POLITECNICO
MILANO 1863

Politecnico di Milano
Department of Civil and Environmental Engineering
Doctoral Programme in Environmental and Infrastructure Engineering

Stochastic analysis of Troposphere's non-hydrostatic refractivity field for small- and mid-scale GNSS networks

Doctoral Dissertation of:
Alessandro Fermi

Supervisor:
Prof.ssa Giovanna Venuti

Tutor:
Prof.ssa Giovanna Venuti

The Chair of the Doctoral Program:
Prof. Alberto Guadagnini

XXIX cycle - 2013/2016

Contents

Abstract

Spatio-temporal Precipitable Water Vapour (PWV) monitoring at regional and small scale is important both to deeper understand the initiation and development of local convection phenomena and to better forecast local heavy rainfall events. Along with scientific reasons, an accurate monitoring of PWV would be important also to support the activities of regional environmental protection agencies. GNSS measurements represent an unexpected but interesting source of information to retrieve absolute and accurate time series of PWV measurements. In fact, the average zenith non-hydrostatic tropospheric delay (i.e. the Zenith Wet Delay) depends (linearly) on PWV. Thus, accurate estimates of GNSS signal tropospheric delays can lead to obtain reliable PWV data.
Our research lay within this field. It mainly focused on the analysis of GNSS observations from different GNSS networks with specific characteristics of receivers' electronics, spatial distribution and acquisition rate, in order to retrieve careful estimates of the non-hydrostatic tropospheric delay and to model its spatio-temporal variations. To this extent, we carried out a comprehensive analysis of different approaches and estimation strategies of GNSS-derived Zenith Wet Delays (ZWDs) and we carried out a re-processing of the GNSS data acquired by the network MIST, deployed within the ESA project METAWAVE. The MIST network consisted of double-frequency geodetic receivers and was a dense network, i.e. a network with inter-distances of nearby stations less than 15 km. We analyzed several aspects both on the relative processing strategy (Double Difference method) and on the absolute processing one (Precise Point Positioning strategy). We have made extensive comparisons among these methods. Along with this study, we gave a closer look at a mid-scale GNSS network, i.e. with nearby stations inter-distances of approximately 50 km. We analyzed two severe rainfall events occurred in Lombardy region on the 22nd and on the 26th of July 2016 . These two severe storms showed a different meteorological nature: the one on the 22nd of July was characterized by a strong advective behavior with a weather front propagating from South-East to North-West, whereas the event on the 26th showed strongly localized convective cells. To analyze whether ZWD time series would give evidence of these phenomena we processed GNSS data from 21st to 29th of July acquired by the interregional GNSS SPIN network. The aim of this analysis was assessing the accuracy of GNSS-derived ZWD and PWV and the potential use of GNSS observations for the determination of (2+1)-D ZWD maps - 2D for the spatial range and 1D for the time variable. We modelled the residual ZWD field – i.e. the ZWDs reduced by a height/time component - as an intrinsic stationary random field. We then computed for each epoch the spatial empirical variograms of the residual field and fitted these empirical spatial variograms with Fractional

Brownian motion models. At the end of the present work, we describe a procedure for the estimation of the mean transport velocity vector. This procedure relies on the conservation equation of the specific humidity and in its simplest form it turns out to be equivalent to Taylor's frozen flow hypothesis.

Chapter 1

Introduction

Spatio-temporal Precipitable Water Vapour (PWV) distribution monitoring at regional and small scale is important both to deeper understand the initiation and development of local convection phenomena and to better fore- and nowcast local heavy rainfall events (cf. [**De Haan(2006)**], [**Brenot et al.(2013)**] and the references therein).
Along with these scientific purposes an accurate monitoring of the distribution of PWV both spatially and temporally would be decisive also to support the activities of regional environmental protection agencies. In fact, the lack of accurate forecasts of localized convective rainfalls may bring inefficiencies in risk alerts and corresponding mitigation activities.
To this regards GNSS measurements represent an unexpected but interesting source of information to retrieve accurate time series of *absolute* PWV measurements at the receiver locations. In fact, the average zenith nonhydrostatic tropospheric delay, experienced by the GNSS signal when crossing Earth's low atmosphere (computed from the slant nonhydrostatic delays by suitable projections), depends (linearly) on the PWV (see for instance [**Bevis et al.(1992)**], [**Smith et al.(1953)**]).
Observations of PWV are hardly available with the time and space resolution needed to properly predict the highly variable behaviour of water vapour tropospheric content. Radiosondes, balloon-borne instrument packages that send temperature, humidity, and pressure data to the ground by radio signals, allowing for a detailed evaluation of the water vapour content along the ballon profile, are rather expensive instruments. Therefore, they are launched only in few locations, thousand of km apart, once or twice a day, thus assuring vapour information at a very low spatial and temporal resolution [**Elliott1991**]. Ground based Water Vapour Radiometers (WVR), currently used for water vapour sounding, allow for a proper time resolution monitoring, but again high operational costs limit their spatial distribution [**Resch1984**]. Satellite borne WVRs, on the other hand, can produce a proper space sampling, but the time resolution is still too poor [**Princea1998**]. Therefore accurate estimates of GNSS signal tropospheric delays may lead to reliable PWV data.
The research presented in this thesis lies within this field. It mainly focused indeed on the analysis of GNSS observations from different GNSS networks in order to retrieve precise estimates of the nonhydrostatic tropospheric delay and to model

its spatio-temporal variations at mid- and small spatial scale. The GNSS networks we analized had specific characteristics about the receivers' electronics, their spatial distribution and acquisition rate (see Chapter 4 and 5 for further details).
To be more precise, we carried out a comprehensive analysis of different approaches and estimation strategies of GPS-derived Zenith Wet Delays (ZWDs). To this regards we carried out a re-processing activity of the GPS data acquired by the network MisT, deployed within the ESA project METAWAVE.
The MisT network consisted of double-frequency geodetic receivers and could be defined as a dense network, i.e. a network with inter-distances of nearby stations less than 15 km. We analyzed several aspects both on the relative processing strategy, i.e. Double Difference method (shortened by the symbol DD), and on the absolute processing one, namely Precise Point Positioning strategy (denoted by PPP). We have made extensive comparisons among these methods.
In particular, within the DD approach we investigated the impact of using single frequency L1 observations with respect to iono-free L3 combinations for each network's station. It turned out that for very short baselines, like the ones characterizing the MisT network, L3 combinations yield noisier and less accurate estimates both of the coordinates and of the ZWDs. Single frequency observations, instead, lead to more repeatable and accurate results (mainly for coordinates). In our results, it was remarkable to notice that in the comparisons between coordinates and ZWDs, obtained by using either L1 or L3 observables, standard deviations range between one-millimeter precision up to a few millimeters value, whereas a bias up to 7/8 mm was present, especially in the North coordinate.
This investigation lead us to conclude that for dense networks the deployment of double frequency receivers for the whole network is unnecessary and L1-only data may be used, showing a more regular behavior. It may be, however, necessary to use at minimum two double frequency receivers for an internal and consistent data quality assessment.
Further on this topic, we made extensive analysis of the estimates obtained by a relative positioning approach and by an absolute strategy as modelled in the PPP. In fact we analyzed first the residuals of UTM coordinates time series, obtained by processing L1-only observations in DD and L3 combinations in PPP with respect to their corresponding regression lines. This showed the repeatability of DD estimates to range between few millimeters, while the repeatability of PPP estimates is of course higher but still below 1 cm.
Thereafter we compared the coordinates and ZWD time series for any MIST station obtained by processing L1-only observations and L3 combinations with a PPP approach computing time series of differences. This has made evident that no significant bias arises in this case and that the standard deviation of the differences is at most 4 mm in planimetry and 9 mm both in the UP coordinates and in the ZWD time series. This variability is very likely a consequence of the observation equation model used in the PPP approach that requires to model all terms involving in the observation model and to introduce a priori estimates of the satellite's clock offset.
This result suggests in any case that as for the estimation of ZWD both strategies are adequate, since the ZWD-derived Precipitable Water Vapor (PWV) is retrieved from the ZWD by a scaling factor of approximately 0.17 ([**Kleijer(2004)**]). By

means of a simple application of the variance propagation law, we then deduce that the variability of the difference time series decreases by a similar factor.
Finally, we computed and compared time series of baselines, i.e. the time series of the baseline of each station to the reference station adjusted in the DD procedure and the time series of the differences of PPP estimates for the same station and reference one. This analysis seemed to suggest that besides PPP residual errors, the time series of PPP-derived ZWD baseline accounts for the same water vapor signal as the time series of DD ZWD baseline (without introducing any a-priori artificial correlation in the estimates inherent to the processing strategy). Moreover, this study lead us to conclude that baselines shorter that 1 km length are not useful to monitor ZWD variations. These results were obtained applying a batch estimation method as implemented in the BERNESE software v5.2 and in its built-in automatic procedure BPE (cf. [**Dach et al.(2015)**]).
Moreover, in order to carry out the analysis we have also developed ad-hoc SW tools that contribute to build a "data warehouse". This software manages the import of the data from separate output BERNESE files and organize these data in a coherent archive and allows to perform different analysis on the dataset by applying several statistical algorithms. This SW has been written in the MATLAB programming language.
Along with the study described above regarding the GNSS network MIST, we have given a closer look at a GNSS network where the inter-distances of nearby stations are approximately 50 km, thus including the case of mid-scale and regional networks.
The case study we analyzed consisted in one week of July 2016, where two severe rainfall events occurred in Lombardy region on the 22nd and 26th. These two severe storms showed a quite different meteorological nature: the one on the 22nd of July was characterized by a strong advective behavior with a weather front propagating from South-East to North-West, whereas the event on the 26th showed strongly localized convective cells.
To analyze whether and how GNSS-derived ZWD time series would give evidence of these phenomena we processed GNSS data from 21st to 29th of July acquired by the interregional GNSS SPIN network. This network consists of 30 dual-frequency geodetic receivers covering quite homogenously the two Italian regions Lombardy and Piedmont. The inter-distances of closeby stations belonging to the SPIN network range between 40 and 50 km.
To analyze this case study we adopted a PPP strategy (thus considering L3 observables) and for each station, we computed a ZWD time series with time resolution of 10 minute. The aim of this analysis was assessing the accuracy of GNSS-derived ZWD (and consequently PWV) and the potential use of GNSS observations for the determination of (2+1)-D water vapor maps - 2D for the spatial range and 1D for the time variable.
We modelled the residual ZWD field – i.e. the estimated ZWDs suitably reduced by an estimate of their mean height component - as an intrinsic stationary random field. We then computed for each epoch the spatial empirical variograms of the residual field and fitted these empirical spatial variograms with Fractional Brownian motion models ([**Shevchenko(2014)**]) of the form

$$\gamma(\|h\|)_t = C_t \cdot (\|h\|)^{\alpha_t}, \tag{1.1}$$

where $C_t \in \mathbb{R}$ and $0 < \alpha_t < 2$, for all available epochs t.
These models seem to retain a physical meaning (at least asymptotically), as they have been used and extensively studied in the turbulence theory of fluids ([**Pope(2000)**]).
Even though the theory predicts the values of the model parameters, in particular of the exponent, we estimated them epoch-wise by applying a least square procedure, with the aim to find the best data-driven model parameters possible. Comparing the theoretical value with our own estimates, we have found that in our case the exponent ranges from about 0.3 to 1.5, thus including values significantly different from that foreseen by the theory (see [**Tatarski(1961)**], [**Pope(2000)**]). On the other side, fractional Brownian motion models seem to represent faithful models for the computed empirical variograms in case of such regional network with stations' inter-distances up to 50 km.
Using these models, we were able to predict the values of the residual ZWD on a uniform grid applying a simple Kriging algorithm (see [**Wackernagel(2003)**]). This yields a stack of 2D residual ZWD maps parametrized by time.
Once again, to perform this analysis we have used and developed further the "data warehouse" SW, implementing several new modules, i.e. for instance the routine for computing spatial empirical variograms, for fitting these variograms with Fractional Brownian motion models, for setting up the model parameters and finally, for creating the maps.
At this point, in order to retrieve accurate PWV maps with absolute values, a last step is required.
We need indeed to re-introduce topography-dependent information on the network area, assessing the mean component of the ZWD due to the height at each grid point.
One possible way to do that would consist in selecting a proper Digital Elevation Model (DEM) of the network area, in order to estimate as accurately as possible the height of each grid point, and then choosing a model that expresses the ZWD as a function of the height component. We could in this way give an estimate of the mean height-dependent component of the ZWD. Adding these mean contribution to the residual ZWD maps would yield absolute estimates of the ZWD values at all grid points taking into account the topography as well.
This procedure could be of some interest since it would allow to compare the resulting absolute ZWD (or equivalently PWV) 2D maps with images acquired by other sensors, in order to figure out whether indeed GNSS-derived maps retain dynamical meteorological information. This step, however, has not been concluded and the comparison of GNSS-derived absolute ZWD water vapor maps with other images represents a future research direction.
In the context of dense and mid-scale GNSS networks, we developed also an algorithm for the estimation of the mean transport velocity vector.
This algorithm relies on the Taylor's frozen flow hypothesis and gives rise to a time-varying but spatially constant advection vector. It is based on a least square estimation approach and on the theoretical variogram models needed to estimate

the 2D ZWD maps.
The purpose of applying this algorithm was to check whether at least the direction of the weather front propagation for the event on the 22nd would have been visible. However, for an orographically complex area, as that the SPIN GNSS network covers, it became clear that such method could be misleading. For these reasons, we thought about a refined dynamical model, which generalizes the model based on the frozen flow hypothesis and is based on the conservation law of specific humidity of a parcel of air. Being both time and space-varying, this model could overcome the problem found with the previous method.

Chapter 2

The Troposphere signal in Global Navigation Satellite Systems

2.1 Introduction

Space geodetic measurement techniques, as for instance VLBI and GNSS, take advantage of the propagation time of signals, in the microwave spectrum, as a measure of the distance. When such a signal travels through the Earth's atmosphere, it affects the electromagnetic wave in three main ways

1. it causes a propagation delay

2. it causes a bending of the ray path of the signal

3. it absorbs part of the energy of the wave

The propagation delay is mainly a consequence of the *refractivity* of the medium where the electromagnetic wave propagates through. In this Chapter we will only deal with the propagation delay in the neutral atmosphere (or Troposphere), i.e. the lower part of Earth's atmoshere that ranges from the ground up to approximately 10/12 km.
We will describe the main components of the delay models typically used for the a-priori correction of the GNSS observations. Furthermore we will also introduce the main techniques and approaches employed for estimating the non-hydrostatic delay. This part of the total delay, indeed, is not easily corrected by means of a priori information or models and it needs to be assessed by statistical inferential methods.
The present Chapter introduces the atmospheric variables we will deal with in the rest of this work. The aim is to give an overview of the most important quantities and results in modelling the GNSS signal's propagation delay. The intention is not to be exhaustive by explaining all the existing models, but only to give an overview to the subject. A more comprehensive introduction can be found in, for example, [**Kleijer(2004)**], [**Dach et al.(2015)**] an the references therein.

2.2 Physics of the neutral atmosphere

This section deals with the physical background that is needed to derive the GNSS signal's troposphere delay models. We will first briefly describe the atmospheric layers and then the governing laws of the dynamical evolution of the neutral atmosphere. After that we will introduce the refractivity of the air, which represents for our purposes the main concept of this sectin. In fact the GNSS signal's propagation delay may be viewed as a linear functional of the refractivity field of air. We will finally present the decomposition of the refractivity into the hydrostatic and the non-hydrostatic components, explaining the main differences between these two parts. This section is based on [**Kleijer(2004)**].
The Earth's atmosphere can be roughly subdivided in different concentric layers, where the common centre is of course the centre of the Earth. The definition and the basic characteristics of these layers depends on the purpose for which the subdivision is made. If the atmosphere is studied from a physical-chemical perspective then the well-known characteristic features are temperature, pressure and ionization. Instead, if it is studied as a medium for electromagnetic waves then an important characteristic is the propagation (cf. [**Kleijer(2004)**] and the references therein). Analyzing the atmosphere by this latter case, i.e. by the way electromagnetic waves propagate through it leads to a subdivision of the atmosphere into two broad layers, namely the troposphere and the ionosphere. The ionosphere is the upper part of the atmosphere and can be taken to be ranging from about 15 km up to approximately 1000 km altitude. For electromagnetic waves with wavelengths ranging from some centimeters to a few millimeters the ionosphere is a dispersive medium, i.e. the propagation delay depends on the frequency of the wave. For the same class of waves, instead, the troposphere turns out to be a nondispersive medium, i.e. a medium for which the propagation delay does not depend on the wave's frequency. For this reason the troposphere is also referred to as *neutral atmosphere*.
In general, the temperature in the troposphere decreases almost linearly with height . At the top of the troposphere, i.e. at a variable height of approximately $9 - 16$ km above the mean sea level, the temperature remains constant for about another 10 km. This stratum of the neutral atmosphere can be referred to as tropopause in order to explicitly highlight this feature.
The neutral atmosphere contains both dry air and water in gaseous or liquid state. By volume dry air contains on average approximately 78% of Nitrogen, 21% of Oxygen, 0.9% of argon and 0.04% of carbon dioxide (together with small amounts to other gases). From a dynamical point of view, dry air shows smooth variations both spatially and temporally. Moreover, it does not show significant changes with latitude. The amount of water vapour, on the contrary, present a highly variable behaviour both spatially and temporally and it is concentrated in the lowest part of the troposphere, especially in the Atmospheric Boundary Layer, up to approximately $5 - 6$ km. Water, both in gaseous and liquid phase, is the most important constituent in relation to weather processes, especially because the large amounts of energy released or absorbed in the condensation and evaporation processes.
For this reason all meteorological phenomena happen generally in the part of the neutral atmosphere ranging from the ground up to approximately 6 km. The dy-

namical evolution of the water vapour concentration and, more generally, of the atmosphere can be expressed by a set of differential equations. These equations are retrieved by applying the conservation of mass, both of dry air and water, energy, momentum and the equation of state for ideal gases (cf. [**Kalnay(2003)**])

$$\frac{d\boldsymbol{v}}{dt} = -\alpha\nabla p - \nabla\phi + F - 2\Omega \times \boldsymbol{v} \tag{2.1}$$

$$\frac{\partial\rho}{\partial t} = -\nabla \cdot (\rho\boldsymbol{v}) \tag{2.2}$$

$$p\alpha = RT \tag{2.3}$$

$$Q = C_p\frac{dT}{dt} - \alpha\frac{dp}{dt} \tag{2.4}$$

$$\frac{\partial\rho q}{\partial t} = -\nabla \cdot (\rho\boldsymbol{v}q) + \rho(E - C) \tag{2.5}$$

where $\boldsymbol{v} = (u, v, w)$ is the flow velocity field of the air, T is the temperature of a parcel of air, p is its the pressure, ρ denotes the mass density of a parcel of (moist) air, α is the specific volume of a parcel of air, C_p represents its specific heat at constant pressure and q stands for the specific humidity of a parcel of air.
The third Equation is the well known equation of state for ideal gases.
The last Equation, instead, represents the conservation equation for the specific humidity q. It simply states that the total amount of water vapor in a parcel of air is conserved as the parcel moves around, except when there are sources (evaporation E) and sinks (condensation C). For our purposes this equation is quite important and we will take advantage of it in Chapter 6 in order to describe an advection-diffusion model for the ZWD field.
When an electromagnetic wave travels through the neutral atmosphere it interacts with all the atmospheric constituents described above and among the other effects mentioned in the Introduction, it suffers a delay compared to its propagation in the vacuum. In fact the propagation speed of an electromagnetic wave in the vacuum is the speed of light c, whereas its propagation velocity in any other medium is $v \leq c$.
Hence the total delay of a radio signal due to the neutral atmosphere depends on the *refraction index* along the traveled path, i.e. the ratio between its actual speed and speed of light

$$n = \frac{c}{v} \tag{2.6}$$

Evidently the refraction index is an adimensional variable. In some cases, n can be modelled as a complex number, where the imaginary part is related to the absorption of energy while the real part reflects the propagation delay and bending (see [**Kleijer(2004)**]). However, for our scope we will always think of the refraction index as a real number, thus considering only the delay and bending part.
Since n is close to 1, the refraction index is usually replaced by the *refractivity* defined as

$$N = (n-1) \cdot 10^6 \tag{2.7}$$

Also the refractivity is an adimensional variable. The values of the refractivity for the atmosphere ranges between 0 and 300.
Taking into account that the atmosphere contains both dry air and water, typically in vapour phase, we may decompose the refractivity into a dry and wet component

$$N = N_d + N_v \tag{2.8}$$

where N_d is the refractivity of dry air, whereas N_v is the refractivity of water vapour.
Both these terms depends on pressure and temperature of the troposphere.
Indeed according to [**Thayer(1974)**] and [**Smith et al.(1953)**] for electromagnetic waves with frequencies up to approximately 20 GHz, an accurate formula for the refractivity of dry air and of water vapour is given by

$$N_d = k_1 \frac{P_d}{T} Z_d^{-1} \tag{2.9}$$

$$N_v = \left[k_2 \frac{e}{T} + k_3 \frac{e}{T^2} \right] Z_w^{-1} \tag{2.10}$$

where

- k_i are constants; $[Kmbar^{-1}], fori = 1, 2$ and $[K^2 mbar^{-1}], fori = 3$
- Z_d is the compressibility factor of dray air $[-]$
- Z_w is the compressibility factor of water vapor $[-]$

The compressibility factors take into account the non-ideal behaviour of dry air and water vapor and under average atmospheric conditions they are close to 1. In order to compute them there exist empirical formulas that let the compressibility factors depend again on temperature and pressure both of dry air and water vapour. The explicit formulas are reported in [**Kleijer(2004)**], pag. 28.
Here we limit ourselves in saying that these factors can be often considered to be equal to the unity with good approximation, which implies the following formulas for the dry and wet refractivity

$$N_d = k_1 \frac{P_d}{T} Z_d^{-1} \tag{2.11}$$

$$N_v = \left[k_2 \frac{e}{T} + k_3 \frac{e}{T^2} \right] Z_w^{-1} \tag{2.12}$$

The constants k_i have been determined empirically and values for them are reported, for instance, in [**Bevis et al.(1992)**].
As explained in [**Kleijer(2004)**] and the references therein, the dry refractivity and the linear term with respect to temperature in the wet refractivity are the result of the molecular polarization induced by the wave when travels through

the neutral atmosphere. The quadratic term, instead, is the result of the dipole structure of the water molecule and it is by far the largest part of the delay induced by water vapour.
Along with the splitting of the refractivity N into its dry and wet part, we may decompose N also in its *hydrostatic* and *nonhydrostatic* components.
More precisely, starting with Equation 2.9 and the equation of state for dry air and water vapour, we may write

$$\begin{aligned} N &= k_1 R_d \rho_d + k_2 R_w \rho_w + k_3 \frac{e}{T^2} Z_w^{-1} \\ &= k_1 R_d \rho_m - k_1 R_d \rho_w + k_2 R_w \rho_w + k_3 \frac{e}{T^2} Z_w^{-1} \\ &= k_1 R_d \rho_m + (k_2 - k_1 \frac{R_d}{R_w}) R_w \rho_w + k_3 \frac{e}{T^2} Z_w^{-1} \end{aligned} \tag{2.13}$$

where we considered the density of moist air given by $\rho_m = \rho_d + \rho_w$.
Setting

$$K_2 = k_2 - k_1 \frac{R_d}{R_w} \quad [Kmbar^{-1}] \tag{2.14}$$

and using the equation of state of water vapor, Equation 2.13 reads

$$\begin{aligned} N &= k_1 R_d \rho_m + (K_2 \frac{e}{T} + k_3 \frac{e}{T^2}) Z_w^{-1} \\ &= N_h + N_w \end{aligned} \tag{2.15}$$

The first term is called hydrostatic refractivity while the second term is the nonhydrostatic refractivity or with a slight abuse of language wet refractivity, even though this term refers to Equation 2.9. Unless otherwise stated, in this work we will always consider the decomposition 2.15 of the refractivity in hydrostatic and nonhydrostic component.
Also in this case the constants may be determined empirically and we refer to [**Kleijer(2004)**] for a table reporting different values for them.

2.3 Propagation of GNSS signals in Earth's troposphere

As already said many times, a radio signal travelling through Earth's neutral atmosphere interacts with the atoms and molecules present, mostly given by dry air and water vapour. This results in a change of the propagation speed and direction, i.e. the signal is refracted. In this section we will discuss in detail the effects of the propagation media on radio signals with frequencies up to approximately 20 GHz, considering therefore in particular GNSS signals. The section is based [**Teunissen and Kleusberg(1998)**], [**Kleijer(2004)**] and the references therein.
The delay experienced by a radio signal travelling through the neutral atmosphere is a linear functional of the refraction index n introduced in the previous section. The most basic physical principle for the propagation of an electromagnetic wave

is Fermat's principle that can be stated as follows

The path taken between two points by a ray of light, or any electomagnetic wave, is the path that can be traversed by the least time.

Keeping this in mind, we may define the elecromagnetic distance between a source point P and a target point Q as (see [**Kleijer(2004)**])

$$L = \int_{l(P,Q)} c \cdot dt = \int_{l(P,Q)} \frac{c}{v} \cdot ds = \int_{l(P,Q)} n(s) \cdot ds \tag{2.16}$$

where $l(P,Q)$ is the effective path followed by the signal, c is the speed of light, v is signal's effective propagation speed and n is the refraction index.
If we define the geometrical distance by

$$R = \int_r dr \tag{2.17}$$

where r is the geometric ideal path of the signal, namely a straight line from the source and target points, then the *slant total delay* of the radio signal is given by

$$\begin{aligned} d^s_{trop} &= L - R = \int_{l(P,Q)} n(s) \cdot ds - \int_r dr \\ &= \int_{l(P,Q)} (n(s) - 1) \cdot ds + \left[\int_{l(P,Q)} ds - \int_r dr \right] \end{aligned} \tag{2.18}$$

The first term on the right-hand side of the equation above is the delay caused by the refraction media, while the second term is the delay caused by ray bending. Introducing the refracted (i.e. apparent) angle θ and the non-refracted (namely the geometric) angle ψ, then the delay experienced by a signal propagating from the top of the neutral atmosphere to the ground can be in first approximation expressed by the formula ([**Teunissen and Kleusberg(1998)**])

$$\int_{r_s}^{r_a} (n(r) - 1) cosec(\theta(r)) dr + \left[\int_{r_s}^{r_a} cosec(\theta(r)) dr - \int_{r_s}^{r_a} cosec(\psi(r)) dr \right] \tag{2.19}$$

where r is the geocentric radius with r_s the radius of Earth's surface and r_a the radius at the top of the troposphere.
This formula holds for a spherically symmetric atmosphere for which the refraction index n depends only on the geocentric radius.
The integral equation 2.19 can be computed if the angles θ and ψ are known together with a vertical profile of the refraction index n. In most applications, however, a profile of the refraction index is not available and therefore one is forced to make use of another approach based on closed-form or approximation series that rely on simplified atmospheric models.
In the context of signals used in satellite ranging systems it is possible to estimate the slant delay by means of approximating analytical functions that models the atmospheric parameters around the receiver. Following Equation 2.15, the delay

due to the hydrostatic component of the refraction index and that due to the nonhydrostatic one are treated separately. Thus similarly to Eq. 2.15 we may decompose the total slant delay as

$$d^s_{trop} = d^s_h + d^s_w \tag{2.20}$$

From the adjustment of GNSS observations, however, it is not possible to deal directly with slant delays. In fact, their introduction in the adjustment model would yield an ill conditioned systems. Thus the neutral atmosphere is assumed to be homogeneously distributed on horizontal layers, i.e. spherically symmetric (cf. for instance [**Teunissen and Kleusberg(1998)**], [**Herring1992**]) and in this way, the path-depending delays can be expressed in terms of a common unknown delay parameter in the zenith direction above the receiver. More precisely each slant delay will be written as the product of a delay in the zenith direction with a proper coefficient, which projects the zenith delay onto the corresponding slant direction

$$d^s_{trop} = ZHD \cdot m_h(\psi) + ZWD \cdot m_w(\psi) \tag{2.21}$$

In the context of GNSS, the total slant delay d^s_{trop} will be more explicitly denoted by T^S_R where R and S denote respectively receiver and satellite (cf. the forthcoming section).
In the above equation ψ is the non-refracted elevation angle of the signal in its path from the satellite to the receiver, whereas the term ZHD stands for Zenith Hydrostatic Delay and ZWD for Zenith Wet Delay. Their sum yields the Zenith Path Delay or total Zenith Delay denoted by

$$ZPD = ZHD + ZWD \tag{2.22}$$

These parameters, with special emphasis on the ZWD, are the most important parameters dealt with in this work and in the forthcoming chapters we will see many of their features, properties and related issues.
The projecting functions m_h and m_w are called *mapping functions*. Among their properties, the mapping functions map the increase in delay with the corresponding decreasing of the signal's elevation angle. Moreover, they take account for average characteristics of the troposphere according to the latitude of the region where the receiver is located and the year's season.
Over the past decades, a variety of model profiles and mapping functions have been developed in order to assess the delay experienced by signals, which propagate in the neutral atmosphere at arbitrary elevation angles.
As explained in [**Teunissen and Kleusberg(1998)**] and the references therein, the elevation angle dependence of the tropospheric delay can be mathematically described by a continued fraction and most of the mapping functions developed so far are based on a truncation of that series. A review of the most common mapping functions can be found in [**Kleijer(2004)**]. Here we would like only to point out that in the present work we extensively used the Niell mapping function within the reprocessing activity carried out for the MisT network (see Chapter 4) and the Global Mapping Function (together with the Global Pressure and Temperature model GPT, as implemented in the GNSS data processing software BERNESE v.

5.2) for the research study realized for the GNSS SPIN network.

2.3.1 Zenith Hydrostatic Delay

The Zenith Hydrostatic Delay (ZHD) is the delay experienced by a signal caused only by the hydrostatic component of the neutral atmosphere.
Using the refractivity N (see Definition 2.7) we have

$$ZHD = 10^{-6} \int_{h_0}^{h_a} N_h(h) dh \tag{2.23}$$

where h_0 is the height of the surface above the mean sea level and h_a is the height of the top of the troposphere.
Applying Equation 2.15 we may express the ZHD as

$$ZHD = 10^{-6} k_1 R_d \int_{h_0}^{h_a} \rho_m(h) dh \tag{2.24}$$

The hydrostatic zenithal term accounts approximately for 90% of the total zenith delay and it can be computed with an accuracy of a few millimeters once surface pressure data are available and the atmosphere is assumed to be in hydrostatic equilibrium.
In general, a fluid is said to be in hydrostatic equilibrium if it is at rest or when its flow velocity at each point is constant over time. This occurs when external forces such as gravity are balanced by the pressure. Mathematically we may state the hydrostatic equilibrium as

$$\frac{dP}{dh} = -g(h)\rho_m(h) \tag{2.25}$$

where P denotes the pressure, g is the gravity acceleration, ρ_m the density of moist air and h the height.
Integrating Equation 2.25 we get

$$\int_{P_0}^{0} dP = -\int_{h_0}^{h_a} \rho_m(h) g(h) dh = -P_0 \tag{2.26}$$

with P_0 the surface pressure exerted by a air.
Introducing a mean gravity term g_m, which represents the gravity acceleration at the center of mass of a column of air, we express the ZHD as

$$ZHD = 10^{-6} k_1 \frac{R_d}{g_m} P_0 \tag{2.27}$$

The mean gravity term is given by

$$g_m = g_m^0 \cdot f(\phi, h) \tag{2.28}$$

where ϕ is the geodetic latitude and $g_m^0 = 9.784 m \cdot s^{-2}$.
As stated in [**Kleijer(2004)**], the function f has the form

$$f(\phi, h) = 1 - 0.00266cos2\phi - 0.00000028h \tag{2.29}$$

and it is very close to 1. Kleijer [**Kleijer(2004)**] reports empirical values for the constants appearing in Equations 2.24 and 2.28 in order to obtain the explicit formula for the computation of the ZHD

$$ZHD = [0.0022768m/mbar] \cdot \frac{P_0}{f(\phi, h)} \tag{2.30}$$

The accuracy of the term expressed in square brackets is $0.5mm/bar$ according to [**Kleijer(2004)**] and the references therein.
Thus, for any point of Earth's surface one can compute the ZHD provided that the pressure at that point is known. No further assumptions about pressure at different heights or on temperature is needed.

2.3.2 Zenith Wet Delay

The Zenith Wet Delay (ZWD) is the delay experienced by a signal caused by the nonhydrostatic part of the neutral atmosphere.
It is a linear functional of the wet refractivity and it is defined as follows

$$ZWD = \int_{h_0}^{h_a} N_w(h)dh \tag{2.31}$$

where h_0 denotes again the height of the surface above the mean sea level, h_a the height of the top of the troposphere and

$$N_w = K_2\frac{e}{T} + k_3\frac{e}{T^2}$$

with unit compressibility factor.
According to [**Kleijer(2004)**], if we define a mean temperature coefficient as

$$T_m = \frac{\int_{h_0}^{h_a} \frac{e}{T}dh}{\int_{h_0}^{h_a} \frac{e}{T^2}dh} \tag{2.32}$$

and use the definition of the nonhydrostatic refractivity the ZWD can be written in the following way

$$ZWD = 10^{-6}(K_2 + \frac{k_3}{T_m}) \int_{h_0}^{h_a} \frac{e}{T}dh \tag{2.33}$$

Thus, from this equation it turns out that one would need water vapour pressure and temperature vertical profiles in order to determine the ZWD.
Whereas the ZHD can be easily and accurately modelled (cf. the previous section), the same is not true for the ZWD. Indeed the water vapour in the troposphere is not well mixed as the dry air constituents and therefore its distribution is spatially and temporally inhomogeneous and highly variable, especially in the atmospheric boundary layer.
For this reason, in the case of satellite ranging systems the ZWD is a parameter that is usually estimated in the adjustment of the GNSS observations. We will

review in the next sections a few approaches to assess the ZWD by means of different adjustment techniques and we will present the results obtained in our research study related to the comparison between these methods for estimating the ZWD. Furthermore, due to its inherent probabilistic nature the spatial and temporal distribution of the ZWD values across a region covered by a GNSS network will be modelled by a stochastic field with suitable stationarity assumptions. We refer to the Chapter 3 for a theoretical introduction to random fields and their basic properties and to Chapter 5 for a thorough analysis of a case study.
For meteorological applications the ZWD is important because it can be directly linked to the Precipitable Water Vapour (PWV). The latter meteorological parameter is defined as

$$PWV = \frac{1}{\rho}\int_{h_0}^{h_a} \rho_w(h)dh \tag{2.34}$$

where ρ_w is the density of water vapour while $\rho = 1000\ kg/m^{-3}$ is the density of (liquid) water.
Using the equation of state of ideal gases we have that

$$\int_{h_0}^{h_a} \frac{e}{T}dh = R_w \int_{h_0}^{h_a} \rho_w(h)dh \tag{2.35}$$

and we may define a factor Q relating the PWV to the ZWD as follows

$$Q = 10^{-6}(K_2 + \frac{k_3}{T_m})R_w\rho_w \tag{2.36}$$

With this definition it results that

$$ZWD = Q \cdot PWV \tag{2.37}$$

that is there exists a linear relation between the PWV and ZWD.
The factor Q is an adimensional quantity that amounts approximately to 6.5. In [**Kleijer(2004)**] or [**Bevis et al.(1992)**], an empirical formula for the mean temperature T_m is given, from which Q can be computed with an accuracy of about 2%. We refer to [**Kleijer(2004)**] and the references therein for further details about this topic.

2.4 GNSS observation equations and estimation strategies

In this section we will briefly introduce the basic observation equation model for GPS signals. Many concepts we are going to present generalize to the other satellite ranging systems already in use or that are being developed, as for instance the Russian system Glonass and the European one GALILEO. Nevertheless, we will focus only on the U.S. system GPS for the sake of simplicity and since it is the GNSS we have used within our research study.
Along with the basic observation equations we will talk about two main GPS observations adjustment strategies, namely the Double Difference (DD) method and the Precise Point Positioning method (PPP), pointing out some of their basic features

we have used in the analysis of our case studies. These methods are extremely important and there is a huge literature on them, explaining all their features and properties. We refer for instance to [**Teunissen and Kleusberg(1998)**], [**Biagi(2009)**], [**Dach et al.(2015)**] and the references therein for a full treatment of these topics.
GPS positioning and monitoring of tropospheric variables is performed on observations made on the signal carrier phases and on the pseudo random codes (e.g. precise or coarse acquisition code) that module the carrier phases themselves. The observation derived from pseudo random codes is usually called *pseudorange* measurement and it is equal to the difference between the receiver time t_i at signal reception and satellite time t_k at signal trasmission scaled by the nominal speed of light c.
The *carrier phase* measurement $\Phi_R^S(t)$, instead, is the difference between the phase ϕ_i of the receiver internally generatd signal at signal reception time with the phase ϕ^k of the satellite generated signal at trasmission time multiplied by the wavelength λ of the carrier. Only the fractional carrier phase can be measured and therefore the integer number M of full cycles is unknown. In the liteature M is called the *carrier phase ambiguity*.
Following [**Teunissen and Kleusberg(1998)**], we obtain the GPS observation equations for the pseudorange and the carrier phase measurements as

$$\begin{aligned}\Phi_R^S(t) = \rho_R^S(t, t-\tau_R^S) - I_R^S + T_R^S + \delta m_R^S + c[dt_i(t) - dt^S(t-\tau_R^S)] + \\ c[\delta_R(t) + \delta^S(t-\tau_R^S)] + \lambda[\phi_R(t_0) - \phi^S(t_0)] + \lambda M_R^S + \epsilon_R^S\end{aligned} \tag{2.38}$$

and

$$\begin{aligned}P_R^S(t) = \rho_R^S(t, t-\tau_R^S) + I_R^S + T_R^S + \delta m_R^S + c[dt_i(t) - dt^S(t-\tau_R^S)] + \\ c[\delta_R(t) + \delta^S(t-\tau_R^S)] + e_R^S\end{aligned} \tag{2.39}$$

where

- τ_R^S is the travel time of the signal on its path from the receiver to the satellite
- $\rho_R^S(t, t-\tau_R^S)$ is the geometric distance between the receiver R and the satellite S, expressed in $[m]$
- I_R^S is the ionospheric delay ($[m]$)
- T_R^S is the troposphere's total slant delay ($[m]$); we have changed the notation of the total slant delay here, compared to the previous section, in order to clearly state the different terms involved in the GPS observation equation model
- $dt_i(t)$ and $dt^S(t-\tau_R^S)$ are respectively the receiver's clock offset and the satellite's clock offset

- $c[\delta_R(t)+\delta^S(t-\tau_R^S)]$ is the term related to the equipment delays, in particular due to electronics of both the receiver and the satellite
- δm_R^S is the error due to multipath effects
- in case of the carrier phase measurement $\Phi_R^S(t)$ we have the additional terms $\lambda[\phi_R(t_0)-\phi^S(t_0)]$ and λM_R^S which result respectively from the non-zero initial phases of the signal emitted by the satellite and that generated internally by the receiver and the carrier phase ambiguity.

Remark. It is widely known that the GPS signal is modulated on two carrier phases denoted by L1 and L2 and, in the modernisation program of the space segment, another carrier will be added, in order to provide a higher accuracy and even more services (see for instance [**Biagi(2009)**]).
For the sake of simplicity we did not explicitly introduce in Eqs 2.38 and 2.39 a dependency on the carrier frequency. However, this dependency must always be kept in mind, since some of the systematic and random errors present in the observation equation models differ slightly according to which carrier frequency they are referred to.
Having more than only one carrier frequency is extremely important because a suitable linear combination of the two carriers allows to drastically reduce the ionospheric delay I_R^S.
This linear combination is called the L3 combination, or iono-free combination, and it is given by the formula

$$L3_R^S(t) = \frac{f_1^2}{f_1^2 - f_2^2}\Phi 1_R^S(t) - \frac{f_2^2}{f_1^2 - f_2^2}\Phi 2_R^S(t) \tag{2.40}$$

with $f_1 = 19.029cm$ is the frequency of the carrier L1 and $f_2 = 24.421cm$ is that of the carrier L2.
This combination allows to remove the first-order effets of the ionosphere, which account for approximately 90% of the total ionospheric delay. The iono-free combination is typically used for high precision positioning purposes and within the GNSS data adjustment method called Precise Point Positioning (PPP).

The GNSS observation equations 2.38 and 2.39 are nonlinear equation. This is because the geometric distance $\rho_R^S(t, t-\tau_R^S)$ between receiver and satellite is a nonlinear function of the coordiantes of the receiver.
Hence in order to apply inferential and regression algorithms to these models, especially least square and Kalman filter algorithms, the observation equations need to be linearized. The linearization process relies mathematically on the Taylor approximation series of the nolinear term truncated at first order. A full treatment is reviewed for instance in [**Teunissen and Kleusberg(1998)**] or [**Biagi(2009)**] and we refer to these sources for all computations, subtleties and consequences of the linearization method.
Along with the assessment of the coordinates of the receiver, the linearized observation equations allow to give an estimate of the tropospheric delay T_R^S.

Recall that we cannot directly estimate the slant delays in the adjustment of GPS observations, rather we express the total slant delay as

$$T_R^s = ZHD \cdot m_h(\psi) + ZWD \cdot m_w(\psi)$$

We refer to the previous section for the explanation of all the terms in the above equation.
Since the tropospheric delay term depends atually linearly on the observation, this is exactly the term appearing in the linearized equations.
The Zenith Hydrostatic Delay ZHD can be evaluated and removed from observations by means of models requiring the knowledge of surface pressure as we have explicitly derived in the Equation 2.30.
On the contrary, due to the high variability of the temporal and spatial distribution of water vapour in the neutral atmosphere, the corresponding Zenith Wet Delay cannot be accurately determined as a function of known parameters. The wet delay is therefore estimated within the general adjustment of GNSS observations.
Different adjustment strategies can be adopted in the analysis of GNSS data to obtain the parameters of interest. For what concerns the ZWD parameter estimation, in this work we have extensively used the following ones:

- least-squares batch adjustment of (L1 observations or L3 iono-free combinations) Double Differences (DD), or Relative Positioning
- least-squares batch adjustment of (iono-free observations) Zero Differences, or Precise Point Positioning (PPP)

In the case studied we are going to present in the next Chapters, we have carried out post-processing data adjustments of GPS-only observations collected from stationary geodetic receivers.
In the Double Difference case, the adjustment is performed on the observations (L1-only or iono-free combination) simultaneously acquired by two stations to the same satellite pair. Using this particular topology this adjustment method allows the removal of some of the systematic effects common to the receivers and the satellites involved. In particular, in the Double Difference the receiver's clock offset $dt_i(t)$ and the satellite's clock offset $dt^S(t - \tau_R^S)$ do not appear anymore. Furthermore in case of phase measurements the fractional phase term disappears as well. Finally, the total atmospheric bias and error in the ephemeris of the satellites are slightly reduced. After some computation, it turns out that the final (nonlinear) equations both for pseudorange and phase measurements are given by the following expressions

$$\Phi_{R_i,R_j}^{S_l,S_k}(t) = \rho_{R_i,R_j}^{S_l,S_k}(t, t - \tau_R^S) - I_{R_i,R_j}^{S_l,S_k} + T_{R_i,R_j}^{S_l,S_k} + \lambda M_{R_i,R_j}^{S_l,S_k} + \epsilon_{R_i,R_j}^{S_l,S_k} \tag{2.41}$$

and

$$PR_i, R_j{}^{S_l,S_k}(t) = \rho R_i, R_j{}^{S_l,S_k}(t, t - \tau_R^S) + I_{R_i,R_j}^{S_l,S_k} + T_{R_i,R_j}^{S_l,S_k} + e_{R_i,R_j}^{S_l,S_k} \tag{2.42}$$

The complete derivation of the above equations and their linearization can be found in [**Teunissen and Kleusberg(1998)**] and [**Biagi(2009)**].
The Double Difference approach implies the relative positioning of one receiver with respect to the other one. In other words it does not allow to estimate separately the coordinates of the two receivers involved, but only the coordinates of the baseline vector between them.
For what concerns the ZWD, instead, when two stations are at a distance shorter than approximately 20 km the quasi parallelism of the signal paths of the two receivers to the same satellite, results in a quasi rank deficiency of the least-squares design matrix. More precisely, the coefficients in Equation 2.21 computed by means of the mapping functions have almost the same value and this implies the least square design matrix to become ill conditioned with a quasi-rank deficiency of 1.
This requires, therefore, the ZWD value of one station to be fixed. Due to this issue, Double Difference adjustment of nearby receivers data yields a relative estimate of the ZWD with respect to a reference one. The same happens in the multi-station adjustment of networks with interdistances among the stations of less than 20 km.
For longer baselines, instead, it is possible to give a separate estimate of the ZWD of two stations since the values of the mapping functions in this case differ significantly, avoiding the quasi-rank deficiency of the regression model.
In the static Precise Point Positioning (PPP) case, single-station (or undifferenced) observations are adjusted. This requires the accurate modelling of a number of systematic errors. Among them, let us mention the satellite attitude effects with respect to the receiver, as for instance phase wind-up correction and the antenna offset, the site displacement effects, such as ocean loading effects and solid Earth tides, Earth Rotation parameters and the relativistic effects. Furthermore, a necessary parameter that has to be introduced in the PPP processing is the estimate of the satellite clock bias. These clock corrections are determined from a globally distributed network of receivers and are made available by the International GPS Service (IGS) with different estimation rate. They are either directly estimated or for higher sampling rate, as for instance 30 sec they need to be interpolated. Nevertheless their behaviour turns out to be quite smooth behavior, allowing them to be interpolated with a high precision.
A further important point to mention is that in case of PPP it becomes essential to remove as much as possible the ionospheric error, since its presence would affect heavily the accuracy of the estimated parameters. Therefore, within this adjustment technique L3 iono-free observations are preferable. In those cases where it is not possible to build the iono-free combination (for instance when single-frequency observations only are available) it is mandatory the removal of the ionospheric delays by using global or local Ionosphere models.
An exact description of the PPP method with all its subtleties is certainly behind the scope of this thesis. A complete description with full details about the Precise Point Positioning can be found in [**Kouba and Heroux(2001)**] and the references therein.

Chapter 3

Stochastic modelling of the non-hydrostatic ZWD field

3.1 Stochastic processes

In this section we will introduce some key notions of Probability theory that will be used throughout this work. In particular we will give the definition of a stochastic, or random, process, which will be one of the most important concepts in the present thesis. Along with that we will recall various basic definitions naturally associated to a random process, in particular generalizing the notions of probability distribution, mean, variance and covariance of two random variables. Finally, we will introduce a wide-sense stationary stochastic process and attempt to outline the weaknesses of such family of random processes in the context of GNSS meteorology. This will motivate the need of a more general family of random processes able to better model meteorological fields. This section is largely based on [**Papoulis et al.(2002)**] and [**Wackernagel(2003)**] and we refer to them for further details.
Heuristically a random variable is a function X that associates a real number $X(\xi)$ to the outcome ξ of a random experiment Ω. A stochastic process is a rule for assigning to any outcome ξ a *function* $X_t(\xi) = X(t, \xi)$. More precisely

Definition 1. *A (real) stochastic process is a collection $X_t(\cdot) = X(t, \cdot)$ of (real) random variables parametrized by $t \in T$, where T is some indexing set. Typically we will assume that T be an interval of $\mathbb{R}^k$ with $k = 1, 2, 3, 4$. In case $T \subseteq \mathbb{R}^k$, with $k > 1$ then we will refer to $X_t(\cdot)$ as a random field.*

If $T \subseteq \mathbb{R}$, then the stochastic process is a *continuous-time* process. Conversely, if $T \subseteq \mathbb{Z}$ then it will be called a *discrete-time* process. Moreover, if the values of the stochastic process $X_t, t \in T$ are countable then it will said to be a *discrete-state* process. Otherwise it will be called a *continuous-state* process, or simply a continuous process. In the following we will always assume to deal with continuous-time and continuous random processes, the definitions for discrete-time and discrete-state processes being easily extended.
Given a continuous stochastic process $X_t, t \in T \subseteq \mathbb{R}$ its first-order distribution is defined as

$$F(x,t) = P(X_t < x) \qquad \forall x \in \mathbb{R} \tag{3.1}$$

Its derivative with respect to x

$$\frac{\partial}{\partial x} F(x,t) = f(x,t) \tag{3.2}$$

is called the first-order density of the stochastic process.
The $n-$th distribution of X_t is the joint distribution function $F(x_1,\ldots,x_n;t_1,\ldots,t_n)$ of the n random variables $X_{t_1},\ldots,X_{t_n}$.
For a complete determination of the statistical properties of a stochastic process X_t the $n-$th distribution function $F(x_1,\ldots,x_n;t_1,\ldots,t_n)$ should be known for all x_i, t_i and $n \in \mathbb{N}$. Nevertheless, in almost all applications one may happily restrict his attention to average quantities that shed light on the statistical behaviour of the process. These quantities depend on the distribution functions of X_t up to its second order and are defined as follows.

Definition 2. *1. Given a stochastic process X_t with (first-order) density $f(x,t)$, we define its mean as*

$$m(t) = \mathbb{E}(X_t) = \int_{\mathbb{R}} x f(x,t) dx \quad \forall t \in T \tag{3.3}$$

2. The autocorrelation function $R(t_1,t_2)$ is

$$R(t_1,t_2) = \mathbb{E}(X_{t_1} X_{t_2}) = \int_{\mathbb{R}^2} x_1 x_2 f(x_1,x_2;t_1,t_2) dx_1 dx_2 \quad \forall t \in T \tag{3.4}$$

3. The autocovariance function, or simpler, covariance function is defined as

$$C(t_1,t_2) = \mathbb{E}(X_{t_1} X_{t_2}) - m(t_1)m(t_2) \quad \forall t \in T \tag{3.5}$$

If $t_1 = t_2$ then $C(t_1,t_1)$ is the variance of the random variable X_{t_1}.

In case of a random field the above definitions can be *verbatim* generalized. A further important second-order statistics of a stochastic process is provided by the so-called structure functions, or variograms, that will be introduced later on in this chapter.

Example 1. *A trivial example of a stochastic process is provided by any deterministic signal $\boldsymbol{s}(t) = g(t)$. In this case*

$$m(t) = \mathbb{E}(g(t)) = g(t) \quad R(t_1,t_2) = g(t_1)g(t_2)$$

Example 2. *The random motion of a particle in a fluid resulting from its mutual collision with the fast-moving molecules in the fluid is typically called Brownian motion and it is described by a continuous-time stochastic process called Wiener process. This process occurs frequently in many branches of Science and Economics and is defined as follows*

Definition 3. *A Wiener process is a stochastic process $W_t, t \geq 0$ such that*

1. *W_0 is identically zero almost surely*
2. *for almost all $\xi \in \Omega$ the function $W_t(\xi)$ is continuous on $(0, +\infty)$*
3. *W_t has independent increments, i.e. for $t_1 < t_2 < \cdots < t_n$ the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent random variables*
4. *W_t has zero-mean normally distributed increments, i.e. $W_{t_2} - W_{t_1} \sim N(0, t_2 - t_1)$*

A Wiener process is a particular case of Gaussian processes*, i.e. those processes $X_t, t \in T$ such that for any finite subset $F \subset T$ the joint distribution of the corresponding set of random variables $X_t, t \in F$ is a multivariate Gaussian distribution. Gaussian random processes are a remarkable class and we will (implicitly) use them throughout this work.*
As already stated in the definition of the Wiener process, the mean of the process is identically zero. The covariance function instead is given by

$$C_W(t_1, t_2) = min(t_1, t_2) \tag{3.6}$$

This implies in particular that the variance $var(W_t) = t$ for all $t \geq 0$.

Among all stochastic processes, an important family is provided by the collection of stationary stochastic processes. Heuristically stationarity is rather an intuitive concept and refer to the idea that the statistical properties of a random process do not change over time. More precisely

Definition 4. *A stochastic process $X_t, t \in T$ is said to be wide-sense or weakly stationary if its mean is constant*

$$\mathbb{E}(X_t) = \mu \quad \forall t \in T \tag{3.7}$$

and its covariance function

$$C(t, s) = C(t - s) \quad \forall t \in T \tag{3.8}$$

Furthermore, if the covariance function only depends on $\tau = \|t - s\|$

$$C(t, s) = C(\tau) \tag{3.9}$$

then the process is called wide-sense (or weakly) stationary and isotropic.

Once again the above definition can be easily carried over to stochastic fields where $T \subseteq \mathbb{R}^k$ with $k = 2, 3, 4$. In this case, however, instead of wide-sense stationarity we will talk about wide-sense *homogeneity* and a random field will be called weakly *homogeneous* and *homogeneous* and *isotropic*.
One of the advantages of having a weakly stationary stochastic process relies on the fact that this assumption makes less involved to compute empirically the covariance function of the process and to fit it with a suitable theoretical model. Once a model for the covariance function is known, there exist statistical inferential methods allowing to make predictions of the stochastic process by means of that covariance model, e.g. collocation formulas (see for instance [**Moritz et al.(1967)**]).

Remark. Along with the weak stationarity just defined, there is also a notion of strong stationarity according to which a random process is strongly stationary if its statistics is invariant to a shift of the origin. In other words, X_t and X_{t+c} must have the same statistical properties for any $c \in \mathbb{R}$ such that $t + c \in T$. However, this notion of stationarity is far too strong and fails to hold true in many practical situations.
Besides the ideas of strong and weak stationarity there are other forms of stationarity. One of which will be treated extensively in the following section. For other notions we refer to [**Papoulis et al.(2002)**].

Example 3. *A continuous stochastic process $U_t, t \in \mathbb{R}$ is said to be white noise if*

$$\mathbb{E}(U_t) = 0 \quad \forall t \in \mathbb{R} \tag{3.10}$$

$$C(t,s) = \mathbb{E}(U_t U_s) = \sigma^2 \delta(t-s) \; \forall t \in \mathbb{R} \tag{3.11}$$

where $\delta(t-s)$ denotes the Dirac delta centered at $t-s$. Thus, from its very definition, it follows that each pair of variables U_t and U_s are uncorrelated and have the same variance σ^2.
If the white noise U_t is normally distributed then it is called Gaussian white noise. According to its definition it is immediate to see that a white noise process is a weakly stationary process. Furthermore, if the white noise is Gaussian then it is also strongly stationary. This is due to the fact that the invariance properties on the second-order statistics of a Gaussian process hand over to its distribution functions.

Example 4. *The Wiener process $W_t, t \geq 0$ is not a weakly stationary process, since its covariance function is $C(t,s) = min(t,s)$, which depends on both t and s and not on their difference $t-s$.*

As already stated in Chapter 2, the atmospheric variables one typically deals with in Meteorology, as well as in GNSS Meteorology, exhibit a stochastic nature and must be treated on a statistical basis. For instance, the ZWD variable at a point (X_0, Y_0, Z_0), where a GNSS receiver is located, in a certain epoch t_0 is typically considered a real random variable whose value must be estimated by means of statistical inference methods. In the next chapters, we will describe some methods to compute these estimates.
Atmospheric fields, i.e. spatio-temporal arrays of atmospheric variables, are instead modelled as random fields parametrized in general by an indexing set $T \subseteq \mathbb{R}^3 \times \mathbb{R}$, where $\mathbb{R}^3$ represents the spatial coordinates, while $\mathbb{R}$ the time variable. It is not possible indeed to handle meteorological problems, from weather forecasts to climate studies, in a deterministic way because of the chaotic nature of the Atmosphere's dynamics (cf. Chapter 2) and the unavoidable errors in its mathematical representations. As an example, the ZWD field will be treated as a spatio-temporal random field parametrized by $T \subseteq \mathbb{R}^2 \times \mathbb{R}$ (since the ZWD value is an integrated value on the vertical component, the spatial coordinates of interest are the planimetric ones).
In order to analyze atmospheric random fields, it would be tempting to assume them weakly homogeneous. However, it turns immediately out that these fields do not have a homogeneous nature as defined in Definition 4 (see [**Tatarski(1961)**]).

Therefore we are forced to look into other forms of stationarity/homogeneity that are general enough to suitably model the atmospheric fields we are dealing with, in particular the ZWD field, and that still allow to apply inferential methods to make predictions of the random field.
In the following section we will introduce the intrinsic stationarity (or homogeneity, in case of random fields) hypothesis, that will give a theoretical framework to tackle this issue.

3.2 Intrinsic stationary stochastic processes

In this section we will define a family of stochastic processes that is larger than that of weakly stationary processes introduced in the previous section, but it still allows to make predictions of the random process. For many aspects this class seems to be also well-suited for modelling the atmospheric fields of interest to us, primarily the ZWD field. In order to define this class, we will introduce the notion of intrinsic stationarity of a stochastic process, which is a second-order statistics based on the increments of the process. For the sake of simplicity, we will phrase all the concepts in this section in terms of stochastic processes. However, all the definitions and properties can be verbatim generalized to the more general case of stochastic field. This section will be based on [**Wackernagel(2003)**] and [**Matheron(1971)**], which we refer to for a complete treatment.
Let $X_t, t \in T$ denote a stochastic process. For any $t \in T$ and h such that $t+h \in T$ we define the *increment* of the stochastic process as

$$Y_h = X_{t+h} - X_t \tag{3.12}$$

In order to generalize the notion of a stationary process the idea is to consider its increments instead of looking at the process itself. Keeping this in mind and recalling that the weak stationarity is built on the second-order statistics of the stochastic process, we define another (second-order) statistics based on the increments

Definition 5. *Given the stochastic process $X_t, t \in T$ the function*

$$\gamma(t_1, t_2) = \frac{1}{2} var(X_{t_1} - X_{t_2}) = \frac{1}{2}\mathbb{E}[((X_{t_1} - m(t_1)) - (X_{t_2} - m(t_2))^2] \tag{3.13}$$

is called the variogram function of the stochastic process X_t.

As for the notion of wide-sense stationarity, the instrinsic stationarity hypothesis follows by imposing regularity properties to the first-order and the second-order statisctics of the increments in the following way

Definition 6. *A stochastic process $X_t, t \in T$ is said to be intrinsically stationary if*

1. *the mean of the increments is supposed to be indetically zero over the definition domain*

$$\mathbb{E}(X_{t+h} - X_t) \equiv 0 \quad \forall t + h, t \in T \tag{3.14}$$

2. *the variance of the increments, i.e. the variogram function, has a finite value and depends only on the increment vector h and no longer on the points $t+h$ and t*

$$var(X_{t+h} - X_t) = 2\gamma(h) \quad \forall t+h, t \in T \tag{3.15}$$

If moreover, the variogram function does only depend on the module of the increment h

$$var(X_{t+h} - X_t) = 2\gamma(\|h\|) \quad \forall t+h, t \in T \tag{3.16}$$

then the stochastic process is called intrinsically stationary and isotropic.

The above definition is again immediately generalized to the case of random fields. In this case, however, we will again talk about an intrinsically *homogeneous* random fields or about an intrinsically *homogeneous* and *isotropic* fields.
Basic properties of the variogram function $\gamma(h)$ of an intrinsically random process are listed below

- the variogram function vanishes at zero

$$\gamma(0) = 0 \tag{3.17}$$

- the variogram function is a positive function

$$\gamma(h) \geq 0 \tag{3.18}$$

- the variogram function is an even function

$$\gamma(h) = \gamma(-h) \tag{3.19}$$

- the variogram funtion has a subquadratic growth, i.e.

$$\lim_{h \to +\infty} \frac{\gamma(h)}{\|h\|^2} = 0 \tag{3.20}$$

Let us now assume that $X_t, t \in T$ be a weakly stationary process. Then its covariance function depend only on the difference of any two points

$$C(h) = \mathbb{E}(X_{t+h} X_t)$$

and it is a bounded function that does not exceed the variance

$$C(h) \leq C(0). \tag{3.21}$$

In this case, it can be proved (see [**Wackernagel(2003)**]) that the variogram function of X_t can be derived from the covariance function by means of the relation

$$\gamma(\cdot) = C(0) - C(\cdot) \tag{3.22}$$

This formula implies immediately that a weakly stationary stochastic process is also intrinsically stationary and that in this case the variogram γ is also bounded by $C(0)$.
However, the converse is not true, since the variogram function $\gamma(h)$ of an intrinsically stationary process need not to be bounded in general. Hence the intrinsic stationarity hypothesis is more general than the weak stationarity and, therefore, the class of intrinsically stationary processes is larger than that of wide-sense stationary processes.

Example 5. *Let us consider the Wiener process $W_t, t \geq 0$ as defined in the previous section.*
It follows from its definition that $\mathbb{E}(W_{t+h} - W_t) \equiv 0$ and that the variogram function reads

$$\gamma(h) = \frac{1}{2} var(W_{t+h} - W_t) = \frac{1}{2}\|h\| \tag{3.23}$$

Thus, the Wiener process is intrinsically stationary and isotropic.

Example 6. *An important family of intrinsically stationary and isotropic stochastic processes (or intrinsically homogeneous and isotropic random fields) is the class of Fractional Brownian motion models. As the name suggests, they are related to the Wiener process and indeed the Wiener process is a particular case of a Fractional Brownian motion.*

Definition 7. *A Gaussian process $B_t^H, t \geq 0$ is said to be a Fractional Brownian motion of Hurst parameter $H \in (0, 1)$ if it has mean zero and the covariance function is equal to*

$$\mathbb{E}(B_t^H B_s^H) = R_H(t, s) = \frac{1}{2} C_B(s^{2H} + t^{2H} - |t - s|^{2H}). \tag{3.24}$$

where $C_B \geq 0$ is a constant.

A Fractional Brownian motion has stationary *and* isotropic *increments, i.e. it is an intrinsically stationary and isotropic random process. In fact, from 3.24 it follows that the increments have a normal distribution with zero mean*

$$\mathbb{E}(B_{t+h}^H - B_t^H) = 0$$

and variance

$$var(B_{t+h}^H - B_t^H) = \mathbb{E}((B_{t+h}^H - B_t^H)^2) = |t - s|^{2H} \tag{3.25}$$

If we choose the Hurst parameter to be $H = \frac{1}{2}$ then we get the Wiener process.

One striking application of the Fractional Brownian motions is in the stochatic description of turbulent flows and, in particular, of the atmosphere's refractivity field. In fact, focusing on this case, Tatarski in [**Tatarski(1961)**] assumed that the refractivity field N of the atmosphere is a locally intrinsically homogeneous random field and based on Kolmogorov's theory of turbulence he stated that its variogram function is modelled by a Fractional Brownian motion with Hurst parameter $H = \frac{1}{3}$

$$\gamma_N(\|h\|) = \mathbb{E}((N(x + h) - N(x))^2) = C_N \cdot \|h\|^{\frac{2}{3}} l_0 \leq h \leq L_0 \tag{3.26}$$

where l_0 and L_0 are the inner and outer scale of turbulence.
Equation 3.26 was first obtained for general turbulent fields by Kolmogorov in his studies of high Reynolds number fluids and it is widely known as the two-thirds law. Kolmogorov's theory of turbulence is a very interesting attempt at finding a

mathematical structure in turbulent phenomena and we refer to [**Pope(2000)**] for a treatment of this theory in all its aspects and drawbacks. In particular a proof of the two-thirds law can be found there. The word "locally" stated above may give rise to different interpretation. Here, we interpret it understanding regional mesoscale meteorological phenomena, i.e. events occurring with a horizontal spatial scale up to approximately 500 km.
Since the ZWD field is a linear functional of the air refractivity field (cf. Chapter 2), in this work we will assume it to be locally intrinsically stationary as well. More precisely, we will suppose that over regions covered by dense and mid-scale GNSS networks - which range from a few km to a few hundred - the ZWD field can be considered intrinsically stationary.
We will present a case study based on Fractional Brownian motion models in Chapter 5.

3.3 Kriging algorithm

In this section we will sketch the typical data-driven strategy used to make predictions of a random field (or equivalently process) in a point not directly observed, by means of a given dataset of measurements. We will introduce the empirical variograms that are usually computed as a first guess of the theoretical variogram functions of the field. Then, once a suitable variogram model has been found by replacing the empirical variograms, we will define the ordinary Kriging system, which will allow to give an estimate of the random process in a an arbitrary point of its definition domain. This procedure will be extensively used in Chapter 5. This section, moreover, is largely inspired by [**Wackernagel(2003)**].
Let $X(\boldsymbol{x}), \boldsymbol{x} \in T \subseteq \mathbb{R}^k (k = 2, 3, 4)$ be a stochastic process. Suppose to have sampled it at different locations $\boldsymbol{x}_\alpha, \alpha \in \Omega \subseteq T$. We measure the *dissimilarity* of two sampled values by computing the empirical variogram

$$\gamma^*(\boldsymbol{x}_\alpha, \boldsymbol{x}_\beta) = \frac{1}{2}(\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta)^2 \tag{3.27}$$

Now, let us further assume $X(\boldsymbol{x})$ to be an intrinsically homogeneous and isotropic random field according to Definition 6. This implies that the dissimilarity γ^* does only depend on the module of the difference $\boldsymbol{h} = \boldsymbol{x}_\alpha - \boldsymbol{x}_\beta$, i.e.

$$\gamma^*(\|\boldsymbol{h}\|) = \frac{1}{2}(\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta)^2 \tag{3.28}$$

Under this assumption then the set of dissimilarities computed for all sampled values can be plotted in a 2-dimensional cartesian diagram with abscissa the module of the vector $\boldsymbol{h}$. This plot is called the *variogram cloud*. An ideal sketch of a generic variogram cloud can be found in [**Wackernagel(2003)**], Chapter 6.
The variogram cloud is an effective tool for investigating the behaviour of a random field. For instance, anomalies, inhomogeneities or outliers can be detected by analysing th vaiogram cloud (cf. [**Wackernagel(2003)**]).
Nevertheless, it is usually clearer and more interesting to average the variogram cloud with respect to a finite set of intervals of the abscissa axis (generally non-

overlapping classes are chosen). To be more precise, let us give the following definition.

Definition 8. *Given a finite collection of classes $\{\mathcal{H}_k\}$ of the abscissa axis, the average dissimilarity with respect to $\mathcal{H}_k$ is*

$$\gamma^*(\mathcal{H}_k) = \frac{1}{2n}\sum_{i=1}^{n}\left(X(\boldsymbol{x}+\boldsymbol{h}) - X(\boldsymbol{x})\right)^2 \qquad \boldsymbol{h} \in \mathcal{H}_k \tag{3.29}$$

The set of all average dissimiliraties is called the empirical variogram of the random field $X(\boldsymbol{x}), \boldsymbol{x} \in T$ with respect to the given dataset.

In Chapter 5 we will give many examples of empirical variograms of the ZWD field with respect to ZWD values estimated at the receiver's position at different epochs.
The empirical variogram is a powerful tool for exploring the features of a random field, since it contains the information of the variogram cloud but averaged on those points that show similar spatial stochastic characteristics. For large spacings the empirical variogram may reach a sill, which approximates the variance of the random field at each point. In this case, moreover, one can deduce that the variogram function has a covariance counterpart and that the random field is weakly homogeneous and isotropic. Also the behaviour of the empirical variogram at small scale is important since it gives evidence to the regularity of the random field, e.g. smooth, continuous but not smooth or discontinuous. For further details on this issue we refer to the work of Matheron [**Matheron(1971)**].
As already mentioned above, the empirical variograms are only a first guess of the behaviour of the random field based on the measurements available. They are replaced by a theoretical variogram model essentially because the theoretical functions can retain a physical meaning of the phenomenon under study and because these models can be used in the stochastic description of the random field giving a sound mathematical basis (cf. [**Tatarski(1961)**] or [**Treuhaft et al.(1987)**]).
A theoretical variogram model is fitted to the empirical variogram. Of course this fit is based on an interpretation of the empirical variogram, especially at the origin, which dictates the behaviour of the random field at small scales, and at large distances, in order to detect a bounded or unbounded behaviour of the empirical variogram. In this approximation, in fact, the most important point to consider is the type of continuity at the origin and the type of homogeneity assumed for the random field. Again a thorough discussion can be found in [**Matheron(1971)**].
Not every function is eligible to be a variogram model. Indeed, a theoretical vaiogram function has to be *conditionally negative* definite, i.e. the following equation must hold

$$[w_\alpha]^T[\gamma(\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta)][w_\alpha] = \boldsymbol{w}^T\Gamma\boldsymbol{w} \leq 0 \tag{3.30}$$

for all $w_\alpha \in \mathbb{R}$ such that $\sum_{\alpha=0}^{n} w_\alpha = 0$.
This condition guarantees that the variance of a linear combination of $n+1$ variables of the random field is nonnegative. A proof of this statement can be found in [**Wackernagel(2003)**]. This condition is mandatory to use the theoretical variogram functions to set up an ordinary Kriging system.

The ordinary Kriging algorithm allows to estimate a value of an intrinsically stationary random field at a point of the domain where the variogram functions are known. The estimates is retrieved by means of the data in the neighborhood of the estimation point. More precisely, if $X(\boldsymbol{x}), \boldsymbol{x} \in T$ is an intrinsically stationary random field we set

$$X^*_{OK}(\boldsymbol{x}_0) = \sum_{\alpha=1}^{n} w_\alpha X(\boldsymbol{x}_\alpha) \tag{3.31}$$

The prediction problem consists in determining the coefficients w_α in order to give an estimate of $X^*_{OK}(\boldsymbol{x}_0)$.
We need to constrain the weights to sum up to 1. In fact, if all data were constant, then the estimated value $X^*_{OK}(\boldsymbol{x}_0)$ should be also equal to that constant. We will assume in the rest of this section that the data are realizations of the intrinsically stationary random field $X(\boldsymbol{x}_\alpha)$ with a variogram model $\gamma(h)$.
The estimation variance σ^2_E is the variance of the linear combination

$$\sigma^2_E = var(X^*_{OK}(\boldsymbol{x}_0) - X(\boldsymbol{x}_0)) = \sum_{\alpha=1}^{n} w_\alpha X(\boldsymbol{x}_\alpha) - 1 \cdot X(\boldsymbol{x}_0) \tag{3.32}$$

. Since $\sum_{\alpha=1}^{n} w_\alpha = 1$ and setting $w_0 = -1$ we get the constraint

$$\sum_{\alpha=0}^{n} w_\alpha = 0 \tag{3.33}$$

Expanding Equation 3.32 with respect to the variogram, we get

$$\begin{aligned} \sigma^2_E &= \mathbb{E}[(X^*_{OK}(\boldsymbol{x}_0) - X(\boldsymbol{x}_0))^2] = \\ &- \gamma(0) - \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} w_\alpha w_\beta \gamma(x_\alpha - x_\beta) + 2\sum_{\alpha=1}^{N} w_\alpha \gamma(x_\alpha - x_\beta) \end{aligned} \tag{3.34}$$

The ordinary Kriging system is obtained by minimizing the estimation variance σ^2_E with the constraint 3.33

$$\begin{pmatrix} \gamma(x_1 - x_1) & \cdots & \gamma(x_1 - x_N) & 1 \\ \vdots & & \vdots & \vdots \\ \gamma(x_N - x_1) & \cdots & \gamma(x_N - x_N) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{pmatrix} = \begin{pmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_1 - x_0) \\ 1 \end{pmatrix} \tag{3.35}$$

where the λ is the Lagrange parameter used to obtain the system.
The left hand side of the system represents the variogram values between the measured points, whereas the right hand side describes the values between the data points and the unknown point $\boldsymbol{x}_0$. If we solve the system 3.35, we get the weights $\boldsymbol{w} = [w_\alpha]$ set in Equation 3.31. This weight vector exists and is unique, being the matrix on the left hand side non-singular.
The estimation variance of the ordinary Kriging system can be estimated by the formula

$$\sigma^2_{OK} = \lambda - \gamma(0) + \sum_{\alpha=1}^{n} w_\alpha^{OK} \gamma(\boldsymbol{x}_\alpha - \boldsymbol{x}_0) \tag{3.36}$$

Together with the linearity, by construction the Kriging predictor

1. minimizes the estimation variance of the prediction error

2. is an exact interpolator, i.e. in case $\boldsymbol{x}_0$ is one of the locations of the measured points then the Kriging estimated value is identical with the data value itself at that point

$$X^*_{OK}(\boldsymbol{x}_0) = X(\boldsymbol{x}_\alpha) \quad \text{with } \boldsymbol{x}_0 = \boldsymbol{x}_\alpha$$

The Kriging algorithm will be used later on in Chapter 5 to obtain a regular grid of predicted ZWD values using ZWD values estimated at the receivers locations of the GNSS SPIN network.

Chapter 4

GNSS data processing of the MIST dense network: impact of relative and absolute strategies on ZWD estimates

4.1 Introduction

In the last years the possibility to increase the spatial resolution of GNSS observations for meteorological purposes has been investigated, with the aim of better describing the highly variable spatial behaviour of atmospheric water vapor content.
In this framework, we carried out a comprehensive analysis of the different approaches and estimation strategies of GNSS-derived ZWDs of a dense netwotk called MisT. In particular we carefully analyzed the influence on the estimation of coordinates and ZWD of the relative processing strategy, i.e. DD, and of the absolute processing one, namely PPP (see Chapter 2 for an introduction to DD and PPP methods).
This chapter is devoted to the description of this activity.

4.2 MisT dense network

We carried out a re-processing activity of the GNSS observation dataset collected by a dense experimental network called MisT.
This network was deployed within the ESA project METAWAVE with the original aim of studying the spatial correlation of GNSS ZWDs estimated from the data collected by the network (cf. ([**Metawave(2011)**])).
MisT covered the urban area of Como, Northern Italy. The area is morphologically complex, including both lowland and mountains.
Eight geodetic receivers were placed around the permanent station of COMO, which belongs to the European Reference Permanent Network (EUREF PN), to form a multi-scale, dense network. COMO station was chosen as reference station in the network adjustment by means of a relative approach, when a relative a

relative positioning strategies was applied to estimate stations coordinates and ZWDs (see below for further details on this point).
The receivers were in fact at distances, with respect to the COMO station, in the order of 100, 1000 and 10000 meters, to form three nested loops which covered in total an area of around 100 km^2. To be more precise the two receivers ANZA, CAST were installed closest to COMO at a distance of approximately 100 m. The receivers BRUN, LAPR and PRCO were instead deployed in the "second loop" of the network, namely at a distance of about 1000 m from COMO. The remaining receivers, i.e. NAND, DANI and MGRA, were the furthest ones, outdistancing around 10000 m from the reference station. It is worth pointing out that the receivers were not on the same height.

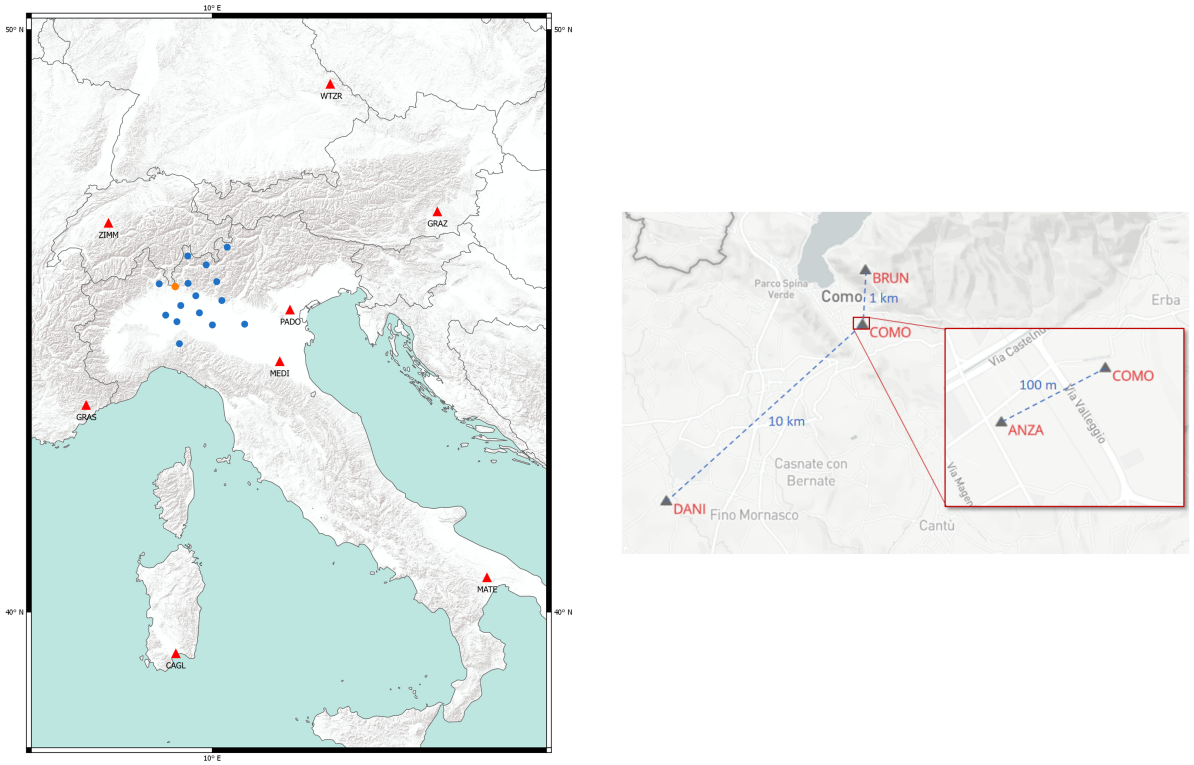

Figure 4.1 – **Left side**: The Lombardy network, given by the blue- and yellow-filled circles. The yellow circle denotes the reference station COMO. The red triangles represent IGS stations. CAGL, MEDI, MATE, GRAS, GRAZ augmented the Lombardy network to compute coordinates and ZWDs of the reference station COMO (for further details see Section 4.3).
Right side: Mist's stations COMO, BRUN, ANZA and DANI, of which we report the results.

The six inner receivers (COMO, ANZA, CAST, BRUN, LAPR, PRCO) collected data continuously for six months (from May to November 2008) at 1 Hz rate, while the furthest three (NAND, DANI and MGRA) worked only for the last five weeks.

For this reason we will report only the results obtained by analyzing the COMO reference station plus three stations ANZA, BRUN and DANI. The location of these stations is shown if Figure 4.1 on the right side.
These three stations, indeed, provide more complete datasets than those collected by the other ones. Furthermore they are representative for the distance levels of the network.
The MisT network was designed, deployed and operated by the Geomatics and Earth Observation Laboratory (GEOlab) of Politecnico di Milano.

4.3 GNSS data processing setup

In this section we will briefly describe the software tools and parameters used to process the GNSS data collected by MisT. The estimates of both coordinates and ZWD time series and, more importantly, the comparison among them are summarized in the next section.
The GNSS observations were used to deeply compare relative (i.e. DD) and absolute stand-alone (i.e. PPP) adjustment strategies. In order to process the data we adopted the BERNESE v5.2 software (throughout this chapter denoted by BSW). This choice implied to carry out a least-square batch adjustment procedure for both strategies, since in the BSW no "real-time" method, e.g. as the Kalman filter algorithm, is implemented. However, as already remarked earlier, the BSW represents the state-of-the-art software available for the GNSS data processing. Therefore, for our study we leant toward BSW to have a sound benchmark point. More precisely, the following analyses accomplished:

1. **Single-Base Double-Difference adjustment of daily sessions -** Single-Base adjustments are performed between each network station and the reference station COMO. The coordinates and ZWD time series of the reference station, used as reference values, were determined in a previous analysis of the MisT data (see [**Visconti et al.(2009)**]) at mm level by the iono-free combinations adjustment of the Lombardy regional network (now part of the SPIN GNSS Piemonte-Lombardia GNSS network). The Lombardy network was augmented with 5 International GNSS Service (IGS) stations, namely CAGL - Cagliari, GRAS - Grasse, GRAZ - Graz, MATE - Matera, MEDI - Medicina.
   The adjustment of this larger network was necessary for the absolute estimate of COMO ZWD time series.
   Figure 4.1 shows the GNSS stations involved in the study (they are denoted by red triangles). On the left the large network used to estimate COMO station position and absolute ZWDs. On the right the stations of the MisT network whose results are reported in this work.
   The processing rate of all MisT stations was set to 30 seconds, to account for COMO sampling rate. The dry and wet mapping functions were chosen in accordance with those used in the adjustment of the augmented Lombardia regional network (i.e. dry/wet Niell mapping functions) and we used 10 degrees of elevation cut-off. The choice of this elevation cut-off might appear

at first glance a quite large value if compared with a setup for positioning purposes only (coordinates are typically estimated on a daily base). However, dealing with a GNSS dense network, we wanted to consider reasonably confined portions of the atmosphere above each receiver that would still contain the highest number possible of signal paths.
Similarly the ZWD estimation rate of the MisT network has been set to 1 hour coherently with the reference ZWD time series of COMO estimated at that rate. Large a-priori standard deviation of the differences between the ZWD parameters of the same station at two consecutive epochs was used to let the ZWD estimates depend on the observations only without introducing any a-priori information.

2. **Precise Point Positioning adjustment of daily sessions -** The processing parameters were chosen in general to be consistent with those of DD processing in view of the comparisons.
In particular the observation sampling rate of all MisT stations was set to 30 seconds in order to be consistent also with the rate of the precise satellite clock estimates necessary for the PPP strategy (Section 2). Moreover, the dry and wet mapping functions were also in this case the Niell mapping functions. Lastly for the a-priori standard deviation of the differences between ZWD estimates of the same station at two consecutive epochs we set the same value used for the DD adjustment.

Further details about the processing settings of the experimental network for both approaches may be found in Table 4.1.

Table 4.1 – MisT - Single Base DD and PPP processing settings

	Relative DD	Absolute PPP
Processing software	BERNESE v5.2	BERNESE v5.2
Processed observations	L1 and iono-free combination	iono-free combination
Processing rate	30 s	30 s
Elevation cutoff	10 degrees	10 degrees
Ambiguity fixing strategy	SIGMA	none
Mapping functions	Niell dry/wet	Niell dry/wet
Gradient estimation model	none	none
Ionosphere model (L1)	CODE GIM	-
A priori error std of COMO crd	$1{\cdot}10^{-5}$ m	none and $1{\cdot}10^{-5}$ m
COMO ZTD hourly time series	fixed	-
ZWD estimation rate	1 hour	1 hour
Ephemeris	CODE (final)	CODE (final)
Satellite clocks	-	CODE 30 seconds (final)
A priori ZWD difference std	5 m	5 m
Satellite/receiver antenna PCV	latest igs08.atx	latest igs08.atx
Ocean loading model	FES2004	FES2004

In order to deal with the great amount of output files that the BSW provides with, we have developed ad-hoc software tools that lead to build a "data warehouse".
We designed in fact a software module that manages the automatic import of the data from separate output BSW files in a coherent archive that can be easily queried and manipulated.
In this regard, we implemented several add-ons that perform the computation of various statistical parameters and allow to carry out stochastic interpolation procedures, i.e. collocation and Kriging algorithms. This software has been extensively used to produce the results we are showing in the next section.

4.4 GNSS data estimation: results and discussion

4.4.1 Ionosphere impact on relative positioning of two close stations

The first comparison we carried out involved double-differenced observations only and in particular we investigated the impact of using single frequency L1 observations with respect to iono-free L3 combinations for the chosen network's stations. In fact if double-differenced L1 observations of dense GNSS networks would lead to accurate ZWD estimates (and therefore accurate PWV estimates), the capillary deployment of such networks for meteorological purposes could become an effective and economically viable way to acquire ground-based data to be used in NWP models (considering of course that single frequency receivers are one order of magnitude more affordable than dual frequancy receivers).
In order to address this issue it was necessary to analyze the ionospheric delay affecting the GNSS signals. When the distance between two stations is shorter than 10 km, the delay affecting the simultaneous observations from these two stations - to the same satellite - is significantly reduced in the difference.
Thus the question we addressed was the following.
Can we use L1-only observations, instead of the iono-free combinations, without accurately modelling the ionosphere? In other words, can we adopt dense networks of single-frequency receivers (e.g. the nowdays available low-cost ones) together with one dual-frequency reference station (whose coordinates and ZWD values are estimated in the adjustment of a permanent regional network) to obtain sufficiently accurate ZWD estimates?
To answer this question we compared the parameters estimated from the adjustment of L1 only observations and those obtained from iono-free L3 combinations, namely station coordinates and ZWDs. The results of such comparisons are shown in Figure 4.2 and summarised in Table 4.2; they refer to the stations ANZA, BRUN and DANI, to account for baseline lengths (with respect to COMO) of 100, 1000, and 10000 metres.
Although all differences have a standard deviations lower than approximately 1 cm, the N and UP coordinate differences are biased, and this holds irrespectively of the baseline length.
However, even though the bias was slightly reduced, it did not disappear showing thereby that it is not directly caused by the factors mentioned above, but rather due to the inconsistency between the COMO reference station coordinates and

ZWD, which are determined by means of iono-free combinations and then used in the adjustment of L1-only data.
ZWD estimates from both techniques do not show instead any evident biases or systematisms.

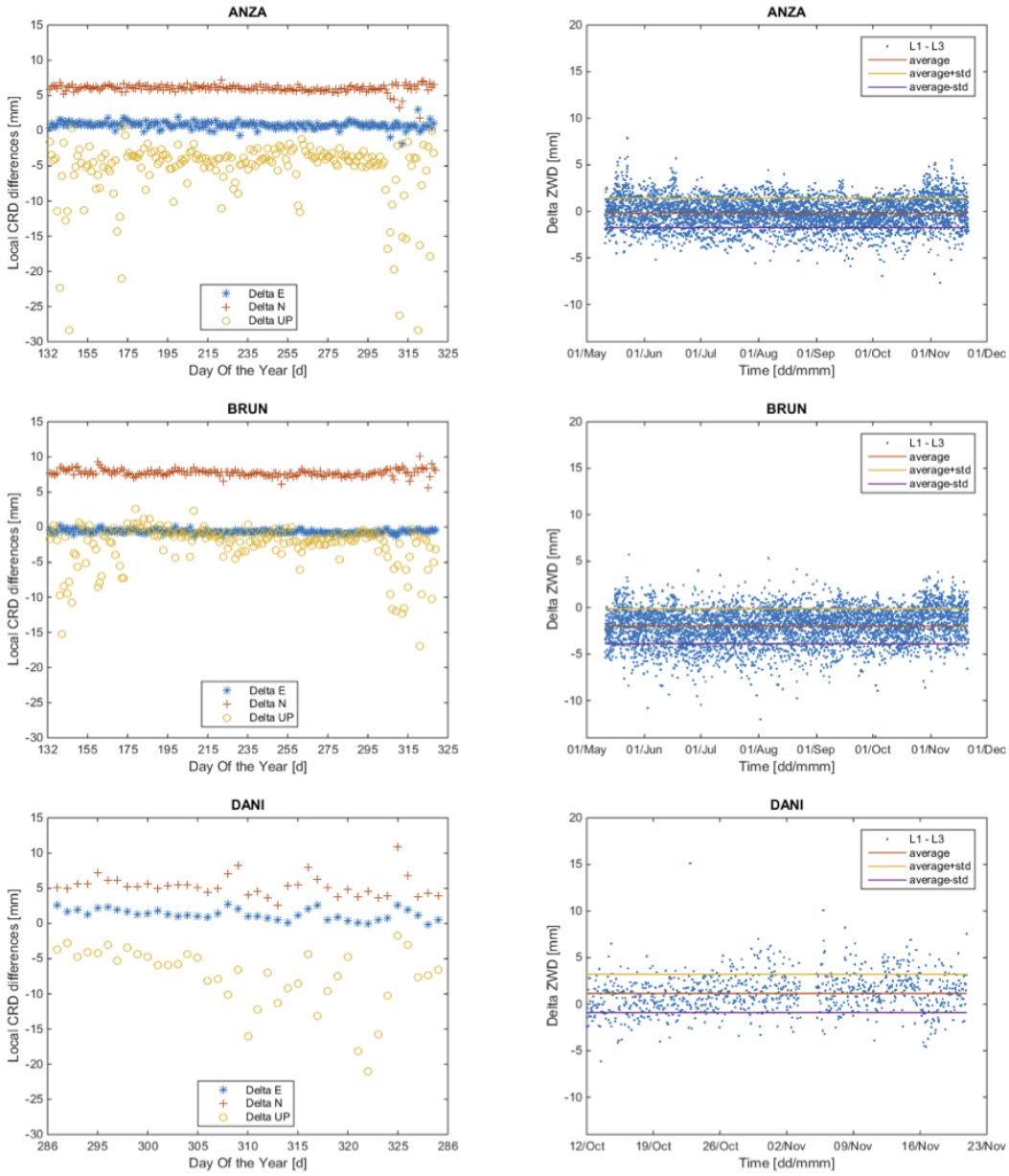

Figure 4.2 – Comparison of coordinates and ZWD from L1 and L3 observations

The three furthest stations of the MisT network (DANI, MGRA and NAND) collected observations for a much shorter period than the inner ones and here we are going to report just one of them, i.e. DANI. Furthermore, this period corresponded to approximately the last month of the experiment and it turned out to be characterised by a large variability in the differences between the L1-only and L3 solution. Thus with the aim of better understanding this variability we looked at the coordinate time series obtained from the adjustment of L1-only, L2-only and iono-free L3 combination of L1 and L2 observations. As the network's stations are in a fixed position, in fact, we would expect a variability of the coordinate time series in a range of millimeters. All the resulting coordinates time series instead show a larger variability in the last month of the experiment, the iono-free combination giving the worse results. We report the standard deviations of the residuals (δE, δN and δUP) with respect to a regression line of the coordinates in Table 4.3 for L1-only, L2-only and iono-free solutions respectively. For short baselines, as those characterizing a dense GNSS network, it seems that the use of iono-free L3 combinations is not only unnecessary but it significantly worsens the accuracy of the coordinates components, especially in the UP direction, with a larger impact on the shorter baselines of ANZA-COMO and BRUN-COMO.
As a matter of fact, according to BSW manual ([**Dach et al.(2015)**], Chapter 13.1.3, Page 312), for small-area high-precision dense networks the use of L1-only data in combination with ionospheric models is recommended although dual-frequency observations are available.
Moreover, we observed that this noisier behaviour of L3-derived coordinate and ZWD estimates was common to all network's stations and eventually we guessed it could be due to low quality observations acquired by the reference station COMO. To verify this hypothesis we performed a relative positioning of BRUN with respect to ANZA station, by fixing ANZA coordinates and ZWD values to those obtained by means of absolute positioning. Although this procedure could not lead to accurate solutions for BRUN (the coordinates and ZWDs time series of the reference station ANZA were affected by PPP residual errors), it had the merit to entirely exclude the observations of COMO station.
It turned out that the variability of the UP coordinate in the last month of acquisition reduced by a factor of approximately 0.6 - irrespectively of the noisier PPP derived coordinates of the reference receiver - thus confirming our hypothesis.
This suggested that for the retrieving ZWD time series from dense single-frequency GNSS networks, at least two dual-frequency stations should be used to allow for redundancy and internal data quality assessment.

Table 4.2 – Statistics of the differences between L1 and iono-free (IF) coordinates and ZWDs

station	variable	n. of data	mean [mm]	std [mm]
ANZA	E_{L1}-E_{IF}	192	0.9	0.3
100 m from COMO	N_{L1}-N_{IF}	192	6.0	0.3
	UP_{L1}-UP_{IF}	192	-5.5	3.4
	ZWD_{L1}-ZWD_{IF}	4610	-0.2	1.6
BRUN	E_{L1}-E_{IF}	192	-0.6	0.2
1000 m form COMO	N_{L1}-N_{IF}	192	7.7	0.5
	UP_{L1}-UP_{IF}	192	-2.5	3.0
	ZWD_{L1}-ZWD_{IF}	4607	-1.9	1.9
DANI	E_{L1}-E_{IF}	37	1.3	0.8
10000 m from COMO	N_{L1}-N_{IF}	37	5.4	1.5
	UP_{L1}-UP_{IF}	37	-7.5	4.5
	ZWD_{L1}-ZWD_{IF}	879	1.2	2.0

Table 4.3 – Standard deviation of the residuals with respect to a regression line of the station coordinates in the last month of the MisT experiment

			L1-only	L2-only	iono-free
station	variable	n. of data	std [mm]	std [mm]	std [mm]
ANZA	δE	34	0.4	0.3	0.8
100 m from COMO	δN	34	0.3	0.2	1.0
	δUP	34	2.5	0.8	7.4
BRUN	δE	34	0.5	0.7	0.6
1000 m from COMO	δN	34	0.7	0.7	0.6
	δUP	34	2.1	1.6	5.8
DANI	δE	33	1.3	1.3	0.9
10000 m from COMO	δN	33	1.7	1.4	1.0
	δUP	33	3.5	4.0	6.1

4.4.2 Relative against absolute positioning

The comparison between the results obtained from the relative and absolute positioning strategies was performed in two ways.
In the first case, we aimed at evaluating basic statistical parameters, in particular mean and standard deviation, of the differences between daily coordinates and hourly ZWDs of each station. These time series have been computed applying both relative and absolute positioning strategies according to a 'best practice' criterion. For the DD adjustment, we acted according to the results of the previous section and thus we used the coordinates and ZWD estimates from L1-only observations and the EPN station of COMO as reference station. The analysis included all the available data, comprising the noisier last days of the experiment (cf. the previous sub-section 4.4.1).
To assess the repeatability level of the two positioning strategies we looked at the coordinates time series determined with the two approaches (which are typically well modelled by a linear trend). The standard deviation of the residuals with respect to a regression line for the three coordinates are in the order of a few mm for the DD case. They are twice as large, instead, in case of the absolute PPP method but still under 1 cm, even in the worst case of the UP component. These results can be seen in Table 4.4 below.
In order to compare the accuracy of the two positioning methodologies we computed time series of differences between the DD and PPP coordinates and ZWDs. The results obtained are shown in Figure 4.3 and the corresponding statistics are reported in Table 4.5.
Almost all the coordinate differences are biased and their variability is always below 1 cm standard deviation. The ZWD differences are biased as well, with a standard deviation below 1 cm. The biases are probably due to the use of L1-only for the DD adjustment and the iono-free L3 combination for the PPP adjustment (cf. the previous sub-section 4.4.1). If expressed in terms of PWV, the std of the differences are below 1.6 mm, which is comparable with the results in [**Niell(2000)**], that is with the uncertainty of our best knowledge of the PWV itself. Moreover, the linear correlation coefficients between the DD and PPP-derived ZWD time series are about 98% for all the three considered stations. This makes the two processing techniques equivalent from the mere accuracy point of view.

Table 4.4 – Standard deviations of the residuals δ with respect to a regression line of the stations coordinates obtained with DD and PPP

			DD	PPP
station	variable	n. of data	std [mm]	std [mm]
ANZA	δE	192	1.3	3.9
100 m from COMO	δN	192	0.3	2.3
	δUP	192	2.2	8.8
BRUN	δE	192	1.2	2.9
1000 m from COMO	δN	192	0.8	2.0
	δUP	192	3.0	6.6
DANI	δE	37	1.2	3.1
10000 m from COMO	δN	37	1.6	2.1
	δUP	37	3.3	6.4

Table 4.5 – Comparison between DD and PPP derived coordinates and ZWDs

station	variable	n. of data	mean [mm]	std [mm]
ANZA	E_{DD}-E_{PPP}	192	3.3	3.6
100 m from COMO	N_{DD}-N_{PPP}	192	5.8	2.3
	UP_{DD}-UP_{PPP}	192	0.7	8.2
	ZWD_{DD}-ZWD_{PPP}	4610	0.9	9.5
BRUN	E_{DD}-E_{PPP}	192	1.0	3.1
1000 m from COMO	N_{DD}-N_{PPP}	192	7.0	2.1
	UP_{DD}-UP_{PPP}	192	-0.5	7.2
	ZWD_{DD}-ZWD_{PPP}	4607	1.2	8.4
DANI	E_{DD}-E_{PPP}	37	1.6	3.2
10000 m from COMO	N_{DD}-N_{PPP}	37	4.4	2.9
	UP_{DD}-UP_{PPP}	37	5.5	5.7
	ZWD_{DD}-ZWD_{PPP}	879	1.1	7.3

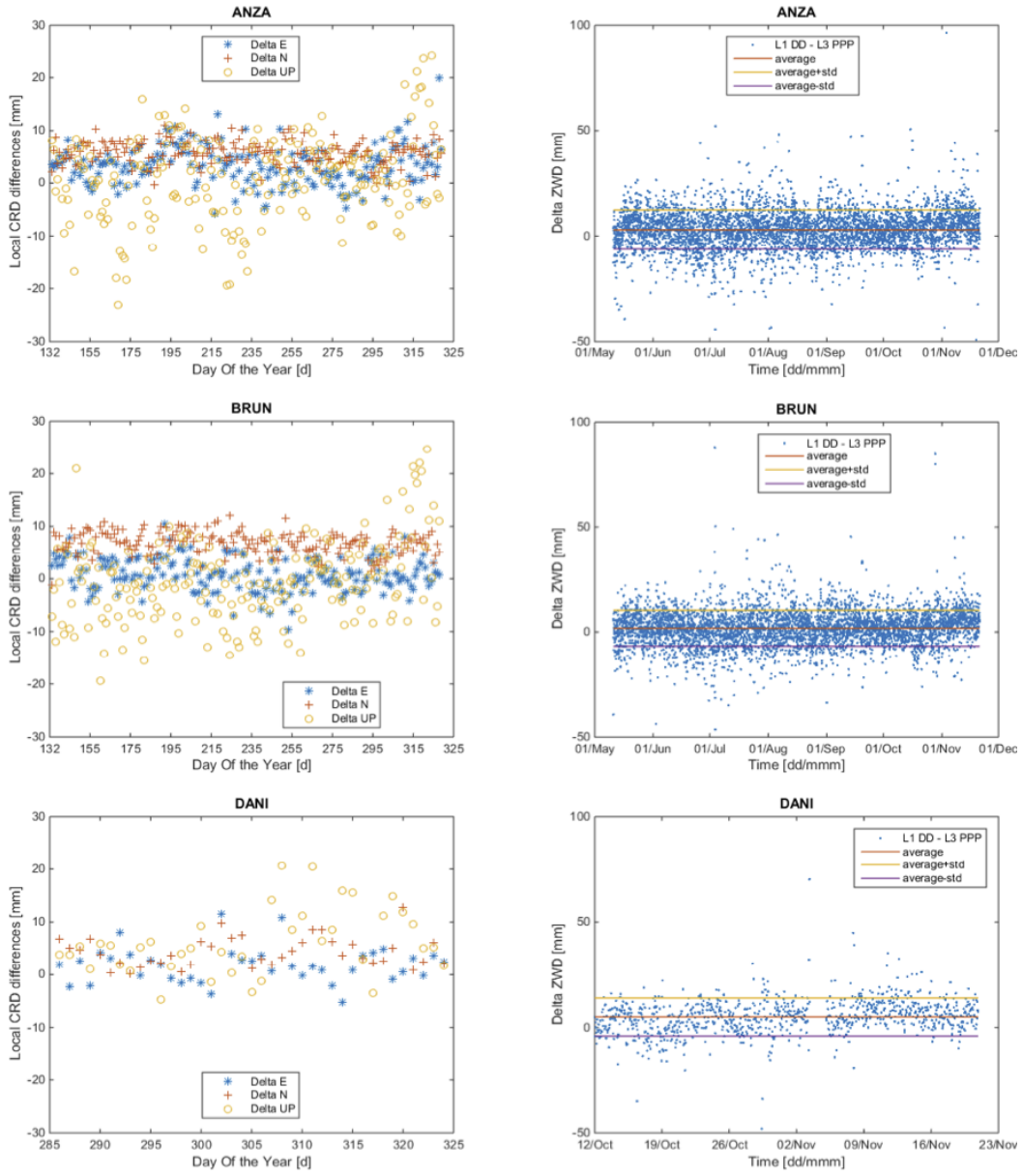

Figure 4.3 – Comparison of coordinates and ZWD from L1 DD and L3 PPP adjustment

The second comparison aimed at evaluating the impact of the different models used in the observation equations within the PPP approach in comparison with DD.
We considered exactly the same data set for both (relative and absolute) adjustment strategies and we used iono-free L3 combinations in both cases, making a comparison between *baseline* components (ΔE, ΔN, ΔUP) and ZWD increments (ΔZWD).

These solutions were directly estimated in the DD approach, whereas the corresponding quantities were computed as difference of the independent PPP adjustments of all network's stations. It is worth noticing that these quantities from both DD and PPP may be computed by means of the same observation set without introducing any external information. In particular the coordinates and ZWD time series of the reference station COMO (obtained from the adjustment of the above mentioned IGS stations)are not included and used anymore. We would expect a high correlation between coordinates and ZWD estimation errors from PPP adjustment for the baseline of 100 m, then decreasing with increasing baseline length. Indeed, the use of L3 combinations guaranteed that the ionospheric delay was drastically reduced. Furthermore, we conjectured that differencing PPP estimates would lead to a reduction of the systematic biases relative to the satellites clocks, satellites attitude effects and site displacement effects (e.g. Earth's solid tides, Earth's Orientation Parameters, ocean loading), thereby making evident a correlation more directly dependent on phenomena due to the variability of the neutral atmosphere above the stations.
However, this hypothesis could not be directly corroborated by the experimental results as it can be seen in 4.6. Here the statistics of the baseline components and ZWD increments separately for the DD and PPP case are reported. The standard deviation of the PPP derived baseline components and ZWD increments seems to be independent of the distance between GNSS stations, when this is smaller than 10 km. Therefore we looked at the differences between DD and PPP derived baseline components and ΔZWD. Such differences are shown in Figure 4.4.

Table 4.6 – Statistics of baseline components and ZWD increments from DD and PPP

			DD	PPP
baseline	variable	n. of data	std [mm]	std [mm]
ANZA - COMO	ΔE	192	1.6	4.5
100 m	ΔN	192	0.4	1.6
	ΔUP	192	2.2	8.8
	ΔZWD	4610	2.8	6.2
BRUN - COMO	ΔE	192	1.3	4.4
1000 m	ΔN	192	0.8	1.9
	ΔUP	192	3.0	7.2
	ΔZWD	4607	6.3	8.2
DANI - COMO	ΔE	37	1.2	5.6
10000 m	ΔN	37	1.7	2.1
	ΔUP	37	3.6	10.5
	ΔZWD	879	3.5	6.9

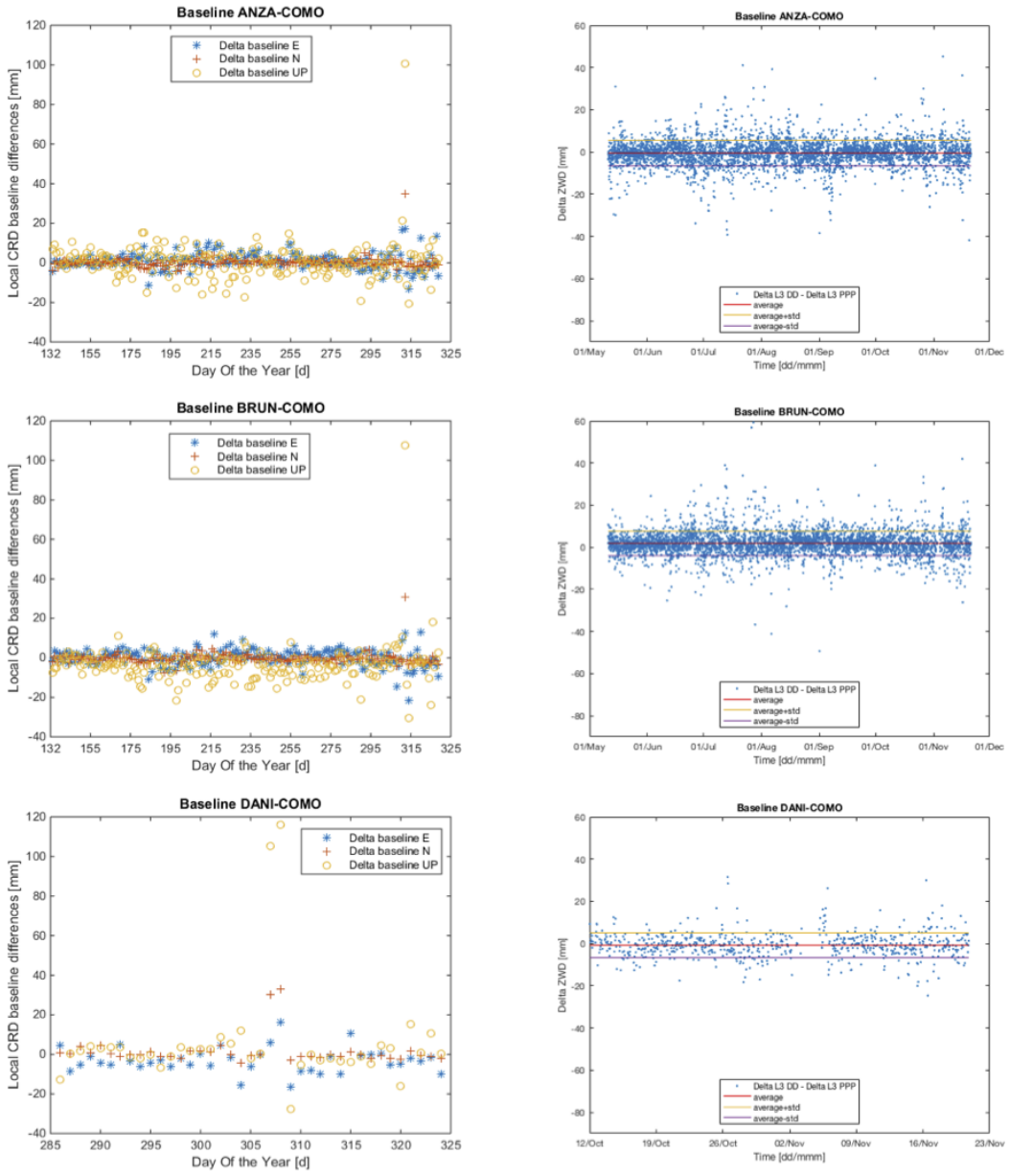

Figure 4.4 – Comparison between DD and PPP estimates. On the left hand side, differences between DD and PPP-retrieved ΔE, ΔN, ΔUP. On the right hand side, differences between DD and PPP-retrieved ΔZWD

Remark. During the baselines analysis, it turned out that the PPP a-posteriori standard deviations of ZWD estimates of COMO station were an order of magnitude higher (centimetres) than those of the other stations (millimetres). We therefore decided to remove the corresponding ZWD estimates, considered as out-

liers, by using a threshold of 2 mm on the a-posteriori standard deviation.

For the sake of completeness the statistics of the COMO PPP-derived coordinates and ZWD, with respect to the reference values used in the DD processing, are reported in Table 4.7. In the same table we report the statistics of ZWD with respect to the reference values after outlier removal.

Table 4.7 – Statistics of the differences between reference and PPP-derived coordinates and ZWDs

station	variable	n. data	mean [mm]	std [mm]
COMO	E_{ref}-E_{PPP}	192	1.6	4.7
	N_{ref}-N_{PPP}	192	-0.3	2.6
	UP_{ref}-UP_{PPP}	192	6.3	8.0
	ZWD_{ref}-ZWD_{PPP}	4610	2.0	21.7
after outlier removal	ZWD_{ref}-ZWD_{PPP}	4219	1.3	8.2

In Table 4.8 we show the statistics of the differences between baseline components and ZWD increments.

Table 4.8 – Comparison of DD estimated and PPP derived baseline components and ZWD increments

baseline	variable	n. of data	mean [mm]	std [mm]
ANZA-COMO	ΔE_{DD}-ΔE_{PPP}	192	1.6	3.8
100 m	ΔN_{DD}-ΔN_{PPP}	192	6.1	1.5
	ΔUP_{DD}-ΔUP_{PPP}	192	-5.5	7.6
	ΔZWD_{DD}-ΔZWD_{PPP}	4198	0.4	6.1
BRUN-COMO	ΔE_{DD}-ΔE_{PPP}	192	-0.6	4.0
1000 m	ΔN_{DD}-ΔN_{PPP}	192	7.3	1.7
	ΔUP_{DD}-Δ UP_{PPP}	192	-6.8	6.3
	ΔZWD_{DD}-ΔZWD_{PPP}	4220	1.8	5.8
DANI-COMO	ΔE_{DD}-ΔE_{PPP}	37	2.4	5.3
10000 m	ΔN_{DD}-ΔN_{PPP}	37	5.3	2.3
	ΔUP_{DD}-ΔUP_{PPP}	37	-7.6	9.0
	ΔZWD_{DD}-ΔZWD_{PPP}	701	-0.7	5.8

As reported in Table 4.8, it turns out that the differences between ΔZWD estimated from DD and those derived from PPP have all a standard deviation of approximately 6 mm.

The linear correlation between DD and PPP-derived ΔZWD time series, reported in Table 4.9, is less than 0.3 for the 100 m long ANZA-COMO baseline, and then it increases for the other two baselines. This could be justified by the fact that the differences between DD and PPP-derived baseline components and ΔZWD for ANZA-COMO account mainly for the PPP residual errors, as water vapour delay variations between the two stations can hardly affect the ZWD parameter.

Table 4.9 – Correlation coefficients between DD-estimated and PPP-derived ΔZWD

baseline	horizontal distance [m]	height difference [m]	correlation coefficient
ANZA-COMO	100	-12	22%
BRUN-COMO	1000	445	70%
DANI-COMO	10000	74	52%

A further investigation on the empirical covariance function of the DD and PPP-derived ΔZWD time series only, reported in Figure 4.5, seems to confirm such hypothesis. In fact, the covariance function of the DD-estimated ΔZWD of ANZA-COMO is pure white noise. However, this is not the case for the PPP-derived ΔZWD, whose empirical covariance function converges to zero after approximately 1 day. We believed that this could be due to the presence of correlated error in the PPP estimates.
To get a deeper understanding of such behaviour, we carried out the following analysis. First of all we assumed that DD ΔZWD time series did not contain correlated errors and therefore the corresponding empirical covariance functions of each baseline would only account for the water vapour signal. Thus, differencing the DD covariance function from the PPP covariance function - containing both signal and noise - we should find only a "pure" noise covariance function, which should appear almost the same covariance function of the PPP-derived ΔZWD of ANZA-COMO. This seems to be the case for the baseline BRUN-COMO, at least approximatively, as it can be seen in Figure 4.6. The signal to noise ratio values of those two baselines, estimated from the empirical covariance functions and reported in Table 4.10, seem to confirm it. This strongly suggests that it is not useful to use baselines shorter than 1 km length, regardless of the processing technique used to derive the ZWD values.
We performed the same analysis also for the DANI-COMO covariance functions, but, due to the poorer data set, the results are not clear. As it can be seen in Figure 4.5, the covariance functions of both DD and PPP-derived ΔZWD time series oscillate more, due to the use of a small number of data, which were collected during the dry season (October-November) when, in addition, related GNSS ob-

servations were noisy. The small water vapour signal in DANI-COMO ΔZWD time series reflects also into the small signal to noise ratio reported in Table 4.10.

Table 4.10 – Signal to noise ratio of DD and PPP-derived ΔZWD

baseline	ANZA-COMO	BRUN-COMO	DANI-COMO
SNR DD	0.13	2.53	0.42
SNR PPP	0.23	0.87	0.30

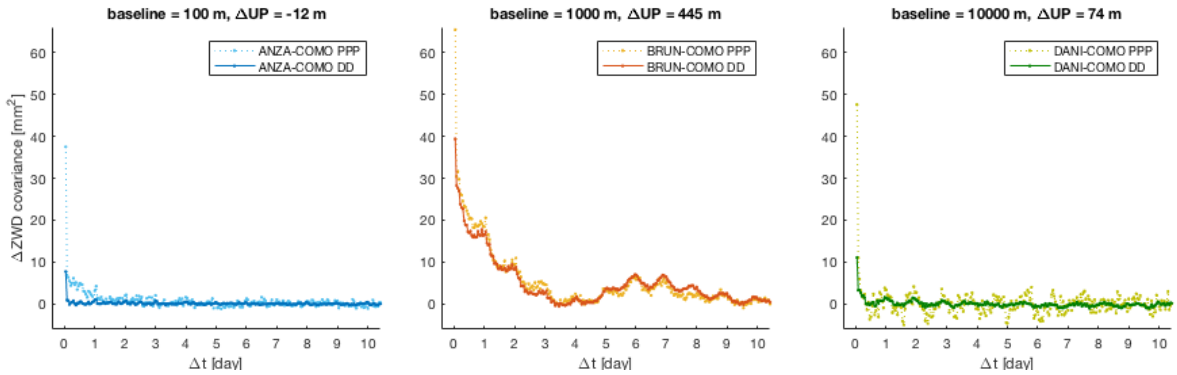

Figure 4.5 – Empirical covariance functions of DD and PPP-derived ZWD increments

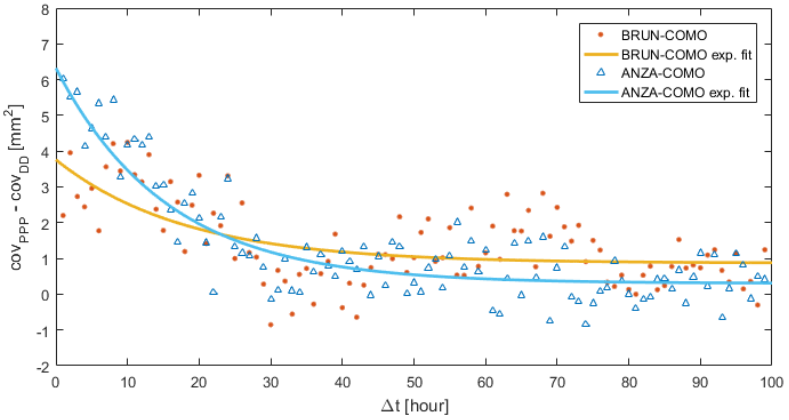

Figure 4.6 – Differences between DD and PPP-derived ΔZWD empirical covariance functions of BRUN-COMO and ANZA-COMO baselines

Chapter 5

GNSS data processing of mid-scale SPIN network: temporal stack of 2D ZWD maps

5.1 Introduction

Along with the study described in Chapter 4 regarding the dense GNSS network MisT, we gave a closer look at GNSS networks where the inter-distances of nearby stations are approximately 50 km, thus including in our research the case of mid-scale networks.
The case study we chose for our analysis consisted in one week of July 2016, where two severe rainfall events occurred in Lombardia region on the 22nd and 26th. These two events, indeed, appeared to have very different features: the one on the 22nd of July was characterized by an advective behavior, while the event on the 26th showed strongly localized rainfall events.
In order to understand to which extent GNSS-derived ZWD time series would give evidence of such behaviour or, more generally, would retain information of the spatio-temporal variability of the water vapor content, we tried to characterize spatially the ZWD field for every epoch available. In this way we obtained a "stack" of 2-dimensional residual ZWD maps (or equivalently residual PWV maps; see Chapter 2) parametrized by the time at the ZWD processing rate.
In this chapter, we will describe in detail the procedure adopted to compute this collection of maps and we will explore their significance regarding water vapor monitoring purposes.

5.2 GNSS data processing setup

In this and next sections we will briefly describe the procedure and the GNSS processing setup for estimating the ZWD time series.
As already mentioned in the Introduction of this chapter, the case studies chosen in this case consisted in one week of July 2016, where two severe rainfall events occurred in Lombardia region on the 22nd and 26th. These two severe storms showed a quite different meteorological nature: the one on the 22nd of July was characterized by a strong advective behavior with a weather front prop-

agating from South-East to North-West, whereas the event on the 26th showed strongly localized convective cells. This behaviour could be easily recognized in the RADAR images shown below.

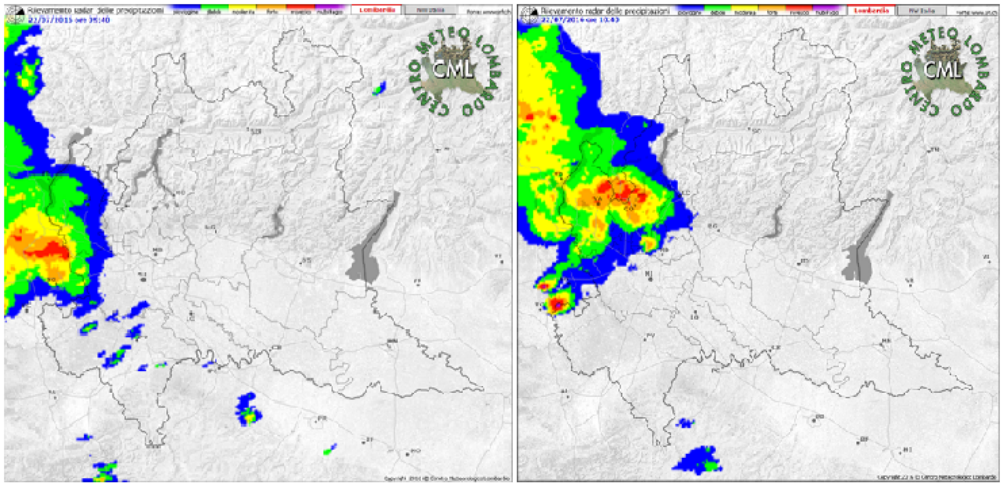

Figure 5.1 – Radar image relative of the event on the 22nd of July at 9:40 and 10:40 AM (Courtesy of ARPA - Agenzia Regionale per la Protezione Ambientale)

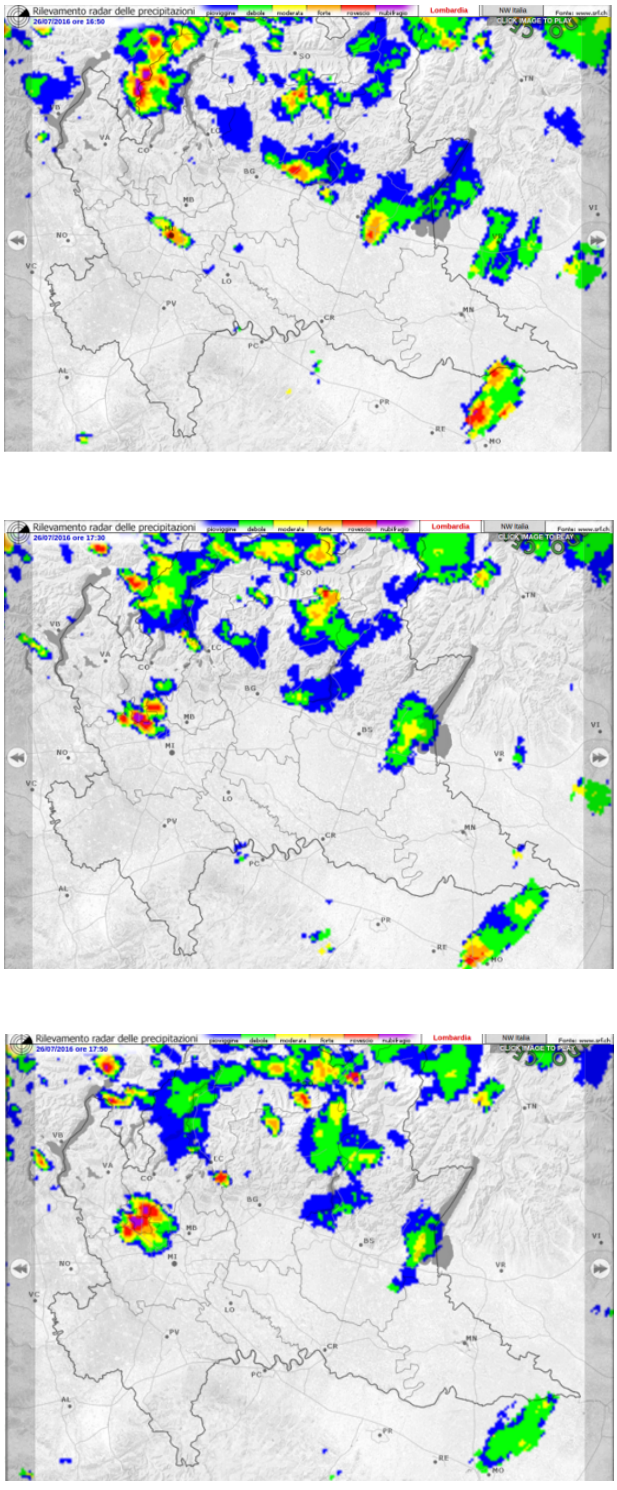

Figure 5.2 – Radar image relative of the event on the 26th of July at 16:50, 17.30 and 17:50 AM (Courtesy of ARPA - Agenzia Regionale per la Protezione Ambientale)

In order to analyze to which extent GNSS-derived ZWD time series would give evidence of these phenomena we processed GNSS data from 21st to 29th of July acquired by the interregional GNSS SPIN network.
This network consists of 30 dual-frequency geodetic receivers covering quite homogeneously the two Italian regions Lombardy and Piedmont. The inter-distances of close by stations of the SPIN network range between 40 and 50 km and it may be considered as a regional network. The signals received are GPS and GLONASS. Among other services provided by this network, RINEX observation files at 30 s acquisition rate can be freely downloaded. For the time period of interest in our case, GPS observations files were available for all stations and without interruptions and we use them for our purposes (we did not use GLONASS data).

Figure 5.3 – The interregional SPIN GNSS network

In order to process the GPS observations, we made use of the BERNESE v5.2 software (throughout this chapter denoted again by BSW) and we adopted a PPP

strategy, thus considering iono-free L3 observables, with a batch least-square based adjustment algorithm (cf. Chapter 4 for further details on this point). Detailed info are reported in Table 5.1

Table 5.1 – SPIN GNSS network - PPP processing setting

Network design	30 dual-frequency GNSS receivers
Received signals	GPS, GLONASS
Processing SW	BSW v5.2
Processing rate	30 s
Elevation cutoff	10 degrees
Processed signals	GPS
Processing method	PPP
Ephemeris	CODE (final)
Satellite clocks	CODE 30 s (final)
Ocean loading model	FES2004
Observables	L3 (iono-free)
Mapping functions	GMF dry/wet
Dry tropospheric model	Saastamoinen + GPT
Gradient estimation model	none
ZWD estimation rate	10 min
A priori ZWD difference std	0.02 m

As we did for processing Mist data, we set an elevation cut off angle of 10 degrees. In this case we dropped instead the use of the Niell mapping function and chose the dry and wet Global mapping functions that lead to an estimation of ZWD based on the Global Pressure Temperature model (cf. Chapter 2).
Once again, to deal with the great amount of BSW output files, we took advantage of the software developed by ourselves that manages the automatic import of all these files and allows manipulation and analysis of the estimated coordinates and ZWD time series. This software has been extensively used to produce the results we are showing in next sections.

5.3 GNSS data estimation

As described in the previous section, we first obtained ZWD time series for all SPIN network's stations, as Zenith Path Delay corrections of the Zenith Dry Delay computed by relying on the Global Pressure Temperature model. This model is a global average representation of the pressure and temperature distribution obtained from long-term gridded monthly Numerical Weather Model pressure and temperature data and it accounts mainly for regional scale variations (cf. 2 and the paper [**Kleijer(2004)**], [**Kouba(2009)**]).
In this case, luckily we had also available some accurate temperature and pressure measurements provided by meteorological stations closeby the GNSS receivers at

the ZWD estimation rate throughout the period of interest. Thus we could compute the a-priori ZHD using the measured pressure data, not relying on the global GPT model. Furthermore the ZHD could be transformed into accurate PWV values taking advantage of the acquired temperature data.
In all cases where the measurements were not available right at receivers locations or at the right time, temperature and pressure data were interpolated by means of deterministic (i.e. spline interpolation) and stochastic methods (i.e. with collocation algorithms) in order to predict them at the station's positions and to have a homogeneous time resolution. These computations have been collected in [**Barindelli et al.(2018)**]. The final ZWD time series were eventually obtained by subtracting these actual and accurate ZHDs from the total Zenith Path Delay, estimated by means of the BSW.
As examples the accurate ZHD and the corresponding ZWD and PWV time series for the three SPIN stations ALSN, COMO and MILA relative to both events are reported below. On the 26th of July, the data started at 6:30 AM.

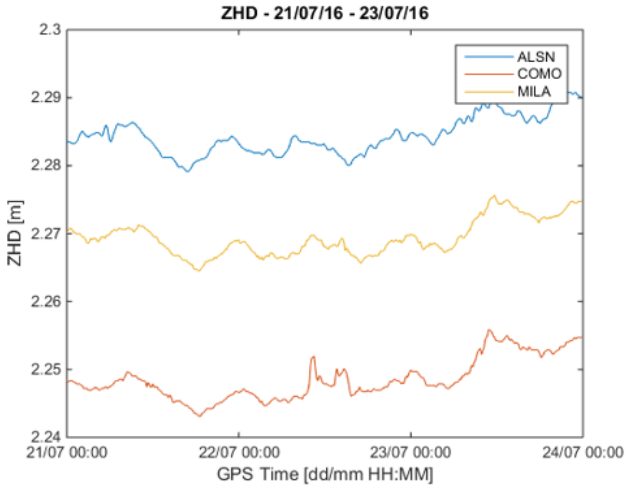

Figure 5.4 – Accurate ZHD time series of stations ALSN, COMO and MILA on 22nd of July

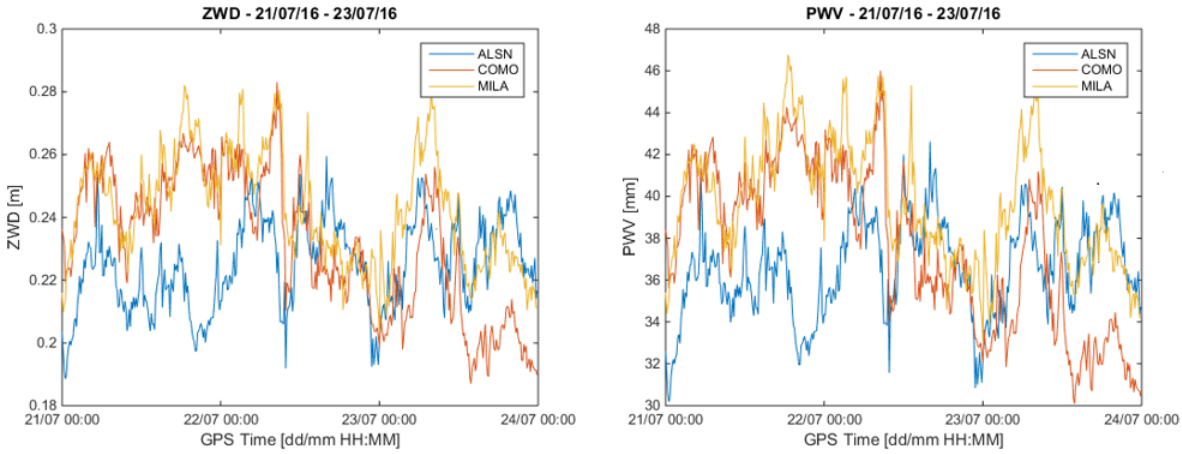

Figure 5.5 – ZWD and PWV time series of stations ALSN, COMO and MILA around 22nd of July

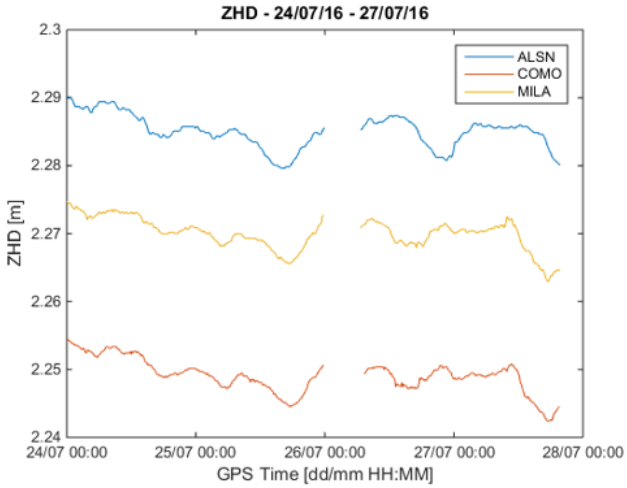

Figure 5.6 – Accurate ZHD time series of stations ALSN, COMO and MILA around 26nd of July

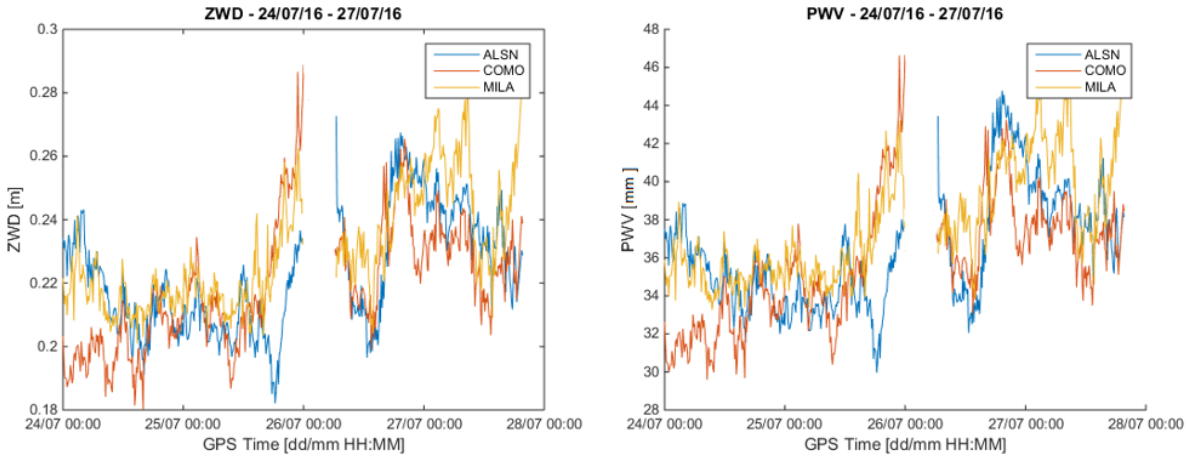

Figure 5.7 – ZWD and PWV time series of stations ALSN, COMO and MILA around 26nd of July

In Figure 5.5 it is easily seen that the ZWD and PWV time series of all stations follow a similar pattern, especially those of COMO and MILA. Furthermore, we may recognize a significant decrease of the ZWD (and equivalently of PWV) of these two stations between 9:30 AM and 9:40 AM of the 22nd of July. It turned out that this corresponded to the weather front passage above these two stations. For what concerns ALSN's time series one could register a smaller drop in the ZWD around the same epochs, again probably due to the weather front propagation.
For the second event occurring on the 26th of July the situation is less clear and it was not possible to link the ZWD time series variability of the stations to the meteorological phenomenon underway.

5.4 Spatial stochastic modelling of the SPIN GNSS network

With the aim of investigating the spatio-temporal variability of the ZWD field on the area covered by the SPIN network during both events, we computed a temporal stack of $2-$dimensional ZWD maps for all epochs available.
In this section and in the next one we will review the procedure followed and the assumptions we needed to make to achieve this. The approach we pursued is that explained in Chapter 3. For basic definitions and properties we refer to this Chapter.
The first fundamental step we did was assuming the ZWD field to be **intrinsically homogeneous** and **isotropic** with respect to the spatial coordinates (please refer to Chapter 3 for details). In particular recall that this means that the mean of the *increments* of the ZWD field are zero

$$\mathbb{E}(ZWD(x+h) - ZWD(x)) = m(\|h\|) = 0 \tag{5.1}$$

and that their halved variance exists and it only depends on the norm $\|h\|$ of the vector h

$$\frac{1}{2} \cdot var(ZWD(x+h) - ZWD(x)) = \gamma(\|h\|) \tag{5.2}$$

As defined in Chapter 3, the functions γ are called variogram functions and will be the object of our study.
The estimated ZWDs for each epoch t_0 are thought of as a realization of the ZWD random field at the epoch t_0 and at particular points, namely the positions of the receivers. Thus they are used to determine empirical variogram functions, describing the spatial dependence and distribution of the ZWD field. However, Equation 5.1 implies that we cannot use the estimated ZWD time series directly but we need first to reduce the mean height-depending component. In fact SPIN's stations have different height and the ZWD depends on this component. If we would not have removed a mean ZWD value due to the height of the station, we would not have fulfilled Equation 5.1. Thus, to overcome this problem and determine the empirical variograms we computed the residual ZWD time series for each station, subtracting its sample mean from the time series itself. Physically this implied that topography-dependent trends were removed before computing the empirical variograms. Ultimately the epoch-wise spatial empirical variogram

functions were estimated from the data by exploiting the intrinsically homogeneity and isotropic assumptions and the methodology presented in Chapter 3. To obtain the empirical variogram from the variogram cloud, we built a 15 points grid with a step of 25 km. The numbers above each point in the variogram count how many dissimilarities contributed to the computation of that average variogram value.
In Figure 5.8 we report some examples. The first column is related to the 22nd July at the two epochs 9:40 and 10:40 AM (when the weather front passed over the stations COMO and MILA). The second column shows the empirical variograms referred to the 26th July at 4:40 and 5:30 PM. Furthemore the last two figures represent the collection of all empirical variograms parametrized by time. More precisely they are the projections of all empirical variograms on a plane with time and spatial lag as, respectively, abscissa and ordinate. The different colors render the value of each variogram for each point in this plane. As one may see these two images are quite difficult to be interpreted and it is not easy to extract a signature of the meteorological phenomena underway.

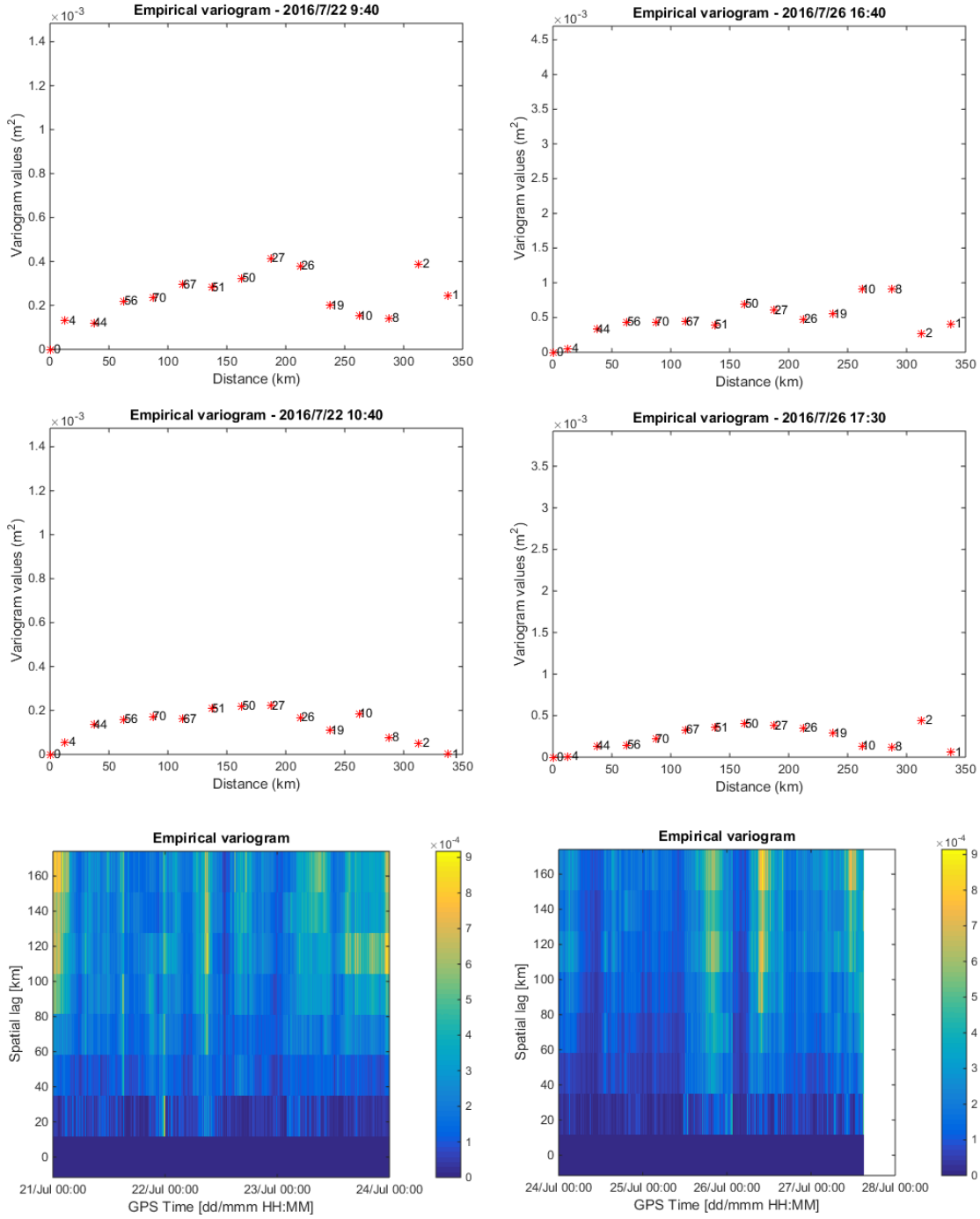

Figure 5.8 – Empirical variograms for both 22nd and 26th of July

As explained in Chapter 3 the empirical variogram is used as a first data-driven estimate of the theoretical variogram needed for spatial interpolation by the Kriging algorithm.
In our case the empirical variograms were fitted with a theoretical model up to about 150 km and the models investigated were Fractional Brownian motion models. These models have a variogram functions of the form

$$\gamma(\|h\|)_t = C_t \cdot (\|h\|)^{\alpha_t}, \tag{5.3}$$

where $C_t \in \mathbb{R}$ and $0 < \alpha_t < 2$, for all available epochs t.

Remark. As mentioned in Chapter 3, these models seem to retain a physical meaning related to atmospheric turbulence as dictated by Kolmogorov's theory where such variogram functions with $\alpha = \frac{2}{3}$ are predicted.

In order to assess their applicability in the case under study, we estimated for each epoch the model parameters (C_t, α_t) using a least square approach and the empirical variograms as observation data.
As an example the estimates $(\hat{C}_t, \hat{\alpha}_t)$ of the model parameters from 24th to 27th of July are reported in Figure 5.9

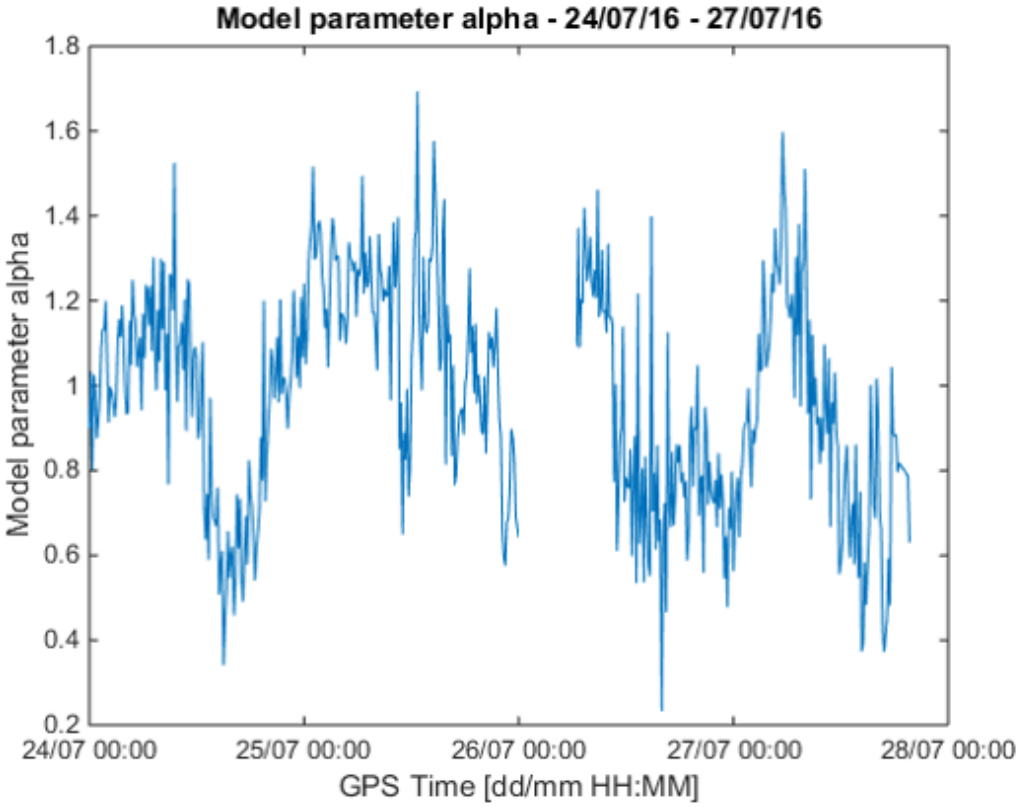

Figure 5.9 – Epoch-wise estimates of the model parameters for the theoretical variograms.

Remark. From this empirical analysis it turned out that the estimates $\hat{\alpha}_t$ of the exponent of a Fractional Brownian motion model differ substantially from the value $\alpha = \frac{2}{3}$ predicted by Kolmogorov's theory. Moreover, the mean of the time series is approximately 0.96. This suggests that, at least for mid-scale GNSS networks, variogram functions as stated in Kolmogorov's theory may lead to an artificial correlation of the random field and therefore to miscalculations of the Kriging estimated points.

Using the solutions $(\hat{C}_t, \hat{\alpha}_t)$ and applying Equation 5.3 we obtained a set of theoretical variogram models indexed by the ZWD estimation rate. These models seem to fit well the empirical variograms for all epochs, thus suggesting that the choice of Fractional Brownian motions is adequate for modelling the phenomena we have focused on (under the assumptions we have made). Some examples of the theoretical variogram functions (blue continuous line) against its corresponding empirical variogram (red points) for both events are shown in Figure 5.10.
It is worth pointing out that many theoretical variograms show an evident unbounded behaviour at infinity implying in particular not to have a covariance function counterpart. This in turn indicates that we could not have possibly modelled the ZWD field as a weakly homogeneous and isotropic stochastic field, since for this class of random fields the covariance functions are naturally bounded (cf. Chapter 3 for further details).

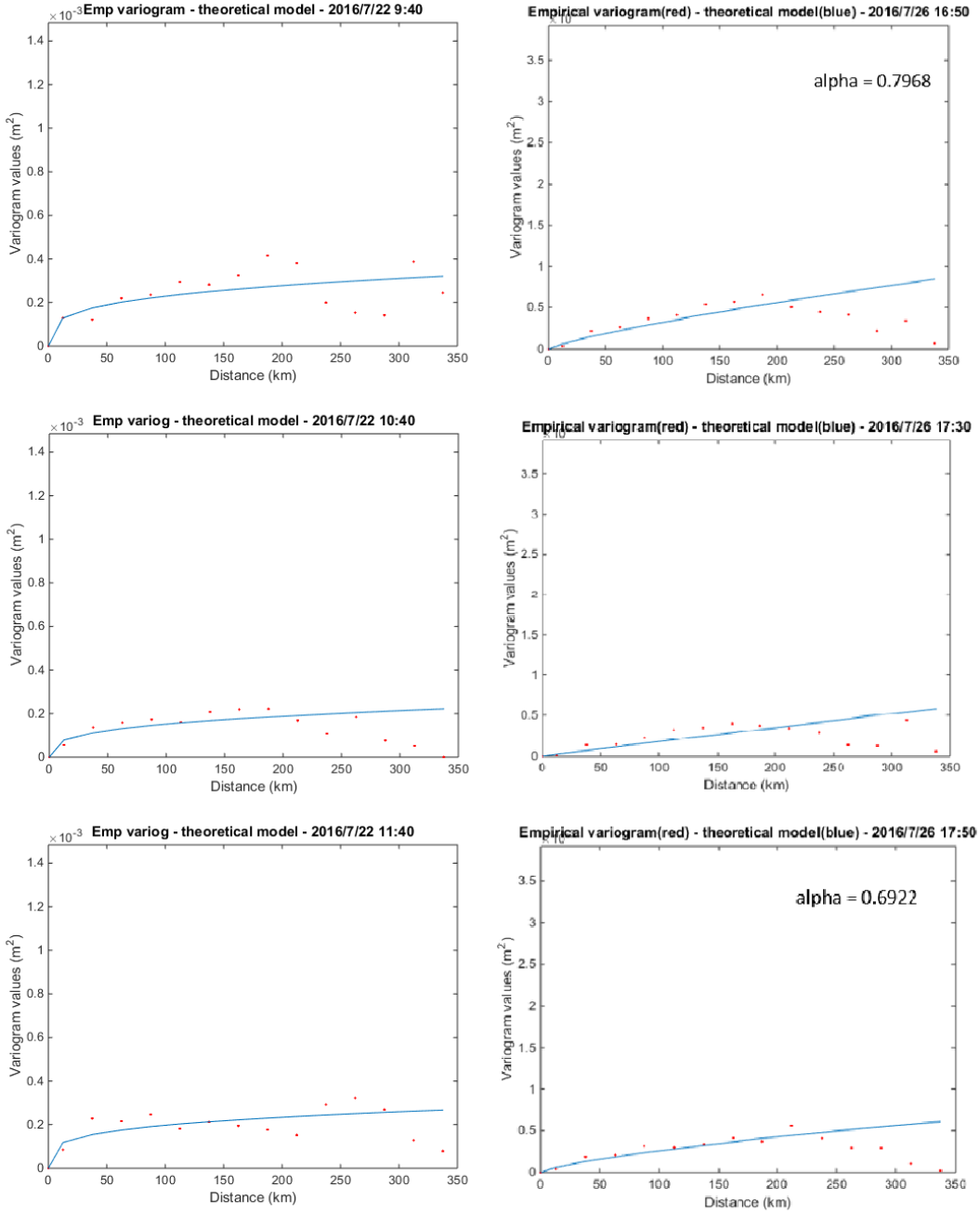

Figure 5.10 – Theoretical variograms for both 22nd and 26th of July

5.5 Temporal stack of 2D ZWD maps: results and discussion

In order to retrieve spatially accurate and high-resolution 2D ZWD maps (or equivalently PWV maps), the SPIN network *residual* ZWDs are interpolated epoch-wise via ordinary Kriging using the theoretical variogram functions introduced in the previous section.
Let us recall briefly the main aspects of the Kriging algorithm applied to the ZWD random field, referring to Chapter 3 and the references therein for a general treatment.
The ordinary Kriging algorithm allows to estimate a value of the residual ZWD random field at a point of its definition region not directly observed but for which a variogram is known. More precisely, the estimate of the residual $ZWD_R(x_0)$ at a point x_0 not directly measured is given by a linear combination of the residual $ZWD_R(x_\alpha)$ direcly observed at each network's station

$$\widehat{ZWD_R}(x_0) = \sum_\alpha \mu_\alpha ZWD_R(x_\alpha) \tag{5.4}$$

Under this assumption and the intrinsic homogeneity one, it follows that

$$\sum_{\alpha=1}^{N} \mu_\alpha = 1 \tag{5.5}$$

and that the estimation variance (cf. Chapter 3) is equal to

$$\sigma_E^2 = \mathbb{E}[(\widehat{ZWD_R}(x_0) - ZWD_R(x_0))^2] = \\ - \gamma(0) - \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} \mu_\alpha\mu_\beta\gamma(x_\alpha - x_\beta) + 2\sum_{\alpha=1}^{N} \mu_\alpha\gamma(x_\alpha - x_\beta) \tag{5.6}$$

By minimizing σ_E^2 with the constraint given by Eq. 5.5, one obtains the ordinary Kriging system

$$\begin{pmatrix} \gamma(x_1 - x_1) & \cdots & \gamma(x_1 - x_N) & 1 \\ \vdots & & \vdots & \vdots \\ \gamma(x_N - x_1) & \cdots & \gamma(x_N - x_N) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \\ \lambda \end{pmatrix} = \begin{pmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_1 - x_0) \\ 1 \end{pmatrix} \tag{5.7}$$

Solving this system yields an estimate of $\widehat{ZWD_R}(x_0)$ at the point x_0.
A stack of 2D residual ZWD maps parametrized by the ZWD estimation rate was obtained by applying the ordinary Kriging algorithm to a homogeneous grid of 21 points built around the SPIN network. The grid points compared to the SPIN network's station positions are shown in Figure 5.11.

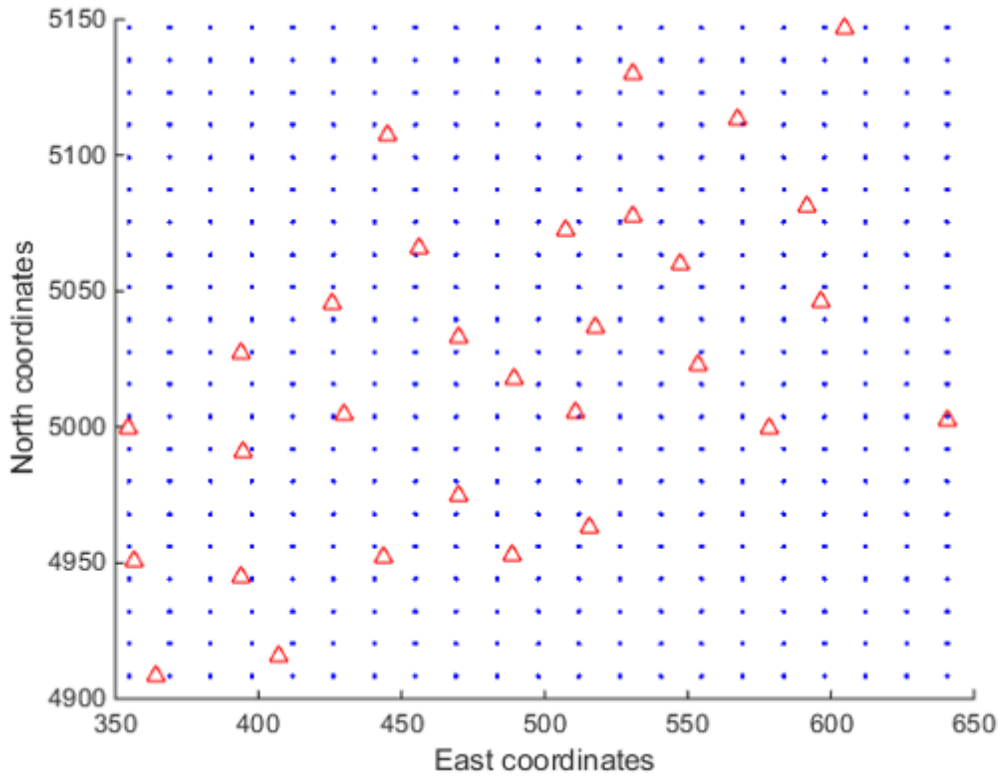

Figure 5.11 – Grid points (blue points) with respect to SPIN network's receivers (red triangles)

Some examples of residual ZWD maps at different epochs for both events are instead presented in Figures 5.12 and 5.13.
It is worth recalling that modelling the ZWD random field as an intrinsic homogeneous and isotropic field implied the removal of topography-dependent trends (cf. Section 5.4). Hence the collection of these images describes the variability of the ZWDs around an average value (mainly depending on the station's height).
The analysis of the stack of the residual ZWD maps relative to the 22nd of July suggested that such GNSS-derived images may withhold some qualitative information about the behaviour of the meteorological phenomenon ongoing. In the attempt to interpret at best these information, we have placed the residual ZWD maps side by side with the the RADAR images available from the regional environmental protection agency.
The weather radar witnessed at 9:40 AM the meteorological front passing from Piedmont region to Lombardy, approaching the SPIN stations of COMO and LECC (cf. Figure 5.12). The residual ZWD maps corresponding to this epoch records a value of approximately $2.0 - 2.2$ cm in the Eastern part of Lombardy that includes the SPIN stations SONP, DARF and BRES. This region is far away from the meteorological event and does not seem to be affected by the weather front approaching.

On the other hand the residual ZWD map shows a very low variability of less than about 5 mm in a region that includes the weather front and extends down to the southern part of Lombardy (closeby the SPIN stations SERR and VARZ). This virtually constant value of the ZWD suggests that in this area there is hardly any water vapour transfer nor rise of moist air above the GNSS receivers. At 10:40 AM the weather front moved and we may observe that the corresponding residual ZWD map exhibits a larger variability ranging from a few mm up to about $1 - 1.5$ cm. This may be due to the pressure gradients in the front itself and to the incoming summer solar radiation, increasing in magnitude and causing therefore a variation of the ZWD spatial distribution. When the weather front passed away or reduced significantly at 11:40 we may finally observe that the residual ZWD map presents an inhomogeneous variability of the ZWD in the order of a few cm on the whole map. One possible reason for this behaviour could be again the incoming solar radiation leveraging a spatial redistribution of the water vapour. A drawback of the above analysis is however that the meteorological event under study is naturally anisotropic but the residual ZWD maps cannot hold any knowledge about this feature. In fact as a consequence of the isotropicity hypothesis only an averaged "spherically symmetric" behaviour of the occurrence is reflected into the residual ZWD maps.
This limitation is even clearer in the case of the phenomenon occurred on the 26th of July. As shown in Figure 5.13, from about 4:30 PM to 6:00 PM of the 26th of July the event was characterized by the development of convective cells mostly across the northern part of Lombardia region, whereas the southern part of the region was not affected by the phenomenon. These cells were virtually not advected even though a slight movement from South-East to North-West can be detected in the radar images. In this case, the residual ZWD maps corresponding to this time window show only a prominent stepwise variability in the upper right part of the map (i.e. approximately the northern part of Lombardia). This may indicate privileged directions of water vapour transfer along the development of the cells themselves. However they do not exhibit any hint about the potential position of the cells nor a clear separation between the northern and the southern part of Lombardia. The reason for this inaccuracy of the ZWD maps can rely on the intrinsic anisotropicity of the event, which cannot be captured by the variogram models used.
Dealing with purely intrinsically homogeneous models without assuming them to be also isotropic raises several issues, in particular the problem of determining the theoretical variograms from the empirical ones, and it represents a novel research perspective.
The last issue we would like to address is the problem of retrieving absolute ZWD maps, i.e. 2D maps with absolute ZWD values instead of residual ones. In order to compute such maps it is necessary to re-introduce topography-dependent information on the network area and to compute the mean height-dependent component of the ZWD at each grid point.
One possible strategy to carry this out consists in the following steps:

1. choose a model that expresses the average component of the ZWD due to the receiver's height

2. select a suitable Digital Elevation Model of the network area, in order to know as accurately as possible the height of each grid point

3. add these mean contributions to the residual ZWD maps

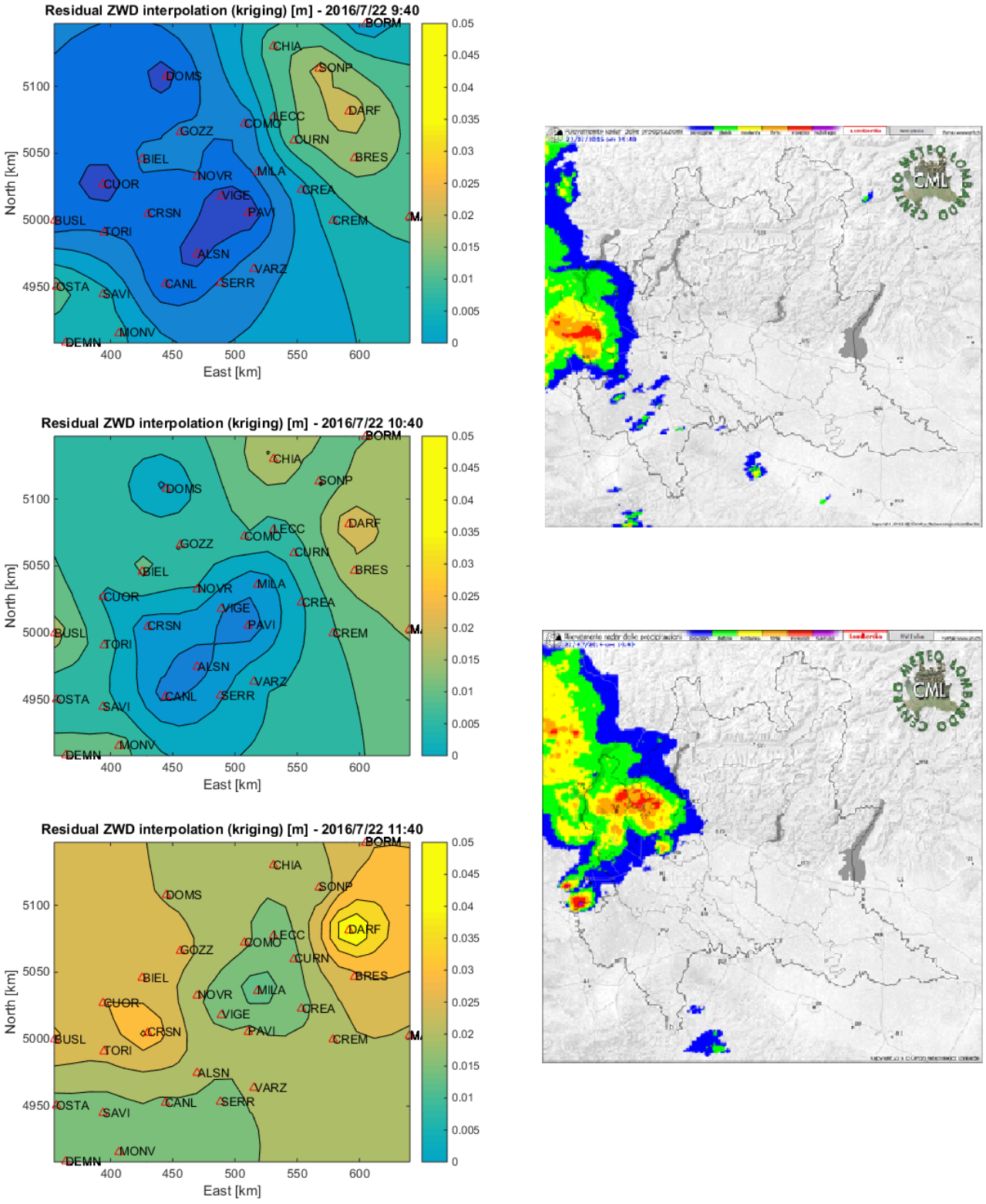

Figure 5.12 – **Left side:** Examples of 2D residual ZWD maps of the event on the 22nd of July at three different epochs
Right side: Radar image relative of the event on the 22nd of July at 9:40 and 10:40 AM (Courtesy of ARPA - Agenzia Regionale per la Protezione Ambientale)

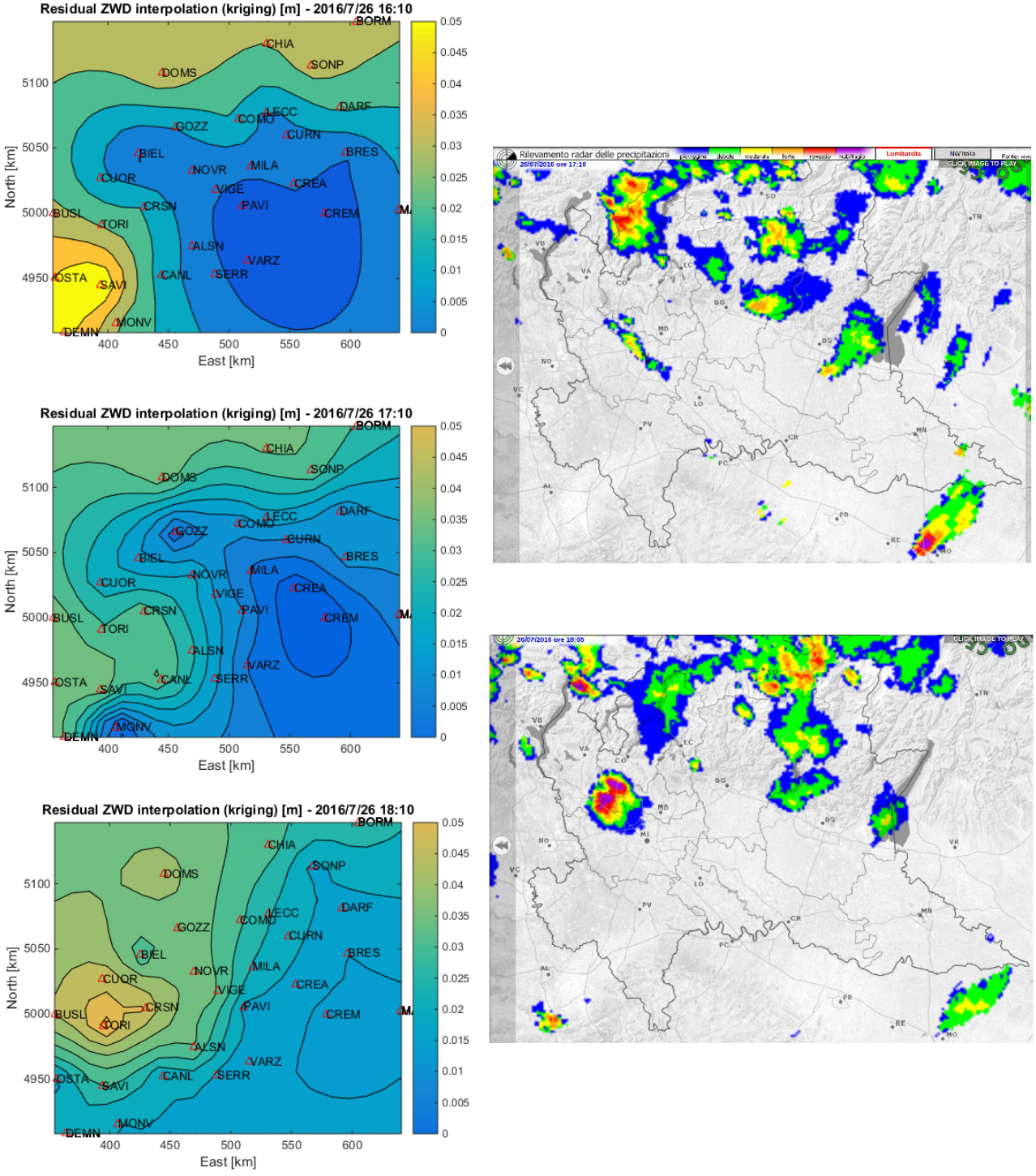

Figure 5.13 – **Left side:** Examples of 2D residual ZWD maps of the event on the 26th of July at three different epochs
Right side: Radar image relative of the event on the 26th of July at 17:10 and 18:05 AM (Courtesy of ARPA - Agenzia Regionale per la Protezione Ambientale)

This would allow to retrieve absolute PWV maps and to analyze them together with independent high-resolution spatial weather sensors, as for instance weather radar, in order to figure out to which extent GNSS-derived maps retain spatio-temporal meteorological information.
This step, however, has not been concluded yet and the study of GNSS-derived absolute PWV water vapor maps with such sensors represents a future research direction.

Chapter 6

Toward spatio-temporal stochastic modelling of the ZWD field

6.1 Introduction

In the previous chapter we have characterized spatially the residual ZWD field and we have obtained a "stack" of 2-dimensional residual ZWD maps for every epoch available. The temporal dimension is not modelled in any way and it only acts as parameter.
In order to investigate the full spatio-temporal structure of the ZWD field and to explore the feasibility of temporal predictions, we need to have a basic understanding of its dynamics. The spatio-temporal evolution of the ZWD field is related to the fluid dynamic properties of the neutral atmosphere and we have detected the contributions of at least two main terms responsible for its evolution. First of all a *"diffusion term"*, for instance due to turbulence effects or evaporation/condensation phenomena. Secondly an *"advection term"* due to a rigid transport of a parcel of air along the fluid velocity field.
In this chapter starting from the conservation equation of the moisture of a parcel of air we will derive an advection-diffusion model for the ZWD field. If the velocity field is assumed to be constant, we obtain as a particular case the so-called *Taylor's frozen flow hypothesis* (cf. [**Treuhaft et al.(1987)**] or [**Tatarski(1961)**]). Under this hypothesis we will present an algorithm for estimating a time-dependent spatially-averaged mean transport field in order to quantitatively detect the advection term. This algorithm could be used in case of dense or mid-scale GNSS networks with no heterogeneous topography.
More generally if the velocity field of the advection-diffusion model is only divergent-free, we are going to present a method for giving an estimate of the mean transport field both time and space-dependent. Moreover, in this case it is also possible in principle to take into account turbulent mixing effects, reckoning quantitatively the diffusion term mentioned above.
This chapter relies heavily on [**Onn et al.(2006)**], where many ideas are taken from. It is also worth pointing out that it only represents a primer to the complex topic of the characterization of the spatio-temporal structure of the ZWD field

and further insights are necessary.

6.2 ZWD advection model

In this section we will derive a general advection model for the ZWD field starting from the conservation equation of moisture of a parcel of air.
This equation indicates that the total amount of water vapour in a parcel of air is conserved as the parcel moves around except when there is a phenomenon of evaporation or condensation

$$\frac{\partial q}{\partial t} + \boldsymbol{v} \cdot \nabla q = E - C \tag{6.1}$$

where the specific humidity of the parcel is denoted by q, the 3D flow velocity field by $\boldsymbol{v}$ of the Troposphere, the condensation by C and the evaporation by E. Multiplying Equation 6.1 by the density ρ of a parcel of (moist) air and adding to it the mass conservation equation multiplied by q

$$q(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v})) = 0 \tag{6.2}$$

one obtains the conservation equation of the specific humidity of a parcel of air

$$\frac{\partial(\rho q)}{\partial t} = -\nabla \cdot (\rho q \boldsymbol{v}) + \rho(E - C) \tag{6.3}$$

Equation 6.3 will be our starting point for constructing the advection model for the ZWD field.
We will assume throughout that

$$\nabla \cdot \boldsymbol{v} = 0 \tag{6.4}$$

This means the fluid be incompressible, which may be assumed for mesoscale weather systems ranging vertically up to approximately 4 km and horizontally up to length scales as those covered by dense and mid-scale GNSS networks. This is because density variations in the lower atmosphere and in relatively small areas are small (see [**Onn et al.(2006)**] and the references therein). Under the incompressibility hypothesis Equation 6.3 is equivalent to

$$\frac{\partial(\rho q)}{\partial t} = -\boldsymbol{v} \cdot \nabla(\rho q) + \rho(E - C) \tag{6.5}$$

The equation above states that time variations of humidity ρq of a parcel of air is given by its net flux through the boundary surface of any volume element. Assuming there is neither condensation nor evaporation and integrating Equation 6.3 (or equivalently Equation 6.5) along the vertical component, one obtains (see also [**Onn et al.(2006)**]) the equation

$$\frac{\partial}{\partial t}(\int_0^H \rho q dz) + \frac{\partial}{\partial x}(\int_0^H u \rho q dz) + \frac{\partial}{\partial y}(\int_0^H v \rho q dz) + (w\rho q)_{|H} - (w\rho q)_{|0} = 0 \tag{6.6}$$

where (u, v, w) are the componenents of the flow velocity field. The boundary terms represent the water vapor exchange between the ground and the upper part of the neutral atmosphere considered (typically the top of the planetary boundary layer).
The Integrated Water Vapor can be expressed as (see Chapter 2)

$$IWV = \int_0^H \rho q dz$$

and let us define the vertically averaged transport field components as

$$U = \frac{\int_0^H u\rho q dz}{\int_0^H \rho q dz}, \quad V = \frac{\int_0^H v\rho q dz}{\int_0^H \rho q dz} \tag{6.7}$$

Substituting these two equations in Equation 6.6, we obtain an advection model for the IWV field

$$\frac{\partial IWV}{\partial t} + \frac{\partial}{\partial x}(U \cdot IWV) + \frac{\partial}{\partial y}(V \cdot IWV) + moisture\ sources = 0 \tag{6.8}$$

This model characterizes the time variations of the water vapour in a column of air as the horizontal net flux of IWV into that column (along the transport field (U, V)).
Taking into account the linear relation $IWV = \kappa ZWD$, as explained in Chapter 2, where κ is assumed to be constant, Equation 6.8 can be straightforwardly expressed in terms of the ZWD, yielding an advection model for the ZWD field.

$$\frac{\partial ZWD}{\partial t} + \frac{\partial}{\partial x}(U \cdot ZWD) + \frac{\partial}{\partial y}(V \cdot ZWD) = 0 \tag{6.9}$$

6.3 Taylor's Frozen flow hypothesis and estimation of the mean transport field

Let us start from the advection model for the ZWD field given by Equation 6.9. If we consider as a particular case that the vertically-averaged transport vector (U, V) is constant and that there are no source or sink terms, Equation 6.9 has the solution

$$ZWD(x, y, t) = ZWD(x - Ut, y - Vt, t_0) \tag{6.10}$$

where ZWD_0 is the initial ZWD distribution at time t_0. Equation 6.10 is classically known as the *Taylor's frozen flow hypothesis*. In general this assumption means that the diffusion contributed by turbulent circulations is small and that the advection of a parcel of air past a given point can be taken to be entirely due to the mean flow velocity field. In particular, in the unusual case of the ZWD field it implies that the ZWD value is rigidly transported around by means of the transport vector (U, V) defined in Equation 6.7.
Throughout this section we will assume the ZWD field to be evolving according to

Equation 6.10. If the transport vector (U, V) would be known, then it would be possible to take advantage of frozen flow hypothesis (when applicable) in order to make predictions of the ZWD values at any point (x, y) and time t. Unfortunately this vector is in general not known and must be estimated.
A possible way to give an evaluation of the mean transport is using Equation 6.10 itself and set an inverse problem for it with unknowns (U, V).
Following this idea and according to Taylor's frozen flow hypothesis, we developed a data-driven procedure to evaluate (U, V). In order to set this inverse problem the input data of the ZWD field may be provided by a collection of known ZWD values at different epochs or by a set of predicted (absolute or residual) ZWD values parametrized by the time, e.g. the collection of Kriging predicted residual ZWD values as derived in Chapter 5.
Hence, let us suppose the spatio-temporal ZWD field to be defined on a region whose dimension is comparable to those covered by a dense or mid-scale GNSS network (for instance as the MisT network or the SPIN network). Moreover let us consider a stack of 2D ZWD maps $M_{ZWD,t}$ parametrized by the ZWD etimation rate t as described in Chapter 5.
Taking *pairs* $(M^{ZWD}_{t_k}, M^{ZWD}_{t_{k+1}})$ of such maps at two consecutive epochs t_k and t_{k+1} the following equation holds for any grid point (x_l, y_l)

$$ZWD(x_l, y_l, t_{k+1}) = ZWD(x_l - U(t_k)\cdot(t_{k+1} - t_k), y_l - V(t_k)\cdot(t_{k+1} - t_k), t_k) \quad (6.11)$$

Hence for any epoch t_k (with the exceptions of the final epoch) Equation 6.11 represents a non-linear system where the spatially-constant time-varying transport field $(U(t_k), V(t_k))$ is the unknown.
In order to effectively solve this system, we linearize Equation 6.11, obtaining a linear system

$$ZWD(x_l, y_l, t_{k+1}) \approx ZWD(x_l, y_l, t_k) - \nabla ZWD(x_l, y_l, t_k) \cdot \begin{pmatrix} U(t_k) \\ V(t_k) \end{pmatrix} (t_{k+1} - t_k) \quad (6.12)$$

The accuracy of the linearization is as accurate as the ZWD estimation rate is high.
A last step we need to carry out to solve the system 6.12 is estimating the gradient $\nabla ZWD(x_l, y_l, t_k)$. We explored two alternative approaches

1. 2° order finite difference approximation of the derivatives

2. stochastic prediction of the derivatives via the Kriging algorithm

The first method consists in computing suitable combinations of differences among the Kriging ZWD estimates at any grid point (we refer for instance to [**Quarteroni(2014)**] for further details). It is worth pointing out that here we must use implicitly the measured ZWD values for every stations and suitable variogram functions as for example those computed in the previous chapter.
The second approach amounts to express directly the gradient of the ZWD field

at any grid point as a linear combination of the measured ZWD values at the receivers locations and evaluate it applying the Kriging ordinary algorithm. In this case, too, it is necessary to take advantage of the variogram functions computed in the above mentioned chapter in order to set up the Kriging system. For further details please see [**Reguzzoni et al.(2005)**].
Once the gradient $\nabla ZWD(x_l, y_l, t_k)$ has been assessed by means of one these two methods, the system 6.12 evaluated at all grid points can be solved epoch-wise through a least squares adjustment approach.
Some examples of the results of this algorithm for the stack of 2D residual ZWD maps computed in Chapter 5 are reported in Figure 6.1. These images refer to the two events occurred on the 22nd of July and on the 26th of July 2016 at three different epochs.

The above algorithm was formulated with the aim to give an estimate of both the magnitude and the direction of the advective term of the dynamical evolution of the ZWD field. In our opinion this represents a first step toward a *joint* spatio-temporal modelling of the ZWD field, that could be useful for meteorological purposes and rainfall events hazards.
The above procedure, moreover, could be of special interest for dense GNSS networks with no heterogeneous topography, since the interdistances between two stations are so small that the mean transport vector $(\bar{U}, \bar{V})$ may be assumed spatially-constant. In this case, indeed, we expect only the variation in time of the advection term to be significant.
A problem regarding the method we tried to tackle was its validation. More precisely, we attempted to verify that the estimates $(\hat{U}, \hat{V})$ of the transport vector obtained by this algorithm can be linked to the meteorological flow velocity field $\vec{v} = (u, v, w)$ and are reliable estimates of Equation 6.7. However, to compute the integral 6.7 and effectively validate our estimates we would need wind and humidity values at different heights of the neutral atmosphere. These values could be either real data provided by meteo sensors, as for instance radiosondes, or simulated data obtained from NWP models. Either way unfortunately these data were not available to us and thus we were not able to proceed further on this point.

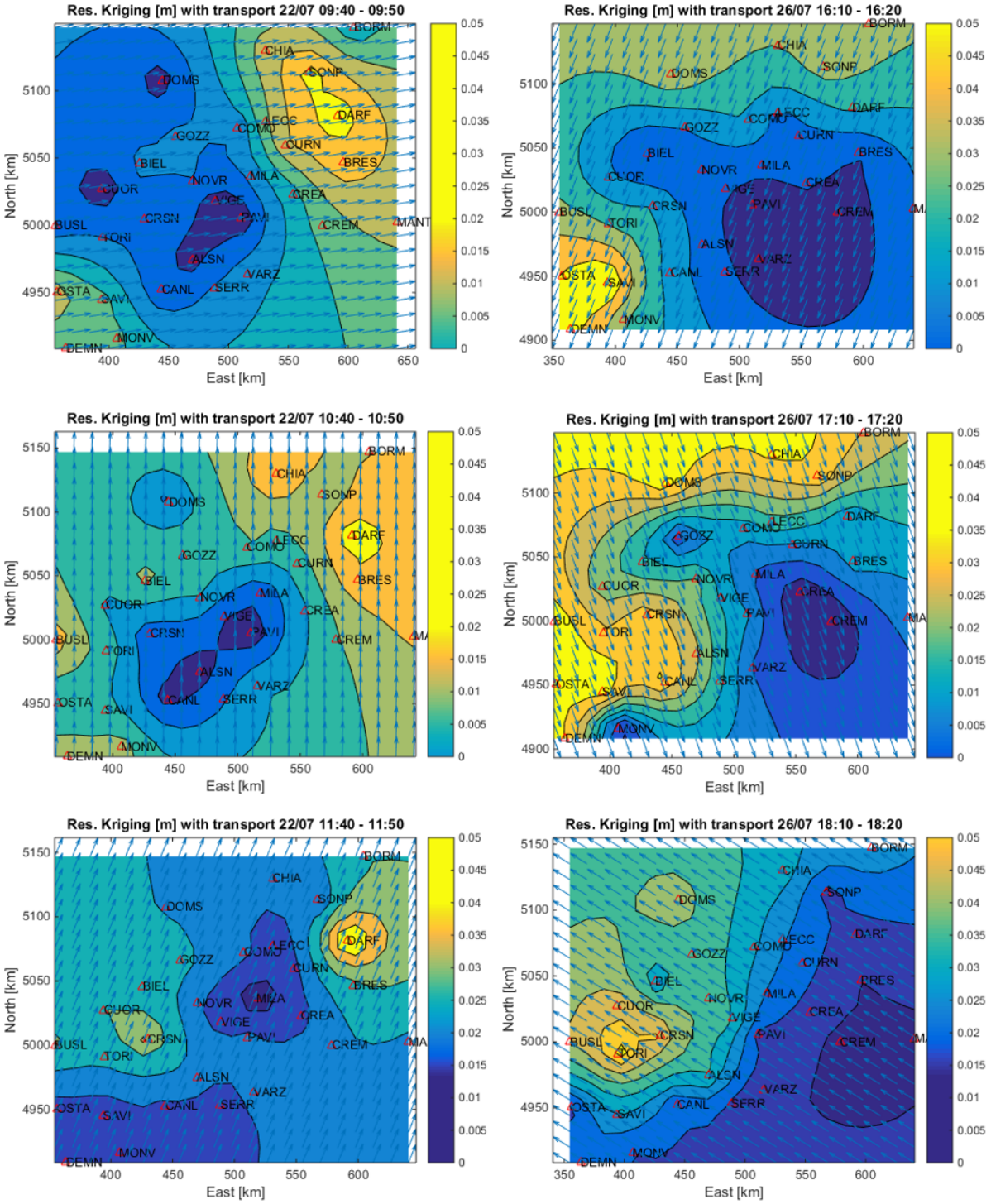

Figure 6.1 – Estimates of the mean transport vector under Taylor's frozen flow hypothesis for the events of the 22nd and 26th of July at three different epochs

6.4 General ZWD advection-diffusion model: discussion and perspectives

In the previous section we considered the transport field (U, V) as being constant throughout the definition area of the ZWD field. However, this may be true only in very few cases and thus it is highly desirable to drop this strong assumption. In this section we are going to derive a general model for the dynamics of the ZWD field that will not only describe its advective behaviour but also the turbulent mixing effects including them in the form of a diffusion term. This section is based on [**Onn et al.(2006)**].
As we have done in Chapter 5 the ZWD field and the transport vector (U, V) can be thought of as being random fields. Hence they may be decomposed into an average component and a random component as follows (cf. [**Pope(2000)**])

$$\begin{aligned} U &= \overline{U} + U' \\ V &= \overline{V} + V' \\ ZWD &= \overline{ZWD} + ZWD' \end{aligned} \tag{6.13}$$

Physically this means that the random fields (U, V) and ZWD are the result of a mean laminar component $(\bar{U}, \bar{V})$ and $Z\bar{W}D$ together with random turbulent fluctuations (U', V') and ZWD' (see again [**Pope(2000)**]). Let us also clarify that practically the mean values for the transport flow $(\bar{U}, \bar{V})$ and wet delay are averaged over an area on the ground, which we will refer to as a grid cell. The size of this grid cell corresponds to the dimensions of the base of a column containing water vapor and it is assumed that the conservation of IWV is averaged over this column.
Substituting the equations 6.13 into the ZWD advection model 6.9, one gets

$$\frac{\partial \overline{ZWD}}{\partial t} + \frac{\partial ZWD'}{\partial t} + \nabla\cdot(\overline{\mathbf{V}}\,\overline{ZWD}) + \nabla\cdot(\overline{\mathbf{V}}ZWD') + \nabla\cdot(\mathbf{V}'\overline{ZWD}) + \nabla\cdot(\mathbf{V}'ZWD') = 0 \tag{6.14}$$

where $\mathbf{V} = (U, V)$ and the differential operator $\nabla\cdot = \frac{\partial}{\partial t} + \frac{\partial}{\partial t}$.
It is not feasible to consider in detail all small-scale random fluctuations appearing in Equation 6.14 and therefore we follow the Reynolds averaging procedure in order to describe only the average state of the ZWD field. This smoothing method is commonly used in Computational Fluid Dynamics for studying the behaviour of turbulent flows and we refer to [**Pope(2000)**] and the references therein for full details about it. Here we would like only to recall that the random components of the ZWD and of the transport field have zero mean, i.e. $\overline{ZWD'} = 0$ and $\overline{\mathbf{V}'} = 0$. Thus we may deduce that

$$\overline{\frac{\partial ZWD'}{\partial t}} = 0$$

$$\overline{\frac{\partial}{\partial t}(\mathbf{V}')} = 0$$

Hence, averaging Equation 6.13 over a grid cell gives

$$\frac{\partial \overline{ZWD}}{\partial t} + \nabla \cdot (\overline{\mathbf{V} ZWD}) + \nabla \cdot \overline{(\mathbf{V}' ZWD')} = 0 \tag{6.15}$$

Equation 6.15 states that time variations of the mean ZWD, averaged on a grid cell, depends both on a laminar mean flow component, described by the advection term $\nabla \cdot (\overline{\mathbf{V} ZWD})$, and on the averaged turbulent flux component $\nabla \cdot \overline{(\mathbf{V}' ZWD')}$. The turbulent flux term appearing in the equation cannot be resolved explicitly with respect to the unknown field the equation is written for and therefore it needs to be modelled (cf. [**Pope(2000)**]). This problem is extremely important and there is a vast literature describing several models of the averaged turbulent flux component (we refer again to [**Pope(2000)**] for a complete treatment of this issue). Here following [**Onn et al.(2006)**], we adopt the so-called *gradient-diffusion* hypothesis according to which the turbulent flux term is directly proportional to the gradient of the mean ZWD field

$$\nabla \cdot \overline{(\mathbf{V}' ZWD')} = \Gamma \cdot \nabla \overline{ZWD} \tag{6.16}$$

The variable Γ is called *turbulent diffusivity* and throughout this chapter we will suppose it to be constant. Its unit of measurement is $[m^2/s]$.

Remark. The gradient-diffusion hypothesis represents the simplest model that can be adopted in order to close Equation 6.15. Its greatest value relies just on its clarity and simplicity and this is exactly the reason we focused on this model. However, it should always be borne in mind its drawbacks that make it not suitable in many cases. For a thorough appraisal of this hypothesis we refer to [**Pope(2000)**], where several alternative models are described.

Substituting the gradient-diffusion hypothesis in Equation 6.15 we eventually get the ZWD advection-diffusion model

$$\frac{\partial \overline{ZWD}}{\partial t} + \nabla \cdot (\overline{\mathbf{V} ZWD}) + \Gamma \Delta \overline{ZWD} = 0 \tag{6.17}$$

We may now assume that the ZWD field evolves according to the general model 6.17. Unlike the model 6.11 obtained applying Taylor's frozen flow hypothesis which is purely advective, Equation 6.17 accounts also for the contribution of turbulence.
Furthermore, in the advection-diffusion model above we need not anymore to assume the transport field $\mathbf{V}$ to be spatially-constant. Indeed, the solution of the inverse problem given by Equation 6.17 would yield a general time and space-varying advection field $\mathbf{V}$. Hence in this respect as well the advection-diffusion eqaution 6.17 generalizes the model 6.11.
There exists several numerical techniques that allow to integrate Equation 6.17 both as direct problem, i.e. assuming the mean laminar flow $\mathbf{V}$ to be known and solving for the unknown $\overline{ZWD}$, and as an inverse problem, i.e. where $\overline{ZWD}$ is known and the unknown is given by the mean transport field.
In [**Onn et al.(2006)**] a numerical integration based on the method of Finite Difference is explained in full details. In particular, the author solves the advection-diffusion equation both as an inverse and as a direct problem describing accurately all terms involved. He proposes a two-step algorithm, whose first step consists in

employing model 6.16 to evalute the mean transport vector $\mathbf{V}$. In this step he makes use of predicted ZWD values in any point of a grid and set the turbulent diffusion coefficient Γ to be zero. In the second stage of his algorithm the estimates $\widehat{\mathbf{V}}$ are exploited to produce 2D ZWD difference maps that take into account the spatial variation of the average ZWD due to advection and turbulent diffusion. Strikingly he gives also an estimate of the turbulent diffusion coefficient. These maps are eventually used as corrections of the turbulently-mixed atmospheric phase of a SAR interferograms (see [**Onn et al.(2006)**] for all details).
Instead of considering Finite Difference methods, we would like to suggest here the possibility of numerical integrate Equation 6.16 by means of Finite Element Methods. Compared to Finite Difference approaches, the Finite Element techniques show some advantages that are worth to be mentioned. First of all they are well suited for problems where the definition domain shows complex geometric features. Moreover, they are suitable when the theoretical solution lacks of regularity (i.e. it is not smooth or not even continuous everywhere) relying on a sound mathematical framework. Finally the Finite Element numerical solution is a *function* approximating the theoretical solution in general better than that given by Finite Difference algorithms.
Independently on the numerical method chosen to integrate Equation 6.16, the great value of this model seems to be its great simplicity compared to general Numerical Weather Prediction models. In fact, lacking of high computing facilities the advection-diffusion model described here could be implemented on a common workstation with limited resources and still capture information about the dynamics of the ZWD field. We think this is a valuable aspect that can motivate further research efforts in the analysis of the model 6.16.

Bibliography

[Baby et al.(1988)] B. H. Baby, P. Golé, J. Lavergnat (1988), "A model for the tropospheric excess path length of radio waves from surface meteorological measurements", Radio Science, 23(6), 1023-1038.

[Barindelli et al.(2018)] S. Barindelli, E. Realini, G. Venuti, A. Fermi and Andrea Gatti, "Detection of water vapor time variations associated with heavy rain in northern Italy by geodetic and low-cost GNSS receivers", Earth, Planets and Space, 70:28.

[Bennitt and Jupp(2012)] G. V. Bennitt, A. Jupp, "Operational Assimilation of GPS Zenith Total Delay Observations into the Met Office Numerical Weather Prediction Models", Monthly Weather Review, Vol. 140, No. 8, 2706-2719.

[Bevis et al.(1992)] M. Bevis, S. Businger, T.A. Herring, C. Rocken,R.A. Anthes, R.H. Ware (1992), "GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System", Journal of Geophysical Research: Atmospheres, 97(D14), 15787-15801.

[Biagi(2009)] L. Biagi (2009), "I fondamentali del GPS", Geomatics Workbooks, Volume 8.

[Brenot et al.(2013)] H. Brenot, J. Nem´eghaire, L. Delobbe, N. Clerbaux, P. De Meutter, A. Deckmyn, A. Delcloo, L. Frappez and M. Van Roozendael1, "Preliminary signs of the initiation of deep convection by GNSS", Atmos. Chem. Phys., 13, 5425–5449.

[Dach et al.(2015)] R. Dach, S. Lutz, P. Walser and P. Fridez, "BERNESE GNSS Software version 5.2", AIUB.

[De Haan(2006)] S. De Haan, "Measuring Atmospheric Stability with GPS", Journal of Applied Meteorology and Climatology, 45

[Deng et al.(2009)] Z. Deng, M. Bender, G. Dick, M. Ge, J. Wickert, M. Ramatschi, and X. Zou1, "Retrieving tropospheric delays from GPS networks densified with single frequency receivers", Geophysical Research Letters, Vol. 36, L19802.

[Elliot and Gaffen(1991)] W.P. Elliott, D.J. Gaffen (1991), "On the Utility of Radiosonde Humidity Archives for Climate Studies", Bulletin of American Meteorological Society, 72,1507-1520.

[Herring(1992)] T.A. Herring (1992), "Modeling atmospheric delays in the analysis of space geodetic data", Journal of Geophysical Research, 74, 4487-4499.

[Hoffman-Wellenhof et al.(2001)] B. Hofmann-Wellenhof, H. Lichtenegger, J. Collins (2001), "Global Positioning System. Theory and Practice", Springer-Verlag Wien. Ed.5

[Kalnay(2003)] E. Kalnay (2003), "Atmospheric modeling, data assimilation and predictability", Cambridge University Press, Cambridge.

[Kleijer(2004)] F. Kleijer (2004), "Troposphere Modeling and Filtering for Precise GPS Leveling", NUGI 816.

[Kolmogorov(1941)] A.N. Kolmogorov (1941), "The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers", Doklady Akad. Nauk SSSR, 30.

[Kouba and Heroux(2001)] J. Kouba, P. Héroux (2001), "Precise Point Positioning using IGS orbit and clock products", GPS Solutions, 5(2), 2-28.

[Kouba(2009)] J. Kouba (2009), "Testing of global pressure/temperature (GPT) model and global mapping function (GMF) in GPS analyses", J. Geodesy, 83, 199-208.

[Matheron(1971)] G. Matheron, "The theory of regionalized variables and its applications", Le cahiers du Centre de morphologie mathematique de Fontainebleau, 5.

[Metawave(2011)] D. Perissin, F. Rocca, M. Pierdicca, E. Pichelli, D. Cimini, G. Venuti and B. Rommen (2011), "The ESA METAWAVE project: Correcting for atmospheric water vapour effects in InSAR products", Geoscience and Remote Sensing Symposium (IGARSS), IEEE International.

[Moritz et al.(1967)] H. Moritz and W.A. Heiskanen (1967), "Physical Geodesy", W.H. Freeman, San Francisco.

[Niell(2000)] A.E. Niell (2000), "Improved atmospheric mapping functions for VLBI and GPS", Earth Planets Space, 52, 699–702.

[Notarpietro et al.(2011)] R. Notarpietro, M. Cucca, M. Gabella, G. Venuti, G. Perona (2011), "Tomographic reconstruction of wet and total refractivity fields from GNSS receiver networks", Advances in Space Research, 47, 898-912.

[Onn et al.(2006)] Fayaz Onn (2006), "Modelling Water vapor using GPS with application to mitigating InSAR atmospheric distorsions", a dissertation submitted to the Department of Electrical Engineering and the Committee on Graduate Studies of Stanford University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

[Papoulis et al.(2002)] A. Papoulis and S. Unnikrishna Pillai (2002), "Probability, Random Variables and Stochstic Processes", McGraw-Hill, Boston.

[Pope(2000)] S.B. Pope (2000), "Turbulent flows", Cambridge University Press, Cambridge.

[Princea et al.(1998)] S.D. Princea, S.J. Goetza, R.O. Dubayaha, K.P. Czajkowskia, M. Thawleya(1998) Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using Advanced Very High-Resolution Radiometer satellite observations: comparison with field observations. Journal of Hydrology. 212-213, 230-249

[Quarteroni(2014)] A. Quarteroni (2014), "Numerical models for Differential Problems", MS&A, Springer, Berlin Heidelberg.

[Reguzzoni et al.(2005)] M. Reguzzoni, F. Sanso´ and G. Venuti, "The theory of general kriging, with applications to the determination of a local geoid", Geophys. J. Int. (2005) 162, 303–314.

[Resch(1984)] G.M. Resch(1984). Water Vapour Radiometry, in Geodetic Applications in Geodetic Refraction, Effects of Electromagnetic Wave Propagation Through the Atmosphere, F.K. Brunner Ed., pp 53-84. Springer-Verlag, New York.

[Shevchenko(2014)] G. Shevchenko, "Fractional Brownian Motion in a nutshell", preprint (available on arXiv:1406.1956v1).

[Shoji et al.(2004)] Y. Shoji, H. Nakamura, T. Iwabuchi, K. Aonashi, H. Seko, K. Mishima, R. Ohtani (2004), "Tsukuba GPS dense net campaign observation: Improvement in GPS analysis of slant path delay by stacking one-way postfit phase residuals", Journal of the Meteorological Society of Japan, 82(1B), 301-314.

[Smith et al.(1953)] E.K. Smith and S. Weintraub (1953), "The constants in the equation for atmospheric refractive index at radio frequencies", in Proc. IRE 41, 1035–1037.

[Tatarski(1961)] V.I. Tatarski(1961), "Wave propagation in a turbulent medium", McGraw-Hill, New York.

[Teunissen and Kleusberg(1998)] P. J. G. Teunissen, A. Kleusberg Eds (1998), "GPS for Geodesy", Springer Berlin Heidelberg.

[Thayer(1974)] G.D. Thayer (1974), "An improved equation for the radio refractive index of air", Radio Science, 9, 803–807.

[Treuhaft et al.(1987)] R.N. Treuhaft and G.E. Lanyi (1987), "The effect of the dynamic wet troposphere on radio interferometric measurements", Radio Science, 22, 251–265.

[Visconti et al.(2009)] L. Biagi, F.Sansò, G. Venuti and M.G. Visconti (2009), "L'esperimento MisT: i dati e le stime degli effetti troposferici in una rete permanente GPS multiscala", 13a Conferenza Nazionale ASITA, 407-408.

[Wackernagel(2003)] H. Wackernagel (2003), "Multivariate Geostatistics, An Introduction with Applications", Springer, Berlin Heidelberg.

List of figures

Tables

Acknowledgements

First of all I would like to thank my supervisor and tutor Prof. Giovanna Venuti for her support and guide into this research topic and, especially, for having made this experience possible.

I would like also to thank heartily "Politecnico di Milano - Polo Territoriale di Como" and Prof. Maria Brovelli for the funding made available for my Ph.D. research.

Furthermore, I would like to thank my first evaluators, Prof. Nazzareno Pierdicca (Università degli Studi di Roma "La Sapienza") and Eng. Bianca Federici (Università degli Studi di Genova), for their suggestions and remarks.

A special thank goes to Eugenio Realini for fruitful discussions on various aspects of GNSS meteorology.

Come ultimo ringraziamento, vorrei esprimere la mia profonda gratitudine a Marco, Florin e Ludovico, amici con i quali ho passato serate indimenticabili tra risate e buon vino, aneddoti e battute. Abbiamo fatto un tratto di vita insieme...per questo vi ringrazio e vi abbraccio forte.