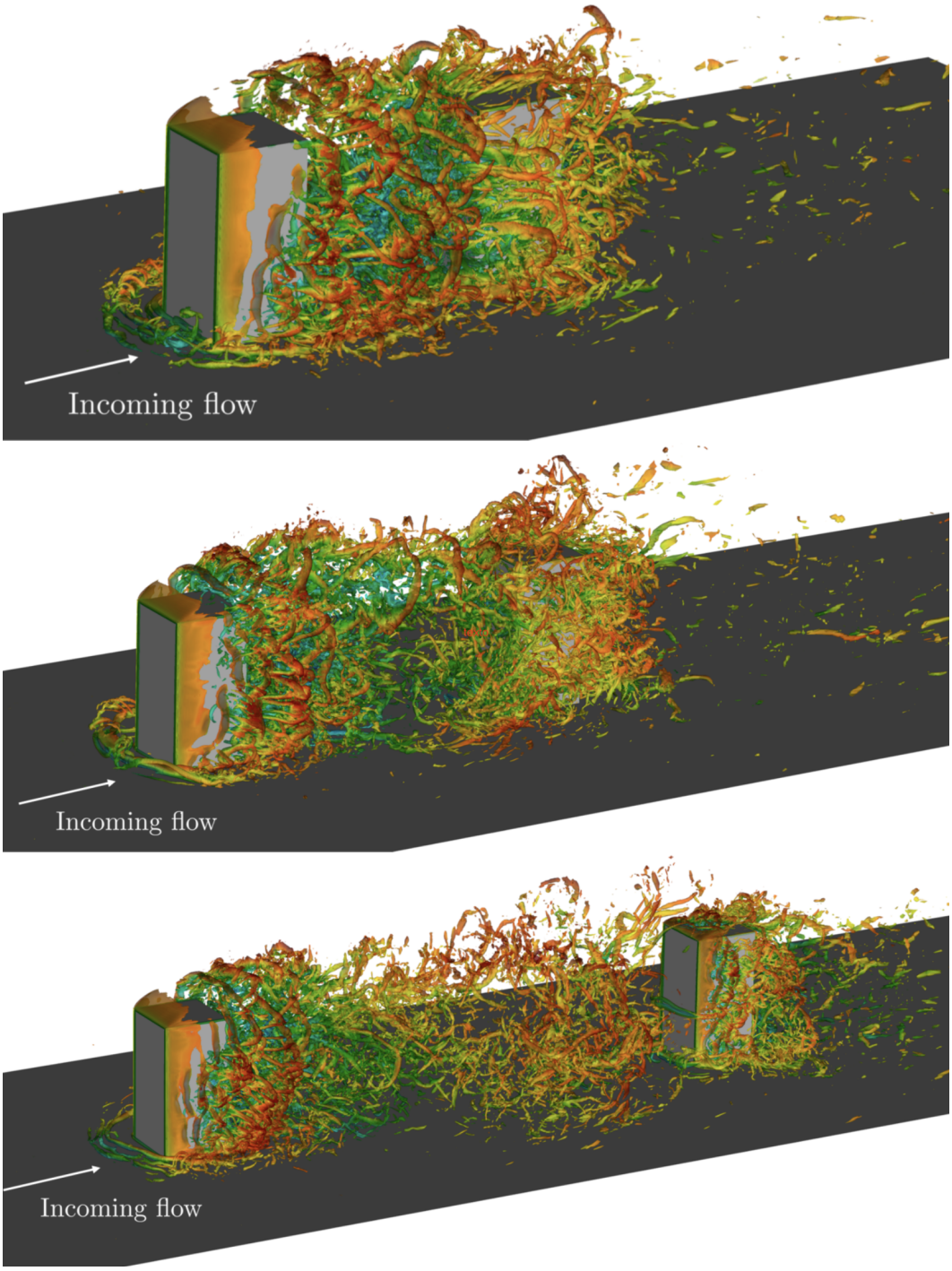


Figure 1.4: Flow regimes based on the obstacle separation, using Oke's criterion [12]: (**top**) skimming flow, (**middle**) wake interference and (**bottom**) isolated roughness. λ_2 method [10] is employed to characterize the vortical structures, which are coloured by streamwise velocity, from (dark blue) -1.2 to (dark red) $+1.8$. Extracted from Torres et al. [11].

bluff bodies, i.e. the two wall-mounted rectangular obstacles. In consequence, a *weak* and a *strong interaction* regime can be defined so that when the interaction between both bodies is practically null, the upwind block will be unaltered by the downwind building and vice versa. These regimes appear, analogously to previous works, as a result of different combinations of the height h, width b and separation ℓ of the blocks.

In such a fashion, for $\ell < 6h$ and $\ell < 3b$, the flow around the downstream block will be firmly influenced by the upwind one and, therefore, the flow regime will be of *strong interaction* type [19]. Whereas this regime will be clearly characterised by a stable vortex spanning the gap between the two blocks, increasing separations will give rise to higher fluctuations towards the ground. Wise and Penwarden [20] supported those verdicts and introduced a resembling analysis based on the height difference: higher speeds near the ground were the result of an increase in height of the downstream block, contrary to the results obtained by Isyumov and Davenport [21]. Britter and Hunt [19] discussed that the apparent contradictions between those results may reside in the understanding and interpretation of the flow measurements, which emphasises the need to account for three-dimensional effects, meaning that each velocity component possesses its own relevance.

1.3.1.1 Atmospheric boundary layer

Even though the turbulent vortical structures may play a prominent role in the pollutant dispersion through urban environments, there are other key factors that need to be considered in some specific applications. This is the case of the atmospheric boundary layer (ABL), which stands for that atmospheric region in contact with the Earth surface, in the same way as any boundary layer. This approach introduces some concepts that need to be analysed when working with ABLs. Fundamentally, the condition of atmospheric stability may change as a result of a given temperature distribution: whereas a stable environment will be characterised by a cooler (and denser) lifted parcel, compared with the surrounding air, which will tend to sink; an unstable environment, as a result of a warmer and less dense lifted parcel, will rise indefinitely. In the middle of those, when no temperature difference is appreciated with the surrounding air, no relative movement will be induced and neutral conditions will be met. The understanding of these conditions is essential for the atmospheric studies, including pollutant dispersion and deposition.

Among all the extensive analyses and discussions conducted with this approach, the work of Counihan [22] presented a thorough review of the meteorological literature on fully developed atmospheric boundary layers. Counihan found that the major anomalies appearing on both the interpretation of full-scale measurements and on the theoretical predictions of some boundary layer characteristics occurred due to the definition of the logarithmic law, with which the overwhelming majority of ABLs are approximated, allowing to its parametrisation as a single layer [22]. Besides, from all the analysed data, Counihan showed how the main boundary layer parameters, concretely those associated with the roughness scale, can be derived from one equation [22]: in the lower part of an ABL, the mean velocity follows the logarithmic behaviour of Equation 1.1.

$$\frac{\overline{u}_z}{u_\tau} = \frac{1}{\kappa} \ln\left(\frac{z-d}{z_0}\right) \tag{1.1}$$

where z_0 is the roughness length[2] and, together with d, depends on the type of terrain, u_τ is the friction velocity and κ the Von Kármán constant [23]. Note that since z_0 may differ from diverse surfaces, this method allows for the particularisation of each urban environment.

[2]The roughness length is a commonly used parameter in computational models to express the roughness of a given surface. It mostly influences the mechanical instability of the flow above the surface, resulting in less exchange between the surface and the atmosphere but higher winds against the ground for low roughness distances. Its value may vary from $z_0 = 0.001$ for ice, open sea or even desert to $z_0 = 0.3$ or $z_0 = 3$ for the suburban area and tall buildings (city centers), respectively.

Blocken et al. [24] also considered the prospect of achieving an atmospheric boundary layer (ABL) model by precise Computational Fluid Dynamics (CFD) simulations. However, at the bottom of the numerical domain, where wall-function roughness modifications, based on experimental results, are added, the precision of such simulations can be severely harmed. They consider easing the criteria in that area as an alternate approach, since it is difficult to meet all requirements for ABL flow simulations. This simplification, though, leaves out critical detail about turbulent coherent structures. This might be posed as the main reason why the literature is mostly concerned with the investigation of turbulent processes rather than the construction of ABL models.

1.3.2 Coherent flow structures

Understanding the chaotic processes that contribute to pollutant-concentration fields, wind patterns and gusts requires the ability to describe the coherent structures guiding the flow dynamics in urban environments. In other words, identifying them allows for the development of physical models, which are crucial in this type of flows due to their inherent complexity. These mechanisms have already been documented in a few experimental studies. However, there is a pressing need for interpretable and generalisable models: models as low dimensional as possible and with as few terms in the dynamics as necessary to explain how these variables interact dynamically in time [25]. As a result, numerical simulations allow the evaluation of their impact in dynamic processes due to their adequate spatial and temporal resolution.

Vinuesa et al. [26] used direct numerical simulations (DNS) to examine the effects of inflow conditions in flow structures around a wall-mounted square cylinder. While the sharp edges of the obstacle maintain the separation position independent of the inflow conditions, the resulting structures and topology of the wake further downstream vary greatly in both situations. The effect of the growing boundary layers on both sides of the obstacle, in particular, results in a different behaviour in both wakes: the wake associated with the laminar inflow is instantaneously broader, while the turbulent one is marginally wider in the time-averaged field. In addition, Vinuesa et al. [26] acknowledged that, in the turbulent-inflow simulation, the horseshoe vortex was modulated by the streamwise fluctuations. Henceforth, in both experimental and numerical studies, this emphasises the importance of having well-established inflow conditions so as to ensure flow reproducibility.

1.3.2.1 Experimental studies of coherent turbulent structures

Keeping this in mind, since the inflow conditions may well be the reason of different flow structures, Hunt et al. [27] reported the structures of laminar and turbulent flows around surface-mounted obstacles by means of flow-visualisation studies. Using a novel oil-film visualisation technique, the authors aimed at recognising the various flow patterns of shear-stress lines around cuboids of distinct shapes and with different incidence angles.

The information derived from the visualisation of the flow, enabled Hunt et al. [27] to clearly represent the general pattern of the streamlines. They found that the number of vortical structures upstream is sensitive to flow parameters, i.e. the aforementioned inflow conditions: up to seven vortices might be observed in some cases. Particularly, the schematic representation of the flow around a cuboid proposed by Hunt et al. [27] is depicted in Figure 1.6, where four different flow patterns have been identified: a) the horseshoe vortex, formed around the obstacle, b) the roof vortex, located in the upper section of the obstacle, c) the

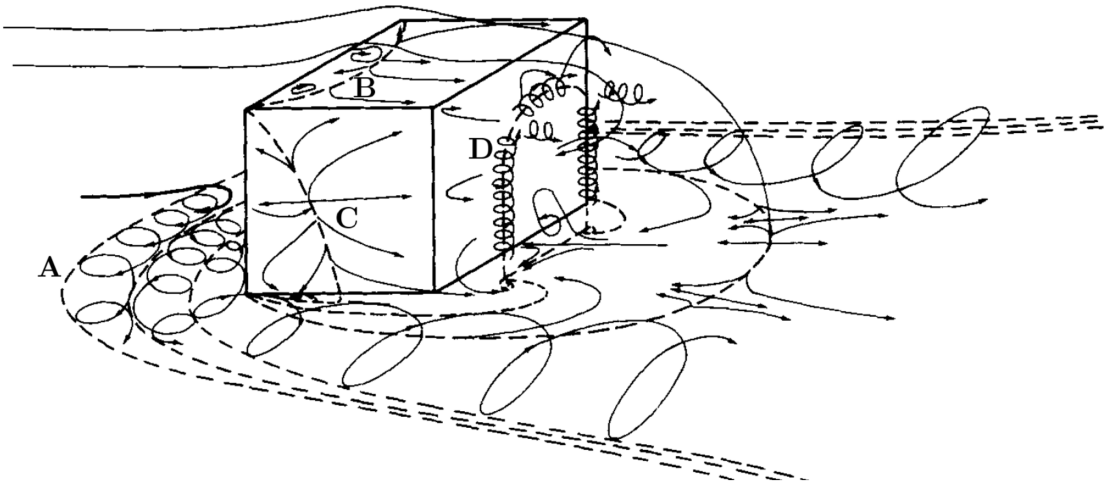


Figure 1.5: Schematic representation of the flow around a surface-mounted rectangular obstacle. Main appearing vortices are highlighted: (**A**) horseshoe vortex, (**B**) roof vortex, (**C**) vortices on the obstacle side and (**D**) arch vortex. Extracted from Hunt et al. [27].

side vortices, created on both lateral parts, having a strong interaction with the wake, and d) the arch vortex, established on the leeward side of the obstacle. Interestingly, with this representation, they proved the absence of a separation bubble or cavity in the wake of a surface-mounted bluff obstacle. Therefore, no closed surface is observed in the wake [27].

Meinders [28] extended the work of flow structures around wall-mounted cubes by analysing the interaction between the obstacles when more than one cuboid was considered. Similarly to the analysis conducted by Torres et al. in Figure 1.4, Meinders qualitatively examined the influence of the separation distance between the obstacles, h/ℓ, on the flow around an in-line tandem disposition of two cubes. For that, the author employed oil-film visualisations together with LDA[3] measurements.

It was shown that the h/ℓ variance only led to a substantial modification of the mean flow and turbulence statistics on the downstream obstacle. The separated shear layer reattached to the downstream side edge of the downstream obstacle for $h/\ell = 0.5$, which led to an inter-obstacle region characterised by an arc-shaped vortex, confined by the side flow and the separating shear layer. In addition, upstream of the front face of the leading cube, a horseshoe vortex was formed and deflected further downstream along the sides due to the presence of the second obstacle [28].

For $h/\ell = 0.25$, the flow reattached in the inter-obstacle region: the shear layer detached from the sides and top edges of the upstream obstacle and breached the inter-obstacle spacing before reattaching on the channel surface. Because of that flow reattachment, a second horseshoe vortex emerged in front of the downstream cube. The imprints of such vortices were clear from the oil-film visualisations, see Figure 1.6. Larger spacing ratios, $h/\ell < 0.125$ were seen to produce flow patterns very close to those around a single obstacle, i.e. the upstream and downstream flow structures were seen to be strongly similar [28]. Despite the fact that the mean flow and turbulence figures around the downstream cube were somewhat similar to

[3]LDA states for Laser Doppler Anemometer and it is an optical well-established technique ideal for non-intrusive velocity measurement in fluid dynamics applications. Its main advantages rely on the high spatial and temporal resolution, the ability to perform measurements in flows that are reversing and the no need for calibration.

those around the upstream cube, the upstream protuberance disrupted the flow, resulting in a very high degree of turbulence intensity as encountered by the downstream cuboid. These results are in good agreement with the flow regimes presented by Oke [12].

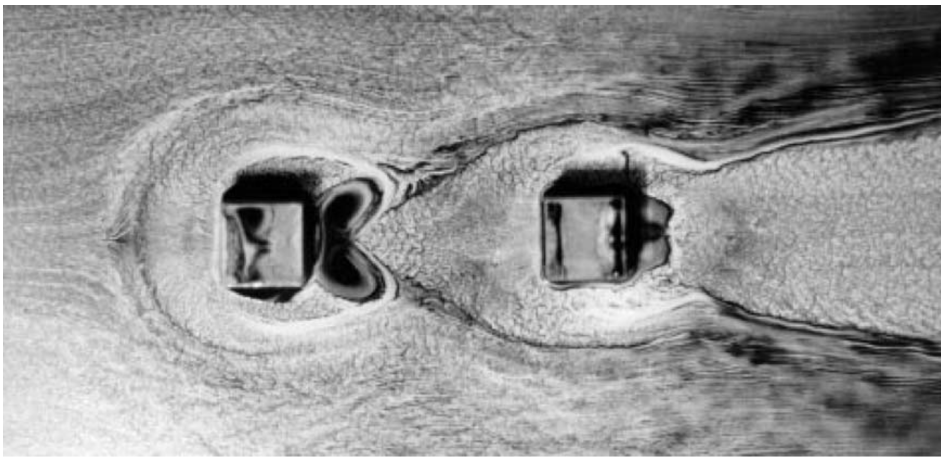


Figure 1.6: Oil-film visualisation of the surface flow pattern of an in-line tandem arrangement of cubes for $h/\ell = 0.25$ ($Re_H = 7000$). Extracted from Meinders and Hanjalić [29].

Additionally, Becker et al. [30] studied experimentally the structure of the flow field around three-dimensional rectangular block obstacles, representing generic shapes of buildings. Such a investigation provided further insights into the dependence of the flow structures on the aspect ratio of the obstacle and the inflow conditions: the type of boundary layer, the incidence angle and the Reynolds number. Using LDA measurements and oil-film visualisations in wind-tunnel experiments, where a suburban boundary layer was employed as inflow conditions, the authors analysed the arch-shaped vortices formed on the leeward side of the obstacle for different values of the incidence angle, AOI. These findings allowed the authors to define the evolution of the arch vortex, which Martinuzzi [31] previously depicted for AOI = 0°, for non-zero incidence angles. For increasing angles of attack, the main observation is the dislocation of one of the vortex's legs, which gradually rotates until it reaches the side of the obstacle for AOI = 45° (Figure 1.7). Curiously, raising the AOI to 60° causes the vortex leg to be displaced to the top of the obstacle, which has a major impact on the momentum-transfer processes within an urban canyon.

1.3.2.2 Methods for vortex identification

A wide range of criteria to identify vortical structures have been developed in the literature. Monnier et al. [32] aimed at identifying the coherent structures of the wind-tunnel flow around the geometry of the Mock Urban Setting Test (MUST) experiment[4] using different criteria. In this urban environment, the authors employed a set of 4×3 buildings of $b/h = 5$ and $w_b/h = 1$ (recall the definition of geometrical parameters in Figure 1.2). The experiment's goal was to identify the well-known arch vortex downstream the obstacles: a vortical structure of two legs and a roof, with each leg consisting of fluid rotating around the vertical axis and the roof rotating around the spanwise axis. As a result, they began by visualising the wall-normal and spanwise mean vorticities, ω_y and ω_z respectively, given by Equation 1.2. Using the modulus

[4]The Mock Urban Setting Test (MUST) was an urban disperse test performed at the US Army Dugway Proving Board (DPG) Horizontal Grid Test Site by the Defense Threat Reduction Agency (DTRA). The aim was to obtain near full-scale meteorological and dispersion data for the production and validation of toxic risk assessment models in urban areas.

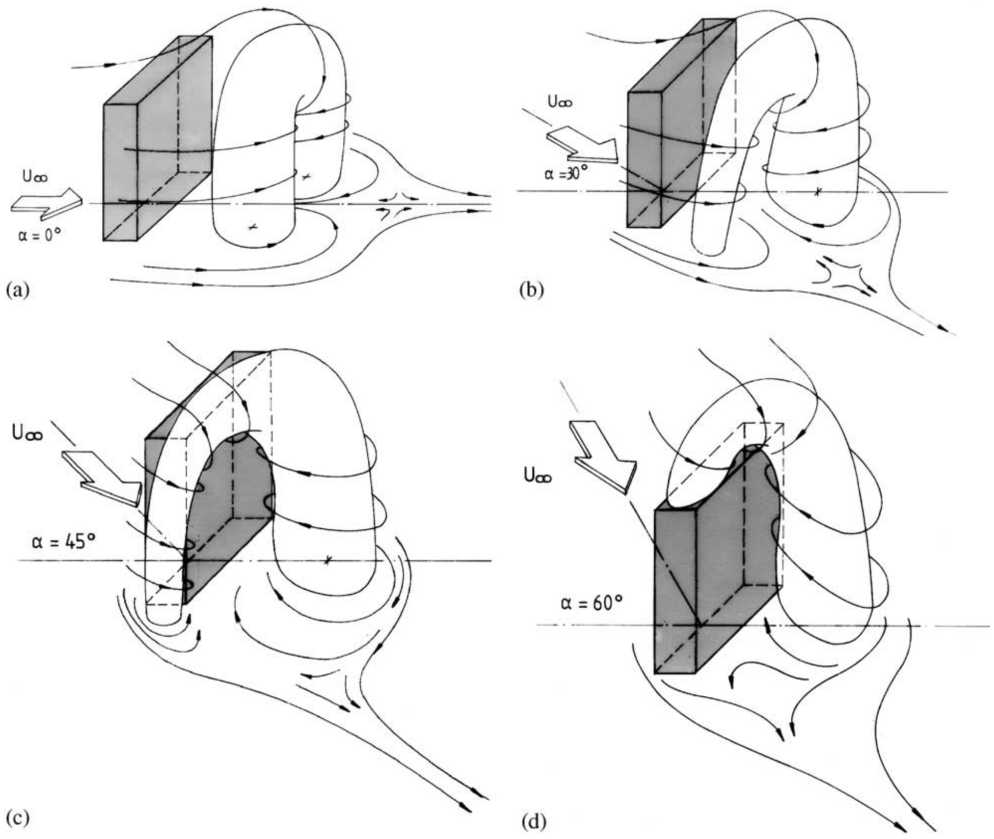


Figure 1.7: Arch vortex schematic representation around a three-dimensional rectangular block with $b/h = 2$ and $w_b/h = 0.29$ depending on the angle of attack: (**A**) AOI = 0°, (**B**) AOI = 30°, (**C**) AOI = 45° and (**D**) AOI = 60°. Extracted from Becker et al. [30].

of the spatially-averaged vorticity vector $\langle\omega\rangle = \sqrt{{\omega_x}^2 + {\omega_y}^2 + {\omega_z}^2}$, the authors were able to define a local threshold for correctly identifying coherent structures.

$$\omega\left(\mathbf{x}, t\right) = \nabla \times \mathbf{U} \tag{1.2}$$

Similarly to what Becker et al. [30] deduced, using various iso-levels for the vorticity quantities, Monnier et al. [32] discovered that when the incidence angle is non-zero with respect to the model, one leg of the arch vortex shrinks in size. Furthermore, the authors used the popular method introduced by Hunt et al. [27], known as the Q-criterion, to classify eddy regions. This approach, which is based on the second invariant of the velocity-gradient tensor, see Equation 1.3, has the advantage of defining coherent structures even when shear loads are large.

$$Q = -\frac{1}{2}\frac{\partial U_i}{\partial x_j}\frac{\partial U_j}{\partial x_i} = -\frac{1}{2}\left(S_{ij}S_{ij} - \Omega_{ij}\Omega_{ij}\right) \tag{1.3}$$

where S_{ij} represents the rate of the strain tensor and Ω_{ij}, the rate of the rotation tensor, which is related to the aforementioned vorticity with $\omega_i = \epsilon_{ijk}\Omega_{kj}$, being ϵ_{ijk} the Levi-Civita symbol [32]. The representation of an iso-Q surface showed two clear recirculation regions: on the entire span of the leeward side of the upstream obstacle and closer to the ground on the windward side of the downstream block [32]. Comparable results were obtained with λ_2 method. In this case, this proposal seeks to analyse the eigenvalues of the $S_{ik}S_{kj} + \Omega_{ik}\Omega_{kj}$ tensor, where the symmetric and anti-symmetric parts of the tensor play a role. Particularly,

only when at least two of its eigenvalues are negative, a point of the velocity field will define a vortex core. Therefore, if $\lambda_1 \leq \lambda_2 \leq \lambda_3$; then, $\lambda_2 < 0$.

Finally, in addition to the velocity gradient techniques, Monnier et al. [32] utilised the normalised angular momentum technique, defined in Equation 1.4, which Sousa [33] used to pinpoint the centre of the vortical structures downstream of a single cuboid obstacle. They concluded that this approach improves large-scale vortical structures identification over vorticity magnitude because, unlike previous methods, it is an integral-based quantity that does not require spatial derivatives of the velocity field.

$$\Gamma_1(x_p) = \frac{1}{D}\int_D \frac{(x - x_p) \times u(x)}{|x - x_p||u(x)|}\, dx \tag{1.4}$$

where x_p is the spatial position at which Γ_1 is computed, D, the domain over which the the integration is performed, $u(x)$, the velocity vector, $|\cdot|$, the norm quantity and $\times$, the vector product.

Monnier et al. [34] used this method to describe the relationship between the arch vortex and high-turbulence areas. Figure 1.8 depicts the resulting isosurface Γ_1 with rms streamwise and spanwise turbulence high intensity areas. Due to shear layers developing on both sides of the block in the AOI = 0° event, see Figure 1.8 (left), two regions of significant streamwise velocity fluctuations emerge. The arch vortex is then found between these two areas. On the other side, the zone of high spanwise velocity found along the downstream block's windward edge is centred immediately after the arch vortex. As a result, Monnier et al. [34] concluded that the arch vortex is situated in high turbulence areas, while its centre is located in a low turbulence zone. As the angle of attack is increased, see Figure 1.8 (right), the arch vortex appears to be rotated with respect to the street axis. Besides, the velocity fluctuations are displaced, too, allowing the arch vortex to stay in a low turbulence area. Please note that in this experiment the obstacles used to simulate buildings in an urban area are relatively high with respect to the street separation, thus presenting features of the skimming flow regime [12].

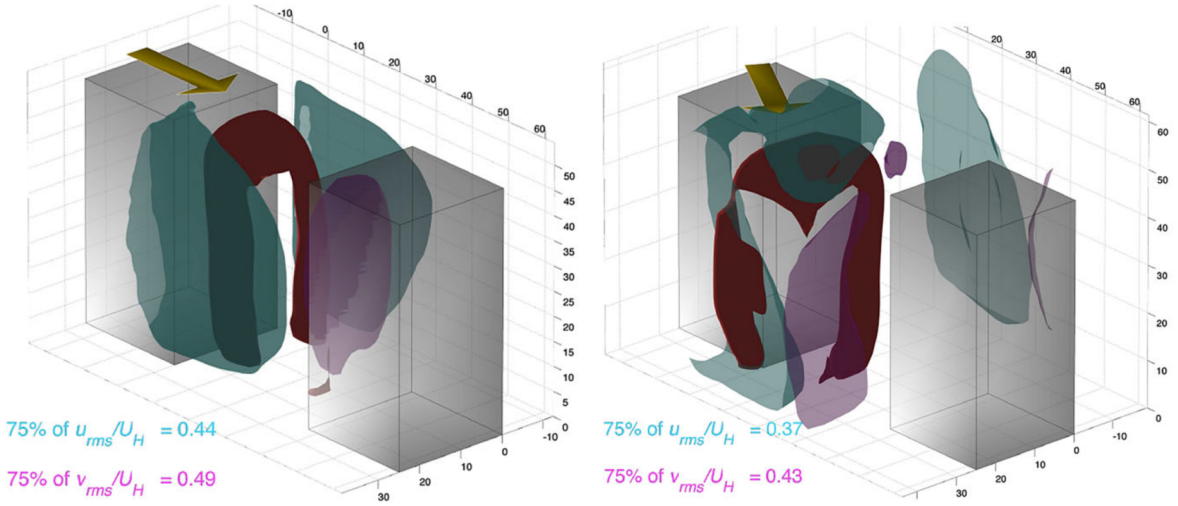


Figure 1.8: Representation of the iso-surfaces of $\Gamma_1 = 0.4$ for (**left**) AOI = 0° and $\Gamma_1 = 0.35$ for (**right**) AOI = 30°, streamwise (green) and spanwise (purple) rms velocity fluctuations, representing the 75% of the maximum value reported in each panel. Extracted from Monnier et al. [32]

1.3.3 Final comments on historical perspective

Finally, to wrap up this part, some final thoughts on the reviewed literature and the current project will be made. Previously, a general overview of research projects conducted with the aim of further understanding the flow inside urbanised areas was portrayed. In this regard, the importance of recognising the spatio-temporal features of the dynamics driving the flow has been emphasised.

There is currently an urgent need to acquire generalisable and interpretable models in order to perform more accurate flow predictions. As a result, the present project will concentrate on the study and identification of the coherent structures responsible for the behaviour discussed throughout this section. To that end, numerical simulations allow data-driven methods to conduct a high-quality modal analysis of three-dimensional instantaneous fields. This will enable us to gain a better understanding of the flow physics. Following sections will present and describe the state-of-the-art techniques that will be used to achieve a better understanding of the complex mechanisms that occur in urban flows.

2

Theoretical background

Having reviewed the past and current state of scientific research on urban flows, the goal of this section is to provide the reader with a sufficient understanding of the physical laws that govern turbulent flows, i.e. the flow within urban environments, and the mathematical tools used throughout the present to gain further insight into the physical processes responsible for these type of complex flows.

2.1 Theory in fluid mechanics

Prior to the description of the mathematical tools that will allow for the obtaining of a reduced-order model (ROM) of the flow simulation data, it is advisable to give some guidelines over the empirical laws governing the fluid physics, which might be posed as the very first step towards the analysis of any kind of numerical simulation.

Flows, in the same way as any physical process, can be classified based on various of their characteristics. For instance, flows can be classified as compressible or incompressible depending on the relevance of density changes; inviscid or viscous based on the importance of inertial effects compared to viscous (friction) effects; or steady and unsteady flows according to their time dependence. However, one of the most common ways of classifying flows is to consider their inherent behaviour, which leads to laminar and turbulent flows. This distinction is critical since it might well be the reason for using different assumptions or problem-solving methods. Henceforth, since the case study will only deal with turbulence, just turbulent flows and their characteristics will be examined from now on.

2.1.1 The nature of turbulent flows

It is common knowledge that turbulent flows may be observed in our everyday surroundings, ranging from water in a river to smoke from a chimney. According to Pope's definition in his book *Turbulent flows* [35], turbulent flows, e.g. those witnessed in a waterfall, are characterised by being unsteady, irregular, apparently random and chaotic, and the velocity of every eddy or droplet is unexpected. This is regarded as a key feature for the study of turbulent flows since it dictates how fluid velocity fluctuates dramatically and randomly in both time and position. However, although the velocity fluctuations over the time-averaged field can be seen to fluctuate considerably (approximately a quarter of the mean velocity in certain

circumstances), far from being unstable, both the fluctuations and the mean can be argued to be *stable*: neither enormous changes of the whole flow field are visible, nor the fluctuations spend lengthy periods of time at values that are notably distant from the mean [35]. Figure 2.1 displays the time history of the axial component of the velocity $U_1(t)$ of a turbulent jet to demonstrate this. It can be seen here that the velocity is not periodic and fluctuates across a wide range of time scales, reinforcing the aforementioned traits of turbulent flows.

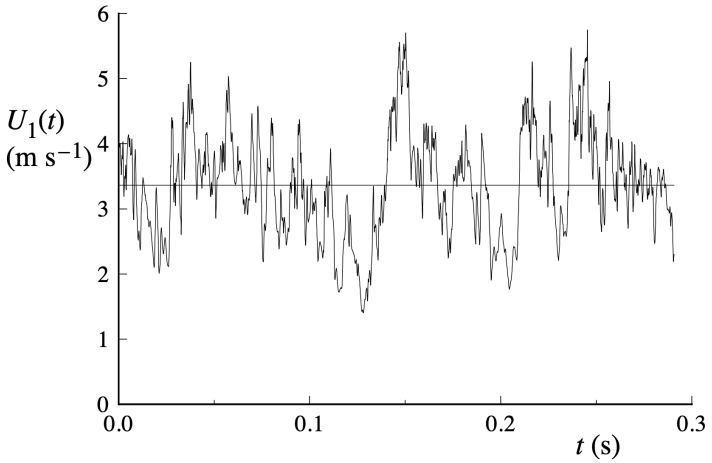


Figure 2.1: Time history of the axial component of velocity $U_1(t)$ on the centre line of a turbulent jet. From the experiment of Tong and Warhaft [36]. Extracted from Pope [35].

Taking this into consideration, another essential feature of turbulence is its capacity to transfer and mix fluid much more efficiently than laminar flows. When various fluid streams must be combined, it is normally preferable for this mixing to occur as quickly as feasible. This is when turbulence comes into play in the case of pollution streams thrown into the atmosphere. Osborne Reynolds vividly proved this experimentally, aiming at analysing the propagation of dye steadily injected on the centre line of a lengthy conduit through which water was flowing. Reynolds discovered that this sort of flow can be described by a single non-dimensional quantity, which is now known as the Reynolds number.

$$Re = \frac{Ul}{v} \tag{2.1}$$

In this sense, it may be interpreted as the ratio of dynamic forces to viscous forces, which aids in defining specific thresholds above which the flow is deemed turbulent and below which the flow is considered laminar. During Reynolds' experiment, if $Re < 2300$, the flow was assumed to be laminar, i.e., the fluid velocity did not vary with time and all streamlines were parallel to the pipe's axis [35]. If, on the other hand, $Re > 4000$, the flow should be turbulent, increasing mixing with the surrounding water and reducing the peak dye concentration downstream of the injection point.

2.1.2 The equations of fluid motion

The Navier-Stokes equations, which govern the flow of Newtonian fluids with constant properties, will be examined briefly in this section. To that purpose, the work once published by Pope in *Turbulent Flows* [35] will be followed. The goal is to present a broad overview of the equations that must be solved in order to represent flow in an urban context. Please

bear in mind that this is merely a high-level overview of the equations; for a more in-depth examination of the numerical implementation, see Ref. [14].

2.1.2.1 Continuum fluid properties

Real fluids are all made up of individual molecules that are far too tiny for us to see directly. Because the molecules have their own motion, they ultimately collide. In fact, fluid motion is the average motion of many particular molecules, and the independent motion of each molecule is what is known as pressure and heat on the macroscopic level. As a result, modelling a fluid as a continuous medium eliminates the need to address the behaviour of each individual molecule, a crucial simplification. However, despite the potency of the continuous concept, it should always be remembered that fluid is made up of molecules. Hence, a fluid parcel is defined as a collection of molecules that occupy a simply interconnected region of space that is substantially larger than the mean free path, i.e. the average distance between molecules.

To support this hypothesis, Pope analysed the length and time scales at both the microscopic and macroscopic levels and concluded that the flow scales outnumber the molecular scales by three or more orders of magnitude, proving the validity of the continuum hypotheses for such types of flows [35]. This difference of length scales is quantified by the Knudsen number:

$$Kn \equiv \lambda/\ell \tag{2.2}$$

where λ is the mean free path and ℓ, the smallest geometrical length scale in a flow. Therefore, if $Kn \ll 1$, the continuum hypothesis is considered to hold, and once invoked, all concepts of the fluid's discrete molecular nature and molecular scales cease to be relevant. This is the approach that will be followed throughout the present work.

2.1.2.2 Continuity equation

The continuity equation is one of the most significant equations in the area of fluid mechanics since it employs the mass-conservation principle, which states that no mass may be generated or destroyed throughout the process. Considering a generic control volume, the mass-conservation or continuity equation may thus be formally stated with Equation 2.3. Please keep in mind that the reader should be already aware of the derivation and interpretation of this equation in terms of control volumes and the Eulerian and Lagrangian fields.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0 \tag{2.3}$$

The time derivative term indicates the accumulation (or loss) of mass within a given system, whereas the divergence term indicates the difference between flow in and flow out in a given control volume. Furthermore, depending on the type of flow being studied, the above general equation allows for many simplifications. In this context, the equation may be simplified such that the velocity field is divergence-free for constant-density flows, i.e. flows where density does not vary with position or time, which is known as the continuity equation for incompressible flows.

2.1.2.3 Momentum equation

Although the continuity equation encompasses a large amount of essential information, an equation that links the forces applied in a specific control volume with its acceleration should now be described. To that aim, the momentum equation, which is based on Newton's second law, connects the fluid particle acceleration to the total of the forces experienced by the fluid. In general, such forces may be split into two categories: surface forces and body forces. On one side, surface forces are of molecular origin and are characterised by the symmetric stress tensor τ_{ij}. Body forces, on the other hand, are those applied to the whole control volume. In that respect, gravity is the body force of interest in the study of turbulent flows. Its mathematical model allows for a wide range of definitions, but based on Pope's formulation [35], a gravitational potential should be specified so that the body force per unit mass is provided by Equation 2.4.

$$\mathbf{g} = -\nabla\Psi \tag{2.4}$$

Consequently, if a constant gravitational field is considered, then, the gravitational potential is given by $\Psi = gz$, where g is the gravitational acceleration. Finally, by direct application of Newton's second law,

$$\rho\frac{\mathrm{D}U_j}{\mathrm{D}t} = \frac{\partial\tau_{ij}}{\partial x_i} - \rho\frac{\partial\Psi}{\partial x_j}\,. \tag{2.5}$$

By substituting the corresponding expression of the stress tensor for Newtonian fluids with constant properties into the general momentum equation, the well-known Navier-Stokes equations are obtained. For their particular derivation, the fact that both the density and the viscosity are constant needs to be exploited together with the continuity equation particularised for incompressible flows.

$$\rho\frac{\mathrm{D}U_j}{\mathrm{D}t} = \mu\frac{\partial^2 U_j}{\partial x_i\partial x_i} - \frac{\partial P}{\partial x_j} - \rho\frac{\partial\Psi}{\partial x_j} \tag{2.6}$$

Finally, introducing the concept of the modified pressure, $p = P + \rho\Psi$, the above equation simplifies to Equation 2.7. Please note that the concept of the total derivative has been used throughout the derivation with the objective of considering both the spatial and temporal variation of the velocity.

$$\frac{\mathrm{D}\mathbf{U}}{\mathrm{D}t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{U} \tag{2.7}$$

2.1.2.4 Passive scalar equation

Hitherto, the continuity and momentum equations have been introduced. Nevertheless, in addition to the velocity, a conservation equation for a conserved passive scalar denoted by $\phi(\mathbf{A}, t)$ can be defined so that it can represent several physical properties of the flow. For instance, it can be a small change in temperature or the concentration of a given trace gas. In those cases, the diffusivity term of Equation 2.8 should change according to the passive scalar considered, e.g. the thermal diffusivity or the molecular diffusivity in the aforementioned cases.

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} = \Gamma\nabla^2\phi \tag{2.8}$$

It should be noted that for the scalar meeting the requirement of passive, its value should have no influence on the material characteristics and hence no influence on the flow. This criterion yields an essential scalar property: its boundedness. If the beginning and boundary values of $\phi(\mathbf{A}, t)$ are within a particular range, then ϕ must be within that range for any location or temporal instant [35].

2.1.2.5 Vorticity equation

The equations described above are seen to be essential in fluid mechanics: they have to be introduced anytime the behaviour of a flow, whether laminar or turbulent, is attempted to be found. However, turbulent flows are generally recognised to be rotating flows, which yields to one of its essential features, i.e. they have non-zero vorticity. Therefore, an extra equation must be provided to finish the exposure on the fluid mechanics equations.

From a formal point of view, vorticity $\omega(\mathbf{x}, t)$ is a measure of a fluid's rotation and may be mathematically represented as the curl of the velocity. It is specifically equivalent to twice the rate of rotation of the fluid at a given location and time.

$$\omega = \nabla \times \mathbf{U} \tag{2.9}$$

Accordingly, the equation for the evolution of the vorticity can be obtained by applying the curl to the Navier-Stokes equations presented in 2.7. Finally, Equation 2.10 states the vorticity equation in turbulent flows.

$$\frac{\mathrm{D}\omega}{\mathrm{D}t} = \nu\nabla^2\omega + \omega \cdot \nabla\mathbf{U} \tag{2.10}$$

2.1.3 Definition of streamlines

An important concept within the field of fluid mechanics concerns the idea of streamlines. A streamline is a path mapped out by a mass-less particle moving with the flow. Because the streamline is drawn out by a moving particle, the velocity is tangent to the path at every point along the path. Furthermore, mass cannot cross a streamline because there is no normal component of velocity along the path. Since no mass can flow through the surface of a particular object, its surface functions as a streamline. Bernouilli's equation confirms that the mass contained between two successive streamlines remains constant throughout the flowfield [37]. Regarding its mathematical derivation, once the velocity field of a given flowfield has been obtained at a certain time instant, a streamline is defined as a set of curves satisfying

$$\frac{\mathrm{d}\vec{x}_s}{\mathrm{d}t} \times \vec{u}(\vec{x}_s) = 0 \tag{2.11}$$

where $\vec{x}_s(s)$ denotes the parametric representation of a single streamline at a given moment in time. In such a fashion, if the velocity field can be decomposed in its components $\vec{u} = (u, v, w)$ and the parametric representation of a streamline, $\vec{x}_s = (x_s, y_s, z_s)$, then, it can be deduced

$$\frac{\mathrm{d}x_s}{u} = \frac{\mathrm{d}y_s}{v} = \frac{\mathrm{d}z_s}{w} \tag{2.12}$$

which demonstrates that the velocity vector is parallel to the streamline. They are also computed instantaneously, which means that they are calculated across the fluid from the instantaneous field at a certain time instant.

2.2 Computational theory

Following the introduction of the governing equations of turbulent flows, a general review of the technique to be followed in any numerical simulation will be now undertaken. Please bear in mind that the goal of this work is not to offer a full analysis of the solution techniques of such governing equations, but to conduct a deep examination of the results acquired from a prior numerical simulation, i.e. utilising a high-fidelity database, and, to that end, certain broad notions of the workflow in computational simulations must be retrieved.

Numerical simulations (CFD) constitute a powerful tool for obtaining detailed information on a specific fluid flow case without the need for intrusive measurements. However, in order to acquire accurate enough results, one must be very meticulous, which demands a good foundation in the linked science, understanding of a specific application and competence in CFD modelling. As a result, CFD processes comprise a number of phases, each of which is critical to the ultimate accuracy of the results. This CFD workflow can be broken into three main stages: pre-processing, calculation (solver) and post-processing.

Pre-processing. The very first step in any numerical simulation is the conceptual model of the problem. During this stage, the project objectives, i.e. the expected results and accuracy, should be established. Additionally, potential problem simplifications, such as problem dimensions, symmetry or periodicity, should be considered, as well as a study of the available literature up to that point. After that, the numerical domain should be defined with the determination of the boundary conditions (BC) far enough away from the region of interest. This region will be characterised by a particular geometry, through which the flow is wanted to be simulated, and needs to be created with special care, removing the non-relevant features which may hinder the final results or the intermediate process.

On the other hand, it is at this point that the mesh topology, i.e. the discrete scheme domain points over which the governing equations will be solved, must be determined. That includes defining the type of mesh components to be used, whether it will be structured or unstructured, and the refinement in high-gradient zones (boundary layer regions). This section is critical since it determines the accuracy and speed of the computation. In reality, if all flow scales are to be solved, i.e. a direct numerical simulation (DNS), the mesh element total number must be exceedingly large. Furthermore, the selection of fluid characteristics and boundary conditions fall within the scope of the pre-processing stage.

Solver. Once the domain has been fully discretised, the governing equations, i.e. differential equations, should be solved across each discretised volume. This necessitates the use of some numerical configuration, which includes the kind of solver, such as pressure or density-based solvers, as well as the order of the method, which is based on spatial and temporal discretisation. The significance of this section is based on its impact on convergence speed and accuracy. Low-order methods, for instance, will be linked with rapid convergence rates but low accuracy. This is identical to what was described during the mesh operation, and it results in two types of errors to account for in fluid numerical simulations.

Post-processing. Having solved the governing equations for the whole domain, the next step is to undertake an analysis to derive some meaningful physical implications from the flow results. Post-processing techniques can vary depending on the application, ranging from simple scene representations, e.g. contours, vectors, or path-lines, to more advanced mathematical analyses.

It is worth noting that data-driven methods are included at this stage since they aim to extract some physical features from high-fidelity databases, which may be acquired through numerical simulations or experiments. Therefore, the following sections will concentrate on the description of the tools employed during this workflow stage.

2.3 Modal-decomposition theory

In the study of fluid mechanics, specific physical properties can be found to be shared across a variety of flows and even across a large range of parameters. As an initial stage in the study of complicated flows, it has become normal practice to seek for and extract physically essential characteristics or modes. Common flow characteristics include von Kármán shedding, Kelvin-Helmholtz instability and vortex merging/pairing. The fact that these traits are frequently easily recognised via basic visual examinations of the flow leads us to believe that they can be retrieved using some mathematical approach [38]. Furthermore, advances in computational tools and experimental observations result in large-scale high-fidelity data, for which establishing reduced-order models to comprehend and describe their dynamical behaviour is critical. It should be noted that these spatio-temporal features or modes can be generated using flow-field data or the governing equations. Because of the intrinsic complexity of urban flows, the goal of this work is to conduct modal-decomposition analyses using flow-field data as input. These methods will be referred to as purely data-based methods.

The purpose of this section is to describe the approaches for performing modal decompositions over complex flows, which will be used to gain a better understanding of the flow through urban environments. Particularly, this section focuses on modal-analysis techniques such as proper orthogonal decomposition (POD), singular value decomposition (SVD), dynamic mode decomposition (DMD) and higher-order DMD (HODMD). The explanation offered here follows the one provided by Le Clainche and Vega in their book *Higher order dynamic mode decomposition and its applications* [39] as well as by Brunton and Kutz in *Data-driven science and engineering* [40]. For a more detailed discussion of modal decompositions, please refer to the cited works.

2.3.1 Singular-value decomposition

Singular-value decomposition (SVD) is a technique that extends eigenvalue decomposition to non-square matrices and is one of the most significant in the area of matrix factorisation. Among its various applications, data reduction should be highlighted since, from a high-dimensional dataset, SVD can extract the essential correlated patterns that improve the level of understanding of the data and allow for reliable low-order matrix approximations [38]. Because of the basic linear algebra on which it is based, as well as its scalability, this mathematical technique has become common in the data-driven world, allowing large corporations such as Google and Facebook to construct certain page rank and facial recognition algorithms.

Let us now consider a matrix $\mathbf{X} \in \mathbb{C}^{J\times K}$ representing a large dataset where each of the columns $x_k \in \mathbb{C}^J$ may represent measurements from simulations or experiments. For instance, in the case of fluid numerical simulations, columns represent the evolution of the flow state in a set of discrete points, i.e. snapshots.

$$\mathbf{X} = \mathbf{V}_1^K = [v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_{K-1}, v_K] \tag{2.13}$$

where v_k represents the field variable evaluated at time instant t_k and is defined as $v_k = v(t_k)$ for brevity. The snapshot matrix has dimensions of $J \times K$, where J represents the spatial degrees of freedom of the evaluated data, i.e. the total number of grid points defining the spatial domain (in three-dimensional computational domains with structured and uniform meshes, $J = N_x \times N_y \times N_z$, where N_x, N_y and N_z represent the number of points along with the streamwise, normal and spanwise directions), and K, as previously mentioned, indicates the number of snapshots.

The SVD allows for obtaining a unique matrix factorisation that exists for any matrix $\mathbf{X} \in \mathbb{C}^{J\times K}$:

$$\mathbf{X} = \mathbf{W}\mathbf{\Sigma}\mathbf{T}^{\mathrm{T}} \tag{2.14}$$

where $\mathbf{W} \in \mathbb{C}^{J\times J}$ and $\mathbf{T} \in \mathbb{C}^{K\times K}$ are known as left and right singular vectors or modes, respectively, and they are unitary matrices[1] such that each column is orthonormal, and $\mathbf{\Sigma} \in \mathbb{R}^{J\times K}$ is a diagonal matrix with non-negative values, known as singular values σ_k, hierarchically ordered. Note that the superscript $(\cdot)^{\mathrm{T}}$ denotes the complex conjugate transpose of a matrix, which corresponds to the regular transpose for real-valued matrices.

$$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_K \geq 0 \tag{2.15}$$

It is worth noting that the columns of the matrix $\mathbf{W}$ are known to have the same shape as columns of $\mathbf{X}$. Therefore, if $\mathbf{X}$ represents the evolution in time of a given flow-field, each u_k can be named as *eigenflowfields* hierarchically arranged so that u_1 is somehow more relevant than u_2 in terms of describing the variance in the columns of $\mathbf{X}$. This means that each u_k can be reshaped into a given flow-field, which will represent a certain flow feature. On the other side, matrix $\mathbf{T}$ represents the *eigentime series* of each *eigenflowfield*, meaning that each spatial mode can be associated with some temporal coefficients.

Aiming at illustrating this, given a matrix $\mathbf{A}$, it can be graphically decomposed using the SVD approach as shown in Figure 2.2. It could happen, and it is usually the case in the modal analysis of fluid flows, that the length of v_k does not coincide with the number of snapshots considered. Figure 2.2 shows an example of $J > K$, where the components in $\mathbf{W}$ encompassed by broken lines are removed from the decomposition since they are multiplied by zeros in $\mathbf{\Sigma}$. Accordingly, the economy SVD is the decomposition that ignores the submatrices in the broken-line boxes, as opposed to the full SVD, which takes into account all of the components involved in the decomposition. Therefore, it is possible to fully represent the matrix $\mathbf{X}$ using the economy SVD.

$$\mathbf{X} = \hat{\mathbf{W}}\hat{\mathbf{\Sigma}}\mathbf{T}^{\mathrm{T}} \tag{2.16}$$

[1] Any square matrix is unitary if it satisfies that $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$.

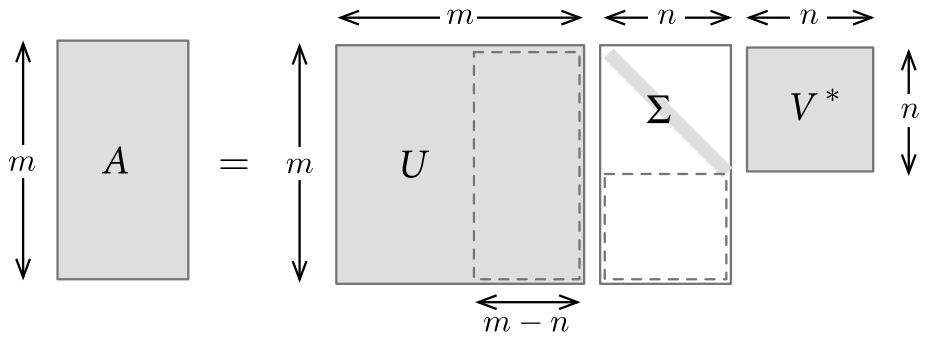


Figure 2.2: Schematic representation of the matrices involved in the full and economy singular value decomposition. Broken lines are used to represent the part of the matrices which is not considered by the economy SVD as they are multiplied by zeros in the singular value matrix. Extracted from Taira et al. [38].

Please note the slight difference between Equations 2.14 and 2.16, where $\hat{\mathbf{W}}$ and $\hat{\mathbf{\Sigma}}$ denote the reduced matrices, which ignore the broken-line boxes of Figure 2.2.

Another variation of the singular value decomposition occurs from truncating the singular value matrix at rank N such that non-zero values are removed.

$$\mathbf{X} \approx \tilde{\mathbf{W}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{T}}^{\mathrm{T}} \tag{2.17}$$

This truncation results in a relative root mean square (RMS) difference between the original and reconstructed matrix $\mathbf{X}^{\mathrm{approx}}$, i.e.

$$\mathrm{RRMSE} = \frac{\|\mathbf{X}^{\mathrm{approx}} - \mathbf{X}\|_2}{\|\mathbf{X}\|_2} = \sqrt{\frac{\sigma_{N+1}^2 + \ldots + \sigma_K^2}{\sigma_1^2 + \ldots + \sigma_K^2}} \tag{2.18}$$

As a result, in order to ensure that the relative RMS is under a set threshold, the truncated SVD allows for the imposition of the following:

$$\frac{\sigma_{N+1}}{\sigma_1} \leq \varepsilon_{\mathrm{SVD}} \tag{2.19}$$

2.3.2 Fourier transforms

Historically, it has been common in the field of engineering mathematics to seek new coordinate systems into which equations may be transformed so as to simplify their calculation and analysis. While researching the theory of heat, Joseph Fourier [41] invented in 1822 what is now perhaps the most widely used coordinate system transformation. Fourier found that sine and cosine functions of increasing frequency may be employed as orthogonal bases for the heat equation, resulting in a revolution in analytical and computational mathematics. Fast forward two centuries, an efficient method for performing the Fourier transform, known as the fast Fourier transform (FFT), has played as important a role in moulding the contemporary world as any other algorithm to date.

The method underlying the Fourier transforms has previously been addressed throughout this chapter in the context of data analysis utilising the SVD matrix factorisation and will

be examined further in this section. The Fourier analysis has a wide range of applications: it may be used to solve partial differential equations (PDEs), i.e. to compute derivatives, to denoise data or to compress audio or picture files. However, hereinafter, the focus will be on the analysis of temporal data in order to extract additional information about the flow dynamics using the Fourier transformation.

2.3.2.1 Fourier series

One of the most relevant results within Fourier's work [41] was the representation of a periodic and piece-wise smooth function $f(x)$ in terms of an infinite sum of sines and cosines of increasing frequency, i.e. the Fourier series. Particularly, if $f(x)$ is an L-periodic function, it can be rewritten as:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[a_k \cos\left(\frac{2\pi k x}{L}\right) + b_k \sin\left(\frac{2\pi k x}{L}\right) \right] \tag{2.20}$$

where the coefficients a_k and b_k are given by:

$$a_k = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{2\pi k x}{L}\right) \mathrm{d}x, \tag{2.21}$$

$$b_k = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{2\pi k x}{L}\right) \mathrm{d}x. \tag{2.22}$$

Furthermore, since cosine and sine functions are being employed, one can rewrite the previous equations in a more compact way using Euler's equation $e^{ikx} = \cos(kx) + i \sin(kx)$:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{2.23}$$

where the coefficient c_k represents:

$$c_k = \frac{1}{2L} \int_{-L}^{L} f(x) e^{-ik\pi x/L} \mathrm{d}x. \tag{2.24}$$

2.3.2.2 Fourier transform

As stated previously, the Fourier series can only be applied to periodic functions repeating itself out of the domain. In this case, the Fourier transform can be interpreted as a Fourier series with a domain length approaching infinity, which allows for the use of functions without repeating, i.e. with a domain corresponding to $(-\infty, \infty)$. Therefore, the Fourier transform is generalised to non-periodic functions. Using Equation 2.20 or 2.23, the Fourier transform can be derived regarding $f(x)$ as a collection of cosines and sines with a discrete set of frequencies. Defining $\omega_k = k\pi/L$ and $\Delta\omega_k = \pi/L$, if the domain length approaches to infinity $L \to \infty$, then $\Delta\omega \to 0$:

$$f(x) = \lim_{\Delta\omega \to 0} \sum_{k=-\infty}^{\infty} \frac{\Delta\omega}{2\pi} \int_{-\pi/\Delta\omega}^{\pi/\Delta\omega} f(\xi) e^{ik\Delta\omega\xi} \mathrm{d}\xi \; e^{ik\Delta\omega x}. \tag{2.25}$$

Once limits have been applied, the integral expression will become the Fourier transform of the function $f(x)$ and will be denoted henceforward as $\hat{f}(\omega) \equiv \mathcal{F}(f(x))$.

$$\hat{f}(\omega) = \mathcal{F}(f(x)) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} \mathrm{d}x \tag{2.26}$$

$$f(x) = \mathcal{F}^{-1}\left(\hat{f}(\omega)\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega x} \mathrm{d}\omega \tag{2.27}$$

Both integrals together are generally referred to as the *Fourier transform pair* and they are helpful owing to a variety of properties, such as linearity, the derivative behaviour in the Fourier transform domain, i.e. frequency, and convolution, which have been widely used, e.g. to solve PDEs in an efficient and accurate manner for computation and data analysis. Due to its importance, the derivative behaviour and the linearity will be briefly discussed here, please consult Ref. [40] for a more detailed explanation of the Fourier transform.

Derivatives of functions. The derivative of a function can be transformed into the Fourier domain using:

$$\mathcal{F}\left[\frac{d}{dx} f(x)\right] = \int_{-\infty}^{\infty} f'(x) e^{-i\omega x} \mathrm{d}x = i\omega \int_{-\infty}^{\infty} f(x) e^{-i\omega x} \mathrm{d}x = i\omega \mathcal{F}[f(x)] \tag{2.28}$$

which is an extremely important property within the field of differential equations which allows for easily transforming a PDE into an ODE.

Linearity. The Fourier transform is a merely linear operator which satisfies that:

$$\mathcal{F}[\alpha f(x) + \beta g(x)] = \alpha \mathcal{F}(f) + \beta \mathcal{F}(g), \tag{2.29}$$

$$\mathcal{F}^{-1}\left[\alpha \hat{f}(\omega) + \beta \hat{g}(\omega)\right] = \alpha \mathcal{F}^{-1}(\hat{f}) + \beta \mathcal{F}^{-1}(\hat{g}). \tag{2.30}$$

2.3.2.3 Discrete Fourier transform

Heretofore, both the Fourier series and the Fourier transform do only consider continuous functions. Nevertheless, in real-world applications data is not continuous; necessitating the development of an algorithm to perform the Fourier transform over a discrete vector of data: the discrete Fourier transform (DFT), which is simply a discretised version of the Fourier series for vector data $\mathbf{f} = [f_1, f_2, ..., f_n]^{\mathrm{T}}$ acquired by discretising a given function $f(x)$ at regular intervals Δt.

Based on the definition of the Fourier series and the aforementioned hypotheses, a simple formulation of the discrete Fourier transform can be determined by:

$$\hat{f}_k = \sum_{j=0}^{n-1} f_j e^{-i2\pi jk/n} \tag{2.31}$$

and the inverse discrete Fourier transform (iDFT) reduces to:

$$f_k = \frac{1}{n}\sum_{j=0}^{n-1} \hat{f}_j e^{i2\pi jk/n}. \tag{2.32}$$

Therefore, defining a fundamental frequency $\omega_n = e^{-2\pi i/n}$, the DFT algorithm can be expressed in matrix form for a given number of points n.

$$\begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \hat{f}_3 \\ \vdots \\ \hat{f}_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_n & \omega_n^2 & \dots & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^4 & \dots & \omega_n^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \dots & \omega_n^{(n-1)^2} \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_n \end{bmatrix} \tag{2.33}$$

Note that the DFT can be regarded as a matrix $\mathbf{F}$, corresponding to a Vandermonde matrix fulfilling $F_{ij} = x_i^{j-1}$, that maps the temporal data points in $\mathbf{f}$ to the frequency domain $\hat{\mathbf{f}}$. It is also worth noting that this transformation comprises the multiplication of a $n \times n$ matrix, which yields to $\mathcal{O}\left(n^2\right)$ operations. Therefore, when the number of samples is substantially large, this algorithm is computationally inefficient. Interestingly, Cooley and Tukey developed one of the most revolutionary algorithms up to date, known as fast Fourier transform (FFT), which, based on the DFT, reduces dramatically the number of operations involved in its computation. The mathematics behind the method will be described in the following section.

2.3.2.4 Fast Fourier transform

As discussed previously, the DFT rests on a linear operator whose multiplication results in $\mathcal{O}\left(n^2\right)$ operations. Conversely, the algorithm proposed by Cooley and Tukey, the fast Fourier transform (FFT), only requires from $\mathcal{O}\left(n \log\left(n\right)\right)$ operations, which represents a major development in a myriad of applications, such as audio and image compression or satellite communications.

For instance, to demonstrate the computational reduction, one can compare the number of operations involved in DFT versus FFT in a real-world scenario. Assuming a 10-second audio file, sampled at 50 kHz, the total number of samples will be $n = 5 \cdot 10^5$. Computing the DFT results in $2.5 \cdot 10^{11}$ operations to be performed, whereas the FFT requires roughly $6.6 \cdot 10^6$. This amounts to a speed-up factor of over 35 000, a marvellous advantage of FFT compared to DFT.

The underlying notion behind the FFT is that, if the amount of samples n is a power of 2, the DFT may be executed considerably more effectively. Considering $n = 1024 = 2^{10}$, the DFT algorithm can be expressed as a combination of lower order matrices for which the number of operations is lower.

$$\hat{\mathbf{f}} = \mathbf{F}_{1024}\mathbf{f} = \begin{bmatrix} \mathbf{I}_{512} & -\mathbf{D}_{512} \\ \mathbf{I}_{512} & -\mathbf{D}_{512} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{512} & 0 \\ 0 & \mathbf{F}_{512} \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\text{even}} \\ \mathbf{f}_{\text{odd}} \end{bmatrix} \tag{2.34}$$

where $\mathbf{f}_{\text{even}}$ correspond to the even index elements of the discrete sample data, $\mathbf{f}_{\text{odd}}$, to the odd ones, $\mathbf{I}_{512}$ represents a 512×512 identity matrix and $\mathbf{D}_{512}$, a diagonal matrix containing

the characteristic frequency hierarchically powered from 0 to 511.

Indeed, the FFT relies on this assumption, dividing the whole computation process into different operations of smaller 2×2 matrices computations, which is what dramatically reduces the number of operations. Please note that when dealing with sample points which are not a power of 2, the resulting vector can be padded with zeros until the condition is met.

2.3.3 Proper orthogonal decomposition

Based on the previous SVD matrix decomposition, a method of modal-decomposition built on the optimisation of the RMS error of the field variable under consideration is the proper orthogonal decomposition (POD). It is one of the most extensively used approaches for the analysis of fluid flows, with a significant number of variations. In POD applications to fluid flows, the analysis begins with a vector field $v(\mathbf{x}, t)$ with its temporal mean removed. The unsteady component can then be decomposed as follows:

$$v(\mathbf{x}, t) - \bar{v}(\mathbf{x}) = \sum_j c_j \phi_j(\mathbf{x}, t), \tag{2.35}$$

where c_j and $\phi_j(\mathbf{x}, t)$ are the expansion coefficients and modes, respectively. For a set of basis functions $\phi_j(\mathbf{x}, t)$, this equation describes the flow-field in terms of a Fourier series. Additionally, modern innovations in the modal-decomposition field have aimed to further separate space and time, requiring just spatial modes, i.e.

$$v(\mathbf{x}, t) - \bar{v}(\mathbf{x}) = \sum_j c_j(t) \phi_j(\mathbf{x}). \tag{2.36}$$

Note that the approach to be followed will strongly depend on the applicability of the problem, i.e. the properties of the flow and the information one wishes to extract [38].

Two types of methods, the standard POD approach, based on the covariance of a vector state varying in time, and the snapshot approach, applied in a similar way to the SVD method, can be used for identifying POD modes. Thereby, while POD is the technique that provides the modal expansion, SVD might be thought of as one of the POD modes algorithms. Both methods are based on the covariance matrix of the snapshot matrix $\mathbf{C} = \mathbf{X}\mathbf{X}^{\mathrm{T}}$, which is known to be symmetric and positive-definite by definition.

The traditional POD method seeks to solve the eigenvalue problem of the covariance matrix, i.e.

$$\mathbf{C}\phi_j = \lambda_j \phi_j \tag{2.37}$$

where the largest eigenvalues λ_j produce the optimum orthogonal modes ϕ_j, named as POD modes, that best represent a given dynamic field. However, when analysing three-dimensional turbulent flows, this technique is prohibitively computationally expensive since it aims at solving the eigenvalue problem of the covariance matrix with dimensions corresponding to the spatial degrees of freedom. Therefore, here we will focus on the SVD approach, which is also known to be more robust against round-off errors [38].

Having clarified this, recalling that the SVD can be applied to a rectangular matrix, a given data matrix, structured as indicated in Equation 2.13, can be directly decomposed with the SVD methodology as $\mathbf{X} = \mathbf{W}\mathbf{\Sigma}\mathbf{T}^{\mathrm{T}}$, where the matrices $\mathbf{W}$ and $\mathbf{T}$ contain the left (spatial) and right (temporal) singular vectors of $\mathbf{X}$ and $\mathbf{\Sigma}$, the singular values. Please keep in mind that the singular vectors in $\mathbf{W}$ and $\mathbf{T}$ are the same as the eigenvectors of AA^{T} and $A^{\mathrm{T}}A$, respectively, and the singular values are connected to the aforementioned eigenvalues through $\sigma_j^2 = \lambda_j$. Hence, both the conventional and the SVD approaches are related so that the POD might be seen as a consequence of the SVD of a snapshots matrix.

The relevance of POD relies on its optimality. POD modes are calculated in an optimal manner regarding the ℓ_2 norm error. For instance, in fluid dynamics, the velocity field is generally used to determine the POD modes. In that case, the modes capture the kinetic energy of the fluid flow in an efficient way. Besides, this technique is masterful in terms of minimising not only the mean-square error between the signal and its truncated reconstruction but also the number of modes required to properly describe the signal for a given threshold.

2.3.4 Higher-order dynamic mode decomposition

Aiming at identifying the spatio-temporal coherent patterns present in high-dimensional flow data, Schmid [42] developed a tool based on proper orthogonal decomposition (POD) which retrieved the spatially correlated structures with comparable behaviour in time. This methodology, known as dynamic mode decomposition (DMD), provides not only a reduction in dimension concerning a reduced set of modes which best reproduce the input flow-field energetically but also a model for the interaction of those modes in time.

The method decomposes the vector field data $v(\mathbf{x}, t)$ in Fourier-type modes, i.e.

$$v(\mathbf{x}, t) \simeq \sum_{m=1}^{M} a_m u_m(\mathbf{x})\, e^{(\delta_m + i\omega_m)t_k}, \tag{2.38}$$

for $k = 1, ..., K$, where u_m represents the DMD modes weighted by an amplitude a_m, ω_m, their associated frequencies and δ_m, their associated growth rates, which symbolise the temporal growth or decay of the u_m modes in time.

Regarding the input data, it should be noted that in order to avoid aliasing[2], the interval between snapshots, Δt, must be considerably smaller than the shorter period involved in Equation 2.38, which demands that the number of snapshots be considerably larger than the number of involved frequencies, i.e. $K \gg N$. Furthermore, in order to have sufficient information about the dynamics, the sampled data time span, T, must be slightly larger than the broader involved period [39]. This method also requires equidistant snapshots in time, i.e. Δt should be kept fixed. These assumptions are of utmost importance for accurately understanding the flow's underlying physics.

[2]Aliasing is a sampling phenomenon that turns separate signals indistinguishable. When the resolution is too poor, it also refers to the difference between a signal reconstructed from samples and the original continuous signal. In general, aliasing is determined by the signal's sampling rate and frequency content. Therefore, if aliasing is wanted to be avoided, according to the Nyquist sampling theorem, the sampling frequency should be at least twice the highest frequency involved in the original signal.

The standard DMD algorithm, introduced by Schmid [42], relies on the linear relationship of two consecutive snapshots matrices using the linear Koopman operator $\mathbf{R}$. To this end, a general snapshot matrix $\mathbf{V}_{k_1}^{k_2}$ can be defined for $k_1 < k_2$ so that its columns represent the snapshots varying equidistantly between k_1 and k_2, namely,

$$\mathbf{V}_{k_1}^{k_2} = [v_{k_1}, v_{k_1+1}, ..., v_{k_2}] \tag{2.39}$$

Therefore, using the previous nomenclature, the standard DMD can be defined based on the Koopman operator.

$$\mathbf{V}_2^K \simeq \mathbf{R}\mathbf{V}_1^{K-1} \tag{2.40}$$

where $\mathbf{V}_2^K$ and $\mathbf{V}_1^{K-1}$ represent here the second to last snapshots and the first to the penultimate snapshots of the data matrix, respectively. Recalling Equation 2.38, this equation might be seen as the simplest equation exhibiting such behaviour [39]. The Koopman matrix $\mathbf{R}$, which is independent of k, contains the dynamical information of the system, ω_m and δ_m represent the computed eigenvalues and $u_m(\mathbf{x})$ DMD modes can be obtained from the eigenvectors of the matrix $\mathbf{R}$. Therefore, combining both expressions, the general solution can be approximated to

$$v_k \simeq \sum_{m=1}^{M} a_m u_m \mu_m^{k-1}, \tag{2.41}$$

for $k = 1, ..., K$, where u_m represent the eigenvectors of the Koopman matrix and μ_m, the associated eigenvalues. This equation allows for a certain relationship of the growth rates and frequencies appearing in Equation 2.38 with the eigenvalues:

$$\delta_m + i\omega_m = \frac{1}{\Delta t} \log \mu_m. \tag{2.42}$$

However, this approach and various of its extended versions are restricted to those cases in which the spatial and spectral complexities coincide, $J = M$, which may not generally be the case. Recently, Le Clainche & Vega [39] extended the DMD method for the analysis of various types of flows, e.g. turbulent, multi-scale or transitional flows and noisy experimental data. Based on Takens' delayed embedded theorem [39], the higher-order dynamic mode decomposition (HODMD) relates d time-delayed snapshots using the higher-order Koopman assumption as:

$$\mathbf{V}_{d+1}^K \simeq \mathbf{R}_1 \mathbf{V}_1^{K-d} + \mathbf{R}_2 \mathbf{V}_2^{K-(d-1)} + ... + \mathbf{R}_d \mathbf{V}_d^{K-1}. \tag{2.43}$$

This more general condition allows the HODMD algorithm for computing the expansion in the following steps.

2.3.4.1 Step 1: dimension reduction

First of all, the singular value decomposition (SVD) technique is employed to reduce spatial redundancy and filter out noise caused by numerical or experimental errors (see § 2.3.1). The truncated SVD allows for the reduction of the original snapshot data into a series of linearly independent vectors of dimension N (where $N < J$ is the spatial complexity), based on a certain tolerance ε_{SVD}.

$$\mathbf{V}_1^K \simeq \mathbf{W}\mathbf{\Sigma}\mathbf{T}^{\mathrm{T}} \tag{2.44}$$

where $\mathbf{\Sigma}$ includes the singular values $\sigma_1, ..., \sigma_N$ and $\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{T}^{\mathrm{T}}\mathbf{T} = I$ are $N \times N$ unitary matrices. Note that the user must adjust the value of the parameter $\varepsilon_{\mathrm{SVD}}$ based on previous information of the simulation or experimental data, e.g. if the noise level of the snapshots is known in advance, then, $\varepsilon_{\mathrm{SVD}}$ may be set to be comparable to that level. Otherwise, the singular-value distribution (σ_n vs. n) can be used to guess a confident value. Above all, this parameter determines the number N of SVD retained modes, as described in Equation 2.19, which is recalled here for the sake of clarity.

$$\frac{\sigma_{N+1}}{\sigma_1} \leq \varepsilon_{\mathrm{SVD}} \tag{2.45}$$

Therefore, the snapshot matrix can be defined, using Equation 2.44, as

$$\mathbf{V}_1^K \simeq \mathbf{W}\mathbf{\Sigma}\mathbf{T}^{\mathrm{T}} \equiv \mathbf{W}\hat{\mathbf{T}}_1^K \tag{2.46}$$

where $\hat{\mathbf{T}}_1^K = \mathbf{\Sigma}\mathbf{T}^{\mathrm{T}}$ represents the dimension-reduced snapshot matrix with dimensions $N \times K$.

2.3.4.2 Step 2: the DMD-d algorithm

The higher-order Koopman assumption, defined in Equation 2.43, is now applied to the reduced snapshot matrix:

$$\hat{\mathbf{V}}_{d+1}^K \simeq \hat{\mathbf{R}}_1\hat{\mathbf{V}}_1^{K-d} + \hat{\mathbf{R}}_2\hat{\mathbf{V}}_2^{K-(d-1)} + ... + \hat{\mathbf{R}}_d\hat{\mathbf{V}}_d^{K-1} \tag{2.47}$$

where $\hat{\mathbf{R}}_k = \mathbf{W}^{\mathrm{T}}\mathbf{R}_k\mathbf{W}$ is used for $k = 1, ..., d$. The snapshot matrix, used in this equation, is divided into d blocks, each storing $K - d$ snapshots. The above equation may be stated in a more generic form by incorporating the modified snapshot matrix $\tilde{\mathbf{V}}_1^{k-d+1}$ and the modified Koopman matrix $\tilde{\mathbf{R}}$.

$$\tilde{\mathbf{V}}_2^{K-d+1} = \tilde{\mathbf{R}}\tilde{\mathbf{V}}_1^{K-d} \tag{2.48}$$

which can also be presented in matrix form for the sake of clarity:

$$\begin{bmatrix} \hat{\mathbf{V}}_2^{K-d+1} \\ \vdots \\ \hat{\mathbf{V}}_d^{K-1} \\ \hat{\mathbf{V}}_{d+1}^{K} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I & 0 \\ \hat{\mathbf{R}}_1 & \hat{\mathbf{R}}_2 & \hat{\mathbf{R}}_3 & \dots & \hat{\mathbf{R}}_{d-1} & \hat{\mathbf{R}}_d \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{V}}_1^{K-d} \\ \hat{\mathbf{V}}_2^{K-d+1} \\ \vdots \\ \hat{\mathbf{V}}_d^{K-1} \end{bmatrix} \tag{2.49}$$

This matrix is also known to present redundancies that must be removed by utilising the truncated SVD algorithm with the aforementioned $\varepsilon_{\mathrm{SVD}}$, varying the number of dimensions to N'.

$$\frac{\tilde{\sigma}_{N'+1}}{\tilde{\sigma}_1} < \varepsilon_{\mathrm{SVD}} \tag{2.50}$$

where $N' > N$ is the number of SVD modes that have been retained. The conclusion that $N' > N$ is based on the partitioning of the input data matrix into snapshot blocks (high-order Koopman assumption), which raises the spatial complexity of the problem from N to N'. As a result, the number of DMD modes is increased, filling in the gaps in information and ensuring that the DMD algorithm performs well in those circumstances when the spatial complexity J is substantially higher than the spatial complexity N (generally the case in the study of three-dimensional turbulent flows). Applying the truncated SVD to the modified snapshot matrix yields to

$$\tilde{\mathbf{V}}_1^{k-d+1} \simeq \tilde{\mathbf{W}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{T}}^{\mathrm{T}} \simeq \tilde{\mathbf{W}}\bar{\mathbf{T}}_1^{k-d+1} \tag{2.51}$$

where $\bar{\mathbf{T}}_1^{k-d+1} = \tilde{\mathbf{\Sigma}}\tilde{\mathbf{T}}^{\mathrm{T}}$, the matrix $\tilde{\mathbf{\Sigma}}$ includes the singular values $\tilde{\sigma}_1, ..., \tilde{\sigma}_{N'}$ and $\tilde{\mathbf{W}}^{\mathrm{T}}\tilde{\mathbf{W}} = \tilde{\mathbf{T}}^{\mathrm{T}}\tilde{\mathbf{T}} = I$ are $N' \times N'$ unitary matrices.

The following step is to introduce the modified snapshot matrix definition of Equation 2.51 in the general expression for the high-order Koopman assumption of Equation 2.48 and premultiply it by $\tilde{\mathbf{U}}^{\mathrm{T}}$. This combination yields to

$$\bar{\mathbf{T}}_2^{k-d+1} = \bar{\mathbf{R}}\bar{\mathbf{T}}_1^{k-d}. \tag{2.52}$$

Here, a new Koopman matrix with dimensions $N' \times N'$ has been introduced, which relates to the modified snapshot matrix by $\bar{\mathbf{R}} \simeq \tilde{\mathbf{W}}^{\mathrm{T}}\tilde{\mathbf{R}}\tilde{\mathbf{W}}$. In order to obtain an expression for this new Koopman matrix, the term $\bar{\mathbf{T}}_1^{k-d}$ in Equation 2.52 is factorised with the SVD methodology

$$\bar{\mathbf{T}}_1^{k-d} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}} \tag{2.53}$$

where $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{V}^{\mathrm{T}}\mathbf{V}$ are $N' \times N'$ unitary matrices and $\mathbf{\Lambda}$ is the diagonal matrix containing the N' singular values. Substituting now Equation 2.53 in 2.52, a general expression for the Koopman matrix is obtained.

$$\bar{\mathbf{R}} = \bar{\mathbf{T}}_2^{k-d+1}\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^{\mathrm{T}} \tag{2.54}$$

Consecutively, the eigenvalue problem can be solved for this matrix so that the eigenvalues μ_m and the eigenvectors $\bar{\mathbf{q}}_m$ can be obtained, thus allowing for computing the reduced DMD expansion:

$$\hat{v}_k = \sum_{m=1}^{M} \hat{a}_m \hat{u}_m e^{(\delta_m + i\omega_m)t_k} \tag{2.55}$$

for $k = 1, .., K$. The reduced DMD modes $\hat{u}_m$ are calculated using $\hat{q}_m = \tilde{\mathbf{U}}\bar{q}_m$ and the frequencies ω_m and the damping rates δ_m are calculated using Equation 2.42. Since the dimension of the reduced snapshots $\hat{v}_k$ is N ($= M$ in the present case) and the eigenvectors $\bar{\mathbf{q}}_m$ are linearly independent, each of these equations uniquely determines the value of the amplitude a_m of each DMD mode. The whole system of equations can be solved using the pseudoinverse, which minimises the least-squares-error of the approximation. Therefore, the Equation 2.55 can be rewritten in matrix form as

$$\mathbf{L}\mathbf{a} = \mathbf{b} \tag{2.56}$$

where the matrix $\mathbf{L}$, the unknown amplitudes vector $\mathbf{a}$ and the forcing term $\mathbf{b}$ can be defined as

$$\mathbf{L} = \begin{bmatrix} \mathbf{Q} \\ \mathbf{Q}\,\mathbf{M} \\ \vdots \\ \mathbf{Q}\mathbf{M}^{K-1} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \vdots \\ \hat{v}_K \end{bmatrix} \tag{2.57}$$

Note that the matrix $\mathbf{L}$ has been defined accordingly to the matrix $\mathbf{Q} = [\mathbf{q}_1, ..., \mathbf{q}_M]$ formed by the eigenvectors and the diagonal matrix $\mathbf{M}$ containing the associated eigenvalues μ_m. Applying the standard SVD without truncation to this matrix, the pseudoinverse can be computed

$$\mathbf{L} = \tilde{\mathbf{U}}_1 \tilde{\mathbf{\Sigma}} \tilde{\mathbf{U}}_2^{\mathrm{T}} \tag{2.58}$$

Combining Equations 2.56 and 2.58, a general expression for the amplitudes vector can be obtained. Note that using the pseudoinverse allows for the minimisation of the least-squares-error of the approximation, which gives a good picture of the effect of each DMD mode on the overall flow dynamics.

$$\mathbf{a} = \tilde{\mathbf{U}}_2 \tilde{\mathbf{\Sigma}}^{-1} \tilde{\mathbf{U}}_1^{\mathrm{T}} \mathbf{b} \tag{2.59}$$

Finally, the spectral complexity M, which denotes the number of retained modes in the Fourier-like expansion of Equation 2.38, can be controlled using a second tolerance $\varepsilon_{\mathrm{DMD}}$, also adjusted by the user.

$$\frac{a_{M+1}}{a_1} \leq \varepsilon_{\mathrm{DMD}} \tag{2.60}$$

Last but not least, similarly to the computed error in the POD case, the error made in the calculations is measured by the root mean square (RMS) error of the HODMD reconstruction (Equation 2.38).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{k=1}^{K} \left\| v_k - v_k^{\mathrm{DMD}} \right\|^2}{\sum_{k=1}^{K} \left\| v_k \right\|^2}} \tag{2.61}$$

2.3.4.3 General remarks on the HODMD algorithm

The HODMD algorithm presented here is based upon the high-order Koopman assumption, see Equation 2.43, which yields to the standard DMD [42] when $d = 1$. Therefore, the HODMD approach may be viewed as an extension to those cases in which the standard DMD fails, i.e. in noisy experimental data or in cases where the spectral complexity M is considerably larger than the spatial complexity N.

In both cases, even though the algorithm is built upon the Koopman linear equation, see Equation 2.52, where the Koopman matrix $\mathbf{R}$ does not vary over time, i.e. it is an invariant matrix in time containing the main information of the data dynamics, it can be applied over

non-linear systems, with the only constraint being to model that dynamical system using Fourier-like modes, as described in Equation 2.38.

The representation of the amplitude and frequency of such modes enables the identification of the prominent spatio-temporal patterns present in the analysed data. This plot will be referred to as DMD spectrum, and will be employed throughout the document to identify the leading modes, namely, those with the greatest amplitude. Furthermore, if the dynamics are real, then, they should also be invariant under $u_m \to \bar{u}_m$ and $\omega_m \to -\omega_m$. Therefore, the $a - \omega$ diagram must be symmetric around $\omega = 0$.

Please note that the primary goal of the present study is to identify the spatio-temporal patterns driving the underlying physics of the flow inside urban contexts. The state-of-the-techniques detailed here enable for gaining further understanding without the need for the governing equations. The HODMD approach, on the other hand, may be used to create some reduced-order models (ROMs) that are fully data-driven, which implies that prior knowledge of the governing equations driving the dynamics is not required. These HODMD-based ROMs are also marketed as an excellent tool for making temporal predictions based solely on available data.

Additionally, this method involves a large number of tunable parameters, which will be now summarised in Table 2.1. Please note that the selection of d is not critical. One can start using $d = 1$ (corresponding to the standard DMD) and then increase its value aiming at minimising the relative RMS errors for the reduced snapshots. It is, however, important to calibrate the parameters ε_{SVD}, ε_{DMD} and d before applying the method, in order to obtain as robust and accurate as possible results. This will be further discussed in the following sections.

Parameter	**Description**
T	Sampled interval of the data, also known as timespan. It should be selected so as to be slightly larger than the highest period involved in the dynamics.
Δt	Temporal distance between snapshots. It should be selected to be considerably smaller compared to the smallest period involved in the dynamics.
ε_{SVD}	Threshold used in dimension-reduction steps, see Equations 2.45 and 2.50. For noisy data, it should be selected comparable to that level and for turbulent and transitory flows, it should be selected so as to remove the small-amplitude modes (singular value distribution).
ε_{DMD}	Threshold used for the mode amplitude truncation, see Equation 2.60. It is normally selected to be slightly smaller than the relative RMS error in the reduced snapshot reconstruction.
d	Number of segments involved in the window shift process.

Table 2.1: Tunable parameters of the HODMD algorithm

2.3.5 HODMD for spatially multidimensional data

During § 2.3.4, the calculation of the DMD modes was performed considering a snapshot matrix with dimensions $J \times K$, where J, the spatial degrees of freedom of the given data, represented the number of mesh points along each direction ($N_x \times N_y \times N_z$) and K, the number of snapshots. However, when dealing with more than one space variable, this snapshot matrix

can be substituted by a snapshot tensor, which may involve more than three indexes. Those cases arise when the available data involves, for instance, more than one velocity component, i.e. the velocity vector of the studied fluid flow is decomposed into its streamwise, normal and spanwise components and discretised over a structured mesh labelled with (j_1, j_2, j_3) for the case of three-dimensional data.

The theory developed for the snapshot matrix applies to the preceding examples as well. However, as noted by Le Clainche & Vega [39], keeping the dependency between snapshots in more than one index may result in superior outcomes. Assuming that all of the velocity components are introduced in the analysis, the correlation between the different snapshots strengthens, because the real flow is treated as a combination of all of these components.

Therefore, one can rewrite Equation 2.38 in such a way that all the variables involved are indexed:

$$v_{jk} \simeq \sum_{m=1}^{M} a_m u_{jm} e^{(\delta_m + i\omega_m)(k-1)\Delta t} \tag{2.62}$$

for $k = 1, ..., K$. This equation allows for an expansion in which j can be separated into the different spatial dimensions of the data and l might represent the number of velocity components.

$$v^l_{j_1 j_2 j_3 k} \simeq \sum_{m=1}^{M} a_m u^l_{j_1 j_2 j_3 k} e^{(\delta_m + i\omega_m)(k-1)\Delta t} \tag{2.63}$$

for $k = 1, ..., K$. Note that the approach of Equation 2.63 has a direct application to the analysis of three-dimensional flows involving three velocity components yielding to $N_j = N_l = 3$. Conversely, for the case of two-dimensional data, N_j reduces to 2.

$$v^l_{j_1 j_2 k} \simeq \sum_{m=1}^{M} a_m u^l_{j_1 j_2 k} e^{(\delta_m + i\omega_m)(k-1)\Delta t} \tag{2.64}$$

Once the shape of the input data has been defined, the procedure to apply the DMD algorithm is analogous but incorporating the necessary tools to work with the snapshot tensor. In that sense, the standard truncated SVD cannot be applied to reduce the dimension of the data, see § 2.3.1, and it should be replaced by a higher-order SVD algorithm, i.e. truncated HOSVD. Hereinafter, Equation 2.64 will be used to illustrate the dimension reduction procedure.

$$v^l_{j_1 j_2 k} \equiv T_{j_1 j_2 l k} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \sum_{p_3=1}^{P_3} \sum_{n=1}^{N} S_{p_1 p_2 p_3 n} \mathbf{W}^{x_1}_{j_1 p_1} \mathbf{W}^{x_2}_{j_2 p_2} \mathbf{C}_{l p_3} \mathbf{T}_{n k} \tag{2.65}$$

where $S_{p_1 p_2 p_3 n}$ represents a fourth-order tensor, the columns of the matrices $\mathbf{W}^{x_1}$ and $\mathbf{W}^{x_2}$, the spatial modes, which in the present case are two, and $\mathbf{T}$, the temporal mode. The reduction of § 2.3.4.1 is then applied to each of these modes and, finally, the DMD-d algorithm of § 2.3.4.2 is applied to the temporal modes tensor $\mathbf{T}$.

2.3.6 Iterative HODMD

It should be noted that the preceding sections' algorithms can be used iteratively across the reconstructed snapshots. This iterative method allows for the removal of noisy artefacts from the data and the generation of highly accurate solutions. In other words, after calculating the DMD modes, frequencies, growth rates and amplitudes, the snapshot matrix or tensor is reconstructed using Equation 2.38. Thereupon, the method is applied to this reconstruction, which comprises the main dynamics of the flow data with noise partially filtered. This procedure is repeated until the number of retained modes coincides in two consecutive iterations.

2.3.7 Summary on modal-decomposition analyses

To conclude with the theoretical analysis of the modal-decomposition methods, some general concepts will be now recovered and compared between methods, aiming at providing a general overview of the strengths and weaknesses of each modal-decomposition method that will be employed in further chapters to analyse the fluid simulation data.

Orthogonality First of all, the POD method is known to return an orthogonal set of basis vectors with the smallest possible dimension [38]. This is especially useful for the creation of a reduced-order model (ROM) of the flowfield data. On the other hand, DMD modes are not orthogonal, which means that, if those modes are utilised for constructing a ROM, it will contain extra terms owing to the non-zero spatial inner product terms between distinct modes. Therefore, while constructing a ROM of the flow data, POD can be applied more efficiently.

Hierarchy of modes In terms of mode hierarchy, whereas POD arranges them in the order of energy contents, it is difficult to arrange DMD modes in the order of physical importance since there is no straightforward method to rank eigenvalues. DMD modes are classified based on their dynamical importance, i.e. their frequency and growth rate parameters. Because of this dynamical behaviour, snapshots for POD analysis are not required to be equidistant in time, as they are for DMD analysis. However, the high-order Koopman condition used in the HODMD algorithm aims to address this problem of the standard DMD method.

Noisy data As far as noisy data is concerned, POD can be used to eliminate incoherent noise from the data if its level is lower than the signal level. Hence, noise emerges in the expansion as high-order modes that can be easily removed. On the other hand, because it changes the dynamics of the system, the DMD method might be sensitive to noisy data [38]. However, due to the use of truncated SVD/HOSVD techniques, noise-resistant algorithms are known to exist.

Linearity and governing equations Both approaches are purely data-driven and provide a linear approximation of a non-linear dynamical system. In other words, prior knowledge of the underlying dynamics of the system, i.e. the governing equations, is not required. Hence, the nonlinear differential equations that must be solved in the domain of fluid flows (see § 2.1) are not required for the analysis. Because of the inherent complexity of the aforementioned equations, this is posed as the major benefit of the method.

3

Problem implementation

The purpose of this chapter will be to conduct a general examination of the numerical implementation of the mathematical tools discussed in the previous chapter. Indeed, it may be viewed as an extension of the theoretical derivation, but in this case, the emphasis will be on providing a comprehensive guide to the practical implementation of the post-processing tools used for the analysis of fluid numerical simulations or experimental data as well as the characterisation of the underlying spatio-temporal features.

The framework that follows will be based on the theory of modal decompositions that was introduced in § 2.3 and it is assumed that the reader is already familiar with the principles introduced there. The discussion will concentrate on the MATLAB routines used to implement those mathematical tools, which are mostly based on the routines provided in the book *Higher-order Dynamic Mode Decomposition* [39]. Please refer to the cited book for a more in-depth explanation of the implementation described here.

3.1 Pre-processing tools

In § 2.2 the significance of each workflow stage within numerical simulations has been established. Here, recalling the purpose of the current analysis, i.e. using post-processing tools to gain a further physical understanding of the flow within urban environments, the pre-processing tools, e.g. the numerical domain or the interpolation meshes, which have been utilised to perform the numerical simulations, will be now briefly defined for the sake of clarity.

Note that this section will only cover those topics that are essential for clearly understanding the following analysis. Please see Ref. [14] for a much more extensive explanation and implementation of the numerical principles discussed here.

3.1.1 Numerical domain

Defining a numerical domain through which the flow may be simulated and analysed is one of the first things to consider during the pre-processing step. As described in § 2.2, this choice will have a substantial influence on subsequent phases, such as the meshing procedure or simulation results. As a consequence, the computational domain must be specified with special care in order to achieve the most precise and efficient solution feasible.

Figure 3.1 depicts a geometrical representation of the xy and xz planes of the numerical domain used for the simulations that will be examined throughout this study. This domain illustrates an in-line arrangement of two wall-mounted obstacles of height h, streamwise length w_b and spanwise length b, around which the flow is intended to replicate that within urban environments. The scheme described here is general, which means that it may be used to analyse the many flow regimes that emerge from varying some of the given parameters. Recalling Oke's classification of flow regimes [12], the most important criteria in determining that regime is the streamwise separation of obstacles ℓ, see Figure 1.3. Therefore, the computational domain for the various case studies will be built depending on this parameter.

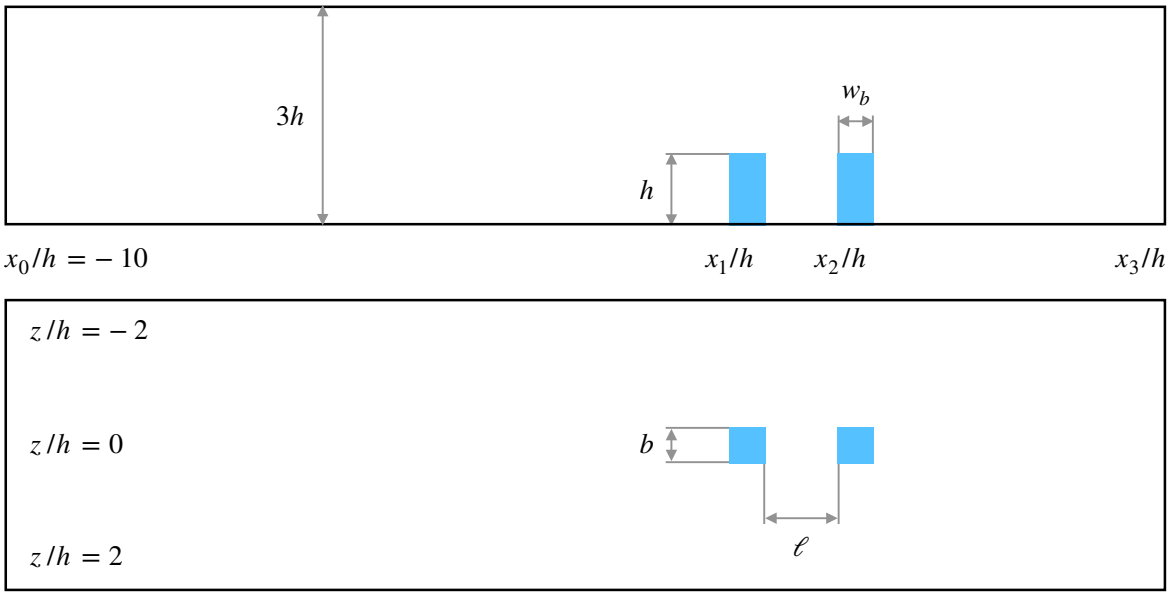


Figure 3.1: Generic geometrical scheme of the numerical domain

The domain origin coincides with the starting point of the first obstacle, i.e. $x_1/h = 0$, allowing for the development of the simplified urban environment inside the domain's positive side. Furthermore, it is worth noting that all of the parameters given in Figure 3.1 are sized using the height of the building-like obstacle, allowing for a rapid re-scaling of the computational domain for different cases by modifying a single parameter [14].

Another critical parameter in the domain specification is the separation between the two obstacles, which has already been proven to be the most relevant in determining the ensuing flow regime. In this case, while the first obstacle's position will be set, the second will be moved when various separation ratios are desired. In other words, the first obstacle will always be at the reference frame's origin, and modifying the streamwise position of the second obstacle will result in varied values of x_2/h. As a result, when this parameter is modified, the domain length in the streamwise direction changes, with the aim of accurately defining the flow behaviour inside the wake for all possible variations. This leads to distinct values for x_3/h, which stands for the streamwise location of the domain's final part in Figure 3.1. Please keep in mind that the only direction changed is the streamwise one; therefore, the remainder of the domain's parameters will remain constant between case studies.

The distinctive parameters for the various flow regimes are depicted in Table 3.1. The only variation between cases is the separation between obstacles ℓ and the domain length L_x, both of which are in the streamwise direction. The remainder of the characteristics remain

unchanged, as previously mentioned.

Parameter	Description	SF	WI	IR
h	Obstacle height	1	1	1
w_b	Obstacle streamwise length	0.5	0.5	0.5
b	Obstacle spanwise length	0.5	0.5	0.5
ℓ	Streamwise separation between obstacles	1	2	4
L_x	Domain length in the streamwise direction	16	17	21
L_y	Domain length in the heightwise direction	3	3	3
L_x	Domain length in the spanwise direction	4	4	4

Table 3.1: Geometrical parameters defining the numerical domain of the three flow regimes identified by Oke [12]: skimming flow (**SF**), wake interference (**WI**) and isolated roughness (**IR**).

The numerical domain specified here is the one through which the simulation was executed. However, for the spatio-temporal analysis, this domain comprises sections that are not very significant and can be omitted when computational efficiency is a priority, e.g. the initial part of the domain, which is defined by the development of the boundary layer. This will be covered in further detail in the next sections.

3.1.2 Mesh design

The mesh issue is actually a separate topic with its own set of tools and investigations. The objective of the work discussed here, however, is not to examine meshes, but rather to examine flows in urban contexts. Some brief comments will now be given in order to provide the reader with sufficient comprehension of the implementation employed. Therefore, if more information on the meshing procedure is required, please consider Refs. [14, 43]. According to Torres [14], this pre-processing stage is both crucial and time-consuming since the mesh resolution must be high enough to adequately represent the desired quantities, while keeping computational cost under control. It should be noted that the simulation was a large-eddy simulation (LES), which requires a high enough resolution to correctly solve the large turbulent scales from the Kolgomorov scale. To that purpose, the mesh is created in partnership with Parallel Works Inc. [44] and implemented in the high-order spectral code Nek5000 [13].

A spectral-element mesh is used in the simulation to discretise the computational domain into different elements (subdomains). Besides, those elements are further divided into smaller grid points using a Gauss-Lobatto-Legendre (GLL) quadrature of $N = 8$, guaranteeing that the resolution is adequate. This is a key characteristic of the spectral-element method; the computational domain is split into a set of elements that include a distribution of points therein, which, in this case, corresponds to an eight-point GLL distribution, allowing for a considerably finer mesh. On top of that, the mesh in the near-obstacle zone is refined in order to precisely solve the flow behaviour within the boundary layer region. This is especially crucial in the domain's initial region, where the boundary layer develops and the transition to turbulent flow is secured. Taking everything into consideration, the total number of mesh elements was around 205,605, which amounts to a total of 105 million grid points [14].

3.1.3 Boundary conditions

The final step in the pre-processing stage is to specify the boundary conditions (BC) that will be used to solve mathematically the partial differential equations (PDEs) that govern fluid flows. The implementation of the boundary conditions presented here is based on Torres' work [14]. In this case, the boundary conditions that must be met will be defined according to the domain's various faces and they are summarised in Table 3.2.

BC	Description
Inflow	Dirichlet condition is applied to simulate the inflow conditions, $\mathbf{u} = f(\mathbf{x})$. Specifically, a Blasius profile is imposed.
Outflow	Extended stabilised version of the standard outflow condition, provided by Dong et al. [45].
Open boundary	Combination of an outflow and a Dirichlet condition: a zero-stress condition is applied in the normal direction, i.e. a standard outflow condition, and a Dirichlet condition in the other two directions.
Wall	The lower part of the domain and the obstacles are modelled with a boundary condition of wall-type, i.e. $\mathbf{u} = 0$.
Periodicity	Periodic conditions are imposed in the spanwise boundaries in order to minimise the influence on the studied region.

Table 3.2: Boundary conditions representing the physical behaviour of the flow simulation. Extracted from Torres [14].

Beginning with the inflow conditions, a Dirichlet condition is applied at $x/h = -10$ so that the incoming flow might well be characterised with a Blasius profile[1]. On the other hand, Torres [14] used an enhanced version of the conventional outflow boundary condition, provided by Dong et al. [45]. This approach enables the truncation length of the outflow boundary to be maximised while causing no detrimental effects in the flow, resulting in a large reduction in computational cost [14]. Closely related to the previous condition, at the upper part of the domain, i.e. at $y/h = 3$, a combination of outflow and Dirichlet conditions is used to simulate an open-air urban environment: a zero-stress condition is applied in the wall-normal direction and a Dirichlet condition in the other two directions. This enables modelling of the domain face as an open boundary. Finally, periodicity is applied in the spanwise direction and the wall condition is applied to the domain's bottom plane and the building-like obstacles.

3.2 Solution process

The following step is to solve the governing equations, using the prior setup configuration, in order to generate the required flow fields over which the post-processing techniques will be applied. Recalling the aforementioned mesh parameters, the governing equations must be solved over a total of 205,605 elements with an eight-GLL point distribution each, yielding about 105 million grid points. The laminar Blasius solution gives an inflow Reynolds number, which is computed as a function of a unitary characteristic length, equivalent to $Re_{\delta^*} = 450$ [14]. Despite the fact that the inflow Reynolds number is low for a turbulent flow simulation, the

[1] A Blasius boundary layer is a well-known concept in the science of laminar flows that describes a stable two-dimensional laminar boundary layer created on a semi-infinite plate parallel to a continuous unidirectional flow.

tripping forces of the initial region of the domain enable turbulent conditions to be reached around the zone of interest. As a result, the inflow Reynolds number can be significantly reduced.

All simulations were run on the Cray XC40 machine "Beskow" at KTH Royal Institute of Technology's PDC Centre for High Performance Computing. This system contains a total of 67,456 cores and 156.4 TB of primary memory, from which up to 512 cores were employed for the present project [14]. For a more detailed discussion of the numerical implementation of the solution process in Nek5000, please refer to previous studies where the LES methodology is used across complicated geometries, namely Refs. [46, 47, 48]. Also, for the calculation of turbulence statistics and validation of the simulation data, consult Ref. [49].

3.3 Post-processing tools

Once the simulation has been carried out, the output data needs to be processed, which covers from validating the obtained results to perform further analyses. Indeed, once the setup and the solution processes have been finished, the post-processing stage assumes great importance and it can take from simple and rapid steps to thorough analyses aiming at gaining a deep understanding of the simulation results. The purpose of this section is to provide the reader with a sufficient understanding of the mathematical tools employed at this stage.

3.3.1 Interpolation mesh

In § 3.1.2, the mesh design used for the subsequent calculations was briefly introduced. Here, the design of the interpolation meshes, also included within the scope of Ref. [14], will be defined for all the different simulation cases. First of all, it is worth noting the difference between the standard and the interpolation meshes. While the first is the mesh employed in the solution process, which implies that the particular governing equations are solved for each grid point, the former is the one used to output and visualise the simulations results. Therefore, although the resolution of the latter will not be as high as that of the calculation mesh, it will be sufficient for the application of the different post-processing tools.

Despite their apparent differences, both types of meshes share some particular features. The interpolation mesh, similarly to the initial mesh, has a progressive distribution in the wall-normal direction, which means that, while the mesh elements are equally spaced in the streamwise and spanwise directions, the elements in the wall-normal direction do not have a linear distribution: they are more concentrated in the lower part of the domain, where the majority of flow phenomena occur. All the same, it is also important to have this element density gradient in the $y-$direction in order to be able to adequately model the behaviour of the boundary layer close to the ground, which is especially relevant at the threshold $-10 \leq x/h \leq -1$. A representation of the interpolation mesh at the plane $z/h = 0$ is shown in Figure 3.2. This illustration shows how the number of points increases as one approaches the bottom horizontal plane and how the resolution drops as one gets away from the ground.

Furthermore, because the spatio-temporal patterns relevant for the current analysis occur within the region of the buildings, the region of interest can be significantly reduced, lowering the computational cost associated with data processing, which, as will be seen later, is criti-

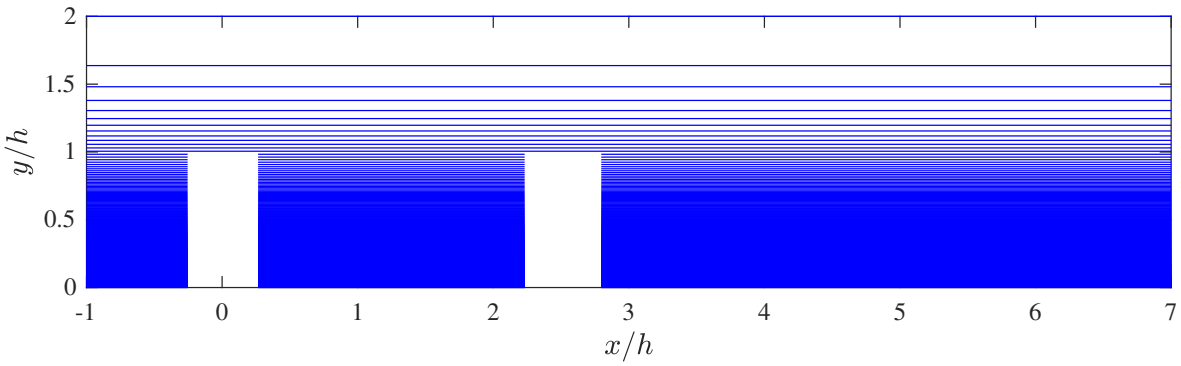


Figure 3.2: Two-dimensional representation of the interpolation mesh at $z/h = 0$, proving the resolution in $y-$direction. The flow regime presented here corresponds to the wake interference (**WI**) case, where $\ell/h = 2$. The interpolation mesh for the rest of the cases is analogous.

cal during the analysis. For that reason, recalling the hole numerical domain represented in Figure 3.1, the interpolation meshes begin at $x/h = -1$ and conclude at the domain's final section, which may vary from case to case, as shown in Table 3.1. As a result of this length difference in the streamwise direction, each case has a different number of mesh cells. The characteristics of each one are summarised in Table 3.3.

Parameter	**Description**	**SF**	**WI**	**IR**
$\Delta x/h$	Streamwise length of the reduced domain	6	8	12
$\Delta y/h$	Heightwise length of the reduced domain	2	2	2
$\Delta z/h$	Spanwise length of the reduced domain	3	3	3
N_x	Number of points in the streamwise direction	100	355	515
N_y	Number of points in the heightwise direction	400	400	400
N_z	Number of points in the spanwise direction	200	200	200
N_T	Total number of points (in millions)	8	28.4	41.2

Table 3.3: Interpolation mesh parameters for the different flow regimes

To finish off, it is worth recalling that the simulation procedure of the described cases is not included within the scope of the present project. Indeed, the results of those simulations will be the ones to be analysed using the post-processing tools defined in § 2.3. The previous sections aimed at providing the reader with a general overview of the simulations to be analysed; for more information on the simulations, refer to Torres et al. [11] and Torres' thesis [14].

3.3.2 Modal decompositions

After having briefly examined all simulation procedures preceding the generation of flow field results, this section aims to provide a practical description of the implementation of the various modal decompositions, which were previously exposed in a theoretical manner in § 2.3. From the singular value decomposition (SVD) to the higher-order dynamic mode decomposition (HODMD), the various algorithms will be thoroughly outlined down to the implementation level.

3.3.2.1 Snapshot matrix

As stated in § 2.3.1, the very first step during a modal-decomposition analysis is to build a considerably large dataset containing the evolution of a given flow field variable in time. Using the nomenclature of Equation 2.13, this dataset is commonly known as snapshot matrix and contains the temporal information, i.e. snapshots, stacked in columns, each of them representing the value of the given variable at different space locations, given by the interpolation mesh. Recalling the total number of mesh elements for the different flow regimes of Table 3.3, it is trite to think that the size of the associated files will be extremely high. Therefore, it will be very important to keep track of the machine limitations and the computational cost in order to build the most appropriate dataset.

From this point on, the concept of dataset will be referred to as the output results of the given flow variables. This data has been obtained and provided by the PDC Centre for High Performance Computing at KTH Royal Institute of Technology. The output files are formatted so that each variable $\mathbf{u}$ is saved in a struct-type file, where $\mathbf{u}\{:\}$ represents the different snapshots with an inner shape equivalent to a vector shaped as $(N_x \times N_y \times N_z)$. Therefore, these datasets can be reshaped so that the dependence on each spatial and temporal dimension is maintained, i.e. $\mathbf{u}\,(x, y, z, t)$. In such a way, considering a given matrix $A(x, y, z, t)$ containing the spatial and temporal information of a given flow field variable, only a couple of MATLAB code lines are needed to transform it into the snapshot-matrix shape.

The temporal parameters of the datasets used for all the flow regimes are depicted in Table 3.4. Here, it should be introduced the concept of convective time units. This a common principle within the analysis of DNS simulations and it can be defined as the ratio of some characteristic length and a velocity. In the present case, time is non-dimensionalised using the free-stream velocity U_∞ and the height of the obstacles, h. All the introduced parameters are expressed in terms of convective time units.

Parameter	Description	SF	WI	IR
T_i	Initial simulation time	38.65	80.40	83.30
T_f	Final simulation time	117.38	144.60	145.60
ΔT	Total simulation timespan	78.73	64.20	62.30
Δt	Time interval between snapshots	0.35	0.3	0.7
N_f	Number of computed fields	225	215	90

Table 3.4: Temporal parameters of the datasets used for the modal-decomposition analysis.

Note that in § 2.3.5 the concept of snapshot tensor was introduced. With this approach, the dependence of the field variable on the different spatial and temporal dimensions is maintained, i.e. Tensor(v, x, y, z, t), where v states for the different velocity components, x, y and z, for the spatial dimensions and t, for the temporal dimension. Therefore, for all those multi-dimensional algorithms, the data should be arranged using the latter approach.

3.3.2.2 Singular value decomposition

The essential theoretical principles and mathematical derivation of the singular value decomposition were introduced in § 2.3.1. This matrix factorisation method's ubiquity stems from

its ease of use and numerous applications in data compression. The goal of this section is to show how simple it is to implement this method, which serves as the foundation for subsequent analysis.

There are several variants of the basic SVD factorisation that rely on the truncation of the singular value matrix based on various criteria. For instance, one may decide to retain only the first K singular values, such that K may be thought of as the approximate rank of a given snapshot matrix A. In this scenario, this version satisfies the criteria $\sigma_n/\sigma_1 \geq 10^{-15}$ and is given the name **compact SVD**. On the other hand, truncation may be used to choose just those modes whose reconstructed matrix has a relative RMS less than a specified threshold ε, see Equation 2.18, i.e.

$$\sqrt{\frac{\sigma_{N+1}^2 + ... + \sigma_K^2}{\sigma_1^2 + ... + \sigma_K^2}} \leq \varepsilon_{\text{SVD}}. \tag{3.1}$$

This is commonly known as **truncated SVD**. Since this algorithm aims to establish an RMS threshold by means of the parameter ε, it will be the one to be employed in further methods, also see § 2.3.4. For that reason, the computational implementation of the algorithm is now presented.

```
function [] = TruncSVD(A,epsilon)

    % Standard SVD calculation of the matrix A
    [U1,S1,V1] = svd(A,'econ');
    sv = diag(S1);

    % Imposed restriction of retained singular values
    ret = 0;
    MM = length(sv);
    av = norm(sv,2);
    for mm = 1:MM-1;
        if norm(sv(mm+1:MM),2)/av > epsilon
            ret = ret+1;
        end
    end
    ret = min(ret,MM);
    sv = sv(1:ret);

    % Only keep the required modes
    S = diag(sv);
    U = U1(:,1:ret);
    V = V1(:,1:ret);

    % Calculation of the truncated version of matrix A
    Aapprox = U*S*V';

end
```

Note that this function computes the left and right modes, represented by U and V, respectively, as well as the singular values S and the reconstructed matrix from a given snapshot matrix A within a set tolerance ε. From the computed standard SVD of matrix A, Equation 3.1 is applied to extract the required singular values. Thereupon, once the appropriate singular values are identified, those modifications should be transferred to the left and right

modes, allowing for the computation of the truncated approximation of the original matrix. In contrast to the preceding methods, this one has the advantage of ensuring that the preserved modes yield a reconstruction matrix with a relative RMS lower than a certain threshold, indicated by ε.

Toy model for the SVD analysis To demonstrate the various versions of the singular value decomposition, let us consider the following toy model described by the function:

$$f(x,y) = 10x\sin(10x+20y)\log(5+y^2+3x^2) - 50(x-1)^2y^{3/2} + e^{xy} \tag{3.2}$$

for $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Note that the complexity of this function prevents an analytic solution in the region of interest from being obtained. As a result, using the many variations of the SVD method, one may approximate this function as a collection of modes arranged hierarchically by their contribution to the system. Given that property of the SVD, representing the distribution of the singular values allows for the identification of those modes that are energetically more important and the establishment of certain thresholds to represent the function in a much more efficient manner, i.e. with a certain relative RMS error.

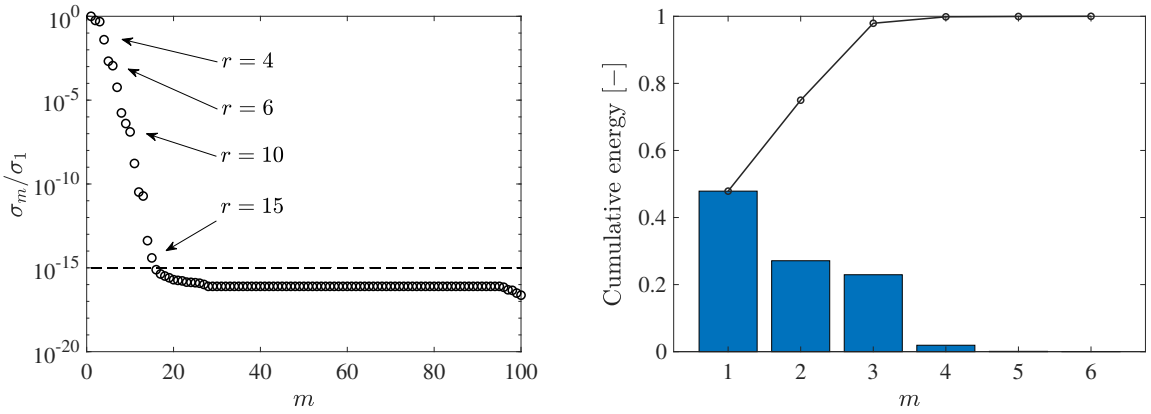


Figure 3.3: Results of SVD applied to a 100×100 matrix for the toy model function in (3.2). (a) Singular value distribution using the economy SVD. The dashed line represents the threshold value of the compact SVD. (b) Energy contribution (bar) and cumulative energy (line) for the latent space $r = 6$.

Figure 3.3 is introduced to demonstrate the minor variations between the SVD versions. When the economic SVD is applied to a given matrix $A(x_i, y_i)$ that contains the toy model data, a total of $\min(N_x, N_y)$ modes are retained. In this particular case, $N_x = N_y = 100$, indicating that the total number of modes is 100. This distribution clearly shows how additional modes ($r > 15$) have a virtually zero-machine contribution to the system ($\sigma_m/\sigma_1 \approx 10^{-15}$). At this stage, the compact SVD outperforms because it only keeps modes that fulfil the criterion $\sigma_m/\sigma_1 > 10^{-15}$, see § 3.3.2.2, with a comparable RMS error. These findings are also shown in Table 3.5, where the relative error for the economy and compact methods is almost similar. Figure 3.3 also depicts a black dashed line indicating the compact threshold, implying that by simply using $r = 15 < \min(N_x, N_y)$ modes, one may achieve results that are reasonably comparable to those produced by maintaining the whole set of modes. Note that it is possible for all modes to make a substantial contribution to the system and, as a consequence, the outcomes of the aforementioned procedures to match. This is usually the case in turbulent

	Tolerance	Modes	Relative RMS
Economy SVD	–	100	9.4×10^{-16}
Compact SVD	10^{-15}	15	9.4×10^{-16}
Truncated SVD	10^{-10}	10	1.7×10^{-9}
	10^{-5}	6	5.8×10^{-6}
	10^{-3}	4	2.1×10^{-4}

Table 3.5: Results of SVD applied to the toy model function in (3.2). The Table shows the number of modes retained with each of the SVD versions and the relative root mean square error of the reconstructed signal.

flows, where smaller scales have a stronger relationship to larger scales than in the current toy model example.

Furthermore, when employing the truncated SVD, a specific parameter, ε_{SVD}, may be used to specify which modes are retained. In fact, with $\varepsilon_{\text{SVD}} = 10^{-3}$ and 10^{-5}, the technique preserves 4 and 6 modes with RMS errors of 2.1×10^{-4} and 5.8×10^{-6}, respectively. In both situations, the RMS error is less than the ε_{SVD} value chosen. Besides, the gaps in the distribution of the modes are worth noticing since they describe modes with substantial fluctuations in energy. In such a way, Figure 3.3 also depicts the energy decay for the first six modes. These two graphs serve as the foundation for efficient mode truncation because, by retaining only the first three modes (notice the gap between the third and fourth modes), the reconstructed field variable may reflect nearly the whole energy of the system.

3.3.2.3 Dynamic mode decomposition

The purpose of this section is to provide sufficient information about the practicalities of the main algorithm to use throughout subsequent analysis phases: the dynamic mode decomposition, namely, the DMD-d algorithm. To that aim, in § 2.3.4, it has been introduced the theoretical concepts and derivation associated with the method. Note that the present section will not be focused on the theoretical perspective of the algorithm, but rather on its implementation and practical applications. It is worth recalling that Le Clainche and Vega [39] developed this higher-order method and presented it in their book Higher-order Dynamic Mode Decomposition. For further details on the method, please refer to the cited book.

Table 3.6 summarises the workflow phases for the DMD-d method, based on the theoretical derivation. First and foremost, the necessary data must be built or reshaped into a snapshot matrix (or tensor depending on the input format of the algorithm to work with). § 3.3.2.1 contains information on how it was implemented in this particular case. Thereupon, this matrix is dimension-reduced using the truncated SVD, with the goal of removing redundancy or noise from the data using a specific threshold, as described in § 3.3.2.2, resulting in a **reduced snapshot matrix**. At this stage, the high-order Koopman condition for the analysis of d-time delayed snapshots is also applied, resulting in an **enlarged-reduced snapshot matrix**, which serves as the foundation for the study of data where the spectral complexity M is significantly greater than the spatial complexity N[2].

[2]Recall that the spectral complexity M denotes the number of terms in Equation 2.38 and the spatial complexity denotes the rank of the system of modes in the aforementioned expansion.

Workflow	Description
Step 0	Build snapshot matrix or tensor.
Step 1	**First dimension reduction** using the truncated SVD algorithm, see § 3.3.2.2, resulting in a reduced snapshot matrix. Over this matrix, the high-order Koopman condition is applied to produce the enlarged-reduced Koopman matrix.
Step 2	**Second dimension reduction**. Truncated SVD is applied over the enlarge-reduced Koopman matrix, yielding to the reduced-enlarged-reduced Koopman matrix.
Step 3	Standard SVD is applied to perform the pseudoinverse over the Koopman condition and obtain a general expression for the Koopman matrix. Computation of DMD modes, growth rates and frequencies solving the **eigenvalue problem** of the Koopman matrix.
Step 4	Computation of the amplitude of the modes, truncating them according to the parameter ε_{DMD}.

Table 3.6: Summary of the workflow stages for the DMD-d algorithm.

Continuing along this line of thought, a second dimension reduction is performed to this enlarged-reduced snapshot matrix, yielding a **reduced-enlarged-reduced snapshot matrix**. Finally, the eigenvalue problem over the Koopman matrix is solved to calculate the DMD modes and their associated amplitudes, which may be modified by the parameter ε_{DMD}. This procedure is addressed in workflow steps 3 and 4. It is important to note that the user must select the tolerances ε_{SVD} and ε_{DMD}, based on the problem objective, and the value of d, which has a major influence on the consequent relative error of the reconstructed field variable.

Toy model for the modified Koopman analysis The modified Koopman analysis provides an extensive range of applications compared to the standard procedure. To illustrate the reader with its various properties, a given toy model function will be used, i.e.

$$f(t) = \frac{2}{2 - \cos(\omega_1 t)\cos(\omega_2 t)}. \tag{3.3}$$

This function can also be rewritten using the Taylor expansion as

$$f(t) = \frac{2}{2 - w(t)} = 1 + \frac{w(t)}{2} + \left[\frac{w(t)}{2}\right]^2 + \left[\frac{w(t)}{2}\right]^3 + \dots \tag{3.4}$$

where

$$w(t) = \cos(\omega_1 t)\cos(\omega_2 t) = \frac{\left(e^{i\omega_1 t} + e^{-i\omega_1 t}\right)\left(e^{i\omega_2 t} + e^{-i\omega_2 t}\right)}{4}. \tag{3.5}$$

Therefore, in a very simple way, the toy model function can be represented in terms of Equation 2.38. This means that the fundamental frequencies of the toy model are commensurable and the combination of the different values that $\omega_2 \pm \omega_1$ may acquire yield to a quasi-periodic behaviour, where $\omega_2 \pm \omega_1$ are the fundamental frequencies with their corresponding harmonics. For the case studied here, $\omega_1 = 1$ and $\omega_2 = \sqrt{2}$.

Accordingly, the function represents the temporal variation of one single point, meaning that the spatial complexity of the toy model is 1, whereas the spectral complexity (number of

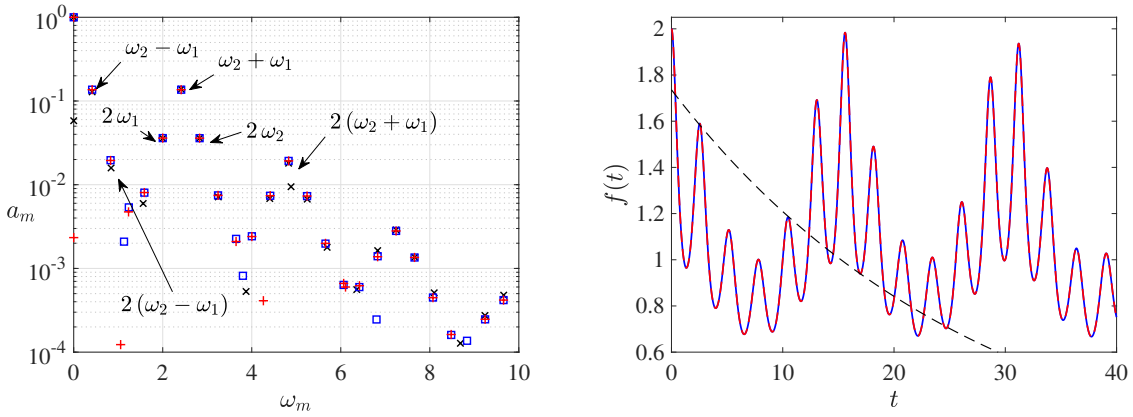


Figure 3.4: Results of DMD-d applied to 5000 snapshots with constant time interval $\Delta t = 8 \times 10^{-3}$ for the toy model in (3.3). Black, blue and red colours correspond to the solution obtained using $d = 1500$, 2000 and 2250. (a) Frequency versus amplitude of the DMD modes. (b) Reconstruction of the (blue) original toy model signal using (red) DMD−2000 and (black) DMD−1.

	Modes	**Relative error**	$\omega_2 - \omega_1$	$\omega_2 + \omega_1$	$2\,\omega_1$	$2\,(\omega_2 - \omega_1)$
$d = 1500$	148	1.0×10^{-2}	0.4141	2.4142	2.0002	0.8398
$d = 1750$	148	3.6×10^{-4}	0.4142	2.4142	1.9999	0.8281
$d = 2000$	77	1.3×10^{-4}	0.4142	2.4142	2.0000	0.8280
$d = 2250$	82	1.6×10^{-4}	0.4142	2.4142	1.9999	0.8284
$d = 2300$	80	1.6×10^{-4}	0.4142	2.4142	1.9999	0.8279

Table 3.7: Results of the DMD-d used to approximate the toy model in (3.3). The data consist of 5000 snapshots with a constant time interval $\Delta t = 8 \times 10^{-3}$. The Table shows the number of modes retained in the calculations, the relative root mean square error of the reconstructed signal and the values of the fundamental frequencies compared to the exact solutions $\omega_2 - \omega_1 \backsimeq 0.4142$, $\omega_2 + \omega_1 \backsimeq 2.4142$, $2\,\omega_1 = 2$ and $2\,(\omega_2 - \omega_1) \backsimeq 0.8284$.

frequencies) can be approximated with the number of retained modes, which may vary with the different values of d. Due to this huge difference in complexity, the standard algorithm cannot cope with the required accuracy and the HODMD should be used instead. As it will be proved later on, the results are quite novel and promising when compared to the standard DMD.

Using this toy model function, one could create a certain data set satisfying different criteria. For instance, the set of data employed in this section is composed of $K = 5000$ snapshots equidistant in time, $\Delta t = 8 \times 10^{-3}$. Applying the HODMD algorithm with the tolerances $\varepsilon_{\text{SVD}} = 10^{-12}$ and $\varepsilon_{\text{DMD}} = 10^{-5}$, the method approximates the original signal with a relative RMS error of $\sim 10^{-4}$ and after having performed a calibration process of the user-controlled parameters, i.e. the tolerances and d. The number of retained modes depends on the prior parameters but it should be noted that the modes leading the dynamical behaviour can be approximated to 13, resulting from the frequencies $\omega_1 \pm \omega_2$, $2\,\omega_1$, $2\,\omega_2$, $2\,(\omega_1 \pm \omega_2)$ and the zero-frequency mode, ω_0, which are in good agreement with the exact ones. These results are depicted in Table 3.7, where the number of modes, the relative RMS error of the reconstructed signal and the values of the fundamental frequencies can be appreciated. Note that the method is able to capture the dominant frequencies of the model at least for the interval of $1500 \leq d \leq 2300$. This is also shown in the graphical representation of the frequency and

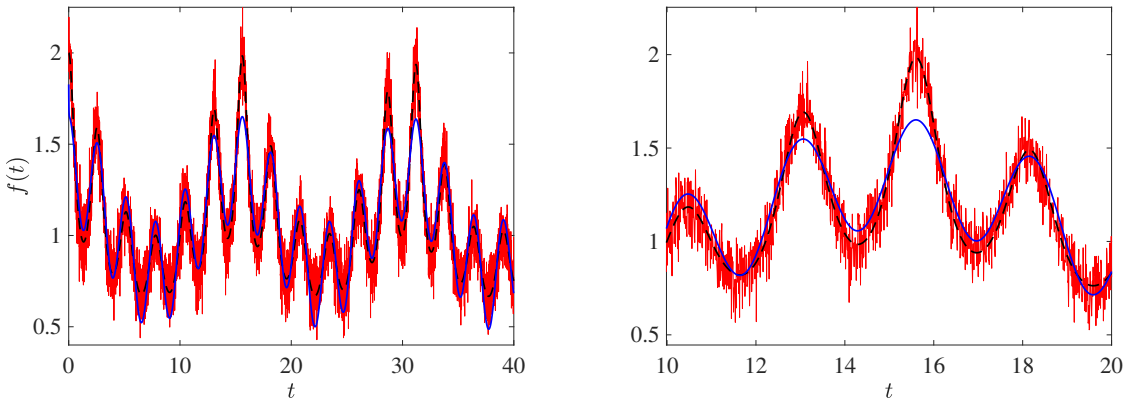


Figure 3.5: Results of DMD-d applied to 5000 snapshots with constant time interval $\Delta t = 8 \times 10^{-3}$ for the toy model in (3.3) with a 10% of random noise. (a) Original toy model signal (black dashed line), original signal with noise (red) and DMD-d reconstruction (blue) using ω_0 and $\omega_2 \pm \omega_1$. (b) Zoomed-in version of panel (a).

	$\omega_2 - \omega_1$	**Relative error**	$\omega_2 + \omega_1$	**Relative error**
$d = 1500$	0.4131	2.6×10^{-3}	2.4097	2.0×10^{-3}
$d = 1750$	0.4137	1.2×10^{-3}	2.4142	1.4×10^{-5}
$d = 2000$	0.4143	1.5×10^{-4}	2.4139	1.3×10^{-4}
$d = 2250$	0.4146	8.5×10^{-4}	2.4139	1.3×10^{-4}
$d = 2300$	0.4144	5.6×10^{-4}	2.4139	1.3×10^{-4}

Table 3.8: Results of the DMD-d used to approximate the toy model in (3.3) with a 10% of random noise. The Table shows the values of the fundamental frequencies compared to the exact solutions $\omega_2 - \omega_1 \backsimeq 0.4142$ and $\omega_2 + \omega_1 \backsimeq 2.4142$ in terms of relative error.

amplitude of the modes of Figure 3.4 and for the different values of d. Note as well that the relative error varies according to the parameter d, which is the main reason for performing a calibration process prior to the final analysis. In this case, for $d = 2000$, the relative error performed during the calculations is minimum, yielding a reconstructed signal approximately equal to the original one, see Figure 3.4 (right). The results are then compared with the reconstructed signal provided by the DMD-1 algorithm, which only recognises one mode with a negative damping rate and null frequency.

The preceding analysis has been replicated adding a 10% of random and normally distributed noise to the toy model function. The results shown in Table 3.8 show that the method is able to capture the underlying phenomena, predicting with small relative errors ($\sim 10^{-4}$) the fundamental frequencies of the model. Due to the noise level introduced, the tolerances ε_{SVD} and ε_{DMD} should be increased in order to remove the low-amplitude modes associated with the noisy data. This also influences the identification of lower amplitude modes of the original signal, but, in general terms, the agreement is considerably large and the method is able to capture the underlying motion. Figure 3.5 shows a comparison between the original signal with and without noise and the reconstruction using DMD-d and the fundamental frequencies $\omega_2 \pm \omega_1$ and ω_0. Particularly, it should be noted that only by retaining 5 modes, the method is able to reproduce the original signal with a relative RMS error of 8×10^{-2}. Recall that the input signal had added random noise, making the results even more promising.

3.3.2.4 General conclusions on modal decompositions

Both techniques have been applied to additional toy model functions with clean and noisy signals and the results are similar to those presented in this section. Overall, the HODMD method used here provides promising results and is capable of capturing complex phenomena within noisy experimental data in a much more efficient manner than the traditional DMD. HODMD can (i) accurately compute the leading modes, i.e. higher-amplitude modes, and (ii) approximate the frequencies of the lower-amplitude modes. This is particularly beneficial for the study of turbulent and multi-scale flows, where the underlying physics is known to be complicated and include a large number of flow scales, as in the case of the current project. The application of such algorithms to data from fluid numerical simulations will be covered in the following chapters.

4

Results and discussion

The previous chapters and sections have dealt with the theoretical concepts and computational implementation of the algorithms that allow for an extensive analysis of the different types of datasets. Particularly, the advantages of HODMD compared to the standard DMD algorithm have been theoretically deduced and proved using a toy model example. The aim of this chapter is to present the results of such techniques applied over three different flow regimes within urban environments, consisting of two building-like wall-mounted obstacles with varying separation ratios; recall Oke's classification [12] on flow regimes.

In § 3.1.1, the numerical domain of the final simulation and the different parameters varying with the flow regime have been defined. Also, in § 3.3.1, a quick review of the interpolation meshes has been conducted. Note that these two concepts are of great importance since they define the data, i.e. the snapshot matrix, over which the subsequent analyses will be performed, see § 3.3.2.1.

First, an analysis over the time-averaged velocity fields will be performed by means of different flow pattern identification methods. Then, different decomposition methods will be applied to a set of instantaneous velocity fields with the objective of extracting the main flow structures driving the dynamics within urban environments.

4.1 Time-averaged velocity fields

The following lines will be dedicated to the analysis of the time-averaged velocity fields at different cut planes for the different flow regimes. These quantities are the result of an averaging process over a considerable timespan for each case study, see Ref. [14], and they are interpolated over the above-described meshes. The resulting files are then processed using MATLAB.

Prior to discussing the results, it is important to emphasise the importance of some theoretical ideas that will serve as the foundation for accurate interpretation of the subsequent figures. Stull [50] defined two boundary layer areas in the context of meteorology: the zone above the mean height of the blocks, known as the **wake layer**, and the region underneath, denoted by the **urban canopy**.

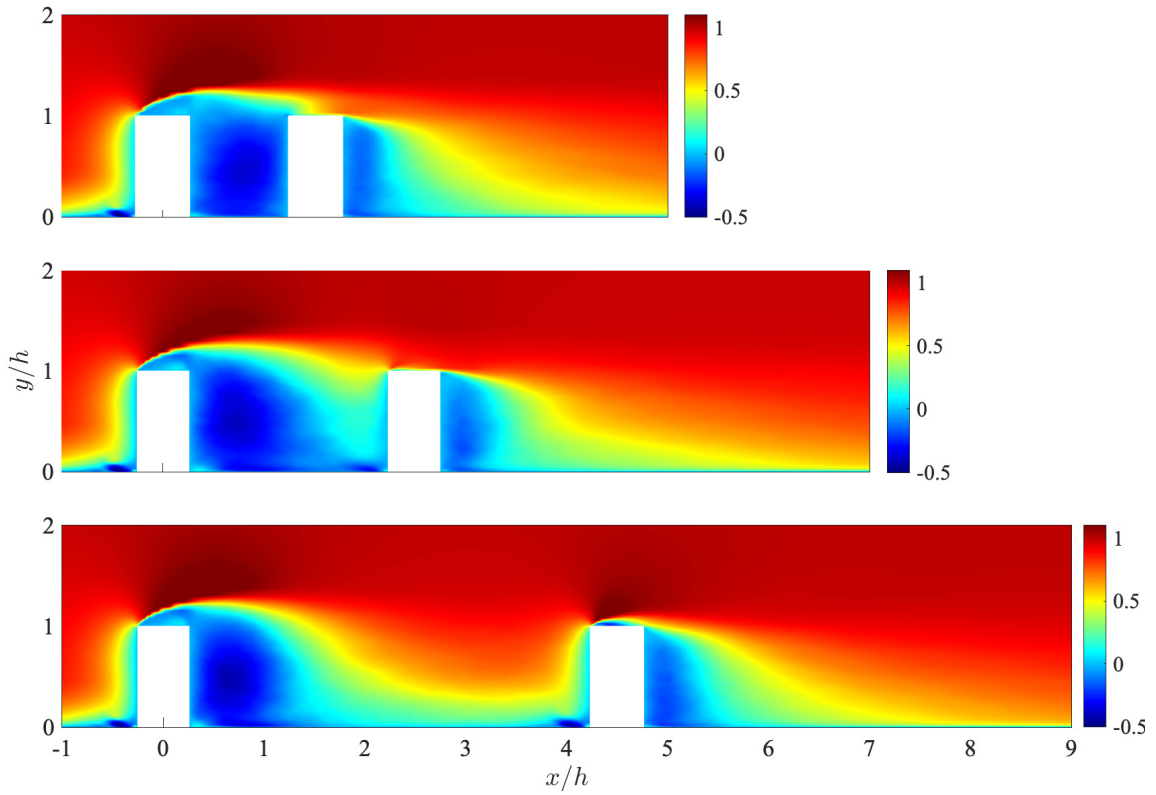


Figure 4.1: Time-averaged streamwise velocity field U normalised with U_∞ at plane $z/h = 0$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes [12].

From Figures 4.1 to 4.3, the time-averaged quantities for all the velocity components are depicted. The non-dimensional streamwise mean velocity component for the three flow regimes is illustrated in Figure 4.1. To begin with, a substantial wall-normal velocity gradient $\partial U/\partial z$ is noticed at the roof level ($0.8 < z/h < 1.2$) for all flow regimes due to flow separation occurring across the upper part of the upstream block. Above this layer, a high-speed freestream velocity profile ($z/h > 1.3$) is visible at the upper-roof level, and closer to the wall, i.e. at $z/h < 0.8$, a nearly zero streamwise mean velocity component with slightly negative values is detected in the intermediate part of the street. Due to the interaction of the flow above the canopy as the separation between the obstacles increases, the flow in between can be separated into two regions: a low streamwise velocity area close to the leeward side of the upstream block and a higher-speed region towards the windward side of the downstream block. As a result, a general assessment reveals that the mean streamwise velocity profiles depicted in Figure 4.1 are obviously impacted by the separation of the obstacles.

These results are in good agreement with the conclusions of Oke [12] and they are posed as the first numerical simulations to characterise the same flow structures. In the upper part of Figure 4.1, the flow above the canopy cannot penetrate the area between the obstacles and it is then characterised by large and negative velocity values. This is the basis of the skimming-flow regime, where the flow in the intermediate section of the obstacles is trapped and spinning, due to the low separation. Slightly dispelling the position of the downstream obstacle yields to the wake-interference regime, where the flow above the canopy scarcely interacts with the area between both obstacles. This interaction is only noticeable on the windward side of the downstream obstacle with larger streamwise velocities and, therefore, this regime might be seen as a mixture of the other flow cases. Finally, decreasing the height-to-separation ratio to

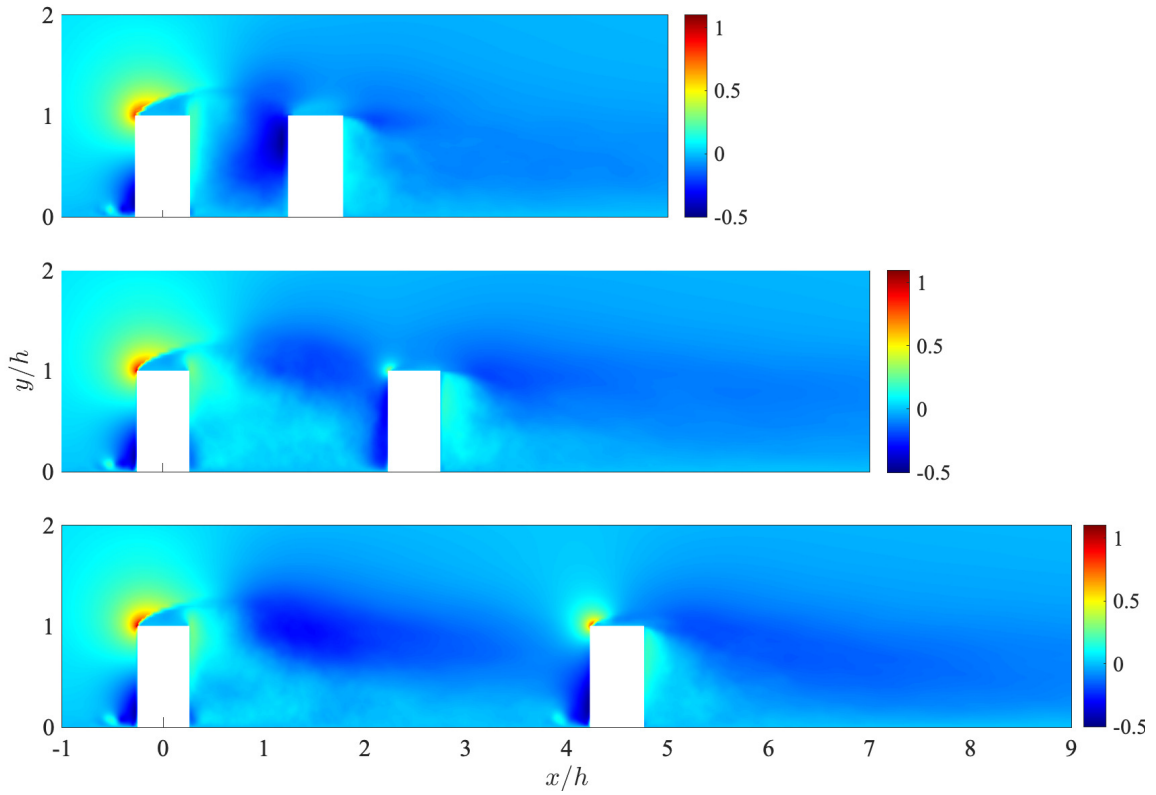


Figure 4.2: Time-averaged wall-normal velocity field V normalised with U_∞ at plane $z/h = 0$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes [12].

$h/\ell = 0.25$ results in the isolated-roughness regime, in which the flow around the obstacles is seen to behave in a much more independent way than in the previous cases. Therefore, the flow structures around both obstacles will be expected to be similar. On the other hand, it is also worth noticing that, for all cases, the leeward side of the downstream obstacle presents a region of large negative velocity values smaller but similar to the one generated downstream of the upstream obstacle. This is due to the absence of an obstacle immediately after the analysed one, which is the case for the upstream obstacle. Note as well that the shear layer over the first obstacle defines the start and height of the aforementioned region and, in the case of the second obstacle, increasing the separation translates into more noticeable shear layers, since it is less affected by the other obstacle.

It is of great importance characterising and understanding these flow features since they lead to different types of vortical structures such as the well-known arch vortex or the horseshoe vortex. Particularly, the regions with large and negative streamwise velocity values (blue) will be associated with regions of rotating flows with a strong relation to vortical structures. For instance, a horseshoe vortex located for all flow regimes almost at $y/h = 0$ on the windward side of the upstream obstacle appears as well around the downstream obstacle for the isolated-roughness regime, thus confirming the independence of the flow structures in this latter case.

Figure 4.2 illustrates the time-averaged wall-normal velocity component at the plane $z/h = 0$. Contrary to the streamwise component, the wall-normal velocity does not experience significant changes within the region between both obstacles. The inflow profile is imposed such that the unperturbed flow has no wall-normal component, i.e. it is fully hori-

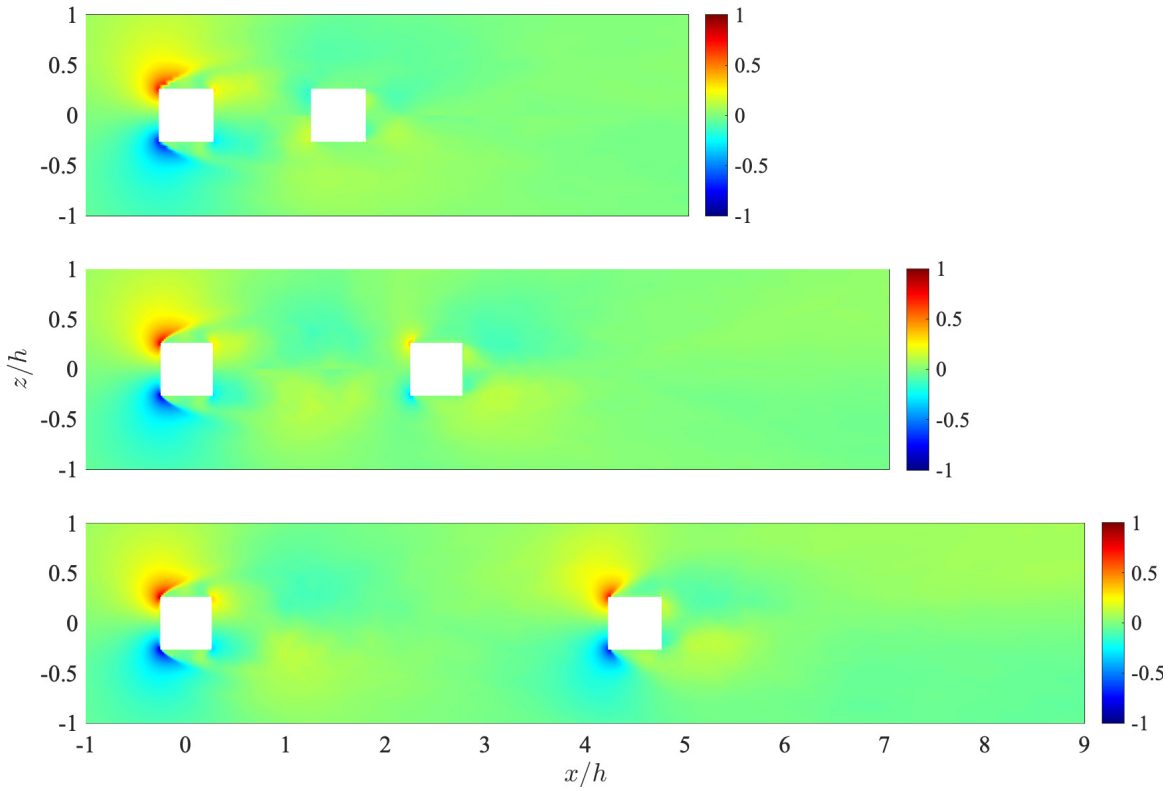
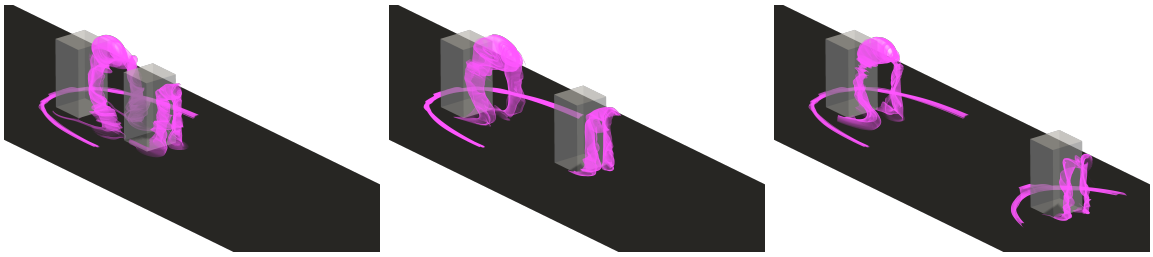


Figure 4.3: Time-averaged spanwise velocity field W normalised with U_∞ at plane $y/h = 0.25$ for (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness flow regimes [12].

zontal. This is the reason why the streamwise velocity field had near-maximum levels prior to the perturbed regions, see Figure 4.1, and near-zero wall-normal velocity values. Notwithstanding the above, the flow experiences a high-velocity region at the top edge of the upstream obstacle, as a result of the impact of the flow over the edge and the shear layer formed over the upper part. Furthermore, this region is also appreciated to a lesser degree on the downstream obstacle of the isolated-roughness regime. On the other hand, it is also true that the wall-normal velocity is negative on the region of the wake, meaning that the flow tends to be directed towards the ground. Once the fluid encounters with the first obstacle, the flow is abruptly deviated to the upper part of the domain and it then rejoins the lower part on the wake.

To conclude the analysis over the time-averaged velocity components, Figure 4.3 depicts the mean spanwise velocity at $y/h = 0.25$ plane. In this case, the analysis is analogous to the previous one, since the inflow conditions do not impose any type of perturbation in the spanwise direction. Therefore, the unperturbed regions of the flow would have a near-zero spanwise velocity, as happened with the wall-normal component. In this regard, some positive and negative values are seen on both lateral sides of the first obstacle and for all flow regimes, due to the impact of the flow over the edges of the obstacles, which deviates the flow towards the outer parts of the domain. Since the flow has been seen to be reattached when reaching the downstream obstacle in the isolated-roughness regime, also see Figure 4.1, the same spanwise structures are appreciated on this obstacle. Note that, while the streamwise and wall-normal velocities would be symmetric over $z/h = 0$ if represented in a constant$-y$ plane, the spanwise component is known to be anti-symmetric.



(a) Skimming flow **(b)** Wake interference **(c)** Isolated roughness

Figure 4.4: Main vortical structures formed around two wall-mounted obstacles with different separation ratios: (**a**) $h/\ell = 1$, (**b**) 0.5 and (**c**) 0.25, visualised by means of the streamlines flow patterns. Note the arch vortex on the leeward side and the horseshoe vortex around the obstacles.

In short, the streamwise component is seen to be the one driving the behaviour within the region enclosed by both obstacles and the wall-normal and spanwise velocities, the ones responsible for the flow behaviour when reaching the obstacles. This means that the time-averaged streamwise component restricts the interaction of the surrounding flow with the one within the canopy, which yields to the appearance of vortical structures within these areas. Recall that the inflow conditions have been imposed in the streamwise direction, thus making the prior conclusions coherent. Note that one possible variation of the inflow conditions is to introduce some angle of incidence (AOI) such that the flow is not fully in the streamwise direction, as Monnier et al. [32] performed.

Finally, depicting the streamline flow patterns over the time-averaged fields enables the characterisation of the key vortical features present in such flowfield. This is exactly what Hunt et al. [27] did throughout their experimental studies, and they discovered four different types of vortices, as shown in Figure 1.6. The objective is to identify the most relevant structures using data from a fluid numerical simulation, which has considerably higher temporal and spatial resolution than a wind-tunnel experiment. Having identified these structures, a more comprehensive analysis of their formation may be carried out in the following sections.

Laying our eyes on Figure 4.4, it is obvious to see an arc-shaped vortical structure on the leeward side of both obstacles. This flow pattern is similar to the well-known arch vortex, which is a two-leg structure found in flows through wall-mounted obstacles. In the skimming-flow regime, its legs are slightly displaced towards the windward side of the downstream obstacle, spanning the gap between the obstacles; in the other two cases, an arch vortex appears immediately after the obstacle and is fully straight. Notice that the downstream arch vortex is not as strong as the upstream one, owing to the flow exchanging momentum with the initial building. On the other hand, the appearance of this structure corresponds with the negative streamwise velocity area on the leeward side of the obstacles (see Figure 4.1).

It also is worth observing the horseshoe vortex created around the upstream obstacle and, exceptionally, around the downstream obstacle in the isolated roughness case. This validates the idea that the flow reattached when it reached the second obstacle in the IR case. This region appears as well as a recirculation bubble on the windward bottom side of the obstacles for the contours shown in Figures 4.1 and 4.2 on the symmetry plane. As a final comment, this figure also illustrates the different symmetries of the velocity components: the streamwise and wall-normal components are seen to be necessarily symmetric whereas the spanwise component, anti-symmetric. This is obviously dependent on the inflow conditions.

In conclusion, the analysis performed on the preceding figures have allowed for the characterisation of the major flow patterns present in the time-averaged fields. These results have been proved to be in good agreement with prior experimental analyses, which may be used to validate the various fluid numerical simulations used here. Furthermore, not only have these flow patterns been characterised, but also their development as the distance between the obstacles has been varied. This leads to a wide range of possibilities for building placement within an urban environment.

4.2 Proper orthogonal decomposition

Once the flow structures present in the time-averaged fields have been identified, the purpose of the following sections is to perform an analysis over the instantaneous velocity fields aiming at characterising the main mechanisms driving the flow dynamics. In such a fashion, the first step is to perform a POD analysis over the datasets of the different flow regimes.

Thereafter, the snapshot matrices over which the POD will be applied should be built based on the description of § 3.3.2.1. In this case, since the number of fields is limited due to computational reasons, convergence can be increased by imposing symmetry over the data to be analysed. To that end, let us consider the initial snapshot of the original data as A_1 and its flipped version as A_{f1}. Then, the final snapshot matrix over which SVD is applied would be $[A_1, \ldots, A_n, A_{f1}, \ldots, A_{fn}]$, where the number of columns will be twice the number of snapshots of the original dataset. In this regard, the spatial complexity of the problem is increased while symmetry is imposed for the resulting POD modes.

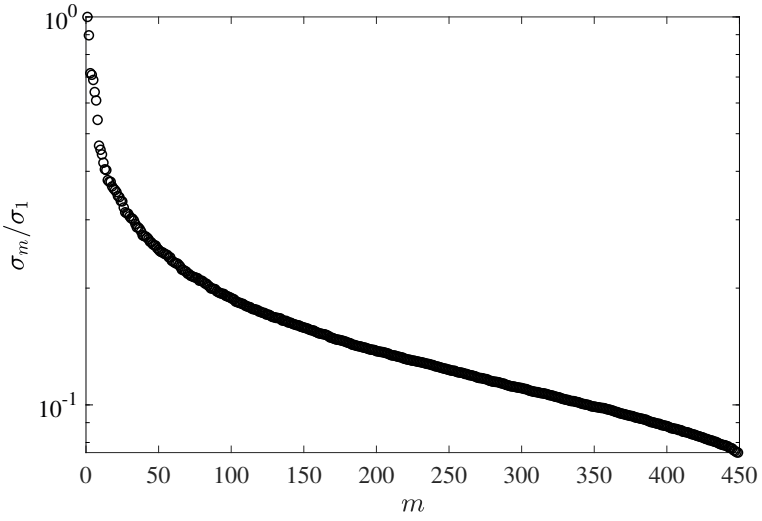


Figure 4.5: Singular-value distribution normalised with its maximum value σ_1 of the POD modes corresponding to the streamwise velocity of the skimming-flow regime and symmetry imposed.

Let us begin the POD analysis with the skimming-flow regime to illustrate the potential of the proper orthogonal decomposition while gaining further insight into the physics of this case study. First of all, a general examination of the singular-value distribution should be conducted to conclude those modes energetically more important. In Figure 4.5 this distribution is depicted for the streamwise velocity of the skimming-flow regime. Despite the gaps in

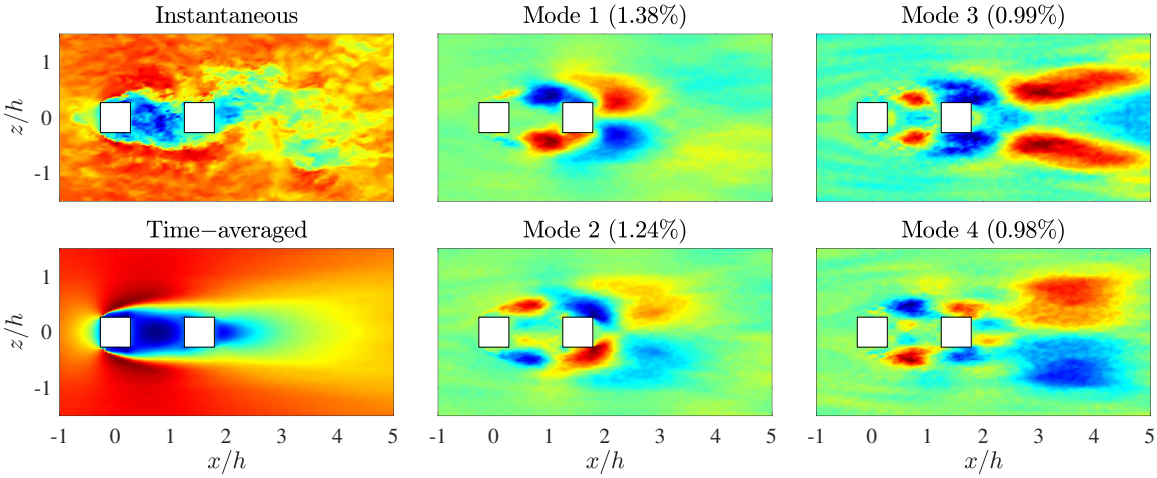


Figure 4.6: POD analysis of the streamwise velocity at $y/h = 0.25$ for the skimming-flow regime ($h/\ell = 1$). Shown are the instantaneous and time-averaged fields and the associated four most dominant POD modes.

energy between the first set of singular values, all modes are seen to have a significant contribution to the final system, meaning that no clear truncation region is appreciated. However, it should be noticed that the objective behind the present is to gain further understanding of the flow dynamics rather than focusing on the reconstruction error (reduced-order models). Therefore, keeping this in mind, the following figures will be dedicated to the analysis of the spatial modes energetically more relevant.

Figure 4.6 shows the classical POD decomposition of the streamwise velocity fields. The whole set of instantaneous fields can be decomposed in their time-averaged quantity and a set of orthogonal POD modes. By inspecting the flowfield, one can observe the formation of an unsteady feature similar to a von Kármán vortex street. Therefore, thanks to this modal decomposition, one could extract those flow features, with their corresponding energy contribution to the system, responsible for the unsteady flow motion. Particularly, the two modes with the largest contribution to the system exhibit some vortex shedding flow patterns close to those of a von Kármán vortex street. However, due to the presence of the second obstacle, these anti-symmetric structures are modified such that their location slightly varies. In this case, the first mode structures are located towards the windward side of the downstream obstacle and the wake and the second one is dominated by structures on the leeward side of both obstacles. Indeed, just retaining these two modes, the reconstructed flowfield presents an RMS error of $\epsilon = 3.32\%$ compared to the reference one, which enhances the contribution of the prior modes to the system. Besides, recalling the position of the arch vortex and the findings of Monnier et al. [32] during their experimental studies, one could associate these flow features to the formation or the destruction of already known vortical structures. This is the case for the third mode, which exhibits certain symmetric structures in the region between the obstacles, where the arch vortex is located. Hence, this mode could be closely related to the formation of such vortical structure. Furthermore, together with the fourth mode, it shows much more prominent structures on the wake than the previous ones. This will also be the main course of action of subsequent analyses.

Despite not being the main purpose of the study, it is also interesting to conduct a general analysis over the reconstructed flowfield when different latent space dimensions are employed. Note that, when increasing this dimension, the number of higher-order modes retained will

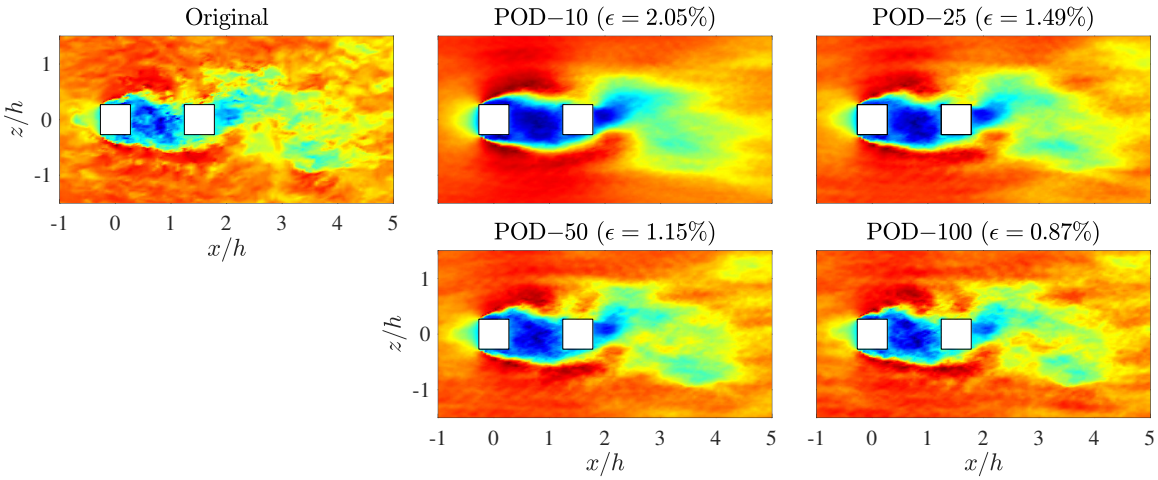


Figure 4.7: Reconstructed streamwise velocity at $y/h = 0.25$ from POD at different latent space dimensions. The corresponding relative ℓ_2-norm of errors (ϵ) are listed in brackets.

consequently rise, thus lowering the final RMS error up to a 0.87% at the latent space dimension of 100 POD modes, see Figure 4.7. These results are promising for the creation of reduced-order models, since only retaining a few modes, the method is able to reconstruct the original field with a relatively low error.

However, it should be noted that the POD algorithm, while efficiently computed in the ℓ_2 sense, is based on determining the linear correlation between the different snapshots. Some state-of-the-art deep learning algorithms outperform in this regard due to the use of non-linear activation functions. Fukami et al. [51] developed a hierarchical autoencoder based on a convolutional neural network for nonlinear mode decomposition of fluid field data. Using this, the method can reconstruct the original flowfield data in a much more efficient manner and with a fewer number of modes. Taking this into consideration, if a reduced-order model of the original data is desired, these cutting-edge machine learning algorithms should also be explored. For the time being, this chapter will concentrate on the physics of these types of flows rather than the optimisation of the relative error incurred during computations and the construction of reduced-order models.

Once the potential and applications of such techniques applied within the context of urban flows for the skimming-flow regime have been introduced, the same analysis will be extended and compared to the other flow cases. Analogously, the first step is to generally examine the distribution of the singular values with the goal of identifying energy gaps that make some modes more important than others in an energetic sense. This distribution is depicted in Figure 4.8 for the three flow regimes. It should be noted that the number of retained modes is strictly dependent on the number of snapshots of the input data, which in the current study varies between cases, see § 3.3.2.1. Although the ideal scenario for properly comparing them would be to have the same exact number of modes for all cases, a general inspection reveals a gap between the second and third modes for all regimes, which is especially noticeable in the wake interference and isolated roughness cases. As a result, the first two modes will be known to contain the more relevant information of the flowfield data.

On the other hand, despite the fact that the number of retained modes in the wake-interference regime is smaller than that of the skimming-flow regime, the latter singular values

($r > 210$) show a clear decay, implying that additional higher-order modes would not introduce more relevant information. This is not the case in the other two cases, where adding more snapshots would almost certainly introduce some relevant data into the analysis.

Closely related to the preceding point, each velocity must make a unique contribution to the system. The energy of the most dominant POD mode for the wall-normal component, in particular, represents approximately 45% of the maximum energy value of the other velocity components. This means that the influence of the wall-normal velocity fluctuations is less important to the overall system than the influence of the other velocity components. This is consistent with the findings of Monnier et al. [32], who discovered that for a zero-incidence angle, the streamwise and spanwise fluctuating components are more significant than the wall-normal component. On the basis of the above, the wall-normal component will be omitted from the analysis and only the streamwise and spanwise components will be further studied for the three flow regimes.

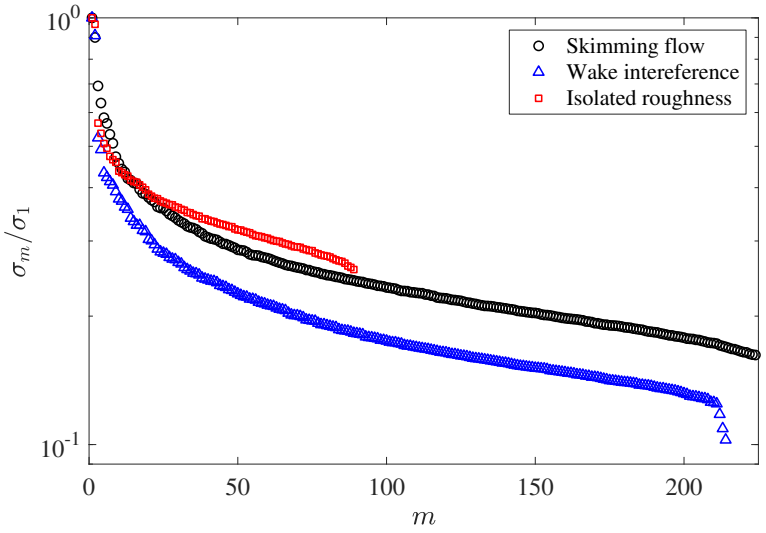
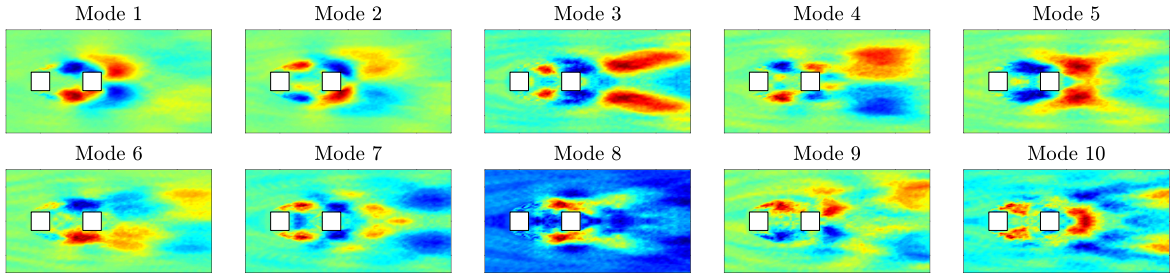
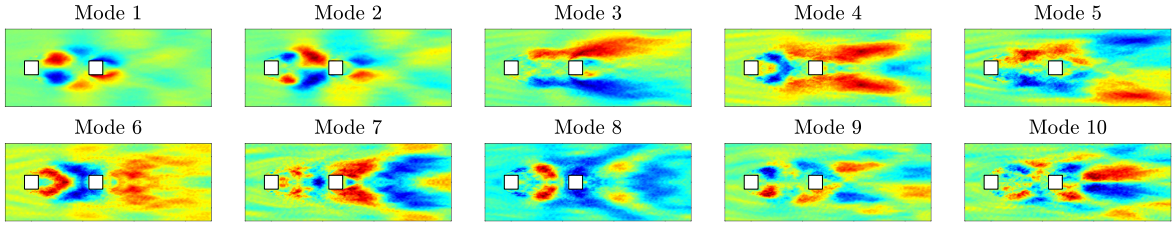


Figure 4.8: Singular-value distribution normalised with its maximum value σ_1 of the POD modes corresponding to the hole set of velocity components (streamwise, wall-normal and spanwise) of the (black) skimming-flow, (blue) wake-interference and (red) isolated-roughness regime. The number of retained modes coincides with the number of columns of the snapshot matrix, i.e. 225, 215 and 90, respectively.

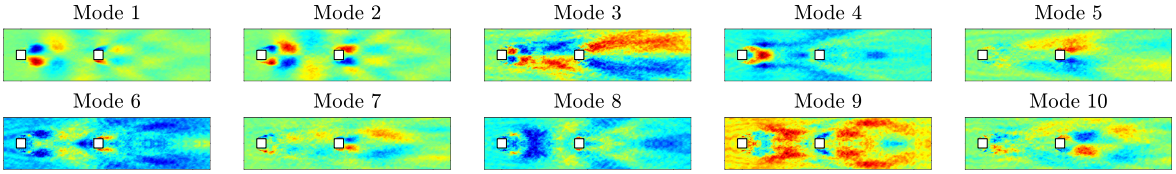
Starting with the streamwise velocity field, Figure 4.9 depicts the orthogonal POD basis for the modes of the three flow regimes. First of all, regarding the first two modes for all the flow regimes, it could be noted their apparent similarity, dominated by fluctuating regions on both sides of the blocks, mainly due to the interaction of the wake region with the shear layer on both sides. These regions nearly match with the high turbulence kinetic energy (TKE) regions of the streamwise component identified by Monnier et al. [32] for an array of building-like blocks. However, the main difference among regimes relies on the position of the secondary structures, which are associated with the downstream block. For instance, while these fluctuating regions span the zone in between the obstacles for the skimming flow and the wake interference cases, in the isolated roughness, they are only located on the immediate leeward side of the upstream block. In such a fashion, it can be concluded that increasing the separation of the obstacles does not yield to more fluctuating regions, at least for the more energetic modes, as it occurs in a vortex shedding case.



(a) Skimming flow



(b) Wake interference



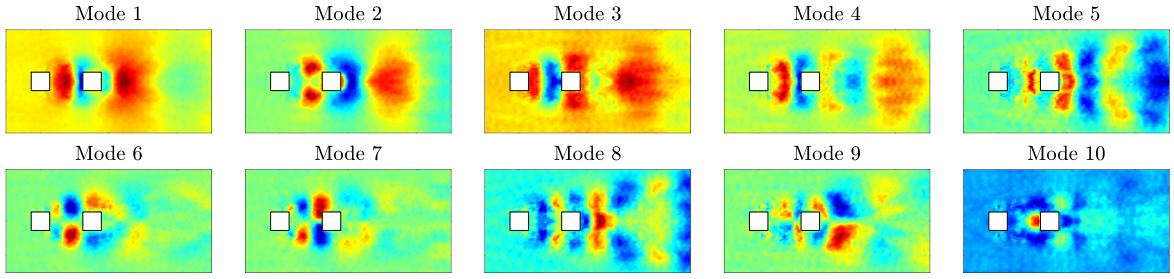
(c) Isolated roughness

Figure 4.9: POD orthogonal basis of the streamwise velocity fields at $y/h = 0.25$ for the different flow regimes.

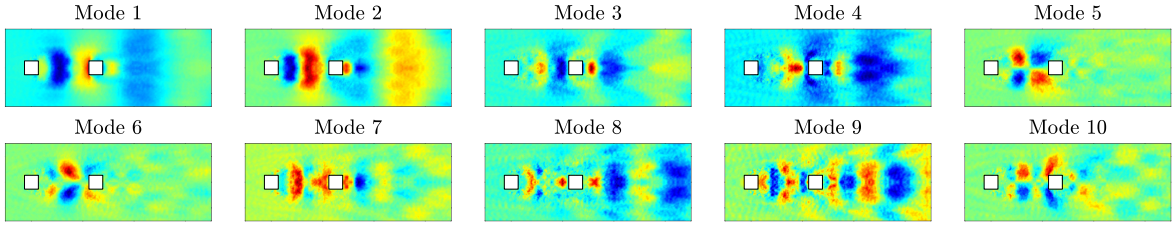
As will be detailed in § 4.3, the structures observed in the first two modes are remarkably similar to those associated with the high-frequency arch-breaking modes. Higher-order modes are more focused on the wake as well as the creation of these vortical structures. Particularly, for all cases, the third and fourth modes are clearly impacted by certain major fluctuating areas on the wake. This pattern is quite similar to that of the time-averaged field, therefore it is reasonable to assume that these modes are low-frequency.

After this number of modes, the correlation begins to diverge among the many cases, each with its own set of properties. In the isolated-roughness regime, modes 5, 6, 7 and 10 present flow structures around the downstream obstacle independently, which supports the independent behaviour of the flow around both obstacles, being the only regime providing such isolated structures. The difference between the skimming-flow and the wake-interference regimes, on the other hand, is based on the intersection of the fluctuating areas in the wake-interference regime owing to the increasing spacing between the obstacles, as shown in modes 4 and 6. Higher-order modes in the case of skimming flow are shown to be harmonics of the dominant ones.

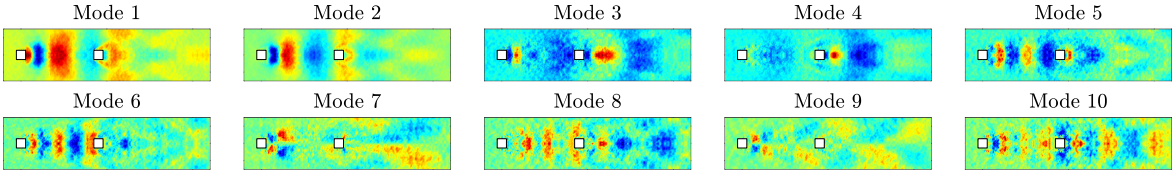
In terms of the spanwise component, Figure 4.10 depicts the set of POD orthogonal modes for all flow regimes. On the basis of the singular value distribution of Figure 4.8, being the most dominant, the first two modes will be firstly analysed. In this situation, unlike the streamwise component scenario, increasing the distance between the obstacles increases the number of spanwise fluctuations in that region. These zones are only due to the interplay of lateral flow within the canopy, which tends to shatter the arch vortex formed on the leeward



(a) Skimming flow



(b) Wake interference



(c) Isolated roughness

Figure 4.10: POD orthogonal basis of the spanwise velocity fields at $y/h = 0.25$ for the different flow regimes.

side of the upstream obstacle. Moving on to the modes that promote the creation of such vortical structures, it is worth noting that modes 6 and 7 are not as uniform as the previous ones; the two separated and symmetric areas correspond to the formation of each of the arch vortex's legs.

The spanwise velocity component, like the streamwise one, reveals some isolated features exclusively for the downstream obstacle of the isolated-roughness regime. As a result of the flow's loss of energy after reaching the first obstacle, the method is able to discern the flow structures around the first and second obstacles.

4.2.1 Temporal analysis using FFT

§ 2.3.1 introduced the concept of temporal coefficients in the context of POD-mode analysis. The goal of this section is to conduct a quantitative analysis of the temporal coefficients associated with the POD spatial modes by characterising their dominant frequencies. The fast Fourier transform (FFT) technique will be used to generate, from this temporal data, the power spectrum revealing the prominent frequencies of each mode.

Let us begin with the analysis of the temporal coefficients of the first two POD modes for the three flow regimes, which have already been shown to be dominant, recall § 4.2. This temporal distribution is shown in Figure 4.11 by two phase-shifted but frequency-similar signals. As a result, the dominating POD modes' frequencies are predicted to be comparable,

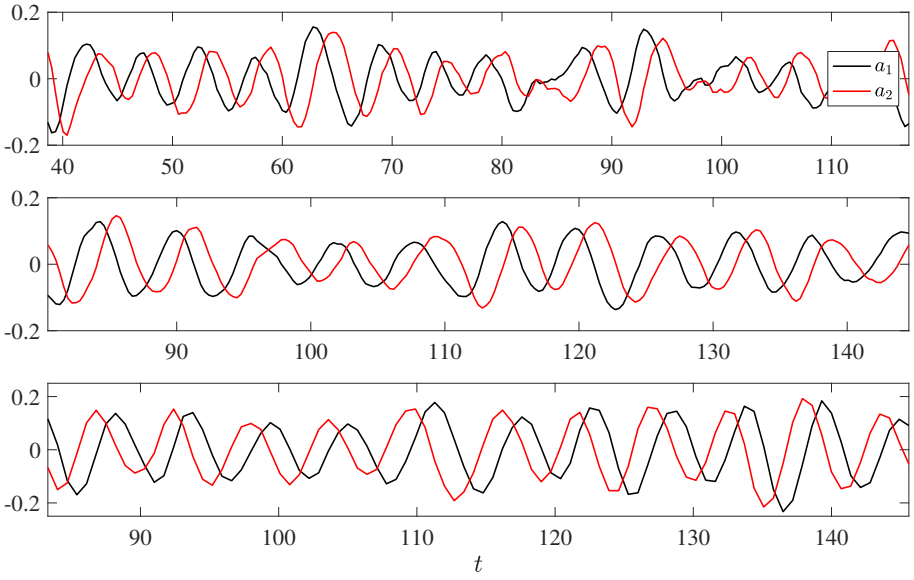


Figure 4.11: Temporal coefficients a_1 and a_2 of the first two POD modes for the (top) skimming-flow, (middle) wake-interference and (bottom) isolated-roughness regime.

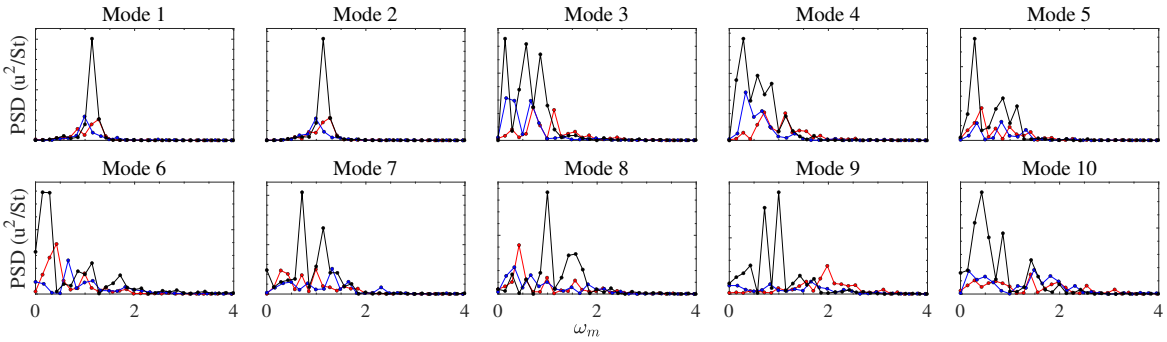


Figure 4.12: Power spectrum of FFT scaled with the Strouhal number St and applied to the temporal coefficients of the POD modes: (red) skimming-flow, (blue) wake-interference and (black) isolated-roughness regime. Peaks show the dominant frequencies associated with each POD mode.

implying that the flow dynamics at that frequency will be a mixture of these POD modes. It is also worth noting that the isolated-roughness regime is the closest to a periodic evolution, with prominent frequencies that can be easily recovered using the FFT technique. In the other cases, some spurious phenomena seem to have been added, thus introducing secondary frequencies into the spectrum.

Keeping this in mind, Figure 4.12 depicts the power spectrum for the first ten POD modes of all flow regimes. At first glance, it is clear that the technique can distinguish between low- and high-frequency modes within the range $[0, 2]$. To begin with, the isolated-roughness regime yields clearer conclusions, and the frequencies are observed to match in all three cases, at least for the initial modes. Confirming the previous claims, the two initial modes are dominated by a peak frequency in the $[1 - 1.2]$ range. Furthermore, modes $3 - 6$ are seen to be of low frequency, supporting the hypothesis expressed in § 4.2 that they are intimately connected to the creation of significant vortical structures such as the arch vortex. Higher-order modes can be thought of as a mixture of the preceding ones and if even higher-order modes ($r > 10$) were studied, certain high-frequency phenomena would be captured due to the smaller turbulent

flow scales.

The results presented hitherto have provided further insight into the physics of the flow within urban environments, being able to differentiate between low- and high-frequency phenomena, which are closely related to the formation or destruction of the time-averaged vortical structures. The following sections will deal with a similar modal-decomposition analysis that also considers the dynamical behaviour of the system. In such a fashion, it would be possible to characterise the flow dynamics based on the frequency of each mode.

4.3 Higher-order dynamic mode decomposition

4.3.1 Initial calibration and parameter selection

Because of the large number of frequencies and spatio-temporal structures present in turbulent flows, identifying flow patterns is challenging. As with the toy model in § 3.3.2.3, it is critical to test the robustness of the results. The modified Koopman operator employed in the HODMD algorithm gives a solution fulfilled by all sub-groups of data (snapshots) simultaneously evaluated and capable of capturing large-scale and large-amplitude features from the highly varied frequencies seen in the flow [52]. Therefore, the DMD-d algorithm is used for this purpose with varied tolerances and d parameter values.

Figure 4.13 shows the frequency versus amplitude of the different modes computed using DMD-d with a variety of parameters for the three flow regimes. Although this technique can be used to analyse datasets with non-equidistant in time snapshots, the intricacy of the flow described here would make modelling much more challenging. As a result, the DMD-d analysis can be conducted considerably more efficiently if the data specifications in Table 3.4 are used. The huge number of modes calculated with each variation necessitates the use of well-established criteria to identify the most robust modes that best characterise the entire system. In this case, the largest-amplitude and lowest-frequency mode will possess the more relevant information about the system and the classification will be made based on this distinction. As highlighted in Figure 4.13, these modes are known to form clusters throughout the spectrum. The amplitude and frequency of the selected modes will be, therefore, the average value of the collection of modes. The number of preserved modes is also a function of the tolerances employed, which were $\epsilon = 10^{-3}$ and 10^{-4} in this study. Similarly, the other user-controlled parameter, d, changes depending on the number of snapshots to be analysed. The skimming-flow and wake-interference regimes, with $K = 225$ and 215 snapshots, respectively, were studied with $d = 10, 20$ and 30, whilst the 90 snapshots of the isolated roughness case were studied with $d = 5, 10$ and 15. Note that, as mentioned during the theoretical derivation of the Koopman operator (see § 2.3.4), as the number of snapshots is reduced, the value of d, which represents the characteristic sliding window process, must also decrease in the same proportion [52].

Furthermore, the relative error made in the calculations remains fenced in the set of tolerances used. The goal of this study, however, is to identify the largest-amplitude modes in order to provide a broad description of the fundamental patterns driving the flow, rather than to build any accurate reduced-order models based on the physical knowledge of the flow. As a result, the relative error will not be examined further in this study.

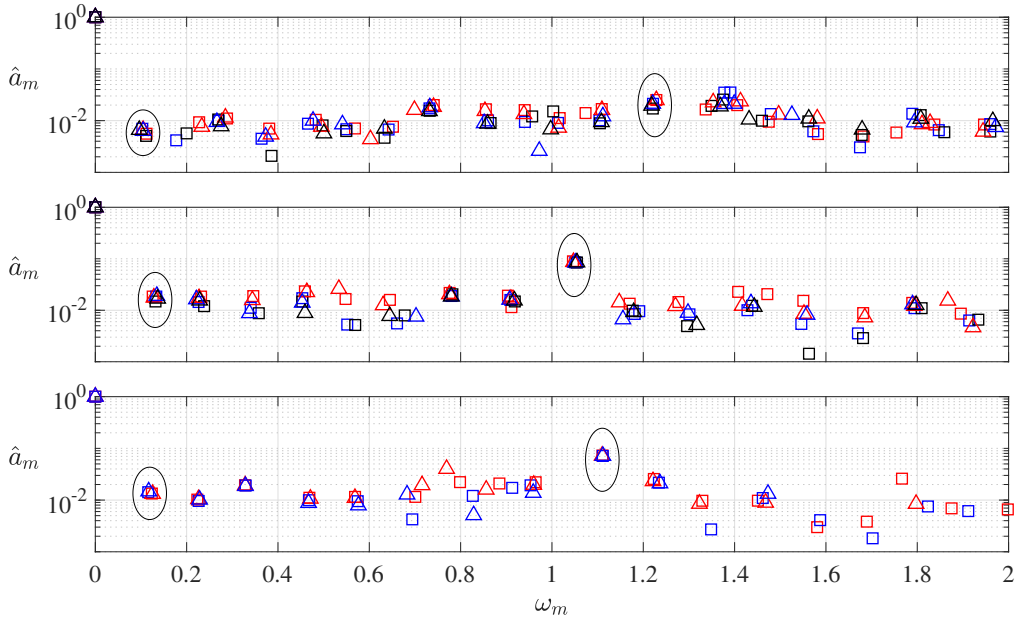


Figure 4.13: DMD-d calibration. Amplitude scaled with its maximum value ($\hat{a}_m = a_m/a_0$) versus frequency ω_m computed with different tolerances for (top) skimming flow, (middle) wake interference and (bottom) isolated roughness. Squares represent $\epsilon_{\text{SVD}} = \epsilon_{\text{DMD}} = 10^{-3}$ and triangles, 10^{-4}. Red, blue and black correspond to $d = 10, 20, 30$ for SF ($K = 225$ snapshots) and WI ($K = 215$ snapshots) and $d = 5, 10$ for IR ($K = 90$ snapshots).

It is worth noting that when the distance between the obstacles increases, the flow complexity decreases, allowing a smaller number of snapshots to be used to estimate the same flow behaviour. Note as well that, despite the fact that the number of snapshots in the isolated-roughness regime is smaller, the time span covered is of the same order of magnitude, as the time step between snapshots has been increased. This is especially essential when dealing with computationally expensive data, but one should bear in mind that in order to accurately represent smaller turbulent scales, a larger number of snapshots with a shorter time step should be utilised instead. However, in general, the HODMD provides a fair balance of computational cost and accuracy.

4.3.2 Global temporal modes

Table 4.1 summarises the spectrum of temporal DMD modes found during the calibration method described in the preceding section. Note that just two modes have been identified: a low-frequency mode and a high-frequency mode, the latter of which relates to the largest-amplitude mode (apart from the zero-frequency mode). Due to their significant interaction with notable vortical structures such as the arch vortex, these modes will be referred to as **arch-generator** and **-breaker** modes, respectively. Their amplitude and frequency values are found to be close throughout the various regimes; however, increasing the distance between the obstacles slightly reduces the frequency of the arch-breaker mode. This analysis was previously carried out in a similar manner utilising POD, which decomposed the flow into low and high-frequency behaviour, which is responsible for the formation and destruction of various flow patterns. The high-frequency mode B, like in the POD case, reveals more information about the breaking vortex phenomenon and is found to be the dominating process. This technique also finds several other modes, which may be thought of as harmonics of the preceding ones and will be discussed briefly in the next sections.

Mode		SF	WI	IR
Arch-generator (G)	ω_m	0.11	0.13	0.12
	$\hat{a}_m$	0.0062	0.018	0.014
Arch breaker (B)	ω_m	1.22	1.05	1.11
	$\hat{a}_m$	0.021	0.086	0.072

Table 4.1: Summary of the frequencies and amplitudes of the DMD modes shown in Figure 4.13. From top to bottom, the modes are ordered from lower to higher frequency.

To begin with, one should recall that the HODMD method decomposes the flow data into a collection of Fourier-like complex conjugate modes associated with the frequencies ω_m and $-\omega_m$, which can have both real and imaginary contributions. These real and imaginary components can be thought of as an oscillation that visits a certain state as time passes. This is why it is important to choose which parts of the study to examine without jeopardising the whole analysis. To illustrate this, Figure 4.14 depicts a two-dimensional representation of the real, imaginary, and absolute sections of the three velocity components for the arch-breaker mode. The goal is to compare the flow structures in each of the parts and create some criteria for future analyses. While the latter case emphasises the fluctuating areas of the flow, implying that the surroundings have no impact, the absolute part may be regarded as a mixture of the preceding components, highlighting both the unperturbed and disturbed regions in a more general fashion. In this approach, the real part of the modes will be investigated further, while keeping in mind that the other components may also bring some physical meaning.

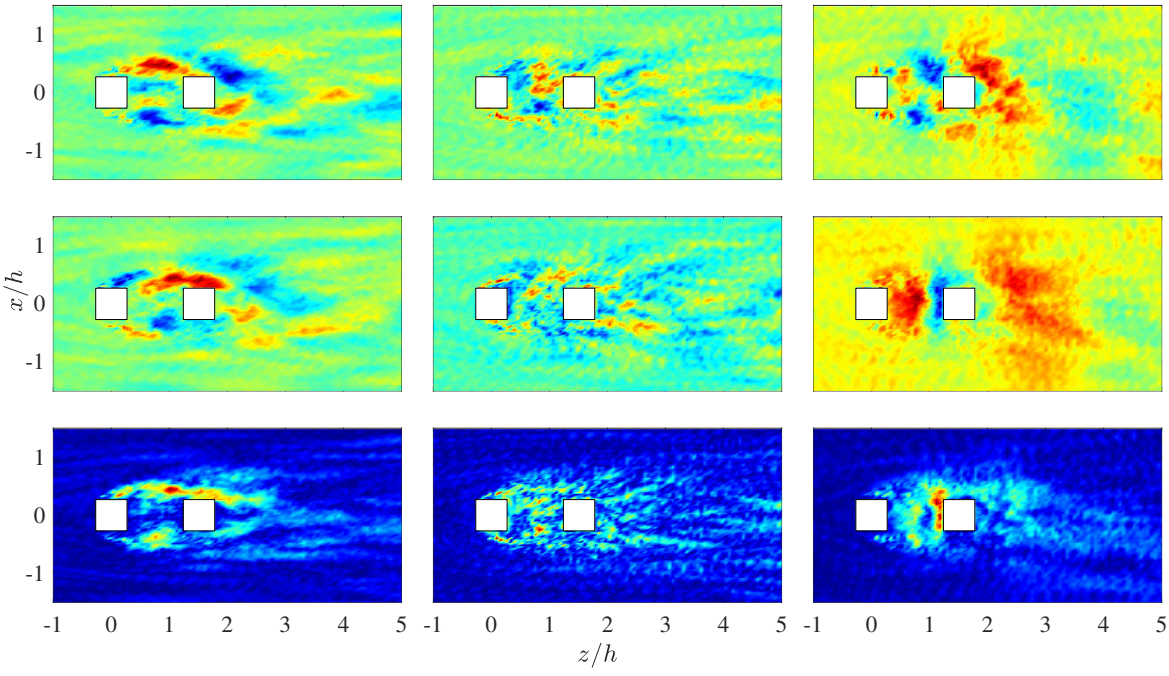


Figure 4.14: Two-dimensional representation at $y/h = 0.25$ of the arch-breaker mode ($\omega_m = 1.22$) for the skimming-flow regime. Contours are scaled with their maximum value. From top to bottom, the real, imaginary and absolute parts of the modes are represented. From left to right, the streamwise, wall-normal and spanwise components are shown.

Using Figure 4.14 as a starting point for the analysis of the DMD modes, namely the arch-breaker mode of the skimming-flow regime, it is possible to observe how these structures relate to the first two POD modes of Figures 4.9 and 4.10. Particularly, both the real and

imaginary parts of the streamwise component show some fluctuating regions towards both windward lateral sides of the downstream obstacle. Besides, the spanwise component is dominated by a high-velocity fluctuation just in front of the downstream obstacle and on the wake. No clear conclusions can be extracted from the wall-normal component, since it does not have a significant contribution in the constant$-y$ plane selected. This supports the idea stated during § 4.2, where the singular values associated with the wall-normal velocity component at the same cut plane did not contribute to the system as much as the other velocity components. However, the following three-dimensional representations will allow for the characterisation of the flow structures in the wall-normal direction.

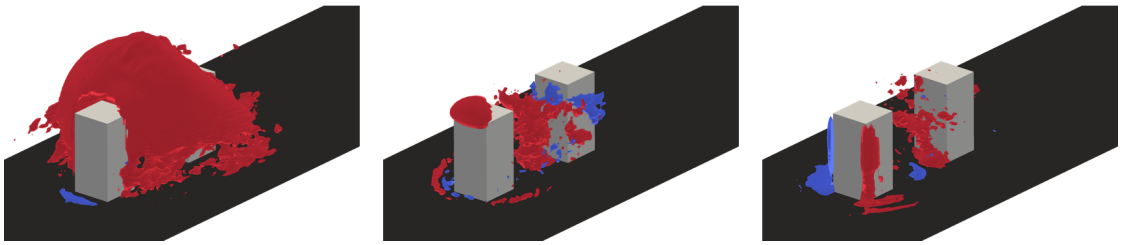
4.3.2.1 Arch-generator modes

Figures 4.15 and 4.16 show a three-dimensional view of the main DMD modes presented in Table 4.1 as a function of the separation ratio between the obstacles, aiming at providing a general overview of the main flow structures related to each mode. Figure 4.15 corresponds to the arch-generator low-frequency mode, whereas the arch-breaker high-frequency mode is depicted in Figure 4.16. In the first case, the arch-generator mode presents some characteristic three-dimensional structures which relate to the formation of the arch and horseshoe vortices in a low-frequency fashion.

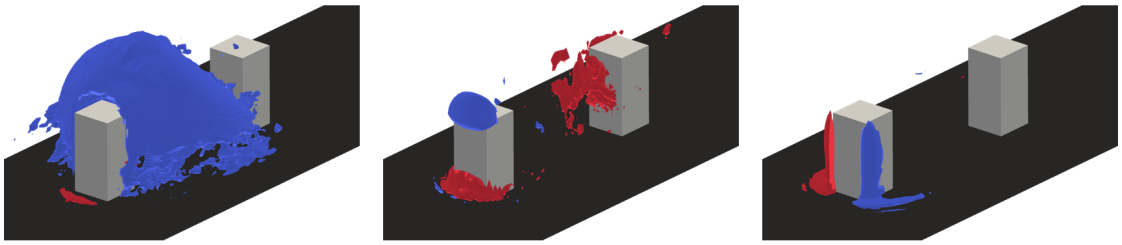
Streamwise velocity Firstly, regarding the velocity in the streamwise direction, see Figure 4.15 (left), a dome-like structure prevails in the intermediate section of the obstacles. Increasing the distance between the obstacles does not result in a reallocation of such flow pattern, which remains in the same place throughout all flow regimes. This indicates that when the flow reaches this region in the WI and IR cases, it interacts with the canopy, which does not happen in the SF regime, where the flow in between the obstacles remains stuck, i.e. the spherical structure covers the same region as the separated zone between the obstacles. Note as well that the dome feature in the IR case is complemented by a structure on the wake which extends through both lateral sides up to the leeward side of the downstream obstacle. These conclusions were already extracted from the analysis of the time-averaged fields in the streamwise direction at the symmetry plane $z/h = 0$, where the increased separation between the buildings yielded a higher interaction of the flow within this zone.

Wall-normal velocity Although not presenting relevant flow structures in the constant$-y$ planes previously studied, the wall-normal velocity component of mode G exhibits a three-dimensional pattern on the upper windward side of the upstream obstacle, which is shared among all the flow regimes. At this location, the flow experiences a high-velocity region due to the impact of the flow over the edge and the shear layer over the upper part of the obstacle. This region is then followed by another fluctuating part in the wall-normal direction in between the obstacles. Albeit to a lesser extent, the downstream obstacle in the IR case also presents a similar flow structure.

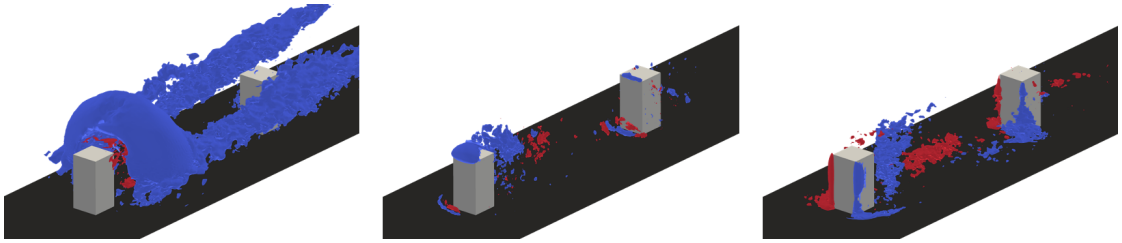
Spanwise velocity Finally, similar to the wall-normal component, the flow encounters a high-spanwise velocity gradient over this zone owing to the effect of the flow over the edges of the first obstacle, which tends to deviate the flow towards the outside regions of the domain. Furthermore, the second obstacle has comparable structures on its windward lateral edges as a result of the flow reattachment that occurs in the IR case. On the other hand, due to the slight interaction of the flow in the spanwise direction within this region, certain fluctuating zones develop in between the obstacles. This characteristic distinguishes this generating mode



(a) Skimming flow



(b) Wake interference



(c) Isolated roughness

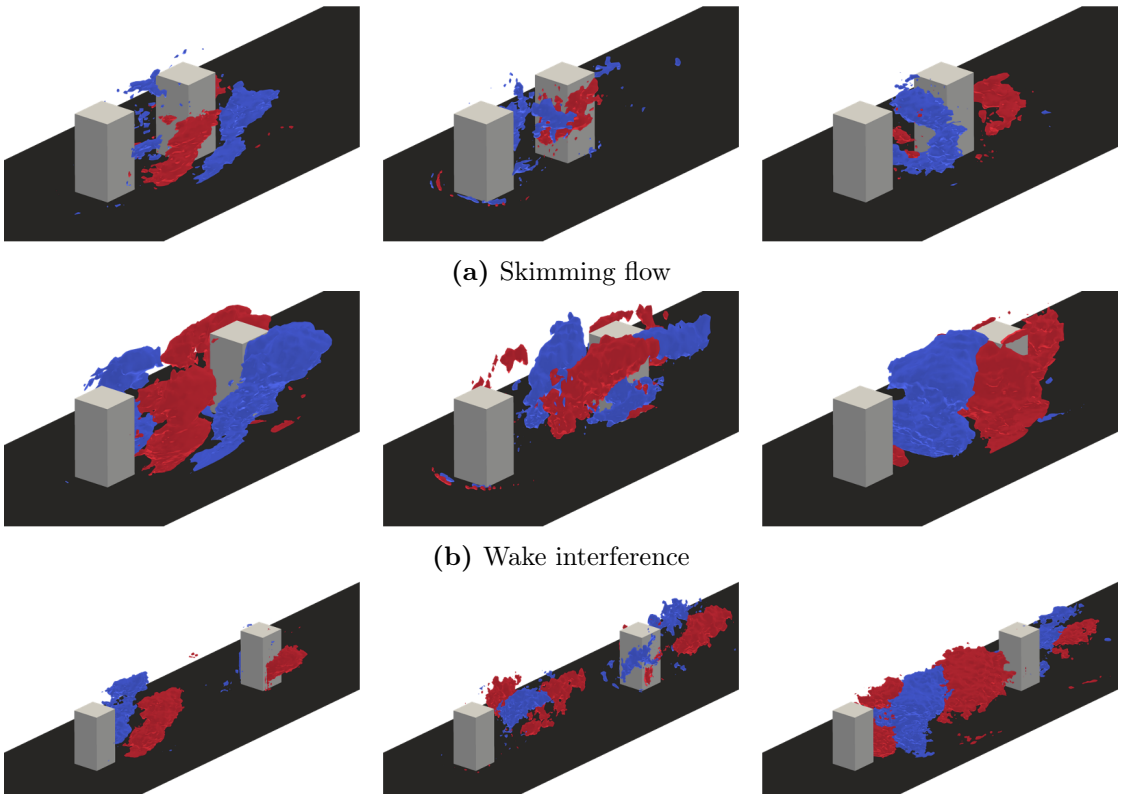
Figure 4.15: Three-dimensional iso-surface of the (left) streamwise, (middle) wall-normal and (right) spanwise velocities of the arch-generator mode shown in Table 4.1. In each panel, the flow moves from left to right. The following velocity iso-values are employed: $a\,U_{\max}$ (blue) and $b\,U_{\min}$ (red), where (left) $a = 0.6$ and $b = 0.7$, (middle) $a = 0.6$ and $b = 0.5$ and (right) $a = b = 0.5$.

from the breaker modes, which exhibit substantial spanwise variations between the obstacles and are thus responsible for the destruction of the vortical structures in this region, as seen in the following figures.

In general, it can be concluded that the flow features that dominate in the arch-generator mode are located close to the first obstacle rather than on the wake, emphasising the idea of being a formation-type mode.

4.3.2.2 Arch-breaker modes

On the other hand, the three-dimensional structures of the arch-breaker mode are illustrated in Figure 4.16. In this case, it is clear that the importance of this mode is reliant on the wake rather than on the first obstacle, as occurred in the latter case. For all of the flow regimes investigated here, the spanwise fluctuating areas are shown to occupy the whole intermediate zone between the obstacles. As a result, increasing the distance between the obstacles leads to a greater number of fluctuation areas in the spanwise direction, see Figure 4.16 (right): in the IR case, up to three alternating structures may be seen, whereas only one can be seen in the SF regime. Conversely, the streamwise component of the present mode does not exhibit the same behaviour; the structures on the leeward side of the upstream obstacle remain un-



(a) Skimming flow

(b) Wake interference

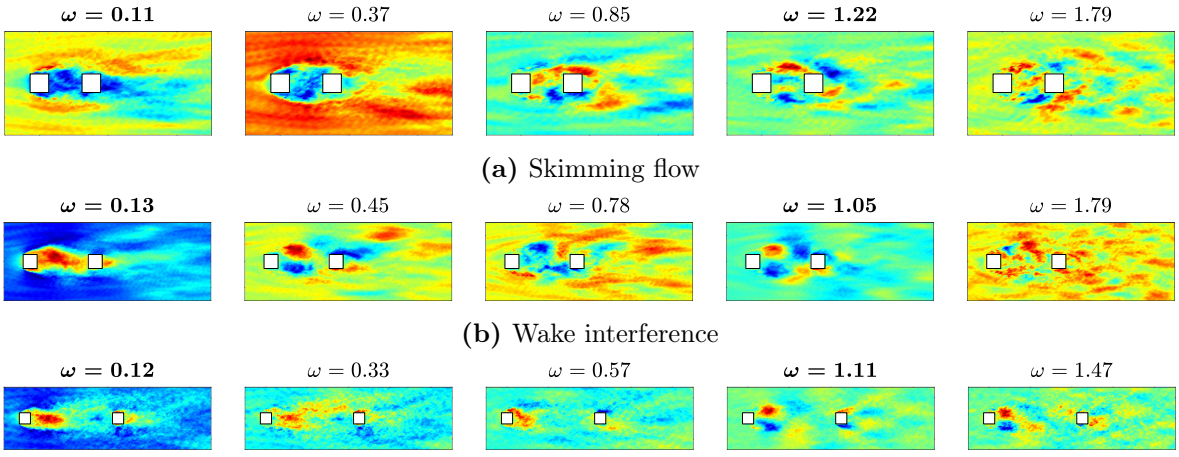
(c) Isolated roughness

Figure 4.16: Three-dimensional iso-surface of the (left) streamwise, (middle) wall-normal and (right) spanwise velocities of the arch-breaker mode shown in Table 4.1. In each panel, the flow moves from left to right. The following velocity iso-values are employed: $a\,U_{\max}$ (blue) and $b\,U_{\min}$ (red), where $a = b = 0.4$ for all panels.

changed for all three flow regimes. However, another oscillating zone arises connected to the downstream obstacle, the position of which is modified among the several regimes. Finally, in terms of velocity in the wall-normal direction, as the separation increases, the flow interacts inside the canopy with considerably more significance, resulting in larger flow characteristics for both WI and IR.

Apart from being quite similar to those of the first two POD modes, these structures are congruent with the results of Monnier et al. [32], with strong streamwise fluctuations on both lateral sides and a high turbulent spanwise region near to the windward side of the downstream obstacle. The arch vortex is known to exist between these regions, and these structures will be related to the process of breaking rather than creation, owing to its location on the wake.

Finally, the three-dimensional structures of the major DMD modes can be compared to those of the various modes identified by the algorithm. Specifically, the following lines will be dedicated to the classification of the modes in arch-generator or -breaker modes based on resemblance to the prior patterns. Only those modes satisfying the criteria $|f_{mi} - f_{mj}| < \epsilon$ have been included, where f_{mi} and f_{mj} represent the frequency value in two different test cases and ϵ is a tolerance set by the user. This allows for selecting the more robust modes identified by HODMD using a varied set of parameters.

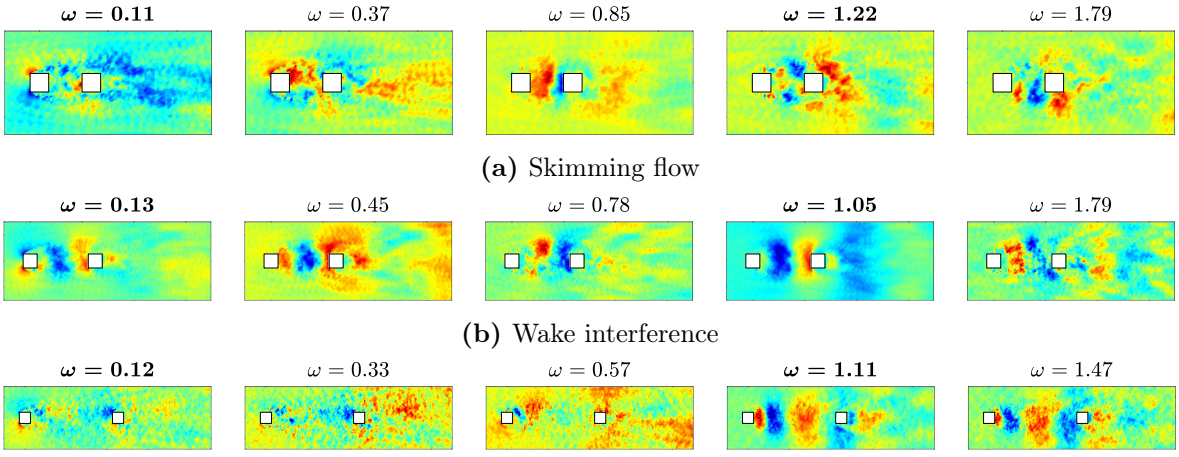


(a) Skimming flow

(b) Wake interference

(c) Isolated roughness

Figure 4.17: DMD modes of the streamwise velocity fields at $y/h = 0.25$ for the different flow regimes.



(a) Skimming flow

(b) Wake interference

(c) Isolated roughness

Figure 4.18: DMD modes of the spanwise velocity fields at $y/h = 0.25$ for the different flow regimes.

Figures 4.17 and 4.18 show a contour representation of the DMD modes presented in Figure 4.1. By general inspection, one could see the structures previously discussed for the main HODMD modes, which are highlighted in bold. From these structures, a limit frequency can be established such that greater frequency values result in flow patterns more focused on the wake and thus being of breaking-type.

For the SF case, the mode $\omega_{\mathrm{SF}} = 0.37$ still presents some flow trapped in between the obstacles and high spanwise fluctuations on the lateral edges of the first obstacle. Even though the flow in this region seems to be modified by the slight interaction with the surrounding flow, this mode can be thought of as a vortex-production mode with a different production mechanism. Further modes ($\omega_{\mathrm{SF}} > 0.85$) are seen to be breaker modes since both the streamwise and spanwise components share the same flow features as the main one ($\omega_{\mathrm{SF}} = 1.22$). Note as well that higher-frequency modes present smaller turbulent scales. This highlights the association of low-frequency modes with large flow scales (dominant patterns) and high-frequency modes with smaller turbulent structures.

A similar conclusion can be extracted from the IR regime, where the threshold value is set for the mode with frequency $\omega_{IR} = 0.57$, which presents some flow structures around the upstream obstacle combined with particular features on the wake. Hence, this mode may be regarded as a transitory mode between the arch-generator and -breaker modes.

Finally, regarding the WI case, apart from the arch-generator mode ($\omega_{WI} = 0.13$), the lowest-frequency mode ($\omega_{WI} = 0.45$) exhibits a flow pattern similar to that of the arch-breaker mode. Therefore, in this situation, the threshold value should be set lower than this frequency, resulting in all modes fulfilling $\omega_{WI} > 0.45$ becoming breaking-type.

4.3.3 Streamline flow patterns

In a similar way to the representation of the main time-averaged vortical structures shown in § 4.1, the dominant flow patterns of the above-described modes can be elucidated by means of streamlines. To that aim, using this visualisation technique, one could clearly relate the previous three-dimensional structures to the arch-generating and -breaking processes.

Figures 4.19 and 4.20 depict the streamline flow patterns of the arch-generator and -breaker modes for the wake-interference regime. The results for the other two regimes are analogous. By general inspection, one can clearly appreciate the two types of modes. On the one hand, in Figure 4.19, the arch-generator mode is shown. Note the apparent resemblance with the arch vortex depicted in Figure 4.4. It can be confirmed, therefore, that this low-frequency mode is responsible for the process of generation of the arch vortex, which, in this case, presents one of its legs displaced. This means that this mode induces a symmetry-breaking process over the vortical structures formed in between the buildings.

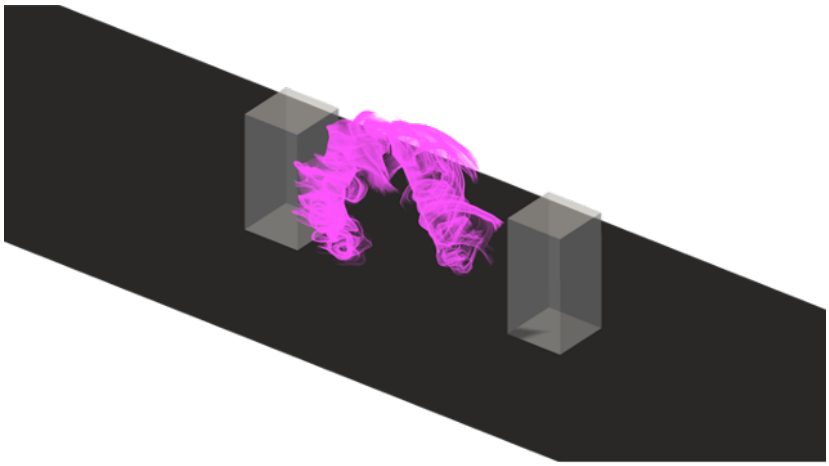


Figure 4.19: Main flow patterns of the arch-generator mode ($\omega = 0.13$) visualised by means of streamlines for the wake-interference regime. The flow moves from left to right. Note the arch-shaped structure on the leeward side of the upstream obstacle.

On the contrary, the flow patterns present in the arch-breaker mode are illustrated in Figure 4.20. When it comes to the arch-breaking process, some helicoidal flow patterns appear in the region in between the obstacles, owing to the increased correlation in the spanwise direction for this type of mode. Indeed, the location of these structures perfectly matches with the gaps in between the velocity fluctuating regions in the spanwise direction (see Figure 4.16). Therefore, since these fluctuating regions define the location of such helicoidal pat-

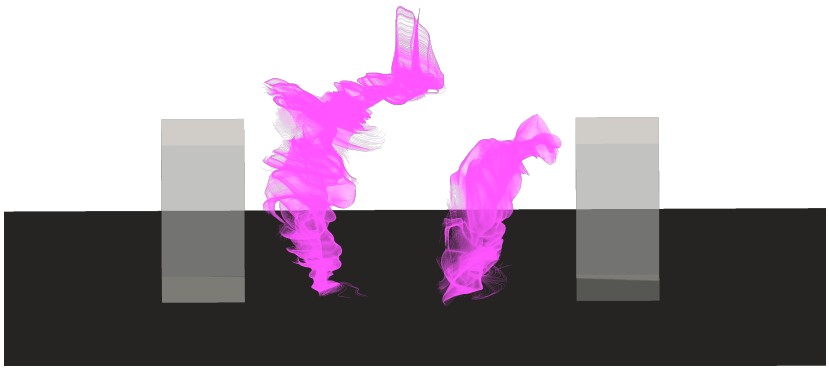


Figure 4.20: Main flow patterns of the arch-breaker mode ($\omega = 1.05$) visualised by means of streamlines for the wake-interference regime. The flow moves from left to right. Note the helicoidal flow structures in between the obstacles.

terns, the number of structures would be modified from case to case: up to three structures are appreciated in between the buildings for the isolated-roughness regime. In such a fashion, the interaction of the flow within this region results in a mixing procedure, leading to the breaking process of the vortical structures. Note that, in this case, there is no resemblance with arc-shaped structures, thus confirming the above-stated ideas, i.e. the arch-breaking process.

To conclude with this section, the three-dimensional structures and flow patterns for the dominant modes identified by HODMD have been characterised for three different flow regimes, representing urban areas with varying separation ratios. Their classification on arch-generator and -breaker modes has been justified using streamline flow patterns.

5

Conclusions

A simplified urban environment model consisting of an array of two buildings with variable spacing ratios was examined using high-fidelity simulation data, in order to provide a full physical description of the fundamental mechanisms controlling the dynamics in different types of urban streets. The aim was to provide an insightful analysis of the physics of the flow within environments that simulate different types of urban areas. The growing expansion of cities boosts the search for physical models capable of reproducing the pollutant and thermal distributions within cities. Here, the three-dimensional flow patterns responsible for the pollutant dispersion have been characterised. Isosurfaces and contour slices were used to demonstrate the complicated flow behaviour of the modes identified by the POD and HODMD algorithms.

The results shown here prove that the flow behaviour can be split into low- and high-frequency phenomena, each with important consequences related to the formation and destruction of vortical structures such as the arch or horseshoe vortices. Interestingly, low- and high-frequency modes are named arch-generator modes since their associated structures are related to the mechanism triggering the formation of the arch vortex formed on the leeward side of both buildings. These structures are particularly noticeable on the windward side of the upstream obstacle for the wall-normal and spanwise velocity components and on the leeward side for the streamwise one, which defines the location and shape of the arch vortex. Furthermore, this location is kept constant among the different flow regimes, which highlights the idea that the process of formation of the arch vortex does not strongly depend on the separation between the obstacles.

HODMD, on the other hand, identifies a high-frequency mode that correlates with the largest-amplitude mode in all cases. Because of the streamline flow patterns in the intermediate section of the obstacles, they are referred to as arch-breaker modes. The large amplitude of these modes emphasises their importance in this sort of urban flow. Indeed, their structures are inextricably linked to the first two highest-in-energy POD modes. Furthermore, as the separation increases, the flow becomes more correlated in the spanwise direction, owing to the larger interaction of the flow with the wake layer inside the canopy, which yields to fluctuating velocity regions occupying the hole section in between the obstacles, thus being associated to the process of destruction of the vortical structures rather than to their formation.

Besides, the arch-breaker modes will be responsible for the apparition of high-TKE regions on both sides of the obstacles and on the windward side of the downstream obstacle. Interestingly, these results are consistent with the wind-tunnel results of Monnier et al. [32], performed on a scaled and slightly more complex urban environment. Therefore, the con-

clusions of the present project can be extrapolated to more realistic urban environments by considering that turbulence levels from street to street are expected to decrease significantly. It is also interesting to note that the results provided by both POD and HODMD show that the wall-normal velocity component does not have a significant influence on the more prominent structures as the streamwise and the spanwise components do.

To finish off, from an environmental point of view, urban areas with highly-separated buildings, namely the isolated-roughness regime, would present much more interaction with clean air sources, thus enabling for a rapid propagation of those pollutants emitted at the street level. However, power plants, commonly located close to urban centres, are also responsible for pollution issues within cities. In those cases, owing to the low interaction of the flow above the urban canopy with the streets, it would be convenient to decrease the separation between buildings, i.e. the skimming-flow regime. In such a fashion, the air at street level would also be in contact with clean sources of air through the arch vortex legs.

Bibliography

[1] European Comission. *Urbanisation worldwide*. URL: https://knowledge4policy.ec.europa.eu/foresight/topic/continuing-urbanisation/urbanisation-worldwide_en (visited on 12/04/2021).

[2] United Nations Climate Change. *The Paris Agreement*. 2015. URL: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (visited on 12/04/2021).

[3] United Nations Climate and Environment. *Cities: a 'cause of and solution to' climate change*. 2019. URL: https://news.un.org/en/story/2019/09/1046662 (visited on 12/04/2021).

[4] Clare Heaviside, Sotiris Vardoulakis, and Xiao-Ming Cai. "Attribution of mortality to the urban heat island during heatwaves in the West Midlands, UK". In: *Environmental Health* 15.1 (Mar. 2016), S27. ISSN: 1476-069X. DOI: 10.1186/s12940-016-0100-9.

[5] Matteo Carpentieri. "Pollutant dispersion in the urban environment". In: *Reviews in Environmental Science and Bio/Technology* 12.1 (Mar. 2013), pp. 5–8. ISSN: 1572-9826. DOI: 10.1007/s11157-012-9305-8.

[6] Ricardo Vinuesa et al. "The role of artificial intelligence in achieving the Sustainable Development Goals". In: *Nature Communications* 11.1 (Jan. 2020), p. 233. ISSN: 2041-1723. DOI: 10.1038/s41467-019-14108-y.

[7] United Nations. Department of Economic and Social Affairs. *The 17 Goals*. URL: https://sdgs.un.org/goals.

[8] R. E. Britter and S. R. Hanna. "Flow and dispersion in urban areas". In: *Annual Review of Fluid Mechanics* 35.1 (2003), pp. 469–496. DOI: 10.1146/annurev.fluid.35.101101.161147.

[9] C. S. B. Grimmond and T. R. Oke. "Turbulent Heat Fluxes in Urban Areas: Observations and a Local-Scale Urban Meteorological Parameterization Scheme (LUMPS)". In: *Journal of Applied Meteorology* 41.7 (1Jul. 2002), pp. 792–810. DOI: 10.1175/1520-0450(2002)041<0792:THFIUA>2.0.CO;2.

[10] Jinhee Jeong and Fazle Hussain. "On the identification of a vortex". In: *Journal of Fluid Mechanics* 285 (1995), pp. 69–94. DOI: 10.1017/S0022112095000462.

[11] Pablo Torres, Soledad Le Clainche, and Ricardo Vinuesa. "On the Experimental, Numerical and Data-Driven Methods to Study Urban Flows". In: *Energies* 14.5 (2021). ISSN: 1996-1073. URL: https://www.mdpi.com/1996-1073/14/5/1310.

[12] T.R. Oke. "Street design and urban canopy layer climate". In: *Energy and Buildings* 11.1 (1988), pp. 103–113. ISSN: 0378-7788. DOI: 10.1016/0378-7788(88)90026-6.

[13] Fisher P; Kruse J; Mullen J; Tufo H; Lottes J; Kerkemeier S. *Nek5000: Open Source Spectral Element CFD Solver*. URL: https://nek5000.mcs.anl.gov (visited on 04/14/2021).

[14] Pablo Torres. "High-Order Spectral Simulations of the Flow in a Simplified Urban Environment". Bachelor's Thesis. Valencia, Spain: Polytechnic University of Valencia, 2020.

[15] D. Zajic et al. "Flow and Turbulence in an Urban Canyon". In: *Journal of Applied Meteorology and Climatology* 50.1 (1Jan. 2011), pp. 203–223. DOI: 10.1175/2010JAMC2525.1.

[16] M. Hussain and B.E. Lee. "A wind tunnel study of the mean pressure forces acting on large groups of low-rise buildings". In: *Journal of Wind Engineering and Industrial Aerodynamics* 6.3 (1980), pp. 207–225. ISSN: 0167-6105. DOI: 10.1016/0167-6105(80)90002-1.

[17] I.Y. Lee and H.M. Park. "Parameterization of the pollutant transport and dispersion in urban street canyons". In: *Atmospheric Environment* 28.14 (1994). Urban Atmosphere, pp. 2343–2349. ISSN: 1352-2310. DOI: 10.1016/1352-2310(94)90488-X.

[18] Jean-François Sini, Sandrine Anquetin, and Patrice G. Mestayer. "Pollutant dispersion and thermal effects in urban street canyons". In: *Atmospheric Environment* 30.15 (1996), pp. 2659–2677. ISSN: 1352-2310. DOI: 10.1016/1352-2310(95)00321-5.

[19] R.E. Britter and J.C.R. Hunt. "Velocity measurements and order of magnitude estimates of the flow between two buildings in a simulated atmospheric boundary layer". In: *Journal of Wind Engineering and Industrial Aerodynamics* 4.2 (1979), pp. 165–182. ISSN: 0167-6105. DOI: 10.1016/0167-6105(79)90044-8.

[20] A.D. Penwarden and A.F.E. Wise. *Wind Environment Around Buildings*. Building Research Establishment report. H.M. Stationery Office, 1975. ISBN: 9780116705334.

[21] N Isyumov and AG Davenport. "The ground level wind environment in built-up areas". In: *Proceedings of the 4th International Conference on Wind Effects on Buildings and Structures, (Heathrow 1975)*. 1975, pp. 403–422.

[22] J. Counihan. "Adiabatic atmospheric boundary layers: A review and analysis of data from the period 1880–1972". In: *Atmospheric Environment (1967)* 9.10 (1975), pp. 871–905. ISSN: 0004-6981. DOI: 10.1016/0004-6981(75)90088-8.

[23] Ted Stathopoulos and Charalambos C Baniotopoulos. *Wind effects on buildings and design of wind-sensitive structures*. Vol. 493. Springer Science & Business Media, 2007.

[24] Bert Blocken, Ted Stathopoulos, and Jan Carmeliet. "CFD simulation of the atmospheric boundary layer: wall function problems". In: *Atmospheric Environment* 41.2 (2007), pp. 238–252. ISSN: 1352-2310. DOI: 10.1016/j.atmosenv.2006.08.019.

[25] Steven L. Brunton, Bernd R. Noack, and Petros Koumoutsakos. "Machine Learning for Fluid Mechanics". In: *Annual Review of Fluid Mechanics* 52.1 (2020), pp. 477–508. DOI: 10.1146/annurev-fluid-010719-060214.

[26] Ricardo Vinuesa et al. "Direct numerical simulation of the flow around a wall-mounted square cylinder under various inflow conditions". In: *Journal of Turbulence* 16.6 (2015), pp. 555–587. DOI: 10.1080/14685248.2014.989232.

[27] J. C. R. Hunt et al. "Kinematical studies of the flows around free or surface-mounted obstacles; applying topology to flow visualization". In: *Journal of Fluid Mechanics* 86.1 (1978), pp. 179–200. DOI: 10.1017/S0022112078001068.

[28] E.R. Meinders. "Experimental study of heat transfer in turbulent flows over wall-mounted cubes". Dissertation Thesis. Netherlands: TU Delft, 1998.

[29] E.R. Meinders and K. Hanjalić. "Experimental study of the convective heat transfer from in-line and staggered configurations of two wall-mounted cubes". In: *International Journal of Heat and Mass Transfer* 45.3 (2002), pp. 465–482. ISSN: 0017-9310. DOI: 10.1016/S0017-9310(01)00180-6.

[30] S. Becker, H. Lienhart, and F. Durst. "Flow around three-dimensional obstacles in boundary layers". In: *Journal of Wind Engineering and Industrial Aerodynamics* 90.4 (2002). Bluff Body Aerodynamics and Applications, pp. 265–279. ISSN: 0167-6105. DOI: 10.1016/S0167-6105(01)00209-4.

[31] R. Martinuzzi. "Experimentelle Untersuchungen der Umströmung Wandgebundener, Rechteckiger, Prismatischer Hindernisses". Ph. D. Thesis. Erlangen, Germany: Universität Erlangen-Nürnberg, 1982.

[32] Bruno Monnier et al. "Turbulent Structure of a Simplified Urban Fluid Flow Studied Through Stereoscopic Particle Image Velocimetry". In: *Boundary-Layer Meteorology* 166.2 (Feb. 2018), pp. 239–268. ISSN: 1573-1472. DOI: 10.1007/s10546-017-0303-9.

[33] J. Sousa. "Turbulent flow around a surface-mounted obstacle using 2D-3C DPIV". In: *Experiments in Fluids* 33.6 (Dec. 2002), pp. 854–862. ISSN: 1432-1114. DOI: 10.1007/s00348-002-0497-5.

[34] Bruno Monnier, Brian Neiswander, and Candace Wark. "Stereoscopic Particle Image Velocimetry Measurements in an Urban-Type Boundary Layer: Insight into Flow Regimes and Incidence Angle Effect". In: *Boundary-Layer Meteorology* 135.2 (May 2010), pp. 243–268. ISSN: 1573-1472. DOI: 10.1007/s10546-010-9470-7.

[35] Stephen B Pope. *Turbulent flows*. Cambridge: Cambridge Univ. Press, 2011.

[36] Chenning Tong and Z. Warhaft. "Passive scalar dispersion and mixing in a turbulent jet". In: *Journal of Fluid Mechanics* 292 (1995), pp. 1–38. DOI: 10.1017/S0022112095001418.

[37] NASA. *Definition of streamlines*. URL: https://www.grc.nasa.gov/WWW/k-12/airplane/stream.html (visited on 03/07/2021).

[38] Kunihiko Taira et al. "Modal Analysis of Fluid Flows: An Overview". In: *AIAA Journal* 55.12 (2017), pp. 4013–4041. DOI: 10.2514/1.J056060.

[39] J.M. Vega and S.L. Clainche. *Higher Order Dynamic Mode Decomposition and Its Applications*. Elsevier Science, 2020. ISBN: 9780128197431.

[40] S.L. Brunton and J.N. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019. ISBN: 9781108386586.

[41] "The Analytical Theory Of Heat". In: *Nature* 18.451 (June 1878), pp. 192–192. ISSN: 1476-4687. DOI: 10.1038/018192a0.

[42] Peter J. Schmid. "Dynamic mode decomposition of numerical and experimental data". In: *Journal of Fluid Mechanics* 656 (2010), pp. 5–28. DOI: 10.1017/S0022112010001217.

[43] Ricardo Vinuesa et al. "Pressure-Gradient Turbulent Boundary Layers Developing Around a Wing Section". In: *Flow, Turbulence and Combustion* 99.3 (Dec. 2017), pp. 613–641. ISSN: 1573-1987. DOI: 10.1007/s10494-017-9840-z.

[44] *Parallel Works*. URL: https://www.parallelworks.com/company/ (visited on 07/06/2021).

[45] S. Dong, G.E. Karniadakis, and C. Chryssostomidis. "A robust and accurate outflow boundary condition for incompressible flow simulations on severely-truncated unbounded domains". In: *Journal of Computational Physics* 261 (2014), pp. 83–105. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2013.12.042.

[46] A. Noorani et al. "Aspect ratio effect on particle transport in turbulent duct flows". In: *Physics of Fluids* 28.11 (2016), p. 115103. DOI: 10.1063/1.4966026.

[47] Ricardo Vinuesa. *Synergetic computational and experimental studies of wall-bounded turbulent flows and their two-dimensionality*. Illinois Institute of Technology, 2013.

[48] R. Vinuesa et al. "Turbulent boundary layers around wing sections up to Rec=1,000,000". In: *International Journal of Heat and Fluid Flow* 72 (2018), pp. 86–99. ISSN: 0142-727X. DOI: 10.1016/j.ijheatfluidflow.2018.04.017.

[49] Ricardo Vinuesa et al. *Turbulence statistics in a spectral element code: a toolbox for High-Fidelity Simulations*. Tech. rep. Argonne National Lab.(ANL), Argonne, IL (United States), 2017.

[50] R.B. Stull. *An Introduction to Boundary Layer Meteorology*. Atmospheric and Oceanographic Sciences Library. Springer Netherlands, 1988. ISBN: 9789027727695.

[51] Kai (深見開) Fukami, Taichi (中村太一) Nakamura, and Koji (深潟康二) Fukagata. "Convolutional neural network based hierarchical autoencoder for nonlinear mode decomposition of fluid field data". In: *Physics of Fluids* 32.9 (2020), p. 095110. DOI: 10.1063/5.0020721.

[52] S. Le Clainche et al. "Coherent structures in the turbulent channel flow of an elastoviscoplastic fluid". In: *Journal of Fluid Mechanics* 888 (2020), A5. DOI: 10.1017/jfm.2020.31.

Part II

Blueprints, solicitation document and budget

6

Plans and blueprints

No plans, drawings, or blueprints are applicable due to the nature of the current project. Therefore, this page has been intentionally left blank.

7

Solicitation document

The purpose of this chapter is to disclose the existing technical and legal rules that apply to the current investigation. The emphasis will be on the work and safety conditions that must be met in a project of this magnitude. The solicitation document adheres to the standards included in the Spanish Law, within the context of computer-oriented projects and the minimal health and safety requirements to be fulfilled in such a workplace.

7.1 Functions of the involved parties

This section aims to provide a basic overview of the major duties and working conditions associated to each of the parties implicated in the project. It should be noted that this part will cover not only legal actions but also advisory recommendations that apply indistinguishably within the scope of the study. The titles of the parties engaged, namely the student, tutor and advisor, will determine the exposition to be followed.

7.1.1 Functions of the student

The engineer or student is the central figure in the success of the project. Among his or her various responsibilities, the student is in charge of implementing the procedures to be utilised and analysing the data obtained. Furthermore, he or she is in charge of acquiring, reviewing, and analysing scientific literature on which the project work might be based. In terms of strategy, he or she must design and discuss a well-founded plan to ensure that the project goals are met. He or she is also in responsibility of drafting the final report and presentation, as well as discussing the paperwork with the University board's assessment committee. To conclude, from an ethical point of view, students must guarantee the quality and reliability of the materials used throughout the project at all stages, as well as adherence to the University board's non-plagiarism policy.

7.1.2 Functions of the supervisor

The student will be expected to rely on the supervision and assistance of an academic tutor who will oversee the project and assist with administrative procedures. The director will be in charge of overseeing the overall development of the project, monitoring the progress of the

student on a regular basis and assisting with serious difficulties at any point. The supervisor will also be responsible for ensuring the quality and originality of the project, as well as supplying the required equipment to achieve the objectives of the study, e.g. computer resources. In addition, the director will be responsible for performing the essential activities within the area of supercomputer administration and validating the quality of the acquired results.

7.1.3 Functions of the advisor

Aside from the guidance of the supervisor, the student may also count on the assistance of an advisor, whose responsibilities vary depending on the situation. Indeed, during the current project, the advisor has oriented the student with computed-related tasks, providing the essential equipment to carry out the required activities as well as the administrative processes required by the University where the study is presented. Note that the number of advisers and their roles are heavily influenced by the project's complexity. As a result, no solid functions can be defined.

7.2 Working environment conditions

The Spanish legal edict, *Real Decreto* 488/1997, collects the minimal dispositions that must be met within the scope of working places with visualisation displays. The purpose of this section is to provide a basic overview of the specific criteria that apply to the current project. To that purpose, the structure that will be followed will be covered by Spanish regulatory institutions.

Working site conditions Due to the nature of the present project, the equipment to be used must comply with certain standards so that its use does not represent a threat to workers. The following points provide the most pertinent information regarding the working environment and equipment conditions.

- The characters on the screen must be well-defined and legible, with proper dimensions. It should be readily orientated and modified to the worker's preferences and its image should be stable and free of sparkles. Furthermore, it must include certain lighting and contrast settings that allow it to be adjusted dependent on the environment.
- To provide a suitable posture, the keyboard must be tiltable and independent of the screen. Its surface should be matte to avoid sparkles and its various keys should be clearly recognisable and readable.
- The working station should be of sufficient size to allow correct placement of the screen and keyboard as well as the worker's comfort. The document support should be stable and adaptable in such a way that head and eye movements are kept to a minimal.
- The working seat should provide the user enough movement, with the height adjustable and the back tiltable.

Lighting and noise The regulations established by the Spanish institutions relating to lighting and noise are exposed in the following lines.

- The lighting conditions of each zone or working site must be tailored to the activity to be carried out in that location. The following lines are dedicated to exposing the minimal illumination standards stated in the Spanish legislative edict, *Real Decreto* 486/1997.
  - The lighting settings must be determined by the dangers to the safety and health of users, as well as the visual requirements of the activity. Working places with high and extremely high visual demands, for instance, require a minimum level of light of 500 and 1000 lux, respectively.
  - Natural lighting should be used whenever feasible, and it should be supplemented by artificial lighting when the first does not offer enough light.
  - On no account must lightning represent a threat to workers when the contrast of the object to visualise and the background is very weak
  - Lighting distribution should be intended to be as uniform as feasible, with binding avoided whenever possible.
- The noise level is controlled by *Real Decreto* 1316/1989, which covers workplace protection.
  - This regulation seeks to decrease noise levels as well as the possible hazards connected with exposure to significantly high noise levels to a minimum level.
  - An annual periodic control should be conducted if the peak noise level exceeds 85 dBA, or every three years if the daily equivalent noise level exceeds 80 dBA.
  - In light of the present project, the noise produced by the equipment installed in each working site should be considered while designing it, and it is especially crucial that it does not disrupt either the attention or the discussion.

Emergency and protection conditions Aside from the above listed criteria, additional emergency conditions must also be satisfied. The entrepreneur (in this case, the University department) must design the corresponding working places in accordance with several security requirements. This includes emergency exits, circulation areas, stairs, handrails, and any other components that might jeopardise user safety or provide assistance in case of an emergency. Furthermore, all installations must comply with all fire and electrical codes.

8

Budget

The purpose of this final chapter is to estimate the cost of the project presented here. The labour of the participants of the project, namely the engineer, supervisor and adviser, will account for the majority of the expense. It should be mentioned, however, that certain expensive tools are used in the development of the project. Indeed, the post-processing activities were performed using a high-performance computer form the UPV, whose cost will be included within the fixed section.

Despite the fact that the project was developed in collaboration with the Swedish University KTH, monetary quantities will be given in Euros. The costs will then be divided into two categories: fixed costs and indirect costs. Table 8.1 shows a breakdown of the expected expenses before and after taxes. It should be noted that the engineer's and thesis director's salaries have been calculated based on a starting engineer with less than one year of experience and an assistant professor. Indirect expenses are considered to account for 25% of total costs. Finally, the entire cost after taxes is calculated using the Spanish VAT, which equals 21% of the whole cost.

Type	**Concept**	**Usage time** (h)	**Cost** (€/h)	**Total cost** (€)
Salaries	Engineer	600	11	6,600
	Thesis director	90	22	1,980
	Advisor	10	20	200
Fixed costs	External computer	–	–	2,880.31
	Expendable goods	–	–	400
Indirect costs	–	–	–	3,015.08
Total cost before taxes (€)				15,075.39
Total cost after taxes (€)				18,241.22

Table 8.1: Breakdown of the main costs before and after taxes of the present project

Finally, the project cost is expected to be 18,241.22 €.