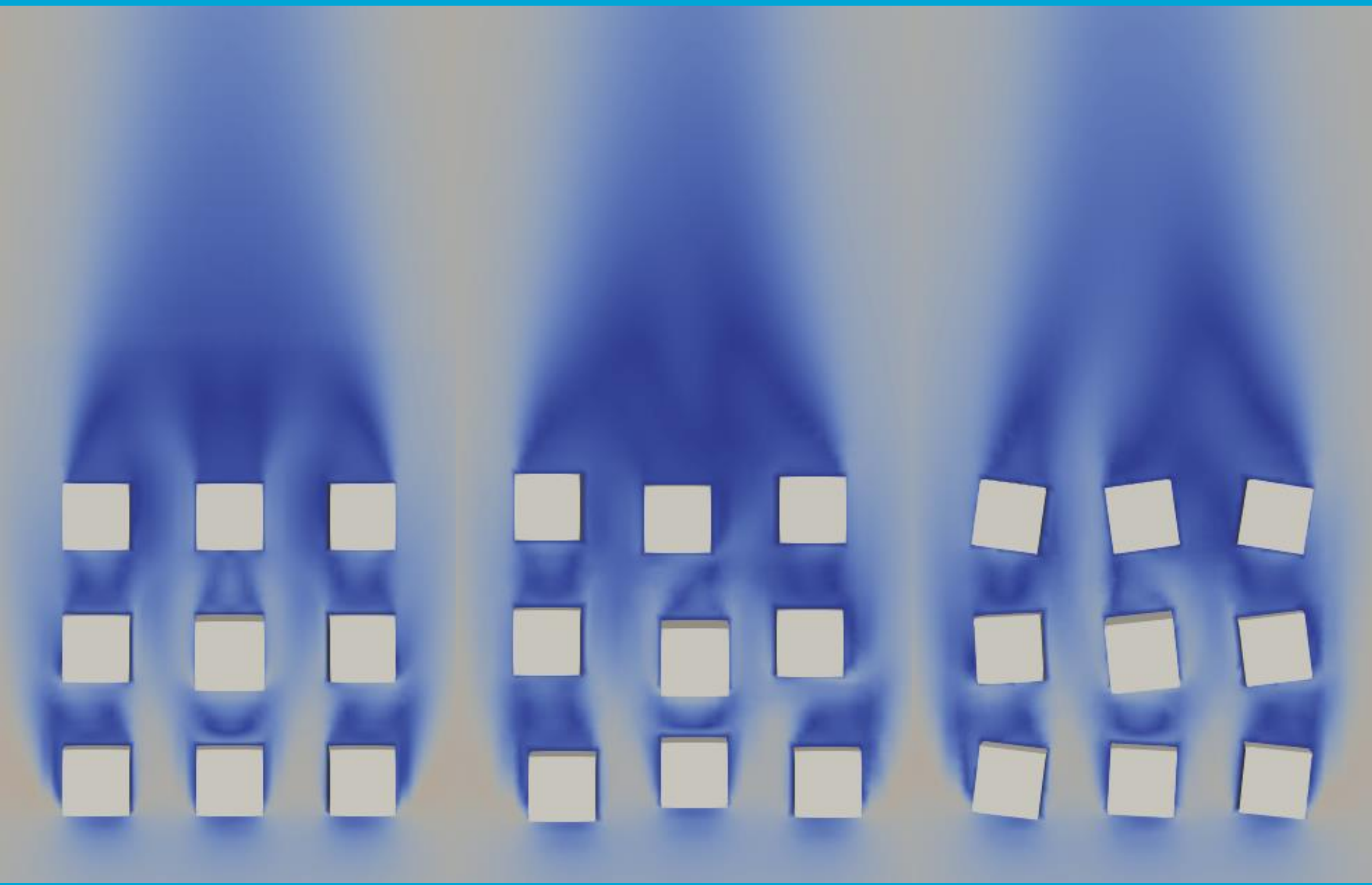
MSc thesis in Geomatics

The effect of building footprint uncertainty on CFD simulations

Christos Chontos

2025



TU Delft

MSc thesis in Geomatics

The effect of building footprint uncertainty on CFD simulations

Christos Chontos

June 2025

A thesis submitted to the Delft University of Technology in partial fulfillment of the requirements for the degree of Master of Science in Geomatics

The work in this thesis was carried out in the:

3D geoinformation group
Delft University of Technology

Supervisors: Dr. Clara García-Sánchez
Akshay Patil
Co-reader: Dr. Hugo Ledoux

Abstract

While reconstructing urban environments, the accuracy of the model is strongly related to the quality of the input data. The initial type of input data for this process could be the footprints of the buildings, which are typically used to reconstruct the buildings, either through manual or automated methods. However, these data could be biased, and when we perform a Computational fluid dynamics (CFD) simulation, they could impact our results. An example of such bias could be either translation or rotation. In our case, simple building footprints were used to define these biases associated with the uncertainty of raw data. Additionally, statistical analysis for representing and quantifying this uncertainty was performed.

The Case C dataset from the Architectural Institute of Japan (AIJ) is an example of a canonical case for our simulations, in which the building footprints could play an important role in the calculation of our results. A CFD simulation on such a canonical case would typically involve the steps of preparing the geometry, generating the mesh, setting boundary and initial conditions, validating the results using experimental data and finally, performing uncertainty analysis. The last step involves various ways of representing the results, like scatter plots, box plots and contour plots. Important aspects of this analysis were the visualization of the flow patterns, the calculation of various quantities of interest such as velocity or turbulent kinetic energy, and the comparison of the simulation results with the experimental data from the AIJ dataset. The effects were examined across multiple wind directions and different footprint uncertainties. This approach could help us to improve the accuracy and reliability of the CFD simulations.

Acknowledgements

First of all, I would like to thank my parents and my brother for supporting me throughout the whole process until I achieved my final educational goal. Their contribution to this is unbelievable and I am sure that they feel for me as proud as I feel for calling them family. I also want to thank the rest of my relatives, uncles, aunts and cousins for caring, supporting and advising me in the most crucial moments. Then I would like to thank the people who, without their assistance, contribution, input, patience and personal time, this thesis wouldn't have been achieved. These people are my supervisors, Dr Clara García-Sánchez and Dr Akshay Patil, whose input, support, ideas and advice were capable of helping me to move forward. In our meetings and hours of discussion and brainstorming, I saw myself improving and changing my way of thinking towards a more structured and organized direction.

Since this thesis marks the end of my student times, I would like to take a moment and thank some of my friends with whom I created some of the strongest memories that will accompany me for the rest of my life. I want to thank my crew from my hometown, Petroupoli, in Greece, with whom we grew up together. Then I would like to thank my fellow surveyors from my university studies in Athens, with whom, until today, we keep a true friendship. Then I would like to thank Kinga Gomułka from Poland, my great support for a big part of my life and especially while I was in Kraków, where I had my Erasmus and also my first international experience from studying abroad. Then I would like to thank my Erasmus friends for their adventurous characters. As the timeline reaches the present, I would like to thank my people from Delft. They made my time here in the Netherlands unforgettable. I got to know many different people, with every single one of them offering me something special with their personality. Specifically, I would like to thank my fellow geomatics classmates, with whom I went through many hours of studying, drinking beers and achieving personal growth through the knowledge that we were gaining. Finally, I would like to thank my partners in crime, Faidon Diakomopoulos, the friend that I have since my day one here in Delft and Pinelopi Kountouri, the girl that we went through big struggles and always winning, the one that motivated me, believing in me and being there for me during my hardest university and thesis times. And finally my party buddies, Ioanna Panagiotidou, the most radical person that I know, my sister and one of my greatest supporters and Eleni Theodoridou, the one who was there to push me when I could not, motivate me, listening to me, discussing my thesis problems, my daily routine, my panic, sad and happy moments.

Having all these people on your side is something unimaginable and I cannot be anything less than deeply grateful.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of the current thesis, the author used, additionally, AI technologies, like Grammarly to proofread, and Chat GPT. After using these tools/services, the authors reviewed and edited the content as needed and would like to assume full responsibility for the content of the publication.

Contents

List of Figures

List of Tables

Acronyms

CFD	Computational fluid dynamics	v
LiDAR	Light Detection and Ranging	1
LoD	level of detail	1
GIS	Geographic Information System	6
OpenFOAM	Open Field Operation And Manipulation	6
OSM	Open Street Map	1
ATKIS	Authorative Topographic-Cartographic Information System	7
DSM	digital surface models	7
BAG	Basisregistratie Adressen en Gebouwen	8
VDI	Verein Deutscher Ingenieure	6
CEDVAL	Compilation of Experimental Data for Validation of Micro-Scale Dispersion Models	8
JU2003	Joint Urban 2003	9
TKE	Turbulent Kinetic Energy	xii
AIJ	Architectural Institute of Japan	v
RANS	Reynolds-averaged Navier–Stokes	8
BR	Blockage Ratio	20
GCI	Grid Convergence Index	32
RMSE	Root Mean Squared Error	26
MAPE	Mean Absolute Percentage Error	34
CSV	Comma-separated value	47
IQR	Interquartile Range	45

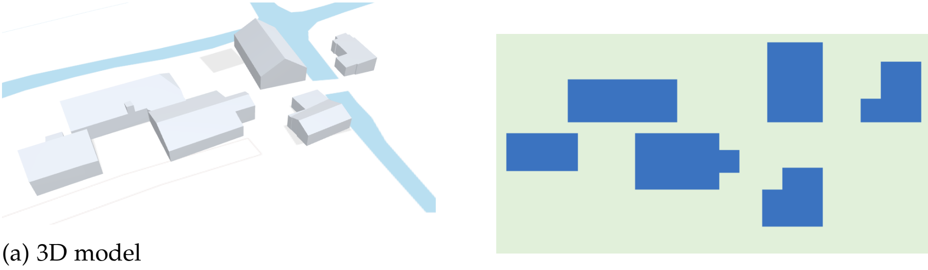
1. Introduction

1.1. Motivation

Computational Fluid Dynamics (CFD) is the scientific field that analyzes fluid flows in various applications. In the field of building design and engineering, CFD simulations play a crucial role in predicting and optimizing airflow patterns. It is a numerical technique in which equations describing the fluid flow are solved on a computer [Franke et al., 2004]. However, when we want to simulate wind flows in urban environments, the accuracy of our simulation and our model is strongly related to the quality of the input data. These data could be the geometric representation of the buildings, or just the footprint data (Figure 1.1). These footprints are typically used to reconstruct the 3D representation of the buildings, through manual or automated methods, representing a fundamental component in CFD simulations. However, in practice, these available data often come from sources such as 3D laser scanners, Light Detection and Ranging (LiDAR), as well as other open sources e.g. Open Street Map (OSM), and may contain biases or uncertainties that can impact the performance of our predictions, and thus need to be addressed [Rottmann et al., 2022; Hecht et al., 2013].

The building footprints define the shape, dimensions, and orientation of the building, influencing the flow patterns and aerodynamic characteristics in the surrounding environment. The quality assessment of the building footprints contains various aspects, based on the primary data. These could be their completeness, position accuracy, shape accuracy, and semantic accuracy [Fan et al., 2014]. Therefore, the process of acquiring building footprints involves various factors of uncertainty and several limitations, and when we perform a CFD simulation, these uncertainties could impact our final results. These uncertainties can arise from inaccuracies in data collection techniques, errors in measurements, errors in reconstruction of the 3D geometry [Pađen et al., 2022], or assumptions made during the modeling process. Typical examples of such bias are translation and rotation (Figure 1.2), which can play a significant role in the final results. Also, the level of detail (LoD) of the building model can vary—higher LoD include roof shapes and facade details, while lower LoD might use only the footprint outline extruded to a certain height. Higher LoDs provide relatively more geometrical realism, but the requirements of complex data and the modeling effort are higher [Pađen et al., 2022].

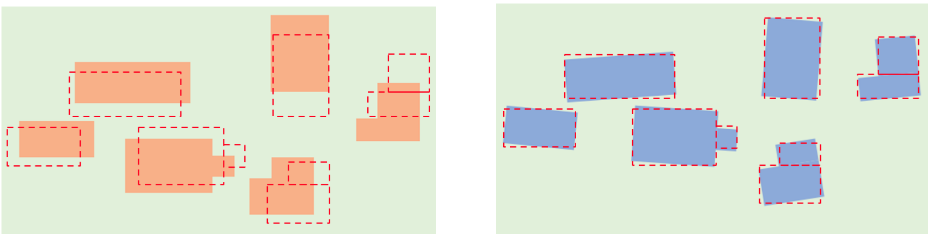
The uncertainties in building footprint representation can have an impact on CFD simulation results, e.g. wind speed distributions, turbulence levels, and pollutant dispersion patterns. Prior studies have noted that even minor geometric discrepancies can lead to notable differences in simulation results. For instance, Ricci et al. [2017] compared CFD simulations against wind-tunnel data for a single building block and then for a whole district, using three levels of geometric detail (simplified, medium, and detailed) and found that the simplifications introduced measurable errors in the predicted wind velocities. Similarly, Hågbo et al. [2021] showed that using an extruded-footprint model (very simplified geometry) in a neighborhood wind simulation led to prominent variations in the wind field compared to



(a) 3D model
Source: https://3dbag.nl/

(b) Building footprints

Figure 1.1.: Actual building footprints.



(a) Translated

(b) Rotated

Figure 1.2.: Example of uncertainties in building footprints.

using a more accurate building model. In their case, a coarse model, based only on 2D footprints, produced different wind speed and turbulence patterns, highlighting how sensitive CFD results are to the fidelity of building representations. In García-Sánchez et al. [2021], the effects of oversimplifying geometries are explored by comparing wind simulations of different levels of detail geometries. Eventually, these different levels of detail lead to diverse wind patterns in built environments. For our case, we will focus on this effect, having as purpose the evaluation of the impact of these uncertainties on the CFD simulation results, determining their spread when the input geometry is uncertain.

Although the possible uncertainties in the data sources used to reconstruct urban areas are acknowledged, there is currently limited understanding of how the footprint uncertainties translate into different CFD results. Consequently, determining their effects and the extent of their impacts can be an important contribution for several applications within the urban wind engineering community. CFD predictions are not deterministic truths but are subject to uncertainty from various sources. This thesis is motivated by this specific need to understand and quantify this uncertainty. Addressing a part of this issue is crucial and by doing so, the aim is to provide guidance regarding data requirements, e.g. when a simple footprint model is sufficient and when a detailed model is warranted. This will help wind engineers, urban planners, CFD modellers, etc., to take more into account the uncertainty in their analyses.

1.2. Research questions

The objective of this thesis is to assess the impact of building footprint uncertainty on CFD simulations. By running multiple simulations with different geometries or wind angles, the research aims to identify how variations in building geometry, position, and orientation influence the flow patterns in our domain. As a result, we aim to understand the magnitude of geometric uncertainty impact and identify the parameters that have the most significant influence on the CFD simulation results. In more detail, the goals and research questions are:

1. Validation of our CFD cases compared to the wind tunnel data.
2. How much do different geometry uncertainties affect the reliability of our CFD simulations compared to the reference cases?
   a) What is the impact of geometrical uncertainties derived from perturbing the buildings' geometry through translation?
   b) What is the impact of geometrical uncertainties derived from perturbing the buildings' geometry through rotation?
   c) How do all these uncertainties affect the model as we modify the incoming wind direction?

1.3. Scope of research

By accomplishing this thesis, the goal is to improve the understanding of uncertainty in CFD simulations. To achieve these objectives, during the research, multiple CFD simulations will be performed. By systematically perturbing building geometries, adjusting footprint orientations, and varying wind directions, this research aims to quantify how such uncertainties propagate and influence the results. The primary goal is to assess the magnitude of uncertainty effects on predicted wind flow patterns, ultimately identifying the most influential geometric parameters. A critical aspect of this study involves validating CFD results against wind tunnel data from AIJ Case C (AIJ-CFD guide website), ensuring an acceptable error margin. Additionally, the research will explore how uncertainty behaves across different ways of altering our geometry—ranging from translation, rotation, and different wind directions—providing insights into how modeling choices can enhance simulation accuracy and reliability. The study's findings will contribute to a better understanding of uncertainty quantification in urban CFD and update the best practices for geometric representation in wind flow modeling.

This research is subject to several limitations. It focuses solely on geometric uncertainties related to building footprints, without modifying the shape itself, and focuses on only a subset (rotation, translation). Also, it does not take into account other uncertainties e.g. inflow conditions. The study is also limited to a single validation case (AIJ Case C), and thus, the generalization of findings to more complex or varied urban configurations may be constrained. This limitation aims to maintain a clear focus on isolating the impact of geometric variability on wind flow predictions. Moreover, only a finite number of translation and rotation scenarios are tested.

1.4. Thesis outline

The structure of the thesis is presented below:

- Chapter 2 presents the related work and theoretical background of the topic. It introduces the different types of uncertainties and guidelines that are followed. Also, other benchmark studies and datasets are presented, concluding in the case C that will be analyzed.
- Chapter 3 focuses on giving us a more general picture of the process that will be followed. First, an introduction is made to our study area and all the subcases that will be studied. Then, the geometry and boundary conditions that will be used for the experiment are analyzed. Afterwards, the construction of our mesh, its resolution and dimensioning are presented. Finally, the terms of uncertainty through translation and rotation are presented, as well as the guidelines that will be followed for the experiment.
- Chapter 4 deals with the application of simulations and the mesh convergence analysis. Through certain criteria, the most optimal mesh was chosen and through comparison with experimental data, the validation process was performed.
- Chapter 5 focuses mainly on the application of the experiments and the methods through which uncertainty is introduced into our model. Thirty simulations are performed. All the metrics, code implementation, statistical approach, and visualization techniques are presented.
- Chapter 6 answers the research questions, regarding the impact of uncertainty and different wind angles in our simulations. Also, the limitations are presented and proposals for future work and improvements are being made.

2. Theoretical background and related work

There are several sources of uncertainty in the results of CFD simulations. To evaluate how well the simulations can predict real-world conditions, it is important to measure the impact of uncertainties in full-scale simulations and validate the results using field experiments. There are two main types of uncertainty: aleatory uncertainty, which comes from natural variations in the system, and epistemic uncertainty, which arises from knowledge gaps or limitations in the model [García-Sánchez et al., 2014]. For example, there might be considerable uncertainty in the prediction of urban flow and dispersion because of variability in the inflow conditions. It is important to make it clear that this is an aleatory uncertainty; it represents physical variability inherent to the system being analyzed [García-Sánchez et al., 2017].

Researchers have been working to identify and quantify how each of these uncertainties influences the results. Numerous studies have compared CFD predictions with experimental data to evaluate overall model accuracy and identify uncertainty sources [Ding et al., 2017; Liu et al., 2017]. As it is presented by Robins et al. [2000], multiple teams simulated gas dispersion around buildings and compared their results to wind tunnel experiments. Substantial differences between the models were reported, stemming from user-related settings, differences in source modeling, boundary conditions, and numerical schemes. Also, it was noted that none of the simulations matched exactly the measurements. Such intercomparison studies underscore the importance of standardized best practices and uncertainty analysis in CFD. In response, the CFD community has developed guidelines e.g. [Franke et al., 2007; Tominaga et al., 2008], to minimize some of these uncertainties by recommending best modeling practices for urban wind simulations. The guidelines provided by Franke et al. [2007] are the set of guidelines that were followed in order to structure the geometry of the current thesis and will be presented in Section 3.6. But, even when following the guidelines, uncertainties cannot be completely eliminated—especially those tied to input data like geometric details or inflow turbulence.

2.1. Geometric uncertainties and CFD

In Blocken and Stathopoulos [2013], it is highlighted: 1) the necessity of increased focus on the assessment of pedestrian-level wind comfort and wind safety instead of only on wind speed conditions (mean velocity and turbulence), 2) continuous sensitivity analysis, validation studies and provision of guideline for CFD simulation of pedestrian-level wind conditions. So, within the broader scope of CFD uncertainty research, the impact of geometric uncertainty has been a focal point in recent years. Ricci et al. [2017] provided a detailed analysis of how geometrical simplifications affect urban wind flow simulations. In their study, a part of a city (a district in Livorno, Italy) was simulated (Figure 2.1) using three different digital models of a representative building block: one with a very simplified footprint and height, one with an intermediate level of detail, and one with a highly detailed shape. By comparing CFD results from these models to wind tunnel measurements, they quantified the deviations caused purely by geometry differences. They found that the simplified model under-predicted wind speed, whereas the detailed model showed much closer agreement with experimental data. This work demonstrated that geometric detail is a significant source of uncertainty: errors in mean wind velocity predictions grew as the building representation became more crude. In a related vein, Hågbo et al. [2021] examined the influence of building geometry input data on CFD simulations for pedestrian wind comfort. Their case study involved a suburban neighborhood in Norway for which four sets of building models were created: varying from a model based on a national Geographic Information System (GIS) building footprint dataset to one derived from high-resolution laser scanning, and even an extreme case where buildings were modeled as simple extruded footprints (Figure 2.2). Their findings highlighted the importance of geometry reliability—the simulation using the basic extruded-footprint model showed notable deviations in wind flow patterns compared to simulations with more accurate building shapes. Also, Hågbo et al. [2021] noted that using the moderately detailed GIS data yielded results not far off from the laser-scanned model, but the lowest-detail model (footprint extrusion) had clear differences. This suggests that there may be a threshold of geometric detail beyond which additional fidelity returns fewer improvements.

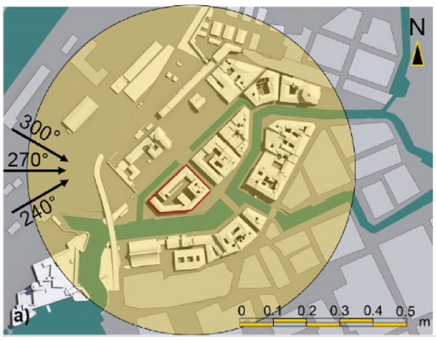


Figure 2.1.: Geometry (Digital model)
Source: Ricci et al. [2017]

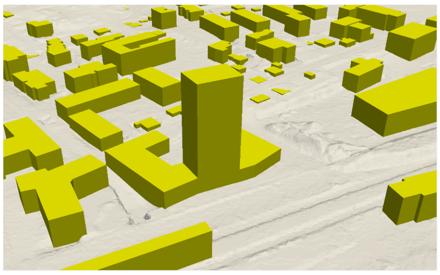


Figure 2.2.: Geometry (Footprint extrusion model)
Source: Hågbo et al. [2021]

Several other works have explored geometric effects on CFD in urban contexts. For example, Franke et al. [2012] conducted a validation of the Open Field Operation And Manipulation (OpenFOAM) CFD code for micro-scale obstacle flows using a standard k–ϵ turbulence model. They followed the German guidelines Verein Deutscher Ingenieure (VDI) for obstacle-resolving models and compared simulation results with laboratory data for simple building configurations. This validation effort confirmed that, with careful setup, CFD can reasonably predict flow around buildings in a controlled environment. However, some discrepancies were observed due to geometrical details that were difficult to resolve on the grid e.g. sharp-edge separation points. The Franke et al. [2012] study is particularly relevant here because it used the OpenFOAM software and the widely-adopted k–ϵ turbulence model, as do many urban flow studies. It highlighted that even when using the same solver and model, differences in geometry modeling (such as how to represent corners or small structures) can lead to differences in outcomes, since geometric representation is an important contributor to uncertainty.

2.2. Building footprints and their importance in urban CFD simulations

Building footprints are the geometric foundation for urban-scale CFD simulations. They are two-dimensional (2D) representations and serve as critical input for generating digital city models. There are several spatial data sources for acquiring building footprint data, and OSM is one of the most commonly used. However, the completeness and positional accuracy of OSM data vary significantly across regions. As mentioned in Section 1.1, Fan et al. [2014] conducted a comprehensive assessment of OSM building footprints and identified inconsistencies in geometry and completeness. In Figure 2.3 we can see a visual comparison between OSM and Authorative Topographic-Cartographic Information System (ATKIS) building footprints. These uncertainties in geometric representation may lead to discrepancies in simulated wind flow patterns.

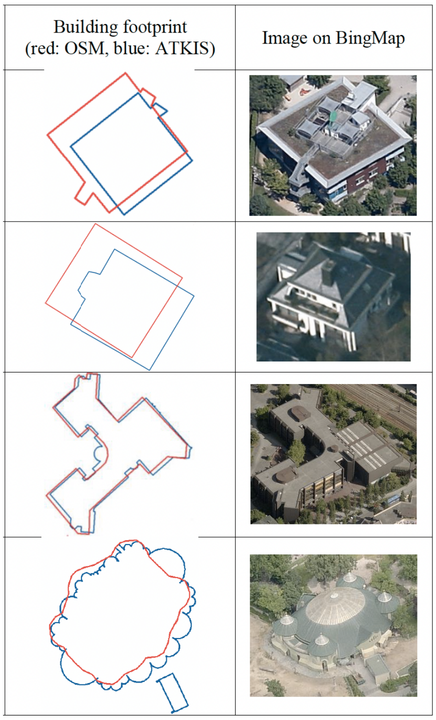


Figure 2.3.: Examples of building footprints in OSM and ATKIS
Source: Fan et al. [2014]

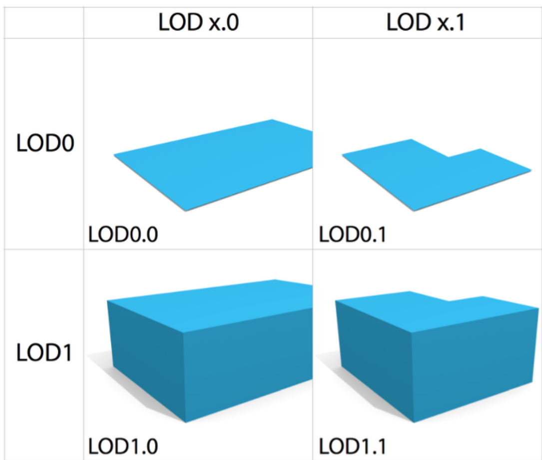


Figure 2.4.: Visual example of the LoDs for a building
Source: Biljecki et al. [2016]

Also, high-resolution LiDAR data is frequently used to extract building footprints by identifying vertical discontinuities and roof structures, often combined with digital surface models (DSM)s for 3D reconstruction. Aerial imagery and photogrammetric techniques also provide valuable inputs, especially when combined with machine learning methods to detect building edges. In countries like the Netherlands, authoritative cadastral datasets, like the Basisregistratie Adressen en Gebouwen (BAG), provide highly accurate footprint geometries.

The concept of LoDs is significant for understanding how building geometry is represented in 3D city models. LoDs vary from simple 2D footprints (LoD0.x) to highly detailed models (LoD4.x) and were originally developed to standardize 3D representations. Biljecki et al. [2015] discussed various applications of 3D city models, emphasizing how the different LoDs are suitable for different tasks, like energy modeling, urban planning and CFD analysis. While LoD1 models (extruded footprints with flat roofs) are commonly used in CFD due to their simplicity and availability, critical features like roof geometry might be missed, potentially affecting the wind flow patterns. To address several limitations, Biljecki et al. [2016] proposed an improved LoD specification, consisting of 16 levels and, except for focusing only on the geometric aspect of the models, the classification is also compatible with different levels of semantic details. As we can see in Figure 2.4, we chose four of these levels that correspond better to the needs of the current thesis. The geometry that is used is simple; therefore LoD0.0 and LoD0.1 can be used to extrude the buildings, but depending on the LoD, it is possible to have an entirely different outcome. The uncertainty in building footprints is a baseline and remains a key variable in simulation accuracy.

2.3. Benchmark studies and databases

Numerous benchmark studies and experimental datasets have been developed to support the assessment and comparison of CFD simulations in urban environments. In this section, an overview of two benchmark studies and datasets that are widely recognized in the literature is provided, highlighting their characteristics, applications, and relevance to model validation and uncertainty quantification.

2.3.1. CEDVAL database

The Compilation of Experimental Data for Validation of Micro-Scale Dispersion Models (CEDVAL) database is an essential resource for validating CFD models in urban wind flow and dispersion studies (Uni-hamburg.de). Developed by the Environmental Wind Tunnel Laboratory at the University of Hamburg, CEDVAL contains a collection of high-quality wind tunnel measurements designed to benchmark numerical simulations. These datasets include velocity profiles, boundary conditions, turbulence parameters, and pollutant dispersion patterns. Researchers have used CEDVAL cases to test the accuracy of their simulations. For instance, one common benchmark is flow and dispersion around a single cubic building (Figure 2.5), where many studies have compared CFD results (using Reynolds-averaged Navier–Stokes (RANS) k–ϵ models) against the measured velocity profiles. Longo et al. [2017] is an example that employed CEDVAL datasets to evaluate the accuracy of advanced turbulence models and boundary conditions in simulating wind flow around different building compositions. Their study demonstrated that while traditional k–ϵ models captured mean velocity fields reasonably well, they struggled with accurately representing TKE distributions, particularly in wake regions. CEDVAL database's role is underscored as a fundamental reference for CFD validation, helping to ensure that numerical models can reliably replicate real-world scenarios.

2.3.2. Joint Urban 2003 field study

The Joint Urban 2003 (JU2003) field study was a large-scale urban atmospheric dispersion experiment conducted in Oklahoma City to investigate wind flow and pollutant transport in a realistic urban environment. The study included tracer gas releases and extensive meteorological measurements, providing valuable data for validating CFD models [Lee et al., 2004]. Scientists simulated these experiments, and one finding from those efforts was that accurately modeling the buildings was crucial. Flaherty et al. [2007] used JU2003 data to evaluate CFD predictions of tracer dispersion. Their findings highlighted that while CFD captured general dispersion patterns, uncertainties were rising from variations in wind direction and source positioning, leading to discrepancies of up to 50% in modeled versus observed concentrations. The study also demonstrated the strong influence of tall buildings (Figure 2.6) on flow separation and mixing, emphasizing the need for accurate urban geometry representation in CFD. Similarly, García-Sánchez et al. [2014] quantified the uncertainty associated with inflow boundary conditions in CFD simulations of Oklahoma City, using JU2003 field measurements as a benchmark. Their research demonstrated that variations in wind speed, wind direction, and surface roughness significantly affected velocity predictions at multiple measurement stations, reinforcing the importance of accurate inflow parameterization in CFD. These findings reinforce the importance of uncertainty quantification in urban CFD.

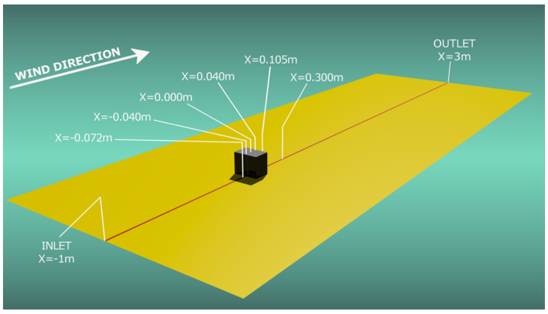


Figure 2.5.: CEDVAL case A1-1
Source: Longo et al. [2017]

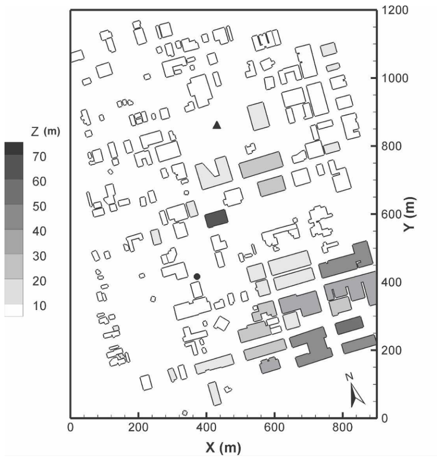


Figure 2.6.: Computational domain (JU2003)
Source: Flaherty et al. [2007]

2.3.3. Guidebook for CFD Predictions of Urban Wind Environment—Architectural Institute of Japan

In Japan, working groups of the AIJ conducted extensive cross-comparisons between CFD simulation results and high-quality wind-tunnel measurements to support the development of guidelines for practical CFD applications. Part of these efforts were reported by Yoshie et al. [2007] and this group intended to propose the guidelines based on the results of their own benchmark tests. In 2008, Tominaga et al. published the "AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings". The feature of these guidelines is that they are based on cross-comparison between CFD predictions, wind tunnel test results and field measurements for seven test cases (Figure 2.7) used to investigate the influence of different computational conditions for various flow fields [Tominaga et al., 2008].

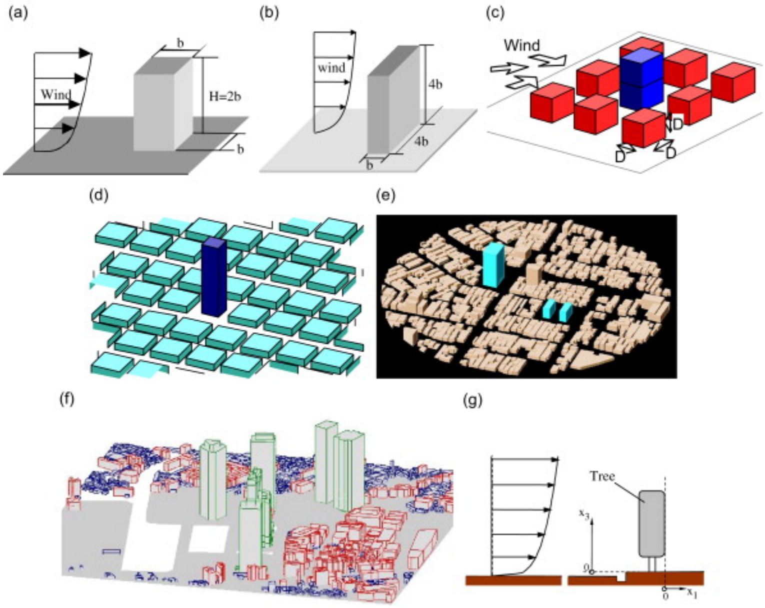


Figure 2.7.: AIJ cases
Source: Tominaga et al. [2008]

According to Tominaga et al. [2005], a comparison between CFD results and wind-tunnel measurements had already been conducted some years earlier, for two different building cases. The first was located at Niigata (Japan), which consisted of low-rise houses, and the second was the Shinjuku sub-central area (Japan), which consisted of high-rise buildings. For the first case overall, CFD matched the wind tunnel results well when assessing the wind environment. For the second case, the high-rise buildings presented a good match, while for some low-rise houses, the accuracy of CFD results depended a lot on grid resolution. Differences appeared when the grid was not fine enough to capture details properly. These kinds of differences we will try to capture in the current thesis.

Another similar comparative study was conducted by Tominaga and Blocken [2015], where velocity was measured in a cross-ventilated flow. The geometry that was used imposes significant similarities with the one that will be used in the current thesis. However, Tominaga et al. [2008] presents a set of guidelines that summarize important points in using the CFD technique for appropriate prediction of pedestrian wind environment. The reason was that the influence of the computational conditions (grid discretization, domain sizes, boundary conditions, etc.) on the prediction accuracy had not been systematically investigated.

2.4. Useful outcomes from the literature review and introduction to our case study

In summary, the related work in this domain establishes that:

- Uncertainty in CFD simulations of urban flow is inevitable and multifaceted.
- Geometric uncertainty is a particularly influential factor for urban wind and dispersion predictions.
- Numerous validation efforts (wind-tunnel and field studies) have helped quantify these uncertainties, showing where and by how much CFD results might deviate.

The case that was chosen to work with is the Case C-Simple building blocks, provided by the AIJ and is presented below (Figure 2.8). It is an example of a canonical case for simulations and is a hypothetical case study that has been created to represent a typical realistic urban case scenario. There are several validation studies regarding the Case C from AIJ (SimScale; Sim-flow; Wang and Ng). In Zheng et al. [2021], the appropriate surrounding building regions are examined in order to achieve accurate prediction of wind flow characteristics around a target building and different building heights, densities, and layouts are considered. A similar building composition exists in our case study, too. The report related to this case can be found in Nonomura et al. [2003].

Although multiple benchmark cases are available, this thesis focuses exclusively on the Case C, due to its well-documented geometry, high-quality wind tunnel measurements, and its widespread acceptance. The insights from previous studies form the foundation for this thesis. The purpose is to contribute to the ongoing effort of uncertainty quantification in CFD, ultimately guiding better modeling practices and more reliable predictions.

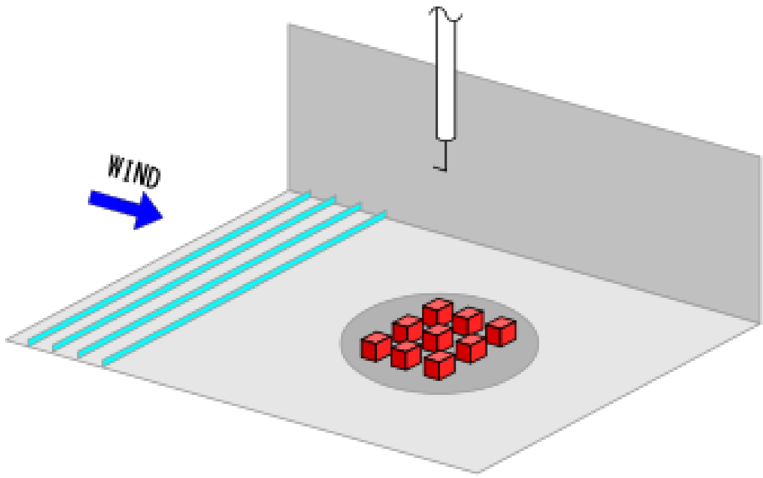
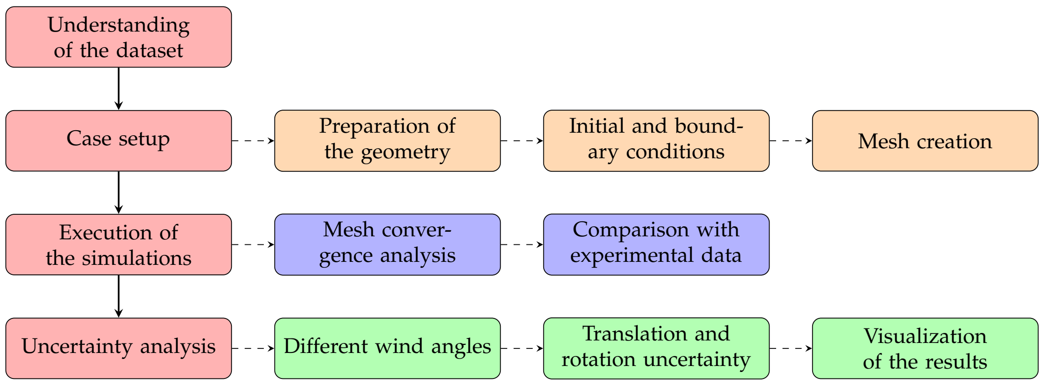


Figure 2.8.: Case C
Source: AIJ-CFD guide website

3. Methodology

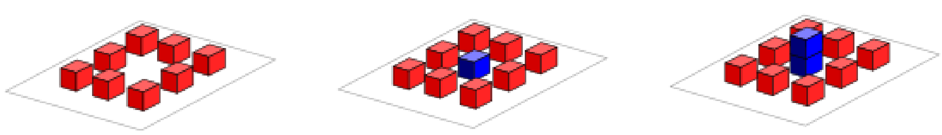
3.1. Research workflow

To set up the case, run the CFD simulations in OpenFOAM and perform our analysis, we facilitated a workflow that consists of the steps below. A general overview of the whole procedure is:



3.2. Dataset

Our dataset consists of nine building blocks in a cubic shape. Each cube has side lengths of 0.2 meters, and the distance between every cube is 0.2 meters. There are three different cases, whose geometries represent simple blocks of buildings, and they are presented below (Figure 3.1). For the three cases, the only building whose height is changing is the middle one. Specifically, for the case 0h, there is no building at its centre, for the case 1h, there is a low-rise building at this location and for the case 2h, there is a high-rise building constructed there.



(a) Case 0h (b) Case 1h (c) Case 2h

Figure 3.1.: Geometry
Source: AIJ-CFD guide website

Besides the given dimensions of our geometry, the inflow angle is also provided and takes three different values: 0°, 22.5° and 45° degrees (Figure 3.3). Below we see a top view of our domain with its dimensions (Figure 3.2). The red points are 120 probes, where the wind velocity was measured experimentally for each of the 9 cases in total (0°, 22.5° and 45° wind angle for 3 different geometries—0h, 1h, 2h). Their height is 0.02m from the ground. The data were acquired from www.aij.or.jp, where the `Case C` is stored.

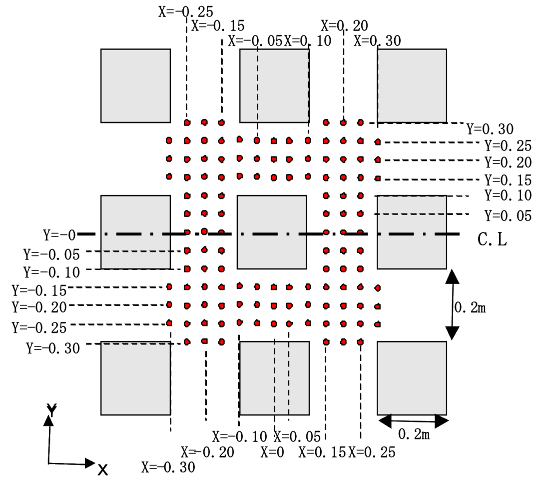


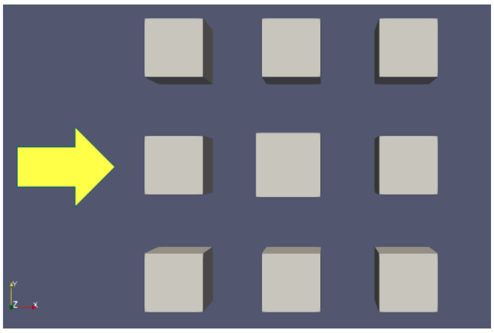
Figure 3.2.: Top view and probes
Source: AIJ-CFD guide website

Furthermore, the initial data provided to us, stem from a wind-tunnel experiment and are the values of inflow velocity $U(m/s)$ that correspond to different heights $Z(m)$, alongside with the standard deviation of the velocity (σ) (Figure 3.4). In the following graph, we plot the points for every U and Z value, and then we drew a curved line that approximates the points. From this graph, it becomes clear that we are dealing with a logarithmic wind velocity profile (Figure 3.5).

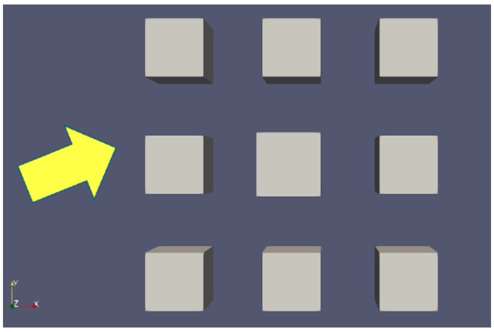
3.3. Preparation of geometry

To begin with, we had to obtain the geometry of our buildings. In our case, the geometry was not given in any presentable format and was just described, so we created an OBJ file named `cubes.obj`. For its creation, we used a simple text editor. According to the given dimension and location of the cubes, we calculated the coordinates of every vertex. Then, we formed the faces. Our geometry is triangulated, and the total number of vertices is 72 (8 for every cube) and triangulated faces: 108 (12 for every cube) (Figure 3.6). It is a simple geometry, so we could structure our initial files manually. The enumeration follows the order in which the entries of each object have been made in the OBJ file.

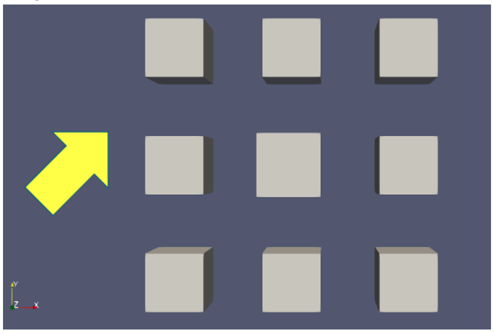
For the cases with different wind angles, we decided to rotate our geometry. After the creation of our 3 initial OBJ files, the command that was used was the '`surfaceTransformPoints -yawPitchRoll "(-45 0 0)" cubes2h0.obj cubes2h45.obj`', for the 45° degrees rotation



(a) Wind angle 0°



(b) Wind angle 22.5°



(c) Wind angle 45°

Figure 3.3.: Different wind angles (top view)

Z(mm)	Z(m)	U (m/s)	σ_u (m/s)
10	0,01	2,372	0,56
20	0,02	2,434	0,564
30	0,03	2,548	0,563
50	0,05	2,912	0,695
70	0,07	3,042	0,71
100	0,1	3,392	0,743
150	0,15	3,483	0,78
200	0,2	3,654	0,801
250	0,25	3,82	0,812
300	0,3	4,019	0,786
350	0,35	4,205	0,83
400	0,4	4,343	0,809
600	0,6	4,985	0,796
800	0,8	5,713	0,587
1000	1	6,12	0,283
1200	1,2	6,201	0,148

Figure 3.4.: Inflow values
Source: AIJ-CFD guide website

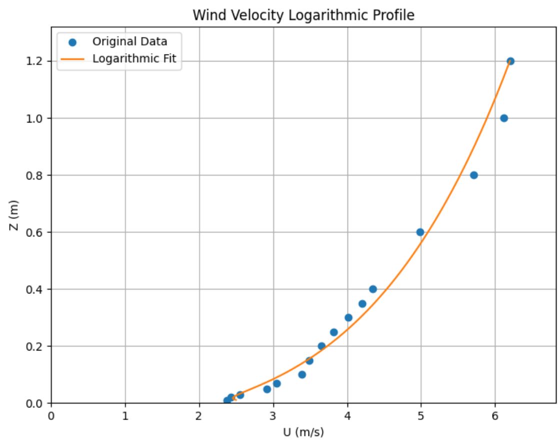
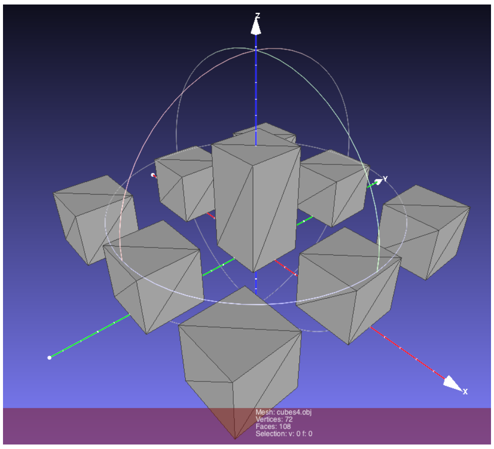
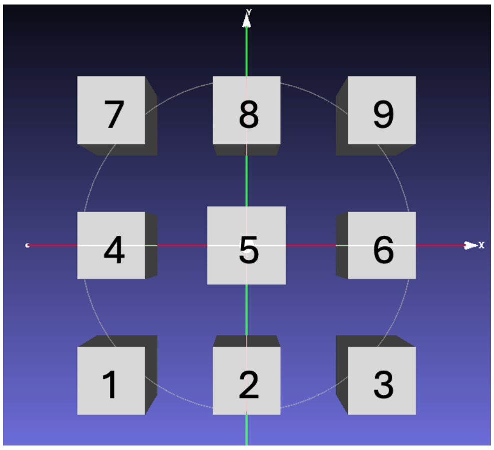


Figure 3.5.: Wind velocity profile

and 22.5° degrees respectively. The command contains −45°, because we needed to perform a clockwise rotation. Finally, we created 9 OBJ files for our different geometries and wind angles, respectively. Alongside the building blocks, with the same command, we also rotated the 120 probes. The clockwise rotated cases are presented below (Figure 3.7).

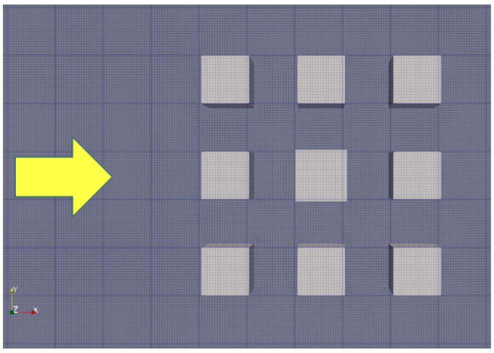


(a) Case2H.obj

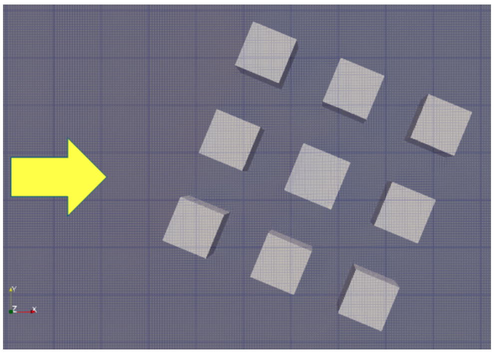


(b) Case C-enumeration

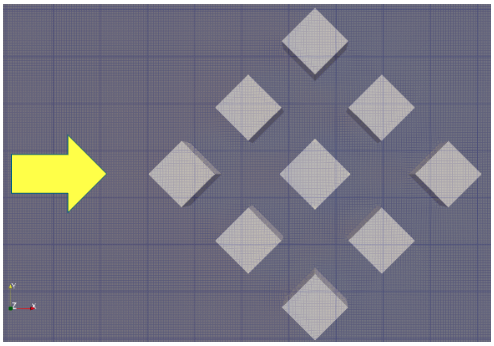
Figure 3.6.: Case C (MeshLab view)



(a) Rotation 0°



(b) Rotation 22.5°



(c) Rotation 45°

Figure 3.7.: Rotated geometry (top view)

3.4. Naming convention

Since multiple simulations will be performed in later stages of the thesis, it was deemed necessary to refer to them to avoid any confusion. To facilitate the procedures and the implementation of all simulations, it was necessary to codify the multiple cases that we will be called upon to manage in the future. For this reason, a codification was created, which presents in a concise but effective manner and illustrates all the future cases on which we will work. The encoding that we used refers both to OBJ files and CFD cases, and is presented below (Figure 3.8) 1-5 with some progressively increasing uncertainty

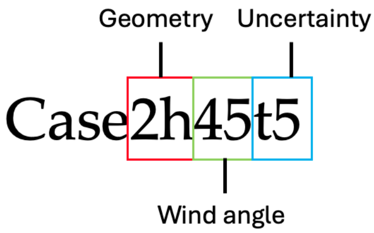


Figure 3.8.: File and case encoding

- Geometry: '0h', '1h', or '2h', depending on the height of the central building block.
- Wind angle: '0', '22', or '45' depending on the wind angle 0°, 22.5° and 45° (units: degrees).
- Uncertainty: 't' for translation and 'r' for rotation. Five progressive translations (t_1, t_2, t_3, t_4, t_5) and rotations (r_1, r_2, r_3, r_4, r_5) will be performed in Chapter 5.

3.5. Initial and boundary conditions

3.5.1. Initial conditions

The next step was to specify our boundary and initial conditions for our simulation. In the RANS family of turbulence models, it is very common to use the k-ϵ turbulence model with a logarithmic velocity profile. In the `case/0` directory, all the physical properties of the wind are located. Specifically, the directory `0/ABLConditions` of our case is the folder responsible for the logarithmic law of the wall (inlet/ABL). The values of the initial conditions are stored here. For their calculation and initialization of the wind flow, the maximum height (z_{max}) of the velocity profile and the corresponding inflow velocity (U_{max}) were chosen as reference. In our case, this was $z_{\text{ref}} = 1,2\,\text{m}$ and the velocity at this height (Figure 3.4), $U(z_{\text{ref}}) = 6,201\,\text{m/s}$, with positive direction, along the x-axis. The reason behind this selection is that we want to capture free-stream conditions. Our tallest object is $H_{\text{max}} = 0.4\,\text{m}$ and above it, the wind flow is less affected by roughness elements and at the same time, we avoid the generation of turbulence and flow separation, which may take place at lower heights.

So, based on the following steady-state, incompressible, RANS equations that were presented by Richards and Hoxey in their pioneering paper [Richards and Hoxey, 1993], we calculated the initial values for the turbulence parameters:

$$U(z) = \frac{u_*}{\kappa} \ln\left(\frac{z + z_0}{z_0}\right) \tag{3.1}$$

$$k(z) = \frac{u_*^2}{\sqrt{C_\mu}} \tag{3.2}$$

$$\epsilon(z) = \frac{u_*^3}{\kappa(z + z_0)} \tag{3.3}$$

where, the von Kármán constant κ = 0.41 and constant C_μ = 0.09. Also, the height of our geometry $z_{\text{ground}} = 0\,\text{m}$, since we don't have any terrain. Finally, the direction of the flow (*flowDir*) was set only along the x-axis (1 0 0), so the wind angle is 0°. It is important to mention here that in later stage of the thesis, for the cases of 22.5° and 45° degrees, these initial values will remain the same, since we chose to rotate our geometry and not to change the direction of the wind.

Another important parameter that we had to consider is the aerodynamic roughness length (z_0), which expresses the roughness of the terrain. Depending on its value, we can obtain information regarding the characteristics of the ground. Similar validation studies of the AIJ case C like the one provided by SimScale, considers as $z_0 = 0.00045\,\text{m}$, which according to the Wieringa [1992] roughness classification, this z_0 value indicates a smooth terrain, like the wind tunnel floor of our case. So for the calculation of u_*, the Equation 3.1 could be rewritten as:

$$from\ (3.1) \rightarrow u_* = \frac{U(z)\kappa}{\ln\left(\frac{z+z_0}{z_0}\right)} \rightarrow u_* = 0.322\,\text{m/s} \tag{3.4}$$

Then, the calculations that followed were the TKE $k = 0.346\ m^2/s^2$ according to Equation 3.2 and the dissipation rate, turbulent epsilon $\epsilon = 0.068\ m^2/s^3$ according to Equation 3.3.

3.5.2. Boundary conditions

The boundary conditions were set in order to describe how the flow behaves at the boundaries of our mesh. Our computational domain consists of boundaries and has a specific volume. From some of these boundaries, the wind flows by entering or leaving them. In general, only the bottom of the domain corresponds to a physical boundary in reality, and the top and side boundaries are artificial [Blocken, 2015]. For that reason, the `atmBoundaryLayerInlet` conditions that were used had as purpose of ensuring a realistic wind profile for our urban case at the inlet. For the buildings, the `kqRWallFunction`, epsilonWallFunction, nutkWallFunction and for the ground, the kqRWallFunction, epsilonz0WallFunction, nutkAtmRoughWallFunction were used, as they are considered rough, solid surfaces. For the sides and top surfaces, the symmetry condition was used. Finally, the inletOutlet condition was used at the outflow, where the wind exits our domain freely. Below (Table 3.1) are presented the boundary conditions that were used.

Patch name	$U(m/s)$	$p(m^2/s^2)$	$k(m^2/s^2)$	$\epsilon((m^2/s^3)$	$nut(m^2/s)$
Inlet	atmBoundaryLayerInletVelocity	zeroGradient	atmBoundaryLayerInletK	atmBoundaryLayerInletEpsilon	calculated
Outlet	inletOutlet	uniformFixedValue	inletOutlet	inletOutlet	calculated
Ground	uniformFixedValue	zeroGradient	kqRWallFunction	epsilonz0WallFunction	nutkAtmRoughWallFunction
Buildings	uniformFixedValue	zeroGradient	kqRWallFunction	epsilonWallFunction	nutkWallFunction
Sides and Top	symmetry	symmetry	symmetry	symmetry	symmetry

Table 3.1.: Boundary conditions

3.5.3. Scheme selection

The scheme selection impacts the accuracy and stability of our CFD simulation and it refers to the numerical method that is used to approximate the differential equations governing the fluid flow (in our case, RANS equations).

The gradient scheme that was used is CellLimited Gauss Linear. For the divergence schemes, the bounded Gauss linearUpwind limited option and bounded Gauss limitedLinear are used, which are also second-order schemes.

The reason behind these selections is that with first-order discretization schemes, numerical diffusion is caused. According to the best practice guidelines Blocken [2015], it is important to use high-quality grids and higher-order discretization schemes, since they allow us to have more accurate results, without having to compromise significantly with the convergence behavior.

3.6. Mesh creation

3.6.1. Creation of the computational domain

The computational domain was created based on the guidelines that were introduced by Franke et al. [2007] and plays an important role in the simulations since it discretizes the geometry into smaller computational cells and also encloses the open flow inside a 'box'. In order to define the dimensions of this background mesh, we took into account the building domain. At first, we found the height of the tallest building, which in our case is H_{max}=0.4m for the 2h case, and 0.2m for the 1h and 0h cases. As mentioned in Blocken [2015], according to Type-1 guidelines by Franke et al. [2004], the ideal domain size depends on H_{max}. For the x-axis, it should be $5H_{\text{max}}$ to the -x direction and $15H_{\text{max}}$ to the +x direction. For the y-axis, it should be on both sides $5H_{\text{max}}$ and for the z-axis, it should be $6H_{\text{max}}$. These distances towards every direction are very specific (Figure 3.9), since the boundaries should be located far enough from the building domain to allow the wind flow to develop freely. The outflow boundary is set at $15H_{\text{max}}$ in order to allow the wake flow behind the buildings to fully develop [Blocken, 2015].

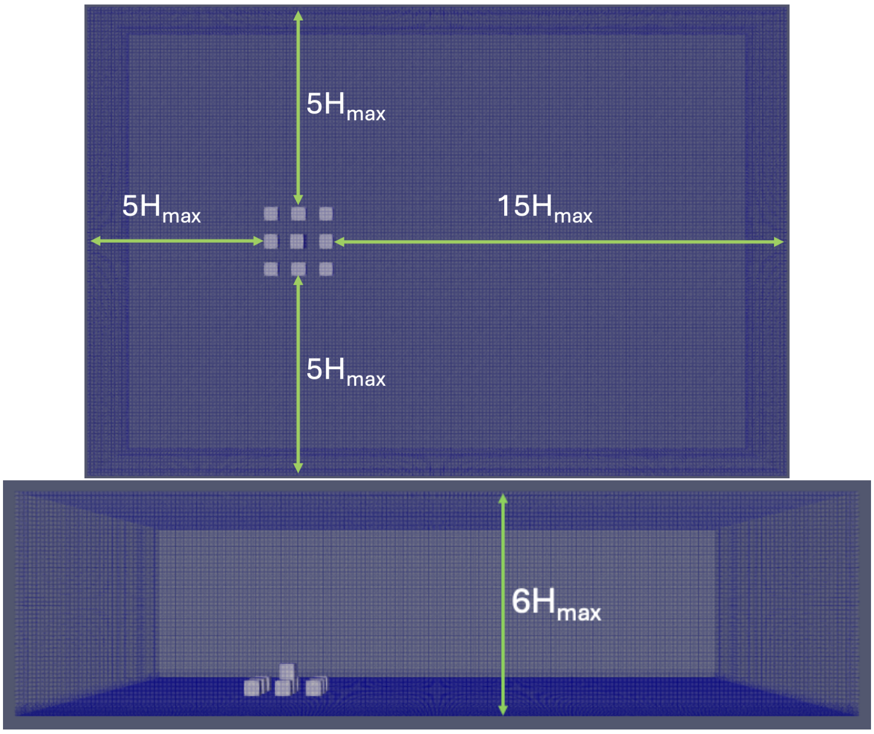


Figure 3.9.: Size of Computational Domain

Another factor that had to be taken into consideration, before we finalize our mesh dimensions, is the Blockage Ratio (BR). According to the type-2 guidelines by Blocken [2015], taking into account the side view of our domain, where the wind enters (inflow), the ratio of the area covered by the building domain, to the total area of the computational domain, should be less than 3%: $BR = \frac{A_{\text{building}}}{A_{\text{domain}}} < 3\%$. Also at Blocken [2015], the type-3 guidelines are introduced, that is a combination of type-1 and 2. Specifically, what should be considered is the Directional BR, where separately the length and height of the buildings should be less than 17% of the length and height of our entire domain: $BR_{\text{L}} = \frac{L_{\text{building}}}{L_{\text{domain}}} = \frac{1m}{6m} < 17\%$ and $BR_{\text{H}} = \frac{H_{\text{building}}}{H_{\text{domain}}} = \frac{0.4m}{2.4m} < 17\%$. This led us to the conclusion that for the cases with wind direction 22.5° and 45°, the computational domain should be larger, since the `min` and `max` coordinates of the building domain are changing as we rotate it. So, the final dimensions of our blockMesh for the 3 wind angles are:

- 0° wind angle: $[9 \times 6 \times 2.4]$ meters
- 22.5° wind angle: $[9.3 \times 7.6 \times 2.4]$ meters
- 45° wind angle: $[9.4 \times 8.4 \times 2.4]$ meters

Below the `min` and `max` coordinates of the building domain and blockmeshes are presented (Table 3.2). Also, a clear view of the `blockMesh` dimensions and the distances of our building domain from the mesh boundaries is given (Figures 3.10, 3.11, 3.12).

Variables	Building Domain 0°(m)	BlockMesh 0°(m)	BlockMesh 22.5°(m)	BlockMesh 45°(m)
x_{min}	-0.5	-2.5	-2.65	-2.7
x_{max}	0.5	6.5	6.65	6.7
y_{min}	-0.5	-3	-3.8	-4.2
y_{max}	0.5	3	3.8	4.2
z_{min}	0	0	0	0
z_{max}	0.4	2.4	2.4	2.4

Table 3.2.: BlockMesh dimensions

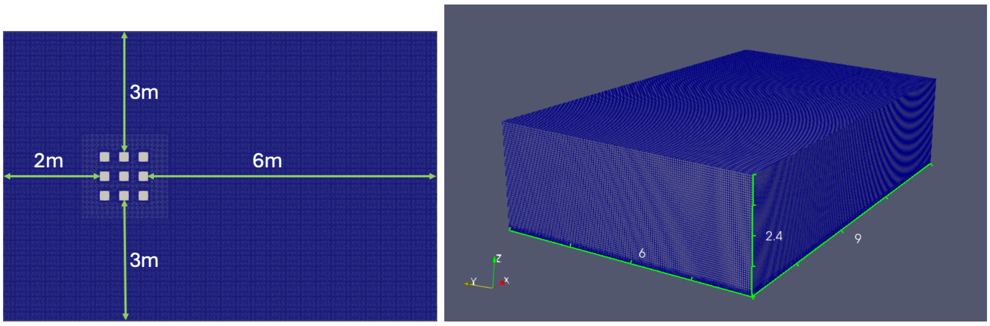


Figure 3.10.: Blockmesh (0° wind angle)

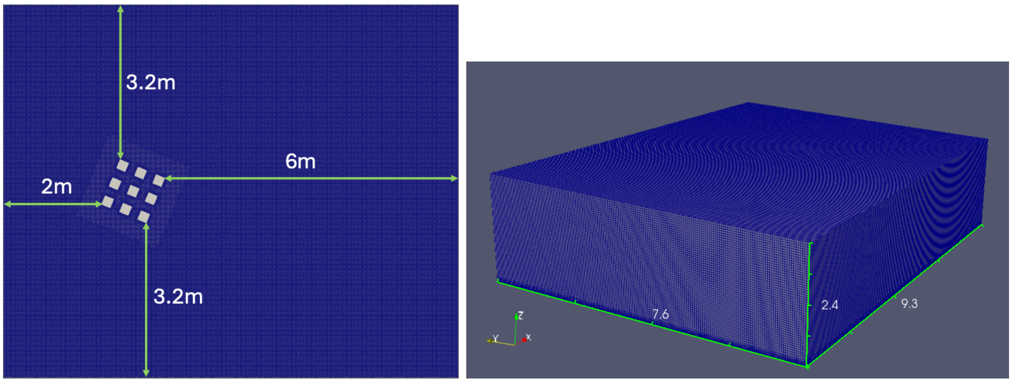


Figure 3.11.: Blockmesh (22.5° wind angle)

3.6.2. Grid resolution

One more functionality of `blockMesh` is to define the resolution of the mesh by dividing the space inside the domain into hexahedral cells. The criteria that were followed for the selection of the appropriate cell size are analyzed below:

- According to Blocken [2015], the grid should consist of at least 10 cells per cube root of the building volume and also between every two buildings. In our dataset, we have 9 cubes, as buildings, with a dimension of 0.2 meters, and an equal distance from each other of 0.2 meters. So, $0.2/10 = 0.02$m cell size can be a first estimation.
- Another guideline that should be followed is that the height of 1st cell from the ground should not be less than our z_0 [Blocken, 2015], which is equal to 0.00045m. In our case,

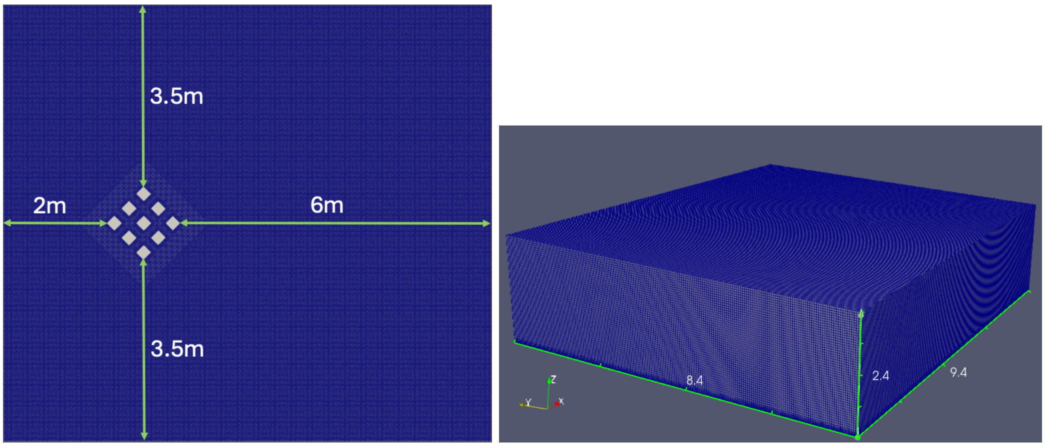


Figure 3.12.: Blockmesh (45° wind angle)

0.02m is higher than 0.00045m.

- This AIJ case simulates an urban environment since these cubes represent simple building blocks. So, according to Blocken [2015], for pedestrian level wind studies, the 3rd or 4th cell should reach the height of the probes (0.02m). It is important not to extract data from the first cell closer to the ground because our wall function can heavily influence our results. So by refining the area close to the ground, the desired height is achieved.

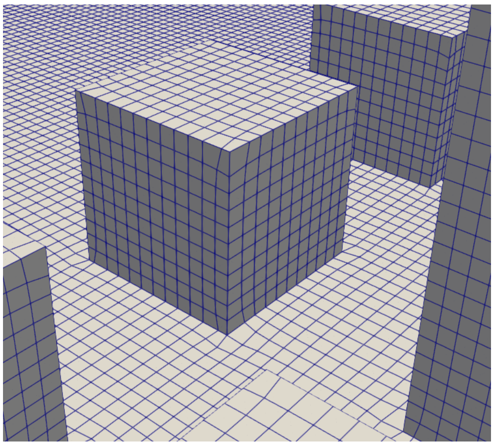
- Another important aspect to be considered is the quality of the computational cells in terms of shape. It is preferable to divide our domain into cubic cells. According to Blocken [2015] guidelines, it is advised that the stretching ratio should be kept below 1.3 in regions of sharp gradients e.g. near buildings, to limit the truncation error. Another reason is to achieve grid independence and convergence. By using cubic cells, we can uniformly refine the grid in all three dimensions without introducing additional complexity. According to Blocken [2015], the systematic refinement of grids is the key to achieving grid-independent results, and cubic grids are ideal for uniform refinement.

To make sure that we will run the simulation having designed the most appropriate mesh, we had to design 3 meshes with progressive refinement: a coarse, a medium and a fine mesh. Then, running the simulation for each of them will finally lead us to decide which one is the most suitable to use. As it is mentioned by Blocken [2015], the resolution of the medium and coarse mesh can be calculated by multiplying our cell size by a factor (grid refinement ratio) of 1.5 consecutively. As a result, we had 0.02*1.5 = 0.03m resolution for the medium mesh and $0.03 * 1.5 = 0.045$m resolution for the coarse mesh.

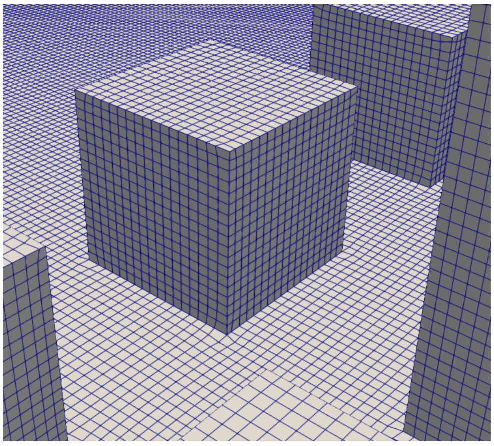
However, it is important to mention that due to the simplicity of the geometry and excessive computational time, we were led to the conclusion that we couldn't proceed with the 0.02m resolution mesh as our fine mesh, so we considered the 0.03m cell size for our fine mesh and we simplified our coarse mesh, 0.045*1.5≃0.07. So, the final cell sizes for each mesh, alongside the number of cells for every dimension, are 0.03m, 0.045m, 0.07m and are presented below (Table 3.3). Also, by using `ParaView` software, we present an example of our geometry with the three different resolutions (Figure 3.13).

Mesh type	Cell size (m)	nCells x	nCells y	nCells z	Total cells
Coarse Mesh	$[0,07 \times 0,07 \times 0,07]$	128	86	34	374.272
Medium Mesh	$[0,045 \times 0,045 \times 0,045]$	200	133	53	1.409.800
Fine Mesh	$[0,03 \times 0,03 \times 0,03]$	300	200	80	4.800.000

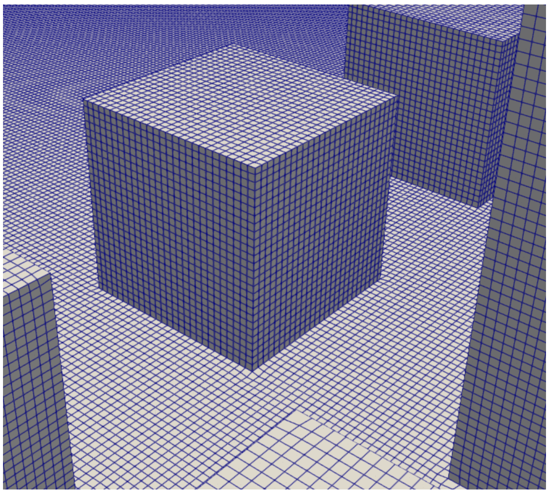
Table 3.3.: Cells per mesh



(a) Coarse Mesh: 0,07m Resolution



(b) Medium Mesh: 0,045m Resolution



(c) Fine Mesh: 0,03m Resolution

Figure 3.13.: The three meshes

In the `blockMeshDict` file, we defined the meshing parameters that are mentioned so far, including the dimensions of our domain and the number of cells. Applying the `blockMesh` command, the background mesh was generated. Then, we used the `surfaceFeatures` command to improve the generation of the mesh as it allows us to extract the outlines of the buildings, which are going to be used in the final mesh generation. Also, this command helps to detect the location of the buildings and their shape.

3.6.3. Mesh refinement

Another important parameter that we should consider for our mesh is to achieve higher resolution in our region of interest. `SnappyHexMesh` was used to perform this refinement and adapt the mesh as close as possible to the real geometry and our building domain. For the mesh refinement, we used the building model boundaries, and we defined 2 levels of refinement. Specifically, we created 2 refinement boxes and 2 refinement levels. They were applied to all three meshes to refine the area of interest, which in our case is the terrain and the building domain, where the probes are located. The refinement regions, their dimensions and the levels of refinement for the coarse, medium and fine meshes of the cases 0h, 1h and 2h are:

- Building surfaces: Level 2
- Refinement box 1: $[-2.5, -3, 0] \times [6.5, 3, 0.04]$ m (Level 2)
- Refinement box 2: $[-0.9, -0.9, 0] \times [0.9, 0.9, 0.8]$ m (Level 1)

Box 1 spreads across our whole computational domain at a relatively small height, enclosing all our probes. It is applied throughout the entire domain and matches the refinement level of the buildings to ensure that the wind profile is not heavily interpolated. Box 2 includes our building site. The height of it was chosen accordingly so as not to exceed the height of $2H_{max}$. The rest of the computational domain is considered level 0. The criteria for choosing these refinement boxes were to be equally distributed in space, to include the whole building domain, our probes, and to have a smooth transition from the sparser grid of the free wind flow to the denser grid that is around the cubes and reaches their walls. Since our geometry consists of plain surfaces and simple shapes like cubes without any complicated composition or abrupt angles, it was deemed appropriate not to carry out further refinement. In Figure 3.14, we can see the refinement boxes and in Figure 3.15, we can see our domain in `Paraview`.

For the cases with 22.5° and 45° wind direction, we adjusted the refinement regions accordingly. The refinement box 1 was rotated through the command: `surfaceTransformPoints -yawPitchRoll "(-45 0 0)" ref_box.obj rotated_ref_box.obj` and the rotated OBJ file was used as a refinement region. The refinement box 2 was adapted to the other blockmeshes, covering their entire terrain respectively.

3.7. Solver type

The simulation is performed by an iterative method, through which the variables U, P, k, ϵ and nut will be calculated by OpenFOAM, for every individual cell of our mesh. The solver is responsible for resolving numerically our set of equations by representing particular physical phenomena. Our cases are set up for a steady-state, incompressible, turbulent flow simulation, so the solver that is used is the `simpleFoam`. The GAMG solver is used for pressure, and a smoothSolver for velocity and turbulence, with a Gauss-Seidel smoother in order to smooth out errors. The convergence limit has been set to 10^{-9} for the residuals, ensuring high accuracy for our model.

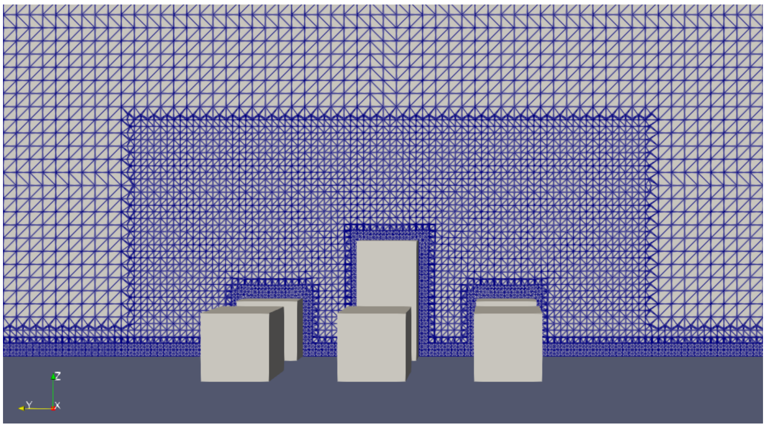


Figure 3.14.: Refinement Boxes (Case 2H)

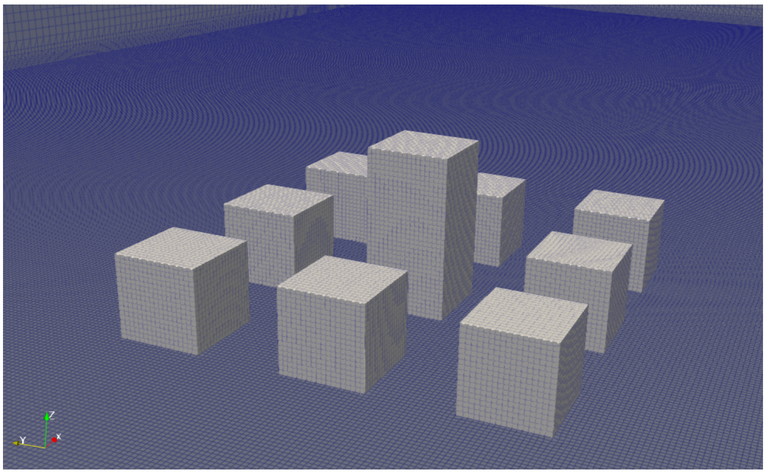


Figure 3.15.: Refined Mesh (Case 2H)

3.8. Uncertainty through translation

The geometric uncertainty is a critical factor, especially when working on CFD simulations around city blocks. Our goal is to quantify the error observed through rotations and translations of our model. An important aspect that should be considered is that our case simulates simple building blocks and represents a real-world scenario. So, we could assume that the height of the 120 probes (0.02 m) represents pedestrian height ($\simeq$ 2 m). Therefore, by scaling up our geometry to 1:100, it represents a realistic urban-case scenario. Using as a reference the height of the probes (pedestrian level), our scaling factor is calculated and set: $\alpha = 100$. As a result, we now have the prototype scale, which represents the wind-tunnel model and the reality scale (i.e. city scale), which is 100 times bigger. In the next steps of our analysis, we will use the prototype scale in order to avoid any confusion or inconsistency.

The direction of the translation could affect our results in different ways. Towards the x-axis (parallel to the wind), it impacts the incoming flow conditions (wake formation is altered) and probes placed behind the buildings will show more changes in velocity. Towards the y-axis (perpendicular to the wind), the geometry comes closer to the boundaries. To quantify this impact, the type of uncertainty that we will study in more detail is the uncertainty that comes from point cloud reconstruction. After related research, specific uncertainty levels are mentioned. In Huang et al. [2022], Root Mean Squared Error (RMSE) was used to quantify the quality of each reconstructed model. The statistics of the quantitative results on the buildings were reported and the obtained reconstruction accuracy for all buildings is between 4cm and 26cm. Also, according to Jarząbek-Rychard and Maas [2023], the positional uncertainty of the reconstructed model vertices, that is specified at the 95% confidence level, was grouped in: <1mm, 1-5mm and 5-10mm classes.

According to the literature review and from a realistic scope, we assumed five, progressively increasing, levels of uncertainty: 1cm, 10cm, 20cm, 50cm, and 100cm. However, these levels of accuracy apply in real-world scenarios. For our case, we need to apply these uncertainties as translations (offsets) in our simulations. To achieve that, we need to scale down these values, in prototype scale, according to α, so we get the following results:

- $\boldsymbol{t_1}$: $1\,\text{cm} = 0.01\,\text{m} \rightarrow 0.01/\alpha = \pm 0.0001\,\text{m}$
- $\boldsymbol{t_2}$: $10\,\text{cm} = 0.1\,\text{m} \rightarrow 0.1/\alpha = \pm 0.001\,\text{m}$
- $\boldsymbol{t_3}$: $20\,\text{cm} = 0.2\,\text{m} \rightarrow 0.2/\alpha = \pm 0.002\,\text{m}$
- $\boldsymbol{t_4}$: $50\,\text{cm} = 0.5\,\text{m} \rightarrow 0.5/\alpha = \pm 0.005\,\text{m}$
- $\boldsymbol{t_5}$: $100\,\text{cm} = 1\,\text{m} \rightarrow 1/\alpha = \pm 0.01\,\text{m}$

So we took into account three base cases, with 0°, 22.5° and 45° wind angle and compared them with the uncertain ones. The minimum distance between the probes and the buildings is 0.05 m, so none of these translations affect them (e.g. by obstructing any of the probes with a building wall). The command that was initially used to import these uncertainties in our geometry is '`SurfaceTransformPoints -translate "(dx dy 0)" cubes.obj translated_cubes.obj`'. This command would transpose our model towards a certain direction and magnitude as a whole. But instead, the goal is to assign different uncertainty values to each building, according to the average uncertainty of our domain and achieve more realistic results. So, five samples (one for each translation value t_i) had to be created, and the following conditions were considered appropriate to be met for the applied part in Section 5.1. They are presented below:

1. The uncertainty will be represented as an offset
2. The offset is radial (could take place towards any direction)
3. There is one translational uncertainty value per building
4. Each sample contains the translation values of the buildings, following the normal distribution $N[\mu, \sigma]$. Regarding the normal distribution of each sample:
    - the mean (μ) is centered around the translation value t_i
    - the coefficient of variation is $CV = 10\% = 0.1$
    - the standard deviation is $\sigma = t_i * CV$

3.9. Uncertainty through rotation

Following a similar method regarding rotational uncertainty, we took into account three base cases, with $0°$, $22.5°$ and $45°$ wind angle and compared them with the uncertain ones. In this type of uncertainty, there is no reason for scaling since we are dealing with angles. Similarly, we noticed that $1° \approx 0.0025m$ offset on the x/y coordinates of our geometry. So, we assumed five, progressively increasing, levels of rotational uncertainty: $0.5°$, $1°$, $2°$, $3°$, $5°$ and we applied these uncertainties as rotations in our simulations. Again, the command that was initially used to import these uncertainties in our geometry is ''`surfaceTransformPoints -yawPitchRoll "(`ϕ` 0 0)" cubes.obj rotated_cubes.obj`'. But, this command would rotate our model towards a certain direction and equally as a whole. Instead, the goal was to assign different rotation values to each building, according to the average rotational uncertainty of our domain and achieve more realistic results. So, five samples (one for each rotation value r_i) had to be created, for the applied part presented in Section 5.2 and we made the following assumptions:

1. The uncertainty will be represented as a rotation
2. The rotation takes place either in the clockwise or counterclockwise direction
3. There is one rotational uncertainty value per building
4. Each sample containing the buildings' rotation values follows the normal distribution $N[\mu, \sigma]$. Regarding the normal distribution of each sample:
   - the mean (μ) is centered around the rotation value r_i
   - the coefficient of variation is $CV = 10\% = 0.1$
   - the standard deviation is $\sigma = r_i * CV$

4. Implementation and verification

4.1. Performing the simulations

We run our case with the 3 different meshes (coarse, medium and fine) and for three different building heights (0h, 1h, 2h). We choose the case with 0° wind angle, since the dimensions of the `blockmesh` are smaller and the running time would be lower. Then we compared the results with the experimental values provided by the dataset. It is important to mention here that the number of processors used both for creating the meshes and the simulations was 24. This number did not change throughout the whole process, in order to compare the running times properly.

The finer the mesh, the more time it needs to converge. For our case, the number of iterations was set to 5000. This is the number of iterations needed so that our fine mesh fully converges. For the comparison, we used the same number of iterations for the coarse and medium meshes. Below, at tables 4.1, 4.2 and 4.3, is presented in detail the number of cells that were created, the time required to create the refined meshes, and the running time of each simulation. Each table represents a different geometry case. As expected, every category increases from coarse to fine mesh. The following numbers were taken from the `.log` files that were used to monitor the simulation.

Mesh type	Cells after refinement	Meshing time (sec)	Running time (sec)
Coarse mesh	891.124	18,15	876
Medium mesh	3.470.800	88,12	4441
Fine mesh	11.733.731	375,81	17360

Table 4.1.: Meshes: Case 0h

Mesh type	Cells after refinement	Meshing time (sec)	Running time (sec)
Coarse mesh	892.871	16,95	861
Medium mesh	3.472.794	83,08	4440
Fine mesh	11.737.254	377,61	17187

Table 4.2.: Meshes: Case 1h

Mesh type	Cells after refinement	Meshing time (sec)	Running time (sec)
Coarse mesh	894.696	16,5	804
Medium mesh	3.475.658	105,93	4120
Fine mesh	11.743.649	396,89	15376

Table 4.3.: Meshes: Case 2h

4.2. Analysis and post-processing

4.2.1. Residuals

The first step related to the post-processing of the results is the convergence of the residuals. In practice, through the folder `Post-processing/Residuals` that was created, we plotted the residuals of the field values and observed if they converged according to our standards. A representative resulting graph of the case 0h is presented below (Figure 4.1). The rest of the resulting graphs are presented in the appendix (Figures B.1, B.2, B.3). Also, we plotted the velocity magnitude of five random probes to verify that the velocity at these points has been stabilized (Figure 4.2).

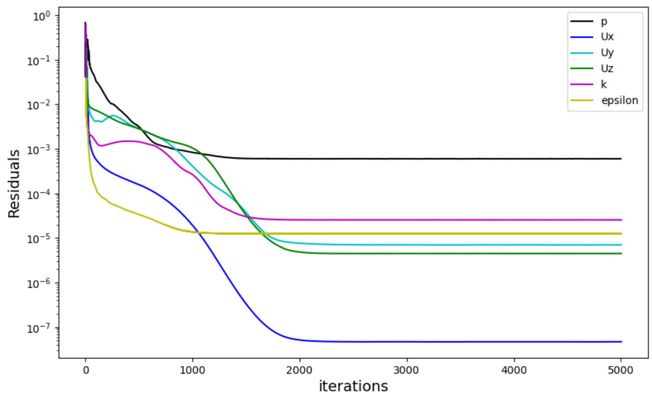


Figure 4.1.: Residuals: Case 0h (Medium Mesh)

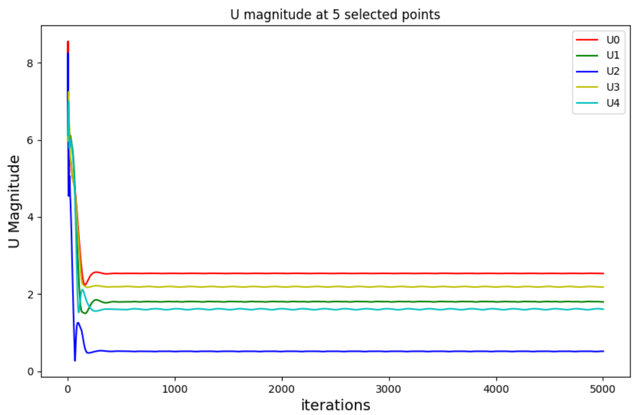


Figure 4.2.: Velocity magnitude over time for five monitoring points

As we can see, the behavior of our mesh is pretty similar in the three cases, since the geometry does not change drastically. As it is expected, the U_x variable is the one that converges faster, since the wind flows parallel to the x-axis and our case simulates an urban environment. In different urban environments, though, where the geometry is characterized by high complexity, it would be harder to achieve small values for the residuals, which explains the results that we obtain in our cases. All of the cases converged, reaching the desired 10^{-6} or lower. We see that the coarse and medium meshes reached the converged solution faster than the fine mesh. For the coarse mesh, approximately 1000 iterations were needed to reach convergence, and for the medium, approximately 2000. For the fine mesh, convergence is achieved within 4000 iterations, much longer than the other two. As expected, the finer the mesh, the more iterations it takes to get closer to zero, since the number of cells is significantly higher. Nevertheless, in terms of time and number of iterations, the coarse mesh and the medium mesh have a clear advantage in all three cases.

4.2.2. Field plots

To compare effectively the three meshes, we created plots of the field values in `Paraview`, which are displayed below. A representative field plot of Ux for the three meshes of the case 0h is presented here (Figure 4.3). The rest of the plots are presented in the appendix (Figures C.1, C.2, C.3). Specifically, the tool that was used was the `'plot over line'` from the data analysis tools in `Paraview`. The line that was chosen was the $[-2.5, 0.15, 0.1], [6.5, 0.15, 0.1]$. The criteria of choosing this line was 1) to be parallel to the wind direction (x-axis), 2) to be as long as the mesh, 3) to be located at a proper height where our field values present some variance and 4) not to be obstructed by the geometry at any point.

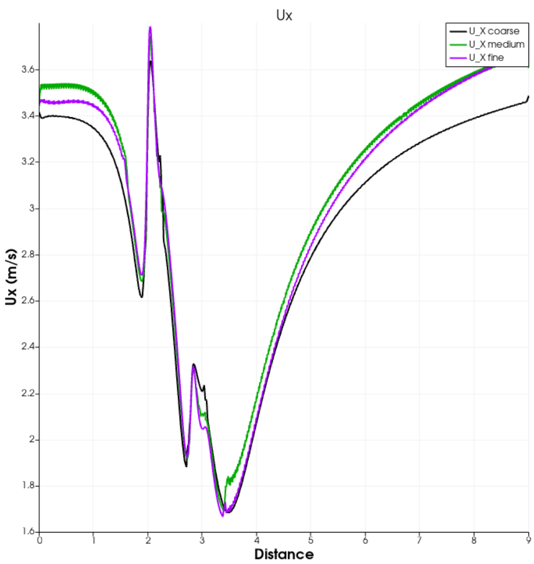


Figure 4.3.: Ux field plot for the three meshes: Case 0h

Through this tool, we plotted the values of our fields of interest Ux, Uy, Uz, p, k and ϵ over this line. As we can see from these plots, the lines are closer in the cases of the fine and medium meshes, while the gap between the coarse and medium meshes seems larger.

4.2.3. Grid Convergence Index

Another criterion for the selection of the most optimal mesh is the comparison of the wind velocity at the points of interest (probes). A meaningful metric to quantify the differences between them was used for this purpose. Grid convergence is often analyzed using the Grid Convergence Index (GCI), a formal method proposed by Roache [1994] for quantifying grid convergence and the level of uncertainty related to grid discretization. It is considered a more accurate and understandable way to present our results, and it involves computing an error estimate based on our solutions at different grid levels. The lower GCI value, the more grid-independent our solution is. Also, a lower GCI value indicates that the solution is less sensitive to further mesh refinement, suggesting grid convergence. The implementation and validation algorithm was based on the procedure outlined by Celik et al. [2008], and was implemented through Python. The key steps involved in the GCI calculation and the results are presented below.

The necessary inputs for the calculation of GCI are the representative cell sizes (h) from Table 3.3, and the quantity of interest (phi) of the three meshes. The key requirement for the selection of phi is that it should be a representative scalar quantity that is sensitive to grid refinement [Celik et al., 2008]. The velocity magnitude at the probes' location, from the folder `postProcessing/Probes`, fits this criterion. The measurement of the discretization error does not necessarily behave linearly across our probes, and we want to avoid mixing regions of different convergence behavior. For that reason, we computed the GCI for each probe separately based on its respective velocity values across the different grids. Then we computed an average to obtain a representative GCI across the entire domains. In order to strengthen the results of this particular method, we included 80 new points in our calculations, and we increased the total number of probes from 120 to 200. The reason behind this decision was the fact that our 120 initial points were gathered close to each other at the center of our building domain, close to the ground. Thus, through Python, we randomly generated 80 sampling points with the assumption: 1) that they are within the boundaries of our study area, 2) that they are not located inside buildings and 3) that they are uniformly dispersed in space. Afterwards, some key steps were the calculation of:

- Grid refinement ratio: $r_{21} = h_2/h_1 = 0.045/0.03 = 1.5$ and $r_{32} = h_3/h_2 = 0.07/0.045 = 1.556$. These ratios represent how much finer one grid is compared to the next coarser one.
- Error estimation: $e_{21} = phi_2 - phi_1$ and $e_{32} = phi_3 - phi_2$. These are the differences in the computed quantities of every probe, between successive grids, for our three cases.
- Apparent order of convergence (p): A measure of how fast the solution of the numerical simulation improves as the grid is refined. It is calculated using Richardson extrapolation.

So, for our three meshes, we calculated two GCIs for every probe: 1) GCI_{21}: between the fine and medium grid and 2) GCI_{32}: between the medium and coarse grid. For the computation, we used Roache's formula, where:

$$GCI_{21} = \frac{1.25 \left| \frac{phi_2 - phi_1}{phi_1} \right|}{r_{21}^p - 1} \times 100\% \tag{4.1}$$

Then, the median method was followed to calculate our final values. The results are presented below (Table 4.4):

	Case 0h	**Case 1h**	**Case 2h**
GCI_{21}:	2.612%,	1.314%,	2.553%
GCI_{32}:	3.069%,	4.518%,	5.640%

Table 4.4.: GCI results

Based on our results, we come to the observation that both GCI values are close to each other in all three cases. This measure expresses how much the refinement improved our results by reaching closer to the true values. For all three cases, the $GCI_{21} < GCI_{32}$, which means that from the coarse to medium mesh our results were improved by $3-5\%$ and from the medium to fine mesh we observe an improvement of around $1-2\%$. The fine mesh is the most accurate, but further refinement might improve our results by an even smaller percentage, with a high computational cost.

These low GCI values indicate that grid independence is approached and the refinement is carried out properly. Finally, based on the results, we have small differences both between coarse-medium and medium-fine grids in all of our cases and we conclude that the medium mesh is the best choice, according to GCI. The medium mesh is more accurate than the coarse mesh and it is considered a balanced choice, since it combines reasonable accuracy, without the high computational cost of the fine mesh.

4.3. Mesh convergence

Achieving grid convergence in CFD indicates that the solution of the discretized equations and our simulation becomes grid independent as the mesh is refined, and tends to the exact solution of the differential equations. It is ensured that the results are not significantly influenced by the size or distribution of the computational cells in the mesh. Also, grid convergence is important for 1) the validation of our results, since it ensures our numerical results are accurate and not dependent on grid resolution and 2) the reproducibility, since it confirms that others can reproduce our results with different grids.

Finally, according to these four criteria: 1) convergence time of the residuals, 2) number of iterations needed, 3) field plots and 4) GCI, the mesh chosen to perform the analysis of the results is the **medium mesh**. It offers a good balance between accuracy and computational cost. In case we need higher accuracy, we could verify our results by running a few key simulations on the fine mesh, but for continuing the full analysis, the medium one is preferred.

4.4. Comparison with experimental data and validation

In the dataset provided by AIJ is available the velocity of the 120 probes for three different geometries and three wind angles. We observe that the velocity values that are experimentally measured (U_{AIJ}) are normalized by the inflow velocity at the same height (0.02m). This means that the velocity values that we have for the probes are expressed as a ratio with respect to the inflow velocity at the same height. We can assume that the formula used for this transformation is (4.2):

$$U_{\text{AIJ(normalized)}} = \frac{U_{\text{AIJ}}}{U_{\text{inflow}}(z)} \tag{4.2}$$

After running the simulations, the resulting velocity values that we have for the probes are at 0.02m height. For that reason, we need to denormalize the velocity ratio based on the Equation 4.2, get the U_{AIJ}, and compare the results. On the other hand, the probes file `U.dat` that was created after running our simulations contains a set of 3 numbers for every probe. These numbers are the components of velocity (U_{CFD}) in the Cartesian coordinate system (U_x, U_y, U_z). So, in order to compute the magnitude of velocity, we used the following formula (4.3):

$$U_{\text{CFD}} = \sqrt{U_x^2 + U_y^2 + U_z^2} \tag{4.3}$$

For comparison, the statistical methods that were used are the RMSE, Mean Absolute Percentage Error (MAPE), Pearson's correlation coefficient (r) and the Best-fit line. RMSE quantifies how far the simulation values are from the actual values on average, while MAPE expresses in percentage the difference between the CFD simulation and the wind-tunnel experiment. r measures the linear correlation between the simulation and experimental values. Creating the best-fit line helps us quantify and visualize bias or trends in our CFD results through the slope (a) and the intercept (b). Using these metrics, we can compare the results and determine which geometry has better results and how our estimations are affected. The equations that describe RMSE and MAPE are presented below (4.4, 4.5):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(U_{\text{AIJi}} - U_{\text{CFDi}})^2} \tag{4.4}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{U_{\text{AIJi}} - U_{\text{CFDi}}}{U_{\text{AIJi}}}\right| 100\% \tag{4.5}$$

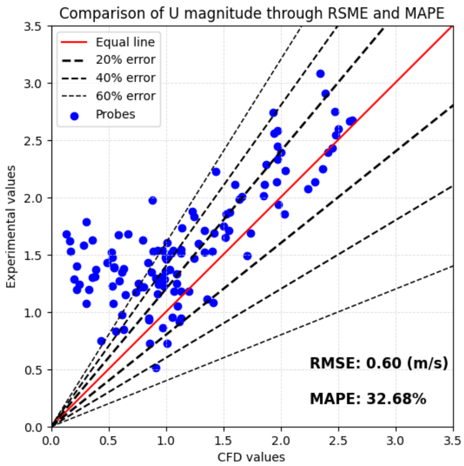
We created a Python script that was tailored to our needs in order to meet the requirements for an accurate visualization of our results and the drawing of conclusions. So, for the evaluation of our simulations against the experimental values, we performed a point by point comparison (scatter plots) by placing the CFD values (U_{CFD}) at the x-axis and the values from the wind-tunnel experiment (U_{AIJ}) at the y-axis (Figures 4.4a, 4.5a, 4.6a). The red line symbolizes the equal line $y = x$ and for the points located on this line $U_{\text{CFD}} = U_{\text{AIJ}}$. The algorithm displays the denormalized values U_{AIJ} alongside the simulation values U_{cfd} of our medium meshes. The plots are presented below, for every case (Figures 4.4, 4.5, 4.6). Also, a table is presented with the r values of every case (Table 4.5).

In addition, velocity slices at probe height (0.02m) and graphs were created that depict the velocity magnitude for each case for the 120 probes (Figures 4.4c, 4.5c, 4.6c). At the x-axis is

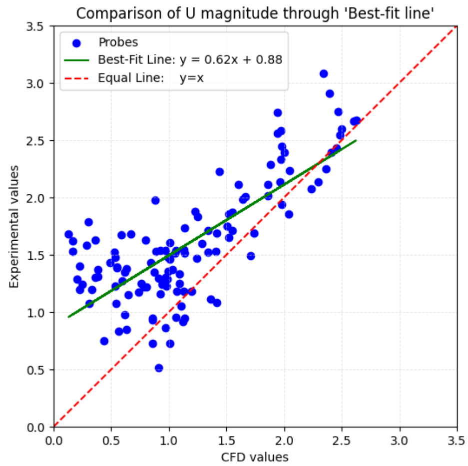
	Case 0h	Case 1h	Case 2h
r:	0.762	0.715	0.831

Table 4.5.: Pearson's r-value

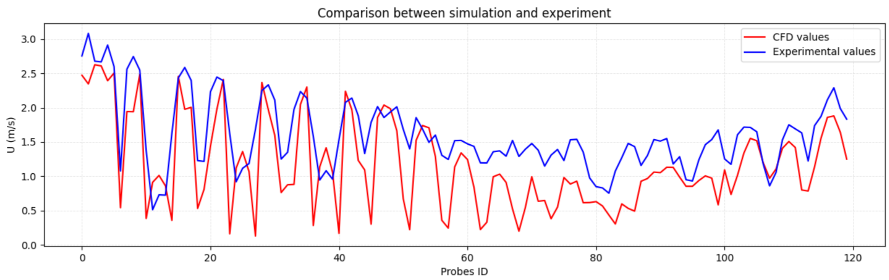
the probe's ID and at the y-axis is the U magnitude. For every ID, there are two U values and the line was used to create a more uniform image of the points that belong either to the CFD simulation or to the experiment. According to the r value, case 2h is more correlated to the experimental data, compared to cases 1h and 0h. Observing the plots, we can also see that the best results appear in the 2h case with $MAPE = 24\%$, while in the other two cases, we see that the middle cube's height increases the error to 32% and 33% respectively. Alongside this difference, it is observed that the experimental values are slightly higher than the values of the simulation. A relative offset is observed in certain locations, while the points of interest with lower velocity magnitude present higher variance compared to points with higher velocity value. A probable reason for this mismatch could be the z_0 value, since it determines the roughness of the terrain and possibly underestimates the wind speed. Given that the best results are from case 2h, it is the case that we will choose to continue with our uncertainty analysis, since different approaches will be considered to further understand the influence of building footprints in the CFD results.



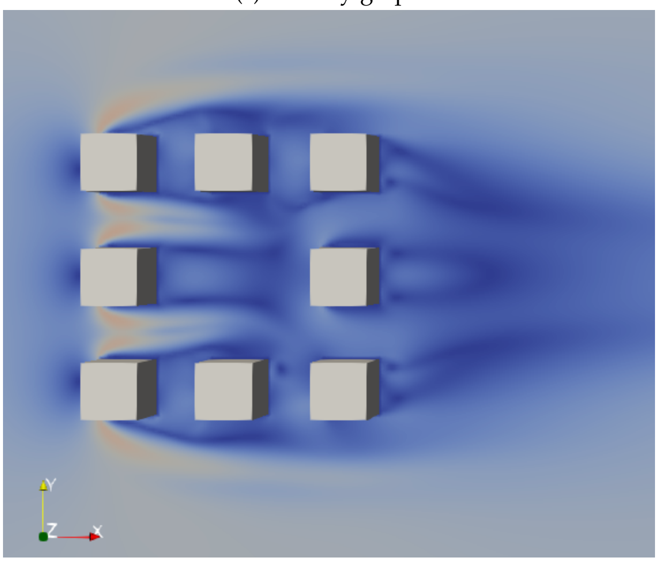
(a) RMSE and MAPE



(b) Best-fit line

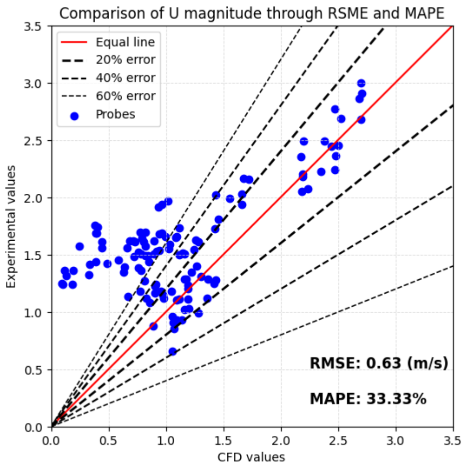


(c) Velocity graph

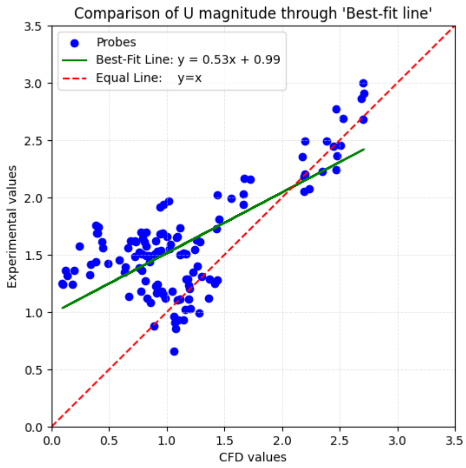


(d) Velocity slice at probes' height (0.02m)

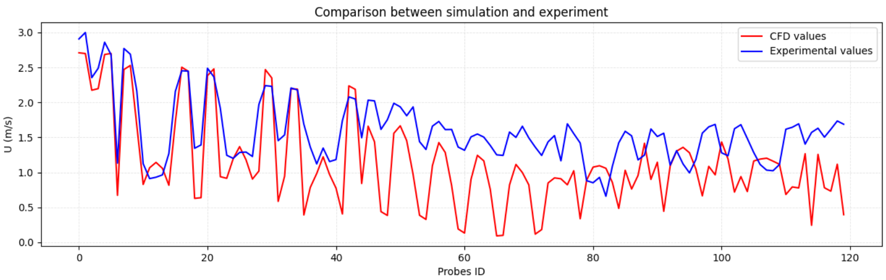
Figure 4.4.: Comparison Wind-tunnel–CFD values (Case 0h)



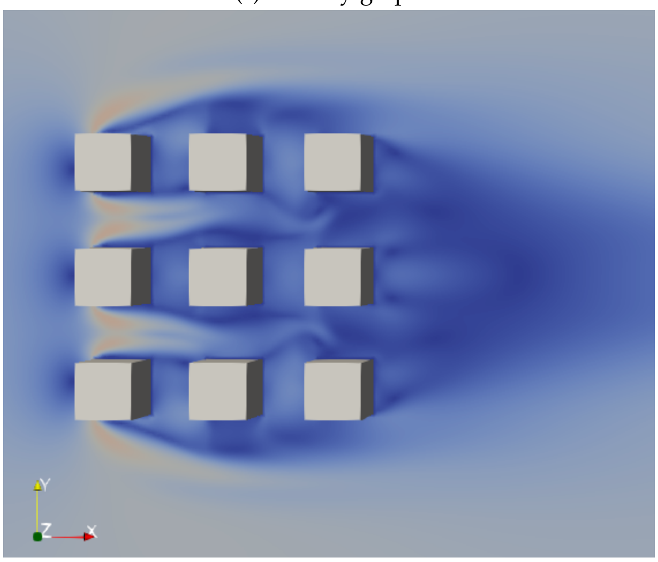
(a) RMSE and MAPE



(b) Best-fit line

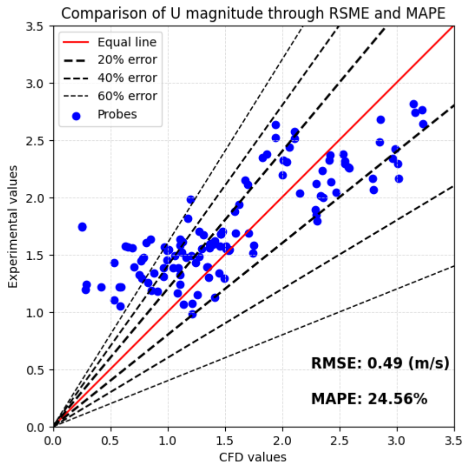


(c) Velocity graph

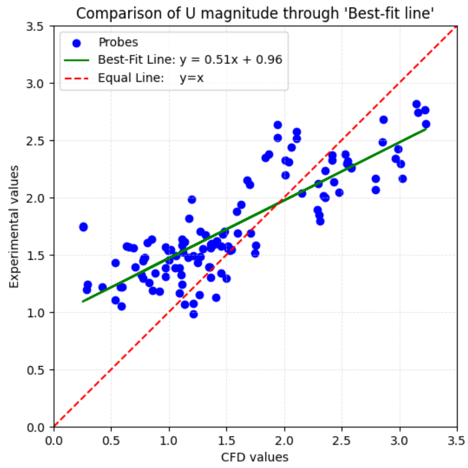


(d) Velocity slice at probes' height (0.02m)

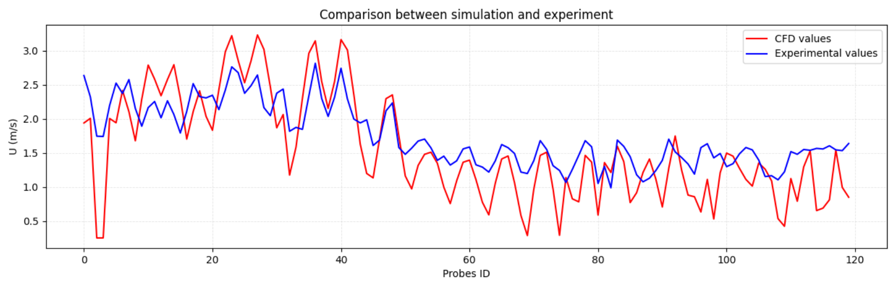
Figure 4.5.: Comparison Wind-tunnel–CFD values (Case 1h)



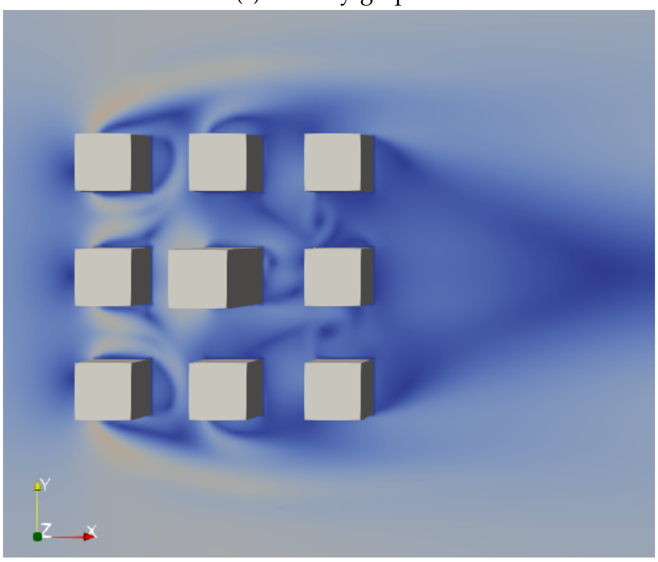
(a) RMSE and MAPE



(b) Best-fit line



(c) Velocity graph



(d) Velocity slice at probes' height (0.02m)

Figure 4.6.: Comparison Wind-tunnel–CFD values (Case 2h)

5. Results and Analysis

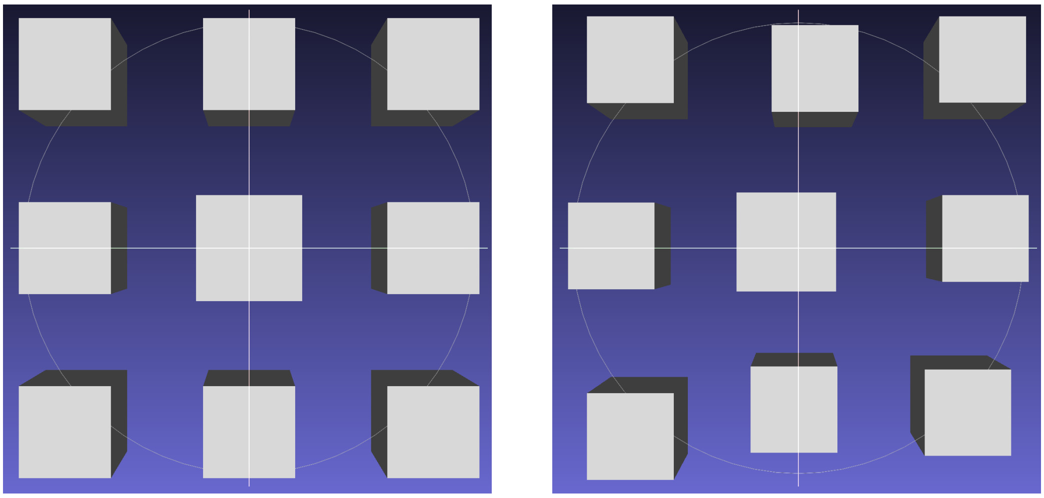
5.1. Uncertainty analysis—Translation

In this section lies the essence of our research, and we will take a more in-depth look at how the results of our simulations change by perturbing their geometry with specific levels of uncertainty. The goal of our research is ultimately to find and quantify the degree to which the results are affected by each level of uncertainty, but also which areas of our domain have undergone the greatest influence. Since we decided to continue our analysis with case 2h (Section 4.4), we also had to run the simulations with the 22.5° and 45° wind angles. The geometry `cubes2h0.obj`, `cubes2h22.obj` and `cubes2h45.obj` (Figure 3.8) will be used. Thus, we will use these three cases as the base cases and compare them with our final results.

To automate this process that is analyzed in Section 3.8 and fulfill all the abovementioned requirements, we created a Python script that performs the following and an example of an output is presented below (Figure 5.1):

1. The first part of the code:
   - Creates a random sample, considering the level of uncertainty t_i (prototype scale), ($n_{sample} = n_{buildings} = 9$ values of uncertainty)
   - The sample values follow the normal distribution with $N[\mu, \sigma]$
   - Creates a sample with direction angles $\theta \in [0, 2\pi)$, which determine the direction of every offset vector and computes the dx and dy offsets (which can also receive negative values) using $\cos\theta$ and $\sin\theta$ respectively
   - Stores dx and dy offsets in numpy arrays
2. The second part of the code:
   - Reads the initial OBJ file
   - Applies the transpose values of the numpy arrays in the x and y coordinates of the vertices following the enumeration of the buildings given in Figure 3.6b. Our geometry is simple, thus, a different offset is applied every 8 vertices (lines starting with "v"), because of the cubic shapes.
   - Creates a new translated OBJ file

Ultimately, we created five samples for every translation t_i. It is important to mention here that the same direction angles (θ) will be used for all five translations, but for the cases with 22.5° and 45° wind angle, we subtracted 22.5°($\simeq$ 0.4 rad) and 45°($\simeq$ 0.8 rad) respectively, so as the relative offsets between the buildings, while we rotate our geometry, to remain the same. Each sample consists of 9 values, as many as the buildings, and they are presented in the following Table 5.1. All five levels of translation were applied in our three base



(a) (b)

Figure 5.1.: Output example of the translated geometry before (a) and after (b)

cases: case2h0, case2h22, case2h45 (Figure 3.8) and finally, we have 15 cases with translated geometry.

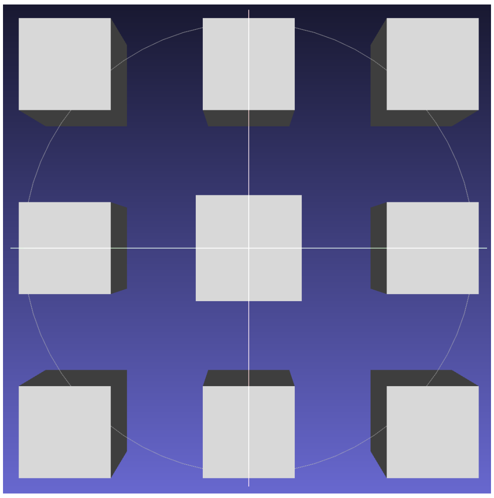
Buildings ID	θ **(rad)**	$\theta -$ **0.4 (rad)**	$\theta -$ **0.8 (rad)**	$t_1 =$ **0.0001 m**	$t_2 =$ **0.001 m**	$t_3 =$ **0.002 m**	$t_4 =$ **0.005 m**	$t_5 =$ **0.01 m**
1	2.7	2.3	1.9	0.00010	0.00114	0.0019	0.0052	0.0082
2	3.4	3.0	2.6	0.00010	0.00101	0.0022	0.0047	0.0100
3	2.8	2.4	2.0	0.00011	0.00108	0.0018	0.0048	0.0103
4	2.5	2.1	1.7	0.00010	0.00102	0.0022	0.0040	0.0105
5	1.9	1.5	1.1	0.00010	0.00093	0.0018	0.0048	0.0118
6	1.1	0.7	0.3	0.00009	0.00092	0.0020	0.0053	0.0091
7	0.5	0.1	6.0	0.00011	0.00111	0.0018	0.0051	0.0111
8	4.3	3.9	3.5	0.00012	0.00101	0.0022	0.0058	0.0097
9	1.7	1.3	0.9	0.00010	0.00100	0.0020	0.0054	0.0114

Table 5.1.: Uncertainties-Translations

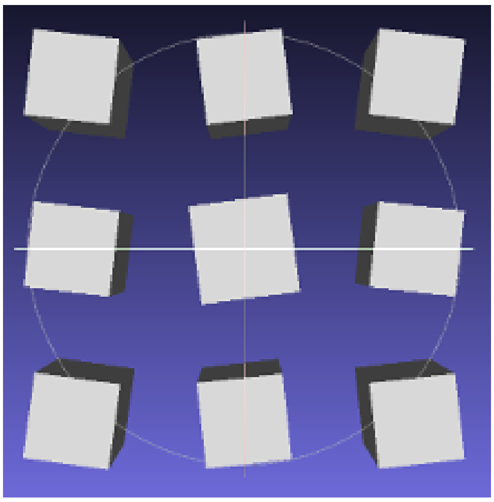
5.2. Uncertainty analysis—Rotation

As in translation, also in rotation, to automate this process that is analyzed in Section 3.9 and fulfill all the abovementioned requirements, we created a Python script that performs the following. The rotation algorithm works similarly to the translation algorithm. An example of an output is presented below (Figure 5.2):

1. The first part of the code:
   - Creates a random sample, considering the level of uncertainty r_i ($n_{sample} = n_{buildings} =$ 9 values of uncertainty)
   - The sample values follow the normal distribution with $N[\mu, \sigma]$
   - Creates a random sample, consisting of 9 values that are either -1 or 1, indicating the direction of the rotation (the same directional vector is kept throughout all the simulations)
   - Each direction is multiplied by an angle
   - Stores ϕ angles in numpy arrays
2. The second part of the code:
   - Reads the initial OBJ file
   - Calculates the centroid of every object (cubic building)
   - Applies the rotation values of the numpy arrays in the x and y coordinates of the vertices around the centroid, following the enumeration of the buildings given in Figure 3.6b. Our geometry is simple, thus, a different rotation is performed every 8 vertices (lines starting with "v"), because of the cubic shapes.
   - Creates a new rotated OBJ file



(a)



(b)

Figure 5.2.: Output example of the rotated geometry before (a) and after (b)

Ultimately, we created five samples for every rotation r_i. Each sample consists of 9 values, as many as the buildings, and they are presented in the following Table 5.2. All five levels of rotation were applied in our three base cases: case2h0, case2h22, case2h45 (Figure 3.8) and finally, we have 15 cases with rotated geometry.

Buildings ID	Directions: -1: cw, 1: counter-cw	$\phi_1 = 0.5°$	$\phi_2 = 1°$	$\phi_3 = 2°$	$\phi_4 = 3°$	$\phi_5 = 5°$
1	1	0.51	0.80	1.92	2.30	5.66
2	-1	-0.43	-0.93	-2.13	-3.34	-4.64
3	1	0.42	1.09	2.08	3.42	4.55
4	-1	-0.52	-0.99	-1.84	-3.22	-5.64
5	-1	-0.49	-1.22	-1.88	-2.63	-6.15
6	-1	-0.47	-0.98	-1.76	-2.87	-4.69
7	1	0.53	1.08	2.21	3.16	4.55
8	1	0.47	1.00	2.32	2.80	3.82
9	1	0.55	1.03	2.0	2.86	4.47

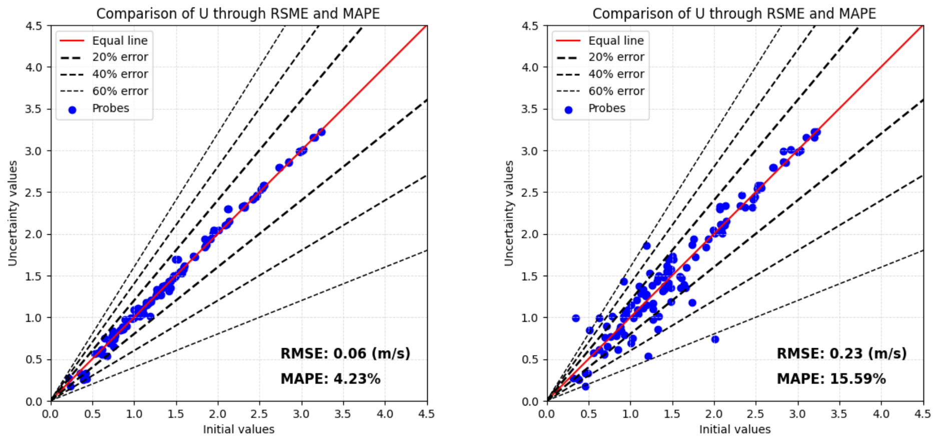
Table 5.2.: Uncertainties-Rotations

5.3. Visualization of the results and comparison

Since we implemented our methodology, validation and applied the uncertainties in the buildings, the appropriate visualization techniques are the best way to draw the right conclusions as well as to observe the phenomenon we studied. Through observation of the results, we will better understand the effects of a possible footprint uncertainty, after a point cloud reconstruction process, on the results of our simulations. Several techniques have been chosen for the visualization of the results. We used Python's matplotlib and we created scatter plots, box plots and contour plots. More detailed results can be found in the Appendices D, E, F. Here we will represent some indicative results, which we will comment on and explain the methodology with which we produced them.

5.3.1. Point-by-point velocity comparison

Similarly to Section 4.4, we created scatter plots for the comparison of the velocity of 120 probes, between the base cases (2h0, 2h22, 2h45) and the case with uncertainty. Through these plots, a point-by-point comparison is performed by placing the values of our base case on the x-axis and the values from our uncertainty case on the y-axis. For presenting here, two representative uncertainty cases were chosen, both with 0° wind angle and translation T2=0.001m and T4=0.005m (Figure 5.3). The red line symbolizes the equal line y = x, and the uncertainty had no impact on the velocity of the points located on it. The total number of plots is 30 (15 for translation and 15 for rotation) and can be found in Appendix D.



(a) T2 = 0.001m

(b) T4 = 0.005m

Figure 5.3.: Point-by-point comparison (0° wind angle)

For these two cases, we can clearly see a progressive perturbation as the uncertainty level increases. In Figure 5.3a, almost all the points are on the equal line, presenting a small deviation. On the contrary, for Figure 5.3b, many points are moved either toward the x-axis (meaning that their velocity decreases) or toward the y-axis (meaning that their velocity increases). This result is normal, as with higher uncertainty, the velocity of the points should change more. This outcome is validated even more in the corresponding Appendix D, where we can see an increase in the disturbance of the points' velocity, through the progressive perturbation of our geometry, i.e.higher uncertainty level, for all the cases.

In addition, certain statistical data were calculated and specifically RMSE and MAPE, to quantify this qualitative difference that can be visually observed through the plots. At the tables 5.3, 5.4 are presented all the statistics for a more direct comparison between different wind angles and between uncertainty levels of the same wind angle. The colour intensity of each cell was decided according to certain classes, starting from lower and reaching higher values. These classes for Table 5.3 are $[0,\ 0.1), [0.1,\ 0.2), [0.2,\ 0.3), [0.3,\ 0.4), [0.4,\ 0.5), [0.5,\ +\infty) m/s$ and for Table 5.4 are $[0,\ 2.5), [2.5,\ 10), [10,\ 20), [20,\ 30), [30,\ 100)\%$.

These tables are accompanied by four representative graphs (two for RMSE and two for MAPE) which illustrate the trend of the statistical quantities we mentioned, showing us how they are affected by the factors of wind angle and uncertainty magnitude (Figures 5.4, 5.5). As we can see in all four graphs, a progressively increasing relationship between the levels of uncertainty and errors is discernible. The greater the uncertainty, the greater the error becomes. At the same time, we observe that in both cases of rotation, the errors are higher, and the physical explanation that could be given consists of two reasons:

1. The first reason is the fact that, while the faces of the building are rotated, the area that blocks the wind becomes bigger, instead of moving parallel to it as in the case of the simple offsets.
2. The second reason is the proportionality between the rotation and translation, regarding the footprint coordinates' offset. As mentioned in Section 3.9, $1° \approx 0.0025m$. For

Uncertainty	Translation-RMSE (m/s)			Uncertainty	Rotation-RMSE (m/s)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	0.02	0.02	0.00	**R1**	0.11	0.15	0.09
T2	0.06	0.12	0.05	**R2**	0.14	0.22	0.32
T3	0.15	0.13	0.11	**R3**	0.24	0.25	0.37
T4	0.23	0.16	0.24	**R4**	0.27	0.29	0.30
T5	0.21	0.29	0.33	**R5**	0.46	0.48	0.65

Table 5.3.: Uncertainties-RMSE

Uncertainty	Translation-MAPE (%)			Uncertainty	Rotation-MAPE (%)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	1.23	0.73	0.17	**R1**	8.44	8.33	3.61
T2	4.23	5.94	1.68	**R2**	11.25	13.98	12.20
T3	12.20	6.68	4.61	**R3**	21.45	14.96	15.37
T4	15.59	7.58	8.78	**R4**	24.48	17.42	12.16
T5	19.91	14.94	12.97	**R5**	41.81	33.57	27.06

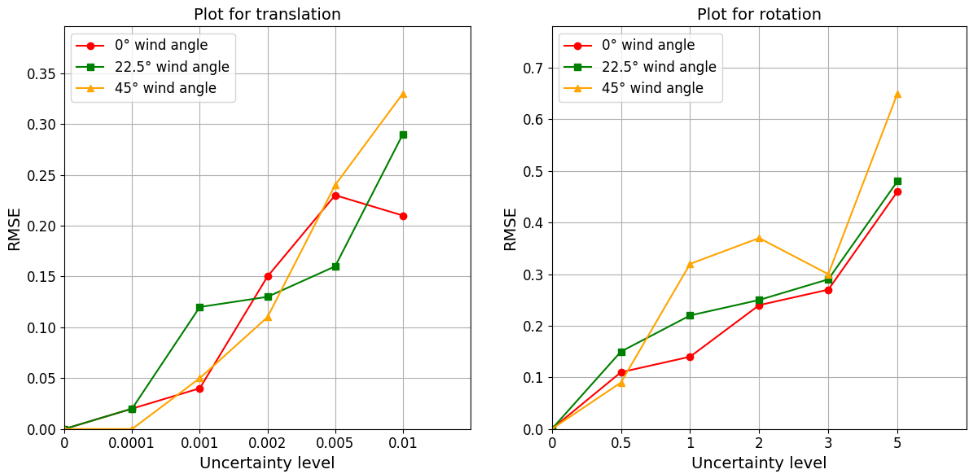
Table 5.4.: Uncertainties-MAPE

the cases of 5°, the impact on the coordinates is 0.0125m, which is higher in comparison with the $t_5 = 0.01$ m.

5.3.2. Box plots

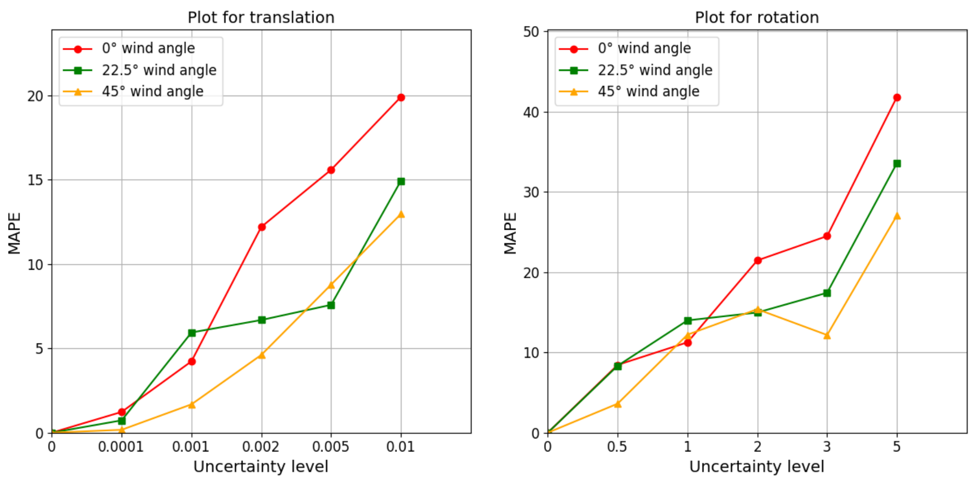
The next visualization technique that is applied is the box plot. Again, using Python, we created graphs consisting of our point IDs on the vertical axis, and the velocity differences from the base case, on the horizontal axis. The x-axis was customized to show every 10th point ID, for clarity, with the first point's ID to be 1 and the last one's to be 120. The total number of box plots is 6, for all three wind angle cases, both with rotational and translational uncertainty, and can be found in Appendix E. All plots were created vertically to fit on the pages. In comparison to the scatter plots that provide us with a more general trend and a statistical comparison, the box plots give us the possibility to identify which specific testing points have been mostly affected by the uncertainty, showing us the magnitude of each point's uncertainty as well as the deviation from the base case.

At first, the five deviations from the base case were calculated for the five cases with uncertainty by subtracting the five velocity values from the base case velocity. These were



(a) Translation (b) Rotation

Figure 5.4.: RMSE graphs



(a) Translation (b) Rotation

Figure 5.5.: MAPE graphs

calculated for all 120 probes. This resulted in a set of 5 differences for each testing point. The advantage of boxplots is that they allow us to observe the variation of these 5 values from our initial one, giving us a clear picture of which points were most affected by the uncertainty. Our five value differences perfectly summarize the box plot's essence based on the five-number summary. There is one box per point ID and each box consists of five numbers in increasing order: minimum, first quartile (Q1), median, third quartile (Q3), and maximum value. The red line in the middle represents our base case (u_0) and each box consists of these five values. A highlighted line within each box represents the median (middle value of the data). The boxes represent the data's Interquartile Range (IQR), the range between Q1 and Q3, and are the bottom and top limits, respectively. The vertical lines that extend from both ends of each box, with black color, connect the minimum and maximum values with the boxes and are called whiskers. Finally, the small, horizontal, black lines at the ends of the whiskers, representing the minimum and maximum values, are the caps [Chia, 2024]. Every box was filled using a color map of red intensity showing how much each set of values deviates from the base case (larger spread means more intense red). So the red intensity visually encoded the variability at each probe. The IQR was the measure of intensity chosen because it is resistant to outliers and ignores extreme values. It reflects how variable these deviations are for each probe, robustly, since it measures the spread of the middle 50% of our data. Also, the IQR is a solid statistical measure of spread, which is easy to interpret (box height).

Additionally, the IDs with the top ten spreads from the base case were found and marked in the box plots. For that reason, probe maps were created, allowing us to easily see the location of each of these ten points, on the xy plane, that present the greatest uncertainty. These plots are similar to the Figure 3.2 and these points are represented with yellow color and are labeled, thus indicating to us the areas that are most affected by each movement of our geometry. As a result, for each box plot, there is also a probe map. Below is presented the box plot with the translational uncertainties of the case with 0° wind angle accompanied by the corresponding probe map (Figure 5.6).

The box plot above shows us that there is a variety of points affected by uncertainty, with the majority of these points being between the last points. A clearer view is given from the probe's map where these are grouped in a certain area, which is the region behind the central building. This area is called the wake region and is characterized by high flow sensitivity, low velocity and high turbulence. In these zones, even small changes in the geometry and minor geometric uncertainty (like small translation or rotation) can lead to large differences in the flow behaviour and cause large variability in velocity at those points. As a result, there are large deviations from the base case, due to a possible change in the wake width. As we can see from the rest of the plots in the Appendix E, the points that are mostly affected are located in similar areas both for translation and rotation. These areas of influence, however, change as the wind direction angle changes and remains behind the buildings, opposite to the wind direction. This comes to reassure our initial observation that the uncertainty impacts mostly the areas behind the buildings, and especially the central one.

A clear disadvantage of this approach is the fact that our study area is limited between the buildings and, specifically, around the central building. Also, there is a limited number of points (120), which we focus on, instead of having a bigger sample and observing the general trend of a certain area. On the other hand, this gives us the possibility to become more precise regarding the measurements and the impact of uncertainty at these points, and computationally more efficient since the number of points is finite. These limitations come to be complemented by the last visualization method, which is presented in the next section.

5.3.3. Contour plots

The two previous ways of visualization were focused on the area around the central building. Our final visualization method is the contour plots, which are used to visualize the spatial distribution of the differences of certain variables between the two simulations (base case and case with uncertainty). This method will give us a more insightful observation of the uncertainty behaviour in our study area, since more points are considered. The variables

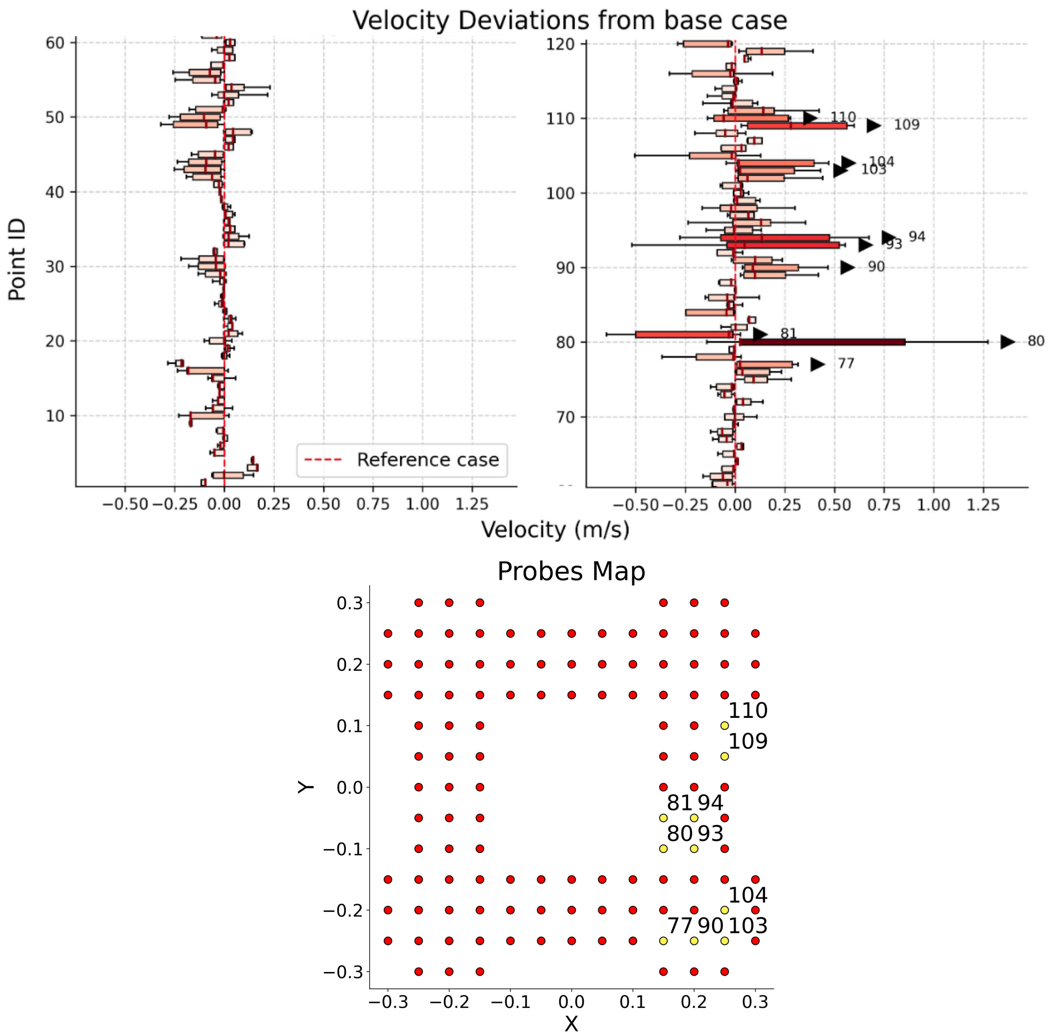
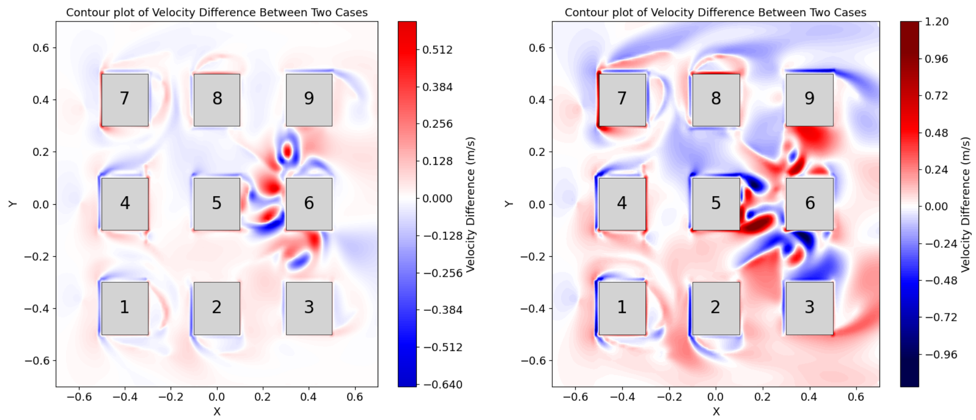


Figure 5.6.: Box plot and probe map of translated cases (0° wind angle)

that we will focus on are the velocity magnitude and TKE. The purpose is the creation of contour plots in the form of heatmaps and to visualize certain areas of our case study region that present the highest variance regarding the two variables of interest. For the creation of the heatmaps, we used Python and the total number of contour plots is 60 (30 for velocity U, 30 for TKE, consisting of one for each uncertainty case of translation and rotation). The rest of the plots can be found in the Appendix F.

What we see in these plots is a top view of our study area, and with light grey are represented our cubes alongside their enumeration (Figure 3.6b). At first, through `Paraview`, we used the tools, first `Clip` and then `Slice`, to isolate the area of interest that contains our geometry, with a buffer zone of 0.2m around our geometry. We exported the contained data of the final slices in Comma-separated value (CSV) format. Ultimately, we had 33 CSV files (3 for the base cases, 15 for the case with rotational uncertainty and 15 for the cases with translational uncertainty). Then we stored our CSV data in dataframes and we used linear interpolation to estimate what the velocity magnitude and TKE from the uncertain case would be at the corresponding points of the base case. Eventually, we computed the velocity and TKE differences. For the contour plots, we used the $x_{\min}, x_{\max}, y_{\min}, y_{\max}$ of the base case and through `linspace` and `meshgrid`, we created a new grid with dimensions 300×300. On this grid, we interpolated linearly the velocity and TKE difference values and a contour plot with a form of heatmap was produced.

For the representation of the variable differences, diverging color maps were used, and the legend of the colormap scale is presented on the right side of the plot. It is centered around zero, with the value range to adapt to each difference map, but the color intensity remains the same in order to make obvious the difference between various uncertainty cases. We subtracted the U_{mag} and TKE values of the base case from the cases with uncertainty. This means that for positive differences, the U_{mag} and TKE increased in comparison to the base case. For negative differences, the U_{mag} and TKE decreased in comparison to the base case. The color map for U_{mag} has red-blue colors and for TKE has green-purple colors. For the cases with U_{mag} differences, the red areas represent regions of acceleration (velocity increased) and the blue areas represent regions of deceleration (velocity dropped). For the cases with TKE differences, the emphasis is given in the green areas, which represent regions where TKE increased and the purple areas, which represent regions where TKE decreased. Below are given four representative cases with 0° wind angle and translations T3 and T4. The first two represent U_{mag} difference (Figure 5.7) and the second two represent TKE difference (Figure 5.8).

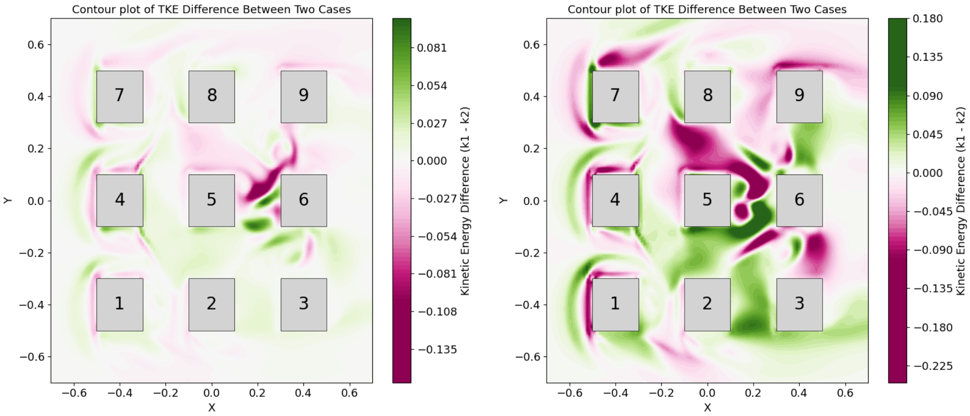


(a) T3 = 0.002m

(b) T4 = 0.005m

Figure 5.7.: U_{mag} difference contour plots

After interpreting these plots, they reveal several information about the CFD simulations. They visualize the differences in U_{mag} and TKE between two simulations—the one as a base case and the other as a comparison (perturbed geometry), emphasizing the regions of the domain where the two variables decreased or increased. Through this method, we could easily highlight the areas where changes in flow occur due to the uncertainty that we imported into our model. These plots provide intuitive insight into the flow field sensitivity to simulation setup changes and help identify local effects caused by the geometry shifts. Also, it becomes obvious how sensitive the U_{mag} and TKE are to small geometric changes.



(a) T3 = 0.002m

(b) T4 = 0.005m

Figure 5.8.: TKE difference contour plots

Essentially, the plots act as a map of sensitivity, helping us identify zones affected by the geometric perturbations. Some key aspects to focus on are the following:

1. Near the building surfaces and their corners are typically areas with flow separation and are the most sensitive ones to geometric changes.
2. In the wake region, especially downstream of our central cube, we see that even small geometric shifts can cause large changes in velocity magnitude due to flow instability.
3. The small positional differences can amplify sharp gradients for transitions from high to low velocity zones.
4. Zones with high absolute differences (shown as red, blue, green, or purple) indicate where geometry shifts mattered the most.

Based on these observations, it is worth noting that the U_{mag} and TKE follow a similar trend. So, for the cases of rotation, the regions that present the greatest decrease in our variables are the ones towards the direction of rotation (blue and purple areas), while it is clear that the regions that show the greatest increase are those that are opposite to the direction of rotation (red and green areas). Correspondingly, for the cases of translation, the areas that present the greatest decrease in both U_{mag} and TKE are the areas to which our geometry was moved (blue and purple areas), while conversely, the regions from which the buildings were moved away show the greatest increase (red and green areas). These observations can be easily elaborated through the extreme uncertainty cases t_5 and r_5, which present the highest difference values. Moreover, by looking at each set of 5 uncertainties per wind angle, there is a gradually increasing perturbation, which indicates a clear dependency between the level of uncertainty and the magnitude of the differences. Finally, some of the details on the plots that it is not necessary to focus on are the undisturbed upstream flow, the regions with low U_{mag} and TKE where the differences are zero and the grid edges, where the interpolation might have introduced some artifacts.

Besides the qualitative observations that could be made by looking at the plots, some statistical indicators for quantitative evaluation of our results were calculated through Python. Specifically, the median absolute difference and the RMSE for U_{mag} and TKE differences were computed. These values are included in the following tables. Each table contains the statistics of translation and rotation for the 3 different wind angles ($0°, 22.5°, 45°$). Finally, we have four tables: RMSE-U_{mag}, RMSE-TKE, median-U_{mag}, median-TKE (Tables 5.5, 5.6, 5.7, 5.8). The colour intensity of each cell in these tables was decided according to certain classes, starting from lower and reaching higher values. These classes are for:

- Table 5.5: $[0,\ 0.01), [0.01,\ 0.1), [0.1,\ 0.2), [0.2,\ 0.3), [0.3,\ 0.4), [0.4,\ +\infty) m/s$,
- Table 5.6: $[0,\ 0.01), [0.01,\ 0.02), [0.02,\ 0.03), [0.03,\ 0.04), [0.04,\ 0.05), [0.05,\ +\infty) m^2/s^2$,
- Table 5.7: $[0,\ 0.01), [0.01,\ 0.03), [0.03,\ 0.05), [0.05,\ 0.07), [0.07,\ 0.1), [0.1,\ +\infty) m/s$,
- Table 5.8: $[0,\ 0.005), [0.005,\ 0.01), [0.01,\ 0.015), [0.015,\ 0.02), [0.02,\ 0.03), [0.03,\ +\infty) m^2/s^2$

Uncertainty	Translation-RMSE (m/s)			**Uncertainty**	Rotation-RMSE (m/s)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	0.0089	0.0273	0.0054	**R1**	0.0620	0.0951	0.1291
T2	0.0255	0.0935	0.0626	**R2**	0.0856	0.1405	0.2402
T3	0.0719	0.1200	0.1121	**R3**	0.1873	0.1887	0.2897
T4	0.1600	0.1960	0.2323	**R4**	0.2175	0.2593	0.2411
T5	0.2376	0.3421	0.3860	**R5**	0.3864	0.3734	0.4933

Table 5.5.: RMSE of U_{mag} difference from the contour plots

Uncertainty	Translation-RMSE (m^2/s^2)			**Uncertainty**	Rotation-RMSE (m^2/s^2)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	0.0019	0.0065	0.0012	**R1**	0.0128	0.0255	0.0209
T2	0.0061	0.0243	0.0156	**R2**	0.0229	0.0366	0.0485
T3	0.0154	0.0291	0.0272	**R3**	0.0468	0.0398	0.0575
T4	0.0375	0.0435	0.0426	**R4**	0.0623	0.0592	0.0447
T5	0.0522	0.0618	0.0517	**R5**	0.1156	0.0750	0.0970

Table 5.6.: RMSE of TKE difference from the contour plots

The tables present a quantitative summary of flow sensitivity, with RMSE capturing the overall error magnitude, but being more influenced by outliers compared to the median, which is more robust to extreme values. Across nearly all the combinations, both our metrics (RMSE, median) increase as the uncertainty grows, indicating that this trend is consistent, since the higher the uncertainty, the larger the flow differences. Some general observations that could be made are:

Uncertainty	Translation-median (m/s)			Uncertainty	Rotation-median (m/s)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	0.0018	0.0025	0.0009	**R1**	0.0230	0.0294	0.0265
T2	0.0088	0.0296	0.0179	**R2**	0.0382	0.0443	0.0754
T3	0.0178	0.0372	0.0348	**R3**	0.0765	0.0550	0.0844
T4	0.0639	0.0592	0.0632	**R4**	0.0934	0.0809	0.0650
T5	0.0930	0.0884	0.0882	**R5**	0.1707	0.1329	0.1551

Table 5.7.: Median U_{mag} difference from the contour plots

Uncertainty	Translation-median (m^2/s^2)			Uncertainty	Rotation-median (m^2/s^2)		
	0°	**22.5°**	**45°**		**0°**	**22.5°**	**45°**
T1	0.0003	0.0006	0.0002	**R1**	0.0054	0.0072	0.0053
T2	0.0022	0.0059	0.0032	**R2**	0.0100	0.0108	0.0111
T3	0.0053	0.0083	0.0063	**R3**	0.0204	0.0147	0.0168
T4	0.0145	0.0146	0.0101	**R4**	0.0232	0.0185	0.0161
T5	0.0229	0.0196	0.0141	**R5**	0.0405	0.0338	0.0347

Table 5.8.: Median TKE difference from the contour plots

- The translation cases generally present more localized but intense differences, particularly near building edges (near the walls) and behind the buildings (wake zones).
- The rotation cases tend to affect broader regions, especially at higher wind angles, due to realignment of the flow separation paths, affecting entirely the distribution of our variables around the buildings.
- Based on these, we also observed that for U_{mag}, the translation presents a sharper increase in RMSE at early uncertainty levels.
- On the other hand, for TKE, rotation dominates, due to the model's sensitivity to flow deflection.

Another aspect that we could approach our metrics is the dependency on the wind angle, which plays a major role in amplifying or suppressing the impact of uncertainty. In many categories, the 45° wind angle shows the largest RMSE and median values, especially for rotation. This makes sense, since at 45°, wind interacts with the buildings diagonally, making the flow more sensitive to small geometry misalignments. Finally, these metrics validate that the impact of uncertainty is also direction-dependent.

6. Conclusions

Through this thesis, we investigated the impact of geometric uncertainty on the accuracy of the CFD simulations in an urban case scenario. The main focus was on perturbation of the geometry through translation, rotation and different wind angles, and the study addressed a significant aspect of CFD model fidelity. The research was conducted through a series of simulations applied in OpenFOAM software, using as reference cases the validated cases, through the wind tunnel data (AIJ Case C). Following this validation, the study introduced a structured methodology for simulating geometric uncertainty in the CFD model. This was achieved by systematically perturbing each building separately through controlled translation and rotation, generating 30 uncertainty scenarios across three wind directions. Each of these scenarios was simulated independently, providing, as a result, a rich dataset from which localized flow behavior was assessed. At first, the 120 probes, distributed around the central building, were used. Both scatter plots and box plots were employed, allowing for a detailed examination of how velocity magnitude U_{mag} deviates from the base case at each probe location. These visualizations revealed the spatial variability in a qualitative way. Also, the RMSE and MAPE were used to reveal the statistical spread, which was introduced by geometric uncertainty. Certain regions were highlighted due to high sensitivity. Afterwards, through the contour plots, a general trend was observed regarding the entire region around the buildings. Metrics such as U_{mag}, TKE, RMSE, and median were used to evaluate the influence of perturbations on flow predictions. A more detailed approach to these three ways to present the results is given below:

1. Scatter Plots (Probe-Based): They show individual comparisons between the base case and each uncertainty case, at the probes' locations. With 120 probes distributed around the central building, scatter plots allowed us to see the variability introduced at each probe due to translation or rotation. Some of the advantages are:

   - High resolution at specific locations
   - Clear visibility of the outliers

   Some of the disadvantages are:

   - Lack of spatial context. We could not see where the probes are located
   - It is hard to identify flow patterns or spatial trends because of their sparsity

2. Box Plots (Probe-Based): They present to us a statistical distribution (min, Q1, median, Q3, max) of U_{mag} across all 30 simulations for each probe. We used them to summarize how each probe behaves across the uncertainty scenarios. By using the probes' map it was easier to spot which probes consistently show high deviation, which ones are stable and where they are located. Some of the advantages of the box plots are:

   - Compact and highly informative
   - Highlights central tendency (median) and spread (IQR)

- Excellent for comparing sensitivity across locations

Some of the disadvantages are:

- They still lack spatial context, without the probes' map
- They do not show continuous flow behavior

3. Contour Plots (Grid-Based): They reveal the spatial distribution of U_{mag} and TKE across the entire domain. The differences are computed for every cell through interpolation. We used them to visualize U_{mag} and TKE differences between the base case and uncertainty cases. Finally, the RMSE and median were computed. Some of the advantages are:

   - The detailed spatial context, since it is obvious where the flow is sensitive
   - It is easier to detect wake shifts, recirculation changes
   - They are proper for understanding global flow impacts

   Some of the disadvantages are:

   - They can be affected by interpolation artifacts near geometry
   - We cannot be based only on qualitative observations of local color changes. The statistical analysis is necessary

The novelty of this thesis is the integrated approach to quantifying geometric uncertainty by combining spatial qualitative methods (scatter plots, box plots, contour plots) with statistical performance metrics (RMSE, MAPE, median) over multiple perturbed cases. Unlike other studies that focus primarily on mesh or turbulence model sensitivity, this work isolates the effects of building footprint inaccuracy (translation, rotation) and investigates how it influences CFD results under specific wind conditions and several repetitive simulations. By applying these uncertainties across three wind angles and evaluating their effects, this thesis provides an approach for assessing geometric robustness in urban CFD modeling. The findings not only confirm the significance of even minor geometric deviations but also emphasize the importance of accounting for directional flow sensitivity in uncertainty-related simulation practices.

6.1. Answer to research questions

- ***Validation of our CFD cases compared to the wind tunnel data***

  The r-values indicate a good match with the experimental data. The 2H case presents the best results and because of that, it was chosen to perform our further analysis. The experimental values are slightly higher than the simulation values. This relative offset that is observed in all three cases is probably because of the z_0 value, which possibly underestimates the wind speed. So, we could tell that the comparison with the wind tunnel data is reasonably accurate, allowing us to proceed with our analysis.

- ***What is the impact of geometrical uncertainties derived from building translation?***

  Translational perturbations primarily impact the local flow field near building surfaces, corners, and downstream wake regions. Small shifts in position lead to changes in blockage ratio, local acceleration, and wake alteration. The effect is highly localized, resulting in sharp spikes in velocity differences, particularly around sharp edges or where flow separation occurs. However, the influence decays downstream unless the translation significantly alters the aerodynamic alignment with the flow.

- ***What is the impact of geometrical uncertainties derived from building rotation?***

  Rotational perturbations produce more spatially distributed changes in the flow field. By altering the orientation of buildings relative to the incoming wind, rotation affects the pressure distribution across the buildings' faces, modifies the separation areas, and changes the shape and size of the wake regions. Unlike translation, rotation has a clearer effect on the flow patterns. This makes rotational uncertainty highly influential.

- ***How do these uncertainties affect the model as we modify the incoming wind direction?***

  The impact of both translation and rotation uncertainties becomes more obvious as the wind direction shifts away from the orthogonal 0° case. At different wind angles e.g. $22.5^\circ, 45^\circ$, buildings go through asymmetric loading and more complex flows, making the model more sensitive to small geometric discrepancies. In particular, rotational perturbations combined with oblique wind can bring significant divergence in velocity and TKE, as seen in the elevated RMSE and median values.

Finally, the abovementioned findings underscore that geometric uncertainties should not be neglected, especially in complex urban building environments where the geometry is characterized by higher complexity. Their effects on CFD accuracy are 1) direction-dependent, 2) localized for translation, and 3) spatially broader for rotation, and they are combined under various wind directions.

6.2. Limitations

A key limitation of this study is that the conclusions of this research are based on a single reference case (AIJ Case C) and a specific validation setup. While this case is well-established and widely accepted, its relatively simple layout, with a very fixed building composition, uniform building heights and spacing, may not reflect the real-life complexity of the urban configurations. As such, the sensitivity trends observed here may vary when applied to more divergent cityscapes or cases with complex terrain, different LoDs and varying building heights.

Another limitation that was noticed during the elaboration of the project is the limited scope of the uncertainty magnitude. The range of perturbation magnitudes applied in this study does not completely cover the entire variety of possible geometric uncertainties. Translation distances and rotation angles were discrete, based on assumed uncertainty thresholds. This stepwise approach may miss critical thresholds where flow behavior presents non-linear sensitivity, or where multiple small perturbations accumulate and produce disproportionate effects.

6.3. Recommendations for future work and improvements

In summary, accurate building footprints are fundamental for reliable urban CFD modeling and based on the current study, certain recommendations for further improvement could be applied. Firstly, the application of this methodology to a more realistic urban geometry is one of them. Future studies could apply the proposed uncertainty framework to more complex urban structures. The geometry could contain irregular building shapes, a variety of different building heights, and non-uniform spacing between the buildings. Through this recommendation is stated the importance of checking the generalizability of the observed sensitivity patterns.

Another recommendation could be the introduction and combination with other uncertainty sources. This thesis focuses on geometric uncertainty, but a factor that could be considered is uncertainty regarding the inflow conditions. In addition, another uncertainty scheme that could be introduced involves different uncertainty values and directions per vertex and not just per building. A combined sensitivity analysis would help prioritize which uncertainties dominate under different urban conditions.

Finally, also, two new interesting research questions arise from the current thesis. The first one is if there is a critical threshold beyond which geometric uncertainty leads to nonlinear flow divergence. The current results imply that the higher the perturbation, with larger the error, but a deeper investigation could identify nonlinear thresholds where minor geometry perturbations suddenly cause major changes in the results. The second question is how the combination of translation and rotation interacts. This thesis isolated the two effects, which was necessary for clarity. However, combined perturbations may interact in non-additive ways. The more in-depth study of these interactions could offer a more complete image regarding this combination of uncertainties.

A. Reproducibility self-assessment

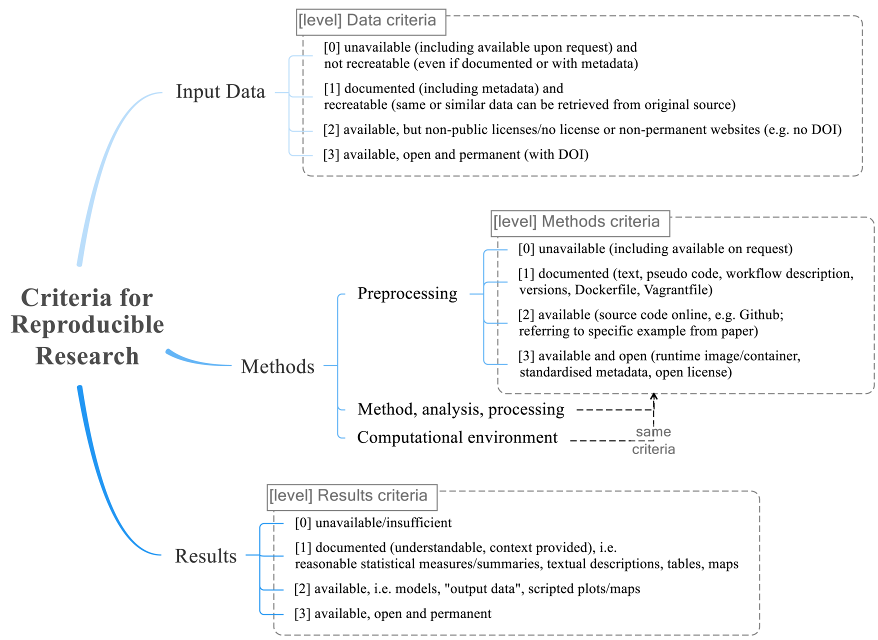


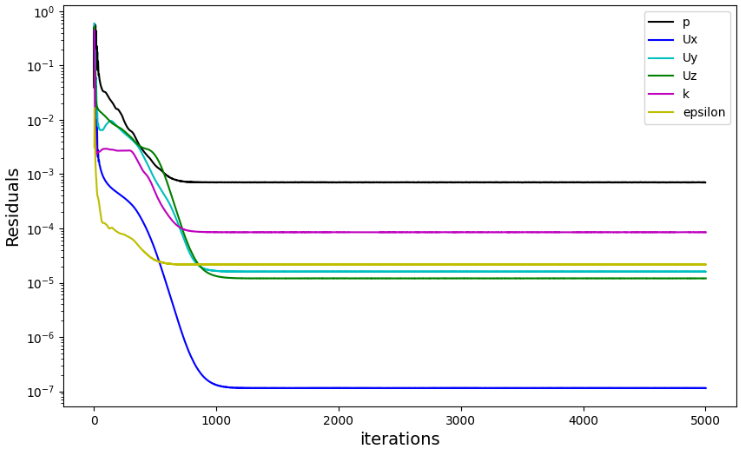
Figure A.1.: Reproducibility criteria to be assessed.

Evaluation for the 5 criteria (giving 0/1/2/3 for each):

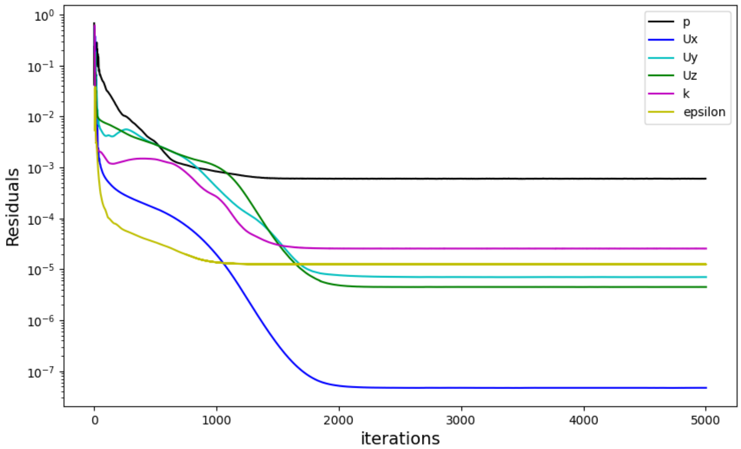
Criteria	Grade	Justification
Input data	3	Can be found in www.aij.or.jp
Preprocessing	3	Available on GitHub
Methods	3	Available on GitHub
Computational environment	2	Open source software was used, with connectivity to Gilfoyle server for higher computational power
Results	2	The cases consist of large files occupying approximately 0.5 TB of total space. Thus, they are not available on GitHub, but the CFD cases and setup file can be found there. The plots are available in this report

Table A.1.: Reproducibility criteria grading for this study. (Project's Github page: `https://github.com/tudelft3d-theses/2023-Chontos`)

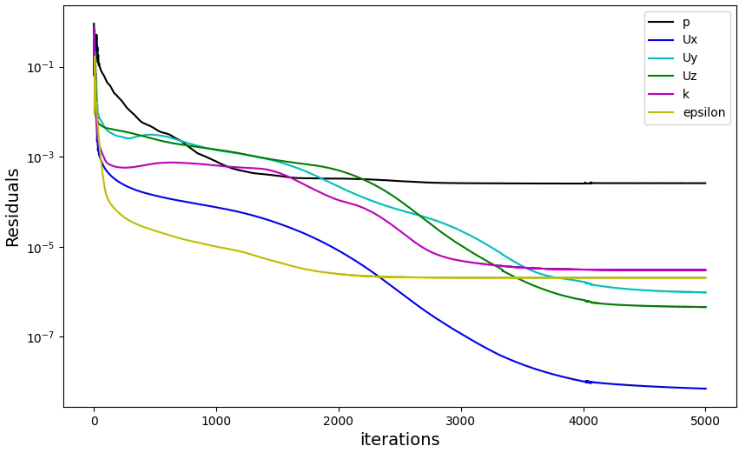
B. Residuals



(a) Coarse Mesh

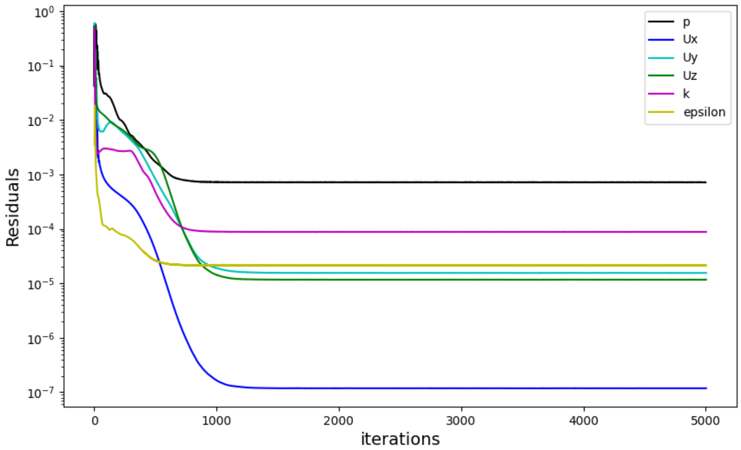


(b) Medium Mesh

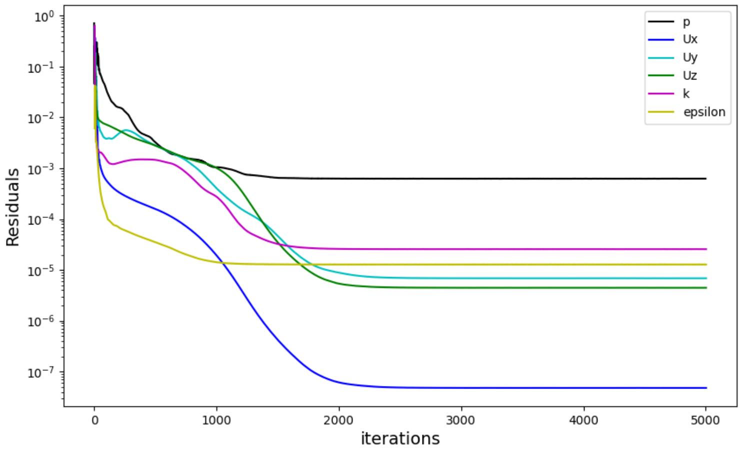


(c) Fine Mesh

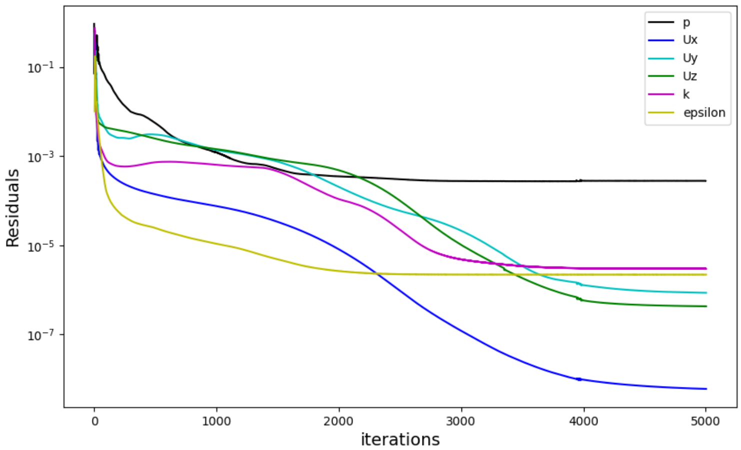
Figure B.1.: Residuals: Case 0H



(a) Coarse Mesh

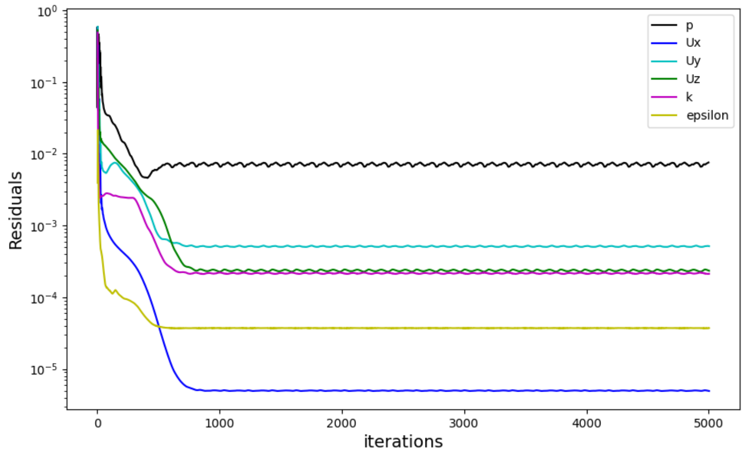


(b) Medium Mesh

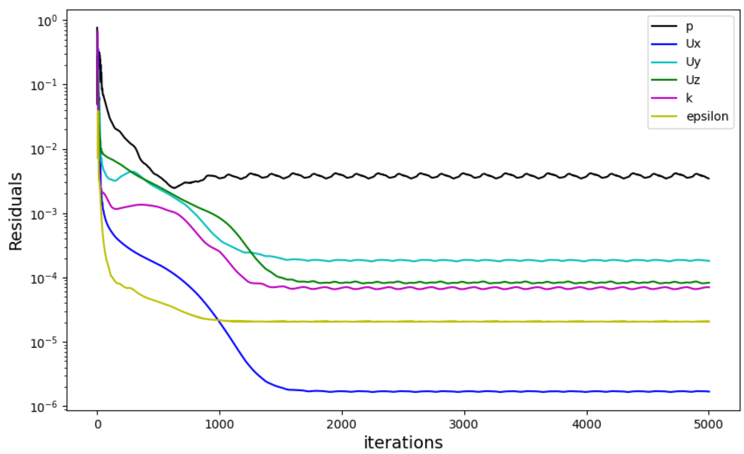


(c) Fine Mesh

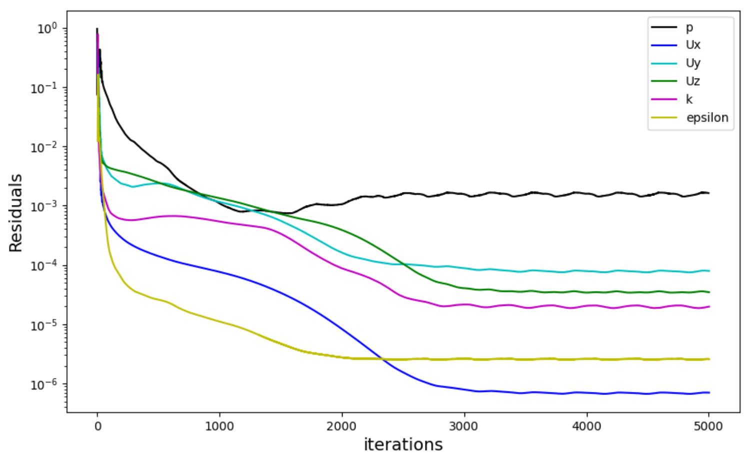
Figure B.2.: Residuals: Case 1H



(a) Coarse Mesh



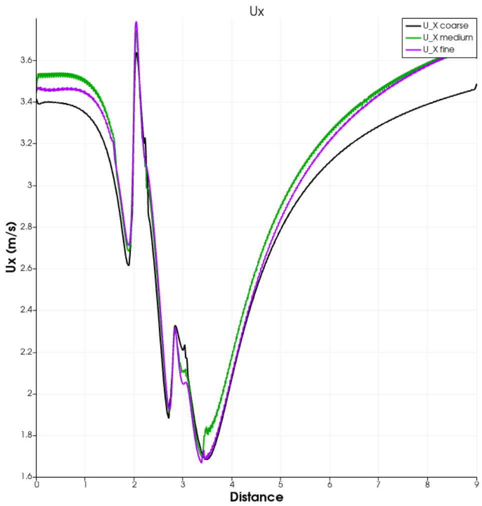
(b) Medium Mesh



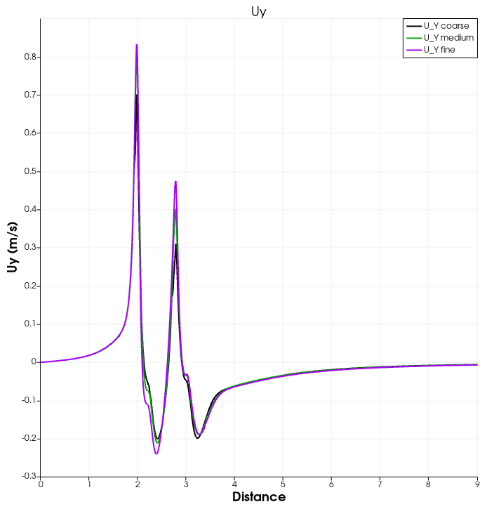
(c) Fine Mesh

Figure B.3.: Residuals: Case 2H

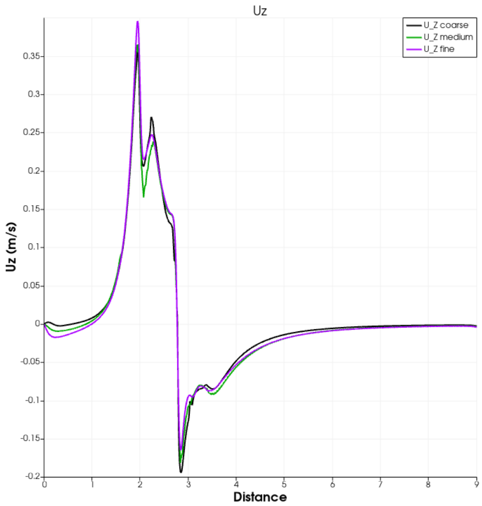
C. Field plots



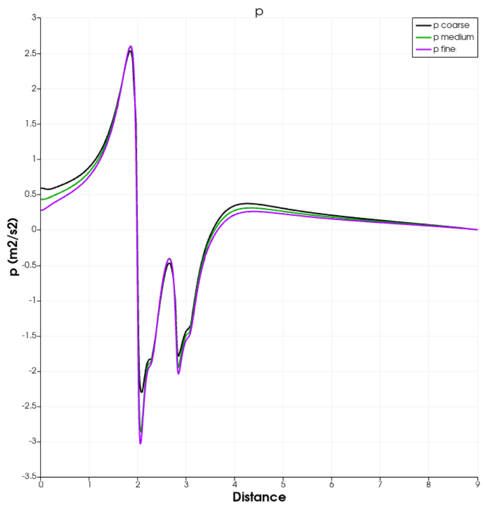
(a) Ux plot for the three meshes



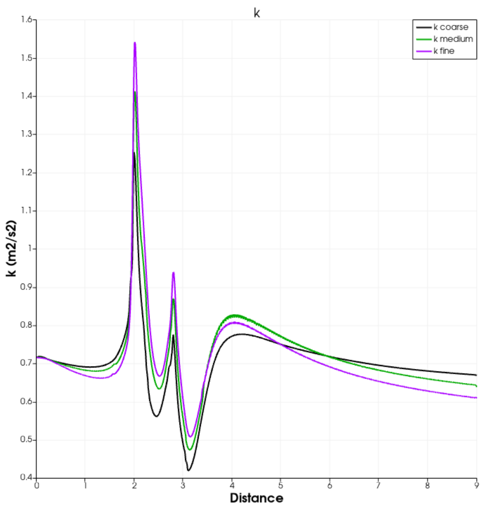
(b) Uy plot for the three meshes



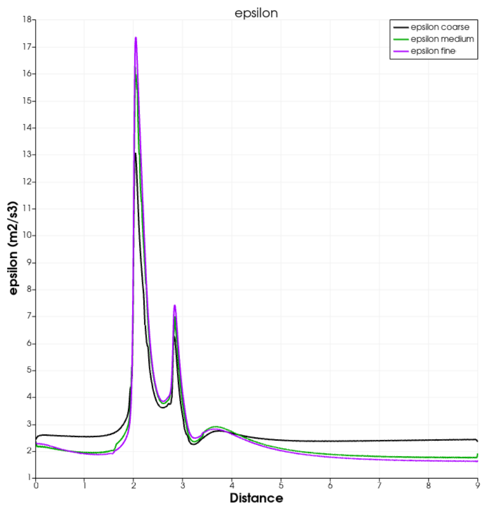
(c) Uz plot for the three meshes



(d) P plot for the three meshes

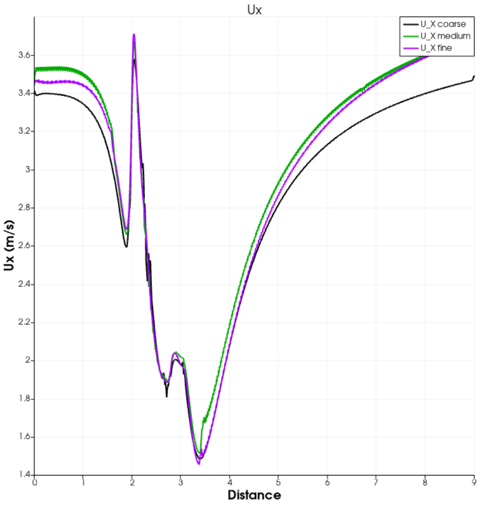


(e) K plot for the three meshes

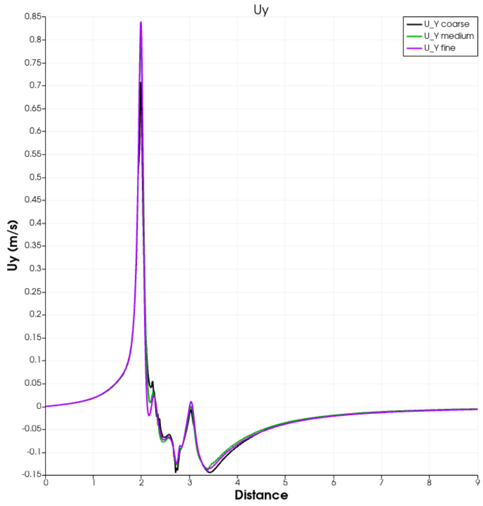


(f) ϵ plot for the three meshes

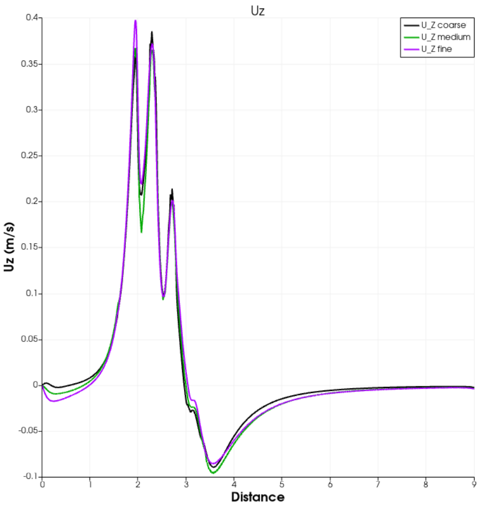
Figure C.1.: Field Plots: Case 0H



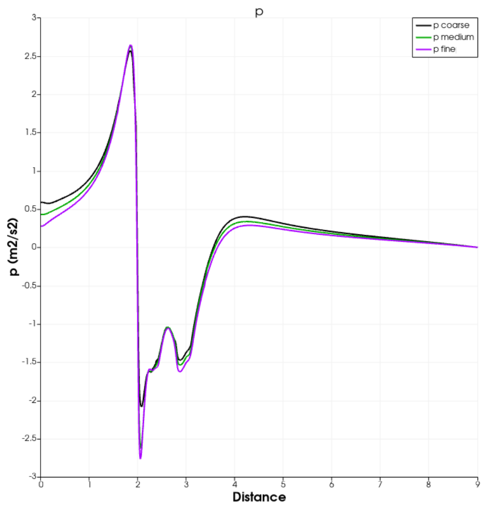
(a) Ux plot for the three meshes



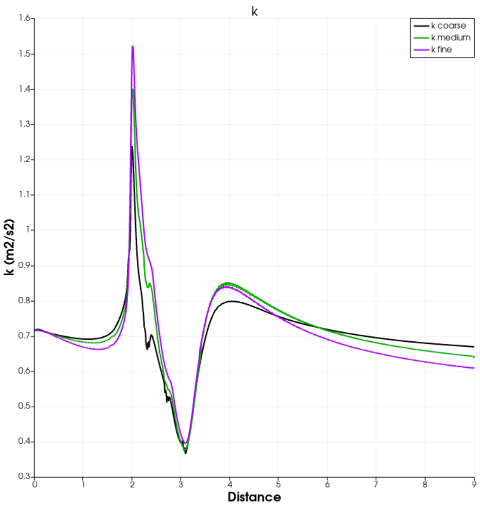
(b) Uy plot for the three meshes



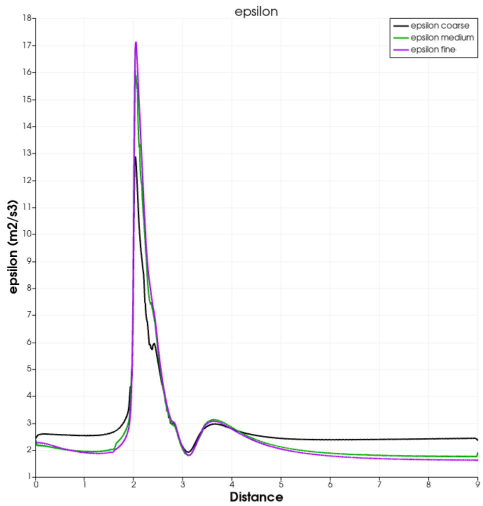
(c) Uz plot for the three meshes



(d) P plot for the three meshes

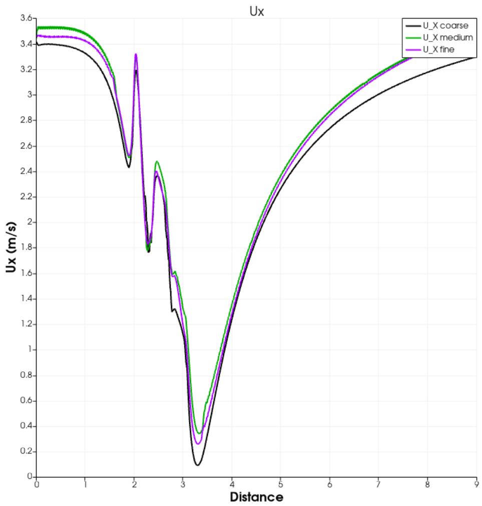


(e) K plot for the three meshes

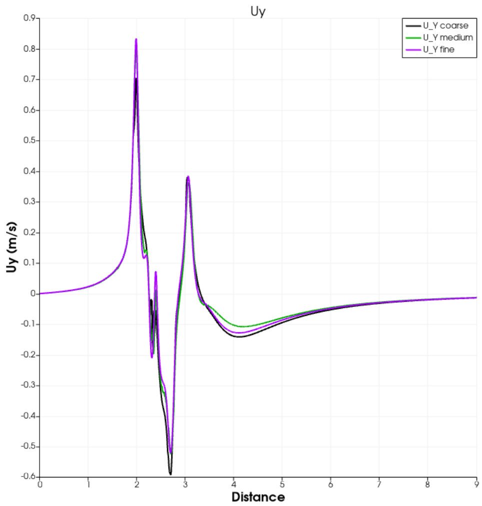


(f) ϵ plot for the three meshes

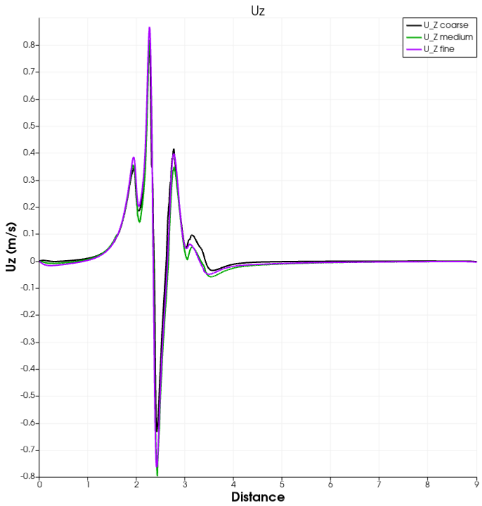
Figure C.2.: Field Plots: Case 1H



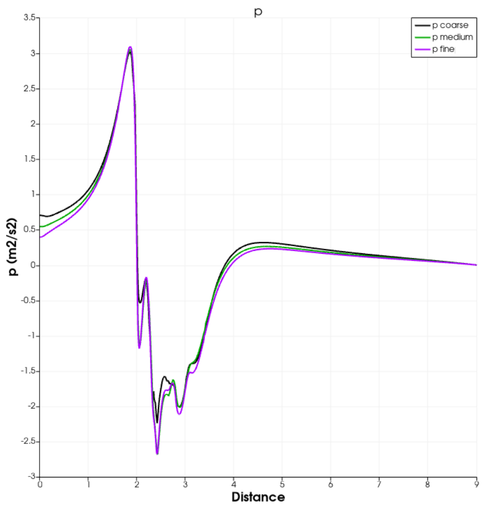
(a) Ux plot for the three meshes



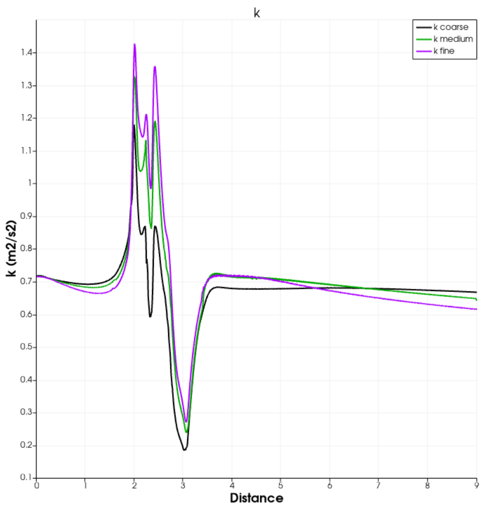
(b) Uy plot for the three meshes



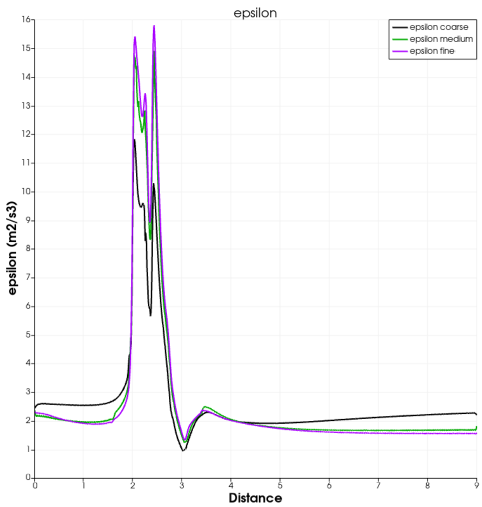
(c) Uz plot for the three meshes



(d) P plot for the three meshes



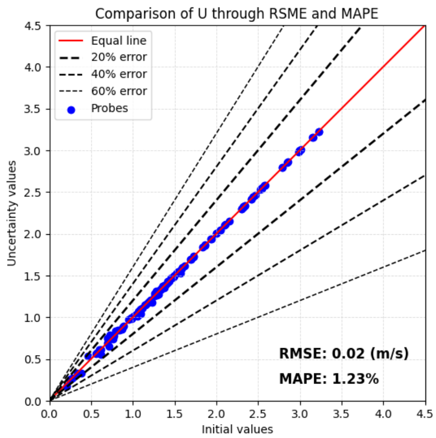
(e) K plot for the three meshes



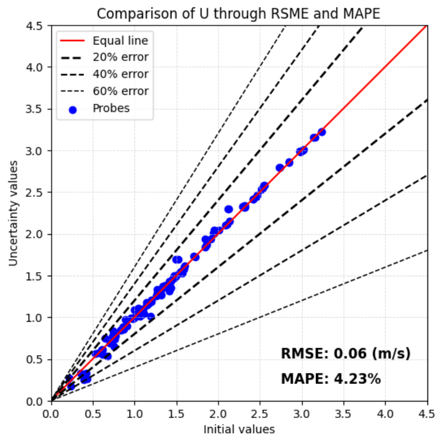
(f) ϵ plot for the three meshes

Figure C.3.: Field Plots: Case 2H

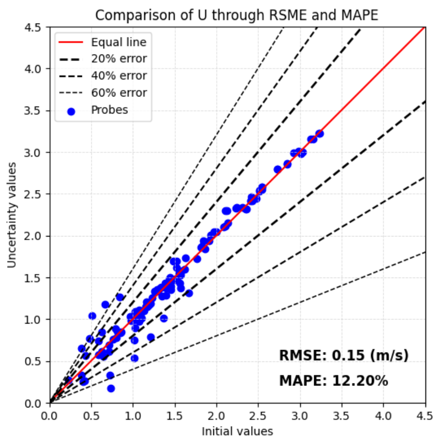
D. Scatter plots



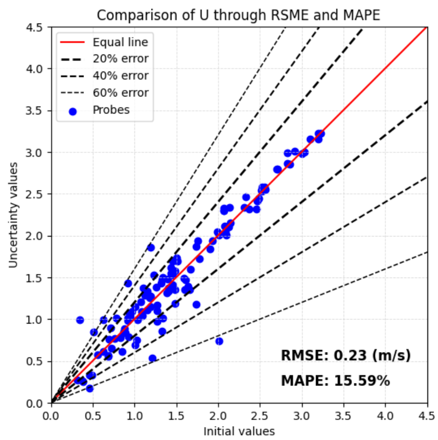
(a) T1 = 0.0001m



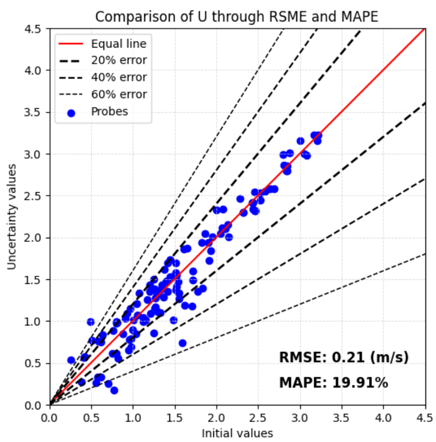
(b) T2 = 0.001m



(c) T3 = 0.002m

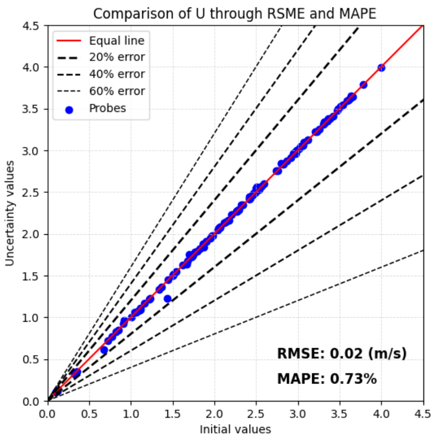


(d) T4 = 0.005m

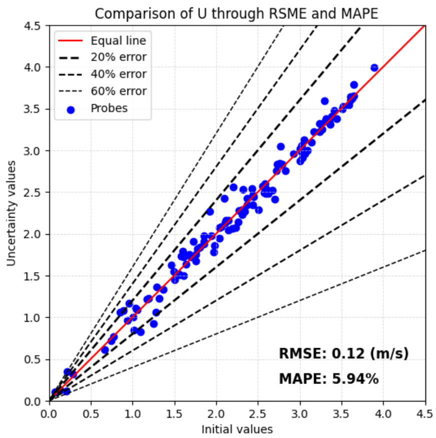


(e) T5 = 0.01m

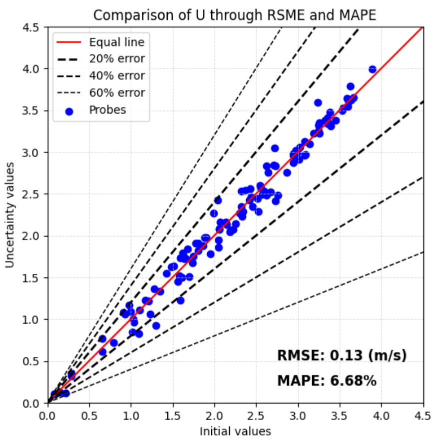
Figure D.1.: Translation scatter plots: Case 0°



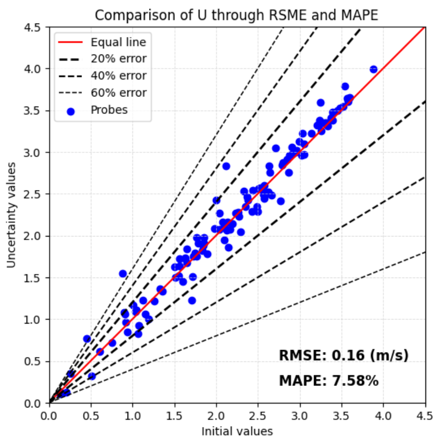
(a) T1 = 0.0001m



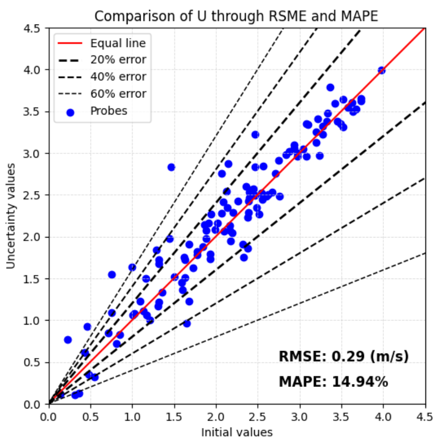
(b) T2 = 0.001m



(c) T3 = 0.002m

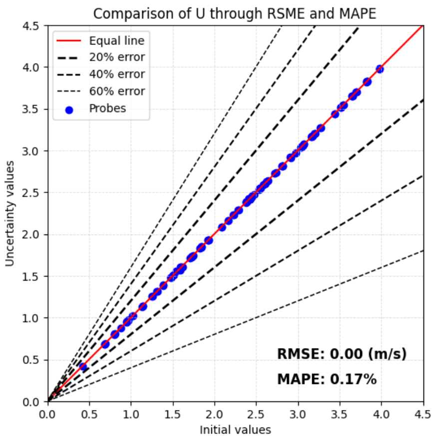


(d) T4 = 0.005m

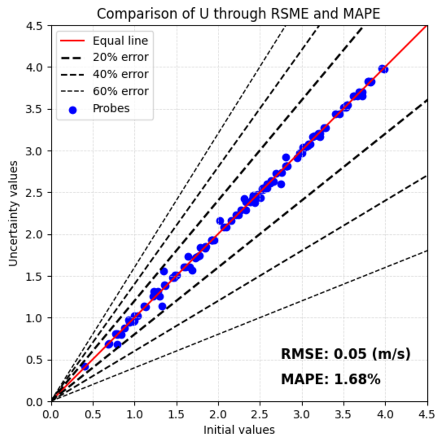


(e) T5 = 0.01m

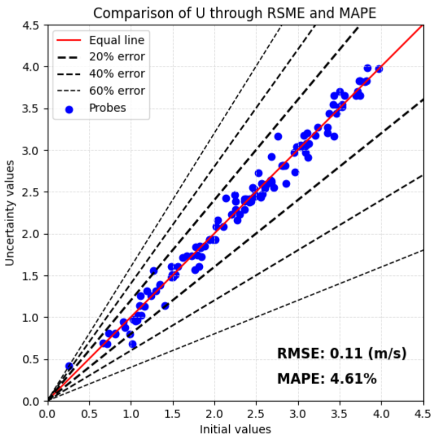
Figure D.2.: Translation scatter plots: Case 22.5°



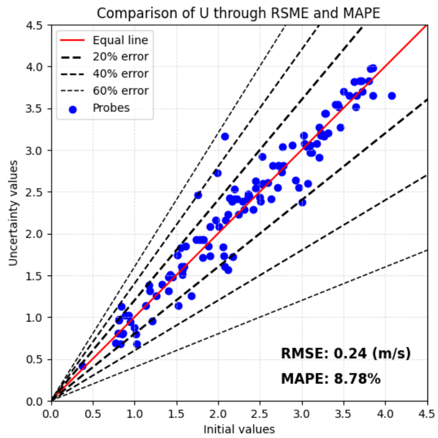
(a) T1 = 0.0001m



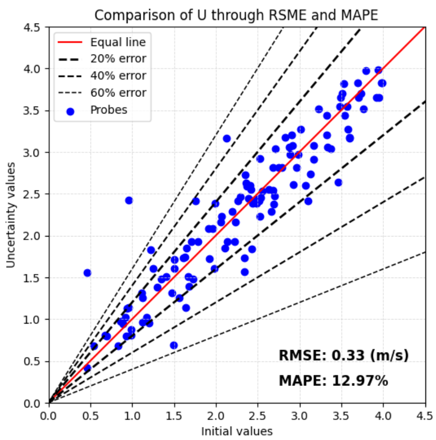
(b) T2 = 0.001m



(c) T3 = 0.002m

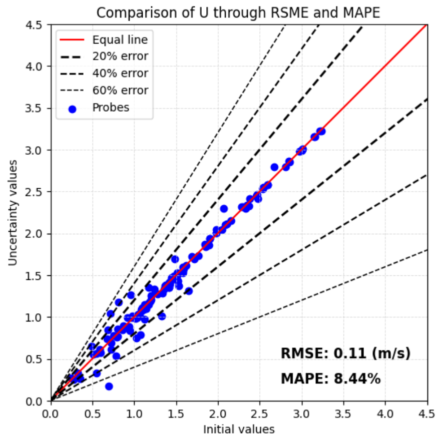


(d) T4 = 0.05m

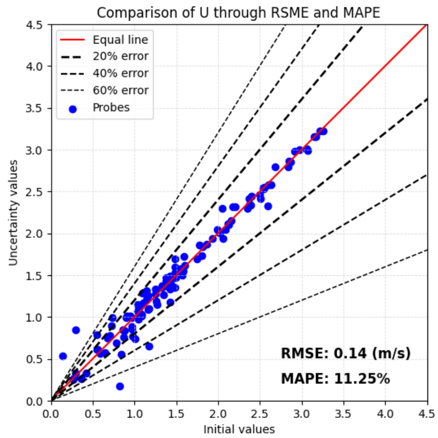


(e) T5 = 0.01m

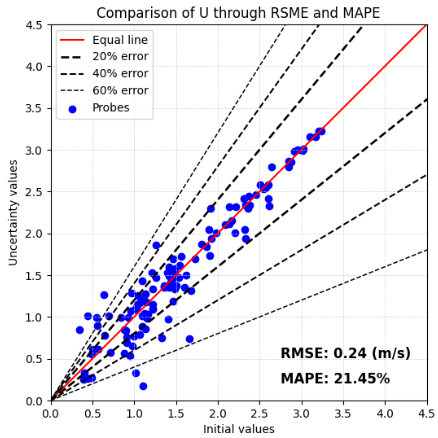
Figure D.3.: Translation scatter plots: Case 45°



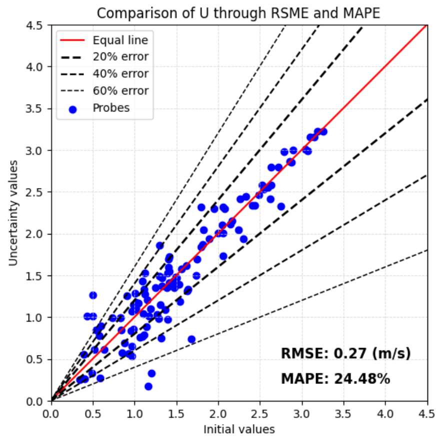
(a) R1 = 0.5°



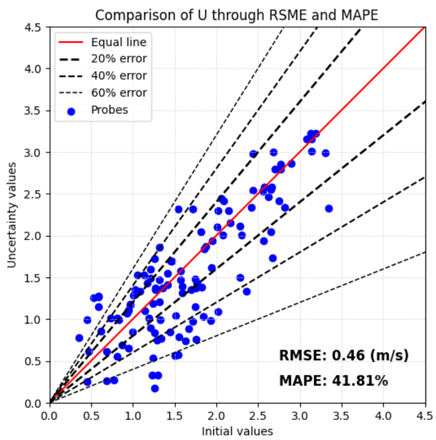
(b) R2 = 1°



(c) R3 = 2°

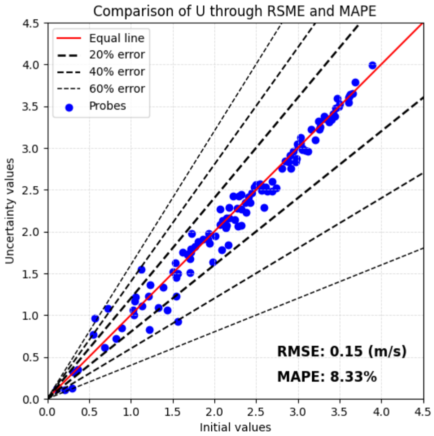


(d) R4 = 3°

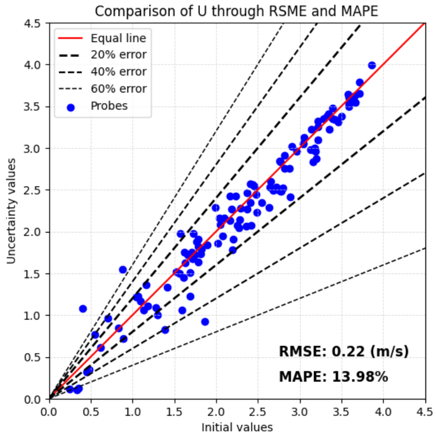


(e) R5 = 5°

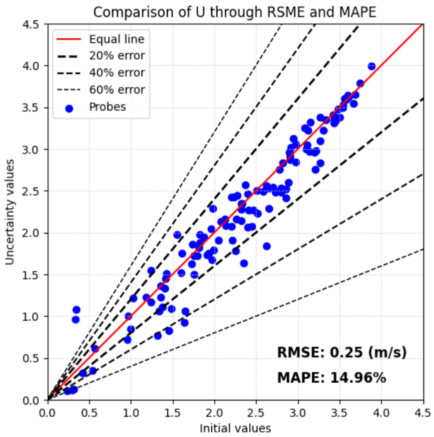
Figure D.4.: Rotation scatter plots: Case 0°



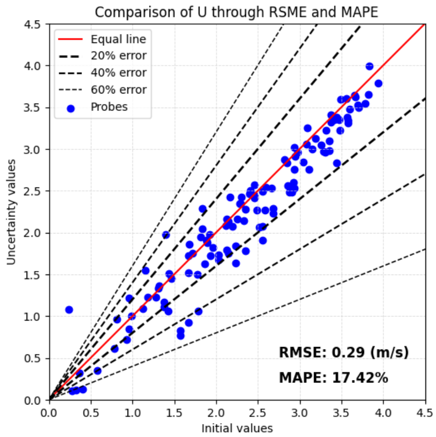
(a) R1 = 0.5°



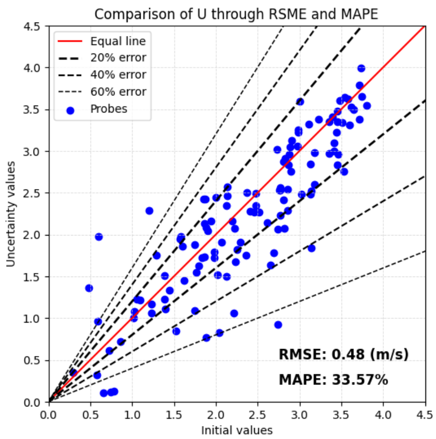
(b) R2 = 1°



(c) R3 = 2°

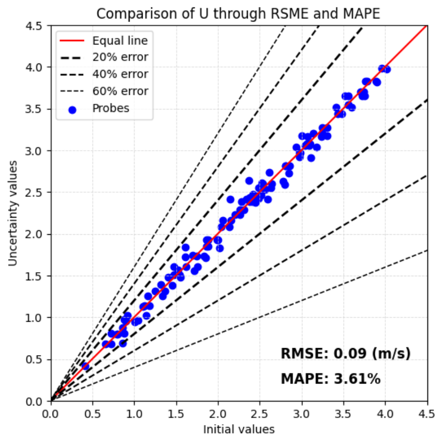


(d) R4 = 3°

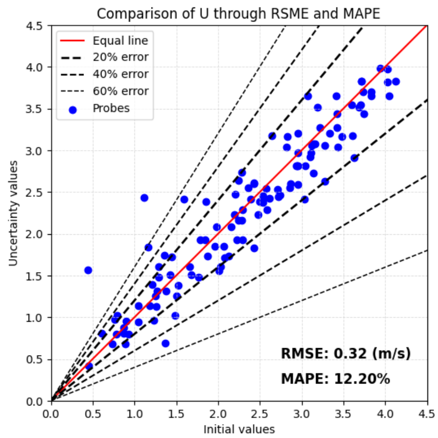


(e) R5 = 5°

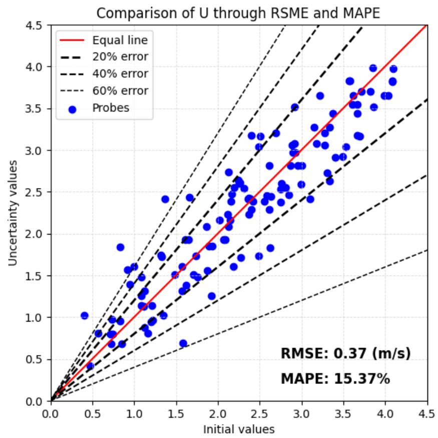
Figure D.5.: Rotation scatter plots: Case 22.5°



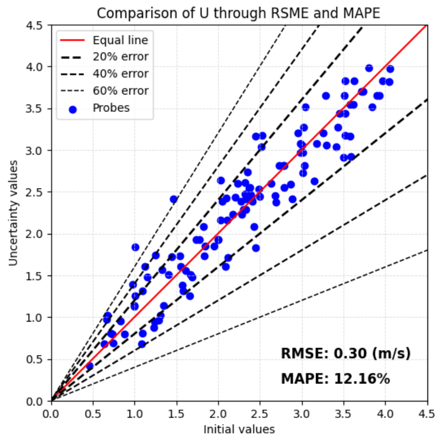
(a) R1 = 0.5°



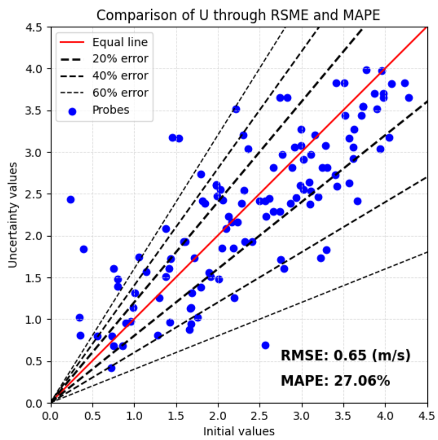
(b) R2 = 1°



(c) R3 = 2°



(d) R4 = 3°



(e) R5 = 5°

Figure D.6.: Rotation scatter plots: Case 45°

E. Box plots

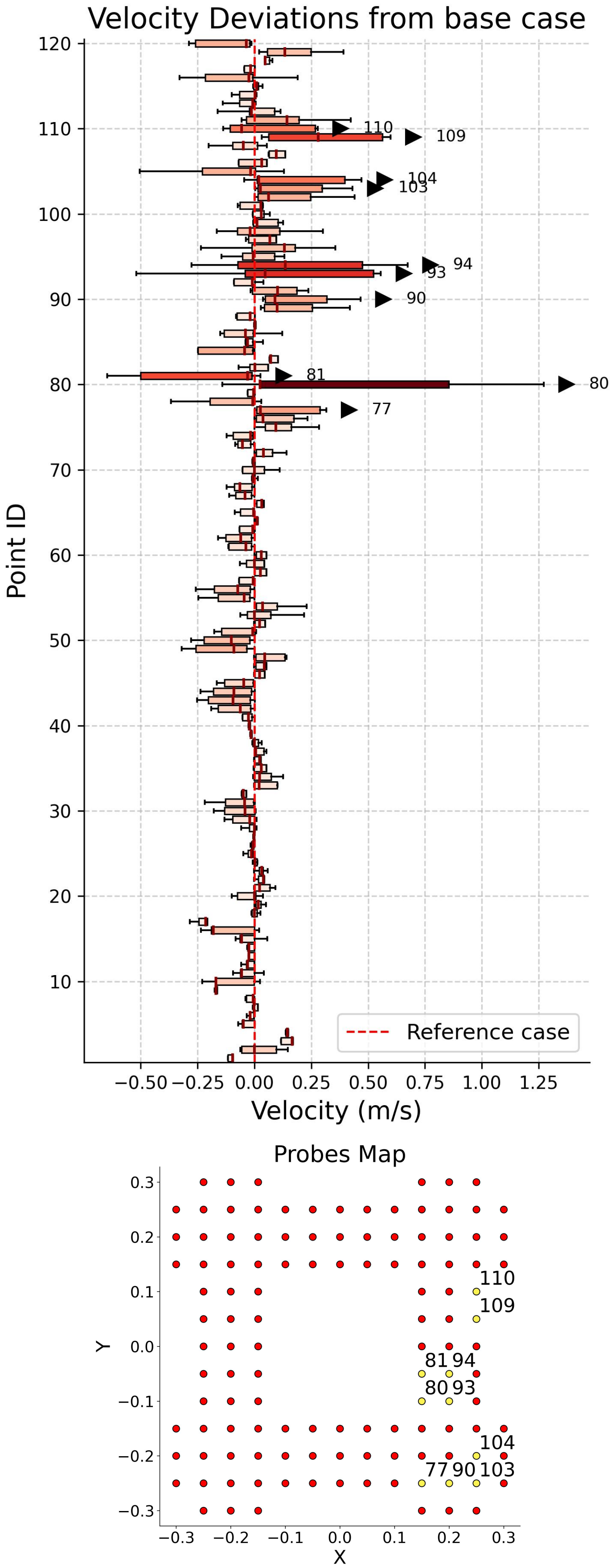
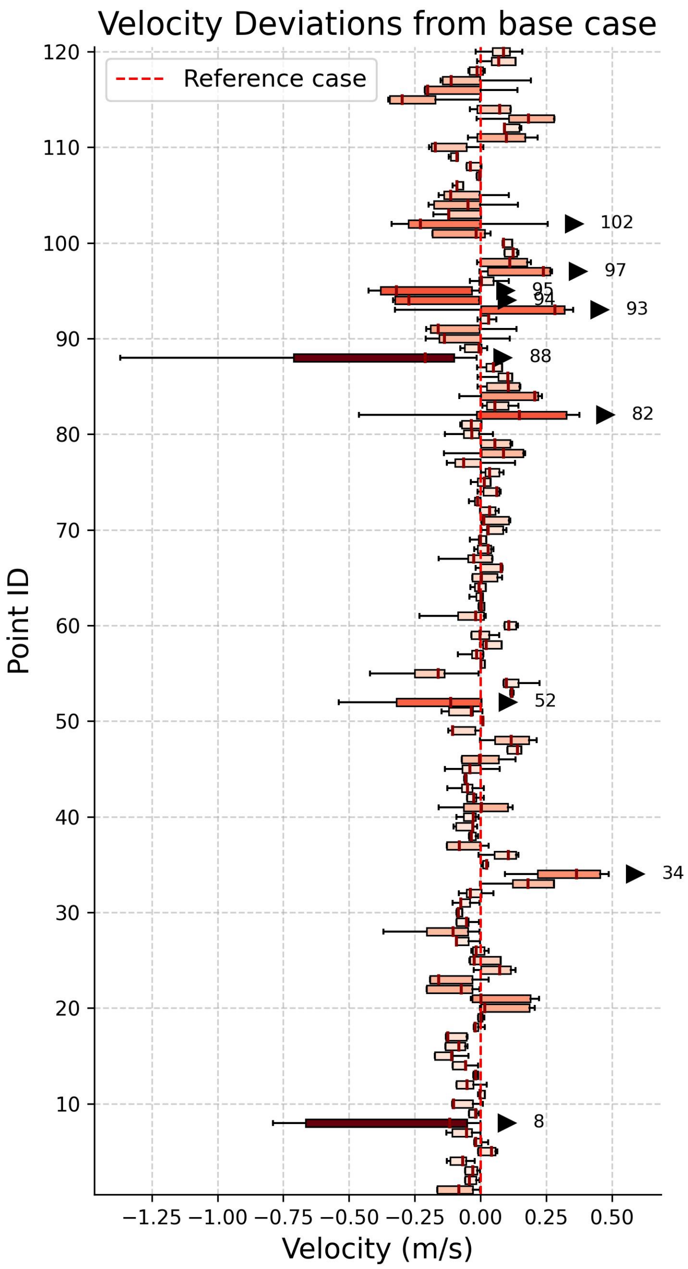


Figure E.1.: Box plots for translation (0° wind angle)



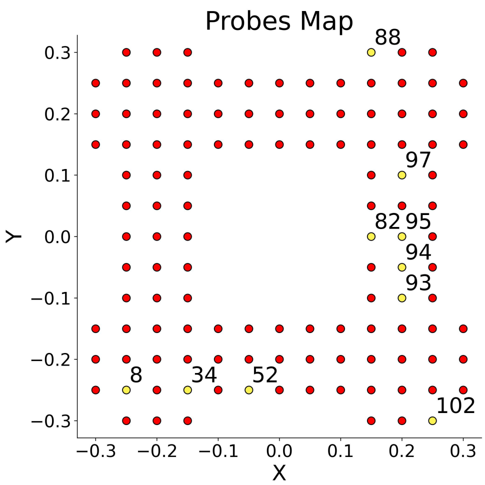


Figure E.2.: Box plots for translation (22.5° wind angle)

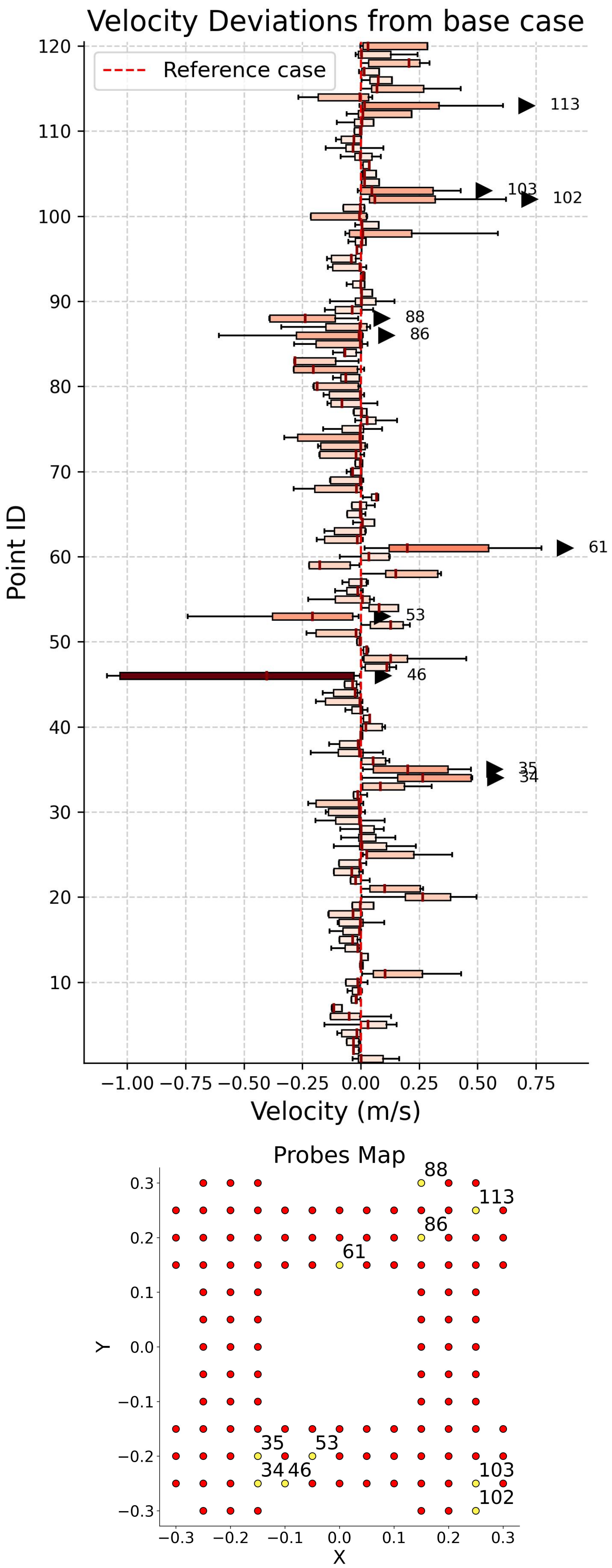


Figure E.3.: Box plots for translation (45° wind angle)

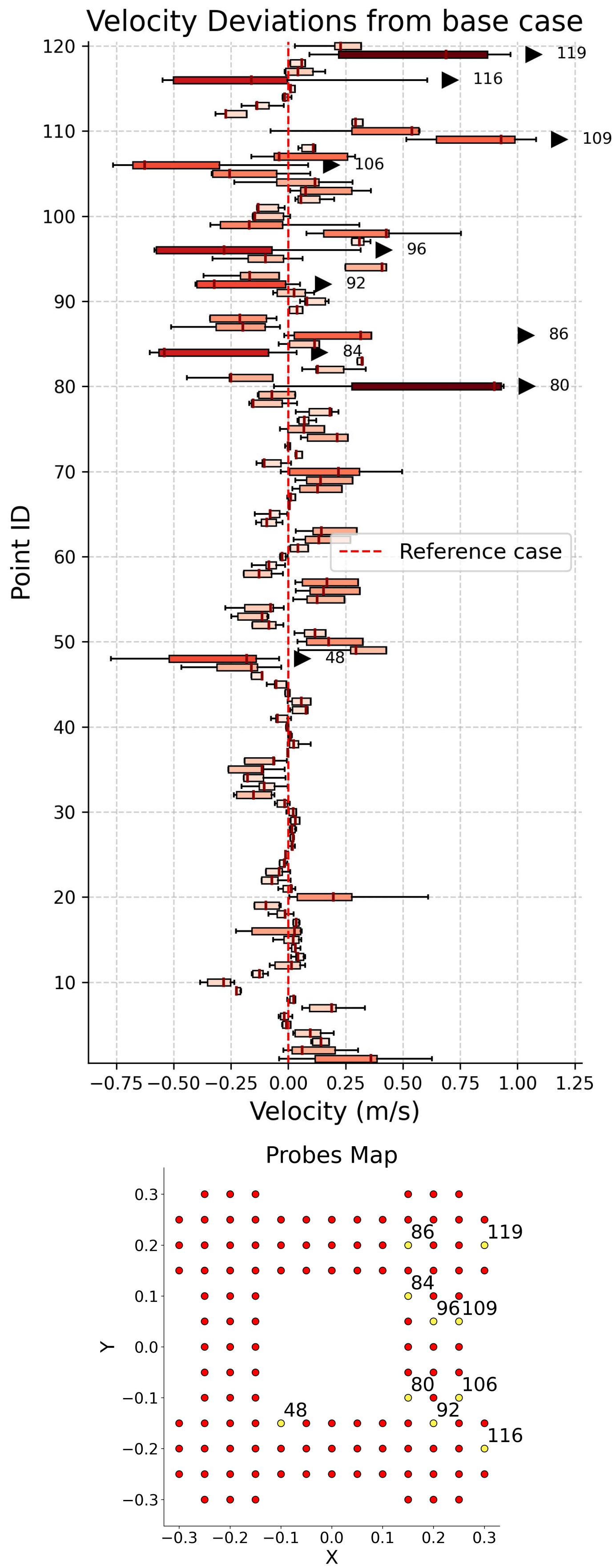


Figure E.4.: Box plots for rotation (0° wind angle)

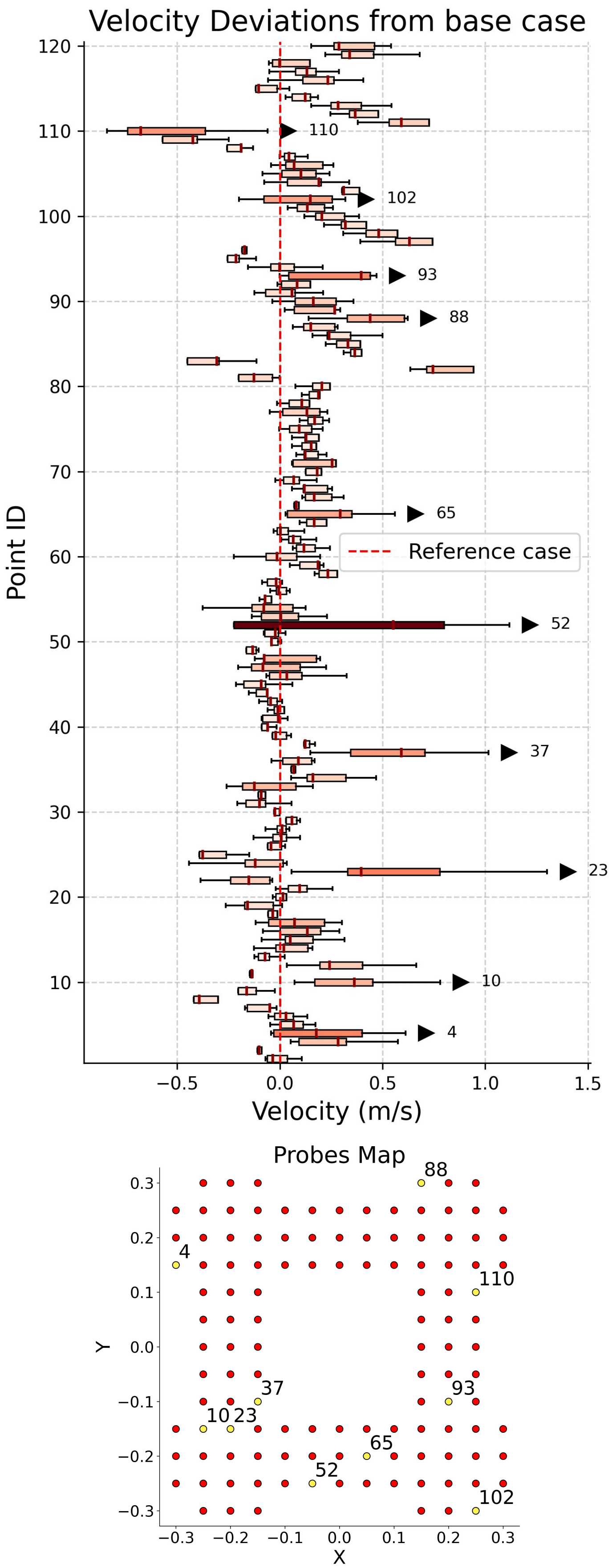


Figure E.5.: Box plots for rotation (22.5° wind angle)

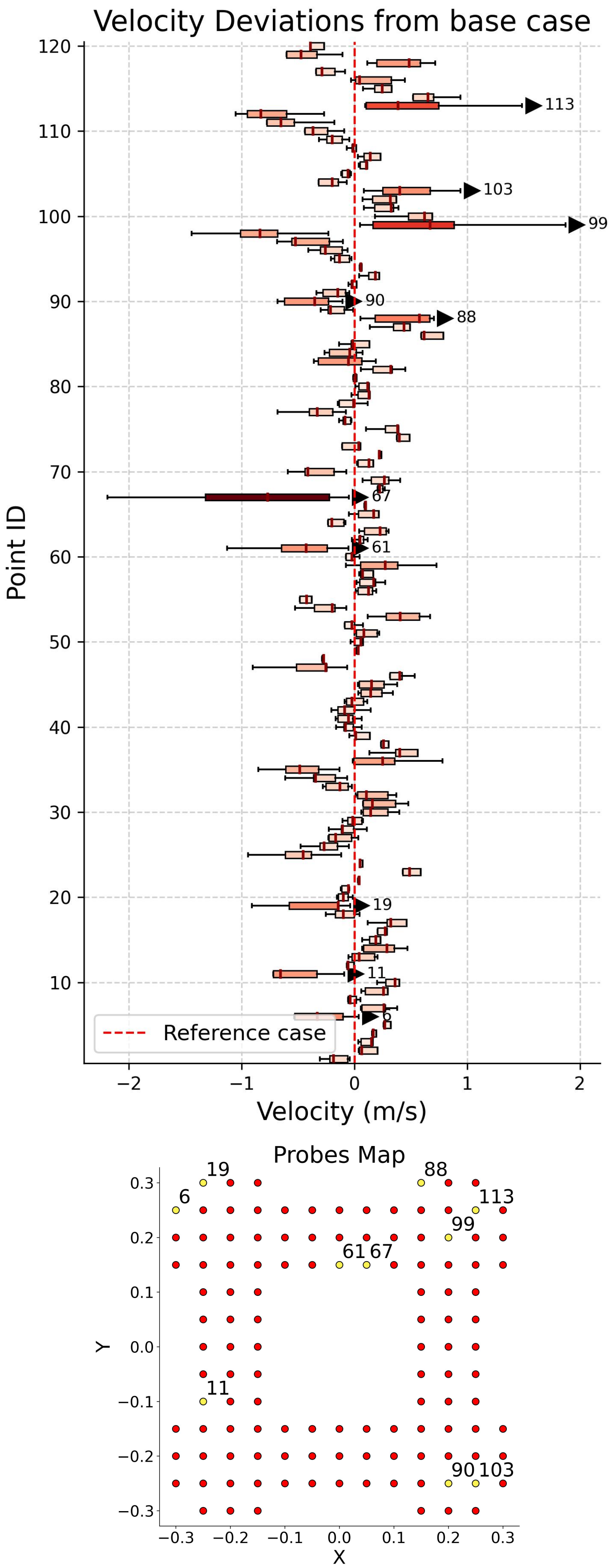
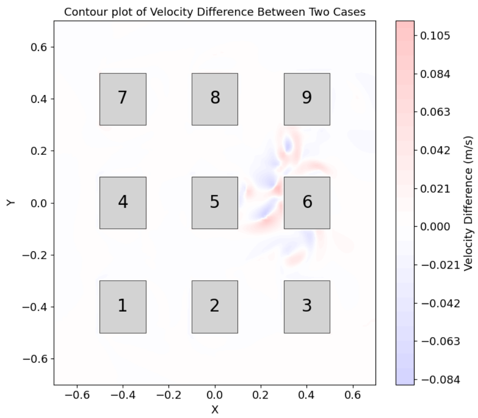
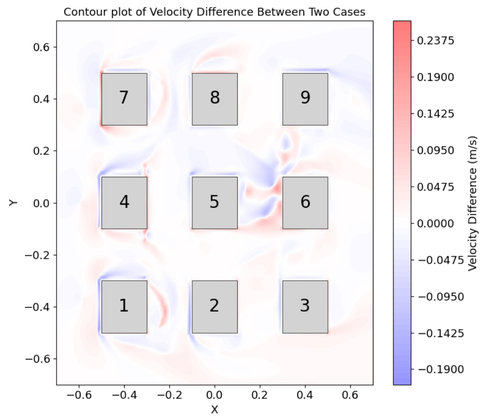


Figure E.6.: Box plots for rotation (45° wind angle)

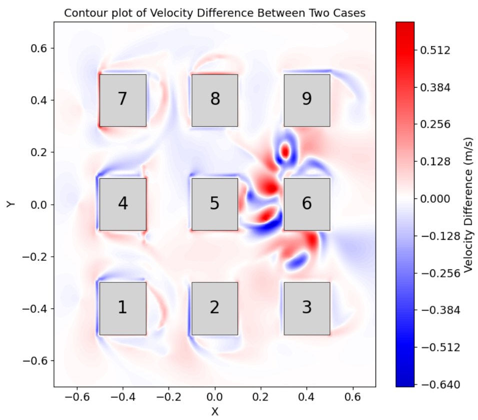
F. Contour plots



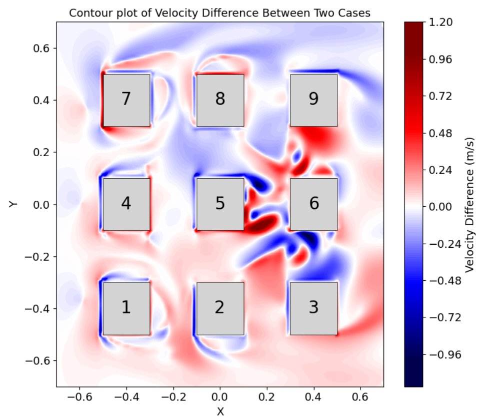
(a) T1 = 0.0001m



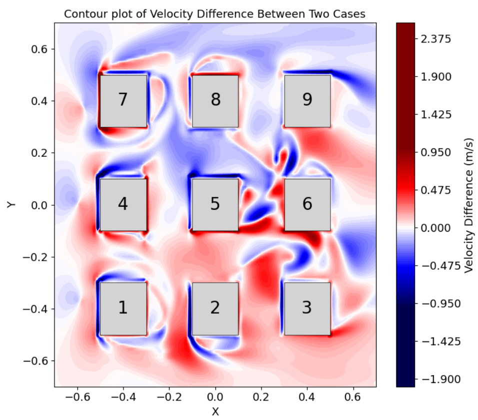
(b) T2 = 0.001m



(c) T3 = 0.002m

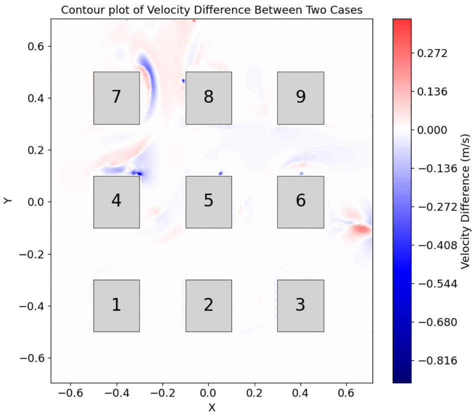


(d) T4 = 0.005m

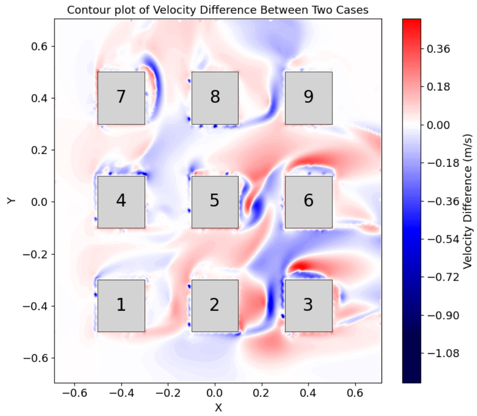


(e) T5 = 0.01m

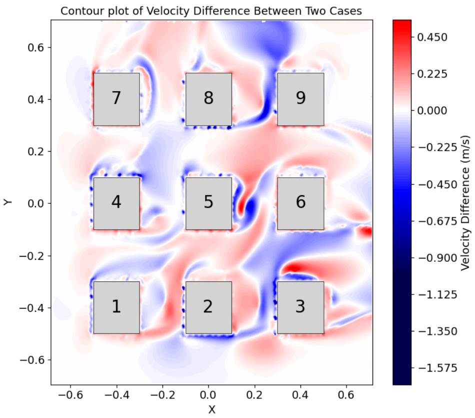
Figure F.1.: Translation contour plots for velocity: Case 0°



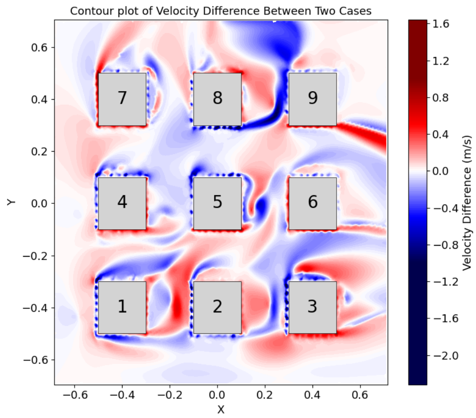
(a) T1 = 0.0001m



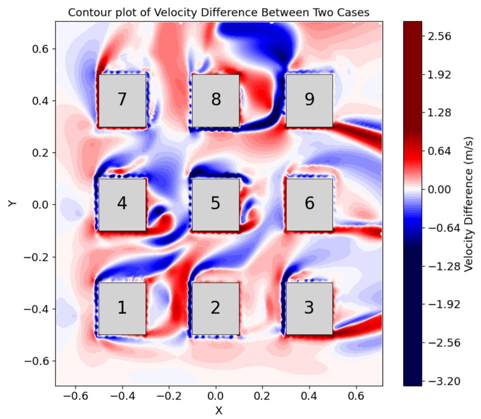
(b) T2 = 0.001m



(c) T3 = 0.002m

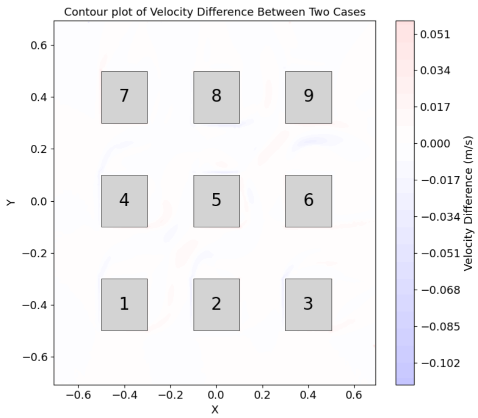


(d) T4 = 0.005m

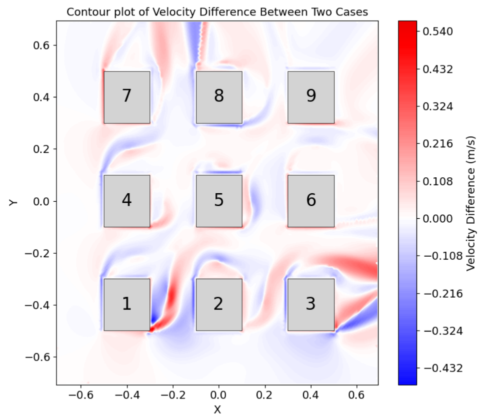


(e) T5 = 0.01m

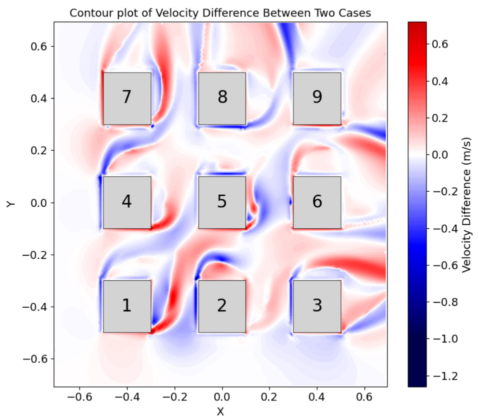
Figure F.2.: Translation contour plots for velocity: Case 22.5°



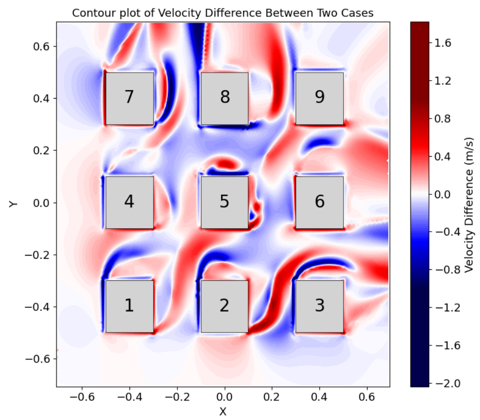
(a) T1 = 0.0001m



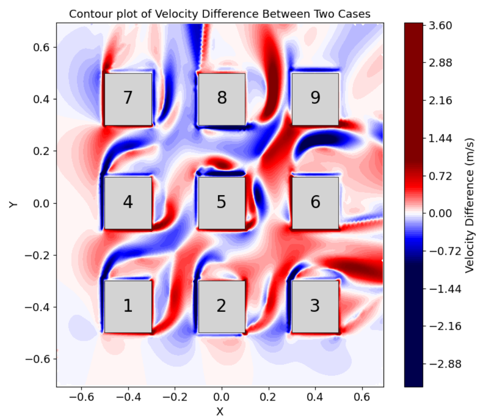
(b) T2 = 0.001m



(c) T3 = 0.002m

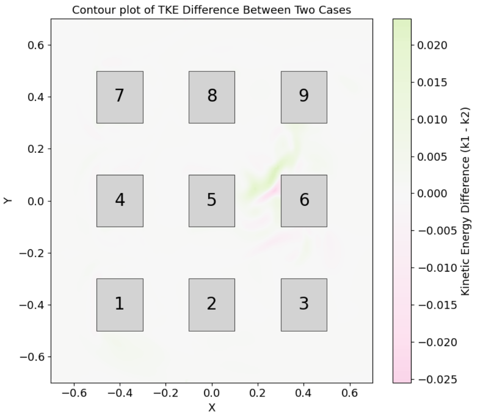


(d) T4 = 0.005m

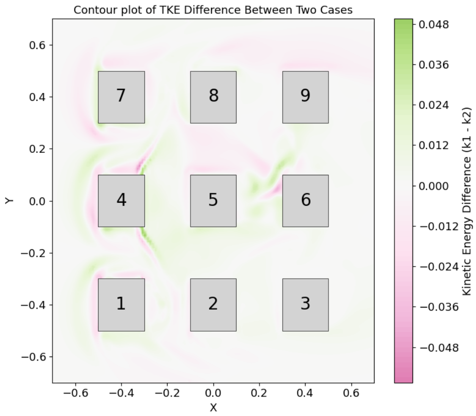


(e) T5 = 0.01m

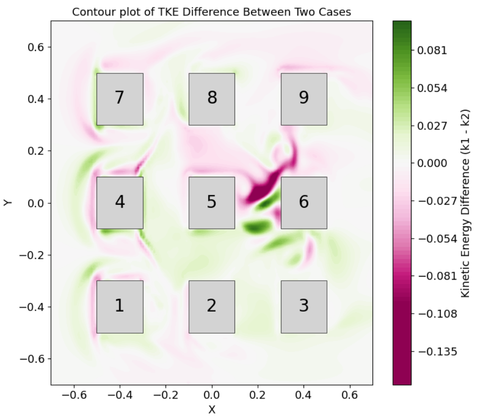
Figure F.3.: Translation contour plots for velocity: Case 45°



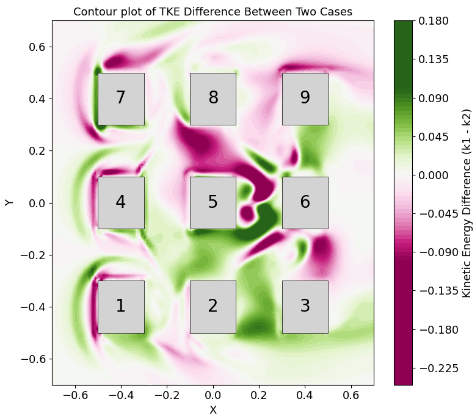
(a) T1 = 0.0001m



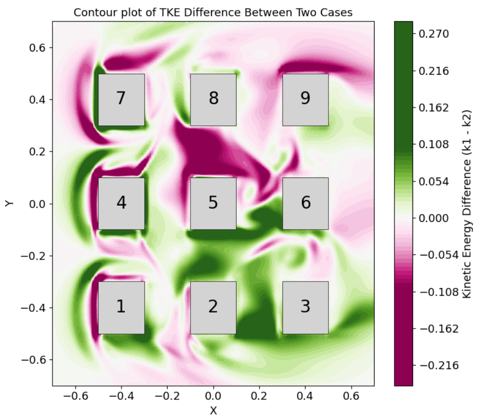
(b) T2 = 0.001m



(c) T3 = 0.002m

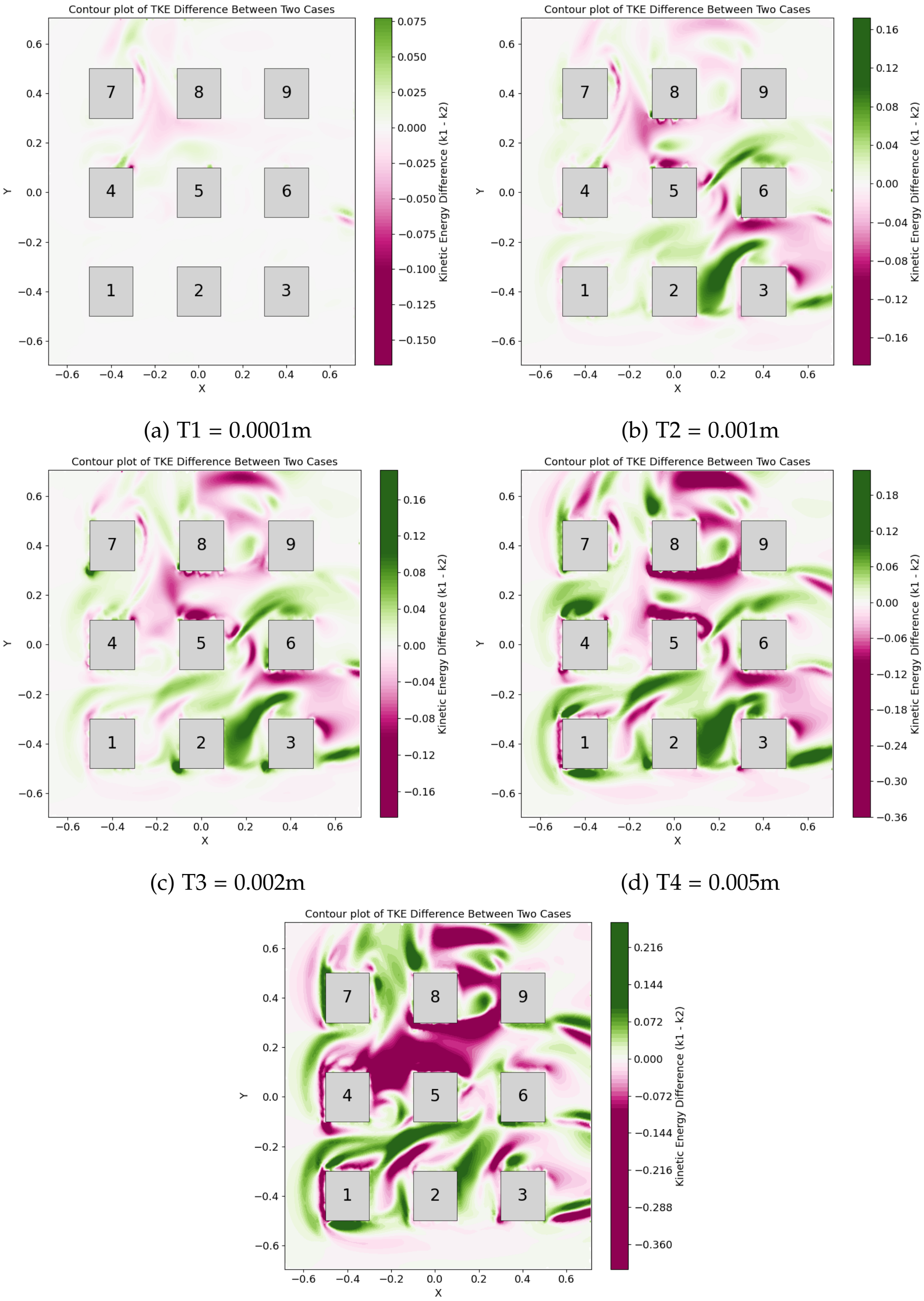


(d) T4 = 0.005m



(e) T5 = 0.01m

Figure F.4.: Translation contour plots for TKE: Case 0°



(a) T1 = 0.0001m

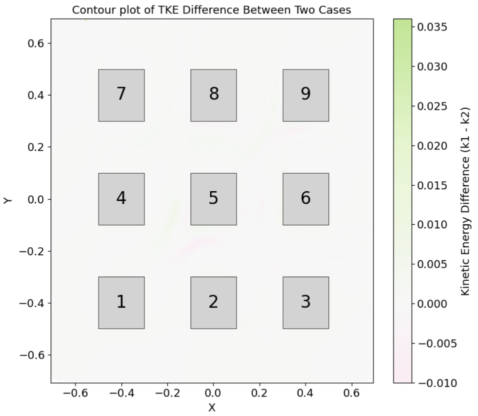
(b) T2 = 0.001m

(c) T3 = 0.002m

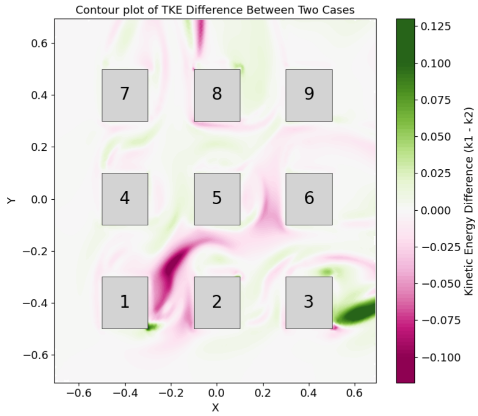
(d) T4 = 0.005m

(e) T5 = 0.01m

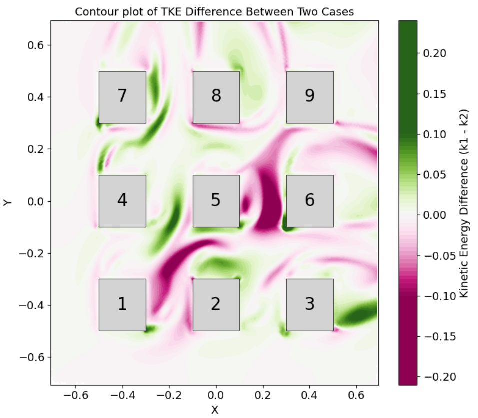
Figure F.5.: Translation contour plots for TKE: Case 22.5°



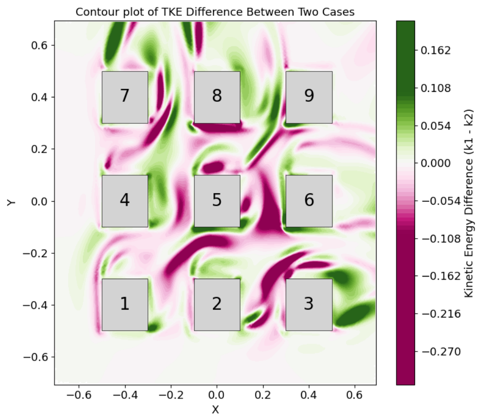
(a) T1 = 0.0001m



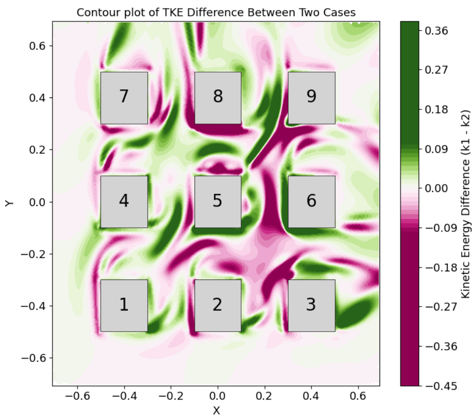
(b) T2 = 0.001m



(c) T3 = 0.002m

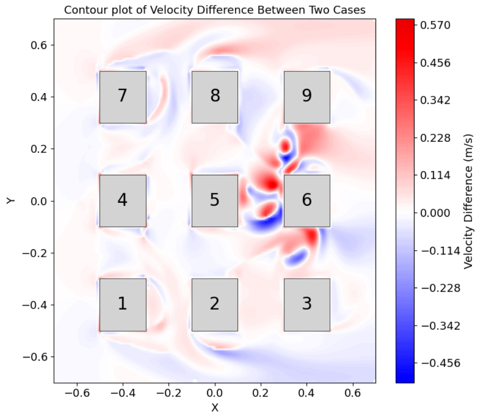


(d) T4 = 0.005m

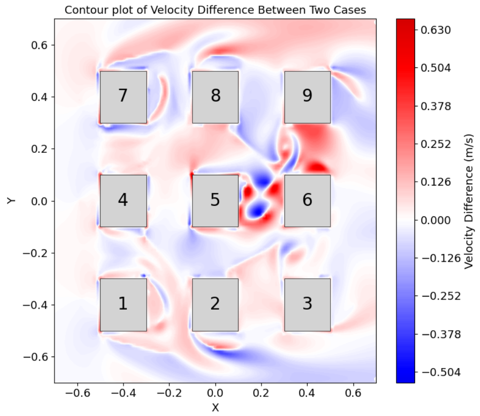


(e) T5 = 0.01m

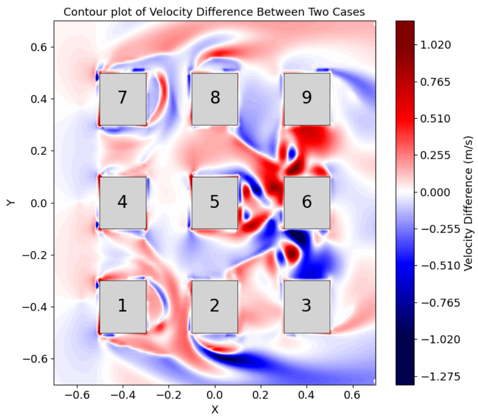
Figure F.6.: Translation contour plots for TKE: Case 45°



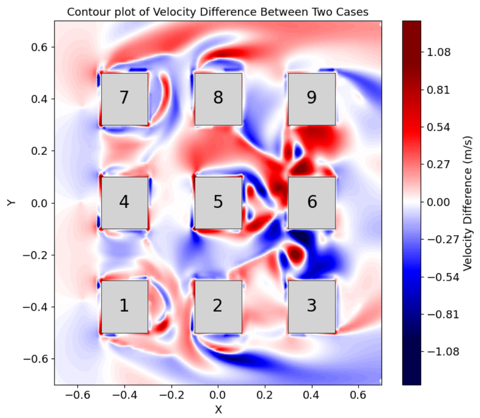
(a) R1 = 0.5°



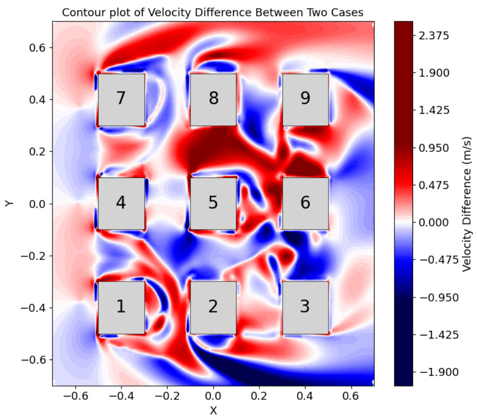
(b) R2 = 1°



(c) R3 = 2°

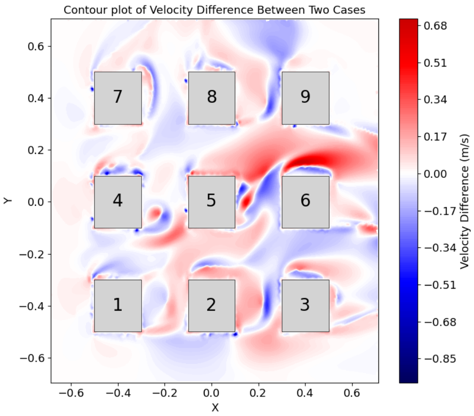


(d) R4 = 3°

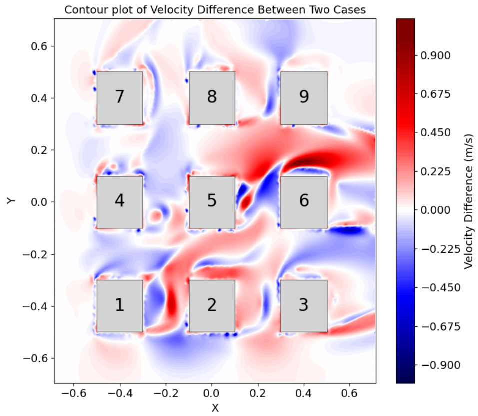


(e) R5 = 5°

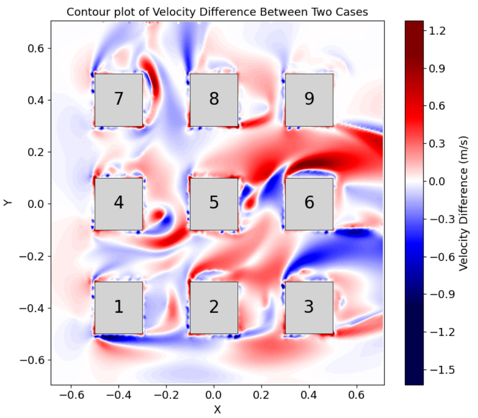
Figure F.7.: Rotation contour plots for velocity: Case 0°



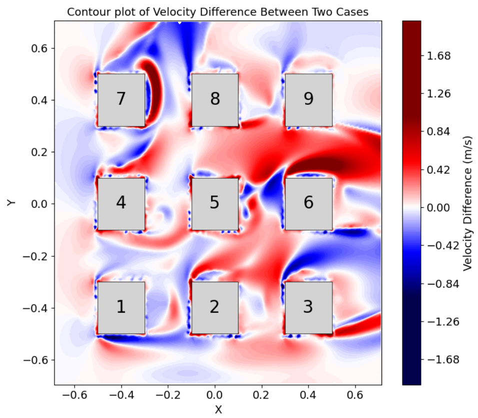
(a) R1 = 0.5°



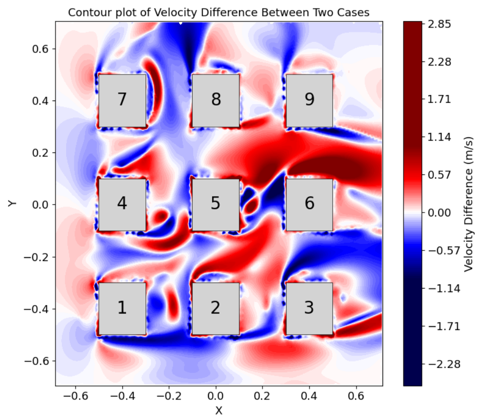
(b) R2 = 1°



(c) R3 = 2°

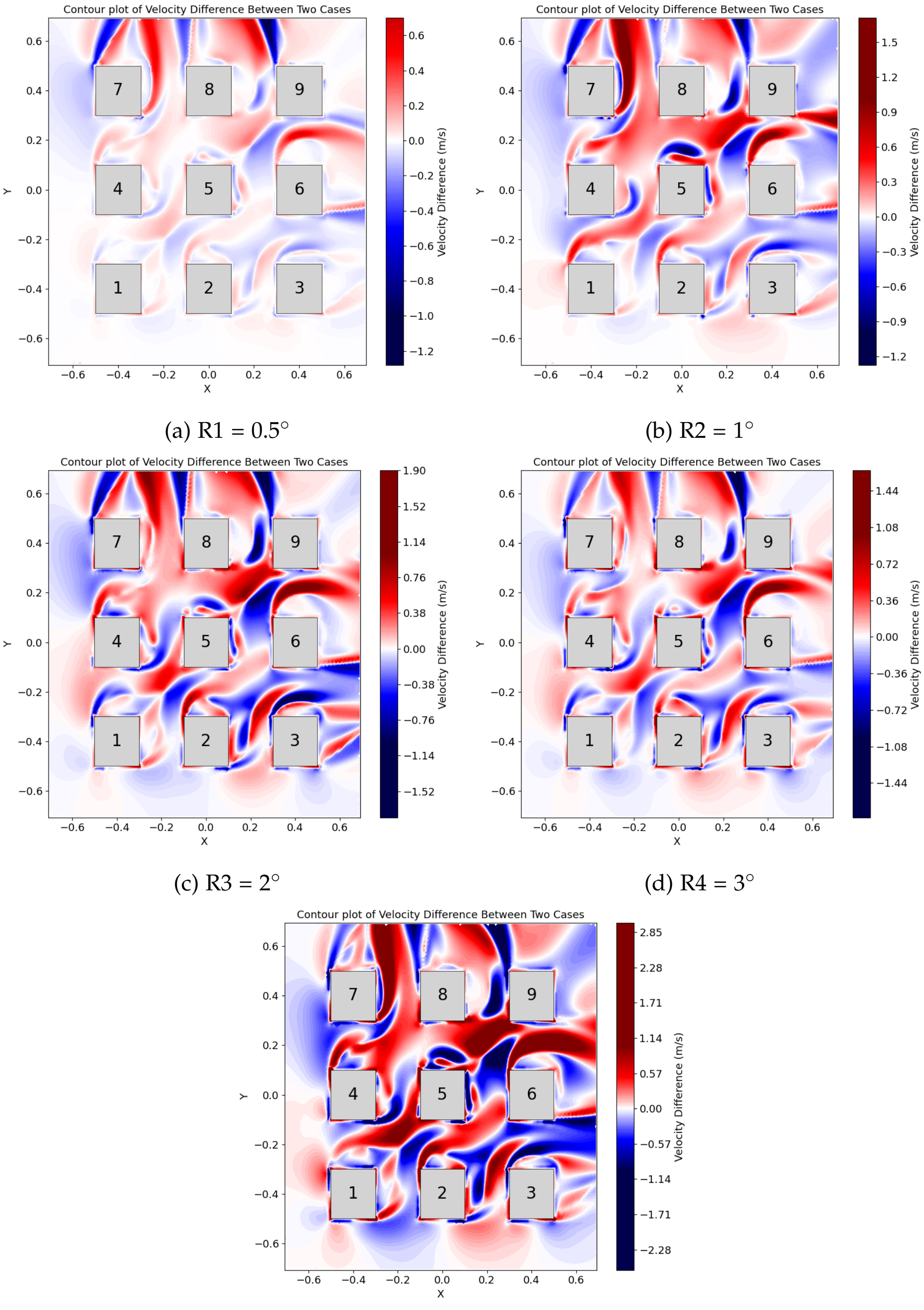


(d) R4 = 3°



(e) R5 = 5°

Figure F.8.: Rotation contour plots for velocity: Case 22.5°



(a) R1 = 0.5°

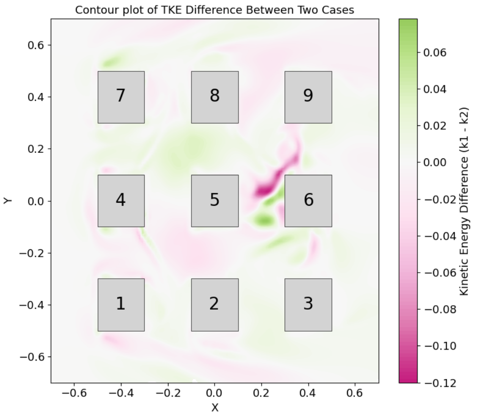
(b) R2 = 1°

(c) R3 = 2°

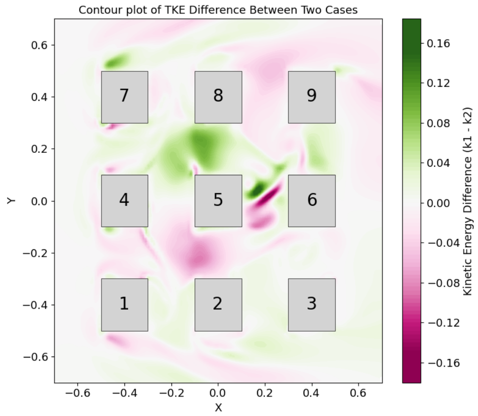
(d) R4 = 3°

(e) R5 = 5°

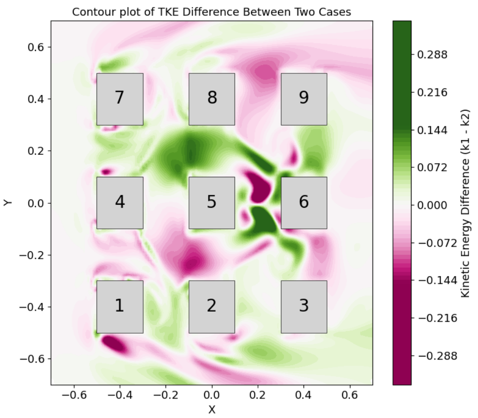
Figure F.9.: Rotation contour plots for velocity: Case 45°



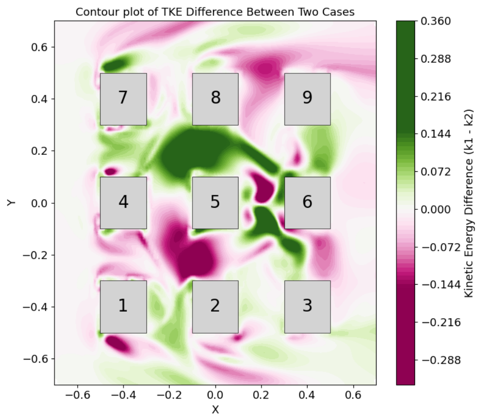
(a) R1 = 0.5°



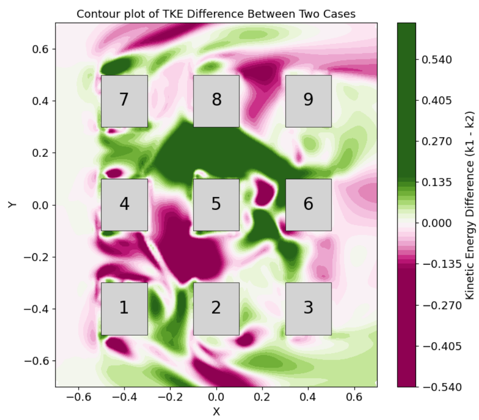
(b) R2 = 1°



(c) R3 = 2°

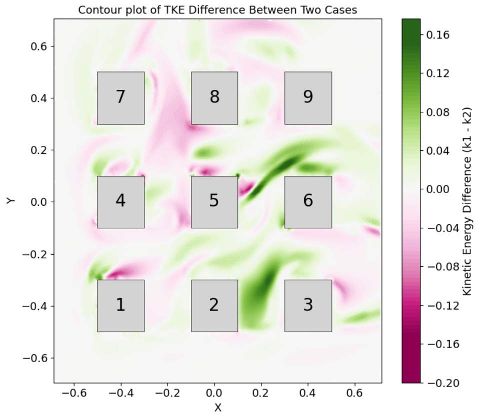


(d) R4 = 3°

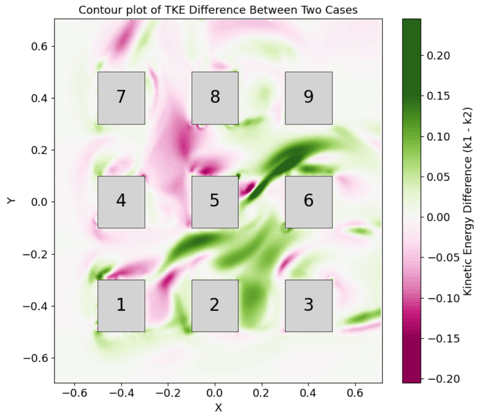


(e) R5 = 5°

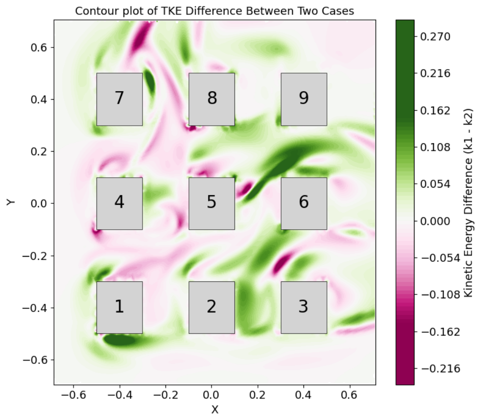
Figure F.10.: Rotation contour plots for TKE: Case 0°



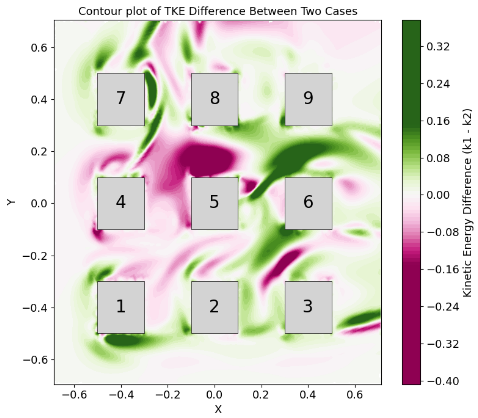
(a) R1 = 0.5°



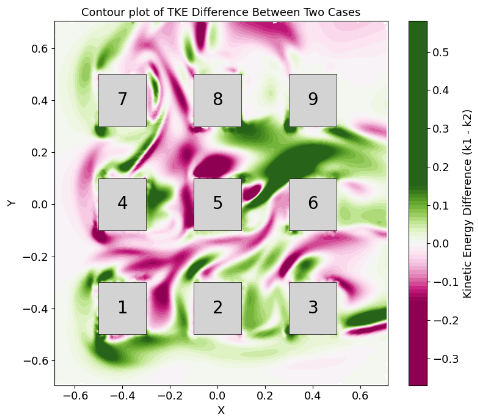
(b) R2 = 1°



(c) R3 = 2°

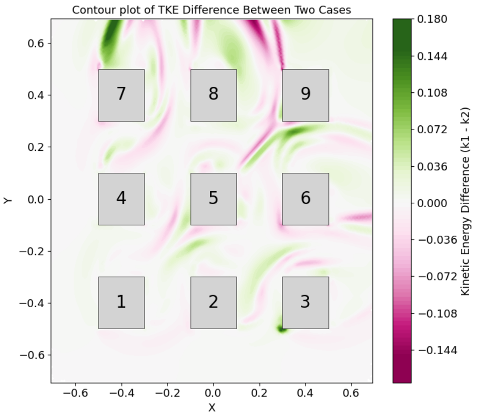


(d) R4 = 3°

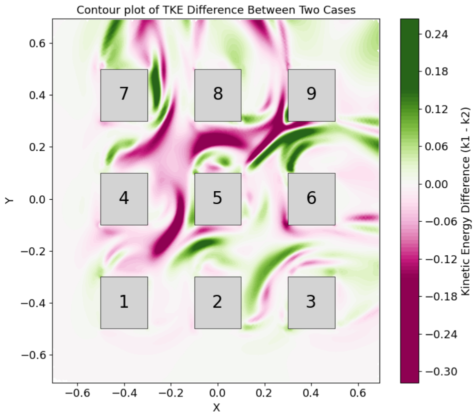


(e) R5 = 5°

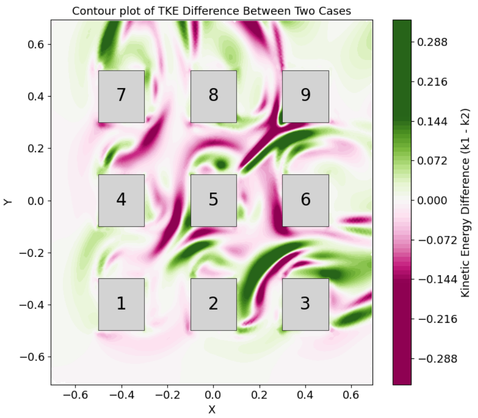
Figure F.11.: Rotation contour plots for TKE: Case 22.5°



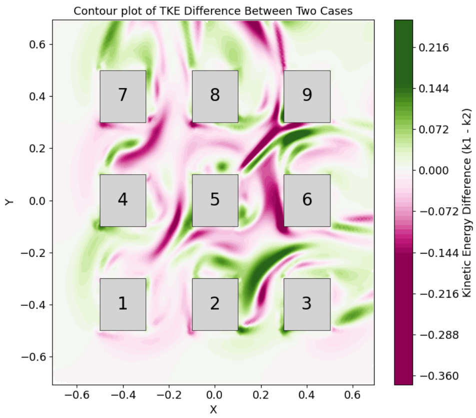
(a) R1 = 0.5°



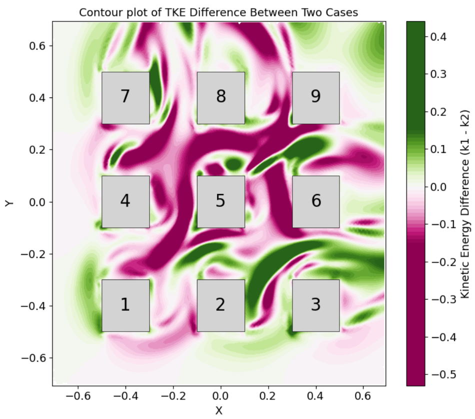
(b) R2 = 1°



(c) R3 = 2°



(d) R4 = 3°



(e) R5 = 5°

Figure F.12.: Rotation contour plots for TKE: Case 45°

Bibliography

Biljecki, F., Ledoux, H., and Stoter, J. (2016). An improved lod specification for 3d building models. *Computers, Environment and Urban Systems*, 59:25–37.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., and Çöltekin, A. (2015). Applications of 3d city models: State of the art review. *ISPRS International Journal of Geo-Information*, 4(4):2842–2889.

Blocken, B. (2015). Computational fluid dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. *Building and Environment*, 91:219–245.

Blocken, B. and Stathopoulos, T. (2013). Cfd simulation of pedestrian-level wind conditions around buildings: Past achievements and prospects. *Journal of Wind Engineering and Industrial Aerodynamics*, 121:138–145.

Celik, I., Ghia, U., Roache, P., Freitas, C., Coleman, H., and P.E., R. (2008). Procedure for estimation and reporting of uncertainty due to discretization in cfd applications. *Journal of Fluids Engineering*, 130(7):078001.

Chia, A. (2024). Python boxplots: A comprehensive guide for beginners. Online. [Last accessed 3 May 2025]. Available at: https://www.datacamp.com/tutorial/python-boxplots.

Ding, C., Lam, K., and Feng, W. (2017). An evaluation index for cross ventilation based on cfd simulations and ventilation prediction model using machine learning algorithms. *Procedia Engineering*, 205:2948–2955.

Fan, H., Zipf, A., Fu, Q., and Neis, P. (2014). Quality assessment for building footprints data on openstreetmap. *International Journal of Geographical Information Science*, 28(4):700–719.

Flaherty, J., Stock, D., and Lamb, B. (2007). Computational fluid dynamic simulations of plume dispersion in urban oklahoma city. *Journal of Applied Meteorology and Climatology*, 46(12):2110–2126.

Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B. (2007). Best practice guideline for the CFD simulation of flows in the urban environment. Technical report, COST European Cooperation in Science and Technology.

Franke, J., Hirsch, C., Jensen, G., Krüs, H., Miles, S., Schatzmann, M., Westbury, P., Wisse, J., and Wright, N. (2004). Recommendations on the use of cfd in wind engineering. Published as part of COST Action C14, Impact of Wind and Storms on City Life and Built Environment, Von Karman Institute, Sint-Genesius-Rode, Belgium.

Franke, J., Sturm, M., and Kalmbach, C. (2012). Validation of openfoam 1.6.x with the german vdi guideline for obstacle resolving micro-scale models. *Journal of Wind Engineering and Industrial Aerodynamics*, 104-106:350–359. 13th International Conference on Wind Engineering.

García-Sánchez, C., Philips, D., and Gorlé, C. (2014). Quantifying inflow uncertainties for cfd simulations of the flow in downtown oklahoma city. *Building and Environment*, 78:118–129.

García-Sánchez, C., Van Tendeloo, G., and Gorlé, C. (2017). Quantifying inflow uncertainties in rans simulations of urban pollutant dispersion. *Atmospheric Environment*, 161:263–273.

García-Sánchez, C., Vitalis, S., Pađen, I., and Stoter, J. (2021). The impact of level of detail in 3d city models for cfd-based wind flow simulations. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLVI-4/W4-2021, pages 67–72.

Hecht, R., Kunze, C., and Hahmann, S. (2013). Measuring completeness of building footprints in openstreetmap over space and time. *ISPRS International Journal of Geo-Information*, 2(4):1066–1091.

Huang, J., Stoter, J., Peters, R., and Nan, L. (2022). City3d: Large-scale building reconstruction from airborne lidar point clouds. *Remote Sensing*, 14(9):2254.

Hågbo, T.-O., Giljarhus, K., and Hjertager, B. (2021). Influence of geometry acquisition method on pedestrian wind simulations. *Journal of Wind Engineering and Industrial Aerodynamics*, 215:104665.

Jarząbek-Rychard, M. and Maas, H.-G. (2023). Modeling of 3d geometry uncertainty in scan-to-bim automatic indoor reconstruction. *Automation in Construction*, 154:105002–105002.

Lee, L., Humphreys, T., and Chan, S. (2004). Cfd simulations of joint urban atmosphere dispersion field study 2003. Technical report, U.S. Department of Energy, OSTI (Office of Scientific and Technical Information). [Last accessed 5 Mar. 2025]. Available at: https://www.osti.gov/biblio/15014068.

Liu, S., Pan, W., Zhang, H., Cheng, X., Long, Z., and Chen, Q. (2017). Cfd simulations of wind distribution in an urban community with a full-scale geometrical model. *Building and Environment*, 117:11–23.

Longo, R., Ferrarotti, M., Sánchez, C., Derudi, M., and Parente, A. (2017). Advanced turbulence models and boundary conditions for flows around different configurations of ground-mounted buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 167:160–182.

Nonomura, Y., Kobayashi, N., Tominaga, Y., and Mochida, A. (2003). The cross comparison of cfd prediction for flow field around building blocks(part 3). *Summaries to Technical Papers of Annual Meeting, Japan Association for Wind Engineering*, 2003:41–41.

Pađen, I., García-Sánchez, C., and Ledoux, H. (2022). Towards automatic reconstruction of 3d city models tailored for urban flow simulations. *Frontiers in Built Environment*, 8.

Ricci, A., Kalkman, I., Blocken, B., Burlando, M., Freda, A., and Repetto, M. (2017). Local-scale forcing effects on wind flows in an urban environment: Impact of geometrical simplifications. *Journal of Wind Engineering and Industrial Aerodynamics*, 170:238–255.

Richards, P. and Hoxey, R. (1993). Appropriate boundary conditions for computational wind engineering models using the k- turbulence model. *Journal of Wind Engineering and Industrial Aerodynamics*, 46-47:145–153.

Roache, P. (1994). Perspective: A method for uniform reporting of grid refinement studies. *Journal of Fluids Engineering*, 116(3):405–413.

Robins, A., Hall, R., Cowan, I., Bartzis, J., and Albergel, A. (2000). Evaluating modelling uncertainty in cfd predictions of building affected dispersion. *International Journal of Environment and Pollution*, 14(1/2/3/4/5/6):52.

Rottmann, P., Haunert, J.-H., and Dehbi, Y. (2022). Automatic building footprint extraction from 3d laserscans. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume X-4/W2-2022, pages 233–240.

Sim-flow (2024). Wind analysis - buildings - validation case - simflow cfd. Online. [Last accessed 3 May 2025]. Available at: `https://help.sim-flow.com/validation/wind-analysis-buildings`.

SimScale (2022). Pedestrian wind comfort: Aij case c - validation case - simscale. Online. [Last accessed 3 May 2025]. Available at: `https://www.simscale.com/docs/validation-cases/aij-case-c/`.

Tominaga, Y. and Blocken, B. (2015). Wind tunnel experiments on cross-ventilation flow of a generic building with contaminant dispersion in unsheltered and sheltered conditions. *Building and Environment*, 92:452–461.

Tominaga, Y., Mochida, A., Yoshie, R., Kataoka, H., Nozu, T., Yoshikawa, M., and Shirasawa, T. (2008). Aij guidelines for practical applications of cfd to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10-11):1749–1761.

Tominaga, Y., Yoshie, R., Mochida, A., Kataoka, H., and Nozu, T. (2005). Cross comparisons of cfd prediction for wind environment at pedestrian level around buildings part 2: Comparison of results for flowfield around building complex in actual urban area. Online. [Last accessed 5 May 2025].

Uni-hamburg.de (2024). data sets. Online. [Last accessed 4 May 2025]. Available at: `https://www.mi.uni-hamburg.de/en/arbeitsgruppen/windkanallabor/data-sets.html`.

Wang, W. and Ng, E. (2018). Large-eddy simulations of air ventilation in parametric scenarios: Comparative studies of urban form and wind direction. *Architectural Science Review*, 61(4):215–225.

Wieringa, J. (1992). Updating the davenport roughness classification. *Journal of Wind Engineering and Industrial Aerodynamics*, 41(1):357–368.

www.aij.or.jp (2025). Guidebook for cfd predictions of urban wind environment. Online. [Last accessed 5 May 2025]. Available at: `https://www.aij.or.jp/jpn/publish/cfdguide/index_e.htm`.

Yoshie, R., Mochida, A., Tominaga, Y., Kataoka, H., Harimoto, K., Nozu, T., and Shirasawa, T. (2007). Cooperative project for cfd prediction of pedestrian wind environment in the architectural institute of japan. *Journal of Wind Engineering and Industrial Aerodynamics*, 95(9-11):1551–1578.

Zheng, S., Wang, Y., Zhai, Z., Xue, Y., and Duanmu, L. (2021). Characteristics of wind flow around a target building with different surrounding building layers predicted by cfd simulation. *Building and Environment*, 201:107962.

Colophon

This document was typeset using LaTeX, using the KOMA-Script class `scrbook`. The main font is Palatino.