Università degli studi di Trieste

Department of physics

Master Degree in Physics

High spatial resolution simulation of pollutants dispersion in air. A lagrangian model application to urban pollution.

Candidate:
Fabio Baldassi

Supervisor:
Prof. Dario Giaiotti

Assistant Supervisor:
Dott. Giovanni Bonafé

Academic year 2017-2018

Università degli studi di Trieste

Dipartimento di fisica

Laurea magistrale in fisica

Simulazione della dispersione di inquinanti in aria ad altissima risoluzione spaziale. Applicazione di un modello lagrangiano a particelle all'inquinamento urbano.

Candidato:
Fabio Baldassi

Relatore:
Prof. Dario Giaiotti

Correlatore:
Dott. Giovanni Bonafé

Anno accademico 2017-2018

Abstract

Pollutants dispersion in the atmosphere and subsequent potential deposition over terrain are crucial topics for the assessment of environment health and air quality, even more so in an urban environment. Many models that simulate atmospheric dispersion are available nowadays; the quality of the dispersion simulation broadly depends on the capability to correctly reproduce micrometeorological fields. Large part of the work then consists in the characterization of the *lagrangian particles* motion (in the case of a Lagrangian dispersion model) or gas or aerosol concentration (in the case of an Eulerian model). Lagrangian particles are not to be confused with the real particles: they consist in a comfortable quantization of the emitted mass, that provides a simplified modellization of a complex phenomenon as it is diffusion. In this thesis work the model MSS (Micro Swift Spray) is used: it is a high spatial and temporal resolution model suite which gives the user a wide selection. The aim of this work is to explore some of the possibilities offered by the suite and to compare its performance to the measures available in the case of a traffic pollution scenario in Udine (Viale San Daniele). Much of the work was dedicated to the preparation of the run, because the suite relies on a cospicuous number of preprocessing tools. Data harvesting and data analysis were carried out with the help of the software R, using the ARPA FVG databases.

Sommario

La dispersione di inquinanti in atmosfera e la potenziale successiva deposizione a terra sono un argomento cruciale per la valutazione dello stato di salute dell'ambiente e in particolare della qualità dell'aria, a maggior ragione in ambiente urbano. Molti modelli che simulano la dispersione di inquinanti sono disponibili oggi; la bontà delle simulazioni dipende in buona parte dalla capacità di riprodurre correttamente i campi micrometeorologici. Dopo aver considerato il meteo, la maggior parte del lavoro è dedicata alla caratterizzazione del moto delle *particelle lagrangiane* (nel caso di modelli Lagrangiani) oppure della concentrazione dell'inquinante (nel caso di modelli Euleriani). Le particelle lagrangiane non vanno confuse con le particelle reali: esse infatti consistono in una quantizzazione di comodo della massa emessa che fornisce una modellizzazione semplificata di un fenomeno complesso come la diffusione. In questo lavoro di tesi è stato usato il modello MSS (Micro Swift Spray): si tratta di una suite modellistica ad alta risoluzione spaziale e temporale che offre all'utente un'ampia gamma di possibilità. Lo scopo di questo lavoro è esplorare alcune di queste possibilità ed eseguire un confronto tra simulazione e misure per un caso di emissioni da traffico nella città di Udine (Viale San Daniele). Molto lavoro è stato dedicato alla preparazione dell'ambiente di calcolo, poiché la suite modellistica fa affidamento su un gran numero di strumenti di preprocessing. La raccolta dei dati e le successive analisi sono state condotte con l'aiuto del software R, avendo a disposizione i database ARPA FVG.

Acknowledgements

I would like to express my gratitude to:

- My supervisor and my assistant supervisor, who patiently guided me in this work always willing to advise and help me.

- The whole CRMA team who were very kind to me and enabled me to accomplish this thesis.

- ARPA FVG for letting me work at their agency.

- My fellow students. I shared many moments of both difficulty and joy with them and the path to this thesis would not have been the same without them.

- My family and friends, who always stood by me. I know I can always count on them.

Without the help of all these people, this work would have never been done.

"I often say that when you can measure what you are speaking about,
and express it in numbers, you know something about it;
but when you cannot measure it, when you cannot express it in numbers,
your knowledge is of a meagre and unsatisfactory kind."

Lord Kelvin

Contents

Introduction

Atmospheric pollution is a very topical issue, both for human health and environment health. Atmospheric pollutants sources may be both natural (e.g. a volcano) and artificial (e.g. gaseous emission from industrial processes). It has already been assessed that during the last centuries the immission of pollutants in the atmosphere due to human activities has been increased, in such a way to produce a noticeable change in the planet climate (see IPCC, 2013). Focusing on a smaller scale, it is important to assess the impact of a certain gaseous/aerosol pollutant source on the environment, and primarily to human health and to activities directly or indirectly connected to it (e.g. deposition of aerosol over a cultivated area, livestock contamination). Starting from the 50s atmosphere scientists began to develop analytical and computational models to describe atmospheric dispersion; now models range from stationary situations to evolving ones and from very simple settings to highly complicated ones (Seinfeld and Pandis, 1998). All the computational models can be grouped in two categories: the Eulerian category and the Lagrangian one. Eulerian models are useful when dealing with chemically active pollutants, Lagrangian models do perform much better in case of massive dispersion over complex terrain and can trivially operate under the parallel paradigm.

This thesis explore the modeling possibilites offered by the model suite MSS, composed by a micrometeorological module (microSWIFT) and by a lagrangian dispersion module (microSPRAY). The diagnostic meteorological model consists of a refined interpolator based mainly on similarity theory relations for the atmospheric boundary layer (ABL) and surface layer, and on the null-divergence condition for the wind field. The lagrangian dispersion model is based on the scheme by Thomson (see Thomson, 1987) and includes some semi-empirical relations to describe peculiar features of a plume (see Briggs, 1969 and Anfossi et al., 1993). Furthermore, it can deal with both dry and wet deposition.

Chapter 1 provides information about the physics underlying the dynamics of the ABL. The main equations of fluid dynamics are presented and the problems arising in their solution are discussed, moreover the topics of turbulence and similarity theory are introduced. Chapter 2 describes the theory which lies under a lagrangian dispersion model and Chapter 3 contains the description of the model suite. Finally, the simulations' preparation and results are reported and commented in Chapter 4. This work was carried out at the Regional Environment Protection Agency of Friuli Venezia Giulia (ARPA FVG), with the support of the Regional Center of Environmental Modeling (CRMA).

Chapter 1

The Atmospheric Boundary Layer

1.1 Introduction

The atmospheric boundary layer is that part of the lower atmosphere directly influenced by the presence of the Earth surface, which acts as a confining element. The physics of the ABL relies on two main fields of study: fluid dynamics and thermodynamics. ABL properties are sensitive to many elements, both geographical and temporal; for example the latitude and the period of the year define the inclination of incoming solar radiation affecting the energy balance. As a consequence, the height of the ABL varies both with time and space. It is important to underline that, given the continuous evolution of the ABL between different states (see section 1.8) and the different scale dimensions which characterize the motion, ABL dynamics can be separated from the dynamics of the upper free atmosphere which instead supplies boundary conditions to the evolution of the ABL. This does not mean that the ABL and the free atmosphere do not influence each other but instead that different phenomena drive their evolution. ABL evolution has a relevant impact on the upper atmosphere at climatic and global scales, so even though over a few hours or over a few days the free atmosphere seems to be scarcely influenced by the ABL, feedback processes have to be parameterized.

In the free atmosphere the wind is usually in agreement with the geostrophic balance and the flow is laminar without turbulence (e.g. Holton, 1979). Instead the air in the ABL is conditioned by the drag of the surface and turbulence is a key characteristic of the dynamics.

In order to understand ABL evolution, the underlying physics has to be studied. Sections

1.2 to 1.7 present a recap of fluid dynamics and thermodynamics while section 1.8 briefly summarizes phenomenological features of the ABL evolution.

This chapter is based on the first chapter of Pierluigi Masai's master degree thesis (see Masai, 2018).

1.2 Elements of fluid dynamics

The atmosphere is a fluid made up of gases, mainly Nitrogen (about 78%), Oxygen (about 21%) and Argon (about 1%). In addition to gases, the atmosphere is also made up of vapors, i.e. fliuds that can undergo a phase transition, whose treatment requires some care. Since the study of the atmosphere, at least in the layers closer to the surface, deals with large masses of air, it is reasonable to rely on continuum mechanics. In reality materials are made up of discrete atoms, separated by space, while a continuum is a body that can be divided in infinitesimal parts that can still be described in terms of continuous functions. In fluid mechanics a good criterion to evaluate the validity of the continuum assumption is given by the Knudsen number, which is defined as the ratio of the molecular mean free path length of the fluid λ and a characteristic physical length scale L of the flow under study:

$$\mathrm{Kn} = \frac{\lambda}{L}$$

The smaller the value of the Knudsen number the more appropriate it is to rely on the continuum assumption. Generally the distinction between a molecular flow and a continuum flow is made for a value of the Knudsen number of 10. In conditions of small Knudsen number an important concept can be introduced: the fluid parcel. A fluid parcel is a small amount of fluid whose dynamical evolution can be completely described and which preserves the properties of a continuum, i.e. density, temperature, velocity and all the other physical fields have defined values. The mass of a fluid parcel is well defined and does not change with time whereas its volume may change (compressible flow) or not (isochoric flow). Parcels are not to be confused with particles: parcels describe properties of fluid particles (atoms and molecules) averaged over a length scale which is large when compared with the molecular mean free path but small

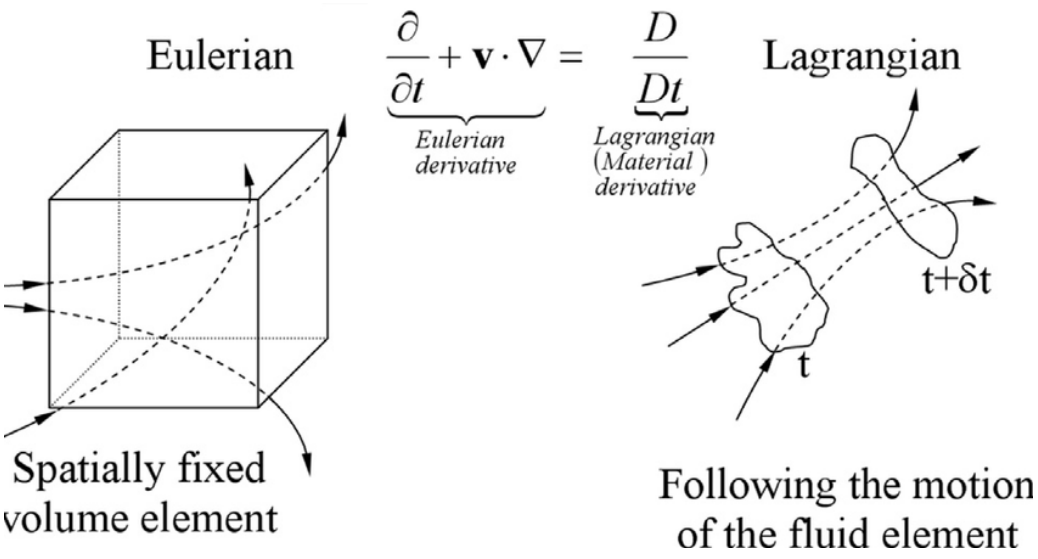


Figure 1.1: A representation of the Eulerian description and the Lagrangian description

when compared with the typical length scales of the motion under consideration.

At the core of continuum mechanics there is the need to describe the evolution of some physical fields which characterize the medium. In fluid mechanics the main interest is on the velocity field and there are two different approaches in building the equations: the Eulerian point of view and the Lagrangian point of view, which are equivalent and lead to the same results anyway (see figure 1.1). The Eulerian description represents a field as a continuous function of time and space and can be interpreted as the point of view of an observer at rest who watches the flow as it passes by and modifies the properties of the medium at any point. On the other hand, the Lagrangian description considers single fluid parcels, each one virtually labeled and distinguished from the others, and follows their motion through time specifying their positions. The Eulerian velocity field $\mathbf{u}$ is related to the Lagrangian position field by the 'material derivative' (also known as 'Lagrangian derivative'):

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + (\mathbf{u}\cdot\nabla)$$

For a fluid the motion of a parcel is described by the Navier-Stokes:

$$\rho\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla P + \mu\,\triangle\mathbf{u} + \mathbf{F}_{body} \tag{1.1}$$

where $\mathbf{u}$ is the velocity field, P is the pressure, μ is the dynamic viscosity, $\triangle$ is the Laplace

operator and $\mathbf{F}_{body}$ is the resultant of the body forces.

NSE alone cannot provide all the information needed to know the motion of the fluid because there are more variables than equations. Aside from NSE which expresses the conservation of momentum, fluid dynamics relies on other equations which express the conservation of energy and the conservation of mass. The conservation of mass is described by the continuity equation:

$$\frac{1}{\rho}\frac{\mathrm{D}\rho}{\mathrm{D}t} + \nabla\cdot\mathbf{u} = 0 \tag{1.2}$$

NSE are a non-linear set of equations and to this day no explicit solution is still known (except for trivial or specific cases). Such a fact shows from the start how the evolution of the atmosphere cannot be simply evaluated: NSE need approximations and experimental corrections to be studied. There is not an official proper way to proceed in such studies but there are many fundamental elements which can guide the decisions. The key idea is to underline the properties of the flow separating them from the properties of the fluid. In fact, different systems which share similar conditions at different scales have similar dynamics. In order to quantify such observations NSE have to be non-dimensionalized. This is a very important step in the analysis of the equations and a few examples are needed. First of all it has to be observed that any physical variable can be expressed as the product of a dimensionless variable and a dimensional value, e.g. $t = \mathrm{t}\cdot\mathrm{T}$ where t is the variable time ($[t] = \mathrm{s}$), t is a dimensionless variable and T is a dimensional value ($[\mathrm{T}] = \mathrm{s}$). The dimensionless variable t keeps the functional properties of t and varies along with it. On the other hand T can be seen as representative of the modulus of t and acts as a constant with respect to functional actions such as derivation or integration. The value T can be used to characterize the scale, i.e. an approximative quantification of the range of variation for the phenomena under consideration; this concept can be associated to the more accurate concept of order of magnitude. Such a representation applies to any physical variable, both scalar and vectorial, and can be extended to operators such as derivatives too. Here a few examples:

$$u_i = \mathrm{u}_i\cdot\mathrm{U} \qquad x_i = \mathrm{x}_i\cdot\mathrm{L} \qquad \frac{\partial}{\partial x_i} = \frac{\partial}{\partial \mathrm{x}_i}\cdot\frac{\mathrm{d}\mathrm{x}_i}{\mathrm{d}x_i} = \frac{1}{L}\cdot\frac{\partial}{\partial \mathrm{x}_i} \qquad \frac{\partial}{\partial t} = \frac{\partial}{\partial \mathrm{t}}\cdot\frac{\mathrm{d}\mathrm{t}}{\mathrm{d}t} = \frac{U}{L}\cdot\frac{\partial}{\partial \mathrm{t}}$$

It is a little bit harder to deal with pressure and density since for the study of the ABL it is more important to pay attention to the little variations (oscillations) of these quantities rather than to their magnitude; usually, a characteristic value P_0 of pressure is considered to define the dimensionless variable $\mathrm{P} = \frac{P-P_0}{\rho \mathrm{U}^2}$. Also, in the study of the atmosphere it is common to consider the Boussinesq approximation, according to which density variations are important only in relation with the gravity term whereas when they multiply the inertia term they can be neglected. A particular scale of motion can then be identified specifying the values of the scale quantities. Many definitions of the atmospheric scales of motion, based on the different phenomena that can take place, have been proposed over the years, perhaps the most important are the ones from Orlanski (Orlanski, 1975) and Fujita (Fujita, 1981). For example, the so-called 'synoptic scale' of motion is defined approximately by $\mathrm{U} = 30\,\mathrm{ms}^{-1}$ and $\mathrm{L} = 10^6\,\mathrm{m}$. Expressing all the variables as described, NSE can be non-dimensionalized and take the form:

$$\frac{\mathrm{U}}{\mathrm{L}}\frac{\partial(\mathrm{u}_i\mathrm{U})}{\partial \mathrm{t}} + \mathrm{U}^2\mathrm{u}_j\frac{1}{\mathrm{L}}\frac{\partial \mathrm{u}_i}{\partial \mathrm{x}_j} = -g\delta_{i3} - \frac{1}{\mathrm{L}}\frac{1}{\rho}\rho\mathrm{U}^2\frac{\partial \mathrm{P}}{\partial \mathrm{x}_i} + \nu\frac{1}{\mathrm{L}^2}\frac{\partial^2(\mathrm{u}_i\mathrm{U})}{\partial \mathrm{x}_j^2} \tag{1.3}$$

Multiplying 1.3 by L/U^2 and considering the operator $\frac{\mathrm{D}}{\mathrm{Dt}} = \frac{\partial}{\partial \mathrm{t}} + \mathrm{u}_i\frac{\partial}{\partial x\, i} \neq \frac{\mathrm{D}}{\mathrm{Dt}}$ we obtain:

$$\frac{\mathrm{Du}_i}{\mathrm{Dt}} = -\frac{1}{\mathrm{Fr}}\delta_{i3} - \frac{\partial \mathrm{P}}{\partial \mathrm{x}_i} + \frac{1}{\mathrm{Re}}\frac{\partial^2 \mathrm{u}_i}{\partial \mathrm{x}_j^2} \tag{1.4}$$

In 1.4 some important dimensionless quantities, analogous the previously described Knudsen number, are introduced:

- the Froude number $\mathrm{Fr} = \frac{\mathrm{U}^2}{g\mathrm{L}}$
- the Reynolds number $\mathrm{Re} = \frac{\mathrm{UL}}{\nu} = \frac{\mathrm{UL}\rho}{\mu}$

These dimensionless parameters define the flow, which means that they characterize a particular configuration of the motion of the fluid regardless of the physical extension of the flow. It is now immediate to understand that in case these quantities maintain the same values the functional form of the solutions of the NSE does not change. It is then possible to say that the motion of a parcel in a river can be the same, from a mathematical point of view, of that of a parcel of air in the upper atmosphere or of a parcel of oil in a pipe, provided that the values of the

aforementioned dimensionless quantities are the same. Moreover, from a physical point of view the actual values of Fr, and Re provide peculiar information on the flow:

- Fr quantifies the importance of gravity. A large value of Fr indicates the presence of stratification in the fluid.
- Re quantifies the importance of friction. A large value of Re indicates that the friction is not important for the flow.

1.3 Laminar and turbulent flows

Among the parameters discussed at the end of section 1.2, the most important to characterize the flow of a fluid is certainly the Reynolds number. In the atmosphere many configurations are to be considered, namely:

- Synoptic scale: Re $2 \cdot 10^{12}$
- Daytime ABL: Re $1 \cdot 10^{7}$
- Nighttime ABL: Re $1 \cdot 10^{-1}$

It is so clear that in the study of the atmosphere the action of friction is important but not at every scale. Re is also a great tool to distinguish two different kinds of flow: the laminar flow and the turbulent flow. Laminar flows are characterized by the motion of the fluid in parallel layers with no disruption in between. There are no eddies or swirls of fluid and, close to a solid surface, particles move orderly in straight lines parallel to the surface itself. Laminar flows are associated with a low Reynolds number. On the other hand, turbulent flows are characterized by chaotic motions in which unsteady vortices of different dimensions form and influence the dynamics. A vortex is a closed, or almost closed, circular trajectory and in general vortices can be seen as zones of the fluid which periodically exhibit the same conditions with some fluctuations. Turbulent flows are associated with a large Reynolds number.

From the dynamical point of view turbulence can develop in two ways: if there is instability in the fluid because of the presence of a gradient for some quantities, such as temperature, which

leads to the formation of convective motions causing mixing and transfer, than it is regarded as 'convective turbulence', whereas if there is a velocity shear which causes the formation of circular turbulent motions, than it is regarded as 'mechanical turbulence'.

It is important to stress that to associate turbulence with chaos can be misleading because turbulence actually has some organized aspects. In fact, turbulence is a mixture of coherent structures distributed in space and time which evolve in a random way. The coherent structures are essentially what we introduced as vortices which show such a coherence at least over the region of their extent, namely a size l. From the experimental point of view, eulerian measures (i.e. performed at a fixed point in space) show time series with fluctuations that appear to be random. In fact, the fluid evolution is deterministic, but chaotic. In case new measures were to be made, the new time series would be different but the average values would be the same, provided that the conditions which determine the system were the same. Such a feature is the consequence of an ergodic behavior, i.e. the time average of the process is the same as its spatial average and as its ensemble average, which is the average over the probability space. We can therefore quantify average values and compare the measurements with the theory.

1.3.1 Taylor's hypothesis and Kolmogorov's theory

It is very important, especially from an experimental point of view, to find a way to link spatial and temporal dimensions of a vortex. To do so, it is possible to rely on Taylor's hypothesis. It consists of introducing the concept of 'frozen turbulence' which refers to systems in which the advection of a vortex past a fixed point occurs in an amount of time significantly shorter than the time scale of existence of the vortex itself, so that the advection can be taken to be entirely due to the mean flow (see figure 1.2). Taylor's hypothesis actually implies that on certain time scales the properties of a vortex are conserved in the motion. This fact can be expressed saying that the lagrangian derivative of a property ξ of the vortex is null:

$$\frac{\partial \xi}{\partial t} = -u\frac{\partial \xi}{\partial x} - v\frac{\partial \xi}{\partial y} - w\frac{\partial \xi}{\partial z} \;\Rightarrow\; \frac{\mathrm{D}\xi}{\mathrm{D}t} = 0 \tag{1.5}$$

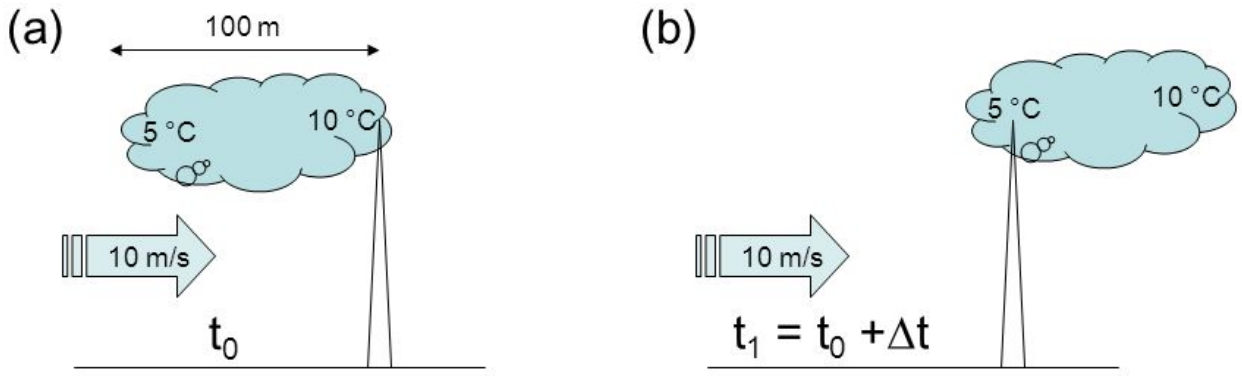


Figure 1.2: A depiction of Taylor hypothesis. An idealized eddy moves rapidly enough not to change noticeably for the sensor: it appears to be frozen. Adapted from Stull, 1988

which furnishes the link between spatial and temporal dimensions of a vortex from an eulerian point of view. A flow in which vortices move according to Taylor's hypothesis generates, in eulerian measurements, signals at fixed frequencies of the order of the various times of existence of a vortex that can be detected if the sample rate is significantly inferior to those time scales.

Figure 1.3 shows the wind power spectral density of some eulerian measures performed at Brookhaven by Van der Hoven (see Van der Hoven, 1957). Some peaks are evident testifying the contribute of different phenomena to the energy balance. Thanks to Taylor's hypothesis we can interpret a determinate frequency as a representative time scale for the existence of a vortex and associate to that vortex a length to estimate its dimensions according to equation 1.5. It is then possible to recognize the action of large scale phenomena, such as cyclones, and distinguish them from turbulent phenomena. A striking fact which can be observed in the power spectrum is the presence of an energy gap between the turbulent phenomena scale and the daily phenomena scale. It is possible to deduce that turbulent motions evolve on time scales which range from milliseconds to minutes. Always thanks to Taylor's hypothesis it is possible to switch from a temporal power spectrum (frequency spectrum) to a spatial power spectrum (wavenumber spectrum). To better understand all the features expressed by such graph it is important to recall Richardson's concept of energy cascade and Kolmogorov's theory which later formalized it. Lewis E. Richardson first expressed in the 1920s the key idea that there is a continuous transfer of energy in a fluid between different scales of motion, in particular vortices form at all scales but energy is injected by the mean flow just at the large scales whereas it is dissipated at the small molecular scales. On the other hand at the intermediate scales no

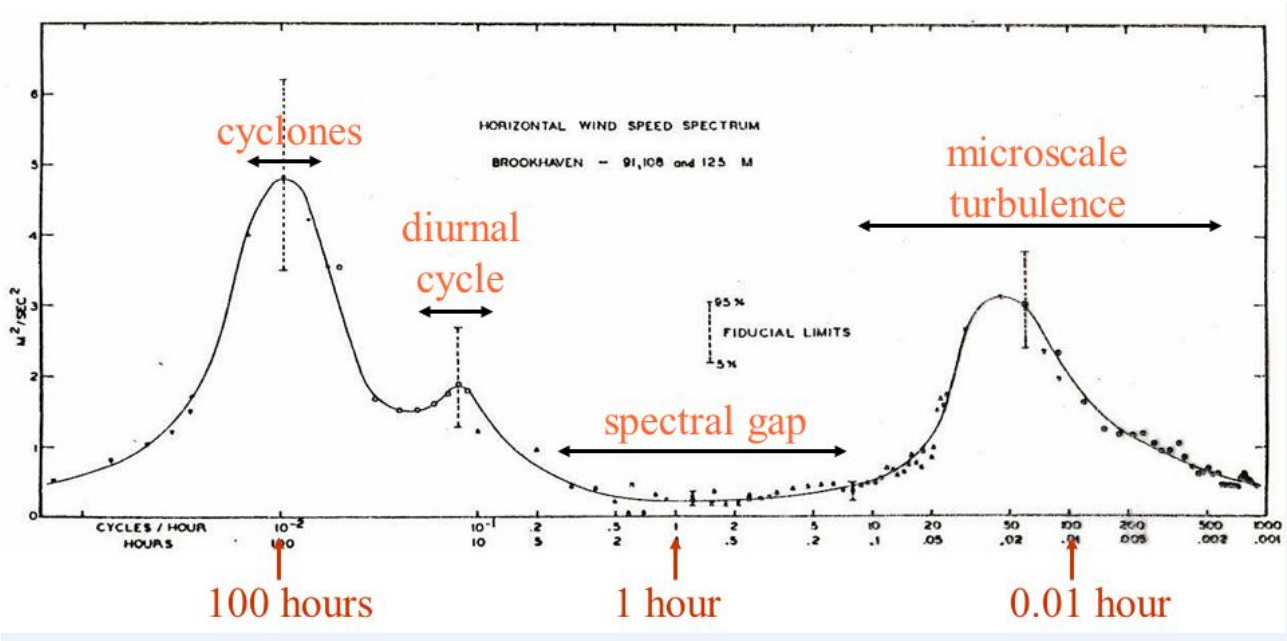


Figure 1.3: The power spectrum of the time series of the wind intensity in the ABL measured at Brookhaven. Adapted from Van der Hoven, 1957.

dissipation or injection of energy take place but a non-linear and non-viscous behaviour enables a continuous transfer of energy from higher to smaller scales like a cascade. Experiments have shown that the continuum assumption still holds for the smallest scale associated to turbulence and hence NSE are valid for a turbulent field.

The key idea introduced in section 1.3 to consider non-dimensionalized equations in order to point out general behaviours of fluid motion can be further taken and leads to the similarity theory. The aim of similarity theory is to find universal relationships between non-dimensionalized variables of fluids, making use of the Buckingham Π-theorem (see Barenblatt, 2003) and of *ad hoc* experiments; it is thus a semi-empirical theory. In the 1940s Andrey N. Kolmogorov used similarity theory and was able to quantify Richardson's intuition. A very important and striking result, known as Kolmogorov $-5/3$ power law, is an equation for $E(k)$ in the inertial scale:

$$E(k) = C \cdot \varepsilon^{2/3} \cdot k^{-5/3}$$

where C is the universal Kolmogorov constant, which was experimentally determined to be $C = 1.5$. Through the years many equations to describe even the dissipation range and the production range have been developed. Without going any further it can just be said that in

general the full spectrum can be expressed with the introduction of two more parameters as:

$$E(k) = C \cdot \varepsilon^{2/3} \cdot k^{-5/3} \cdot f_L \cdot f_\eta$$

1.4 Reynolds decomposition

In the study of the atmosphere, and in particular of the ABL, it is reasonable to consider turbulence to be homogeneous and stationary, i.e. statistically not changing over time. Therefore, it is possible to accept the ergodic condition and so to consider that space, time and ensemble averages coincide for different measurements taken over similar systems, i.e. systems with the same physical conditions. These considerations suggest to rely on statistics in order to study the evolution of the physical fields. The goal would be to have a set of prognostic equations, i.e. which predict the time evolution, for all the physical variables. Reynolds developed a powerful method to do so. The basic idea is to express every physical field as the sum of its expectation value and a deviation which represents the action of turbulence. For example, for the temperature T we would have:

$$T = \overline{T} + T' \qquad \text{with} \qquad \overline{T} = E[T] \qquad \text{and} \qquad E[T'] = 0$$

$E[T]$ is the expectation value of T and for real measurements has to be computed with a statistical estimator such as the arithmetic average (the notation $\overline{T}$ denotes a generic average). It is important to say that the expectation values are taken for physical fields that are functions of time and space and so they also depend on time and space. Once every variable is decomposed according to Reynolds' idea, new equations can be found. It has to be underlined that the expectation values are intended to provide information about the mean flow and so, from the experimental point of view, an appropriate physical system has to be under study, i.e. these ideas apply to flows large enough to consider the turbulent contributions as random variations so that the expectation value of these variations can be regarded as null. It is therefore important to take an appropriate interval of time or space when averages are taken with experimental measurements otherwise turbulent phenomena would not be outlined. Obviously,

these considerations can be related to the spectrum discussed in the previous sections (see figure 1.3). Applying an average operator to the fluid equations whose variables are rewritten in terms of the Reynolds decomposition one obtains equations containing second order correlation terms. Iteration of this procedure upon the second order correlation terms leads to prognostic equations for third order correlation terms, requiring new equations containing fourth order correlations and so on. It seems therefore that there is no end to such calculations and the system cannot be closed. In order to overcome these difficulties many solutions to the so-called 'closure problem' have been proposed. Now, terms which involve the product of a component of the velocity and a generic quantity α express the flux, i.e. transport, of that quantity in the direction of the velocity, which is to say that, for example, a term like UT expresses the transport of temperature in the x direction through an advective process. Reynolds decomposition makes it possible to distinguish between two different kind of fluxes: kinematic fluxes which are driven by the mean flow and are represented by terms like $\overline{U} \cdot \overline{\alpha}$, and eddy fluxes which are driven by turbulent motions and are represented by terms like $\overline{u'\alpha'}$.

1.5 Scaling

The term *scaling* describes a seemingly very simple situation: the existence of a power law relationship between certain variables y and x,

$$y = Ax^{\alpha},$$

where A, α are constants. Such relations often appear in the mathematical modelling of various phenomena, not only in physics but also in biology, economics and engineering. However, scaling laws are not merely some particularly simple cases of more general relations. Scaling laws always reveal an important property of the phenomenon under consideration: its *self-similarity*. This term means that a phenomenon reproduces itself on different time and/or space scales.

Simple but non-trivial considerations on physical dimensions (see Barenblatt, 2003) allow to state that any function $f(a_1, \ldots, a_k, b_1, \ldots b_m)$ that defines some physical relationship and whose

arguments are some dimensionally independent parameters (the a_i) and some parameters b_j whose dimensions instead do depend on those of the a_i, possesses the property of a generalized homogeneity, i.e. it can be written in terms of a function of a smaller number of variables and is of the following special form:

$$f(a_1, \ldots, a_k, b_1, \ldots b_m) = {a_1}^p \cdots {a_k}^r \Phi \left(\frac{b_1}{{a_1}^{p_1} \cdots {a_k}^{r_1}}, \ldots, \frac{b_m}{{a_1}^{p_m} \cdots {a_k}^{r_m}} \right) \tag{1.6}$$

This result leads to the central theorem in dimensional analysis, the so called Π theorem which states the following: *a physical relationship between some dimensional (generally speaking) quantity and several dimensional governing parameters can be rewritten as a relationship between a dimensionless parameter and several dimensionless products of the governing parameters; the number of dimensionless products is equal to the total number of governing parameters minus the number of governing parameters with independent dimensions.*

In mathematical form this is:

$$\Pi = \Phi(\Pi_1, \ldots, \Pi_m) \tag{1.7}$$

where

$$\Pi = \frac{f(a_1, \ldots, a_k, b_1, \ldots b_m)}{{a_1}^p \cdots {a_k}^r} \qquad \text{and} \qquad \Pi_i = \frac{b_i}{{a_1}^{p_i} \cdots {a_k}^{r_i}}.$$

Dimensional analysis is a powerful tool for studying certain equations whose analytic solutions are unknown, and also allows to determine the relevant parameters to be kept under control in an experiment. It is thus a precious instrument for both computer-oriented and experimentally-devoted physicians. This is the reason of the wide success of similarity theory in the study of the atmosphere. Moreover, similarity theory is a type of zero-order closure technique and could be applied in atmospheric computational models.

For meteorological purposes, similarity solutions are to be applied to stationary situations: rarely is time included as one of the relevant variables. Some variables, such as depth of the boundary layer, are so strongly dependent on time that no successful similarity expression have been found to diagnose them. Instead, boundary layer depth must be calculated or measured using other techniques. This depth is used as input into dimesionless groups to diagnose other variables that do reach a quasi-steady state.

In all this cases the *self-similarity* has to be intended as a similiarity between different experiments with the same experimental situation, which define a particular similarity class, that it to say, a well defined set of scaling variables.

The most important similarity classes (according to Stull, 1988) are:

- Monin-Obukhov similarity: works for non-calm wind situations, describes variables in the surface layer;
- Mixed-layer similarity: applies to free convection, assuming calm or light winds;
- Local similarity: works for statically stable boundary layers, where local fluxes, shears and stability are more important than surface fluxes;
- Local free convection similarity: works for statically unstable boundary layers, where local effects are more relevant on buoyancy.

1.6 The energy balance

Earth atmosphere can be regarded as a huge thermodynamic machine which basically converts solar radiative energy into heat and long wave radiation energy. The incoming solar radiation is both reflected and absorbed, depending on the kind of objects it encounters across its path. Matter-radiation interactions establish the behaviour of the incoming and outgoing spectra at the various heights of the atmosphere; the well-known atmospherical thermal profiles are the result of the available energy re-distribution by means of three different mechanisms: thermal conduction, convection and radiative processes. Observing those diagrams one can extract a lot of information.

The energy balance at the planetary surface can be expressed through an equation derived from the first law of thermodynamics and the ideal gas law. Without being too specific, the variation of temperature in a given volume of air in time is related to the heat flux through the same volume in time. Many processes contribute to the net heat flux: solar radiation, earth radiation, turbulence, advection, friction, plant transpiration, evaporation and others. Considering the diurnal evolution of the ABL, which is the main topic of section 1.8, the

radiation contributes are the most important. Usually the net radiation term is split into four components. According to Stull (Stull, 1988):

$$Q* = K\uparrow + K\downarrow + I\uparrow + I\downarrow$$

where the single components represent respectively:

- $K\uparrow$ = incoming reflected short wave (solar) radiation
- $K\downarrow$ = outgoing shortwave radiation transmitted through the air
- $I\uparrow$ = longwave (infrared, IR) radiation emitted up
- $I\downarrow$ = longwave diffusive IR radiation down

The distinction between shortwave (approximately from 300 nm to 800 nm) and longwave radiation (essentially from 4 μm to 100 μm), so that only two wavelength bands are considered, is possible because the solar spectrum has a peak at the normal visible light wavelengths and the earth/atmosphere system emits infrared radiation characteristic of its absolute temperature (the usual range varies approximately from 280K at the surface to 245K at the top of the atmosphere). Furthermore, since there are no other bodies near the earth which could contribute in a significant way, it is possible to consider just those two bands.

1.7 Vertical stability

In the study of the ABL the vertical direction exhibits different features because of the action of gravity, which tends to stratify the fluid. Anyhow, the presence of wind shear in relation to the height and the radiative processes which heat or cool the air near the surface can generate convective turbulent motions, i.e. advective motions which tend to mix the fluid. From now on, 'stability' has to be intended in relation to vertical motions. It is important to introduce some parameters or variables to quantify the degree of the stability of a fluid both for theoretical and experimental reasons, especially in the study of the ABL. The main quantities which are usually considered are potential temperature, the Brunt-Väisälä frequency and the Richardson

number. These quantities are meant to provide information about the state and the evolution of a column of fluid, therefore it is more important their variation in different parts of the fluid rather than their value at a specified point.

Potential temperature is introduced to take into consideration the pressure variations of the environment in which a parcel of air moves adiabatically. It would be intuitive to say that warmer air raises up while colder air sinks down but this is not true in general because if pressure diminishes while the height increases there can be stratification with warmer air beneath colder air. Potential temperature is defined as:

$$\theta = T\left(\frac{P_0}{P}\right)^{\frac{R}{c_P}}$$

where P_0 is a reference pressure (usually 1000 hPa). The variation of the potential temperature with height provides an excellent tool to evaluate vertical stability (see figure 1.4). It is important to say that θ does not take into consideration the presence of water vapor in the air; however, water vapor can be easily treated by introducing the virtual potential temperature θ_V which is associated to the potential temperature in an analogous way to that in which the virtual temperature T_V, which is the temperature that dry air must have to equal the density of moist air at the same pressure, is associated to the absolute temperature T:

$$T_V = T \cdot (1 + 0.61 \cdot r) \qquad \theta_V = \theta \cdot (1 + 0.61 \cdot r)$$

where r is the mixing ratio, i.e. the ratio in a given volume of air between the mass of water vapor and dry air. For saturated (cloudy) air '$0.61 \cdot r$' has to be replaced by '$0.61 \cdot r_{sat} - r_L$' where r_{sat} is the water-vapor saturation mixing ratio and r_L is the liquid-water mixing ratio.

The Brunt-Väisälä frequency N, also known as the buoyancy frequency, is the frequency of oscillation of a parcel displaced vertically in a stratified environment. In fact, the study of the motion of a parcel of density ρ_0 in an environment with a density $\rho(z)$ varying with height leads to the equation for the vertical displacement z' (e.g. Holton, 1972):

$$\frac{\mathrm{D}^2 z'}{\mathrm{D}t^2} - \frac{g}{\rho_0}\frac{\partial \rho(z)}{\partial z} z' = \frac{\mathrm{D}^2 z'}{\mathrm{D}t^2} + N^2 z' = 0 \tag{1.8}$$

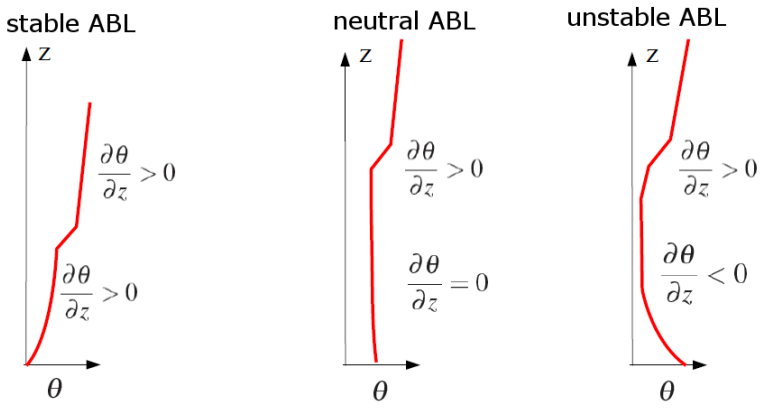


Figure 1.4: The vertical profile of the potential temperature as a criterion for vertical stability. In each case the upper part is stable to testify the presence of the free atmosphere, which is seldom unstable, above the ABL.

Equation 1.8 has the mathematical structure of the equation of an harmonic oscillator and so N can effectively be interpreted as a frequency. If $N^2 > 0$, then a parcel vertically displaced oscillates back towards its starting position, therefore the stratification is stable. Otherwise, if $N^2 < 0$, equation 1.8 has exponential solutions and a parcel displaced vertically is accelerated away from its initial position, therefore there is instability and convective motions can form. N is differently specified in the atmosphere (where it is function of the potential temperature θ profile) and in the ocean (where it is function of the potential density ρ_θ profile):

$$N = \left(\frac{g}{\theta}\frac{\partial\theta}{\partial z}\right)^{\frac{1}{2}} \qquad \text{or} \qquad N = \left(-\frac{g}{\rho_\theta}\frac{\partial\rho_\theta}{\partial z}\right)^{\frac{1}{2}}$$

This new expression for the Brunt-Väisälä frequency connects with the potential temperature and validates the criterion shown in figure 1.4.

In a stable environment, vertical turbulent motions act against gravity. In fact, gravity tends to stratify a fluid while turbulence tends to mix a fluid and homogenize its properties. Therefore the strength of gravity with respect to the strength of turbulence is an indicator of stability. To quantify these observation the Richardson number Ri_f has been introduced. Its definition relies on considerations upon the order of magnitude of the terms appearing in the following equation for the turbulent kinetic energy (denoted $\overline{e}$)

$$\frac{\partial\overline{e}}{\partial t} + \overline{u_i}\frac{\partial\overline{e}}{\partial x_i} = \delta_{i3} g \frac{\overline{u_i'\theta_V'}}{\overline{\theta_V}} - \overline{u_i'u_j'}\frac{\partial\overline{u_i}}{\partial x_j} - \frac{\partial\overline{u_i'e}}{\partial x_i} - \frac{1}{\overline{\rho}}\frac{\partial\overline{u_i'P'}}{\partial x_i} - \varepsilon. \tag{1.9}$$

The Richardson number is defined as the ratio of the first and the second terms on the right side and is usually simplified assuming horizontal homogeneity and neglecting subsidence (in

this case Ri_f is usually referred to as Richardson flux number). From its definition, Ri_f is a dimensionless quantity like Re and provides information about the properties of the flow rather than local properties of the fluid. Moreover it can be further simplified to get the so-called gradient Richardson number Ri which can be related to the Brunt-Väisälä frequency:

$$\mathrm{Ri}_f = \frac{\frac{g}{\overline{\theta_V}}\left(\overline{w'\theta'_V}\right)}{\left(\overline{u'_i v'_j}\right)\frac{\partial \overline{u_i}}{\partial x_j}} \cong \frac{\frac{g}{\overline{\theta_V}}\left(\overline{w'\theta'_V}\right)}{\left(\overline{u'w'}\right)\frac{\partial \overline{u}}{\partial z} + \left(\overline{v'w'}\right)\frac{\partial \overline{v}}{\partial z}} \qquad \mathrm{Ri} = \frac{\frac{g}{\overline{\theta_V}}\frac{\partial \overline{\theta_V}}{\partial z}}{\left(\frac{\partial \overline{u}}{\partial z}\right)^2 + \left(\frac{\partial \overline{v}}{\partial z}\right)^2} = \frac{N^2}{\left(\frac{\partial \overline{u}}{\partial z}\right)^2 + \left(\frac{\partial \overline{v}}{\partial z}\right)^2}$$

Ri is introduced because Ri_f cannot be easily evaluated from measurements in non-turbulent fluxes due to the correlation terms. If Ri$>>1$ gravity dominates and the kinetic energy is not enough to homogenize the fluid that gets stratified. On the other hand, if Ri<0 then N^2 is negative as well, the fluid is unstable and the flow is turbulent. So, high positive values of Ri are associated with laminar flows while negative values of Ri are associated with turbulent flows. Small positive values of Ri characterize a flow that is potentially unstable, i.e. turbulence has not fully developed but some eddies are forming from the mean flow.

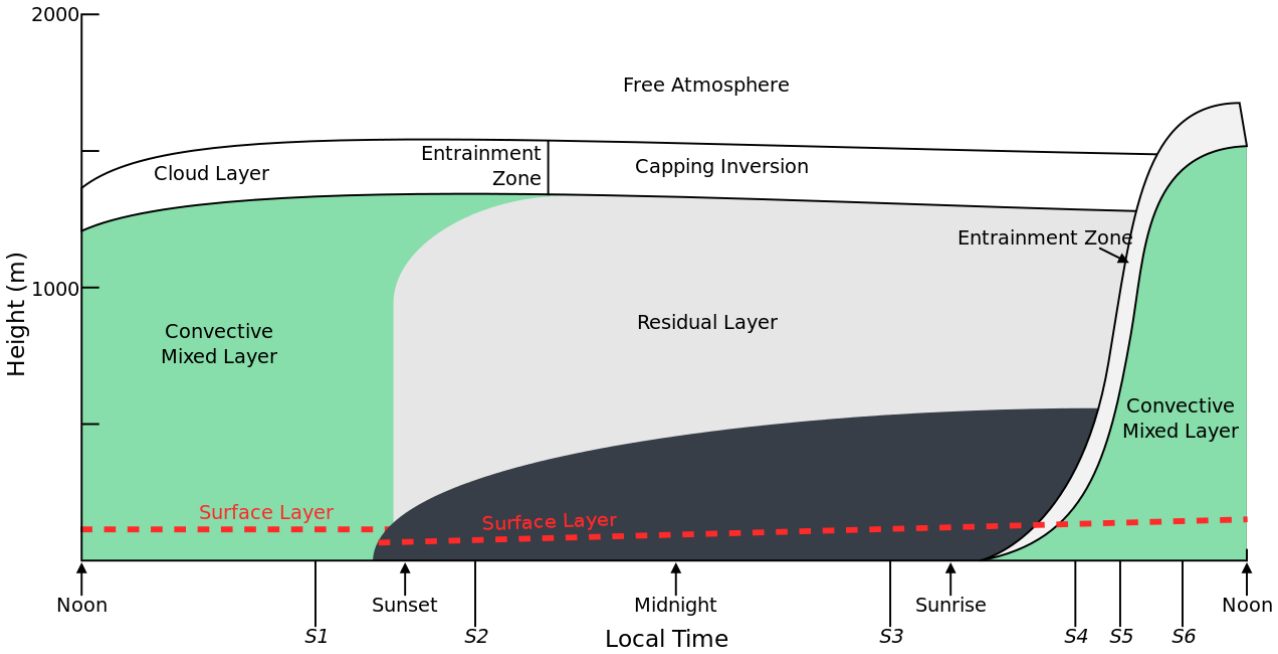


Figure 1.5: A schematic representation of the typical evolution of the ABL. Adapted from Stull, 1988.

1.8 The diurnal evolution of the atmospheric boundary layer

The dynamics of the ABL is quite complicate since the system itself is chaotic and characterized by the presence of many non-linear phenomena. Anyhow, from the phenomenological point of view some periodical aspects can be highlighted. Given the alternation of night and day, without considering the polar regions, a diurnal cycle takes place (see figure 1.5). From sunrise until sundown the Sun radiates energy which heats the air. The balance of incoming and outgoing radiation implies the formation of warmer (colder) air parcels near the surface during daytime (night time). As a consequence during the day parcels tend to float from the surface towards the sky and convective motion take place mixing the air vertically whereas during the night parcels tend to sink and air in the ABL gets stratified.

1.8.1 Daytime evolution

During the day, the heating effect of solar incoming radiation drives the dynamics. Regardless of the presence of clouds, after dawn the air near the ground gets heated so that it gets

lighter and starts to rise. The air which takes the place of the rising parcels will then be heated as well and the process continues giving birth to convection. The formation of convective motions keeps moving air along the vertical direction mixing all the properties of the fluid, i.e. because of this motion the fluid gets homogenized and a column of fluid will tend to exhibit an approximately constant value of the concentration of tracers (e.g. the mixing ratio). This is a behaviour typical of turbulent flows. Schematically, in its daily configuration the ABL can be divided into three parts: the surface layer (SL), the mixed layer (ML) and the entrainment zone (EZ). The SL is the closest part to the surface which includes all those morphological elements which strongly influence the air due to friction and exchanges of properties, e.g. vegetation and urban structures. The SL tends to exhibit a super-adiabatic profile, i.e. temperature diminishes strongly with the height (and so does potential temperature), more than a parcel adiabatically raised would do; this is a feature of instability. Friction at the earth surface usually causes the SL to develop a strong wind shear (see figure 1.6); it has to be kept in mind that ideally the no-slip condition at the surface should hold. The mixed layer, sometimes referred to as the convective layer (CL), is the body of the ABL where the larger convective motions take place. The mixed layer is also the part of the ABL in which the properties are more homogenized, in particular potential temperature and humidity are nearly constant with height (see figure 1.6). The more the surface is heated and energy is given to the air, the more the mixed layer will grow in height. The entrainment zone separates the ML from the free atmosphere (FA). The EZ is characterized by the presence of parcels which sink from the FA and parcels which end their ascent from the ML. In the EZ exchanges of properties between the FA and the ML take place and clouds may form.

The height of the ABL, which can be defined in many ways, e.g. the point at which the Richardson number changes sign or the base of the entrainment zone, can reach values up to 4000 m on a desert during summertime while the usual value at the mid latitudes oscillates around 2000 m. Convective motions start diminishing their intensity in the afternoon and even before twilight they tend to cease because the net flux of energy in the ABL gets negative and the energy lost by radiative transmission is more than that gained from solar radiation. Therefore, as time passes by the daily ABL leaves room to the nocturnal ABL.

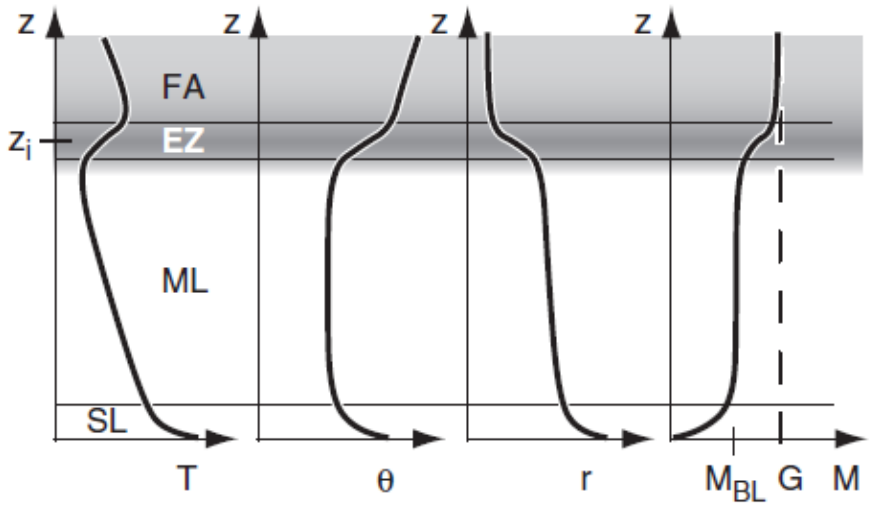


Figure 1.6: The vertical profiles of some physical quantities in the daytime ABL: temperature (T), potential temperature (θ), mixing ratio (r) and wind module (M). The shading delineates statically unstable (white) to very stable (black) layers. The G values indicates the geostrophic wind in the FA. The θ profile testifies a super-adiabatic profile in the SL. Adapted from Stull, 1988.

1.8.2 Night-time evolution

After sunset, the SL rapidly cools and parcels near the surface, which are less dense than the above ones, are prevented from rising up. Turbulence ceases to form in about an hour (recalling figure 1.3, the lifetime of the largest vortices is about half an hour). Consequently, without considering exceptional cases, e.g. nights with strong winds, air tends to stratify and the nocturnal boundary layer (NL) forms. The NL starts forming from the ground "infiltrating" under the daily ABL but cannot reach the heights of the daily ABL so that a well mixed residual layer (RL) remains above the NL, disconnected from the surface (see figure 1.5). The RL is separated at the top from the FA by the so-called capping inversion (CI) which is characterized by a strong inversion and prevents air from the ABL to raise into the FA. The NL is characterized by thermal inversions, i.e. the temperature profile increases with the height near the surface so that cooler air is below warmer air and a stable configuration holds. Typically the NL reaches heights of 200 m to 400 m; over sea during winter time it can get down to 80 m. Turbulence hardly develops, if not for strong wind shears associated with particular synoptic conditions, e.g. the passage of a front, and consequently surface winds at night are much lower and less gusty than during the day. On the other hand, low-level jets (LLJ), known as nocturnal

jets, usually located 100 to 300 m above the ground, may form. LLJ are essentially fast moving masses of air, whose wind speed can reach values up to $20\,\mathrm{m\,s^{-1}}$.

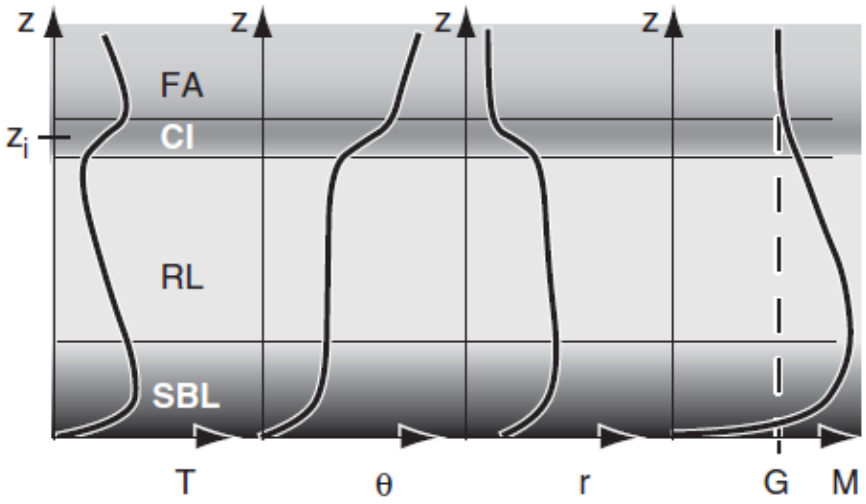


Figure 1.7: The vertical profiles of some physical quantities in the night time ABL: temperature (T), potential temperature (θ), mixing ratio (r) and wind module (M). The shading delineates statically unstable (white) to very stable (black) layers. The G value indicates the geostrophic wind in the FA. The M profile testifies the presence of LLJ. Adapted from Stull, 1988.

Chapter 2

Lagrangian Dispersion Models

2.1 Introduction

When studying the dispersion of pollutants in the atmosphere, one can choose to adopt an Eulerian point of view or, instead, a Lagrangian perspective. The Eulerian approach consists of studying the physical fields of a flux inside a test-volume, fixed in space; in the Lagrangian view, instead, the test-volume moves along with the flux, i.e. it is a parcel of fluid, and the comoving properties of the parcel are investigated. The Lagrangian approach appears far more complicated than the Eulerian one, but when fully developed under the proper modelling assumptions, it proves to be an extremely simple-to-implement and low-cost-CPU-time strategy. In a lagrangian particle model, the dispersion of an airborne pollutant is modeled by "virtual" particles referred to as lagrangian particles, where a particle represents a fixed amount of pollutant mass. The instantaneous pollutant concentration can be determined simply by following the motion of the pollutant particles within a certain spatial domain and then averaging over a proper test-volume.

Section 2.2 presents a brief outline of the Lagrangian theory (see Sozzi, 2003) and sections 2.3 to 2.5 introduce the theory of stochastic processes: in fact it will be clear that, under certain assumptions, the motion of a pollutant lagrangian particle in the atmosphere can be regarded as a Markovian stochastic process. In section 2.6 some atmospherical applications will be briefly discussed (see Sozzi, 2003). All the material in this chapter has been selected from the works of Wio, Rodean and Risken (see Wio et al., 2012, Rodean and Risken, 1996) except when

otherwise specified.

2.2 The lagrangian approach

The equation of a conserved scalar quantity, as the concentration of a pollutant, reads:

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \triangle C + S_C(x, y, z, t) \tag{2.1}$$

where S_C represents a source or sink term. For atmospheric purposes the diffusion term can be neglected, and making the hypothesis of the pollutant to be a non-interacting dust, also the source/sink term can be discarded, and the equation now reads:

$$\frac{\mathrm{D}C}{\mathrm{D}t} = 0 \tag{2.2}$$

This is simply a statement of the conservation properties of a parcel. Such a formula has no practical applications, but highlights the peculiarity of the lagrangian approach: at the core of the lagrangian view stands the trajectory.

In the lagrangian approach the atmosphere is modeled as a huge set of particles; each particle interacts with the other ones and the history of such interactions determines the trajectory of a particle. Following the motion of every particle is a hard task and so it is convenient to adopt a statistical view of the phenomenon, as it happens in the kinetic theory of gases.

Let $\psi(\mathbf{x}, t)d\mathbf{x}$ be the probability that a particle is found in the small volume $d\mathbf{x}$ around position $\mathbf{x}$. So $\psi(\mathbf{x}, t)$ is the associated probability density function and it must satisfy the normalization condition:

$$\int_{\mathbf{R}^3} \psi(\mathbf{x}', t)\, d\mathbf{x}' = 1 \tag{2.3}$$

By definition of a quantity called *transition probability* $Q(\mathbf{x}, t|\mathbf{x}', t')$, which defines the probability that a particle at position $\mathbf{x}'$ at time t' is found at $t > t'$ at the new position $\mathbf{x}$, it is possible to state the following recursive relation for ψ:

$$\psi(\mathbf{x}, t) = \int_{\mathbf{R}^3} Q(\mathbf{x}, t|\mathbf{x}', t')\psi(\mathbf{x}', t')\, d\mathbf{x}' \tag{2.4}$$

If there are m particles, the average concentration in $\mathbf{x}$ is:

$$\langle C(\mathbf{x},t)\rangle = \sum_{i=1}^{m} \psi_i(\mathbf{x},t) \tag{2.5}$$

Applying the summation to Eq (2.4) and considering a source that creates new particles for $t > t_0$ one gets:

$$\langle C(\mathbf{x},t)\rangle = \int_{\mathbf{R}^3} Q(\mathbf{x},t|\mathbf{x}_0,t_0)\langle C(\mathbf{x}_0,t_0)\rangle \, d\mathbf{x}_0 + \int_{\mathbf{R}^3}\int_{t_0}^{t} Q(\mathbf{x},t|\mathbf{x}',t')\langle S(\mathbf{x}',t')\rangle \, d\mathbf{x}' \, dt' \tag{2.6}$$

Eq (2.6) is the statement of the Lagrangian approach in the case of non-reactive particles. Given the initial concentration $\langle C(\mathbf{x_0},t_0)\rangle$ and the emission function $S(\mathbf{x},t)$, the concentration at a time t and at a point $\mathbf{x}$ is determined by the transition probability Q. Thus the problem is reduced to the determination of Q and that is precisely the aim of the Lagrangian models. In general the transition probability is build up using the statistical knowledge on atmospheric turbulence: the more accurate is this knowledge the better will be the capability to portray diffusion processes.

2.3 Stochastic processes

Once a stochastic variable X has been defined, an infinity of other stochastic variables derive from it, namely, all the quantities Y defined as functions of X by some mapping. These quantities Y may be any kind of mathematical object; in particular, also functions of an auxiliary variable t:

$$Y_X(t) = f(X,t),$$

where t could be the time or some other parameter. Such a quantity $Y_X(t)$ is called a *stochastic process*. On inserting for X one of its possible values x, one gets an ordinary function of t, $Y_x(t) = f(x,t)$, called a *sample function* or *realisation* of the process. In physical language, one regards the stochastic process as the "ensemble" of these sample functions.

It is easy to form averages on the basis of the underlying probability density function $P_X(x)$

(PDF). For instance, one can take n values $t_1, ..., t_n$ and form the nth moment:

$$\langle Y(t_1) \cdots Y(t_n) \rangle = \int Y_x(t_1) \cdots Y_x(t_n) P_X(x) dx \tag{2.7}$$

A stochastic process is called *stationary* when the moments are not affected by a shift in time, that is, when

$$\langle Y(t_1 + \tau) \cdots Y(t_n + \tau) \rangle = \langle Y(t_1) \cdots Y(t_n) \rangle. \tag{2.8}$$

In particular, $\langle Y(t) \rangle$ is independent of the time, and the auto-correlation function $\mathcal{K}(t_1, t_2) = cov[Y(t_1), Y(t_2)]$ only depends on the time difference $|t_1 - t_2|$. Often there exist a constant τ_c such that $\mathcal{K}(t_1, t_2) \simeq 0$ for $|t_1 - t_2| > \tau_c$; one then calls τ_c the *auto-correlation time* of the stationary stochastic process.

2.3.1 The hierarchy of probability density functions

The probability distribution for $Y_X(t)$ to take the value y at time t is

$$P_1(y, t) = \int \delta[y - Y_x(t)] P_X(x) dx. \tag{2.9}$$

Similarly, the joint probability distribution that Y has the value y_1 at t_1, and also the value y_2 at t_2, and so on, up to t_n, is

$$P_n(y_1, t_1; \ldots; y_n, t_n) = \int \delta[y_1 - Y_x(t_1)] \cdots \delta[y_n - Y_x(t_n)] P_X(x) dx. \tag{2.10}$$

In this way an infinite hierarchy of probability distributions $P_n, n = 1, 2, \ldots$ is defined. They allow the computation of all the averages already introduced, e.g.,

$$\langle Y(t_1) \cdots Y(t_n) \rangle = \int y_1 \cdots y_n P_n(y_1, t_1; \ldots; y_n, t_n) dy_1 \cdots dy_n \tag{2.11}$$

The hierarchy of probability distributions P_n then obeys the following consistency conditions:

1. $P_n \geq 0$;

2. P_n is invariant under permutations of two pairs (y_i, t_i) and (y_j, t_j);

3. $\int P_n(y_1, t_1; \ldots; y_n, t_n) dy_n = P_{n-1}(y_1, t_1; \ldots; y_{n-1}, t_{n-1})$;

4. $\int P_1(y_1, t_1) dy_1 = 1$.

It is possible to demonstrate that a certain PDF describes a stochastic process if and only if it satisfies the four conditions aforementioned (Kolmogorov).

2.3.2 Conditional probabilities

The notion of conditional probability for multivariate distributions can be applied to stochastic processes, via the hierarchy of probability distributions introduced above. For instance, the conditional probability $P_{1|1}(y_2, t_2|y_1, t_1)$ represents the probability that Y takes the value y_2 at t_2, given that its value at t_1 "was" y_1. It can be constructed as follows: from all sample functions $Y_x(t)$ of the ensemble representing the stochastic process, select those passing through the point y_1 at the time t_1; the fraction of this *sub-ensemble* that goes through the gate $(y_2, y_2 + dy_2)$ at the time t_2 is precisely $P_{1|1}(y_2, t_2|y_1, t_1)dy_2$. More generally, one may fix the values of Y at n different times $t_1, \ldots, t_n$ and ask for the joint probability at m other times $t_{n+1}, \ldots, t_{n+m}$. This leads to the general definition of the conditional probability $P_{m|n}$ by Bayes' rule:

$$P_{m|n}(y_{n+1}, t_{n+1}; \ldots; y_{n+m}, t_{n+m}|y_1, t_1; \ldots; y_n, t_n) = \frac{P_{n+m}(y_1, t_1; \ldots; y_{n+m}, t_{n+m})}{P_n(y_1, t_1; \ldots; y_n, t_n)} \tag{2.12}$$

2.4 Markov processes

Among the many possible classes of stochastic processes, there is one that merits a special treatment: the so-called Markov process.

Recall that, for a stochastic process $Y(t)$, the conditional probability $P_{1|1}(y_2, t_2|y_1, t_1)$, is the probability that $Y(t_2)$ takes the value y_2, provided $Y(t_1)$ has taken the value y_1. In terms of this quantity one can express P_2 as

$$P_2(y_1, t_1; y_2, t_2) = P_1(y_1, t_1) P_{1|1}(y_2, t_2|y_1, t_1). \tag{2.13}$$

However, to construct the higher-order P_n one needs transition probabilities $P_{n|m}$ of higher order, e.g. $P_3(y_1,t_1;y_2,t_2;y_3,t_3) =$
$= P_2(y_1,t_1;y_2,t_2)P_{1|2}(y_3,t_3|y_1,t_1;y_2,t_2)$. A stochastic process is called a *Markov process*, if for any set of n successive times $t_1 < t_2 < \cdots < t_n$, one has

$$P_{1|n-1}(y_n,t_n|y_1,t_1;\ldots;y_{n-1},t_{n-1}) = P_{1|1}(y_n,t_n|y_{n-1},t_{n-1}). \tag{2.14}$$

In words: the conditional probability distribution of y_n at t_n, given the value y_{n-1} at t_{n-1}, is uniquely determined, and is not affected by any knowledge of the values at earlier times.
A markov process is therefore fully determined by the two PDFs $P_1(y,t)$ and $P_{1|1}(y',t'|y,t)$, from which the entire hierarchy $P_n(y_1,t_1;\ldots;y_n,t_n)$ can be constructed. For instance, consider $t_1 < t_2 < t_3$; P_3 can be written as

$$\begin{aligned} P_3(y_1,t_1;y_2,t_2;y_3,t_3) =& P_2(y_1,t_1;y_2,t_2)P_{1|2}(y_3,t_3|y_1,t_1;y_2,t_2) \\ =& P_2(y_1,t_1;y_2,t_2)P_{1|1}(y_3,t_3|y_2,t_2) \\ =& P_1(y_1,t_1)P_{1|1}(y_2,t_2|y_1,t_1)P_{1|1}(y_3,t_3|y_2,t_2) \end{aligned} \tag{2.15}$$

From now on only Markov processes will be treated and $P_{1|1}(y',t'|y,t)$ will be addressed as the *transition probability*. Subscript "1|1" is no longer necessary and thus will be omitted in the following.

2.4.1 The Chapman-Kolmogorov equation

On integrating Eq (2.15) over y_2, one obtains ($t_1 < t_2 < t_3$)

$$P_2(y_1,t_1;y_3,t_3) = P_1(y_1,t_1)\int P(y_2,t_2|y_1,t_1)P(y_3,t_3|y_2,t_2)dy_2, \tag{2.16}$$

where the consistency condition 3 of the hierarchy of distribution functions P_n has been used to write the left-hand side. Now, on dividing both sides by $P_1(y_1,t_1)$ and using the special case

of Bayes' rule Eq (2.13), one gets:

$$P(y_3, t_3; y_1, t_1) = \int P(y_2, t_2|y_1, t_1) P(y_3, t_3|y_2, t_2) dy_2, \tag{2.17}$$

which is called the *Chapman-Kolmogorov equation*. Time ordering is essential.

2.4.2 The Master equation

The master equation is a differential equation for the transition probability. Accordingly, in order to derive it, one needs first to ascertain how the transition probability behaves for short time differences.

First, on inspecting the Chapman-Kolmogorov equation for equal time arguments one finds the natural result

$$P(y_3, t_3; y_1, t_1) = \int P(y_2, t|y_1, t) P(y_3, t_3|y_2, t) dy_2 \tag{2.18}$$

$\Rightarrow P(y_2, t|y_1, t) = \delta(y_2 - y_1)$,

which is the zeroth-order term in the short time behaviour of $P(\cdot|\cdot)$. Keeping this in mind one *adopts* the following expression for the short-time transition probability:

$$P(y_2, t + \Delta t|y_1, t) = \delta(y_2 - y_1)[1 - a^{(0)}(y_1, t)\Delta t] + W_t(y_2|y_1)\Delta t + \mathcal{O}\big((\Delta t)^2\big), \tag{2.19}$$

where W_t is interpreted as the *transition probability per unit time* from y_1 to y_2 at time t. Then, the coefficient $1 - a^{(0)}(y_1, t)\Delta t$ is to be interpreted as the probability that no "transition" takes place during Δt. Indeed, from the normalization of $P(y_2, t_2|y_1, t_1)$ one has:

$$1 = \int P(y_2, t + \Delta t|y_1, t) \simeq 1 - a^{(0)}(y_1, t)\Delta t + \int W_t(y_2|y_1)\Delta t \, dy_2. \tag{2.20}$$

Therefore, to first order in Δt, one gets

$$a^{(0)}(y_1, t) = \int W_t(y_2|y_1) dy_2 \tag{2.21}$$

which substantiates the interpretation mentioned: $a^{(0)}(y_1,t)\Delta t$ is the total probability of escape from y_1 in the time interval $(t, t+\Delta t)$ and, thus, $1 - a^{(0)}(y_1,t)\Delta t$ is the probability that no transition takes place during this time.

Insertion of the above short-time expansion for the transition probability into the Chapman-Kolmogorov equation, yields

$$\begin{aligned} P(y_3,t_2+\Delta t|y_1,t_1) \simeq [1 - a^{(0)}(y_3,t_2)\Delta t]P(y_3,t_2|y_1,t_1) \\ + \Delta t \int W_{t_2}(y_3|y_2)P(y_2,t_2|y_1,t_1)dy_2. \end{aligned} \tag{2.22}$$

Next, on using Eq (2.21) to write $a^{(0)}(y_3,t_2)$ in terms of $W_{t_2}(y_2|y_3)$, one has:

$$\begin{aligned} &\frac{1}{\Delta t}[P(y_3,t_2+\Delta t|y_1,t_1) - P(y_3,t_2|y_1,t_1)] \\ &\quad \simeq \int [W_{t_2}(y_3|y_2)P(y_2,t_2|y_1,t_1) \\ &\quad - W_{t_2}(y_2|y_3)P(y_3,t_2|y_1,t_1)]dy_2. \end{aligned} \tag{2.23}$$

which in the limit $\Delta t \to 0$ yields, after some changes in notation ($y_1,t_1 \longrightarrow y_0,t_0$; $y_2,t_2 \longrightarrow y',t$ and $y_3 \longrightarrow y$), the *master equation*

$$\frac{\partial}{\partial t}P(y,t|y_0,t_0) = \int [W_t(y|y')P(y',t|y_0,t_0) - W_t(y'|y)P(y,t|y_0,t_0)]dy'. \tag{2.24}$$

which is an integro-differential equation. The master equation serves to determine the time evolution of the system over long time periods, at the expense of *assuming* the Markov property.

2.4.3 The Kramers-Moyal expansion and the Fokker-Planck equation

The dependencies of the transition probability rate W_t can be reformulated as follows:

$$W(y|y') = W(y';r), \qquad r = y - y'. \tag{2.25}$$

Assuming that the changes on y occur via small jumps, i.e. that $W(y';r)$ is a sharply peaked function of r but varies slowly enough with y', and that $P(y,t)$ itself varies slowly with y, it is

possible to Taylor-expand the term $W(y-r;r)$ on the right-hand side of the master equation. Making the additional assumption of dealing with problems to which the boundary is irrelevant (integration extends to infinity) one finally obtains the *Kramers-Moyal expansion of the master equation*:

$$\frac{\partial}{\partial t}P(y,t) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \frac{\partial^m}{\partial y^m}[a^{(m)}(y,t)P(y,t)] \tag{2.26}$$

where $a^{(m)}(y,t)$ is the mth *jump moment* [1]:

$$a^{(m)}(y,t) = \int r^m W(y;r)dr \tag{2.27}$$

The Kramers-Moyal expansion has recast an integro-differential equation (the master equation) into a infinite-order differential equation (Eq (2.24)). Formally the two equations are identical and is therefore not easier to deal with Eq (2.24), but that form suggests that one may break off after a suitable number of terms. For instance, there could be situations where, for $m > 2$, $a^{(m)}(y,t)$ is identically zero or negligible. In this case one is left with

$$\frac{\partial}{\partial t}P(y,t) = -\frac{\partial}{\partial y}[a^{(1)}(y,t)P(y,t)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[a^{(2)}(y,t)P(y,t)] \tag{2.28}$$

which is the *Fokker-Planck equation*. The first term is called the *drift* or *transport* term and the second one the *diffusion* term, while $a^{(1)}(y,t)$ and $a^{(2)}(y,t)$ are the drift and diffusion "coefficients".

2.5 The Langevin equation

In this section a different approach is adopted; a particular form of the differential equation for the motion of the pollutant particles is examined: the Langevin equation, named after the French physician Paul Langevin who first proposed it to describe Brownian motion. It is a stochastic differential equation, whose behaviour is fully and rigorously explained in the framework of the Itō calculus.

[1] one can alternatively express the jump moments as $\lim_{\Delta t \to 0}\langle [Y(t+\Delta t) - Y(t)]^m \rangle|_{Y(t)=y}$

The Langevin equation for one variable is a "differential equation" of the form

$$dy = A(y,t)\,dt + B(y,t)\xi(t)\,dt, \tag{2.29}$$

where $\xi(t)$ is a given stochastic process. The choice for $\xi(t)$ that renders $y(t)$ a Markov process is that of the "Langevin process" (white noise), which is Gaussian and its statistical properties are

$$\begin{aligned} \langle \xi(t) \rangle &= 0, \\ \langle \xi(t_1)\xi(t_2) \rangle &= 2D\delta(t_1 - t_2). \end{aligned} \tag{2.30}$$

Since Eq (2.29) is a *first order* differential equation, for each sample function (realization) of $\xi(t)$, it determines $y(t)$ uniquely when $y(t_0)$ is given. In addition, the values of the fluctuating term at different times are statistically independent, due to the delta-correlated nature of $\xi(t)$. Therefore, the values of $\xi(t)$ at previous times, say $t' < t_0$, cannot influence the conditional probabilities at times $t > t_0$. From these arguments it follows the Markovian character of the solution of the Langevin equation.

The terms $A(y,t)$ and $B(y,t)\xi(t)$ are often referred to as the *drift* and *diffusion* terms, respectively. To solve a Langevin equation means to determine the statistical properties of the process $y(t)$.

2.5.1 The Fokker-Planck equation for the Langevin Equation

Here it is shown how the Langevin equation can be recast in the form of a Fokker-Planck eqaution. First, cast the Langevin differential equation into the form of an integral equation:

$$y(t+\Delta t) - y = \int_t^{t+\Delta t} A[y(t_1), t_1]\,dt_1 + \int_t^{t+\Delta t} B[y(t_1), t_1]\xi(t_1)\,dt_1 \tag{2.31}$$

where y stands for the initial value $y(t)$. Then, Taylor-expand A and B terms, as follows:

$$\begin{aligned} A[y(t_1), t_1] &= A(y, t_1) + A'(y, t_1)[y(t_1) - y] + \ldots \\ B[y(t_1), t_1] &= B(y, t_1) + B'(y, t_1)[y(t_1) - y] + \ldots \end{aligned}$$

After substitution of the above series expansion in Eq (2.31) one obtains $y(t+\Delta t)-y$ as a function of integral terms, some of which contain the dependency on $y(t_1)-y$. Iteration of the same integral expansion on these terms and average of both sides of the equation yields:

$$\langle y(t+\Delta t)-y\rangle = \int_t^{t+\Delta t} A[y,t_1]\,dt_1 + \int_t^{t+\Delta t} A'[y,t_1]\left(\int_t^{t_1} A(y,t_2)\,dt_2\right)dt_1 + 2D\int_t^{t+\Delta t} B'[y,t_1]\left(\int_t^{t_1} B(y,t_2)\delta(t_2-t_2)\,dt_2\right)dt_1+\ldots \quad (2.32)$$

Next, applying the properties of the Dirac delta and performing the limit for $\Delta t \longrightarrow 0$ one finally obtains

$$a^{(1)}(y,t) = A(y,t) + DB(y,t)\frac{\partial B(y,t)}{\partial y}. \quad (2.33)$$

Other integrals not written down in the above formulae do not contribute in the limit $\Delta t \longrightarrow 0$. On using the same type of arguments to identify some vanishing integrals one can compute the second coefficient in the Kramers-Moyal expansion, obtaining

$$a^{(2)}(y,t) = 2DB^2(y,t), \quad (2.34)$$

whereas all the coefficients $a^{(m)}$ vanish for $m \geq 3$. Thus, the final results are:

$$\begin{aligned} a^{(1)}(y,t) &= A(y,t) + DB(y,t)\frac{\partial B(y,t)}{\partial y}, \\ a^{(2)}(y,t) &= 2DB^2(y,t), \\ a^{(m)}(y,t) &= 0, \quad \text{for } m \geq 3 \end{aligned} \quad (2.35)$$

and the Fokker-Planck equation for the Langevin equation is

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial y}\left[\left(A(y,t) + DB(y,t)\frac{\partial B(y,t)}{\partial y}\right)P\right] + D\frac{\partial^2}{\partial y^2}[B^2(y,t)P]. \quad (2.36)$$

It is quite easy to extend all the above derivation to the multi-variable case.

2.5.2 Final considerations

Under all the mathematical passages explained in the preceding sections lies the complementary dichotomy of the Lagrangian and Eulerian approaches. Langevin equation is focused on the motion of a particular particle, and so it could be regarded as a Lagrangian statement of the physical system; the associated Fokker-Planck equation, instead, deals with the PDF, i.e. it is focused on a particular point of the phase space whose values can be assumed by *any* particle of the physical system. In other words, the Langevin equation has an individualistic (Lagrangian) character while the Fokker-Planck reformulation is a collective description of the system (Eulerian). The two approaches are not completely equivalent, as it will be clear in the following, and one needs both points of view in order to develop a complete model of the atmospheric dispersion.

2.6 Atmospheric models

The situation is summed up in the following four points:

1. Langevin equation gives the evolution of a Markovian field, but does the velocity of an atmospheric particle represent a Markovian process? In the atmosphere, viscous forces act on timescales of the order of the Kolmogorov timescale, that is $\tau_\eta \simeq 10^{-2}$ s, for high Reynolds number flows. So a Lagrangian discrete model of order 1 (i.e. velocity is a Markovian variable) could be adopted if the sampling time is greater than the Kolmogorov timescale.

2. The ideal model that should be used to describe the particle motion (1D) is the stochastic differential system (derived from the Langevin equation):

$$\begin{aligned} dw &= A_w(z,t)dt + B_w(z,t)\xi(t)dt \\ dz &= wdt \end{aligned} \tag{2.37}$$

where w is the velocity. Similar relations hold for the motion in the other directions. The

coefficients A and B are unknown.

3. The same coefficients appear in the associated Fokker-Planck equation.

4. In the Fokker-Planck equation also appears the PDF ($P(z, w, t)$) that represents the statistical properties of the motion of the pollutant particles; such properties are not accessible.

Actually, the available statistical properties are those related to the turbulent motion of the *air* in which the pollutant particles are dispersed. It would be desirable to find a link between the statistics of the fluid particles and that of the pollutant particles. That is precisely what the *Well Mixed Condition* stands for.

2.6.1 The well mixed condition

Invoked when dealing with particle tracers with the same properties of the fluid particles in which they are immersed, the Well Mixed Condition requires that, if at any time the particles are homogeneously distributed within the fluid, in average, then for any subsequent time they remain homogeneously distributed. It is nothing but a closure criterion, expressed by the relation

$$C(z, w, t) = \frac{P_{poll}(z, w, t)}{P_{fluid}(z, w, t)} = const. \tag{2.38}$$

where P_{poll} and P_{fluid} are the pollutant and fluid PDFs respectively.

Using Eq (2.38) and expressing the system in Eq (2.37) as a multi-variable Fokker-Planck equation one obtains:

$$\frac{\partial P_{fluid}}{\partial t} + \frac{\partial}{\partial z}(wP_{fluid}) = -\frac{\partial}{\partial w}(A_w P_{fluid}) + \frac{1}{2}\frac{\partial^2}{\partial w^2}({B_w}^2 P_{fluid}). \tag{2.39}$$

The strategy now is to use system Eq (2.37) to describe the particle motion and to exploit Eq (2.39) to determine the unknown coefficients A_w and B_w. It is important to stress that the vertical direction has a privileged role; that is because of gravity, which affects vertical motions,

and of significant temperature gradients, mainly oriented along the z axis. Thus, in general, statistical description of turbulence along vertical and horizontal directions will be different.

2.6.2 Determination of the drift and diffusion coefficients

From the theory of Kolmogorov it is known that the statistic of a fluid in the inertial subrange is independent from its dynamic viscosity ν; moreover, it depends *only* on the kinetic energy dissipation rate ε. Monin and Yaglom (1975) demonstrated that the Lagrangian velocity structure function $D(dt)$ does admit a similarity relation:

$$D(dt) = dw^2 = \langle [w(t+dt) - w(t)]^2 \rangle = C_0 \varepsilon dt \tag{2.40}$$

where C_0 is a universal constant whose value is approximately 2. The relation 2.40 holds for $\tau_\eta << dt << T^L$, where τ_η is the Kolmogorov time scale and T^L is the integral Lagrangian length scale. From the Langevin equation one also has [2]

$$D(dt) = dw^2 = B_w^2 dt, \tag{2.41}$$

and comparison of Eq (2.40) and Eq (2.41) gives the estimation of the *diffusion coefficient*:

$$B_w = \sqrt{C_0 \varepsilon} \tag{2.42}$$

Determination of the *drift* coefficient is not so straight-forward. To achieve this task one has to employ the Fokker-Planck equation i.e. the statistical properties of the fluid, condensed in the probability density function. Statistical knowledge of the PBL fluxes could be provided by a numerical model or by some similarity relations. Here only two cases (the most relevant ones) will be examined, namely:

1. a turbulence well described by a *gaussian* PDF, representative of stable, adiabatic situations;

[2] Proof: by comparison with the jump moment $D(dt) = a^{(2)}(y,t)$ of the associated Fokker-Planck equation, or by direct computation of dw^2 in the Langevin equation.

2. a *non-gaussian* turbulence, which describes strongly convective situations.

Gaussian Turbulence

By choosing a Gaussian PDF for w in the Fokker-Planck equation and resolving for A_w one gets

$$A_w(z,w,t) = -\left(\frac{C_0\varepsilon}{2\sigma_w^2}\right) w + \frac{\phi}{P_{fluid}} \tag{2.43}$$

where

$$\frac{\phi}{P_{fluid}} = \frac{1}{2}\frac{\partial\sigma_w^2}{\partial z} + \frac{1}{2\sigma_w^2}\left(\frac{\partial\sigma_w^2}{\partial t}\right) w + \frac{1}{2\sigma_w^2}\left(\frac{\partial\sigma_w^2}{\partial z}\right) w^2. \tag{2.44}$$

Both w and σ_w can vary with height and time. In many practical cases the hypothesis of stationarity holds and σ_w^2 does not depend on time any longer, thus yielding

$$A_w(z,w,t) = -\left(\frac{C_0\varepsilon}{2\sigma_w^2}\right) w + \frac{1}{2}\left[1 + \left(\frac{w}{\sigma_w}\right)^2\right]\frac{\partial\sigma_w^2}{\partial z}. \tag{2.45}$$

Non-Gaussian Turbulence

Hurley and Physick (1993) observed that in convective situations the *Mixed Layer* occupies a very large percent in portion of the PBL and invoked the assumption of homogeneity for the entire PBL. Zones in which there still are vertical gradients are the *Surface Layer* and the *Entrainment Layer* but neglection of these thin portions does not introduce a significant error. So the hypothesis of homogeneity applies and the drift coefficient takes the following form:

$$A_w(z,w,t) = -\left(\frac{C_0\varepsilon}{2\sigma_w^2}\right) \cdot \frac{A_1 \cdot N(m_1,\sigma_1)/\sigma_1^2 + A_2 \cdot N(m_2,\sigma_2) \cdot (w-m_2)/\sigma_2^2}{A_1 \cdot N(m_1,\sigma_1) + A_2 \cdot N(m_2,\sigma_2)}. \tag{2.46}$$

In this relation the functions $N(m_i,\sigma_i)$ are gaussians with expectation value m_i and variance σ_i^2. The coefficients A_1 and A_2 are the areas occupied by the updrafts and the downdrafts respectively while m_i and σ_i are the associated velocity and standard deviation respectively. The parameters appearing in Eq (2.46) can be determined by similarity relations which depend on the *Skewness* of the PDF.

Chapter 3

Models Description

3.1 Introduction

The two models used in this thesis to describe pollutant dispersion in the atmosphere are products of ARIA TECHNOLOGIES SA and ARIANET Srl. One model, microSWIFT, computes high-resolution atmospherical fields (wind speed, temperature, etc); The other model, microSPRAY, is a Lagrangian Stochastic model which simulates the dispersion of the pollutant particles. microSPRAY inputs are the atmospheric fields generated by microSWIFT. Together they form the micrometeorological suite MSS.

Time evolution of the model is run by an external meteorological model.

3.2 The microSWIFT model

microSWIFT is a diagnostic model which provides atmospheric fields at a typical grid spacing of about 5 m at local (radius of 5-50 km) and regional (radius of 50-500 km) scales. The main parameters at these scales are:

- local wind (near ground and up to 1000 m above ground)
- topography and terrain roughness
- atmospheric turbulence

- temperature, humidity

microSWIFT is designed to rapidly compute wind fields from on-site observations or simulated data (at least one point-like site measurement and one sounding). These comply with the first Navier-Stokes equation, the mass conservation, to account for terrain effect on the flow structure. The influence of atmospheric stability on wind flow over terrain is modeled using a weighting factor α (ratio of the horizontal wind component to the vertical wind component). Similarity theory plays a fundamental role in interpolating surface layer quantities.

If obstacles such as buildings are included in a local scale simulation, their influence is modeled using a first guess prescription on the flow structure, then mass consistency and impermeability are applied.

As the other Navier-Stokes equations have not been confirmed, energy and momentum may not be conserved.

3.2.1 Space grid

Wind fields are computed on a 3D space grid in a geometrical domain with a rectangular base. The x-axis lies in a West-East direction and the y-axis lies in a South-North direction. For the horizontal levels, the grid uses a Cartesian coordinates system, with fixed cell sizes. The vertical grid levels are expressed in "sigma" coordinates (terrain following near the ground, flat near the top of the domain); the vertical coordinate z^* is related to the corresponding Cartesian coordinate z by the following equation:

$$z^* = H\frac{z - z_g}{H - z_g} \tag{3.1}$$

where H is the altitude at domain top (in m), $z_g = z_g(x, y)$ is the ground elevation (in m) at x, y points on the horizontal grid and z is the actual altitude (in m above the sea level) of the current grid point. The value $h(x, y) = H - z_g(x, y)$ represents the thickness of the atmospheric layer between the ground and domain top (elevation H). Points on the vertical grid can be evenly spaced in order to give a better detail in particular regions, such as close to the ground, where vertical gradients generally achieve highest values.

3.2.2 Interpolation of data

Three types of interpolation procedures can be defined for the above data categories and grid types:

- interpolation of surface network data (first grid level above the ground)
- interpolation of the upper-air profiles by grid level; these procedures are associated with interpolation of surface network data
- 3D interpolation using data from both categories (surface stations and upper-air profiles), without special processing of surface data.

Selecting the right method requires knowledge of network geometry. In most cases, upper-air data for a given site are likely to be fewer and less frequent than surface data.

Interpolation of wind field

The values at the measurement points must first be corrected to describe effective grid point quantities.

Wind field vertical correction

Let $h_m(x, y)$ be the above-ground height of the *first* microSWIFT level above the ground. Elevation varies with: the horizontal position, altitude at domain top (H) and the value of z^* of the *first* microSWIFT level. The measured value of the horizontal module $U(M)$ at point $M(x, y, h_s)$ can be corrected and converted to $U_C(M)$ (corrected wind), by the equations

$$\begin{aligned} U_C(h_m) &= u^* F\left(\frac{h_m}{z_0}, \frac{h_m}{L_{mo}}\right), \\ U(h_s) &= u^* F\left(\frac{h_s}{z_0}, \frac{h_s}{L_{mo}}\right), \end{aligned} \tag{3.2}$$

leading to

$$U_C(h_m) = U(h_s)\frac{F\left(\frac{h_m}{z_0}, \frac{h_m}{L_{mo}}\right)}{F\left(\frac{h_s}{z_0}, \frac{h_s}{L_{mo}}\right)} \tag{3.3}$$

where u^* is the friction velocity, F is the universal function for the surface layer and L_{mo} is the Monin-Obukhov length.

Wind field horizontal correction

The underlying idea is that horizontal speeds measured at grid points located at $M_k(x_k, y_k)$ on the horizontal grid of the surface layer are influenced by the local roughness length $z_0(x_k, y_k)$, even after being assigned to the first upper-air grid level $h_m(x_k, y_k)$. The method can be summarized in three operations:

1. "uninfluence" via a z_{red} (a uniform arbitrary value for the roughness length),

$$U_{red} = U_C\frac{F\left(\frac{h_m(x_k,y_k)}{z_{red}}, \frac{h_m(x_k,y_k)}{L_{mo}}\right)}{F\left(\frac{h_m(x_k,y_k)}{z_0(x_k,y_k)}, \frac{h_m(x_k,y_k)}{L_{mo}(x_k,y_k)}\right)} \tag{3.4}$$

2. 2D interpolation ($U_{red} \longrightarrow U_{int}$),

3. "influence" via $z_0(x_k, y_k)$

$$U_f = U_{int}\frac{F\left(\frac{h_m(x_k,y_k)}{z_0(x_k,y_k)}, \frac{h_m(x_k,y_k)}{L_{mo}}\right)}{F\left(\frac{h_m(x_k,y_k)}{z_{red}}, \frac{h_m(x_k,y_k)}{L_{mo}}\right)}. \tag{3.5}$$

It is thus also known as the "influence-uninfluence" method. The universal function F must be specified by the user between the following possibilities:

- Businger-Dyer
- Businger-Dyer for stable atmosphere and Louis for unstable atmosphere

- Louis for stable and unstable conditions.

An analogous principle is adopted when dealing with topography: in that case the influence function is

$$h(x, y) = H - z_g(x, y) \tag{3.6}$$

instead of

$$F\left(\frac{h_m(x_k, y_k)}{z_0(x_k, y_k)}, \frac{h_m(x_k, y_k)}{L_{mo}}\right). \tag{3.7}$$

Interpolation for low resolution wind profiles

With regard to profiles, the associated first microSWIFT-above-ground level (or more generally the associated first lower levels) is determined from the first two profile measurements. Options for this passage are:

- natural interpolation in $\log\left(\frac{h_2}{h_1}\right)$ for all $h \leq h_2$,
- same method except for cases where $U_1 > U_2$, for which $U(h) = U_2$,
- interpolation in $\log\left(\frac{h_1}{z_0}\right)$ for all $h \leq h_1$,
- interpolation in $\log\left(\frac{h_2}{z_0}\right)$ for all $h \leq h_2$.

Wind direction is obtained through linear interpolation of the angles associated to the two wind vectors at h_1, h_2 by minimizing the rotation between them.

Interpolation of surface wind data

After having corrected surface data one can proceed with 2D interpolation. The following possibilities are available:

- 2D Cressman method. It is based on a weighted interpolation of the measurements according to the distance of the selected grid point from the ground station. The interpolated i-th component of the wind at a point with (x, y) coordinates satisfies the following equation

$$U_i(x,y) = \frac{\sum_{k=1}^{N} U_i(x_k, y_k) P_k(x,y)}{\sum_{k=1}^{N} P_k(x,y)} \tag{3.8}$$

where N is the number of surface station considered, $P_k(x,y) = 1/r_k^2$ is the weighting function (with $r_k = R_k/P_h$, R_k distance of the grid point (x,y) from the station point (x_k, y_k) and P_h horizontal range parameter for the selected data).

- MacLain method. The horizontal plane is overlaid with disjointed triangles, whose vertices are the points of the surface data network. In order to make the triangular mesh cover the entire domain, a set of fictitious stations is created. Data values at these points are obtained by means of a Cressman interpolation with high influence radius P_k. Then all grid points are automatically contained in one of the triangles as they overlay the entire domain. The interpolated value at a grid point is computed using the following equation:

$$U_i(x,y) = \frac{\sum_{k=1}^{3} \phi_k(x,y) d_k}{\sum_{k=1}^{3} d_k} \tag{3.9}$$

where d_k is the distance of the target grid point from the k-indexed side of the triangle containing the target grid point and ϕ_k is the value of the function associated with the station at the k-indexed side of the triangle; it can have two forms: $\phi_k = U_i^k$ wind component at station k or a parabolic function closest to the values of the wind component at the three vertices of the triangle.

- Spline method. This is a well-known interpolation method, whose mathematical state-

ment is the following system of equations:

$$\begin{cases} U_i(x,y) = x + \beta y + \gamma + \sum_{k=1}^{N} \delta_k r_k^2 \log(r_k) \\ U_i(x_k,y_k) = U_i^k \\ \sum_{k=1}^{N} \delta_k = 0 \\ \sum_{k=1}^{N} \delta_k x_k = 0 \\ \sum_{k=1}^{N} \delta_k y_k = 0 \end{cases} \tag{3.10}$$

where r_k has the same meaning as in the Cressman method.

Interpolation at the upper microSWIFT levels

The wind field values at the upper microSWIFT levels are determined by a 2D Cressman level interpolation. The first step is to interpolate the *profiles* linearly on vertical grid levels. For grid points lying above the last measurement level, the profile is interpolated between the last measurement level and the geostrophic wind (given in data file). Each grid level now has an interpolated measurement and a Cressman method is then applied within the same level values.

Another possibility is to use directly a 3D Cressman interpolation method. The weighted distance between the grid points and a measurement point is given by the following equation

$$r_k^2 = \frac{(x-x_k)^2 + (y-y_k)^2}{P_h^2} + \frac{(z-z_k)^2}{P_v^2} \tag{3.11}$$

where P_h is the horizontal range and P_v is the vertical range.

Large differences can generally be obtained along the same vertical profile between surface values ($z^*(2)$) and upper-air values ($z^*(k), k > 2$), except with a 3D Cressman method. The user can select linear matching (which distributes the differences between the surface and the top level H linearly) or exponential matching (for which a vertical smoothing scale H_r must be supplied) of the values.

Fast interpolation

This interpolation is designed for cases with numerous profiles, such as gridded output of larger scale models. In such case, each point of the mesh is influenced by a very large number of seeds, even if the seeds have very little influence on the value of this point (see Cressman method Eq (3.8)). To be able to limit the number of seeds to compute values at mesh points, the strategy retained is to interpolate only from neighboring points.

Interpolation of temperature and humidity

Temperature and Humidity fields (here denoted by X) can be determined directly with a 3D Cressman interpolation or proceeding with a 2D interpolation by levels. If the users chooses the 2D interpolation by levels, than a surface correction to the station measurements must be applied, in order to reproduce a situation in which points at the same absolute altitude have nearly the same values. To obtain this result, the influence of measurement altitude is first removed from the X_k values by creating a set of X_k^* values brought to the same altitude, using the formula:

$$X_k^* = X_k - \Gamma \cdot (z_r - z_k) \tag{3.12}$$

where z_r is any reference altitude (e.g. $z_r = 0$) and Γ is an arbitrary X-gradient other than zero. Then a 2D interpolation is applied to obtain values (and in particular grid point values) $X^*(x, y)$ whose virtual vertical location is z_r. Surface field at level $z^*(2)$ is obtained by applying the following formula:

$$X(x, y) = X^*(x, y) - \Gamma \cdot (z_r - z_2(x, y)) \tag{3.13}$$

where $z_2(x, y)$ is the altitude of grid points at level $z^*(2)$:

$$z_2(x, y) = z_g(x, y) + z^*(2)\frac{H - z_g(x, y)}{H} \tag{3.14}$$

3.2.3 Adjustment of wind field

The interpolated wind then has to be corrected to account for incompressibility and boundary constraints. The basic notations are:

$$\mathbf{V}_0 = \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \tag{3.15}$$

is the field interpolated from original measurements

$$\mathbf{V} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \tag{3.16}$$

is the final field with zero divergence.

Incompressibility is introduced by means of a Lagrangian multiplier; moreover, the constraint for the adjusted wind to be as close as possible to the initial interpolated wind, plus the lateral edges flux constraint, have to be considered, thus leading to the following form of the functional to be minimized:

$$J(\mathbf{V}, \lambda) = \int_\Omega \left[(u-u_0)^2 + (v-v_0)^2 + \frac{1}{\alpha}(w-w_0)^2 \right] d\Omega + \\ + \int_\Omega \left[\lambda \boldsymbol{\nabla} \cdot \mathbf{V} \right] d\Omega + \beta \int_\Gamma (\hat{\mathbf{n}} \cdot (\mathbf{V} - \mathbf{V}_0))^2 \, d\Gamma \tag{3.17}$$

where λ is the Lagrangian multiplier, α is the stability parameter, β is the "lateral flux" parameter and Ω and Γ are the domain volume and the domain boundary respectively. Minimization of the functional gives the following set of equations:

$$\begin{cases} \boldsymbol{\nabla} \cdot \left(\mathrm{A}^{-1}\boldsymbol{\nabla}\lambda\right) = -\boldsymbol{\nabla} \cdot \mathbf{V}_0, & \text{within the domain} \\ \hat{\mathbf{n}} \cdot (\mathrm{A}^{-1}\boldsymbol{\nabla}\lambda) = -\hat{\mathbf{n}} \cdot \mathbf{V}_0, & \text{at ground and domain top} \\ \beta\hat{\mathbf{n}} \cdot (\mathrm{A}^{-1}\boldsymbol{\nabla}\lambda) = -\lambda, & \text{at lateral edges} \end{cases} \tag{3.18}$$

where

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha^{-1} \end{pmatrix} \tag{3.19}$$

Resolving with respect to λ and substitution in the formula

$$\mathbf{V} = \mathbf{V}_0 + \mathrm{A}^{-1}\boldsymbol{\nabla}\lambda$$

gives the desired adjusted wind field.

Determination of α and β

Determination of α is carried out bearing in mind the following considerations:

- in stable conditions, wind tends to bypass terrain obstacles and vertical wind is very weak compared with horizontal wind ($\alpha \to 0$)
- in unstable conditions, wind tends to cross the obstacle and vertical wind is stronger ($\alpha \to 1$).

microSWIFT offers several methods for determining the α factor; among them the Geai and the Moussiopoulos methods are worth mentioning. Both these algorithms relate on energy considerations of the flux and involve computation of the local Froude number.

- Geai method gives:

$$\alpha = \mathrm{Fr}\frac{h}{L}$$

where Fr is the local Froude number

$$\mathrm{Fr} = \frac{U}{Nh} = U\left[\left(\frac{g}{\theta}\frac{\partial\theta}{\partial z}\right)^{0.5} h\right]^{-1},$$

h is the difference in altitude between the two levels bordering the computation point and $L = \sqrt{\Delta x \Delta y}$, where the Δs are the horizontal dimensions of the cell containing the computation point.

- Moussiopoulos method gives:

$$\begin{cases} \alpha = 1 - \mathrm{Str}^4/2(\sqrt{1+4\,\mathrm{Str}^{-4}} - 1), & \text{for Str} \geq 0 \\ \alpha = 1, & \text{for Str} < 0 \end{cases}$$

and Str is the Strouhal number, inverse of the Froude number.

Determination of β is carried out by recasting equation (Eq (3.18)) in the form:

$$\lambda + \beta\frac{\partial\lambda}{\partial n} = 0 \tag{3.20}$$

where n is the direction of the unit vector normal to the boundary. Then, the following consideration on the limiting behaviour is made:

- $\beta \to 0$: the equation becomes $\lambda \to 0$; the normal wind component at the boundary is adjusted by applying a Dirichlet condition on λ
- $\beta \to +\infty$: the equation becomes $\partial\lambda/\partial n \to 0$; components parallel to the boundaries are adjusted by applying a Neuman condition on λ

The Dirichlet condition on λ is used in microSWIFT.

Interpolation and adjustment with obstacles

microSWIFT has the capability to handle obstacles such as buildings or bridges. The philosophy of the approach retained is to get 90% of the solution in 1% of the time needed by a

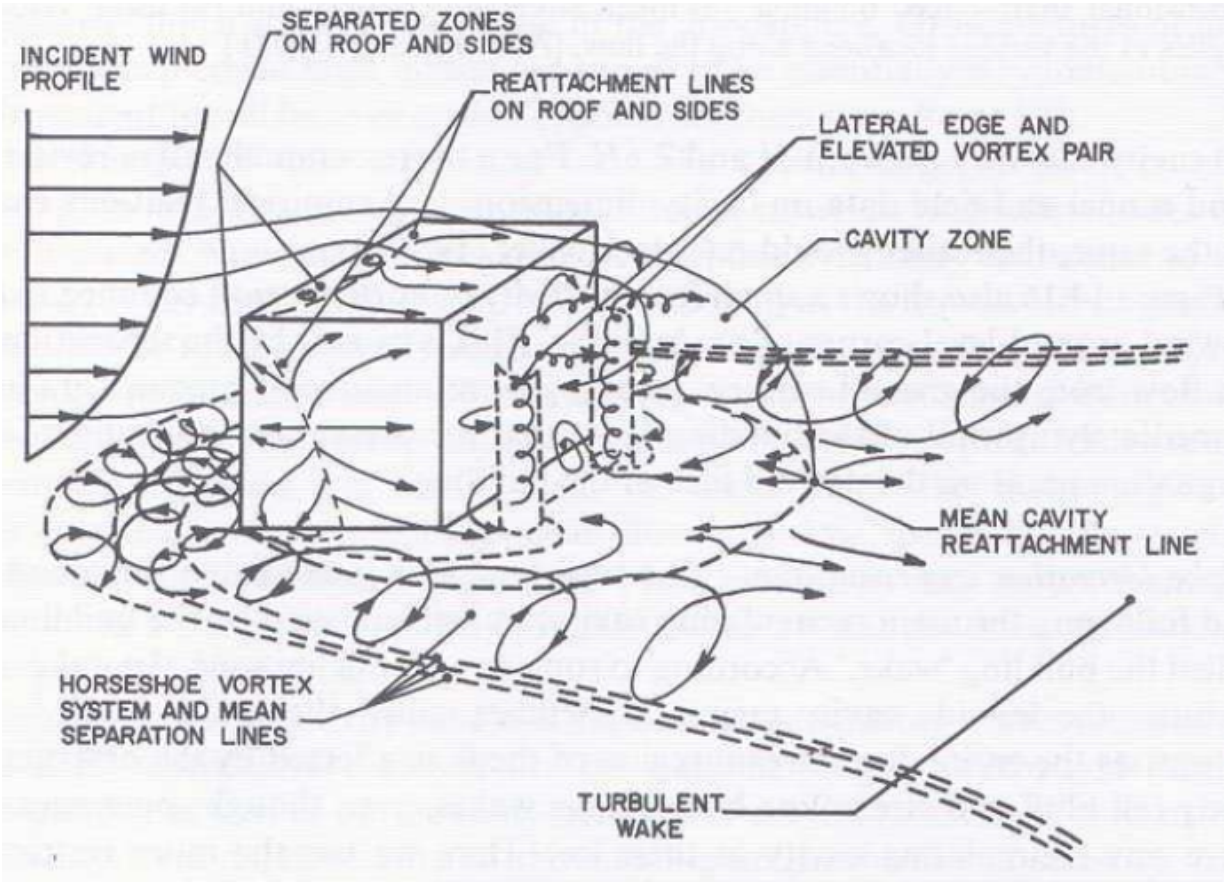


Figure 3.1: Main features of an obstacle-induced wind field.

full CFD calculation. The main steps are:

1. modify the interpolated wind field according to prescribed zones surrounding obstacles

2. adjust the wind field including the impermeability condition on walls and roof of buildings.

microSWIFT can handle prism-shaped obstacles with both rectangular or triangular base and is able to deal with the interaction of the induced-flux of two or more obstacles.

At the end of the interpolation step, and before the adjustment step, buildings are added in the mesh. At this stage, the wind inside the shapes is set to zero, then analytical zones are attached to obstacles. These prescribed zones are defined according to wind tunnel simulations and literature data. Inside these zones, the wind is modified analytically.

3.2.4 Computation of turbulence quantities

In the *surface layer* turbulent fluxes are expressed as follows, according to the Monin-Obukhov theory:

$$\begin{aligned} -\rho\overline{u'w'} &= const. \quad \text{momentum flux} \\ C_p\rho\overline{\theta'w'} &= const. \quad \text{heat flux} \end{aligned} \tag{3.21}$$

Profiles of average quantities are obtained through the formulas:

$$\begin{aligned} \overline{u}(z) - \overline{u}(z_0) &= \frac{u_*}{kL_{mo}} F\left(\frac{z}{L_{mo}}\right) \\ \overline{\theta}(z) - \overline{\theta}(z_0) &= \frac{\theta_*}{kL_{mo}} G\left(\frac{z}{L_{mo}}\right) \end{aligned} \tag{3.22}$$

where F and G are universal functions of the surface layer.

In the *boundary layer* a vertical flux of a scalar quantity is expressed by the relation:

$$\overline{a'w'} = -K_a \frac{\partial a}{\partial z} \tag{3.23}$$

where K_a is the turbulent diffusivity coefficient for variable a which could be a velocity component or the potential temperature. In microSWIFT, the turbulent diffusion coefficients (or velocity variances and Lagrangian timescales, Hanna scheme) are computed by the following first-order schemes, which contain parameterized functions of wind and temperature vertical profiles:

- O'Brien scheme
- Louis scheme
- Hanna scheme

O'Brien profiles are obtained through a polynomial fit of the available turbulence data on a domain which extends form the surface layer level up to the top of the ABL; Louis profiles

instead are determined by using the following formula:

$$K = l^2 \left[\left(\frac{\partial \overline{u}}{\partial z} \right)^2 + \left(\frac{\partial \overline{v}}{\partial z} \right)^2 \right]^{0.5} f(\mathrm{Ri}) \tag{3.24}$$

where l is the *mixing length* (not height) and $f(\mathrm{Ri})$ is a function of the local Richardson number Ri, which allows introduction of local effects of static and dynamic stability.

Hanna profiles are obtained through subdivison of the boundary layer into three layers for which different parameterizations of the various profiles are encoded. In particular, for the first layer above ground, the following distinction is made:

$$\begin{aligned} &\text{stable conditions:} && 0\,\mathrm{m} < L_{mo} < 300\,\mathrm{m} \\ &\text{neutral conditions:} && L_{mo} \geq 300\,\mathrm{m} \quad L_{mo} \leq -300\,\mathrm{m} \\ &\text{unstable conditions:} && -300\,\mathrm{m} < L_{mo} < 0\,\mathrm{m}. \end{aligned} \tag{3.25}$$

where L_{mo} is the monin-Obukhov length. Inside the S_2 layer (whenever existing) the code proceeds as in the neutral case, interpolating the equations with the results obtained on the layer below in the region between S_1 and S_2. Into S_3 layer variances are linearly brought to zero at the domain top level, whereas Lagrangian time scales are kept constant.

Turbulence on terrain with obstacles

In this case, a 3D field of local turbulence can be computed in the following way:

$$\sigma_i = \left(\frac{1}{2} C_0 K_i \varepsilon \right)^{0.25} \quad \text{and} \quad T_{Li} = \frac{2\sigma_i^2}{C_0 \varepsilon} \tag{3.26}$$

where K_i are the turbulent diffusion coefficients, ε is the turbulent kinetic energy dissipation rate and C_0 is a universal constant.

Sum of local and background (Mellor-Yamada level 2, SCIPUFF formulation or Hanna schemes, analogous to O'Brien and Louis schemes, - these last two are not available -) turbulence kinetic energy can also be computed on the whole domain, either in zone affected by

obstacles.

3.3 The microSPRAY model

microSPRAY is a 3D particle lagrangian dispersion model. The program simulates the advection and diffusion of gaseous species in the atmosphere, producing 3D concentrations, dosages and 2D dry or wet depositions. microSPRAY can perform simulations at urban microscale with a resolution lower than 10 meters, where obstacles are represented by filled cells.

3.3.1 Thomson's scheme

This scheme is based on a non-linear form of the Langevin stochastic equation; the motion of a given particle is defined by the following equations:

$$\begin{aligned} d\mathbf{x}(t) &= \mathbf{u}(t)dt \\ d\mathbf{u} &= \mathbf{a}(\mathbf{x},\mathbf{u})dt + \mathbf{b_0}(\mathbf{x})d\mu dt \end{aligned} \tag{3.27}$$

where

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \tag{3.28}$$

represents the position vector of the particle defined on a fixed cartesian reference frame and

$$\mathbf{u}(t) = \begin{pmatrix} u_x(t) \\ u_y(t) \\ u_z(t) \end{pmatrix} \tag{3.29}$$

represents the velocity vector.

$\mathbf{a}$ and $\mathbf{b_0}$ are generally functions of position and time, while $d\mu$ is a stochastic standardized Gaussian term (zero mean and unit variance). It has been demonstrated that the two

Lagrangian equations are actually equivalent to an Eulerian equation (the Fokker-Planck equation) for the spatial PDF ($P(\mathbf{x}, \mathbf{u}, t)$) in the phase space.

$$u_x \frac{\partial P}{\partial x} + u_y \frac{\partial P}{\partial y} + u_z \frac{\partial P}{\partial z} = \\ = -\left(\frac{\partial (a_x P)}{\partial u_x} + \frac{\partial (a_y P)}{\partial u_y} + \frac{\partial (a_z P)}{\partial u_z} \right) + b_{0x} \frac{\partial^2 P}{\partial u_x^2} + b_{0y} \frac{\partial^2 P}{\partial u_y^2} + b_{0z} \frac{\partial^2 P}{\partial u_z^2} \quad (3.30)$$

As for the $\mathbf{b_0}$ term, Thomson suggests the following expression for the vertical component:

$$b_{0z} = \frac{\sqrt{C_0 \varepsilon}}{2} = \frac{\sqrt{\overline{u_z^2}}}{T_{Lz}} \quad (3.31)$$

where C_0 is a universal constant, ε is the dissipation rate of turbulent kinetic energy, T_{Lz} is the vertical Lagrangian timescale and $\overline{u_z^2} = \sigma_z^2$ is the second moment of the distribution for vertical velocities.

These considerations can be extended also to the horizontal components, thus obtaining:

$$\mathbf{b_0} = \begin{pmatrix} \frac{\sigma_x}{T_{Lx}} \\ \\ \frac{\sigma_y}{T_{Lx}} \\ \\ \frac{\sigma_z}{T_{Lz}} \end{pmatrix}. \quad (3.32)$$

Turbulence related quantities (σ_i^2 and T_{Li}) generally depend on position and time; they are provided by an external model (such as microSWIFT) or computed by microSPRAY. The term $\mathbf{a}$ is determined by imposing the *well mixed condition* i.e. by substitution of the desired PDF in the Fokker-Planck equation (Eq (3.30)) and then resolving with respect to $\mathbf{a}$.

3.3.2 Implemented equations

Equations of motion for each particle are a finite difference discretization of Eqs (3.27). In a cartesian reference frame, the position of any single particle

$$\mathbf{x}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}} \tag{3.33}$$

evolves on the basis of the equation:

$$\begin{aligned} x(t+\Delta t) &= x(t) + u_x(t)\Delta t \\ y(t+\Delta t) &= y(t) + u_y(t)\Delta t \\ z(t+\Delta t) &= z(t) + u_z(t)\Delta t \end{aligned} \tag{3.34}$$

while the components of velocity

$$\mathbf{u}(t) = u_x(t)\hat{\mathbf{i}} + u_y(t)\hat{\mathbf{j}} + u_z(t)\hat{\mathbf{k}} \tag{3.35}$$

are obtained as a sum of a mean value and a fluctuation:

$$u_i = \overline{u}_i(\mathbf{x}, t) + u_i'(t) \quad \text{with } i = x, y, z. \tag{3.36}$$

The equations describing the evolution of turbulent *horizontal* velocities ($i = x, y$) have got the following form:

$$u_i'(t+\Delta t) = \left(1 - \frac{\Delta t}{T_{Li}(\mathbf{x}, t)}\right) u_i'(t) + \frac{1}{2}\frac{\partial \sigma_i^2}{\partial x^2}\left(1 + \frac{u_i'(t)^2}{\sigma_i^2}\right)\Delta t + \mu_i'(t)\sigma_i\sqrt{2\frac{\Delta t}{T_{Li}}} \tag{3.37}$$

These equations are the well mixed solution for the *horizontal* part of the Fokker-Planck equation assuming Gaussian turbulence as inhomogeneous. In this way, it is possible to take into account the inhomogeneities due to complex topography. The general form of the equation for the *vertical* component is:

$$u_z'(t+\Delta t) = a_z(u_z', \mathbf{x}) + \mu_z'(t)\sigma_z\sqrt{2\frac{\Delta t}{T_{Lz}}} \tag{3.38}$$

In the above equations the term μ' is a random number drawn from a standardized Gaussian distribution. With regard to the vertical velocity, the user can choose among a bi-Gaussian PDF or a Gram-Charlier PDF.

microSPRAY can further differentiate the motion above the PBL by using the equation:

$$dx_j(t) = \frac{\partial K_j}{\partial x_j} dt + \sqrt{2K_j} d\mu_j \tag{3.39}$$

where K_j are the turbulent diffusion coefficients, which must be supplied as input.

Concentrations are then computed by counting the number of particles contained in a given control volume centred on the nodes of a user-defined grid, multiplying this number by the lagrangian particle mass and dividing by the control volume. The user can also define the time window to be used for the computation of the average concentration.

3.3.3 Topography and coordinates

microSPRAY can also simulate the presence of complex topography, which could determine such phenomena as slope flows, channeling effects, wakes, that could have a substantial influence on the concentration of an emitted pollutant substance.

In microSPRAY topography is described by a continuous function $z_g(x, y)$ constituting the inferior boundary of the spatial 3D domain where particles are dispersed. Topography is generally given as a set of values defined on a grid, but lagrangian particles are localized in uncostrained points of the space, so the problem of knowing if a particle is under or above the topography level arises. Topography in a cell is guessed by means of a bilinear function of the form:

$$z = ax + by + cxy + d \tag{3.40}$$

and this surface is used to establish particle behaviour (bounce-back or deposition).

All the meteorological quantities of microSPRAY are defined on a reduced "sigma" coordinate system (S'):

$$\begin{aligned} x' &= x \\ y' &= y \\ z' &= \frac{z - z_g(x, y)}{H - z_g(x, y)} \end{aligned} \tag{3.41}$$

where H is the domain top level. z' values can be evenly spaced in order to give a better detail in particular regions, such as close to the ground, where vertical gradients generally achieve highest values.

3.3.4 Physical effects

microSPRAY can take into account some physical effects affecting the dynamics of the particle plume. These dynamical effects are generated by external forces acting on the plume and cannot be described directly by the same velocities derived from Langevin equations, responsible only of the motion due to the atmospheric mean transport and turbulence. For this reason, new equations are used to derive extra velocities to be considered by each particle. Effects currently available include:

- plume rise, simulated by bulk equations, and divided into
  - vertical plume rise for hot sources from chimneys
  - mechanical vertical plume rise for neutral sources from chimneys
- buoyant plume behaviour simulated by prognostic equations
- dense gas spreading at ground

3.3.5 Deposition

Deposition consists of the removal of the emitted material due to interactions with the ground. Various removal mechanisms, such as the interaction with a surface roughness or canopy layers, are generally grouped under a single phenomenon named "dry deposition". Wet deposition is a physical mechanism removing material from the emitted gaseous or particle plumes by means of scavenging, or washout, due to the rain.

Dry deposition

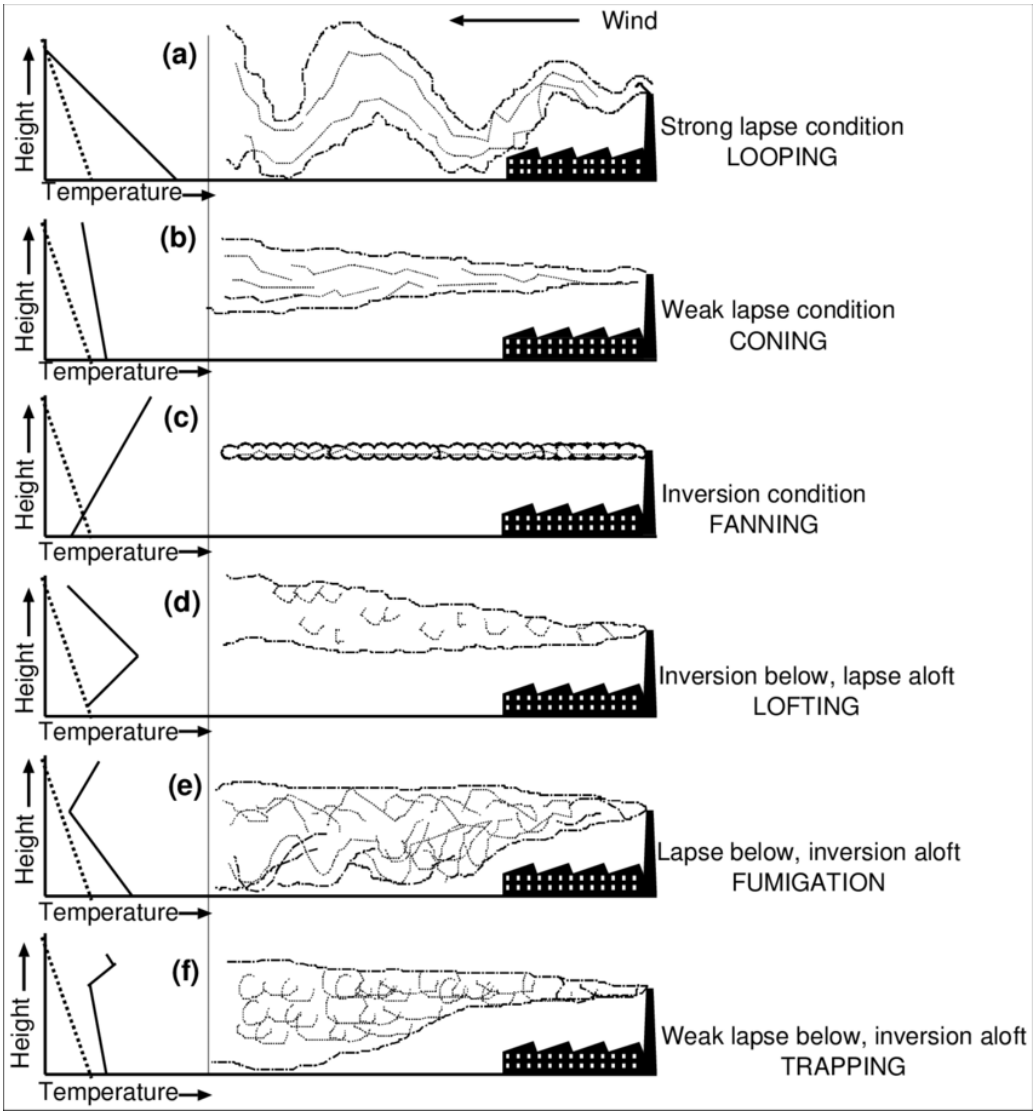


Figure 3.2: Different behaviours of a plume.

This mechanism should be simulated through a rate of removal, or deposition flux, proportional to the ground level concentration in a following way:

$$F_d = Cw_d$$

where the w_d term is named dry deposition velocity. A particle-oriented removal mechanism derived from a solution of the Fokker-Planck equation may be implemented in order to ensure that, during a time interval Δt, the deposition flux is proportional to the ground level concentration. This method can be extended also to the general case of particles experiencing vertical gravitational settling. It is based on the computation of a transitional probability that

a particle lying at a level z with respect to the ground is absorbed during a certain time inerval Δt. This transitional probability $P_d(z, \Delta t, w_d, w_s)$ is a function of both the depostion velocity and the vertical settling velocity w_s, rapidly tending to zero moving away from the ground level, being the probability of interaction with the ground negligible here. As for the settling velocity, the following formula holds:

$$w_s = \frac{g\rho d^2 C_c}{18\mu}$$

where g is the gravitational acceleration, ρ is the particle density in SI units, d is the particle diameter expressed in microns, μ is the air dynamic viscosity and C_c is the Cunningham slip-flow correction factor.

If a particle belongs to a gaseous emission, different transitional probabilities P_{ds} are firstly computed for each species, depending on the species-dependent w_d values, then an amount equal to $P_{ds}m_s$ is removed from the particle mass m_s referring to one species s. If a particle comes from an aerosol emission (in this case it is supposed that only one species is carrying a mass) a particle is entirely removed through a probabilistic method. If a random number extracted from a [0, 1] uniform distribution is less than P_c then the particle is totally absorbed, otherwise it is treated as if no absorption would take place, reflecting eventually. Species-dependent deposition velocities are given as input to the model. microSPRAY accepts either homogeneous and stationary values or 2D non-uniform and non-stationary deposition velocity fields generated by an external model. Masses removed from particles by the dry deposition process are cumulated into cells below particle depositions, having the same horizontal structure of the user-concentration grid. At the end of average periods defined for concentration computations, 2D total dry deposition and deposition flux fields of the considered species are stored on the gridded concentration/deposition output file.

Wet deposition

This effect may be treated using species-dependent washout coefficients S, defined in such a way that, for a species s:

$$\frac{dm_s}{dt} = -Sm_s$$

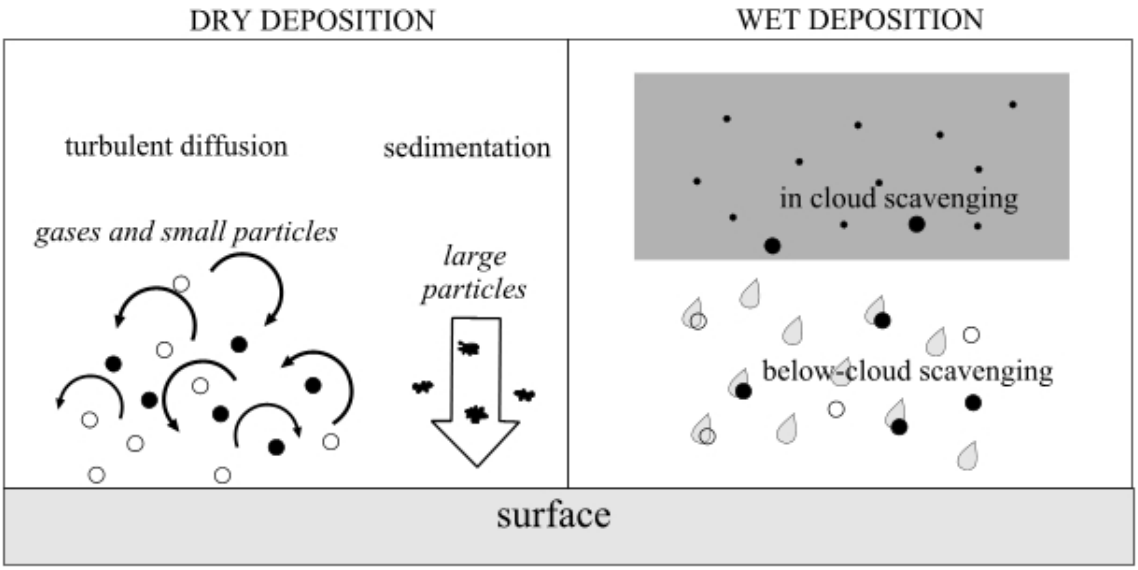


Figure 3.3: Pictoric representation of dry and wet depostion phenomena.

where m_s is the particle mass of species s. This leads to a time-dependent exponential decay for the particle mass that, in a step Δt reads:

$$m_s(t + \Delta t) = m_s(t)e^{-S\Delta t}.$$

Species-dependent washout coefficient S increase proportionally to the precipitation, being defined as follows:

$$S = S_1 R$$

where R is the precipitation rate (in mm/h) and S_1 is the species-dependent washout coefficient referred to a standard precipitation of 1 mm/h. These two values are given as input to the microSPRAY code.

3.4 WRF model

The Weather Research and Forecasting (WRF) model is a numerical model developed and improved in the last two decades for purposes of both research and operational forecasting use. WRF is a limited area model designed to work on many different computing platforms (it is particularly suited for parallel computing environments) and is available

via free download. The model simulates many physical processes and can resolve a wide range of scales both in time and space. Each run requires initial and boundary conditions that are usually supplied by global simulations. Users have the possibility of defining the domain and many physical parameterizations. All these properties make the WRF model very flexible. WRF considers the atmosphere as a fully compressible and non-hydrostatic fluid. Equations are integrated applying Runge-Kutta methods. For further information see www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/contents.html.

Chapter 4

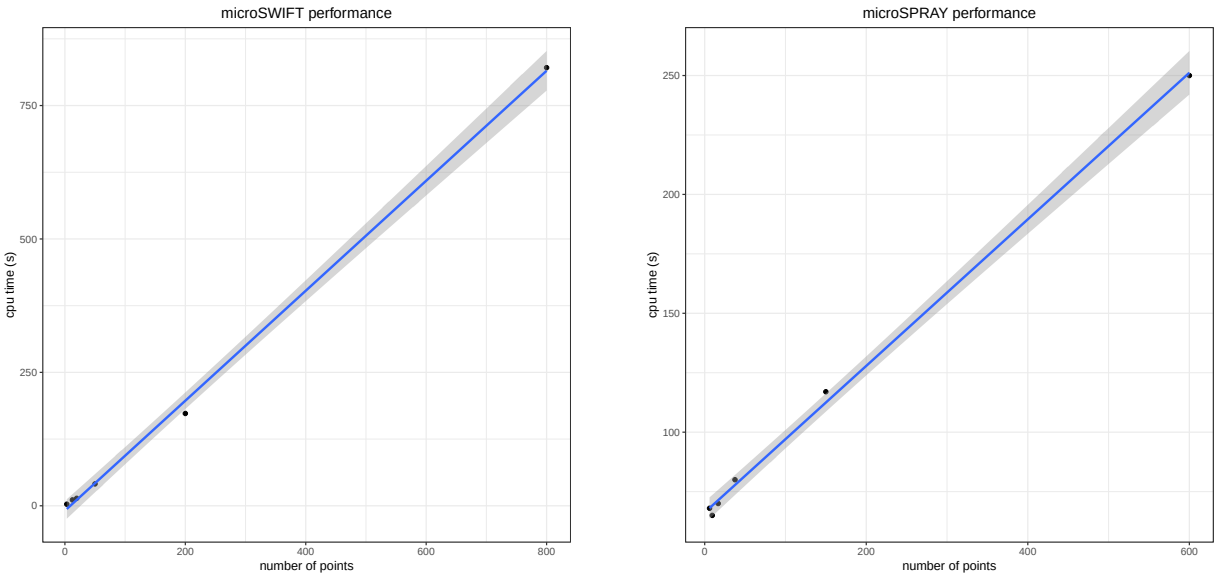
Simulations

4.1 Introduction

Simulations with the suite MSS were performed on the cluster FENICE (located in Amaro (UD)); commands were provided using a terminal at ARPA FVG head quarter, in Palmanova (UD), connected to the cluster. The preparation of the computational environment reqiured a lot of time, due to the moltitude of preprocessing tools on which MSS relies. The first section of this chapter deals with the operational issues of running a computational model. In the second section the preparation of the emissive input is explained while in section 4.4 the adopted climatological criteria is presented. Section 4.5 reports the results about simulations, with a space-time analysis of the output and a representativeness study.

4.2 Some preliminary operations

Before running the model one has to consider various computational aspects. First, it is desirable to give a rough estimate of the time needed for the suite to complete a specific job; this should in principle be assessed using theoretical arguments based on the fundamental computation time dependencies. For microSWIFT the serial computation time should depend linearly on the 3D resolution (given a fixed domain): it increases if a more accurate adjusted field is desired. For microSPRAY the serial computation time (still with a fixed domain size) should increase linearly with the number of particles emitted, or, if the emissions and the required



(a) microSWIFT performance (b) microSPRAY performance

Figure 4.1: CPU time characterization of microSWIFT and microSPRAY. Fit with linear model, shaded region corresponds to 95% confidence level.

concentration resolution are the same for each simulation, should increase linearly with the inverse spatial resolution of the concentration grid (this is because microSPRAY chooses a number of particles in such a way to give the desired concentration resolution) and with the number of temporal substeps required by the user. Things then can become more complicated when computing the concentration: again, the time required to count particles in a control volume should depend linearly on the 3D resolution of the control grid, given a fixed emission field and a fixed domain size. Actually, the time needed to complete a job depends also on the activated model scheme; it is thus clear that the overall computation time is the sum of different processes that have different time dependencies.

microSWIFT and microSPRAY have been tested on a domain of 800m x 800m with different resolutions and same emissions: figure 4.1 shows the expected linear behaviour of the computation time with respect to the number of grid points both for microSWIFT and microSPRAY. Due to the fluctuating data/process traffic on the cluster the computation time has been considered in pretty standard condition of activity on the cluster node, and the mode or median has been taken, in order to reduce the influence of outliers. Jobs with a small computation time do not show large fluctuations or they don't even show fluctuations at all, and thus can be considered as pivoting entities for a fit.

4.2.1 I/O structure of the models

As anticipated in the previous sections, both microSWIFT and microSPRAY do relate on a bunch of preprocessing tools. As a consequence of the computational flux complexity, a lot of work has been devoted to project and write the scripts and the codes required to drive data and executable functions along the preprocessing steps and the simulations. Since all the computation has been carried on the High Performance Computational facility of ARPA FVG, which is a HPC Linux cluster, the BASH interpreter and the PBSPro queue software have been the main communication and programming languages. Furthermore, the openMPI libraries have been the communication protocol for massive computation. All this stuff required a careful verification activity because of the sensitivity of the data flow on the matching between output and the related inputs for the next step. The I/O structure of the models is shown in figure 4.2 (for microSWIFT) and in figure 4.3 (for microSPRAY) where the arrows represent the whole programming activity that was spent to implement the simulation chain.

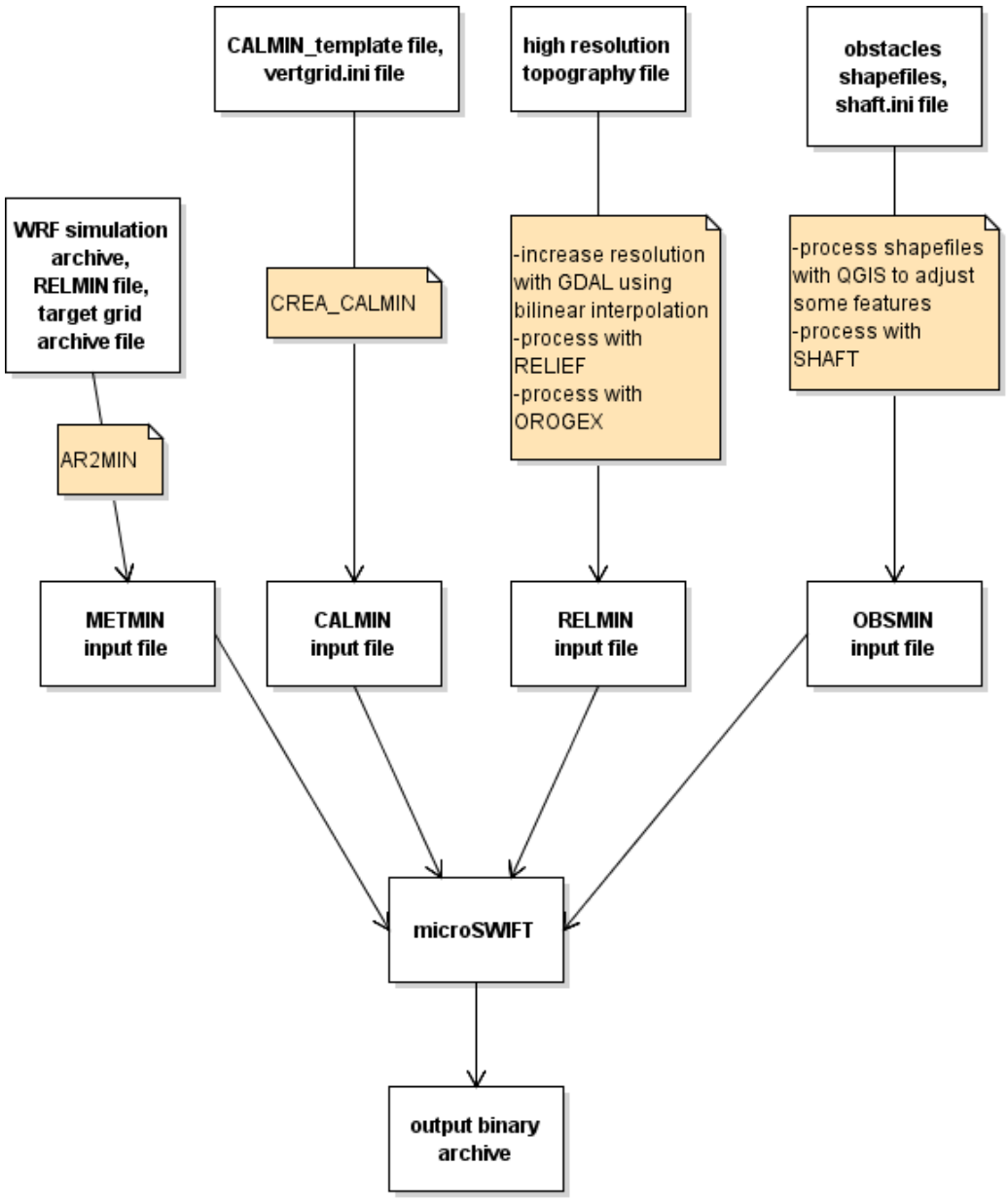


Figure 4.2: I/O structure of microSWIFT.

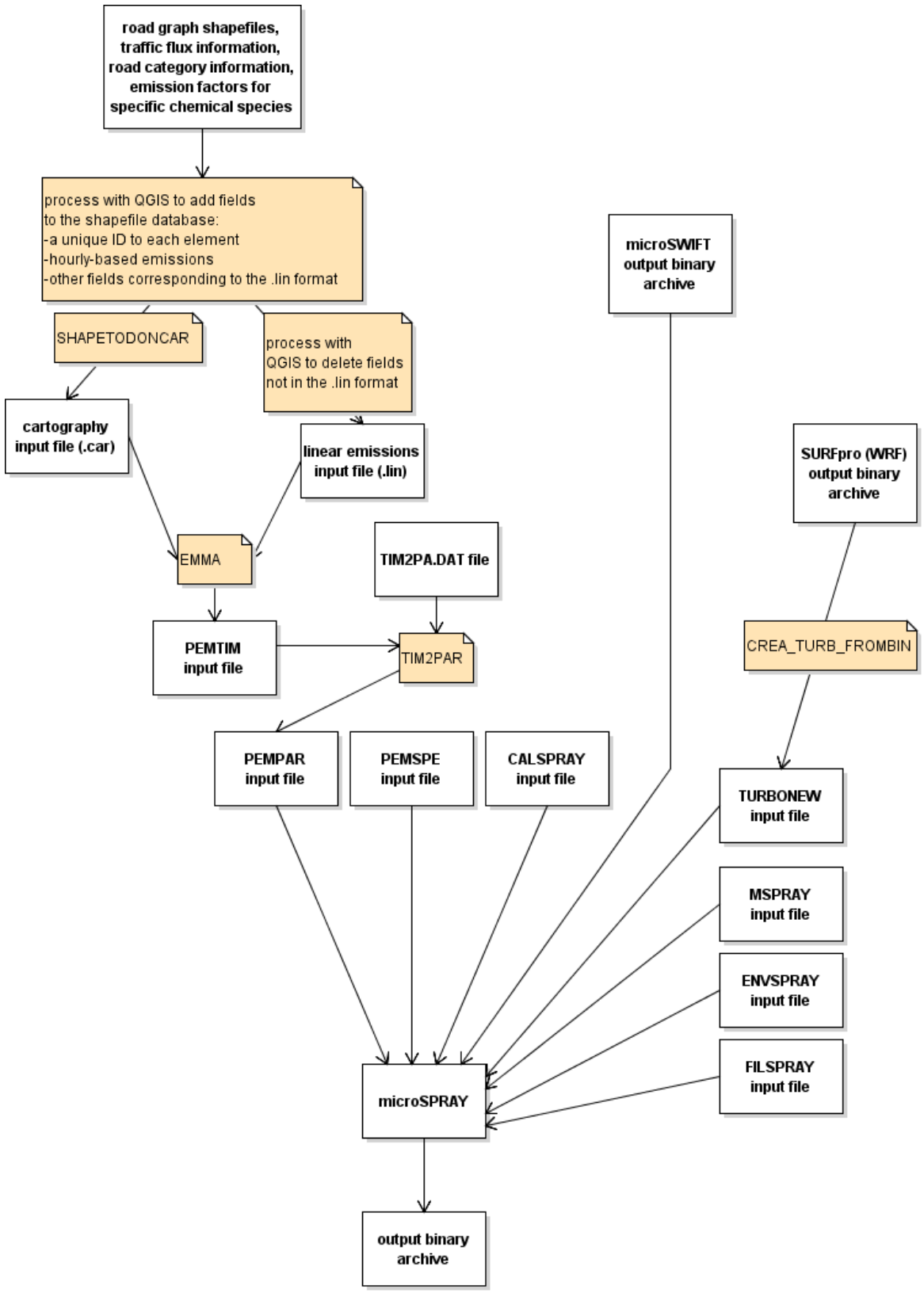


Figure 4.3: I/O structure of microSPRAY.

4.3 Preparation of the emissions

The subject of the analysis will be the concentration of traffic pollutants in a monitored area: a domain of 747m x 747m centred on the air quality station of Viale San Daniele (located in the centre of Udine town) has been chosen (see Fig. 4.4) with a resolution of 3m x 3m. Temporal resolution is 1 hour for microSWIFT and 10 seconds for microSPRAY, although in this last case the term *synchronization time* is to be preferred. In order to reconstruct the traffic flux in that area, a trial with air quality data has been made, assuming a very fast response of some pollutants' concentration to the traffic flux; comparison of these data with real traffic data (2007 ARPA FVG campaign) highlighted inadequacities of such a choice as 4.5 shows. Normalized profiles of the traffic flux during the weekdays have been estimated and compared, revealing that there is not much difference between two available nearby roads (Via del Cotonificio and Via Martignacco) even if they were classified as two different categories (OpenStreetMap classification, https://overpass-turbo.eu/), hence the need of a direct measurement of the difference between traffic flux in secondary and primary roads. A trivial count of the vehicles circulating during the rush-hour revealed that secondary road traffic flux is about 1/26 of the primary road traffic flux.

The data from OpenStreetMap provided all the information neccessary to the construction of an emission-magnitude category index, in particular:

- primary roads coefficients were assessed after traffic conservation considerations, given that the north-east stretch of Viale San Daniele had a magnitude of 1
- one way secondary roads has been assigned a cofficient of 1/26
- two lanes secondary roads were assigned a coefficient of 2/26
- service roads were assigned a coefficient of $2/3 \cdot 1/26$

Figure 4.4: The chosen domain. The white dot is the position of the air quality station; in UTM33 coordinates (363109, 5103452)m.

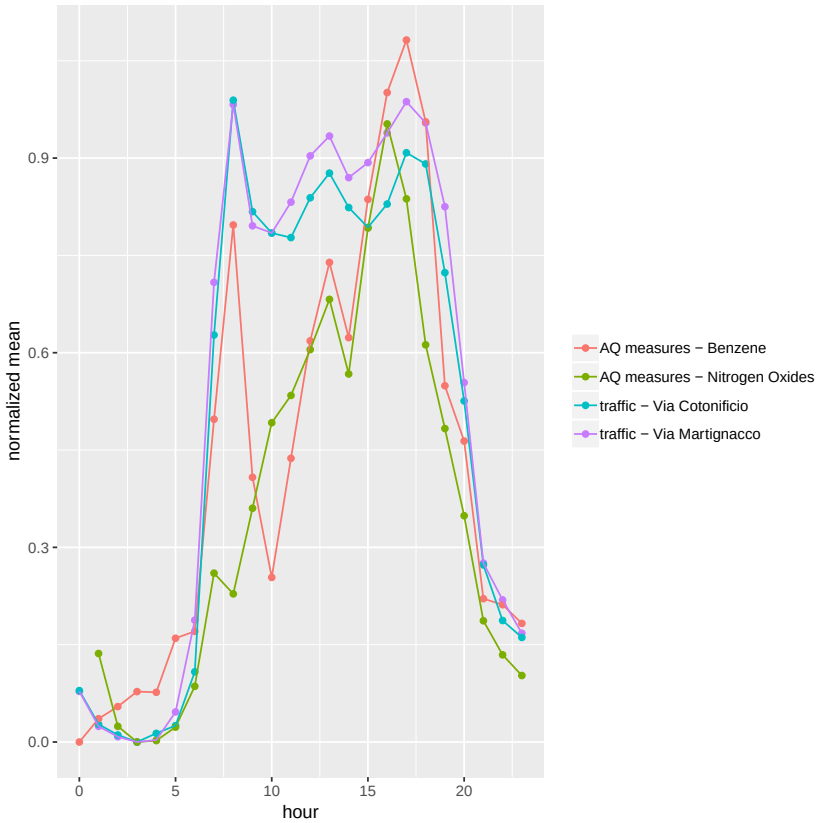


Figure 4.5: Comparison of air quality normalized mean concentration measures profiles and traffic flux normalized profiles.

Figure 4.5 shows how much the response of a pollutant concentration is fast with respect to the traffic flux. Benzene concentration in particular has a very high correlation with the traffic flux, in fact it is basically a passive pollutant, unlike nitric oxide (NO) and nitrogen dioxide (NO_2) which instead are affected by chemical and photo-chemical reactions. The dispersion model (microSPRAY) can deal with chemical reactions but activation of the chemical module comes at cost of an increase in complexity and in time needed for the computation. Thus benzene appears to be a more appealing pollutant to feed to a simulation in order to produce realistic concentration fields.

The mean emission factors for Italy (referred to year 2016) can be downloaded from the ISPRA[1] website www.sinanet.isprambiente.it/it/sia-ispra/fetransp/, which provides factors in grams per kilometer for many pollutants and for many vehicle categories (see 4.6). An estimate of the average emission for a certain road is thus given by the sum of the number of vehicles weighted by the specific emission factors (ISPRA), then summing again over the weekdays (and possibly over the lanes) and eventually dividing by the product (weekdays·lanes·r) with r=24 to obtain hourly based data. In order to obtain emissions in a format suitable for microSPRAY the hourly based emission in grams per kilometer (as previously calculated) has to be multiplied by the length of the street arc and by the category index (built with graph information from OpenStreetMap).

The daily emission profile has been derived from the weekday traffic data rearranged in a way that the hourly values sum up to 24 (requirement imposed by the EMMA software structure). Then information about the emissions time modulation, about the emissions features and about the road graph cartography were provided to a software (EMMA: EMission MAnager) which transforms those inputs in a file PEMPAR, containing the history of the emissions.

[1] Istituto Superiore per la Protezione e la Ricerca Ambientale

(a) Hourly emissions linear density. (b) Hourly emissions.

Figure 4.7: Road graph with *line width-emission magnitude* correspondence.

Category	Benzene 2016 g/km TOTALE	Benzene 2016 t/TJ TOTALE
Passenger Cars	0,002504321	0,001031224
Light Commercial Vehicles	0,001477127	0,000434738
Heavy Duty Trucks	0,00013385	1,54413E-05
Buses	0,000147955	1,45381E-05
Mopeds	0,03202891	0,047572824
Motorcycles	0,010626187	0,008208775

Figure 4.6: Emission factors from the ISPRA inventory.

4.4 Wind regimes

The microscale meteorological fields of a certain area have a tendency to periodicity; in fact the seasonal cycle and the planetary waves periods do affect the weather in a cyclical way. Further, certain wind structures tend to be more frequent than other, with no apparent periodic behaviour. This is a somewhat obvious statement: in other words it just says that the wind field of a certain site on Earth cannot even coursely match the wind field of a different place,

but it surely matches (more or less accurately) the wind field of that same site at a certain moment.

This is the principle for a clustering algorithm to be applied to a certain meteorological dataset. In this work, in fact, real meteorological conditions has been used as driving forces; these are *not* idealized simulations. In the case under study it has been decided to divide three years of hourly-based horizontal wind speed observations into four different classes. A PAM (Partition Around Medoids, see Reynolds et al., 2006) algorithm has been used to select the most representative elements among the days rearranged in eight intervals of three hours each (3 hours-based day: a set of eight numbers). The clustering revealed the presence of a summery cluster (number 4 in the figure below 4.8) which shows a pronounced breeze oscillation and the presence of a high speed north-east wind cluster (4.9); the other two clusters have no evident signature moves.

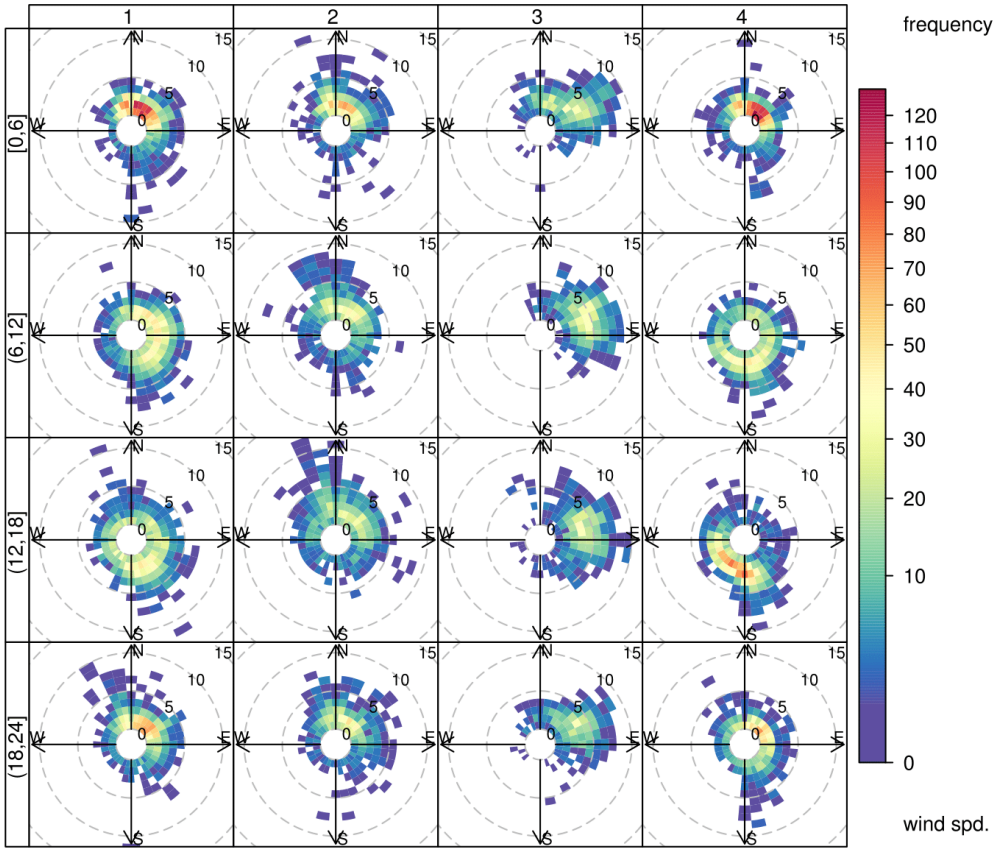


Figure 4.8: Frequency-based windroses for the four clusters. 6 hours-based days.

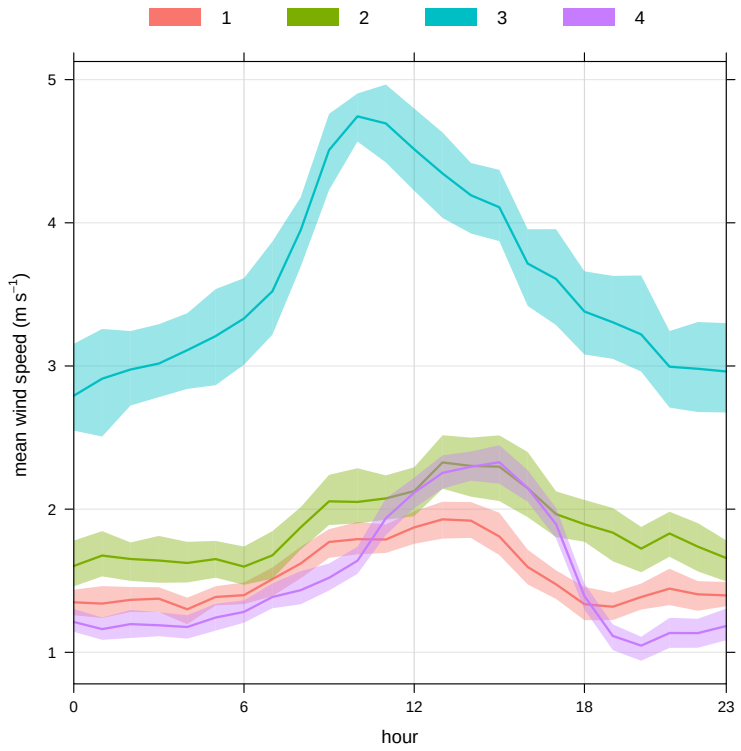


Figure 4.9: Hour-based mean wind module for each cluster.

A comparison with data from air quality measurements should be useful to confirm or reject some trivial hypothesys that can be made on dispersion processes, in particular: cluster 3 should be the least polluted, followed by cluster 2 and then by cluster 1 and 4 that should show accumulation evidence. Before starting to process the data one has to filter out non-winterly months; in fact, benzene displays a seasonal oscillation as shown in 4.10 due to the dynamics of the ABL. Figure 4.11 confirms all the trivial assumptions made on pollution for the four clusters, but sheds light on interesting daily dynamics: peak mean concentrations are associated to late afternoon hours, with a typical pattern that repeats through each cluster; cluster 4 has the most wide range of mean concentrations, even if most of time it remains under the concentration values of clusters 1 and 2; cluster 3 instead has values in a small range of mean concentration.

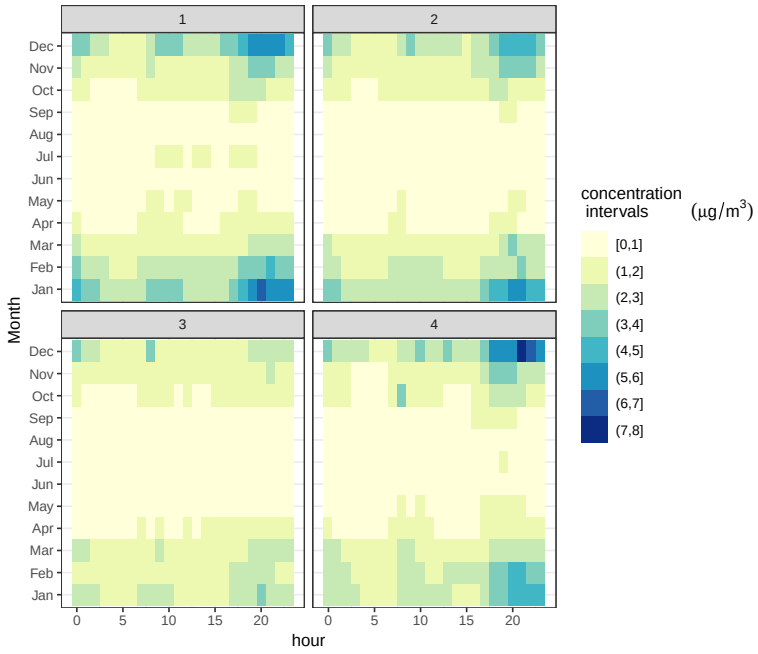


Figure 4.10: Annual and daily variation of benzene concentration.

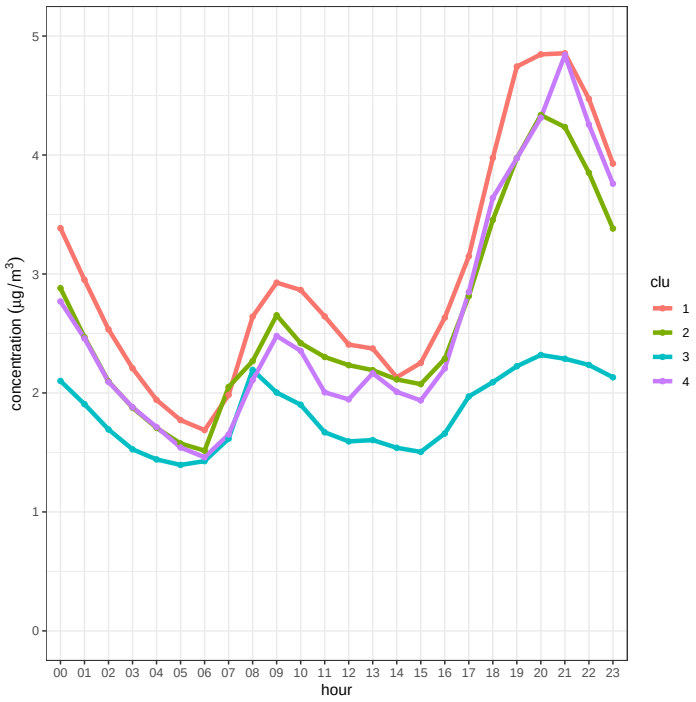


Figure 4.11: Daily benzene concentration measurements using a GAM smoothing (Generalized Additive Model).

As anticipated, the MSS suite can deal with both meteorological sparse data and simulated gridded data. Treatment of meteorological data presented a lot of technical problems and so

data from WRF simulations has been used. In order to identify a day whose actual meteorological conditions and simulated ones match to a good extent, a "goodness" ranking has to be produced, using a certain similarity distance definition. In this case the variance has been chosen and four winterly weekdays has been selected to represent each cluster class. As an example, two plots showing the comparison between WRF and observed data for the day 20-02-2017 are reported in figure 4.12.

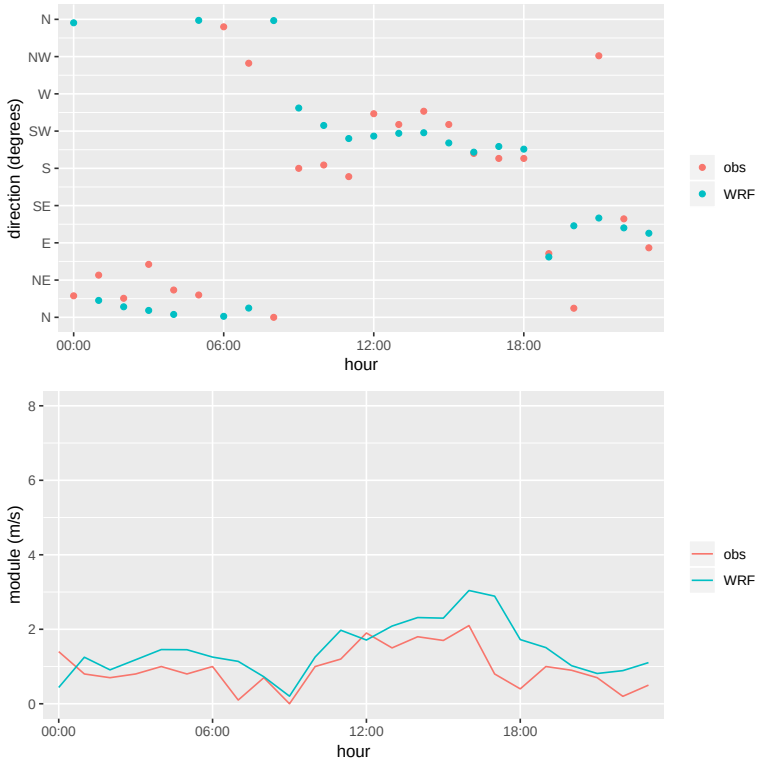


Figure 4.12: Comparison between WRF and observed data for the day 20-02-2017, a high quality WRF element of the breezy cluster.

4.5 Results

In this section the results about the four simulations (one for each selected day) are reported. In figures 4.13 and 4.14 four levels of the null-divergence wind field are shown for the day 26-01-2016 at 12:00 UTC+1 and for the day 20-02-2017 at 12:00 UTC+1. In both cases canyon effect zones and recirculation areas can be noticed. This is praticularly evident in figure 4.13, where the higher wind speed values affect a canyon zone corresponding to a street aligned with the wind in the upper layers.

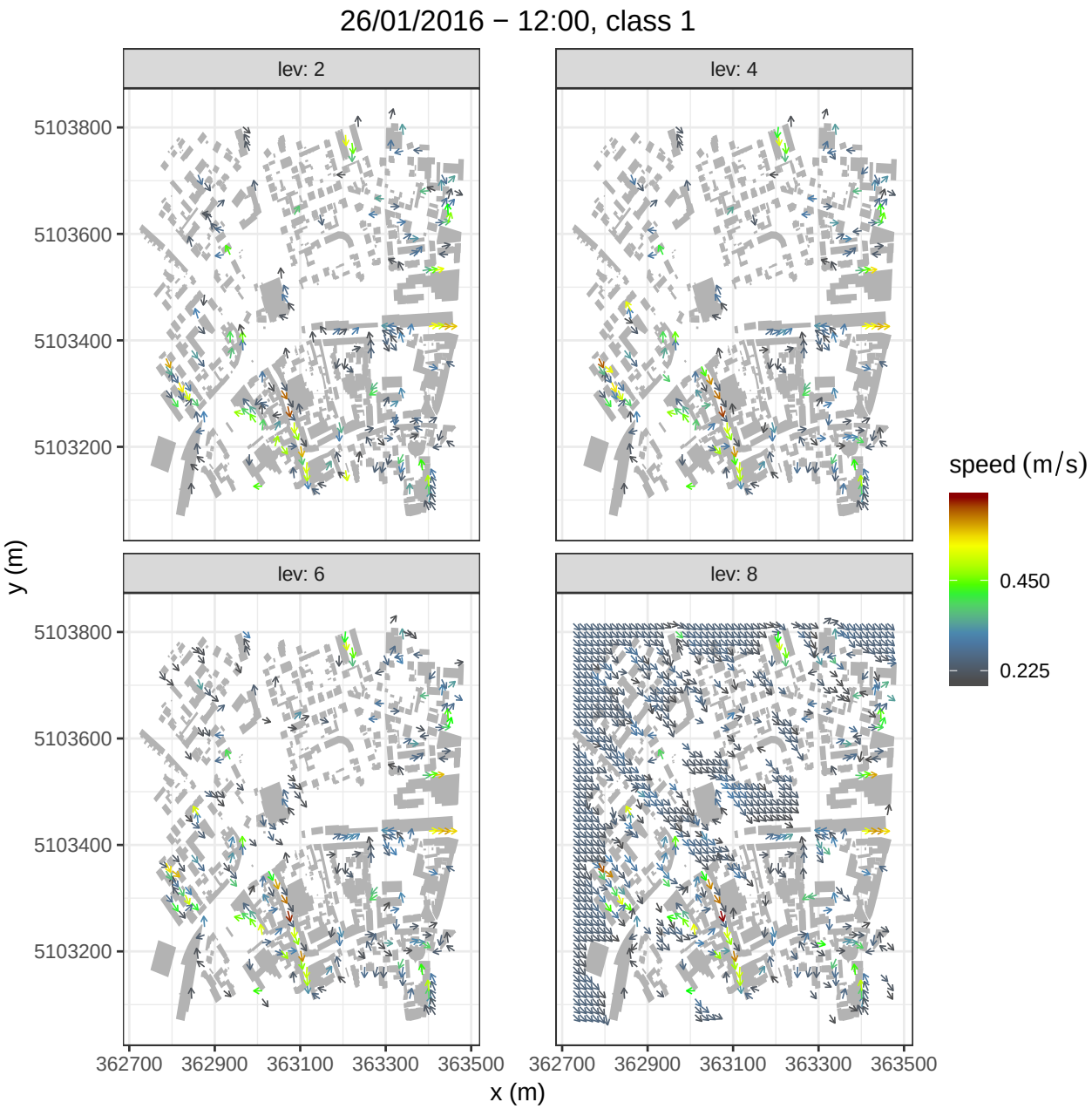


Figure 4.13: Four wind levels for the day 26-01-2016 at 12:00 UTC+1. Level 2 corresponds to microSWIFT coordinate (sigma coordinates) 1m, level 4 to 3m, level 6 to 5m, level 8 to 7m. Arrows correspond to one value every five on a row and a cut off has been set on lower values.

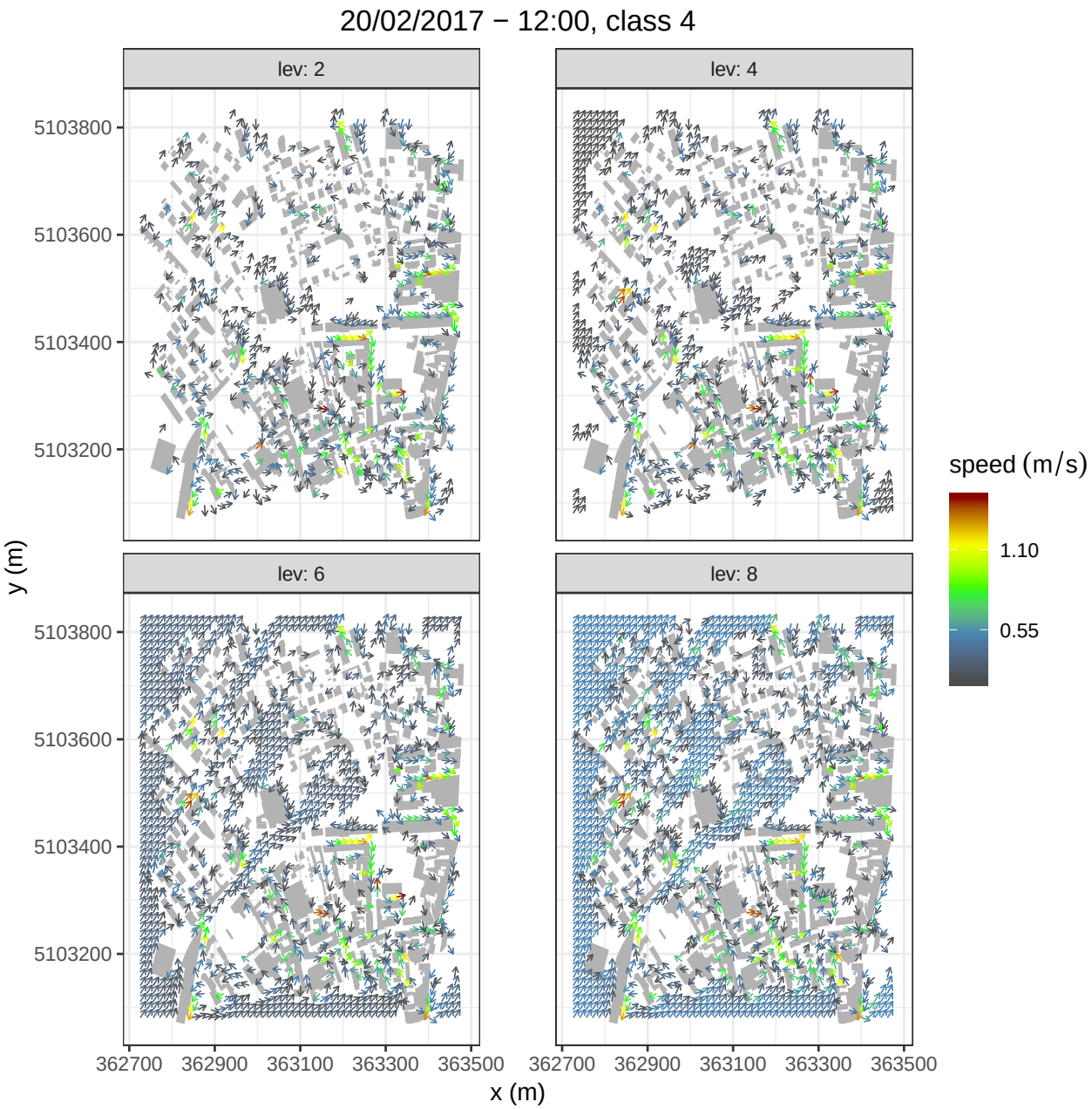


Figure 4.14: Four wind levels for the day 20-02-2017 at 12:00 UTC+1. Level 2 corresponds to microSWIFT coordinate (sigma coordinates) 1m, level 4 to 3m, level 6 to 5m, level 8 to 7m. Arrows correspond to one value every five on a row and a cut off has been set on lower values.

The concentration maps produced with the adjusted wind field confirm the expectations: days belonging to class 1 and 2 show larger concentration values, while class 3 has the smallest maximum values; class 4 has an intermediate behaviour, with some zones near the road graph displaying very high values and some other near zones showing low concentration values. Two cases were investigated, corresponding to relatively high magnitude simulated values recorded at the traffic station site, i.e. maps at 12:00 and at 18:00 (UTC+1) reported in figures 4.15 and

4.16. It is interesting to notice how even at 18:00 the breezy day still displays high concentration in the eastern and southern regions of the domain, where persistent stagnation zones are likely to take place, while the other days show a strongly dispersive character.

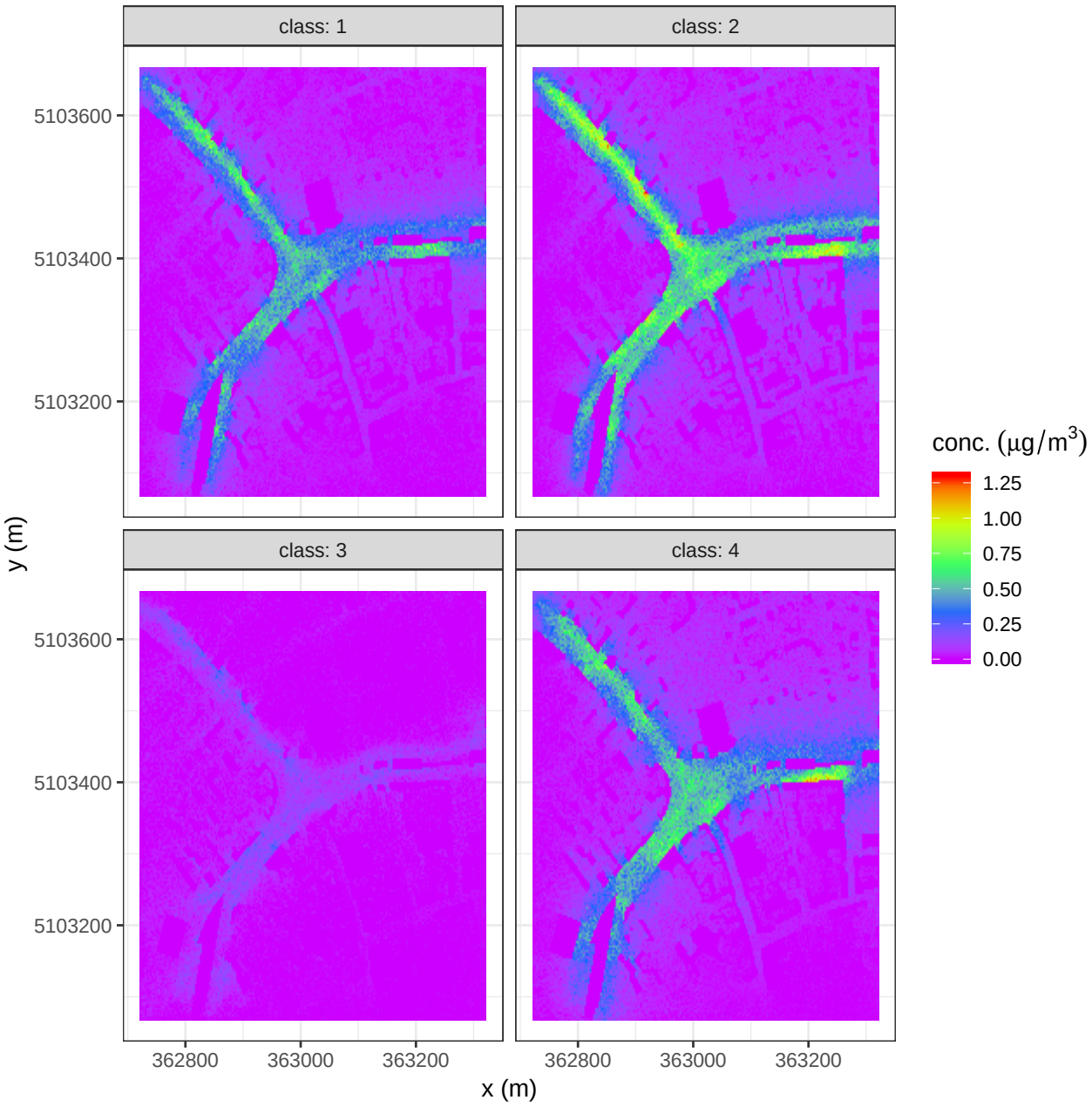


Figure 4.15: Concentration maps for the first microSPRAY level (1.65m with control volumes of 27m^3) at 12:00 UTC+1.

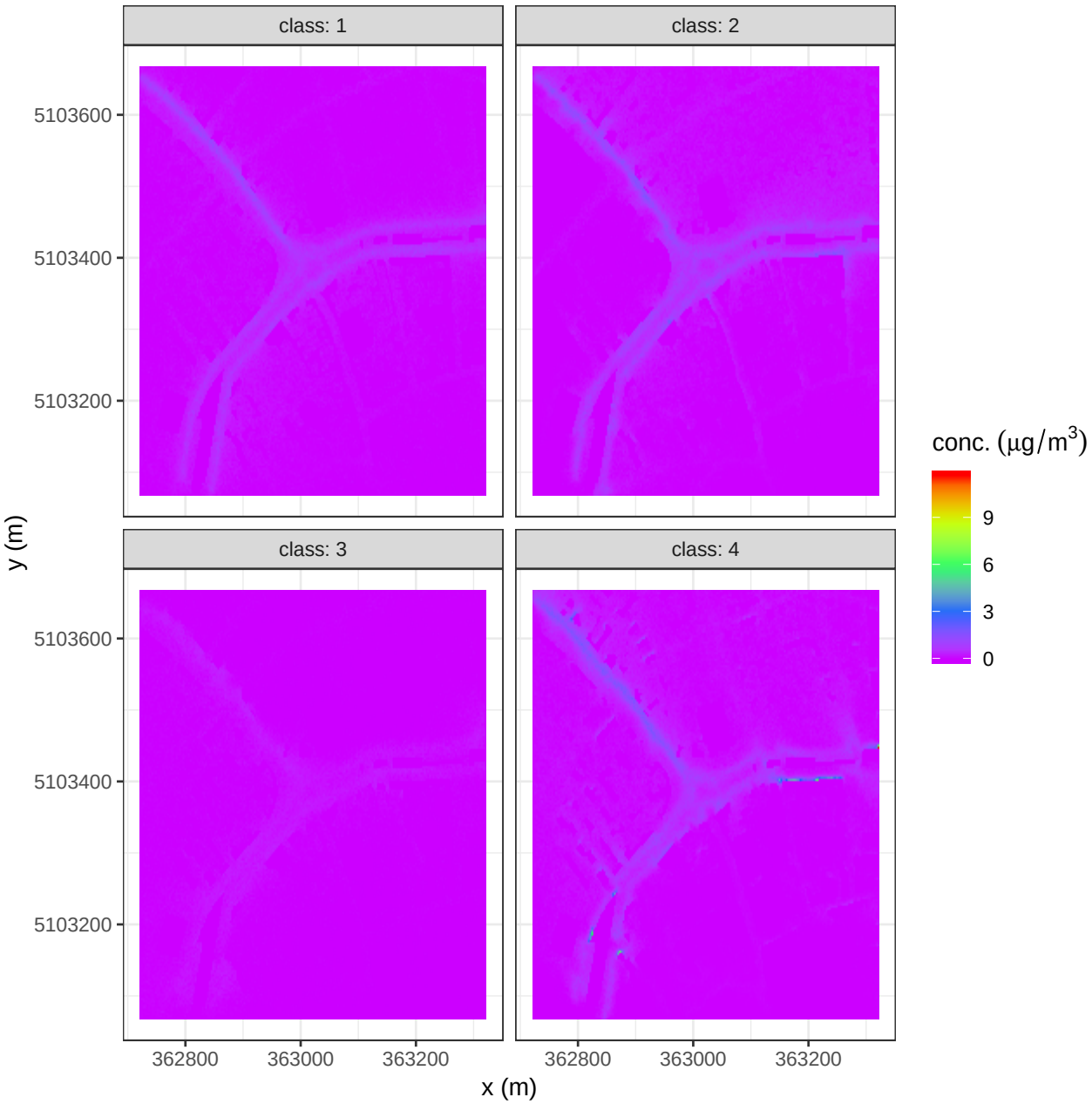
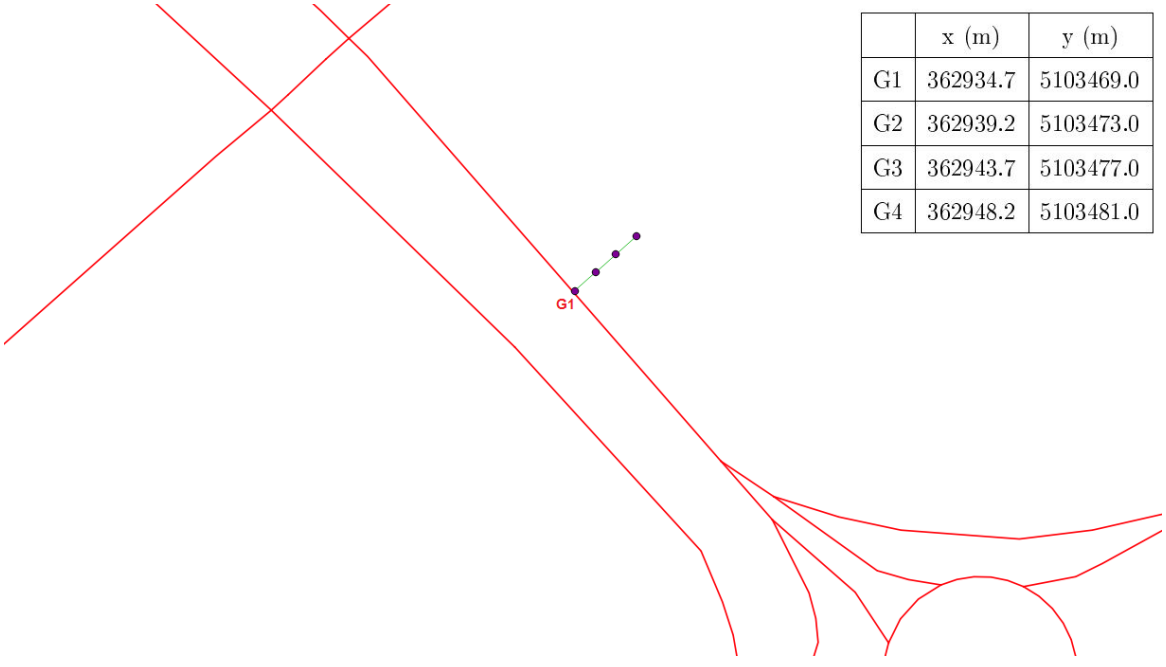


Figure 4.16: Concentration maps for the first microSPRAY level (1.65m with control volumes of 27m^3) at 18:00 UTC+1.

4.5.1 Concentration profiles across roads

In order to better characterize the concentration fluctuations a spatial/temporal analysis has been carried out on a line that crosses Via Martignacco. Four points with a step of 6m has been considered (see figure 4.17) and the concentration time series has been extracted for each selected day. The four plots (figure 4.18) shed light on the concentration spatial gradient in very different meteorological conditions: the day 19-12-2016 shows an initial accumulation when the flow at night is still stable; then, the signal from the source is lost due to the strong wind conditions; the other days, especially 26-01-2016 and 20-02-2017, show a bi-modal concentration pattern, with two peaks at 09:00 and at 18:00 for class 1 and 4, and two peaks at 09:00 and at 16:00 for class 2. Thus, as expected, there's a high correlation with the traffic flux curve (see figure 4.5 and the concentration curves at the source (G1 site) for stable conditions. This result agrees with other studies carried on by means of gaussian models applied to linear sources (see Cardozo and Sanchez, 2017 and Montanari, 2017) : this is particularly evident looking at the 18:00 UTC+1 concentrations, when simulations produce quite different values moving across the road (see 4.1).



	x (m)	y (m)
G1	362934.7	5103469.0
G2	362939.2	5103473.0
G3	362943.7	5103477.0
G4	362948.2	5103481.0

Figure 4.17: The four points considered in the gradient analysis. Coordinates are expressed in the UTM33 coordinate reference system.

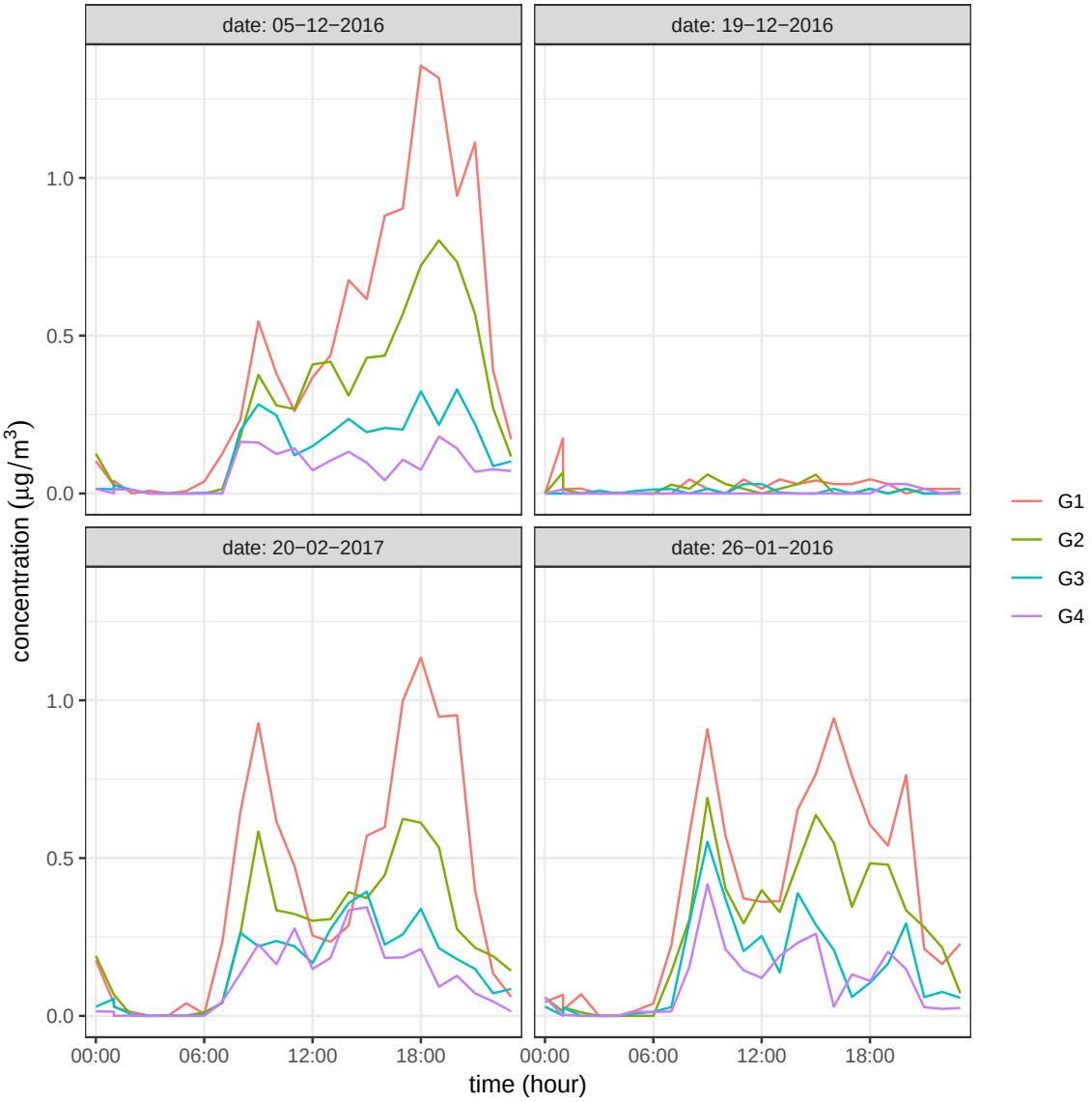


Figure 4.18: Concentration time series extracted from four points lying on a line crossing Via Martignacco, each 6m apart one from another. G1 is a point on the street.

time	G1 (μg/m^3)	G2 (μg/m^3)	G3 (μg/m^3)	G4 (μg/m^3)
05-12-2016 18:00	1.36	0.72	0.32	0.075
20-02-2017 18:00	1.13	0.61	0.34	0.211

Table 4.1: Hourly concentrations at points crossing the street. G1 is loacted in the road, that is over the emission line, while G2, G3 and G4 represent points displaced far from the street, each one 3 meters apart the closest, and G4 is the farther.

4.5.2 Concentration fluctuations on a grid

Since the concentration field fluctuates with high frequency and it presents strong anisotropic gradients, station measurements report a special detail of the field evolution only. Moreover, single grid point simulation data attempt to reproduce the concentration behaviour according to the spatial and temporal scales set in the computational resolution. According to the still present measurement errors and the limits in the numerical model performance, the weakness of the single point time series comparison it is straightforward. To avoid this problem and to let introduction of of the simulation fluctuations time series from a grid of 81 points (resolution of 3m x 3m) (see figure 4.19) around the station site has been extracted and the statistic of the concentration fluctuations has been examined. Figure 4.20 shows that maximum values in an area of 24m x 24m around the traffic station do not reach very high values, at least not as high as the ones recorded by the traffic station (for comparison see 4.11). It also shows that on day 05-12-2016 concentration values have the most wide range, signaling a strongly dispersive behaviour. Class 3 has a very narrow range of small concentration values (as expected from a windy day). Class 4 and 1 have very similar behaviour with three peaks emerging from the concentration time series.

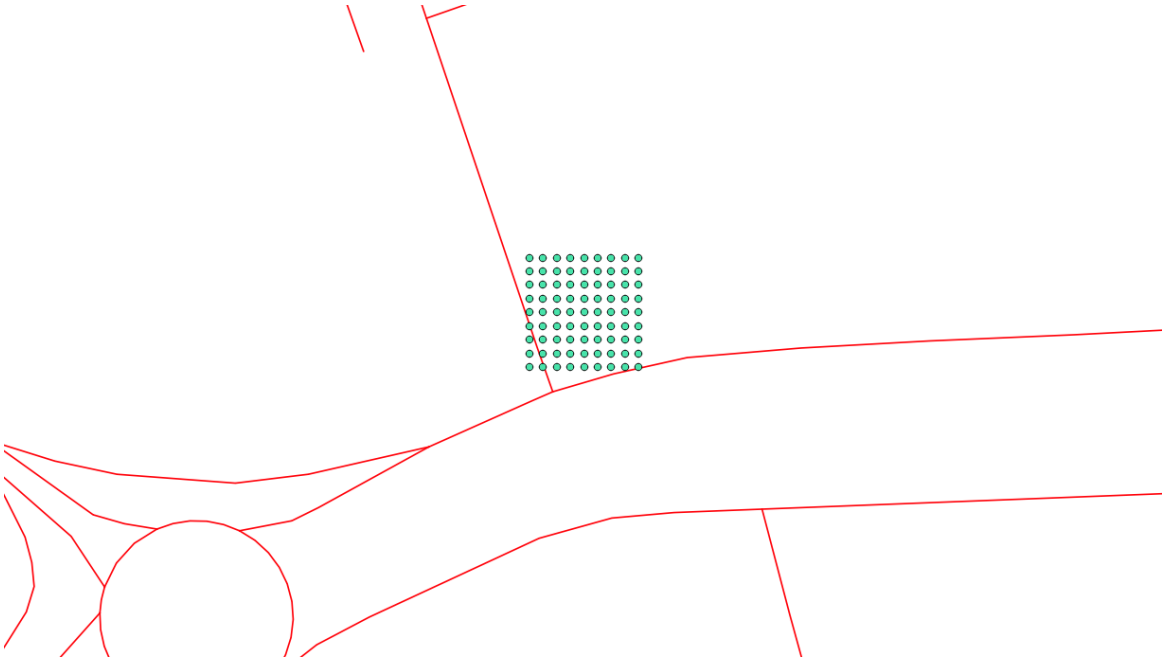


Figure 4.19: The grid points considered in the fluctuations analysis.

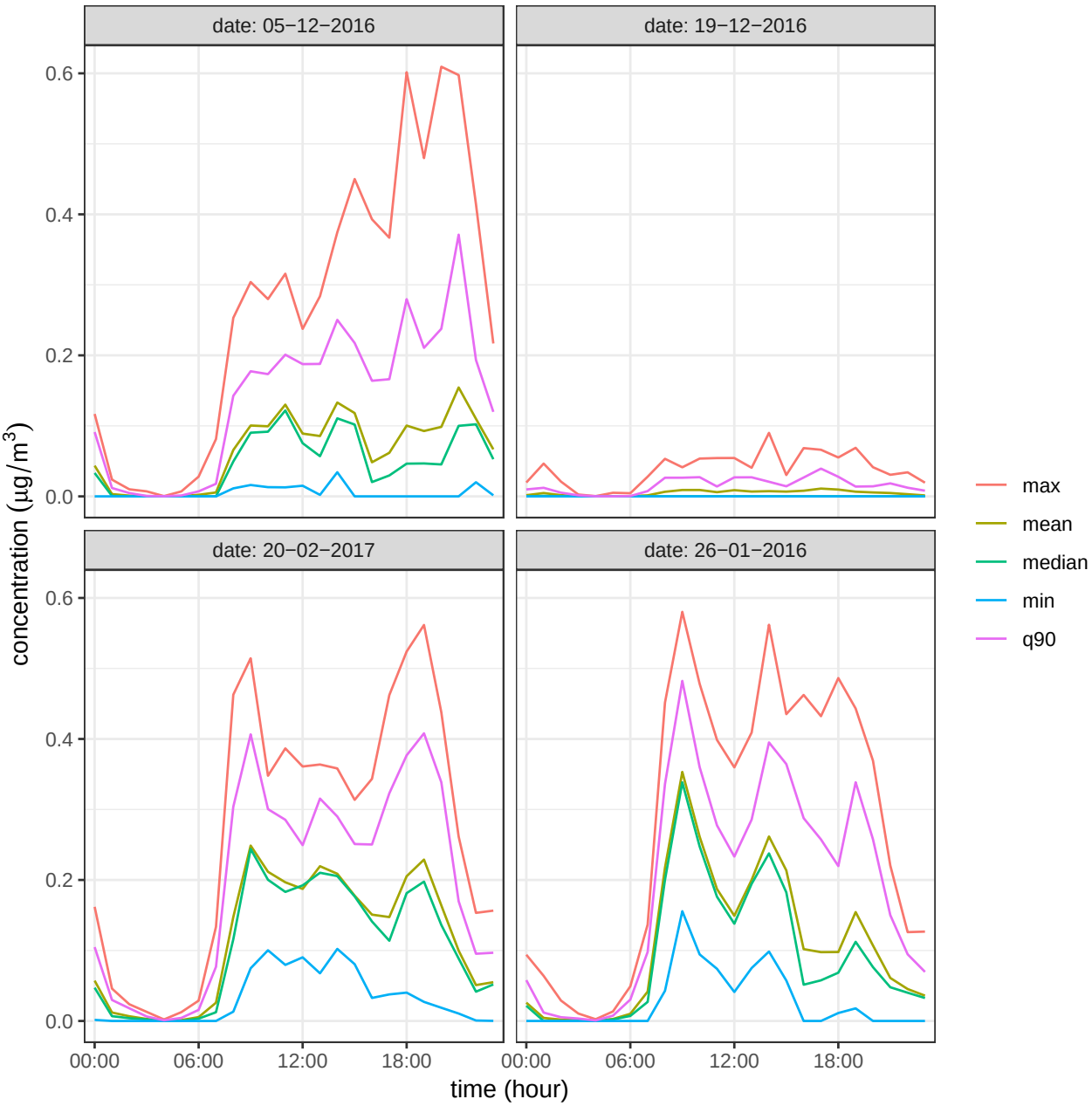


Figure 4.20: Time series for the concentration extracted on a grid covering an area of 576m^2 around the traffic station of Viale San Daniele (q90 is the 90th quantile).

4.5.3 Representativeness study

Representativeness, in the study of air pollution, is defined as an adimensional function of the inverse of the distance between the mean behaviour of a certain domain and the behaviour of a specific point in the domain. There is no restriction upon the kind of measure to be applied, except for it has to produce a positive number. In the case under investigation the distance has been defined as the root mean square of the difference of hour-based point data

between the hour-based average over all the domain points. On iterating the same procedure for every point in the domain one obtains a representativeness map, a useful tool for assessing the presence of pollutant stagnation/recirculation areas and basically for assessing how much a point is a good candidate to host an air quality station. For the case under investigation a class-frequency weighted map has been produced, in order to aggregate the most relevant meteorological regimes and produce a single point representativeness datum. The result is shown in figure 4.21.

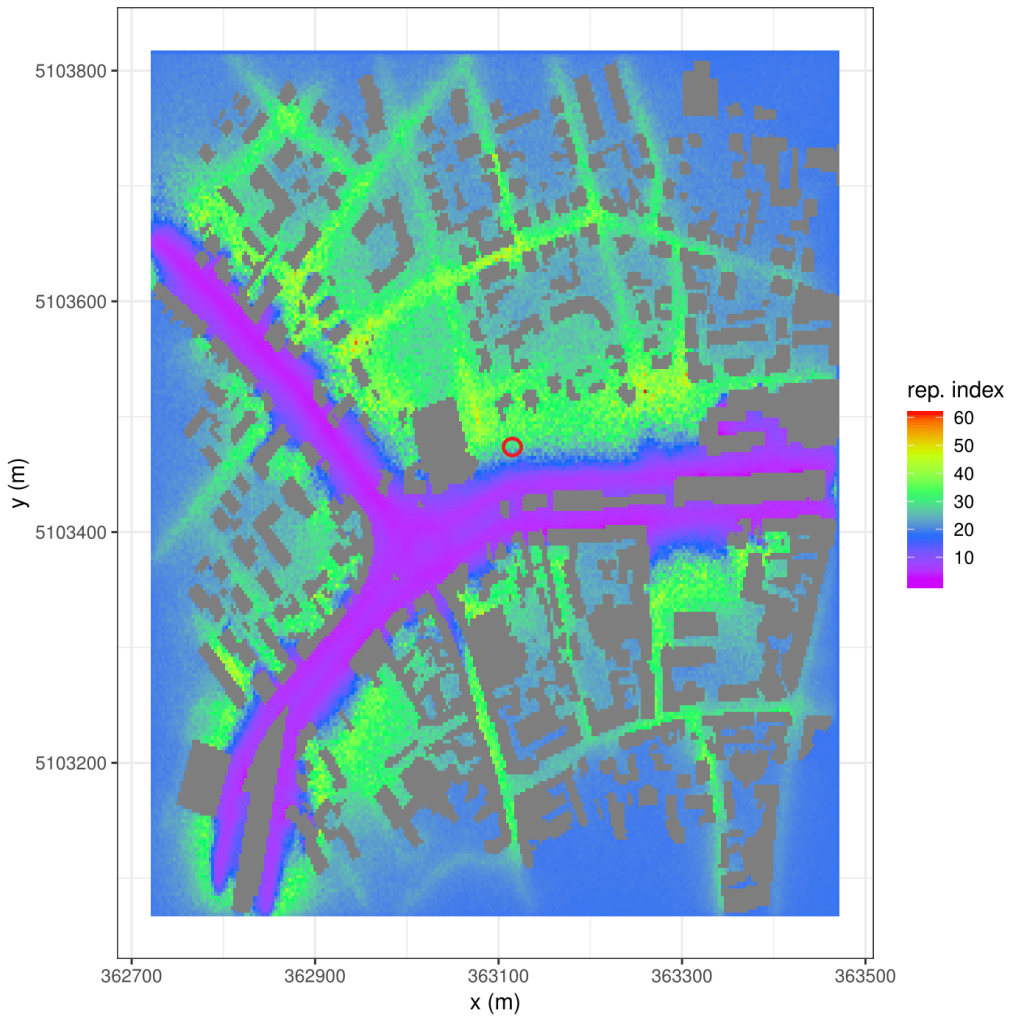


Figure 4.21: Class-frequency weighted representativeness map. Representativeness index in arbitrary units.

Fortunately the traffic air quality station stays on a medium representativeness point.

4.5.4 Validation issues

The MSS suite has already been validated either on field (air quality/traffic measurements see Trini Castelli et al., 2017 and Ghermandi et al., 2016) and using CFD models (see Anfossi et al., 2008). This thesis mainly focuses on the operative aspects of running a model, in fact a detailed modeling of emissions goes beyond this work. Nonetheless, a trial has been made using microSPRAY to model the difference between the traffic station measures and some background station measures. The two stations considered were that of Viale San Daniele (traffic station) and that of Via Cairoli (background station). Unfortunately, the magnitude of the differences between such two stations the most of time is smaller than twice the instrumental error (0.5μg): in a set of three years hourly data differences (Viale San Daniele datum - Via Cairoli datum) the 77.4% was found to be comparable within the intrumental error and only the 5.1% was found to be positive and larger than twice the instrumental error. Things don't change much considering only half of the compatibility stripe: only the 15% of the data differences is larger than 0.5μg while 4% is smaller than 0.5μg. It is thus clear that the two stations cannot be considered as a couple of background-traffic air quality benzene measurements.

4.5.5 Orders of magnitude

It has already been remarked in the previous sections that microSPRAY has not reproduced the same order of magnitude of the measured concentration values (see figures 4.11, 4.20 and 4.18). This is related to the different scales involved in the time series, that, according to a well established model (see Kakosimos et al., 2010) include a regional background signal, an urban background signal and a very local signal (figure 4.22). Benzene has a poor documentation about typical magnitudes of the fluctuations' periods and amplitudes related to the various scales, and thus a comparison with this work is interdicted. Simulations of one day over a limited spatial domain do not reproduce enough accumulation in order to justify the non zero nightly measured concentrations, but keep track of the traffic flux information at points at a relatively large distance from the sources (in pretty stable meteorological conditions) and indeed they show a local feature with a high spatial concentration gradient for the signal amplitude, but the term "local" should be better characterised in the framework of atmospheric pollution.

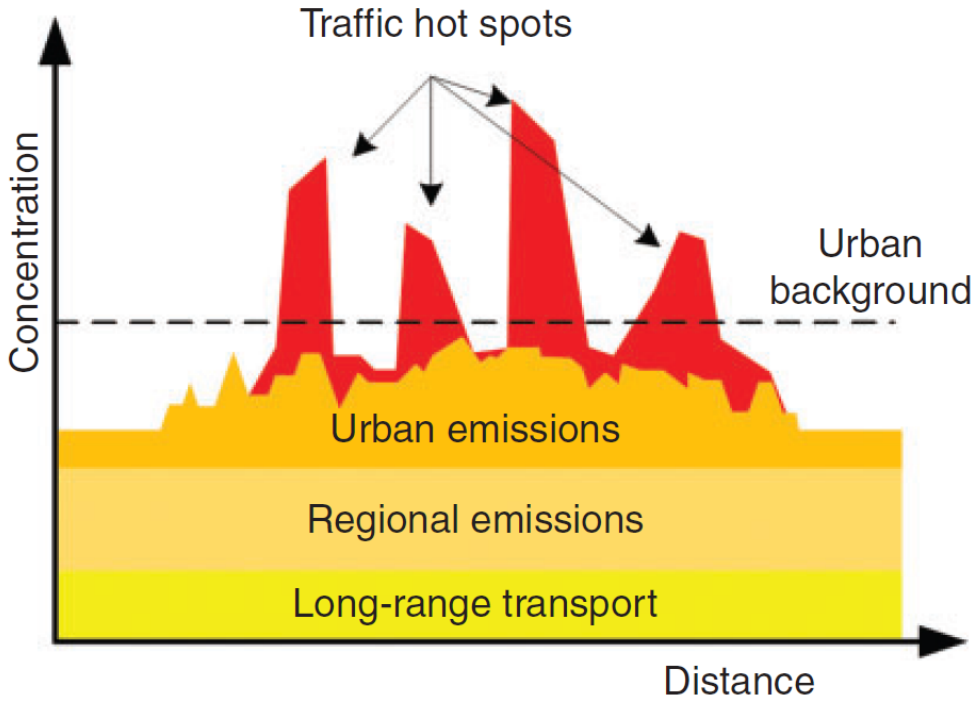


Figure 4.22: Spatial and temporal scales involved in the pollution phenomena. Adapted from Kakosimos et al., 2010.

Chapter 5

Conclusions

5.1 Summary of Achievements

This work has proved that the microSPRAY model gives a good representation of the turbulent diffusion process at the microscale, although there exist some limits, mainly due to the necessary reduction of the domain to a computation grid and to the temporal resolution. These limits affect directly the micrometeorological model, microSWIFT and are in some sense resolved by the Lagrangian dispersion model by interpolating meteorological fields linearly in space and time, so as to allow complete freedom of the Lagrangian particels' motion (up to the theoretical limits imposed by the turbulent motion scheme). Sharply peaked concentration values near the source have been reproduced in the case of meteorologically stable days, according to many experiments and simulations with other models, also at a lower resolution. In unstable meteorological conditions the model correctly simulates the strong dispersive character, with a breakdown of the spatial concentration gradient. Comparison of the simulated data with measurements highlighted the lack of a background component in the simulated concentrations (see figure 4.11) that persists at night, and revealed that traffic hot-spots give a small contribution to the overall concentration, or, better, a contribution at least comparable with the sum of the urban and regional background concentrations. It must be said that in the last decade benzene emissions due to the urban transport have undergone a strong abatement, and that, by consequence, measurements of a real background concentration and of a real local concentration are a mutable concept over time. However, a representativeness study is independent on the

true magnitude of the concentration: what matters in this case is the comparison between the points included the domain. In this work a representativeness index for a benzene monitoring station has been produced. Results of this analysis are satisfactory.

5.2 Future Work

The microSPRAY model has proven to be a reliable instrument to describe atmospherical dispersion; microSWIFT instead is limited by the input meteorological resolutions and thus time evolution is implemented in the Lagrangian dispersion model by a simple linear interpolation between low resolution values. A point of strength of microSWIFT is the easy assimilation of meteorologic input both from sparse true measurements points and from gridded simulated values. An improvement of microSWIFT conserving energy and momentum exists but it's not yet operational. The limits of microSWIFT in resolving very local and limited in time wind fluctuations are considered hard to fix since basics of the codes do not allow to reproduce small eddies. For a better reproduction of high resolution wind field features a Large Eddy Simulation (LES) is considered a more appropriate approach. Anyway, because of the very high computational cost of a LES simulation in comparison to those carried on by means of microSWIFT and microSPRAY, the last are commonly adopted. A separate work on the background and local character (in the sense of concentration measurements) of air quality stations should be carried out for future works on atmospheric dispersion; this should also provide a deeper understanding on representativeness studies.

Appendices

Appendix A

Eulerian and Lagrangian timescales, frozen turbulence hypothesis

Assuming an ensemble of particles moving in a turbulent flow, the displacement in the jth direction, at a time t after the release, is defined as

$$\overline{x_j'^2(t)} = \overline{(x_j^i(t) - \overline{x_j^i(t)})^2}, \tag{A.1}$$

where $x_j^i(t)$ is the position of the ith particle, and the overline represents the average over all the particles. Following the classical analysis of Taylor (1921), this displacement is expressed as a function of the properties of the turbulent flow according to

$$\overline{x_j'^2(t)} = 2\sigma_j^2 \int_0^t \int_0^{t'} R_j^L(\tau)\, d\tau\, dt', \tag{A.2}$$

where σ_j is the (square root of the) velocity variance, and $R_j^L(\tau)$ is the Lagrangian autocorrelation function, defined as

$$R_j^L(\tau) = \frac{\overline{u_j'(t)u_j'(t+\tau)}}{\sigma_j^2}, \tag{A.3}$$

Here, $u_j'(t) = u_j^i(t) - \overline{u_j^i(t)}$ is the velocity fluctuation of the ith particle at time t, and τ is the time lag. Relationship A.2 has two analytical limits for short and large times, respectively,

$$\begin{aligned} \overline{x_j'^2(t)} &= \sigma_j^2 t^2 \qquad t << T_j^L \quad \text{and} \\ \overline{x_j'^2(t)} &= 2\sigma_j^2 T_j^L t \quad t >> T_j^L, \end{aligned} \tag{A.4}$$

where the Lagrangian integral time scale T_j^L is defined as

$$T_j^L = \int_0^\infty R_j^L(\tau)\, d\tau$$

There is a large uncertaintiy in the value of the Lagrangian time scale and its dependence to other variables of the ABL, with values that range from 80s up to 10000s. Lagrangian statistics are seldom measured experimentally in the ABL, and T_j^L is normally inferred from Eulerian statistics using the following relationship:

$$T_j^L = \beta_j T_j^E$$

where T_j^E is the Eulerian integral time scale, and β is the ratio of the Lagrangian to Eulerian time scales. Both atmospheric measurements and numerical experiments have produced a value of β_j ranging between 3 and 5.

Following Pasquill (1974), Taylor's frozen turbulence hypothesis is applied to autocorrelations as follows:

$$R(t) = R(x) \qquad \text{if} \quad x = Ut.$$

This in turn leads to the relationship between Eulerian length and time scales, $UT^E = \lambda^E$ where U is the mean wind in the direction along which data are collected. From numerical experiments with LES simulations (see Dosio et al., 2005 and Anfossi et al., 2006) the frozen turbulence hypothesis has been verified to a good extent and the equivalence between Lagrangian particle displacement $\overline{x_j'^2}$ and the Eulerian dispersion parameter $\sigma_{x_j}^2$ has been established (for practical purposes) thus legitmizing the use of σ_i together with T_{Li} in the diffusion coefficient appearing in the microSPRAY algorithm (see section 3.3).

Appendix B

Basics of Itō calculus

Material has been selected from Calin, 2012 and Maurizi, 2013.

B.1 Wiener process

A Wiener process (or Brownian motion process) is a stochastic process W_t which satisfies the following conditions:

1. the process starts at the origin, $W_0 = 0$
2. W_t has stationary, independent increments
3. the process W_t is continuous in t
4. the increments $W_t - W_s$ are normally distributed with mean zero and variance $|t - s|$,

$$W_t - W_s \sim N(0, |t - s|)$$

B.2 Nonanticipating processes

Consider the Wiener process W_t. A process F_t is called a *nonanticipating process* if F_t is independent of any future increment $W_{t'} - W_t$ for any t and t' with $t < t'$. Consequently, the process F_t is independent of the behaviour of the Wiener process in the future, i.e. it cannot anticipate the future. For instance, $W_t, e^{W_t}, W_t^2 - t$ are examples of nonanticipating processes,

while $W_{t+1}, \frac{1}{2}(W_{t+1} - W_t)^2$ are not. Nonanticipating processes are important because the Itō integral concept applies only to them.

B.3 The Itō integral

Consider $0 \leq a < b$ and let $F_t = f(W_t, t)$ be a nonanticipating process with

$$E\left[\int_a^b F_t^2 dt\right] < \infty.$$

Divide the interval $[a, b]$ into n subintervals using the partition points

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b,$$

and consider the partial sums

$$S_n = \sum_{i=0}^{n-1} F_{t_i}(W_{t_i+1} - W_{t_i}).$$

It is worth noting that the intermediate points are the left endpoints of each interval, and this is the way they should be always chosen. Since the process F_t is nonanticipating the random variables F_{t_i} and $W_{t_i+1} - W_{t_i}$ are independent; this is an important feature in the definition of the Itō integral.

The Itō integral is the limit (in the mean square sense) of the partial sums S_n

$$ms \lim_{n\to\infty} S_n = \int_a^b F_t \, dW_t,$$

provided the limit exists.

B.4 Itō differentiation

Most stochastic processes are not differentiable. For instance, the Wiener process W_t is a continuous pocess which is nowhere differentiable. Hence, derivatives like $\frac{dW_t}{dt}$ do not make sense

in stochastic calculus. The only quantities allowed to be used are the infinitesimal changes of the process, in our case, dW_t.

Basic properties of a Wiener process (with $s < t$) are:

- $E[(W_t - W_s)^2] = t - s$
- $Var[(W_t - W_s)^2] = 2(t - s)^2$

which are easily demonstrated recalling that $\frac{W_t - W_s}{\sqrt{t-s}} \sim N(0, 1)$.The infinitesimal version of the previous results is obtained by replacing $t - s$ with dt

- $E[dW_t^2] = dt$
- $Var[dW_t^2] = 2dt^2$

Some very useful properties for manipulating stochastic differentials can be derived from the above relations, namely $(dW_t)^2 = dt$ and $dW_t dt = 0$

B.5 The Kolmogorov and Monin-Yaglom relations for the local structure of turbulence

Kolmogorov and Monin-Yaglom equations were the first two equations of the "dynamic theory" of the local structure of turbulence. The name "dynamic theory" was originated by Monin and Yaglom to mean the derivation of the equations relating structure functions by the use of the Navier-Stokes equation and/or the scalar conservation equation, and the investigation of the resulting statistical equations. Here below the two fundamental equations are reported.

$$\begin{aligned} S_E^{(2)} &:= \overline{((\mathbf{u}(\mathbf{x} + \Delta\mathbf{r}) - \mathbf{u}(\mathbf{x})) \cdot \Delta\mathbf{r}/\Delta r)^2} = C_k(\varepsilon\Delta r)^{2/3} \quad \text{for} \quad \eta << \Delta r << L_E \\ S_{Li}^{(2)} &:= \overline{(v_i(t + \Delta t) - v_i(t))^2} = C_0(\varepsilon\Delta t) \qquad \text{for} \quad \tau_\eta << \Delta t << T_i^L \end{aligned} \tag{B.1}$$

where η and τ_η are the Kolmogorov length scale and time scale respectively, whereas L_E and T_i^L are the integral Eulerian length scale and Lagrangian time scale respectively. The second equation of the set B.1 is the non-differential form of equation 2.40 in section 2.6.

Bibliography

D. Anfossi, E. Ferrero, G. Brusasca, A. Marzorati, and G. Tinarelli. A simple way of computing buoyant plume rise in lagrangian stochastic dispersion models. *Atmospheric Environment. Part A. General Topics*, 27(9):1443 – 1451, 1993. ISSN 0960-1686. doi: https://doi.org/10.1016/0960-1686(93)90130-Q. URL http://www.sciencedirect.com/science/article/pii/096016869390130Q.

D. Anfossi, U. Rizza, D. G. Mangia, C., and E. Pereira Marques Filho. Estimation of the ration between the lagrangian and eulerian time scales in an atmospheric boundary layer generated by large eddy simulation. pages 326–337, Jan 2006. URL https://doi.org/10.1016/j.atmosenv.2005.09.041.

D. Anfossi, G. Tinarelli, S. Trini Castelli, J. Commanay, and M. Nibart. Microspray simulation of dense gas dispersion in complex terrain. 2008.

G. Barenblatt. *Scaling*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2003. ISBN 9780521533942. URL https://books.google.it/books?id=05zBYET6tR0C.

G. A. Briggs. Plume rise. *USAEC Critical Review Series*, page 81, 1969. URL https://doi.org/10.1098/rsta.1969.0048.

O. Calin. *An Introduction to Stochastic Calculus with Applications to Finance*. 2012. URL https://people.emich.edu/ocalin/Teaching_files/D18N.pdf.

J. Cardozo and D. Sanchez. Air pollution near arterial roads: An experimental and modelling study. *Atmos. Chem. Phys. Discuss.*, 2017. URL https://doi.org/10.5194/acp-2017-753-AC1.

A. Dosio, J. De Arellano, A. Holtslag, and P. Builtjes. Relating eulerian and lagrangian statistics for the turbulent dispersion in the atmospheric convective boundary layer. *Journal of the Atmospheric Sciences*, 2005. ISSN 1520-0469. URL https://doi.org/10.1175/JAS3393.1.

T. T. Fujita. Tornadoes and downbursts in the context of generalized planetary scales. *Journal of the Atmospheric Sciences*, 38(8):1511–1534, 1981. doi: 10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2.

G. Ghermandi, S. Fabbi, A. Bigi, and L. Torreggiani. Dispersione atmosferica a microscala di emissioni veicolari da flussi di traffico rilevati automaticamente e confronto con misure di qualitÀ dell'aria. *Ingegneria dell'Ambiente*, 3(3), 2016. ISSN 2420-8256. URL https://www.ledijournals.com/ojs/index.php/IngegneriadellAmbiente/article/view/711.

J. R. Holton. *An Introduction to Dynamic Meteorology*. An Introduction to Dynamic Meteorology. Academic Press, 1979. ISBN 9780123543608. URL https://books.google.it/books?id=ejZRAAAAMAAJ.

IPCC. *IPCC Fifth Assessment Report (AR5)*. WMO, IPCC Secretariat, 2013. URL https://books.google.it/books?id=eGntrQEACAAJ.

K. Kakosimos, O. Hertel, M. Ketzel, and B. Ruwim. Operational street pollution model (ospm) - a review performed application and validation studies, and future prospects. *Environmental Chemistry*, pages 485–503, 2010. URL https://doi.org/10.1071/EN10070.

P. Masai. Valutazione della qualità di diverse parametrizzazioni dello strato limite atmosferico implementate nel modello wrf. un caso studio annuale sul friuli venezia giulia. Technical report, Mar 2018.

A. Maurizi. Are lagrangian stochastic models at odds with statistical theories of relative dispersion? Jul 2013.

F. Montanari. Relazione qualita' dell'aria, gorizia. Technical report, ARPA FVG, Feb 2017.

I. Orlanski. A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, 56:527–530, 1975.

A. P. Reynolds, G. Richards, B. de la Iglesia, and V. J. Rayward-Smith. Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, 5(4):475–504, Dec 2006. ISSN 1572-9214. doi: 10.1007/s10852-005-9022-1. URL https://doi.org/10.1007/s10852-005-9022-1.

H. Risken. *Fokker-Planck Equation*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996. ISBN 978-3-642-61544-3. doi: 10.1007/978-3-642-61544-3_4. URL https://doi.org/10.1007/978-3-642-61544-3_4.

H. Rodean. Notes on the langevin model for turbulent diffusion of "marked" particles. doi: 10.2172/10131402.

J. Seinfeld and S. Pandis. *Atmospheric chemistry and physics: from air pollution to climate change*. A Wiley interscience publication. Wiley, 1998. ISBN 9780471178156. URL https://books.google.it/books?id=lK8PAQAAMAAJ.

R. Sozzi. *La micrometeorologia e la dispersione di inquinanti in aria*. APAT-ACE, 2003.

R. B. Stull. *An Introduction to Boundary Layer Meteorology*. Atmospheric and Oceanographic Sciences Library. Springer Netherlands, 1988. ISBN 9789027727695. URL https://books.google.it/books?id=eRRz9RNvNOkC.

D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180:529–556, 1987. doi: 10.1017/S0022112087001940.

S. Trini Castelli, G. Tinarelli, and T. G. Reisin. Comparison of atmospheric modelling systems simulating the flow, turbulence and dispersion at the microscale within obstacles. *Environmental Fluid Mechanics*, 17(5):879–901, Oct 2017. ISSN 1573-1510. doi: 10.1007/s10652-017-9520-5. URL https://doi.org/10.1007/s10652-017-9520-5.

I. Van der Hoven. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *Journal of Meteorology*, 14(2):160–164, 1957. doi:

10.1175/1520-0469(1957)014<0160:PSOHWS>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1957)014<0160:PSOHWS>2.0.CO;2.

H. S. Wio, R. R. Deza, and J. M. Lopez. *An introduction to stochastic processes and nonequilibrium statistical physics*. World Scientific Publishing Co. Pte. Ltd., 2012. ISBN 9789814374798. doi: 10.1142/9789814374798_0001.