# COST 732 MODEL EVALUATION CASE STUDIES: APPROACH AND RESULTS

**DRAFT May 31**

## COST Action 732

## QUALITY ASSURANCE AND IMPROVEMENT OF MICROSCALE METEOROLOGICAL MODELS

**May 2009**

# Contents

# 1 Introduction

## 1.1 Motivation for the work

Urban emissions occur mainly within or slightly above the canopy layer i.e. within a zone where the atmospheric flow is heavily disturbed by buildings and other obstacles. It is well-known that buildings can deflect plumes and result in sometimes increased, sometimes decreased dilution, depending on the density, height and shape of the obstacle array and its orientation towards the wind. In comparison to unobstructed terrain, local concentrations can change by more than an order of magnitude. As a consequence, it is inappropriate to consider buildings within a surface roughness parameterisation only, particularly if predictions on the scale of a few streets or city blocks are being made. These facts, in conjunction with increasing computer power, have promoted the development of obstacle-resolving or obstacle-accommodating prognostic and diagnostic models, subsequently called micro-scale meteorological models. Nowadays these models are commercially available and widely used in environmental impact studies.

The increasing use of these models is paralleled by a growing awareness that the majority of these models have not been the subject of systematic evaluation. Consequently, a lack of confidence in the modelled results might arise. To cast doubt on the results is justified, as was shown by systematic studies in which applications of the same model by different modellers to a given problem (Hall, 1997) and applications of different models by either the same or different modellers to the same problem (Ketzel et al., 2001) revealed significant differences. Nevertheless, these models are used in the preparation of decisions with profound economic and political consequences.

The reason that most of the models lack quality assurance is not due to insufficient efforts made by the model developers. It is mainly caused by

- a lack of a commonly accepted quality assurance procedure for such models, and
- a lack of data sets that are quality checked and commonly accepted as a standard for model validation purposes.

In order to assure the quality of models for flow and dispersion predictions in urban and industrial areas, a European COST initiative (see http:/www.cost.esf.org) was launched. COST is an intergovernmental European framework for international cooperation between nationally funded research activities. COST creates scientific networks and enables scientists to collaborate in a wide spectrum of activities in research and technology. Extra research money is not provided.

The objective of action COST 732 which comprised scientists from 22 European countries (see http://www.mi.uni-hamburg.de/Home.484.0.html) was to improve and to assure the quality of micro-scale meteorological models that are applied for predicting flow and transport processes in urban or industrial environments. The steps towards the realisation of this objective are given in Fig. 1.

## 1.2 Types of Models

For the short-range local problems (up to about 5 km) at stake here, simple Gaussian type models have been and are still being used. These models are applicable for pollutant emissions into stationary and uniform atmospheric flows (for example tall stack releases in flat, unobstructed terrain). It is accepted that in urban areas these models can still be applied when clouds of pollutants disperse above the buildings or when the dimensions of the cloud are much bigger than the dimensions of the obstacles. Within the urban canopy and in the near field of the source where the flow conditions are usually far from uniform or stationary their usefulness is restricted to certain purposes.

**The objective** of action COST 732 is to improve and **assure the quality of** micro-scale meteorological **models** that are applied **for predicting flow and transport processes in urban or industrial environments.**
The steps towards realising this objective are

- develop a coherent and structured **quality assurance procedure**,
- provide appropriate **data for model validation,**
- **build a consensus** within the community,
- stimulate the preparation of **quality assurance protocols**, and
- establish a **consensus on 'best practises'** in current model use.

Fig. 1: Major objective and steps towards its realisation

At least in the developed countries air quality protection policy has achieved that emissions from most of the big elevated sources (power plant stacks etc.) has been substantially reduced. Concern has shifted to sources that emit close to the ground. The emissions result from traffic and domestic or industrial sources. Other types of releases that deserve attention are toxic or flammable gases that escape during accidents or that are deliberately released in the context of terrorist attacks.

It is common in all these cases that the dispersion takes place within or slightly above the urban or industrial canopy layer. In this layer, the flow is significantly disturbed by buildings and other obstacles. The meteorological and concentration fields are typically inhomogeneous and vary rapidly with time. Under such conditions Gaussian models may still play a role as screening tools but lack the potential for realistic predictions of the spatial distribution of the concentration field.

Tools which have the potential to accommodate obstacles in a reasonable way are micro-scale meteorological models of prognostic type. These models are based on the Navier-Stokes equation. To directly solve the equation in a turbulent flow requires a very fine grid to capture all the relevant scales, down to the so-called Kolmogorov scale (usually less than a mm). Furthermore, a time-dependent solution over sufficiently long periods is needed to yield stable time averages of the flow variables. This approach is called direct numerical simulation (DNS). As its computational demand is too high for the Reynolds numbers typically encountered in atmospheric boundary layers, DNS is not applicable here.

The computational demand can be substantially reduced when the time-dependent equations are solved on a grid that is fine enough (less than a few m in the case of a city quarter) to resolve the larger atmospheric eddies. This approach is called large eddy simulation (LES). The small scales are formally removed from the flow variables by spatially filtering the Navier-Stokes equations. The influence of the small scales then appears as sub-filter stresses in the momentum equation. Since the large eddies are always unsteady, LES models require input conditions which are time dependent as well. Whilst being less demanding than DNS, LES still requires significant computer resources which go beyond the capabilities most users presently have.

In view of this the still most wide-spread method used for the computation of turbulent atmospheric flows is the Reynolds-Averaged Navier-Stokes (RANS) approach. Within this approach the equations are averaged in time over all turbulent scales, to directly yield the statistically steady solution of the mean and turbulent flow variables. Like LES the averaging leads to additional terms in the momentum equation known as the Reynolds stresses. They represent the effects of the turbulent fluctuations on the averaged flow and have to be parameterised. This is the task of turbulence closure models. In most models which are presently in use, the Reynolds stresses are assumed to depend linearly on the strain rate, as do the molecular stresses. The eddy viscosity appears as a proportionality factor that can be calculated using additional differential equations for the various order moments. Many modellers regard the 2-equations turbulence closure schemes which solve differential equations for the turbulent kinetic energy k and the dissipation rate ε a good compromise between universal validity and operating expense. In particular, the standard k-ε model is widely used in engineering and micro- meteorological applications, despite the fact that it produces too much turbulent kinetic energy in regions of stagnant flow etc. Several modifications have been proposed which ease this problem but most times on the expense of the quality of other flow property predictions.

Another option is the use of standard turbulence models from the RANS approach in time-dependent simulations. Contrary to LES the averages are defined as ensemble or as time averages over small time intervals, although the later definition leads to more additional terms in the momentum equation than in the case of ensemble averaging. This approach is also known as unsteady RANS (URANS). This type of model is used in standard meteorological meso- and macro-scale applications (weather forecasts etc.). URANS models are driven by time-dependent boundary conditions. They account, e.g., for different land uses with different radiation budgets and provide time dependent predictions. In the context of urban applications, URANS models are under development. However, much more powerful computers than presently available and substantial research efforts will still be needed before the first reliable unsteady obstacle resolving predictions for urban-scale dispersion problems will become available.

A further class of models still in use are diagnostic models. They do not use the Navier-Stokes Equation but are based solely on mass conservation. The important influence of pressure gradients and forces on the flow development can only indirectly (= empirically) be taken into account. These models start with a first guess of the 3-dimensional flow field that is subsequently modified until the divergence of the flow falls below a chosen limit. For a given obstacle array many different mass conserving flow fields can be found, depending on the particular choice of the initial flow field and the ‘tuning of knobs’ inside the model which, e.g., determine whether the fluid at a specific position moves over or goes around an obstacle. Although practitioners tend to ignore that; diagnostic models may be helpful in analysing known cases, in particular when a good set of observed data is available for the area of interest. However, they can not themselves be regarded as tools for the prediction of new cases, even if they can be used for downscaling meteorological fields provided by prognostic models to higher resolutions.

Lagrangian models, finally, are pure dispersion models, i.e. they follow individual plume parcels and model their paths on the basis of a random walk process. They need a complete mean and turbulent flow field as model input, which is usually delivered in form of 3D-gridded fields by either a diagnostic or prognostic model. Consequently, the validation of Lagrangian models is closely linked to the validation of the model that was chosen to provide the flow field, and the performance of the Lagrangian highly depends on the quality of the meteorological input data. A further validation is then required for the Lagrangian model itself, since several features can differ in the core of these models. Among them, the interpolation techniques applied to derive, from the Eulerian 3D-gridded data, the fields in any point of the Lagrangian grid-free domain at any time frame; the choice of the probability density function for the vertical wind velocity fluctuations; the integration of the Langevin equation, that may be solved with a one-dimensional approach up to its fully 3D version; the bouncing of the particles at the lower boundary and on the obstacles; the numerical aspects and the implementa-

tion of other parameterizations, such as dry and wet deposition, radiation decay or the treatment of non-neutral gases.

Since the time of COST 732 was limited, the action had to be selective concerning the choice of models. Focus was laid on models mostly applied in the EU. These are the RANS-CFDs and Non-CFD models of mainly Gaussian type.

### 1.3 Methodology

The action started in July 2005 with a joint ESF/COST 732 Exploratory Workshop on 'Quality Assurance of Micro-Scale Meteorological Models'. About 45 invited scientists from Europe and the US attended the workshop. The workshop proceedings (Schatzmann and Britter, Eds., 2005) contain a state of the art report on former quality assurance initiatives in the field of micro-scale meteorological models and recommendations that were given to the COST action.

The next logical step was to draft a first version of the evaluation procedure and its underlying motivation in order to provide the basis for subsequent discussions within the scientific community. This was done in form of two related documents: A rather lengthy

- Background and justification document to support the model evaluation guidance and protocol document (Britter and Schatzmann, Eds., 2007)

and a much shorter

- Model evaluation guidance and protocol document (Britter and Schatzmann, Eds., 2007)

with step-by-step guidance for the evaluation of a model.

The major steps of the evaluation procedure as recommended in the reports are sketched in Fig.2. Five of the six steps of the evaluation procedure described are relatively straightforward. Only the model validation step is complex and requires more attention.

In practise the quality of model output depends not only on the accuracy of the model itself and the model input. Likewise important is the qualification of the person running a model. Numerical simulation is a knowledge-based activity. Appropriate knowledge can be transferred to users by recommendations concerning the proper use of models. For obstacle resolving CFD codes such recommendations are not straightforward. COST 732 tried to respond to that problem by releasing a fourth document, the

- Best practice guideline for the CFD simulation of flows in the urban environment [6].

All 4 reports are available to the general public and can be downloaded from the home page of the action under http://www.mi.uni-hamburg.de/Home.484.0.html.

## 1.4 Role of this report

In order to test the recommendations in practice, model evaluation exercises have been carried out by the members of the action themselves. The data sets which were selected represent an idealised urban situation: the container village used in the Mock Urban Setting Tests in the desert of Utah, and a real challenging urban landscape; the business district of Oklahoma City (see chapter 2).

With the present document it is intended to

- show a few examples of how a model evaluation exercise can be conducted,
- highlight problems and experiences gained during the course of the exercise,
- give an indication of the state of the art in CFD and non-CFD modelling for micro-

scale meteorological applications,

- demonstrate the potential of certain tools which were developed within the action in order to perform exploratory analyses of model performance, and quantify it in terms of metrics.
- to provide specific guidance so that other modellers in future can put their model results into the same framework that has been used within the COST group.

The structure of the document is as follows:

(will be added later)

**MODEL EVALUATION PROCEDURE**

- **Model Description:** this should be a brief description of the characteristics of the model, the intended range of applicability, the theoretical background on which the model development was based, the software and hardware requirements, etc.
- **Database Description:** a complete description of the database that has been employed for the evaluation of the model, including the reasons why this specific database was chosen. An estimation of the data variability is required.
- **Scientific Evaluation:** this is a description of the equations employed to describe the physical and chemical processes that the model has been designed to include. If appropriate it should justify the choice of the numerical modelling procedures and it should clearly state the limits with respect to the intended applications.
- **(Code) Verification:** this process is to verify that the model produces results that are in accordance with the actual physics and mathematics that have been employed. This is to identify, quantify and reduce errors in the transcription of the mathematical model into a computational model and the solution (analytical or numerical) of the model.
- **Model Validation:** this is a structured comparison of model predictions with experimental data and is based on statistical analyses of selected variables. It seeks to identify and quantify the difference between the model predictions and the evaluation datasets; it provides evidence as to how well the model approximates to reality. A quantification of the uncertainty of the model predictions should be produced.
- **User-Oriented Assessment:** is there a readable, comprehensive documentation of the code including technical description, user manual and evaluation documentation? The range of applicability of the model, the computing requirements, installation procedures, and troubleshooting advice should be available.

Fig. 2: Steps that should be completed during the evaluation process

# 2 Validation Data

The validation of mathematical flow and dispersion models relies markedly on the availability of experimental data for model testing and for the evaluation of model quality. There has been no clear definition of what should be called 'validation data' and which quality requirements should be met by validation data. Consequently, various sources of atmospheric flow and dispersion data have been used, and - in many cases – misused for validation purposes.

The quality and completeness of the test data significantly affects the outcome of the validation process, and it must be stated clearly that the result of any model validation activity will critically depend on the quality of the reference data.

## 2.1 Validation Data Requirements

The validation data requirements are essentially based on the type of problem that is being considered and the type of model that is being used to simulate the problem. The model types as they were introduced in Chapter 1.2 need validation data which match the complexity of the models. E.g. for micro-scale meteorological URANS- models the ideal data set would be composed of data from a dense array of locations at which time series of the major flow and dispersion properties are simultaneously measured by hopefully flawless instruments. The inflow boundary condition have to represent the unsteady behaviour of the approach flow in sufficient spatial detail, and the lower boundary the varying surface and heat exchange characteristics at the ground. Gaussian models which can only handle steady uniform flows, on the other hand, would be much less demanding and only require a reliable mean wind velocity plus a sufficiently dense array of concentration samplers downwind from the source.

COST action 732 dealt mainly with Urban RANS CFD codes and Non-CFD models since these types of models are presently most widely used in the COST member states. Both model types apply steady boundary conditions, i.e. they deliver one solution for one specific meteorological situation. The variability of atmospheric situations is indirectly accounted for by repeating the runs for different wind directions, stabilities etc., each of these runs being steady state again.

Steady state models require validation data from experiments which were taken under steady state atmospheric conditions, and here the problem begins. The atmosphere is intrinsically time dependent and never steady state. The weather is continuously changing; both due to the atmospheric circulation and due to the diurnal cycle. The fluctuations and gusts in the incoming wind interact with equally important vortex shedding from buildings to control the local flow fields and the dispersion of pollutants. To overcome this problem quasi-steady situations were defined which are composed of mean values averaged over e.g. 20 or 30 min. However, as was shown in Schatzmann et al. (2009) at the example of a detailed analysis of data from urban field experiments, the time scale of the naturally occurring turbulent structures significantly exceeds such short averaging periods. As a consequence, the commonly determined short-time mean values measured inside the urban canopy layer have the character of random samples, i.e. snapshots, only. Depending on the wind direction, the variability between seemingly identical cases can be large. To simply increase the sampling time would not solve but worsen the problem since over periods longer than 30 min a systematic trend in meteorological conditions due to the diurnal cycle has to be expected. Wind tunnel and LES simulations show that many hours of constant weather conditions would be needed until the time series become ergodic and deliver always the same mean values. In view of this it has to be concluded that episodic field measurements cannot be representative at locations with highly fluctuating flow properties.

A second problem caused by the naturally occurring atmospheric variability concerns the determination of representative input conditions for the model runs. In most field experiments only one reference station exists that can be used to characterise the meteorological situa-

tion and the model input. Fortunately, however, there are a few exceptions, most prominent among them the Joint Urban 2003 experiment in Oklahoma City (Allwine et al., 2004). In this experiment the wind vector was measured with high temporal resolution at multiple locations in parallel. Two of the instruments with which the wind was measured were mounted onto poles well above the rooftops of buildings located outside the Business District and about 1 km apart from each other. The data from these two sites (PNNL and PWIDS) are often selected to determine the inflow profiles for numerical model runs. At least for southerly and westerly wind directions the flow approaching the instruments is regarded as being undisturbed.

Hertwig (2008) compared for several periods the wind velocities and directions which were measured simultaneously at the two stations. A typical result is shown in Fig. 2.1. The data are presented as they were measured, i.e. no corrections were made for the different measurement heights (PWIDS 49m and PNLL 37 m above ground). In order to reduce the scatter, averages over 5 minutes have been made.

As can be seen, there is significant variability in the wind velocities whereas the wind direction is comparatively constant. This is not always the case; other examples show just the opposite or a variation of both. Not very surprising, the curves in Fig. 2.1 differ from each other, and a correction of the data for different measurement heights would not diminish but increase the differences. Although not the raw data but already 5 min mean values are presented, the data do not change in synchrony, which indicates that there are large scale structures superimposed to the mean flow.

Assumed that 20 min averages would be used to determine the input for a model run, it becomes very obvious that the models would predict quite distinct flow and, above all, concentration fields, depending on from which measurement site the input data were chosen. In the Oklahoma field experiment more than 100 anemometers, Sodars etc. were deployed and measured in parallel. The data of many of them would likewise be suitable to be used as reference velocities. All of them deliver different model inputs, and the differences are much larger than the instrument uncertainty. A model user could select the input which leads to the best fit. To regard such a test as a serious proof for model quality would certainly not be justified.

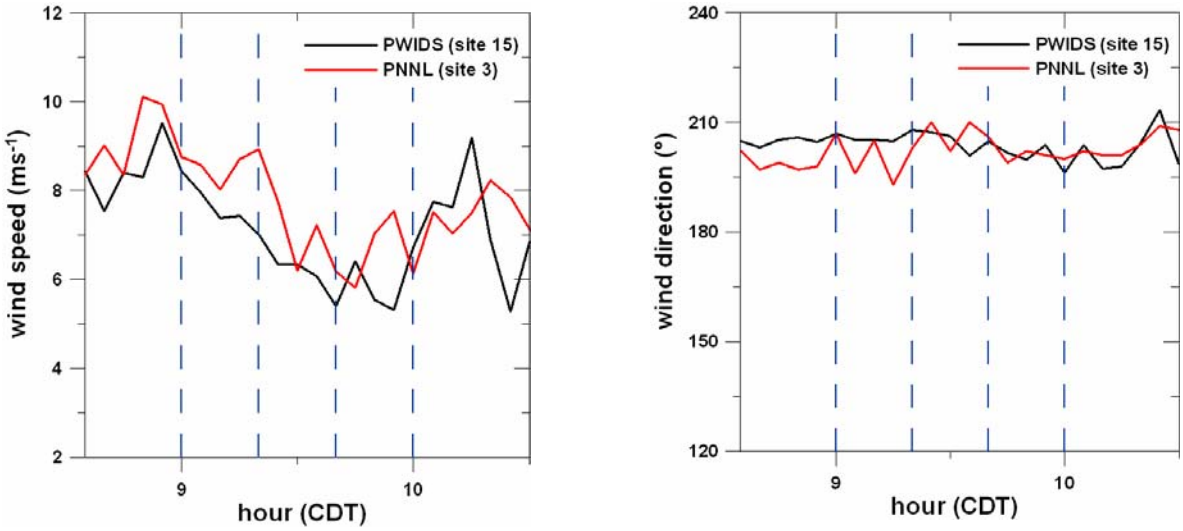



Fig. 2.1: 5 minute averaged wind speeds (left) and directions (right) measured at the stations PWIDS 15 and PNNL 3 during the Joint Urban 2003 campaign IOP3. The dashed lines mark 20 minute intervals (from Hertwig, 2008).

The ambiguity in the right choice of input data for model simulations concerns every field data set, not only those in which the model input can be selected from several meteorological towers. In experiments with one reference station only it simply remains unknown how representative this tower measurement really is. The large turbulent elements embedded in the mean flow create changes from location to location and from averaging interval to averaging interval which could be determined only if the mean weather remained constant and a sufficiently large ensemble of short time means could be collected and statistically analysed.

A third difficulty which has been encountered in the search for reliable validation data sets originates from the fact that resources are limited and field measurements usually are made at a few selected points only. A CFD model, on the other hand, requires at the inflow boundary of the domain a complete data array, and this not only for the mean velocity but also for more complex properties like the turbulent kinetic energy or the dissipation rate. Provided the measurement was made with a high-resolution instrument that delivered a velocity time series (which is usually not the case in older field data sets) this information is available for the measurement position only. It needs to be extrapolated laterally and vertically which requires additional assumptions to be employed (e.g. assessment of a roughness length, a logarithmic wind profile in the surface layer, lateral homogeneity etc.). In view of what was explained before, all these assumptions are most likely not justified over short averaging intervals but presume the existence of long periods of constant weather, the existence of a boundary layer in equilibrium with the underlying surface etc. As a consequence, the model is fed with input which most likely does not sufficiently correspond to the flow or concentration measurements which are used in the validation process to check the accuracy of the model output.

Since it is unfair to blame a model for “flaws” contained in the data and which originate from the inherent variability of the atmosphere, and since all models which were tested by the action were run in steady state mode, action COST 732 came, after lengthy discussions, to the conclusion to base the validation work predominantly on wind tunnel data (for exceptions from the rule see Chapter 4.4). In order to provide the opportunity to compare the wind tunnel data with reality, it was decided to accept only data from laboratory experiments in which a real field situation was replicated. Furthermore it was required that the wind tunnel boundary layer characteristics were known and documented in sufficient detail (see e.g. Leitl, 2000, or Pascheke et al., 2003).

## 2.2 Validation Data Sets

Since COST 732 was tasked to provide guidance for the evaluation of urban models, the action focused on experiments in which flow and dispersion processes in boundary layers disturbed by a large number of obstacles were investigated. It was assumed that participating models were already tested at simpler cases with only one or a few obstacles. For such basic tests data are available e.g. from the internet platform CEDVAL (http://www.mi.uni-hamburg.de/CEDVAL-Validation-Data.427.0.html). The use of these data was recommended already in other guidelines as, e.g. the VDI Guideline 3783, part 9 (2002).

Following the general data concept, the new test data were assembled explicitly from test campaigns for which field data as well as corresponding laboratory results were available. Regarding data quality, data density and documentation, the Mock Urban Setting Test (MUST) field experiment (Biltoft 2001, Yee and Biltoft, 2004) and the Joint Urban 2003 Oklahoma City (OKC) Atmospheric Dispersion Study (Allwine (2004)) were found to be most suitable for use as exemplary complex test cases[1]. The set-up of the MUST experiment was relatively simple. 120 ship containers of identical size were arranged in a regular pattern on flat ground in the desert of Utah. In contrast to this idealised situation, the OKC experiment which was carried out in the Business District of Oklahoma City was really a challenge for the numerical models (Fig 2.2). Only a limited number of COST participants had sufficient re-

[1] The members of the Action wish to express their gratitude for getting permission to extract and use relevant information and data from corresponding field trials. Support provided by Dr. Bowers (DPG) and Dr. Fry (US DTRA/JSTO) is gratefully acknowledged.

sources to carry through the simulations. For both field experiments corresponding wind tunnel data were available from studies carried out in the Environmental Wind Tunnel Laboratory at Hamburg University (Leitl et al, 2003, Bezpalcova and Harms, 2005, Bezpalcova, 2006).


Fig. 2.2: Pictures from the two test sites; the MUST container village in the desert of Utah (left hand side) and the Business District of Oklahoma City (right hand side).

In order to simplify both compiling the data sets and using them for validation purposes, a straightforward approach was chosen. Data documentation is provided as PDF-files and the numerical data is stored in simple Excel tables. Each of the data sets consists of four branches of information, providing

(1) **data documentation** files such as descriptive information on the experiments, data reports and related published papers relevant for understanding the experiments and using the data safely ('Documentation' folder),

(2) **geometry data** relevant for replicating the setup of urban structures in a corresponding computational domain ('Geometry' folder),

(3) **inflow data** to be used for characterizing the inflow boundary conditions at the upwind edge of the computational domain ('Inflow' folder) and

(4) the actual **test data** to be compared with the results from numerical simulations ('Test Data' folder)

Wherever applicable, the experimental data were harmonised regarding the reference coordinate systems used, the non-dimensionalisation of results and with respect to the labelling of variables. In contrast to usual field data sets which use an obstacle-referenced or geo-referenced frame, the reference coordinate systems in the COST database were defined as right-handed systems with the positive x-axis to be aligned with the corresponding mean wind direction. Whereas the location of obstacles and measurement points can be easily transferred to a fixed, geo-referenced coordinate system, the transformation of wind velocity components measured in a wind tunnel is a non-trivial task since the laboratory data has been collected from independent 2-dimensional LDA (Laser Doppler Anemometer) wind velocity measurements.

Regarding a quantitative model validation, a crucial problem remains the confidence interval of the reference data. The 'accuracy' of the reference data to be specified must include both, the uncertainty of the instrumentation used and the uncertainty due to varying boundary conditions (field tests) and finite sampling periods/ensemble sizes (laboratory and field data). For properly generated wind tunnel test data, it is possible to define the confidence interval of the experimental results based on repetitive, independent realizations of identical experi-

ments. The confidence range is expected to be bigger than the measurement accuracy of the instrumentation but must still be sufficiently small for the test data to be used for validation purposes. Defining a similar confidence interval for field data would be far more laborious or is even impossible, when only a limited set of test data is available. Chapter 4.4 discusses the specific problems in characterizing the uncertainty of field data in more detail. In the test data sets compiled, the confidence interval is provided for all reference data.

Following a careful analysis of the laboratory test data, a number of measurements were excluded from the final reference data set because of specific limitations of the instrumentation used at certain measurement locations. For example, due to the shape and the small but finite size of the measurement control volume of an LDA system, the measured U-velocity component for consecutive UV- and UW-measurements at exactly the same location differs because the measurement volume has a different orientation in space. Particularly in regions with strong velocity gradients, the differences caused by spatial averaging can be bigger than accepted for a 'reference data set'. The removal of parts of the test data from the Oklahoma City wind tunnel experiments is complicating the validation process as it excludes some model regions close to the ground from the comparison. Strictly spoken, this introduces a not yet quantifiable bias to the validation process which cannot be avoided with the data currently available.

The laboratory data were generated in a neutrally stratified boundary layer wind tunnel, thus they are clearly limited in scope. Compared to field data, on the other hand, the data come with complete inflow boundary conditions. Since the “mean weather” in the wind tunnel can be kept constant over long periods of time, data can be generated which have the same statistical relevance as the output of RANS models which are run in steady-state mode. Although it requires significant measurement effort, the experimental results can be delivered with a similarly high spatial density as the model output. This makes the results directly comparable with each other.

Whereas it was possible to combine both, field results and laboratory data, in one data set for the COST-MUST exercise, for the COST-OKC data set based on the Joint Urban 2003 field tests only the relevant wind tunnel data are provided. The corresponding field data are available at https://ju2003-dpg.dpg.army.mil . Because of the large amount of field data available, selecting, reviewing and harmonizing the OKC field data would have been a project of its own and was not feasible within an action like COST 732 with funds for networking only.

Both test data sets, MUST and OKC are published on the COST732 website under http://www.mi.uni-hamburg.de/Home.484.0.html. It is intended to maintain the data at least for the next 5 to 10 years. In addition, the reference data were included in Excel workbooks to facilitate a more efficient exploratory analysis of validation results (see chapter 4.2.2 for details).

## 2.3 Remaining Problems and Challenges

Even with access to high quality field and laboratory data sets, or just because of the high quality of the data, work performed within the Action revealed a number of problems in generating qualified validation data sets. The Action clearly proofs that validation data must not be ordinary data from usual field and laboratory experiments. If an experiment is not precisely enough designed and complete with respect to the specific input needs of a certain model type, the results require a significant amount of post-experimental interpretation. Without the 'commonly accepted opinion' developed by a larger community as it was present in COST Action 732, the use of one and the same set of (incomplete) experimental data will leave sufficient room for a subjective interpretation of the reference data, thus complicating an objective model validation. If, for example, in the MUST wind tunnel test the importance of turbulent flux measurements would have been known earlier and given a higher priority, the missing data could have been measured directly. Similarly, for the Joint Urban 2003 field campaign it would have been desirable to improve the experimental characterization of the

approach flow conditions by additional, goal-oriented field measurements. However, since the desired information is not available, there are gaps in data coverage which can not belatedly be fixed anymore. As the discussions within the COST 732 evidenced, the best platform for preparing reference data sets is a community consisting of both numerical modellers and experimentalists. Both groups have different experiences with data. They can control but also help each other in gaining insight into the data problems and into their potentials and shortcomings.

Regarding complex, urban-type flow and dispersion situations, the number of potential validation data sets is still very limited. There is a severe lack of test cases representing an 'intermediate complexity' between idealized urban roughness configurations such as the MUST experiment and individual test cases of high geometrical complexity such as the Joint Urban 2003 trials. If a model 'passes' the MUST test and fails in replicating the very complex OKC situation, it does not necessarily mean that the model would not deliver a decent simulation of a more uniform urban building structure as it is typical for most of the European cities. For example, test configurations including different complexity of terrain or cases with sudden changes in the urban roughness pattern (lake/river within a city) are basically missing.

With the continuously increasing computer power and better spatial resolution of numerical models, a precise enough definition of the obstacle/building geometry is found to become more important. Experience with the existing data sets shows that for example GIS building data, a wind tunnel model and the real urban structure may not be in sufficient agreement with each other. Whereas the geometric simplification of buildings introduced by GIS and the geometric detail of a wind tunnel model can be sufficiently matched, the exact location of the real buildings is sometimes difficult to obtain. A higher accuracy in building data and their provision in a sufficiently accurate form is crucial for future test cases.

An even bigger challenge in generating qualified validation data will be the provision of test data suitable for testing models exceeding the capabilities of RANS models. Eddy resolving flow and dispersion models (LES) will require not just mean flow and dispersion data to be provided but will need access to time resolved, spatially correlated reference data.

# 3 Models participating in the exercises

## 3.1 Non-CFD Models

The use of non-CFD models, sometimes called in the literature operational or integral dispersion models, is widespread for urban air quality studies and dispersion problems. That is the reason for applying COST 732 validation measures to non-CFD models, even though the main interest of COST Action 732 has been towards CFD models. Non-CFD models do not explicitly resolve fluid-dynamics equations but physical processes are parameterized. For example, non-CFD models accommodate flow and dispersion aspects of the urban canopy in a statistical manner rather than to resolve them explicitly. As stated in the Background Document, non-CFD dispersion models do not calculate the flow around a single building or obstacle when applied to an urban-like geometry, but the effect of a group of obstacles is taken into account through an increased surface roughness value or by a coarse resolution of the buildings. This is also typically done in many meteorological mesoscale models to represent urban areas.

It is recognized that non-CFD dispersion models are in general less complex and easier to run than typical CFD models and require much shorter calculation time. One main advantage is that non-CFD models may be run over a long series of input data to represent different meteorological conditions. The validation procedure for a non-CFD model requires that (i) the purpose of its application has been clearly stated, (ii) metrics for “state of the art/science” models when applied to particular data sets are available, and (iii) a specified protocol for the specific application is available.

In this context the recommendations of the Guidance and Protocol Document which refers mainly to CFD models are followed as much as possible for non-CFD models. In particular the performance of the non-CFD models in a certain setting according to different statistical measures is investigated. There is a large variety of non-CFD approaches (from very simple to relatively complex), and they are applied for different purposes. Therefore the Best Practice Guideline, which is mainly intended for CFD models, is only in parts applicable to the simpler non-CFD models.

### 3.1.1 -45 degree MUST wind tunnel experiment

Two main types of non-CFD models are applied to simulate the -45° MUST case: *Gaussian type models* of varying type and *Lagrangian type models* (including Puff models).

In particular, the Gaussian type models ESCAPE, ADMS, OML, Lagrangian puff model CALPUFF and the Lagrangian particle dispersion model LASAT took part in the exercise. These models are currently used in Europe and elsewhere for several dispersion type applications. A brief description of these models is presented in the following.

The **ESCAPE** model (http://www.fmi.fi/research_air/air_55.html) evaluates the releases, source terms and atmospheric dispersion of dangerous materials. It is applicable to both continuous and instantaneous releases of toxic and flammable gases. The released substance and the location of a release point are required as input data. Meteorological input data has to include wind speed, wind direction, atmospheric stability, ambient temperature and the relative humidity of ambient air.

The ESCAPE-model is not applicable to the estimation of the atmospheric dispersion of pollutants emitted from typical fires in warehouses and chemical stores. Terrain effects are only considered in their effect on the atmospheric conditions through the surface roughness parameter. The model is not suitable for transient releases.

The **OML** model is a modern Gaussian plume model (based on boundary layer scaling instead of relying on Pasquill stability classification) intended to be used for distances up to about 20 km from a number of sources (http://oml-international.dmu.dk). The model can be used for both high and low sources; it is not suitable for applications in complex terrain. The model requires information on emission and meteorology on an hourly basis.

The handling of building effects is based on simple methods, whereas in reality, aerodynamics in the wake of a building is an extremely complex matter. The primary intent of the building effect algorithm used in OML-Point is to improve concentration estimates applicable for distances beyond approx. five building heights downwind.

**ADMS-Urban** (CERC, 2006; http://www.cerc.co.uk) is an advanced atmospheric pollution dispersion model for calculating concentrations of atmospheric pollutants emitted continuously from point, line, volume and area sources or intermittently from point sources. It is able to describe in details what happens in a range of scales, from the street scale to the city-wide scale. It incorporates the latest understanding of the boundary layer structure. The model uses advanced algorithms for the height-dependence of wind speed, turbulence and stability to produce improved predictions. The model also includes algorithms which take into account: downwash effects of nearby buildings within the path of the dispersing pollution plume; effects of complex terrain; wet deposition, gravitational settling and dry deposition; short term fluctuations in pollutant concentration; chemical reactions; pollution plume rise as a function of distance; averaging time ranging from very short to annual. The system also includes a meteorological data input pre-processor. Included in the model is also a linearised perturbation model which allows for the calculation of flow fields over changing terrain or roughness.

**CALPUFF** (http://www.src.com/calpuff/calpuff1.htm) is a multi-layer, multi-species, non-steady state puff dispersion model which can simulate the effects of time- and space-varying

meteorological conditions on pollutant transport, transformation and removal. CALPUFF can use the three-dimensional meteorological fields computed by the CALMET model, or simpler, single station data with varying degree of complexity. CALPUFF contains algorithms from near-source effects such as building downwash, transitional plume rise, partial plume penetration and sub grid scale terrain interactions, as well as longer range effects such as pollutant transformation, vertical wind shear, and overwater transport. Most of the algorithms contain options to treat the physical processes at different levels of detail, depending on the model application. Some examples of applications for which CALPUFF may be suitable include: near-field impacts in complex flow or dispersion situations (complex terrain, stagnation, inversion, recirculation, and fumigation conditions, overwater transport and coastal conditions; light wind speed and calm wind conditions), long range transport, visibility assessments and class I area impact studies, criteria pollutant modelling, including application to State Implementation Plan (SIP) development, secondary pollutant formation and particulate matter modelling, buoyant area and line sources (e.g., forest fires and aluminium reduction facilities).

The dispersion model **LASAT** (Janicke Consulting, 2007; http://www.janicke.de) simulates the dispersion and the transport of a representative sample of tracer particles utilizing a random walk process (Lagrangian simulation). It computes the transport of passive trace substances in the lower atmosphere (up to heights of about 2000 m) on a local and regional scale (up to distances of about 150 km). A number of physical processes, including time dependencies, are simulated, such as transport by the mean wind field, dispersion in the atmosphere, sedimentation of heavy aerosols, deposition on the ground (dry deposition), washout of trace substances by rain and wet deposition, first order chemical reactions. The quality of the results achievable by Lagrangian models mainly depends on the wind field they are based on, as will be shown in this investigation.

In contrast to CFD models flow within non-CFD models is not resolved at the building scale. Generally, different grades of approximations and simplifications to the primitive equations are used when calculating concentrations. Inherent to all Gaussian type models (even if a boundary-layer parameterization is included such as within the ADMS model family) is their inability to allow for meteorological changes (direction, speed, atmospheric stability) within space and the time interval in which the concentration field is calculated (usually half an hour or one hour) so that a steady state for this period is assumed. Gaussian type models might incorporate chemical models of various complexities and may allow for a large number of sources making them suitable for large air quality studies. Lagrangian type models can use meteorological input data from one (representative) station only, but usually use three-dimensional wind fields. Thus meandering and other shear effects as well as topographical effects can in principle be simulated.

Both types of models have strengths and weaknesses depending, as already stated, on the application. In this exercise the models where used by several groups and in several setups. Details are given in Table 3.1.

### 3.1.2 Choice of model configuration with respect to obstacles

In non-CFD models, buildings are taken into account in the model domain in several ways. The simplest way is *increasing the roughness length* by using a single spatially-averaged value or using a distribution of roughness lengths. The choice of the roughness length has a significant effect on the calculated concentrations. In the case a distribution of roughness values is used, a change of roughness and a certain degree of flow distribution is allowed that affects concentration results. An overview of non-CFD model runs for the -45° MUST experiment is given in the following table. Apart from model names and types as well as the institutions running the models, this table includes also information on the roughness length used or the incorporation of obstacles.

| model type | Model | Group | z0 [m] | obstacles | comment |
|---|---|---|---|---|---|
| Gaussian | ADMS a | ZAMG | 0.1 | 1 'effective' building | |
| | ADMS b | ZAMG | 0.1 | no | |
| | ADMS c | Universita del Salento | 0.269 | no | 1) The area occupied by the buildings is replaced by a single value of surface roughness $z_0$=0.269m, calculated by morphometric method (Macdonald et al, 1998) 2) The same $z_0$ as in wind tunnel experiments is used in the area upwind of the buildings. |
| | ADMS d | University of Cyprus | 0.381 | no | |
| | ESCAPE | FMI | 1/10 building height | no | |
| | OML a | NERI | 0.037 | no | For illustration purposes: Roughness based on wind profile above the source |
| | OML b | NERI | 0.3 | no | Realistic roughness, based on profile in the middle of the building array. |
| Puff model | CALPUFF | SHMU | building area: 0.1 | no | Roughness length based on Lettau formula |
| Lagrangian | LASAT a | ZAMG | 0.1 | field of obstacles | grid according to north-east direction, diagnostic wind field with building parameterization |
| | LASAT b | ZAMG | 0.1 | field of obstacles | grid parallel to building walls, diagnostic wind field with building parameterization |
| | LASAT c | ZAMG | 0.1 | field of obstacles | grid parallel to building walls, wind field calculated with MISKAM |

Table 3.1 Overview of non-CFD model runs and input parameters for the -45° MUST case. Additional information on the participating groups and model runs can be found in Appendix A.

Different groups have represented the buildings in several ways according to the possibilities of the model used. For example **ADMS** was used by 4 groups and different results were obtained according to how buildings were chosen to be represented.

In the model run *ADMS a* buildings were represented by using the 'Building module'.

A validation of the *ADMS building module* is described in Carruthers et al. (1999). The buildings module in ADMS is able to consider up to 25 buildings in one model run. For each source, the program combines the buildings to a single effective building for each wind direction. The height of this is equal to the height of the 'main' building for this source which is predefined by the user. The disturbed flow field consists of a re-circulating flow region (cav-

ity) in the lee of the building with a diminishing turbulent wake downwind. The influence of the buildings is considered in the dispersion calculation out to a distance of about 60 building heights (up to about 130 m in the MUST case). Concentrations within the well-mixed cavity zone are uniform and calculated as a fraction of the release which is entrained. The concentrations further downwind are calculated as the sum of a ground-level plume from the cavity zone and the elevated plume from the non-entrained remainder. The turbulent wake reduces the plume height and increases the turbulent spread.

For sensitivity analysis, model run *ADMS b* is conducted without any consideration of the obstacles – neither using the ADMS building module nor increasing the roughness length.

In model run *ADMS c* the individual buildings of the array were not explicitly modelled, but the area occupied by the buildings was parameterized using a single value of the surface roughness $z_0$. In particular, the same $z_0$ as in the CFD modelling in the area upwind of the array was used, and $z_0$=0.269m in the area occupied by the array. This value is derived from the morphometric method proposed by Macdonald et al. (1998):

$$\frac{z_d}{H} = 1 + \alpha^{-\lambda_p}(\lambda_p - 1)$$

$$\frac{z_0}{H} = (1 - \frac{z_d}{H}) \exp\left\{ -\left[ 0.5\beta \frac{C_D}{k^2}\left(1 - \frac{z_d}{H}\right)\lambda_f \right]^{-0.5} \right\}$$

$z_d$ indicates the zero-plane displacement height; $C_D$ is the drag coefficient (1.2), α an empirical coefficient (4.43) and β a correction factor for the drag coefficient (1.0). $\lambda_p$ and $\lambda_f$ are the plan area index and the frontal area index, respectively. The estimated value for $\lambda_p$, which is independent from the wind direction, is 0.095. For the 0° approaching flow case, $\lambda_f$ is equal to 0.101. For the -45° case, by projecting the frontal areas of the buildings along the wind direction axis, we have estimated a value of $\lambda_f$ equal to 0.085.

The structure of the buildings array is too complex to be handled explicitly in the **OML** model, so the presence of buildings has been simulated simply by using *an appropriate (large) roughness length*. The roughness and other meteorological parameters were deduced from wind tunnel measurements as explained in detail below.

Vertical profiles of wind speed were available at 18 locations. Data from these profiles were considered for heights above 5m (e.g., approximately 2 building heights). Subsequently, for each tower the values of $z_0$ and $u_*$ were calculated as the best fit to a logarithmic wind profile (no displacement height). The roughness length determined in this manner is not the same for all profiles, because an internal boundary layer grows from the upwind edge of the building array. This is illustrated in Figure 3.1 It appears that a representative roughness value for the middle of the building array is 0.3 m.

To illustrate the effect of roughness, two runs with OML were performed: One with data derived from the source location, and another with data estimated from the middle of the array. The latter is the most realistic when the intention is to obtain an estimate of concentrations within the building array. The data for the two runs were:

a) Based on vertical profile at the source (not representative for the pollution plume). $z_0$=0.037m; $u_*$=0.521m/s; $u_*/u_{ref}$ =0.076; Heat Flux=0. The reference velocity $u_{ref}$ was calculated from a log profile at z=7.29m.

b) Based on vertical profile in the middle of the buildings' array. $z_0$=0.3m; $u_*$=0.812m/s; $u_*/u_{ref}$ =0.125; Heat Flux=0. The reference velocity $u_{ref}$ was calculated from a log profile at z=7.29m.

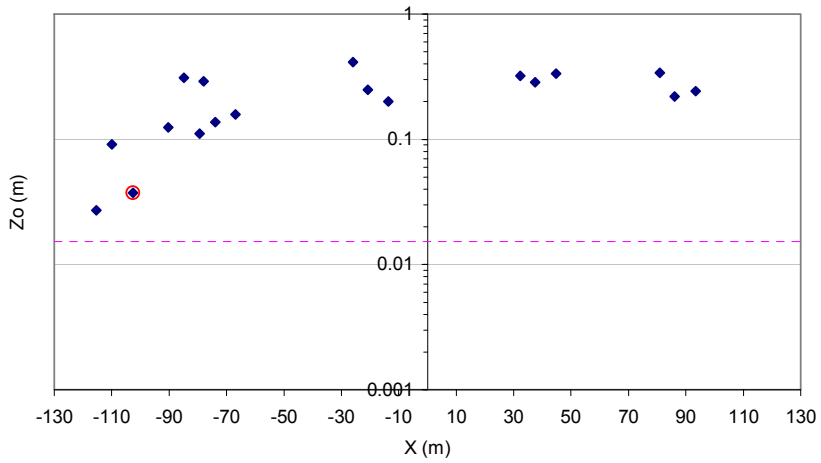



Figure 3.1: $z_0$ derived from 18 vertical profiles. The horizontal axis indicates the x coordinate of the profile. The approach flow comes from the left, so that the leftmost point (x=-115) represents a profile in the first row of buildings. The point with a circle represents a profile at the source location, which is close to the upwind edge of the building array. Within the middle of the building array $z_0$ reaches a relatively constant value of 0.3 m. The dashed line indicates $z_0$ of the approach flow.

The **CALPUFF** model run was performed using single station data including boundary layer parameters calculated with the diagnostic wind field model CALMET and the vertical profile of wind and turbulence based on measured data at the emission source position.

Although CALMET is capable of computing three-dimensional wind fields in complex terrain, it is technically not capable to incorporate rectangular buildings on the fine mesh. Therefore, a CALMET simulation based on the approach flow wind profile, using $z_0$=0.02 outside the building area and $z_0$ = 0.1 inside, based on the formula by Lettau, was used to determine the boundary layer parameters u*, L and h inside the building area.

For CALPUFF the sensitivity to the roughness length and grid resolution has been tested. CALPUFF model runs for the -45° MUST case were performed for roughness lengths $z_0$ = 0.05 m, 0.15 m, 0.2 m, 0.3 m and 0.5 m. Maximum concentrations varied only by about 5%. The maximum concentration for $z_0$ = 0.5m was closest to the observed maximum. However, all the statistical parameters have been considerably worse than the results for $z_0$= 0.1 m (Hit rate = 0.28, FA2 = 0.20). It seems that $z_0$ calculated using the Lettau formula quite adequately represents the roughness caused by the obstacles in the MUST case. A domain with 30 x 30 cells in three 10m thick layers has been used in the simulation.

The full building configuration of the MUST experiment can be parameterized with Lagrangian models like **LASAT** based on a diagnostic wind field model which includes a building module (Janicke Consulting, 2007). The choice of how to incorporate buildings thus depends crucially on the model used. In LASAT, the wind field in the wake area is calculated from the exact building shapes. Here it is important that building fronts not aligned with the calculation grid are defined as such and not as a zigzag line following the coordinate directions. In model run *LASAT a* a grid according to the usual North-East orientation and in model run *LASAT b* a grid aligned to the building walls was used. A limitation of the building module is that the calculation assumes that every building is subject to an undisturbed incoming flow, implying that situations with one building located in the close wake of another building are not modelled correctly. LASAT is verified among others according to the guideline VDI 3945 Part 3.

An advanced way to apply LASAT for the MUST case is to base the Lagrangian dispersion calculation on CFD flow fields which are previously simulated with MISKAM (model run *LASAT c* in this study).

### 3.1.3 Morphometry analysis to derive inputs for non-CFD dispersion models

In this section, a morphometry analysis is proposed in order to derive inputs for non-CFD dispersion models in the OKLAHOMA exercise as well as to obtain some building statistics for this urban site. The methodology is described in some details and some comments are made on the use of those derived parameters within non-CFD models

Only one non-CFD model is applied, i.e. ADMS-Urban. Similarly to the methodology followed in the MUST exercise, to model the OKLAHOMA case study, individual buildings are not explicitly represented as in the CFD modelling, but the area occupied by the buildings is replaced with a single value of surface roughness. This is equivalent to model the flow over a change of surface roughness, which can be either a step change or a smooth variation of surface roughness. The latter can be considered as a caveat within non-CFD models to take into account the effect of flow adjustment over the urban surface. In the OKLAHOMA exercise both methods are used.

#### **3.1.3.1** Methodology for the calculation of $z_0$

The calculation of the surface roughness is based on the construction of the Digital Elevation Model (DEM). The methodology followed for the DEM construction and analysis is based on buildings' height information. DEM can be represented as a raster image that can be easily interpreted using for instance Matlab© (MATLAB, 1997) image analysis Toolbox. To obtain DEM for OKLAHOMA several steps can be recognized.

The starting point was the CAD geometry of the study area. For the present study, a DWG file was available from the COST Action 732 itself. From the building analysis, it was possible to reconstruct the full DEM in a gray scale that corresponds to the second step of this methodology.

The initial OKLAHOMA 3D CAD map was in fact redrawn on a new vector layer, where the basic units to be mapped were structures of different heights. Every building was described by a 2D polygon with base and rooftop elevation attributes. As information about elevation was lost, the height value available from the DWG file was assigned to each building, one by one.

Vector DEM of the study area was so generated and then accessed by ESRI ArcView Gis 3.2 by means of appropriate conversion tools, such as CAD Reader, EDTools, Spatial Analyst and Spatial Tools. ArcView allowed to interpolate vector data, retaining the original value of the sample data points in the resulting surface, and to obtain DEM in raster form, i.e. in the appropriate form to apply image processing technique.

Indeed, the computation of urban morphometric parameters by the Matlab© Image Processing Toolbox needs:

- images represented by square matrices;
- pixels coded using 8 bits of information, and the objects i.e. the buildings, defined by a pixel-cluster with specific size and shape showing a low internal grey level variation and a strong variation in its close neighbourhood;
- matrices of the same map with various resolutions, to improve the sensitivity of the method.

The last issues was solved using a software for digital image processing.

The final product, i.e. the raster DEM, was 2D, static, greyscale images, where each pixel was related to building height and 256 level of brightness were allowed. The value 255 (displayed as white on a raster device) was assigned to the road level (0m), instead the value 0 (i.e. black) was assigned to the maximum value of the height in relief in the study area (Figure 3.2)

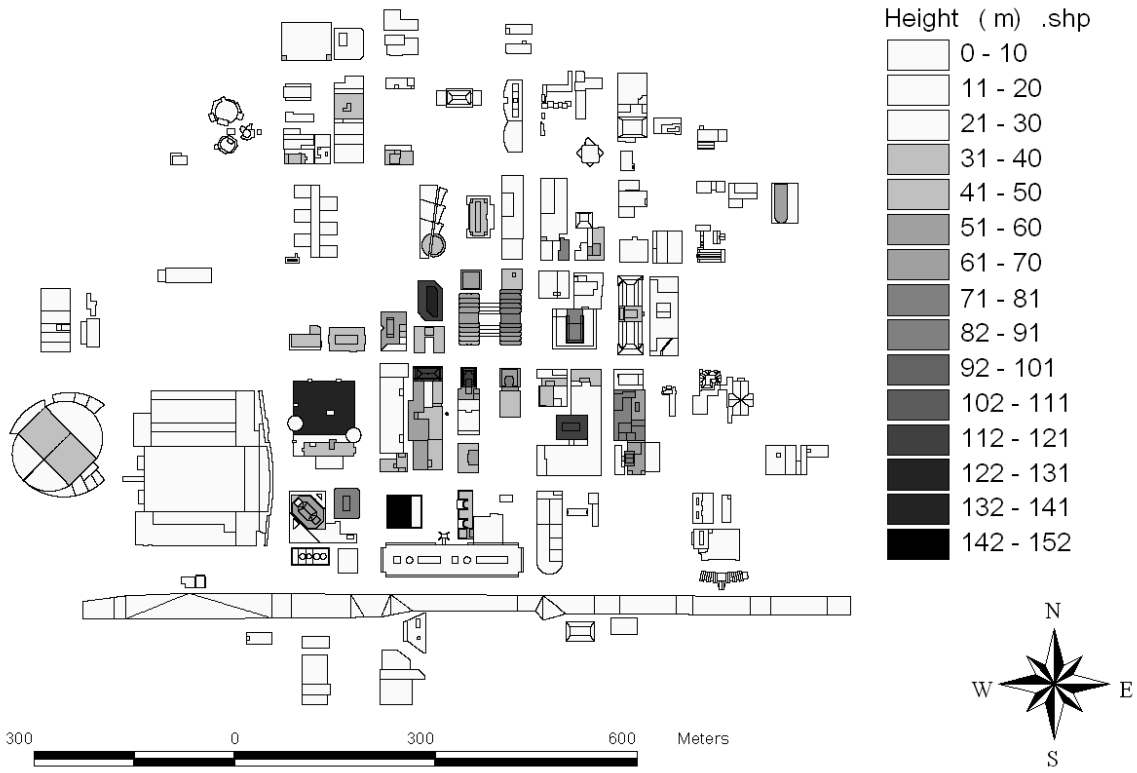



Figure 3.2: DEM of the area of OKLAHOMA City

The next step was that of estimating morphometric parameters by using image analysis performed in the MATLAB environment.

As for the MUST case procedure adopted for non-CFD calculations within ADMS-Urban c. Image processing algorithms used in the OKLAHOMA case were based on those originally developed and discussed in detail by Ratti et al. (2002, 2006). In particular $\lambda_p$, $\lambda_f$ and $z_H$ (height weighted with the values of building frontal area) (Grimmond and Oke, 1999) parameters were calculated as:

$$\lambda_p = \frac{\sum_i A_{p,i}}{A_T}$$

$$\lambda_f(\theta) = \frac{\sum_i A_{f,i}(\theta)}{A_T}$$

$$z_H(\theta) = \frac{\sum_i H_i A_{f,i}(\theta)}{\sum_i A_{f,i}(\theta)}$$

where $H_i$ and $A_{p,i}$ are respectively the height and the planar area related to ith building, while $A_{f,i}(\theta)$ is its frontal area exposed to a wind along θ direction; $A_T$ is the total site planar area.

Using formulae above $\lambda_p$ from the OKLAHOMA DEM is equal to 0.19. It should be reminded that $\lambda_p$ does not depend on wind direction, while $\lambda_f$ does as shown in Figure 3.3 (left). Wind tunnel experiments were performed with winds coming from the left in Figure 3.2, corresponding to wind from the South. The corresponding $\lambda_f$ value for this wind direction is equal to 0.15.

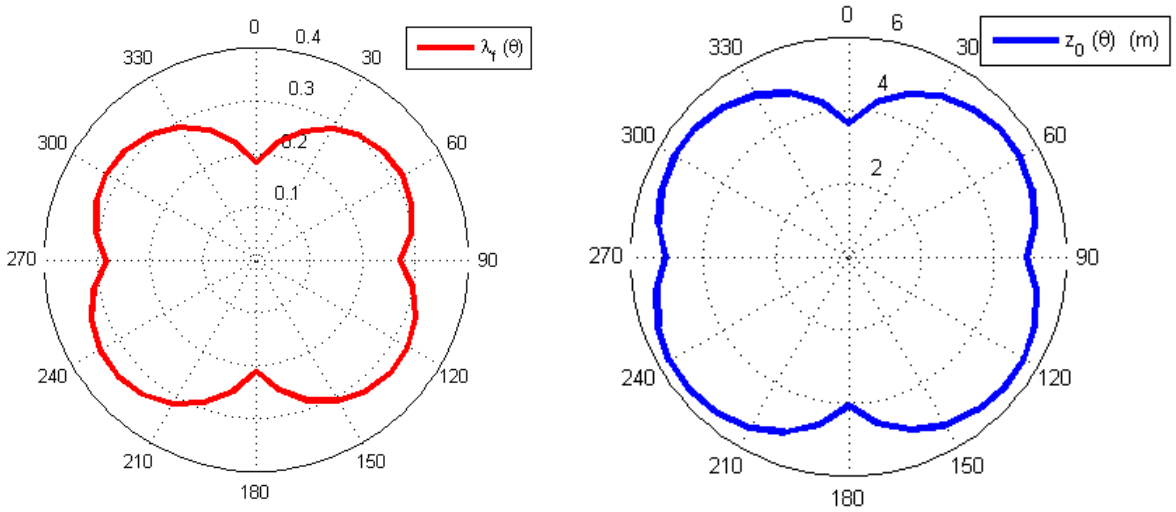



Figure 3.3: Polar diagram showing the variation of $\lambda_f$ (left) and $z_0$ (right) with wind direction.

Starting from these parameters, $z_d$ and $z_0$ were calculated using the equations described in the previous section (Macdonald et al., 1998). The $z_0$ value strongly depends on the value of the mean building height in those equations which is univocally defined for cube arrays but not for a real city as it depends on the planar area chosen to calculate it (Figure 3.3, right). Table 3.2 shows all building statistics for the OKLAHOMA DEM.

| $A_T$ | $\lambda_p$ | $\lambda_f$ | $\overline{H}$ (building mean height) | $H_{max}$ | $\sigma_H$ | $z_d$ | $z_0$ |
|---|---|---|---|---|---|---|---|
| 1300m x 1050m | 0.19 | 0.15 | 32 | 152 | 32 | 12.5 | 3 |

Table 3.2: Summary of the building statistics for OKLAHOMA. $\sigma_H$ is the standard deviation of building heights.

It is worth mentioning that planar and frontal area indeces have been defined for array made of cubic buildings of uniform height as discussed before. Indeed real city are made of buildings arranged irregularly and characterized by a non uniform height. Macdonald's formulae are strictly valid for cubes arrays but doubts remain about the application to complex geometries like real cities. For this reason, it might be preferable to extract $z_0$ and $z_d$ parameters

as a function of elevation z in Macdonald's formulae. Therefore the above procedure should be revised when applied to real urban geometries such as the OKLAHOMA case.

First of all it is advisable to use, instead of the mean value of $\lambda_f$ , its variation with the elevation that is $\lambda_f(z)$ defined as:

$$\lambda_f(z) = \frac{\sum_i dA_{f,i}(z)}{A_T}$$

Given a constant increment of elevation $\Delta z$, $dA_{f,i}(z)$ indicates the portion of the building frontal area in the region between z and z+ $\Delta z$. In other words, $dA_{f,i}(z)=W_{,f,i}\,\Delta z$ where $W_{f,i}$ is the width of the ith building at the elevation z and in the direction perpendicular to the wind. Obviously, the sum of $\lambda_f(z)$ values on all elevations z gives the total frontal area index. A similar procedure can be applied also to calculate $\lambda_p(z)$. Figure 3.4 shows the calculated $\lambda_f(z)$ and $\lambda_p(z)$ profiles.

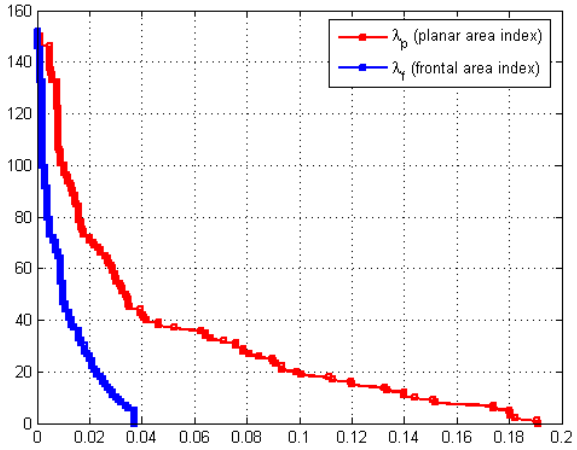



Figure 3.4: Profiles of $\lambda_f(z)$ and $\lambda_p(z)$.

Similarly to the methodology applied for $\lambda_f(z)$, given a constant increment of elevation $\Delta z$, $z_0(z)$ indicates the value of $z_0$ in the layer between z and z+ $\Delta z$. If an increment $\Delta z$ approximately equal to the average width $\overline{w}$ of buildings is considered, what the incoming wind "sees" within each single layer is a geometry similar to cubes arrays. Consequently, the Macdonald's formulae might be considered more adequate when applied to layers of height equal to $\overline{w}$ . Obviously, the average of $z_0(z)$ values on all elevations z gives the total roughness length related to the study area.

Following the above methodology, $z_d(z)$ and $z_0(z)$ values were calculated for OKLAHOMA city by choosing $\Delta z$ approximately equal to the average width of buildings.

In order to explain how $\overline{w}$ was calculated, one may start by considering the ith building of height $H_i$, length $L_i$ and width $W_i$. Its planar area $A_{p,i}$ is $L_i\,W_i$. For simplicity, one may assume that the building is positioned in such a way that the frontal area is just $A_{f,i}(\theta) = H_i W_i$ (this simplification is not too stringent because one can always find a wind direction θ for which most buildings have a façade in the direction perpendicular to the wind). Similarly, for the direction perpendicular to θ, identified as $\theta_\perp$, the frontal area $A_{f,i}(\theta_\perp)$ is $H_i\,L_i$. Then one can replace $A_{p,i}$ and $A_{f,i}$ with those last expressions within the $\lambda_p$ and $\lambda_f$ original formulations. In this way, it is easy to recognize that:

$$\frac{\lambda_p}{\lambda_f(\theta_\perp)} = \frac{\sum_i L_i W_i}{\sum_i H_i L_i} = \frac{\overline{W_L}}{\overline{H_L}}$$

where $\overline{H_L}$ and $\overline{W_L}$ are respectively the average height and the average width of buildings weighted with their length. Instead of $A_f(\theta_\perp)$, one may consider the profile $A_f(z, \theta_\perp)$ with a unitary increment ($\Delta z$=1m) and calculate its value in the first layer, that is between z=0 m and z=1m, where all buildings (or more precisely a portion of them) are seen by the wind. In this way the above ratio becomes:

$$\frac{\lambda_p}{\lambda_f(z=0,\theta_\perp)} = \frac{\sum_i L_i W_i}{\Delta z \sum_i L_i} = \overline{W_L}$$

This allowed to define $z_0$ as function of H(z), $\lambda_f(z)$ and $\lambda_p(z)$ in Macdonald's formulae. This only apparent complicated methodology allows for refinement of $z_0$ values obtained by Macdonald's formulae. This is important as $z_0$ is probably the most important parameter within non-CFD models. The methodology applied the OKLAHOMA case gave a $z_0$ value equal to 2.5m.

#### 3.1.3.2 *Results and discussion*

Dispersion calculations for the OKLAHOMA case using ADMS-Urban were performed in full scale. The point source was positioned at (x,y,z) = (-6.22m, -80.08m, 4m) with an internal diameter of 2.1m. $z_0$=0.24m was used in the area upwind of the array, while in the area occupied by the array we used both $z_0$=3m and $z_0$=2.5 m. As mentioned in Section 3.1.3 three different runs were performed. The difference between the runs is mainly the way we modelled the surface roughness change. As in the case of ADMS-Urban c one can make use of the FLOWSTAR model within ADMS-Urban and carry out dispersion calculations over a change of roughness using a linerarised perturbation theory approach similarly to that used for flow over hills (Jackson and Hunt, 1975).

Concentrations were normalized as follows:

$$C^* = \frac{C U_{ref} H^2}{Q}$$

where C is the calculated concentration, Q the emission rate and H=300m. The characteristics of the three runs were:

1) the first run was done as a simple roughness step change from $z_0$=0.24m to $z_0$=3m

2) the second run was performed as a smooth change of roughness using Macdonald's formulae. That is instead of a single value of $z_0$ we use a smoothed surface of $z_0$ as illustrated in Figure 3.5 with the maximum value of $z_0$=3m

3) The third run is similar to the second run but with $z_0$=2.5 m

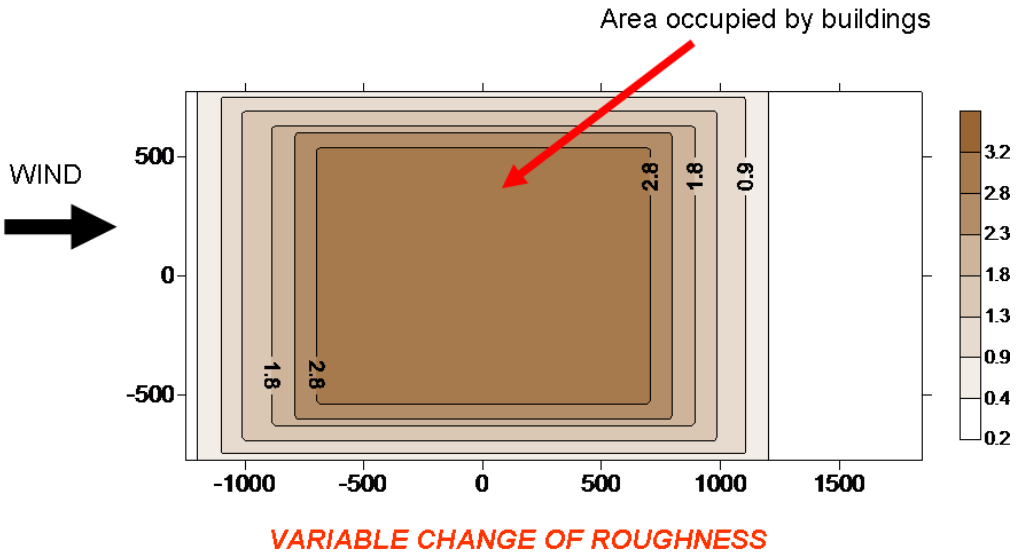



Figure 3.5: Smooth change of roughness method applied to ADM-Urban simulations.

As an example, Figure 3.6 shows dimensionless concentration profiles along the y direction (perpendicular to the wind) at three heights, namely z=3m, z=6m and z=12m. ADMS-Urban results from the three runs are compared with wind tunnel data.

Qualitatively, the shape of contours is similar for the three runs at all heights, even though maximum values are different. In particular, near the ground, the step change run gave the maximum value along the centreline, while the other two runs are very similar to each other. At larger heights the situation is reversed, with maximum values experienced by the smoothed roughness runs. We should keep in mind that in ADMS-Urban simulations the source was positioned about four times higher than in wind tunnel as the method just used would require a source height larger than the roughness length. On the other hand wind tunnel measurements refer to a street canyon. Larger concentrations should be expected here. It would be worth to combine the methodology above with a street canyon module for a point source. This is not available within ADMS-Urban but could in principle be implemented. This would improve the results substantially.

Results obtained with the smooth change of roughness are in better agreement with measurements even though only roughly. These calculations show an example of what can be expected from a non-CFD model. The value of the maximum is captured somehow but not the width of the plume. On the other hand the very rough comparison with CFD calculations in similar positions show that CFD models describe the width of the plume but not in this case the value of the maximum. Those comparisons are not sufficient to make a model validation but only hint on some problems that can be encountered in applying different models.

Non-CFD models have some limitations when used in the frame of dispersion studies for complex urban environment. Application of those models to the relatively simple MUST case allowed us to make statements about the validity of the results of those case.

The MUST case exercise showed that non-CFD models when applied to answer specific questions (maximum concentration, concentration decay from the source) may be used with a certain degree of confidence. The change of roughness proved to be one of the best approach. In fact, ADMS-Urban performed well when used in this way.

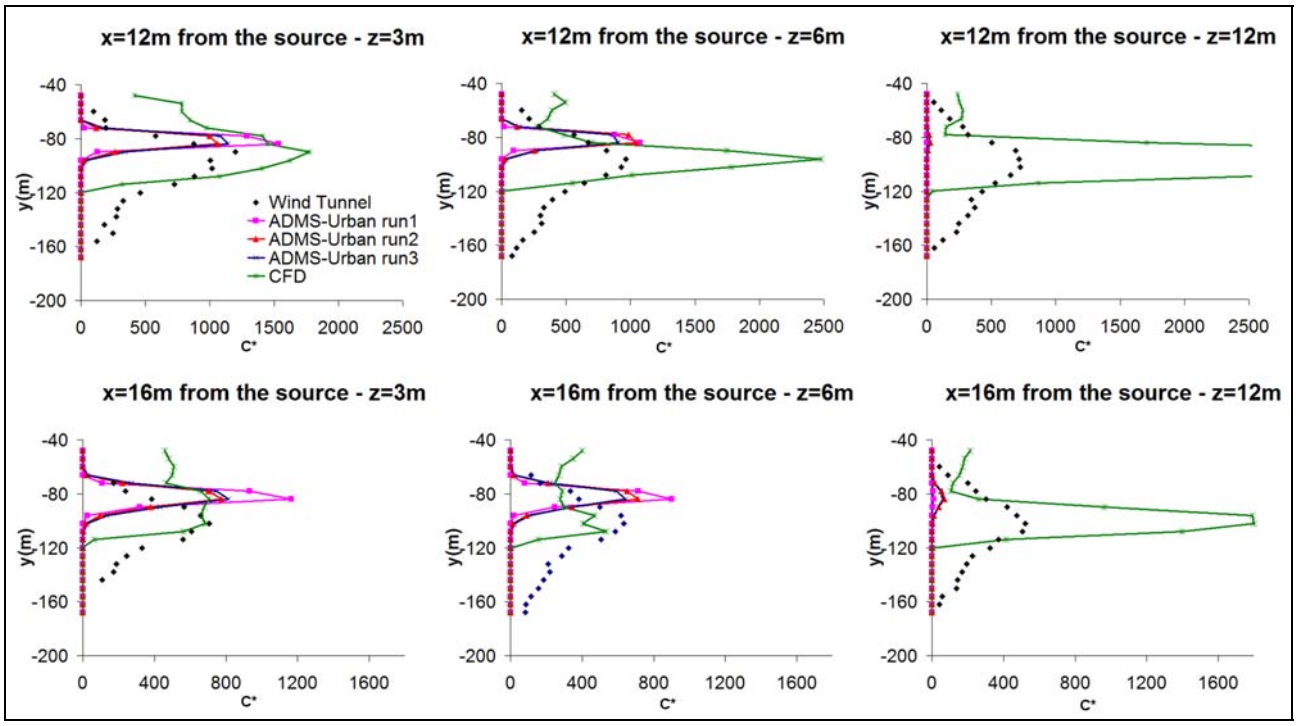


Figure 3.6 Concentration profiles along the y direction at z=3m (left), z=6m (middle) and z=12m (right) for the three runs.

The change of roughness approach is believed to be more critical for real cities, while the step change approach is too restrictive for real cases as the OKLAHOMA one. The analysis of building morphometry is important for the derivation of the roughness variation through the city.

Current formulae do not take into account building height variability and we are currently addressing this problem. However comparison of presented results with measured quantities are still preliminary and verification is needed before making final statements about the proposed formulas for the surface roughness.

## 3.2 CFD RANS and micro-scale obstacle accommodating Models

The main focus of the COST 732 action is the validation of CFD and micro-scale obstacle considering models. Contrary to non-CFD models described above, CFD and micro-scale obstacle accommodating models are in principle able to resolve the 3D flow and pollution dispersion around individual buildings in a prognostic way.

The models that were used in the validation exercises are briefly introduced in the following chapters together with details on the simulation set ups. More details on the individual model set up can be found in Appendix A. All simulations used the Reynolds Averaged Navier Stokes (RANS) equations to compute the statistically steady flow and concentration distributions. In Table 3.3 the research groups and models used for the MUST validation exercise are listed and Table 3.4 provides the corresponding information for the Oklahoma city case.

| PARTICIPANT | NUMERICAL CODE | TEST CASES |
|---|---|---|
| Aristotle University, Laboratory of Heat Transfer and Environmental Engineering, Greece (AUTh - LHTEE). | CFX | 0° case, -45° flow, -45° dispersion |
| Budapest University of Technology and Economics, Department of Fluid Mechanics, Hungary (BME). | FLUENT, MISKAM | |
| CEREA, Université Paris Est. | CODE-SATURNE | |
| Danish Meteorological Institute, Denmark (DMI). | M2UE | |
| Department of Fluid - and Thermodynamics, University of Siegen, Germany (USG). | FLUENT | |
| Dipartimento di Scienza dei Materiali, University of Lecce, Italy (USL). | FLUENT | |
| Environment Department, Research Centre for Energy, Environment and Technology (CIEMAT), Madrid, Spain. | FLUENT | |
| Helsinki University of Technology, Finland (TKK). | FINFLO | |
| (NERI), Denmark. | MISKAM | |
| University of Bielsko-Biala, Poland (UBP). | STAR-CD | |
| Central Institute for Meteorology and Geodynamics (ZAMG), Vienna, Austria. | MISKAM | |
| Department of Mechanical Engineering, University of Western Macedonia, Greece (UOWM). | ADREA | 0° case, -45° flow, -45° dispersion |
| | STAR-CD | 0° case, -45° flow |
| University of Aveiro, Portugal (UAVR). | VADIS | -45° flow, -45° dispersion |
| Meteorological Institute, (ZMAW), University of Hamburg, Germany. | MITRAS | 0° case |
| Institute of Atmospheric Sciences and Climate(ISAC/CNR, Torino, Italy) and SOREQ Nuclear Research Centre (Yavne, Israel), with the collaboration of ARIANET s.r.l (Milano, Italy) *(simulations performed in parallel to the validation exercise)* | RAMS | 0° case, -45° flow, -45° dispersion, -41° field: flow, -41° field: dispersion |
| | µRMS | -45° dispersion, -41° field: dispersion |

Table 3.3: Groups and models participating in the MUST validation exercise.

| PARTICIPANT | NUMERICAL CODE | TEST CASES |
|---|---|---|
| Budapest University of Technology and Economics, Department of Fluid Mechanics, Hungary (BME). | MISKAM | 180.7 flow and dispersion |
| CEREA, Université Paris Est. | CODE-SATURNE | |
| Department of Fluid - and Thermodynamics, University of Siegen, Germany (USG). | FLUENT | |
| Dipartimento di Scienza dei Materiali, University of Lecce, Italy (USL). | FLUENT | |
| Environmental & Applied Fluid Dynamics Department, Von Karman Institute for Fluid Dynamics (VKI). | FLUENT | |
| Department of Mechanical Engineering, University of Western Macedonia, Greece (UOWM). | STAR-CD | 180.7 flow |

Table 3.4: Groups and models participating in the Oklahoma City validation exercise.

### 3.2.1 Short description of models

The following is a very brief description of the basic ingredients of most of the models used in the exercises. Only the main numerical approaches and the available turbulence models are given. More detailed information can be found in the references.

**ANSYS™ CFX** is a general purpose CFD analysis system. Within ANSYS™ CFX, the conservation equations for mass, momentum and scalar quantities like temperature, turbulent kinetic energy and any number of species are solved. The numerical solution is based on first-order in time and second-order in space discretisation (the discretisation method is based on the Finite Volume approach). The basic discretisation technique adopted in ANSYS™ CFX is a conventional Upwind Difference Scheme (UDS) with Numerical Advection Correction (NAC) for the advection terms in the momentum and energy equations. Numerical modelling of turbulence plays a crucial role in providing accurate predictions. In ANSYS™ CFX, the Reynolds stresses and turbulent fluxes of scalar quantities can be calculated by several linear and nonlinear turbulence models.

**FINFLO** is a block-structured, cell-centred finite-volume flow solver developed at Helsinki University of Technology. It has originally been developed for compressible-flow aeronautical applications, but the latest versions can also be applied to incompressible flows using an artificial compressibility method. In the MUST computations the flow was, however, assumed slightly compressible, since the verification of the incompressible code was still undergoing at the time of these computations. The stationary solution is achieved by a McCormack-type implicit solution of the governing RANS-equations and multigrid-method is used to speed up the convergence towards the steady state. A second-order upwind biased interpolation scheme is used for the inviscid fluxes. For the viscous terms, a second-order thin-layer algorithm is applied. Turbulence is modelled using a linear two-equation model which is essentially the standard $k$-ε model implemented in the $k$-ε formulation. The pollutant concentration is treated as a passive scalar quantity for which a transport equation is solved.

**M2UE (Microscale Model for Urban Environment)** is a structured, cell-centred finite-volume flow solver developed at Tomsk State University and Danish Meteorological Institute. The steady solution is achieved by a fully implicit solution of the governing RANS equations. A second-order upwind based interpolation scheme with flux limiters is used for the advection terms. For the viscous terms, second-order central differences are used. There are several options for turbulence closure scheme that are linear and non-linear eddy viscosity k-ε models. The pollutant concentration is treated as a passive scalar quantity for which a transport equation is solved.

The CFD model **VADIS** (Borrego et al., 2003) was developed at the Department of Environment and Planning of the University of Aveiro as a tool to estimate flow and pollutants dispersion resulting from accidental and uncontrolled releases of chemical products to the atmosphere, under unfavourable and complex dispersion conditions (low wind speed), near a single obstacle. Since its first version, VADIS has been in continuous development. It was adapted to the calculation of urban air pollution due to traffic road emissions in urban built-up areas by improving its capability to support multi-obstacle and multi-source description, as well as, time varying flow fields and time varying emissions. These updates allowed the evaluation of maximum short-term local concentrations of traffic related pollutants in urban geometries, especially under low wind speed conditions. VADIS structure is based on two modules, FLOW and DISPER. The first module, FLOW, is a Reynolds Averaged Navier-Stokes (RANS) prognostic model with a standard k-ε turbulence closure that calculates the wind components, the turbulent viscosity, the pressure, the turbulent kinetic energy and the temperature 3D fields through the finite volume method. The second module, DISPER, applies the Lagrangian approach to the computation of the 3D concentration field of inert pollutants using the wind field estimated by FLOW.

**ANSYS™ FLUENT** (Fluent, 2006) is a general purpose CFD solver. Within FLUENT, the integral equations for the conservation of mass and momentum, and, when appropriate, for energy and other scalars such as turbulence and chemical species are solved. The solver

offers a choice between an implicit first or second order scheme for the temporal discretisation and first or second order upwind schemes for the spatial discretisation of the convective terms in the scalar transport equations. Discretisation of the diffusion terms is central-differenced and always second order accurate. For the discretisation of the momentum equation FLUENT uses a collocated scheme, where pressure and velocity are both stored at cell centres. The default pressure-velocity coupling is the SIMPLE algorithm, which solves the pressure-correction equation using an algebraic multigrid method. Modelling of turbulence can be performed using several turbulence models, including the two-equation standard k-ε model and its variants (RNG and realizable).

**Code_Saturne** is a general purpose computational fluid dynamics free software. Developed since 1997 at EDF R&D, Code_Saturne is distributed under the GNU GPL licence (http://www.code-saturne.org). It is based on a co-located Finite Volume approach that accepts meshes with any type of cell (tetrahedral, hexahedral, prismatic, pyramidal, polyhedral…) and any type of grid structure (unstructured, block structured, hybrid, conforming or with hanging nodes…). Its basic capabilities enable the handling of either incompressible or expandable flows with or without heat transfer and turbulence (mixing length, 2-equation models, v2f, Reynolds stress models, Large eddy simulation…). Dedicated modules are available for specific physics such as radiative heat transfer, combustion (gas, coal, heavy fuel oil, …), magneto-hydro dynamics, compressible flows, two-phase flows (Euler-Lagrange approach with two-way coupling) and extensions to specific applications. For example the extension. for the atmospheric environment is developped by the CEREA (http://www.enpc.fr/cerea). Code_Saturne can be coupled to thermal software SYRTHES for conjugate heat transfer. It can also be used jointly with structural analysis software Code_Aster, in particular in the Salomé platform. SYRTHES and Code_Aster are developed by EDF and distributed under the GNU GPL licence.

**ADREA** is a three dimensional RANS (Reynolds Averaged Navier Stokes) CFD (Computational Fluid Dynamics) code. It was originally developed at Environmental Research Laboratory of NCSR “Demokritos”, Greece, as a transient, non-hydrostatic, fully compressible model, to be applied for fluid flow and vapour cloud dispersion predictions, especially designed for terrain of high complexity. It uses the conservation laws of mass, momentum, energy, humidity and in-ground heat conduction. The turbulence closure modelling is limited for now to eddy viscosity/ diffusivity concept. Anisotropic effects are also included. Zero, one and two-equation schemes are available. ADREA utilizes a finite volumes methodology for the numerical solution of the conservation equations, with a staggered grid for the velocities. It is fully implicit in time while it uses the upwind scheme for the convective terms. The SIMPLER/ADREA algorithm is adopted, which consists of transforming the mixture mass conservation equation into a full pressure equation, overall solution per time step by an iterative procedure, solution per variable by the various numerical methods including Biconjugate Gradient Stabilized Method and automatic time step selection based on convergence error bands. The equations are solved on a Cartesian, non equidistant grid. The topography is described by triangular surfaces which are allowed to cross the computational cells (volume porosity and surface permeability concepts). Treatment of complex domains is achieved by allowing a boundary surface, of arbitrary irregularity, to cross a calculation cell. An increase in the terrain complexity does not generally translate in increase of problem complexity. The geometrical input data can be given either manually (in simple cases) or using the geometrical input processor DELTA_B.

**STAR-CD** is a three dimensional RANS (Reynolds Averaged Navier Stokes) CFD (Computational Fluid Dynamics) code. It was originally developed as a thermo- fluid analysis system for the calculation of fluid flow, heat and mass transfer and chemical reaction in industrial and environmental applications. Basic differential equations are solved for flow, heat and mass transfer in the fluid. General forms are used that are appropriate for both laminar and turbulent flow. STAR-CD offers a wide variety of turbulence modelling capabilities (Eddy Viscosity models, Reynolds Stress models, Large Eddy Simulation models). It employs a highly flexible mesh system. At its heart lies the STAR-CD solver's ability to perform a numerical analysis in a mesh consisting of arbitrary polyhedral cells,

which may be used for fitting local geometrical features to a high degree of fidelity or for facilitating further mesh generation or optimization. Discretisation of the differential equations is achieved using the finite volume method. STAR-CD contains built-in boundary condition options that cover the majority of practical situations. Any physically-consistent mix of boundary conditions is admissible. The code currently incorporates two implicit algorithms, namely: A variant of the well known SIMPLE method, the PISO method, specially adapted to the STAR-CD requirements. Special modelling situations include calculation of flow, heat and mass transfer in porous media and heat exchangers, radiant heat interchange between surfaces, chemical reaction, including combustion, internal combustion engine simulation, two-phase flows, moving meshes and rotating systems, heat transfer in solids, compressible flows etc.

**MISKAM** was developed and is maintained by J. Eichhorn (list of references and downloads under http://www.lohmeyer.de/eng/Software/winmiskam.htm) and is now implemented about 100 times in Europe. MISKAM consists of a 3-dimensional non hydrostatic flow model and an Eulerian dispersion model. The physical basis are the complete 3-dimensional equations (RANS) of motion of the flow field and the advection-diffusion equation to determine the concentrations of substances with neutral density. It implements the k-$\varepsilon$ turbulence closure and provides an option of Smolarkiewicz correction steps to reduce the numerical diffusion of upwind scheme. The calculated result is the stationary flow- and pressure-field, diffusion coefficients and the concentration field in an area of typically 1000 m x 1000 m (200 x 200 x 100 cells or more, non equidistant grid). The main application is in built-up areas where neutral atmospheric stratifications are dominant. Input parameters are: Information about the grid, the buildings, the aerodynamic roughness of the areas between the buildings and on the building surfaces as well as position and strength of the sources. The present MISKAM version (5.02) has been thoroughly examined according to the requirements of the VDI-Code 3783/9 (See report http://lohmeyer.de/data/MISKAM-Vali.htm)

Other models that have been used for simulating both the meteorological flow and the tracer dispersion are atmospheric-micrometeorological models run in a CFD-like setup configuration. As in CFD models, atmospheric models at the micrometeorological scale (Atmo-RANS hereafter) solve the full set of the Reynolds-averaged equations of the atmosphere. The main advantage of this approach for applications in real atmosphere lays in the possibility of naturally interfacing the microscale system to larger scale meteorological models, in particular for determining the initial and boundary conditions, as well as in implementing data assimilation procedures. The soil models, parameterizations for the land use and several other physical processes (boundary layer, precipitation, scavenging, solar radiation, etc.) support a reliable reproduction of real atmospheric conditions. Furthermore, these models are implemented to produce time dependent solutions, which is an important aspect when considering the environmental assessment in a forecasting mode.

On the other hand, at present Atmo-RANS models (called U-RANS in Chapter 1.2) are not able to deal with very complex geometries of obstacles and buildings and are computationally demanding.

In the frame of COST732 Action, the Atmo-RANS models RAMS and the meteo-Lagrangian particle modelling system µRMS (composed by the modelling system RAMS-MIRS-MicroSPRAY for the microscale) have been tested for the MUST Experiment, both for the wind tunnel and the field experiment datasets. A discussion of these approaches and of some illustrative results can be found on the COST732 web site and in Reisin et al. (2007), Trini Castelli and Reisin (2008) and Trini Castelli et al. (2008).

The goal of these studies was to evaluate the possibility of adapting and using a regional meteorological modelling system to describe the flow and pollutants dispersion in an urban environment, characterized by short time (minutes) and space (tens to hundreds of meters) scales and by the presence of obstacles. The twofold aim was to test both RAMS model, used in a CFD-like configuration for simulating the flow dynamics and tracer dispersion, with a resolution of the order of 1 m in presence of obstacles, and MicroSPRAY in cascade to RAMS.

### 3.2.2 Set up of model runs

For the MUST wind tunnel experiment the 120 obstacles of the full scale experiment were modelled at a 1:75 scale and positioned on the wind tunnel's turn table as shown in Figure 3.7. The approach flow conditions in the wind tunnel were adjusted to meet the approach flow measured in full scale. In Figure 3.8 the measured mean velocity profile and the measured profile for the turbulent kinetic energy are shown. H = 2.54m is the height of the containers in full scale.

Figure 3.7: Setup for the MUST wind tunnel experiment (Bezpalcova and Harms, 2005).

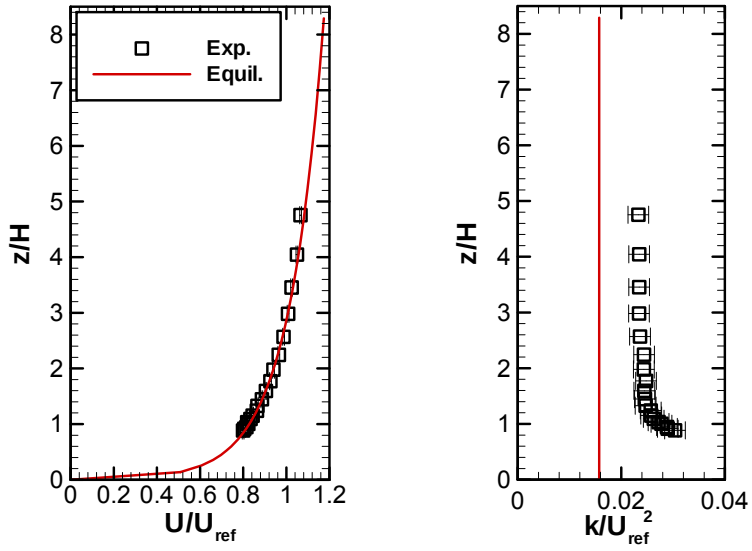



Figure 3.8: Measured velocity (left) and turbulent kinetic (right) profile of the approach flow in the wind tunnel. Curves are the equilibrium profiles derived from the measured roughness $z_0$ = 0.0165m.

Also shown are the profiles resulting for an equilibrium boundary layer over a rough wall (Richards and Hoxey, 1993) which are often used in CFD modelling. While the equilibrium velocity profile agrees very well with the experiments, the measured turbulent kinetic energy is much higher than the equilibrium value, especially close to the ground at container height. The reason for this discrepancy is that in the equilibrium profile turbulence is only created on the rough wall, while in the wind tunnel turbulence generators are used over nearly the entire height of the tunnel, cf. Figure 3.7.

The experimental approach flow profiles were the same for each investigated wind direction. For COST 732 two wind directions have been used, the 0 degree and the –45 degree case. While flow measurements are available for both cases, concentrations were only measured for the –45 degree case. The setup of the different simulation runs for these two wind directions is presented in the following chapters.

#### 3.2.2.1 *0 degree MUST wind tunnel case*

For the 0 degree MUST wind tunnel case more than 40 simulations were performed by the 15 research groups listed in Table 3.3. For 32 of these simulations results are available for the public inside the Excel workbooks described in Chapter 4.2. The corresponding Run_IDs and the short names of these 32 simulations are given in Table 3.5.

| Run_ID | Short Name |
|---|---|
| FI00_001 | MISKAM_Ketzel |
| FI00_002 | MISKAMcoarse_Ketzel |
| FI00_003 | MISKAMcoarse20layer_Ketzel |
| FI00_004 | MISKAM1mres_Goricsan |
| FI00_005 | MISKAM08mres_Goricsan |
| FI00_006 | MISKAM05mres_Goricsan |
| FI00_007 | MISKAM_ZAMG |
| FI00_008 | FLUENTske_DiSabatino |
| FI00_009 | FLUENTrsm_DiSabatino |
| FI00_010 | FLUENTskeBCsym_Santiago |
| FI00_011 | FLUENTskeBCwall_Santiago |
| FI00_012 | FLUENTrkeBCsym_Santiago |
| FI00_013 | ADREA_Bartzis |
| FI00_014 | STAR_CD_Bartzis |
| FI00_015 | STAR_CD_doublegrid_Bartzis |
| FI00_016 | Fluent_mskesp_Franke |
| FI00_017 | Fluent_mskespudf_Franke |
| FI00_018 | Fluent_mskespudftsym_Franke |
| FI00_019 | FluentFS1_Goricsan |
| FI00_020 | FluentFS10_Goricsan |
| FI00_023 | FINFLO_Hellsten |
| FI00_024 | MITRAS2-10oC_Schluenzen |
| FI00_026 | CFX-unstr-ke_Fotios |
| FI00_027 | CFX-str-ke_Fotios |
| FI00_028 | CFX-str-sst_Fotios |
| FI00_029 | CFX-str-ssg_Fotios |
| FI00_030 | CFX-unstr-ke_final_Fotios |
| FI00_031 | STARCD_coarse_Brzozowski |
| FI00_032 | STARCD_medium_Brzozowski |
| FI00_033 | STARCD_fine_Brzozowski |
| FI00_034 | M2UE_Nuterman_Baklanov |
| FI00_038 | Code_Saturne |

Table 3.5: Run_IDs and short name of the simulations for the 0 degree flow case.

The detailed description of the model set up for each simulation can be found in Appendix A. Here only summarised information on the set up is presented.

In Figure 3.9 a typical computational domain is shown. The wind direction is parallel to the x-axis. The height of this computational domain is six times the height $H_{max}$ of the highest building, the so called VIP building, which is located roughly in the centre of the array and has a smaller ground area than the surrounding containers. The choice of this domain height is based on the *Best Practice Guideline for the CFD simulation of flows in the urban environment* (Franke et al., 2007) and was used in most of the simulations, but some simulations used a higher domain.

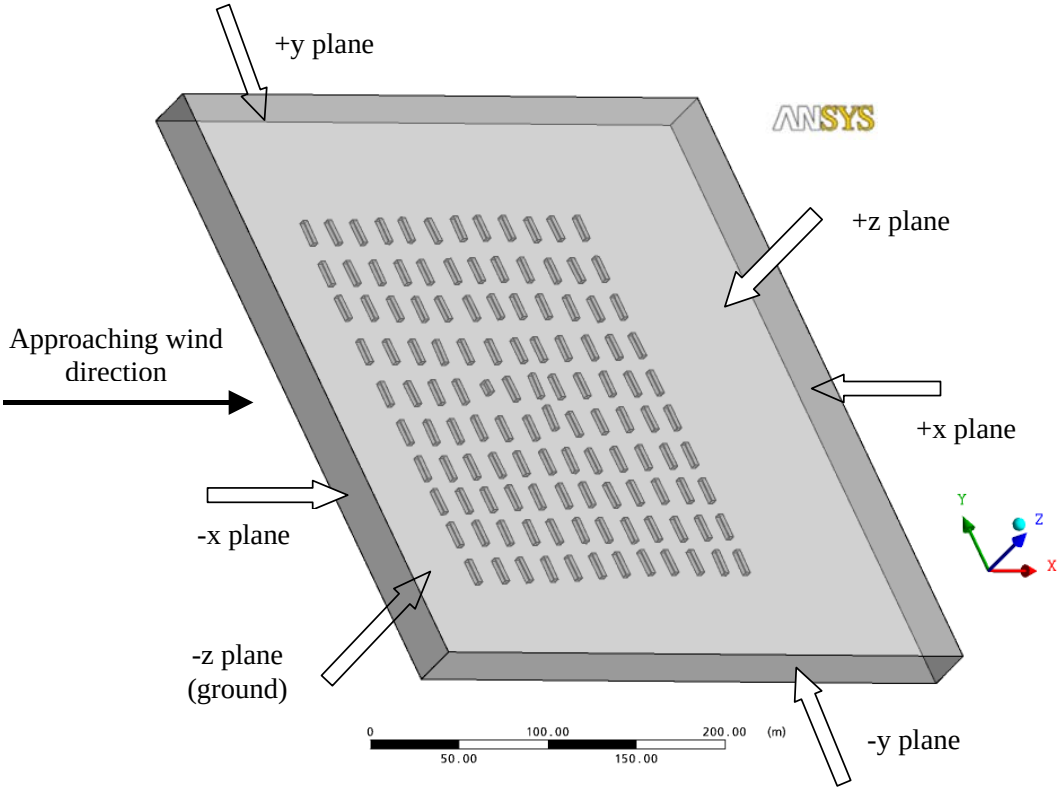



Figure 3.9: Typical computational domain for the 0 degree case.

The lateral extent of the domain was most often chosen as the depth of the wind tunnel's test section shown in Figure 3.7. Only a few simulations used a larger extent in the lateral direction. The +x and –x boundaries were mostly positioned further away from the obstacles as recommended in the aforementioned Best Practise Guideline, with a typical distance of 15 $H_{max}$ for the +x plane and 25$H_{max}$ for –x plane. Some simulations used larger distances.

The models ADREA, MISKAM, MITRAS and VADIS use a Cartesian grid, therefore the detailed position of the obstacles could not be reproduced. All other simulations except the CFX-unstr-XXX runs used block structured meshes with the exact position and size of the blocks. The CFX-unstr-XXX runs also used the exact geometry but an unstructured mesh made of tetrahedral cells. Some groups did test the grid sensitivity of the results by using different resolutions. One group did a systematic grid convergence study which is presented in Chapter 4.2.4.

For the conditions prescribed at the boundaries of the computational domain the following choices were made. At the inlet plane (-x) either equilibrium profiles as shown in Figure 3.3

were used, or directly the measured profiles. One run used a radiative or open (outflow) boundary condition at the -x plane. This kind of boundary condition was also used in many simulations for the +x plane. The other simulations used a constant pressure there.

At the lateral boundaries (+y, -y) three different kinds of boundary conditions were used. A few simulations did also employ open or radiative boundary conditions, some simulations prescribed walls and the majority used symmetry conditions. The symmetry condition was also often used for the top of the domain, the +z plane. Another common choice was to fix the values of the velocity components and turbulence quantities with the aid of the approach flow profiles. Two simulations used wall boundary conditions.

The ground floor (-z) and the building walls were always treated as walls. Differences exist in the choice of the roughness. The ground floor was modelled as smooth, partly rough or entirely rough, mostly with the roughness derived from the measurements. The building walls were either smooth or rough.

The turbulence was most often modelled with the standard k-$\varepsilon$ model, but also other linear two-equation models and differential stress models were used. One simulation employed a one-equation model. For modelling the turbulence and the flow close to the wall, standard wall functions were used nearly exclusively.

The numerical approximations for the advective terms in the transport equations employed either the first or the second order upwind scheme. Differences did also exist in the stopping criteria for the iterative solution process or the time integration towards a steady state. While some iterative solutions were run until the residuals stopped decreasing below an already small value, other iterations were stopped at a relatively high level of the residuals.

#### 3.2.2.2 *-45 degree MUST wind tunnel case*

For the -45 degree MUST wind tunnel case more than 30 simulations were performed by the 15 research groups listed in Table 3.3. For 28 of these simulations results are available for the public inside the Excel workbooks described in Chapter 4.2. The corresponding Run_ID and the short name of these 28 simulations is given in Table 3.6. The variations in the set up are very similar to the ones described in the previous chapter for the 0 degree flow case. The only additional variation concerns the orientation of the computational domain towards the approach flow direction. Many simulations, especially those using Cartesian grids did simply rotate the entire computational domain and defined the -x and +y planes as inlet boundaries with the same profiles that were used in the 0 degree flow case, and the +x and -y planes as outlet boundaries, see Figure 3.10.

With this approach the lateral walls of the wind tunnel test section can no longer be taken into account. Therefore other simulations kept the lateral extent of the computational domain and rotated the obstacles within the domain, like it was done in the wind tunnel. Due to mesh quality requirements the distances between the obstacles and the +x and -x planes had to be increased. Figure 3.11 shows this modified setup, for which the boundary conditions used for the 0 degree case could be kept.

The flow results were used for the calculation of the dispersion from a ground source, which was most often treated as a volumetric source. The pollutant was normally treated as passive scalar in an Eulerian framework. The turbulent scalar flux in the transport equation was in all simulations modelled with the gradient diffusion hypothesis. With this approach the turbulent diffusion is proportional to the turbulent viscosity. The constant of proportionality is the turbulent Schmidt number, which is given in Table 3.7 together with the Run_ID and the short name of each simulation. Note that the turbulent Schmidt number varies by a factor of two between 0.5 and 1.0, with most simulations using a value around 0.7. Altogether there are 22 simulations for the –45 degree dispersion case.

| Run_ID | Short Name |
|---|---|
| Fl45_001 | MISKAM_Ketzel |
| Fl45_002 | MISKAM_Ketzel_varRoughness |
| Fl45_003 | Miskam1mres_Goricsan |
| Fl45_004 | Miskam08mres_Goricsan |
| Fl45_005 | Miskam05mres_Goricsan |
| Fl45_006 | MISKAM_ZAMG |
| Fl45_007 | FLUENT_Santiago |
| Fl45_008 | FLUENTske_DiSabatino |
| Fl45_009 | Fluent_mrotskespudf_Franke |
| Fl45_010 | Fluent_mskesp_Franke |
| Fl45_011 | Fluent_mskespudf_Franke |
| Fl45_012 | Fluent_mskespudftsym_Franke |
| Fl45_013 | Fluent_ke_Goricsan |
| Fl45_014 | Fluent_RSM_Goricsan |
| Fl45_015 | STAR_CD_Bartzis |
| Fl45_016 | CFX-unstr-ke_fine_Fotios |
| Fl45_017 | FINFLO_Hellsten1 |
| Fl45_018 | FINFLO_Hellsten2 |
| Fl45_019 | FINFLO_Hellsten3 |
| Fl45_020 | M2UE_Nuterman_Baklanov |
| Fl45_021 | VADIS_Costa_2m |
| Fl45_022 | VADIS_Costa_1m |
| Fl45_023 | VADIS_Costa_0_5m |
| Fl45_024 | STARCD_coarse_Brzozowski |
| Fl45_025 | STARCD_medium_Brzozowski |
| Fl45_026 | STARCD_fine_Brzozowski |
| Fl45_027 | ADREA_Bartzis |
| Fl45_029 | Code_Saturne |

Table 3.6: Run_ID and short name of the simulations for the –45 degree flow case available for the public.

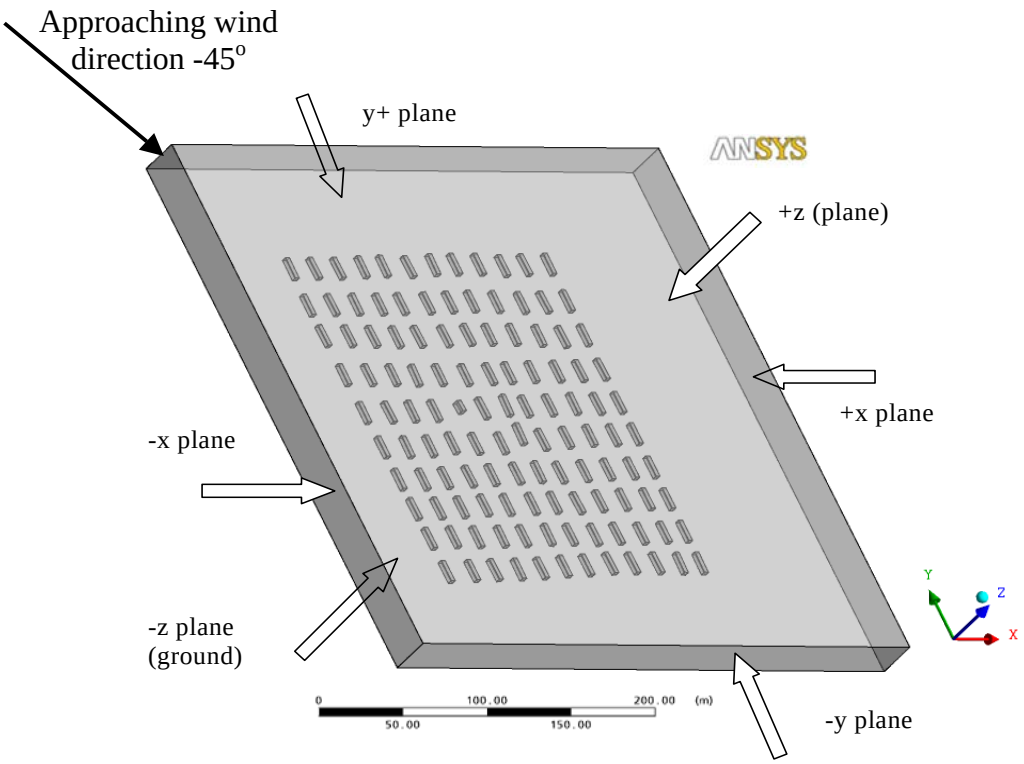



Figure 3.10: Computational domain for the –45 degree case obtained by rotating the domain of the 0 degree case.

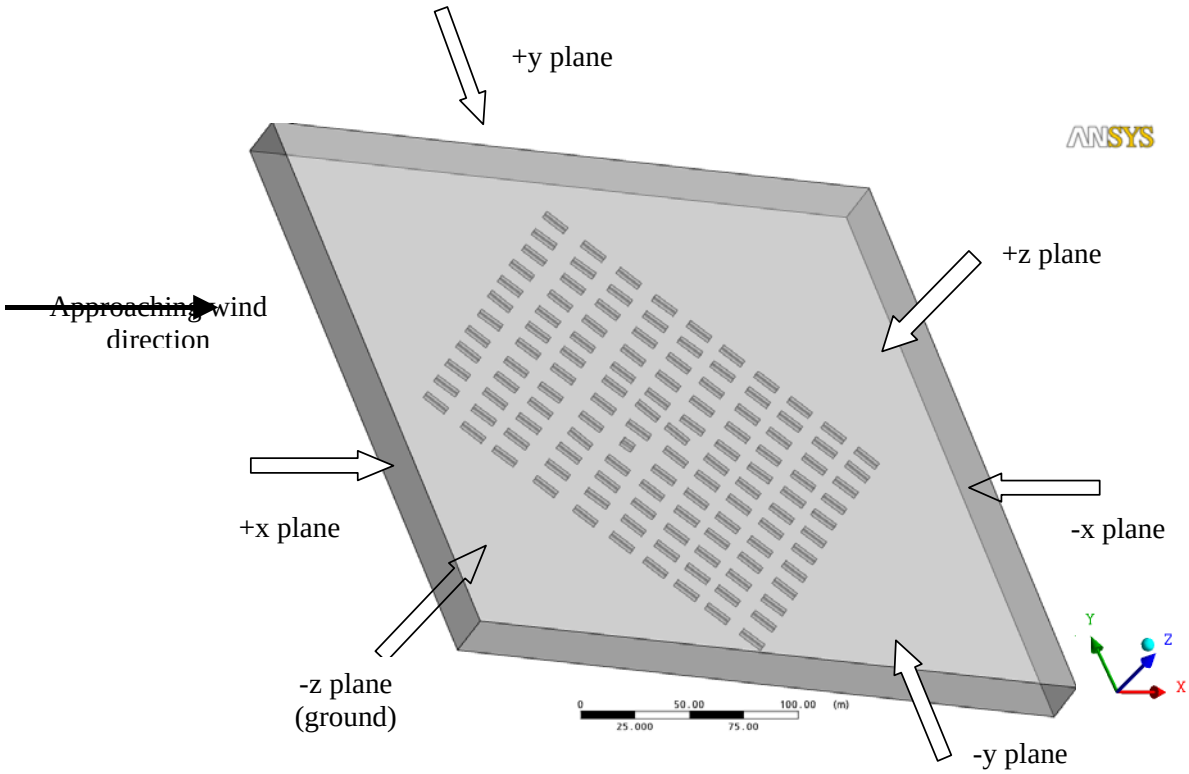



Figure 3.11: Computational domain for the –45 degree case obtained by rotating the domain of the 0 degree case.

| Run_ID | Short Name | Turbulent Schmidt number |
|---|---|---|
| Disp_001 | MISKAM_Ketzel | 0.74 |
| Disp_002 | MISKAM_Ketzel_varRoughness | 0.74 |
| Disp_003 | MISKAM1mres_Goricsan | 0.74 |
| Disp_004 | MISKAM08mres_Goricsan | 0.74 |
| Disp_005 | MISKAM05mres_Goricsan | 0.74 |
| Disp_006 | MISKAM_ZAMG | 0.74 |
| Disp_007 | Fluent_ke_Goricsan | 0.70 |
| Disp_008 | Fluent_RSM_Goricsan | 0.70 |
| Disp_009 | FLUENTske_DiSabatino | 0.70 |
| Disp_010 | FLUENT_Santiago | 0.90 |
| Disp_011 | Fluent_mskespudf_Franke | 0.70 |
| Disp_012 | CFX-unstr-ke_fine_Fotios | 0.77 |
| Disp_013 | FINFLO_Hellsten1 | 0.80 |
| Disp_014 | FINFLO_Hellsten2 | 0.80 |
| Disp_015 | FINFLO_Hellsten3 | 0.80 |
| Disp_016 | M2UE_Nuterman_Baklanov | 0.50 |
| Disp_017 | VADIS_Costa_2m | 0.70 |
| Disp_018 | VADIS_Costa_1m | 0.70 |
| Disp_019 | VADIS_Costa_0_5m | 0.70 |
| Disp_020 | STARCD_fine_Brzozowski | 0.90 |
| Disp_023 | Code_Saturne | 1.00 |
| Disp_025 | ADREA_Bartzis | 0.74 |

Table 3.7: Run_ID, short name and turbulent Schmidt number of the simulations for the –45 degree dispersion case available for the public.

#### 3.2.2.3 *180.7 degree Oklahoma city wind tunnel case*

The Oklahoma City wind tunnel experiment is the second case used for testing the *Model Evaluation and Protocol Document*. Started towards the end of the action, only a limited number of groups took part in this exercise, see Table 3.4. Being a real city, the geometry of Oklahoma City is rather complex, making the CAD geometry generation, the mesh generation and the simulations more complicated than for the MUST case. Therefore it was decided that in addition to partially test the protocol for this realistic urban geometry, the exercise should mainly serve to analyse

- the influence of modelling detail of the parking garages, and
- the influence of modelling detail of the source

on the results. Also, the influence of the discretisation scheme for the advective term in the transport equation for the pollutant was investigated. The results help to identify important parameters and are an important preparation for a detailed validation.

The set up of the simulations of Oklahoma City was as follows. The geometry and positions of the buildings were converted from 2D CAD drawings of the wind tunnel experiment to 3D. These buildings are shown in Figure 3.12. The computational domain around the buildings is shown in Figure 3.13. The wind blows in positive x-direction. The lateral extent of the computational domain corresponds to the depth of the test section in the wind tunnel. The empty space before, behind and above the buildings was chosen according to the Best Practise Guideline, based on the height of the tallest building, $H_{max}$ = 152m. Note

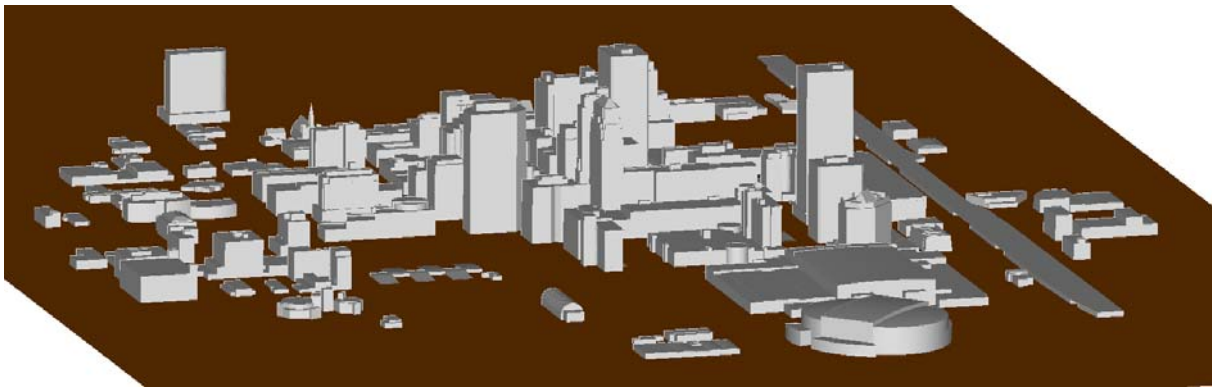


Fig. 3.12: Buildings used to model the centre of Oklahoma City for the 180.7 degree case.

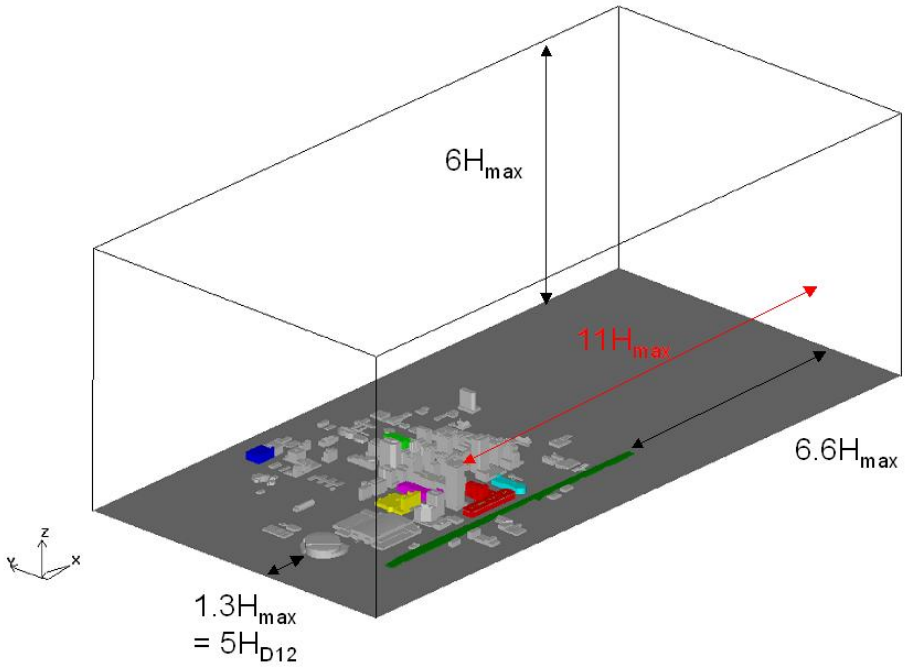



Fig. 3.13: Computational domain for the 180.7 degree Oklahoma City wind tunnel case. $H_{D12}$ is the height of the oval building closest to the inflow plane at $x_{min}$. and $H_{max}$ the height of the tallest building in the domain.

that the oval shaped building D12 closest to the inflow plane at $x_{min}$ is five times its height away from this boundary.

Except for the dark green railway dam parallel to the x-axis, the other coloured buildings in Figure 3.13 are parking garages. These parking garages have openings so that the wind can pass through them. Figure 3.14 shows the seven parking garages and their location in the built area. Several of the garages are close to the measurement locations in Park Avenue, which is the street where the ground source S5 is located. From the wind tunnel measurements it is known that closing the openings in the garages has a large influence on the flow and concentrations, but this has not been examined for the measurement positions used in the present exercise. Therefore it was decided to analyse the influence of the garages on the flow and concentrations in Park Avenue by performing numerical simulations with the openings in all garages and numerical simulations with closing all openings in the garages.

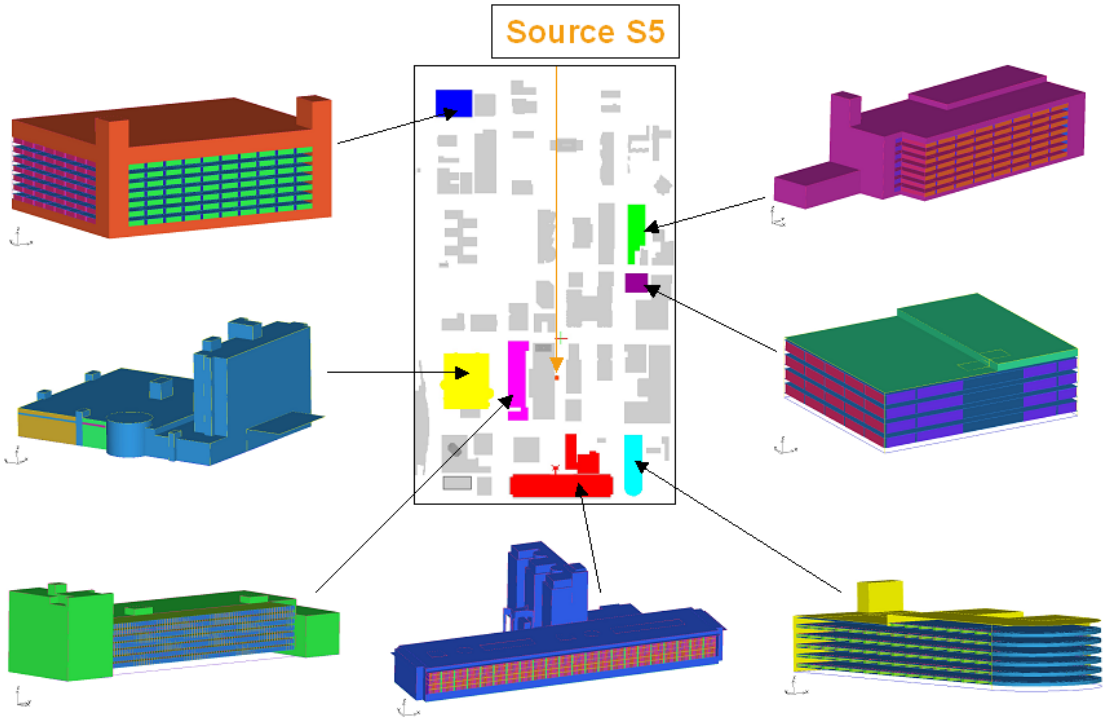



Fig. 3.14: The seven parking garages and their location in the 180.7 degree Oklahoma city case. The ground source S5 is located in Park Avenue, where the measurement positions are located.

While closing all openings in the garages is straightforward, the open garages are a challenge for the simulations with regards to the grid generation. One can either use conformal grids in- and outside of the garages with correspondingly small cells in regions far from the measurement locations, which could otherwise be meshed with coarser cells. Or independent grids can be generated in- and outside of the garages with additional interpolation between the two grids at the interfaces. The simplest option is to define the garages as porous zones, but this would require information about the direction dependent pressure loss coefficients in each garage. In the present simulations the approach with independent grids and coupling at the interfaces was used. Inside the garages a block-structured grid with hexahedra was created and outside an unstructured grid with tetrahedra. The total number of cells is $7.9 \cdot 10^5$ hexahedral cells and $14.5 \cdot 10^6$ tetrahedral elements. In Figure 3.15 some details of the grid are shown. All simulations presented in Chapter 4.5 were done using the same grid.

The second parameter examined is the influence of the source modelling on the results. In Figure 3.16a the real geometry of the ground source S5 (cf. Figure 3.14) is shown. The pollutant enters through a small pipe, above which a circular plate is placed to minimise the vertical momentum. This plate was also kept in the second modelling approach, where the pollutant does not enter through the pipe, but is released as a volume source in the cells at the pipe exit, see Figure 3.16b Two volume flow rates were used, the experimental minimum 0.3 l/h and the experimental maximum 2.5 l/h.

All other parameters for the simulations were identical. The standard k-ε turbulence model was used with standard wall functions. The turbulent scalar flux in the transport equation for the passive scalar was modelled with the gradient diffusion assumption, with a turbulent Schmidt number of $Sc_t = 0.7$. The boundary conditions for the computational domain shown in Figure 3.13 were symmetry at $y_{min}$, $y_{max}$ and $z_{max}$, and constant pressure a $x_{max}$. At $x_{min}$ the velocity component, turbulent kinetic energy (TKE) and turbulent dissipation (ε) were prescribed. They were derived from measurements, yielding a power law for the velocity component in x-direction,

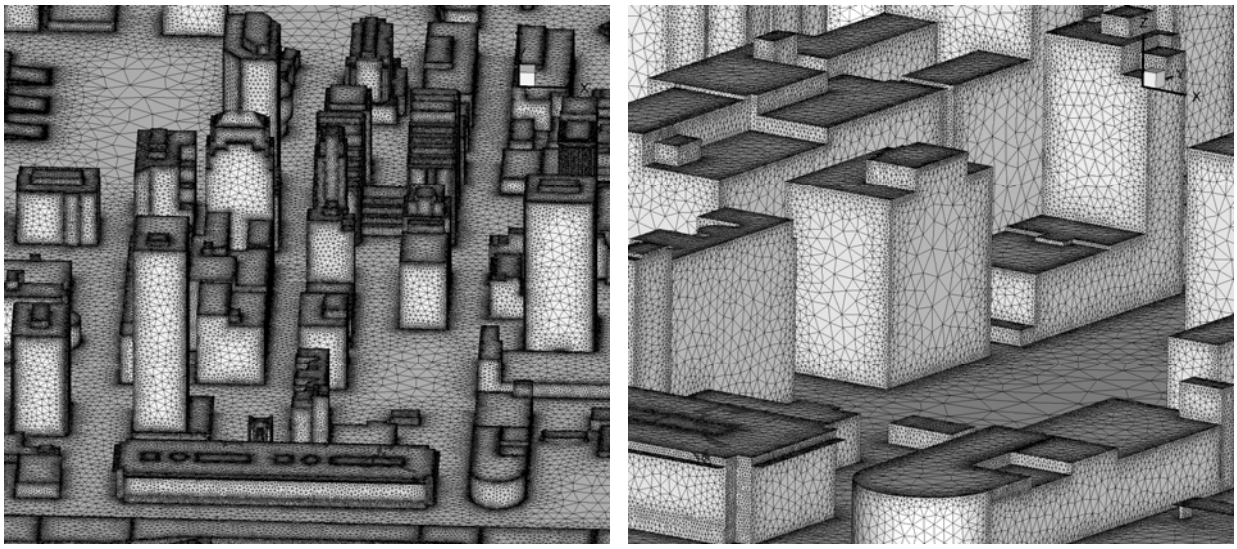


Fig. 3.15: Triangular surface mesh details of the 180.7 degree Oklahoma city case.

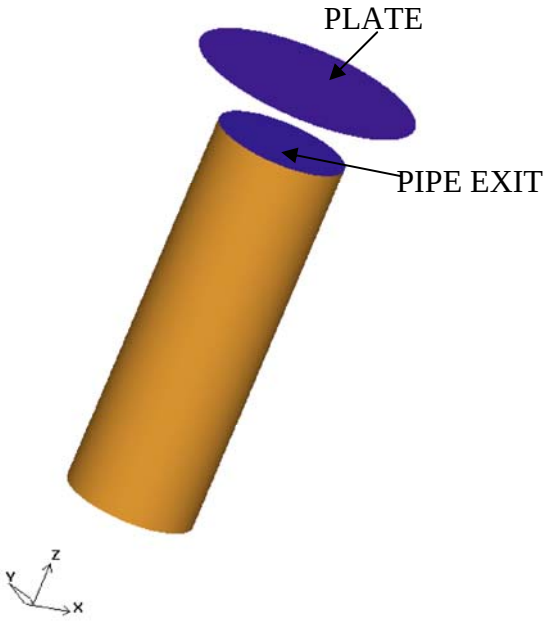



Fig. 3.16a: Real ground source geometry for numerical simulation.

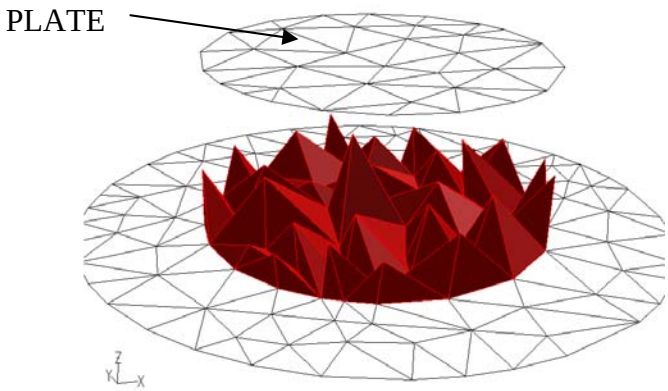



Fig. 3.16b: Cells with volume source for numerical simulation.

$$U(z) = U_{ref}\left(\frac{z}{z_{ref}}\right)^{0.18} \quad with \quad U_{ref} = 8.84\frac{m}{s} \quad and \quad z_{ref} = 40.2m$$

and a second order polynomial for TKE up to z = 386m and constant above. The turbulent dissipation was set equal to the production of TKE.

Except for the turbulence equations the 2nd order upwind approximation was used for the advective terms in the transport equations. Simulations were run until the scaled residuals attained a relatively constant value.

# 4 Validation Procedures and Results

## 4.1 Overview

The *Model Evaluation Guidance and Protocol Document* outlines a long list of items to be considered when performing a model validation activity. The Guidance document represents an attempt to be comprehensive, so that the document can be used for a wide variety of situations. Thus, it does not go into detail, but outlines a number of questions to be considered in the model validation process.

This is where the present document comes in. After an introduction to the validation objectives and metrics, the chapter “Validation Procedures and Results” provides details for four specific cases at hand:

- CFD models validated with MUST wind tunnel data
- Non-CFD models validated with MUST wind tunnel data
- Validation attempts with MUST field data
- CFD models validated with Oklahoma City wind tunnel data

### 4.1.1 Validation objective

The Guidance Document states that *validation objectives* have to be defined. When a model is to be applied for a specific purpose, specific validation objectives should be defined.

In the case of the exercises within COST 732 the objective was to acquire a general insight into the state of art in model performance. Therefore it has not been possible to define so specific objectives as if a particular application were in focus.

#### 4.1.1.1 *Validation objectives - MUST wind tunnel data*

The prediction of the following parameters has been in focus for CFD models:

- Velocity components u, v, w (normalised by a reference velocity).
- Turbulent Kinetic Energy as modelled and as assessed from measurements. Note that the measurements were performed for no more than two velocity components simultaneously, so assessment of TKE involves an assumption on the third component.
- Concentration c (normalised)

For non-CFD models the main focus has been on concentration. A diagnostic mass-consistent wind field model (LPRWND) which is usually used to compute wind fields for the Lagrangian particle dispersion model LASAT is also validated against the MUST wind tunnel flow data.

One may distinguish between tasks that are more or less challenging for a model. In the subsequent chapters validation results for MUST are presented that involve data from all measurement positions. However, the task of predicting flow in the 'streets' between containers is much more challenging than the task of predicting flow well above the roof of the containers. In the analyses, several subsets of data are considered. It can therefore be investigated separately how well models replicate tasks of various complexity.

Concerning concentration, CFD models are presumed to be able to predict the development of the concentration pattern in space and time. Therefore, for CFD models it makes sense to consider pairs of concentrations as observed and as modelled, paired in space.

On the other hand, such a comparison makes less sense for the large group of non-CFD models that assume a straight centreline for any plume. For such models it is *a priori* known that they are unable to predict the position of a plume correctly if the geometry is complex. Therefore, for such models it is not reasonable to set up an evaluation objective based on data paired in space. However, non-CFD models may be used for purposes where they are asked to predict the *magnitude* of concentrations. An evaluation objective for such models can therefore be based on maximum arc-wise concentrations.

#### 4.1.1.2 Validation objectives - Oklahoma City wind tunnel data

The prediction of the following parameters has been in focus for CFD models:

- Velocity components u, v, w (normalised by a reference velocity).
- Concentration c (normalised)

Due to the short time available for this case no validation with regard to the turbulence variables was performed. Corresponding data are however available from the measurements and the simulation runs and will be incorporated in the analysis in the future.

In Chapter 4.5, validation results are presented that involve data from all measurement positions. As only CFD models were used for the simulations the observed and predicted results are again paired in space.

### 4.1.2 Validation metrics

After a paired set of experimental data and model predictions has been obtained and the exploratory data analysis has been performed, a quantitative comparison of experimental and model results has to be done. To that end several standard metrics are available that were also used in this action and are briefly introduced in the following, starting with the not so common hit rate as used by the German VDI guideline (VDI, 2005) for the comparison of velocities. Then the other metrics used mainly for the concentrations are defined. These are essentially the metrics also available in the BOOT software (*URL 1*).

#### 4.1.2.1 *Hit rate (q)*

The *Hit Rate* (q) is a metric that has been found useful for velocity models. It is applied in the German VDI Guideline on prognostic micro-scale wind field models.

***Definition of Hit Rate***

From normalised model results $P_i$ and normalised comparison data $O_i$ a hit rate q is calculated from the equation below, which specifies the fraction of model results that differ within an allowed range D from the comparison data. D accounts for the relative uncertainty of the comparison data. Only those differences are counted that are above a threshold value W, which describes the repeatability of the comparison data.

$$q = \frac{N}{n} = \frac{1}{n}\sum_{i=1}^{n} N_i \quad \text{with} \quad N_i = \begin{cases} 1 \text{ for } \left|\frac{P_i - O_i}{O_i}\right| \le D \text{ or } \left|P_i - O_i\right| \le W \\ 0 \text{ else} \end{cases}$$

In other words, a hit is characterised in the following way: Assume we have an observed value $O_i$ and a predicted value $P_i$. We have a hit if one of the following conditions is fulfilled:

- Is Abs($P_i$ - $O_i$) smaller than the allowed absolute deviation W?
- Is Abs(($P_i$ - $O_i$)/$O_i$) smaller than the allowed fractional deviation D?

The Hit rate can be interpreted graphically. This is illustrated in Fig. 4.1.

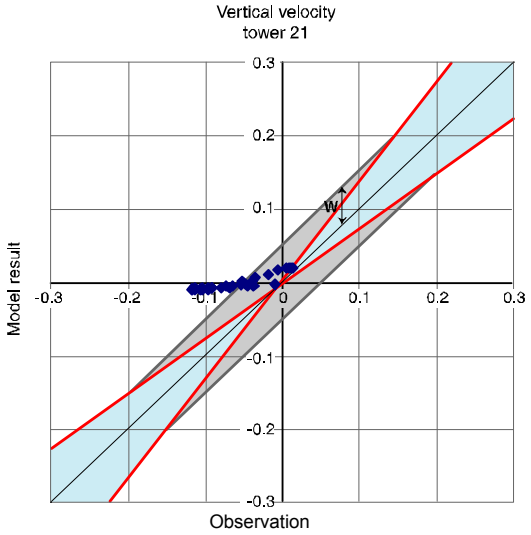



Fig. 4.1 Illustration of Hit Rate. Two (grey) lines parallel to the one-to-one line define a band of acceptance. These lines are displaced ±W in the vertical direction from the one-to-one line. Points that are within this band are counted as hits. Also, points within the (blue) sectors confined by the red lines are included as hits (there are none in the example).

Results are accepted as hits if thy are *either within a band of acceptance* defined by two lines, parallel to the one-to-one line, *or within one of the two narrow sectors* defined by the red lines representing fractional deviation (blue area). From Figure 4.1 it is obvious that the selected values for W and D have a large influence on the hit rate. While the allowed relative difference was always 25%, i.e. D=0.25 as it was used by VDI (2005), the allowed absolute difference W was different for each variable, based on the experimental uncertainty. For an ideal model prediction one would have q=1.

#### 4.1.2.2 *Factor of two of observations (FAC2)*

The Factor of two observations (FAC2) is a metric also used in the BOOT software, but the present definition is differing from the implementation in BOOT. FAC2 is similar to the hit rate as it counts the fraction of data points, where the predictions are within a factor of two of the observations, based on the ratio of the predicted and observed value. This general definition raises the question what to do with data points where the observations are exactly zero. In the BOOT software these data points are simply not taken into account, i.e. the number of total data points is reduced. Contrary to this the present implementation of FAC2 uses a case differentiation based on the threshold W, also used in the hit rate. If the observation is below the threshold W then it is checked, if the prediction is also below W. If so, the data point is counted as within a factor of two.

$$FAC2 = \frac{N}{n} = \frac{1}{n}\sum_{i=1}^{n} N_i \quad \text{with} \quad N_i = \begin{cases} 1 \quad \text{for} \quad 0.5 \le \dfrac{P_i}{O_i} \le 2.0 \\ 1 \quad \text{for} \quad O_i \le W \text{ and } P_i \le W \\ 0 \quad \text{else} \end{cases}$$

FAC2 is the most robust measure concerning the influence of infrequently occurring high or low observations and predictions for the concentrations. For an ideal model prediction one would have FAC2=1.

#### 4.1.2.3 *Fractional bias (FB)*

The fractional bias (FB) is a metric also used in the BOOT software and implemented in the same way in the present exercise.

$$FB = \frac{\langle O \rangle - \langle P \rangle}{0.5\left(\langle O \rangle + \langle P \rangle\right)}$$

Here angular brackets denote the average over all data points.

FB is a linear measure of the mean bias and reveals systematic errors. It is known to be strongly influenced by infrequently occurring high observations and predictions for the concentrations. Another shortcoming of FB is that it is based on the mean bias, so that over-prediction and under-prediction of observations may cancel each other, leading to FB=0, the result for an ideal prediction. Chang and Hanna (2005) therefore recommend to separate FB into its over- and under-predictions. They show that FB can be expressed as

$$FB = FB_{fn} - FB_{fp},$$

where $FB_{fn}$ (false negative) contains all under-predictions and $FB_{fp}$ (false positive) all over-predictions. From this definition it also follows that a negative FB indicates a systematic over-prediction and vice versa, as $FB_{fn}$ and $FB_{fp}$ are both positive quantities.

FB makes no sense for parameters that can take both positive and negative values, such as velocity components.

#### 4.1.2.4 *Geometric mean bias (MG)*

The geometric mean bias (MG) is a metric also used in the BOOT software and implemented in the same way in the present exercise.

$$MG = \exp\left(\left\langle \ln \tilde{O} \right\rangle - \left\langle \ln \tilde{P} \right\rangle\right) \quad \text{with} \quad \tilde{O} = \max(W, O) \quad \text{and} \quad \tilde{P} = \max(W, P)$$

Here angular brackets again denote the average over all data points. The usage of a threshold W for the data values is necessary as the logarithm is on one hand not defined for concentrations being exactly zero. Furthermore, MG as logarithmic measure of the mean bias is strongly influenced by extremely low observations and predictions for the concentrations. However it is not so strongly influenced by infrequently occurring high observations and predictions for the concentrations, as FB.

From its definition it is obvious that in MG also the over- and under-predictions may cancel and yield the results for an ideal model, MG=1. This problem can be again solved by looking at the under-predictions (false negative) and over-predictions (false positive) separately.

$$MG = \frac{MG_{fn}}{MG_{fp}}$$

A systematic over-prediction will now lead to MG<1 and a systematic under-prediction to MG>1.

#### 4.1.2.5 *Normalised mean square error (NMSE)*

The normalised mean square error (NMSE) is a metric also used in the BOOT software and implemented in the same way in the present exercise.

$$NMSE = \frac{\left\langle (O-P)^2 \right\rangle}{\langle O \rangle \langle P \rangle}$$

Here angular brackets again denote the average over all data points.

NMSE is, like FB, strongly influenced by infrequently occurring high observations and predictions for the concentrations. NMSE is a measure of the scatter of the data and therefore indicates both, systematic and random errors. For an ideal model prediction one would have NMSE=0.

NMSE makes no sense for parameters that can take both positive and negative values, such as velocity components.

#### 4.1.2.6 *Geometric variance (VG)*

The geometric variance (VG) is a metric also used in the BOOT software and implemented in the same way in the present exercise.

$$VG = \exp\left[\left\langle \left(\ln \tilde{O} - \ln \tilde{P}\right)^2 \right\rangle\right] \quad \text{with} \quad \tilde{O} = \max(W, O) \quad \text{and} \quad \tilde{P} = \max(W, P)$$

Here angular brackets again denote the average over all data points and the tilde indicates that a threshold has been used for the data, similar to MG. This threshold removes problems with zero or extremely small concentrations within the logarithms.

Like NMSE, VG shows the scatter in the data and therefore contains both, systematic and random errors. For an ideal model prediction one would have VG=1.

### 4.1.3 Quality acceptance criteria for metrics

The validation exercises provide qualitative and quantitative views of the model quality. To decide whether the model performance is “acceptable” with regards to the metrics, it is necessary to specify limit values for the metrics as performance measures. E.g. the German VDI guideline (VDI, 2005) requires for successful validation a hit rate q $\geq$ 0.66 for all velocity components of the five validation test cases. As stated in Chapter 4.1.2.1 the hit rate depends strongly on the values chosen for the allowed relative deviation D and allowed absolute deviation W. While in the present exercise the same D=0.25 of VDI (2005) is used, the present W values are approximately only one tenth of the values of the VDI guideline. Therefore the direct transfer of the limit value q $\geq$ 0.66 to the present exercise is at least problematic.

For the metrics used in BOOT and introduced in the previous chapters the situation is similar. Hanna et al. (2004) used for research-grade field experiments the following limit values: FAC2 > 0.5, |FB| < 0.3, 0.7 < MG < 1.3, NMSE < 4 and VG < 1.6. This means that the fraction of predictions within a factor of two of observations is higher than 50%, the mean bias is within $\pm$ 30% of the mean and the random scatter is lower than approximately a factor of two of the mean. Concerning the transfer of these limits to the present exercise it must be noted that Hanna et al. (2004) use these values for comparisons of maximum concentrations on arcs, i.e. unpaired in space. As the present exercise is based on a point by point compari-

son, which is more stringent, less strict criteria might be used. The values of metrics can differ widely, depending on what the data underlying the metrics actually represent.

The limit values stated above for all metrics will be used as guidance in the following exercises. They will be compared to the actual outcome of the validation simulations, which themselves are used to define the present “state-of-the-art”.

## 4.2 CFD models validated with MUST wind tunnel data

The present chapter concerns model validation, based on the MUST wind tunnel data. The focus here is on CFD models, although some of the tools described here are relevant for non-CFD models as well. A separate chapter (Chapter 4.3) is devoted to validation of non-CFD models.

Within the COST action and specifically for the MUST case, tools in the form of Excel workbooks have been developed to perform exploratory analyses of model performance and present results graphically, while also allowing for computation of metrics. These workbooks are the basis for much of the work on exploratory analysis and comparison of metrics, so they are discussed in considerable detail. The material related to the MUST experiment - the data set and the Excel tools - provide a unique opportunity to identify and explore patterns within model performance for a multitude of models and setups. They have been used within the COST action, but can also provide a framework for many future studies.

It seems pertinent here to highlight one particular recommendation from the *Model Evaluation Guidance and Protocol Document*: It is recommended to use 'exploratory data analysis' as one of the elements in model validation. It was confirmed repeatedly throughout the work with the exercise that such exploratory analysis is crucial to reveal shortcomings of models that might otherwise pass unnoticed. Thus, exploratory analysis is a vital element in quality assurance of models. The following sections will illustrate this by several examples.

The present chapter has the following subsections

- Overview of available material
- Introduction to the Excel tools
- Exploratory analysis
- Model performance in terms of metrics

Disclaimer: Note that the model results presented in this chapter represent a snapshot in time, in the sense that the results were delivered before October 31, 2008. In future, new model versions will appear, and changes to model setups and input parameters may prove pertinent, resulting in different model performance. In future, updated information may be found at the COST web site (see *URL 2*, http://www.mi.uni-hamburg.de/Home.484.0.html), or through the Atmospheric Dispersion Wiki, where anybody can contribute with comments (*URL 3*).

### 4.2.1 Overview of available material

The present chapter is supplemented by much additional material.

- Appendix A is an overview of the models that have participated in the MUST exercise (only models whose results were made public are included).

- Appendix B provides a large number of graphs with many results produced by different models and modelling teams, with the use of the Excel tools.

- The previously mentioned Excel-based tools that allow exploratory analyses of model performance. These files are available through the COST 732 home page or directly at web site *URL 4*. Besides the actual Excel files, the latter web site contains other

related information.

- There is a technical documentation on the Excel tools in the 16-page note "*Guide to Excel sheets for MUST exercise"* (Olesen and Berkowicz, 2008) found at *URL 4*. This can be regarded as a supplement to the present document. The note is recommended reading if you wish to prepare your own data for inclusion in the sheets. It discusses the coordinate system and the layout of the measurements, shows samples of the data illustrating the formats applied, and gives practical advice on how to use the tools. It includes an FAQ.

- A mechanism exists so that in future – after the end of the COST action – experiences on the MUST exercise can be reported and pooled. A Wiki on atmospheric dispersion modelling has a page prepared for this purpose (*URL 3*).

Other related documents are a number of papers presented at the 11th and 12th Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes (*URL 5* and *URL 6*).

### 4.2.2 Introduction to the Excel tools

As indicated above, tools in the form of Excel workbooks have been developed to allow exploratory analyses of model performance. They can be used by anybody interested in exploring existing results or as a framework to put new results into. There is a basic group of Excel files that allows *easy graphical inspection of the details of every case for all of the models*. Based on these 'full workbooks', another group of Excel workbooks have been derived, containing essential information extracted from the full workbooks. This extract is presented as summary plots and metrics.

There are a few additional files, e.g. one – referring to the case with dispersion – that uses Gaussian fitting to describe the pollution cloud in terms of Gaussian plume parameters. This makes it easy to obtain an impression of main features of the plume, as modelled and as observed.

An overview of the files mentioned in the following can be found in Table 4.1. There are three groups of workbooks:

- 'Full workbooks'
- 'Saved-metrics' workbooks
- Additional workbooks

There are multiple workbooks within each group, because three cases are simulated by the models: Two representing flow, and one representing dispersion. They are referred to as, respectively, the 0 degree flow case, the -45 degree flow case, and the -45 degree dispersion case. The layout of the two flow cases is shown in *Figure 4.2*.

#### **4.2.2.1** *'Full workbooks'*

The main group of Excel files is a set of files that allows *easy graphical inspection of the details of every case for all of the models* (here called**:** the *Full workbooks*). These files are very large. The largest (with results of dispersion modelling) has a size of more than 100 MB. The workbooks allow computation of metrics.

The Full workbooks are:

- A workbook for the 0 degree flow case, where flow results from a large number of models are collected (an example of the name of this workbook is **UVW-tke_0degree_**11nov08.xls – the date when the workbook was last modified is included in the name).

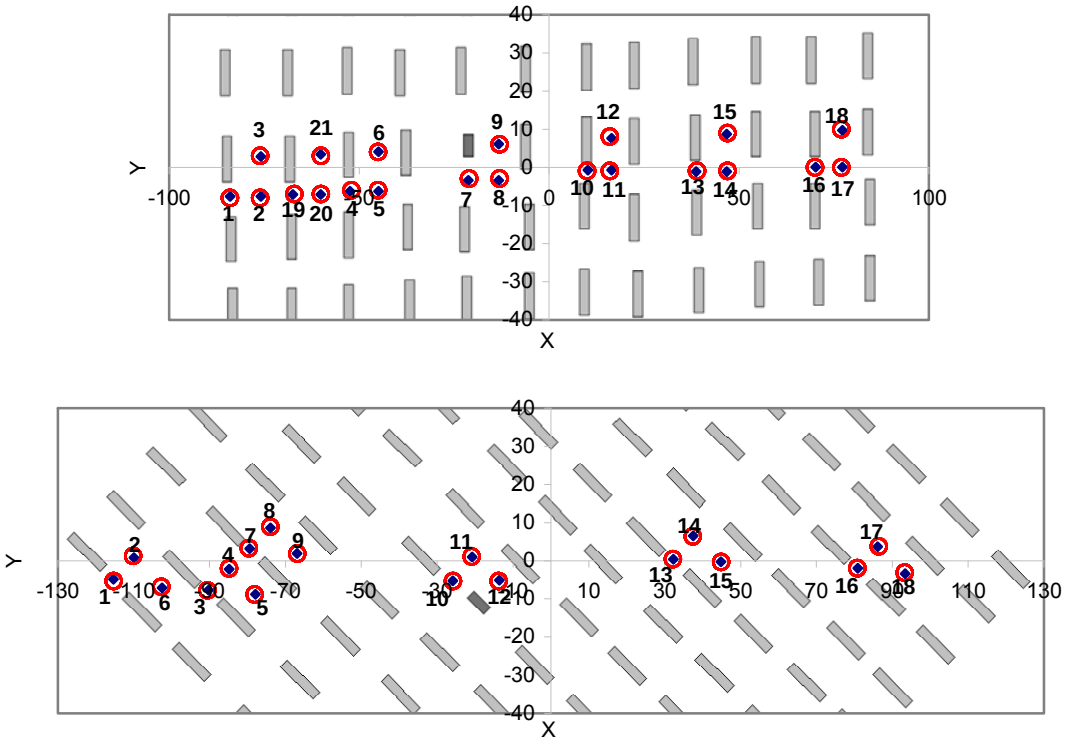



*Figure 4.2* Layout of measurements and the buildings for the flow cases. The vertical profiles of wind and turbulence components are measured at the locations ('towers') shown. The wind blows from the left.
Top: the 0 degree case. The points can be classified into three groups, so Tower 1 represents a 'Narrow Street', Tower 2 a 'Crossing', and Tower 3 a 'Wide Street'.
Bottom: the -45 degree case. Tower 1 represents a 'Narrow Street', Tower 2 a 'Crossing', and Tower 6 a 'Wide Street'.

- A workbook for the -45 degree flow case, where flow results from several models are collected (an example of the name of this workbook is **UVW-tke_45degree_**11nov08.xls).

- Two workbooks for the -45 dispersion degree case, where concentration results from several models are collected. The amount of data is huge, so one workbook is devoted to results from CFD models, while another is devoted to results from non-CFD models. (Examples of the names are **Dispersion_45degree_**11Nov08.xls and **Dispersion_nonCFD_**7Mar08.xls).

#### 4.2.2.2 *'Saved-metrics' workbooks*

Based on the 'full workbooks', another group of Excel workbooks have been derived, containing essential information extracted from the full workbooks. These are called the *'Saved-metrics' workbooks***.** They are smaller in size than the 'Full workbooks'.

For each of the 'Full workbooks', there are one or two workbooks with Saved Metrics, such as SavedMetrics_**Dispersion_45degree_**26May08.xls.

#### 4.2.2.3 *Additional workbooks*

Some additional workbooks exist. There are workbooks that have been used for submission of data to the person who collected results. It is convenient to submit data with these special workbooks, because they are more manageable in size than the Full workbooks.

Further, referring to the case with dispersion, there is a group of files called Plume_xxx.xls. In these files Gaussian fitting is used to describe the pollution cloud in terms of Gaussian plume parameters. This makes it easy to obtain an impression of main features of the plume, as modelled and as observed. This approach is possible in the case of the MUST wind tunnel experiment, because the plume is so coherent that Gaussian fitting to measurements and model results makes sense.

| **Full workbook** | **Sheet in Full workbook used to activate metrics** | **Sheet with metrics in Full workbook** | **Corresponding workbook in Saved-Metrics package** |
|---|---|---|---|
| UVWtke_0degree_xxx.xls | PlotsZ | MatchVertical | SavedMetrics_UWtke_0degree.xls |
| UVWtke_45degree_xxx.xls | PlotsZ | MatchVertical | SavedMetrics_UWtke_45degree.xls |
| UVWtke_45degree_xxx.xls | PlotsX | MatchHorisontal | SavedMetrics_UVtke_45degree.xls |
| Dispersion_45degree_xxx.xls | Plots | MatchConc | SavedMetrics_Dispersion_45degree.xls |
| Dispersion_nonCFD_xxx.xls | Plots | MatchConc | *Has not been produced*. |

Table 4.1 Relationship between 'Full workbooks' and 'Saved Metrics' workbooks.

#### 4.2.2.4 *Example of workbook contents: Detailed case studies*

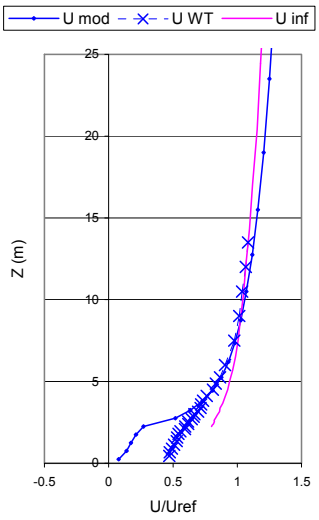
The workbooks allow inspection of vertical and/or horizontal profiles for many parameters, as predicted by a model and as observed in the wind tunnel. One example is shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

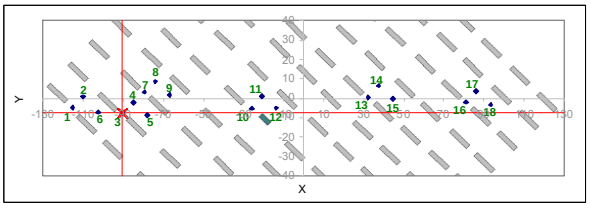
This figure shows a vertical profile of the u wind speed component (relative to the reference velocity $u_{ref}$) along a tower, taken from the workbook UVWtke_45degree_1apr08.xls. The Excel workbook indicates the tower position on the 'map' in the lower part of the figure. The measured values are indicated by crosses, while the other blue line represents model results for a certain model. The inflow profile is indicated by a thin violet line. The example reveals a problem with the specific model: It deviates strongly from the measured profile, and it has a kink at a height of approximately 2.5 m. A further analysis reveals that this is a symptom of a misplaced building in the model simulation. The Excel tool allows users to quickly browse through a series of profiles as the one illustrated; either model-by-model or tower-by-tower. Such browsing allows users to spot anomalies quickly.

#### 4.2.2.5 *Example of workbook contents: Summary of results, model by model*

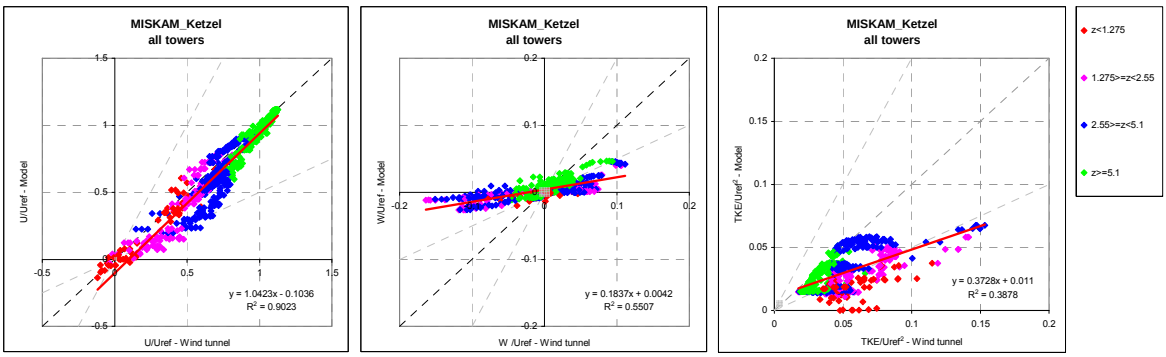
The Excel tools allow you to create a summary of information for all cases in the form of scatter plots and computation of metrics. Such summaries can be generated in the full workbook, and they are collected in the Saved-Metrics workbooks. *Figure 4.4* shows an example, referring to the 0 degree flow case. For one model - in this case MISKAM - information for all towers is summarised. There are 21 towers, with altogether 566 data pairs (observed, measured) for the parameters $u/u_{ref}$, $w/u_{ref}$ and $TKE/u_{ref}^2$. These are shown as scatter plots. Further, basic metrics have been computed and there is a 'map' showing the

location of measurements. The metrics have been computed with threshold values as specified in Table 4.6 (also indicated in the figure under the heading 'Settings').




.




| Metrics | | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 | FB | NMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | U | 566 | 0.684571822 | 0.609994092 | 0.283196 | 0.31076 | 0.413683 | 0.509448 | 0.9023 | 0.750883 | 0.897527 | N/A | N/A |
| | W | 566 | -0.00531686 | 0.003217093 | 0.054936 | 0.013599 | -10.3325 | 4.226981 | 0.5507 | 0.213781 | 0.293286 | N/A | N/A |
| | TKE | 566 | 0.051082543 | 0.030053474 | 0.024081 | 0.014417 | 0.471418 | 0.479728 | 0.3878 | 0.335689 | 0.671378 | 0.518366 | 0.519125 |

| Settings | W | D |
|---|---|---|
| U | 0.008 | 0.25 |
| W | 0.007 | 0.25 |
| TKE | 0.005 | 0.25 |

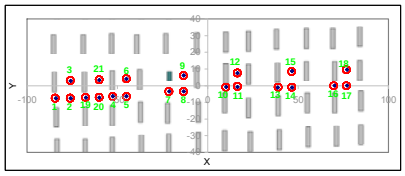
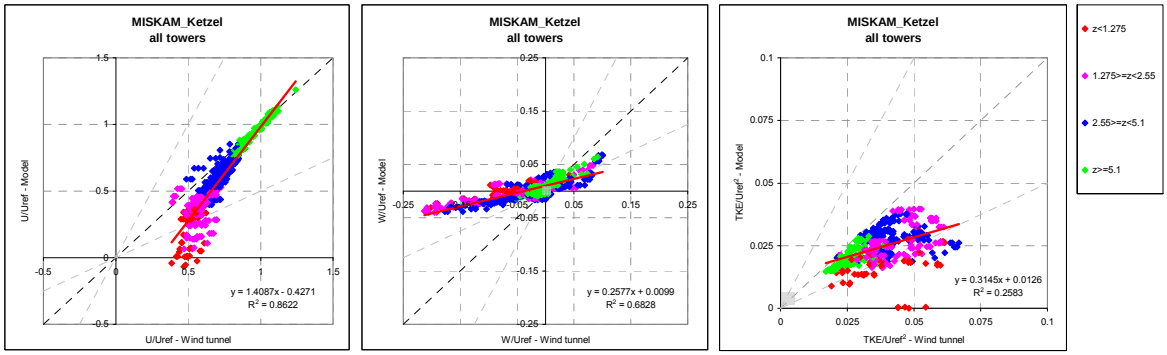


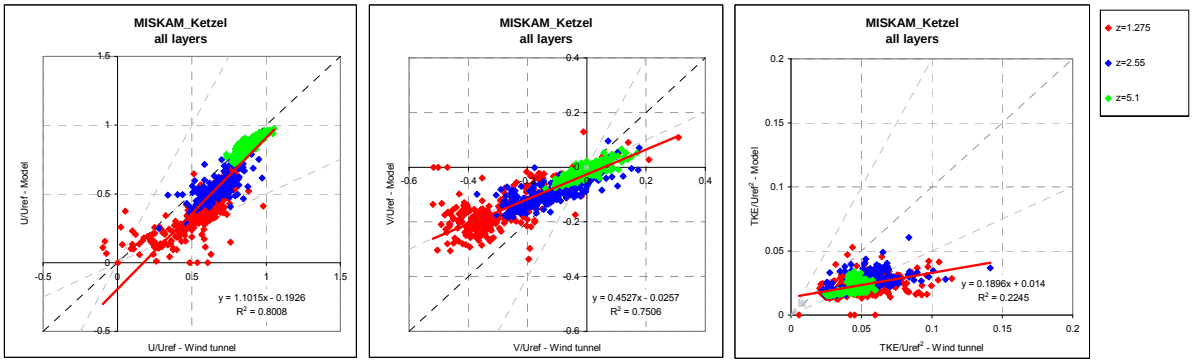

Figure 4.4: 0 degree flow case. Example of scatter plots and metrics for velocities for the MISKAM model, based on data from the towers. Parameters $u/u_{ref}$, $w/w_{ref}$ and $TKE/u_{ref}^2$. The figure is from SavedMetrics_UWtke_0degree_28Nov08.xls




| Metrics | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 | FB | NMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| U | 498 | 0.739746313 | 0.615004401 | 0.183552 | 0.278477 | 0.248128 | 0.452805 | 0.8622 | 0.692771 | 0.893574 | N/A | N/A |
| W | 498 | -0.033506719 | 0.001249413 | 0.064328 | 0.02006 | -1.91987 | 16.05564 | 0.6828 | 0.154618 | 0.23494 | N/A | N/A |
| TKE | 498 | 0.036350438 | 0.024079307 | 0.011115 | 0.006877 | 0.305778 | 0.285591 | 0.2583 | 0.483936 | 0.849398 | 0.406129 | 0.278229 |

| Settings | W | D |
|---|---|---|
| U | 0.008 | 0.25 |
| W | 0.007 | 0.25 |
| TKE | 0.005 | 0.25 |

Figure4.5: -45 degree flow case. Example of scatter plots and metrics for velocities for the MISKAM model, based on data from the towers. Parameters $u/u_{ref}$, $w/w_{ref}$ and $TKE/u_{ref}^2$. The figure is from Saved Metrics UWtke_45degree_28Nov08.xls




| Metrics | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 | FB | NMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| U | 787 | 0.691798333 | 0.569392714 | 0.193931 | 0.238704 | 0.280329 | 0.419226 | 0.8008 | 0.635324 | 0.936468 | N/A | N/A |
| V | 787 | -0.140079885 | -0.089132346 | 0.153602 | 0.080265 | -1.09653 | -0.90052 | 0.7506 | 0.198221 | 0.589581 | N/A | N/A |
| TKE | 787 | 0.048670019 | 0.023193262 | 0.016234 | 0.006496 | 0.33355 | 0.280072 | 0.2245 | 0.047014 | 0.496823 | 0.709034 | 0.75708 |

| Settings | W | D |
|---|---|---|
| U | 0.008 | 0.25 |
| V | 0.007 | 0.25 |
| TKE | 0.005 | 0.25 |

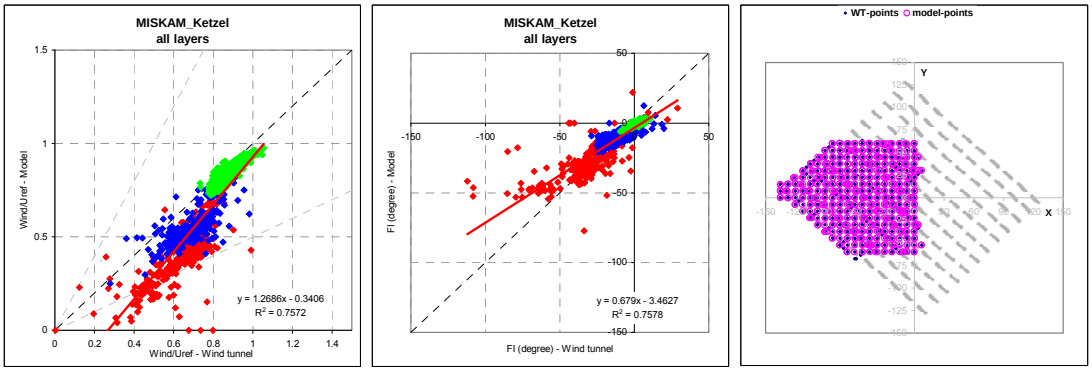
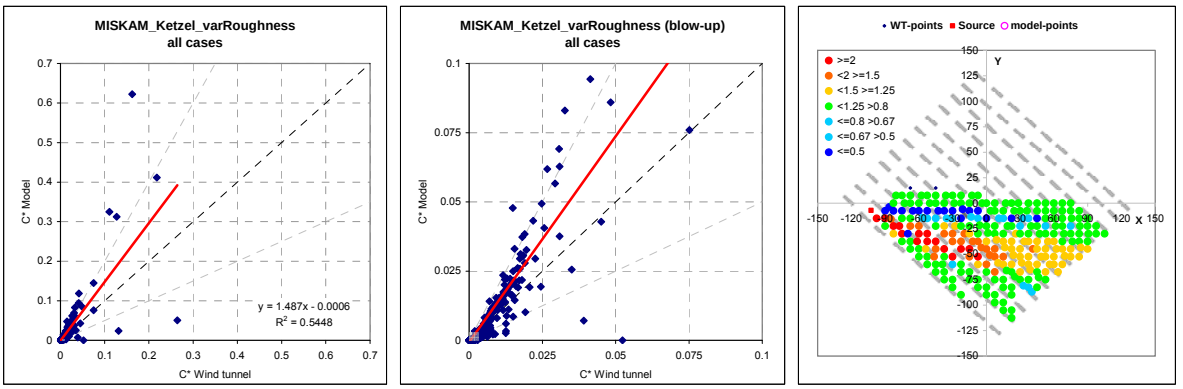



Figure 4.6: -45 degree flow case. Example of scatter plots and metrics for velocities for the MISKAM model, based on data from 'layers' – i.e. grids of measuring positions at three heights. In the first row parameters $u/u_{ref}$, $v/u_{ref}$ and $TKE/u_{ref}^2$. In the second row parameters $wind/u_{ref}$ and angle. 'wind' is wind speed derived from the u and v components. The figure is from SavedMetrics_UVtke_45degree_28Nov.xls

Appendix B shows such plots for a large number of models. The appendix allows you to gain an overview of similarities and differences in model performance. A closer analysis is also possible with the Excel tools. You may impose a filter on the data, thereby selecting a subset of data according to your choice. E.g., one can study the differences between model behaviour in 'Narrow Streets' versus 'Wide Streets' (defined in Figure 4.2). This question is pursued in Section 4.2.3.1. Many models share the same problems in reproducing the measurement results.

The use of the Excel tool is explained in a guide to the tool on the Web (Olesen and Berkowicz, 2008).

Summaries of results similar to those shown above can be produced for the case of -45 degree flow (examples in Figure4.5 and Figure 4.6), and for the case of -45 degree dispersion (example in Figure 4.7).




| Metrics | | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 | FB | NMSE | FB_FP | FB_FN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C* | 254 | 0.012732348 | 0.018308194 | 0.027915 | 0.056236 | 2.192477 | 3.071637 | 0.5448 | 0.649606 | 0.870079 | -0.359262 | 7.073891 | 0.512549 | 0.153287 |

| MG | VG | MG_FP | MG_FN |
|---|---|---|---|
| 0.928167 | 1.231622 | 1.191653 | 1.106052 |

| Settings | | W | D |
|---|---|---|---|
| | C | 0.003 | 0.25 |

Figure 4.7: -45 degree dispersion case. Example of scatter plots and metrics for concentrations computed with the MISKAM model. The figure in the middle is a blow-up of the figure on the left, in order to show behaviour for small concentrations more clearly. The figure to the right indicates measuring positions. They are colour-coded to indicate the ratio between model results and observations. The figure is from Saved Metrics_ Dispersion _45degree_xxx.xls

### 4.2.3 Exploratory analyses

Exploratory analysis should never be omitted when performing quality assurance of models. This has been confirmed in the process of the MUST model evaluation exercise. Even experienced modellers do sometimes perform errors, and these can be overlooked unless one inspects data, presented in graphical form. Conditions are best for detecting anomalies if you have a situation where *several models are put into a common framework*. This is the case for the MUST exercise. Such a framework makes it also possible to identify features common to several models, and thereby derive an indication of the state of art.

The following sections provide some examples of exploratory analysis applied to the MUST exercise.

As for all results presented in this chapter, note that the model results represent a snapshot in time, in the sense that the results were delivered before October 31, 2008. In future, new model versions will appear, and changes to model setups and input parameters may prove pertinent.

#### 4.2.3.1 *Flow modelling, -45 degree case*

Figure 4.2 shows the setup of the measuring positions where vertical profiles of u and w were measured. The locations can be classified into three groups according to the geometry: "Narrow streets" as exemplified by tower 1, "Crossings" exemplified by point 2, and "Wide streets" by tower 6.

Figure 4.8 shows for one model (Fluent) a scatter plot of measured versus modelled values for the horizontal, along-wind velocity component u, and the vertical component w. All data points (498) in the vertical profiles are included. The buildings generally have a height of 2.5 m. There are measurement positions both below and above roof height. It is important to note that the appearance of scatter plots like Figure 4.8 will depend very much on the proportion of data at various heights, and on how well the various measuring positions are represented. A similar statement applies to metrics such as Hit Rate. A value of a certain metric tells nothing unless it appears in a context.

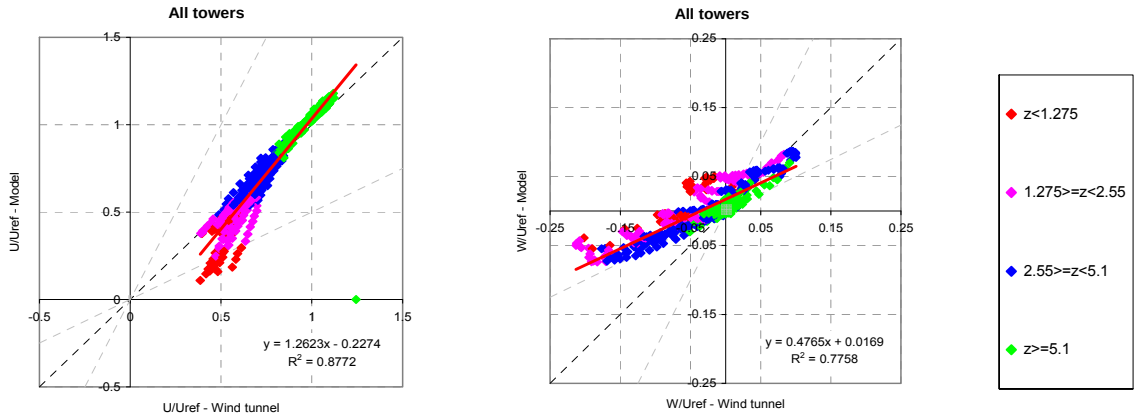



Figure 4.8: Flow, -45 degree case. Scatter plots for velocity components u (left) and w (right) for one particular model (Fluent). The points are colour-coded, with red indicating the lowest heights (below 1.275 m). The figure is from Saved Metrics _UWtke_45degree_04july.xls.

Therefore the results underlying the figure should be further examined. For the u component Figure 4.9: shows the same data as Figure 4.8, but split into three classes of data: "Wide streets", "Crossings", "Narrow Streets". The model appears much less successful in performance of u for Narrow Streets than for the other locations. Close to the ground the modelled u comes close to zero, whereas this is not the case for the measured values. However, one should also note, that due to the technical set-up of the measurements, only very few measuring points below 1.275 m exist in the case of Wide Streets and Crossings.

Concerning the vertical velocity w, Figure 4.8 shows that the range of modelled values is smaller than the range of measured values. In particular, predicted downward velocities are numerically much smaller than measured. The available material permits one to determine whether this behaviour of w is peculiar to the model examined or whether it is a common feature for many models.

Figure 4.10 shows the behaviour of w for a variety of models and model setups, performed by 16 different modelling groups. There are 3 sets of MISKAM runs, 5 of Fluent, and a variety of other models.

The figure reveals that it is a common feature for nearly all models that for downward vertical velocities, the numerical value of w is under-predicted. We may further note - although this is not illustrated - that the behaviour of w is quite distinct for the three groups of measurement positions.

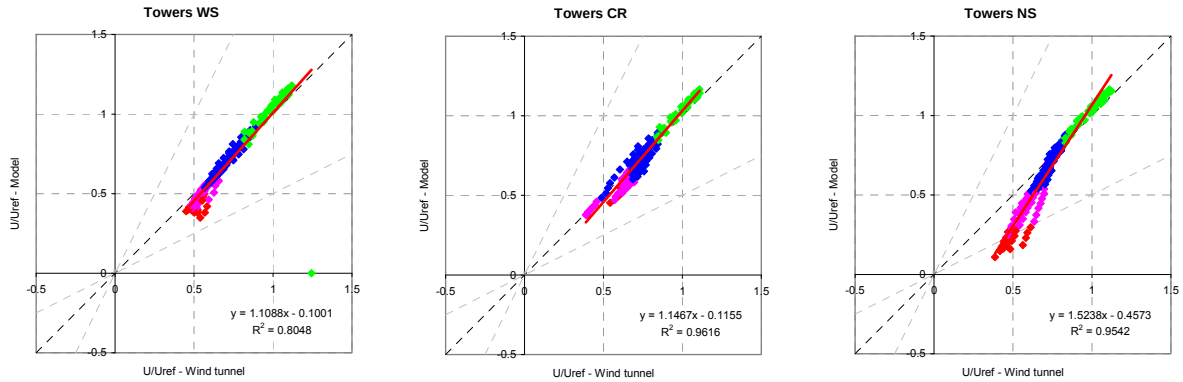



Figure 4.9: Flow, -45 degree case. Same data as Figure 4.8 for the u component, but split into 3 classes of data: "Wide streets" (WS), "Crossings" (CR), "Narrow Streets" (NS). The figure is from SavedMetrics_UWtke_45degree_04july.xls with a filter applied (as explained by Olesen and Berkowicz, 2008).

The models' ability to predict turbulent kinetic energy is in general poor. This can be seen by inspecting the figures in Appendix B.

#### 4.2.3.2 *Flow, 0 degree case*

Figure 4.2 shows the setup for the 0 degree flow case. Patterns of model behaviour in terms of prediction of velocity components and turbulent kinetic energy can be studied in a manner similar to the example discussed for the -45 degree case. There are plots for all models in Appendix B. One general conclusion on the state of the art is that all models have some difficulty in predicting u for *Wide Streets* – they generally predict too low a wind speed close to the roof top and slightly above it. Otherwise, prediction of u is generally good. Concerning w, the range of predicted w values is smaller than the range observed – similarly to the -45 degree case.

#### 4.2.3.3 *Dispersion, -45 degree case*

Figure 4.11 a shows the setup of the -45 degree dispersion case. In the wind tunnel the source was placed at ground level, while the concentration measurements were performed at a height of 1.275 m. Although the net of monitors at first sight looks dense, the sparseness of monitors is a problem when results are interpreted. Figure 4.11 b illustrates the difficulty in comparing measurements and modelled results. It shows a cross-section relatively close to the source (18 m from it). The modelled values are zero within a certain interval where there is a building. It is seen that the measurements lack a sufficiently detailed resolution to properly evaluate model performance. The presence of buildings complicates matters.

Although difficult to interpret, point-by-point comparisons do carry some information. For each measurement point, the measured value is compared to the modelled value at the closest point in the modelling grid. This will normally be a point outside buildings, unless the grid has been chosen with too poor resolution. However, for some of the models the data were provided interpolated exactly to the measuring points; in those cases interpolated values were used. Due to the discontinuity at the building walls (see e.g. Figure 4.11 b), such interpolation is dangerous and requires care if it shall not produce misleading values. Figure 4.12 shows a rather typical scatter plot of modelled versus measured concentrations for all measuring points (256). Despite the obvious mal-predictions for some high concentrations (close to the source), the Hit Rate is as high as 0.74 because there are many points with low concentration values where the model fits fairly well (further away from the source). 34% of the observations are below the threshold value, and such values are not so difficult to predict.

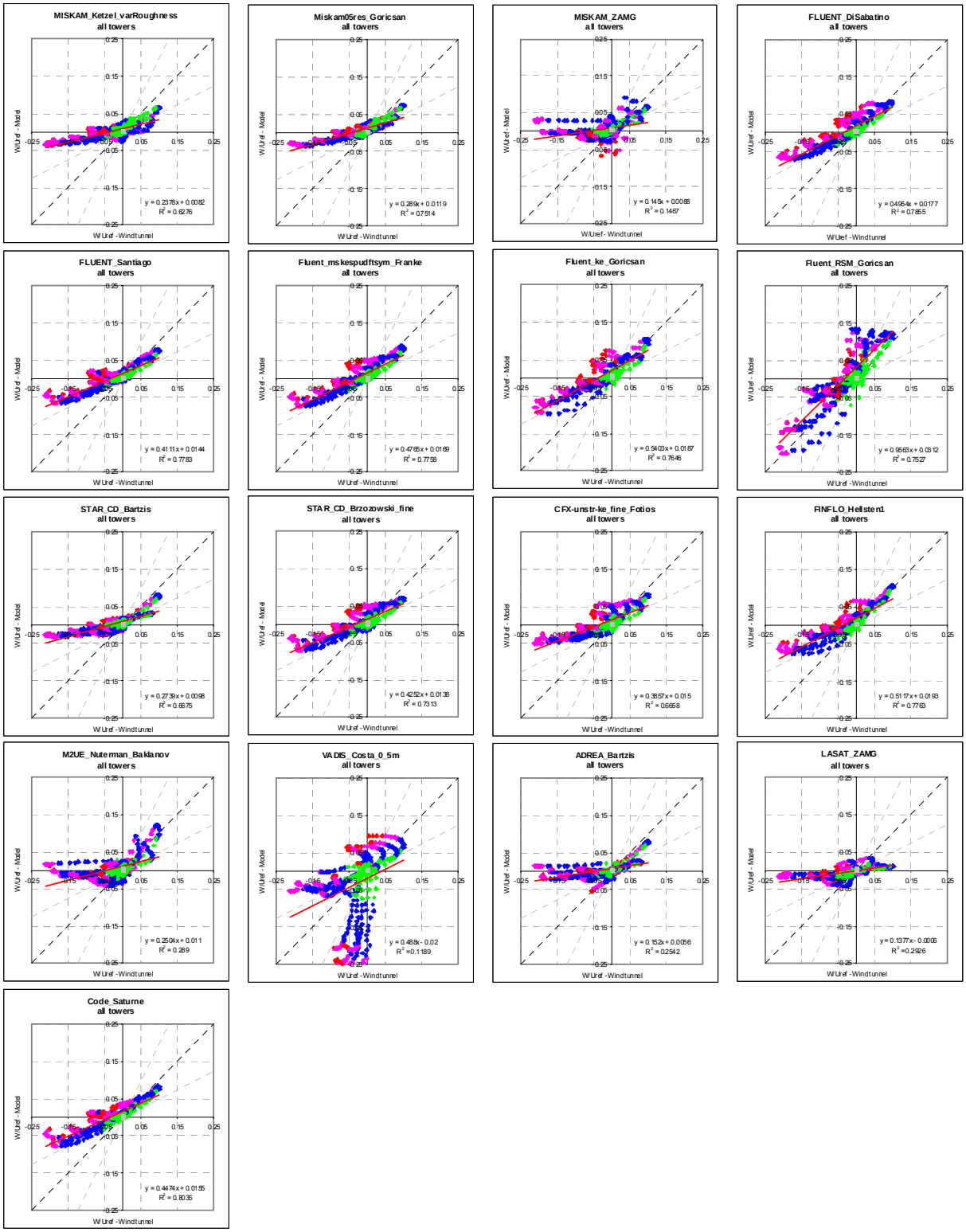



Figure 4.10: Flow, -45 degree case. Scatter plot of w for 17 models. Each model is only represented by one run by the same group, except that two of the Fluent runs are by the same group.

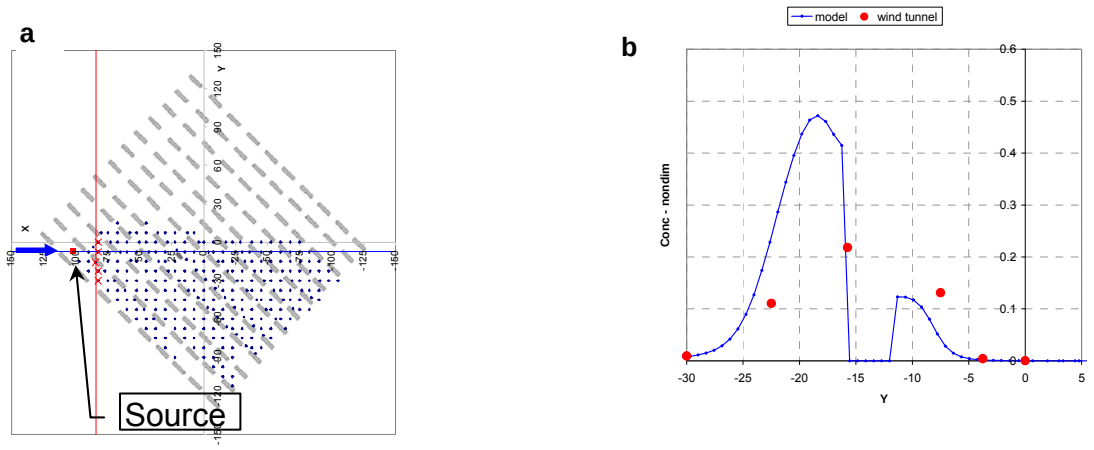



Figure 4.11 (a): Layout of dispersion experiment. The flow is from the left. Monitors are located at blue dots, the source at the red. (b): Profile of concentrations in cross-section at 18 m from the source, along the red line in (a). The modelled concentration is zero where there is a building.

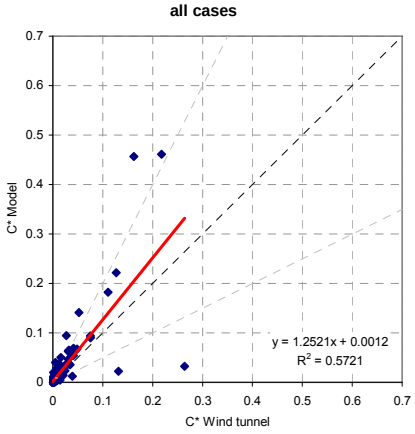



Figure 4.12: Scatter plots of concentrations for one of the Fluent runs.

Figure 4.13 is a map showing the plume trajectories according to the measurements with those predicted by 10 models, relatively close to the source. The trajectory has been determined by defining a series of cross-sections and fitting a Gaussian curve to each of these. Points within buildings are excluded from the fitting. This method was applied for both measurements and model results. There are only five points based on measurements, because only five cross-sections could be defined from available measurements in the area shown. Each cross-section with measured data include 4-10 data points, except for the cross-section closest to the source (x=-90 m), which is based on 3 points only.

The inspection of cross-sections shows that the plume is relatively coherent, so a Gaussian fit is a reasonable approximation in order to describe some main features of the pollution cloud. There is no a priori guarantee that the technique of using Gaussian fits is reasonable. In the given geometric configuration, the source is not placed in the immediate wake of a building, but mid-way between several buildings, and therefore the plume is coherent. With a different positioning of the source, the pollution cloud might have been less coherent.

Figure 4.13 shows that most models predict the plume trajectory well, but problems are revealed quite clearly. The figure illustrates well the usefulness of an exploratory approach

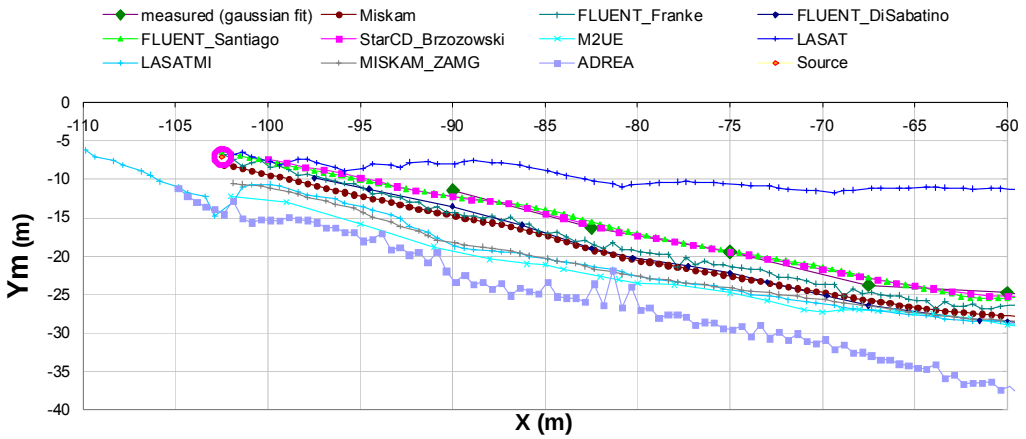



Figure 4.13: Map indicating plume trajectories predicted by 10 models in comparison to measurements. The source is inside the violet circle. There are 5 points based on measurements (large green diamonds for x values from -90 and larger). The map does not cover the entire array of containers, but extends only around 40 meters downwind of the source.

with graphical presentation of results. For instance, one of the models (M2UE) has a problem in the first part of the trajectory, close to the release point, while it converges towards measurements at further distances. It is interesting to note that in terms of Hit Rate, this model has the second-best score of all the 24 available model runs. This is so because the Hit Rate is based on 256 data pairs, while the displaced part of the plume only affects around 20 of these. This fact illustrates the point that metrics alone do not assure quality and that they can not be a substitute for looking at the data. Otherwise, it can be noted that the run labelled LASAT is the only model based on a diagnostic flow field model. This model does not properly predict the deflection of the plume trajectory by the buildings.

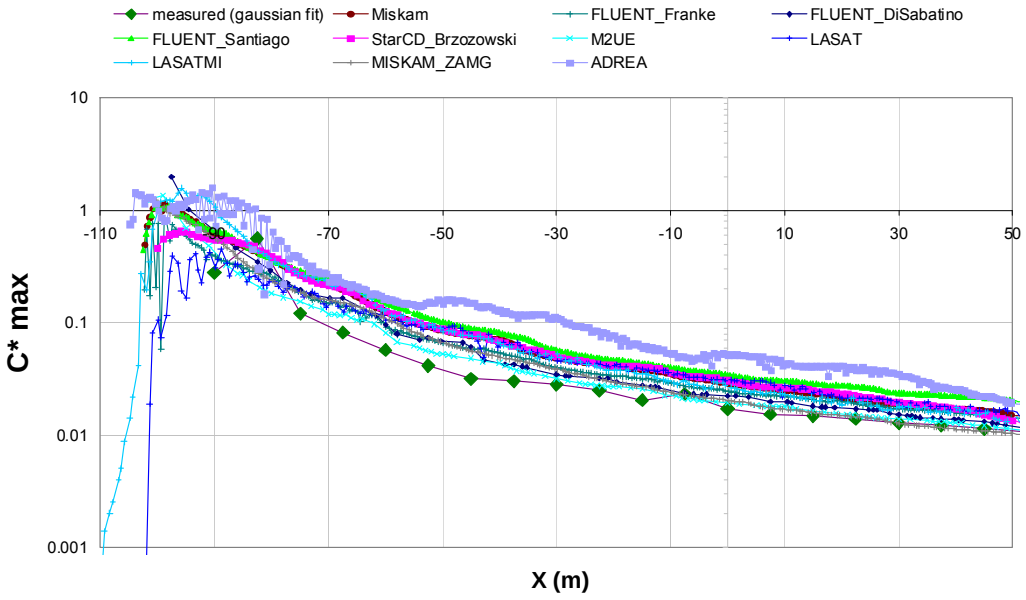



Figure 4.14: Estimated maximum plume concentrations using Gaussian fitting. Note that the vertical scale is logarithmic. The horizontal axis indicates the x coordinate; it extends further downwind than the corresponding axis in Figure 4.13.

Figure 4.14 shows the estimated maximum concentration in the plume, according to measurements and 10 models. The estimation of the maximum is obtained by Gaussian fitting. The estimate is rather uncertain at the two cross-sections closest to the release point where the plume is too narrow to be adequately resolved by measurements. Anyhow, at greater distances the figure shows a general trend which can also be discerned when looking at data in other ways: There is a common tendency for the models to predict too high values along the centre line.

### 4.2.4 Numerical error estimation

As stated in the Model Evaluation Guidance and Protocol Document (Britter and Schatzmann, 2007b) the verification of models and their results is an important part of the model evaluation. While verification of the models is used to find and reduce programming and logical errors in the models, the verification of the results is used to estimate the numerical errors present in the solutions that shall be used for validation. Verification of models is known as code verification and numerical error estimation as solution verification.

Both methods use generalized Richardson Extrapolation (RE) as explained briefly in the Background and Justification Document (Britter and Schatzmann, 2007a) and in more depth in the Best Practise Guideline of the action (Franke et al., 2007). For steady simulations with negligible errors from incomplete convergence, either iterative or towards a steady state, generalized RE is used to estimate the spatial discretisation error due to the finite resolution of the flow field with the computational grid. The procedure presented in the following was proposed by Eça and Hoekstra (2004) and used by Franke and Frank (2008) for the flow across an asymmetric street intersection.

#### 4.2.4.1 *Guide to interpretation of error estimates*

Generalised RE estimates the spatial discretisation error $\delta_i$ of a variable $f_i$ with the aid of a truncated Taylor expansion

$$\delta_i = f_i - f_{ex} \approx g h_i^p \qquad (1)$$

The three unknowns (exact solution $f_{ex}$, constant g and observed order of accuracy p) require solutions on three grids with characteristic grid spacings $h_i$. Therefore one has a fine solution (i=1), medium solution (i=2) and coarse solution (i=3).

Based on the ratio of solution changes $R=(f_2-f_1)/(f_3-f_2)$ the outcome of generalized RE is classified as (Eça and Hoekstra, 2004)

- $0 < R < 1$: monotonic convergence
- $-1 < R < 0$: oscillatory convergence
- $R < -1$: oscillatory divergence
- $R > 1$: monotonic divergence

In addition the case of converged solutions should be taken into account. Following the proposal of Cadafalch et al. (2002) a fifth point was included

- $(f_3/f_{3,max}-f_2/f_{2,max})\cdot (f_2/f_{2,max}-f_1/f_{1,max}) \leq C_0$

The maximum of the magnitude of the corresponding variable was used and $C_0$ = 1e-28.

In the case of monotonic convergence the observed order of accuracy p can be calculated from the implicit relation

$$p = \frac{\ln(1/R) - \left[\ln\left(r_{32}^p - 1\right) - \ln\left(r_{21}^p - 1\right)\right]}{\ln\left(r_{21}\right)} \qquad (2)$$

with the refinement ratios $r_{32} = h_3/h_2$ and $r_{21} = h_2/h_1$. In the present work the *regula falsi* was used for the iterative solution of (2). Iterations were stopped when the equation was either solved exactly or the change in p between two iterations was less than 1e-10. When the iterative solution did not converge, no error estimate was made at this measurement position.

Solution of (2) does not always yield values for p that are reasonable and could be directly inserted into (1) to calculate the discretisation error estimate. Therefore a further subdivision of the error band estimation for the solution on the finest grid ($\pm E_1$), based on the value of p, is necessary.

- $0.5 < p \leq 2$: $E_1 = 1.25 \mid \delta_1 \mid$
- $2 < p \leq 3$: $E_1 = 1.25 \max(\mid \delta_1 \mid, \mid \delta'_1 \mid)$

here $\delta'_1$ is the discretisation error estimate computed from

$$\delta'_i = f_i - f_{ex} \approx g_a h_i + g_b h_i^2 \qquad (3)$$

with the aid of the solution on the three grids (Roy, 2003; Eça and Hoekstra, 2004). Unknown are now the constants $g_a$ and $g_b$, and the exact solution $f_{ex}$.

- $0 < p \leq 0.5$ or $3 < p \leq 10$: $E_1 = 3 \max(\mid f_2 - f_1 \mid, \mid f_3 - f_2 \mid)$

When p is outside the range $0 < p \leq 10$, no error bar estimate is made.

In Table 4.2 the possible outcomes of the generalized RE are summarized together with the corresponding calculation of the error band estimation ($\pm E_1$).

| | Converged **(IV)** | $0 < R < 1$ | $-1 < R < 0$ **(V)** | $R < -1$ **(VI)** | $R > 1$ **(VII)** |
|---|---|---|---|---|---|
| $0.5 < p \leq 2$ | $E_1 = 0$ | $E_1 = 1.25 \mid \delta_1 \mid$ **(I)** | $E_1 = 3 \max(\mid f_2 - f_1 \mid, \mid f_3 - f_2 \mid)$ | | |
| $2 < p \leq 3$ | | $E_1 = 1.25 \max(\mid \delta_1 \mid, \mid \delta'_1 \mid)$ **(II)** | | | |
| $0 < p \leq 0.5$ or $3 < p \leq 10$ | | $E_1 = 3 \max(\mid f_2 - f_1 \mid, \mid f_3 - f_2 \mid)$ **(III)** | | | |
| $p \leq 0$ or $p > 10$ | | | | | |
| no iterative solution for p **(VIII)** | | | | | |

Table 4.2: Definition of node classifications and estimation of the numerical discretisation error bar ($\pm E_1$).

An error band is therefore only estimated for conditions I – V. For the other conditions no error band estimate is possible.

Average and RMS values are calculated for the observed order of accuracy p and the error band estimate $E_1$. For the average order of accuracy only positions with condition I and II are taken into account, while for the average error band positions with conditions I to V are used. In addition the average error band and its RMS value are normalized by $f_{1,max}$.

The procedure described above has been applied with the three systematically refined grids shown in Figure 4.15. The coarse mesh has $N_3$ = 751 194, the medium mesh $N_2$ = 1552792 and the fine mesh $N_1$ = 3213845 hexahedral cells. This leads to refinement ratios of $r_{21} = (N_1/N_2)^{1/3}$ = 1.27439451 and $r_{32} = (N_2/N_3)^{1/3}$ = 1.27385616. The ratios differ only slightly but require the iterative solution for p.

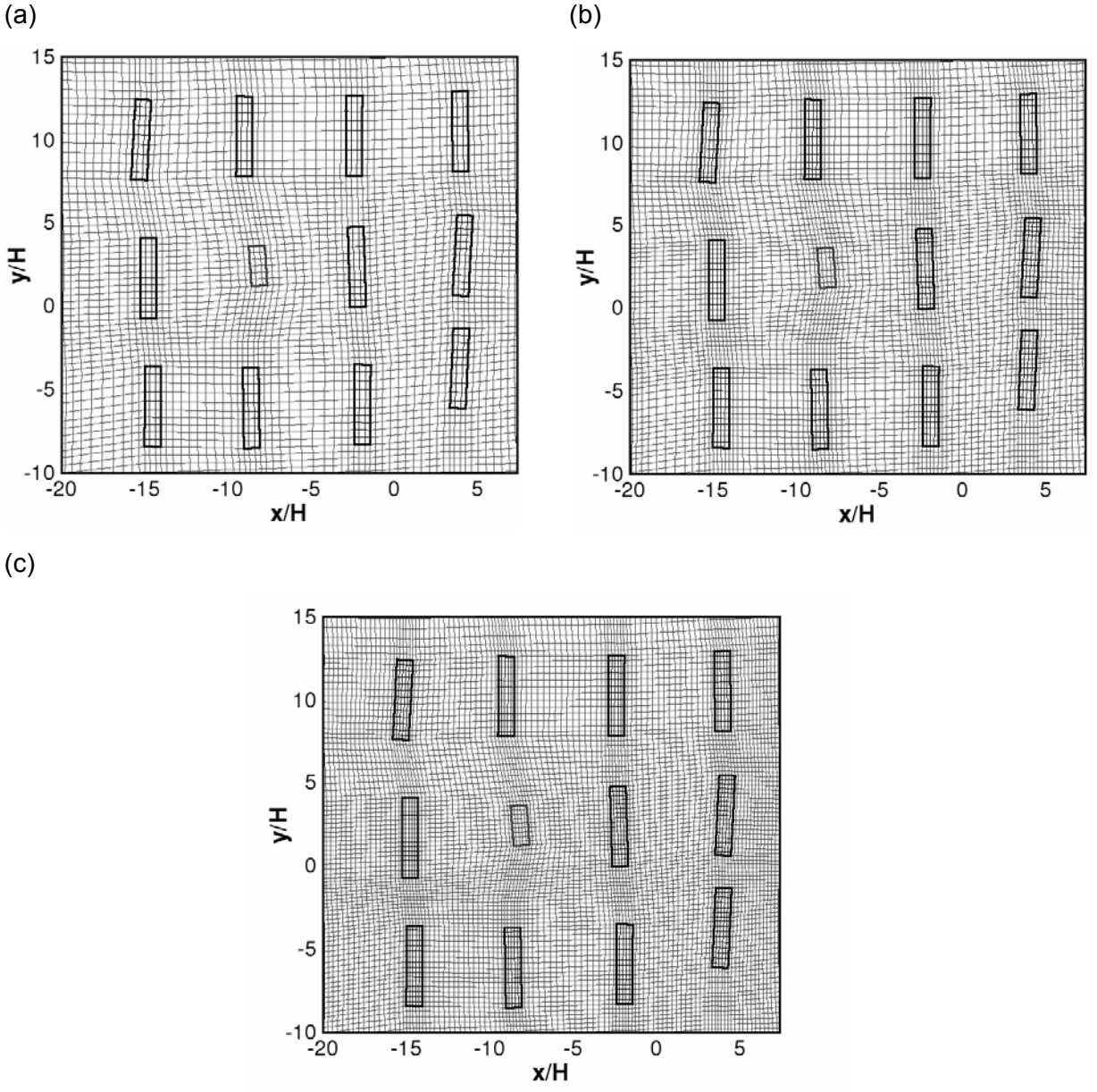



Figure 4.15: Detail of the three grids used for Richardson Extrapolation. (a) coarse grid, (b) medium grid, (c) fine grid.

#### 4.2.4.2 *Flow, 0 degree case*

This method was applied to the computed variables at the measurement positions. For the 0° case the coarse and fine uv measurement zones were examined, and the profile measurements. The fine uw zone was not yet analysed. The overall results are shown in Table 4.3.

| Variable | I [%] | II [%] | III [%] | IV [%] | V [%] | VI [%] | VII [%] | VIII [%] | <p> | $p_{rms}$ | $<E_1>$ [%] | $E_{1,rms}$ [%] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coarse and fine uv (N = 1179 measurement points) | | | | | | | | | | | | |
| $U/U_{ref}$ | 6.45 | 6.19 | 24.2 | 0.08 | 33.5 | 11.7 | 13.7 | 4.24 | 1.87 | 0.74 | 4.34 | 3.74 |
| $V/U_{ref}$ | 7.46 | 4.58 | 22.5 | 0.08 | 31.8 | 10.9 | 17.2 | 5.51 | 1.73 | 0.75 | 3.53 | 3.67 |
| $U_{rms}/U_{ref}$ | 7.21 | 3.33 | 18.1 | 0.00 | 29.9 | 14.9 | 20.1 | 6.53 | 1.60 | 0.68 | 4.77 | 4.63 |
| $V_{rms}/U_{ref}$ | 5.00 | 2.46 | 16.3 | 0.08 | 32.1 | 16.3 | 21.5 | 6.28 | 1.62 | 0.76 | 4.93 | 4.40 |
| $u'v'/U_{ref}^2$ | 6.62 | 4.41 | 15.4 | 0.08 | 29.3 | 19.5 | 20.3 | 4.41 | 1.75 | 0.72 | 7.28 | 7.54 |
| Profiles (N = 566 measurement points) | | | | | | | | | | | | |
| $U/U_{ref}$ | 8.66 | 8.30 | 23.0 | 0.18 | 18.4 | 10.8 | 23.9 | 6.89 | 1.84 | 0.69 | 2.18 | 1.88 |
| $W/U_{ref}$ | 8.13 | 6.01 | 15.7 | 0.18 | 18.4 | 8.13 | 42.0 | 1.41 | 1.76 | 0.71 | 7.46 | 5.42 |
| $U_{rms}/U_{ref}$ | 6.89 | 1.77 | 17.3 | 0.18 | 22.1 | 18.0 | 27.9 | 5.83 | 1.49 | 0.66 | 2.65 | 1.49 |
| $W_{rms}/U_{ref}$ | 7.07 | 4.59 | 17.1 | 0.18 | 19.6 | 17.0 | 30.6 | 3.89 | 1.83 | 0.73 | 2.49 | 1.75 |
| $u'w'/U_{ref}^2$ | 7.24 | 4.77 | 15.0 | 0.18 | 21.6 | 19.4 | 29.2 | 2.65 | 1.67 | 0.72 | 3.14 | 3.83 |

Table 4.3: Results for the estimates of the numerical errors of the 0 degree case. Percentages of different regimes as defined in section 4.2.4.1, mean observed order of accuracy and its RMS value, as well as normalised mean numerical error and its RMS value.

Total divergence of RE (VI-VIII) was obtained for 29.6% (U), 33.6% (V), 41.6% (Urms), 44.0% (Vrms) and 44.2% (u'v') of all measurement positions for the coarse and fine uv zones. For the profiles the corresponding numbers are 41.5% (U), 51.6% (W), 51.8% (Urms), 51.4% (Wrms) and 51.2% (u'w'). Better grid convergence is therefore obtained for the coarse and fine uv zones.

The values of the mean order of accuracy <p> do not differ much from each other. The same is true for the RMS values of p, although they are quite large. The number of measurement positions that were used for their calculation is however small, not exceeding 17% of all measurement positions (I-II). The average uncertainty on the other hand is moderate and with two exceptions always less than 5% of the maximum of the corresponding variable's magnitude.

In Figure 4.16 the mean velocities computed on the three grids are shown at three towers, which are representative of 'narrow streets' (Tower 1), 'crossings' (Tower 2) and 'wide streets' (Tower 3), cf. Figure 4.2 in Chapter 4.2.2. The experimental results and the estimated uncertainties on the fine grid are also shown as error bars. Error bars are only available at those measurement positions where an uncertainty estimate was possible.

For $U/U_{ref}$ all simulations are very close, especially above 2.5 times the container height H. The corresponding uncertainty is also small. Below 2.5H it increases, although the simulations also do not differ much. The largest uncertainties are obtained in the 'wide streets' where recirculation regions behind the containers exist.

For $W/U_{ref}$ the differences and uncertainty intervals are also small. The largest uncertainties are here however obtained for Tower 1 in the 'narrow street'. At Tower 3, where the largest differences between simulation and experiment exist, the solution on the fine grid is slightly closer to the experiments, but the uncertainty indicates that further refinement of the much will not lead to substantial improvements.

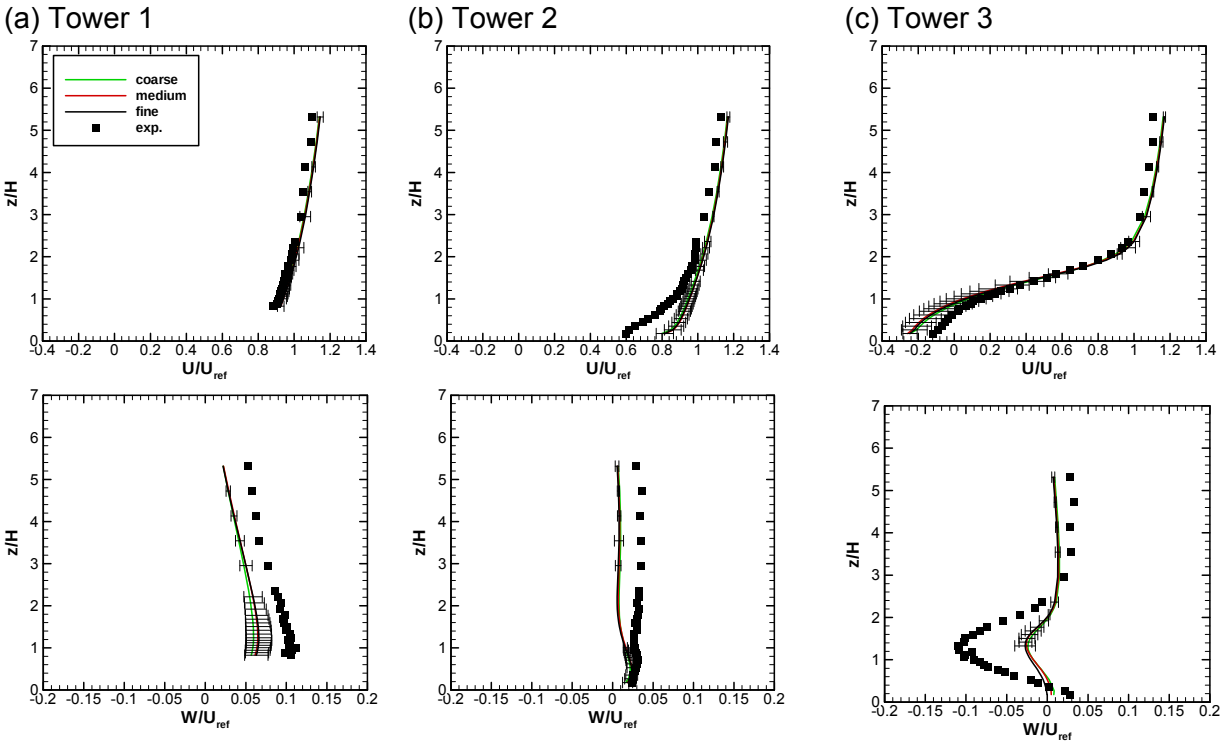



Figure 4.16: Profiles of $U/U_{ref}$ and $W/U_{ref}$ for the fine, medium and coarse mesh with numerical discretisation error bars for the solution on the fine grid. The results are representative for 'narrow streets' (Tower 1), 'crossings' (Tower 2) and 'wide streets' (Tower 3), as introduced in Figure 4.2 in Chapter 4.2.2.

Note that already the visual comparison of the solutions on the three grids leads to the conclusion that the medium and especially the fine mesh have a sufficient resolution, i.e. the solutions will not change substantially with further refinement. This qualitative statement based on a sensitivity study is then quantified by the estimate of the uncertainty. The size of the error bars can not be deduced from the graphical comparison of the three solutions.

#### 4.2.4.3 *Flow, -45 degree case*

The corresponding results of the uncertainty estimates for the –45° case are listed in Table 4.4.

| Variable | I [%] | II [%] | III [%] | IV [%] | V [%] | VI [%] | VII [%] | VIII [%] | <p> | $p_{rms}$ | $<E_1>$ [%] | $E_{1,rms}$ [%] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coarse and fine uv (N = 1219 measurement points) | | | | | | | | | | | | |
| $U/U_{ref}$ | 10.2 | 10.0 | 29.3 | 0.08 | 15.8 | 12.0 | 17.6 | 5.00 | 1.86 | 0.72 | 3.50 | 5.02 |
| $V/U_{ref}$ | 14.2 | 7.55 | 25.0 | 0.08 | 11.5 | 9.43 | 28.7 | 3.53 | 1.67 | 0.70 | 5.21 | 7.20 |
| $U_{rms}/U_{ref}$ | 12.9 | 9.02 | 29.7 | 0.08 | 14.0 | 9.76 | 18.9 | 5.58 | 1.81 | 0.70 | 6.27 | 7.08 |
| $V_{rms}/U_{ref}$ | 10.3 | 7.05 | 25.0 | 0.08 | 20.8 | 9.84 | 22.1 | 4.84 | 1.77 | 0.71 | 7.17 | 9.88 |
| $u'v'/U_{ref}^2$ | 13.9 | 7.22 | 24.0 | 0.08 | 17.1 | 8.12 | 26.0 | 3.45 | 1.65 | 0.73 | 11.5 | 12.3 |
| Profiles (N = 498 measurement points) | | | | | | | | | | | | |
| $U/U_{ref}$ | 10.8 | 7.83 | 27.7 | 0.04 | 15.5 | 8.63 | 26.9 | 2.21 | 1.77 | 0.72 | 2.45 | 1.82 |
| $W/U_{ref}$ | 17.9 | 10.0 | 14.9 | 0.04 | 9.64 | 6.43 | 38.8 | 2.01 | 1.82 | 0.63 | 7.89 | 4.40 |
| $U_{rms}/U_{ref}$ | 8.43 | 3.01 | 22.7 | 0.04 | 15.9 | 9.24 | 36.9 | 3.41 | 1.50 | 0.65 | 4.08 | 4.17 |
| $W_{rms}/U_{ref}$ | 5.82 | 6.83 | 22.5 | 0.04 | 12.0 | 14.1 | 35.9 | 2.41 | 1.87 | 0.73 | 4.51 | 4.59 |
| $u'w'/U_{ref}^2$ | 14.1 | 11.4 | 24.3 | 0.02 | 9.04 | 6.43 | 29.7 | 4.82 | 1.76 | 0.78 | 9.33 | 13.6 |

Table 4.4: Results for the estimates of the numerical errors of the –45 degree case. Percentages of different regimes as defined in section 4.2.4.1, mean observed order of

accuracy and its RMS value, as well as normalised mean numerical error and its RMS value.

The results are similar to the ones obtained for the 0 degree case. The total divergence at the coarse and fine uv planes is 34.6% (U), 41.7% (V), 34.3% (Urms), 36.8% (Vrms), 37.6% (u'v'). Compared with the 0 degree case the divergence is therefore slighty larger for the mean velocities and slightly smaller for the turbulence quantities. For the profiles all variables except Wrms have a smaller total divergence, 37.8% (U), 47.2% (W), 49.6% (Urms), 52.4% (Wrms), 41.0% (u'w').

The mean order of accuracy and its RMS value are also similar to the 0 degree results. But the average uncertainty and its RMS value are different. Except for $U/U_{ref}$ at the coarse and fine uv planes, all average uncertainties are larger than for the 0 degree case. The largest uncertainties exist for the Reynolds shear stresses, but also the average uncertainty for W/Uref at the towers (Profiles) is app. 8%. In addition nearly all RMS values also increase.

In Figure 4.17 results for the estimated error bars are shown at three towers for U and W. The results are representative for 'narrow streets' (Tower 1), 'crossings' (Tower 6) and 'wide streets' (Tower 2), cf. Figure 4.2 in Chapter 4.2.2.

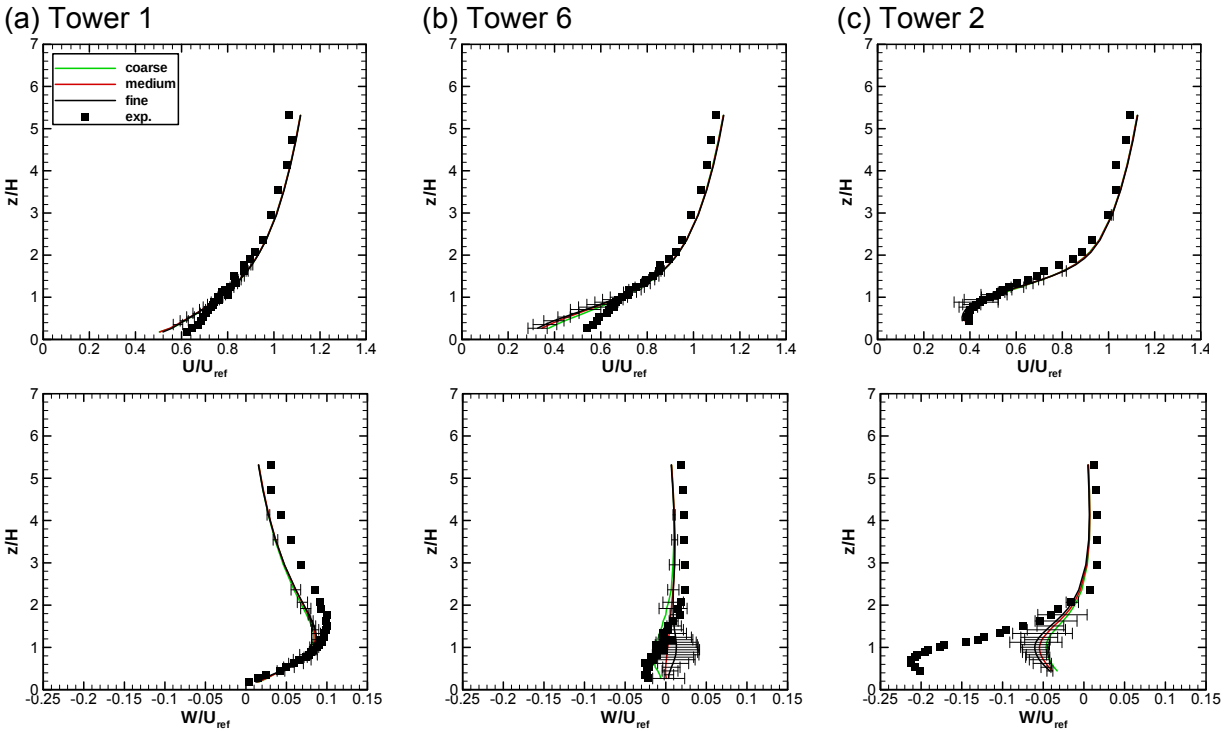



Figure 4.17: Profiles of U and W for the fine, medium and coarse mesh with numerical discretisation error bars for the solution on the fine grid. The results are representative for 'narrow streets' (Tower 1), 'crossings' (Tower 6) and 'wide streets' (Tower 2), as introduced in Figure 4.2 in Chapter 4.2.2.

While for $U/U_{ref}$ the results are similar to the ones of the 0 degree case shown in Figure 4.16, the uncertainty of W/Uref is much larger than in the 0 degree case, especially for the 'crossings' (Tower 6) and 'wide streets' (Tower 2). For Tower 6 the simulation results without error bars indicate that with further grid refinement the numerical simulation will differ more from the experiments. The large error bars do however show that with further refinement the numerical solution might as well come closer to the experiments. For Tower 2 one can deduce from the error bars that a further refinement might improve the agreement between simulation and experiment, but substantial differences will still remain.

#### 4.2.4.4 *Dispersion, -45 degree case*

Finally for the concentrations the following results for the uncertainty estimates are obtained, Table 4.5.

| Variable | I [%] | II [%] | III [%] | IV [%] | V [%] | VI [%] | VII [%] | VIII [%] | <p> | $p_{rms}$ | $<E_1>$ [%] | $E_{1,rms}$ [%] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C* | 0.08 | 0.00 | 2.73 | 0.39 | 33.6 | 58.2 | 3.52 | 0.78 | 1.31 | 0.97 | 2.67 | 8.38 |

Table 4.5: Results for the estimates of the numerical errors of the normalised concentrations for the -45° case at N = 256 measurement positions. Percentages of different regimes as defined in section 4.2.4.1, mean observed order of accuracy and its RMS value, as well as normalised mean numerical error and its RMS value.

These results are much worse than the results for the flow with a total of 62.5% of the measurement positions with divergence. They can be also seen in Figure 4.18. Positions with |R| > 1 are diverged, when one omits the small number of nodes (0.78%) where the iterative solution for p did not converge. Note that the small average uncertainty is due to the fact that uncertainty estimates could be made at only 37.5% of the positions.

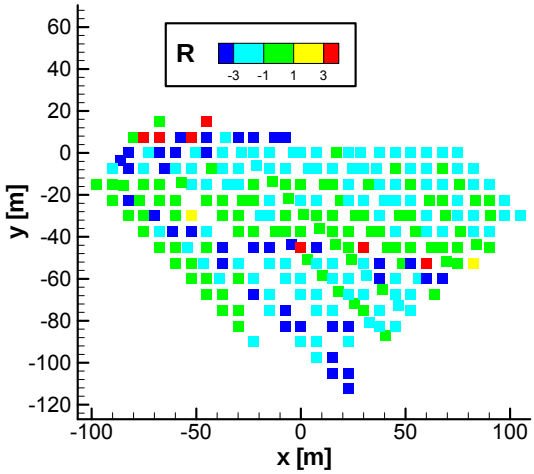



Figure 4.18: Ratio of solution changes (R) at concentration measurement positions. Only positions with |R| < 1 (green) allow for an error estimation in the context of the presented procedure.

The reason for the large number of positions without error estimate is most probably the fact that the source volume used for the three grids could not be kept constant due to the employed refinement ratios and the fact that the source volume was not defined as a geometric entity in the beginning, like in the Oklahoma city case, see Chapter 3.2.2.3. Therefore not only the mesh resolution did change with refinement, but also the source volume. For a correct application of Richardson extrapolation the source volume must be kept constant.

### 4.2.5 Model performance in terms of metrics

#### 4.2.5.1 Guide to interpretation of metrics

There are some things to keep in mind when considering model performance metrics: In general, it can be stated that *by just looking at metrics one cannot assure quality. Metrics carry some information, but they do not tell the whole story. They are no substitute for looking at data.*

Further, one should keep in mind that the results as presented here represent a snapshot in time, and if the MUST exercise is repeated at a later date, model performance may very well be different from that exposed here.

In general there is an important point to keep in mind before considering the values of any metrics: When comparing observations with model results one will often be tempted to assume that for a good model the bias should be as close to zero as possible. *However, this should not always be taken for granted*. In the present context there has been concern about the fact that *'TKE' derived from wind tunnel measurements is not the true TKE*, but an artefact, because only two of the three turbulent wind components were measured. This problem has been investigated with the aid of several simulation results from three different models and it was found that the influence on the results is, in general, negligible. Franke (2009b) shows this comparison for the hit rates of the 0 degree and –45 degree flow case, based on his FLUENT results. Note that the performance measures *Fractional bias (FB) and Normalised Mean Square Error (NMSE) make no sense for variables that can take both negative and positive values*, such as velocity components. Although they can be computed for such variables, they should not be used. In the Excel workbooks dated earlier than November 2008, FB and NMSE were computed and displayed for all variables, but should be disregarded for the velocity components u, v and w.

When comparing metrics for several models *it is essential that the comparison takes place for the same set of measurement points*. This statement implies that for model intercomparison metrics should not be computed by each modeller individually, but with a tool such as the Excel workbooks developed for the COST action. Use of the workbooks makes it transparent which data are used to compute the metrics and helps in ensuring consistency. This does, however, require that simulation results are made available at all measurement positions.

In the following chapters metrics are therefore only presented for those simulations that provided results at all measurement positions inside the computational domain. In addition simulation results that are obviously deficient were left out. The metrics for all simulations listed in Tables 3.4 – 3.6 in Chapter 3.2.2 can be extracted from the Excel workbook. The number of results left after excluding several simulations due to the aforementioned reasons is still large enough to perform a statistical analysis of the ensemble of individual results. From the individual metrics the mean ($Y$) and the median ($Z$) are computed and given together with the corresponding standard deviations ($S$, $T$):

$$Y = 1/M \sum_{j=1}^{M} X_j \,,$$

$$S = \left[1/(M-1)\sum_{j=1}^{M}\left(X_j - Y\right)^2\right]^{1/2},$$

$$T = 1/0.6745\sqrt{M/(M-1)}\,median\left(\left|X_i - Z\right|\right)$$

$T$ is the standard deviation based on the sample median absolute deviation (MAD), see e.g. Hemsch (2000). $X_j$ are the individual metrics and $M$ is the number of simulation results. From the standard deviations scatter limits are computed, corresponding to 95% confidence intervals around the mean, $Y \pm P_{S,95}$ and $Z \pm P_{T,95}$, based on the Student t distribution (Coleman and Steele, 1999).

The scatter of the results can then be regarded as the *reproducibility* of the computational

process. In experiments the *reproducibility* is obtained by replicating measurements under changed conditions and called N-th Order replication. In numerical simulation this procedure is known as N-Version testing and the changed conditions correspond to the different codes, computational domains, boundary conditions, grids, turbulence models, numerical parameters, observers, etc. The outcome of this procedure is a collective result, the mean or median. The small scatter around this collective result, the 95% confidence interval, corresponds to noise. Then only results outside the confidence intervals are significant for further analysis (Hemsch, 2000).

The average metrics and corresponding confidence intervals can then also be viewed as indication of the present "state-of-the-art" of the models, when applied to the validation test cases of COST action 732.

#### **4.2.5.2** *Threshold values*

As noted in the *Model Evaluation Guidance and Protocol Document* (Britter and Schatzmann, 2007b) for the factor of 2 (FAC2) and the geometric metrics MG and VG, treatment of zeros or relatively small data requires attention. There it was stated that for the calculation of FAC2, MG and VG the allowed absolute deviation W of the hit rate is used in the present exercises. This threshold is prescribed for each of the investigated variables, i.e. the non-dimensional velocity components, the non-dimensional turbulent kinetic energy and the scaled concentrations C*. The scaled concentrations are defined as

$$C^* = \frac{C \cdot u_{ref} \cdot H^2}{Q_{source}}$$

with $C$ = concentration, $u_{ref}$ = reference velocity in x-direction at (x,y,z) = (-144m,-2.25m,7.29m), the container height $H$ = 2.54m and $Q_{source}$ the volumetric source flow rate.

The values for W given in Table 4.6 are the experimentally determined uncertainties of the measured variables. They do represent the repeatability of the experiments. Contrary to the VDI guideline (VDI, 2005), the estimated error due to interpolation onto the measurement locations has not been taken into account in the present exercises. As the VDI guideline uses 0.05 for this contribution, the present threshold is approximately one order of magnitude smaller, which must be kept in mind, when relating the present results for the hit rate to the results of the validation cases of the VDI guideline. For the allowed relative difference D the same value as in the VDI guideline was used, i.e. D = 0.25.

Furthermore it should be noted that the influence of different interpolation methods has been investigated for this exercise. It was found that the interpolation method has only a small influence on the results, and that the differences between the interpolation methods decrease with increasing grid resolution.

| | Allowed absolute deviation W, referring to Hit rate. The same value is used as threshold when FAC2 is determined and, in the case of concentrations, for MG and VG. | Allowed fractional deviation D, referring to Hit rate |
|---|---|---|
| $u/u_{ref}$ (velocity component) | 0.008 | 0.25 |
| $v/u_{ref}$ (velocity component) | 0.007 | 0.25 |
| $w/u_{ref}$ (velocity component) | 0.007 | 0.25 |
| $TKE/u^2_{ref}$ | 0.005 | 0.25 |
| C* (concentration) | 0.003 | 0.25 |

Table 4.6: Threshold values for Hit Rate, FAC2, MG and VG Metrics for the 0 degree flow case

#### 4.2.5.3 Metrics for the 0 degree flow case

For the 0 degree flow case metrics for the velocity components and the turbulent kinetic energy are available from the Excel 'Saved-metrics' workbook only for the towers, shown in Figure 4.2. Metrics for the horizontal layers, including also the Reynolds stress components, can be found e.g. in Franke (2009b, 2009a).

In Figure 4.19 the running record of the hit rate for the velocity components is shown, together with the mean, median and the 95% confidence interval. While the velocity component $u/u_{ref}$ is predicted very well by nearly all simulations, the vertical velocity component $w/u_{ref}$ has much lower hit rates, with a maximum at $q$ = 0.31. In the exploratory data analysis it was said that all runs failed to predict the strong downward flow, leading to these low hit rates. For the horizontal component the high hit rates are mostly due to the good agreement between simulation and experiment above roof height, while in between the containers the hit rates are much smaller (Franke, 2009a).

The mean and median hit rates are $Y$ = 0.72, $Z$ = 0.73 and $Y$ = 0.21, $Z$ = 0.20 for the horizontal and vertical component, respectively. The corresponding 95% uncertainty limits are $P_{S,95}$ = 0.08 = 0.11$Y$ and $P_{T,95}$ = 0.10 = 0.14$Z$ for the horizontal velocity component and $P_{S,95}$ = 0.14 = 0.67$Y$ and $P_{T,95}$ = 0.15 = 0.75$Z$ for the vertical velocity component, see also Table 4.7. While for the horizontal velocity component the relative uncertainty range is rather narrow, it is wide for the vertical velocity component. However, there is only one outlier for the vertical velocity component. This outlier corresponds to the simulation FI00_019, which is identical to the simulation FI00_020 except for the fact that the reference velocity at the inlet in the former run is only one tenth of the one used for FI00_020.

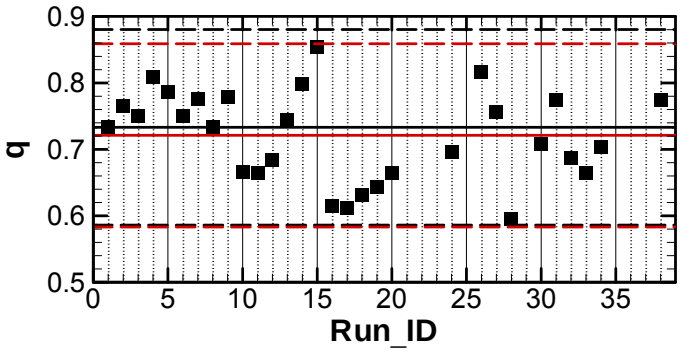
The reason for the large difference in the w component is not quite clear, as other groups did also check the influence of the magnitude of the reference velocity at the inlet and found no change for the non dimensional velocities. The other groups did however use standard wall functions, while in run FI00_019 non equilibrium wall functions were used, see Appendix A. This might be the reason for the differences in the w component. For other quantities the differences are smaller.

In Figure 4.20 the metrics for the turbulent kinetic energy $TKE/u_{ref}^2$ are shown. Due to $TKE/u_{ref}^2$ being a positive quantity also the factor of two (FAC2) and the fractional bias (FB) are shown together with the hit rate (q). The hit rates are again rather low, which is mainly due to the differences between the measured and prescribed turbulent kinetic energy at the inlet.

Most simulation runs were performed with a constant $TKE/u_{ref}^2$ at the inlet, derived from the assumption of an equilibrium boundary layer flow. As shown in Figure 3.3, Chapter 3.2.2, this $TKE/u_{ref}^2$ is much lower than the measured one. The too low $TKE/u_{ref}^2$ is then transported nearly unaltered towards the measurement locations. Especially at the measurement towers in the 'narrow streets' and 'crossings' (cf. Figure 4.2 in Chapter 4.2.2) the differences between the observed and predicted $TKE/u_{ref}^2$ are large. In the 'wide streets' the agreement is better, due to the turbulence produced in front and behind the containers. The under-prediction of $TKE/u_{ref}^2$ is also very well visible in the values of FB. All simulations have FB > 0 and therefore predict a too low $TKE/u_{ref}^2$.

The results for FAC2 show a more positive picture than the results for the hit rate. For FAC2 all simulations have FAC2 > 0.5. The median and mean are even $Z$ = 0.77 and $Y$ = 0.74, respectively. The corresponding 95% uncertainty limits are $P_{S,95}$ = 0.20 = 0.27$Y$ and $P_{T,95}$ = 0.23 = 0.30$Z$, see Table 4.7. For FB one has $Z$ = 0.40, $Y$ = 0.43, $P_{S,95}$ = 0.23 = 0.53$Y$ and $P_{T,95}$ = 0.31 = 0.78$Z$, and for the hit rate $Z$ = 0.35, $Y$ = 0.35, $P_{S,95}$ = 0.20 = 0.57$Y$ and $P_{T,95}$ = 0.28 = 0.80$Z$. Except for FAC2 there is a substantial scatter in the individual metrics, but no outliers can be identified.

a)




b)

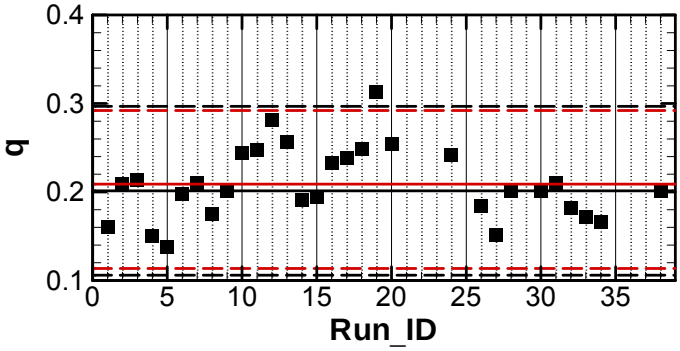
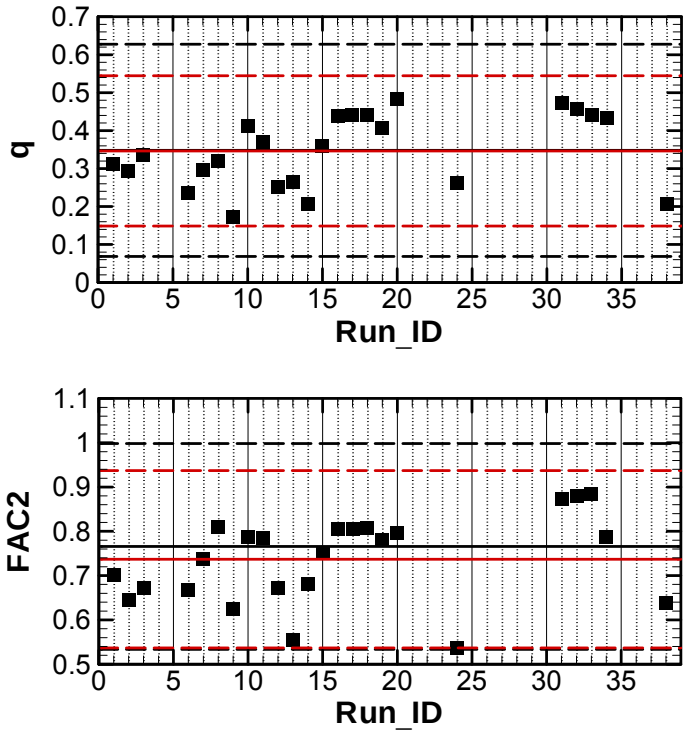



Figure 4.19: Running record of velocity component hit rates for the 0 degree case. a) $u/u_{ref}$, b) $w/u_{ref}$. ■ : individual metrics $X_j$; — : median $Z$; — : mean $Y$; - - - : 95% uncertainty interval $Y \pm P_{T,95}$; - - - : 95% uncertainty interval $Y \pm P_{S,95}$.



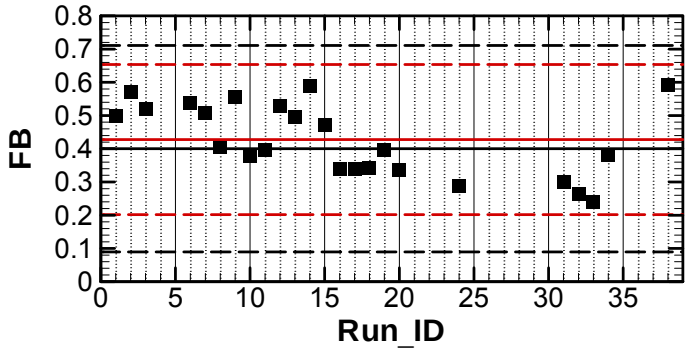



Figure 4.20: Running record of 0 degree turbulent kinetic energy hit rate (q), factor of two (FAC2) and fractional bias (FB), from top to bottom. ■ : individual metrics $X_j$; — : median $Z$; — : mean $Y$; - - - : 95% uncertainty interval $Y \pm P_{T,95}$; - - - : 95% uncertainty interval $Y \pm P_{S,95}$.

| | $u/u_{ref}$ | $w/u_{ref}$ | $TKE/U_{ref}^2$ | | | |
|---|---|---|---|---|---|---|
| | q | q | q | FAC2 | FB | NMSE |
| Median | 0.73 | 0.20 | 0.35 | 0.77 | 0.40 | 0.39 |
| T | 0.07 | 0.05 | 0.14 | 0.11 | 0.15 | 0.23 |
| Mean | 0.72 | 0.21 | 0.35 | 0.74 | 0.43 | 0.39 |
| S | 0.07 | 0.04 | 0.10 | 0.10 | 0.11 | 0.15 |

Table 4.7: Median (Z), mean (Y) and corresponding standard deviations (T, S) for metrics of flow variables at towers for the 0 degree case. 95% confidence intervals are calculated from M = 25 (TKE), M = 30 (w) and M = 30 (u) runs, see Figure 4.19 and 4.20.

Looking at the mean and median metrics of the ensemble the following observations can be made. For $u/u_{ref}$ the hit rate is well above 0.66, the quality acceptance criteria defined by the VDI guideline (VDI, 2005). Contrary to this the ensemble hit rates for $w/u_{ref}$ and $TKE/u_{ref}^2$ are much smaller. These quantities do however have a much smaller magnitude as $u/u_{ref}$ with many positions being close to zero, so that the value prescribed for the allowed absolute deviation is more important for these quantities. As stated in Chapter 4.1.2.1, the present exercise used a much smaller allowed absolute deviation than the VDI guideline. Therefore the acceptance criterion of 0.66 cannot be transferred to the present case. Instead the ensemble metrics should be used to define the “state-of-the-art”, taking also the “best” results from the presented simulations into account.

#### 4.2.5.4 Metrics for the –45 degree flow case

As stated in Chapter 4.2.2.2, metrics for the velocity components and the turbulent kinetic energy are available for the –45 degree flow case from the Excel ‘Saved-metrics’ workbook for the towers, shown in Figure 4.2, and most of the horizontal layers, cf. Figure 4.6. Results are however only shown for the towers, like for the 0 degree case. Detailed metrics for the horizontal layers, including also the Reynolds stress components, can be found e.g. in Franke (2009a, 2009b).

In Figure 4.21 the metrics for the turbulent kinetic energy at the towers are shown. First it can be noted that for this case the hit rate and FAC2 are much higher than for the 0 degree case, although for the –45 degree case the differences of the turbulent kinetic energy at the inlet between most simulations and the experiment also exists. But here the entire approach

flow is affected by the containers, with turbulence production in front of and behind the containers before the measurement towers are reached. The generated turbulence is higher than the approaching turbulence, leading to a better agreement between prediction and observation at the towers. Consequently the fractional biases are reduced, but most simulations still predict too low $TKE/u_{ref}^2$ values (FB > 0).

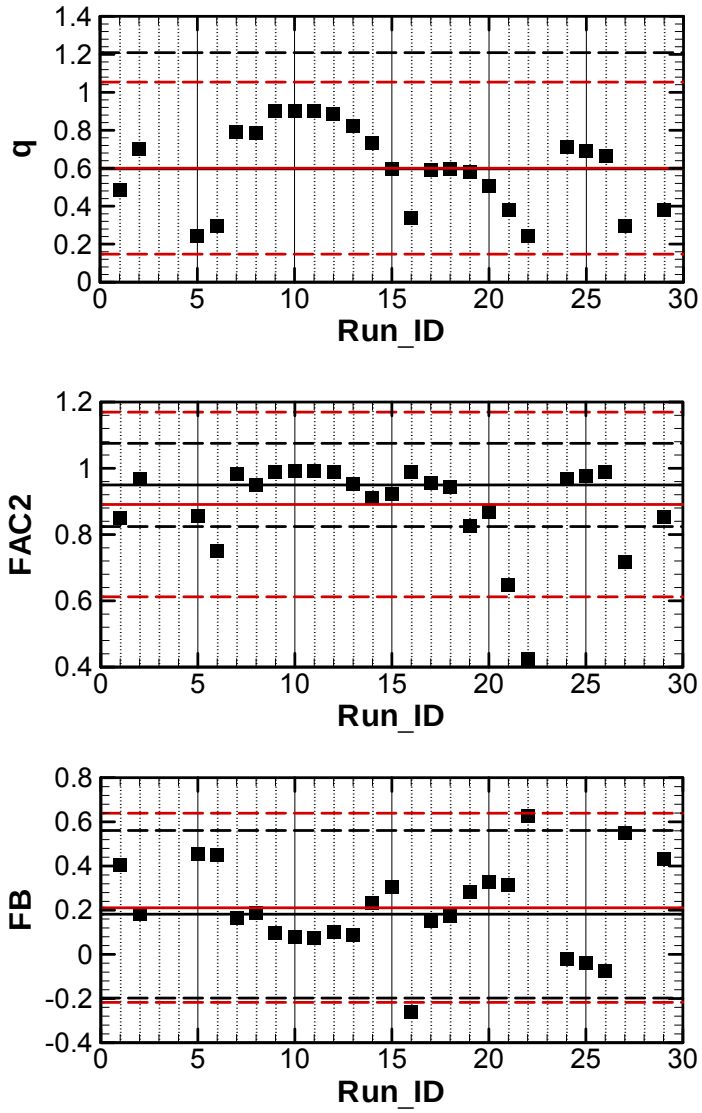



Figure 4.21: Running record of -45 degree turbulent kinetic energy hit rate (q), factor of two (FAC2) and fractional bias (FB), from top to bottom. ■ : individual metrics $X_j$; — : median $Z$; — : mean $Y$; - - - : 95% uncertainty interval $Y \pm P_{T,95}$; - - - : 95% uncertainty interval $Y \pm P_{S,95}$.

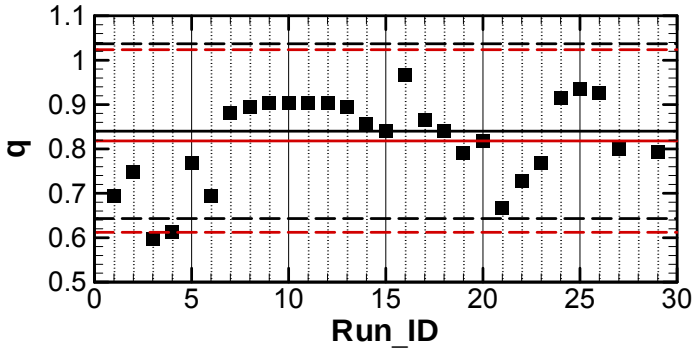
In Table 4.8 the median and mean are shown, together with the ranges defining the 95% uncertainty intervals. The median and mean of the metrics are also much better than for the 0 degree case, i.e. higher values for q and FAC2 and lower values for FB and NMSE. The 95% uncertainty intervals for q are $P_{T,95}$ = 0.61 = 1.02$Z$ and $P_{S,95}$ = 0.45 = 0.75$Y$, for FAC2 $P_{T,95}$ = 0.13 = 0.14$Z$, $P_{S,95}$ = 0.28 = 0.31$Y$, for FB $P_{T,95}$ = 0.38 = 2.11$Z$, $P_{S,95}$ = 0.43 = 2.05$Y$ and for NMSE $P_{T,95}$ = 0.23 = 1.77$Z$, $P_{S,95}$ = 0.34 = 1.79$Y$. Except for FAC2 these intervals are much larger than the corresponding ones for the 0 degree case, indicating the larger scatter in the results. Despite the large confidence intervals outliers can be identified for FAC2, FB and NMSE (not shown here). Although the description presented in Appendix A for each run is relatively complete, it was not possible to come up with clear conclusions why some of the

runs are outside the confidence intervals. E.g. FAC2 for runs Fl45_021 and Fl45_022 show that the agreement between the VADIS results and experiments gets worse when the grid is refined. Similar results are obtained with STAR CD, runs Fl45_024 to Fl45_026, but with FINFLO the opposite is true, see runs Fl45_017 to Fl45_019. Also for FB there is run Fl45_016 as outlier, which predicts too high turbulent kinetic energy (FB < 0) contrary to most of the other simulations. An explanation for this could be that Fl45_016 used the measured turbulent kinetic energy profile as inlet condition, but the STAR CD results Fl45_024 to Fl45_026 also have FB < 0 with a smaller and constant turbulent kinetic energy.

For the velocity components at the towers similar results as in the 0 degree case are obtained for the hit rate. From Figure 4.22 it is visible that $u/u_{ref}$ is again very well predicted, with most of the simulations having hit rates above 0.66. The mean and median are therefore also high, with relatively narrow 95% confidence intervals, see Table 4.8. Like in the 0 degree case the good overall agreement is due to the good prediction of velocity components above roof height, as shown in the exploratory analysis presented in Chapter 4.2.3.1.

Also for $w/u_{ref}$ results similar to the 0 degree case are obtained. The hit rate is again much lower than for u/uref, with mean $Y$ = 0.19 and median $Z$ = 0.18. The low hit rates are due to the inability of the simulations to reproduce the strong downward flow between the containers, as shown in Figure 4.10 in Chapter 4.2.3.1.

a)




b)

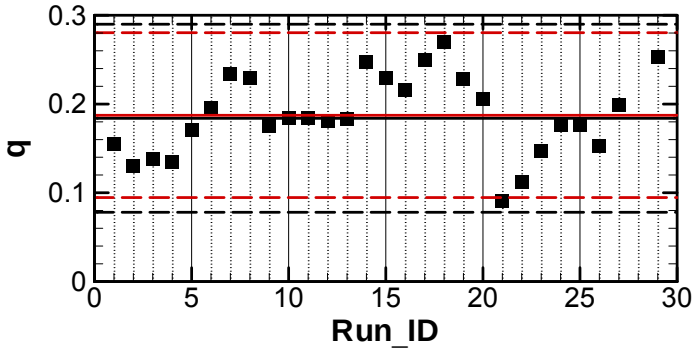



Figure 4.22: Running record of velocity component hit rates for the -45 degree case. a) $u/u_{ref}$, b) $w/u_{ref}$. ■ : individual metrics $X_j$; — : median $Z$; — : mean $Y$; - - - : 95% uncertainty interval $Y \pm P_{T,95}$; - - - : 95% uncertainty interval $Y \pm P_{S,95}$.

For the mean and median as indicators of the “state-of-the-art” the same observations hold as for the 0 degree flow case with the exception that here the hit rate for the turbulent kinetic energy is much higher.

| | U/Uref | W/Uref | TKE/Uref2 | | | |
|---|---|---|---|---|---|---|
| | q | q | q | FAC2 | FB | NMSE |
| Median | 0.84 | 0.18 | 0.60 | 0.95 | 0.18 | 0.13 |
| T | 0.10 | 0.05 | 0.30 | 0.06 | 0.18 | 0.11 |
| Mean | 0.82 | 0.19 | 0.60 | 0.89 | 0.21 | 0.19 |
| S | 0.10 | 0.05 | 0.22 | 0.14 | 0.21 | 0.16 |

Table 4.8: Median (Z), mean (Y) and corresponding standard deviations (T, S) for metrics of flow variables at towers for the -45 degree case. 95% confidence intervals are calculated from M = 25 (TKE), M = 28 (w) and M = 28 (u) runs, see Figure 4.21 and 4.22.

As stated above, results for nearly all horizontal layers are also available for this case. The metrics are similar to the results for the profiles with no good agreement for $v/u_{ref}$, which has a small magnitude like $w/u_{ref}$ at the towers. The hit rate for $u/u_{ref}$ is again rather high, due to the good prediction above the buildings. Selected results for these metrics can also be found in Franke (2009a).

#### 4.2.5.5 Metrics for the –45 degree dispersion case

In Figure 4.23 the results for all metrics evaluated for the concentrations are shown. The individual metrics are presented together with the mean, median and corresponding 95% confidence intervals. The average hit rate q and FAC2 are relatively high (see also Table 4.9), indicating already that most of the simulations did well predict the concentrations. From FB it is visible that all simulations except four predict too high concentrations on the average, as FB < 0.

In Chapter 4.1.2.3 it was explained that good results for FB, i.e. FB close to zero can be obtained although the differences between simulation and experiment are large. The reason for this is that in FB over- and under-predictions can cancel each other. Therefore it was recommended to split FB into false positive ($FB_{fp}$) and false negative ($FB_{fn}$) contributions and present the result as shown in Figure 4.24 (Chang and Hanna, 2005), where $FB_{fp}$ is plotted as function of $FB_{fn}$. All results on the solid diagonal will have FB = 0, as over- and under-predictions will cancel. Only results close to origin at (0,0) will have a low FB due to the right reasons. For the present simulation results it is seen that most of the them over-predict the concentrations, with substantially larger $FB_{fp}$ than $FB_{fn}$. Lying above the diagonal FB = 0 these results have FB < 0. Correspondingly, the four simulation results below the diagonal have FB > 0. Values lying between the dashed diagonals have |FB| < 0.3, the limit value introduced in Chapter 4.1.3. Four simulations fulfil this criterion for the present test case.

Another way of checking FB is to look at NMSE. While over- and under-predictions can cancel in FB, they will lead to a high NMSE. E.g. run Disp_006 has FB = -0.33 but NMSE = 10.34. An efficient way to combine these two metrics is to plot NMSE as function of FB as shown in Figure 4.25a (Chang and Hanna, 2005). The small rectangle is defined by $-0.3 \leq FB \leq 0.3$ and NMSE = 4, which are the limit values for these variables, introduced in Chapter 4.1.3. In this Figure it can be easily seen that Disp_006 (▲) has a reasonable FB, but too high NMSE. The four simulations results that have |FB| < 0.3 do also have NMSE < 4 and therefore fulfil both criteria.

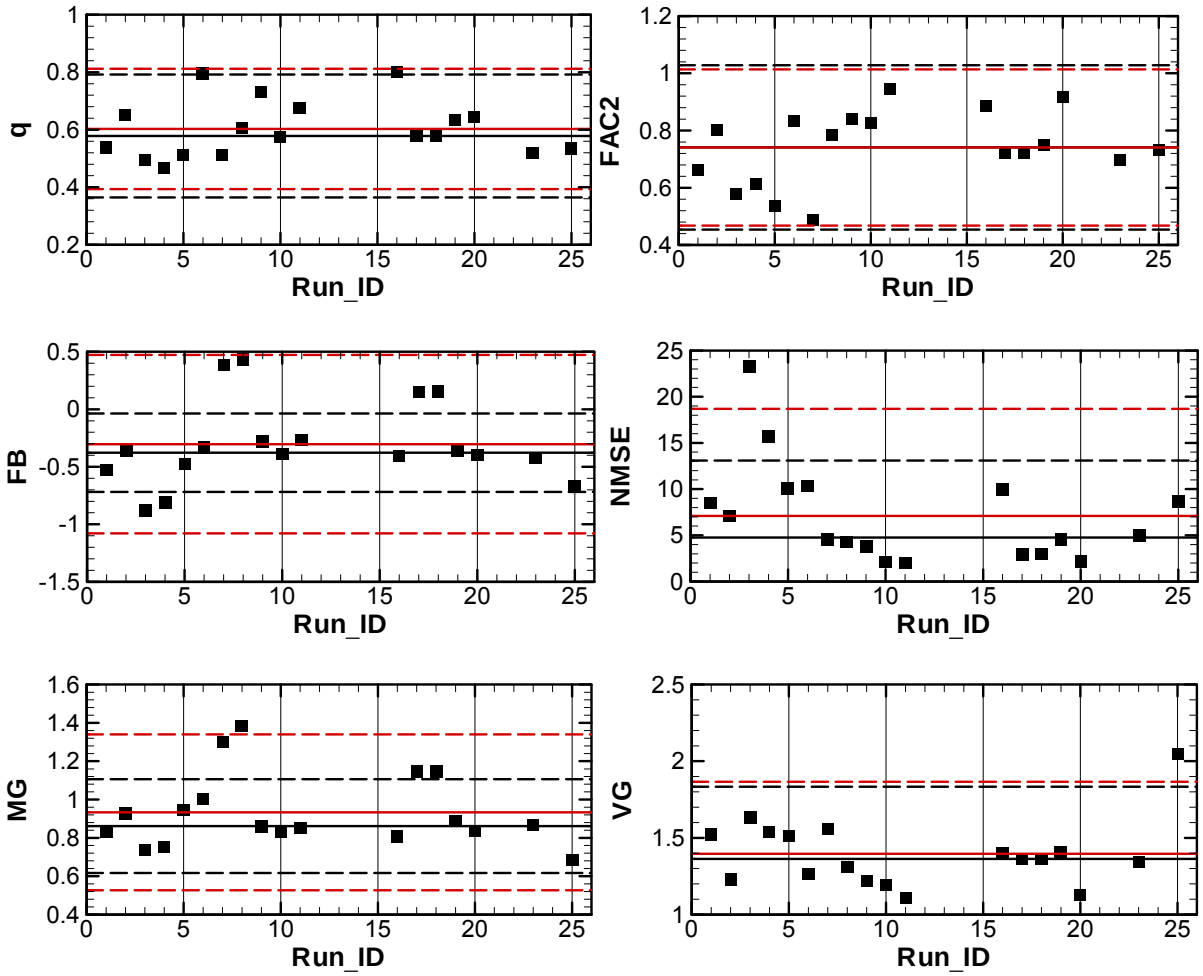



Figure 4.23: Running record of dispersion metrics for the -45 degree case. ■ : individual metrics $X_j$; — : median $Z$; — : mean $Y$; - - - : 95% uncertainty interval $Y \pm P_{T,95}$; - - - : 95% uncertainty interval $Y \pm P_{S,95}$.

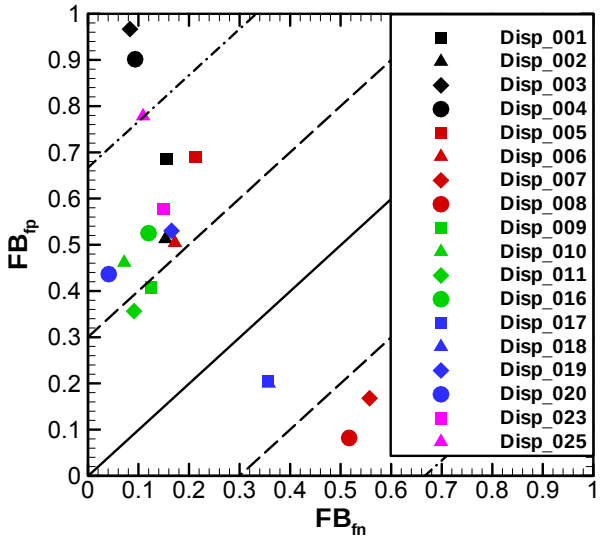
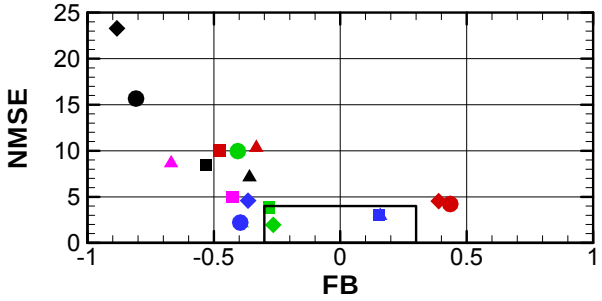



Figure 4.24: False positive ($FB_{fp}$) contributions to FB as function of false negative ($FB_{fn}$) contributions. Solid diagonal shows FB = 0, dashed diagonals |FB| = 0.3 and dash dotted diagonals |FB| = 0.67.

The discussion of FB and NMSE can be directly transferred to the geometric metrics MG and VG. They should also be presented and analysed together, as over- and under-predictions can cancel in MG, but not in VG. In Figure 4.25b VG is therefore presented as function of MG. Note that splitting MG in false positive and false negative contributions is also possible, as shown in Chapter 4.1.2.4. After the splitting, plots similar to Figure 4.24 can be generated, see Chang and Hanna (2005).

a)




b)

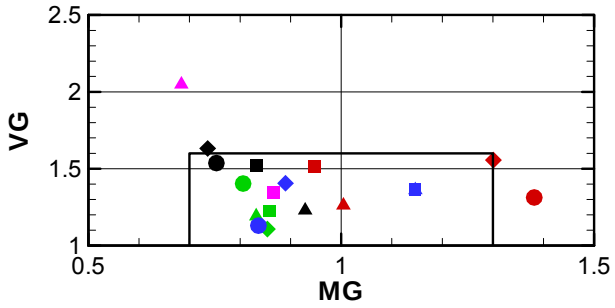



Figure 4.25: a) NMSE as function of FB. b) VG as function of MG. Symbols as in Figure 4.24.

Figure 4.25b also contains a rectangle defined by $0.7 \leq MG \leq 1.3$ and $VG = 1.6$, which are the limit values for these variables, introduced in Chapter 4.1.3. All simulation results except four are within the rectangle and therefore meet the criteria for both variables. This good agreement is due to the fact that MG and VG are not overly influenced by infrequently occurring high concentrations, like FB and NMSE. The latter suffer from the fact that most simulations predict too high concentrations close too the source. MG and VG reflect the good agreement for the small concentrations outside the plume, as shown in Figure 4.7 in Chapter 4.2.2.5 and in Appendix B.

The running records of MG and VG are shown in Figure 4.23. Most of the simulations under-predict the concentrations, having $MG < 1$. Therefore also the mean and median MG are smaller than one, see also Table 4.9. Based on the median, all four simulations that over-predict are outside the 95% confidence interval. The same is true for the FB of these simulations. Based on the mean there is no outlier for FB and one outlier for MG. This demonstrates the better suitability of the median to identify outliers. For the mean the outliers mainly result in larger standard deviations and therefore confidence intervals.

From the ensemble metrics listed in Table 4.9 FAC2, MG and VG are within the limits presented in Chapter 4.1.3 to define the “state-of-the-art”. FB and NMSE are more or less slightly above the limits. For the hit rate q the mean and median are below the limit of 0.66, used in the VDI guideline (VDI, 2005), but for velocity components. In addition it must be kept in mind that the allowed absolute deviation (see Chapter 4.1.2.1) of the VDI guideline is substantially larger than the experimental uncertainty, which was used in this exercise.

Taking all this into account it can be stated that the limits $q \geq 0.66$, FAC2 > 0.5, |FB| < 0.3, 0.7 < MG < 1.3, NMSE < 4 and VG < 1.6 could be used as quality acceptance criteria, where the limit for FAC2 might even be increased to 0.7 or 0.75 as this metric is less strict than the hit rate. This is substantiated by the fact that two of the 18 simulations do meet all these criteria.

| | q | FAC2 | FB | NMSE | MG | VG |
|---|---|---|---|---|---|---|
| Median | 0.58 | 0.74 | -0.38 | 4.76 | 0.86 | 1.36 |
| T | 0.10 | 0.14 | 0.16 | 3.95 | 0.12 | 0.22 |
| Mean | 0.60 | 0.74 | -0.30 | 7.10 | 0.93 | 1.40 |
| S | 0.10 | 0.13 | 0.37 | 5.49 | 0.19 | 0.22 |

Table 4.9: Median (Z), mean (Y) and corresponding standard deviations (T, S) for concentrations for the -45 degree case. 95% confidence intervals are calculated from M = 18, see Figure 4.23.

## 4.3 Non-CFD models validated with MUST wind tunnel data

### 4.3.1 Exploratory data analysis

#### 4.3.1.1 *Flow, -45 degree and 0 degree case*

Most of the non-CFD models participating in the MUST exercise have a very simple parameterisation of the flow. However, one particular model - the particle dispersion model LASAT - has been applied with a diagnostic wind field model. For this model it makes sense to evaluate flow results in the same way as for the CFD flow simulations. The diagnostic wind field results for the three wind components calculated with the wind field model are therefore included in the files for 0 degree and -45 degree flow that were discussed in sections 4.2.2 and 4.2.3.

Figure 4.10 shows the behaviour of w for a variety of CFD models for the -45 degree case. Results for the LASAT diagnostic wind field model have been included in that figure.

#### 4.3.1.2 *Dispersion, -45 degree case*

As explained previously in section 4.1.1.1, point-to-point comparisons considering pairs of modelled and measured concentration in space do not make sense for those non-CFD models which assume a straight centreline of the plume. As the main purpose of these non-CFD models is to predict the magnitude of concentration in terms of maximum and percentiles, the validation of these models can be well based on a comparison of modelled and measured concentrations unpaired in space, as will be demonstrated later in this section. Further, it is possible to explore model performance by considering main features of the plume, as also explained.

A comparison of concentrations between wind tunnel and models in the form of *scatter plots* calculated by the Excel workbook (Section 4.2.2) is nevertheless given for information in Figure 4.26 for all the non-CFD model runs. The interpretation of the paired comparison is well justified for the results of the particle diffusion model LASAT which is based on a three-dimensional diagnostic wind field simulation. Model behaviour for the other models is also commented, although this type of comparison should not be over-interpreted.

Run *ADMS a* is the one where an equivalent building is included, whereas all the other ADMS model runs are without buildings and just differ in the roughness length chosen. The inclusion of a building shows the most severe over-estimation of all the model runs with

ADMS due to the high concentrations in the cavity (not shown). Apparently the best results are achieved with run *ADMS c* with a roughness length of 0.269 m within the buildings array. From the visual inspection of the scatter plots, the overall results for the simple ADMS model are certainly not worse than those for LASAT (model runs *LASAT a* and *LASAT b*) and show a comparable variability from run to run.

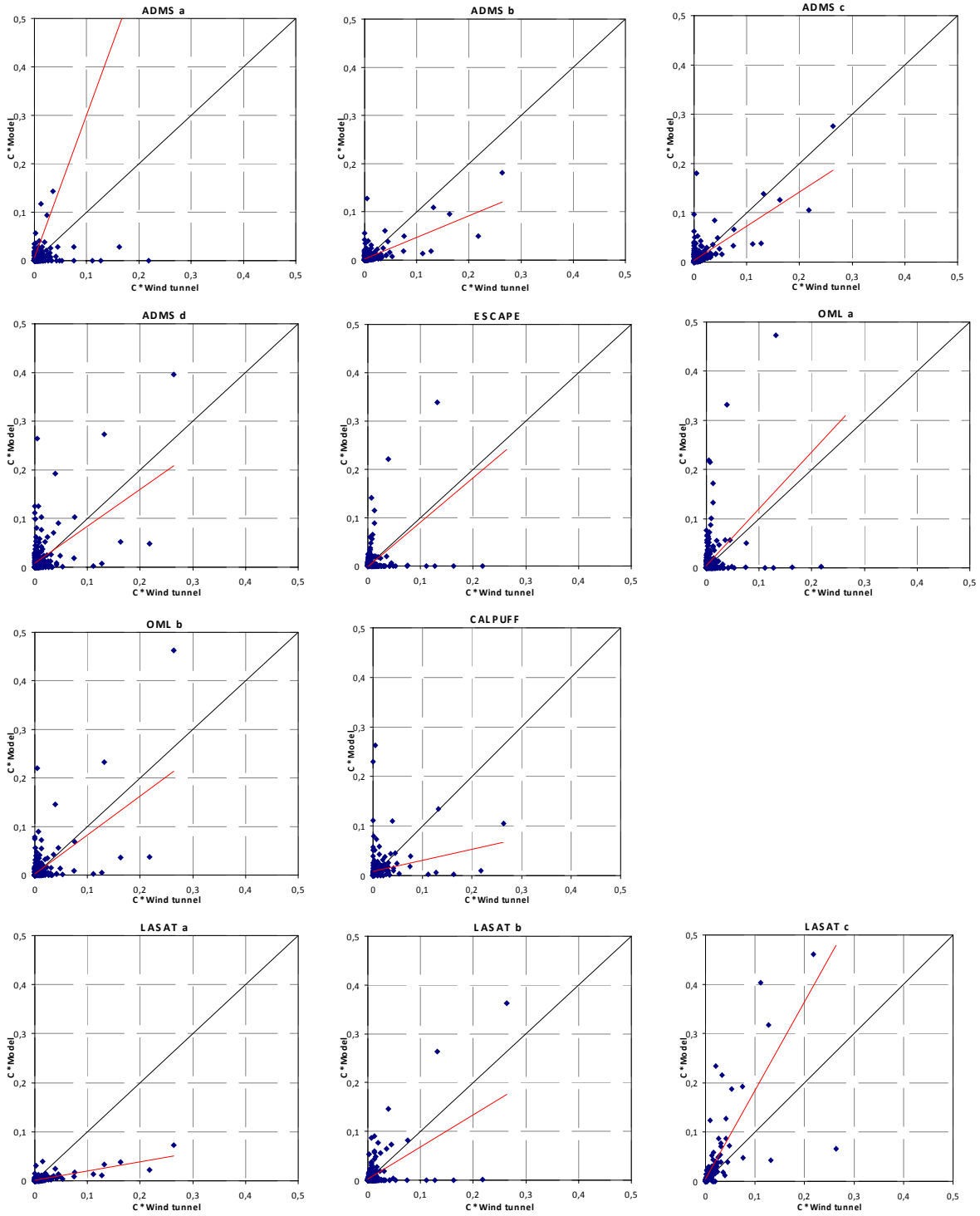



Figure 4.26: Scatter plots: non-CFD model results against wind tunnel measurements for the -45° MUST case.

The two OML runs differ in the values for u* and $z_0$. Model run *OML b* is considered to represent the more realistic roughness length, and it shows slightly better results than run *OML a*. The results are similar to *LASAT b* or *ESCAPE*. For *ESCAPE* (as well as for *LASAT b* or *ADMS d*), high concentrations in the wind tunnel occur with low modelled concentrations and vice versa, so the overall correlation is very poor. The *CALPUFF* run using wind and turbulence profiles at the source position and boundary layer parameters calculated by CALMET shows results comparable to *ADMS c*.

Run *LASAT a* with a grid oriented in North-East directions gives a severe underestimation of concentrations, well in contrast to runs *LASAT b* and *LASAT c*, where the grid is parallel to the building walls. For run *LASAT c*, the wind field is calculated with MISKAM. This run clearly shows the best performance of all the LASAT runs. LASAT results were also included among the CFD results shown in Section **Fehler! Verweisquelle konnte nicht gefunden werden.**.

A comparison of modelled and measured concentrations **unpaired in space** in the form of *quantile-quantile plots* is given in Figure 4.27 for the non-CFD model runs. These plots are derived by ranking both modelled and measured concentrations from the highest to the lowest value and plotting these "ranked" concentration pairs. In this way, only the range of values can be compared, without considering their co-incidence in time and space. As stated before, this approach is adequate to judge the quality of a non-CFD model, because the applicant of such a model has to be sure that especially the highest concentrations or high percentiles are calculated reliably, thus enabling him to state if air pollution thresholds will be violated or not. It is less important where and exactly when such violations will occur. To be on the safe side, a model with a tendency for over-prediction is more welcome than one with a systematic under-prediction.

Run *ADMS a* is the one where an equivalent building is included, whereas all the other *ADMS* runs are without buildings and just differ in the roughness length chosen. Run *ADMS a* shows the strongest over-estimation of concentrations of all the ADMS runs which is due to the high concentrations in the cavity of the "effective" building (not shown); thus the explicit inclusion of a building shows the worst results. Apparently the best results are again achieved with run *ADMS c* with a roughness length of 0,269 m. Run *ADMS b* slightly underestimates, runs *ADMS d* and *ESCAPE* show a tendency to slightly over-estimate the concentrations.

The two *OML* runs differ in the values for u* and $z_0$. *Run b* is intended to be the more realistic, while *run a* has been included to illustrate the effect of roughness parameterisation (see Section 3.1.2). *Run b* shows a very good agreement with the wind tunnel concentrations whereas *run a* over-estimates considerably. The *CALPUFF* run, using wind and turbulence profiles at the source position and boundary layer parameters calculated with CALMET, shows the best results compared to all the other models in this quantile-quantile comparison.

Run *LASAT a* with a grid oriented in North-East directions gives a severe underestimation of concentrations, well in contrast to runs *LASAT b* and *LASAT c*, where the grid is parallel to the building walls. For *run c*, the wind field is calculated with MISKAM. This run shows an over-estimation compared to the wind tunnel data. *Run b* here clearly shows the best performance of all the *LASAT* runs.

Figure 4.29 shows the residuals of the observed and predicted concentration fields in analogy to Figure 4.7. Dark red points show areas of over-estimation, dark blue points areas of under-estimation of the wind tunnel concentrations by the models. Areas in light green show the best agreement. There is a remarkable increase in the quality of the LASAT results from run *LASAT a* to *LASAT c*; the grid orientation along the building walls and even more the use of MISKAM wind fields improve the results considerably. The diagnostic wind field model of LASAT is apparently not able to reproduce correctly the deflection of the winds and thus the concentrations into the container array as revealed by the large array of blue dots for *LASAT a* and *LASAT b*; this deficiency, already partly compensated for by the grid orien-

tation (compare *LASAT a* and *LASAT b*) is clearly overcome by the use of the MISKAM wind field.

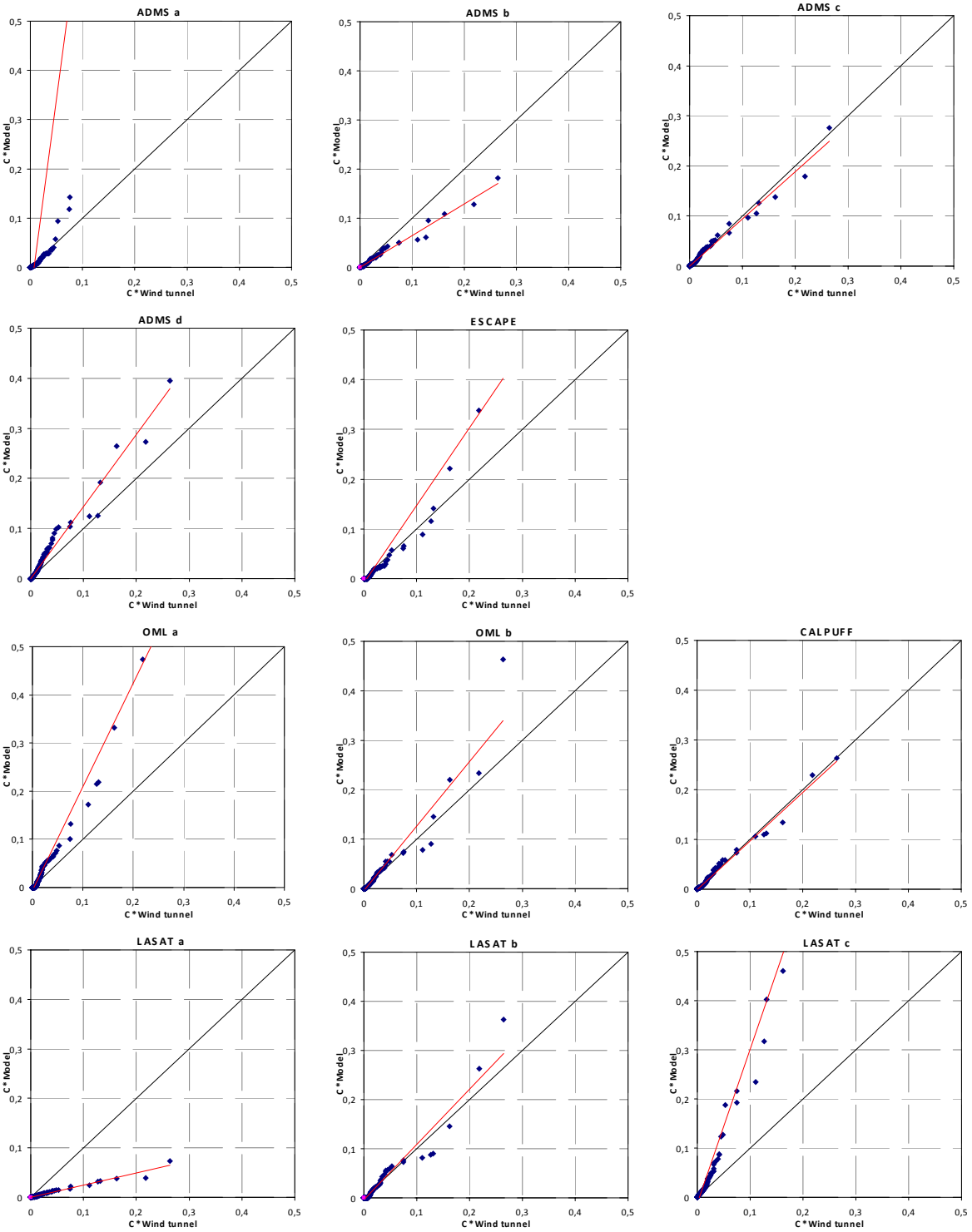



Figure 4.27: Quantile-quantile plots: non-CFD model result against wind tunnel measurements for the -45° MUST case.

As stated earlier, plume propagation in Gaussian models for a selected time interval is along a straight line; this is well reflected in Figure 4.28, where the Gaussian runs show a strong over-estimation of concentrations along the x-axis and a strong under-estimation south-east of the source where the plume really is. The model runs *ADMS c* and *ADMS d* show slightly more points with smaller deviations from the wind tunnel concentrations than

the other Gaussian model runs. The *CALPUFF* run shows comparatively small areas with large over- or under-estimation and is thus superior to the other Gaussian runs.

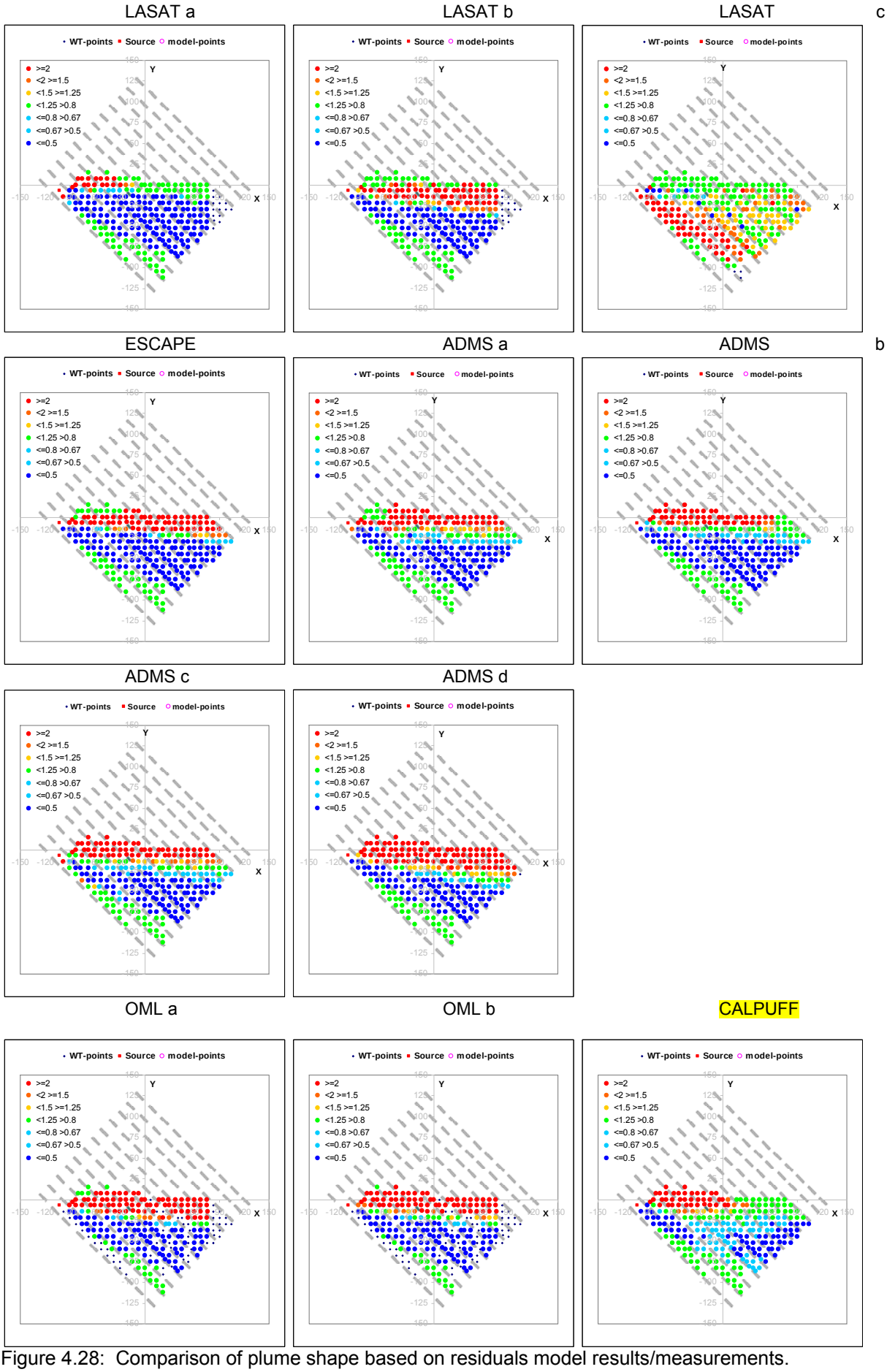



Figure 4.28: Comparison of plume shape based on residuals model results/measurements.

An alternative way for non-CFD models is a *residual plot* as suggested by Chang and Hanna (2004) in Figure 4.29. In this case, the residuals (modelled value to observation) are derived based on the sorted data-sets (therefore unpaired in space). The lower and upper ends of the vertical bars represent the 2nd and 98th percentile of the residuals. The lower and upper limits of the boxes indicate the 16th and 84th percentiles of the residuals, the horizontal line in between the median. Chang and Hanna (2004) mention as a characteristic of good performing models (based on a variety of studies) that about 50% of model predictions are within a factor of 2 compared to the observations, when they consider data unpaired in space. The dotted lines in Figure 4.29 indicate a factor of two, and it appears from the figure that many model runs fulfil this criterion, namely all *ADMS* runs, *LASAT c*, *OML b* and *CALPUFF*. Model run *LASAT a* (not rotated grid) significantly underestimates the concentrations (most of the residuals below 1). The model runs *LASAT b*, *ESCAPE* and *OML a* show a rather large range of residuals indicating that model results can differ by a factor of 10 from the observations. The best agreement between model and observations is found for the model runs *ADMS c*, *ADMS d*, *LASAT c* (based on MISKAM wind fields, with a tendency for over-estimation) and *CALPUFF*.

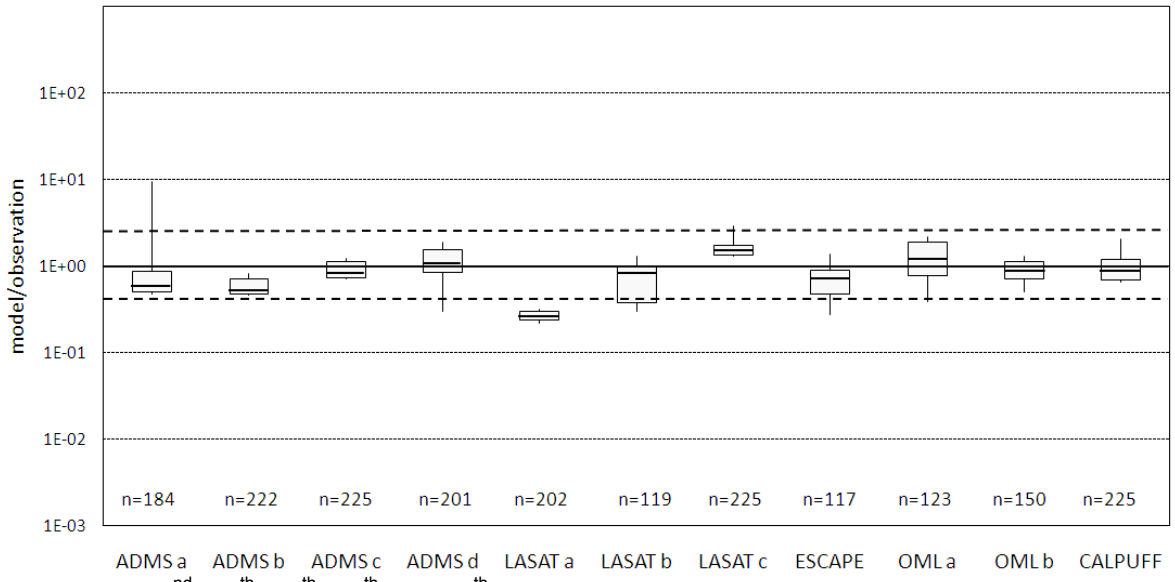



Figure 4.29: 2nd, 16th, 50th, 84th and 98th percentiles of residuals (predicted to observed concentrations unpaired in space, data threshold: 0.003).

Model performance can be explored in a further way, which avoids point-by-point comparisons. Descriptive parameters of the pollution plume, as measured and as modelled, can be deduced by Gaussian fitting, similarly to what was done for CFD models in section **Fehler! Verweisquelle konnte nicht gefunden werden.**.

Figure 4.30 shows plume width, and Figure 4.31 the estimated maximum concentration in the plume according to measurements and selected non-CFD model runs. Figure 4.31 is analogous to Figure 4.14 for the CFD models. Measured values are obtained by Gaussian fitting. For plume width, model runs *ADMS d* and *OML b* are close to observations; *ADMS c* is over-estimating, *OML a* is somewhat under-estimating the plume width. As for the CFD models (Figure 4.14), the estimate of maximum concentrations in Figure 4.31 is partly uncertain close to the release point where the plume is too narrow to be adequately resolved by measurements. At larger distances, *ADMS d* and *OML b* show smaller deviations to the measurements than *ADMS c* and *OML a*. In accordance with plume width in Figure 4.30, *ADMS c* is under-estimating, *OML b* is over-estimating plume centreline concentrations.

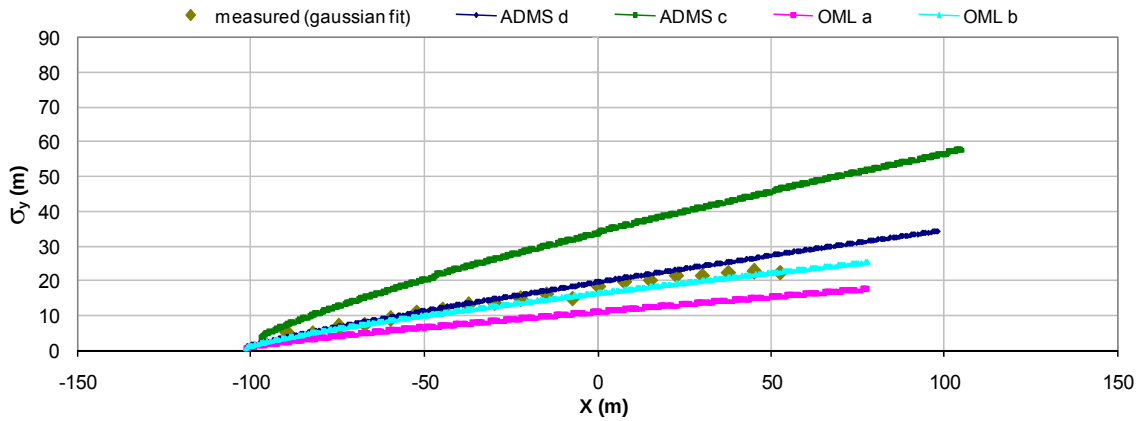



Figure 4.30: Estimated plume width, using Gaussian fitting, with distance x from the source for selected non-CFD models

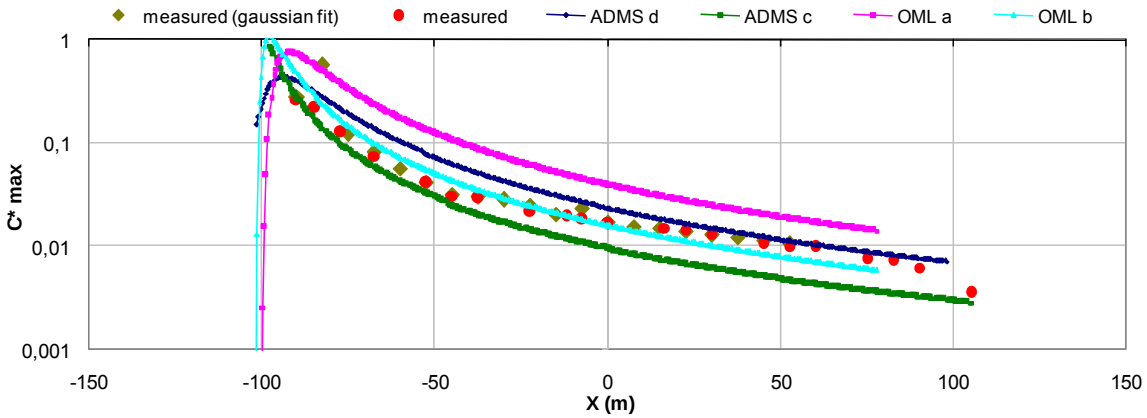



Figure 4.31: Estimated maximum concentration at plume centreline, using Gaussian fitting, with distance x from the source for selected non-CFD models

## 4.3.2 Model performance in terms of metrics

### 4.3.2.1 *Dispersion, -45 degree case*

In the following, the metrics which are recommended for the validation of CFD models are also given for the non-CFD runs for information. These values are based on a point-to-point comparison (**paired in space**). Therefore they cannot be used to judge whether a model is fit for purpose where it purports only to predict the size of concentrations, not their exact location - which is the case for most non-CFD models with the exception of LASAT, as stated earlier.

In Table 4.10 the best result for each measure is highlighted in bold. The model runs *LASAT c* and *CALPUFF* show the best performance with respect to the hit rate and FAC2. There is almost no variation of the hit rate and FAC2 between the different ADMS runs; use of the roughness length instead of an “effective building” does not improve the results. *LASAT a* and *LASAT b* are comparable to the *ADMS* runs with respect to hit rate and FAC2, results of *ESCAPE* and *OML* are less favourable.

An ‘acceptable’ geometric mean in the MUST experiment is calculated for ADMS c and d, and for the model runs *OML a*, *OML b* and *CALPUFF*. MG for the model runs *ADMS a* and *LASAT c* is at the upper and lower limit of acceptability, respectively. If FB is compared to the proposed limits, only *ADMS c, LASAT b, ESCAPE, OML b* and *CALPUFF* are accepted.

| | ADMS a | ADMS b | ADMS c | ADMS d | LASAT a | LASAT b | LASAT c | ESCAPE | OML a | OML b | CALPUFF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Hit rate | 0.35 | 0.33 | 0.36 | 0.36 | 0.29 | 0.35 | **0.58** | 0.24 | 0.15 | 0.19 | 0.43 |
| FAC2 | 0.38 | 0.36 | 0.34 | 0.39 | 0.35 | 0.33 | **0.76** | 0.26 | 0.21 | 0.25 | 0.59 |
| MG | 1.3 | 1.4 | 1.1 | 0.9 | 2.0 | 1.4 | 0.7 | 1.4 | **1.0** | 1.1 | 1.2 |
| VG | 5.2 | 3.0 | 2.7 | 6.1 | 3.0 | 4.8 | **1.6** | 6.6 | 16.1 | 6.4 | 2.9 |
| FB | -1.1 | 0.5 | **0.03** | -0.3 | 1.2 | 0.2 | -0.7 | 0.2 | -0.5 | **-0.03** | 0.1 |
| NMSE | 100.8 | 4.3 | **2.6** | 5.7 | 14.3 | 6.0 | 17.5 | 10.7 | 10.6 | 5.6 | 8.0 |

Table 4.10: Hit rate, factor of 2 (FAC2), geometric mean (MG), geometric variance (VG), fractional bias (FB) and normalized mean square error (NMSE).

The FB of ADMS d is at the lower limit. The geometric variance is larger than the proposed limit of 1.6 for all model runs of the MUST experiment depicted in Table 4.11. Only *LASAT c* (LASAT based on MISKAM wind fields) achieves a VG of 1.6. Nevertheless, the NMSE for this model run is relatively large (17.5). *ADMS c* is in this case the only model run with an 'acceptable' NMSE. The by far largest NMSE is obtained for *ADMS a*, reflecting the very large scatter in the data for this model run.

Metrics which are based on a comparison of model data and measurements **unpaired in space** are the bias SB and the difference between the standard deviations of the simulations and measurements SDSD.

$$SB = \left(\overline{C_p} - \overline{C_o}\right)^2$$

$$SDSD = \left( \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_p - \overline{C_p}\right)^2} - \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_o - \overline{C_o}\right)^2} \right)^2$$

The squared bias SB describes the difference of the simulated and observed mean values. A larger SDSD indicates that the model fails to simulate the magnitude of fluctuations of the observational data.

| | ADMS a | ADMS b | ADMS c | ADMS d | LASAT a | LASAT b | LASAT c | ESCAPE | OML a | OML b | CALPUFF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SB | 0,0007 | 0,0000 | 0,0000 | 0,0000 | 0,0001 | 0,0000 | 0,0001 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| SDSD | 0,0382 | 0,0001 | 0,0000 | 0,0002 | 0,0004 | 0,0000 | 0,0008 | 0,0003 | 0,0011 | 0,0001 | 0,0000 |

Table 4.11. Squared bias SB and difference of simulated and observed standard deviation SDSD.

For the four ADMS runs it is obvious that the performance of the model is strongly dependent on the chosen input configuration and the selection of input parameters as e.g. roughness length: *ADMS a* (with parameterization of the flow around buildings, roughness length = 0.1 m) gives the highest SB and SDSD due to the high concentrations in the cavity of the "effective" building. For runs *ADMS b* (without buildings, roughness length = 0.1m) and *ADMS c* (no buildings, roughness length = 0.269 m), the results are obviously much improved. Run ADMS d with a higher roughness length and a prescribed MOL of 30 m gives a somewhat higher SDSD.

LASAT is applied to this case with the grid oriented according to the north and east direction (*LASAT a*), in a second run with rotated grid parallel to the side wall of the buildings (*LASAT b*) and based on three-dimensional wind fields calculated with the CFD model MISKAM (*LASAT c*). The difference in standard deviations (SDSD) and the difference between the mean values (SB) are decreased significantly from *LASAT a* to *LASAT b*. *LASAT c* achieves a better agreement concerning plume shape, but the difference in standard deviations is much larger for this run.

Comparing model run *OML a* with a very small roughness length of 0.037m (B) and *OML b* a roughness length of 0.3 m (C) it turns out that the latter is obviously the better choice according to the SDSD.

Both metrics indicate a good agreement in terms of bias and difference of standard deviation for *CALPUFF*.

## 4.4 Model Validation with MUST field data

The “perfect model” is in principle the one that can deal with all the complexity and variability of the atmospheric motions with an acceptable accuracy.

As for variability, it means that it should be able to provide prognostic fields to forecast unknown or future atmospheric conditions. That is to produce unsteady-type simulations, accounting for time-dependent boundary conditions of the flow. To properly describe the complexity of meteorology, when resolution does not allow explicitly computing the thermo- and dynamical fields, the most advanced and physically based parameterisations should be used.

As for other aspects of complexity, the model should be advanced enough to be able to describe the full range of inhomogeneities encountered in complex terrain: topography, land-use variability, complex canopies and intricated urban areas.

The perfect model should be based on DNS approach, where dynamical equations are solved for all variables on all scales, enabling the ‘real world’ to enter in all its details.

As anticipated in the introduction, this approach cannot be widely practiced yet due to the enormous computational power required. At present, modellers’ community is following two main approaches: CFD and (micro-) atmospheric models, both based on the RANS approach, and LES models. In this COST Action, CFD models and Atmospheric RANS models, here referred to also as Unsteady RANS (URANS), have been applied.

CFD models are able to deal with extremely complex geometries at very high resolutions, but at present, in their vast majority, they provide mainly steady state solutions. Furthermore, users of CFD models often use simplified initial conditions and systematic assimilation of 3D fields from larger-scale meteorological models is not implemented yet.

Micro-scale atmospheric models are derived from meteorological models used at larger resolutions. They are conceived to consider all the atmospheric processes in real topography, can be driven by larger-scale models through data assimilation and provide time-dependent solutions, describing for instance a daily cycle. However, they are not yet able to include complex geometries but only relatively simple obstacles.

The models being only a ‘representation’ of reality, their evaluation and validation are a requisite of a proper scientific approach. Therefore, it is important to perform experiments and make observed data available to test the performance and reliability of the model. Two main kinds of experiments can be conducted: real field experimental campaigns and controlled experiments in wind tunnel or water tank. While field experiments capture the full variability of the atmosphere, controlled experiments permit repeatability and exclusion of external factors influencing the essence of the experiment, related for example to changes in external meteorological conditions.

Field data from wind and dispersion measurements represent – for almost any realistic averaging period – snapshots of temporally and spatially variable meteorological phenomena at all time scales. Hence, a significant amount of scatter is expected to be inherent in any field result, even if the data has been averaged over a significant amount of time. Unfortunately, quantifying the expected variability or the confidence interval of given field results is

not trivial because of the continuously changing meteorological conditions during a field test. Quasi-stationary experimental conditions, as they would be required for a statistical analysis of the scatter in 'short-term' averaged turbulent flow and dispersion measurements cannot be achieved at full scale and a variability analysis based on field data only will always be based on non-stationary mean quantities.

Another way of estimating the uncertainty inherent in field data is to analyze laboratory data from controlled experiments in a boundary layer wind tunnel. Although the wind tunnel data represent a more or less idealized image of the reality, they can provide the data quality needed for a statistical uncertainty analysis. Assuming, that wind tunnel experiments have been carried out under carefully controlled boundary conditions matching the full scale situation as close as possible, at least the minimum expected scatter inherent in field data, caused by the presence of turbulent motions in the atmospheric boundary layer, can be quantified.

The methodology of quantifying the scatter is usually straightforward. For a sufficiently long, statistically representative time series or a large enough representative ensemble of data mapping the entire set of turbulent motions, the 'true' mean value for each of the measured quantities is calculated. The 'true' mean value should be repeatable within the range of accuracy of the measurement for independent experiments, carried out under the same boundary conditions. Whereas this precondition for the subsequent scatter analysis can be achieved in a carefully controlled laboratory experiment, it is impossible to calculate an equivalent 'true/global' mean from field data. From a physical point of view, an (accumulated) 'true' mean value calculated from field time series is characterized by a significant (strictly spoken non-quantifiable) variability itself, because of the presence of large-scale turbulent motions, the synoptic variability of the boundary conditions and the resulting 'under-representation' of these effects in time traces of up to a few hours. In order to reach a representative mean value, the averaging time must be several times larger than the characteristic time scale associated to large-scale turbulent fluctuations.

Once the 'true' representative mean of a time series captured under quasi-stationary boundary conditions is known, the same time trace can be divided in (consecutive, overlapping or non-overlapping) sub-ensembles of different sizes, representing different averaging time periods. For each of the sub-ensembles, the corresponding mean (and, if required, higher order moments) can be calculated. If the sub-ensemble mean values are plotted as a function of the averaging time, the increasing 'scatter' in short-term average values for decreasing averaging times becomes visible.

The problem will be discussed based on selected field data taken from trial 11 of the MUST field data as reported in Yee and Biltoft (2004). These measurements were taken on 25 September 2001 (Julian date : 268, time 18:29). In this trial the wind angle relative to the axis of the container array is -41° with a nearly neutral atmospheric stratification. The trial is selected as close as possible to corresponding wind tunnel data used in the first COST 732 validation exercise.

### 4.4.1 Using field data measurements for model validation

The ultimate goal of transport and dispersion model is to reproduce the conditions actually occurring in the real atmosphere, for any site and atmospheric conditions. From this perspective, using field measurements for the purpose of model validation would be the correct and conclusive approach, as long as the data cover all temporal and spatial variability. However, in the process of validation of micrometeorological models it is often necessary to reduce, as much as possible, the influences of varying external conditions in order to correctly evaluate the model performance within the region of interest fort precisely defined boundary conditions. The dynamics of the atmosphere is usually very complex and rarely steady. Steady state conditions can be considered being too improbable to be measured in the real

atmosphere. This bears severe constraints in the use of field measurements for model validation because of two main principle problems:

1) Field experiments are not repeatable, therefore a certain measured quantity can be related, at best, to a certain "physical phenomenon or situation" but certainly includes "sampling uncertainty". Because the uniqueness of the measurement is not possible to distinguish between the two based on repetitive measurements.

2) As stated before, the conditions in real atmosphere are not stationary and any field measurement represents a superposition of its intrinsic variability in time and transient changes in the "external" boundary conditions.

Fortunately, in some field experiments the boundary conditions present lower variability, and therefore may be considered to be better qualified for evaluating micro-scale numerical models. For example, in the MUST field experiment, flow and dispersion measurements were carried out within an idealized "suburban roughness" formed by an array of shipping containers. The field tests were performed in the Utah desert in a flat open terrain with a low surface roughness.

The variability inherent in one of the MUST field experiments is quantified. The results are compared with corresponding data derived from measurements in a boundary layer wind tunnel. Ideally, the combination of results of a variability analysis applied to field data and laboratory data enables the effect of varying external boundary conditions to be estimated.

In the atmosphere observed variations and fluctuations are irregular. In order to deal quantitatively with the corresponding scatter in field test data it is necessary to employ a statistical approach and tools of probability. When comparing sample mean values, a key assumption of statistics is the independence of individual observations comprising each of the sample means. This corresponds to a mutual independence of the variable's individual values.

Often, atmospheric data do not satisfy the criterion of statistical independence, since averages are often time averaged values and the inherent time dependence (*persistence*) conflicts the assumption. Meteorological persistence implies that the variance of a time average is larger for statistically independent samples (*reference?*).

Stationary processes are characterised by the *repeatability* of independent measurements under identical boundary conditions, and the statistical uncertainty of measurements is estimated as the *variance* of the individually measured values. However, since in the atmosphere measurements cannot be repeated under identical boundary conditions, in a statistical sense, no multiple samples may be collected. In this case, the variance does not represent the uncertainty related to the estimate of the 'real' value of the variables from repeated measurement, but it describes the variability of the atmospheric quantity in a given time period. This *fluctuation* is itself a time-dependent variable.

In controlled wind tunnel experiments, instead, quasi-stationary conditions with regard to long-term mean boundary conditions can be generated. Individual series of measurements corresponding to a set of field samples can be captured for non-stationary boundary conditions which are a subset of one and the same mean boundary conditions. Moreover, in a proper simulation of turbulent flows, also the turbulent fluctuation of the variable *in each sample* is accounted for. Thus, in wind tunnel studies both the turbulent fluctuation and the measurement uncertainty can be represented in the variance if sufficiently long, statistically independent samples are recorded.

In order to transform the meteorological variability into a "scatter" similar to what is available from boundary layer wind tunnel tests, one can try to separate the variability of field data in two components, an "*intrinsic*" one, characterising the specific physical process such as turbulent motions, and an "*extrinsic*" one, dominated by the outer boundary conditions and external forcing at larger time and space scales. The wind tunnel modelling can provide

an estimate of the *intrinsic* values of the variability only, since the local turbulence of interest may be singled out and statistically analysed. This estimate of the intrinsic variability is expected to represent a lower limit of the corresponding field values and, as such, could be applied as threshold for field results. On the other hand, comparing the full-scale variability with a corresponding wind tunnel estimate can provide basic quantitative information on effects resulting from the *extrinsic* variability caused by larger scales phenomena present at full-scale.

### 4.4.2 Analysis of field data – the MUST case as an example

A key feature of validation data compiled in COST Action 732 is that combinations of field datasets with corresponding boundary layer wind tunnel measurements are intended to be used for the validation of micrometeorological models. As one of the few test cases available, the MUST case offers both sources of information.

The objective of the MUST (Mock Urban Setting Test) trials, conducted at Dugway Proving Ground on September 6-27 2001, was to acquire near full-scale dispersion and meteorological data for the purpose of urban dispersion model development and validation. The Mock Urban Setting Test is one of the most successful idealized field experiments and it supplies a rich and comprehensive dataset. It has been largely used by the scientific community because of the detailed information on transient tracer dispersion and turbulence phenomena within an idealized suburban roughness.

Within COST 732, a working group re-processed selected field results available in order to prepare a specific data set qualified for validating CFD and non-CFD codes. The additional information is intended to complement the data available from the University of Hamburg boundary layer wind tunnel (Bezpalcova, 2005), which are used mostly by COST732 participants.

#### **4.4.2.1** *Examination of Existing Meteorological and Concentration Data Sets*

There are two main sets of data acquired during the field trials, namely: the dispersion dataset, which was obtained using 74 high-speed photo ionization detectors (48 DPIDs and 26 UVICs) within the obstacle array; and the meteorological dataset (i.e., the wind velocity and sonic temperature), which was obtained using 22 sonic anemometers (14 DPG and 8 DSTL) within or surrounding the obstacle array.

For the exemplary analysis presented here, a small sub-set of data, collected during two days (25 and 26 September 2001) under neutrally stratified atmospheric surface layer (ASL) conditions (according to Monin Obukhov Length) was selected. The results presented were obtained with data from the experiment with the release starting at 18:30 and ending at 18:45 of September 25, 2001 (Efthimiou et al, 2008).

#### **4.4.2.2** *Processing of Velocity and Concentration Time Series - Statistics*

The concentration time series were acquired over sampling periods of 15 minutes for most of the continuous release experiments. Because background meteorological conditions may be changing over the 15-minute sampling period, conditional sampling was applied to the concentration time series in order to extract the 3- to 5- minute period from each record of 15-minute duration with the least variation of mean wind direction and use this 3- to 5-minute period as the standard sampling period for computation of the plume concentration statistics.

In this context, two periods from trials on September 25, 2001 were chosen for detailed analysis. The two periods (100-900 seconds and 300-500 seconds from the start of the re-

lease) were originally chosen by Yee and Biltoft (2004) because of reasonable stationarity in the approach flow conditions (i.e. wind speed and direction). Wind velocities and measured temperatures were also analyzed within a 30 minutes period and a time trace of data averaged over one minute was generated. For trace gas concentrations a similar analysis was applied except that the analyzed time period was reduced to 17 minutes.

In the following a summary of the considered periods and of the calculated variables is reported. For a detailed description of MUST Field Experiment we refer to Biltoft (2001) and Yee and Biltoft (2004).

**PERIODS**

**Velocities – Temperatures**

- 100-900 seconds (18:30:40 – 18:44:00), values averaged over 800 s
- 300-500 seconds (18:34:00 – 18:37:20), values averaged over 200 s
- 30 minutes period (18:15:00 – 18:45:00), time series of data averaged over 1 minute

**Concentrations**

- 100-900 seconds (18:30:40 – 18:44:00) , values averaged over 800 s
- 300-500 seconds (18:34:00 – 18:37:20) , values averaged over 200 s
- 15 minutes period (18:29:00 – 18:46:00), time series of data averaged over 1 minute

**VARIABLES**

**Velocities – Temperatures**

For each data record from each sonic anemometer, we computed the following quantities:

Mean values for the three velocity components
Mean wind velocity direction
Standard deviations of the velocity fluctuations in the x, y, z directions
Turbulence kinetic energy
Mean temperature
Covariances
Temperature flux
Friction velocity
Local free convection velocity scale
Monin-Obukhov length
Sensible heat flux

**Concentrations**

After the conditional sampling of concentration, we computed the following concentration statistics:

1. Mean concentration (ppm).
2. Standard deviation of the concentration fluctuation
3. Concentration fluctuation intensity

#### 4.4.2.3 *Data analysis: Moving and Increasing Windows of Meteorological and Concentration Observations*

To analyze the variability of field measurements within the two selected time periods, 300-500 s (200 s period) and 100-900 s (800 s period), averages were calculated in so called *moving* and *increasing* time-windows.

In *moving windows*, a fixed averaging period of 200 s, as in 300-500 s case, is kept but the start and end times are varied inside the 100-900 s larger period. When using *increasing windows*, the flow and turbulence quantities were calculated starting from the sub-period 300-500 s average and enlarging the averaging interval up to the largest 100-900 s period.

The comparison of the values obtained from the moving window approach allows evaluating both, the representativeness of the sub-interval 300-500 s and the variability of the 200s-samples within the 100-900 s period, together with the actual stationarity of the measured values around the 300-500 s interval. Varying and increasing the averaging windows gives an estimate of the variability of the averaged variables inside the second reference period, 100-900s.

The intervals listed in Table 4.12 were considered for calculating the averages with *moving* and *increasing* windows (in seconds):

| **Moving windows** | | **Increasing windows** | |
|---|---|---|---|
| 300-500 s | 1st reference period | 300 - 500 s | 1st reference period, 200 s average |
| 350-550 s | around the 1st ref. period 300-500 s | 250 - 550 s | 300 s average |
| 400-600 s | " | 200 - 600 s | 400 s average |
| 250-450 s | " | 150 - 650 s | 500 s average |
| 200-400 s | " | 100 - 700 s | 600 s average |
| 100-300 s | tail of the 2nd reference period 100-900 | 100 - 800 s | 700 s average |
| 700-900 s | " | 100 - 900 s | 2nd reference period, 800 s average |

Table 4.12: Intervals which were used for calculating the averages with *moving* and *increasing* windows

Afterwards the data collected at the South Tower, and the Main Tower were analyzed. South Tower data are expected to provide the inflow conditions for the selected "trial 11" and the Main Tower results describe the flow disturbed by the local array of obstacles.

The sequence of 1-minute averages are plotted in Figure 4.32, illustrating the data variability within the reference time intervals.

#### **4.4.2.4** *Moving windows for meteorological data*

In Figure 4.33 (left column) the vertical profiles characterize the incoming flow during the selected trail. A rather large variability of the incoming wind flow and turbulence can be noticed when averaging over 200 s periods for different intervals inside the 2nd reference period (100-900s). The biggest deviation from the other profiles is observed for the tail period 700-900s. The 1st reference period, 300-500s, is located more or less in the mean range with respect to the periods considered for the wind velocity components and direction. Unlike mean wind and direction, for the turbulence quantities the 300-500s average shows the highest values and, considering in particular the tail periods 100-300s and 700-900s the data seems to be dominated by a transient peak of turbulence.

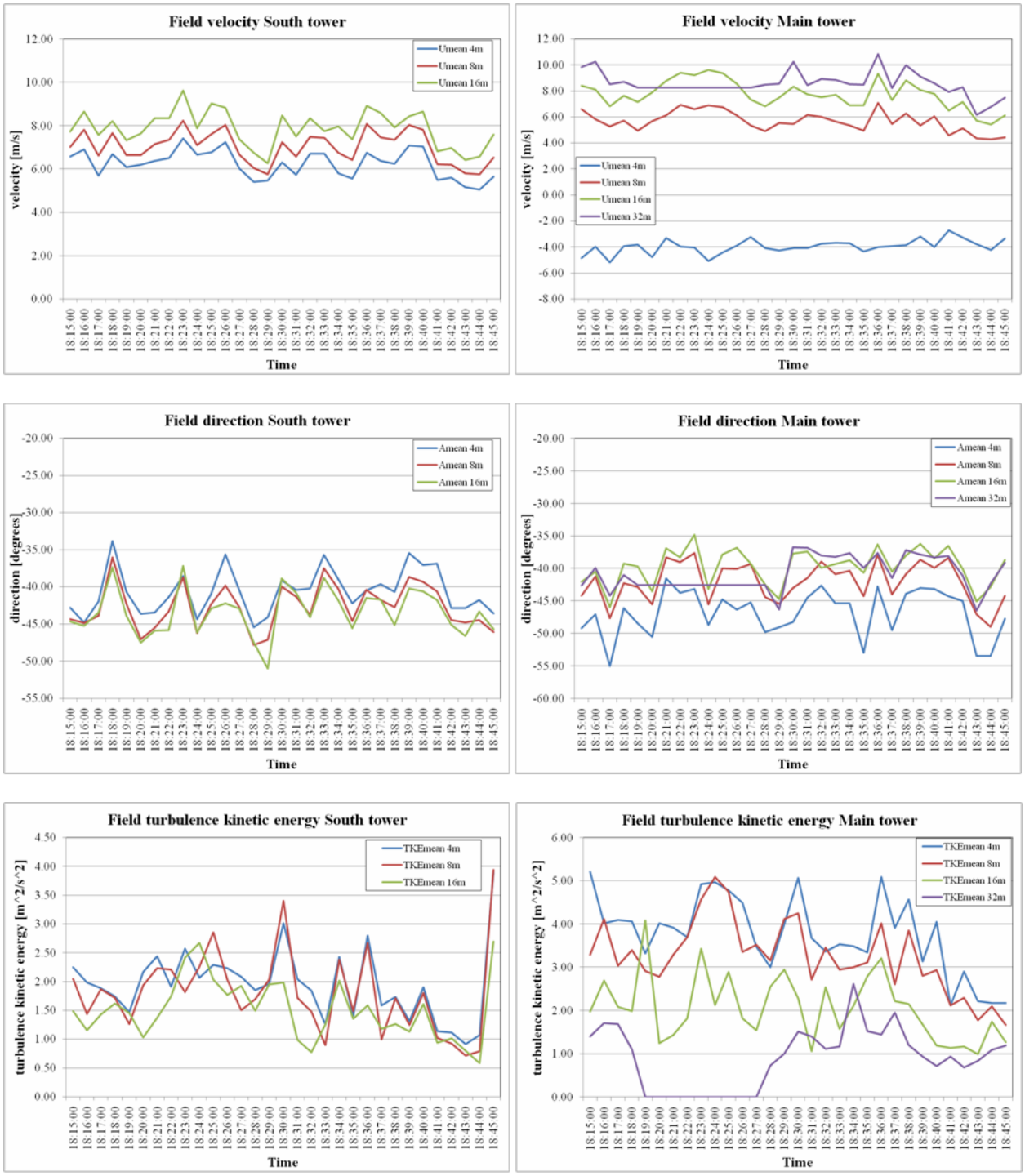


Figure 4.32. Sequence of 60s sub-periods averaged meteorological variables in the full sampling period. South (left) and Main (right) Towers.

The variable profiles at Main Tower (Figure 4.33, right column) document the wind flow above the obstacles as measured in the centre of the container array. All mean flow profiles show nearly the same shape but the magnitude of measured winds consistently differs and fluctuates without a clear trend. Compared to the South Tower data, the sequence of growth in turbulent profiles changes. The transitional turbulence change seems to be shifted or extended in time, resulting in close values for 300-500s and 350-550s averages.

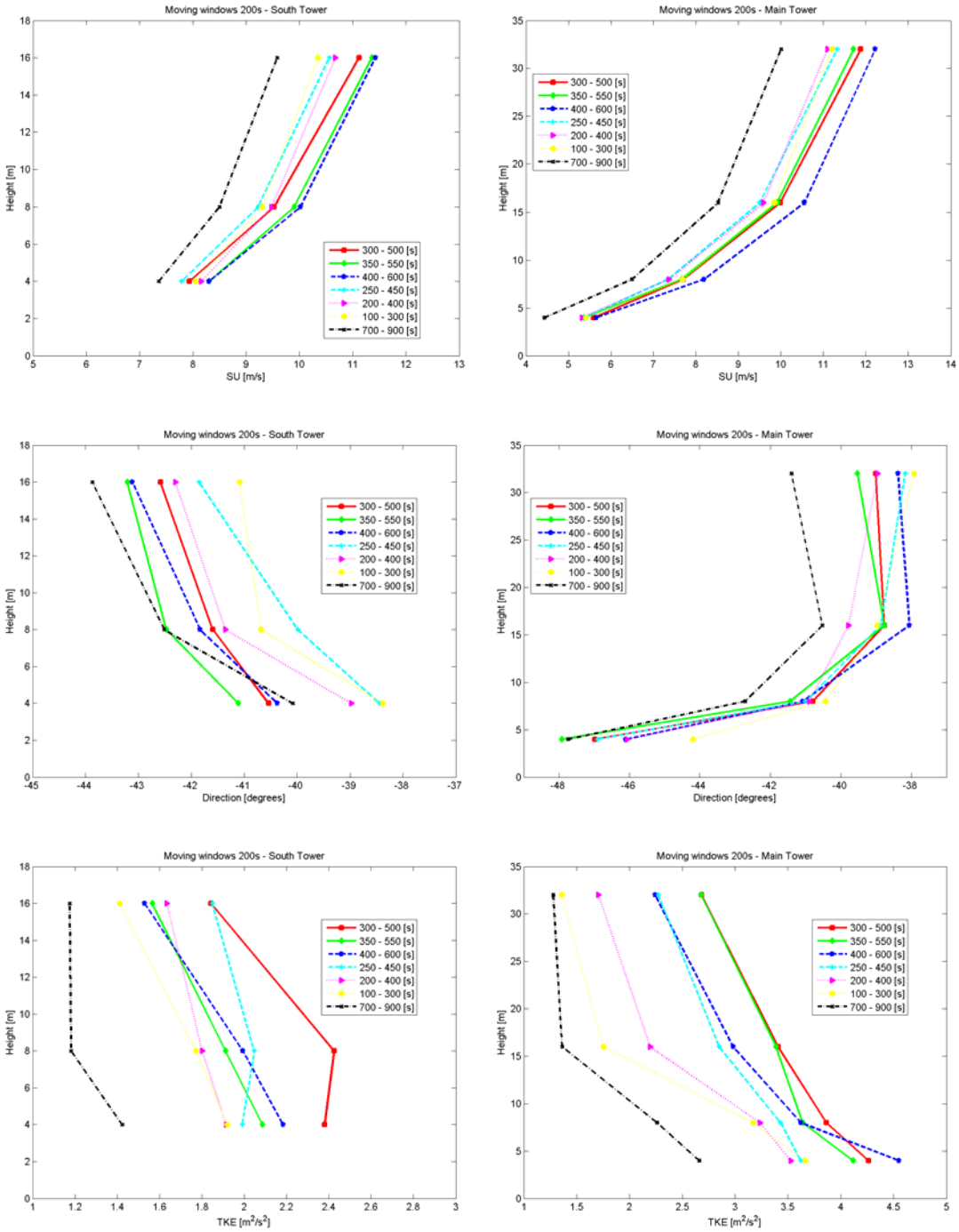



Figure 4.33:. Moving windows for meteorological variables, profiles at South (left) and Main (right) Towers

The transient and/or temporal features are also visible from the sequential plots of the averaging periods. Figures 4.34 shows the corresponding wind velocity and turbulent kinetic energy time traces at the three heights measured at the South Tower.

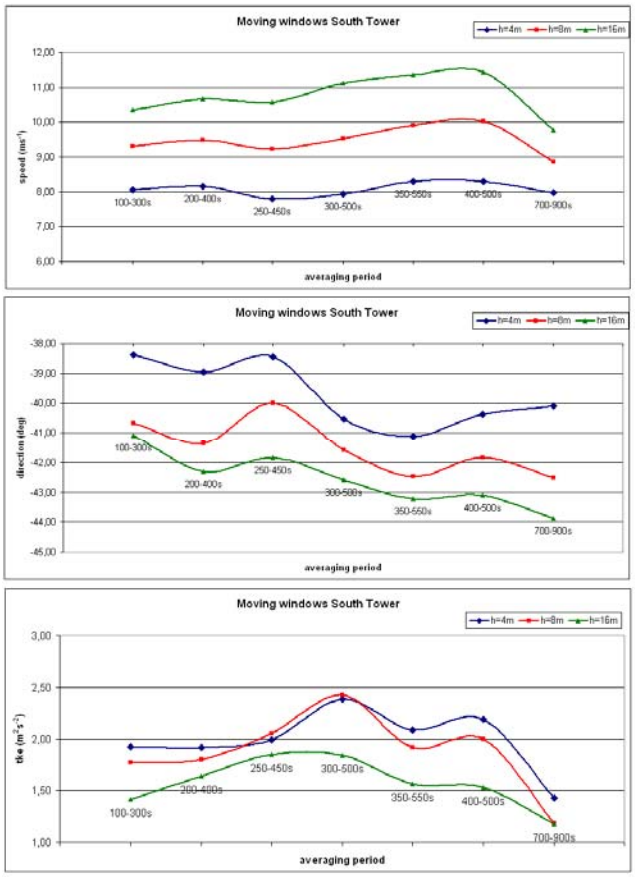



Figure 4.34. Sequence of 200s sub-periods averaged meteorological variables in the full sampling period at three heights.

#### 4.4.2.5 *Increasing windows for meteorological data*

With respect to the moving window, it can be noticed that for an increasing window size the variability is reduced consistently. At the South tower (Figure 4.35, left column) the longitudinal component of wind velocity and the direction of the 1st reference period 300-500s slightly differ from the other profiles. The peculiarity of the 300-500s average for the turbulence variables is again clearly visible, as for the moving windows analysis. For larger averaging periods the turbulence profiles change and fluctuate with respect to the 300-500s case and they tend to group together. Consistent differences between the two reference periods, 300-500s and 100-900s can be identified. Similarly, for the Main Tower (Figure 4.35, right column) the variability of the profiles reduces and the mean flow profiles of the 1st reference period 300-500s are located among the lowest values, while the corresponding turbulence profile shows the largest value. Increasing the averaging period results in a decrease of turbulence, which confirms the peculiarity of the 300-500s period also for the flow modified by the obstacles.

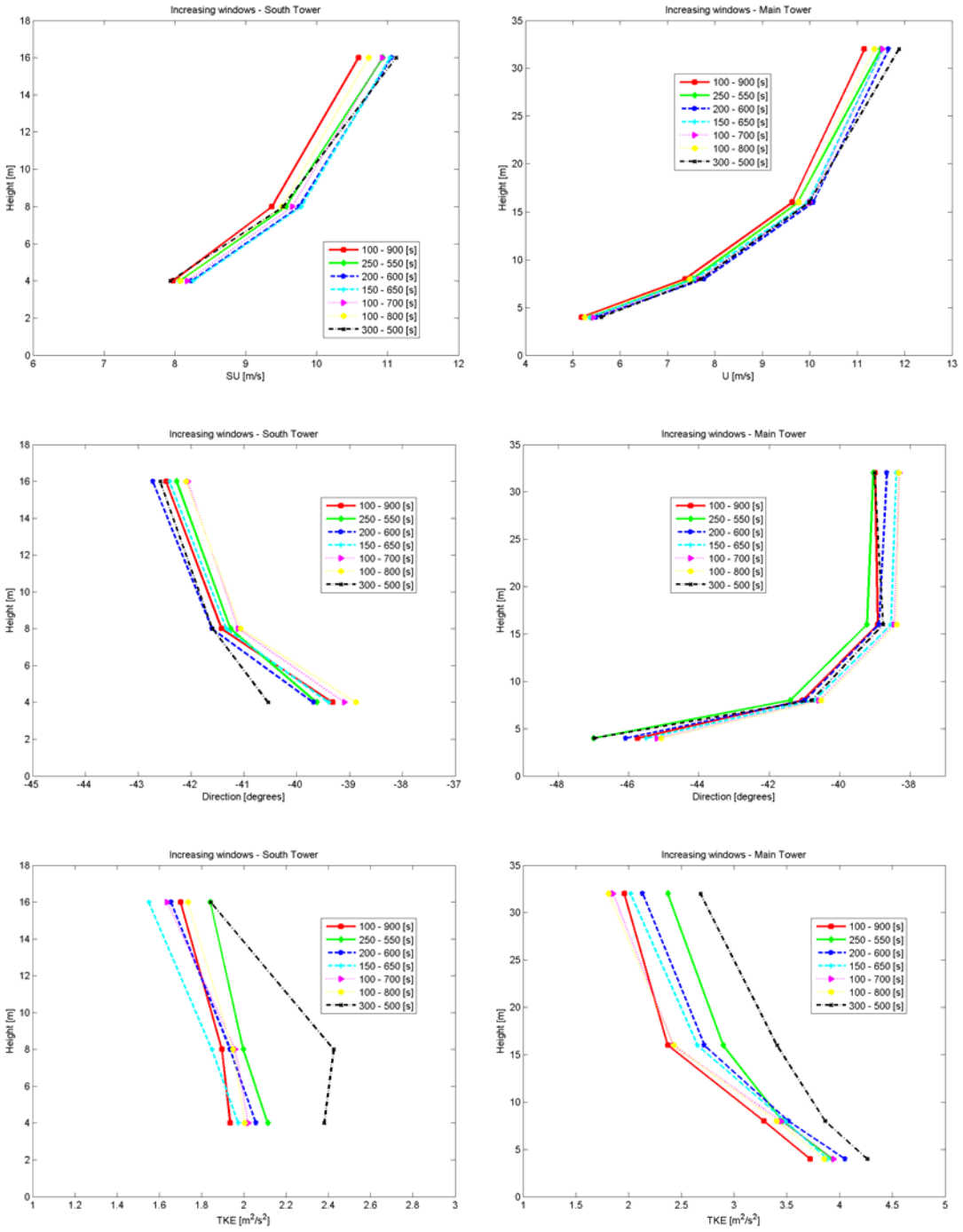



Figure 4.35: Increasing windows for meteorological variables, profiles at South (left) and Main (right) Towers

#### 4.4.2.6 *Moving windows for concentration*

The profiles of concentrations at masts B, C and D and at the Main tower (examples in Figure 4.36) also show a dependence on the averaging time window and a relative variability. The profiles of the 1st reference period 300-500s take more or less mean values also for the set of concentration profiles.

An unexpected large concentration value can be noticed at the highest level of 6 m at mast B.

Considering the concentrations at the 1.6 m samplers along the rows between the containers (Figure 4.37), the larger spread among the moving averaging periods is recorded in the row closest to the source. It is worth also to note that the concentrations measured at the row-samplers took rather different values with respect to the ones recorded at C and D masts.

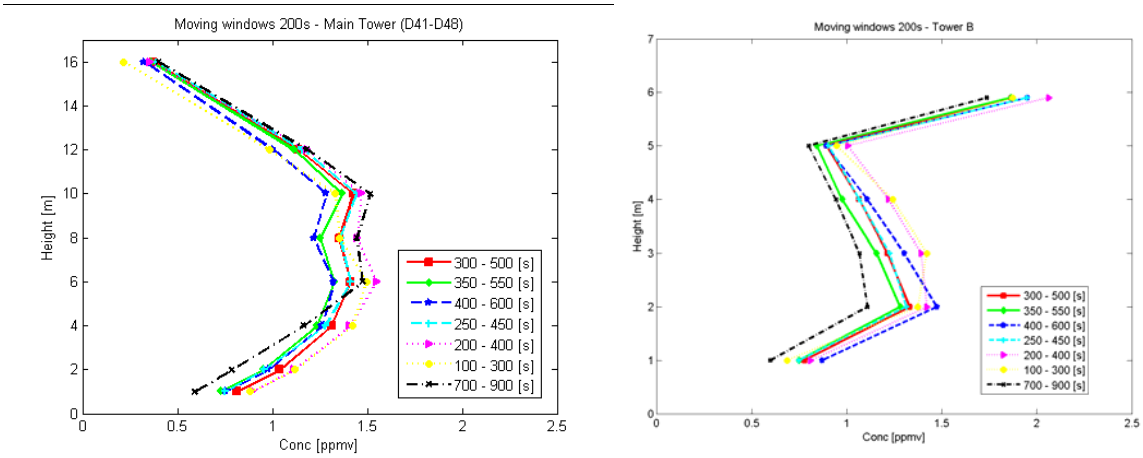



Figure 4.36: Moving windows for concentrations, profiles at Main Tower (left) and B mast (right)

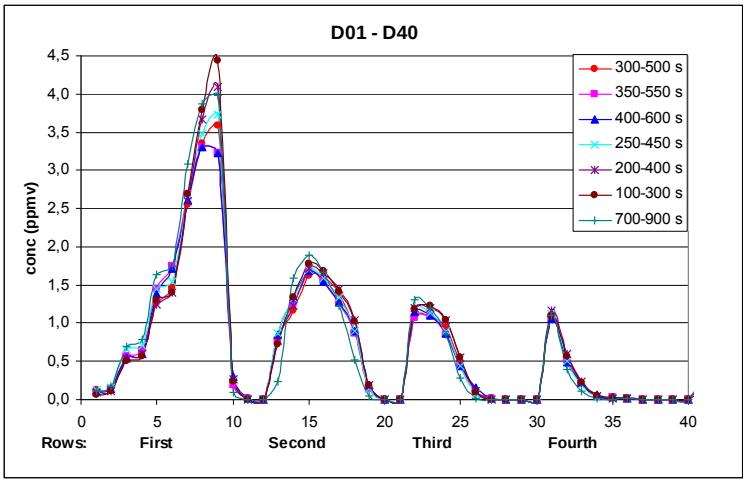



Figure 4.37: Moving windows for concentrations at samplers along rows between obstacles

### 4.4.2.7 *Increasing windows for concentration*

In the case of increasing windows, analogously as for the meteorology, also for the concentration the fluctuations related to the changing averaging periods reduce (Figure 4.38).

This trend is stronger for the ground-level 1.6 m samplers, where, apart from the row closest to the source, the concentration lines group towards very close values, as in Figure 4.39.

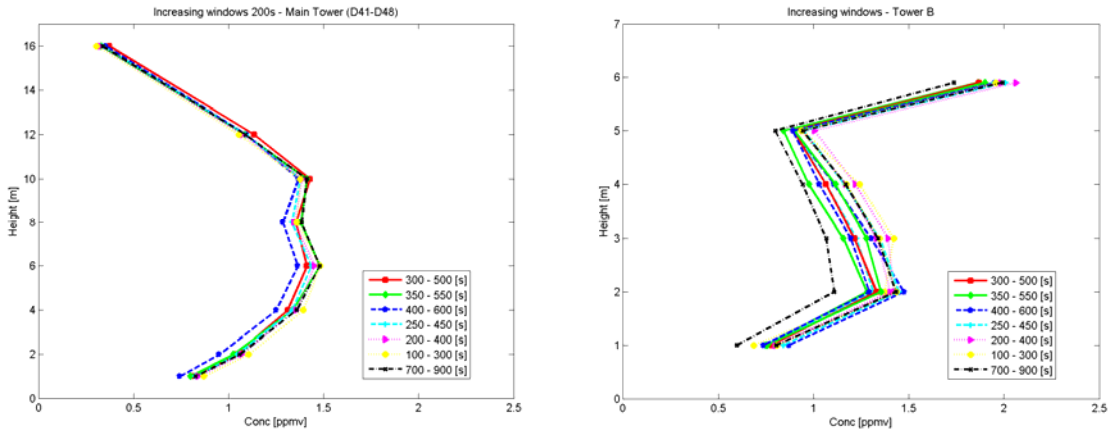



Figure 4.38: Increasing windows for concentrations, profiles at Main Tower (left) and B mast (right)

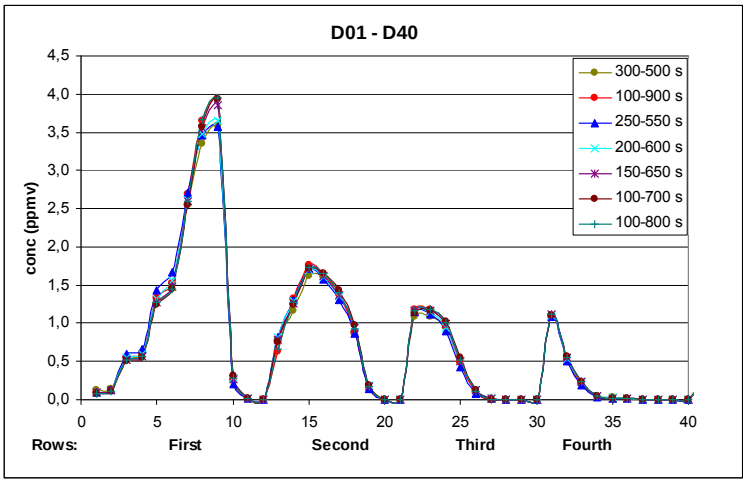



Figure 4.39: Increasing windows for concentrations at samplers along rows between obstacles

#### 4.4.2.8 *Data analysis: Boxplots, Moving and Increasing Windows of Meteorological and Concentration Observations*

Looking at the time series of 1-minute averaged concentration, for instance at the first and second rows of receptors with respect to the source location (Figure 4.40), we notice that the 'tails' of the 15-minutes emission periods are excluded by the larger 2nd reference period, 100-900s. This correctly allows to filter the transient related to the start and end of the emission, so to refer to an approximate "steady-state condition" also for the evaluation of concentration.

To summarize the various values from our analysis with moving and increasing windows, we can use box plots to represent the distribution of mean values that we have obtained with these different windows during the same release. As an example, we have used these plots for the concentration at towers B, C and D, as in Figure 4.41: the sensor number is indicated on the x axis. Similar box plots can of course be drawn for the concentration variances and fluctuation intensity also.

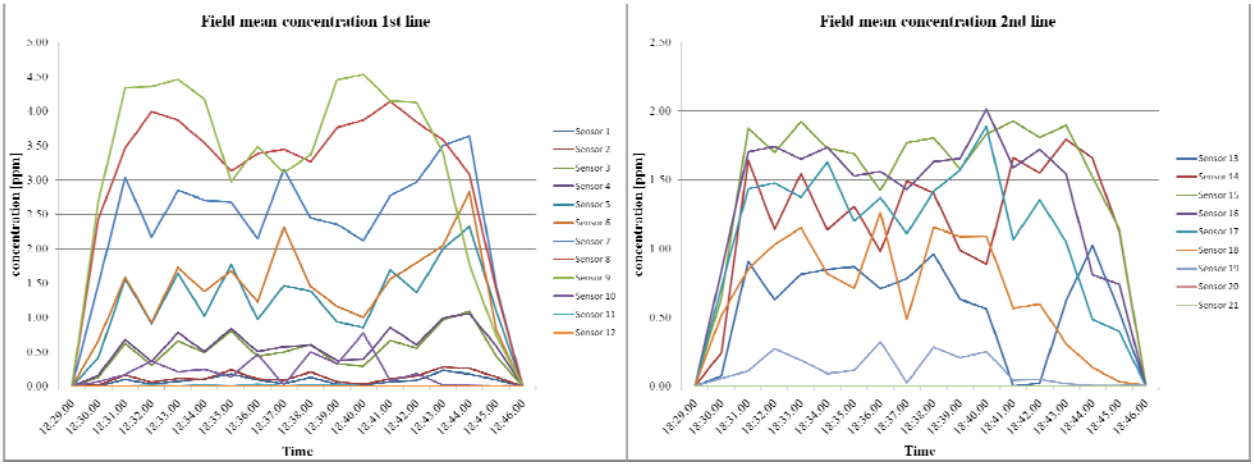



Figure 4.40: Sequence of 60s sub-periods averaged meteorological variables in the full sampling period. South (left) and Main (right) Towers.

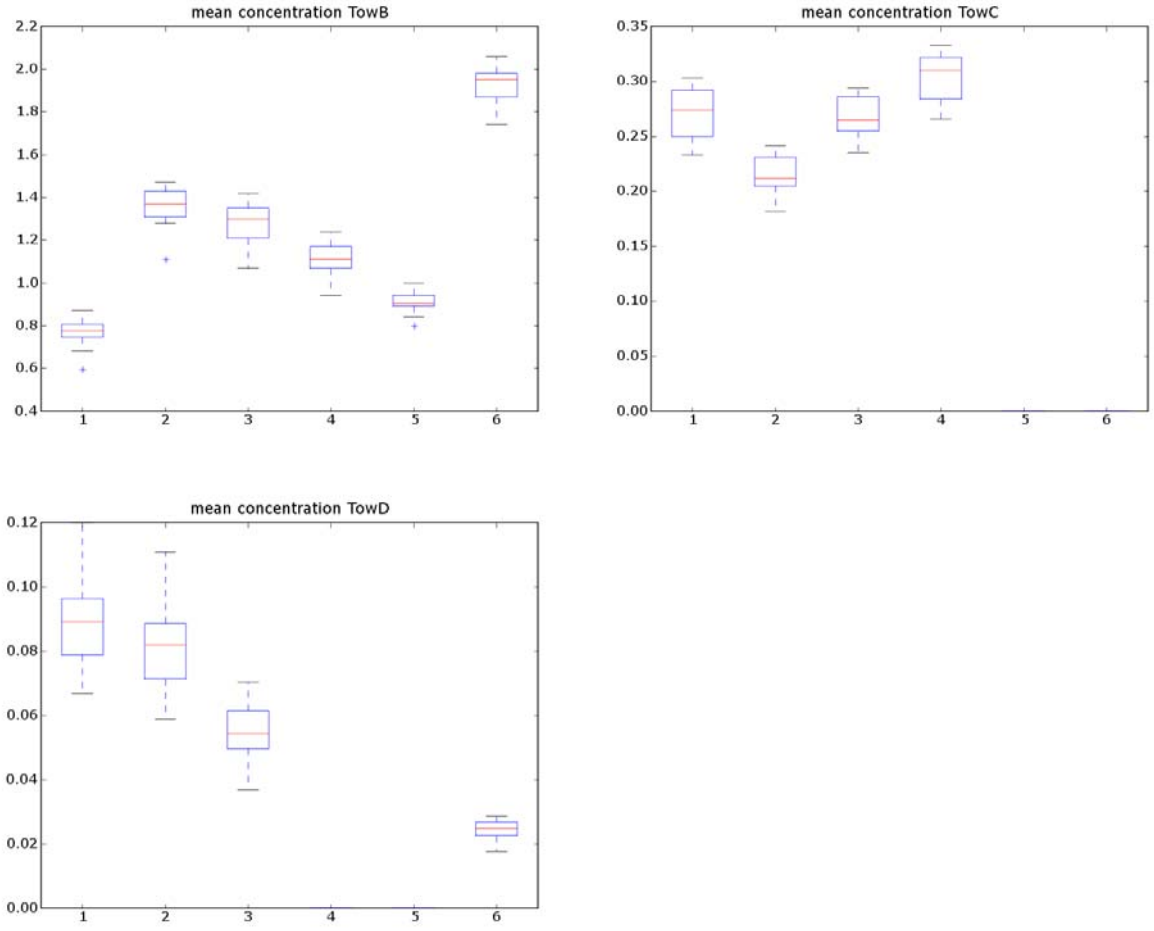



Figure 4.41: Example of box plots for the mean values of concentrations using different time windows.

In the box plots, the top and bottom of the box represent the 25 and 75 percentile, respectively. The red line is the sample median. The whiskers are lines drawn from the ends of the box and with a length of 1.5 interquartile range. Values outside are outliers and indicated by Symbol+. If the box is small, the variability is small around the median value, the red line.

Based on the same sample of values for each level and each tower, one can also estimate the confidence intervals and calculate the min and max of the sample, analogously to what was done for the wind tunnel as a function of sample time. For the towers B,C and D we estimated the values as reported in Table 4.13.

| | | **h=1m** | **h=2m** | **h=3m** | **h=4m** | **h=5m** | **h=6m** |
|---|---|---|---|---|---|---|---|
| **TowB** | p5 | 0.73 | 1.31 | 1.24 | 1.07 | 0.88 | 1.89 |
| | p95 | 0.79 | 1.39 | 1.32 | 1.15 | 0.93 | 1.96 |
| | mmin | 0.60 | 1.11 | 1.07 | 0.94 | 0.80 | 1.74 |
| | mmax | 0.87 | 1.47 | 1.42 | 1.24 | 1.00 | 2.06 |
| **TowC** | p5 | 0.26 | 0.21 | 0.26 | 0.29 | | |
| | p95 | 0.28 | 0.22 | 0.28 | 0.31 | | |
| | mmin | 0.23 | 0.18 | 0.24 | 0.27 | | |
| | mmax | 0.30 | 0.24 | 0.29 | 0.33 | | |
| **TowD** | p5 | 0.08 | 0.08 | 0.05 | | | 0.02 |
| | p95 | 0.09 | 0.09 | 0.06 | | | 0.03 |
| | mmin | 0.07 | 0.06 | 0.04 | | | 0.02 |
| | mmax | 0.12 | 0.11 | 0.07 | | | 0.03 |

Table 4.13: Confidence intervals for observed data at the towers.

### 4.4.3 Comparative analysis with Wind Tunnel Data

To connect our analysis to observations collected in the wind tunnel, we considered three time series of concentration for the Wind Tunnel case where the inflow wind direction was -45deg, measured at 1.725 m above ground. The three time series are briefly described in the following Table 4.14.

| time series | $X_{FS}$ [m] | $Y_{FS}$ [m] | $Z_{FS}$ [m] | Uref@8m [m/s] |
|---|---|---|---|---|
| MUST_C1_103.txt.ts#2 | -91.5 | -9.75 | 1.725 | 5.37 |
| MUST_C1_104.txt.ts#1 | -75 | -15.75 | 1.725 | 4.78 |
| MUST_C1_108.txt.ts#1 | 33.75 | -55.5 | 1.725 | 5.01 |

Table 4.14: Details of the Wind Tunnel series considered for the analysis.

Here the “FS” subscript refers to the equivalent full scale value. The time series 103 and 104 are 240 seconds long, providing 80000 measurements for a total time (equivalent full scale) of 18000s (i.e. 5 hours). The time series 108 correspond to an observational period of 180 seconds, that is about 60000 measurements, i.e. close to 4h at full scale. The statistics calculated for these three full time series are reported in the following Table 4.15.

| | C1_103 | C1_104 | C1_108 |
|---|---|---|---|
| mean | 18.6 | 40.1 | 34.3 |
| Std | 38.0 | 29.9 | 14.3 |
| median | 3.6 | 33.7 | 33.4 |
| percentile 5 | 0 | 5.7 | 12.5 |
| percentile 95 | 87.6 | 96.4 | 58.8 |

Table 4.15: Statistics calculated for the 3 Wind Tunnel time series referred to in Table 4.14.

In order to apply a similar methodology as was done above for the field data, a moving-windows analysis was applied also to these time series, considering three increasing window sizes : 900s (4000 measurements in each subset); 450s (2000 measurements in each subset) and 225s (1000 measurements in each subset). In the following Figure 4.42 the results of the analysis with different windows are plotted for Δt =900s (left) and 225s (right), as examples of the variability of the means obtained when considering different averaging periods.

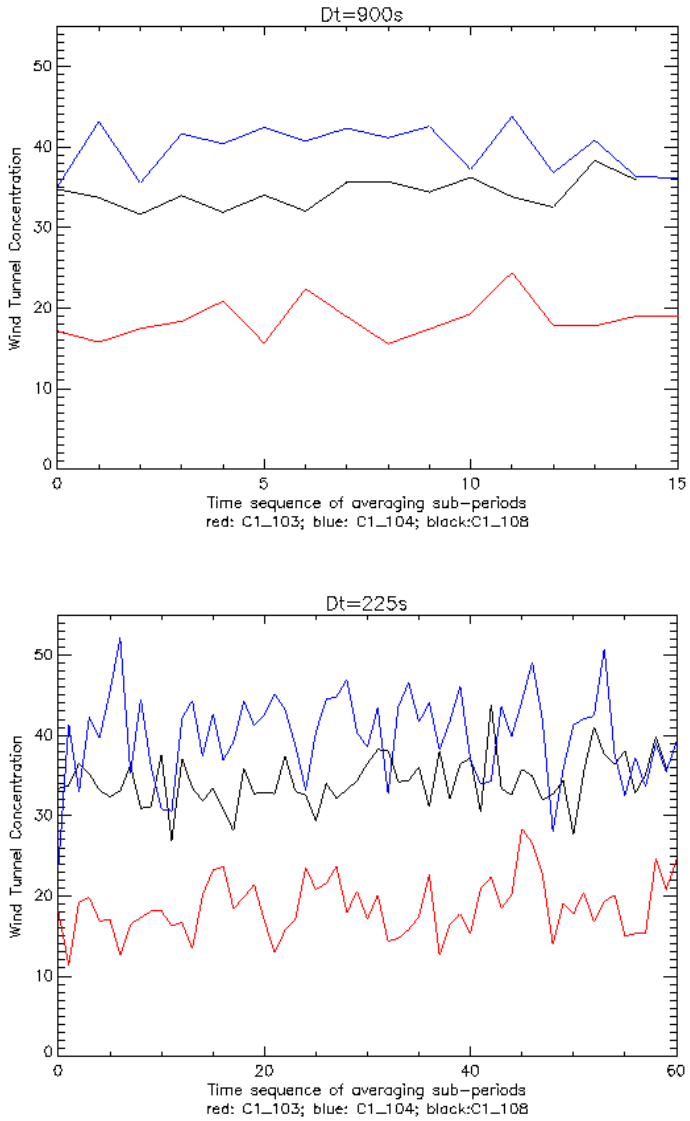



Figure 4.42: Sequence of sub-periods averaged concentrations in the full sampling period: 900s (top) and 225s (right) for WT C1_103 (red), C1_104 (blue) and C1_108 (black) series.

This figure shows that the variability of the 900s mean is small in the wind tunnel and if we had only one 900s measurement time series it would give a reliable estimate of the true mean (5h time serie). For the 225s averaging period, the variability is clearly larger.

In the following figures we plot histograms of normalized concentrations C/Cmax, where Cmax is the maximum concentration. In the first set of Figures 4.43, the full period of the Wind Tunnel dataset for C1_103, C1_104 and C1_108 is considered (4 or 5 hour), while in the second group of Figures 4.44 the histograms reflect the distributions during the first 800 s sub-period. For comparison, similar plots are presented for the 800s period of field data (Figure 4.45) at three samplers located in comparable positions inside the row array in the field data experiment.

DN01: centered between rows J1, J2, I1, I2

DN04: centered between J3, J4, I3, I4

DN08: centered between J5, J6, I5, and I6

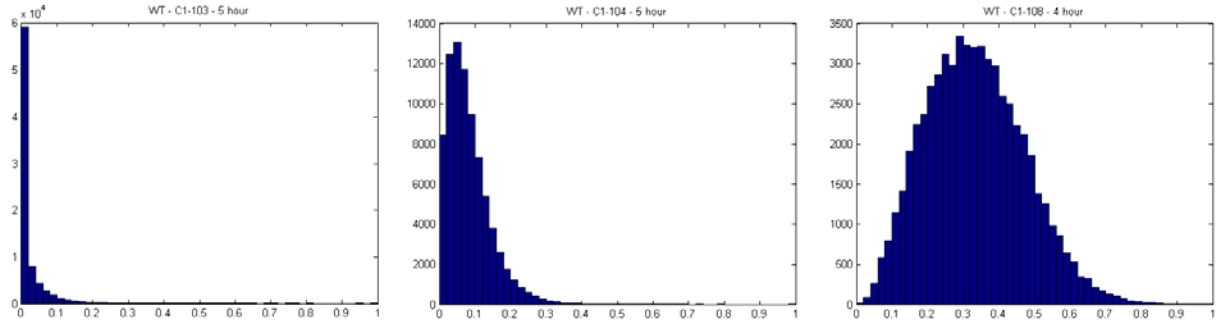



Figure 4.43: Histograms of the wind tunnel data : full time period

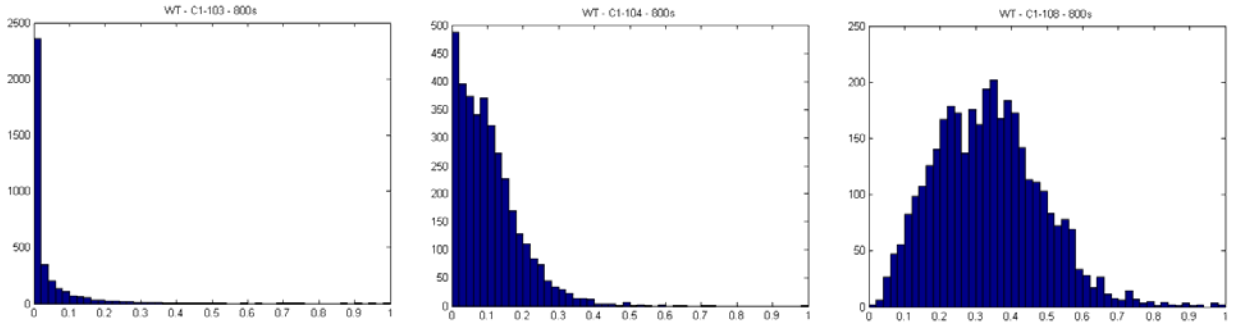



Figure 4.44: Histograms of the wind tunnel data : 800 s time period

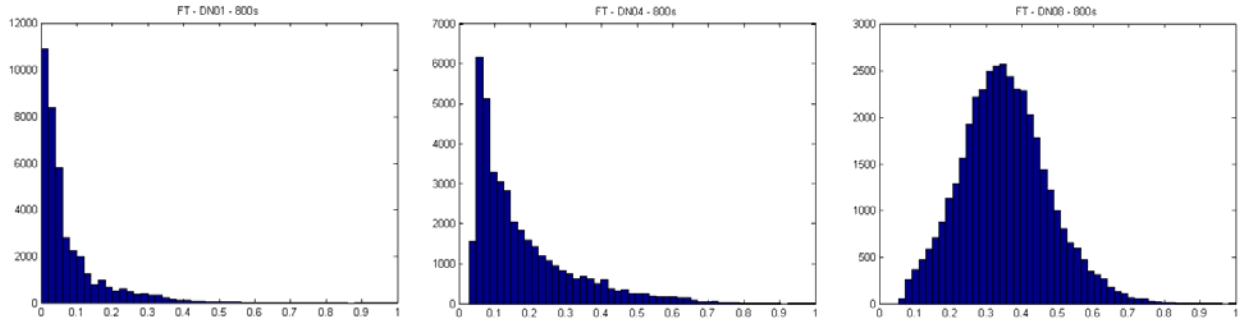



Figure 4.45: Histograms of the field data : 800 s time period

From left to right, the distance from the source is increasing. The concentration distribution is found log-normal near the source and increasingly Gaussian further away.

The distribution estimated from a short sequence (800s) in the wind tunnel provides a reasonable approximation to the “true” distribution obtained after 5h. Similar properties can also be found in the field data for the 800s period. This suggests that also the field data may be used with a reasonable confidence to test the numerical models.

### 4.4.4 Discussion and Conclusion

The original MUST database provides already some averages and fluctuation values that had previously been used in model validation (for example Milliez and Carissimo, 2007).

This chapter has addressed the issue of estimating the uncertainties / variability surrounding these values. To this end various possibilities to compute the means by moving and increasing the averaging window have been explored. A similar analysis has been performed with some long time series obtained in the wind tunnel (4 to 5h long). By taking a subset of these wind tunnel time series, it has been shown that the short sampling time 200-800s can give a reasonable estimate of the “true” mean defined over the full duration.

Although these field estimates are certainly not as rigorous as those obtained in the wind tunnel where the experiment can be fully controlled, they however give some insight which periods of the field measurements may provide consistent averages (small differences with moving/increasing windows, little scatter in the plots, small box plots, small confidence intervals…) and could be used for validation purposes.

## 4.5 Validation of CFD models with Oklahoma City wind tunnel data

### 4.5.1 Introduction

As stated in Chapter 3.2.2.3, the Oklahoma City test case was chosen as second validation test towards the end of COST action 732. Because of the short time and the geometrical complexity validation could only be started and is not finished yet. Being aware of this problem it was decided in the very beginning that a sensitivity study should be made to assess the influence of the source treatment and the parking garages on the results. Simulations were therefore done with and without the interior of the parking garages shown in Figure 3.9 in Chapter 3.2.2.3. For the assessment of the source treatment the real source geometry was once modelled in detail with an inlet boundary condition, and simplified as a volumetric source.

During the exercise an error was detected in the CAD geometry of the wind tunnel model. It was noticed that fifteen of the measurement points for the flow field were inside a building. This was due to the fact that the final set up in the wind tunnel, which was used to generate the CAD geometry, did not comply with the set up when the measurements were conducted. During the measurements building F8a and F6 were aligned, but in the CAD geometry their position differed by 1.82m, see Figure 4.46.

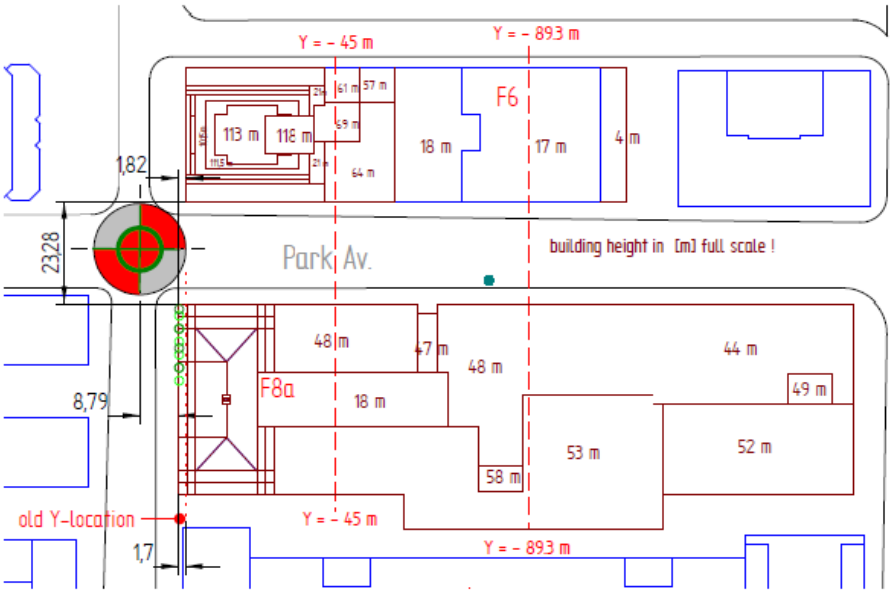



Figure 4.46: Misplacement of building F8a in the last CAD geometry.

Due to this error another parameter could be tested, the sensitivity of the simulation results to a small error in the position of the building upwind of the measurement region in Park Avenue. Note that the full scale misplacement of 1.82m corresponds to 6.05mm in wind tunnel scale (1:300).

### 4.5.2 Exploratory analyses

Like for the MUST wind tunnel case two Excel workbooks have been developed for the Oklahoma City case to enable an easy and consistent analysis of the simulation results. The data are organised in pretty much the same way as for the MUST case. With the workbooks profiles of flow variables and concentrations can be created to visually compare the simulation results with the experiments. The computation of metrics and the generation of scatter plots is possible as well. To date the workbooks only consider the mean velocities and the mean concentrations, but they will be extended to also deliver information about the turbulence quantities like RMS velocities, two-dimensional turbulent kinetic energy and concentration fluctuations.

#### 4.5.2.1 Flow modelling, 180.7 degree

All measurements have been performed close to Park Avenue, where the ground source is located too, cf. Figure 4.46 and Figure 3.9 in Chapter 3.2.2.3. The general location of the flow measurement planes is shown in Figure 4.47, while Figure 4.48 gives an impression of the detailed measurement positions. The latter Figure is taken from the Excel workbook for flow variables. Altogether there are 2839 Measurement positions, each with two components of the mean velocities and three components of the Reynolds stress tensor.

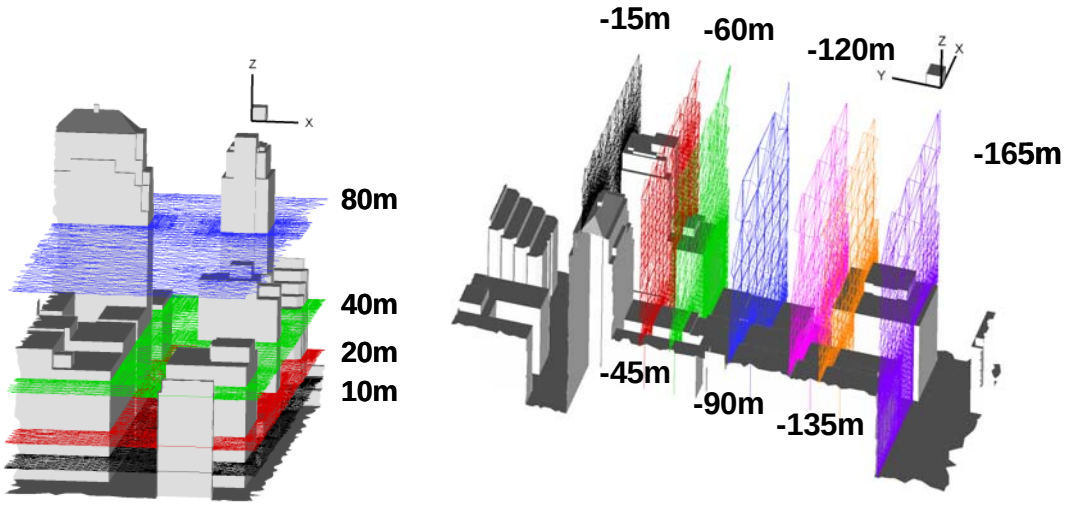



Figure 4.47: Velocity measurement planes in Park Avenue. Flow is in positive x-direction.

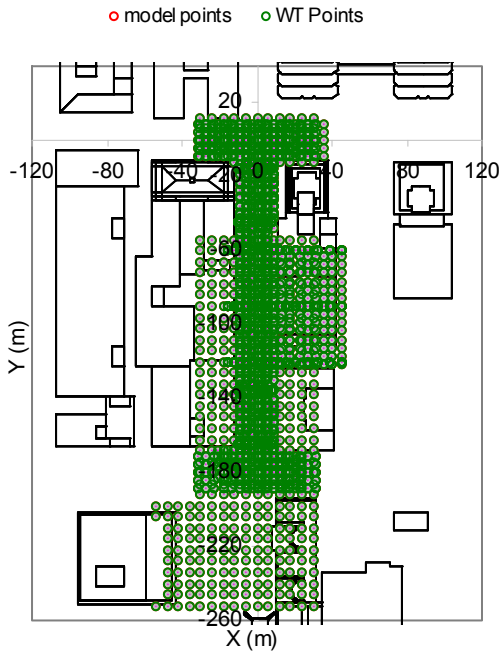



Figure 4.48: Velocity measurement positions in Park Avenue. Flow is in positive x-direction.

As the pollution dispersion is most strongly influenced by the flow close to the source location, velocity results are first shown for the measurement plane z = 10m. The measured flow field is shown in Figure 4.49 as vector plot. In this plane the fluid enters Park Avenue from both sides. At the corners weak recirculation regions are visible, but part of the flow is already parallel to the building walls aligned with the y-axis. This parallel flow in negative y-direction quickly extends over the entire street width. Above the ground source, shown with a red cross, the flow direction is towards negative y-values.

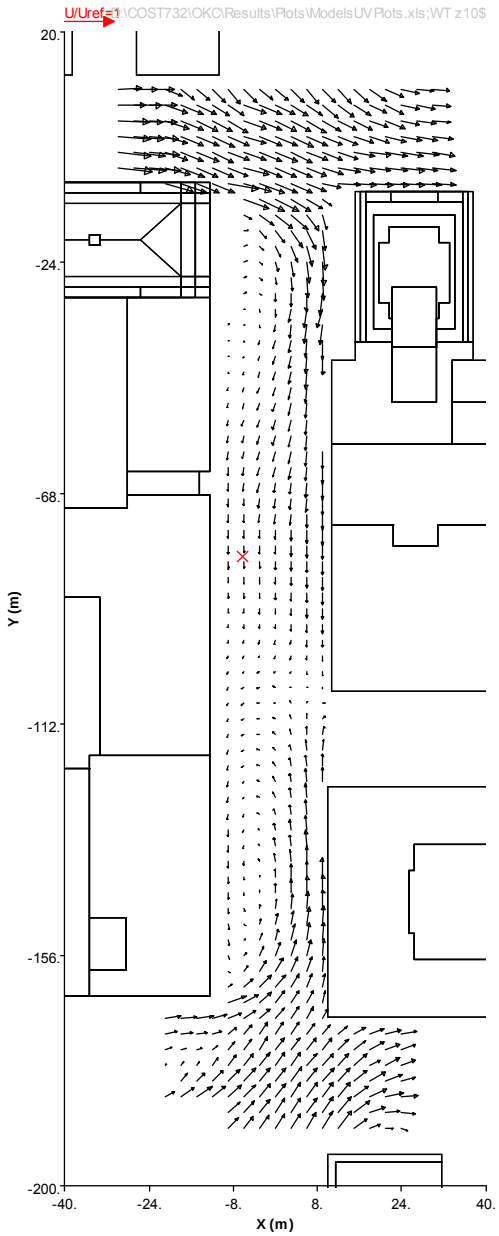



Figure 4.49: Measured velocity field at z = 10m. The red cross indicates the ground source location.

From the simulations a completely different flow field is obtained at z = 10m, Figure 4.50. For the old geometry, shown on the left, the footprints of three well established vortical structures can be discerned. Also the flow entering the street in positive y-direction differs much from the experiment. For the corrected new geometry, shown on the right in Figure 4.50, again a different flow field is obtained. Remember that the only difference between the two simulations is a 1.82m shift of the building on the left side of Park Avenue in negative y-direction, Figure 4.46. This is less than 10% of the street width, but leads to a substantially different flow field. Compared to the result with the old geometry the upper vortex gets smaller and stronger. In the lower part of the street the flow field is closer to the measurements, cf. Figure 4.49. And also the flow field close to the street between the buildings on the right side of Park Avenue is in better agreement with the measurements. However, above the ground source, shown with a red cross, the flow in the corrected geometry has the opposite direction as the experimentally determined flow.

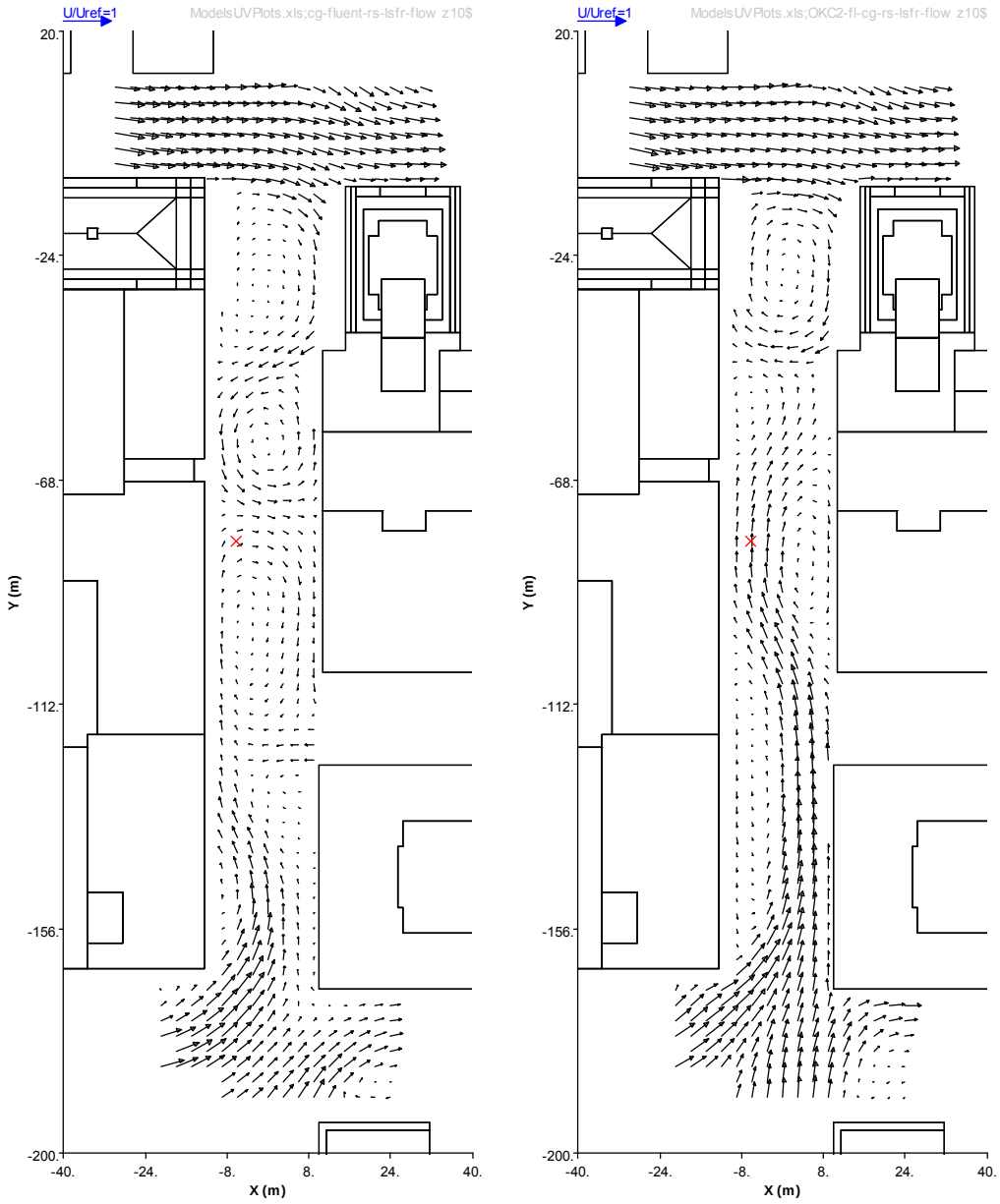



Figure 4.50: Velocity field at z = 10m for the wrong (left) and corrected (right) geometry. Garages open, real source. The red cross indicates the ground source location.

The second sensitivity test with regards to the geometrical representation concerns the treatment of the open parking garages. As explained in Chapter 3.2.2.3 these garages were for one kind of simulations not at all taken into account by treating every exterior wall as solid. For the other simulations the interior of the garages was reproduced and meshed so that the flow through them was computed. Figure 4.51 shows the influence of the open garages on the flow field at z = 10m. On the right side the vector field for the case without garages is shown. Now this flow field is again different from all the other flow fields shown and discussed above. While in the upper part of Park Avenue the flow field resembles very much the flow field in the wrong geometry, the lower part has more similarity with the flow field in the corrected geometry. This part therefore agrees better with the experiments.

On the left of Figure 4.51 the influence of the source modelling on the flow field is shown. In this case a volumetric source is used in the corrected geometry with modelled garages. This flow field should be compared with the one on the right in Figure 4.50. There the real source is taken into account, with a small volume flow rate entering into the computational

domain. The other parameters are identical. As expected there are no visible differences in the flow pattern.

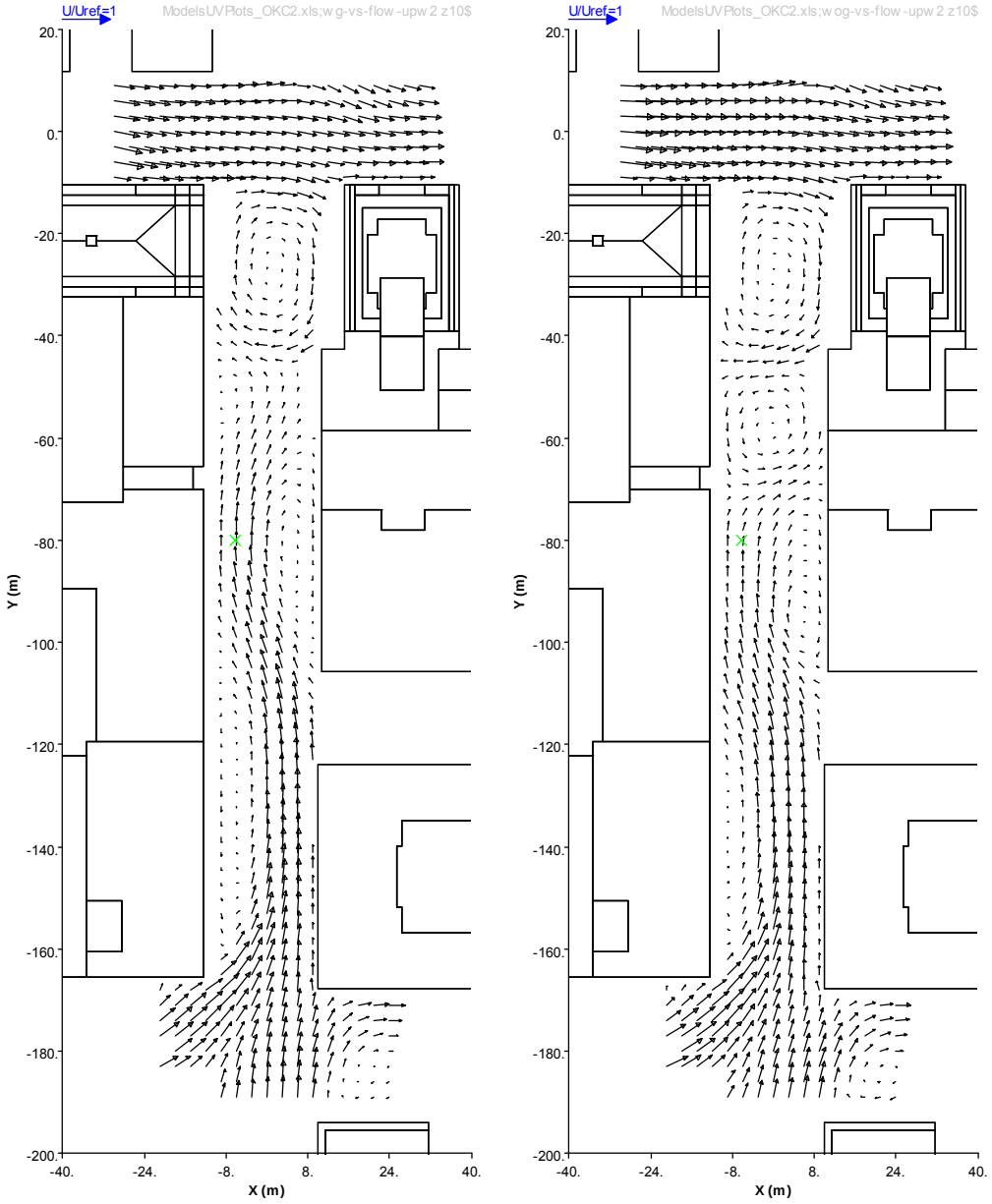



Figure 4.51: Velocity field at z = 10m for the open garages (left) and no garages (right). Corrected geometry, volumetric source. The green cross indicates the ground source location.

With increasing height the differences between the measured and calculated horizontal flow fields get smaller. This can be seen in Figure 4.52 where the experimental flow field is compared with the simulated one in the wrong and the corrected geometry at z = 20, 40 and 80m. At z = 20m there are even large differences between the computed and measured flow fields outside Park Avenue. These differences get smaller with increasing height and at z = 80m the agreement between both simulations and the measurements is good.

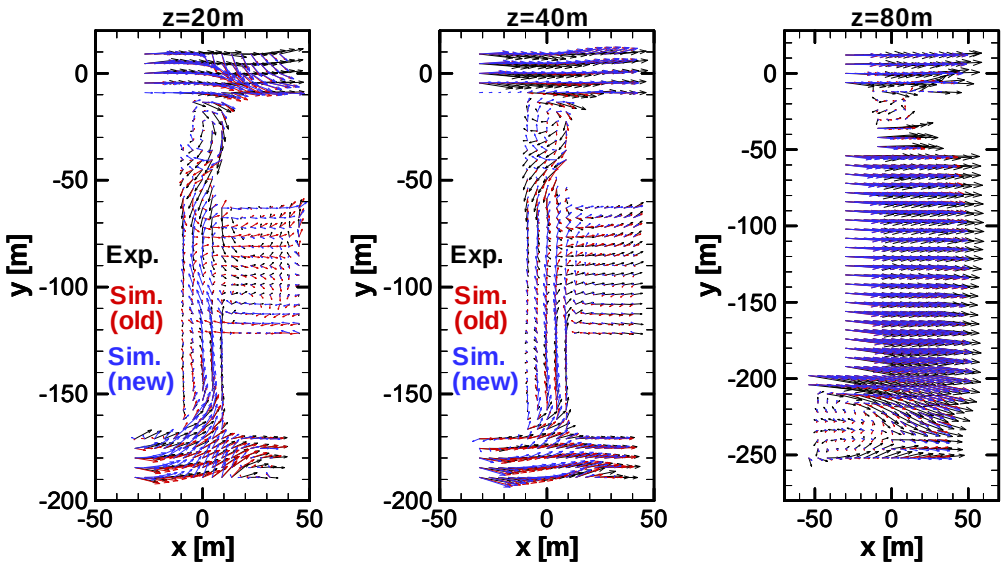
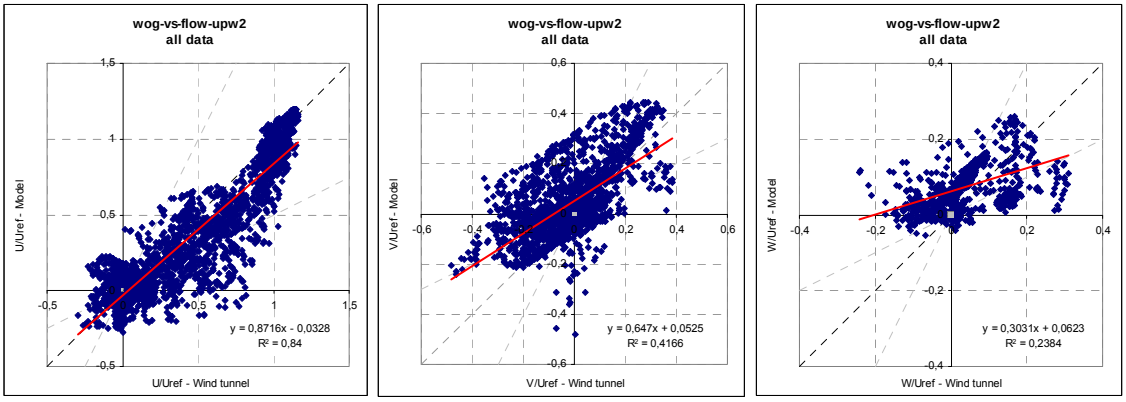



Figure 4.52: Velocity vectors at z = 20, 40 and 80m. Measured (Exp.) flow field and simulated flow field in the wrong (old) and corrected (new) geometry. Garages open, volumetric source.

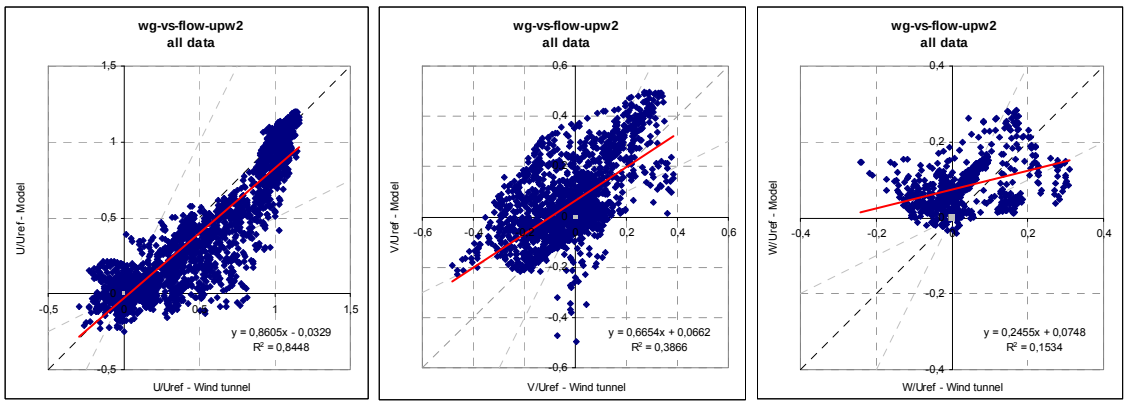
Finally scatter plots and validation metrics are shown for the mean velocity components at all measurement positions. The results in Figure 4.53 are for the case of the corrected geometry, no garages and volumetric source. As expected the hit rates are lower than for the MUST case, although the threshold values of the MUST case were used in this preliminary analysis. Although, because the measurement uncertainties are smaller for Oklahoma city than for MUST.




| Metrics | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 |
|---|---|---|---|---|---|---|---|---|---|---|
| U | 2828 | 0,409157684 | 0,323792579 | 0,423514 | 0,402769 | 1,035088 | 1,24391 | 0,84 | 0,360325 | 0,56082 |
| V | 1928 | -0,02386576 | 0,037069001 | 0,146208 | 0,146565 | -6,126281 | 3,953846 | 0,4166 | 0,171162 | 0,43361 |
| W | 900 | 0,036353496 | 0,073331162 | 0,090631 | 0,056274 | 2,493052 | 0,767393 | 0,2384 | 0,068889 | 0,204444 |

Figure 4.53: Scatter plots and metrics for the velocity components at the measurement positions. Corrected geometry, no garages, volumetric source.

In Figure 4.54 the corresponding results are shown for the simulation with corrected geometry, open garages and volumetric source. Comparing both results it can be seen that the metrics for the case without garages, which does not reproduce the setup of the wind tunnel experiment, are slightly higher than the ones shown in Figure 4.54 for the simulation which best reproduces the wind tunnel measurements of the flow. The difference are however quite small.




| Metrics | #Points | Average obs. | Average mod. | Std. obs. | Std. mod. | CV obs. | CV mod. | R2 | Hit rate | FAC2 |
|---|---|---|---|---|---|---|---|---|---|---|
| U | 2828 | 0,409157684 | 0,319164558 | 0,423514 | 0,396508 | 1,035088 | 1,24233 | 0,8448 | 0,360325 | 0,541018 |
| V | 1928 | -0,02386576 | 0,05034572 | 0,146208 | 0,156483 | -6,126281 | 3,108166 | 0,3866 | 0,145228 | 0,400934 |
| W | 900 | 0,036353496 | 0,083677982 | 0,090631 | 0,056821 | 2,493052 | 0,679049 | 0,1534 | 0,04 | 0,193333 |

Figure 4.54: Scatter plots and metrics for the velocity components at the measurement positions. Corrected geometry, open garages, volumetric source.

#### 4.5.2.2 Dispersion modelling, 180.7 degree

Like for the MUST case a second Excel workbook exists for the concentrations measurements. These measurements are available at the planes and locations shown in Figure 4.55. There are five horizontal planes, starting at z = 3m and one vertical plane. Altogether there are 683 measurement positions.

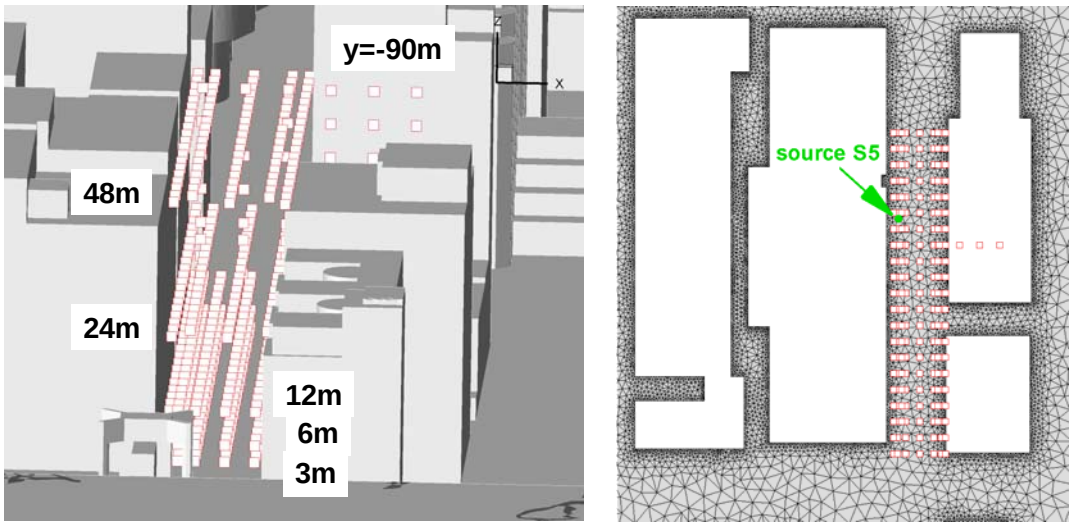



Figure 4.55: Concentration measurement positions in Park Avenue. Flow is in positive x-direction.

With the large disagreement between the simulated and measured flow fields close to the ground where the tracer is released one cannot expect a good agreement for the concentrations. Due to the qualitative wrong prediction of the flow field at z = 10m the tracer is transported in the opposite direction in the simulation. In Figure 4.56 it can be seen that in the experiment the maximum concentration at z = 3m is observed downwind of the source close to the building on the left. Contrary to this the simulations predict the concentration maximum close to the building on the right side. Accordingly the metrics are very low and not shown here.

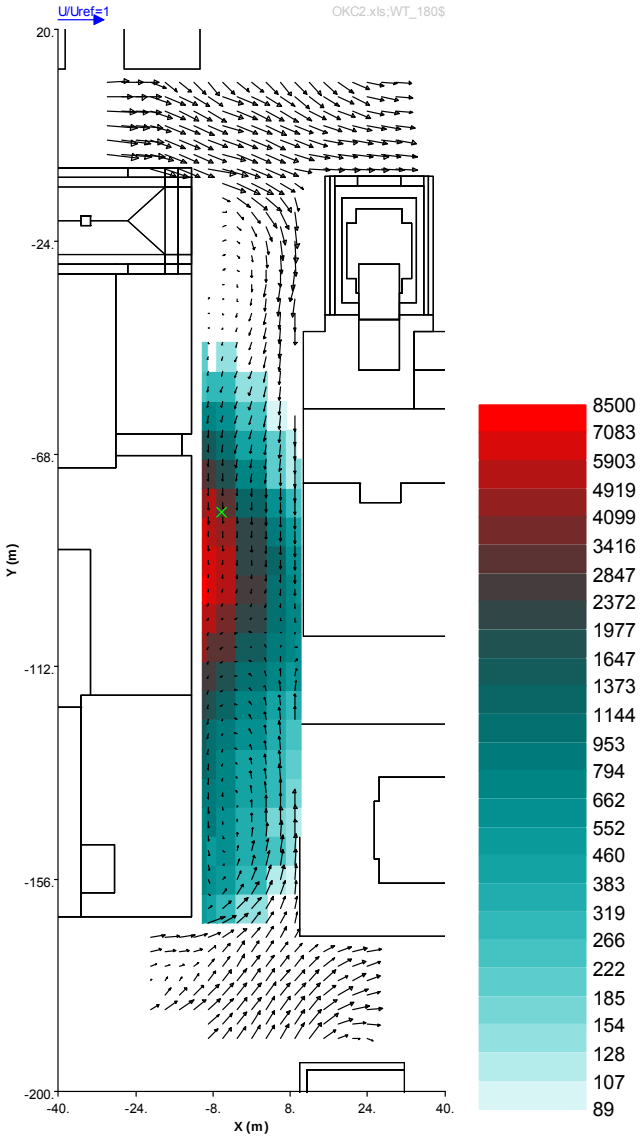



Figure 4.56: Measured flow field at z = 10m and dimensionless concentration C* at z = 3m.

### 4.5.3 Conclusions

The sensitivity analysis of the influence of the treatment of the open garages and the treatment of the source did yield the following results. The modelling of the garages has a clearly visible influence on the flow field in Park Avenue, where the measurements are taken. Therefore they should be modelled as in the wind tunnel experiment as open structures. Contrary to this the flow field and dispersion is not strongly influenced by the treatment of the source. With the real geometry with an inlet boundary one obtains the same results as with a volumetric source, if the volumetric source is defined in close agreement with the real source geometry.

Another finding is that the computed flow field is very sensitive to even small changes in the positions of the buildings. This substantiates the need for well documented and reproducible setups in the wind tunnel.

However, none of the computed flow fields showed a good agreement with the measurements. At the measurement positions close to the source even a qualitatively wrong

flow field was predicted with negative consequences for the dispersion predictions. The most probable reason for the bad agreement between the measured and computed flow field in Park Avenue is the grid that has been used. On one hand it is already relatively coarse in a short distance from the buildings. On the other hand it is purely tetrahedral outside of the buildings. Therefore the boundary layers on the building walls are not accurately predicted. The only chance to substantiate this assumption is the generation of an improved mesh with a higher mesh density in the area close to the measurement region and above all with at least prismatic layers on the walls, cf. Franke et al. (2007).

## 4.6 Some ideas concerning the fitness for purpose

### 4.6.1 Introduction

Microscale meteorological models are used for a variety of purposes. The optimal model should be applicable in the full range of meteorological conditions and dispersion scenarios. However, the choice of the model typology and its further use often depends on the specific purpose and scale of the problem treated, and the performances of the model may vary in different case studies. Thus, the validity of the models is subjected to the purpose for which they are used.

Comparison between model results and measurements is a common practice to judge if a model is valid for a specific purpose or not. This judgement is a quantitative one – by stating that a model is valid for a certain purpose, we are saying that we are confident that the model is able to estimate the value of a certain variable (relevant for the purpose), within a certain range of error (considered acceptable). Moreover, two points must be considered: model results depend on a series of set-up choices (initial and boundary conditions, grid resolution, physical options, etc.); the ability of the model to estimate the relevant variable (e. g. the error committed), varies with the case study (different errors can be expected if the model is used to run the MUST case, or the Oklahoma case). As a consequence it is important that a validation activity is performed for every application of the model, and in this sense it would be more correct to speak about validation of model runs, or simulations.

The criteria for the comparison should address the following points:

1. From a precise definition of the purposes of the simulation, a *variable of interest* for the specific purpose must be defined. For example if the purpose of the simulation is the evaluation of a strategy to improve the air quality in a certain area, the variable of interest may be the maximum pollutant concentration below 2m height in the area, or the size of the region where the concentration is above a certain threshold.
2. A metric to measure the *distance* between a “representation of the real world” (e. g. simulation results) and the “real world” (both when it is possible to ‘measure’ it in practice or not) must be defined. For example this can be the difference between the maximum of pollutant concentration below 2m height computed by the model, and the “real world” maximum in a certain region. We specify that with the word “metric” here we indicate any technique that is able to quantify the distance between a “representation of the real world” and the “real world”, or between two “representations of the real word”. It should not be interpreted strictly in mathematical sense.
3. A *quantitative criterion of acceptance* must be defined. For example, the simulation fits the purpose if the distance (computed with the metric defined in point 2) is less than a certain threshold value *K*. Again this should be dependent on the purposes. This will also be seen as the uncertainty we can expect (and accept) in model results for the variable of interest and that we may consider acceptable on the basis of some established criteria.

Another important element to qualify the appropriateness of a model for a certain purpose

(for example emergency response) is the computational demand and the simulation time needed to obtain the results, but this point is not addressed here.

If the variable of interest is directly measured, the methodology described above can be easily applied. If it is not, strictly speaking the methodology described above cannot be used. This second case is, in general, very common. This is because the variable of interest is in general 'global': maximum of concentration over a whole region, area with concentration above a certain threshold, etc., and cannot be derived from point measurements since they do not cover the whole domain.

The question to be addressed here is how to build confidence that a model is able to simulate a variable of interest that is not measured, with certain accuracy, by comparing model results with measured variables. More broadly, this is equivalent to look for the relationship between different metrics involving the same magnitude (for example pollutant concentration), and different metrics involving different magnitudes (for example pollutant concentration and mean wind speed). Moreover, as a by product of this, interesting insights, that can guide future model development, can be expected.

### 4.6.2 Methodology

As explained above, a fundamental in the present paradigm is the recognition that for a large number of purposes, the relevant information, that is the variable of interest, cannot be measured, but can be derived from model's output. This is because models provide 3D fields, and can be used to derive the global values that are needed to estimate the variable of interest. The criteria to decide if a simulation is fit for a certain purpose should be based on the difference between the 'real world' value of the variable of interest, and the correspondent value derived from model's output. This difference can be 'measured' using a metric, called $d_{purpose}$. If $d_{purpose}$ cannot be computed, because the variable of interest is not measured, another metric should be defined, called $d_X$, based on the measured variables, that estimates the distance between model's outputs and the measured aspects of the 'real world' (RMSE, FB, Hit rate, etc.). These metrics can be computed, because they are based on measured variables.

The question is, then, which is the best $d_X$ that can surrogate $d_{purpose}$. The idea proposed in this contribution, is to choose the best $d_X$ for a specific $d_{purpose}$, by analyzing the similarities between $d_{purpose}$ and $d_X$ in estimating the differences among the simulations. The advantage of this approach is that *both* $d_{purpose}$ and $d_X$ can be used to evaluate the differences between simulation results.

More formally, let's assume that we went through points 1), 2) and 3) above, and we have identified a metric $d_{purpose}$ and a quantitative criterion *D,* that is able to discriminate between simulations that fit the purpose and simulations that do not, or that for a model *i*

$$d_{purpose}\left(M_i^R, R\right)\begin{cases} \leq D \Rightarrow \textbf{\textit{fit}} \\ > D \Rightarrow \textbf{\textit{do not fit}} \end{cases}$$

Where $d_{purpose}\left(M_i^R, R\right)$ is the "distance", based on the metric $d_{purpose}$, between the modelled value $M_i^R$, and the real world value $R$, of the variable of interest. In cases where the "real world" *R* measurements of the variable of interest are not available, the question is how to find a surrogate metric of $d_{purpose}\left(M_i^R, R\right)$ using some variables *O* that instead have been measured. We are looking, then, for a metric $d_X\left(M_i^O, O\right)$ that can

approximate $d_{purpose}\left(M_i^R, R\right)$. Many metrics involving measured variables $d_X\left(M_i^O, O\right)$ can be defined (RMSE, FB, Hit rate, etc.). The problem is to find the best surrogate among the many possible metrics.

The basic idea is to use model intercomparison assuming that a sufficient number of simulations are available, from different models used by the same user, different users that use the same model, different set ups of the same model. Simulation results, in fact, give an estimate of the variable of interest jointly with an estimate of the values of other variables at the measurements points. It is possible, then, to estimate both $d_{purpose}\left(M_i^R, M_j^R\right)$ and $d_X\left(M_i^O, M_j^O\right)$. Making the assumption that simulations are, to some extent, a realistic representation of the real world, by exploring the relationship between $d_{purpose}\left(M_i^R, M_j^R\right)$ and $d_X\left(M_i^O, M_j^O\right)$, it is possible then to make an inference on the relationship between $d_{purpose}\left(M_i^R, R\right)$ and $d_X\left(M_i^O, O\right)$. In order to ensure that the simulations are a sufficiently realistic representation of the real world it is important that the physical formulations of the models and the simulations set up fulfils the recommendation of the Best Practice Guidelines.

The problem then, is how to compare $d_{purpose}\left(M_i^R, M_j^R\right)$and $d_X\left(M_i^O, M_j^O\right)$. We propose to adopt a technique based on the Kendall's tau and the lift curve, and we defined an approach (separation value) to translate the criterion of acceptance of $d_{purpose}$ in a criterion valid for the surrogate metrics $d_X$. This technique has been tested using 17 simulations (different models, different set-up, and different users) of the COST732-MUST exercise (the details of the technique and the analysis can be found in an extended document that will be available in the web page of COST732, http://www.mi.uni-hamburg.de/COST_732.464.0.html). Firstly, to validate the approach, two examples of $d_{purpose}$ were defined based on a variable that was actually measured. The intercomparison technique was carried on, and at the end, the findings were compared to those obtained by directly computing $d_{purpose}$ using the measurements. For one of the two $d_{purpose}$ defined, the results obtained from the simulation-to-simulation intercomparison were in agreement with those obtained from the simulation-to measurements comparison (Figure 4.57), but for the second one it was not. The important aspect, however, is that for the unsuccessful case, the values of the Kendall's tau and lift curve parameters were significantly lower than those obtained in the successful case, meaning that it is possible to decide, only from the simulation-to-simulation intercomparison if the technique can work or not.

Finally the whole approach has been repeated for two non measured variables (maxima over the whole domain at two different heights). The results obtained are satisfactory and the technique seems to be quite robust. It is important to remember that what is presented here is just an example. The numerical values are not relevant in themselves; rather the important thing is that such values exist and that it is possible to find them.

### 4.6.3 Conclusions

In principle the whole technique should be applied for each case study (city geometry, arrangement of measurements, etc.). A consequence of this is the following: while the quantitative criteria that applies to $d_{purpose}$ (the metrics involving the relevant variable for the purpose, and that cannot be computed) is only a function of the purpose, and can be the same for any case study and arrangement of measurement points, the quantitative criteria

for the metrics involving the measured variables ($d_X$, the one that finally will be used in practice), is dependent on the case study and the distribution of the measurement points.

To perform the analysis, the important thing is to have a set of model runs large enough (at least 10-15). These simulations can be obtained also by running the same model slightly varying the Initial or Boundary conditions, or with different physical options, using the techniques adopted in general for Model Ensemble studies.

The technique presented in this study, although tested for microscale models, can be applied to models at any scale.

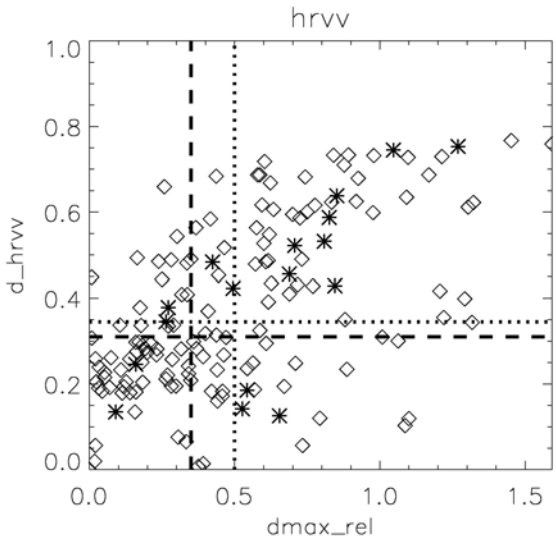



Figure 4.57: Distance between model results (diamonds), and between model and measurements (asterisks), based on hit rate for horizontal wind speed (d_hrvv, vertical axe, in this case considered as $d_X$) vs distances based on relative differences between maximum of concentration at the measurement points (dmax_rel, horizontal axe, in this case considered as $d_{purpose}$). The shape of the cloud of points based on the model-to-measurements distance (asterisks) is similar to the shape of the cloud of points based on the model-to-model distances (diamonds). For this case, a quantitative criteria to separate between fit and not fit for purposes simulations, can be based on the surrogate metrics $d_X$ based on hit rates for horizontal velocities.

# 5 Conclusions

This chapter presents a number of conclusions on model evaluation, evaluation data sets and modelling. Some of the conclusions are derived from problems that the COST action has encountered during its evaluation activities. Since lessons can be learned from these experiences they should be kept in mind. Although the groups who participated in the exercise consisted of experienced modellers/and or experimentalists, similar problems can arise in other contexts when the performance of models is evaluated.

## 5.1 On evaluation in general

- Within the COST action, the entire process of carrying out model performance

evaluation has been valuable as a means to increase the quality of model simulations, performed by the participants. The process has increased awareness of problems, and errors of many kinds might have passed unnoticed if the COST action had never taken place.

- One of the recommendations from the *Model Evaluation Guidance and Protocol Document* is to use 'exploratory data analysis' as one of the elements in model validation. Throughout the work with the data sets in the current COST action it was confirmed repeatedly that such exploratory analysis is crucial to reveal shortcomings of data sets, model setups and models. It is our experience that even experienced modellers sometimes do perform errors, and that these can be overlooked unless one inspects data, presented in graphical form. Conditions are best for detecting anomalies if you have a situation where *several models are put into a common framework*. Such a framework makes it also possible to identify features common to several models, and thereby derive an indication of the state of art. The tools developed for the MUST wind tunnel exercise provides one such excellent framework.
- When evaluating model performance, it is important to consider the representativeness of the underlying data. Section 4.2.3.1 provides illustrative examples of the effect of considering different subsets of data in the case of MUST wind flow simulations. The examples make it clear that a value of a certain metric tells nothing unless it appears in a context.
- When comparing metrics for several models it is essential that the comparison takes place for the same set of measurement points. Therefore it is recommended that metrics should not be computed by each modeller individually, but with a common tool such as the Excel workbooks developed for the MUST case of the COST action. Use of the workbooks makes it transparent which data are used to compute the metrics and helps in ensuring consistency.
- Occasionally modellers deliver results without checking the plausibility of the complete model output. Above all the more complex models deliver huge amounts of information. It is not sufficient to look only at the (usually small) subset of properties for which measurements are available. Recommendation: *Look at the complete model output!* In general it must be recommended to use graphical tools to get an overview of the results, and in the case of the MUST exercise such tools are readily available in the form of the Excel tools presented in section 4.2.2. Lots of things can go wrong and, in practice, they actually do.
- In the first phase of the exercise within the COST action, the magnitude of threshold values for computing hit rate etc. was chosen without carefully considering the size of data values. E.g., for the vertical component of wind speed, the threshold value was inappropriate: it was so high that even very bad model results were accepted as 'hits'. *A lesson to be learned is to carefully consider the implications of choosing the threshold value*.

## 5.2 On the data sets

- Regarding the MUST -45 degree case for dispersion in a wind tunnel, it turned out that the pollution cloud is relatively coherent. This has made possible certain kinds of useful analyses (Gaussian fitting to the plume, section **Fehler! Verweisquelle konnte nicht gefunden werden.**). However, this technique is not universally applicable. If the source had been positioned closer to a building, the pollution cloud might have been less coherent, and both modelling and model evaluation might have been more difficult.

- During the course of the action several versions of the MUST wind tunnel data set has been used for the exercise. This has led to confusion about which data a modeller should use, and to confusion about which basis was used when one considered a particular set of results. There have been compelling reasons to change from one version of the data set to a new: Errors in coordinates were detected; incorrect information on source conditions appeared in one version; some data points were omitted from the data set for while, before it was decided that they were okay for use. Doublets occurred in some wind tunnel data. For example, concentration data originally contained 11 doublets and one triplet. Doublets cause problems in data presentation and data processing, so data have been cleaned for doublets. Steps that can be taken to minimize confusion in future when conducting similar model validation exercises are:
    - *Indicate clearly in files when they were created or modified*, and preferably include a log of changes. It can be helpful to use a naming convention where the date enters into file names.
    - *Communicate to all interested parties whenever there is a change in files*.
    - To the extent practicable, let a limited group of persons take responsibility and *perform a test of data or and tools before they are distributed widely*.

(What else ????)

## 5.3 On CFD modelling

The present document includes results from a large number of models for the MUST wind tunnel case in the form of graphs and metrics, some of which are in the main document, while additional graphs can be found in Appendix B. This material gives an indication of the state of the art when both CFD and non-CFD models are applied to the case of the MUST wind tunnel simulations.

Some notable features from simulations for the MUST case are:

MUST 0 degree flow: Models have some difficulty in predicting horizontal wind speed u for "Wide streets". Otherwise, u prediction is good. Concerning w, the range of predicted w values is smaller than the range observed.

MUST -45 degree flow: In general, models have difficulty in predicting horizontal wind speed in 'Narrow streets' at low heights - they produce too low values. For other positions, u prediction is much better. For the vertical velocity w, it is a common feature that the models predict too low values (numerically) for the downward velocity. The vertical turbulent flux <u'w'> is apparently better predicted than the corresponding mean flux UW, but nevertheless, it can not be concluded that turbulence modelling is not responsible of poor prediction of W since both flux components strongly influence each other. Moreover, the models' ability to predict TKE is in general poor.

MUST -45 degree dispersion: Most models predict the plume trajectory well, but there are some exceptions. It is a common feature that models tend to predict too high values for the maximum concentration in the plume at distances more than 30 m downwind of the source (at the general measurement height of 1.25 m).

[There ought to be more conclusions on CFD modelling....

## 5.4 On non-CFD modelling

Whenever CFD modelling is available to simulate the complex flow structure influenced by an array of several buildings like in the MUST case, the application of non-CFD models is a priori not advisable. The MUST configuration is a regular array of buildings, and non-CFD models do not resolve building structures explicitly. Instead, they use rather simple parameterizations (increase in roughness length or parameterization of a cavity zone in the lee of a building and the influence of the wake behind on the dispersion).

In view of these reduced abilities of non-CFD models to deal with obstacles, the overall impression when looking at the results of the comparison with the wind tunnel data of the MUST experiment is nevertheless that of a surprisingly good agreement in most cases, statistically (see Figures 4.27 and 4.29 as well as Table 4.11). This conclusion is drawn on a comparison with wind tunnel data **unpaired in space** (see Section 4.2.1.1).

Looking closer at the results it is clearly seen that care and experience is necessary when applying even a simple model. The results for LASAT, for example, show a significant improvement from the grid with a North-East orientation (*LASAT a*) to the run where the grid is parallel to the building walls (*LASAT b*) and finally to the use of a MISKAM wind field (*LASAT c*). Another issue is the proper choice of a value for the roughness length which has an influence not to be neglected when looking at the ADMS or OML results. The use of a proper roughness length instead of including obstacles in the flow can apparently be a good alternative when applying simple models; however, one cannot expect to determine the exact localization of the concentration maximum with these model approaches. This is also the case when applying Gaussian models to an obstacle array like in MUST as they cannot treat flow distortions due to buildings. However, the range of concentrations and the concentration maxima are reproduced surprisingly well by almost all the model runs.

In light of the application of non-CFD models to the MUST exercise, the following lessons have been learned:

- As most of the non-CFD models have a very simple parameterisation of the flow (mostly by only increasing the roughness length), their evaluation with observations has to be done **unpaired in space**. Evaluation of concentrations is best done via quantile-quantile plots or a comparison of centreline concentrations and plume width. For the Lagrangian particle dispersion model LASAT which has been applied with a diagnostic wind field model, it makes sense to evaluate flow results in the same way as for the CFD flow simulations.

- The evaluation has shown that also non-CFD models are sensitive to grid orientation and input parameters, like the roughness length. In the specific case at hand, a proper choice of the roughness length in Gauss models like ADMS can deliver results superior to the inclusion of a building which generates a wake with concentrations much higher than observed within the realistic array of buildings. If obstacles are included, and a rectangular grid is used, the model grid has to be oriented along the building walls (this is also true for CFD models).

- If there is interest to reconstruct the concentration field more accurately, choice of the wind field is crucial, and the dispersion module has to take into account for flow distortions due to buildings/topography, as can be done with Lagrangian models and Gaussian puff models. The combination of CFD flow fields with Lagrangian dispersion calculation as possible with a new LASAT module (reading MISKAM wind fields or another meteorological model) offers a promising way to combine the

advantages of a realistic flow simulation around obstacles with the advantages of Lagrangian dispersion modelling.

- The evaluation of non-CFD models is crucial as they are widely used in air pollution expertises. Due to generally short calculation times, they may be run over longer time series of input data to represent different meteorological conditions to give e. g. yearly averages or percentiles of concentrations. The results for the -45° MUST case give insight in strengths and weaknesses of these models for the specific case at hand.

# References


Allwine K. J. (2004) Overview of JOINT URBAN 2003 - an atmospheric dispersion study in Oklahoma City, Proceedings of the AMS Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone, January 11-15, Seattle, Washington, USA.

Bezpalcova, K. and Harms, F. (2005) EWTL Data Report / Part I: Summarized Test Description Mock Urban Setting Test. Environmental Wind Tunnel Laboratory, Centre for Marine and Atmospheric Research, University of Hamburg

Bezpalcova, K. (2006): Physical Modelling of Flow and Dispersion in Urban Canopy, Doctoral Thesis, Charles University in Prague, Faculty of Mathematics and Physics, 193 pp.

Biltoft, C. A.: (2001) Customer Report for Mock Urban Setting Test, DPG Document No. WDTCFR-01-121, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway, Utah, 58 pp.

Borrego, C., Tchepel, O., Costa, A.M., Amorim, J.H. and Miranda, A.I. (2003): Emission and dispersion modelling of Lisbon air quality at local scale. Atmospheric Environment, Vol. 37, p. 5197-5205.

Britter, R., and Schatzmann, M. (Eds.) (2007a): Background and justification document to support the model evaluation guidance and protocol document. COST Office Brussels, ISBN 3-00-018312-4.

Britter, R., and Schatzmann, M. (Eds.) (2007b): Model evaluation guidance and protocol document. COST Office Brussels, ISBN 3-00-018312-4.

Cadafalch, J., Pérez-Segarra, C.D., Cònsul, R. and Oliva, A. (2002): Verification of Finite Volume Computations on Steady-State Fluid Flow and Heat Transfer, J. Fluids Eng., Vol. 124, 11-21.

Carruthers D.J., A.M. McKeown, D.J. Hall, S. Porter (1999): Validation of ADMS against wind tunnel data of dispersion from chemical warehouse fires. Atmospheric Environment, Vol. 33, 12, 1937-1953.

Coleman, H.W. and Steele, W.G. (1999) Experimentation and Uncertainty Analysis for Engineers, 2nd Edition, John Wiley & Sons, USA, 275pp.

CERC (2006) 'ADMS-Urban, USER Guide', available from Cambridge Environmental Research Consultants, Cambridge, UK.

Chang , J.C. and S.R. Hanna (2004): Air quality model performance evaluation. Meteorol. Atmos. Phys., Vol. 87, 1-3

Chang, J.C and Hanna, S.R. (2005): Technical Descriptions and User's Guide for the BOOT Statistical Model evaluation Software Package. Available through www.harmo.org/ki

Efthimiou G., Trini Castelli S., Reisin T. (2008): Field Experiment MUST- Short Term Scientific Mission, COST 732. Scientific Report, 17 pages.

Eça, L. and M. Hoekstra, M. (2004): A verification exercise for two 2-D steady incompressible turbulent flows, In P. Neittaanmäki, T. Rossi, K. Majava and O. Pironneau (Eds.), Proc. ECCOMAS 2004, Jyväskylä.

FLUENT (2006): FLUENT V6.3 User's guide. Lebanon, New Hampshire: Fluent Inc.

Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B. (Eds.) (2007) Best Practice Guideline for the CFD simulation of flows in the urban environment. COST Office Brussels, ISBN 3-00-018312-4.

Franke, J., and Frank, W. (2008): Application of generalised Richardson extrapolation to the computation of the flow across an asymmetric street intersection. Journal of Wind Engineering and Industrial Aerodynamics, Vol. 96, 1616-1628.

Franke, J. (2009a). "Validation of a CFD model for the prediction of flow and dispersion in the urban environment", Proceedings of the 4th International Building Physics Conference, 15-18 June, Istanbul, Turkey.

Franke, J. (2009b). "Validation metrics of Reynolds stresses and turbulent kinetic energy for the MUST wind tunnel case of COST action 732", Proceedings of the 5th European & African Conference on Wind Engineering (EACWE 5), 19-23 July, Florence, Italy.

Grimmond, C., and Oke, T., Aerodynamic properties of urban areas derived from analysis of surface form', J. Appl. Meteorol., Vol. 38, pp. 1261-1292, 1999.

Hall, R.C. (Ed.) (1997) Evaluation of modelling uncertainty - CFD modelling of nearfield atmospheric dispersion. EU Project EV5V-CT94-0531, Final Report. WS Atkins Consultants Ltd., Woodcote Grove, Ashley Road, Epsom, Surrey KT18 5BW, UK.

Hanna, S.R., J. Chang, R. Britter, M. Neophytou (2003): Overview of Model Evaluation History and Procedures in the Atmospheric Air Quality Area. QNET-CFD Network Newsletter, Vol. 2, 5, 1-4.

Hanna, S. R., Hansen, O. R. & Dharmavaram, S. (2004): FLACS CFD air quality model performance evaluation with Kit Fox, MUST, Prairie Grass, and EMU observations. Atmospheric Environment, Vol. 38, 4675-4687.

Hemsch, M. (2002): Statistical analysis of CFD solutions from the drag prediction workshop, AIAA paper 2002-0842.

Hertwig, D. (2008) Dispersion in an urban environment with a focus on puff releases. Study Project, University of Hamburg, Meteorological Institute.

Jackson, P.S., and Hunt, J.C.R., Turbulent wind flow over a low hill. Q. J. Roy. Meteorol. Soc., Vol. 101, pp. 929–955, 1975

Janicke Consulting (2007) LASAT Reference book for Version 3.0., www.janicke.de.

Ketzel, M., Louka, P., Sahm, P., Guilloteau, E., Sini, J.F., and Moussiopoulos, N. (2001) Inter-Comparison of Numerical Urban Dispersion Models. Proceedings, 3rd Int. Conf. on Urban Air Quality, Loutraki, Greece.

Kobayashi K. and M. U. Salam (2000): Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92, 345-352.

Leitl, B. (2000): Validation Data for Microscale Dispersion Modelling. EUROTRAC Newsletter, 22, pp. 28-32.

Leitl, B, Pascheke, F., Schatzmann, M. and Kastner-Klein, P. (2003): Wind tunnel experiments within the scope of the Oklahoma City tracer experiments. Proc. Int. Workshop on

Physical Modelling of Flow and Dispersion Phenomena PHYSMOD2003, Prato, Italy, September 3-5, ISBN 88-8453-095-4.

Macdonald, R., Griffiths, R. and Hall, D (1998): An improved method for estimation of surface roughness of obstacle arrays. Atmospheric Environment, Vol. 32, 1857–1864.

MATLAB, The Math Works Inc., 1997. http://www.mathworks.com/

Milliez M. and Carissimo, B. (2007) Numerical simulations of pollutant dispersion in an idealized urban area, for different meteorological conditions. Bound. Layer Meteor., 122, 321–342

Olesen, H.R. and Berkowicz, R. (2008): Guide to Excel sheets for MUST experiment http://www2.dmu.dk/atmosphericenvironment/Docs/SpreadsheetInfo.pdf

Pascheke, F., Leitl, B., and Schatzmann, M. (2003) Results from Recent Observations in an Urban Boundary Layer. Proceedings, COST 715 Workshop on Urban Boundary Layer Parameterisations, European Communities, ISBN 92-894-4143-7.

Ratti, C, Di Sabatino, S, and Britter, R.E., Analysis of 3-D urban databases with respect to air pollution dispersion for a number of European and American cities. Water, Air Soil Poll: Focus, 2, pp. 459–469, 2002.

Ratti, C., Di Sabatino, S., and Britter, R., Urban texture analysis with image processing techniques: winds and dispersion, Theor. Appl. Climatol, Vol. 84, pp. 77–90, 2006.

Reisin T., Altaratz Stollar O. and Trini Castelli S. (2007): Numerical simulations of microscale urban flow using the RAMS model. Air Pollution Modeling and its Applications XVIII. Developments in Environmental Science, Borrego C. and Renner E. Eds, vol. 6, 32-44.

Richards, P. J., Hoxey, R. P., 1993. Appropriate boundary conditions for computational wind engineering models using the $k$-$\varepsilon$ model. Journal of Wind Engineering and Industrial Aerodynamics, Vol. 46&47, 145-153.

Roy, C. (2003): , Grid Convergence Error Analysis for Mixed-Order Numerical Schemes, AIAA Journal, Vol. 41, 595 – 604.

Schatzmann, M., and Leitl, B. (2009) Evaluation of Numerical Flow and Dispersion Models for Applications in Industrial and Urban Areas. Chem. Eng. Technol., 32, pp. 241-246.

Schatzmann, M., and Britter, R. (Eds.) (2005): Proceedings from the International Workshop on 'Quality assurance of microscale meteorological models'. European Science Foundation, ISBN 3-00-018312-4.

Trini Castelli S. and Reisin T. G. (2008): Application of a modified version of RAMS model to simulate the flow and turbulence in presence of buildings: the MUST COST732 exercise, International Journal of Environment and Pollution, in press

Trini Castelli S., Reisin T. G. and Tinarelli G. (2008): Development and Application of MicroRMS Modelling System to Simulate the Flow, Turbulence and Dispersion in the Presence of Buildings. Air Pollution Modeling and its Application XIX, Borrego C. and Miranda A.I. Eds., Springer Ed., 81-89.

VDI (2005). Environmental meteorology – Prognostic microscale windfield models – Evaluation for flow around buildings and obstacles. VDI guideline 3783, Part 9. Beuth Verlag, Berlin.

Yee, E., and Biltoft, C.A. (2004) Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles. Boundary-Layer Meteorology, 111pp. 363–415.

## References to web sites

URL 1: http://www.harmo.org/kit/BOOT_details.asp
Access to the BOOT software.

URL 2: http://www.mi.uni-hamburg.de/index.php?id=464
Official web site of COST 732

URL 3: http://atmosphericdispersion.wikia.com/wiki/COST_732_forum
The Wiki on Atmospheric Dispersion Modelling has a particular forum for COST 732, with possibility to report MUST experiences.

URL 4: http://www.dmu.dk/International/Air/Models/Background/MUST
Web site with Excel tools used for the MUST exercise.

URL 5:
http://www.harmo.org/conferences/proceedings/_Cambridge/Cambridge_proceedings.asp
On-line access to papers from the 11th conference on *Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes* in Cambridge, July 2007.

URL 6: http://www.harmo.org/conferences/proceedings/_Cavtat/Cavtat_proceedings.asp
On-line access to papers from the 12th conference on *Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes* in Cavtat, Croatia, October 2008.

# Appendices

The appendices are available as pdf files. They are currently available at the web addresses indicated below. They may later be relocated to the central COST 732 web site.

A. Overview table specifying modelling groups, model types and model runs. Can be downloaded as http://www.dmu.dk/atmosphericenvironment/Docs/AppendixA.pdf

B. Compilation of plots presenting results. Can be downloaded as http://www.dmu.dk/atmosphericenvironment/Docs/AppendixB.pdf