

1. Introduction

Those among us who are unwilling to expose their ideas to the hazard of refutation do not take part in the scientific game.

Sir Karl Popper (1902-1994)

This chapter provides a brief background and overview of space geodetic observations and the associated errors. The focus is on understanding and characterisation of geodetic data from the point of view of appropriately applying robust quantitative analysis strategies of tropospheric water vapour and tropospheric delay due to water vapour time series. The aim, objectives, research questions and scope of the study are also described.

1.1. Background

Measures of physical observables (which change over time) to study natural phenomena is of extraordinary importance for comprehending and characterisation of the underlying physical process. In particular, measurements from space geodetic techniques provide observational data that are used to depict the global picture of the Earth. The geodetic observations have a wide range of scientific applications such as fluctuations of the Earth (including precession, polar motion, nutation etc.) and all types of height changes and deformation due to mass transfer between the solid Earth, the atmosphere and hydrosphere. However, these measurements are influenced by many geophysical processes such as weather and climate encompassing the atmospheric structure and dynamics, mass fluctuations, large water mass circulations and sea level fluctuations, postglacial rebound and, tide of solid Earth and Oceans; are often embedded with measurement error and are often corrupted by unknown noise sources. All the physical processes and noise introduce biases to the geophysical signal structure that are embedded in the parameters derived from data. It is crucial that these processes are therefore understood in data analysis. Key auxiliary parameters associated with all the primary geodetic target parameters are the tropospheric WV or the tropospheric delay of radio signals induced by WV and meteorological parameters. Geodetic time series often exhibits irregular, nonstationary and wide-band signals due to the complex interaction between different components of the Earth system. In the majority of the applications of studying time series of geodetic data, the underlying structure is often assumed to be

stationary. It is unlikely therefore that analysis techniques developed to reliably display the spatial-temporal variability of geodetic data are robust. Applying non-linear time series techniques to geodetic observations provide new information about the complex dynamics of the atmosphere-Earth system. In order to extract and characterise the relevant geophysical signals embedded in the data, a variety of measures (especially those that are data adaptive) are welcome.

1.2. Significance of the study

The significance of this research is in the realm of atmospheric modelling for space geodetic applications. The concept of non-linear and nonstationary time series analysis in space geodesy is relatively new. Therefore, research which explores data adaptive analysis methods will contribute to the understanding of the signal structure embedded in geodetic data from the point of view of displaying the spatial-temporal features present in the data. The analysis framework reported in this thesis envisage geodetic time series in different frequency bands thereby allowing for individual modes of oscillation to be linked to the different geophysical processes (or systematic variations of technical origin) by use of a variety of different measures that characterise their dynamical characteristics.

The goal of tropospheric geodetic modelling is to mitigate the contribution of tropospheric delays to the delay observable. Estimating the deterministic and stochastic components of the tropospheric delay due to WV using data adaptive techniques contributes to a more realistic robust geodetic parameter estimates. Therefore, the findings from this research have the capacity to impact on the tropospheric modelling strategies taking cognisance of the evolutionary second-order structure and self-similar behaviour in the geodetic data. In addition, this work creates awareness of the non-linearity and nonstationarity problem, as linearity and stationarity is always assumed within current models. A further logical step, which is briefly explored in the current work, will be to incorporate a non-linear component as an estimate and additional tropospheric correction, for instance to the range component of Satellite SLR. This should improve the observed minus computed residuals (as demonstrated herein) and allow for improved precise orbit determination and consequently, an improvement in the estimates of other modelled variables. This is true in general for all the space geodetic techniques, therefore this work paves the way forward for the development and inclusion into analysis software, of advanced

models which can be realised in software algorithms, thereby contributing to global geodetic science and atmospheric/geodetic water vapour modelling.

1.3. Aim and objectives

The overall aim of this study is to apply pragmatic analyses strategies to quantitatively investigate and understand the variability of the geodetic tropospheric WV over southern Africa and other parameters associated to the geodetic delay observable. In order to achieve this goal, the following specific objectives of the research are upheld;

- a. To investigate the nature of geodetic WV time series (e.g., linear trends, periodic and non-periodic transients) and determine the appropriate model that describes the variability pattern.
- b. To characterise the multiscale behaviour of tropospheric WV in southern Africa.
- c. To detect and characterise the non-linear and nonstationary properties of tropospheric geodetic WV and understand the associated geophysical causes.
- d. Incorporate non-linearity in atmospheric range correction and therefore evaluate its contribution to the Observed minus Computed (O-C) residuals in geodetic (e.g., SLR) analysis.

In this work, when the term *geodetic* WV is used, it refers to the contribution of atmospheric water vapour to the geodetic delay observable and does not refer to water vapour as a meteorological variable. However when the term water vapour (WV) is used, it refers to meteorological water vapour, which can be determined by meteorological instruments or can be inferred from the geodetic delay models, i.e., by separating the true delay from the total delay, so that the difference is in effect the water vapour contribution.

1.4. Research questions

There have been much theoretical and modelling work done on the methods and validation of geodetic WV derived from different data sources. Statistical methods have been used to analyse the spatial-temporal structure of long-term variations of WV. Additionally, a number of different studies to measure trends in WV have been reported by Nilsson and Elgered, (2008) and references therein. Despite these efforts, this research sought to address the following key questions;

- a) What is the appropriate model that could describe a stochastic geodetic WV?
- b) What is the scaling behaviour of geodetic WV data records?
- c) Are the fluctuations of geodetic WV driven by stationary or nonstationary non-linear tropospheric processes?
- d) If non-linearity is inferred, what would be the contribution for incorporating non-linear models to the correction of atmospheric range bias, in geodetic data analysis?

It is vital for these questions to be answered to enable the space geodesy research community to fine-tune the existing tropospheric models so as to accommodate the dynamic characteristics and contributions of WV to the geodetic delay observable estimation. Our interest in this research reported herein lies with the usage of time series analysis and the description of the general time series characteristic of WV based on an automatic data-driven model selection process.

1.5. Hypothesis

This study hypothesizes that tropospheric WV and other geodetic parameters exhibit non-linear and nonstationary properties with memory.

1.6. Scope of the study

The only International VLBI Service (IVS) station in Africa (Hartebeesthoek Radio Astronomy Observatory, HartRAO) is located at 25.89° S; 27.090° E, South Africa. This is a fiducial geodetic station that has been contributing to the space geodesy community geodetic observations for nearly three decades. The current research covers quantitative analyses of stochastic, non-linear and nonstationary properties of geodetic WV and associated parameter tropospheric delay due to WV (e.g., tropospheric delay and tropospheric delay linear gradients) over and around HartRAO and WV derived from NWP model simulations. In the analysis, the non-parametric time series analysis methodologies such as Detrended Fluctuation Analysis (DFA), Wavelet Transform (WT) and EEMD are used. In order to evaluate the theoretical concepts developed throughout this work within the context of space geodetic techniques, an application was developed for SLR. This approach utilised an enhancement of existing software (Combrinck, 2010) to include non-linear effects not provided for in the current modelling of the additional range delay due to the atmosphere.

2. Tropospheric modelling for space geodetic applications

*... by observing the most distant objects of the Universe
(Quasars), we can learn things that happen around the corner,
here on Earth.*
Miguel T. (1999).

A review of the literature on geodetic tropospheric modelling is presented in this Chapter. In particular, the existing literature focussing on the analyses strategies for the derived tropospheric parameters from geodetic measurements and recent developments in tropospheric TD/WV modelling is reviewed.

2.1. Introduction

Over the last three decades, Space Geodetic Techniques (SGT) such as VLBI, GNSS, SLR, Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) and Water Vapour Radiometry (WVR) have continued to improve the accuracy, timely and continuous provision of geodetic products, which have benefits to both science and society (Niell, 2005). Some of the unique products from SGT, such as Earth Orientation Parameters (EOP), Earth gravity fields, sea surface and sea level changes, solid surface geometry and kinematics have been provided reliably and consistently. These products form the basis for scientific application in areas such as geodynamics, global change, land management, engineering and navigation. In addition, these products have been used in surveying, global spatial data infrastructure as well as in rural and urban development. Geodetic data have applications in monitoring surface deformation (e.g., GPS, VLBI and SLR), navigation (e.g., GPS), natural hazards like volcanic eruptions and stress levels for earthquake hazard assessment (see for example Nyst *et al.*, 2006). A summary of geodetic parameters related to the system Earth reported by Rothacher, (2002) is tabulated in Table 2.1.

Table 2.1. Geodetic parameter groups related to system Earth as reported in Rothacher, (2002).

Parameter type		VLBI	GPS	SLR	DORIS
RS	Quasar positions	x			
	Orbits (Satellites, Moon)		x	x	x
EOP	Nutation	x	x		
	Nutation rates	x	x	x	x
	UT1-UTC	x			
	LOD	x	x	x	x
	Polar motion	x	x	x	x
TRF	Station positions	x	x	x	x
Gravity field	Geocenter		x	x	x
	Low degree		x	x	x
Atmosphere	Troposphere	x	x		x
	Ionosphere	x	x		x

The scientific application of geodetic data has been fundamental in the understanding of the structure and deformations of the Earth's crust (Sansò and Gil, 2006), mantle and core. Geodetic data have also been applied in the analysis of geophysical fluids (Wolfgang and Richter, 2008), the coupling between the free atmosphere and solid earth. For instance, Davies *et al.*, (2004), in an analysis of GRACE and GPS data, identified a climate-driven deformation signal of the solid earth. Seasonal variations in gravity fields were positively correlated with climate-driven fluxes of surface water.

Current VLBI community under the auspices of the IVS working groups have redefined the target accuracy goals in their products. For instance, the next generation VLBI system (termed VLBI2010) ought to provide a 1 mm positional and 1 mm per annum velocity accuracies within the ITRF. The VLBI network is depicted in Figure 2.1 where a number of these stations have collocated GNSS, SLR or DORIS instruments, which have applications in

geodesy and atmospheric studies. Continuous measurement of EOP and the rapid generation and distribution of geodetic products (Niell, 2005) are additional targets. Since the atmospheric structure and dynamics influences the accuracy of the estimated geodetic products, these applications require that atmospheric biases in the geodetic products ought to be minimised.



Figure 2.1 The current global VLBI network.

Research that focuses on geodetic tropospheric modelling involves the development of strategies for improving the atmospheric models that reduce random and systematic bias components of the Tropospheric Delay (TD) observable. The TD observable is the extra path length covered by the radio wave due to a delay caused by the change in the direction of propagation of the radio signal in the troposphere (ionosphere, see Ho *et al.*, (1997); Hobiger *et al.*, (2006)) and troposphere (Saastamoinen, 1972; Haase *et al.*, 2003).

Modelling and analyses strategies of the tropospheric contribution to the delay observable ought to be improved if the goal of one millimetre accuracy in the estimation of

EOP, station positions and velocities is to be achieved. To measure the improvement in modelling accuracy, the secular signals associated with crustal deformations are monitored regularly. Other non-secular geophysical deformations (resulting from atmospheric loading, groundwater level seasonal fluctuations etc.) may be seen in the repeatability of baseline length between two stations.

The TD component is the most dominant bias contributor in GNSS and VLBI geodetic measurements (Gradinarsky *et al.*, 2000; Behrend *et al.*, 2008). As the signal traverses the neutral atmosphere, it experiences a change in the effective path length and direction of propagation. The change in the geometric path of the microwave signal occurs due to spatial-temporal variability of the refractive index. One major contributor to variability of the refractive index is tropospheric WV. In the troposphere, WV induces the highly fluctuating tropospheric delay observable, commonly referred to as the wet delay.

Current tropospheric modelling strategies do not accurately capture actual stochastic and non-homogeneous properties of geodetic WV or TD due to WV. Although a lot of effort has been made to improve the modelling and analysis of the atmospheric contribution to the geodetic TD observable, the envisioned accuracy constraints, as reported by for instance Niell, (2005), require better and in addition regional (rather than just global) modelling techniques that could correct for the random and systematic components of the geodetic TD observable biases (e.g., troposphere and instrumental). There is need for improved geodetic site specific model analysis strategies of the anisotropic atmosphere and error budgets.

The VLBI2010 project in particular focuses on addressing key constraints, which undermine the attempts to meet mm accuracy geodetic station positions through simulations (Wresnik *et al.*, 2008). Research in appropriate geometry, sky coverage, system constraints, tropospheric modelling and software have been reported by Niell, (2005); Boehm *et al.*, (2006); Nilsson and Haas, (2008). Notwithstanding all these efforts, the current modelling strategies do not consider possible stand-alone station dependent strategies that might constrain (to some extent) the empirical and systematic biases in geodetic parameters. The next section describes measurements and empirical formulation of tropospheric parameters (e.g., delay gradients, WV) that influence the accuracy of the geodetic TD observable.

2.2. Structure of the atmosphere

Different conventions or physical parameters are used to describe atmosphere layers. In space geodesy, the atmosphere would be classified into troposphere; mesosphere, stratosphere and

ionosphere (see Table 2.2). The troposphere is the neutral part of the atmosphere which includes the troposphere, tropopause and stratosphere as defined in meteorology. The troposphere has a quasi-frequency independent refractive index at microwave signals below 30 GHz. The neutral layer (0 to 15 km height) has a high concentration of neutral particles. This part of the atmosphere is dominated by physical processes which characterise the dynamics of the atmosphere determined by turbulent molecular viscosity. Due to spatial-temporal variability of temperature and pressure, mass re-distribution in the atmosphere causes local and global circulation systems which are affected by Earth rotation, large water bodies and other extra-terrestrial forces. This coupling is often captured in the correlation of EOP and the Atmospheric Angular Momentum (AAM), see for example Zhang *et al.*, (2003) and references therein.

Table 2.2. Geodetic view of the atmosphere

Height, km	0 to 50	≥ 50	≤ 80	≥ 80
Temperature	Troposphere Stratosphere	Stratopause Mesosphere		Thermosphere Exosphere
Refractivity	Troposphere			
Mixing ratio			Homosphere	Heterosphere
Magneto-electronic structure	Neutral atmosphere	Ionosphere		Ionosphere Plasmasphere Magnetosphere

The troposphere exhibits turbulent processes and therefore the mixing ratio of its constituent's undergoes subtle changes, except for WV which experiences a marked change. However, at high altitudes, > 80 km, the kinetic molecular processes dominate the turbulence processes. Below 80 km, the atmosphere is called the homosphere. The heterosphere is part of the atmosphere medium above 80 km. Furthermore, based on the dynamic processes, the troposphere can be divided into the atmospheric boundary layer (layer adjacent to the Earth's surface with characteristic high Reynolds number, implying high turbulent flow) and the free troposphere.

The ionosphere layer consists of mainly ionised atmospheric constituents. In this layer, the Ultra Violet (UV) and the X-ray radiation produce non-vanishing ionisation density. The ionosphere is conventionally characterised by the Total Electron Content (TEC) and scale height (H). These parameters exhibit spatial-temporal variability caused by

underlying processes. The speed and attenuation of microwave signals traversing the ionosphere and troposphere are affected by the refractive index of the atmosphere. In the neutral atmosphere, the optical and radio frequencies are related to the $n(\mathbf{r}, t)$ using the Cauchy Equation given in Equation (1) as reported by Griffiths, 1999):

$$n(\mathbf{r}, t) = 1 + N_a \left(1 + \frac{B}{\lambda^2} \right); \quad (1)$$

where N_a and B are the refractivity and dispersion coefficient respectively. Following Riepl and Schluter, (2000), $B = 1.7 \times 10^{14} \text{ m}^2$ for the atmosphere. Note that, at optical and radio wavelengths, $\lambda^2 \sim B$ and $B < \lambda$ or $B \ll \lambda^2$. As a result, the signal propagation in the atmosphere will experience a time delay relative to its propagation in vacuum due to $n < 1$.

In Space Geodetic measurements, the microwave signals experience bending due to the gradients, see Equation (2), of the troposphere (Davis, 1992; Davis *et al.*, 1993) and this is commonly called the geometric delay. Further, the microwave signals are slowed down and therefore causes an excess delay which is a function of the refractivity given in Equation (2);

$$n(\{\mathbf{r}, z\}, t) = n_0(z) + \vec{\xi}(z) \times \mathbf{r} + \dots; \quad (2)$$

where $n_0(z) = n(r=0, z)$ is the horizontal invariant component of the refractive index and $\vec{\xi}(z) = \nabla_r n(\{\mathbf{r}, z\}, r=0, t)$ is the horizontal gradient. According to Rocken *et al.*, (1993), the summation of the geometric and excess delay could be expressed as shown in Equation (3);

$$ZTD = \int_l n(\{\mathbf{r}, z\}, t) dr - \tau_G; \quad (3)$$

where $r \in l$ and ZTD , l and τ_G are the total tropospheric delay in the zenith direction, the curved ray path and the straight-geometric delay through the atmosphere. Equation (3) can also be expressed in terms of the slowing and bending terms in Equation (4),

$$ZTD = \int_l [n(\{\mathbf{r}, z\}, t)] dr - [S - \tau_G]; \quad (4)$$

where S is the curved path length along L . The slowing and bending components are given by first and second terms of the right hand side of Equation (4). The bending term ($[S - \tau_G] \lesssim 10 \text{ mm}$) is elevation dependent and vanishes if the ray path is in the zenith (in the absence of gradients).

The temperature and density of the tropospheric constituents affect the geodetic ZTD observable, which in turn determines its spatial-temporal distribution. The stochastic behaviour of ZTD is still one of the limiting factors that restrict the accuracy of space

geodetic techniques as reported by Boehm *et al.*, (2006). Since ZTD is caused by the dry and wet part of the atmospheric constituents, the delay could be separated into the hydrostatic (referred to as the dry part in the literature as it is caused by the refractivity of the dry gases in the troposphere, but in fact contains the non-dipole component of water vapour refractivity) and wet (non-hydrostatic) components:

$$ZTD(\varepsilon) = Y_h(\varepsilon)TD_h^z + Y_w(\varepsilon)TD_w^z; \quad (5)$$

where $Y_{w,h}$ are the elevation (ε) dependent (and azimuth independent) mapping functions which are determined independently from numerical weather prediction models, see Boehm *et al.*, (2006) or radiosonde data (Niell, 2000; Niell, 2001). The predictable ZTD are often estimated from empirical models using unbiased surface meteorological data to mm accuracy (Janes *et al.*, 1991). The wet delay component TD_w^z is related to the quantity of WV in the atmosphere and is often used as a passive tracer for atmospheric studies. For instance it can be used to reveal the structure of the atmosphere at all temporal and spatial scales. In the next section, we focus on the characteristics of the troposphere using refractivity, $(N(\mathbf{r}, z), t)$ and the relationship between N , TD and tropospheric precipitable WV.

2.3. Refractivity, tropospheric delay and precipitable WV

The refractive index, $n(\{\mathbf{r}, z\}, t)$ of the troposphere is often expressed in terms of $N(\{\mathbf{r}, z\}, t)$ i.e. $N(\{\mathbf{r}, z\}, t) \equiv 10^6 \times n(\{\mathbf{r}, z\}, t)$. Expressed as a function of temperature, T , the partial pressure of dry air, p_d and the partial pressure of WV, p_v (e.g., Smith and Weintraub 1953; Thayer, 1974; Bevis *et al.*, 1994), the compact form of $N(\{\mathbf{r}, z\}, t)$ is given in Equation form as,

$$N = \frac{k_1 \cdot Z_d^{-1} \cdot p_d}{T} + \frac{k_2 \cdot Z_v^{-1} \cdot p_v}{T} + \frac{k_3 \cdot Z_v^{-1} \cdot p_v}{T^2} + \frac{k_4 \cdot p_c}{T^2}; \quad (6)$$

where the symbols denoting the spatial-temporal dependence have been dropped for simplicity. The corrections from the departures of moist air from the ideal gas are given by the inverse compressibility factors, Z_d^{-1} and Z_v^{-1} for air and water respectively. The coefficients $k_i, \forall i=1,2,3$ and their associated errors are given in Table 2.3 according to Bevis *et al.*, (1994). If temperatures T and t are measured in Kelvin and Celsius and using James, (1967), the compressibility factors can be expressed as:

$$Z_d^{-1} = 1 + p_d \left\{ 57.97 \cdot 10^{-8} \cdot \left(\frac{T + 0.52}{T} \right) - 94611 \cdot 10^{-4} \cdot \frac{t}{T^2} \right\},$$

$$Z_v^{-1} = 1 + \frac{1650 \cdot p_v}{T^3} \left\{ 1 - 0.01317 \cdot t + 1.75 \cdot 10^{-4} \cdot t^2 + 1.44 \cdot 10^{-6} \cdot t^3 \right\}. \quad (7)$$

Table 2.3. Nominal atmospheric refractivity constants and their standard errors.

Refractivity constant	k_1 [Kmb^{-1}]	k_2 [Kmb^{-1}]	k_3 [K^2mb^{-1}] $\times 10^5$
Smith & Weintraub, (1953)	77.607 \pm 0.013	71.60 \pm 8.500	3.7470 \pm 0.0310
Thayer, (1974)	77.604 \pm 0.014	64.79 \pm 0.080	3.7760 \pm 0.0040
Bevis <i>et al.</i>, (1994)	77.600 \pm 0.005	70.40 \pm 2.200	3.7390 \pm 0.0012

From Equation (6), the dry and wet refractivity can also be obtained, which is analogous to the TD_h^z and TD_w^z derived from the geodetic TD observable. The first term on the right hand side of Equation (6) denotes dry refractivity, N_d and is often compared to the TD_h^z . The error in N_d is less than 0.5% (ITU, 2003). The second and last terms of Equation (6) are called the wet refractivity N_w . While N_d is about 0.75 N , N_w is the largest contributor to the variability of tropospheric refractivity. The zenith delay components given in Equation (5) can then be obtained by the integration of the vertical profiles of $N_{w,d}$ (see Equations (8) and (9)) which could be computed from radiosonde and numerical weather prediction simulation data sets.

$$TD_h^z = 10^{-6} \times \int_r N_d \mathbf{dr},$$

$$= 10^{-6} \times \int_r k_1 R_d \rho \mathbf{dr}; \quad (8)$$

$$TD_w^z = 10^{-6} \times \int_r N_w \mathbf{dr},$$

$$= 10^{-6} \times \int_r \left(\frac{k_2 p_v Z_v^{-1}}{T} + \frac{k_3 p_v Z_v^{-1}}{T^2} \right); \quad (9)$$

where $\mathbf{r} = [0 \ 0 \ z]$ and $M_d k_2' = M_d k_2 - k_1 M_w = 22.1 \pm 2.2 \text{ K mb}^{-1}$ (Bevis *et al.*, 1994). M_w and M_d are the molar weight of wet and dry air respectively. Here, R_d and ρ are the specific gas constant and the mass density of dry air respectively.

Using the hydrostatic equilibrium assumption reported by Houghton, (1986); $dp = -g(z) \times \rho(z) dz$, (here, z denotes the vertical coordinate) and a standard atmosphere profile, TD_h^z can be re-written as:

$$TD_h^z = \frac{2.2768 \times 10^{-3} \pm 5 \times 10^{-7} \times p_s}{1 - 0.00266 \times \cos 2\lambda - 0.00028H}, \quad (10)$$

where λ and H is the latitude and height, in km over the geoid respectively. The bias in TD_h^z due to the surface pressure, p_s hPa, λ and H can be derived using Equation (10) as follows; the delay, surface pressure, latitude and height biases are defined as $\sigma_\tau, \sigma_p, \sigma_H$ respectively such that:

$$\begin{aligned} \sigma_\tau &\propto \sigma_p, \\ \sigma_\tau &\propto \sigma_\lambda, \\ \sigma_\tau &\propto \sigma_H. \end{aligned}$$

Tropospheric integrated WV (hereafter *IWV*) kg/m^2 of the column above height z_0 can be derived from TD_w^z using mean temperature, T_m of WV in the atmospheric column using Equations (11) and (12) respectively, see for instance Askne and Nordius, (1987) and Bevis *et al.*, (1992).¹

$$T_m = \frac{\int_{z_0}^{z_T} \rho dz}{\int_{z_0}^{z_T} \rho dz / T}, \quad (11)$$

$$IWV = \frac{10^6 TD_w^z}{R_v \left\{ -k_1 \epsilon + k_2 + \frac{k_3}{T_m} \right\}}. \quad (12)$$

The *IWV* derived from pressure and mean tropospheric temperature measurements suffer from measurement errors.² One way to eliminate propagating these errors is to use TD_w^z in the data assimilation for numerical weather prediction. In geodetic applications, *IWV* is only a

¹ Radiosonde profiles have been used by Bevis *et al.*, (1992) and Emardson and Derks, (1998), to derive linear and quadratic relation between surface temperature T_s and T_m

² Using Equation (12), $ZWD \propto IWV$ and therefore $\sigma_{p_s} \sim 1 \text{ mbar} \sim 0.4 \text{ mm IWV}$.

derived quantity while TD_w^z is more directly used to improve on the accuracy of the delay observable.³

2.4. Overview of geodetic VLBI, GNSS and SLR

In the geodetic VLBI technique, extragalactic radio signals (at 2.3 and 8.36 GHz) from Quasars are received simultaneously by a network of terrestrial telescopes, from which very accurate estimates of distances between telescopes and estimates of Earth rotation can be made. Radio signals from Quasars or Quasi-Stellar Objects (QSO) are highly variable and do not suffer from parallax or proper motion since the radio sources are very distant. In fact, they are the furthest known objects. In astronomy, the VLBI measurement principle is based on interferometry where pairs of telescopes from various locations observe/receive and combine simultaneously the signals *in-phase* from the same object to reveal detail structure of the object. For the purposes of geodesy, sources are selected that exhibit little or no structure (even with the high resolution possible with VLBI) so as to use the quasars as point sources, effectively fixed points realising the International Celestial Reference Frame (ICRF). Simultaneous combination of geodetic VLBI signals in real time is not possible; instead, the signals measured at each station are recorded onto some storage media (magnetic tapes, disks) together with a very accurate time-keeping signal from a high-precision atomic clock. The recorded observations and the time signal are later retrieved at various correlators stations at selected data centers.

The time shift in the arrival time of the signal (hereafter the delay) at the relative positions of the VLBI telescopes is used to determine, with high accuracy, the positions and velocities of each station relative to each other and the positions of the radio sources. These measurements have application in the determination and maintenance of the International Terrestrial Reference Frame (ITRF) and ICRF, EOP and atmospheric parameters. In addition, during further processing of the time history of baselines between stations, three dimensional motions of the stations are obtained. Horizontal motion carries information regarding plate tectonics and, in addition, the vertical component could contain information related to geophysical processes such as post-glacial rebound.

Unfortunately, geodetic VLBI measurements contain complex signals that could be translated as *biases*⁴. Some of these errors are as a result of non-synchronised measurements,

³ For normal atmospheric conditions, it is expected that $10 \text{ kg/m}^2 \sim 65 \text{ mm}$ of the zenith tropospheric delay.

hardware constraints like:-(i) changes in the hardware or (ii) mechanical response of the hardware to weather; e.g. thermal expansion (Wresnik *et al.*, 2005), wind, precipitation and atmospheric delay due to variations in the refractivity (Haase *et al.*, 2003; Boehm *et al.*, 2006) and geophysical factors such as ocean loading (Ray, 1999; Scherneck, 1991), crustal motion (Haas *et al.*, 2000; Haas *et al.*, 2003) and atmospheric pressure loading (Petrov and Boy, 2004). To determine geodetic VLBI parameters with very high accuracy, all these signals ought to be modelled and accounted for during processing of the data.

In general, the working principle of GNSS (e.g. GPS, GLONASS, GALILEO and the planned COMPASS system) is similar to VLBI (e.g. radio waves traversing through the ionosphere). However, the major difference is that GNSS uses artificial satellites, while the VLBI uses distant radio sources. In VLBI the main observable is the phase delay, whereas in GNSS it is the range. The American GPS navigation system consists of a constellation of more than 30 satellites orbiting at approximately 20 000 km at 55° inclination. This geometry allows simultaneous visibility of at least four satellites by any receiver on the surface of the Earth (Hofmann-Wellenhoff *et al.* 1997). The Russian GLONASS currently has about 20 satellites in its constellation. The GLONASS satellites orbit at a 64.8° inclination and at an altitude of about 19 000 km.

Radio signals (at different frequencies) from transmitters onboard GNSS satellites are continuously transmitted and received by GNSS receivers. GPS uses different codes; CDMA⁵ to separate signals from different satellites, while GLONASS satellites transmit their radio signals at different frequencies, the FDMA⁶. The GPS satellites transmit right-handed polarized radio signals at two frequencies; $L_1 \sim 1.57542$ GHz and $L_2 \sim 1.2276$ GHz, that are modulated with two Pseudorandom noise, PRN codes (C/A and P(Y)) and with a navigation message. GLONASS transmit right-hand circular polarized radio signals with band specifications of $1.6025625 < L_1 < 1.6155$ and $1.240 < L_2 < 1.260$ GHz as well as the C/A and P carrier band widths of 1.0 and 10.0 MHz respectively.

The travel time or distance between the receiver and the satellite is calculated by comparing the time the signal was transmitted to the time the signal was received (based on

⁴ The description of errors in geodetic analysis should be treated with caution. Most of the signals that were initially treated as biases have recently found significant scientific applications. In geodetic VLBI, for astrometric and geodetic applications, the delays other than the geometric delay are regarded as biases.

⁵ Code Division Multiple Access

⁶ Frequency Division Multiple Access

the receiver clock). The Pseudo distances calculated in this way have ~ 10 m accuracy. Using the carrier phases (measured with an order of accuracy of millimetres), centimetre level of position accuracy could be obtained but for the inherent integer ambiguity (Hofmann-Wellenhoff *et al.*, 2001). Just like the measurement errors in VLBI, GNSS error sources such as clock, orbital, receiver noise, multipath, ionospheric and tropospheric errors have to be accounted for to obtain an optimal delay observable. The ionospheric delay is often removed based on the concept of double-differencing⁷.

2.4.1. GPS delay observable

In GPS observations, measurements are often carried out using the pseudo-range (or code range) and carrier phase. The primary measurement is the phase measurement, which has applications for high precision positioning. Pseudo-ranges are only accurate to a meter and are therefore considered as ancillary observations to be used for eliminating synchronization, clock, integer ambiguity and cycle slip biases. The basic form of carrier phase observation (also zero difference observation, ρ_z) is given by Equation (13)

$$\rho_z = \rho_0 - \lambda N - \tau_{ion} + \tau_{trop} + \tau_{cl} + \tau_{hw} + \tau_{syn} + \tau_{or} + \tau_{apc} + \tau_{rel} + \zeta, \quad (13)$$

where ρ_0 , λ and N are the geometrical distance from the satellite to the receiver, wavelength of the carrier signal and the ambiguity integer. The ionospheric delay is τ_{ion} and the tropospheric delay is τ_{trop} , with τ_{cl} , τ_{hw} , τ_{syn} , τ_{or} and τ_{apc} the combined receiver and transmitter clock biases, the hardware bias of the receiver, synchronisation error, the receiver/transmitter antenna orientation error and the antenna phase centre offset respectively. A relativistic contribution is denoted by τ_{rel} while ζ describes residual errors.

The satellite transmitter and receiver clock errors in Equation (13) require double-differencing, see Equation (15), for their elimination (Alber *et al.*, 2000). The satellite clock error is often eliminated using single-differences, which are formed by differencing the simultaneous one-way measurements from the satellite to two ground receivers given by Equation (14). Furthermore, ionospheric and tropospheric delays are also eliminated (or at

⁷ In double-differencing, single differences are first formed by subtracting observation equations from two separate receivers to a single satellite. Taking the difference between these two single differences for a specific receiver pair gives the carrier phase double difference (Alber *et al.*, 2000).

least reduced) if the receivers are closely spaced. In double-differencing, the difference between two single-differences is computed and this eliminates the two receiver clock biases.

$$\begin{aligned}\Delta\tau_{r_1,r_2}^1 &= \tau_{r_2}^1 - \tau_{r_1}^1, \\ \Delta\tau_{r_1,r_2}^2 &= \tau_{r_2}^2 - \tau_{r_1}^2,\end{aligned}\tag{14}$$

$$\Delta\tau_{dd} = \Delta\tau_{r_1,r_2}^2 - \Delta\tau_{r_1,r_2}^1\tag{15}$$

Here $\tau_{r_1}^1$ and $\tau_{r_1}^2$ are the observations of satellites 1 and 2 by receiver r_1 and $\tau_{r_2}^1$, and $\tau_{r_2}^2$ the observations of satellites 1 and 2 by receiver r_2 , $\Delta\tau_{dd}$ is the double-difference delay observation component. Optimal correction of receiver clock errors is only possible, firstly, if the measurements are taken at the same time or a priori knowledge of antenna position, satellite position, and the pseudo-range measurements are used to constrain the offset of the station clock to within a μ s.

For original carrier phases τ_{L_1} and τ_{L_2} , a combined phase measurement with combination factors, κ_1 and κ_2 is given by Equation (16),

$$\tau_{L_{1,2}} = \kappa_1\tau_{L_1} + \kappa_2\tau_{L_2}\tag{16}$$

If the carriers phases τ_{L_1} and τ_{L_2} have systematic errors $\delta\tau_{L_1}$ and $\delta\tau_{L_2}$, the combined observation will have a systematic error given by Equation (17);

$$\delta\tau_{L_{1,2}} = \kappa_1\delta\tau_{L_1} + \kappa_2\delta\tau_{L_2}\tag{17}$$

For a dispersive ionosphere, $\delta\tau_{L_2}\cdot f_2^2 = \delta\tau_{L_1}\cdot f_1^2$, for which Equation (18) could be written as

$$\delta\tau_{L_{1,2}} = \Gamma_{ion}\delta\tau_{L_1},\tag{18}$$

Where $f_2^2\cdot\Gamma_{ion} = f_2^2\cdot\kappa_1 + \kappa_2\cdot f_1^2$, quantifies the first-order ionospheric contribution in linear combination. Ionospheric contribution would therefore be reduced with the selection of optimal values of κ_1 and κ_2 ; e.g. in GPS observations, an ionospheric free measurement would be described based on $f_1^2 - f_2^2\cdot\kappa_1 = f_1^2 \sim 2.25$ and $f_1^2 - f_2^2\cdot\kappa_2 = -f_2^2 \sim -1.55$. In this case the combined noise level could be given by 2.7322σ .⁸

⁸ According to error propagation law, given the σ of independent measurements $\{x, y, z\}$, related as $r = ax + by + cz$, the combined error in r , σ_r is given by $\sigma_r = \sqrt{(a\sigma)^2 + (b\sigma)^2 + (c\sigma)^2}$ or $\sigma_r = \sigma\sqrt{a^2 + b^2 + c^2}$.

2.4.2. The geodetic VLBI delay observable

The VLBI observables (Shapiro 1976; Cannon, 1978; Thompson *et al.*, 2001), such as the phase delay, delay rate and the group delay, carry positional information of extraterrestrial radio sources and terrestrial telescopes, which have vital geodetic and astrometry applications. The phase delay, though unambiguous, is the most accurate VLBI observable. It is used to determine the quasi-inertial reference frame to sub-milliarcsecond (mas) precision using the method of VLBI differential astrometry.⁹

The physical system of the geodetic VLBI could be represented by the schematic diagram shown in Figure 2.2 (Cannon, 1978). In the Figure 2.2, vector \mathbf{r} is a unit vector, LO is the local oscillator and RF AMP is the radio frequency amplifier. As depicted in the Figure the signal (with frequency, ν) from a distance radio source arrives at antenna 2 after a time delay, τ_g with respect to antenna 1.

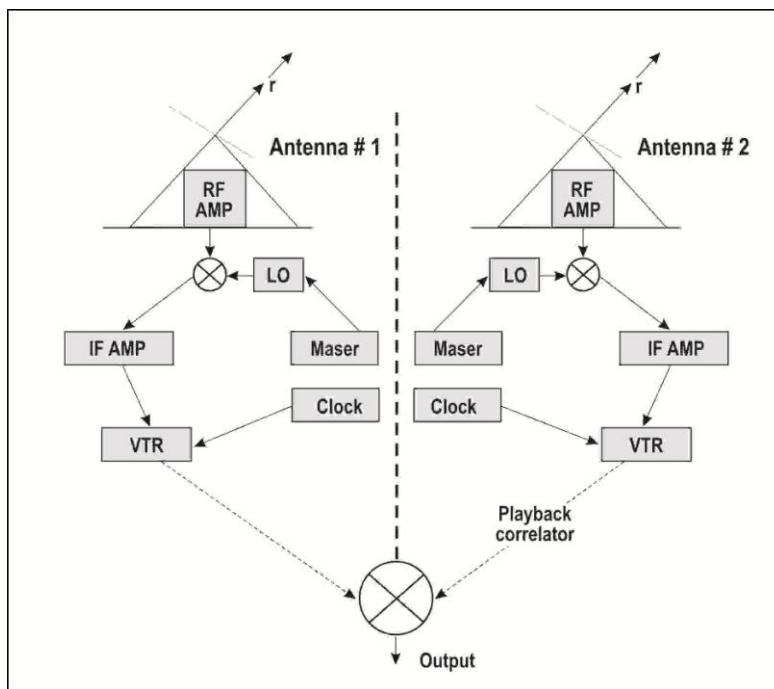


Figure 2.2. Schematic of a long baseline interferometer

⁹ In VLBI differential astrometry, the interferometric phases of two radio sources are alternately sampled once every few seconds for a given period of time. The different contributions to the phase delays are removed, thereby isolating the geometric contribution, which is finally modelled using the weighted-least squares algorithm of the geometrical parameter adjustment.

In an ideal case, contributions from the propagation medium and instrumentation are neglected and therefore the geometric delay is given by Equation (19),

$$\tau_g = \frac{\mathbf{B}}{c} \hat{\mathbf{r}} \quad (19)$$

where c , \mathbf{B} and $\hat{\mathbf{r}}$ are the speed of light, the baseline vector and direction unit vector of the radio source from the centre of the Earth. The interferometer response, $\Lambda(\nu, \mathbf{r})$ assuming accurate delay tracking with respect to angular position, \mathbf{r}_0 , is given by:

$$\Lambda(\nu, \mathbf{r}) = S_\nu e^{\left\{j2\pi\frac{\nu_0}{c}\mathbf{B}\cdot\mathbf{r}_0\right\}} \times e^{\left\{j2\pi\frac{\nu}{c}\mathbf{B}\cdot(\mathbf{r}-\mathbf{r}_0)\right\}}, \quad (20)$$

where S_ν is the total flux density of the source at the observing frequency, ν .

For extended radio sources, the interferometer response will have two components given in Equation (21); a rapidly varying and time dependent response (this is equivalent to the point source response) but centred at \mathbf{r}_0 at the observing frequency, ν_0 , and a slow-varying component called the visibility function, $V(\mathbf{B})$, which describes the amplitude and phase offsets. The visibility function is a complex function that depends on the bandwidth of the radio source, the emission content of the radio sources and the baseline geometry.

$$\Lambda(\nu, \mathbf{r}) = S_\nu e^{\left\{j2\pi\frac{\nu_0}{c}\mathbf{B}\cdot\mathbf{r}_0\right\}} V(\mathbf{B}). \quad (21)$$

However, for simplicity, Equation (20) could be interpreted as sinusoidal response patterns (also called the fringes), which are formed as a result of the continuous variation of the geometric delay caused by the diurnal motion of the point radio sources, and therefore the interferometer response function could be expressed as:

$$\Lambda(\nu, \mathbf{r}) = S_\nu \cos(2\pi\mathbf{B}_\lambda \cdot \hat{\mathbf{r}}), \quad (22)$$

where the interferometer fringe phase (due to the quasi-sinusoidal response) is defined as $\phi_g = 2\pi\nu\tau_g$.¹⁰

In general, the interferometer phase equation, taking into account the contributions from the source structure and contributions from the propagation medium (e.g. the atmosphere) as reported in Thompson *et al.* (2001), is given by:

$$\phi(t) = \frac{2\pi L}{\lambda} \left\{ \cos D \cos \delta \cos [A_0 + \omega t - \alpha] + \sin D \sin \delta \right\} + \phi^{at}(t) + \phi^{ins}(t) + \phi^{str}(t), \quad (23)$$

¹⁰ Fringe phases are the spatial-temporal patterns of the interferometer response

where $L = |B|$, A_0 and D are the right ascension and declination of the baseline vector, \mathbf{B} , respectively. The angular velocity of the Earth is denoted as ω rad/sec while α and δ are the right ascension and declination of the radio source respectively. The propagation medium, the instrumentation and source structure contributions to the fringe phase pattern are denoted by ϕ^{at} , ϕ^{ins} and ϕ^{str} respectively. From Equation(23), the geodetic parameters relevant to the thesis reported here are the \mathbf{B} , A_0 , D and ϕ^{at} .

The phase delay τ_ϕ , is the ratio of the observed fringe phase (with phase ambiguities), which is an integral number of 2π 's and the reference angular frequency is given by Equation (24).

$$\tau_\phi = \frac{\phi(t) + 2\pi n}{\omega}; \quad \omega = 2\pi\nu. \quad (24)$$

The group delay τ_g is the derivative of the fringe phase with respect to angular frequency and is given by Equation (25),

$$\begin{aligned} \tau_g d\omega &= d\phi, \\ &= \frac{1}{2\pi} \frac{d\phi}{2\nu}. \end{aligned} \quad (25)$$

The time derivative of the phase delay is given by:

$$\begin{aligned} \dot{\tau}_\phi &= \frac{d\tau_\phi}{dt}, \\ &= \frac{1}{2\pi\nu} \frac{d\phi}{dt}. \end{aligned} \quad (26)$$

The uncertainties in the VLBI observations are dependent on temperature of the antenna (T_{ant}) and of the receiver (T_s), the observing frequency and band width ($\Delta\nu$) of the radio signal. According to Thompson *et al.* (2001), the uncertainty in the phase delay σ_ϕ , group delay σ_g , and the phase delay rate, σ_{τ_ϕ} are given by Equations (27), (28) and (29) respectively.

$$\sigma_{\tau_\phi} = \frac{T_s}{2\pi T_{ant} \sqrt{2}} \left\{ \frac{N}{\nu^2 \Delta\nu \Delta t} \right\}. \quad (27)$$

$$\sigma_{\tau_g} = \frac{T_s}{2\pi T_{ant} \sqrt{2}} \left\{ \frac{1}{\Delta v_{rms}^2 \Delta v \Delta t} \right\}. \quad (28)$$

$$\sigma_{\tau_\phi} = \frac{T_s}{\pi T_{ant}} \left[\sqrt{\frac{3}{2}} \right] \left\{ \frac{1}{\sqrt{\Delta v \Delta t}} \right\}. \quad (29)$$

Here, Δt and N are the single source scan period and sampling frequency over Δt . The root mean square band width is $v_{rms} \sim \frac{\Delta v}{12v}$. Using Equations (27) and (28), the relationship between phase delay and group delay standard deviation could be written as:

$$\frac{\sigma_{\tau_\phi}}{\sigma_{\tau_g}} \sim \frac{\Delta v}{12v} \sqrt{N}. \quad (30)$$

2.4.3. The geodetic SLR delay observable

Geodetic VLBI and GNSS systems described above operate in the radio wavelengths. The SLR is the only SGT that operates in the optical region with a good ranging accuracy of 1 to 2 cm (Combrinck, 2010). Some of the applications of SLR data include measurement of small scale geodetic station position variations that arise from geophysical processes, the contribution to the development of gravity field models, and the establishment and maintenance of ITRF.

The SLR observable is the Time-of-Flight (ToF) of a laser pulse between the SLR station and target satellite that can be translated to the range to the target satellite; this is often corrected for system delay. The range to the satellite is used to derive other parameters such as EOP, station position, gravity coefficients, etc. The range Equation (31), often used in SLR processing, takes into account atmospheric effects (Δa), Centre-of-Mass (ΔCoM), correction of the satellite, station range bias (ΔR_b) and a relativistic correction (ΔGR).

$$NPR_i = \left\{ \frac{c \times NP_{TOF_i}}{10^{12}} \right\} \left\{ 2 - \Delta a_i + \Delta CoM_i - \Delta R_{b_i} - \Delta GR_i - \Delta \epsilon_i \right\}^{-1} \quad (31)$$

Here, NPR and $\Delta \epsilon$ are the normal point range or the observed range and the correction for unknown systematic and random errors respectively. The systematic errors (which are mainly from tropospheric influence) in the SLR observations are contained in Equation (31) and therefore critically influence the absolute ranging accuracy.

SLR operates at optical wavelengths where the atmosphere is dispersive. Therefore, during the processing of SLR data, a correction for the additional delay due to the atmosphere is required. Mendes and Pavlis (2004) reported closed-form expressions in Equation (32) and suitable for calculating the additional zenith SLR ranges due to troposphere;

$$\begin{aligned}\tau_{atm}^z &= \tau_h^z + \tau_{nh}^z \\ \tau_h^z &= 0.002416579 \times \frac{f_h(\lambda)}{f_s(\phi, H)} \times P_s \\ f_h(\lambda) &= \left[k_1^* \frac{k_0 + \lambda^{-2}}{k_0 + \lambda^{-2}} + k_3^* \frac{k_2 + \lambda^{-2}}{k_2 + \lambda^{-2}} \right] \times 9.9995995 \times 10^{-3} \\ f_s(\phi, H) &= 1 - 0.00266 \cos 2\phi - 0.00000028H\end{aligned}\tag{32}$$

$$\begin{aligned}\tau_{nh}^z &= 10^{-4} \frac{e_s}{f_s(\phi, H)} [5.316 f_{nh}(\lambda) - 3.759 f_h(\lambda)] \\ e_s &= 10^{-2} RH \left[6.1078 \times \left(\frac{7.5t}{237.3 + t} \right)^{10} \right] \\ f_{nh}(\lambda) &= 0.003101 (\omega_0 + 3\omega_1 \lambda^{-2} + 5\omega_2 \lambda^{-4} + 7\omega_3 \lambda^{-6})\end{aligned}\tag{33}$$

Here, P_s and $f_h(\lambda)$ are the surface barometric pressure and the hydrostatic dispersion equation respectively. The geodetic latitude and height are denoted by ϕ and H respectively while $k_0=238.0185 \mu\text{m}^{-2}$, $k_2=57.562 \mu\text{m}^{-2}$; k_1^* and k_3^* are $19990.975 \mu\text{m}^{-2}$ and $579.55174 \mu\text{m}^{-2}$ respectively. The WV pressure (e_s) at the surface is calculated from Relative Humidity (RH). Further, the dispersion component of the non-hydrostatic is denoted by $f_{nh}(\lambda)$ and $\omega_{0,1,2\&3}$ are given as 295.235 , $2.6422 \mu\text{m}^2$, $-0.032380 \mu\text{m}^4$ and $0.004028 \mu\text{m}^6$.

The atmospheric model of Mendes and Pavlis (2004) adopted by the Analysis Working Group of the International Laser Ranging Service (ILRS) is currently used in the SLR data satellite analysis software package developed at HartRAO (Combrinck, 2010). In order to demonstrate the bias contribution of the troposphere to the range measurements in SLR, the tropospheric zenith delay derived from LAsER GEODynamics Satellite-1 (LAGEOS I) satellite data is plotted as a function of elevation (left panel) and azimuth (right panel) in Figure 2.3. Figure 2.3 illustrates a band of corrections derived from laser wavelengths 423 nm, 532 nm and 846 nm and a range of ground level measurements of relative humidity (%), pressure and temperature. In this example, the elevation angle of 15° ; where the delays were more than 8 m, was set as the cut-off angle. As depicted in the right panel, the azimuth angle

only influences the geometry of the satellite orientation as determined by their orbital parameters.

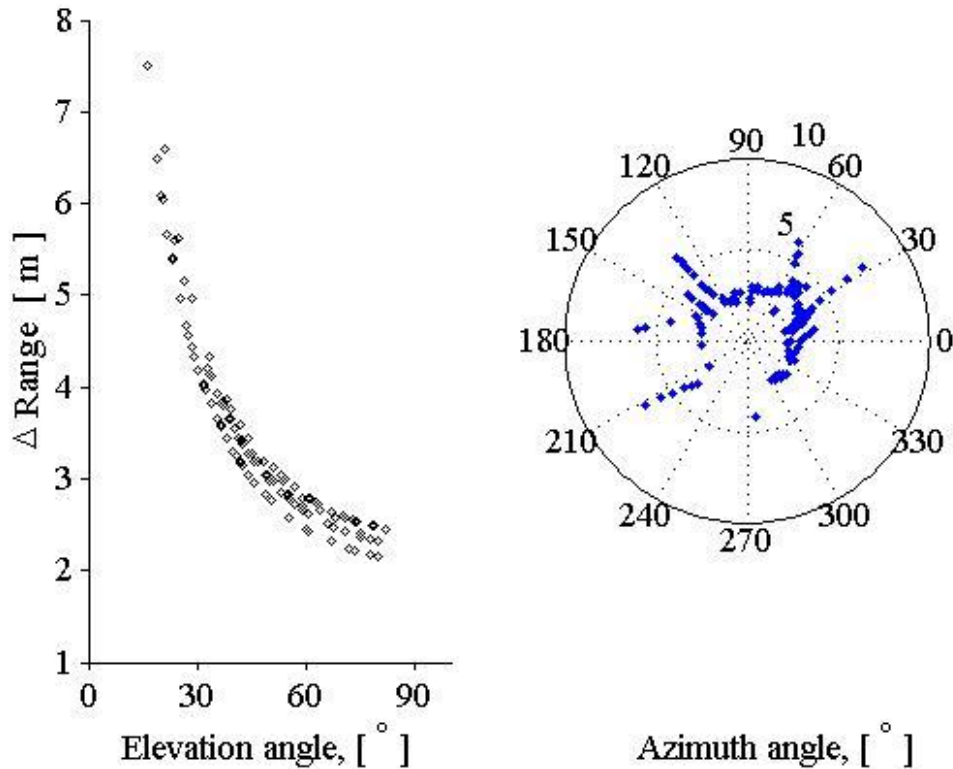


Figure 2.3. Mendes and Pavlis (2004) model of the absolute change in the SLR range measurement due to atmospheric delays

2.5. Derived tropospheric parameters in geodetic analyses

The presence of the dynamic troposphere along the lines of sight from each antenna to the radio source will affect the propagation of the radio signal due to variations of the refractive index of the traversed medium. This geodetic radio signal interaction with troposphere induces a bias in the geometric delay measurement. In addition, the rate of change of phase delay over short time spans will also be affected. Therefore, geodetic parameters such as baseline length and orientation, position of radio signal receivers and transmitters as well as clock offsets; which are estimated from the group delay and phase delay rate observables would be biased due to the fluctuating troposphere.

Tropospheric delay parameters are often estimated in most GNSS and intercontinental VLBI data reduction. In the analysis strategies employed by various geodetic analysis groups, only the spatial-temporal average-troposphere-parameter component (e.g., *ZTD*) is estimated

for each station over the measurement period. As a result, the TD_w^z fluctuations around these averages are the dominant tropospheric errors, which in turn map into biases in geodetic parameters. To improve the accuracies in the geodetic observables, the deviations from the spatial-temporal TD_w^z average ought to be identified and quantified so as to understand the character and effect of WV fluctuations on the estimated geodetic parameters.

Biases due to the modelling of tropospheric delays of microwave signals from satellites or radio sources due to the neutral atmosphere lower the accuracy of the delay observable in GPS and VLBI analyses. The parameterisations often used to compute the total delay, TD is given by Equation (34) according to Davis *et al.*, (1985).

$$ZTD(\mathcal{G}, \varepsilon) = Y_h(\varepsilon)\{TD_h^z\} + Y_w(\varepsilon)\{TD_w^z\} + Y_g(\varepsilon) \times \cot(\varepsilon) \{G_n \times \cos(\mathcal{G}) + G_e \times \sin(\mathcal{G})\} \quad (34)$$

Here Y_g is either the wet or hydrostatic gradient mapping function, G_n and G_e are the North and East gradients respectively. The tropospheric mapping functions and gradients form a set of derived parameters that influence proper modelling of the tropospheric delay. As a result, the mapping functions and gradients affect the accuracy of the geodetic parameters such as station coordinates and velocities computed from space geodetic techniques (Petrov *et al.*, 2009). In particular, biases in the station height component are directly related to tropospheric biases using a rule-of-thumb reported in Niell *et al.*, (2001) and Boehm and Schuh (2004). Both of these derived parameters aid in characterising the azimuthally symmetric component of the tropospheric delay.

Due to the Earth's surface (and therefore troposphere) asymmetry, the second-order terms of the refractivity, $\xi(z) = \nabla_r n(\{\mathbf{r}, z\}|_{r,r=0})$, given by Equation (2) often emerge. Though their contribution to the delay observable is arguably minimal, correcting them would certainly play an important role in meeting the goals of 1 mm and 0.1 mm/year accuracy of station positions and velocities respectively. These goals are described in the Global Geodetic Observing System (GGOS) of the International Association of Geodesy (IAG).

Mapping functions

Many mapping functions have been suggested (Niell 2000; Niell 2001; Boehm *et al.*, 2006) in many geodetic tropospheric modelling works. Among them, the Niell Mapping Function (NMF) is the most common mapping function used in many geodetic software packages.

Mapping functions $Y_h(\varepsilon)$ and $Y_w(\varepsilon)$ are computed by fitting the coefficients a, b and c in Equation (35) to the standard atmospheres, to in-situ radiosonde measurements or to NWP models (Niell, 2001; Boehm and Schuh, 2007).

$$Y_{h,w}(\varepsilon) = \frac{1 + \frac{a}{1 + \frac{b}{1 + c}}}{\sin(\varepsilon) + \frac{a}{\sin(\varepsilon) + \frac{b}{\sin(\varepsilon) + c}}}. \quad (35)$$

The NMF are provided at five latitude bands and are quasi-symmetric (they have a seasonal dependency with inherent half-year phase shift) with an annual signal. Tropospheric delays are usually separated into hydrostatic and wet components; computed as a product of the delay in the zenith direction and the corresponding hydrostatic and wet mapping functions. Later, mapping functions based on the NWP models, the Isobaric Mapping Functions (IMF) were developed (Niell, 2000). The hydrostatic IMF ($Y_h(\varepsilon)$) uses a height of 200 hPa pressure level while the wet IMF ($Y_w(\varepsilon)$) is based on a coarse ray-trace at 3.3° elevation through the NWP pressure levels.

Boehm and Schuh, (2007) reported on the new Vienna Mapping Function (VMF) which are dependent only on the elevation angle, ε , assuming a symmetric atmosphere around the stations. The values of the b and c coefficients of Equation (35) are obtained from the IMF and NMF for the hydrostatic and wet components respectively. Updated VMFs (hereafter VMF1) were developed based on new values of b and c coefficients of the Y_h . The coefficients c were derived from ray tracing and fitted to a function of latitude and day of year to remove systematic errors. The systematic station height improvement of up to ~10 mm (which is equivalent to 2 mm improvement in the station height) is obtained on application of VMF1 (Boehm *et al.*, 2006; Niell, 2006). Therefore, using VMF1 improves the precision of geodetic parameters. Global Mapping Functions (GMF) that are comparable to the VMF1 have also been developed based on the global ECMWF numerical weather model data sets. The coefficients b and c are obtained from the spherical harmonics expansion of the VMF1 and then mapped onto a global grid.

Tropospheric gradient

To consider the four dimensional structure of the tropospheric parameters, G_n and G_e are used (Boehm and Schuh, 2007). The gradient components G_n and G_e are mapped by $Y_g(\vartheta)$ to the slant direction of the observation. In this way, the tropospheric delay and gradient asymmetries are accounted for. Current geodetic analysis approaches model G_n and G_e using the assumption that the atmosphere is driven by stationary processes and two gradients per station which are the North-South and East-West components. These components are used to describe the tilting of the zenith as described by the mapping functions.

The gradient components G_n and G_e are estimated using two methods. These are using priori hydrostatic gradients that have no temporal dependence or by determining the time dependent gradients from a 200 hPa tilted pressure level (see for instance in Boehm and Schuh, (2007) and others therein). Notice that $G_\alpha = \{G_n, G_e\}$ could also be determined from NWP as the vertical integral of $\xi(z)$ weighted with height (hereafter linear horizontal gradients (Boehm and Schuh, 2007; Davis *et al.*, 1993), see Equation (36)). The linear horizontal gradients are computed based on the assumption that the vertical refractivity gradient $\xi(z)$ is constant over some finite distance around a geodetic station.

$$G_g = 10^{-6} \times \int_{z=0}^{\infty} \xi(z) \times z \times dz. \quad (36)$$

Figure 2.4 illustrates that the hydrostatic linear horizontal gradients (North-South and East-West) have very minimal fluctuations of > 0.5 mm. However, the North-South linear horizontal gradient contains fluctuations of ~ 1.5 mm. Similarly, the East-West linear horizontal gradient exhibits amplitude of ~ 1.0 mm.

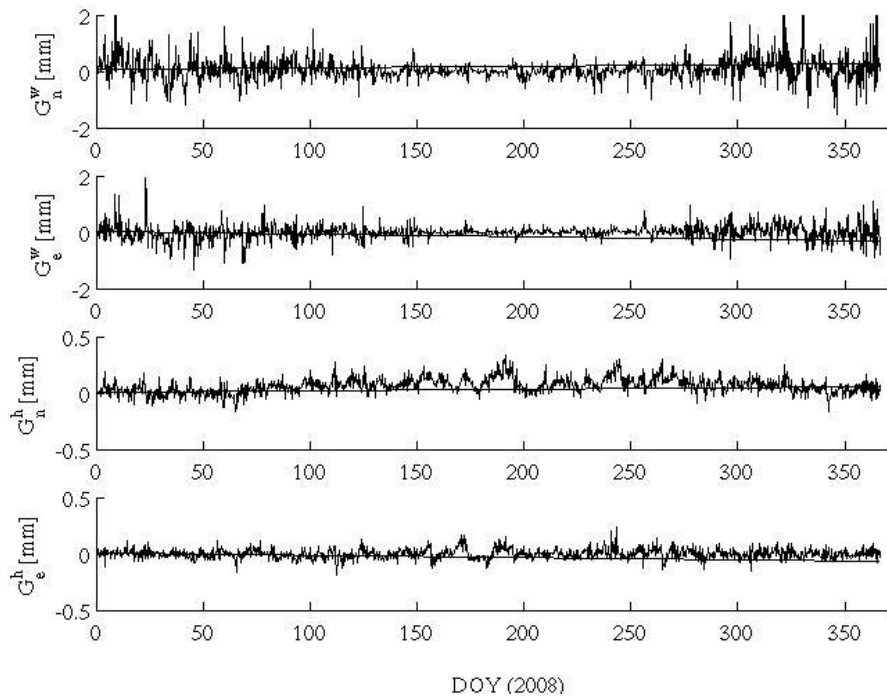


Figure 2.4. The linear horizontal gradients at HartRAO estimated from ECMWF data.

2.6. Recent developments in modelling TD and WV

Ray-tracing has been used to derive mapping functions and also assess the theoretical models used in many geodetic applications. Over a long-term period, the theoretical assessments of these models exhibit reasonable accuracy for both the TD_w^z and TD_h^z assuming accurate surface meteorological measurements. There is however larger scatter in the geodetic path delay and position estimates in the short-term atmospheric fluctuations. A number of factors could be attributed to this scatter. Firstly, though the mm-level accuracy in the delay observable is achievable with precise meteorological sensors, surface measurements are often not representative of the vertical profiles through the whole troposphere; for instance in the boundary layer, the humidity and temperature reveal strong evolution during the day and are stable in the free atmosphere (Stull, 1994). Secondly, empirical meteorological models yield better results with standard atmosphere parameters instead of with measured meteorological parameters (Wang *et al.*, 2008). Furthermore, the presence of horizontal gradients in the refractivity field induces the errors when mapping the delay observable onto the line-of-sight. To address this gap, parametric estimation and external correction strategies have been suggested.

Arguably, the meteorological models have not been able to predict the geodetic delay observable with sufficient accuracy. As a result, the parameter estimation approach has been developed which involves a post acquisition processing approach that uses geodetic measurements to constrain the parameters of a meteorological model for determining the geodetic TD observable (Tralli, 1988; Bock and Doerflinger, 2001). In this approach, the path delay is corrected for *a priori* with a meteorological model using *ZTD* and mapping functions. This methodology has inherent theoretical accuracy limitations due to clock offsets, orbit and multipath offsets (e.g., in GNSS) in the delay observable.

Deterministic parameter estimation by least squares adjustment and stochastic estimation based on Kalman filtering (Pacione and Vespe, 2003; Jin and Park, 2005) are the two common strategies used for the parameter estimation of the TD. In deterministic parameter estimation, the tropospheric delay is modelled either as a series of piecewise independent constant terms or correlated parameter which are closely related to a random model. This approach is computationally demanding in large geodetic networks or over long observing periods.

In the Kalman filtering, arbitrary values of the TD with high turn-around time can be estimated. However, this method suffers from increased multipath and tropospheric mis-modelling of the gradients and small-scale in-homogeneities due to low elevation angle observations. Nevertheless, parameter estimation has improved the estimation of the TD observable; e.g., a few parts in $\sim 10^8$; improvement in the VLBI baseline repeatabilities (Tralli, 1988). Recently, improvement in parameter estimation is based on the turbulent atmosphere model (Nilsson and Haas, 2008). In this model, the *ZTD* is simulated through a turbulent atmosphere. The delays are simulated to vary both as a function of direction of observation and time. This approach yields more realistic delays as compared to those simulated from random walk processes.

For high accuracy geodetic positioning, external correction strategies are used. External correction strategies rely on the use of independent techniques for measurements of the TD_h^z and TD_w^z (Bock and Doerflinger, 2001). In this strategy, the wet path delays are retrieved by remote sensing the troposphere, in order to correct the geodetic delay observable *a priori*. The TD_h^z is modelled by empirical meteorological models and evaluated from either surface meteorological measurements or standard atmosphere data. Thereafter, this

component is mapped onto the path of the geodetic radio signal using a mapping function. The wet path delay is derived from integrated water vapour content that is remotely sensed. The WV in the zenith direction is expressed as either the *IWV* or Precipitable Water Vapour (*PWV*) and are formulated in Equations (37):

$$\begin{aligned} IWV &= \int \rho_v dz \left[\frac{\text{g}}{\text{cm}^2} \right]; \\ PWV &= \frac{IWV}{\rho_l} \text{ [cm]}. \end{aligned} \quad (37)$$

Here, ρ_l is the density of liquid air. WV radiometers and infrared solar hygrometers were the early instruments developed mainly for meteorologists and astronomers respectively for remote sensing *PWV*. Ground microwave radiometers have been extensively developed for correcting tropospheric path delay in geodetic analysis (Nilsson *et al.*, 2006). These radiometers are however limited to sensing the wet part of the delay only. Ground-based WVRs measure the sky brightness temperatures at a given frequency, ν , see Equation (38).

$$T_{b,\nu} = T_{b0,\nu} e^{\tau(\nu,\infty)} + \int_S T(s) \alpha(\nu, s) e^{-\tau(\nu,s)} ds, \quad (38)$$

where S is the straight path from the radiometer to the top of the atmosphere. The frequency, the cosmic background temperature and total attenuation coefficient (due to WV, oxygen and liquid water) are denoted by $\nu, T_{b0,\nu} \sim 2.8\text{K}$ and $\alpha(\nu, s)$ respectively. Whereupon $\tau(\nu, s) = \int_0^s \alpha(\nu, s') ds'$ is the opacity between location s and the radiometer. The total atmospheric attenuation along path S is denoted by $e^{\tau(\infty,\nu)}$.

In the estimation of WV, an observational frequency where the WV content influences the brightness temperature; in most cases one of the water absorption lines is often used. Further, contribution by other gases and liquid water to the brightness temperature ought to be accounted for to obtain accurate WV. Accounting for the cloud liquid water is not an easy task. Nevertheless, liquid water contribution can be removed by combining two measurements at two frequencies (22.225 GHz and 30 GHz WV line) since the contribution to the absorption coefficient $\alpha \propto \nu^2; \forall r \ll f^{-2}$.

The relation between brightness temperature and WV is fairly non-linear due to the $e^{-\tau(\infty,\nu)}$. Using a linearised brightness temperature ($T_{b,\nu}$) or introducing a new parameter expressed in terms of the brightness temperature solves the problem of separating water

vapour from brightness temperature. Here, $\tau(\infty, \nu)$ is often derived from Equation (39) using the effective temperature T_e :

$$\tau(\infty, \nu) = -\ln \frac{T_e - T_{b,\nu}}{T_e - T_{b0,\nu}}; \quad (39)$$

where $T_e(\nu) \times 1 - e^{-\tau(\infty,1)} = \int_S T(s) \alpha(\nu, s) e^{-\tau(\nu, s)} ds$. An approximate model expressing T_e as a function of surface temperature and elevation angle of the path S can be used thereafter and the amount of WV in the atmosphere would be retrieved from $\tau(\infty, \nu)$. Some of the biases in the estimated WV could therefore be as a result of error propagation from the T_e bias. Similarly, by defining $T_{b,\nu'}$ and using the opacities, the linearised effective temperature $T_{e'}$ could be obtained based on Equations (40):

$$\begin{aligned} T_{b,\nu'} &= T_{b0,\nu} \{1 - \tau(\infty, \nu)\} + \int_S T(s) \alpha(\nu, s) e^{-\tau(\nu, s)} ds, \\ T_{b,\nu'} &= T_{b0,\nu} + \{T_{e'}(\nu) - T_{b0,\nu}\} \tau(\infty, \nu), \\ T_{e'} &= \frac{\int_S T(s) \alpha(\nu, s) e^{-\tau(\nu, s)} ds}{\tau(\infty, \nu)}. \end{aligned} \quad (40)$$

The relationship between opacity and WV has been reported by Deuber *et al.*, (2005). As can be deduced, surface temperature could therefore be used to derive $T_{e'}(\nu)$. Biases in $T_{b,\nu'}$ due to T_e are cancelled out by the errors in $T_{e'}(\nu)$ due to the correlation between $T_{e'}(\nu)$ and $T_{b,\nu'}$ (Jarlemark, 1997). As reported by Jarlemark, (1997), the model that is used to convert the linearised brightness temperature to TD_w^z is given by Equation (41):

$$TD_w^z = c_b \left\{ \left[\frac{\nu_2}{\nu_1} \right]^2 T_{b(\nu_1')} - T_{b(\nu_2')} - T_{bo_2} \right\}. \quad (41)$$

Here, T_{bo_2} is the cosmic microwave background radiation and oxygen contribution to the linearised brightness temperatures. Contribution of the cloud liquid water to the linearised brightness temperature is eliminated in the two frequencies by double differencing during the analysis process. The conversion factors c_b are determined in two possible ways. Firstly, simultaneous measurements of brightness temperatures with the radiometer are used together with TD_w^z from either GNSS or radiosondes measurements. Secondly, TD_w^z and brightness temperatures could be simulated from radiosonde data and models of the attenuation coefficients and thereafter use these values to determine c_b . As a result, the brightness

temperature and TD_w^z biases emanate from the systematic errors in radiosonde measurements. These errors are minimal owing to the fact that they cancel out when estimating c_b and T_{bO_2} . However, models that are used to compute $\alpha(\nu, s)$ from radiosonde measurements may introduce systematic biases in the derived algorithm parameters.

Space-borne microwave and Infrared (IR) radiometers have also been used for measuring WV and water profiles; these provide accurate tropospheric vertical WV structures. In addition, Raman LIDARS have also been used for external correction of both the wet and hydrostatic path delay by measuring the inelastic backscatter of WV incident on the laser pulse over large scales (Tarniewicz *et al.*, 2002). The Raman LIDARS have been particularly useful in sensing the lower atmosphere where gradients are highly pronounced. In general, active and passive remote sensing techniques (on different platforms) have been used to obtain n-D fields of WV each employing different retrieval methods (e.g., GPS occultation (Ao, 2007) and rapid WV retrieval using Raman and differential absorption LIDAR (Dinoyev *et al.*, 2006). Some of the global monitoring campaigns of WV have been carried out on space-borne platforms such as Moderate Resolution Imaging Spectroradiometer (MODIS), CHALLENGING Minisatellite Payload (CHAMP)- Constellation Observing System for Meteorology, Ionosphere & Climate (COSMIC) and Upper Atmosphere Research Satellite (UARS) (Banks *et al.*, 1978) while others are terrestrial based (e.g., SHADOZ). Retrieval of WV over regional scales and short time scales based on dedicated campaigns has recently increased.

All the efforts (this includes the development of measurement systems and retrieval techniques) to monitor tropospheric WV take cognisance of the difficulties in modelling the spatial temporal variability and evolution of WV. In particular, the fluctuations of WV exhibit complex modes, each associated with different (coupled) physical processes that act as a feedback system in Earth's climate system (useful in climate modelling and meteorology) and imposes the accuracy limitation to the geodetic delay observable (vital to the space geodesy research community).

2.7. Analysis strategies for TD/WV in space geodesy

Tropospheric WV rarely attains a permanent hydrostatic equilibrium but continuously changes into or from ice, liquid and WV (which dominates). It is primarily in the vapour phase that water is globally transported into the air. Though WV constitutes 1 part per million

of water on Earth, it transports and redistributes huge amount of moisture and energy (latent heat). The two week life-span of WV in the atmosphere is characterised by vertical and horizontal transports, mixing, condensation, precipitation and evaporation. WV is therefore a key element in climate of the Earth and the hydrological cycle. It is the most variable amongst the major components of the Earth and a vital element in numerical weather prediction as reported in Cucurull, (2000).

Treuhaft and Lanyi, (1987) used a statistical model to quantify the effect of the dynamic wet atmosphere on radio interferometry measurements. The statistical model employed had two major assumptions: a) that a simplified Kolomogorov theory could be used to approximate the spatial structure of the refractivity fluctuations and, b) the temporal fluctuations are caused by spatial patterns driven by wind. In addition, the model assumed that the WV spatial structure and the wind vector were independent of atmosphere height up to some predefined effective height. The structure function given by Equation (42) was then used to describe the spatial characteristics of the wet troposphere.

$$D_{\text{wv}}(\mathbf{r}, \Delta\mathbf{r}) = \left\langle \left[\text{wv}(\mathbf{r} + \Delta\mathbf{r}) - \text{wv}(\mathbf{r}) \right]^2 \right\rangle \quad (42)$$

In Equation (42), $D_{\text{wv}}(\cdot) \Leftrightarrow D_{\text{wv}}(|\Delta\mathbf{r}|)$ and therefore can be written in the form

$$\begin{aligned} D_{\text{wv}}(\Delta\mathbf{r}) &= \left\langle \left[\text{wv}(\mathbf{r} + \Delta\mathbf{r}) - \text{wv}(\mathbf{r}) \right]^2 \right\rangle \\ &= C^2 \Delta\mathbf{r}^\alpha \end{aligned} \quad (43)$$

where the spatial in-homogeneity of WV is characterised by C . For geodetic VLBI, the general expression for the spatial structure function of the slant delay, $\tau_{\theta,\phi}$ (here, $\{\theta;\phi\}$ are elevation and azimuth respectively) between two VLBI antennas separated by baseline vector \mathbf{b} for an atmosphere of effective height, h is given by Equation (44), see Treuhaft and Lanyi, (1987) for further details.

$$\begin{aligned} D_{\tau_{\theta,\phi}}^b &= \left(\frac{1}{\sin \theta} \right)^2 \int_0^h \int_0^h dz dz' \times \left\{ D_f \left(\left[b^2 + 2(z-z')b \cot \theta \cos \phi + A^2 \right]^{0.5} \right) - D_f A \right\} \\ A &= \left(\frac{z-z'}{\sin \phi} \right) \end{aligned} \quad (44)$$

where $D_{\text{wv}} \in D_f$. Additionally, based on the *frozen* spatial structure assumption, the temporal structure function given by Equation (45) could be derived by setting $b = vt$, where v is the wind speed at time t .

$$D_{\tau_{\theta, \phi}}^{\Delta t} = \left\langle \left[\tau_{\theta, \phi}(\mathbf{r}, t + \Delta t) - \tau_{\theta, \phi}(\mathbf{r}, t) \right]^2 \right\rangle \quad (45)$$

Using Equations (44) and (45), the differential fluctuations between two geodetic stations and the per-station dependent temporal fluctuations can be captured based on the equality constraint given by Equation (46).

$$D^{\Delta t} = D^{b=v\Delta t} \quad (46)$$

In the above description of WV fluctuations, the Kolmogorov turbulence model is suited for representing the local spectrum of WV fluctuations. Jarlemark *et al.*, (1998) used the structure function given by Equation (47) to describe changes over time of the zenith total delay depending on the time lag, t_τ . The linear dependence between $D_{\tau^{ztd}}$ and τ was suggested as a special case of random walk process.

$$D_{\tau^{ztd}}(t_\tau) \stackrel{\text{det}}{=} \left\langle \left[\tau_{t+t_\tau}^{ztd} - \tau_t^{ztd} \right] \right\rangle \quad (47)$$

This representation has several complicating properties such as nonstationarity and the passage of fronts. As a result, it is difficult to reconcile the changes associated with the passage of synoptic scale systems with the apparent observed Kolmogorov behaviour. In Hogg *et al.*, (1981), the determinations of WV from most geodetic techniques are modelled in terms of atmospheric turbulence. As a result, WV could be viewed as a passive tracer that is blown turbulently and the methods used to analyze the WV observations are therefore statistical. The spatial characteristics of WV fluctuations based on the inherent statistical properties could be probed directly using a network of instruments such as GPS network. In addition, Zhang *et al.*, (2003) surveyed satellite and *in situ* observations and reported that the probability distribution functions of the troposphere WV in the tropics was predominantly bimodal due to the spatial-temporal gradients components.

Bevis *et al.*, (1992) reported that the geodetic TD_w^z could be estimated through two approaches. Firstly, a simple estimation method where the geodetic WV could be kept constant for a given time interval and its value obtained as part of the overall least-squares inversion. In this approach, the geodetic WV was assumed constant for a time period ranging from 1 to 24-hours. This deterministic approach implies that the WV is constrained to some value and its space-time derivative kept over some bounds. Secondly, Bevis *et al.*, (1992) used an estimation method where the analysis of geodetic WV utilises the statistical properties of the spatio-temporal variability of geodetic WV. In this approach, the fluctuation of geodetic WV is assumed to be driven by a stochastic process. This implies that process

parameters could be estimated using a Kalman filter or other related class of optimal filters based on the state-space time domain formulation.

It worthy to note that, geodetic WV and/or alternative geodetic troposphere quantities estimated via stochastic filtering would require a specific class of stochastic processes to be selected to capture the inherent fluctuations. This option would be based on *a priori* knowledge of the underlying process. However, the current practice of selecting the representative class of stochastic processes involves the visual inspection of the power spectrum. Documented evidence on the early efforts to model the variability of TD_w^z/WV using a random walk process or a first-order Gauss-Markov process was reported by Herring *et al.*, (1990) and other references therein. As reported in the literature, the variation of TD_w^z is space-time constrained to 1-20 mm/year using a specific stochastic process noise model.

Stoev and Elgered, (2005) used realisations of random walk stochastic process to characterise the spatial-temporal variation of TD using a ground network of GPS receivers in Europe with a time interval of 1 to 3-hours. In their study, using monthly TD data between 1997 and 1998, a standard deviation < 50 mm without a clearly visible seasonal component was reported. To capture the excursions present in the data, temporal structure functions were computed. It is vital to remark that using random walk processes to model the nature of TD above a given geodetic site is a first order approximation that is useful for geodetic inter-technique comparisons. The absorbing barriers model (see Grimmett and Stirzaker, 2001) was introduced as an extension to the random walk paradigm to capture local TD fluctuations associated with the passage of atmospheric fronts, thunderstorms and other local weather systems.

Recently, the work reported by Boehm *et al.*, (2007) used the turbulence strategy reported by Nilsson *et al.*, (2007) to estimate TD_w^z via simulations. In the simulations, the asymmetry in TD_w^z variations is taken into account by the covariance information between all observations at each station. As a result, a time series of equivalent TD_w^z are derived that includes the elevation and azimuth dependency as opposed to the random walk or Gauss-Markov simulations. The turbulent framework reported hinged on the following factors; a) initial zenith wet delay ($TD_{w,0}^z$), b) the wind speed and direction, $\{v;\theta\}$, c) the structure constant parameter (C_n), d) the troposphere effective height (h) and e) the height increment (Δh). A typical simulation scenario of TD_w^z and clock biases based at HartRAO, as described

by Boehm *et al.*, (2007) is given in Figure 2.5. The parameterisations in the simulation of the 24-hour equivalent TD_w^z variability illustrated in Figure 2.5 were as follows: a) $\tau_0^{zwd} = 128$, b) $C_n = 2.4 \times 10^{-7} \text{ m}^{-1/3}$, c) $\{h; \Delta h\} = \{1000; 100\} \text{ m}$ and d) $\{\nu; \theta\} = \{12 \text{ ms}^{-1}; 180^\circ\}$. In the simulations of integrated and random walk clock excitations, the Allan standard deviation, $\sigma_{asd} (\sim 2 \times 10^{-15})$ set at 50 minutes was used. The Allan variance (defined as the average fractional deviation stability) is used here to characterise the fluctuations of the noise contribution from the geodetic system clock offsets.

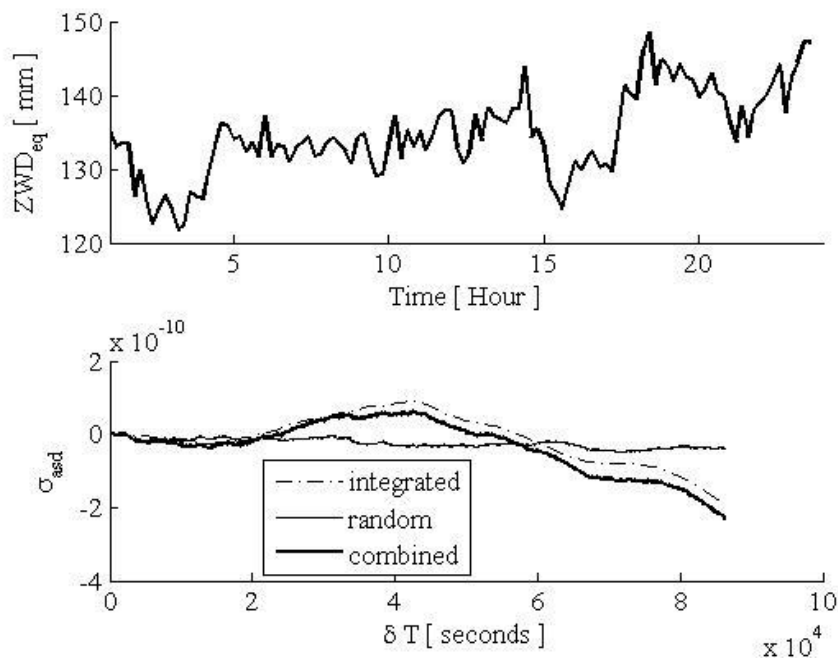


Figure 2.5. Simulation of equivalent zenith wet delay (top panel) and clock (bottom) variability using the random walk process.

A stochastic mathematical model of the combined solution of TD was formulated by Heinkelmann *et al.*, (2007). This methodology was based on the assumption of zero correlation between the solutions of the individual VLBI analysis centres (ACs); implying that the solutions of the individual ACs are independent. However, the presence of inhomogeneous structure of the standard deviations of the ACs imply that weighting of individual observations among the ACs would be unrealistic and therefore the standard deviations ought to be ignored in further analysis. As an alternative to the stochastic model, a functional approach was proposed. In this functional form, the trend and seasonal components of TD_w^z time series is incorporated into the model. The combined IVS time series

of TD given by Equation (48) are determined for each station separately by a weighting mean of hourly delay values of the ACs using the relative weighting factors derived from the variance component estimation given by Equation (48)

$$w_k = \frac{N \times \frac{1}{\hat{\sigma}_k^2}}{\sum_{k=1}^N \frac{1}{\sigma_k^2}} \quad (48)$$

$$ZTD_c = \sum_{k=1}^N \frac{w_k}{\sum w} \times ZTD_k \Big|_{N>2}.$$

Here, k and N are the individual AC and total number of ACs respectively, which contribute troposphere delay solutions, to the combined solution. The weighting scheme follows the rigorous Bounded Influence by Standardised Residual (BIBER) estimator reported by Wicki, (2001). The BIBER estimator reported (Heinkelmann *et al.*, 2007) neglects the standard deviation computed by each AC.