Optimization of Tropospheric Delay Retrieval from Numerical Weather Prediction Models and Assimilation of Zenith Path Delays from Surrounding Reference Stations

eingereicht von

Anthony Joseph Kiroe

Vollständiger Abdruck der bei der Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines Doktor der Ingenieurwissenschaften (Dr.-Ing.) eingereichten Dissertation.

Vorsitzender:	Prof. Dr.-Ing. Steffen Marburg
1. Berichterstatter:	Prof. Dr.-Ing. habil. Torben Schüler Universität der Bundeswehr München Fakultät für Luft- and Raumfahrttechnik
2. Berichterstatter:	Prof. Dr. techn. Johannes Böhm Technische Universität Wien Department of Geodesy and Geoinformation

Die Dissertation wurde am 25.07.2014 bei der Universitat der Bundeswehr Munchen,Werner-Heisenberg-Weg 39, D-85577 Neubiberg eingereicht.

Tag der mundlichen Prufung: 08.06.2015

Abstract

The Global Navigation Satellite System (GNSS) is essential for the monitoring of earth deformation processes the results of which serve as vital inputs for the realization of geodetic reference systems. However, the accuracy and precision of GNSS positioning is compromised by tropospheric wet propagation delays.

Compared to other GNSS error sources, wet tropospheric delay is small in magnitude. However, it is difficult to model due to its dependence on the temporally and spatially variable atmospheric water vapor. It therefore continuous to be the most challenging source of error on GNSS radio signals and the situation is likely to continue for years to come as the average global temperatures continue to rise. Consequently, solutions must be provided to supply sufficiently precise troposphere delay corrections.

The use of surface meteorological data does not allow the precise prediction of wet delays. In fact, current troposphere delay models leave unmodelled errors of about 3 cm in the zenith direction. Numerical weather prediction (NWP) models can be used to integrate the refractivity profiles at an accuracy level of 1 cm in zenith direction.

GNSS troposphere delay observations contain information on atmospheric water vapor. This can be used to derive the integrated water vapor above a given point. Owing to the water vapor information in the delay observations, the delay data can as well be ingested into an NWP model to improve forecasts, particularly precipitation forecasts.

This thesis presents the results and discussions from several studies that seek to revisit the topics discussed in the preceding paragraphs. The accuracy of NWP-derived troposphere delays is assessed and precipitable water vapor estimated using different methods including GNSS, NWP and radiosonde is analyzed. Part of this study includes an assessment of the accuracy of the weighted mean temperature, an important parameter in the determination water vapor using GNSS observations. The study also seeks to assess the impacts of assimilating ground-based GNSS troposphere delay observations on short-range NWP forecasts. The entire study, therefore, involves lots of accuracy assessments that are done through validation studies by comparing NWP data with similar observations sourced from meteorological sensors and radiosonde.

The studies discussed in the preceding paragraph are conducted on a region that falls in the mid-latitudes, i.e. central Europe, and are also implemented on a region falling in the African equatorial region, i.e. Kenya.

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

3D/4D Var	Three-/four-dimensional Variational system
AFREF	African geodetic reference frame
AGARD	Advisory Group for Aerospace Research and Development
ARW	Advanced Research Weather Research and Forecasting system
ASCII	American Standard Code for Information Interchange
asl	Above sea level
BUFR	Binary Universal Form for the Representation of meteorological data
DA	Data Assimilation
DAAD	*Deutsche Akademischer Austauschdienst*
DOY	Day-of-Year
ECMWF	European Center for Medium-Range Forecasts
EGM96	Earth Gravitational Model 1996
EMS	Environmental Monitoring System
EPN	EUREF Permanent GNSS Network
EUREF	European Reference Frame
ext	Extended mean parameters model
ftp	File Transfer Protocol
Galileo	European Global Satellite Navigation System
GDAS	Global Data Assimilation System
GFS	Global Forecast System
GLONASS	*GLObalnaya Navigatsionnaya Sputnikovaya Sistema*
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GrADS	Gridded Analysis and Display System
GRIB	GRIdded Binary data format
hPa	Hectopascal
http	HyperText Transfer Protocol
IGRA	Integrated Global Radiosonde Archive
IGS	International GNSS Service
ISTA	Institute of Space Technology and Space Applications
IWV	Integrated Water Vapor
LACs	Local Analysis Centers
LIDAR	Light Detection And Ranging
MODIS	Moderate Resolution Imaging Spectroradiometer

NACOSTI	National Commission for Science, Technology and Innovation
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Predictions
NCL	NCAR Command Language
NetCDF	Network Common Data Format
nfs	Network File System
NOAA	National Oceanic and Atmospheric Administration
NMM	Non-hydrostatic Mesoscale Model
num_int	Numerical integration
NWP	Numerical weather prediction
NWS	National Weather Service (US)
PWV	Precipitable Water Vapor
RINEX	Receiver INdependent EXchange file format
rms	Root mean square
saas	Saastamoinen hydrostatic model
SHADOZ	Southern Hemisphere Additional OZonesondes
SINEX	Software Independent EXchange file format
SRTM	Shuttle Radar Topography Mission
std	Standard processing
stdev	Standard deviation
STRC	Science and Training Resource Center
SuomiNet	A GPS network designed for real-time remote sensing
TOUGH	Targeting Optimal Use of GPS Humidity measurements in meteorology
UNAVCO	University NAVstar COnsortium
UTC	Universal Coordinated Time
WGS84	World Geodetic System 1984
WMO	World Meteorological Organization
WRF	Weather Research and Forecasting
WRFDA	WRF data assimilation
ZHD	Zenith Hydrostatic Delay
ZTD	Total Zenith Path Delay
ZWD	Zenith Wet Delay

List of Symbols

B, R	error covariance matrices of a background NWP model field, defined in Equation (2.45)
c	speed of light in vacuum
c_p	specific heat capacity at constant pressure
t_s, t_r	time on satellite clock and time on receiver clock respectively
d	radio signal propagation delay
dp	differential change in pressure
e	partial pressure of water vapor
E	evaporation and sublimation, defined in Equation (1.7)
f	Coriolis force, defined in Equations (1.3) and (1.4)
F_x, F_y	surface friction and turbulent mixing acting on the west-to-east and south-to-north winds respectively, defined in Equations (1.3) and (1.4) respectively
g	acceleration due to gravity
h, z	geometric height
H	geopotential height
H	radiation, condensation and mixing processes, defined in Equation (1.6)
H	observation operator, defined in Equation (2.45)
i, j, k	indices in x-, y- and z- coordinates respectively
J	cost function, defined in Equation (2.45)
$J\,K^{-1}\,mol^{-1}$	Joule per Kelvin per mole
k	refraction constant
K	kelvin
kg	kilogram
kgm^{-2}	kilogram per square meter
kgm^{-3}	kilogram per cubic meter
$kg\,kmol^{-1}$	kilogram per kilomole
l	length, signal path length
m	number of model half mass levels above a GPS antenna height
m	meter
mm	millimeter
M	molar weight
msl	mean sea level
N	cycle ambiguity, defined in Equation (1.2)
N	atmospheric refractivity

n	refractive index
n	total number of NWP model half mass levels
p_t, p_s	pressure evaluated at model top and at the surface respectively
p_b, p_p	base state and perturbation pressure respectively
P	total pressure
P	code pseudorange observable, defined in Equation (1.1)
P	condensation and subsequent precipitation
q	humidity
r	satellite-to-receiver geometric range
r	water vapor mixing ratio
R, R_d, R_w	gas constants; universal, dry gas and water vapor respectively
R_e	radius of the Earth
R_{max}, R_{min}	semi-major axis (= 6378.137 km) and semi-minor axis (= 6356.752 km) respectively
t	time
T	absolute temperature
$T_m, \bar{T}_m, \tilde{T}_m$	weighted mean temperature, average weighted mean temperature and amplitude of the weighted mean temperature respectively, defined in Equation (2.18)
Δt	change in time
u, v, w	west-east, south-north and vertical wind velocity components
x, y, z	cartesian coordinates directed eastward, northward and upward
z	zenith direction
Z^{-1}	inverse compressibility factor
X	forecast variable, defined in Equation (1.11)
ΔX	change in forecast variable, defined in Equation (1.11)
$\partial/\partial t$	partial derivative with respect to time
d/dt	total derivative
dt_s, dt_r, $d\rho$	satellite clock error, receiver clock error and orbit error respectively
Δl	change in signal path length
dl	an infinitesimal displacement along a signal path
dz	an infinitesimal displacement along a signal path in the zenith direction
ε	ratio of the molar weight of water vapor to that of dry air
ε_p	unmodelled error effects including receiver noise and multipath effects, defined in Equation (1.1)
Φ	carrier phase observable, defined in Equation (1.2)
λ	wavelength

η	Eta vertical coordinate system, defined in Equation (1.10)
σ	Sigma vertical coordinate system, defined in Equation (1.9)
Θ	potential temperature, defined in Equation (3.5)
θ	perturbation potential temperature, defined in Equation (3.5)
ω	= Dp/Dt defined in Equation (1.4)
ρ	density
π	Pi (≈ 3.14159)
φ	latitude
κ	= R_d/c_p, defined in Equation (3.6)

Subscripts

a or ant	antenna
act	actual
avg	average
b	background (initial) NWP model field
d	dry
e	Earth
hyd	hydrostatic
wet	wet (non-hydrostatic)
iono	ionosphere
l	liquid water
tropo	troposphere
ref	reference
s	surface
vac	vacuum
w	wet, water vapor

1. General introduction

The Global Navigation Satellite System (GNSS) currently consists of three main systems: USA's Global Positioning System (GPS), Russia's *GLObalnaya Navigatsionnaya Sputnikovaya Sistema* (GLONASS) and Europe's Galileo[1]. These systems are continuously modernized and updates are always made available on many sources and forums including the internet.

The GPS system signals have been used in this study and this introductory chapter guides us through the subject matter of this research by making brief descriptions of the main topics involved.

1.1. GPS error sources

GPS satellites orbit the earth at an overhead height of approximately 20,200 km (Misra and Enge, 2006). Radio signal transmitters onboard the GPS satellites continuously transmit radio signals at varying frequencies including L1 - 1575.42 MHz, L2 - 1227.60 MHz, L3 - 1381.05 MHz, L4 - 1379.13 MHz, and L5 - 1176.45 MHz that have to transverse a space-to-earth path to a GPS receiver that then computes its position solution[2]. The receiver is however unable to compute the exact position solution due to a number of error sources.

A GPS receiver receives a signal from a GPS satellite that is tagged with, among other information, the time the signal was transmitted by the satellite. It then computes the difference between the time it received the signal and the time the signal was transmitted by the satellite in order to estimate the signal transit time[3]. The GPS receiver then computes the geometric satellite-to-receiver distance (or range) by multiplying the signal transit time by the supposed speed of the signals (assumed to be the speed of light in free space i.e. 3×10^8 m/s). Ideally, for a GPS satellite that is directly overhead, the computed range is expected to be approximately 20,200 km in which case the transit time could have been approximately 0.07 seconds.

In reality, however, the space-to-earth signal path is not entirely a vacuum. The signals therefore do not travel at the speed of light in vacuum. Also by the time the receiver computes the range, the satellite has already moved a certain distance away from its position when it sent the signal. Further, the satellite and receiver clocks are not synchronized resulting to errors in the computed transit time. All these factors act as sources of error for GPS signals as illustrated in Figure 1.1, and the computed satellite-to-receiver distance is never a true range (geometric distance) but rather a pseudorange.

---

[1] The system is not yet fully operational.

[2] The receiver also computes its velocity and time.

[3] Transmit time is the difference between the time the signal was received (as determined by the receiver clock) and the time the signal was transmitted. The expected transmit time varies depending on the position of the satellite in space with respect to that of a receiver on the ground.

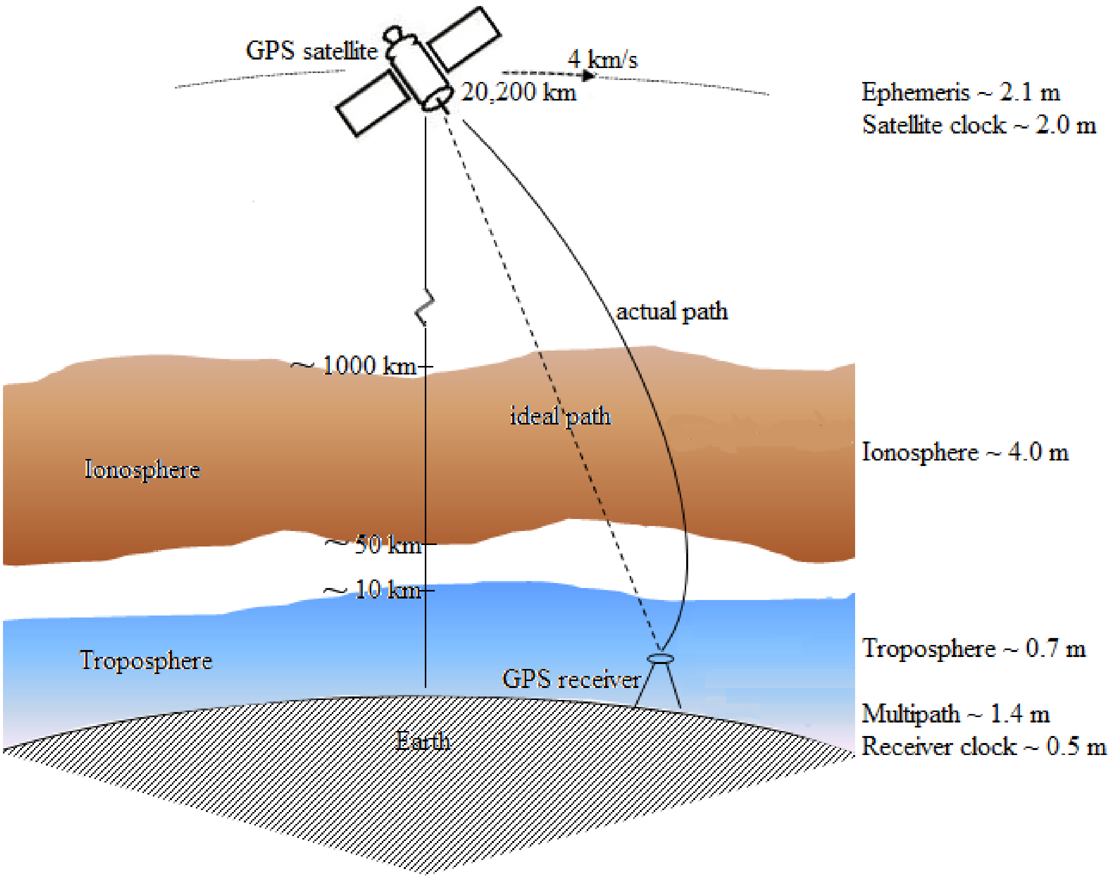


Figure 1.1: Typical GPS error sources and budget. The errors are one-sigma bias errors (Varner, 2000).

GPS errors, therefore, mainly originate from various sources, including: errors in the transmitted satellite position (ephemeris errors) and satellite clock, errors in the signal's propagation through the atmosphere (due to free electrons in the ionosphere and matter in the troposphere), and errors due to reflected signals entering the receiver (multipath) and noise in the receiver.

Following Misra and Enge (2006) GPS observables that are generated by a GPS receiver are expressed as,

- code pseudorange:

$$P = r + c \cdot (dt_s - dt_r) + d\rho + d_{iono} + d_{tropo} + \varepsilon_P \tag{1.1}$$

Where,

P ≡ code pseudorange observable

r ≡ satellite-to-receiver geometric range

c ≡ speed of light in vacuum

dt_s and dt_r ≡ satellite and receiver clock errors respectively

$d\rho$ ≡ orbit error

d_{iono} and $d_{tropo} \equiv$ ionospheric and tropospheric delay respectively

ε_p $\equiv$ unmodelled effects (including receiver noise and multipath effects).

- carrier phase observable:

$$\Phi = r + c \cdot (dt_s - dt_r) + d\rho + \lambda \cdot N - d_{iono} + d_{tropo} + \varepsilon_p \tag{1.2}$$

Where,

Φ $\equiv$ carrier phase observable

λ $\equiv$ wavelength of the carrier signal

N $\equiv$ cycle ambiguity

r, c, dt_s, dt_r, $d\rho$, d_{iono}, d_{tropo}, and ε_p are as earlier defined.

As seen in the observables equations, a significant amount of effort has been made to include all the possible errors in estimating the range to a GPS satellite. Errors, however, still remain even after the receiver corrects what it can. These errors are normally depicted in a GPS error budget as illustrated in Figure 1.1.

From the error budget, it is evident that the errors due to the atmosphere are relatively large in magnitude. They are therefore the most significant source of GPS errors. Ionosphere and troposphere delays are errors induced on a GPS signal during signal propagation through the earth's atmosphere. The overall effect is that the distance travelled by the signal is longer than the actual geometric distance between a GPS satellite and a receiver as the signal propagation speed is decreased and it follows a curved path as shown in Figure 1.1.

Ionosphere delay is caused by the presence of free electrons in the ionosphere which is the ionized portion of the upper atmosphere and extends from a height of about 50 km up to approximately 1000 km. The effect of the charged ionosphere is a nonlinear dispersion and refraction of the GPS radio signals propagating through it such that the carrier signal experiences a phase advance[1] while the code experiences a group delay.

The dispersion of GPS radio signals by the ionosphere forms the basis of one of the ionosphere delay mitigation techniques. Since the ionosphere is dispersive[2] for radio signals with frequencies above 300 MHz, GPS radio signals propagating at different frequencies follow different paths in the ionosphere

---

[1] A particular phase of the carrier signal arrives at a receiver earlier than it would have had the signal travelled from the satellite in a complete vacuum.

[2] Dispersion here refers to frequency dispersion. L-band frequencies are in the atomic frequency range of the ionosphere's electrically charged medium. The resulting resonance affects the radio signals in a frequency-dependent manner (Thessin, 2005).

and arrive at a GPS receiver at different times. By making measurements on the two frequencies (using dual frequency receivers) and combining them, first order ionospheric effects can be eliminated in what is referred to as the ionosphere-free linear combinations method (Schueler, 2001; Alizadeh *et al.*, 2013). This enables the removal of approximately 99% of ionosphere delay in most cases. Second order effects are much smaller (a few mm).

Though the largest error source seems to be the ionosphere, it is relatively easy to model and methods exist that help to mitigate the ionospheric effects. In single-frequency receivers a model developed by Klobuchar (1987) is used to eliminate approximately 50% of the error in most cases.

Troposphere delay, on the other hand, is caused by the presence of matter in the earth's troposphere that contains about 80 % of the total molecular mass of the atmosphere and extends from the earth's surface up to an average of 12 km in altitude[1]. Upon interaction with matter, the GPS radio signals undergo the mechanisms associated with electromagnetic signal interaction with matter. The overall effect mainly depends upon the troposphere constituent that is the most abundant and this happens to be water vapor.

Atmospheric water vapor is highly variable both in space and time and this makes it difficult to properly model troposphere delay. This generally makes troposphere delay models ineffective particularly in high accuracy GPS positioning applications. According to a study by Mendes (1999), even the best global troposphere delay models do leave unmodelled errors of about 3 cm in the zenith.

The nondispersive nature of the troposphere for radio signals up to 30 GHz (obviously includes GPS signals) further complicates troposphere delay mitigation efforts. This is because GPS radio signals propagating through the troposphere undergo similar effects despite differences in frequency making it impossible to eliminate troposphere delay through dual frequency measurements (AGARD, 1970; Langley, 1996; Thessin, 2005).

The challenges presented by troposphere delay in GPS positioning applications makes it an ongoing and interesting research topic aimed at mitigating the tropospheric effects. Efforts to achieve this have included the development of troposphere delay models and the use of numerical weather prediction (NWP) models in a bid to estimate troposphere delays as accurately as possible and hence provide the necessary troposphere delay corrections for precise GPS positioning.

In troposphere delay modeling, analytical models are used to estimate the troposphere delay in the zenith direction. The zenith delay is then mapped to the elevation angle of the ray-path of the arriving signal by use of a mapping function. Detailed discussions on troposphere delay modeling/models and mapping functions can be found in e.g. Mendes (1999) and Schueler (2001).

---

[1] The troposphere height is a function of latitude and reaches a maximum in tropical regions.

Meanwhile, despite being a nuisance in GPS positioning applications troposphere delay has turned out to be a delight in meteorology applications. This is further discussed in the section that follows.

1.2. Troposphere delay: a nuisance in GPS positioning, a blessing in meteorology

The distribution of atmospheric constituents in space is heterogeneous and since each constituent is characterized by a certain refractive index, it therefore follows that a GPS signal goes through regions of varying refractive indices as it propagates from a satellite to a receiver.

The atmospheric radio refractive index mainly depends on atmospheric pressure, absolute temperature and water vapor pressure. Figure 1.2 illustrates the spatial distribution of these variables in the atmosphere and is important in providing an insight into the spatial variability of atmospheric water vapor compared to pressure and temperature particularly in the lower atmosphere. The profiles were deduced from an NWP model simulation over the Wettzell geodetic observatory in Germany on 29th May 2013 18 UTC.

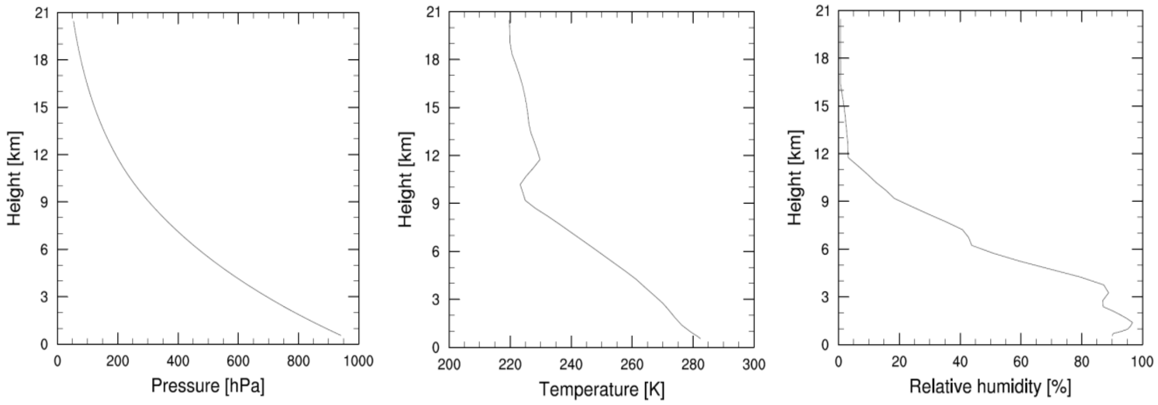


Figure 1.2: Vertical profiles of NWP simulations of pressure, temperature and relative humidity over the Wettzell geodetic observatory for 29th May 2013 18 UTC.

The vertical profiles go up to the maximum NWP model height of approximately 20 km. Pressure (left panel) is seen to decrease with height as the atmosphere varies in density and composition above the earth's surface. The temperature (center panel) averages 12°C near the surface and decreases rapidly and almost linearly with increasing height (at approximately 6.5°C/km on average) up to a minimum value of approximately -50°C, which defined the troposphere's upper boundary, the tropopause, at approximately 9 – 10 km altitude[1]. The region at which the temperature remains nearly constant marks the start of the stratosphere. Relative humidity (right panel) is seen to vary greatly near the surface after which it starts decreasing with height.

Being an electromagnetic signal, a GPS signal is sensitive to the vertical variability of the atmospheric refractive index which is mainly due to the spatial (right panel of Figure 1.2) and temporal (Figure 1.3) variability in the atmospheric water vapor content. This is what constitutes troposphere delay and

---

[1] Tropopause height varies geographically.

methods have been devised to extract the water vapor information from it as seen in e.g. Mendes (1999) and Schueler (2001) and further discussed in Section 2.1.

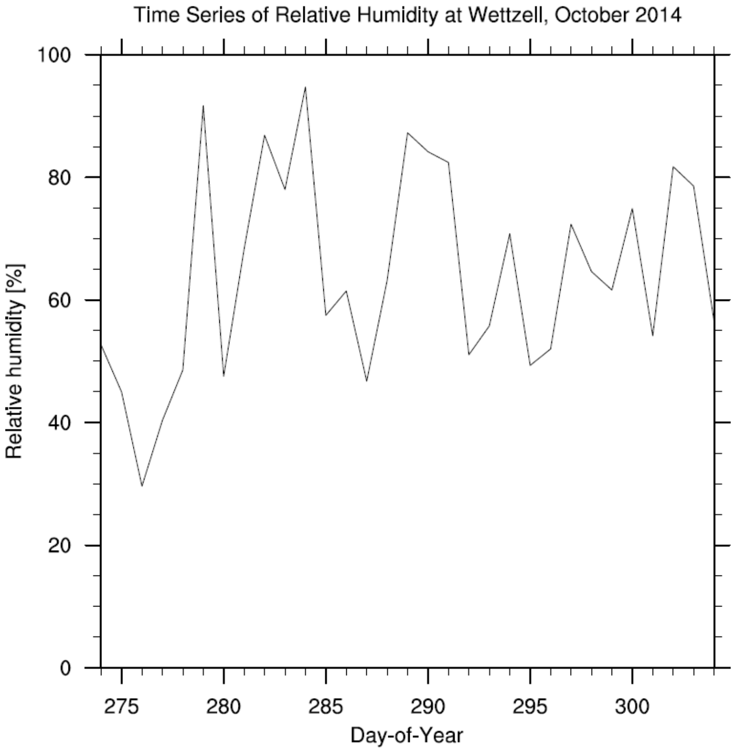


Figure 1.3: Trend in relative humidity atWettzell geodetic observatoryin October 2013.

Troposphere delay, therefore, has turned out to be yet another source of information on atmospheric water content that is a vital ingredient in meteorology applications (Haan, 2008; Jones, 2010). The quality of GPS-derived precipitable water vapor (PWV) has been ascertained and found to be of high accuracy as seen in e.g. Christian *et al.*(1995). Therefore, though seen as a nuisance in GPS positioning applications, troposphere delay is seen as a blessing in meteorology applications.

The use of GPS in the remote sensing of atmospheric water vapor is what entails GPS meteorology. The high temporal and spatial variability of atmospheric water vapor causes difficulties in modeling water vapor-related atmospheric processes (Hall and Manabe, 1999). This includes efforts on troposphere delay modeling. Consequently, any source leading to information concerning atmospheric water vapor is treated with utmost importance. Such is the case with GPS meteorology.

Traditional standard systems used for measuring atmospheric water vapor, notably radiosonde and meteorological satellites, have not effectively monitored it due to limitations linked to its variability in space and time. Details on this including other water vapor measurement systems such as ground-based water vapor radiometers, LIDAR[1] systems and sun-photometers can be found in e.g. Gaffen *et*

---

[1] Light Detection And Ranging is a remote sensing technology that emits intense focused light and measures the time it takes for the reflections to be detected by a sensor (Carter, *et al.*, 2012).

al.(1991), Han *et al.*(1994), Christian *et al.*(1995), Elgered and Jarlemark (1998), Plana-Fattori *et al.*(1998), Karbou *et al.*(2005), Jones (2010), and Neely and Thayer (2011).

The GPS system, with its global coverage, high temporal resolution (from a few minutes up to 2 hours) and all-weather availability, has come to be widely used as an alternative source for water vapor observations (Bevis *et al.*, 1992; Yuan *et al.*, 1993; Hagemann *et al.*, 2003; Sibylle *et al.*, 2010).

Currently, GPS, radiosonde and ground-/space-based microwave radiometers are the most commonly used atmospheric water vapor estimation systems. The GPS system emerges as the most superior of these methods (Table 1.1).

Table 1.1: Comparison of the main atmospheric water vapor estimation methods (adapted from Guerova (2003)).

	Radiosonde	**Microwave radiometers**	**GPS**
Temporal resolution	Low	High	High
Spatial resolution	Low	Low	High
Vertical resolution	High	Low	Low
All-weather availability	No	No	Yes
Cost	High	High	Medium
Dataset length	Since the 1940s	Since 1990	Since 1990

Many applications have been found for the GPS water vapor observations including, validating radiosonde observations (Köpken, 2001) and satellite water vapor observations (Rama Varma Raja *et al.*,2006), monitoring climate change (Yuan *et al.*, 1993), and assimilating into numerical weather prediction models in order to improve weather prediction (Guo *et al.*, 2004).

The accuracy of GPS water vapor observations, however, depends to a large extent on the vertically averaged temperature of the atmosphere. This parameter is water-vapor weighted and is therefore variable from one geographical location to another. The usual practice is to use a weighted mean temperature model in the estimation of water vapor observations using GPS. These models have been developed and are suitable for global, regional and site-specific applications. The commonly used method for the development of such models has been the use of radiosonde data. Numerical weather prediction models can also be used to develop the weighted mean temperature models. This subject is discussed further in Section 2.2 and forms one of the subjects of the studies presented in Chapter 5 and Chapter 7.

1.3. Numerical weather prediction (NWP) models with relevance to troposphere delay

Numerical Weather Prediction (NWP) is the use of numerical methods to solve mathematical equations pertaining to the fluid dynamics of the atmosphere that are derived from conservation laws of mass, momentum, thermodynamic energy, and moisture in order to predict the state of the atmosphere at a specified time in future (Mesinger and Arakawa, 1976).

The topic of numerical weather prediction is quite wide and deep engagement into the topic falls beyond the context of this study. A brief overview of NWP modeling, however, is presented in this section.

1.3.1. The fundamental equations in NWP

The principal NWP mathematical equations include the hydrostatic equation, the thermodynamic equation, moisture equation, equation of state, continuity equation and the equations of motion (uk.sci.weather resources, 2014). They are generally referred to as primitive equations in the sense that they describe the fundamental processes that occur in the atmosphere. Using the pressure coordinates (x, y, p) and the basic forecast variables: the three wind components (u,v,w), temperature (T), humidity (q) and height/pressure (z), the primitive equations are expressed as follows (Popovic, 2006; MetEd, 2014),

- Wind forecast equations:

$$\frac{\partial u}{\partial t} = -u \cdot \frac{\partial u}{\partial x} - v \cdot \frac{\partial u}{\partial y} - \omega \cdot \frac{\partial u}{\partial p} + f \cdot v - g \cdot \frac{\partial z}{\partial x} + F_x \tag{1.3}$$

$$\frac{\partial v}{\partial t} = -u \cdot \frac{\partial v}{\partial x} - v \cdot \frac{\partial v}{\partial y} - \omega \cdot \frac{\partial v}{\partial p} + f \cdot u - g \cdot \frac{\partial z}{\partial y} + F_y \tag{1.4}$$

Where,

Equation (1.3) and Equation (1.5) ≡ the wind forecast equations for the west-to-east wind velocity component (u) and the south-to-north component (v) respectively,

Both are prognostic equations in that they use a time derivative in solving for the two horizontal wind components. u and v are therefore prognostic variables,

ω ≡ pressure change following the motion ($\omega = \frac{Dp}{Dt}$, where $\frac{D}{Dt} = \frac{\partial}{\partial t} + u \cdot \frac{\partial}{\partial x} + v \cdot \frac{\partial}{\partial y} + w \cdot \frac{\partial}{\partial z}$, w is the vertical wind component),

f ≡ the Coriolis force[1],

F_x and F_y ≡ surface friction and turbulent mixing acting on the west-to-east and south-to-north winds respectively,

g ≡ acceleration due to gravity.

---

[1] The Coriolis effect causes a deflection (when viewed from a rotating reference frame) in wind patterns when a force acts perpendicular to the direction of motion of the wind. It is important in this situation because Newton's 2nd law of motion is valid only in an inertial reference frame, whereas the Earth's frame is non-inertial.

- Continuity equation:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial p}=0 \tag{1.5}$$

- Temperature forecast equation:

$$\frac{\partial T}{\partial t}=-u\cdot\frac{\partial T}{\partial x}-v\cdot\frac{\partial T}{\partial y}-w\cdot\left(\frac{\partial T}{\partial p}-\frac{RT}{c_p p}\right)+\frac{H}{c_p} \tag{1.6}$$

Where,

T ≡ a prognostic variable,

R ≡ the ideal gas constant,

$\frac{H}{c_p}$ ≡ other processes that cannot be forecasted directly including radiation, condensation and mixing,

c_p ≡ the specific heat at constant pressure.

- Moisture forecast equation:

$$\frac{\partial q}{\partial t}=-u\cdot\frac{\partial q}{\partial x}-v\cdot\frac{\partial q}{\partial y}-w\cdot\frac{\partial q}{\partial p}+E-P \tag{1.7}$$

Where,

q ≡ a prognostic variable,

E ≡ evaporation and sublimation,

P ≡ condensation and subsequent precipitation.

- Hydrostatic equation:

$$\frac{\partial z}{\partial p}=-\frac{RT}{pg} \tag{1.8}$$

In summary, Equations (1.3), (1.4), (1.6) and (1.7) are prognostic equations while Equations (1.5) and (1.8) are diagnostic (in that they don’t involve time derivatives). Consequently, the two horizontal wind components, temperature and humidity are prognostic variables and they must therefore be assigned initial conditions (Randall, 2004).The diagnostic equations use prognostic variables to solve for diagnostic variables including pressure and the vertical wind component.

An NWP model solves the primitive equations[1]numerically in either of two techniques, grid-point or spectral, and hence representing the variables in a gridded form (grid-point model) or spectral form (spectral model) respectively. While a grid-point model represents variables at discrete fixed grid points, a spectral model represents them in continuous wave forms (MetEd, 2014). Therefore, although both model types are based on the same primitive equations, each solves them differently and hence represents data in its own unique way with varying magnitudes in forecast errors. Also, since the spectral method requires less computing resources than the grid-point method, it is primarily used in global (areal coverage) NWP models. Examples of spectral models include the GFS (Global Forecast System) and the ECMWF (European Centre for Medium-Range Weather Forecasts) models. The grid-point method, on the other hand is mostly used in limited-area models with a few global models using it and this tends to increase as computing resources get advanced. Examples of grid-point models include the WRF (Weather Research and Forecasting), COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System), and NAM (North American Mesoscale Forecast System).The spectral method is not further discussed here. A detailed discussion, however, can be found in MetEd (2014).

1.3.2. Grid-point NWP model description

A grid-point model lays out the atmosphere into a 3-dimensional grid by dividing it into a grid of horizontal boxes that extend vertically into a number of vertical layers to form a large grid composed of several 3-dimensional boxes (Figure 1.4). Variables are therefore computed at discrete grid points.

Figure 1.5 (a) shows a grid box with a grid point at each corner. Grid points on the same horizontal surface form a vertical level and each of these grid points contains values that are representative of the mean values of the variables in the box of air surrounding it (Figure 1.5 (b)). The horizontal spacing between grid points (grid spacing/length) defines its horizontal resolution in that it determines the scale of the features that can be simulated with a grid-point NWP model. NWP model resolution is further discussed in Section 1.3.3.

A single grid box, however, represents a very small portion of the entire atmosphere. Small scale features and/or physical processes that occur within the confines of such a single grid box cannot properly be resolved by the NWP model. They must be parameterized to allow their average contribution on a large scale to be included in the NWP model (Eumetcal, 2014). This is achieved by use of parameterization schemes which are models that are embedded into the NWP model for the purpose mentioned. Parameterized features and processes include land surface processes (conduction, evaporation, etc.), convection, radiation, precipitation and clouds (Figure 1.6).

---

[1] NWP models can be hydrostatic or non-hydrostatic where the former assumes hydrostatic equilibrium and hence uses hydrostatic prognostic equations while the latter uses non-hydrostatic prognostic equations.

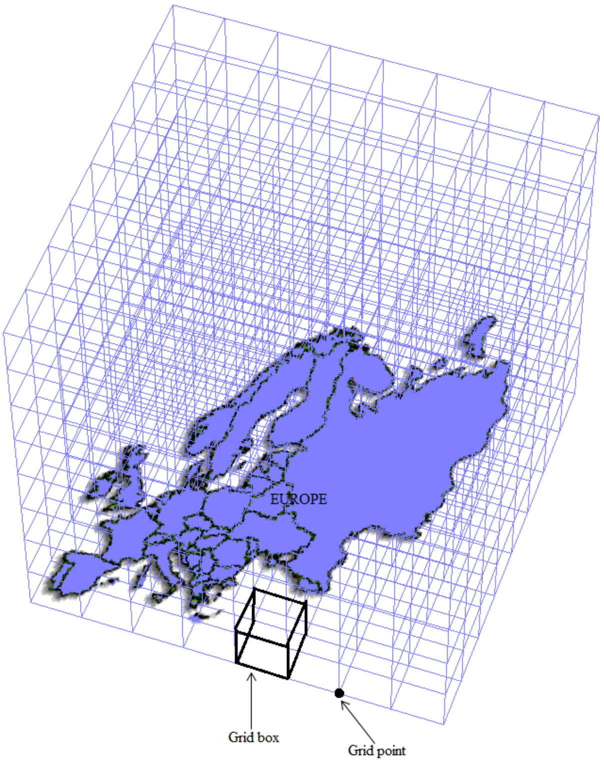


Figure 1.4: Typical grid of an NWP model. The illustration depicts a grid-point NWP model that covers Europe.

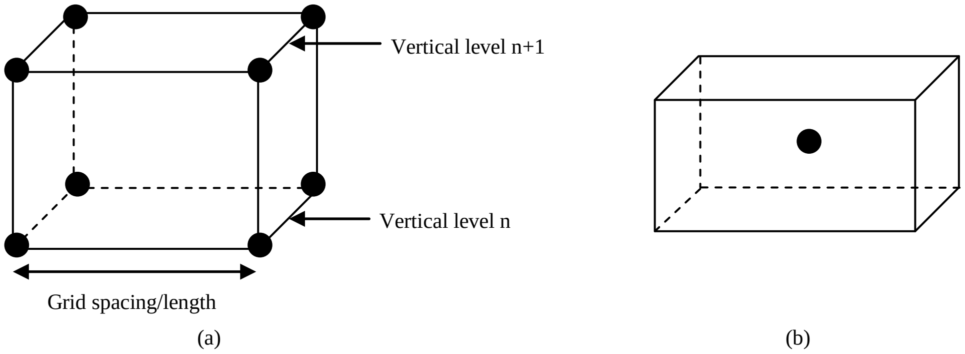


Figure 1.5: Illustration of grid points and grid spacing (adapted from Eumetcal (2014)).

A given NWP model supports several parameterization schemes examples of which includes the Arakawa-Schubert scheme, Betts-Miller-Janjic scheme, Kuo scheme, etc. A description of these models can be found in e.g. Betts (1996) and MetEd (2014). Each of them is often optimized for a certain range of grid-spacing in the model (Eumetcal, 2014).

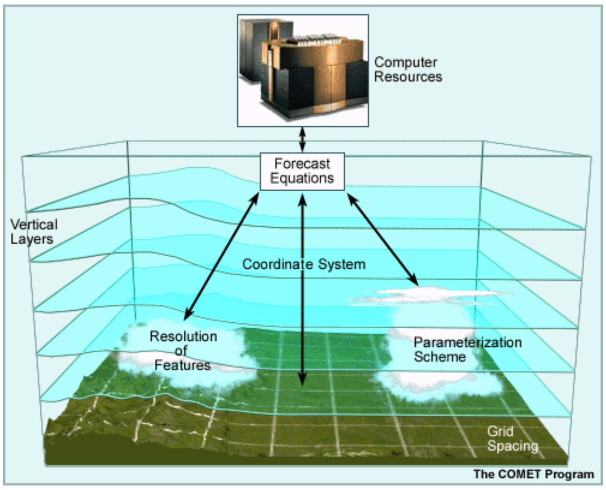


Figure 1.6: Typical NWP model configurations (adapted from MetEd (2014)).

1.3.3. NWP model resolution

Grid spacing/length is related to the horizontal resolution of an NWP model and it is a major factor in determining NWP model resolution such that the smaller the grid spacing the better (or higher) the resolution and vice versa. Specifying it mainly requires a consideration of the available computing resources and the area to be covered.

The arrangement of prognostic variables within an NWP model's grids determines how the model calculates variables. They may be arranged in such a way that all the variable types are defined at the same point (unstaggered grid, as depicted in Figure 1.7 (a)) or in such a way that we have more than one point containing variables of different types (staggered grid, as depicted in Figure 1.7 (b - e)).Grid staggering, as defined in the latter case, offers advantages including an increase in grid resolution. An explanation of this can be found in literature including e.g. Collins, *et al*. (2013). A summarized overview of the so-called Arakawa grids is shown in Figure 1.7. The indices i-1, i and i+1 represent grid points in the east-west direction whereas indices j-1, j and j+1 represent grid points in the south-north direction. The A grid (Figure 1.7 (a)) is the only unstaggered grid. The rest (Figure 1.7 (b - e)) are staggered.

In the staggered grids the values of the wind components (u, v) and the thermodynamic variables (represented by h, e.g. temperature and pressure) are calculated at different points. This is seen in the B, C, D and E grids of Figure 1.7. However, the wind components are at the same point in the B and E grids and at different points in the C and D grids. A model's resolution is defined as the average distance between adjacent grid points with the same variables. The grid type therefore has an effect on the resolution.

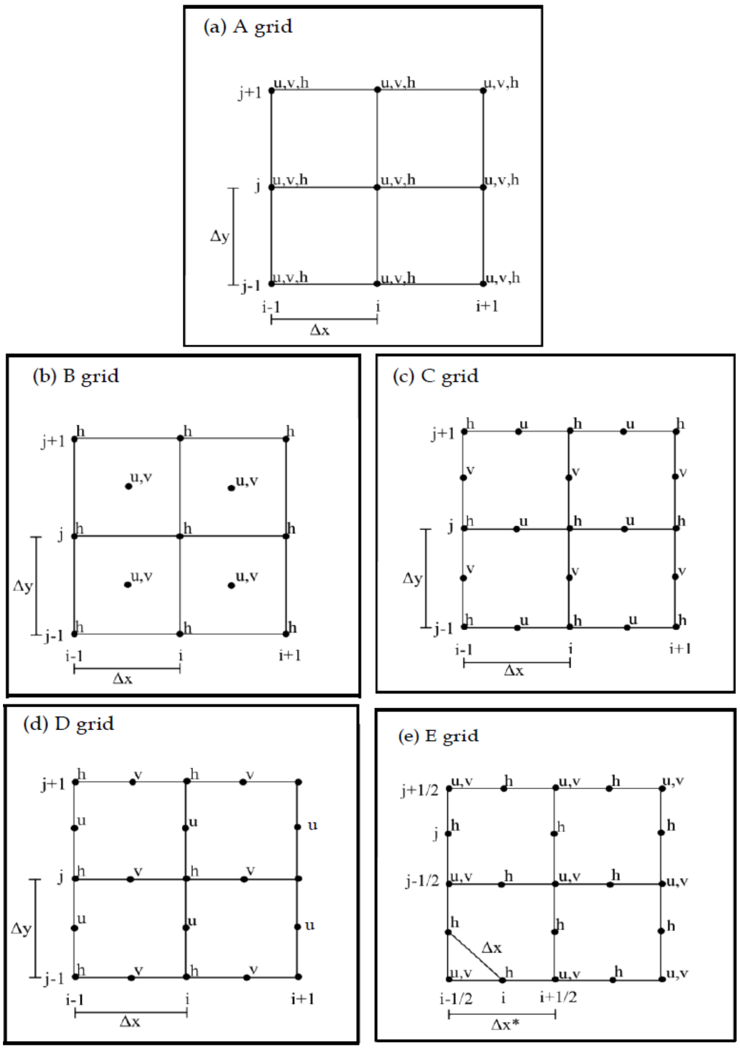


Figure 1.7: Grid types based on horizontal staggering, namely, (a) the A grid, (b) the B grid, (c) the C grid, (d) the D grid and (e) the E grid. Δx is the east-west resolution and Δy is the north-south resolution (adapted from Collins, *et al*. (2013)).

In the vertical direction, the atmosphere is divided into several vertical layers that encompass its vertical extent. As such, vertical resolution is mainly determined by the number of vertical layers and the model height. The resulting atmospheric layers are defined by a vertical coordinate system that is mainly specified in terms of pressure levels (or height). An NWP model therefore forecasts for the average over an atmospheric layer between the vertical coordinate surfaces and not on the surfaces themselves (MetEd, 2014).

The most commonly used systems, however, mainly take the form of the terrain-following coordinate systems, sigma (σ) and eta (η), or their variants and hybrid systems that combine different systems in order to capitalize on the advantages of each while handling their associated limitations.

Figure 1.8 shows a simplified overview of the sigma coordinate system. The lower levels of the sigma coordinate system tend to follow the terrain and this tendency progressively flattens as the pressure decreases towards the top of the model atmosphere. The coordinate surfaces are therefore sloped where the terrain is sloped and flattened where the terrain is flat.

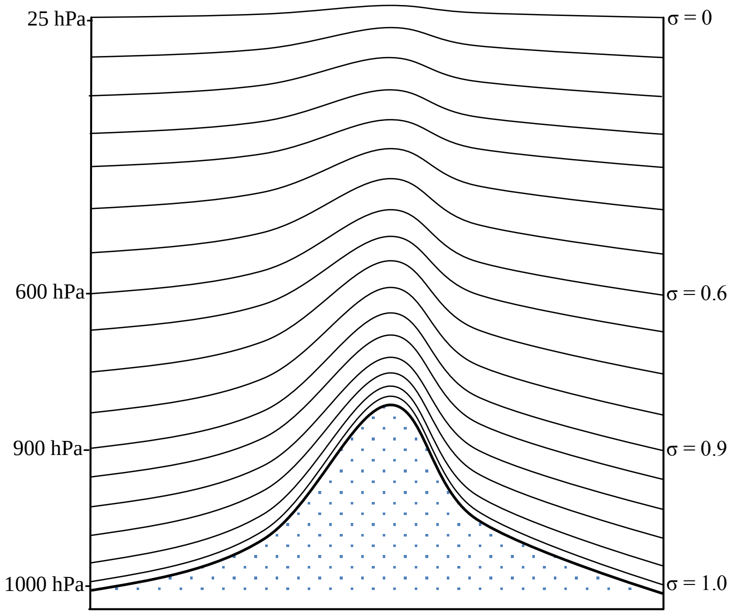


Figure 1.8: Simplified illustration of the sigma vertical coordinate system.

In the sigma system, the vertical position of a surface in the atmosphere is defined as the ratio of the pressure difference between that point and the top of the model domain to that of the pressure difference between the surface and the top of the model domain. It is defined by (Popovic, 2006; MetEd, 2014),

$$\sigma = \frac{p - p_t}{p_s - p_t} \tag{1.9}$$

Where, p, p_t and p_s are pressure at a point, pressure at the top of the model domain and pressure at the surface respectively.

The lowest sigma coordinate surface mimics the actual terrain and it is labeled $\sigma = 1$ while that at the top of the model domain, typically placed well above the tropopause at between 25 and 1 hPa, is labeled $\sigma = 0$ (MetEd, 2014). The other sigma surfaces therefore take values between 0 and 1.

Surfaces on the eta coordinate system, on the other hand, are relatively horizontal so that the bottom atmospheric layer is presented in a step-like fashion (Figure 1.9).

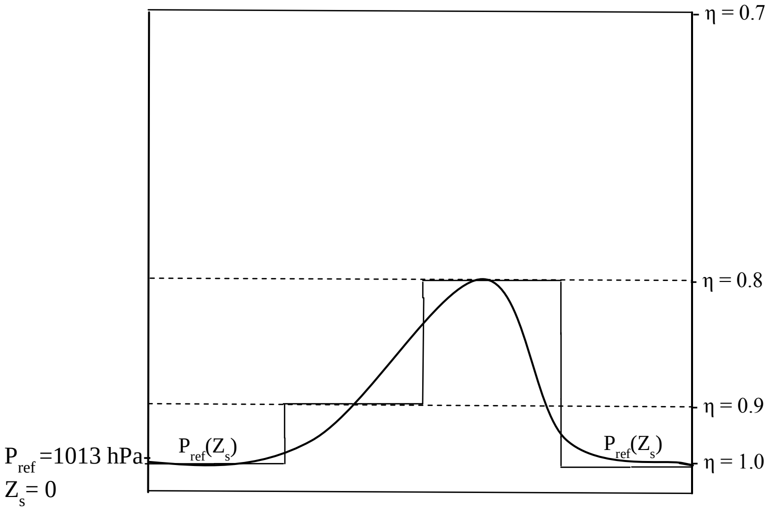


Figure 1.9: Simplified illustration of the eta vertical coordinate system.

In the eta system, the vertical position of a surface in the atmosphere is defined as the ratio of the pressure difference between that point and the top of the model domain to that of the pressure difference between the mean sea level and the top of the model domain. It is defined by (Popovic, 2006; MetEd, 2014),

$$\eta_s = \frac{p_{ref}(z_s) - p_t}{p_{ref}(z=0) - p_t} \tag{1.10}$$

Where, p_t ≡ pressure at the top of the model domain

$p_{ref}(z=0)$ ≡ mean sea level pressure (= 1013 hPa)

$p_{ref}(z_s)$ ≡ standard atmospheric pressure at surface height z_s.

Like the sigma system, the eta system is usually labeled from 0 to 1 from the top of the model domain to the sea level.

Both the sigma and eta coordinate systems are terrain-following and each has advantages and disadvantages. Further discussion on this topic can be found in e.g. Kasahara (1974), Popovic (2006) and MetEd (2014).

There are also advantages associated with the staggering of grids in the vertical direction. A discussion on this can be found in Fox-Rabinovitz (1994). The most commonly used vertical staggering grids include the Lorenz grid (Lorenz, 1960) and the Charney-Phillips grid (Charney and Phillips, 1953). In the Lorenz grid, a thermodynamic variable is stored at the same level as the horizontal wind velocity components and the vertical wind component is stored at intermediate levels (Figure 1.10 (a)), whereas in the Charney-Phillips grid a thermodynamic variable and the vertical wind velocity component are stored on the same level while the horizontal wind components are stored in intermediate levels (Figure 1.10 (b)).

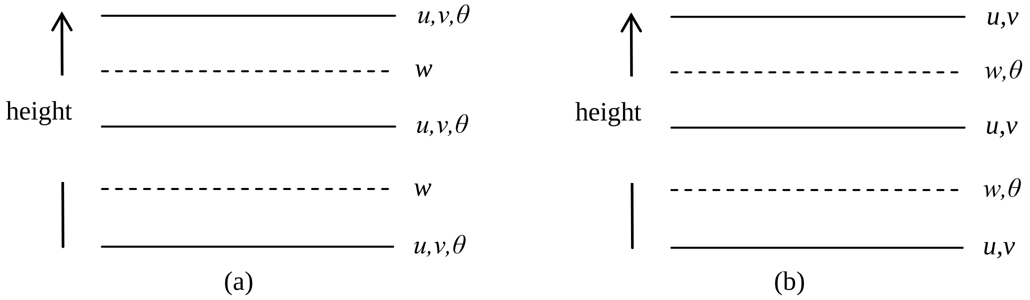


Figure 1.10: Schematic depiction of the (a) Lorent grid and (b) Charney-Phillips grid.

Most NWP models implement the Lorenz grid. Details concerning the advantages and disadvantages of the two mentioned vertical grid staggering methods can be found in e.g. Fox-Rabinovitz (1994). An illustration of vertical staggering in the WRF model is discussed later in Section 2.3.

The horizontal and vertical resolutions of an NWP model define its spatial resolution. Given that computing power is the main constraint in NWP simulations, the frequency at which computations are performed also matters a lot. This then defines a time resolution that is specified by a time step so that a model simulation proceeds in a series of distinct steps (defined by constant time steps). Each time step[1] specifies a time interval and it is within this interval that the model equations have to be solved before moving on to the next time step (Eumetcal, 2014). This then enables the NWP model to integrate forward in time.

Setting an appropriate time step is driven by the expected wind speeds and grid spacing in a simulation. For modeling efficiency, NWP models implement the so-called, time integration schemes that reduce the user-configurable features while ensuring model stability. The WRF model, for instance, uses the 3rd order Runge-Kutta time integration scheme (Skamarock, *et al.*, 2008). Other time integration schemes used in NWP models include the Leap-frog scheme and the Adam-Bashforth schemes (Durran, 1998; Smith R. , 2010).

1.3.4. NWP model simulation

A model simulation starts by defining the initial state of the atmosphere usually by the use of a forecast from a previous analysis (as well as local observations), in the so-called model initialization process. It therefore defines the initial and boundary conditions of the model by putting variables into each grid point. Normally, a limited-area model uses a global model to define the initial and boundary

---

[1] The time step is user-configurable. Setting an appropriate time step is driven by the expected wind speeds and grid spacing in a simulation.

conditions. Boundary conditions are meteorological conditions at the edges of the area of the model orography, including the lateral, top and bottom edges as illustrated in Figure 1.11 (MetEd, 2014).

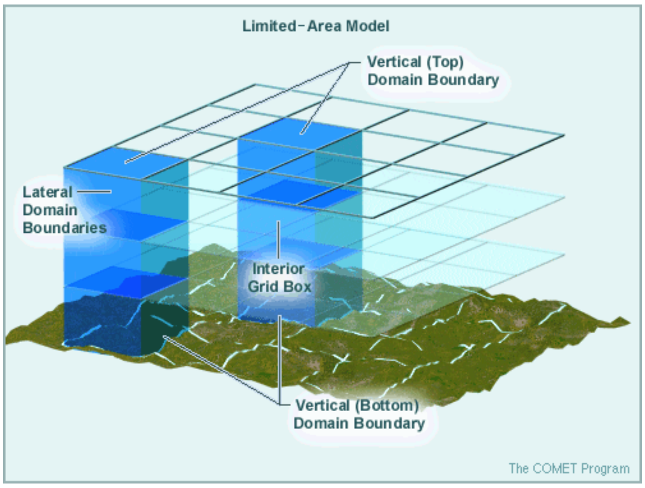


Figure 1.11: Lateral and vertical domain boundaries in a limited-area model (adapted from MetEd (2014)).

After initialization, model equations are then integrated forward in time by making some changes to the variables and resulting into forecast variables that are then used to predict the state of the atmosphere at a specified time in future. Thus, a forecast variable X at a given grid point is the sum of its initial value and the change in the variable (MetEd, 2014),

$$X^{forecast} = X^{initial} + \Delta X \tag{1.11}$$

Where,

$X^{forecast}$ ≡ forecast variable

$X^{initial}$ ≡ initial value of the forecast variable

ΔX ≡ change in the forecast variable and is given by,

$\Delta X = F(X) \cdot \Delta t$

Where,

$F(X)$ ≡ a combination of all kinds of forcing that can cause changes in X

Δt ≡change in time (how far into the future we are forecasting).

NWP is, therefore, an initial and boundary (surface and lateral) value problem and its accuracy is determined by, among other factors, the accuracy of the initial and boundary conditions. However, global models used in the initialization of regional models have coarser resolution compared to the regional models. Also local observations are never enough to adequately determine the initial and boundary conditions. Consequently, under-determination is a major problem in NWP and often leads to low quality forecasts (Schwitalla, 2012). To solve this problem, supplementary observations are used to improve and update the accuracy of the initial and boundary conditions in order to give the best estimate of the true atmosphere. This forms the basis of data assimilation (DA) process.

1.3.5. Assimilation of local observations including troposphere delays

Data assimilation (DA) is a process through which observations are incorporated into model initial fields (first guess) to provide an improved forecast (analysis) (Nichols, 2003). Traditionally, observations used in data assimilation processes have usually come from, among other sources, satellites, radar, upper-air soundings, surface meteorological stations and ship. Due to the water vapor information contained in GPS ZTD observations (as mentioned in Section 1.2), it’s now common practice to assimilate GPS ZTD (or even precipitable water vapor (PWV) derived from ZTD and surface meteorological variables) observations into NWP in order to adjust the vertical humidity profile (Bevis *et al.*, 1994; Christian *et al.*, 1995; Schwitalla, 2012).

Most of the observations used in data assimilation including ZTD observations, are indirect observation types in that they do not directly give prognostic variables. They are assimilated into NWP models using a variational method[1] that performs a least square fit by minimizing the “distance” between a model state (background) and the available observations (Parrish and Derbe, 1992; Gauthier *et al.,* 1999; Laing and Evans, 2011; Schwitalla, 2012). This amounts to what is referred to as a cost function and is further discussed in Section 2.6 in relation to ZTD.

Other than GPS ZTD observations, GPS derived PWV observations can also be assimilated into an NWP model. However, as pointed out by Elgered *et al.*(2005), Macpherson *et al.*(2008), Boniface *et al.*(2009) and Schwitalla (2012), the assimilation of ZTD is preferable in variational data assimilation systems over PWV. This is mainly related to the fact that variational data assimilation assumes unbiased errors which are more difficult to compute in the case of PWV because of the assumption of a mean temperature when it is being derived from ZTD.

The impact of assimilating ZTD observations from ground-based GNSS reference stations has been studied over the years (e.g. Cucurull *et al.*(2004), Macpherson *et al.*(2008), Yan *et al.*(2008)). Most of them were targeted towards assessing the impacts on short-range weather forecasts, especially the impacts on precipitation. Results from these studies (and many others) are varied from neutral (forecast errors unchanged) through modest (very small changes in forecast errors) to positive (forecast errors reduced) (e.g. Guerova (2003), TOUGH Project (2006), Macpherson *et al.*(2008)).

---

[1] This can be the 3D-Var or 4D-Var method. During a given time window all data points are analyzed once in 3D-Var while in 4D-Var there is a continuous cycle of data assimilation during the given time window.

1.3.6. Using NWP model data to estimate troposphere delays

Troposphere delay is a function of the meteorological parameters pressure, temperature and humidity. These parameters can be obtained from an NWP model grid (or from radiosonde data) and then used to estimate troposphere delay (a detailed discussion on this topic is presented in Section 2.1).

Studies by Schueler (2001), Jensen *et al.*(2002) and Cove (2002), showed that tropospheric delay can be estimated using NWP models with a higher accuracy than global troposphere delay models. The high accuracy may be attributed to, among other factors, the high temporal resolution in the NWP model data and the fact model output can be improved by the assimilation of additional observations and adjusting model configurations such as the resolution. Additionally, NWP model data can be used to estimate troposphere delays at almost every point of a given area. This, together with the high temporal resolution makes NWP models quite advantageous in the estimation of troposphere delays when compared to other methods such as the use of radiosonde observations that can only be obtained at sparsely located sites.

1.4. Ground-based sources of troposphere delay observations

Troposphere delay observations can be sourced from ground-based networks of continuously operating GNSS reference stations. There are several networks currently in existence all over the world. They are mainly classified based on their areal coverage. Such networks include the global International GNSS Service (IGS) and the regional EUREF (European REference Frame) Permanent GNSS Network (EPN).

By the year 2012, the IGS consisted of a network of 440 reference stations with a global coverage 84% of which were actively transmitting data (Dach and Jean, 2013). Most of the other GNSS networks are, however, regional or local with some of them doubling up as IGS reference stations.

The EPN is a well distributed and dense regional network of continuously operating GNSS reference stations that homogenously covers the European continent. It consisted of 244 reference stations by the end of 2012 with 34% of them also belonging to the IGS (Dach and Jean, 2013). Maps of IGS and EPN networks are shown in Appendix A.

The tropospheric zenith path delays at the EPN sites are processed at Local Analysis Centers (LACs). This is achieved by use of Bernese[1] GNSS software which implements troposphere delay models (including mapping functions) that derive their input (pressure, temperature and humidity) from a standard atmosphere model or from meteorological values sourced from radiosonde or meteorological sensors. EPN also makes available ZTD time series derived from radiosonde data of several stations in the vicinity of EPN stations.

---

[1] Bernese GNSS software is developed at the Astronomical Institute of the University of Bern. Other software used for the same purpose includes MicroCosm® (a registered trademark of Van Martin Systems, Inc) and Gypsy (developed at the Jet Propulsion Laboratory).

1.5. Research goal and objectives

Most of the subject matter briefly discussed in this introductory chapter has been (and continues to be) subjected to intensive studies as seen in e.g. Schueler (2001), TOUGH Project (2006), Haan (2008), and many others. However, due to the ever evolving research methods and increasing global atmospheric water vapor content, studies on these subjects have become an ongoing process. It was, therefore, the intension of the researcher to address certain questions in addition to the obvious ones relating to the expected findings, shortcomings and rationale of the research work. The following questions are addressed:

1. What will determine the accuracy of NWP-derived troposphere delays?
2. Which model for the weighted mean temperature is most suitable for the derivation of water vapor observations using GPS troposphere delay data and how does GPS-derived precipitable water vapor compare with similar observations derived using NWP and radiosonde?
3. What will determine meaningful impact of the assimilation of GPS total path delays into numerical weather prediction models?
4. What could be the troposphere delays in Kenya and what is the appropriate weighted mean temperature model for the region suitable for estimating water vapor using GPS troposphere delay data?

Based on the stated research questions, the following research objectives were formulated and hoped to be achieved at the end of the research:

i. To install and run systems for numerical weather prediction and data assimilation and all their related dependencies.
ii. To extract meteorological variables necessary for the estimation of troposphere delays from NWP model grids using methods that would be as accurate as possible and validate them using similar observations coming from meteorological sensors and radiosonde data.
iii. To estimate troposphere delays using the NWP-derived meteorological variables (from ii. above) and validate the results using troposphere delays sourced from ground-based GPS observations and those derived from radiosonde data.
iv. To analyze weighted mean temperature models.
v. To estimate and compare NWP and GPS precipitable water vapor observations.
vi. To assess the impact of assimilating GPS total zenith path delay observations on short-range NWP forecasts.

The main study evolved around a region in central Europe and an almost similar implementation in an African equatorial region (specifically, Kenya).

1.6. Structure of the thesis

The introductory chapter has given us an idea of why the study was conducted and the objectives that were hoped to be achieved at the end of the study. This thesis, therefore, summarizes the findings and conclusions made. It has been delivered in a somewhat simplified form that only includes what has been considered as essential and basic. Where deemed necessary, the reader is referred to the works of other authors and in some instances an explanation is presented in the appendices.

The entire research has been divided into a series of smaller studies. Chapter 2 presents a background study on some of the topics found relevant for this research. The concept of troposphere delay is discussed and relevant mathematical formulations presented from a theoretical point of view. This is a subject that has been extensively covered by many other researchers and therefore includes many important references. Precipitable water vapor (PWV) is also discussed. The weighted mean temperature forms a significant chunk of the study as it plays quite a vital and central role in the accuracy of PWV estimates. This chapter also presents a description of an NWP model grid with the view of modifying the mathematical formulations of troposphere delays and precipitable water vapor to make them suitable for application in NWP.

Chapter 3 presents the general methodology used in carrying out the research. The methodology presented here is very general with the specifics presented under each of the studies presented in Chapter 3, 4, 5, 6 and 7.

In Chapter 3, the accuracies obtained in NWP-derived meteorological variables are presented and discussed. These are surface or near-surface variables of pressure, temperature and relative humidity (or water vapor mixing ratio). They are then used to estimate troposphere delays that are then validated and the results presented in Chapter 4.

A study on precipitable water vapor is presented in Chapter 5. The importance of GPS PWV observations is discussed here through the comparison of NWP-derived PWV estimates with the GPS PWV estimates. The weighted mean temperature is studied in detail and analyzed in several ways including comparison with similar observations derived from radiosonde data.

The water vapor information contained in the wet component of the troposphere delay is expected to affect the errors in short-range forecasts of certain parameters by reducing, increasing or having no effect on them. The impact of assimilating total zenith path delay observations into NWP is therefore studied in Chapter 7.

There are challenges that would be associated with carrying out this research on a region encompassing the African equatorial including the lack of sufficient data. The researcher, however, comes from this region and Chapter 7 is, therefore, meant to make this research relevant to him.

Additionally, it is hoped that the study will serve as a basis upon which similar studies will be formed in future.

Each study is summarized at the end of its respective chapter. An overall conclusion and relevant recommendations are given in one of the closing section of the thesis.

1.7. Summary

The main sources of the errors encountered in GPS measurements have been briefly discussed and the atmospheric errors have been seen to be the most dominant. Tropospheric delay has been seen to be difficult to model and mitigate due to its dependence on the highly variable atmospheric water vapor.

Techniques used to mitigate tropospheric delay have been briefly discussed with a greater focus on numerical weather prediction models. This discussion has concentrated on the use of numerical weather prediction models not only for estimating troposphere delays but also in regards to the assimilation of troposphere delay observations in order to improve weather forecasts.

Troposphere delays to be assimilated into a numerical weather prediction model should somehow come from somewhere. A brief mention of ground-based sources of GPS troposphere delay observations has therefore been presented.

The broad aim for carrying out the research has been mentioned together with the research objectives that were then laid out. The chapter ends with a discussion of the order in which this thesis is presented.

2. Background study

A background study on the subject matter was undertaken and is presented in this chapter thereby providing a general idea of the main building blocks constituting the entire study.

The applicability of geometrical optics to the study of GPS signal propagation is often assumed. This together with an understanding of atmospheric dynamics forms the basis of the concepts of troposphere delay. This section therefore discusses the basic concepts of troposphere delay and the supporting mathematical formulations that have led to its quantitative description.

The scope of this study does not allow us to delve deeply into the numerical methods involved in NWP. However, some basic formulations are discussed including; estimation of troposphere delay and precipitable water vapor using NWP-derived parameters, troposphere delay and precipitable water vapor formulation with regards to data assimilation.

Understanding, interpreting, and correctly extracting the necessary variables from NWP model output files requires an understanding of the 3D grid of the NWP model itself. This can contribute greatly to the accuracy and precision of NWP-estimated troposphere delays and precipitable water vapor. A description of the structure of the ARW model, therefore, forms part of this chapter.

Following a descriptive study of the ARW model’s structure, a method is devised to numerically integrate the dry and wet refractivities now expressed in terms of NWP model variables and hence mathematical formulations for zenith hydrostatic delay and zenith wet delay in terms of NWP-derived parameters. These parameters henceforth also referred to as variables, can be directly or indirectly (after derivation from other variables) extracted from NWP model output files.

2.1. Troposphere delay

The troposphere is mainly made up of dry gases, hydrometeors and particulate matter (Léna *et al.*, 2012). Consequently, a GPS signal propagating through it and impinging upon its constituents is likely to reduce in speed and strength as it undergoes some of the mechanisms associated with an electromagnetic signal’s interaction with matter, notably: refraction, reflection and diffraction. The combined effect of these mechanisms lead to a scattered radio signal which not only reduces its strength at the receiver (i.e. the signal is attenuated), but also alters its original propagation path (i.e. the signal is refracted). Tropospheric attenuation is, however, relatively small compared to free space loss. For instance, tropospheric attenuation on the GPS L1 signal is found to be approximately 0.4 dB while free space loss on the same signal is found to be approximately 185.4 dB (Hein, *et al.*, 2007).

The ratio of the speed of a radio signal in free space to its speed as it propagates through a given medium has a direct implication on the amount of departure from the original propagation path (refraction) to be undergone by the signal. This is expressed as an index, the refractive index.

The nature (chemical composition and physical structure) and distribution of tropospheric constituents vary greatly in both space and time resulting into a heterogeneous atmosphere with varying refractive indices. Assuming a horizontally stratified atmosphere the refractive index is assumed uniform horizontally and variable vertically. The vertically varying refractive index causes longer signal paths (as the signal is refracted) and greater transit times[1] (as the signal speed is reduced).

Deviation from the original path (ideally supposed to be the geometric path) leads to an excess path length. This causes range errors in GPS positioning applications hence compromising GPS positioning accuracy.

Principles of optics are applied in the assessment GPS radio signals because of their proximity to light signals in the electromagnetic spectrum. Therefore, following Fermat's principle[2], the path traversed by a GPS radio signal as it propagates through space is the path that takes the least time to get to the receiver (as depicted in Figure 2.1).

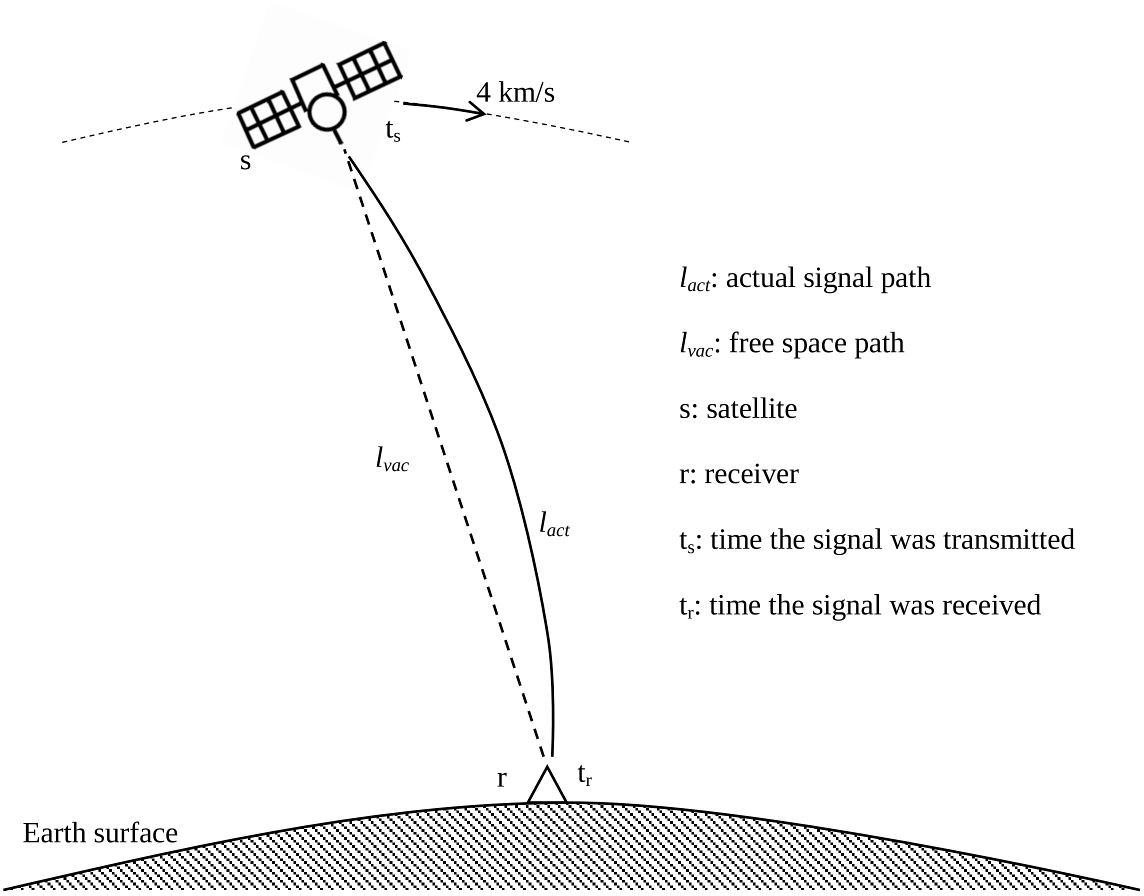


Figure 2.1: Depiction of the ideal and actual satellite-to-receiver path of a GPS signal.

---

[1] Transmit time is the difference between the time the signal was received (as determined by the receiver clock) and the time the signal was transmitted (as marked on the signal). The expected transmit time varies depending on the position of the satellite in space with respect to that of a receiver on the ground.

[2] The path taken by a ray of light between two points is such that there are many other paths nearby which take almost exactly the same time.

In Figure 2.1, the actual path that is likely to be taken by a GPS radio signal as it propagates through space from a satellite to a receiver is depicted by the curved path denoted by l_{act}. The path that would otherwise be taken by the signal if the entire path was a vacuum is denoted by the straight discontinuous line l_{vac}.

Because of the signal's deviation from the path that it would ideally take, it ends up covering an excess path length. It can be shown that this excess path length approximates to,

$$\Delta l = l_{\mathrm{act}} - l_{\mathrm{vac}} = \int_{l_{act}} n \cdot dl_{act} - \int_{l_{vac}} dl_{vac} \tag{2.1}$$

Where, the integration is along the signal path, and

n ≡ refractive index of a medium encountered in the signal path

dl_{act} and dl_{vac} ≡ an infinitesimal displacement along the actual and ideal signal paths respectively.

Equation (2.1) shows that the deviation of a GPS signal's path from an ideal path is caused by the continuous variation of the refractive index. Following Mendes (1999) this equation may as well be expressed in the form,

$$\Delta l = \int_{l_{act}} n \cdot dl_{act} - \int_{l_{act}} dl_{act} + \left[\int_{l_{act}} dl_{act} - \int_{l_{vac}} dl_{vac} \right] \tag{2.2}$$

For a spherically symmetrical atmosphere, the bracketed term becomes zero if the propagation paths are taken in the zenith direction. This is because the two paths would be identical. Consequently, given the zenith considerations, the excess path length amounts to,

$$\Delta l_z = \int_z (n-1) \cdot dz \tag{2.3}$$

Where, the subscript z denotes a zenith path, and dz is an infinitesimal displacement along the signal path in the zenith direction.

The signal would, therefore, be delayed (in terms of arriving to the receiver) when the actual path length is compared to the ideal path length. The excess path length with zenith considerations is therefore termed as the zenith path delay.

Since the numerical value of the refractive index is only slightly larger than one, using it in its usual form makes it relatively difficult to describe its spatial and temporal variations. Consequently, it is expressed in terms of the atmospheric refractivity, N^1, the difference of n from unity in parts per million. Thus, (Thayer, 1961; Boumis *et al.*, 2002; Agnew, 2002),

$$N = (n-1)10^6 \tag{2.4}$$

The spatial variation of N (and hence n) over the Wettzell geodetic observatory is illustrated in Figure 2.2. The atmospheric refractivity profile is based on an NWP model simulation for 29th May 2013 18 UTC. It is seen to decrease exponentially with height with a value of approximately 280 N-units near the surface.

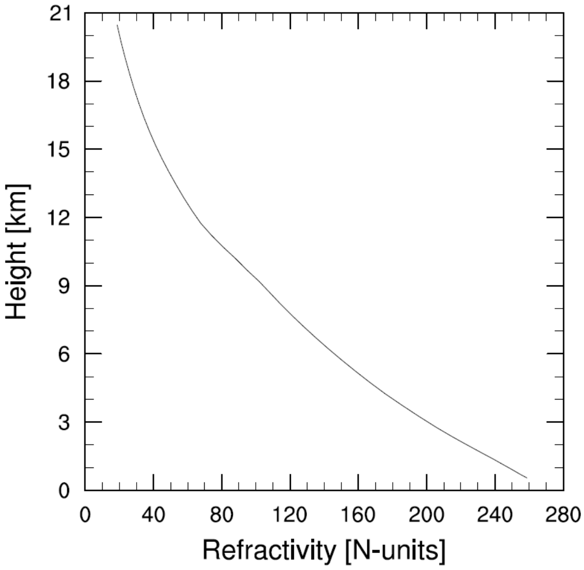


Figure 2.2: Atmospheric refractivity profile over the Wettzell geodetic observatory on 29th May 2013 18 UTC. The refractivity, N, is expressed in N-units where $N = (\text{n}-1)\cdot 10^6$.

Taking the integral in Equation (2.3) from the GPS antenna height to a point in infinity where the effects of the atmosphere on the signal can be assumed to be zero, the total zenith path delay (henceforth, ZTD), may therefore be expressed as,

$$ZTD = 10^{-6}\int_h^{\infty} N \cdot dh \tag{2.5}$$

Where, $h \equiv$ GPS antenna height.

Atmospheric refractivity is a function of pressure, temperature and humidity and following Essen and Froome (1951); Smith and Weintraub (1953); Boudouris (1963); Hasegawa and Stokeberry (1975); Davis *et al.*(1985); Mendes (1999) and Schueler (2001), it is given by the three-term formula,

---

[1] N is dimensionless but is commonly referred to in N-units.

$$N = k_1 \cdot \frac{P_d}{T} \cdot Z_d^{-1} + k_2 \cdot \frac{e}{T} \cdot Z_w^{-1} + k_3 \cdot \frac{e}{T^2} \cdot Z_w^{-1} \tag{2.6}$$

Where, the first term represents the sum of the effects of the induced dipole moments of dry gas molecules, the second term represents the sum of the effects of the induced dipole moments of water vapor, and the third term represents the dipole orientation effects of the permanent dipole moment of water vapor molecules, and

P_d ≡ pressure of dry air in hectopascals (hPa)

e ≡ partial pressure of water vapor in hectopascals (hPa)

T ≡ absolute temperature in Kelvin (K)

Z_d^{-1}, Z_w^{-1} ≡ inverse compressibility factors of dry gases and water vapor respectively to cater for deviation from the ideal gas law

$k_{1,\ldots,3}$ ≡ refraction constants.

The refraction constants are based on experimental values. Over the years different experimental approaches have been applied to estimate refraction constants with varying accuracies. Based on these results, researchers have computed mean values of refraction constants and their associated uncertainties for general usage also with varying accuracies as seen in Davis *et al.*(1985); Bevis *et al.*(1994); Mendes (1999) and Rüeger (2002). The weighted mean values of the refraction constants and their associated uncertainties (standard deviations) as given by Rüeger (2002) are $k_1 = 77.6848 \pm 0.0094\,\mathrm{K/hPa}$, $k_2 = 71.2152 \pm 1.3\,\mathrm{K/hPa}$ and $k_3 = 375463 \pm 0.76\,\mathrm{K^2/hPa}$.

The atmospheric refractivity can be expressed as a sum of a hydrostatic and wet refractivity. Thus, following Schueler (2001),

$$N = N_{hyd} + N_{wet} = k_1 \cdot R_d \cdot \rho + \left[k_2' \cdot \frac{e}{T} + k_3 \cdot \frac{e}{T^2} \right] \cdot Z_w^{-1} \tag{2.7}$$

Where,

N_{hyd} ≡ hydrostatic refractivity, $N_{\mathrm{hyd}} = k_1 \cdot R_d \cdot \rho$

Where,

k_1 ≡ refraction constant (= 77.6848±0.0094 K/hPa)

R_d ≡ gas constant of dry air (= 287.054 $\mathrm{JK^{-1}\,kg^{-1}}$)

ρ ≡ total density of air in $\mathrm{kgm^{-3}}$.

N_{wet} ≡ wet refractivity, $N_{wet} = \left[k_2' \cdot \frac{e}{T} + k_3 \cdot \frac{e}{T^2} \right] \cdot Z_w^{-1}$

Where,

$k_2' = k_2 - k_1 \cdot \frac{M_w}{M_d}$

k_2 ≡ refraction constant (=71.2152±1.3 K/hPa), and k_1 is as earlier defined

M_w and M_d ≡ molar weights of water vapor (= 18.0152kg kmol^{-1}) and dry air (= 28.9644 kg kmol^{-1}) respectively

T, e and Z_w^{-1} are as earlier defined in Equation (2.6) and the partial pressure of water vapor is related to the density of water vapor by the ideal gas law. Thus,

$$e = \rho_w \cdot R_w \cdot T \qquad (2.8)$$

The total zenith path delay is, therefore, expressed as,

$$ZTD = 10^{-6} \int_h^{\infty} \left(k_1 \cdot R_d \cdot \rho \right) \cdot dh + 10^{-6} \int_h^{\infty} \left[k_2' \cdot \frac{e}{T} + k_3 \cdot \frac{e}{T^2} \right] \cdot Z_w^{-1} \cdot dh \qquad (2.9)$$

Where, the first term represents the zenith hydrostatic delay (ZHD) while the second term represents the zenith wet delay (ZWD).

Derivation of the compressibility for water vapor, Z_w^{-1}, can be found in Owens (1967). However, in a study by Mendes (1999), its omission changed the ZWD by only 0.1 to 0.2 mm. The approximation $Z_w^{-1} \approx 1$ is, therefore, found to be sufficient in terms of accuracy in most cases (Schueler, 2001).

Assuming the atmosphere to be in hydrostatic equilibrium, the vertical coordinate can be transformed from height to pressure. Therefore, it can be shown that ZTD can be expressed as,

$$ZTD = -10^{-6} \int_{p_a}^{0} k_1 \cdot R_d \cdot \frac{1}{g} \cdot dp \quad + \quad -10^{-6} \int_{p_a}^{0} \frac{1}{\rho \cdot g} \cdot \left(k_2' + \frac{k_3}{T} \right) \cdot \frac{e}{T} \cdot dp \qquad (2.10)$$

Where, Z_w^{-1} has been assumed to be approximately 1, and

p_a ≡ atmospheric pressure at GPS antenna height, pressure at infinity is taken to be zero

dp ≡ differential change in pressure, $dp = -\rho \cdot g \cdot dh$

Where,

dh ≡ differential change in height

g ≡ gravity acceleration, a function of height and latitude

ρ ≡ total density of air

k_1, k_2', k_3 are as earlier defined.

By using the standard values of surface pressure (= 1013.25 hPa) and gravity acceleration (= 9.80665 m/s^2) and taking k_1 = 77.6848 K/hPa and R_d = 287.054 $JK^{-1}kg^{-1}$, ZHD is found to be approximately 2.3 m (approximately 8 ns delay). This is the typical ZHD value at sea level and can reach approximately 23 m at 5° elevation (Rocken *et al.*, 2001). Hydrostatic delay can, therefore, be easily modeled and estimated by use of surface measurements.

ZWD mainly depends on water vapor which in turn happens to be strongly dependent on air temperature. For this reason the greatest water vapor concentrations are found in the tropics and decrease towards the Polar Regions. Additionally, since air temperature decreases with altitude (by approximately 6.5°C/km on average), larger amounts of water vapor concentration are found near the Earth's surface than higher up in the atmosphere of which approximately 99% of the total amount is found in the troposphere. According to Bevis *et al.* (1992) and Zhenhong (2005), ZWD typically varies from a few cm to about 0.2 m (approximately 0.7 ns delay) during the year at mid-latitudes and to about 0.4 m (approximately 1 ns delay) in tropical and humid regions. The spatial and temporal variability of atmospheric water vapor makes it difficult to consistently measure and model the ZWD.

Assuming a ZTD of approximately 2.6 m (sum of ZHD and ZWD), therefore, hydrostatic delays typically account for approximately 90% of the total delay with the remaining 10% being wet delays. Though the percentage contribution of wet delays to the total troposphere delay is quite small when compared to that of hydrostatic delays, it is difficult to accurately model and estimate due to its dependence on water vapor. Hence, wet delays present a significant problem in the accuracy of GPS positioning and thus cannot be ignored.

2.2. Precipitable water vapor and weighted mean temperature

Integrated water vapor (IWV) is the integral of water vapor density between an antenna height *h* and a point at infinity. Thus,

$$IWV = \int_h^\infty \rho_w \cdot dh \tag{2.11}$$

Where, ρ_w ≡ density of water vapor in kgm^{-3}. IWV is therefore in kgm^{-2}.

The amount of atmospheric water vapor is also described using precipitable water vapor (PWV), which is IWV weighted by the density of liquid water. It is equivalent to the amount of water that would be contained in a vertical column of unit cross section extending from an antenna to infinity. Thus,

$$PWV = \frac{IWV}{\rho_l} = \frac{1}{\rho_l}\int_h^\infty \rho_w \cdot dh \tag{2.12}$$

Where, ρ_l ≡ density of liquid water (= 1000 kgm^{-3}).

Studies have shown that PWV can be derived from the ZWD (Davis *et al.*, 1985; Askne and Nordius, 1987; Bevis *et al.*, 1994; Schueler, 2001; Schueler *et al.*, 2002). Thus, the relationship between PWV and ZWD can be expressed as,

$$PWV = \frac{ZWD}{10^{-6} \cdot R_w \cdot \rho_l \cdot \left[k_2' + \frac{k_3}{T_m}\right]} \tag{2.13}$$

Where, T_m ≡ the weighted mean temperature of the atmosphere in kelvin, and all the other terms are as earlier defined.

T_m is a water-vapor weighted vertically averaged temperature of the atmosphere. It is given by Davis *et al.* (1985) and Bevis *et al.* (1994) as,

$$T_m = \frac{\int_h^\infty \frac{e}{T} \cdot dh}{\int_h^\infty \frac{e}{T^2} \cdot dh} \tag{2.14}$$

Where, e is partial pressure of water vapor in hPa, given by,

$$e = \rho_w \cdot R_w \cdot T \tag{2.15}$$

Where,

ρ_w ≡ density of water vapor in kgm^{-3}

R_w ≡ gas constant of water vapor (= 461.524 Jkg^{-1}K^{-1})

T ≡ absolute temperature in K.

T_m, therefore, mainly depends on the vertical profiles of temperature and humidity. This makes it a key parameter in the estimation of PWV given ZWD observations such that together with the accuracy of the ZWD observations, its accuracy greatly determines the accuracy of the PWV estimates.

ZWD can be obtained from ZTD observations by using a hydrostatic delay model and surface pressure observations to separate ZHD from ZTD (i.e. ZWD = ZTD - ZHD). T_m, on the other hand, can be estimated by numerically integrating vertical temperature and humidity profiles (based on Equation (2.14)) obtained from radiosonde (or NWP models) or it can as well be taken to be a constant value. The latter approach, however, is likely to lead to large errors in the estimation of PWV since several studies have shown that T_m varies diurnally, seasonally and geographically (Askne and Nordius, 1987; Bevis *et al.*, 1992; Emardson and Derks, 2000; Wang *et al.*, 2005).

Using the T_m estimates at radiosonde sites (or grid points in case of NWP) and the corresponding surface temperature T_s observations, a linear relationship between T_m and T_s is established in a regression analysis method. This relationship, the so-called T_m-T_s relation, leads to the development of linear surface temperature models for T_m that can then be used to estimate T_m at any point within the region covered. T_m models therefore are mainly site-specific and regional in nature with the site-specific models being more superior (Ross and Rosenfeld, 1997).

T_m models developed using radiosonde data are seen in e.g. Bevis *et al.*(1992), Ross and Rosenfeld (1997), Emardson and Derks (2000) and Singh *et al.* (2013) for the United States (US), several global sites, Europe and India respectively. The use NWP data offers a greater advantage over radiosonde data in terms spatial and temporal availability and therefore leads to greater accuracy in the T_m (Wang, *et al.*, 2005).

The Bevis T_m-T_s relation has come to be the most widely accepted model for global use by the scientific community. It was developed following a study that used almost 9000 radiosonde profiles from 13 sites in the United States to derive a relationship between T_m and surface temperature T_s which was found to be largely linear and is expressed as,

$$T_m = 70.2 + 0.72 \cdot T_s \tag{2.16}$$

Where, $T_s \equiv$ the surface temperature in K.

The site-/regional-specific T_m models mainly vary in the model coefficients resulting from relating the T_m estimates to T_s. In Schueler *et al.* (2002) a linear surface temperature model is discussed where the functioncoefficients represent a global mean. Thus,

$$T_m = 86.9 + 0.647 \cdot T_s \tag{2.17}$$

Where, T_s is surface temperature in K.

Additionally, in Schueler *et al.* (2002), a model for T_m is discussed that combines together the linear surface temperature model and a periodic model that takes into account diurnal and seasonal variations. The resulting mixed linear surface temperature/periodic model is thus,

$$T_m = \bar{T}_m + \tilde{T}_m \cdot \cos\left[2\pi \cdot \frac{DoY - DoY_w}{365.25[days]}\right] + q_T \cdot T_s \tag{2.18}$$

Where,

T_m ≡ mean tropospheric temperature (in K) at the requested day of year (DoY)

$\bar{T}_m$ ≡ average mean tropospheric temperature in K

$\tilde{T}_m$ ≡ amplitude of mean temperature in K

DoY_w ≡ day of “maximum winter” (28 days for northern and 211 days for southern hemisphere)

q_T ≡ temperature amplifier

T_s ≡ surface temperature in K.

The T_m model of Equation (2.18) was found to overcome some limitations observed in each of the individual models used. The coefficients of this function are derived using data (such as mean tropospheric temperature and surface temperature) coming from the GDAS (Global Data Assimilation System) weather fields on a 1°×1° global grid resulting into a database refered to as TropGrid (and more recently TropGrid2 as discussed in Schueler (2014)). Site-specific coefficients are also computed for most IGS and EPN sites resulting into a database refered to as TropSite (Schueler *et al.*, 2002).

Using over 120,000 radiosonde profiles from 38 different sites in Europe, Emardson and Derks (2000) estimated ZWD and PWV at the sites and developed a formula that is applicable over Europe and that can be used to estimate PWV from ground-based ZWD observations without necessarily having T_m data (Haase *et al.*, 2003). This may be seen to be particularly important owing to the sensitivity in the accuracy of T_m. The formula is a polynomial function given as (Emardson and Derks, 2000),

$$\frac{ZWD}{PWV} = a_0 + a_1 \cdot \Delta T + a_2 \cdot \Delta T^2 \tag{2.19}$$

Where,

a_0= 6.458 m^3kg^{-1}, a_1 = -1.78×10^{-2} $m^3kg^{-1}K^{-2}$, a_2 = -2.2×10^{-5} $m^3kg^{-1}K^{-2}$

$\Delta T = T_s - T_{avg}$

Where, T_s and T_{avg} ≡ surface temperature (in K) and mean surface temperature (= 283.49 K) respectively. T_s is obtained from the radiosonde data.

The coefficients a_0, a_1 and $a_2$are considered valid for most parts of the European continent (Schueler *et al.*, 2002).

2.3. Advanced Research WRF (ARW) model grid

The Advanced Research WRF (ARW) is a hydrostatic NWP model, one of the two dynamical cores of the Weather Research and Forecasting (WRF) system, the other one being the Non-hydrostatic Mesoscale Model (NMM). It is the main model used in this study and hence a discussion of a few of its main features is presented here. A detailed description can be found in Skamarock *et al.*(2008).

A description of the structure of the ARW model grid is seen as a prerequisite for understanding and correctly extracting the necessary variables contained in the model's output files. The model implements a terrain-following hybrid hydrostatic pressure vertical coordinate with a constant pressure surface at the top (Figure 2.3). According to Skamarock *et al.*(2008) and Boone (2013), this vertical coordinate, η, is defined as a function of the surface pressure (p_s) and atmospheric pressure (p) and it decreases monotonically such that $\eta = 1$ at the surface and $\eta = 0$ at the top of the modeled atmosphere. Thus, if p_t is pressure at the model top,

$$\begin{aligned} p(\eta = 1) &= p_s \\ p(\eta = 0) &= p_t \quad (p_t \geq 0) \end{aligned} \tag{2.20}$$

And η is therefore expressed as,

$$\eta = \frac{p_h - p_{ht}}{p_{hs} - p_{ht}} \tag{2.21}$$

Where, p_h ≡ the hydrostatic component[1] of the pressure (as denoted by the subscript h) that varies between the value on the surface (i.e. p_{hs}) and the value at the top (i.e. p_{ht} , a constant value, normally 50 hPa). The denominator therefore is proportional to the mass of the air in a vertical column of the model domain. From Equation (2.21), we see that the ARW vertical coordinate system is essentially the sigma system definition (as was discussed in Section 1.3.3) implemented in a hydrostatic NWP model.

Grid staggering in the ARW model is the Arakawa C-grid staggering for variables in the horizontal direction and the grids in the vertical direction are staggered with prognostic variables at the center of pressure layers and the vertical wind component at the boundary of the layers.

---

[1] As opposed to a non-hydrostatic model.

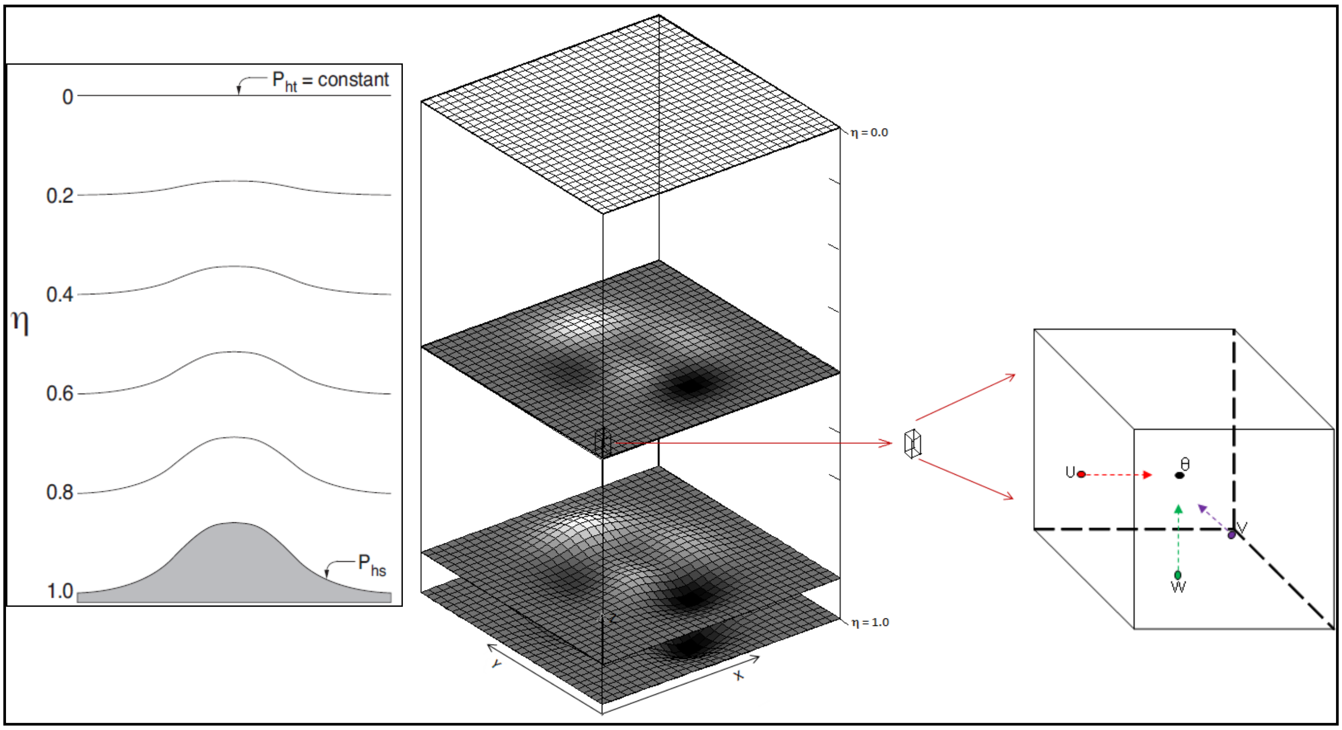


Figure 2.3: Schematic of the hydrostatic pressure vertical coordinate of the ARW model (adapted from Skamarock *et al.* (2008)). Diagram on the right hand side provides a three-dimensional overview.

Variable points in a grid box within the WRF model grid are as illustrated in Figure 2.4. Except the geopotential height, all scalar variables (e.g. temperature, pressure and humidity) are stored at the so-called theta points that are centered in each grid box. The value of a variable at this point represents the average conditions throughout the grid box. The vector variables (wind velocity components) are stored at the u, v and w points located at the centers of the grid faces and directed inwards where the u and v points are in adjacent side faces and directed in the west-east and north-south directions respectively, while the w point is on the bottom face and directed upwards.

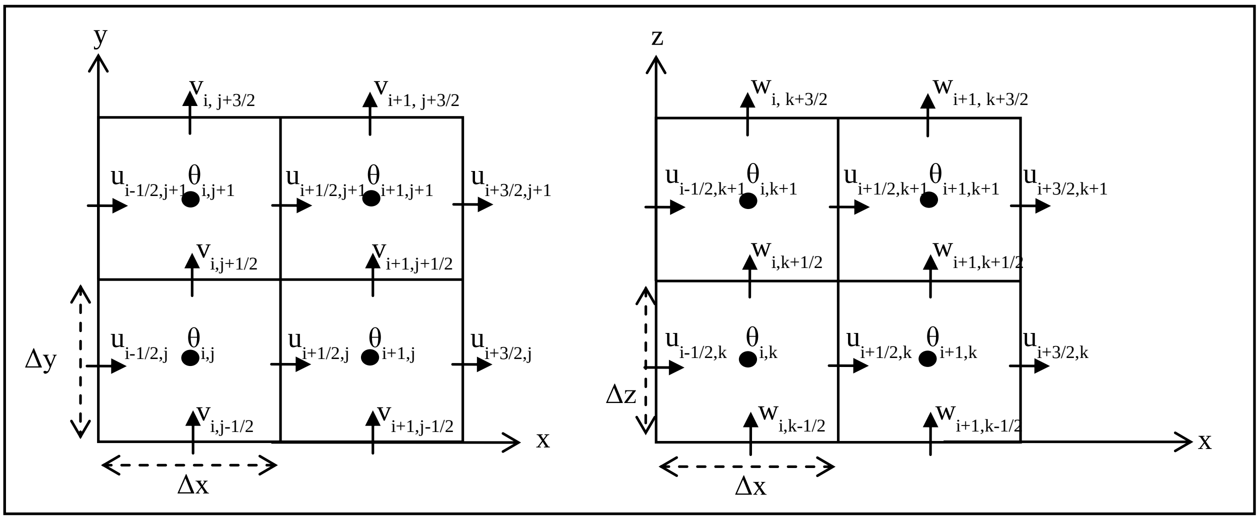


Figure 2.4: Horizontal (left hand side) and vertical (right hand side) grid staggering in the ARW model (adapted from Skamarock *et al.* (2008)).

Each point is indexed with respect its position in space (three-dimensional) in terms of horizontal and vertical positions. As illustrated in Figure 2.4, the position of a point in the west-east direction (x direction) is indexed by i, that of a point in the north-south direction (y direction) is indexed by j, and that of a point in the z direction indexed by k. The horizontal wind velocity components u and v are therefore staggered in the x and y directions respectively while the wind velocity component w and the geopotential height are both staggered in the z direction. The theta points are unstaggered.

A structure of the ARW model vertical grid is illustrated in Figure 2.5. The vertical levels (horizontal lines) represent the so-called full levels (solid lines denoted by p_j^f for the j^{th} full level) and half levels (dashed lines denoted by p_j^h for the j^{th} half level). n is the number of full levels[1]. Two full levels enclose a pressure layer and they are not necessarily evenly spaced.

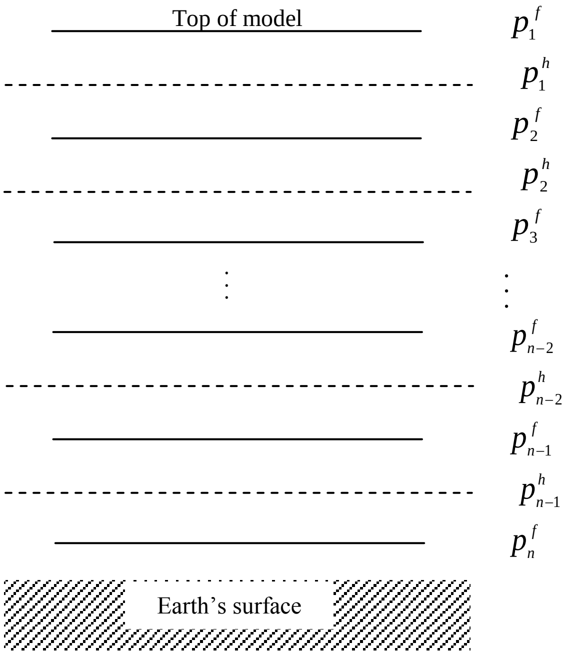


Figure 2.5: NWP model orography description. Solid lines represent full levels while the dashed lines represent the half levels.

The thermodynamic variables (at the theta points) are found along the half levels except the geopotential height that is defined on full levels. Pressure is available on both full and half levels. Pressure at the top of the model is fixed at 50 hPa and hence the model extends to a vertical height of approximately 20 km (more or less).

Time discretization (or time-splitting) in the ARW model is achieved by use of the 3rd order Runge-Kutta time integration scheme (first mentioned in Section 1.3.3).

---

[1] The number of vertical levels is usually user configurable. The levels are also referred to as mass levels. Hence, full mass levels and half mass levels.

Main prognostic variables include all the three wind velocity components, perturbation potential temperature, perturbation geopotential, perturbation surface pressure of dry air, and water vapor mixing ratio.

2.4. Troposphere delay formulation using ARW NWP model variables

Although the horizontal coordinate of the ARW model mimics the local terrain, it is highly unlikely that the antenna coincides exactly with a model level at all locations. Consequently, situations arise where the antenna height could be below the model orography (Figure 2.6 (a)) or it could as well be falling somewhere within the model orography (Figure 2.6 (b)).

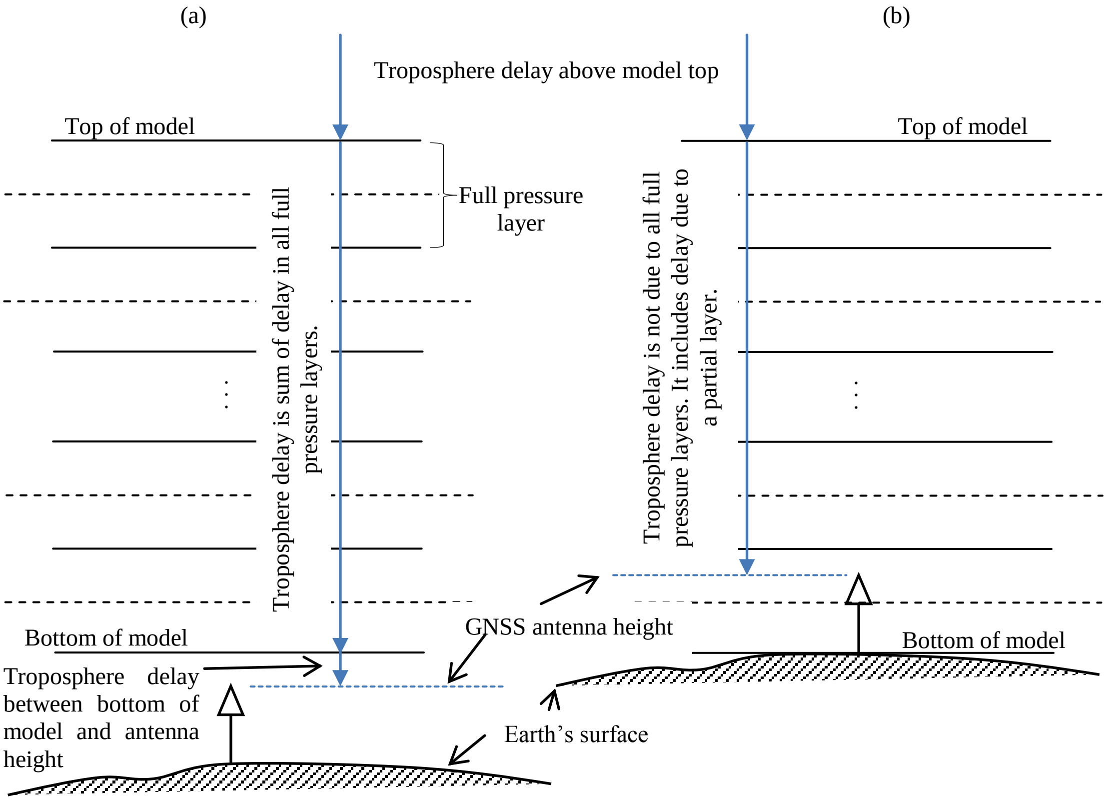


Figure 2.6: Possible antenna/model orography vertical position scenarios and the corresponding troposphere delay computations, (a) antenna height below the model bottom height, (b) antenna height falling somewhere within the model orography.

Given that the model extends to a vertical height of approximately 20 km some portion of the upper atmosphere is not covered by the model orography. This is depicted in both scenarios in Figure 2.6. Additionally, in the scenario depicted in Figure 2.6 (a), a portion of the lower atmosphere is not covered by the model orography because the antenna is below the orography. On the other hand, as depicted in Figure 2.6 (b), the antenna is within the model orography. This calls for caution in computing troposphere delays in both scenarios so that the method used tries to account for the delays due to the portions of the atmosphere not covered by the model orography.

If the antenna is below the model orography, troposphere delay is a sum of the delay above the model top, the delay within the entire model orography and the delay due to a portion that appears between the antenna height and the height of the model bottom (Figure 2.6 (a)). The delay within the entire model orography in this case, therefore, is due to all the model pressure layers.

If the antenna falls within model orography, troposphere delay is a sum of the delay above the model top and the delay due to the portion bounded by the pressure layer at the model top and the pressure layers just above the antenna height (Figure 2.6 (b)). The latter delay term, therefore, includes a delay due to a partial pressure layer between the antenna height and the half mass level just above it.

In formulating troposphere delays using NWP model variables, reference is made to the following formulations of ZHD and ZWD (Equation (2.10)),

$$ZHD = -k_1 \cdot R_d \cdot 10^{-6} \int_{p_a}^{0} \frac{1}{g} \cdot dp \tag{2.22}$$

$$ZWD = -10^{-6} \int_{p_a}^{0} \frac{1}{\rho \cdot g} \cdot \left(k_2' + \frac{k_3}{T} \right) \cdot \frac{e}{T} \cdot dp \tag{2.23}$$

2.4.1. Troposphere delays above NWP model orography

The portion of the atmosphere above the NWP model orography may contribute to the magnitude of the total delay be it in a small way. The overall effect of this portion may differ from one NWP model to another depending on the maximum vertical extent of the model. Probably that’s why it has been previously largely ignored as seen in Haan (2008) and Jones (2010). In the case of the ARW model, however, the pressure value at the top is a constant value of 50 hPa at a height of approximately 20 km. This pressure at model top translates to a ZHD of approximately 0.1 m.

The integrals of Equation (2.22) and (2.23) have, therefore, been evaluated from an undefined point in the atmosphere to the NWP model’s topmost level, yielding to Equations (2.24) and (2.25) for the ZHD and ZWD respectively,

$$ZHD_{upper} \approx 10^{-6} \cdot k_1 \cdot R_d \cdot \frac{1}{g_1} \cdot p_1 \tag{2.24}$$

$$ZWD_{upper} \approx 10^{-6} \cdot \frac{R}{M_w} \cdot \frac{1}{g_1} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_1} \right) \cdot \frac{r_1}{r_1 + 1} \cdot p_1 \tag{2.25}$$

Where,

$\varepsilon \equiv$ ratio of the molar weight of water vapor to that of dry air, i.e.

$$\varepsilon = \frac{M_w}{M_d} \tag{2.26}$$

p_1, T_1, r_1 and g_1 ≡pressure, temperature, water vapor mixing ratio and the local gravity acceleration respectively at the topmost model level, and all the other terms are as earlier defined.

2.4.2. Troposphere delays within NWP model orography

We now evaluate the delays due to the portion of the atmosphere covered by the NWP model orography. Evaluating the integrals of Equation (2.22) and (2.23) should start from the topmost NWP model level to the GPS antenna height if it falls within model orography or to the lowermost model level if the antenna does not fall within model orography.

Assuming a linear variation of variables within the model layers, Equation (2.22) can be approximated to,

$$ZHD_{within} \approx 10^{-6} \cdot k_1 \cdot R_d \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \Delta p^h \tag{2.27}$$

Where,

$g_j \equiv$ local gravity acceleration at the j^{th} half-level

$\Delta p^h = p_{j+1}^h - p_j^h \equiv$ difference in pressure at adjacent half-levels $j+1$ and j respectively

m ≡ the number of model half mass levels above the GPS antenna height. If $m = n$ (n is the total number of half mass levels), the antenna height is below the model orography. Otherwise, if $m < n$, the antenna height falls within the model orography.

Moving on to the ZWD, we start by making use of the equation for partial pressure of water vapor (Equation (2.8)) and that of the molar weight ratio between water vapor and dry air (Equation (2.26)) to re-write Equation(2.23). Thus,

$$\begin{aligned} ZWD &= 10^{-6} \int_0^{p_a} \frac{1}{g} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T} \right) \cdot \frac{\rho_w \cdot R_w \cdot T}{\rho} \cdot \frac{1}{T} \cdot dp \\ &= 10^{-6} \int_0^{p_a} \frac{1}{g} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T} \right) \cdot \frac{\rho_w}{\rho} \cdot R_w \cdot dp \end{aligned} \tag{2.28}$$

We make further definitions as follows:

- Specific humidity and water vapor mixing ratio,

$$q = \frac{\rho_w}{\rho} \tag{2.29}$$

$$r = \frac{\rho_w}{\rho_d} \tag{2.30}$$

Where,

q and r ≡ specific humidity and water vapor mixing ratio respectively

ρ ≡ total air density, given by $\rho = \rho_d + \rho_w$, ρ_d and ρ_w ≡ densities of dry air and water vapor respectively

We also express q in terms of r as follows,

$$q = \frac{\rho_w}{\rho} = \frac{\rho_w}{\rho_w + \rho_d} = \frac{\rho_w / \rho_d}{\rho_w / \rho_d + \rho_d / \rho_d} = \frac{r}{r+1} \tag{2.31}$$

- The gas constant for water vapor,

$$R_w = \frac{R}{M_w} \tag{2.32}$$

Where,

R ≡ universal gas constant (= 8.31434 $\mathrm{JK^{-1}mol^{-1}}$)

M_w ≡ molar weight of water vapor (= 18.0152kg $\mathrm{kmol^{-1}}$).

Equation (2.25) therefore becomes,

$$ZWD = 10^{-6} \int_0^{p_a} \frac{R}{M_w} \cdot \frac{1}{g} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T} \right) \cdot \frac{r}{r+1} \cdot dp \tag{2.33}$$

Once again assuming a linear variation of variables within a model layers Equation (2.33) can be approximated to,

$$ZWD_{within} \approx 10^{-6} \cdot \frac{R}{M_w} \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_j} \right) \cdot \frac{r_j}{r_j + 1} \cdot \Delta p^h \tag{2.34}$$

2.4.3. Troposphere delays below NWP model orography

Finally, we evaluate the delay due to the portion of the atmosphere between the NWP model level just above the GPS antenna and the actual antenna height. The model level just above the antenna could be the lower-most model level (antenna is below model orography) or any other level in case the antenna falls within the model orography. In this case,

$$ZHD_{lower} \approx 10^{-6} \cdot k_1 \cdot R_d \cdot \frac{1}{g_a} \cdot \left(p_a - p_j^h \right) \tag{2.35}$$

$$ZWD_{lower} \approx 10^{-6} \cdot \frac{R}{M_w} \cdot \frac{1}{g_a} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_a} \right) \cdot \frac{r_a}{r_a + 1} \cdot \left(p_a - p_j^h \right) \tag{2.36}$$

Where,

p_a, T_a, r_a and $g_a \equiv$ pressure, temperature, water vapor mixing ratio and the local gravity acceleration respectively at the GPS antenna height

p_j^h ≡ the half-level pressure at the model level just above the GPS antenna height.

2.4.4. Local acceleration due to gravity

In all the formulations of troposphere delay described in Sections 2.4.1, 2.4.2 and 2.4.3, acceleration due to gravity is a variable that is expressed as a function of latitude and height. Thus, following Lide (2005), the reference formula for gravity acceleration (in m/s^2) at latitude φ and height h (in km) above the earth's surface is,

$$g(\varphi, h) = g_e(1 + 0.0052885 \cdot \sin^2 \varphi - 0.0000059 \cdot \sin^2 2\varphi) - 0.003086 \cdot h \tag{2.37}$$

Where,

g_e ≡ gravity acceleration at the earth's equator (= 9.780356 ms^{-2})

h ≡ geometric height in kilometers.

2.4.5. Total zenith path delay in terms of NWP model variables

The total zenith path hydrostatic and wet delays are, therefore, obtained by summing up the delay contributions expressed in Equation (2.24), (2.27) and (2.35) for total zenith path hydrostatic delay and Equation(2.25), (2.34) and (2.36) for total zenith path wet delay. The sum of these two quantities gives the total zenith path delay in terms of NWP model variables. Thus,

$$ZTD_{NWP} = ZHD_{NWP} + ZWD_{NWP} \tag{2.38}$$

Where,

$$\begin{aligned} ZHD_{NWP} &\approx ZHD_{upper} + ZHD_{within} + ZHD_{lower} \\ &\approx 10^{-6} \cdot k_1 \cdot R_d \cdot \frac{1}{g_1} \cdot p_1 + 10^{-6} \cdot k_1 \cdot R_d \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \Delta p^h + 10^{-6} \cdot k_1 \cdot R_d \cdot \frac{1}{g_a} \cdot \left(p_a - p_j^h \right) \end{aligned} \tag{2.39}$$

$$\begin{aligned} ZWD_{NWP} &\approx ZWD_{upper} + ZWD_{within} + ZWD_{lower} \\ &\approx 10^{-6} \cdot \frac{R}{M_w} \cdot \frac{1}{g_1} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_1} \right) \cdot \frac{r_1}{r_1 + 1} \cdot p_1 \\ &\quad + 10^{-6} \cdot \frac{R}{M_w} \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_j} \right) \cdot \frac{r_j}{r_j + 1} \cdot \Delta p^h \\ &\quad + 10^{-6} \cdot \frac{R}{M_w} \cdot \frac{1}{g_a} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_a} \right) \cdot \frac{r_a}{r_a + 1} \cdot \left(p_a - p_j^h \right) \end{aligned} \tag{2.40}$$

Equation (2.39) and (2.40) are having three similar terms; the first term accounts for troposphere delay above the model orography, the second term accounts for troposphere delay within the model orography, and the third term accounts for troposphere delay for a model layer bounded by the model level just above the GPS antenna and the actual antenna height.

2.5. Precipitable water vapor formulation using NWP model variables

Reference is hereby made to Equation (2.12). Thus,

$$PWV = \frac{1}{\rho_l} \int_h^{\infty} \rho_w \cdot dh \tag{2.41}$$

Where, ρ_l and $\rho_w \equiv$ the densities (in kgm^{-3}) of liquid water and water vapor respectively.

Assuming the atmosphere to be in hydrostatic equilibrium, $dp = -\rho \cdot g \cdot dh$, and therefore Equation (2.41) becomes,

$$PWV = \frac{1}{\rho_l} \int_0^{p_a} \frac{\rho_w}{\rho} \cdot \frac{1}{g} \cdot dp \tag{2.42}$$

Where, $\rho \equiv$ total air density, $\rho = \rho_d + \rho_w$, $\rho_d \equiv$ density of dry air, $\rho_w \equiv$ density of water vapor, $\rho_l \equiv$ density of liquid water.

The integrand includes a ratio of water vapor density to total air density. This has already been shown to be equal to specific humidity which has then been expressed in terms of the water vapor mixing ratio in Equation (2.29).

Given that most water vapor is located in the lower 2 - 3 km of the atmosphere (Bevis *et al.*, 1994), the term that would otherwise represent PWV on a column above the NWP model's top-most level can be safely ignored. Thus, the total precipitable water vapor becomes,

$$PWV_{NWP} \approx \frac{1}{\rho_l} \cdot \frac{1}{g_a} \cdot \frac{r_a}{r_a + 1} \cdot (p_a - p_j^h) + \frac{1}{\rho_l} \cdot \sum_{j=l}^{m} \frac{1}{g_j} \cdot \frac{r_j}{r_j + 1} \cdot \Delta p^h \tag{2.43}$$

Where, all the constants and variables are as earlier defined.

Similarly, the weighted mean temperature may be estimated by,

$$\begin{aligned} T_m &= \frac{\int_h^\infty \frac{e}{T} \cdot dh}{\int_h^\infty \frac{e}{T^2} \cdot dh} \\ &\approx \frac{\frac{1}{g_a} \cdot \frac{r_a}{r_a + 1} \cdot \frac{1}{T_a} \cdot (p_a - p_j^h) \quad + \quad \sum_{j=1}^{m} \frac{1}{g_j} \cdot \frac{r_j}{r_j + 1} \cdot \frac{1}{T_j} \cdot \Delta p^h}{\frac{1}{g_a} \cdot \frac{r_a}{r_a + 1} \cdot \frac{1}{T_a^2} \cdot (p_a - p_j^h) \quad + \quad \sum_{j=1}^{m} \frac{1}{g_j} \cdot \frac{r_j}{r_j + 1} \cdot \frac{1}{T_j^2} \cdot \Delta p^h} \end{aligned} \tag{2.44}$$

Where, all the terms are as earlier defined.

2.6. Total zenith path delay observations operator

A variational data assimilation (DA) system attempts to iteratively minimize the differences between a model initial fields and observations (mentioned in Section 1.3) in order to get an improved analysis. For a 3D-Var system, the resulting cost function *J*, consists of two terms; J_b, which measures the distance between the resulting analysis and the initial field and J_o, which measures the distance between the resulting analysis and the observations. Following Elgered *et al.*, (2005); Schwitalla (2012) and NCAR (2014), this cost function is given by,

$$J_{3D-Var} = J_b + J_o = \frac{1}{2} \cdot (x - x_b)^T \cdot B^{-1} \cdot (x - x_b) + \frac{1}{2} \cdot \left(H(x) - y\right)^T \cdot R^{-1} \cdot \left(H(x) - y\right) \tag{2.45}$$

Where,

x ≡ the model state vector to be determined by the minimization

x_b ≡ the background (short range forecast derived from a previous analysis)

B and R ≡ the error covariance matrices of the background field (a constant in 3D-Var) and the observations respectively

H ≡ observation operator which transforms the NWP model variables into the observed quantities

y ≡ a vector containing all observations.

The observations operator, *H*, allows for comparison of observations and model initial fields. The differences are important in determining the most likely corrections to the model initial fields which are then added to the initial fields to give a more accurate analysis.

For an observation corresponding to a variable of an NWP model, an observation operator performs interpolation in the model field to the location of the observation. However, some observations, such as total zenith path delay (ZTD) observations do not correspond to a variable of an NWP model. A ZTD observation operator, therefore, has to be computed.

A ZTD observations operator should determine the estimated value of a ZTD observable given the NWP model state. However, since ZTD observations do not correspond to any one particular NWP model variable, but rather on a number of variables (pressure, temperature, and humidity), the estimated value of the ZTD observable is determined given the model state of atmospheric parameters at a GPS site.

The ZTD observations operator is expressed as (Saastamoinen, 1972; Huang and Vedel, 2000),

$$ZTD = 2.2767\times10^{-3}\left[\mathrm{m/hPa}\right]\cdot\frac{p_a}{f(\varphi,h)} \quad + \quad \frac{1}{g(\varphi)}\cdot\sum_{j=1}^{n} q_j\cdot\Delta p^h \tag{2.46}$$

Where,

$$f(\varphi,h) = 1 - 2.66\times10^{-3}\cdot\cos 2\varphi - 2.8\times10^{-7}\left[\mathrm{m}^{-1}\right]\cdot h_a$$

Where, h_a ≡ antenna height in m

p_a ≡ pressure at the GPS antenna height in hPa

$g(\varphi)$ ≡ gravity acceleration depending on latitude, φ

q_j ≡ humidity at the j^{th} half-level

$\Delta p^h = p^h_{j+1} - p^h_j$ ≡difference in pressure at adjacent half-levels $j+1$ and j.

The first term of Equation (2.46) represents the observations operator for ZHD and is based on the Saastamoinen hydrostatic model (a detailed description of the model can be found in Schueler (2001), while the second term represents the observations operator for ZWD.

2.7. Summary

This chapter has presented some of the background concepts that were found necessary in this research. Important mathematical formulations have been laid out. Fundamental derivations have been avoided particularly in situations where the same can be found in other scientific texts. Appropriate references have therefore been provided.

Most of the formulations revolved around troposphere delays and precipitable water vapor with relevance to NWP models. The concepts laid out here shall be applied and expanded further (if need be) in the studies presented in this thesis.

3. General methodology

The environment necessary to carry out the research was laid out. This included the acquisition and installation of all the necessary software including the Environmental Monitoring System (EMS) with which the ARW model was run, and the 3D-Var WRF data assimilation (WRFDA) system. It was deemed unnecessary to include details on how these systems were set up and implemented in the main write up of this thesis. However, a summarized description is presented in Appendix B.

One of the main factors that determine the accuracy in data extracted from NWP model output files is the strategy employed in the data extraction process. The method that was used in this research is presented here.

After setting up the NWP and data assimilation (DA) systems, and several other related software, several test simulations were done in order to test among other things, the optimum NWP model resolution needed to make a successful model run with the available computing resources. After making enough test runs that tested software against hardware, some simulations were made the results of which form a bulk of what is presented in this thesis.

The main study involved an NWP model simulation that was made on an area covering Central Europe. The results were analyzed in several different ways as was deemed necessary. Another study involved an area along the African equator specifically covering Kenya.

3.1. Variables retrieval from NWP model output files

The output files from the EMS are in (by default) NetCDF[1] format. The basic structure of a NetCDF file consists of a header and a main body. The header contains dimensions and variables while the main body contains the real data. Dimensions represent physical quantities that have a name and a length such as latitude, longitude, time and height while variables represent arrays of values of the same type. A variable therefore has a name, data type (real, character, etc.), and a shape (defined by a set of dimensions). Attributes may also be associated with a variable to specify such properties as units, maximum and minimum valid values, etc.

Meteorological variables in the atmosphere are located in space (as specified by latitude, longitude and height/or pressure level) and time. For a given time step n, data entry at grid points can be defined in terms of either only the horizontal dimensions (in x and y) for surface variables or a combination of the horizontal dimensions and a vertical dimension (vertical pressure level in z) for all the other variables falling above the surface or within the three-dimensional grid of the NWP model. Consequently, for a

---

[1] Network Common Data Format (NetCDF) is a machine-independent data format that supports the creation, access and sharing of array-oriented scientific data.

given variable, say *q*, its values are stored in the NetCDF file as the entry *q*(i,j,k,n) for the values associated with the grid box centered at indices[1] (i,j,k) at time step n.

The data-containing grid points and the station-specific longitude/latitude points are not collinear. Neither are the vertical pressure levels and station (or antenna) heights. This calls for the use of a combination of several methods (basically including interpolation or extrapolation procedures) in the extraction of the variables at the station-specific parameters, namely; longitude, latitude, and height.

From the studies by Schueler (2001) and Jones (2010), it appears that the methods used to extract these variables are somehow similar in strategy but different in implementation. In Schueler (2001) the variables at four nearest neighboring grid points are height-reduced and then interpolated horizontally to the user position. Jones (2010) uses linear interpolation to compute variables to the antenna height. He, however, does not elaborate how variables are deduced in the horizontal (to the user position).

The strategy used in this study to extract variables from the NWP model output files is illustrated in Figure 3.1. A GNSS reference site is identified in terms of its longitude and latitude coordinates (denoted as *x* and *y* respectively in Figure 3.1). A pair of indices corresponding to the latitude and longitude variables is then extracted from the NetCDF files. This is our point of interest, denoted as P in Figure 3.1, and it is highly unlikely that it corresponds to any grid point. Consequently no variables exist at the point.

Four nearest neighboring grid points are then identified denoted as Q1,..,Q4 in Figure 3.1. This gives us a 2-dimensional plane with the point of interest at the center and a grid point at each of the four corners. At these grid points we find all the variables that we would want. Longitude and latitude variables at each of these four grid points are extracted and then used to extract a target variable at each of the four grid points.

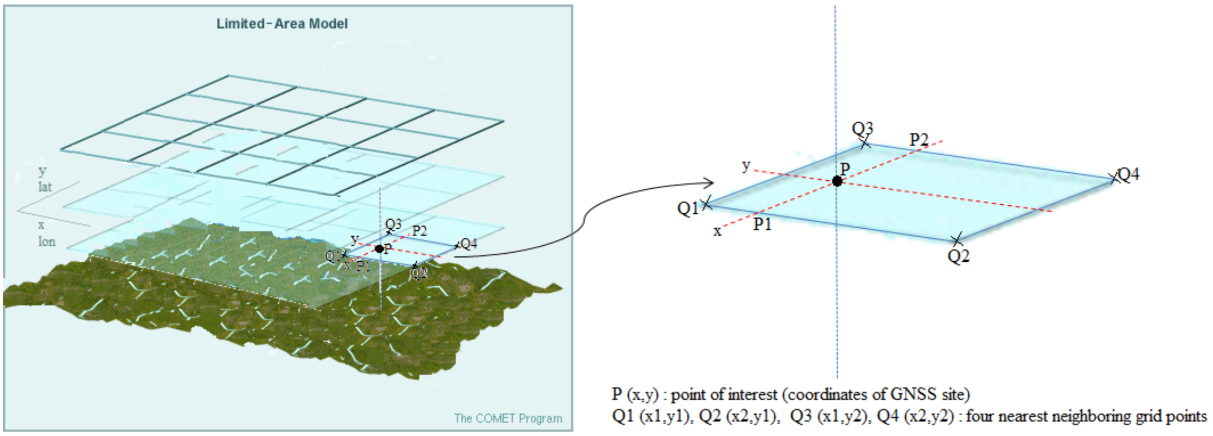


Figure 3.1: Variables retrieval strategy from the NWP model 3D grid. Part of the diagram was adapted from MetEd (2014).

---

[1] Section 2.3.

Notice that we have not yet specified the half mass level at which the variables reside. This means that so far we have a series of 2-dimensional planes extending vertically at each of the model half mass levels.

By use of the longitude and latitude variables at the point of interest and at the four nearest neighboring grid points, the target variable (contained at each of the four neighboring grid points) is bilinearly interpolated to the point of interest. The bilinear interpolation process is elaborated in Appendix C. The accuracy of the estimated variable at the point of interest is a trade-off between grid spacing and the available computing resources. The smaller the grid spacing, the smaller will be the 2-dimensional planes (Figure 3.1) and a better accuracy in the estimated variable at the point of interest. This, however, is subject to the available computing power.

Now we have the target variable not only at the four nearest neighboring grid points but also at the point of interest on each of the 2-dimensional planes. If we were to join the points of interest on each plane, we would have a vertical line with nodes denoting the half mass levels.

Finally, given that each half mass level is characterized by a given geometric height, we can linearly interpolate the target variable to the geometric height of the antenna at the GNSS site. Two things are likely to happen: the antenna could be below the NWP model orography (taken to mean below the lower-most half mass level), or it could as well be somewhere within model orography (as was illustrated in Section 2.4).

In a situation where the antenna is below model orography basic target variables are estimated by use of near surface variables and the standard atmosphere model as further discussed in Section3.1.1. In case the antenna is within the model orography linear interpolation is performed because the model variables above and below the antenna position would be known. This is further discussed in Section 3.1.2.

Variables at model half mass levels are readily extracted as discussed in Section 3.1.3.

In the context of this work, the main target variables are pressure, temperature and water vapor mixing ratio. Other variables, for instance, local gravity acceleration at a point of measurement have to be derived.

3.1.1. Antenna below model orography

In order to determine variables at the antenna height, extrapolation of model level variables above the antenna can be performed. Alternatively, the variables at the antenna height (mainly pressure and temperature) can be estimated from the standard atmosphere model (Derks *et al.*, 1997; Cavcar, 2013). Thus,

$$p_{ant} \approx p_{msl}\left(1-0.0000226\cdot h\right)^{5.225} \tag{3.1}$$

$$T_{ant} \approx T_{msl} - 0.0065\cdot h \tag{3.2}$$

Where,

P_{ant} and T_{ant} ≡ antenna height pressure (in hPa) and temperature (in K) respectively

h ≡ antenna height in m

P_{msl} and T_{msl} ≡ mean sea level pressure and temperature respectively

Where, $T_{msl} \approx 288.15$ K and $P_{msl} \approx 1013.25$ hPa

P_{msl} can also be derived from the NWP model output files. The standard value of temperature lapse rate, i.e. 6.5 K/1000m is assumed.

Use of the standard pressure and temperature models was preferred over the extrapolation procedures because it helps us avoid errors associated with extrapolation.

Water vapor mixing ratio was taken to be the value at 2 m above the surface. This variable is readily extracted from the NWP output files.

Local gravity acceleration at the antenna height was estimated using the latitude and elevation variable formula,

$$g_{ant} = g_e\cdot\left(1+0.0052885\cdot\sin^2\varphi-0.0000059\cdot\sin^2\left(2\cdot\varphi\right)\right)-0.003086\cdot h_{ant} \tag{3.3}$$

Where,

g_e ≡ gravity acceleration at the earth's equator (= 9.780356 $\mathrm{ms^{-2}}$)

φ ≡ the latitude of the reference station

h_{ant} ≡ the antenna height in km.

3.1.2. Antenna within model orography

Though the antenna could be well within the model orography, there is still a possibility that the vertical position of the antenna does not coincide with a half mass level. For this reason, since variables above and below the vertical position of the antenna are known, vertical linear interpolation of the variables to the antenna height is performed in order to estimate variables at the antenna height.

Variables at model half mass levels used in the interpolation are directly extracted from the NWP output files as discussed in Section 3.1.3.

3.1.3. Variables at model half mass levels

The water vapor mixing ratio is readily available at all the model levels. However, pressure and temperature variables have to be computed from other variables as discussed here.

The model output files contain base state pressure and perturbation pressure. The base state pressure is considered as the mean value of pressure and a departure from this mean is the perturbation pressure. Pressure is therefore computed as a sum of these of these two pressure variables that are directly extracted from model half mass levels. Thus,

$$P = P_b + P_p \tag{3.4}$$

Where, P_b and P_p are the base state and perturbation pressure respectively.

The temperature variable contained in the model output is the perturbation potential temperature from 300 K. Potential temperature[1] (in Kelvin) is, therefore, computed from,

$$\Theta = \theta + 300 \tag{3.5}$$

Where, Θ and θ are the potential and perturbation potential temperatures respectively.

The absolute temperature is then computed using Poisson’s equation for potential temperature as,

$$T = \Theta \cdot \left(\frac{P}{P_o} \right)^{\kappa} \tag{3.6}$$

Where,

P ≡ pressure as determined in Equation (3.4)

P_o ≡ standard pressure (=1013.25 hPa)

$$\kappa = \frac{R_d}{c_p}$$

Where,

R_d ≡ gas constant of dry air (= 287.054 JK^{-1} kg^{-1})

c_p = specific heat capacity at constant pressure (=1.01 kJ/kg K).

---

[1] This is the temperature a parcel of air would have if it were brought adiabatically to some reference pressure.

Local gravity acceleration at model half mass levels needed extra attention in its computation. Since geopotential heights in the ARW model grid are staggered in the z direction (stored in the full levels), their values at half levels is computed by averaging the values at adjacent full levels. The geopotential heights at half levels are then converted to the corresponding geometric heights for implementation in the formula for gravity acceleration.

Following OFCM (1997), with knowledge of the geopotential height, the geometric height at the j^{th} half level can be computed from,

$$h_j = \frac{H_j \cdot R_e}{G \cdot R_e - H_j} \tag{3.7}$$

Where,

h_j ≡ the geometric height (in km) at half mass level *j*

H_j ≡ the geopotential height (in km) corresponding to h_j

R_e ≡ radius of the earth (in km) at latitude φ, given by,

$$R_e = \left(\frac{\cos^2 \varphi}{R_{\max}^2} + \frac{\sin^2 \varphi}{R_{\min}^2} \right)^{-\frac{1}{2}}$$

Where,

$R_{\max}$ ≡ the semi-major axis/earth's equatorial radius (= 6378.137 km)

$R_{\min}$ ≡ the semi-minor axis/earth's polar radius (= 6356.752 km)

G ≡ gravity ratio[1], given by,

$$G = \frac{g\left(\varphi, h\right)}{g_o}$$

Where,

$g\left(\varphi, h\right)$ ≡ gravity at the point of measurement

g_o ≡ standard gravity at mean sea level (= 9.80665 ms^{-2})

Thus, local gravity at half level *j* is estimated as,

---

[1] The gravity ratio has been taken to be unity in this study.

$$g_j = g_e \cdot \left(1 + 0.0052885 \cdot \sin^2 \varphi - 0.0000059 \cdot \sin^2 \left(2 \cdot \varphi\right)\right) - 0.003086 \cdot h_j \qquad (3.8)$$

Where, all the terms are as earlier defined.

3.2. Study case: Central Europe

The actual study area, henceforth interchangeably referred to as study domain, was centered on a geographical location that covers a region in Central Europe including Germany, Switzerland, Poland, Slovakia, Czech Republic, Austria and Hungary (Figure 3.2).

The study period included the last three days and the first two days of May and June respectively of the year 2013. Heavy precipitation was experienced in Central Europe during this period with some weather stations recording over 200 mm precipitation (Schrier, 2013).

The main configurations of the NWP model computational domain (Figure 3.2) extended from a longitude of approximately 5°E to 24°E and latitude of approximately 46°N to 55°N. It included an array of 208 by 158 grid points defining a horizontal dimension in *x*- and *y*-axis respectively, a grid spacing of 6 km (resolution or scale of the features to be simulated), and the Lambert Conformal map projection type. The choice of the Lambert Conformal map projection meant that among the two WRF dynamical cores, NMM and ARW, it is the ARW model that was used. The ARW model is particularly suited for regional climate modeling and idealized simulations.

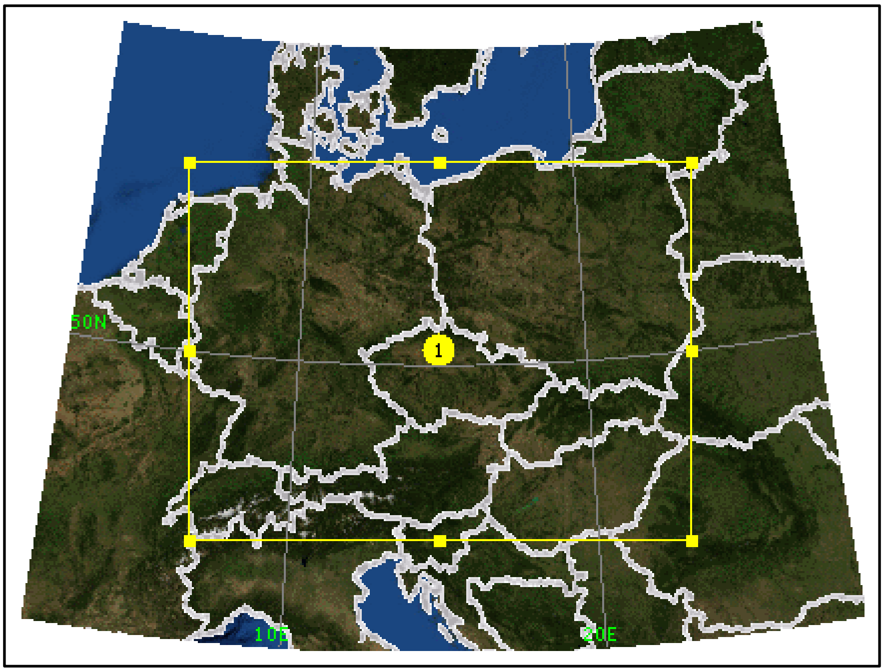


Figure 3.2: The study domain covers a region around Central Europe.

45 vertical pressure levels were used to define the vertical structure of the atmosphere. For microphysics and parameterizations, the Kessler and the Kain-Fritsch schemes were used respectively. Details on these schemes can be found in Skamarock *et al.* (2008).

Although the ARW can be used on a global scale, it was implemented on a limited-area scale while the Global Forecast System (GFS) 0.5 degree analysis fields were identified to provide the initial and boundary information for the NWP simulation. The EMS achieves this through an initialization routine. This is then followed by the actual model simulation that starts by creating the initial and boundary conditions and then runs the actual simulation (details on this in Appendix B).

The model simulation involved a forecast length of 24 hours with the first simulation running from 2013-05-29_18:00 UTC to 2013-05-30_18:00 UTC and followed by four other consecutive simulations with the last one running from 2013-06-02_18:00 UTC to 2013-06-03_18:00 UTC. This resulted into 25 hourly NetCDF output files for each simulation.

Because the forecast times run from one doy (day of year) to the next (not really defining an actual doy), the following definitions were made:

- 2013-05-29_18 UTC to 2013-05-30_18 UTC : doy 149/150
- 2013-05-30_18 UTC to 2013-05-31_18 UTC : doy 150/151
- 2013-05-31_18 UTC to 2013-06-01_18 UTC : doy 151/152
- 2013-06-01_18 UTC to 2013-06-02_18 UTC : doy 152/153
- 2013-06-02_18 UTC to 2013-06-03_18 UTC : doy 153/154

Pressure, temperature and relative humidity variables were extracted from the NWP model output files and subjected to a validation study. Further discussion on this and the results are presented in Chapter4. Following the validation study, troposphere delays and precipitable water vapor were estimated using NWP model meteorological variables and the results were validated by comparison with similar observations coming from ground-based GPS reference stations and as derived from radiosonde data at selected radiosonde sites. This forms the subject of the studies discussed in Chapter5 and 6.

GPS ZTD observations were then ingested into the initial analysis file from the NWP model simulation by use of the WRF 3D-Var DA system. The DA system was also used to update the boundary conditions file so that it was consistent with the new initial analysis file. After assimilation of the ZTD observations, another round of NWP model simulations was done by maintaining the same configurations that were made during the first round of model simulations and using the initial and boundary conditions files coming from the data assimilation system.

The NWP model output files that were obtained before assimilating ZTD observations were then compared to those obtained after ZTD assimilation with an aim of assessing the impact of ZTD data assimilation on short range forecasts of surface pressure, and near-surface temperature and relative humidity. Discussions on this together with the method used in the ZTD data assimilation process are presented in Chapter 7.

3.3. Study case: Kenya

Finally, a model simulation was made for a domain covering Kenya, an equatorial region. Most of the studies conducted in the Central Europe case were repeated here. However, due to the scarcity of ground-based GPS ZTD observations in this region, a study on ZTD data assimilation impact assessment was not done. The results and discussions of the study are presented in Chapter 8.

3.4. Summary

The method used in the extraction of variables from an NWP model grid is vital for the accurate estimation of some target parameter.

The general methodology adopted in several different studies presented in this thesis has been provided. The specific methods used to carry out each individual study have not been exhaustively discussed here as these are discussed further under the methodology in each of the studies presented in the chapters that follow.

Majority of these studies are centered on an NWP model simulation that covered a region that was and has been henceforth referred to “Central Europe”. Some of these studies have been replicated on a region covering Kenya, an equatorial region.

4. Validation and reliability assessment of NWP-derived surface meteorological data

4.1. Introduction

Three-term expressions necessary for estimating zenith hydrostatic delay (ZHD), zenith wet delay (ZWD) and precipitable water vapor (PWV) based on meteorological variables that are available in the 3D grid of an NWP model were formulated (Equation (2.39), (2.40), and (2.43)). The accuracy of these target quantities shall depend to a large extent on the accuracy of the NWP-derived meteorological variables. These accuracies will primarily serve to assess the reliability[1] of using the NWP-derived variables in the computation of troposphere delays and precipitable water vapor.

In this chapter, results and discussions are presented as deduced from a study in which NWP-derived surface or near-surface[2] meteorological variables, namely: pressure, temperature and relative humidity were validated by comparison with similar observations sourced from selected EPN GNSS reference stations that are equipped with meteorological sensors for the said variables. Validation was also done using radiosonde data sourced from several radiosondes launch sites falling within the study domain.

4.2. Methodology

NWP model simulations were been performed based on the computational domain discussed in Section3.2. The target meteorological variables, namely; pressure, temperature and relative humidity, were retrieved from the NWP model output files as described in Section 3.1.

The NWP-derived meteorological variables were subjected to a validation study using similar observations sourced from meteorological sensors at selected EPN stations. Selection of the meteorological stations was based on the location (must fall within the defined study domain) and data availability (data must be available at all times during the study period). The data was therefore sourced from the stations marked in green as shown in Figure 4.1. A list of the stations can be found in Appendix E.

Only the observations coinciding or nearly coinciding with the NWP forecast times described in Section 3.2 were extracted from the meteorological data files.

The average accuracies (accounting for both bias and precision errors) of the meteorological sensors at EPN stations are claimed to be 0.2 hPa to 0.5 hPa for pressure, 0.2 °C to 0.5 °C for temperature and 3% to 5% for humidity (Bosy and Rohm, 2007).

---

[1] In order to answer the question on the accuracy of NWP-derived troposphere delays, we need to know whether the variables used in their computation are reliable in terms of accuracy. For instance, 1.0 hPa error in pressure would translate to about 2.3 mm error in zenith hydrostatic delay.

[2] Refers to variable estimate approximately 2 meters above the surface.

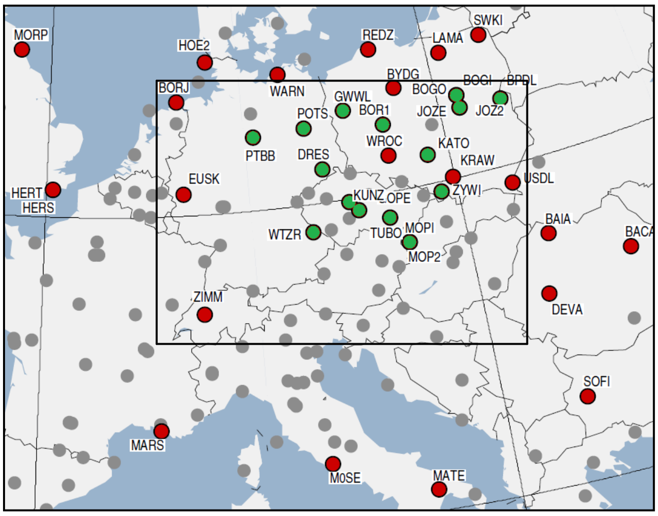


Figure 4.1: EPN sites that make available meteorological data are marked in red and green. The inset rectangle denotes the study domain. Sites marked in green were used in the current study.

Most of the meteorological sensors at the EPN stations are installed close to but not necessarily at the same height as, the GNSS antenna[1]. Therefore, extracting a variable in question from the NWP model grid was done with respect to the height corresponding to its respective sensor (i.e. the pressure, temperature or relative humidity sensor). But, situations arise where a sensor height is below the NWP model orography or where the sensor height lies within the model orography, i.e. above the lower-most model level (refer to illustration in Figure 2.6). The sensor/model orography height differences are likely to lead to errors particularly when a sensor is a long distance below the model orography. A comparison study of these heights was done with a view of having an insight into these height differences and possibly relating this to the magnitude of errors to be obtained.

Validation of the NWP-derived surface meteorological variables at the GNSS reference sites by comparison with radiosonde data is possible if the differences in vertical height and horizontal distance between the GNSS and radiosonde sites are taken care of. However, given that the NWP model can produce surface variables over its entire area of coverage, then it is also possible to compare these variables with those from radiosonde at any given point within the study domain.

NWP-derived meteorological variables were validated by use of radiosonde data sourced from Wyoming University’s radiosonde database (University of Wyoming, 2014). The seven radiosonde launch sites shown in Table 4.1 were found appropriate for the validation study. The criteria for their selection were that they fell within the study domain and that data was available at least twice a day.

---

[1] This information can be found in a station‘s log file. The approximate sensor heights are specified with respect to the GNSS antenna height.

Table 4.1: Radiosonde sites in Central Europe used in validation studies.

Radiosonde site name	Site identifier (WMO)	Longitude [deg]	Latitude [deg]	Geopotential height [m]	Launch times [UTC]
Wroclaw I	12425	16.88	51.78	122	00, 12
Prostejov	11747	17.13	49.45	216	00, 12
Praha-Libus	11520	14.45	50	303	00, 06, 12
Kuemmersbruck ETGK	10771	11.9	49.43	418	00, 06, 12, 18
Meiningen	10548	10.38	50.56	453	00, 12
Lindenberg	10393	14.11	52.21	115	00, 06, 12, 18
Bergen ETGB	10238	9.93	52.81	69	00, 06, 12, 18

All validations were done in terms of bias, bias-reduced standard deviation and root mean square errors computations and analysis where the EPN and radiosonde observations were used as the reference quantities. Formulas used for the bias and standard deviation error computations are shown in Appendix D.

4.3. Comparison of NWP and meteorological sensors observations at EPN sites

Comparison is only possible at specific points in the study domain defined by the EPN sites' coordinates. With the NWP model, however, it is possible to simulate a target variable with a much better horizontal resolution defined by the grid points dotting the horizontal extent of the model. Therefore, surface pressure, surface temperature, and near-surface relative humidity variables were extracted from the NWP model grid. The mean estimates of these variables across the entire study domain based on an NWP model simulation for 29th May 2013 18 UTC is shown in Figure 4.2in order to give us an insight into the magnitude of the target variables across the study domain.

From Figure 4.2, it is seen that most areas had a mean surface pressure greater than 950 hPa and temperatures above 7.5 °C. The eastern part of the domain had relative humidity ranging from approximately 68% to 84% and an average of 92% on the western part with some situations of relative humidity 0.1 to 0.2% greater than 100%[1] at several areas as will be seen later in the study. This study, therefore, compared the magnitude of the three variables as simulated by the NWP model at selected EPN sites. Results from the comparison expressed as an overall mean of bias, standard deviation and root mean square (rms) errors are shown in Table 4.2.

---

[1] This can happen. A condition referred to as supersaturation. Supersaturation in the atmosphere is, however, rare (Eumetcal, 2014).

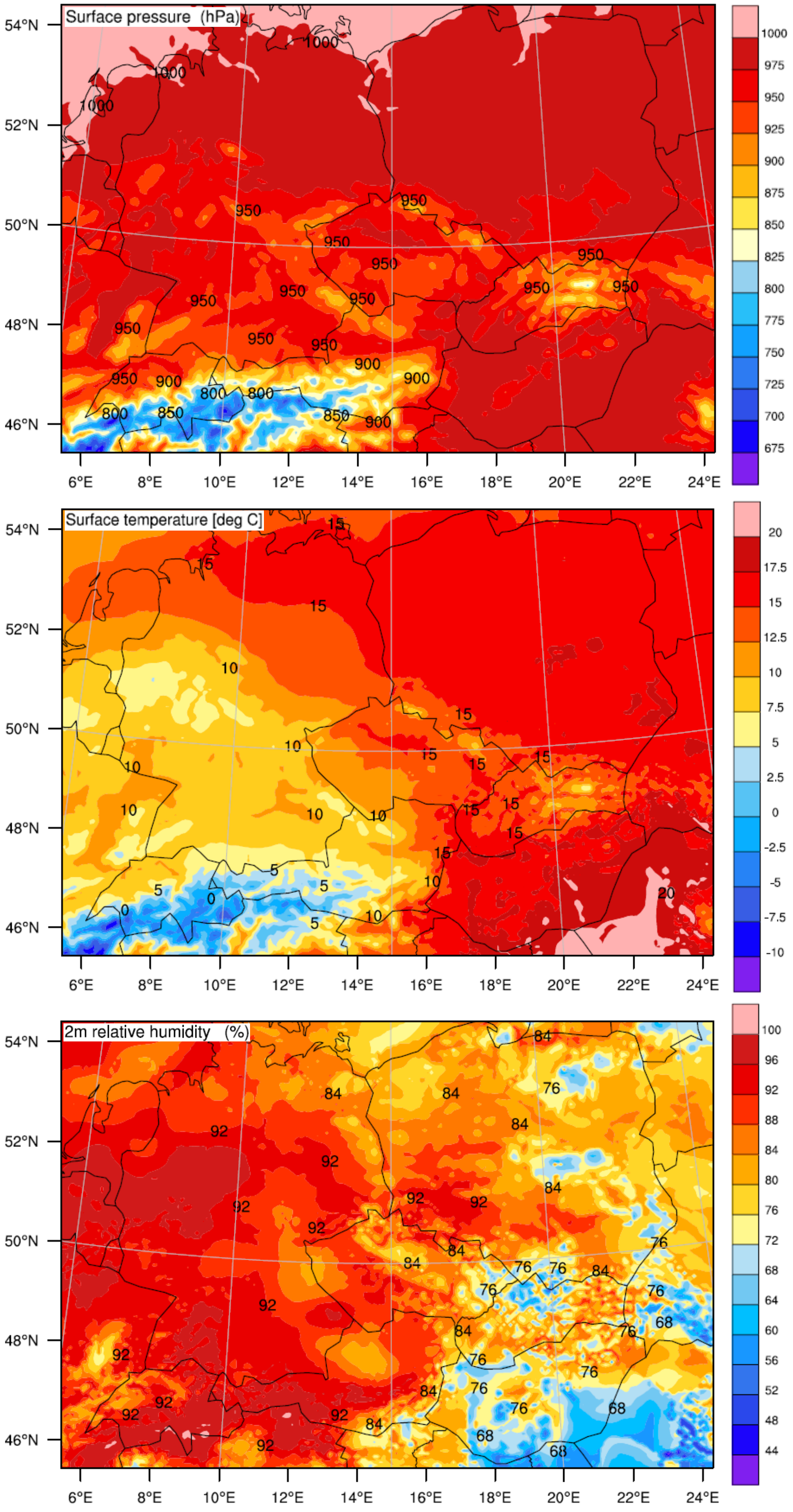


Figure 4.2: NWP simulated observations of surface pressure (top panel), surface temperature (middle panel), and near-surface relative humidity (bottom panel) across the entire study domain for 29th May 2013 18 UTC.

Table 4.2: Mean accuracy and precision errors of surface meteorological observations during the study period[1] at selected EPN sites.

Station id.	Samples	pressure [hPa] bias	stdev	rms	temperature [°C] bias	stdev	rms	relative humidity [%] bias	stdev	rms
BOGI	75	0.1	1.0	1.0	-0.6	2.0	2.2	**11.9**	**10.5**	**15.9**
BOR1	125	**-4.6**	0.6	**4.6**	**4.4**	**5.7**	**7.8**	**13.2**	8.2	**15.6**
BPDL	125	-0.1	0.6	0.7	1.5	2.0	2.7	2.2	**10.1**	11.6
DRES	125	-0.4	0.6	0.8	1.2	1.7	2.3	9.1	6.5	11.6
GOPE	125	0.2	0.6	0.7	0.1	0.9	1.0	6.5	5.4	8.6
GWWL	125	0.3	0.6	0.9	0.1	1.2	1.4	9.8	6.4	12.3
JOZE	125	0.6	0.8	1.1	0.6	**2.8**	2.9	**10.6**	**14.4**	**17.8**
KATO	125	-0.5	0.7	0.9	**3.8**	1.8	**4.3**	-2.3	**12.4**	13.5
KUNZ	125	1.0	0.5	1.2	0.3	1.0	1.1	-4.1	5.0	7.3
MOPI	125	0.1	0.4	0.5	-0.7	1.2	1.8	**11.2**	8.8	**14.6**
POTS	125	0.3	0.5	0.8	-0.1	1.4	1.6	1.4	7.2	8.3
PTBB	125	-0.3	0.7	0.9	0.4	1.5	1.9	**13.2**	9.0	**16.7**
TUBO	125	-0.5	0.6	0.8	2.0	1.4	2.6	1.0	8.3	9.1
WTZR	100	-0.3	1.0	1.2	0.2	1.1	1.2	5.1	5.7	8.3
ZYWI	125	-0.4	0.6	0.8	**3.5**	2.0	**4.1**	3.3	**11.9**	12.9
mean		**-0.30**	**0.66**	**1.13**	**1.11**	**1.86**	**2.60**	**6.15**	**8.65**	**12.29**

The surface pressure observations were found to have an overall negative mean bias of approximately 0.3 hPa, a standard deviation that was slightly less than 1 hPa and an rms of approximately 1.1 hPa with an obvious outlier in the bias and rms errors at BOR1 which is further investigated in Section 4.5.

The surface temperature observations were found to have an overall mean bias of approximately 1.0 °C, a standard deviation error of approximately 2.0 °C and an rms error of approximately 2.6 °C. Outliers are noted at BOR1, KATO and ZYWI. These are further investigated in Section 4.6.

The overall mean bias in the relative humidity observations was found to be approximately 6% with approximately 9.0% and 12% standard deviation and rms errors respectively. Bias errors greater than 10% were found mainly at BOGI, BOR1, MOPI and PTBB. These are further investigated in Section 4.7.

The rms errors are not bias-corrected. They therefore reflect the forecast errors associated with the bias. This then explains why the rms errors are large where the corresponding bias errors are large (as seen in the bold print bias and rms entries in Table 4.2).

---

[1] This period was defined in Section 3.2. Results covering a longer period of time, one month, January 2014, are shown in Appendix F. This was done much later and was therefore not included in the main text.

A possible source of the error could be due to the meteorological sensor/model orography height differences at the stations considered. This is especially expected with bias errors and in situations where the sensor position is a long distance below model orography (taken to mean model's lower-most mass level).

4.4. Meteorological sensor/model orography heights comparison

In most situations, the three sensors in question, pressure, temperature and humidity sensor, at the EPN stations from which meteorological data has been sourced are housed within a weather station. Also, in most cases, small sensor height differences exist and are well documented in the station log files. Meteorological sensor heights are specified with respect to the GNSS antenna height. Since the latter height is given as the ellipsoidal height, conversion to the corresponding orthometric height[1] was done based on the EGM96[2] geoid undulation model (Lemoine, et al., 1998; UNAVCO, 2013). These heights were then compared to geometric heights corresponding to the NWP model's lower-most level (henceforth simply, model orography height).

A significant amount of effort was made to verify the meteorological sensor heights using the information provided in the station log files and clarification from the necessary agencies particularly in situations where the information in the log files was not clear[3].

First, model orography heights (for a simulation on 29th May 2013 18 UTC) were compared with 90 m resolution SRTM[4] elevation data. The latter data was taken to represent the actual elevations. As shown in Table 4.3, the model simulated orography is above the actual orography at BOR1, DRES, KATO, TUBO and ZYWI, and it is below the actual orography at GOPE, GWWL, MOPI, POTS, PTBB and WTZR. The elevations at BOGI, BPDL and JOZE could not be deduced due to voids in the SRTM data.

Extremely large heights differences are particularly seen at GOPE, MOPI and WTZR. Possible reasons for this were found to originate from the model simulation as discussed in the paragraphs that follow.

Table 4.3 also shows the differences in height for each of the sensors (pressure, temperature and humidity) when compared to the model orography heights. A negative height difference indicates a sensor that is below the lower-most mass level of the NWP model while a positive height difference indicates a sensor that is above the lower-most mass level of the NWP model (lying within the NWP model orography).

---

[1] Orthometric height = GPS ellipsoidal height – geoid height.

[2] Earth Gravitational Model 1996.

[3] For instance, concerning missing data for doy 153 and 154 at BOGI, and confirmation of sensor heights at BPDL.

[4] Shuttle Radar Topography Mission (Gorokhovich and Voustianiouk, 2006).

Table 4.3: Comparison of the model orography heights with SRTM elevations and meteorological sensor heights.

Station id.	**NWP ht. [m]**	**SRTM ht. - NWP ht. [m]**	**P_{sensor} ht. - NWP ht.[m]**	**T_{sensor} ht. - NWP ht.[m]**	**H_{sensor} ht. - NWP ht.[m]**
BOGI	71.7	-	37.4	37.4	37.4
BOR1	67.2	-3.2	16.4	16.4	16.4
BPDL	162.8	-	5.1	5.1	5.1
DRES	171.5	-28.5	**-17.9**	**-13.4**	**-14.4**
GOPE	409.2	**100.8**	**137.8**	**137.7**	**137.7**
GWWL	43.8	28.2	47.5	47.5	47.5
JOZE	98.3	-	11.5	0.4	0.4
KATO	283.1	-13.1	8.5	8.5	8.5
KUNZ	607.8	-1.8	40.9	40.9	40.8
MOPI	362.3	**199.7**	**173.0**	**173.0**	**173.0**
POTS	38.3	26.7	54.2	63.8	63.8
PTBB	73.6	5.4	13.5	10.5	10.5
TUBO	270.0	-20.0	8.4	8.4	8.4
WTZR	520.7	**60.3**	**88.0**	**98.5**	**98.5**
ZYWI	415.9	**-44.9**	**-45.2**	**-45.2**	**-45.2**

Positive height differences should not be a concern because this would mean that the antenna is lying within the NWP model orography and could be close to any of the other lower mass levels of the model in which case interpolation involves variables found within model mass levels. On the other hand, negative height differences should be a concern particularly where large differences are found. This is because the interpolation would likely involve near surface (or even mean sea level) variables that may be assumed to correspond to the model's bottom which is again assumed to mimic the local terrain. However, the NWP model may not truly capture local terrain effects at the specified grid spacing[1].

The pressure, temperature and humidity sensors at some of the stations considered are not at the same height as seen for DRES, JOZE, POTS, PTBB and WTZR. The antenna is above the model orography at majority of the stations considered apart from DRES and ZYWI where it is below model orography with a significant difference observed at ZYWI. Such a large difference may introduce errors.

Some situations where extremely large height differences are observed could be due to the model failing to capture the local terrain effects, say due to valleys and trees, at the specified grid spacing. The large height difference at MOPI could be due to the environment surrounding the site which consists of a forest with rocky formations(Hefty, 2001). The relatively large height differences at GOPE, POTS and WTZR could also be as a result of surrounding forests.

---

[1] Increasing the resolution by decreasing the grid spacing adds more strain to the computing power.

If the errors in Table 4.2 originated from the model orography/antenna height differences, then we would expect to find the largest errors only at ZYWI because the antenna here appears to be a large distance (approximately 45 m) below the model orography. The sections that follow, therefore, attempts to analyze the errors observed in each variable with focus on the most problematic stations: pressure at BOR1, temperature at BOR1, KATO and ZYWI, relative humidity at BOGI[1], BOR1, JOZE, MOPI and PTBB.

There were no meteorological observations in the EPN RINEX[2] meteorological files for BOGI/BOGO on the doy earlier defined as 152/153 and 153/154 and for WTZR on doy 151/152.

4.5. Pressure error analysis

As shown in Table 4.2, the pressure bias at most of the stations considered is less than 0.5 hPa. The pressure bias at station BOR1 is conspicuously large (approximately 5 hPa). An assessment of the pressure bias error at each station on each day of the study period has been done (Figure 4.3).

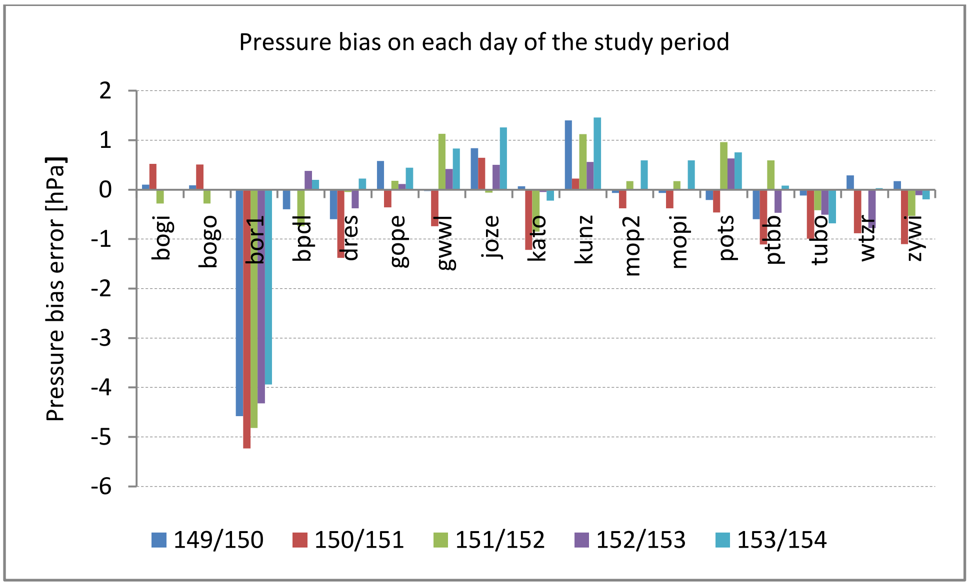


Figure 4.3: Mean pressure bias errors on each day of the study period at all the stations.

There are significantly large negative biases in pressure observations at BOR1 on all the days considered. Sensor and NWP hourly pressure observations appear to follow similar trends as illustrated in Figure 4.4. However, pressure variables derived from the NWP are significantly greater than the sensor values with mean values of 991 hPa and 995 hPa for the sensor and NWP respectively, hence the strong negative bias.

---

[1] BOGI and BOGO are practically at the same location. The same is with MOPI and MOP2. Results from either of the station pairs are similar. Therefore, only BOGI and MOPI have been discussed in the analysis.

[2] Receiver Independent EXchange format.

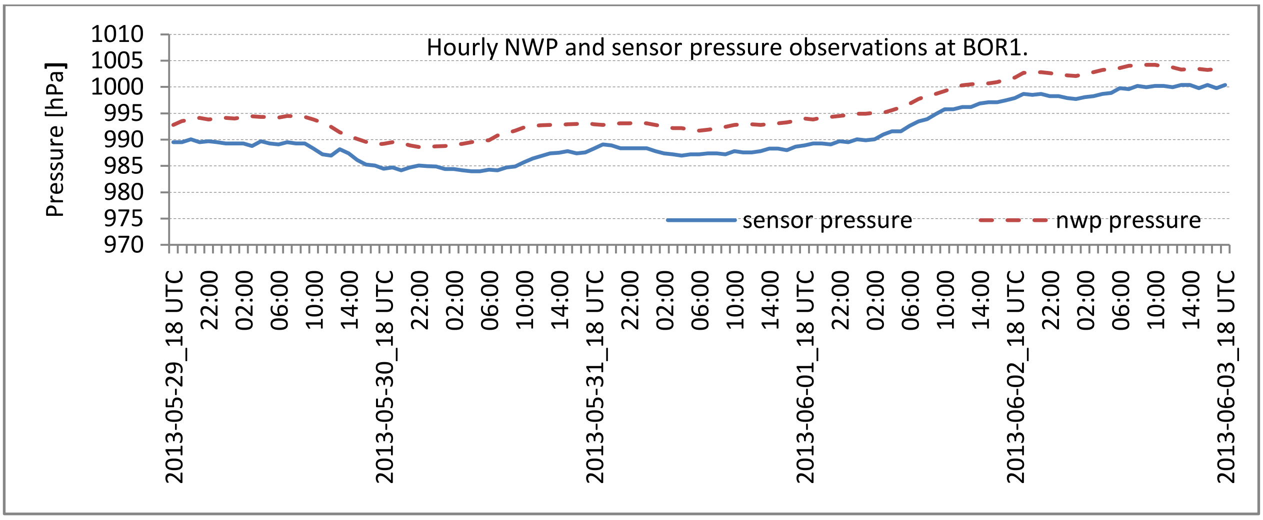


Figure 4.4: Trends in NWP and sensor pressure observations at BOR1.

BOR1, whose longitude/latitude coordinates are (17.07°, 52.1°)is geographically close to GWWL (15.21°, 52.74°) and JOZE(21.03°, 52.1°). Assuming the atmospheric conditions do not change much between the sites, their pressure observations can be compared.

In Figure 4.5, the hourly NWP and sensor pressure observations are seen to closely follow each other at GWWL and JOZE with mean values for both the sensor and NWP of approximately 996 hPa and 992 hPa at GWWL and JOZE respectively.

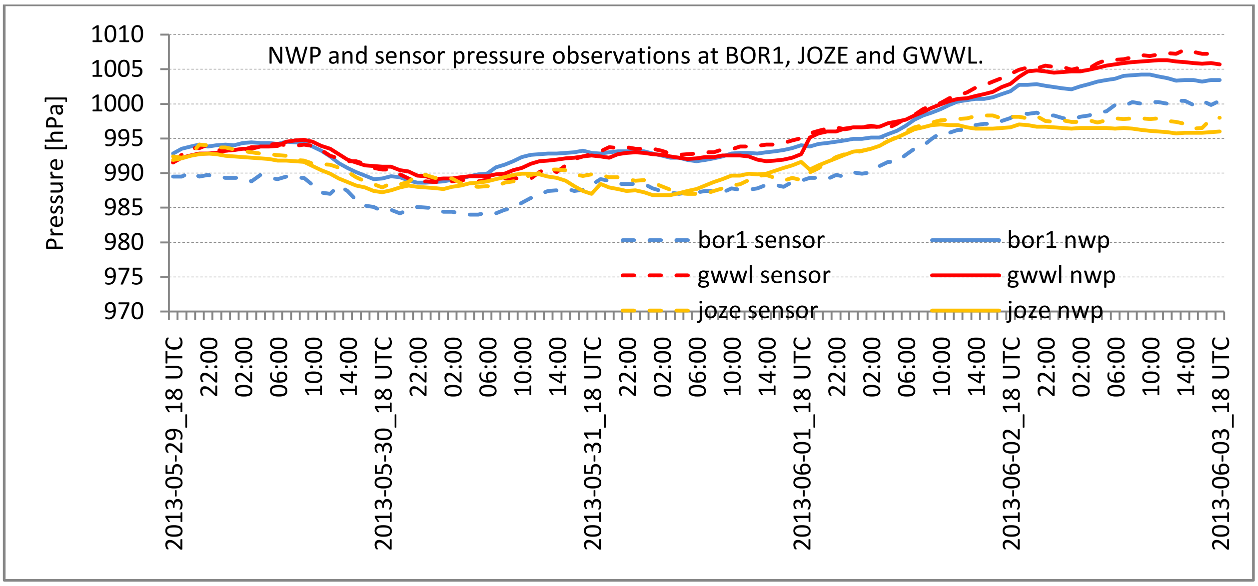


Figure 4.5: Hourly NWP and sensor pressure observations at BOR1, JOZE and GWWL.

Following the discussion in the preceding paragraphs, the pressure biases at BOR1 could either be originating from a malfunctioning (or uncalibrated) sensor or failure of the model to correctly estimate pressure at the site.

The overall precision error is considered small and the same result is reflected in the daily assessment of the standard deviations at most of the stations considered on each day of the study period (Figure 4.6). WTZR, however, shows a large precision error on doy 149/150.

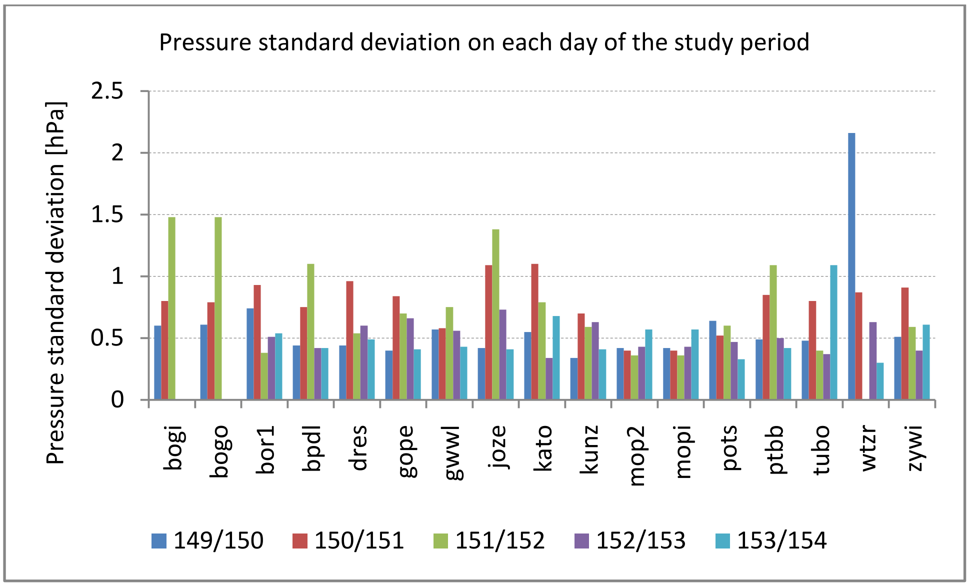


Figure 4.6: Mean pressure standard deviation errors on each day of the study period at all the stations.

An analysis on the errors in the surface pressure variable for a period of one month in January 2014 showed a mean bias error of approximately -0.03 hPa and a mean standard deviation error of approximately 1.75 hPa (results shown in Appendix F). The largest bias error during this period was seen at BYDG and it is seen to results from pressure sensor calibration problems (Figure F.1 in Appendix F).

4.6. Temperature error analysis

The mean temperature bias error at majority of the stations studied is less than 1 °C (Table 4.2). BOR1 shows a significantly large bias error with an overall mean of approximately 4.5 °C. The sensor lies within model orography (positive height difference of 16.4 m, see Section 4.4), hence reducing the chance of the error originating from the height difference. The overall mean bias error could, therefore, be due to the conspicuously large bias error of approximately 14 °C on doy 150/151 and relatively small, though significant, bias errors on the next two days (Figure 4.7).

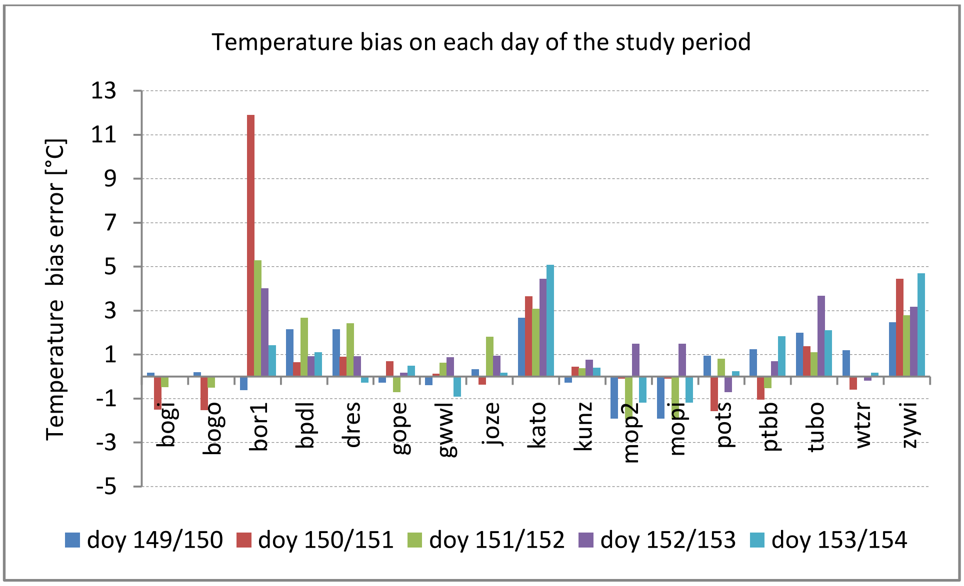


Figure 4.7: Mean temperature bias errors on each day of the study period at all the stations.

The cause for the large bias error at BOR1 was traced back to the temperature observations contained in the RINEX meteorological files bor11500.13m and bor11510.13m. Unrealistic observations are seen on 2013-05-30 at 22:00 UTC and 23:00 UTC, and on 2013-05-31 at 07:00 UTC, 12:00 UTC, 13:00 UTC, and 14:00 UTC (Figure 4.8). These observations fall on the period of time that defines doy 150/151.

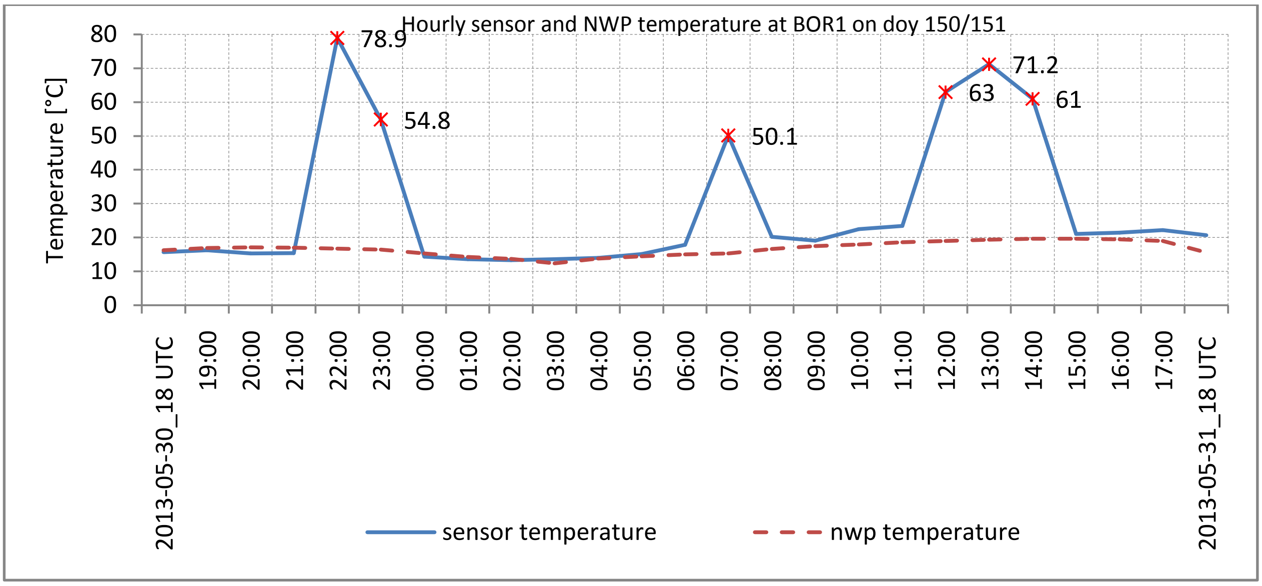


Figure 4.8: Comparison of sensor and NWP temperature observations at BOR1 between 2013-05-30_18 UTC and 2013-05-31_18 UTC.

It is clearly evident that this was a case of either a malfunctioning temperature sensor or a problem in the data handling procedure at BOR1 on doy 150/151 at the specified times that led to the erroneous observations.

KATO and ZYWI also show significantly large temperature biases of approximately 4 °C. It is possible that the error at ZWYI originates from the relatively large negative height difference seen in

Section 4.4. The height difference at KATO is not an issue; it is positive and relatively small (approximately 9 m). However, a study of the sensor and NWP temperature observations at the station shows significant differences in magnitude as shown in the hourly trend of Figure 4.9 with the NWP observations being significantly lower most of the times. This is the cause for the large overall mean bias error at the station.

The NWP simulation appears to have given lower temperature observations most of the times considered apart from about two instances noted, such as at the times between 2013-05-29 06 UTC and 2013-05-29 16 UTC, where the observations are almost equal in magnitude. The meteorological sensor, on the other hand, appears to be functioning normally. However, an error analysis in temperature variable in January 2014 seems to give a similar behavior as seen in Figure F.2. This together with a spike between 23rd and 26th could possibly point to problems in temperature sensor calibration at KATO.

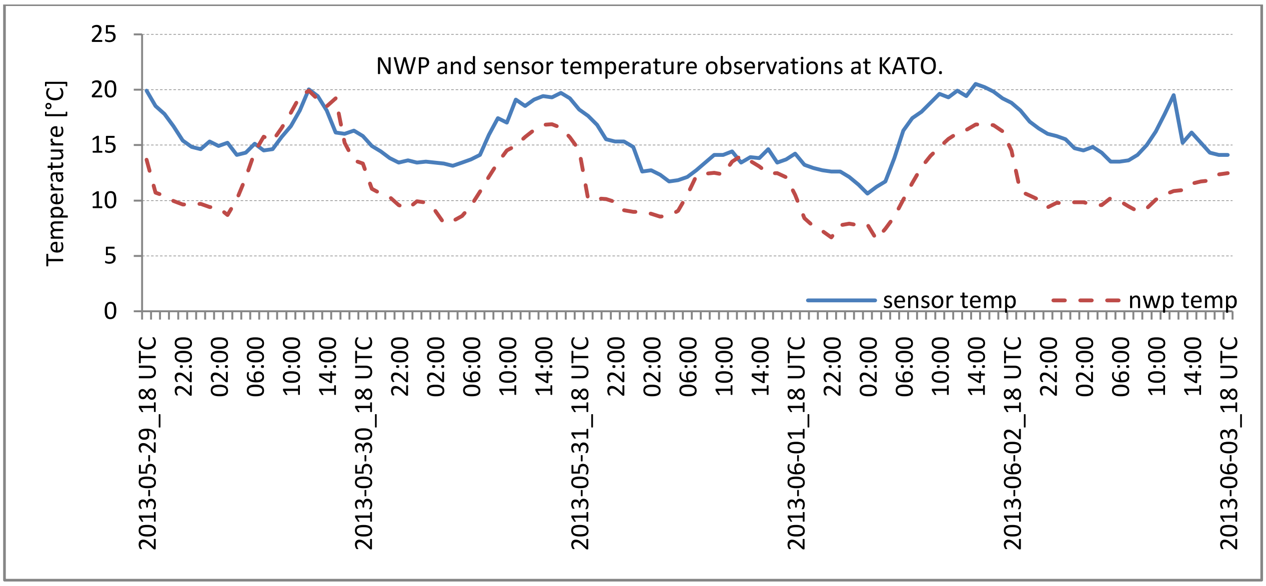


Figure 4.9: Temperature observations at KATO over the entire study period.

Precision errors at most of the stations are less than 2 °C. When compared to the sensor accuracy limits (0.2 °C to 0.5 °C), this seems to be significantly high. Once again, the precision error at BOR1 on doy 150/151 is significantly large (Figure 4.10).

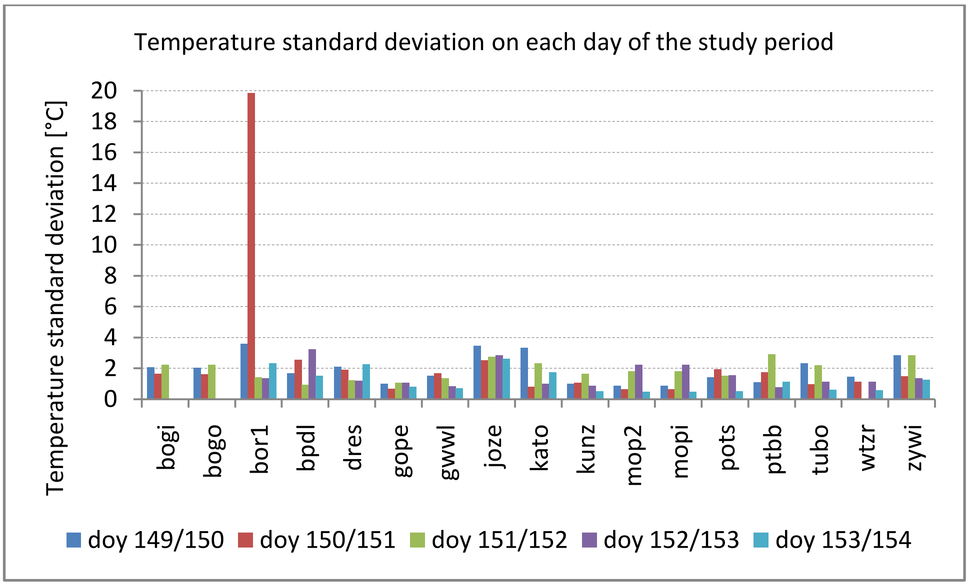


Figure 4.10: Mean temperature standard deviation errors on each day of the study period at all the stations.

An analysis on the errors in the near-surface temperature variable for a period of one month in January 2014 showed a mean bias error of approximately 1.3°C and a mean standard deviation error of approximately 3.0°C (results shown in Appendix F). The largest bias error during this period was seen at BYDG and it is seen to results from pressure sensor calibration problems (Figure F.2 in Appendix F).

4.7. Relative humidity error analysis

As shown in Table 4.2, the overall bias and standard deviation errors for the relative humidity variable are significantly large when compared to the 5% accuracy limit of the humidity sensors. While a number of stations are showing bias errors slightly greater than 10%, an almost equal number of stations are having bias errors less than 5% leading to an overall bias error of approximately 7%.

The precision errors are consistently large at all the stations with most of them showing standard deviations between 5% and 10% resulting to a mean value of approximately 9%.

Although the overall mean bias error was found to be approximately 7%, a study of the mean biases at each station during the study period reveals quite some large bias errors on some of the days studied with most of the biases being positive.

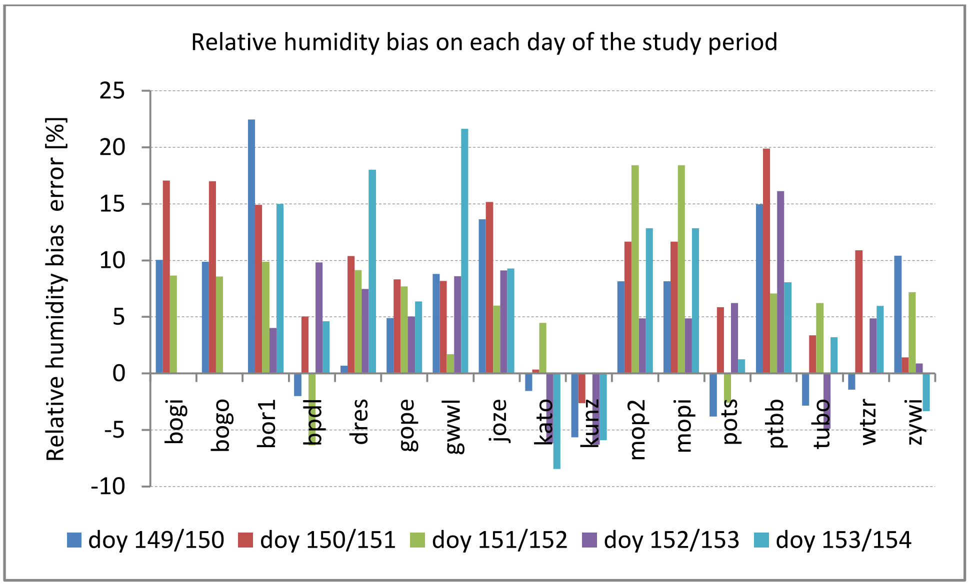


Figure 4.11: Mean relative humidity bias errors on each day of the study period at all the stations.

Significant mean bias errors were observed at BOGI, BOR1, JOZE, MOPI and PTBB. Humidity sensors at these sites are falling within the model orography; hence height difference is not an issue here.

Relative humidity observations at BOR1 show constant values (approximately 100%) in sensor observations most of the times during the study period as shown in Figure 4.12.

As shown in Figure 4.12, hourly trends in sensor and NWP relative humidity observations at BOR1 over the entire study period are largely similar. However, the NWP observations are smaller in magnitude with the sensor observations showing constant values (mostly 0.1 to 0.2% greater than 100%) most of the times, hence a relatively large overall mean bias error (of approximately 13% as seen in Table 4.2) when compared to the bias errors at the other sites. This was thought to originate from a problem in the calibration of the humidity sensor at the site.

NWP and sensor humidity observations at MOPI show relative humidity values 0.1 to 0.2 % greater than 100% most of the times (Figure 4.13). Unlike the case of BOR1 where the NWP observations were seen to be less than sensor observations, at MOPI the observations are either the same or NWP observations are slightly larger the sensor observations.

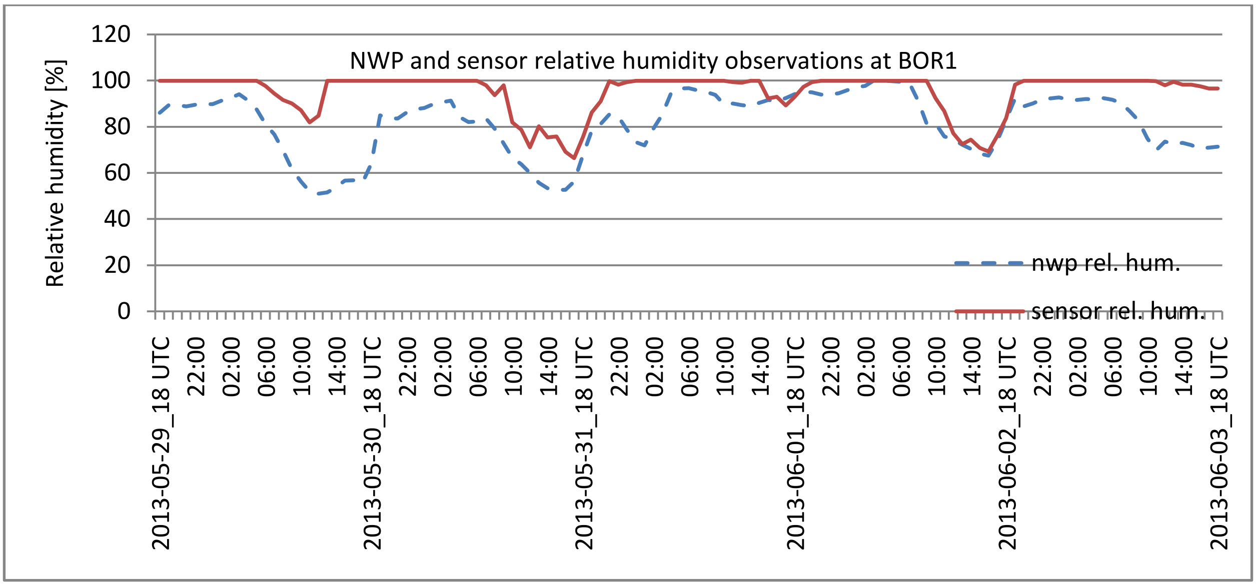


Figure 4.12: Hourly sensor and NWP relative humidity observations at BOR1 over the entire study period.

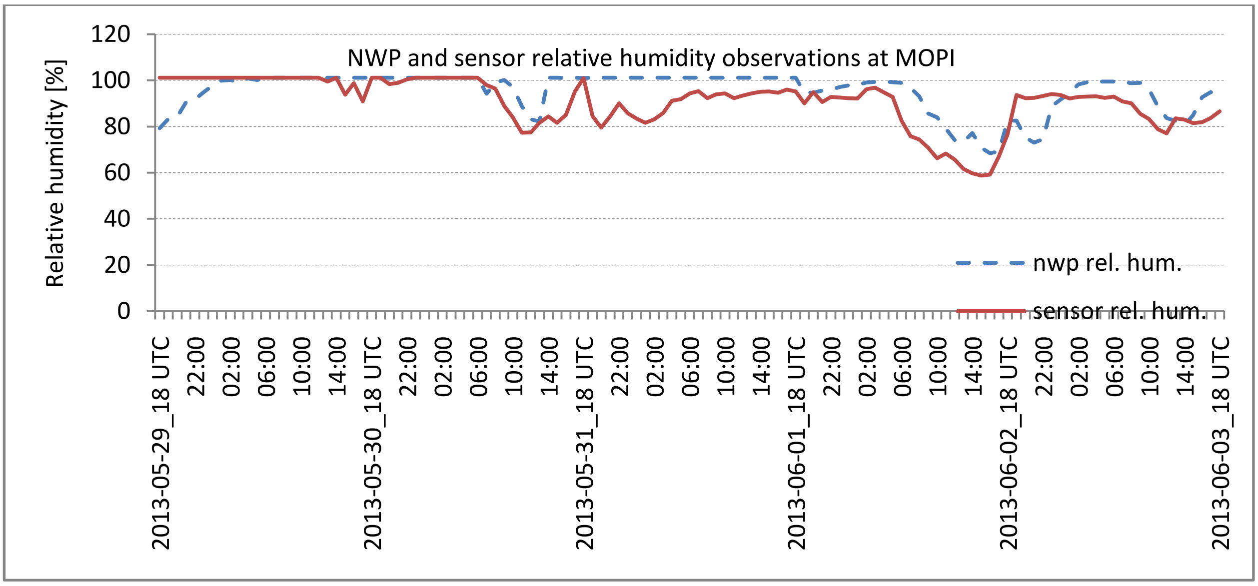


Figure 4.13: Hourly sensor and NWP relative humidity observations at MOPI over the entire study period.

From Figure 4.13, a larger bias is observed particularly on the last three days where the hourly trends seem to be somehow different. This could have led to the bias error at the site.

The bias errors in relative humidity at BOGI and JOZE are almost equal in magnitude as the error at MOPI. The conclusion made for MOPI can, therefore, be assumed for BOGI and JOZE.

In Figure 4.14, the NWP and sensor relative humidity observations at PTBB are shown. The hourly trends are not exactly similar. Moreover, the NWP observations are significantly less than sensor observations most of the times hence the relatively large bias error. It is difficult to say which of the two, model or sensor, gives the correct observations.

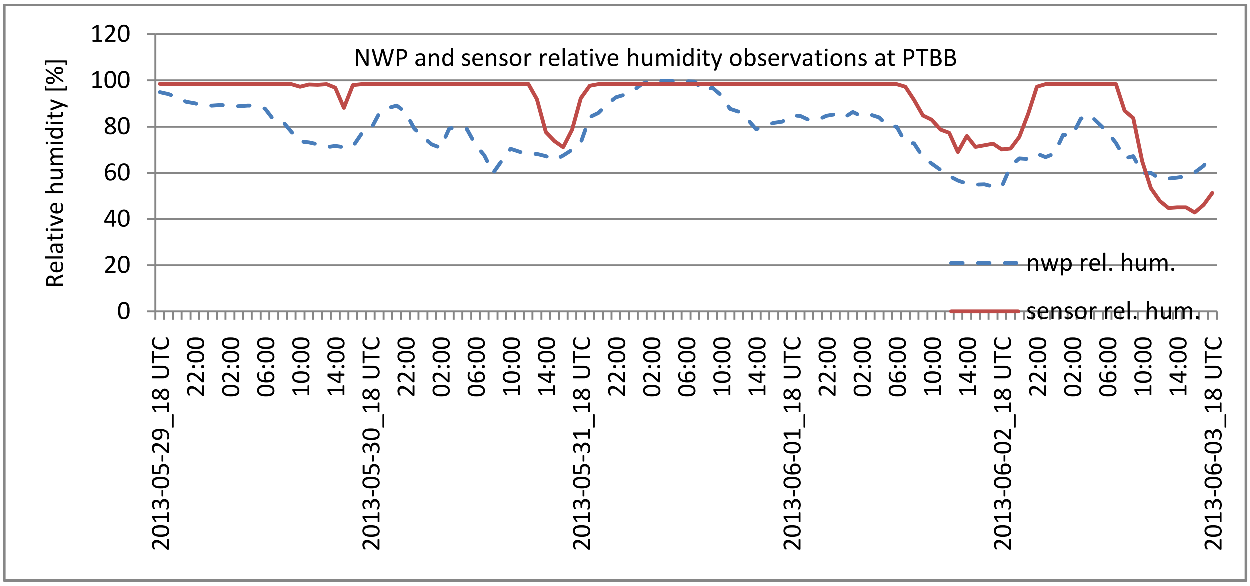


Figure 4.14: Sensor and NWP relative humidity observations at PTBB over the entire study period.

The overall mean standard deviation errors were found to be less than 10% at majority of the sites. JOZE shows mean precision errors of values greater than 10% on all the days considered (Figure 4.15).

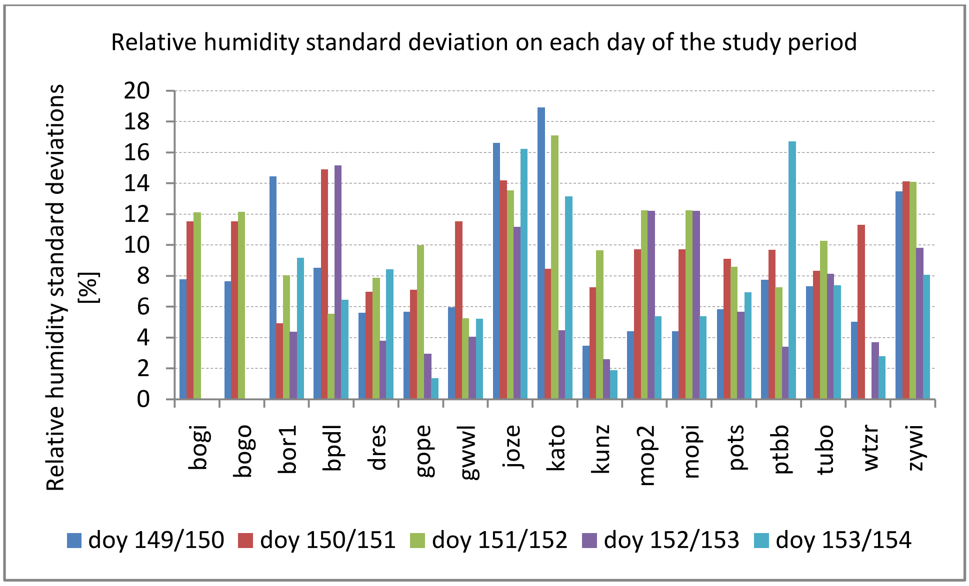


Figure 4.15: Mean relative humidity standard deviation errors on each day of the study period at all the stations.

An analysis on the errors in the near-surface relative humidity variable for a period of one month in January 2014 showed a mean bias error of approximately 3.0% and a mean standard deviation error of approximately 10% (results shown in Appendix F).

4.8. Comparison of NWP and radiosonde surface meteorological observations

Table 4.4 shows the overall mean results of the accuracy assessment of the NWP-derived surface meteorological variables after comparison with radiosonde data sourced from selected radiosonde launch sites (Table 4.1).

Table 4.4: Overall mean bias and standard deviation errors in surface meteorological observations at selected radiosonde stations.

Radiosonde WMO site identifier	pressure [hPa]		temperature [°C]		relative humidity [%]	
	bias	stdev	bias	stdev	bias	stdev
12425	0.2	1.0	-0.5	1.8	4.0	15.4
11747	-0.2	0.7	0.2	1.8	-2.7	9.5
11520	**-35.8**	1.0	-0.5	1.2	6.0	7.4
10771	-0.1	1.0	-0.2	1.8	3.3	5.6
10548	0.8	0.8	-0.5	1.6	2.7	6.3
10393	**-13.9**	0.8	-0.5	1.6	8.3	12.1
10238	0.6	0.9	-0.5	2.3	5.3	9.3
mean	**0.25**	**0.90**	**-0.34**	**1.75**	**3.86**	**9.39**

Some significantly large bias errors in pressure are seen at Praha-Libus (11520) and Lindenberg (10393). The NWP model seems to have over-estimated the pressure observations at two sites particularly at Praha-Libus as seen in Figure 4.16. The large bias errors were therefore thought to have originated from a modeling problem in the NWP model.

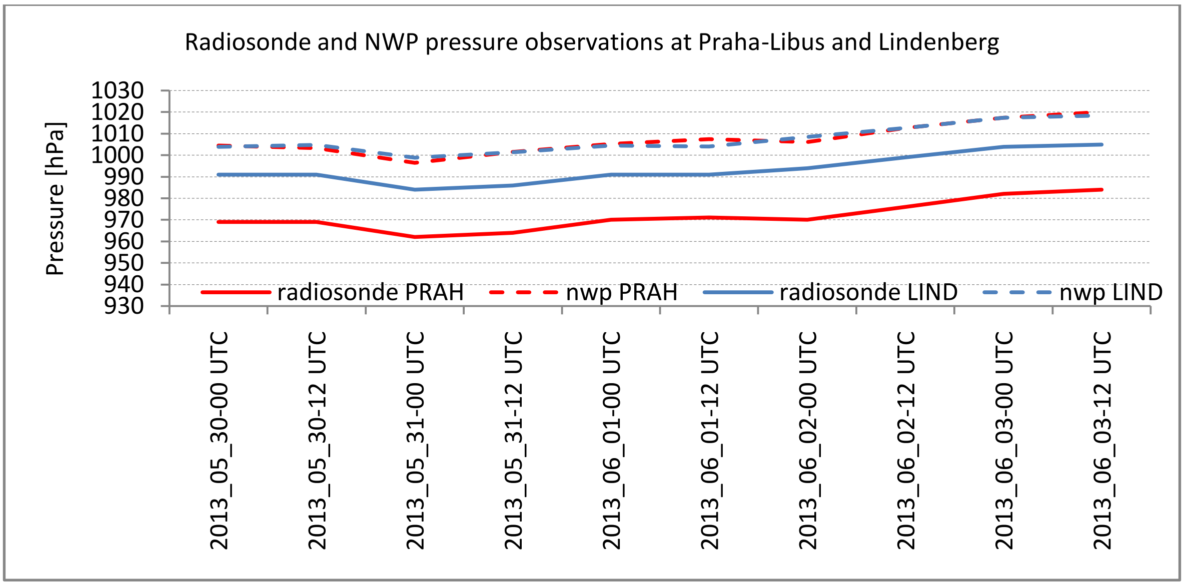


Figure 4.16: Radiosonde and NWP pressure observations at Praha-Libus and Lindenberg.

The pressure bias errors at the Praha-Libus and Lindenberg radiosonde sites were not included in the final analysis of the overall mean bias error in pressure. The overall mean bias and standard deviation errors in the three target variables at the selected radiosonde sites are therefore summarized as follows:

- surface pressure: approximately 0.3 hPa and 1 hPa respectively
- surface temperature: approximately 0.3 °C and 2 °C respectively
- near-surface relative humidity: approximately 2.5% and 10% respectively.

4.9. Summary

Accuracy assessment of the NWP model simulated variables of surface pressure, surface temperature and near-surface relative humidity was done at several locations falling within the study domain. The reference observations were sourced from meteorological sensors at EPN sites and radiosonde sites.

The overall mean bias and standard deviation errors respectively at the EPN sites were found to be approximately -0.2 hPa and 0.7 hPa for pressure, approximately 0.5 °C and 2.0 °C for temperature, and approximately 6.8% and 8.8% for relative humidity.

The station BOR1 featured a lot in the error analysis. The systematic errors in temperature were attributed to either a malfunctioning sensor or a data handling problem while the cause of the systematic errors in pressure was suspected to originate from an NWP modeling problem. The small pressure biases at most of the stations considered compensated for the large pressure bias at BOR1 leading to a small overall bias error. The humidity sensor at the site was thought to have undergone a calibration drift (i.e. it could have fallen out of calibration). Thus, explaining the largely constant relative humidity sensor observations slightly greater than 100%.

Several situations were noted in which the relative humidity was 0.1 to 0.2% greater than 100% during the study period and across the entire study domain. This period, May/June 2013, is known to have experienced extremely heavy precipitation that even lead to flooding in central Europe (probably due to the melting of snow following a prolonged winter). The NWP model, however, appears to have under-estimated the relative humidity observations at some times during the study period.

Precision seems to be the dominant error behavior for the three target variables. However, given that the precision with which a meteorological variable should be measured varies with the specific purpose for which the measurement is required, this precision is found adequate for onward application in the estimation of troposphere delays and precipitable water vapor.

Differences in height between the sensors’ and model orography’s positions were found to cause significant errors only when the difference is large and the sensor is below the model orography.

Similarly, at the radiosonde sites, the overall mean bias and standard deviation errors respectively were found to be approximately 0.3 hPa and 1 hPa for pressure, approximately 0.3 °C and 2 °C for temperature, and approximately 2.5% and 10%for relative humidity. Since the number of radiosonde sites was fewer than that of EPN sites, the reliability assessment was based on the validation with EPN meteorological data.

By assuming the EPN meteorological observations to be of high quality and drawing from the findings of the error analysis of the three meteorological variables considered in the validation study, a high level of reliability was developed for the NWP-derived surface and near-surface meteorological variables. The same reliability was therefore assumed for the meteorological variables to be derived from the upper NWP model levels. The accuracy levels for the three variables were therefore found to be sufficient for this study.

5. Accuracy assessment of NWP-derived troposphere delays

5.1. Introduction

The reliability of the NWP-derived near-surface meteorological data for use in the precise estimation of hydrostatic delays was ascertained in Chapter 4. The accuracy levels of the surface (or near-surface) observations were found to be sufficient for this study and the reliability achieved was also assumed for the observations coming from the vertical extent of the NWP model.

In the current study, zenith hydrostatic and wet delays were computed using NWP and radiosonde data. The accuracies of the resulting propagation delays were checked using the methods discussed in the next section.

5.2. Methodology

The configurations defining the computational domain for the NWP model simulation were discussed in Section 3.2.

GNSS antenna heights were used as the reference heights in the estimation of the troposphere delays using NWP data. However, elevation differences over the study domain are obviously expected to result into different magnitudes in troposphere delays at different regions across the domain. The terrain heights over the entire domain were therefore extracted and used to give an overview of the expected troposphere delays.

By computing troposphere delays at the EPN sites under study with and without the portion of the atmosphere above model orography, the effect of ignoring this portion was investigated and the results are discussed in this study.

It was mentioned earlier that meteorological sensors at EPN stations are mounted very close to GNSS antennae with just a few height differences at some stations. Therefore, the conclusions made from the comparison of sensor/model orography height differences can be assumed for GNSS antenna/model orography height differences (see Section 4.4).

Several EPN sites spread across the study domain and that happened to be equipped with meteorological sensors were selected for this study (Figure 5.1; see also a list of the stations in Appendix E). The sensor pressure at a given site and time was extracted from the corresponding RINEX meteorological file. This pressure quantity together with site parameters of height and latitude were then implemented into the Saastamoinen hydrostatic model (first term of Equation (2.46)) in order to compute the zenith hydrostatic delay at the site at the specified time. The EPN combined troposphere delay solution (which normally has a precision of 6 mm or better) at the site and time corresponding to those discussed in the preceding sentence was extracted from the corresponding

SINEX[1] troposphere delay file. The zenith wet delay at the site at the specified time was then computed by subtracting the Saastamoinen zenith hydrostatic delay from the EPN total zenith path delay. The zenith hydrostatic, wet and total delays discussed in this paragraph are henceforth abbreviated as $\mathrm{ZHD}_{\mathrm{saas}}$, $\mathrm{ZWD}_{\mathrm{epn/saas}}$ and $\mathrm{ZTD}_{\mathrm{epn}}$ respectively.

Another approach used to determine the zenith path delays involved the use of NWP-derived variables. Using the same site parameters and times as mentioned in the preceding paragraph, zenith hydrostatic and wet delays were estimated by numerically integrating the hydrostatic (Equation (2.7)) and wet refractivities (Equation (2.7)) respectively. The NWP-derived total zenith path delay was then computed as a sum of the NWP-derived zenith hydrostatic and wet delays. The zenith hydrostatic, wet and total delays estimated by the numerical integration method are henceforth abbreviated as $\mathrm{ZHD}_{\mathrm{num\_int}}$, $\mathrm{ZWD}_{\mathrm{num\_int}}$ and $\mathrm{ZTD}_{\mathrm{num\_int}}$ respectively where "num_int" stands for numerical integration.

Zenith path troposphere delay (hydrostatic, wet and total) data at several EPN sites computed using the TropGrid2 tropospheric model were also made available[2] for comparison with the troposphere delays that were discussed in the preceding two paragraphs. TropGrid2 models zenith hydrostatic and wet delays as special harmonic functions taking seasonal and diurnal variations into consideration. It uses a database consisting of more than 9 years of 3D NWP fields of the GDAS weather model (Schueler, 2014). Surface fluxes are used to compute the grid-point values that are height reduced and can be interpolated to user position. A more detailed explanation can be found in e.g. Schueler *et al.*(2002), Pósfay *et al.*(2003) and Schueler (2014).

Data from the TropGrid2 model came in two sets; one set comprising delays computed using the weighted mean temperature model comprising both the linear and periodic terms that were mentioned in Section 2.2, and a second set that did not comprise the linear term. The two sets are henceforth referred to as the extended mean temperature model (*ext*) and standard processing (*std*) respectively.

For comparison purposes, a reference troposphere delay (hydrostatic, wet and total) had to be chosen. This quantity would act as a basis upon which the other delays (preferably *num_int*, *ext* and *std*) would be compared. The $\mathrm{ZHD}_{\mathrm{saas}}$, $\mathrm{ZWD}_{\mathrm{epn/saas}}$ and $\mathrm{ZTD}_{\mathrm{epn}}$ were taken as the reference quantities and the accuracies and precisions in the *num_int*, *ext* and *std* quantities were therefore computed by comparison with the reference quantities in terms of bias, bias-reduced standard deviation and root mean square (rms) error computations using the formulas that are shown in Appendix D.

---

[1] Software Independent EXchange file format.

[2] Courtesy of Prof. Torben Schueler.

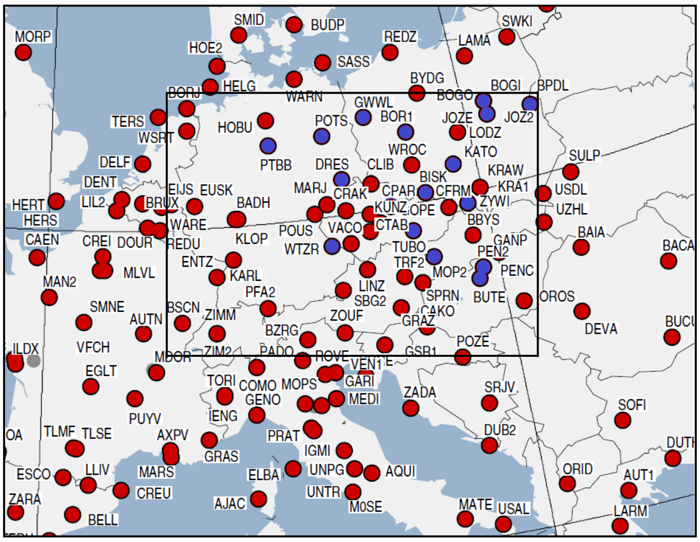


Figure 5.1: EPN stations from which GPS ZTD observations were sourced are marked in blue. Inset rectangle shows the study domain.

Finally, by numerically integrating hydrostatic and wet refractivities determined using radiosonde data, zenith hydrostatic and wet delays (and the corresponding total path zenith delays) were estimated at several radiosonde locations falling within the study domain. Using a similar method the delays at the radiosonde sites were computed using NWP data. The resulting NWP/radiosonde troposphere delay estimates were then compared in terms of bias, bias-reduced standard deviation rms computations.

5.3. Terrain heights across the study domain

A study of the terrain through the terrain heights would serve to provide an overview of the expected troposphere delays. Terrain variations are seen particularly in the lower left region of the study domain due to inclusion of the Alps (Figure 5.2). These regions are characterized by high altitudes.

Most of the other regions average between roughly 150 and 700 m in height. We, therefore, generally expect the least amounts of troposphere delays over the Alps and the surrounding regions, and greater amounts over the regions covered in a blue shade mostly in the upper and lower-right parts of the domain.

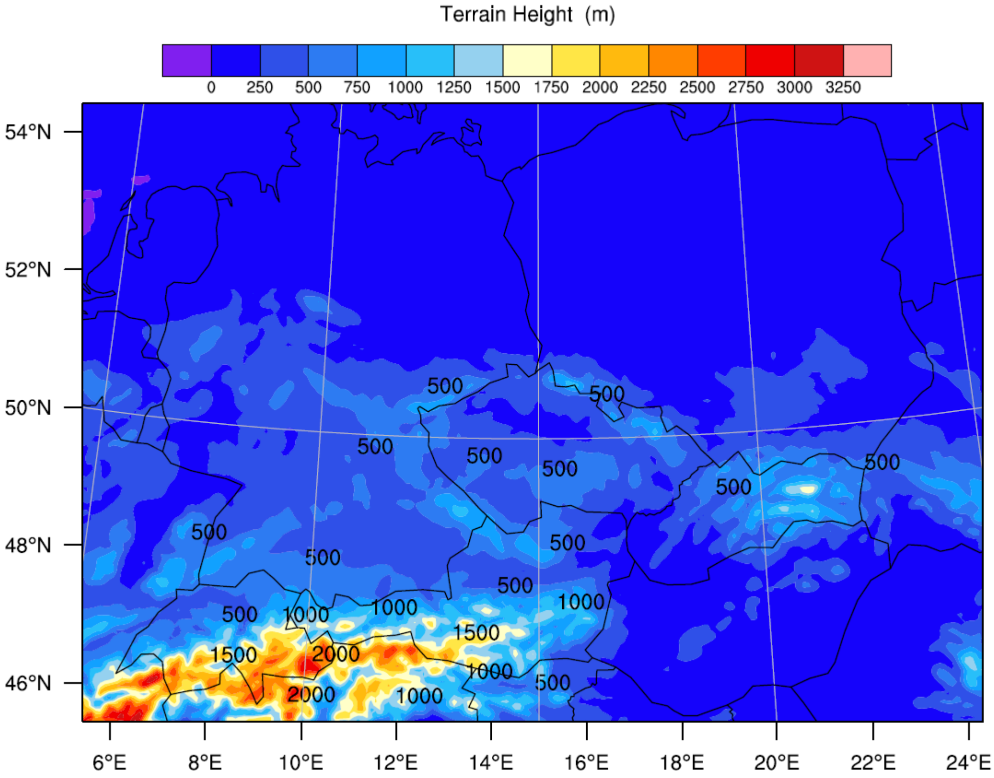


Figure 5.2: NWP terrain heights simualtion over the entire study domain for 29th May 2013 18 UTC.

5.4. Contribution of the upper atmosphere to troposphere delays

Upper atmosphere here refers to the portion above the NWP model top as was discussed in Section 2.4.1. The effect of ignoring this portion was investigated with regards to the magnitudes of the NWP-derived zenith hydrostatic and wet delays at the EPN sites under study.

Consistent with the discussion of Section 2.4.1, it was established that ignoring the upper atmosphere resulted into NWP-derived ZHD estimates at all the EPN sites considered that were smaller in magnitude by approximately 0.12 m. The NWP-derived ZWD estimates, on the other hand, were not affected by omission of the upper atmosphere which was attributed to the fact that the atmospheric water vapor content is mainly significant in the lower atmosphere.

The overall effect of omitting the upper portion in the computation of troposphere delays was therefore reflected into the ZTD estimates as shown in Figure 5.3. The ZTD estimates computed without the upper atmosphere are seen to be smaller by approximately 0.1 m at all the EPN sites considered.

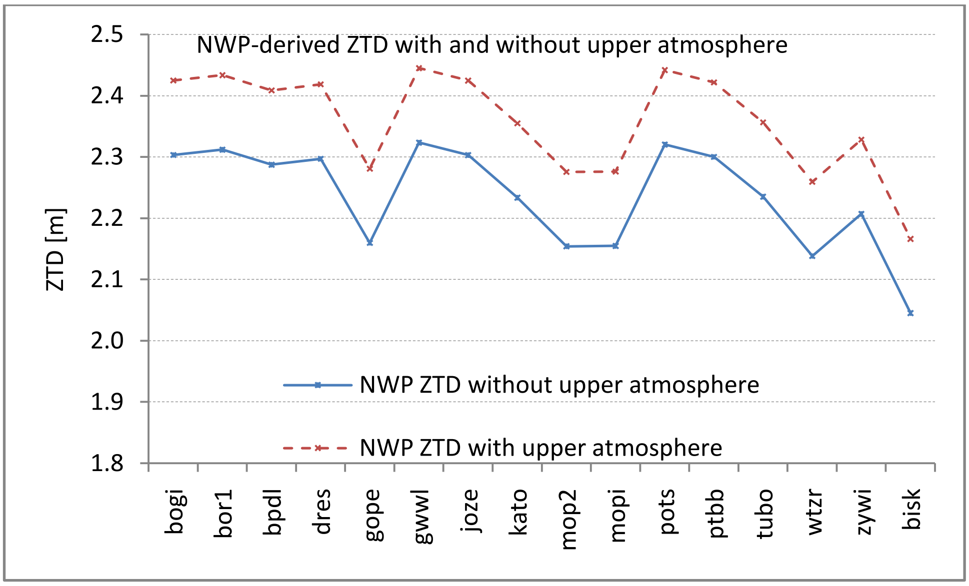


Figure 5.3: Overall mean NWP ZTD estimates (with and without the upper atmosphere) all the EPN sites under study.

Figure 5.3 also serves to compare the magnitudes of the NWP-derived ZTD estimates at the EPN sites under study. The largest values are seen mainly at POTS, GWWL and BOR1 while the least values are seen at WTZR and BISK.

The zenith hydrostatic, wet and total path delays are each analyzed in the sections that follow.

5.5. Zenith hydrostatic delay analysis

The total zenith hydrostatic delay (ZHD) at a given location is a function of the local atmospheric pressure. Its magnitude above every point on the surface is, therefore, largely dependent on elevation as illustrated in Figure 5.4.

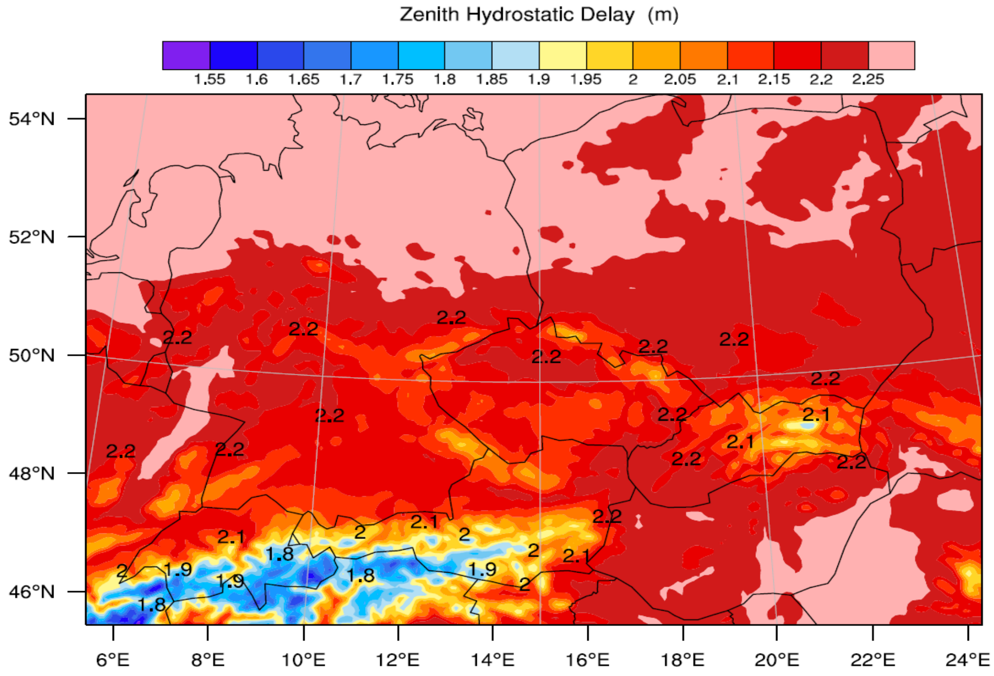


Figure 5.4: Contour map of NWP ZHD simulations for 30th May 2013 18 UTC over the entire study domain.

The least zenith hydrostatic delays averaging about 1.8 m are seen in a region around the Alps where the elevation is highest, while the greatest amounts averaging about 2.3 m are seen mostly in the upper and lower-right parts that are characterized by low elevations.

ZHD estimates at selected EPN sites showed an overall mean value of approximately 2.2 m when computed using each of the four methods mentioned in the methodology (Section 5.2), .i.e. the Saastamoinen ZHD (ZHD_{saas}), the numerical integration ZHD (ZHD_{num_int}), the extended mean parameter model ZHD (ZHD_{ext}) and the standard processing ZHD (ZHD_{std}). A comparison of these ZHD estimates at each of the selected sites is shown in Table 5.1. ZHD_{saas} was used as the reference value so that any other ZHD observation was subtracted from it in order to compute the bias.

Table 5.1: Comparison results of zenith hydrostatic delay at selected EPN sites.

		ZHD_{saas} [m]	**ZHD bias [mm]**		**ZHD stdev [mm]**		**ZHD rms [mm]**	
Stn id.	**Samples**		**num_int**	**ext/std**	**num_int**	**ext/std**	**num_int**	**ext/std**
BOGI	75	2.2539	0.87	-0.91	1.58	0.88	1.80	1.28
BOR1	125	2.2543	**-9.18**	**-7.88**	1.41	1.03	**9.29**	**7.95**
BPDL	125	2.2417	-0.09	-0.28	1.40	1.03	1.65	1.12
DRES	125	2.2513	0.66	-0.50	1.38	0.91	1.92	1.14
GOPE	125	2.1473	0.73	0.58	1.37	0.85	1.73	1.14
GWWL	125	2.2666	0.89	0.76	1.32	0.92	2.09	1.22
JOZE	125	2.2581	1.61	-0.02	1.84	1.42	2.71	1.57
KATO	125	2.2085	-0.86	-0.15	1.56	0.77	1.91	0.80
MOP2	125	2.1482	0.97	2.54	1.00	0.95	1.46	2.72
MOPI	125	2.1482	0.43	2.00	1.00	0.95	1.23	2.22
POTS	125	2.2706	**4.05**	**4.61**	1.16	0.82	**4.24**	**4.68**
PTBB	125	2.2778	-0.51	0.97	1.52	1.23	1.99	1.61
TUBO	125	2.2134	-0.73	-2.05	1.44	1.01	1.65	2.29
WTZR	125	2.1355	1.91	2.47	1.22	0.74	2.44	2.59
ZYWI	125	2.1885	-0.62	-2.83	1.38	0.89	1.66	2.97

Though the overall mean value in the ZHD estimates computed using the four approaches was approximately the same (2.2 m), there were site-to-site differences as reflected in the bias results in Table 5.1. ZHD estimates from the *ext* and *std* methods were found to be similar and are therefore analyzed together.

The *ext*/*std* bias errors were found to be largely negative and slightly larger than those of *num_int* at some of the sites considered with the overall mean biases of approximately -0.05 mm and 0.01 mm for the *ext*/*std* and *num_int* methods respectively. Two conspicuously large bias errors were also observed at BOR1 and POTS for both the *ext*/*std* and *num_int* methods.

While the cause for the large bias error at BOR1 could be traced back to the sensor pressures (refer to the discussion in Section 4.5), that for POTS had nothing to do with the sensor pressures as shown in the sensor and NWP pressure time series (Figure 5.5).

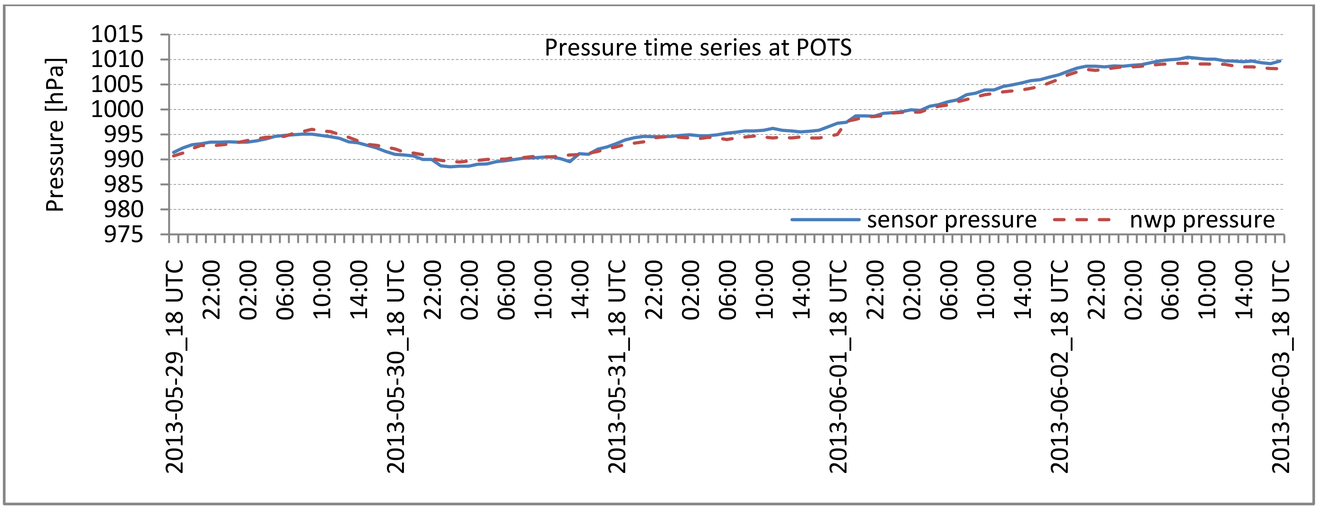


Figure 5.5: Trends in sensor and NWP (*num_int*) pressure at POTS over the entire study period.

From Figure 5.5, the hourly trends of the sensor and NWP pressure estimates at POTS are largely similar and, additionally, their overall mean values over the entire period are 998.0 hPa and 997.7 hPa respectively, a difference of about 0.3 hPa. The errors were therefore thought to be due to a large difference in ZHD computations at the site by the NWP models compared to the Saastamoinen (using sensor pressure as input).

The sites, MOP2 and MOPI, are more or less at the same geographical location with a very small difference in their ellipsoidal heights (Appendix E). Results from these two sites are therefore expected to be essentially the same or if not the same the difference should be very small. However, the results in Table 5.1 show some differences in the errors at the two sites.

The overall mean Saastamoinen ZHD estimates at MOP2 and MOPI are both approximately 2.1482 mm (because sensor pressures at the two sites as extracted from their corresponding RINEX meteorological files were the same) and slightly greater than the *num_int*, *ext* and *std* ZHD estimates. Hence the positive biases. Additionally, the bias errors in ZHD at MOPI were found to be approximately 0.54 mm greater than those at MOP2 for both the *ext*/*std* and *num_int* methods. This could be attributed to the fact that while the site pressures (at MOP2 and MOPI) implemented in the Saastamoinen should be assumed to be the same (despite the small height difference at the two sites), the *ext*/*std* and *num_int* methods, on the other hand, use slightly different pressure values for the heights corresponding to the two sites.

The mean standard deviations vary between approximately 1.0 mm and1.58 mm for *num_int* and between 0.74 mm and 1.42 mm for *ext*/*std*. The *ext*/*std* methods showed a slightly greater precision in ZHD than *num_int* with an overall mean standard deviation of approximately 1.0 mm for *ext*/*std* and approximately 1.4 mm for *num_int*.

The observed bias errors are manifested in the rms errors and hence the large rms errors at BOR1 and POTS. The overall mean rms errors were found to be approximately 2.4 mm for *ext/std* and approximately 2.5 mm for *num_int*.

A graphical overview of the bias and standard deviation errors in ZHD as discussed in the preceding paragraphs is shown in Figure 5.6.

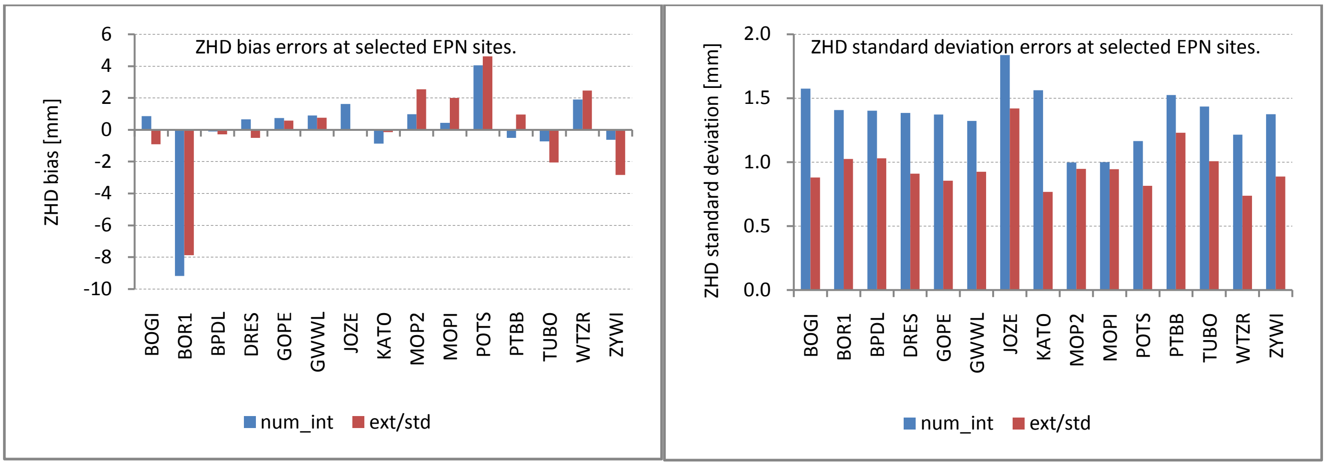


Figure 5.6: Mean accuracies and precisions in ZHD at selected EPN sites.

5.6. Zenith wet delay analysis

The total NWP zenith wet delay simulations for 30th May 2013 18 UTC (same time as that of ZHD in Figure 5.4) across the entire study domain are shown in Figure 5.7.

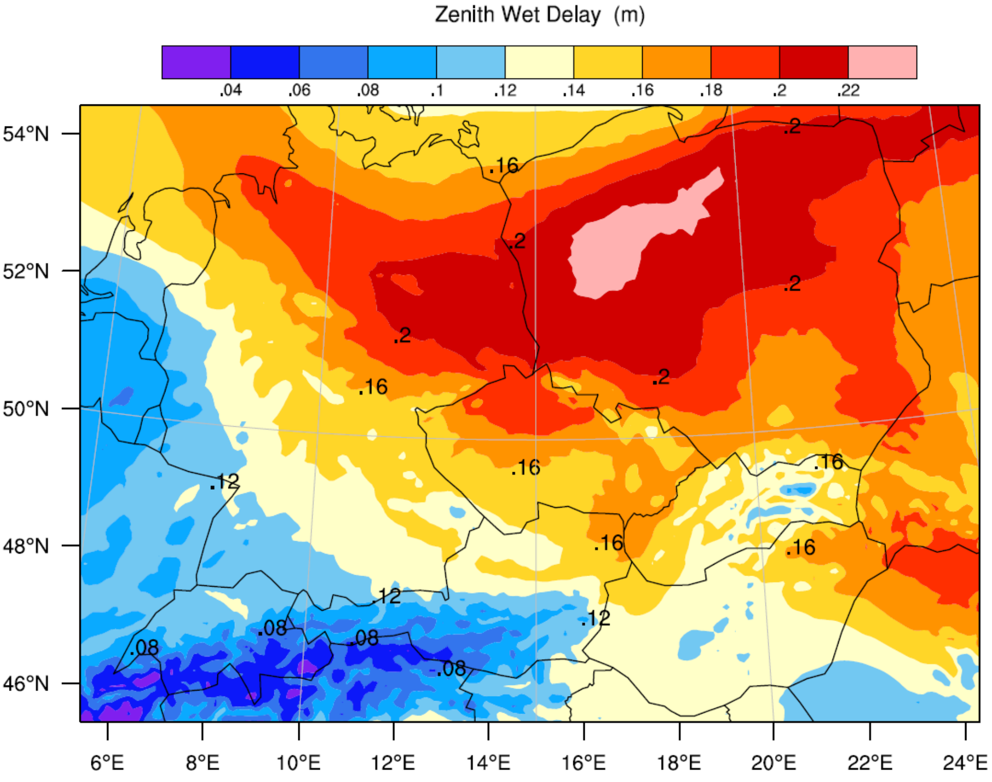


Figure 5.7: Contour map of NWP ZWD simulations for 30th May 2013 18 UTC over the entire study domain.

The lower-left region shows the least amounts of ZWD, approximately 0.06 m delay, particularly in an area surrounding the Alps. The highest amounts of ZWD average about 0.18 m as can be seen in the lower-right region.

Expressed to two significant figures, ZWD$_{epn/saas}$ estimates (ZTD$_{epn}$ minus ZHD$_{saas}$), ZWD$_{ext}$ and ZWD$_{std}$ were all found to amount to an overall mean value of approximately 0.1 m, while ZWD$_{num\_int}$ resulted to an overall mean value of approximately 0.2 m. In Table 5.2, a comparison of the ZWD$_{epn/saas}$ with ZWD$_{num\_int}$, ZWD$_{ext}$ and ZWD$_{std}$ at the selected sites is shown, where ZWD$_{epn/saas}$ was used as the reference value.

Table 5.2: Comparison results of zenith wet delay at selected EPN sites.

		ZWD$_{epn/saas}$ [m]	**ZWD bias [mm]**			**ZWD stdev [mm]**			**ZWD rms [mm]**		
Stn id.	**Samples**		**num_int**	**ext**	**std**	**num_int**	**ext**	**std**	**num_int**	**ext**	**std**
BOGI	75	0.1649	-9.08	-0.39	-3.44	18.42	8.57	8.92	20.26	9.13	9.77
BOR1	125	0.1631	-6.56	-0.59	2.05	**28.11**	**27.43**	12.23	**29.86**	**28.55**	13.12
BPDL	125	0.1585	-8.34	-1.60	-3.84	13.08	9.05	9.11	16.20	9.39	10.10
DRES	125	0.1589	-8.94	-7.34	-7.31	8.81	7.26	7.61	12.84	11.17	11.28
GOPE	125	0.1226	-11.60	-4.12	-3.85	10.56	6.22	5.97	17.54	7.38	7.03
GWWL	125	0.1729	-6.24	3.44	1.40	10.82	10.75	10.95	15.70	12.17	12.19
JOZE	125	0.1569	-11.07	-2.15	-4.63	18.33	12.94	12.95	21.54	14.03	14.66
KATO	125	0.1407	-4.72	-0.63	-1.60	15.37	14.21	14.10	18.06	14.47	14.58
MOP2	125	0.1277	-0.57	11.55	8.24	17.90	16.52	7.80	20.04	20.51	11.60
MOPI	125	0.1102	**-18.12**	-6.01	-6.46	8.95	7.28	7.00	20.38	9.76	9.82
POTS	125	0.1621	-13.00	-4.75	-6.15	10.94	8.95	9.20	17.38	10.41	11.31
PTBB	125	0.1442	0.90	6.93	6.20	8.40	7.79	7.96	10.64	11.34	10.95
TUBO	125	0.1386	-3.69	2.38	2.34	14.13	13.02	12.66	15.99	13.49	13.24
WTZR	125	0.1208	-6.25	-3.38	-2.42	8.57	6.13	6.19	13.84	8.75	8.39
ZYWI	125	0.1350	-7.03	-1.81	-6.50	13.75	20.48	10.78	16.21	22.22	12.87

The bias errors were found to negative at most of the sites considered for all the methods used. The magnitude of the mean bias errors were found to vary from 0.6 mm to 18.0 mm for *num_int*, 0.4 mm to 12.0 mm for *ext* and 1.4 mm to 8.2 mm for *std*. The overall mean bias errors were found to be largely negative with magnitudes of approximately 7.6 mm, 0.6 mm and 1.7 mm for *num_int*, *ext* and *std* methods respectively. The *ext* and *std* ZWD computations were therefore found to be more accurate in terms of bias than the *num_int* ZWD computations.

The overall mean standard deviations were found to be approximately 14.0 mm, 12.0 mm and 10.0 mm for the *num_int*, *ext* and *std* ZWD computations respectively. BOR1 showed the least precision in *num_int*- and *ext*-computed ZWD with standard deviations of approximately 28.0 mm and 27.0 mm respectively. The rms error was also found to be largest at this site for *num_int*- and *ext*-computed ZWD with values of approximately 30.0 mm and 29.0 mm respectively. The reason for this could be the sensor calibration/data handling issues at the site for the period under study (as mentioned in Sections 4.6 and 4.7).

The large *num_int* bias error in ZWD at MOPI was attributed to a computational error arising from either some inconsistencies or missing values in the EPN ZTD data. This is further illustrated in Figure 5.10.

Figure 5.8 gives a graphical view of the accuracies and precisions in the ZWD estimates as discussed in the preceding paragraphs.

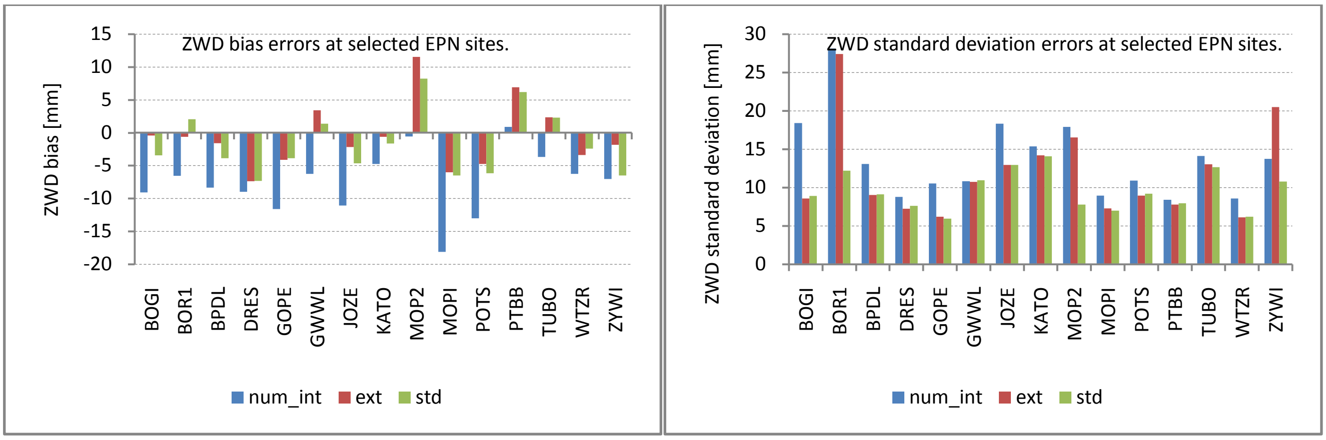


Figure 5.8: Mean accuracies and precisions in ZWD at selected EPN sites.

5.7. Total zenith path delay analysis

The total zenith path delay is essentially a sum of the NWP-derived hydrostatic and wet delays. Henceforth denoted as $\mathrm{ZTD}_{\mathrm{num\_int}}$, its mean estimates for 30th May 2013 18 UTC (same time as that of ZHD and ZWD in Figure 5.4 and Figure 5.7 respectively) across the entire study domain as simulated by the NWP model are shown in Figure 5.9.

Like ZHD and ZWD, the magnitude of the ZTD is also dominated by the terrain heights of the area covered. High ZTD regions average approximately 2.4 m delay and can be seen in the upper and lower-right parts of the study domain. Low ZTD regions are found around the Alps and the surrounding areas and average approximately 1.8 m delay.

The ZTD estimates at selected EPN sites showed an overall mean value of approximately 2.3 m for $\mathrm{ZTD}_{\mathrm{epn}}$, $\mathrm{ZTD}_{\mathrm{num\_int}}$, $\mathrm{ZTD}_{\mathrm{ext}}$ and $\mathrm{ZTD}_{\mathrm{std}}$. A comparison of these ZTD estimates at each of the selected sites is shown in Table 5.3 where $\mathrm{ZTD}_{\mathrm{epn}}$ was used as the reference value.

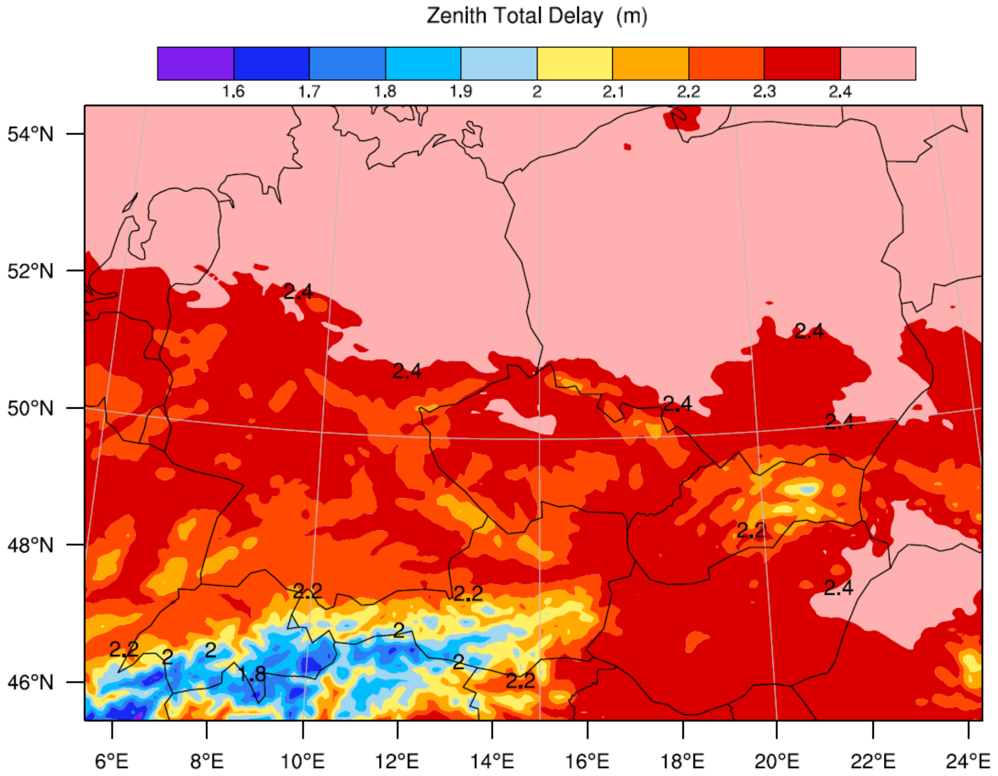


Figure 5.9: Contour map of NWP ZTD simulations for 30th May 2013 18 UTC over the entire study domain.

Table 5.3: Comparison results of total zenith path delay at selected EPN sites.

Stn id.	Samples	ZTD$_{epn}$ [m]	ZTD bias [mm]			ZTD stdev [mm]			ZTD rms [mm]		
			num_int	ext	std	num_int	ext	std	num_int	ext	std
BOGI	75	2.4187	-8.21	-1.31	-4.36	17.70	8.66	9.00	19.27	9.33	10.20
BOR1	125	2.4174	-15.75	-8.48	-5.91	**28.19**	**27.39**	12.10	**33.24**	**28.77**	13.98
BPDL	125	2.4001	-8.43	-1.87	-4.12	12.75	9.07	9.07	16.18	9.50	10.25
DRES	125	2.4101	-8.28	-7.85	-7.83	8.47	7.26	7.68	12.23	11.46	11.60
GOPE	125	2.2699	-10.87	-3.54	-3.27	9.85	6.26	5.98	16.38	7.11	6.72
GWWL	125	2.4395	-5.35	4.21	2.17	10.57	10.78	10.98	15.71	12.49	12.38
JOZE	125	2.4150	-9.46	-2.17	-4.65	18.13	12.67	12.68	20.64	13.73	14.35
KATO	125	2.3492	-5.57	-0.78	-1.75	14.87	14.24	14.13	17.53	14.50	14.63
MOP2	125	2.2758	0.40	14.10	10.60	18.17	16.60	7.86	19.88	22.27	13.39
MOPI	125	2.2584	**-17.69**	-4.01	-4.47	9.20	7.53	7.19	20.10	8.83	8.78
POTS	125	2.4327	-8.95	-0.14	-1.54	11.07	9.00	9.24	15.04	9.43	9.79
PTBB	125	2.4221	0.40	7.89	7.16	8.24	7.95	8.12	10.05	12.08	11.67
TUBO	125	2.3520	-4.41	0.32	0.29	13.78	13.28	12.88	15.69	13.52	13.23
WTZR	125	2.2535	-5.82	-2.52	-1.57	8.28	5.65	5.71	12.11	6.66	6.49
ZYWI	125	2.3119	**-16.32**	**-18.09**	-9.36	**36.86**	**34.46**	11.18	**40.39**	**39.06**	14.91

As shown in Table 5.3, the accuracy and precision errors in ZTD at the selected EPN sites follow what has been seen in the error analysis of Sections 5.5 and 5.6. However, the cause for the errors at ZYWI was found to originate from inconsistencies in the EPN ZTD solutions at certain times during the study period particularly on day-of-year 151/152. The same is observed with EPN ZTD solutions at MOPI which is essentially at the same location as MOP2 (Figure 5.10).

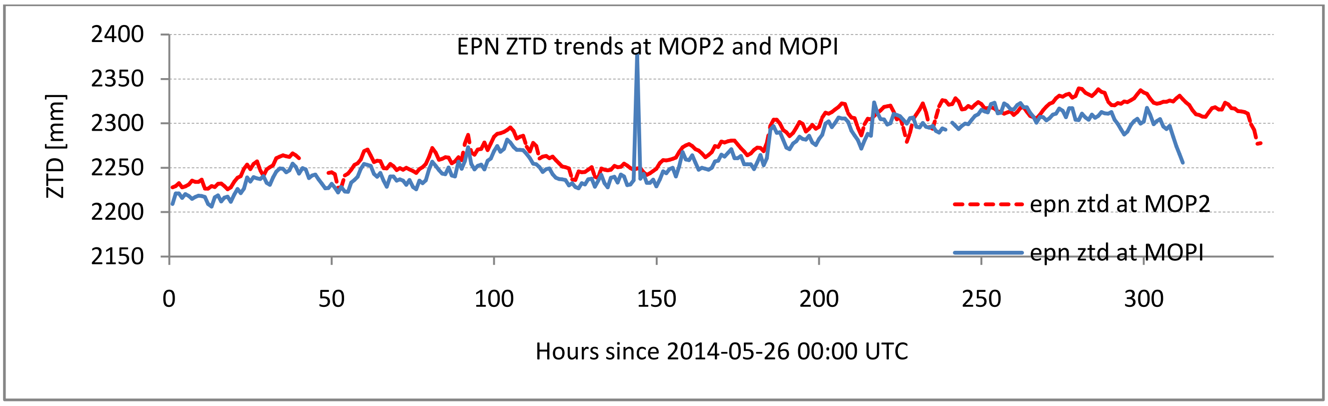


Figure 5.10: ZTD time series at MOP2 and MOPI. The spike corresponds to day-of-year 150.

A graphical overview to summarize the discussions in the preceding paragraphs is shown in Figure 5.11.

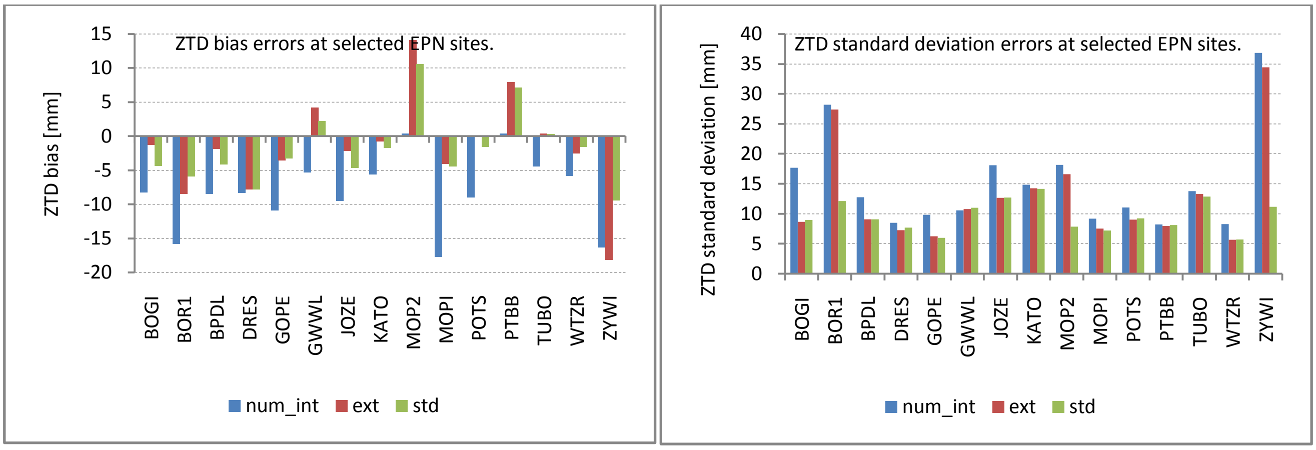


Figure 5.11: Mean accuracies and precisions in ZTD at selected EPN sites.

5.8. Comparison of NWP- and radiosonde-derived troposphere delays

Synoptic stations can also be used to source meteorological data necessary for estimating troposphere delays. Radiosonde observations from the radiosonde sites of Table 4.1 were used to estimate troposphere delays by numerical integration of refractivity along the radiosonde profiles. Using the same site parameters of coordinates and heights, troposphere delays were also estimated from the NWP model by numerical integration of refractivity along the model pressure levels.

The maximum radiosonde ascent height (approximately 30 km) was found to be greater than the approximate height of the NWP model top of about 20 km. The same depth of the atmosphere (approximately 20 km) was therefore used to compute troposphere delays from the NWP model and radiosonde profiles. Consequently, the following modified forms of Equation (2.39) and (2.40) were used to compute the NWP-derived ZHD and ZWD estimates respectively by leaving out the upper components of the delays,

$$ZHD_{NWP} \approx 10^{-6} \cdot k_1 \cdot R_d \cdot \frac{1}{g_1} \cdot p_1 + 10^{-6} \cdot k_1 \cdot R_d \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \Delta p^h \tag{5.1}$$

$$\begin{aligned} ZWD_{NWP} \approx\; & 10^{-6} \cdot \frac{R}{M_w} \cdot \frac{1}{g_1} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_1} \right) \cdot \frac{r_1}{r_1 + 1} \cdot p_1 \\ & + 10^{-6} \cdot \frac{R}{M_w} \cdot \sum_{j=1}^{m} \frac{1}{g_j} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_j} \right) \cdot \frac{r_j}{r_j + 1} \cdot \Delta p^h \end{aligned} \tag{5.2}$$

Where, all the terms are as earlier defined.

On the other hand, radiosonde ZHD and ZWD estimates were computed by using the following approximations,

$$ZHD_{radiosonde} \approx 10^{-6} \cdot k_1 \cdot R_d \cdot \sum_{i=1}^{n} \frac{1}{g_i} \cdot \Delta p_i \tag{5.3}$$

$$ZWD_{radiosonde} \approx 10^{-6} \cdot \frac{R}{M_w} \cdot \sum_{i=1}^{n} \frac{1}{g_i} \cdot \left(k_2 - k_1 \cdot \varepsilon + \frac{k_3}{T_i} \right) \cdot \frac{r_i}{r_i + 1} \cdot \Delta p_i \tag{5.4}$$

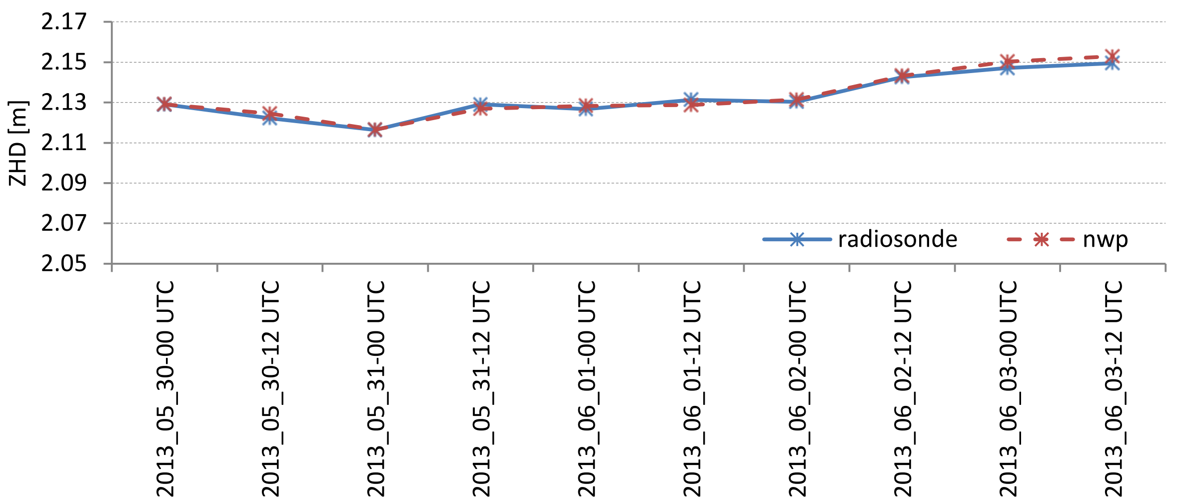
Where, subscript *i* denote a radiosonde profile while all the other terms are as earlier defined.

Radiosonde total zenith path delay estimates, $\mathrm{ZTD}_{\mathrm{radiosonde}}$, were then expressed as a sum of $\mathrm{ZHD}_{\mathrm{radiosonde}}$ and $\mathrm{ZWD}_{\mathrm{radiosonde}}$.

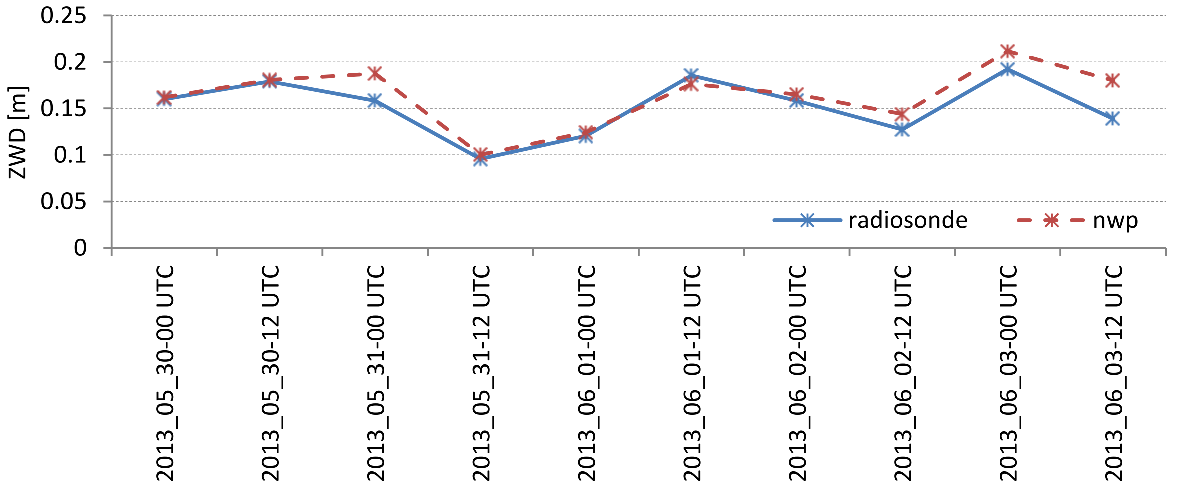
Figure 5.12 gives us an insight into the relationship between the NWP- and radiosonde-derived estimates of troposphere delays at the Wroclaw I (12425) radiosonde site over the study period. These are 12-hourly trends in the radiosonde and NWP estimates.

All the trends appear to have a somewhat similar pattern. However, while the radiosonde and NWP ZHD estimates (Figure 5.12 (a)) seem to be of equal magnitudes at all times, this is not the case in the ZWD estimates (Figure 5.12 (b)). There are also greater variations in the ZWD estimates compared to the ZHD estimates. For instance, on 31st May 2013 between 00 UTC and 12 UTC, radiosonde ZWD changed by approximately 60 mm while ZHD changed by only about 10 mm. ZTD estimates appear to have taken the variation effects of the wet delays (Figure 5.12 (c)).

(a) 12-hourly trends in radiosonde and NWP estimates of ZHD at Wroclaw I (12425)



(b) 12-hourly trends in radiosonde and NWP estimates of ZWD at Wroclaw I (12425)



(c) 12-hourly trends in radiosonde and NWP estimates of ZTD at Wroclaw I (12425)

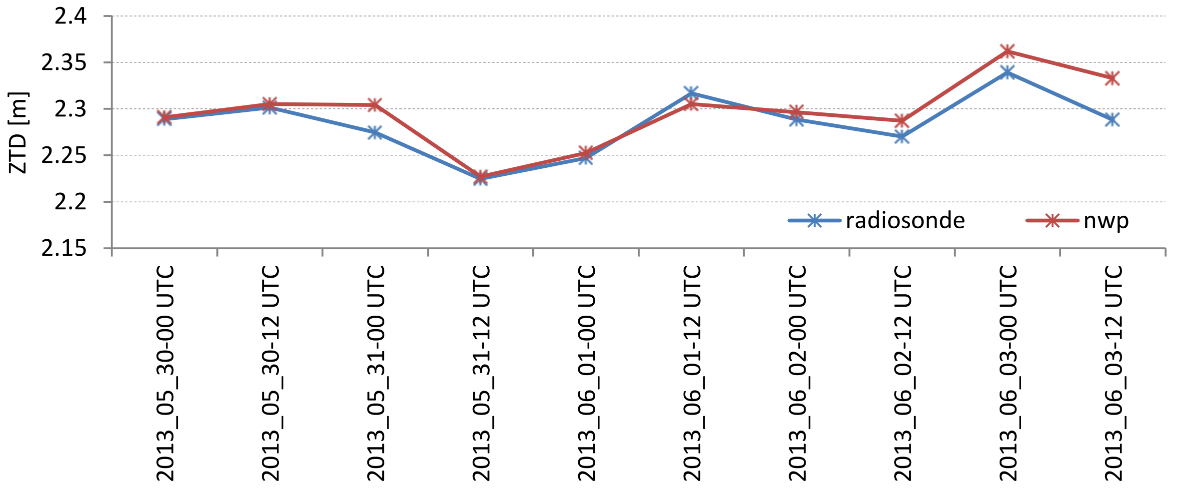


Figure 5.12: 12-hourly trends in radiosonde and NWP estimates at Wroclaw I for (a) ZHD, (b) ZWD and (c) ZTD.

The radiosonde-derived troposphere delays were found to have overall mean values of approximately 2.1 m, 0.2 m and 2.2 m for the zenith hydrostatic, wet and total path delays respectively. A comparison with the NWP-derived delays resulted into the mean accuracies shown in Table 5.4.

Table 5.4: Comparison of radiosonde- and NWP-derived troposphere delays.

Radiosonde site identifier	**ZHD [mm]**		**ZWD [mm]**		**ZTD [mm]**	
	bias	**stdev**	**bias**	**stdev**	**bias**	**stdev**
12425	0.8	2.0	0.9	18.8	1.0	20.2
11747	-1.8	5.4	1.0	22.5	2.3	20.2
11520	0.7	2.2	0.8	21.6	1.5	21.5
10771	0.0	3.7	3.9	14.0	3.5	15.8
10548	0.4	3.7	2.4	14.9	2.7	13.2
10393	-0.9	2.7	2.2	16.5	2.0	15.7
10238	0.6	4.2	2.4	16.8	2.5	18.2
mean	**-0.05**	**3.42**	**1.94**	**17.86**	**2.20**	**17.84**

Most bias errors were found to be positive with two relatively large negative biases that resulted into an overall mean bias of approximately -0.1 mm for ZHD. As for the ZWD and ZTD the overall mean biases were found to be approximately 2.0 mm and 2.2 mm respectively.

ZHD estimates were found to be more precise than the ZWD estimates with a standard deviation error of approximately 3.0 mm for the former and 18 mm for the latter. The precision errors in the ZTD estimates were also found to be approximately 18 mm, probably due to the large precision error in the in the wet delays.

5.9. Summary

Zenith hydrostatic and wet delays were estimated using NWP and radiosonde data by numerical integration of the hydrostatic and wet refractivity profiles in the NWP model levels and radiosonde vertical profiles.

Ignoring the portion of the atmosphere above the NWP model top was found to result into zenith hydrostatic delays that were smaller in magnitude by approximately 0.1 m while the zenith wet delays remained unaffected.

The zenith hydrostatic delays were also computed by use of the Saastamoinen hydrostatic model with the sensor pressure at selected EPN sites as the main input. The results were compared to similar observations derived from the NWP model (*num_int*) and those from an extended mean parameters model (*ext*) and standard processing (*std*) as discussed in the main text. All methods yielded zenith hydrostatic delay estimates of an equal magnitude of approximately 2.2 m. The overall mean accuracies in terms of bias errors were found to be 0.01 mm and -0.05 mm for *num_int* and *ext*/*std* respectively. The overall mean precisions in terms of standard deviations were found to be approximately 1.37 mm and 0.96 mm for *num_int* and *ext*/*std* respectively, while the root mean square errors were found to be approximately 2.52 mm and 2.35 mm for *num_int* and *ext*/*std* respectively.

The NWP-derived hydrostatic delay estimates were further validated using similar observations obtained from radiosonde data that gave an overall mean hydrostatic delay estimate of approximately 2.1 m. The resulting mean accuracies were approximately 0.3 mm and 4.0 mm for the bias and standard deviation errors respectively.

Numerical integration wet delays from NWP data resulted into a mean value of approximately 0.2 m. Zenith wet delays were also estimated from the difference of total zenith path delay observations sourced from selected EPN stations and the Saastamoinen hydrostatic delays (discussed in the preceding paragraphs) resulting into wet delays amounting to an overall negative mean with a magnitude of approximately 0.1 m. A comparison of the two sets of data resulted into overall mean bias, standard deviation and root mean square errors of approximately -7.6 mm, 14.0 mm and 18.0 mm respectively. Further comparison between the EPN/Saastamoinen ZWD with *ext* and *std* ZWD estimates resulted into an overall mean bias, standard deviation and root mean square error of approximately -0.6 mm, 12.0 mm and 14.0 mm for *ext* and approximately -1.7 mm, 10.0 mm and 11.0 mm for *std* respectively.

Validation of the NWP numerical integration wet delay estimates with similar estimates from radiosonde data resulted into mean accuracies of approximately 2.0 mm and 18 mm in bias and standard deviation errors respectively.

Total zenith path delays were estimated from NWP data by numerical integration of the total refractivity resulting into an overall mean value of approximately 2.4 m. Similar observations sourced from selected EPN stations were also found to have an overall mean of approximately 2.4 m. The EPN total zenith path delay estimates were compared with those derived by numerical integration from the NWP and those from *ext* and *std* (discussed in preceding paragraphs). The overall mean bias errors were found to be approximately -8.3 mm, -1.6 mm and -1.9 mm for *num_int*, *ext* and *std* respectively, the overall mean standard deviations were found to be approximately 15.0 mm, 13.0 mm and 10.0 mm for *num_int*, *ext* and *std* respectively, and the root mean square errors were found to be approximately 19.0 mm, 15.0 mm and 11.0 mm for *num_int*, *ext* and *std* respectively. A similar validation using radiosonde data resulted into bias and precision errors of approximately 2.0 mm and 18 mm respectively.

In a nutshell, numerically integrating the atmospheric refractivities through the NWP model mass levels in order to compute troposphere delays was found to be more accurate than the *ext* and *std* methods in the ZHD computations while the latter methods were found to be more precise. The *ext* and *std* methods were also found to be more accurate and precise in ZWD and ZTD computations than the numerical integration method.

6. Analysis of GPS-, NWP- and radiosonde-derived precipitable water vapor

6.1. Introduction

It has already been mentioned that ZTD observations contain information on atmospheric water vapor. Formulations for precipitable water vapor (PWV) have then been made using two methods; one (Section 2.2) from observations (ZWD and surface temperature) that can be obtained from ground-based sources (GNSS receivers and meteorological sensors respectively) and the other (Section 2.5) from variables (water vapor mixing ratio and atmospheric pressure) to be extracted from an NWP model.

The accuracy levels of the meteorological variables extracted/derived from the 3D grid of the NWP model have been discussed in a previous study (Chapter 4). The present study therefore made use ground-based GPS ZTD observations and the NWP-derived meteorological variables to estimate PWV using the two approaches mentioned in the preceding paragraph. Additionally, PWV has been estimated using radiosonde data.

An accuracy assessment of the NWP-derived PWV estimates has been done through comparisons with similar observations from GPS and radiosonde estimates.

6.2. Methodology

This section gives an overview of the methods used in the estimation of PWV using NWP and radiosonde data and the validation procedure. The configurations defining the computational domain for the NWP model simulation are just as they were discussed in Section 3.2.

The weighted mean temperature T_m was discussed in Section 2.2 including the significance of its accuracy in the precise estimation of PWV using GPS-derived ZWD observations. In the current study, T_m estimates across the entire study domain were computed by numerical integration (*num_int*) of vertical profiles of temperature and humidity through NWP model levels and by use of the Bevis T_m-T_s relation (Equation (2.16)). Additionally, Tropsite2 T_m (Section 2.2) data was made available for comparison purposes.

T_m estimates were also computed at the radiosonde sites described in Table 4.1. T_m estimated using the Bevis T_m-T_s relation involved the use of surface temperature observations sourced from the NWP model. Consequently, the accuracies in the numerical integration T_m (using NWP data) and the Bevis T_m were computed by comparison with radiosonde T_m. The resulting accuracies were then compared. The NWP-derived T_m estimates were then checked against the Bevis and radiosonde T_m estimates.

PWV across the entire study domain was estimated by numerical integration of water vapor density through the NWP model levels and by use of surface temperature and surface pressure observations in order to estimate T_m (from the Bevis T_m-T_s relation) and ZWD (from ZTD - ZHD_{saas}) respectively and hence compute PWV using Equation (2.13).

PWV estimates were also computed at selected radiosonde sites that were shown in Table 4.1 by numerically integrating the vertical profiles of temperature and humidity in the NWP model mass levels and radiosonde profiles, and by using the Emardson/Derks PWV formula that was discussed in Section 2.2 (also refer to Emardson and Derks (2000) for details). The NWP- and Emardson/Derks-derived PWV estimates were then checked for correctness by comparing them with the GPS and radiosonde estimates.

Finally, GPS PWV observations that are made available by SuomiNet, a GPS network designed for real-time remote sensing (Ware *et al.*, 2000), were compared with the NWP- and GPS-derived PWV observations.

6.3. Assessment of the weighted mean temperature over the study domain

The weighted mean temperature T_m was estimated by numerically integrating humidity and temperature variables in the vertical extent of the NWP model via the water vapor mixing ratio, temperature and pressure variables as expressed in the equation of the form (Section 2.5),

$$T_m \approx \frac{\frac{1}{g_a}\cdot\frac{r_a}{r_a+1}\cdot\frac{1}{T_a}\cdot(p_a-p_j^h) \quad + \quad \sum_{j=1}^{m}\frac{1}{g_j}\cdot\frac{r_j}{r_j+1}\cdot\frac{1}{T_j}\cdot\Delta p^h}{\frac{1}{g_a}\cdot\frac{r_a}{r_a+1}\cdot\frac{1}{T_a^2}\cdot(p_a-p_j^h) \quad + \quad \sum_{j=1}^{m}\frac{1}{g_j}\cdot\frac{r_j}{r_j+1}\cdot\frac{1}{T_j^2}\cdot\Delta p^h} \tag{6.1}$$

T_m over the study period can be modeled using NWP model (or radiosonde) data. The resulting regression equation can then be compared to the Bevis T_m-T_s relation (discussed in Section 2.2), i.e.

$$T_m = 70.2 + 0.72\cdot T_s \tag{6.2}$$

Where, T_s is the surface temperature in Kelvin.

However, given that the study period is relatively short, the NWP model data may not be sufficient to convincingly model T_m over the study domain. This may require data covering a much longer period, say at least two years data. Consequently, a comparison of the regression equations may not be appropriate at this point. Nevertheless, T_m/T_s[1] data points simulated using the NWP model over the entire study period, including 25 hourly forecasts over each day in the 5-day study period, were made to run through a best-fit line resulting into the scatter plot shown in Figure 6.1. This was done in order to provide an insight into the T_m-T_s relation across the study domain. The R-squared value of

---

[1] T_s was taken to be the 2 m temperature.

approximately 0.89 was found to be a good fit of the T_m/T_s data. The resulting empirical relation between T_m and T_s takes the form,

$$T_m = 26.12 + 0.87 \cdot T_s \tag{6.3}$$

Where, $T_s \equiv$ the surface temperature in K.

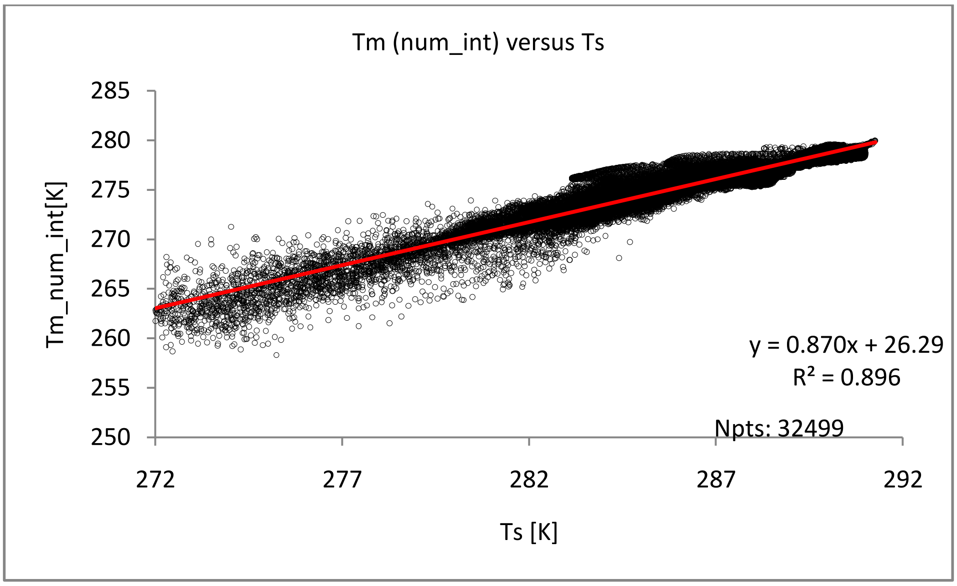


Figure 6.1: NWP-derived T_m-T_s relation over the entire study period and domain covering central Europe.

From Figure 6.1, the T_m estimates vary from approximately 260 K to 280 K corresponding to a variation in T_s of approximately 272 K to 292 K. Most of the data points are aligned in a straight line revealing a linear relationship between the weighted mean temperature and surface temperature estimates. There is a particularly high density of the data points for $T_s > 282$ K and a low density for $T_s < 278$ K.

The only variable required in order to estimate T_m using the Bevis T_m-T_s relation is the NWP-derived surface temperature field. The plots of Figure 6.2 provide a visual comparison of T_m estimates simulated for 30th May 2013 00 UTC across the entire study domain using the two methods: numerical integration and Bevis T_m-T_s relation.

From Figure 6.2, it appears that the numerical integration method predicts lower mean weighted temperatures (as low as -12°C) over a wider area surrounding the Alps compared to the Bevis T_m-T_s relation. It also predicts higher mean weighted temperatures (as high as 8°C) for a better part of the eastern part of the study domain. The accuracies of these temperatures, however, were checked by comparison with weighted mean temperature computed using radiosonde data at specified sites Section 6.4).

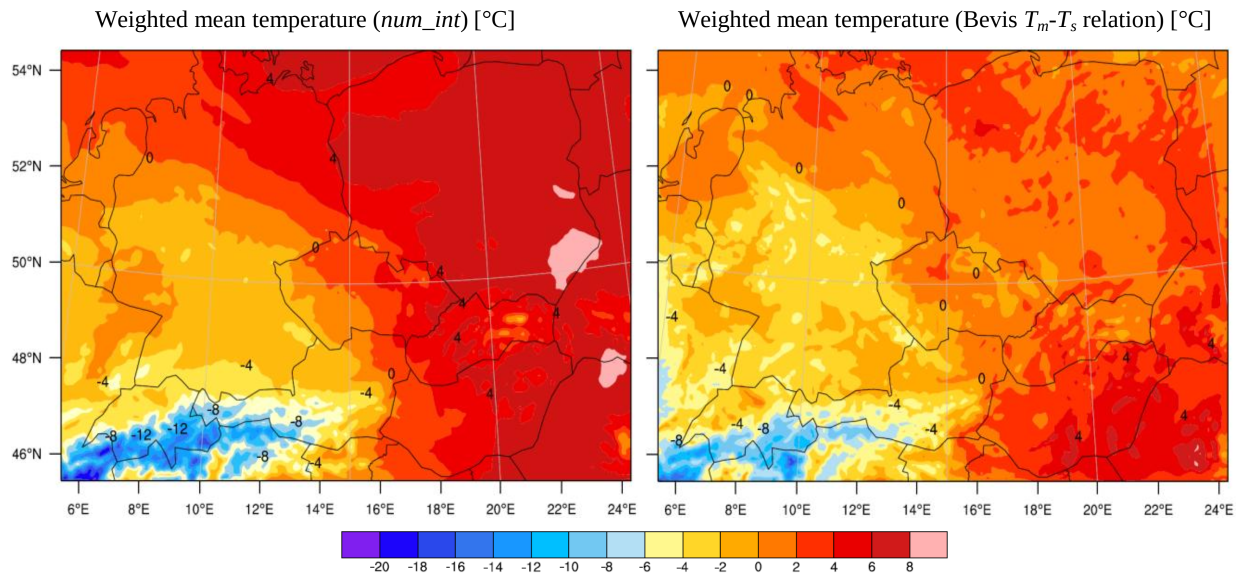


Figure 6.2: Numerical integration (*num_int*) T_m and Bevis T_m-T_s relation T_m for 30th May 2013 00 UTC over the entire study domain.

6.4. Validation of NWP-derived weighted mean temperature estimates at radiosonde sites

T_m was estimated by integration of humidity and temperature values found in the vertical profiles of radiosonde data sourced from Wyoming University's radiosonde data archive (University of Wyoming, 2014). The radiosonde sites, their coordinates and launch times were summarized in Table 4.1.

The current study, therefore, involved estimating T_m at the radiosonde sites using NWP model data (numerical integration and Bevis T_m-T_s relation) and radiosonde data. The numerical integration T_m at the radiosonde sites was then compared to the Bevis T_m and radiosonde T_m.

With regards to estimating T_m using radiosonde data, Equation (6.1) was modified to suit the current application. T_m was therefore approximated by,

$$T_m = \frac{\int_h^\infty \frac{e}{T}\cdot dh}{\int_h^\infty \frac{e}{T^2}\cdot dh} \approx \frac{\sum_{i=1}^{n} \frac{e}{T}\cdot \Delta h_i}{\sum_{i=1}^{n} \frac{e}{T^2}\cdot \Delta h_i} \qquad (6.4)$$

Where,

$\Delta h_i = h_{i+1} - h_i$ ≡ height difference at adjacent radiosonde profiles, i being the current profile and i+1 being the nearest upward profile defined by the geopotential heights h_i and h_{i+1} respectively.

But, by assuming the atmosphere to be in hydrostatic equilibrium and expressing the water vapor density in terms of the water vapor mixing ratio (refer to Equation (2.29)), Equation (6.4) can be expressed as,

$$T_m \approx \frac{\sum_{i=1}^{n} \frac{1}{g_i} \cdot \frac{r_i}{r_i+1} \cdot \frac{1}{T_i} \cdot \Delta p_i}{\sum_{i=1}^{n} \frac{1}{g_i} \cdot \frac{r_i}{r_i+1} \cdot \frac{1}{T_i^2} \cdot \Delta p_i} \tag{6.5}$$

Where,

$\Delta p_i = p_i - p_{i+1}$ ≡ atmospheric pressure difference at adjacent radiosonde profiles defined by the pressures p_i (corresponding to the geopotential height h_i) and p_{i+1} (corresponding to the geopotential height h_{i+1}), and $p_i > p_{i+1}$

n ≡ number of radiosonde profiles.

All the variables appearing in Equation (6.5) can be obtained from radiosonde profiles while the local gravity acceleration at the radiosonde profiles can be derived using Equation (2.37). T_m was then estimated using numerical integration and the Bevis T_m-T_s relation at all the radiosonde launch locations specified in Table 4.1.

A 12-hourly trend in T_m estimates from radiosonde, numerical integration and the Bevis T_m-T_s relation for two of the radiosonde sites considered, Meiningen (10548) and Lindenberg (10393), are illustrated in Figure 6.3.

From Figure 6.3, trends in T_m appear to be similar most of the times. But, the magnitudes of T_m estimated using the Bevis T_m-T_s relation are different from those of numerical integration and radiosonde some of the times. Differences in magnitude between numerical integration and radiosonde T_m are very small most of the times.

Accuracy in T_m obtained using numerical integration, the Bevis T_m-T_s relation and TropGrid2 was then computed by using T_m estimates from radiosonde data as reference. The results are shown in Table 6.1.

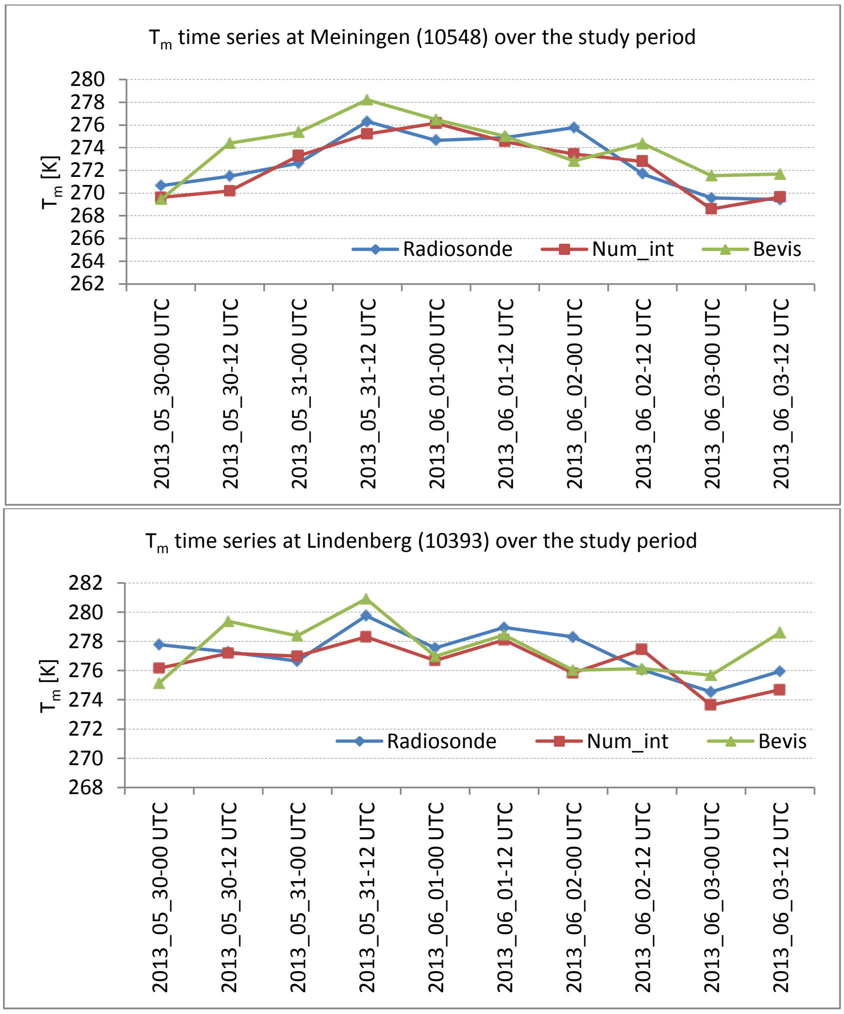


Figure 6.3: 12-hourly trend in the weighted mean temperature at (a) Meiningen (10548) and (b) Lindenberg (10393).

Table 6.1: Mean accuracy and precision errors for the weighted mean temperature at radiosonde sites over Central Europe.

Radiosonde site identifier (WMO)	Mean accuracies for T_m [K]								
	Bias [K]			Standard dev. [K]			RMS [K]		
	num_int	Bevis	TropSite2	num_int	Bevis	TropSite2	num_int	Bevis	TropSite2
12425	-0.1	0.2	1.2	1.0	0.9	2.6	1.0	1.0	2.8
11747	0.6	0.4	2.8	1.1	1.5	2.2	1.2	1.5	3.5
11520	0.5	-1.2	3.8	0.8	0.7	1.9	0.9	1.3	4.3
10771	0.8	0.9	2.7	0.7	1.3	3.0	1.1	1.6	3.9
10548	0.4	-1.2	1.7	1.2	1.9	2.8	1.2	2.2	3.2
10393	0.8	-0.3	2.2	1.1	1.8	2.0	1.3	1.7	2.9
10238	0.3	0.8	1.4	1.4	1.8	1.7	1.4	1.9	2.1

The bias errors are largely positive except for a few sites such as Wroclaw I (12425) that shows a negative bias of 0.1 K for *num_int*, and almost half of the sites that show negative biases of varying magnitudes for Bevis. The *num_int* bias errors are lower in magnitude at majority of the sites considered. Overall, therefore, the *num_int* method computed better T_m estimates than Bevis and TropSite2 at the sites considered. The bias error results are illustrated in Figure 6.4.

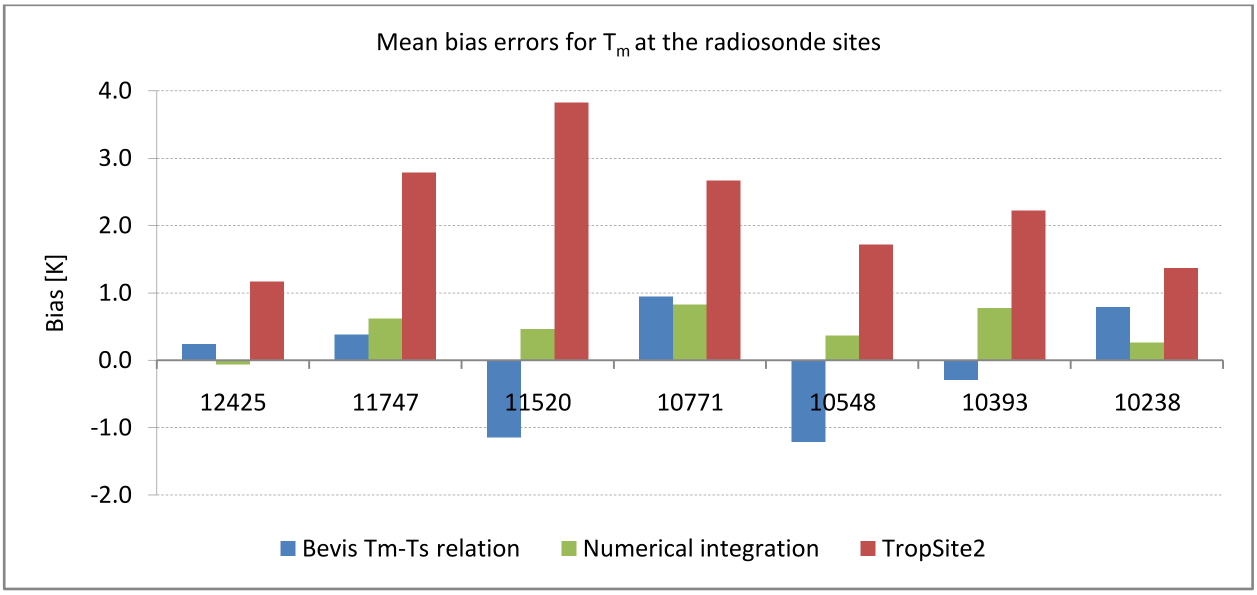


Figure 6.4: Mean accuracies in the weighted mean temperature at the radiosonde sites.

The bias-reduced standard deviation errors from *num_int* average approximately 1.0 K, those obtained using the Bevis T_m-T_s relation average approximately 1.4 K, while those from TropSite2 average approximately 2.3 K. TropSite2 standard deviation errors are generally greater at majority of the sites except at Bergen (10238) where the error is less than that of the Bevis T_m-T_s relation (Figure 6.5).

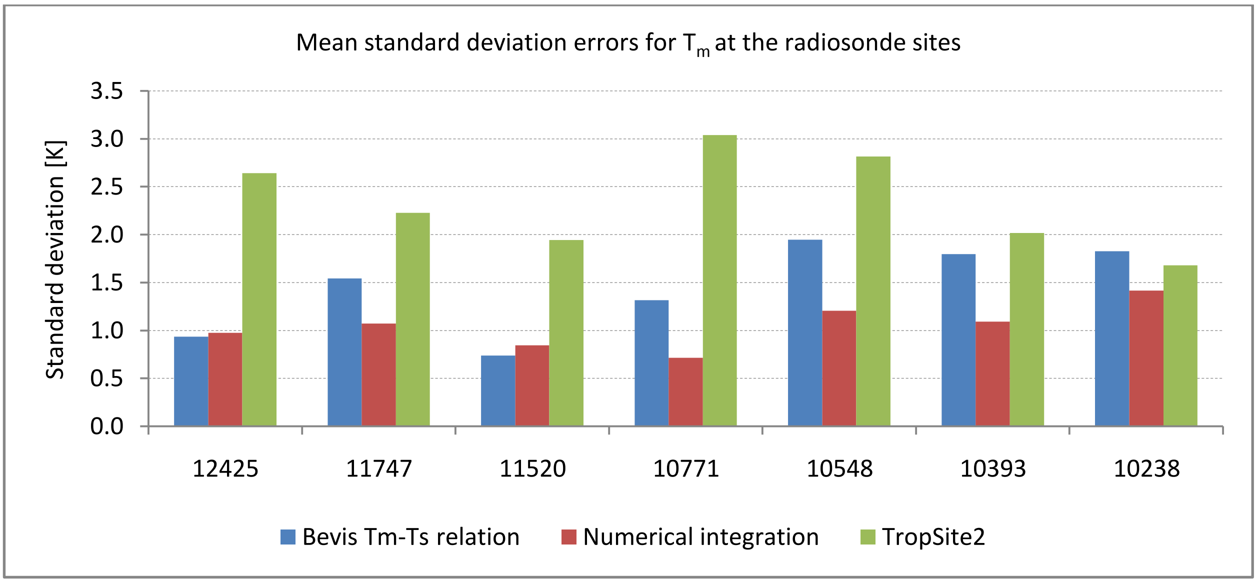


Figure 6.5: Mean precision errors in the weighted mean temperature at the radiosonde sites.

An analysis of the root mean square (rms) errors, which include the bias errors, shows significantly large errors for TropSite2 compared to *num_int* and Bevis (Figure 6.6).

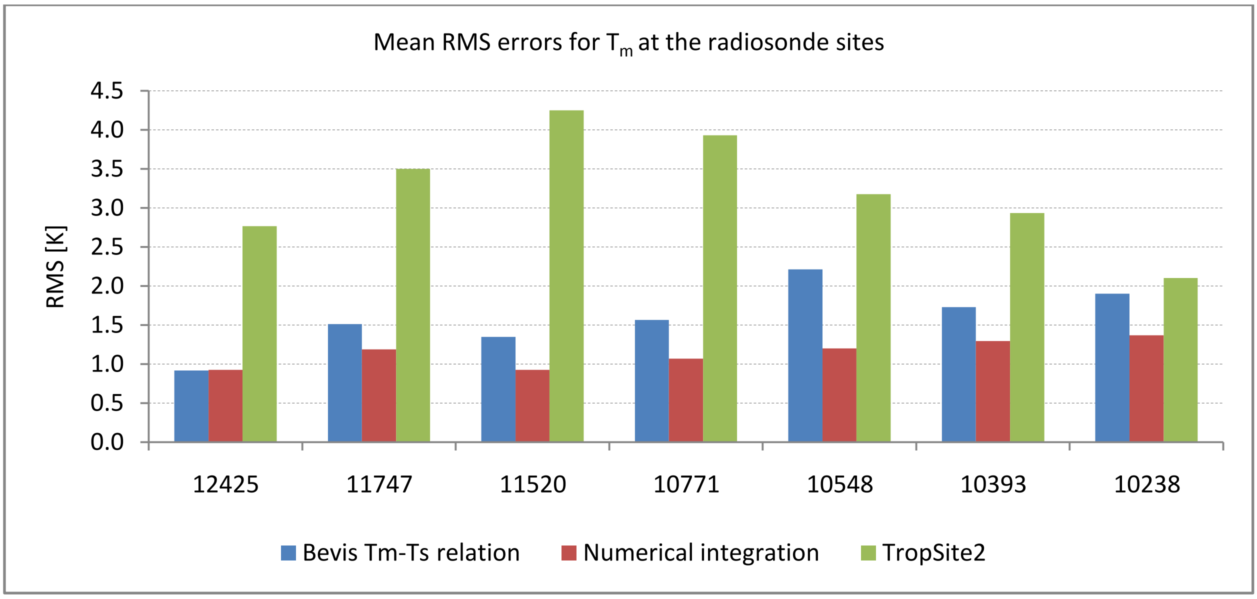


Figure 6.6: Mean root mean square errors in the weighted mean temperature at the radiosonde sites.

The error analysis for T_m suggests that the numerical integration method was more precise than the Bevis T_m-T_s relation and TropSite2 in the estimation of the weighted mean temperature T_m. This stems from the fact that the *num_int* generally exhibited errors that were lower in magnitude than the other methods. Since the function coefficients (as mentioned in Section 2.2) for a specified site involve interpolation from grid-point values, it is possible that the errors observed in TropSite2 originated from this. Finally, though the Bevis T_m-T_s relation is widely accepted for use in the estimation of T_m, it has to be used with caution, as it may not be suitable for use in all locations particularly when high precision is required.

6.5. Comparison of NWP- Emardson/Derks- and radiosonde-derived PWV

An insight into NWP PWV estimates across the entire study domain is provided in Figure 6.7.

From Figure 6.7, NWP-derived PWV estimates were found to be lowest in the regions around the Alps. A larger part of the study domain had PWV values averaging between 18 - 30 mm.

The aim of this activity was to compare PWV estimates from radiosonde data and those of the NWP model. The coordinates and heights of the radiosonde sites were used to retrieve data from the NWP model. PWV was then estimated by integration of the humidity variable in the vertical levels defining the NWP model data and the vertical profiles of the radiosonde data. PWV results from the two sets of data were then compared.

An effort was made to ensure that the PWV estimates were calculated over the same depth of the atmosphere covered by the NWP model and radiosonde vertical profiles. This is not only because radiosonde heights (approximately 30 km and more) are different from those of NWP model (approximately 20 km), but also because the water content in the atmosphere is mainly significant in the lower atmosphere as further illustrated in Figure 6.8.

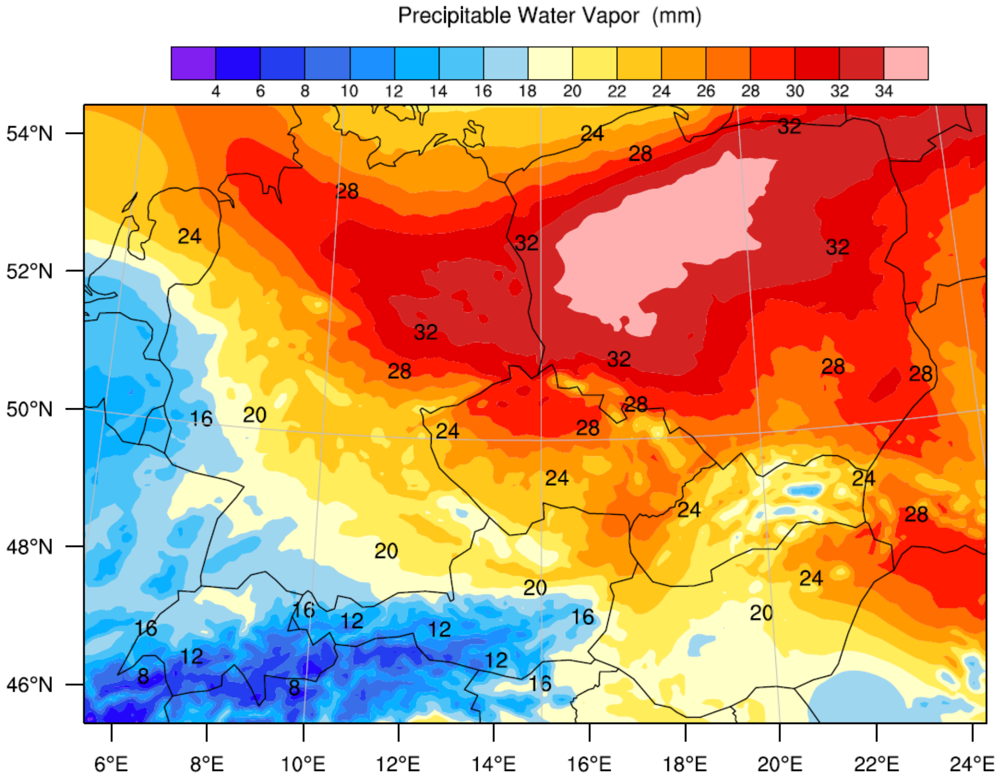


Figure 6.7: NWP PWV over the entire study domain simulated for 30th May 2013 18 UTC.

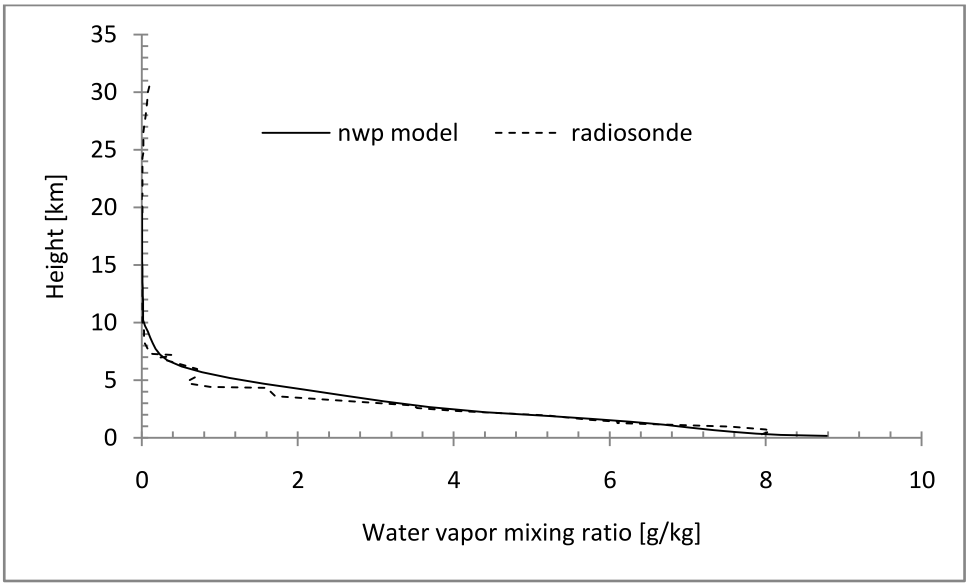


Figure 6.8: Vertical profiles (dashed for radiosonde and solid for nwp model) of water vapor mixing ratio for the radiosonde site Wroclaw I (12425) on 29th May 2013 18 UTC.

In Figure 6.8, we see that humidity is only significant closer to the surface up to a vertical height of approximately 3 km after which it falls off to insignificant levels. Therefore, water vapor mixing ratio is a function of altitude. It is also found to be strongly dependent on both time and latitude (Léna *et al.*, 2012).

NWP PWV and radiosonde PWV estimates were therefore computed to a height of 20 km (the approximate height of the NWP model top) using the following approximations,

$$PWV_{NWP} \approx \frac{1}{\rho_l} \cdot \frac{1}{g_a} \cdot \frac{r_a}{r_a + 1} \cdot (p_a - p_j^h) + \frac{1}{\rho_l} \cdot \sum_{j=l}^{m} \frac{1}{g_j} \cdot \frac{r_j}{r_j + 1} \cdot \Delta\, p^h \tag{6.6}$$

$$PWV_{radiosonde} \approx \frac{1}{\rho_l} \cdot \sum_{i=l}^{n} \frac{1}{g_i} \cdot \frac{r_i}{r_i + 1} \cdot \Delta\, p_i \tag{6.7}$$

Where, subscripts *i* and *j* denote the NWP model levels and radiosonde profiles respectively while all the other terms are as earlier defined.

The Emardson and Derks formula for PWV was also used to compute PWV at the radiosonde sites considered. The surface temperature input (Equation (2.19)) was sourced from the radiosonde data.

A sample of the results is illustrated with the 12-hourly trends of the PWV at Wroclaw I (12425) and Prostejov (11747) over the entire study period as shown in Figure 6.9.

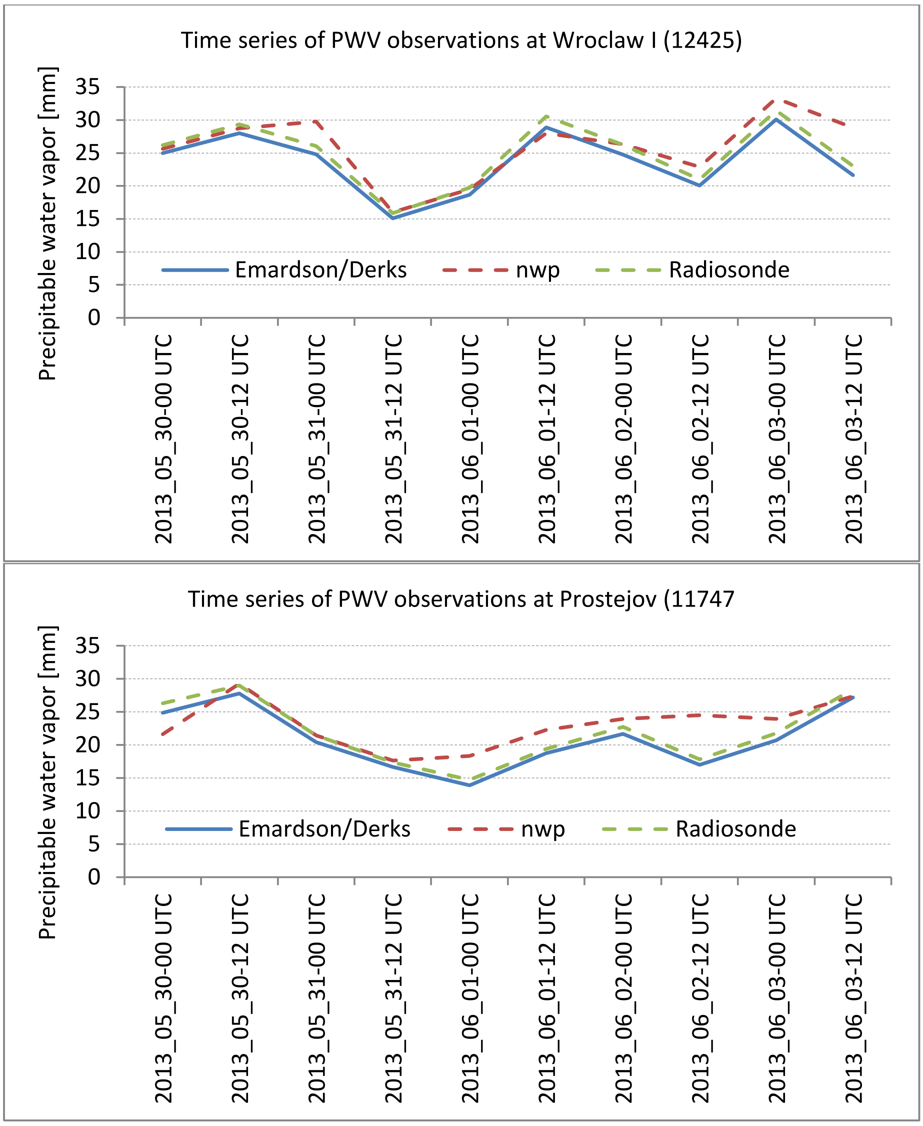


Figure 6.9: NWP, radiosonde and Emardson/Derks PWV estimates at Wroclaw I (upper panel) and Prostejov (lower panel) over the entire study period.

There is a very close relationship between the radiosonde and Emardson/Derks PWV estimates. This could be because of the surface temperature input in the latter formulation that was sourced from the radiosonde data. There is also a close relationship with the NWP-derived PWV most of the times at Wroclaw I. This is not the case at Prostejov such as the times between 2013_05_31-12 UTC and 2013_06_03-00 UTC where significant differences are observed with a particularly huge difference greater than 5 mm at 2013_06_02-12 UTC.

The mean accuracy and precision errors at all the radiosonde sites are shown in Table 6.2. Radiosonde PWV was used as reference and the results from the Emardson/Derks formulation are simply expressed as Emardson.

Table 6.2: Mean accuracy and precisions in PWV for selected radiosonde sites in Central Europe.

Radiosonde site WMO code	**bias [mm]**		**stdev [mm]**		**rms [mm]**	
	NWP	**Emardson**	**NWP**	**Emardson**	**NWP**	**Emardson**
12425	-1.0	1.2	2.4	0.3	2.5	1.3
11747	-1.1	1.0	3.0	0.3	3.1	1.0
11520	0.8	1.1	3.3	0.2	3.3	1.2
10771	-1.4	0.9	1.4	0.2	1.9	0.9
10548	-0.4	0.9	2.1	0.3	2.1	0.9
10393	-0.3	1.3	2.2	0.2	2.3	1.3
10238	-0.9	0.9	2.2	0.4	2.3	1.0

A comparison with NWP PWV estimates resulted into mostly negative bias errors with the highest (approximately 1.4 mm) being observed at Kuemmersbruck (10771) and the least error (approximately 0.3 mm) at Lindenberg (10393). A comparison with the Emardson/Derks PWV estimates, on the other hand, showed positive biases.

The magnitude of the overall mean bias error for NWP PWV is slightly lower (approximately 0.6 mm) than that of the Emardson/Derks PWV (approximately 1.1 mm). The standard deviation and root mean square errors are, however, greater for NWP PWV than Emardson/Derks PWV.

6.6. Comparison of NWP- and GPS-derived PWV

GPS-sensed PWV was estimated from surface observations, namely; ZTD, surface pressure and surface temperature observations that are available at selected GPS reference stations falling within the study domain.

Surface pressure observations at a given site were used to estimate ZHD using the Saastamoinen hydrostatic model which was then subtracted from the GPS ZTD observations for the site to obtain the ZWD. The Bevis T_m-T_s relation[1] was used to calculate the weighted mean temperature by use of the

---

[1] The Bevis T_m-T_s relation was used to compute T_m instead of using the num_int T_m based on the assumption that surface pressure, surface temperature and GPS ZTD observations are available at the EPN site.

surface temperature observations at the site. ZWD and T_m were then used to estimate PWV as described in Section 2.2.

NWP-derived PWV estimates were computed as was discussed in the preceding section (Section 6.5). The results were then compared with the GPS-derived PWV estimates in terms of bias and standard deviation errors computations with the GPS PWV taken as the reference quantity. The mean accuracies are shown in Table 6.3.

Table 6.3: Comparison of NWP- and GPS-derived PWV at EPN sites.

	mean accuracies [mm]	
Site ID	**bias**	**std**
BOGI	-1.0	2.4
BOR1	0.0	2.3
BPDL	-1.2	2.1
DRES	-1.5	1.4
GOPE	-1.8	1.7
GWWL	-1.2	1.7
JOZE	-1.7	2.9
KATO	-0.6	2.4
MOPI	-2.8	1.4
POTS	-2.3	1.7
PTBB	0.0	1.3
TUBO	-0.5	2.2
WTZR	-1.0	1.5
ZYWI	-1.0	2.2

An overall negative mean bias error in the PWV estimates of approximately 1.0 mm was obtained, while the standard deviation error was found to be approximately 2.0 mm. The largest bias errors greater than 2.0 mm were seen at MOPI and POTS.

An insight into the PWV trends at MOPI is shown in Figure 6.10.

The NWP and GPS PWV trends at MOPI appear to have a somewhat similar pattern. However, the NWP PWV estimates are slightly greater in magnitude than the GPS PWV estimates most of the times.

GPS-sensed PWV observations were also sourced directly from the SuomiNet database for a few sites that happen to fall within the study domain. A comparison of these observations with NWP-derived PWV was done at five sites common to both the study domain and SuomiNet. The mean accuracies in the PWV estimates are summarized in Table 6.4, where the SuomiNet GPS PWV estimates were taken as the reference quantities.

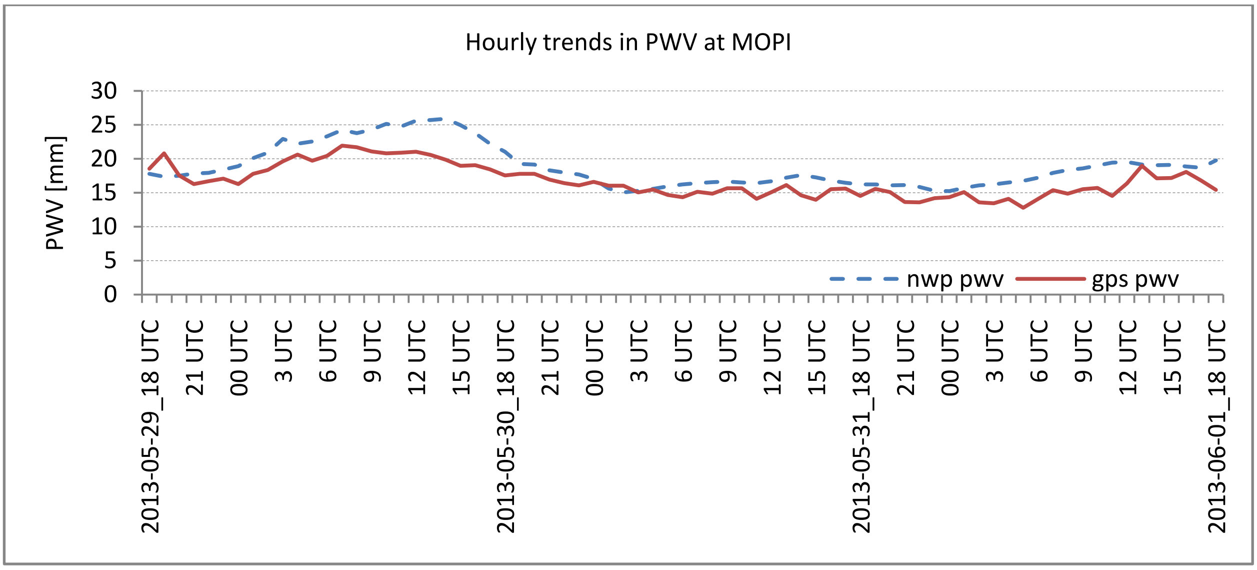


Figure 6.10: Hourly trends in the NWP and GPS PWV estimates at MOPI.

Table 6.4: Comparison of NWP PWV and SuomiNet GPS PWV at EPN sites.

	mean accuracies [mm]	
Site ID	**bias**	**std**
BOR1	-0.2	1.9
GOPE	-2.4	1.6
JOZE	-2.5	2.5
PTBB	-0.5	1.1
WTZR	-1.6	1.0

All the stations considered showed negative bias errors with an overall negative mean bias error of approximately 1.4 mm, while the overall mean standard deviation error was approximately 1.6 mm.

GOPE and JOZE had the largest bias errors. Hourly trends of PWV at JOZE were produced with an aim of looking into the NWP and SuomiNet PWV evolutions over time.

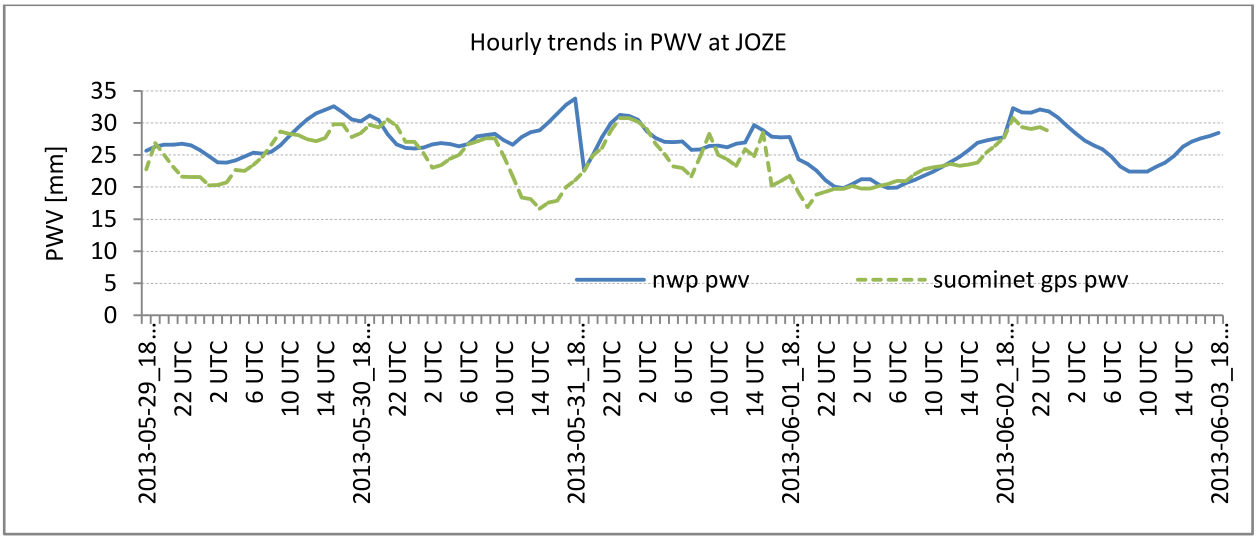


Figure 6.11: Trends in NWP and SuomiNet PWV observations at JOZE.

The trends are similar in pattern most of the times with small differences in magnitude. However, a breakdown in the NWP model simulation is seen from around 30th May 2013 11 UTC to around 31st May 2013 19 UTC.

6.7. Summary

NWP simulated weighted mean temperature estimates across the study domain were found to range from approximately 260 K to 280 K. On comparison with radiosonde-derived weighted mean temperature estimates, the NWP-derived estimates were found to have a mainly negative bias error at the radiosonde sites ranging from approximately 0.1 K to 0.8 K, while the Bevis T_m-T_s relation-derived estimates were found to have a mainly positive bias error ranging from approximately 0.2 K to 1.2 K.

Assuming high accuracy in the radiosonde data, the precision errors in the weighted mean temperature were found to be slightly better for the NWP-derived estimates compared to the estimates derived using the Bevis T_m-T_s relation.

NWP-derived PWV estimates were found to range from approximately 18 to 30 mm in magnitude. A comparison of these estimates with the radiosonde-derived estimates at selected radiosonde sites showed negative bias errors ranging from approximately 0.3 mm to 1.4 mm and an overall mean standard deviation of approximately 2.4 mm. A comparison with PWV derived using surface observations of GPS ZTD, pressure and temperature also yielded a mainly negative bias error with a mean value of approximately 1.0 mm and an overall mean standard deviation of approximately 2.0 mm. Almost similar results were obtained from the comparison with SuomiNet GPS PWV observations.

7. Impact assessment of GPS ZTD data assimilation on short-range NWP forecasts

7.1. Introduction

An insight into data assimilation (henceforth sometimes simply, DA) was given in Sections 1.3 and 2.6. The relationship between troposphere delay and water vapor content in the atmosphere was shown in Section 2.2. The study presented in this chapter discusses an activity (and the results) that involved ingesting ZTD observations from some stations in the domain of study into an NWP model in order to assess the impact that this would have on short-range forecasts of surface pressure, 2m temperature, and near-surface relative humidity, and hence add onto the on-going efforts by researchers to investigate the usefulness of ground-based GPS ZTD observations in meteorology.

Whether or not the ZTD observations to be ingested into the NWP model would make any impact in weather forecasts depends on several factors as revealed in the studies briefly mentioned in Section 1.3.5. The general outcome of those studies, however, resulted into neutral impacts (no effect on NWP forecasts), modest impacts (little effect on NWP forecasts) or positive impacts (improvement in the NWP forecasts).

Impact assessment mainly involved comparing the errors obtained in the forecasted variables before and after ZTD data assimilation. The current study, therefore, discusses the method used and the results obtained.

7.2. Methodology

The specification of the NWP model computation domain was discussed in Section 3.2. The data assimilation process was accomplished by use of the WRF 3D-Var data assimilation system.

Hourly GPS ZTD observations that were assimilated into the NWP model data were sourced from about 28 EPN reference stations that are marked in black in Figure 7.1.

The strategy used in the data assimilation process is well documented in Appendix B. Generally, the method involved making NWP simulations first without the GPS ZTD observations, then running the DA system to ingest the ZTD observations into the NWP model's initial condition files, and finally making a second NWP simulation that includes the ingested ZTD observations.

Meteorological variables of surface pressure, near-surface temperature and relative humidity were then extracted from the NWP model output files generated before and after ZTD DA. They were first extracted across the entire study domain for the two scenarios; before and after ZTD DA. The results of this procedure were presented graphically in order to make a visual comparison of the surface fields' forecasts before and after DA.

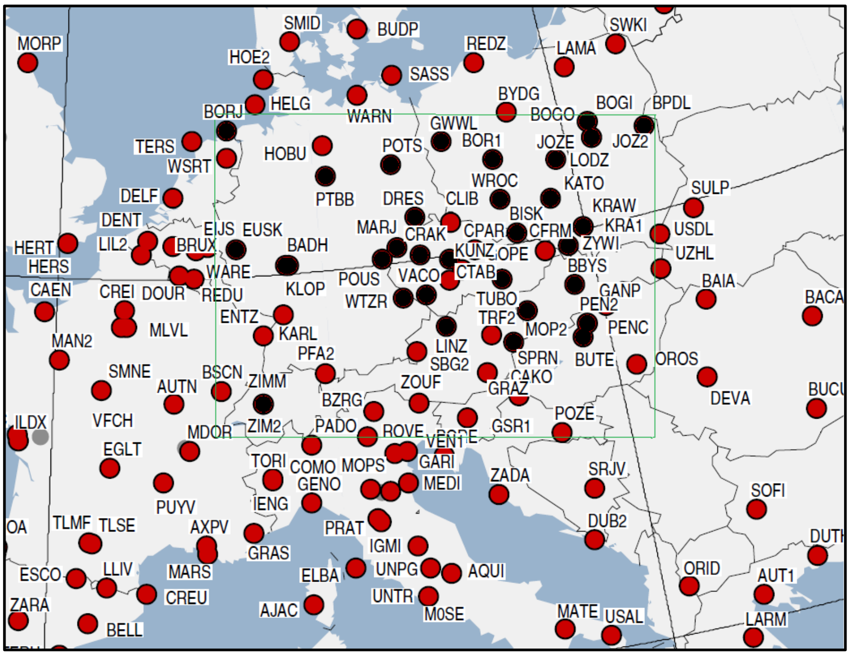


Figure 7.1: EPN GNSS reference sites from which ZTD data for data assimilation was sourced are marked in black. Inset rectangle marks the study domain (adapted from EPN (2013)).

The target observations were also extracted at specific EPN sites that are equipped with meteorological sensors and then subjected to an accuracy assessment exercise similar to the one discussed in Chapter 4 by comparison with EPN meteorological observations.

The comparisons in terms of bias and standard deviation errors as obtained for the data before DA and after DA were then compared in order to assess the impact of the assimilated ZTD data on the NWP simulations of the said variables.

7.3. Data assimilation results and analysis

Figure 7.2 (a) shows the mean bias errors in the surface pressure variable obtained from the NWP simulation before and after assimilating GPS ZTD observations. The bias errors before and after DA were largely modest at majority of the sites considered. This means that there were relatively small reductions in the bias error of the pressure forecasts after assimilating GPS ZTD observations. The impact was neutral at a few of the stations, with the errors remaining the same.

The change in the surface pressure standard deviation error is negligible at most of the sites considered (Figure 7.2 (b)). The error is, however, seen to increase slightly at DRES and decrease slightly GOPE and KUNZ.

The impact of the GPS ZTD data assimilation on pressure forecasts is seen in the visual comparison of the pressure forecasts over the entire study domain shown in the contour plots of Figure 7.3. The plots hardly reveal any changes in the pressure forecasts.

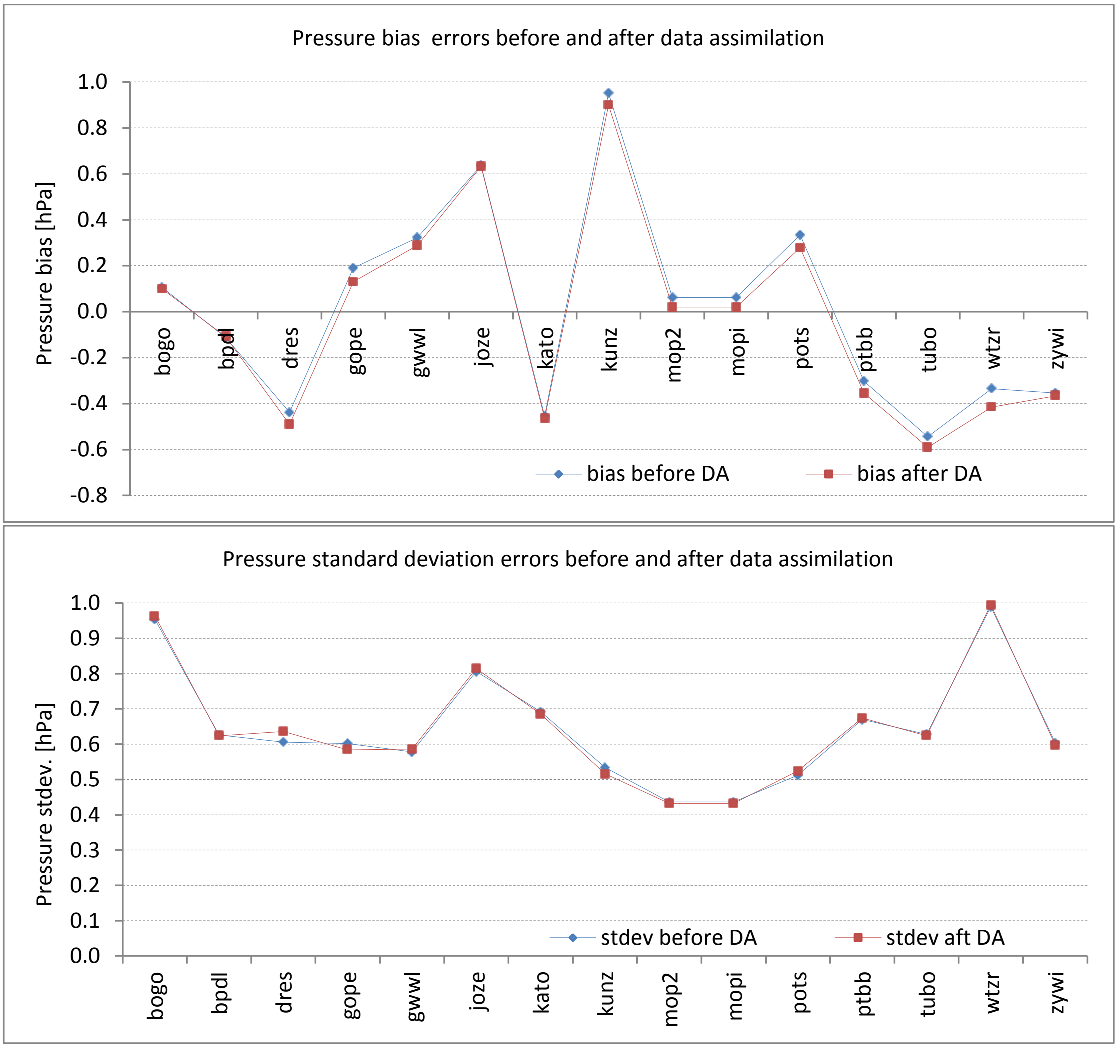


Figure 7.2: Pressure mean accuracies before and after ZTD data assimilation at EPN sites, (a) bias errors and (b) standard deviations.

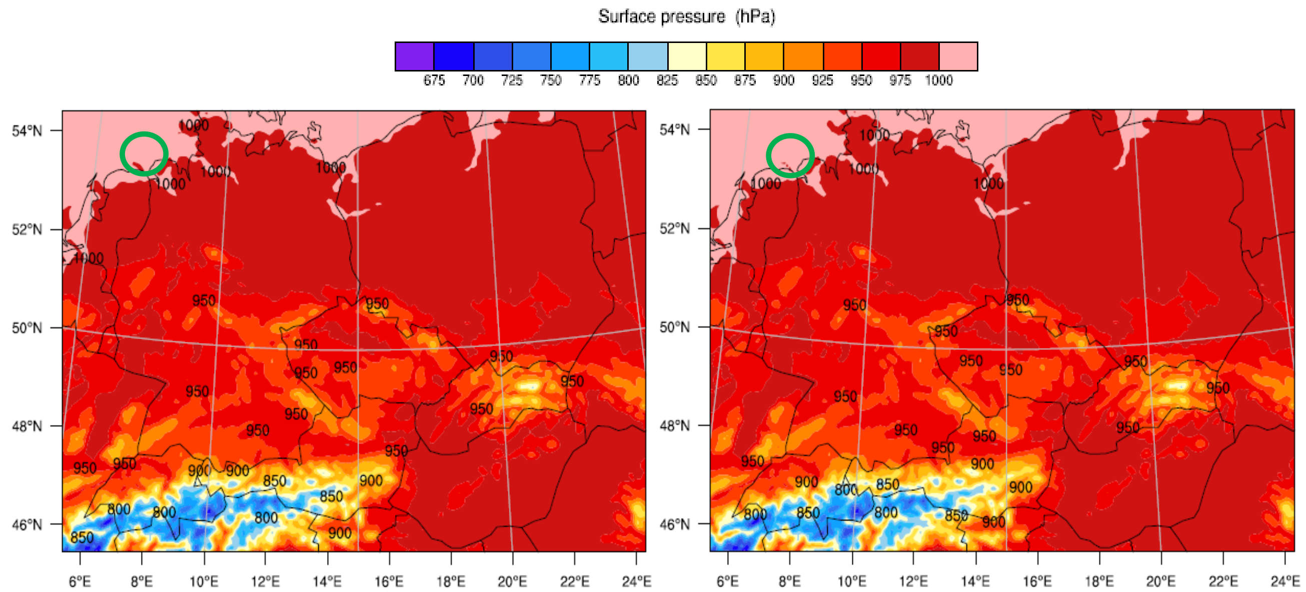
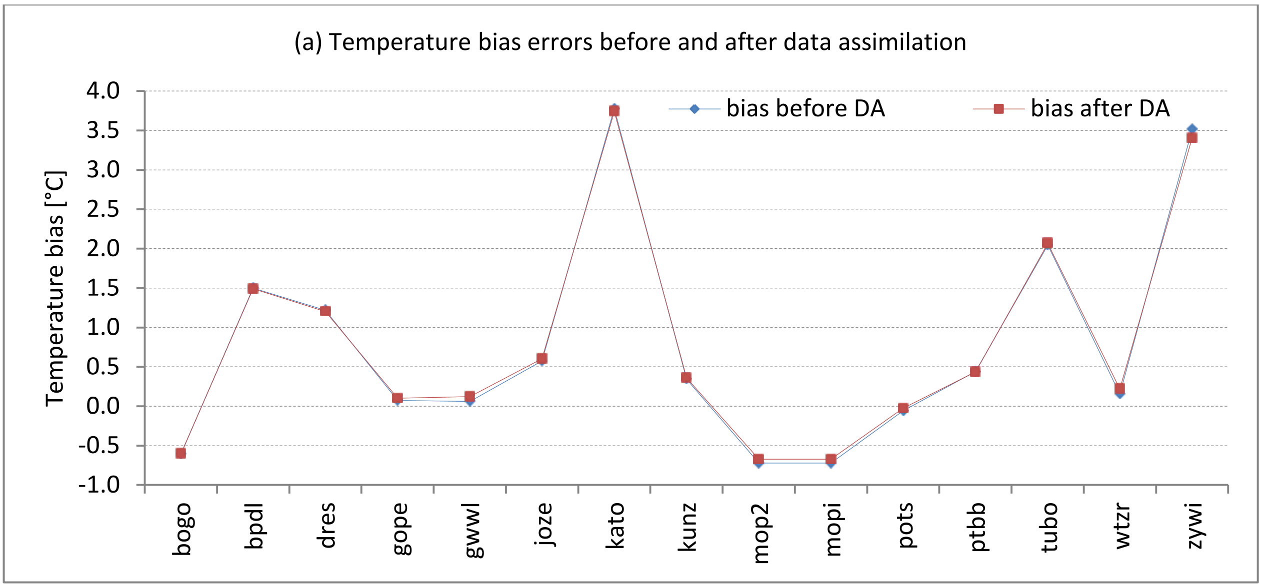


Figure 7.3: Contour plots of the surface pressure field over the entire domain on 29th May 2013 19 UTC. The panel on the left-hand side shows pressure forecast before GPS ZTD DA, while the one on the right-hand side shows the forecast after DA. An instance where a change is seen, albeit small, is circled in green.

The mean bias errors in the 2 m temperature variable obtained before and after assimilating the GPS ZTD observations are shown in Figure 7.4 (a). Unlike the surface pressure variable, the impact of GPS ZTD DA on 2 m temperature forecasts is seen to be largely neutral at majority of the stations considered in the study. A few stations, however, are showing an increase in the bias errors, notably, GWWL, MOP2 and MOPI, indicating a negative impact on temperature forecasts due to the assimilated ZTD observations. The 2 m temperature standard deviation error, however, remains unchanged (Figure 7.4 (b)).


(a) Temperature bias errors before and after data assimilation

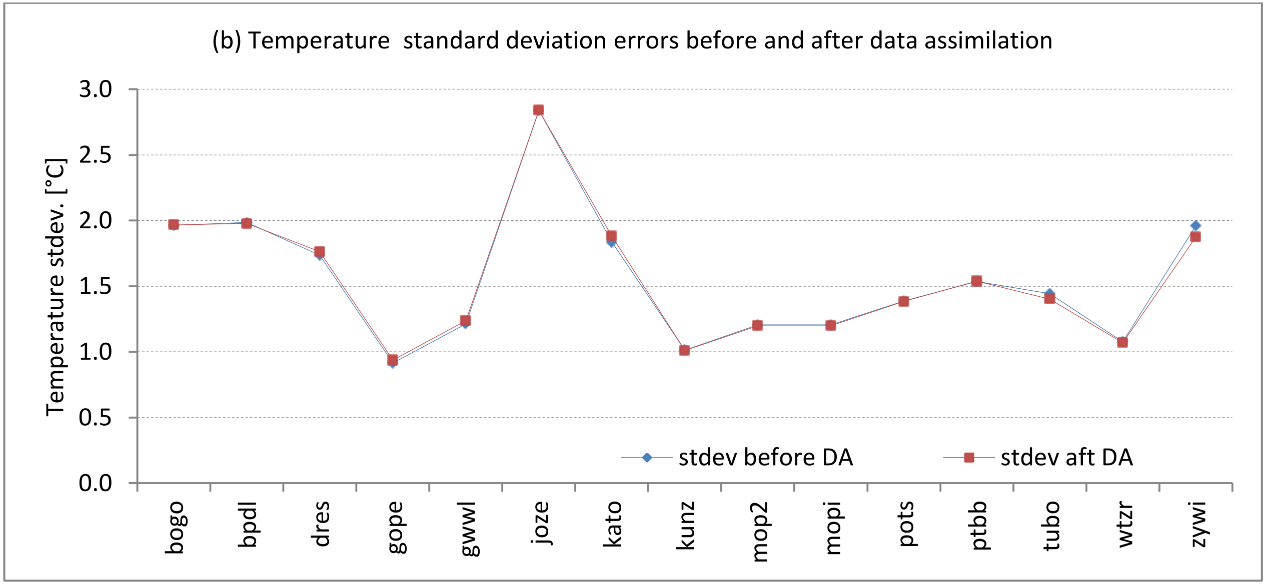

(b) Temperature standard deviation errors before and after data assimilation

Figure 7.4: Near-surface temperature mean accuracies before and after ZTD data assimilation at EPN sites, (a) bias errors and (b) standard deviations.

A visual illustration of the impact on short-range surface temperature forecasts over the entire domain is shown in Figure 7.5. Just like in the case of the pressure variable, changes are hardly noticeable. The area circled in green is a region close to the site GWWL.

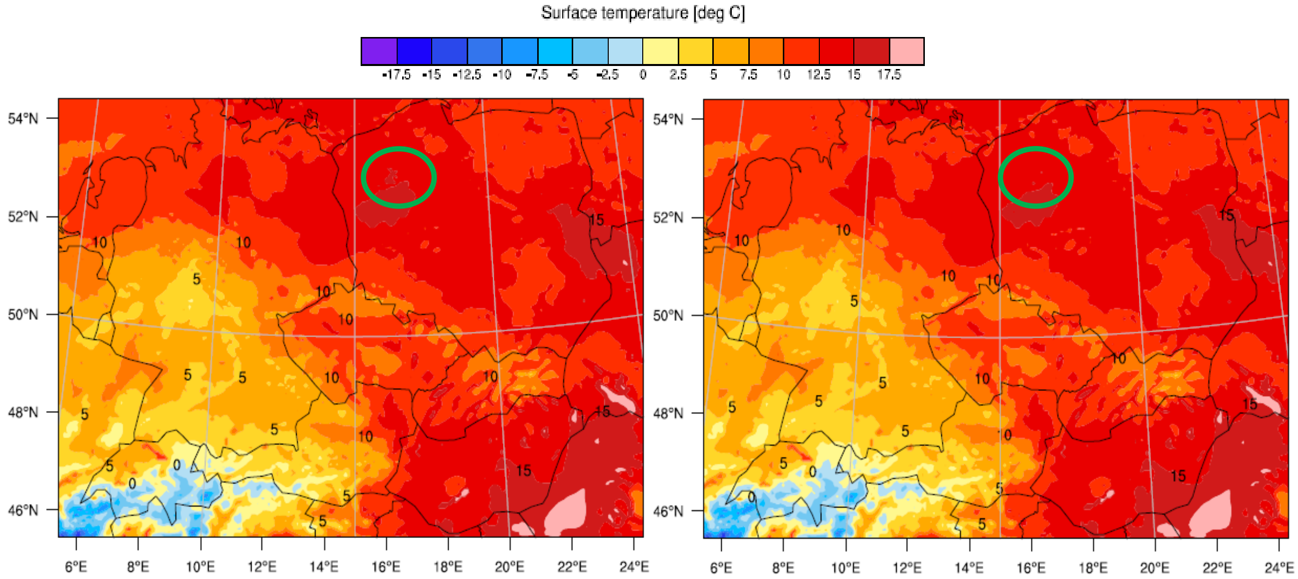


Figure 7.5: Contour plots of the 2 m temperature field over the entire domain on 29th May 2013 19 UTC. The panel on the left-hand side shows temperature forecast before GPS ZTD DA, while the one on the right-hand side shows the forecast after DA.

Some changes, albeit small, in the near-surface relative humidity bias errors are noticed at GWWL, JOZE, MOP2, MOPI and ZYWI (Figure 7.6 (a)). Most of these changes are small reductions in bias errors (a modest impact of the assimilated ZTD observations) with values as high as 0.4% at MOPI.

Unlike the standard deviation errors in pressure and temperature that showed largely negligible changes after the assimilation of GPS ZTD observations, the humidity variable showed changes in the standard deviation errors, albeit small, at majority of the sites considered (Figure 7.6 (b)).

Changes in the relative humidity forecasts across the study domain are noticeable as illustrated in the plots of Figure 7.7. For instance, a larger region of the domain showed relative humidity forecasts mostly ranging from 56% to 60% before GPS ZTD data assimilation. This changed to mostly 60% to 64% relative humidity forecasts after DA.

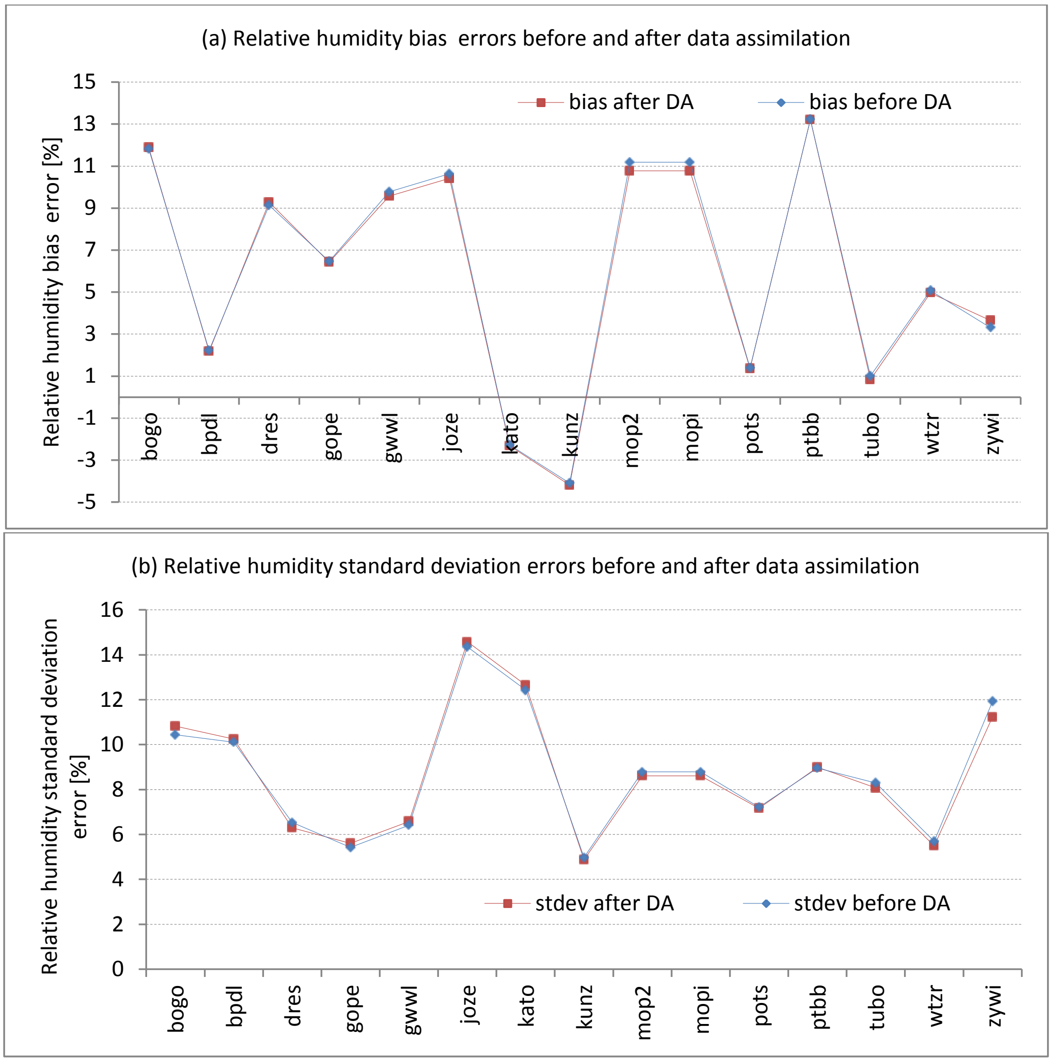


Figure 7.6: Near-surface mean accuracies in relative humidity before and after ZTD data assimilation at EPN sites, (a) bias errors and (b) standard deviations.

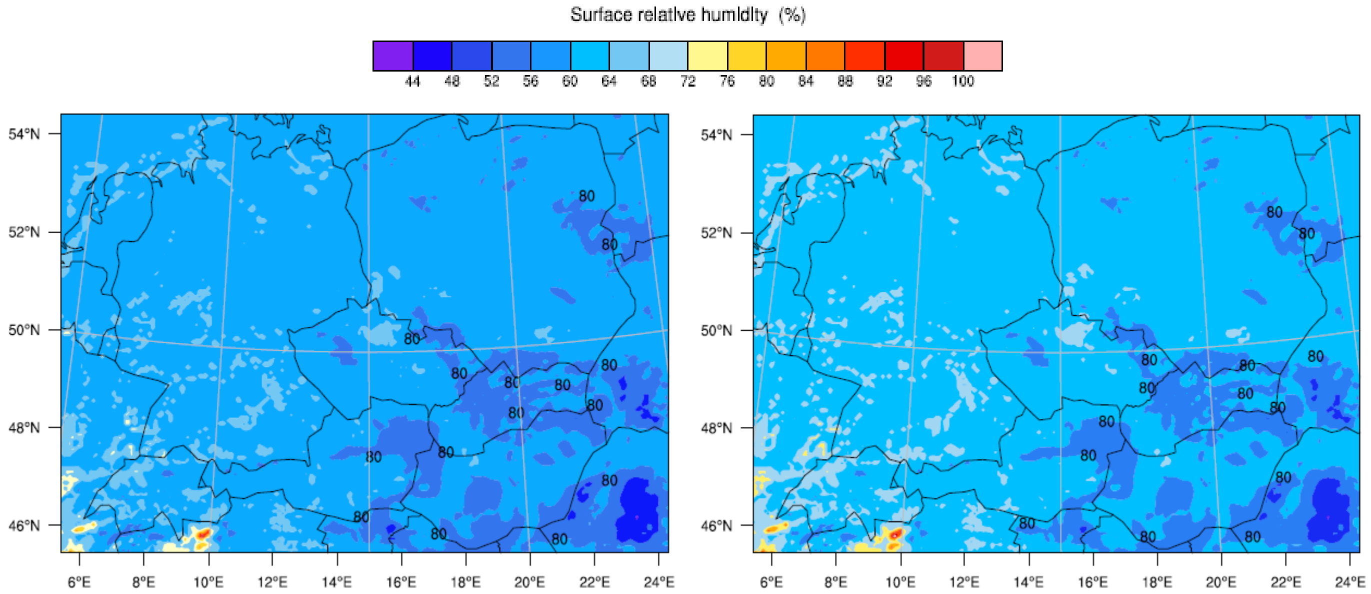


Figure 7.7: Contour plots of the near-surface relative humidity field over the entire domain on 29th May 2013 20 UTC. The panel on the left-hand side shows relative humidity forecast before GPS ZTD DA, while the one on the right-hand side shows the forecast after DA. The left panel is slightly darker than the right panel.

7.4. Summary

Ground-based GPS ZTD observations from 28 EPN stations spread across the study domain have been assimilated into an NWP model. The impact of this activity on short-range forecasts of surface pressure, 2 m temperature and near-surface relative humidity has then been assessed by comparing the errors obtained before and after assimilating the GPS ZTD observations.

A reduction in the errors was termed as a positive impact, an increase was termed as a negative impact, a significantly small change was termed as a modest impact and a neutral impact was taken to mean a situation where the errors before and after DA remained almost equal.

The impact on short-range forecasts of the pressure and temperature surface fields was found to be largely neutral. A modest impact was observed for the relative humidity variable indicating a small reduction in forecast errors due to the assimilated ZTD observations.

Generally, the impacts were found to be site-specific such as the 0.4% reduction in bias error for the relative humidity observations at MOPI (and MOP2) and barely any change at BPDL, GOPE and POTS.

8. Tropospheric delays and precipitable water vapor estimates over Kenya

8.1. Introduction

Kenya lies in the African equatorial region with the equator dividing the country in two almost equal halves that extend approximately 5° north and south of the equator. The equatorial region is mainly characterized by a sun angle that remains relatively high in the sky throughout the year. As a consequence, the equatorial atmosphere holds more water vapor than any other region on the earth (Figure 8.1). Additionally, this is a region where short periodic variations of water vapor occur (Musa *et al.*, 2005). The equatorial weather, therefore, is mainly characterized by hot and wet conditions.

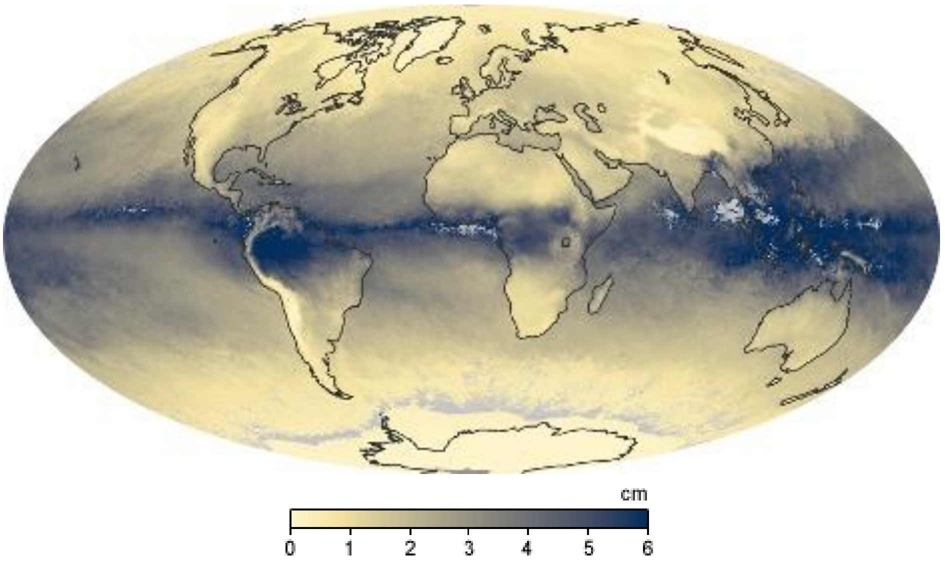


Figure 8.1: The map shows global average amounts of precipitable water vapor for May 2013 collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on NASA's Aqua satellite(NASA, 2014).

Knowing that a large amount of water vapor is found in the equatorial region, a better understanding of tropospheric effect on GNSS positioning accuracy is needed. However, owing to lack of equipment and a sparse network of GNSS reference sites particularly in the African equatorial region (Appendix A); little has been done in terms of troposphere delay studies.

It was mentioned in Section 1.2 that the GPS has come to be widely used as an alternative source for water vapor observations which are then used in numerical weather prediction (NWP) models to improve precipitation forecasts. The accuracy of these observations depends to a large extent on the weighted mean temperature, T_m. This parameter is usually estimated using radiosonde data and using simulated data from NWP models in situations where radiosonde data is missing. More commonly, however, it is computed using T_m models which are regional or site-specific and have been developed using radiosonde data.

In the present study, the Bevis T_m model has been assessed for suitability in a region covering Kenya by comparing it to site-specific and regional models developed using both radiosonde and NWP model data. The accuracy in GPS-derived PWV computed using the Bevis T_m model has also been assessed by comparing it with that derived from numerical integration of the appropriate variables in NWP model.

8.2. Methodology

The NWP model computational domain was defined as shown in Figure 8.2. It extends from approximately 34°E to 42°E in longitude and approximately 5°N to 5°S in latitude. The horizontal dimension consists of 180 by 216 grid points in x- and y-axis respectively with a grid spacing of 5 km. The vertical dimension consists of 45 vertical levels. The Mercator map projection type was used implying the use of the ARW model as the WRF dynamical core for the simulation. The Kessler and Kain-Fritsch schemes were used for microphysics and cumulus parameterization respectively.

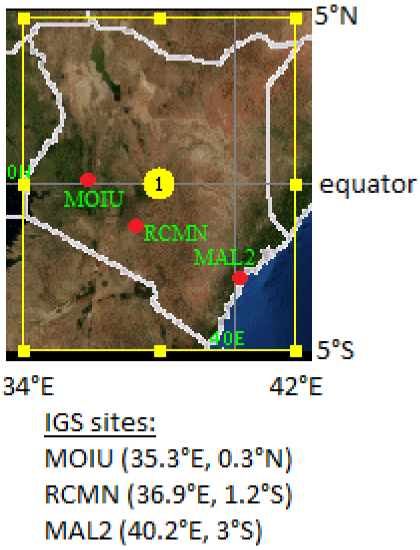


Figure 8.2: Study domain covering Kenya. Red marks show the relative locations of the IGS stations. The computational domain is enclosed in the yellow rectangle.

Initialization was done by use of the GFS 0.5 degree global model and simulations were made for each day of January 2014 with each simulation having a forecast length of 24 hours.

Using the techniques described in Section 3.1, variables were extracted from the NWP model output files and used to compute troposphere delays and precipitable water vapor.

Hourly final troposphere zenith path delay observations (accuracy of 4 mm) were sourced from the IGS archive for the entire month of January 2014. This period encompasses the 4th day of GPS week 1773 and the 6th day of GPS week 1777.

The study targeted the derivation and validation of various quantities including hydrostatic delays, wet delays, total path zenith delays, weighted mean temperatures (T_m) and precipitable water vapor (PWV). Details of the methods used to achieve these tasks are similar to those discussed in Chapters 5 and 6.

8.3. Brief description of the Kenyan terrain and data availability

Kenya is characterized by a varied terrain that is defined by the East African rift system that includes volcanoes such as Mount Kenya (approximately 5300 m) and the slopes of Mount Kilimanjaro (close to the border with Tanzania) and Mount Elgon (approximately 4300 m high, on the border with Uganda). The elevation is lowest close to and along the Indian Ocean coastline (which stretches for about 563 km) and some of the regions lying in the rift valley including regions around Lake Turkana. This results into a highly diversified terrain with climatic conditions that range from moist to arid.

The estimation of troposphere delay and precipitable water vapor by use of NWP model data was discussed in Sections 2.4 and 2.5 respectively. A validation of NWP-derived troposphere delay estimates using GPS ZTD observations is currently possible at only three IGS reference sites in Kenya whose details are summarized in Table 8.1 and their locations shown in Figure 8.2. These stations are all found in the central and lower parts of the country with none in the upper part[1].

Table 8.1: IGS stations currently in Kenya (IGS Central Bureau, 2014).

Station id and city/municipality	Longitude (E) [degrees]	Latitude (N) [degrees]	GPS ellipsoidal height [m]
RCMN (Nairobi)	36.8938	-1.2210	1607.5400
MAL2 (Malindi)	40.1940	-2.9960	-20.4000
MOIU (Eldoret)	35.2900	0.2883	2201.5321

All the three sites are currently active and GPS ZTD observations for the stations are available at the IGS archives. MAL2 is close to seashore (the Indian Ocean) hence the negative ellipsoidal height indicates that the ellipsoid (WGS84[2]) surface at the site is overhead.

Radiosonde data for Kenya is currently freely available from the IGRA[3](IGRA, 2008) and SHADOZ[4] (Thompson, *et al*., 2007) archives (Table 8.2). While most of the data is found in the IGRA archive, the most recent records are mainly available in the SHADOZ archive with Nairobi (IGRA up to 2012, SHADOZ up to 2014) and Malindi (SHADOZ up to 2006) stations containing the most recent data.

Figure 8.3 shows positions where the IGS and radiosonde stations in Kenya are roughly located.

---

[1] This is the least inhabited part of Kenya owing to its semi-arid nature.

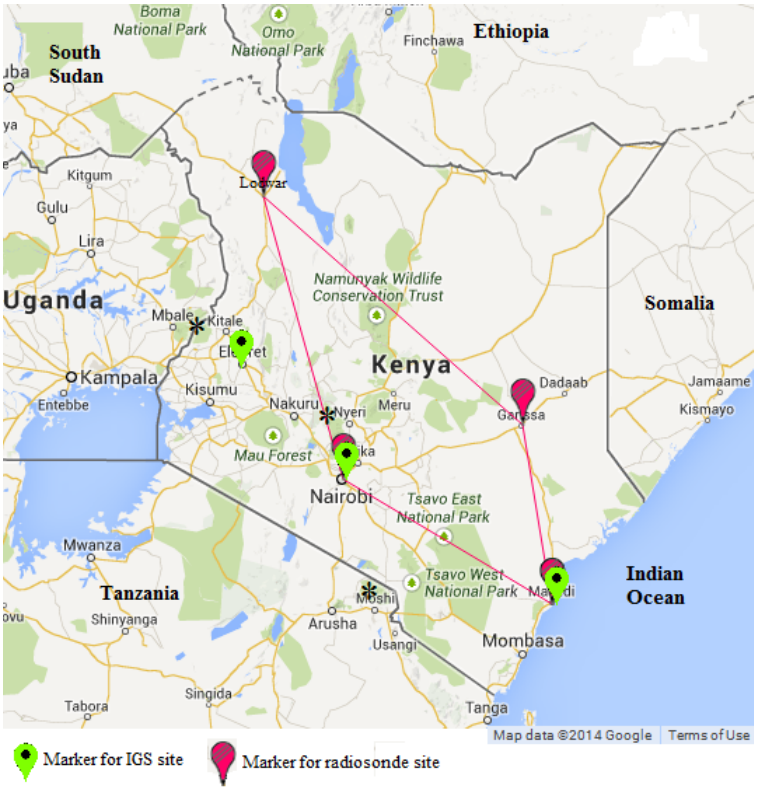
[2] World Geodetic System 1984.

[3] Integrated Global Radiosonde Archive.

[4] Southern Hemisphere Additional OZonesondes.

Table 8.2: Radiosonde data availability in Kenya (IGRA, 2008; Thompson, 2014). Entries in bold print (last column) signify the records that happen to be found in the SHADOZ archive.

WMO Code	Location	Lat. (N) [degrees]	Lon. (E) [degrees]	Elevation (asl) [m]	Availability (first and last year of record)
63612	LODWAR	3.12	35.62	515	1975, 1999
63624	MANDERA	3.93	41.87	230	1973, 1990
63714	NAKURU	-0.27	36.07	1871	1973, 1999
63723	GARISSA	-0.47	39.63	147	1973, 1999
63741	NAIROBI	-1.30	36.75	1798	1966, 2012 **1998, 2014**
63766	MAKINDU	-2.28	37.83	998	1973, 1996
63793	VOI	-3.40	38.57	560	1973, 1995
63799	MALINDI	-3.23	40.10	24	1973, 1988 **1999, 2006**



Approximate distances with respect to Nairobi: Malindi - 424 km,Garissa - 324 km, Lodwar - 510 km Eldoret - 265 km

Asterix indicate Mt. Kenya (central Kenya), Mt. Elgon (border with Uganda) and Mt. Kilimanjaro (close to border with Tanzania)

Source: http://www.daftlogic.com/projects-google-maps-distance-calculator.htm

Figure 8.3: IGS and radiosonde sites in Kenya.

8.4. Troposphere delays over Kenya

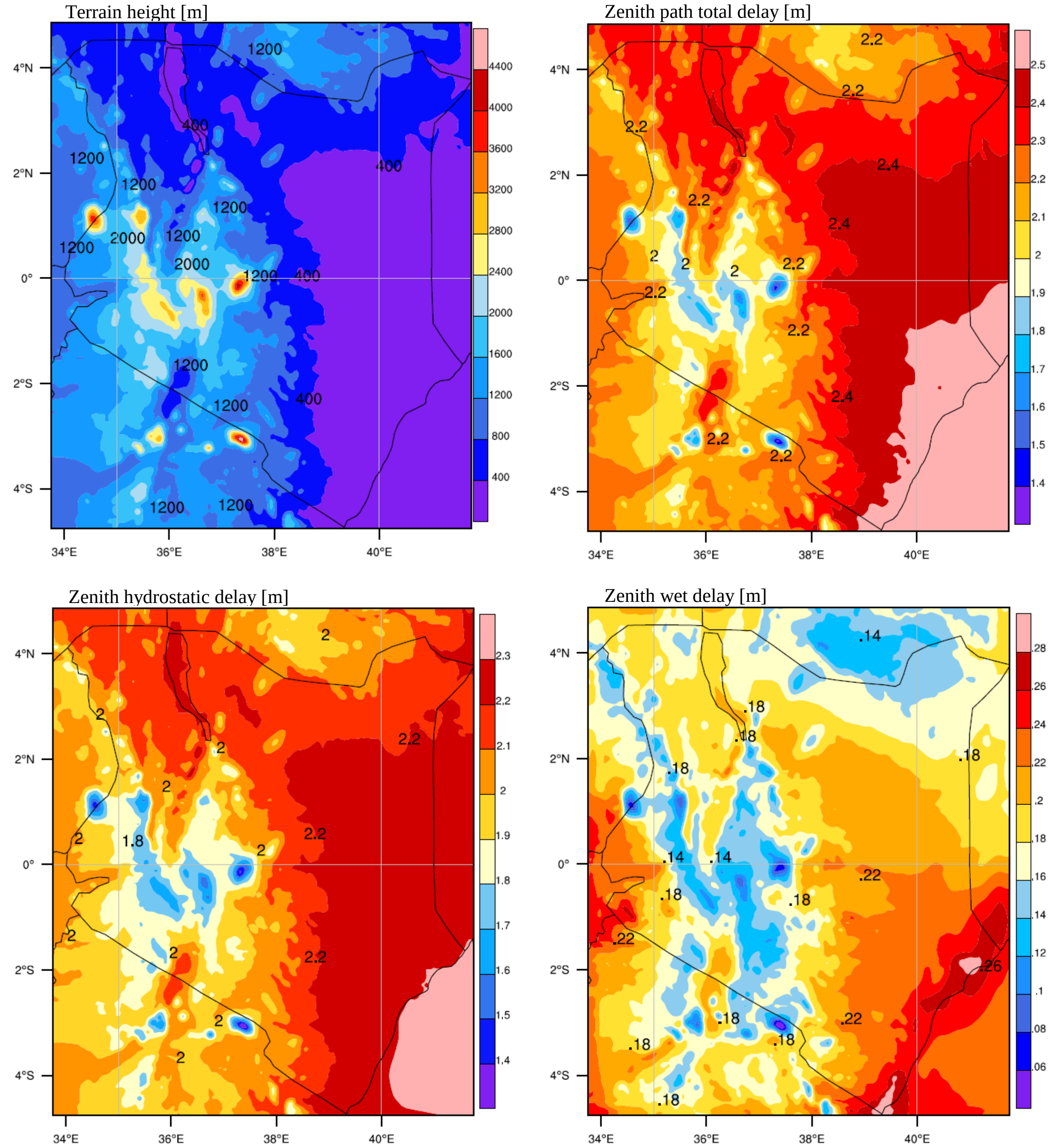
The terrain heights across Kenya computed from the NWP model on 1st January 2014 00 UTC and shown in the top-left panel of Figure 8.4 provides us with an insight of the expected magnitudes of the troposphere delays.

The lowest regions average below 400 m in elevation while the highest regions average approximately 2000 to 3200 m in elevation. A few other regions are having elevations greater than 4000 m. The highest elevations greater than 4000 m above mean sea level are seen close to the peaks of Mt. Kenya, Mt. Elgon and on the slopes of Mt. Kilimanjaro.

From the top-right panel of Figure 8.4, total zenith path delays for Kenya as derived from the NWP on the same day and time as the terrain heights (top-left panel of Figure 8.4), show that the ZTD estimates are largest, approximately 2.5 to 2.6 m, at regions of low elevation. This includes regions close to the Indian Ocean, eastern and northern Kenya. The central and western parts of the country are showing the least amounts (less than 2.0 m) of ZTD particularly the regions surrounding Mt. Kenya and Mt. Elgon.

A pattern similar to that of ZTD is shown for ZHD in the bottom-left panel of Figure 8.4 with the greatest amounts being observed along the coastline, northern and eastern Kenya and averaging about 2.4 m while the least amounts are observed in regions around central and western Kenya and are less than 2.2 m.

The ZWD in the bottom-right panel of Figure 8.4 is seen to be greatest along the coastline and in the eastern part of the country where the mean wet delays are seen to range from approximately 0.24 m to 0.3 m.


Terrain height [m]

Zenith path total delay [m]

Zenith hydrostatic delay [m]

Zenith wet delay [m]

Figure 8.4: Contour plots showing NWP-derived terrain heights (top-left panel), ZTD (top-right panel), ZHD (bottom-left panel) and ZWD (bottom-right panel) over Kenya on 1st January 2014 00 UTC.

From the description provided in the preceding paragraphs and with reference to the location of the three IGS sites as shown in Figure 8.3, it is expected that MAL2 would have the greatest amounts of troposphere delays and precipitable water vapor estimates followed by RCMN and MOIU in that order.

8.5. Validation of NWP-derived troposphere delays at the IGS sites in Kenya.

Geoid heights, computed using the EGM96 geoid undulation model, were used to calculate the orthometric heights for the three IGS sites in Kenya and were then compared to the geometric heights corresponding to the NWP model's lower-most mass level in order to get height difference estimates between the vertical positions of the NWP model's lower-most mass level and the antenna at the three IGS stations. The results are shown in Table 8.3.

Table 8.3: NWP model orography/GNSS antenna height comparison.

Station id.	**NWP ht. [m]**	**$\mathrm{GNSS_{ant}}$ ht. - NWP ht.[m]**
MAL2	30.6	-20.5
MOIU	2178.4	36.8
RCMN	1676.2	-52.8

The antenna is below the model orography at MAL2 and RCMN and it lies within the model orography at MOIU. Consequently, we are likely to have errors due to interpolation procedures at MAL2 and RCMN.

Figure 8.5 (also see Appendix F) shows the GPS and NWP estimates of ZTD at the three sites for the month of January 2014. ZTD estimates are indeed largest at MAL2 (average of 2.5 m for GPS and 2.6 m for NWP), followed by RCMN (average of 2.0 m for both GPS and NWP) and MOIU (average of 1.9 m for both GPS and NWP).

There were no GPS ZTD data at RCMN on DOY 12, 13, 25 and 26. A closer look into evolution of the troposphere delays at each of the IGS stations helps us see the trends better. This is shown in Figure 8.6.

The patterns in the ZTD trends seem to agree fairly well at all the sites. However, the NWP-derived ZTD estimates for MAL2 are significantly greater than those of GPS at all times (Figure 8.6 (a)). The ZTD trends and magnitudes at MOIU are largely similar (Figure 8.6 (b)).

In Figure 8.6 (c), the GPS and NWP ZTD trends appear to be generally similar most of the times with small differences in ZTD magnitudes some of the times. Instances of the missing GPS ZTD observations are seen in the discontinuities in the GPS ZTD trend. This serves to demonstrate that the NWP-derived ZTD estimates may be relevant in situations where GPS ZTD observations are not available.

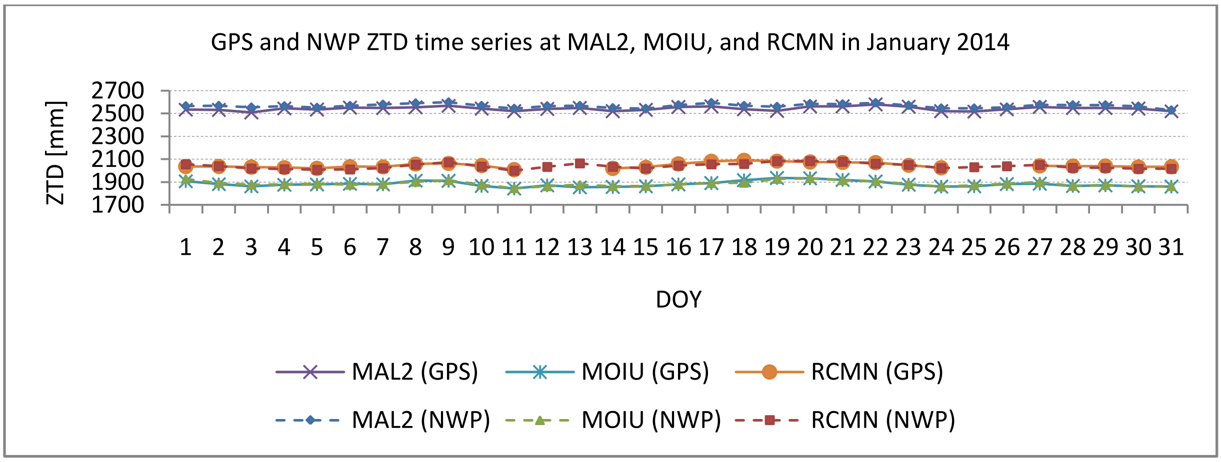
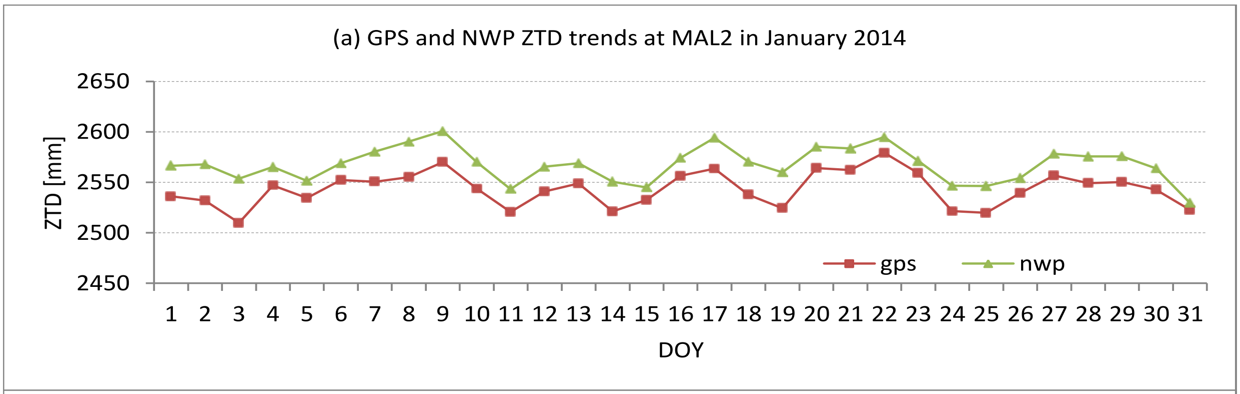


Figure 8.5: Time series for GPS and NWP ZTD estimates at IGS sites in Kenya for January 2014.



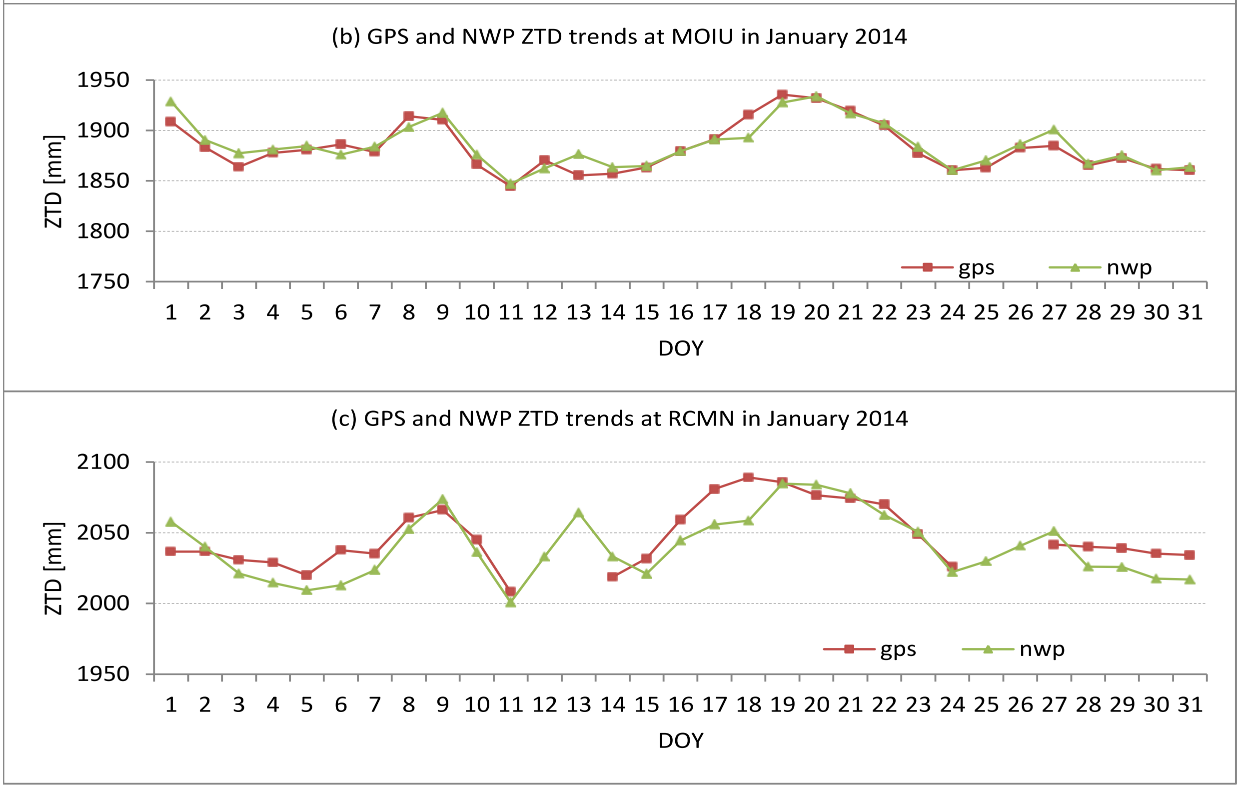


Figure 8.6: Times series for GPS and NWP ZTD at MAL2 (a), MOIU (b) and RCMN (c) in January 2014.

Based on the GPS ZTD observations and the NWP simualtions for ZTD at the IGS sites discussed in the preceding paragraphs for the month of January 2014, an accuracy assessment was done in terms of bias error computations with the GPS ZTD observations as the reference values and bias-reduced standard devaition error computations. The resulting day-to-day mean accuracies are shown in Figure 8.7.

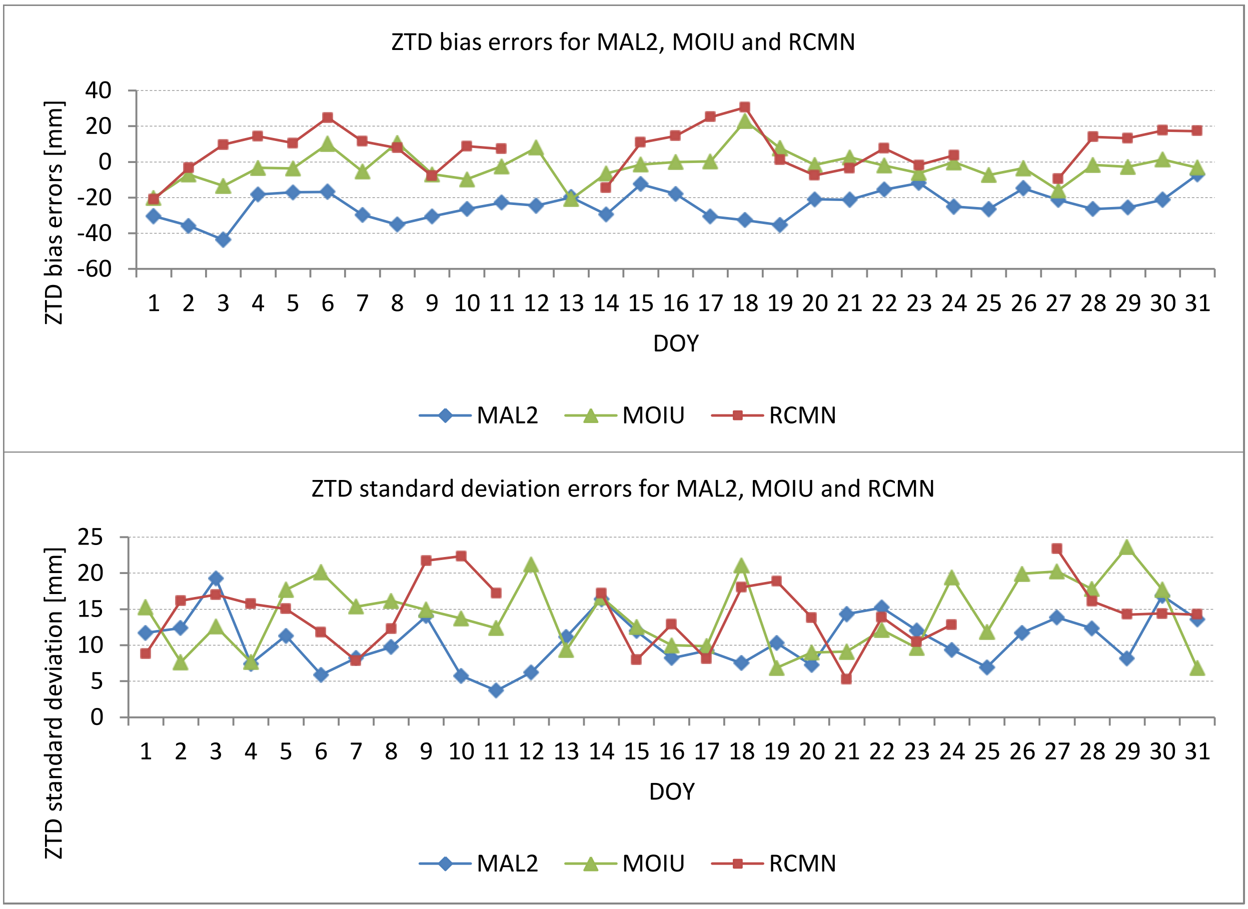


Figure 8.7: Mean ZTD (a) bias and (b) standard deviation errors for MAL2, MOIU, and RCMN in January 2014.

Consistent with the ZTD observations seen in Figure 8.7, the largest bias errors were observed at MAL2 followed by RCMN and MOIU. Bias errors at MAL2 were found to be negative during the entire study period with a maximum of approximately 44 mm on day 3 and a minimum of approximately 7 mm on doy 31. The resulting overall negative mean bias error at MAL2 was found to be approximately 24 mm.

The bias errors at MOIU over the study period are both positive and negative with an overall negative mean value of approximately 3 mm. RCMN shows positive biases most of the times with slightly greater magnitudes than those observed at MOIU and smaller than those observed at MAL2. The overall mean bias error at RCMN was, therefore, found to be approximately 7 mm.

On the other hand, a smaller precision error is observed at MAL2 with a mean accuracy of approximately 11 mm compared to the mean accuracy of approximately 14 mm at MOIU and RCMN.

In Table 8.3, the GNSS antenna/model orography height difference was found not only to be negative (antenna below model orography) but also smaller in magnitude at MAL2 compared to RCMN. The systematic errors in the NWP-derived delays are, however, larger at MAL2 compared to RCMN. Errors at MAL2 may, therefore, be attributed to a combination of factors including the negative height difference and proximity to the Indian Ocean.

8.6. Weighted mean temperature estimates over Kenya

The weighted mean temperature T_m was estimated by use of NWP data in a way similar to the method discussed in Section 6.3. The T_m estimates were plotted against NWP near-surface temperature (2 m) simulations for the entire study period including hourly simulations for each day in the 31-day study period. The resulting scatter plots of the T_m/T_s data points and the regression line are shown in Figure 8.8.

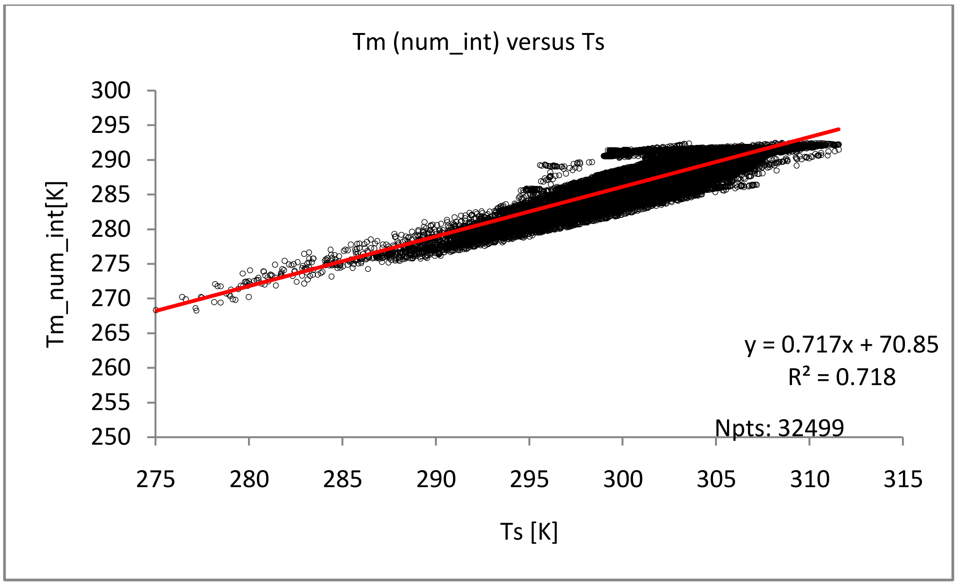


Figure 8.8: NWP-derived T_m-T_s relation over the entire study period and domain covering Kenya.

The resulting T_m regression model is of the form,

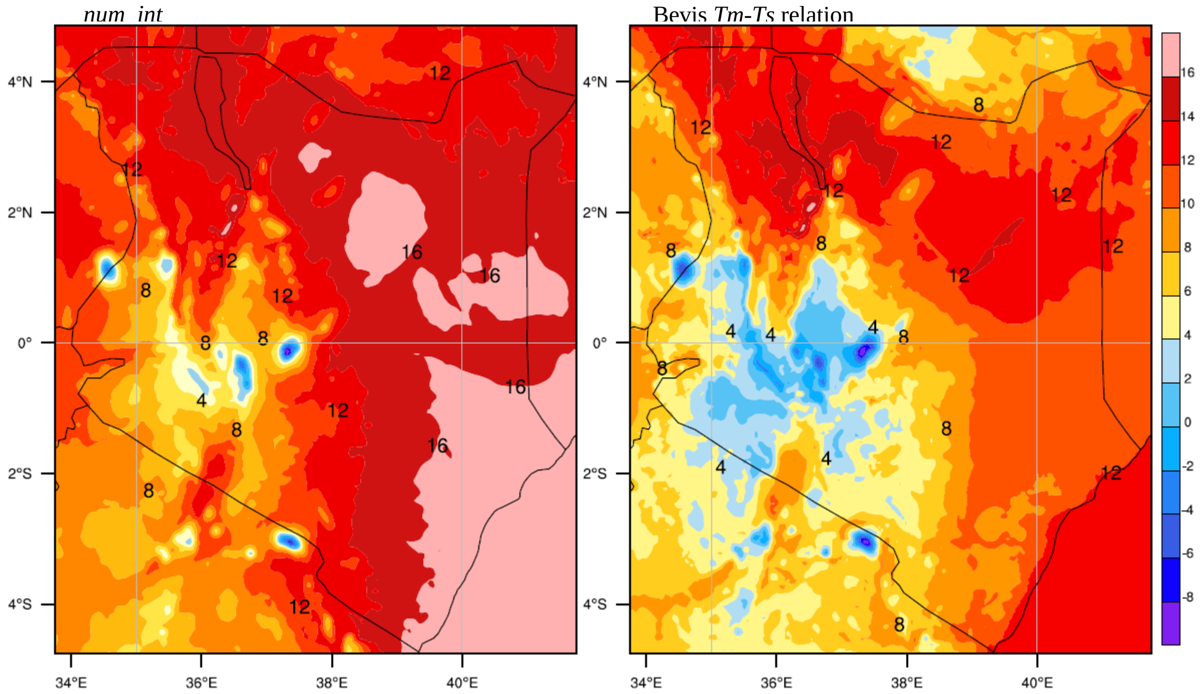
$$T_m = 71.66 + 0.71 \cdot T_s \tag{8.1}$$

Where, $T_s \equiv$ the surface temperature in K.

From Figure 8.8, the model is seen to be a good fit as evidenced by the R-squared value of approximately 71.7%. The T_m estimates vary from approximately 268 K to 293 K corresponding to a variation in T_s of approximately 275 K to 312 K. The relationship is mostly linear with a particularly high density of the data points for $T_s > 290$ K.

A visual comparison of T_m estimates simulated for 15th May 2013 00 UTC and 12 UTC across the entire study domain covering Kenya and using the numerical integration method and the Bevis T_m-T_s relation is shown in Figure 8.9.

(a) *num_int* and Bevis T_m estimates across Kenya for 15th January 2014 00 UTC



(b) *num_int* and Bevis T_m estimates across Kenya for 15th January 2014 12 UTC

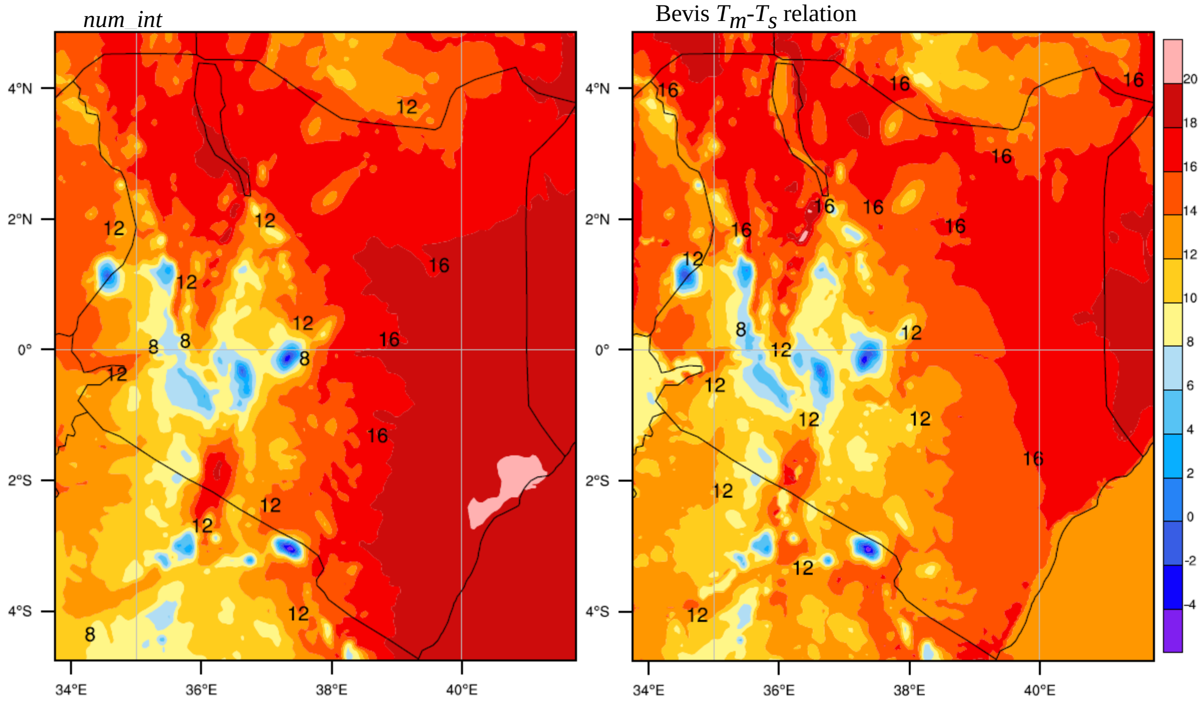


Figure 8.9: Night- and day-time numerical integration (*num_int*) and Bevis T_m estimates over Kenya on 15th January 2013.

From the night-time simulations (Figure 8.9 (a)), weighted mean temperature estimates using the numerical integration (*num_int*) method (left panel) appear to be slightly greater than those computed using the Bevis T_m-T_s relation (right panel). Day-time simulations (Figure 8.9 (b)), however, appear to produce almost similar results for both the numerical integration method (left panel) and the Bevis T_m-T_s relation (right panel).

Although there seems to be some visual differences in the results of the numerical integration and Bevis T_m simulations illustrated in Figure 8.9, the actual magnitudes of the differences may not be significant. This point is strengthened by the fact that the regression equations from the two methods are almost similar, i.e.

- Num_int: $T_m = 71.7 + 0.71 \cdot T_s$
- Bevis: $T_m = 70.2 + 0.72 \cdot T_s$

The largest T_m estimates are observed on the eastern part of the domain with the least amount being observed on regions at and surrounding Mt. Kenya (central Kenya), Mt. Elgon (boarder with Uganda) and Mt. Kilimanjaro (boarder with Tanzania).

Though the NWP-derived T_m model compares relatively well with the Bevis T_m model, we note that it is only based on a one month simulated data (the Bevis model is based on actual radiosonde observations for a period spanning approximately two years). It may therefore not be a true representative of a T_m model that is suitable for use in Kenya. The comparison therefore shows that the Bevis T_m model is suitable for use in Kenya.

8.7. Radiosonde-derived T_m estimates at selected radiosonde sites in Kenya

The relationship between the weighted mean temperature and surface temperature was investigated at four radiosonde sites in Kenya for which data was available at both the IGRA and SHADOZ archives. These sites are Lodwar, Garissa, Malindi and Nairobi (Table 8.2). Data records available for all the stations are once per day records and at varied frequencies per month with an average of four records per month for Lodwar, Garissa and Malindi and five records per month for Nairobi.

Data records for Lodwar and Garissa were also found to be quite inconsistent with huge gaps in between. For instance, at Lodwar, though the data records are stated as 1989 to 1999, there are no data records between 1991 and 1997 with only a single entry in 1992.

Nairobi and Malindi have the most recent data records. Data spanning the period 1990 - 2005 was considered for the two stations. The data is more consistent at the two stations compared to Lodwar and Garissa with the data for Nairobi being at a much greater frequency.

Weighted mean temperature estimates at the four radiosonde stations were computed in a method similar to the one described in Section 6.4. Based on the findings in the brief review discussed in the preceding paragraphs and the data periods considered, there are 218 samples for Lodwar and Garissa, 107 samples for Malindi and 323 samples for Nairobi.

Plots showing the relationship between the weighted mean temperature estimates (T_m) and surface temperature observations (T_s) for Lodwar and Garissa are shown in Figure 8.10.

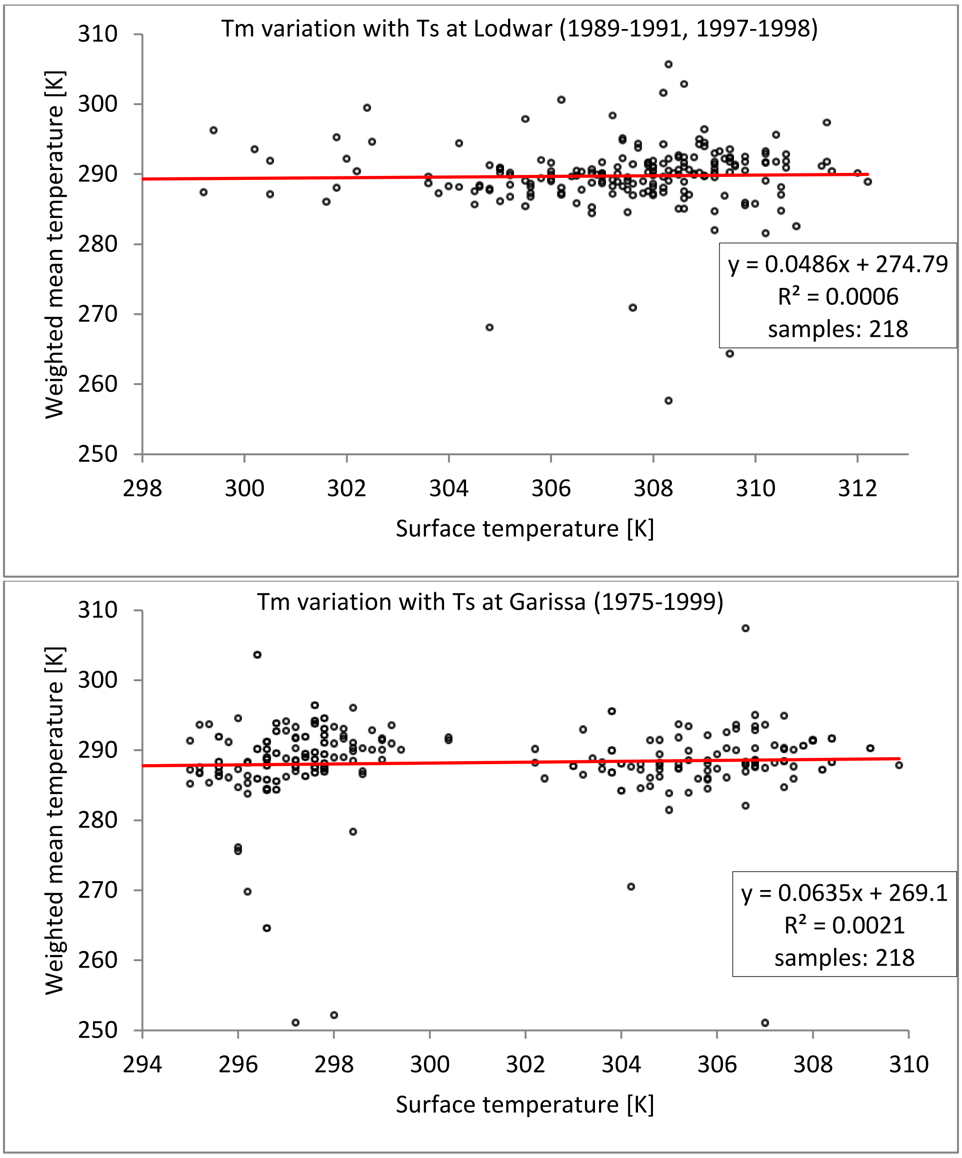


Figure 8.10: Site-specific T_m models at the Lodwar and Garissa radiosonde sites.

From Figure 8.10, the T_m models for Lodwar and Garissa respectively are,

- Lodwar: $T_m = 274.8 + 0.05 \cdot T_s$
- Garissa: $T_m = 269.1 + 0.06 \cdot T_s$

The models, however, do not properly fit the data, as evidenced by the extremely low R-squared values). This tells us that the radiosonde data available at these sites is not sufficient to deduce the site-specific weighted mean temperature models. It may also be due to some erroneous values in the variables used in the computation of T_m.

As shown in Figure 8.11, the Malindi and Nairobi radiosonde sites show better T_m models, as evidenced by the 20.2% and 11.3% R-squared values for Malindi and Nairobi respectively, and the corresponding site-specific T_m models,

- Malindi: $T_m = 176.5 + 0.37 \cdot T_s$
- Nairobi: $T_m = 165.7 + 0.39 \cdot T_s$

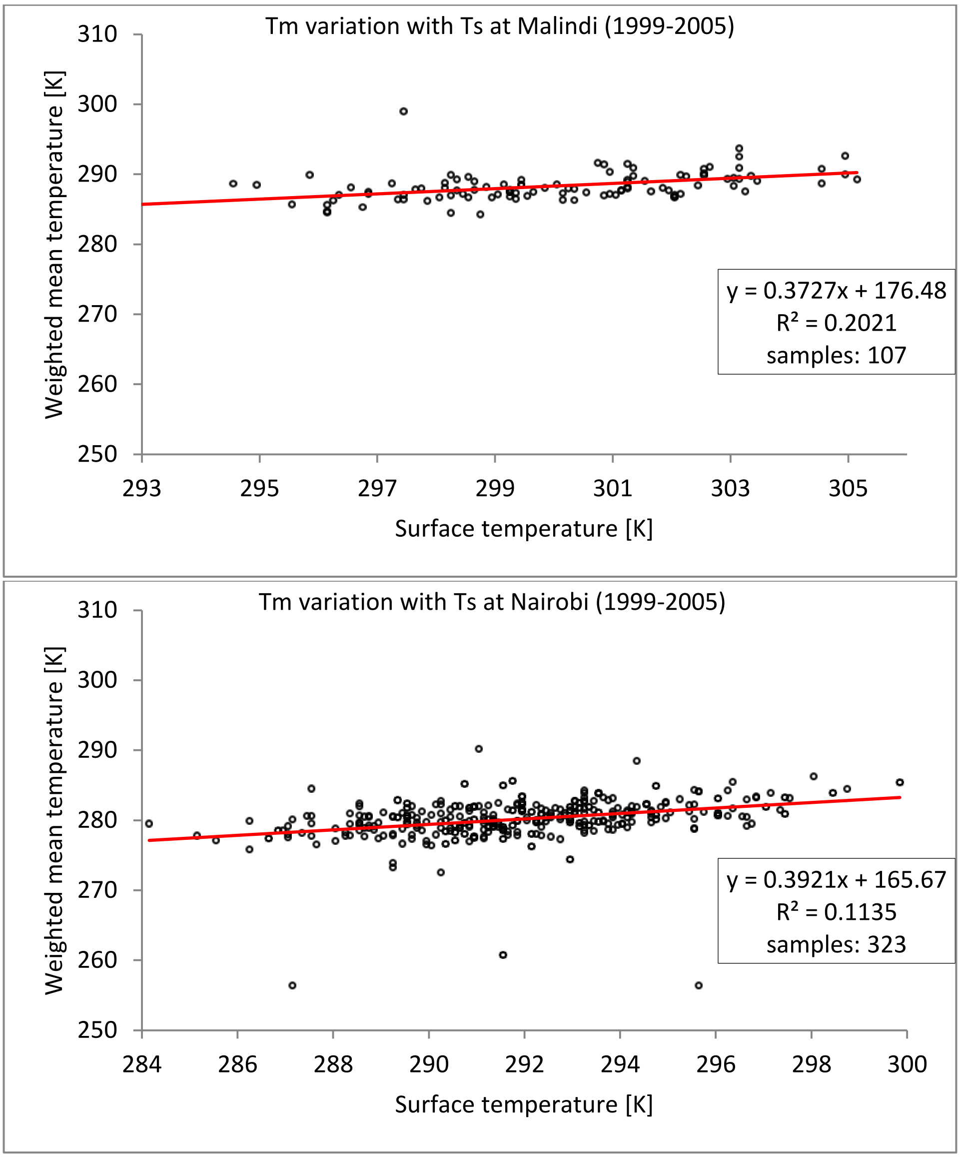


Figure 8.11: Site-specific T_m models at the Malindi and Nairobi radiosonde sites.

The site-specific T_m model pairs Lodwar/Garissa and Malindi/Nairobi are nearly similar. This nature of the regression equations at the station pairs may be attributed to the fact that the climatic conditions at Lodwar and Garissa are similar, i.e. semi-arid, while those at Malindi and Nairobi may be said to be humid and moist respectively.

Based on the site-specific T_m models at the four radiosonde sites considered, regional T_m models suitable for regions between Lodwar and Garissa and between Malindi and Nairobi are shown in Figure 8.12 (a) and (b) respectively.

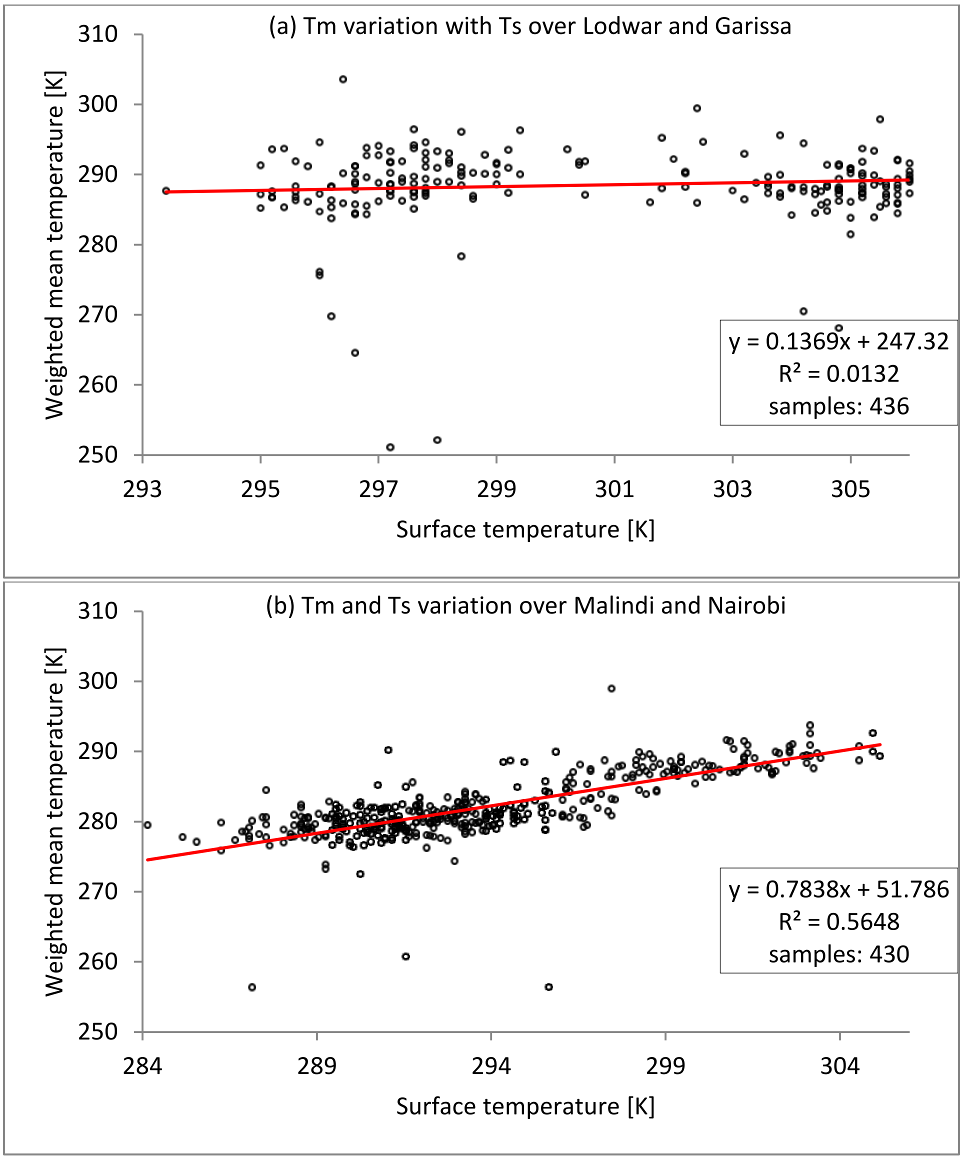


Figure 8.12: Regional T_m models (a) over Lodwar and Garissa and (b) over Malindi and Nairobi.

As shown in Figure 8.12 (a), a T_m model for a region between Lodwar and Garissa yields a lower value of R-squared compared to the values obtained for the site-specific T_m models. The resulting relation connecting T_m and T_s is of the form,

$$T_m = 247.3 + 0.14 \cdot T_s \qquad (8.2)$$

T_s are seen to vary from approximately 295 K to 306 K resulting into variations in T_m of approximately 289 K TO 290 K. The highest density of points is observed for T_s values between approximately 295 K and 299 K and between approximately 304 K and 308 K. The observed T_s values are typical of this region.

The plot of Figure 8.12 (b) is taken to be a representative regional T_m model for the region between the coast of Kenya (Malindi) and Nairobi. T_s are seen to vary from approximately 287 K to 305 K resulting into variations in T_m of approximately 275 K TO 290 K. The resulting empirical relation connecting T_m and T_s is of the form,

$$T_m = 51.79 + 0.78 \cdot T_s \tag{8.3}$$

The data points fit relatively well to the regression line with an R-squared value of approximately 56.5%. The highest density of points is observed for T_s values between 288 K and 296 K with most of them belonging to Nairobi which lies at approximately 1700 m above mean sea level.

A comparison of the NWP and radiosonde T_m estimates was only possible for the Malindi and Nairobi radiosonde sites. However, the most recent radiosonde data for Malindi is for a launch on 2005-12-05_09 UTC (available at the SHADOZ archives). In order to make an NWP simulation for this date using the EMS, it required a reanalysis using the ECMWF[1] model (as a historical dataset) for initialization. The simulated NWP T_m estimate was found to be approximately 286.9 K while the radiosonde estimate was found to be 287.5 K.

At the time of running the experiment the radiosonde data file (from SHADOZ) for the Nairobi radiosonde site was for 9th January 2014 with the launch time stated as 8.26 UTC. T_m estimates were then computed using NWP and radiosonde data for this site at the stated launch time. They were found to be 280.5 K and 282.3 K for NWP and radiosonde respectively.

8.8. Precipitable water vapor estimates over Kenya

PWV was computed by numerically integrating the humidity variable (and all other appropriate variables) in the NWP model pressure level. Figure 8.13 shows the resulting NWP PWV estimates over the entire study domain covering Kenya based on an NWP simulation for 1st January 2014 12 UTC. The PWV estimates were found to be lowest in the high altitude regions around the slopes of Mt. Kenya, Mt. Elgon and Mt. Kilimanjaro. The eastern part shows the largest amounts of PWV with maximum observations averaging approximately 44 mm along the coast. Most parts in the western half average approximately 24 mm to 32 mm of PWV.

The MAL2 site is the only IGS site that happens to have radiosonde launches and data. In Section 8.7, NWP and radiosonde estimates of T_m at this site were found to be almost equal. Consequently, NWP and radiosonde PWV estimates at the site may as well be assumed to be equally comparable.

GPS ZTD observations at the IGS sites in Kenya were also used in the computation of PWV. The important parameter, T_m, was obtained from NWP model data and by use of the Bevis T_m global model. Therefore, methods similar to those discussed in Sections 6.5 and 6.6 were used to estimate

---

[1] European Centre for Medium-Range Weather Forecasts

PWV using NWP data and ground-based observations including surface pressure and ZTD observations at the IGS sites.

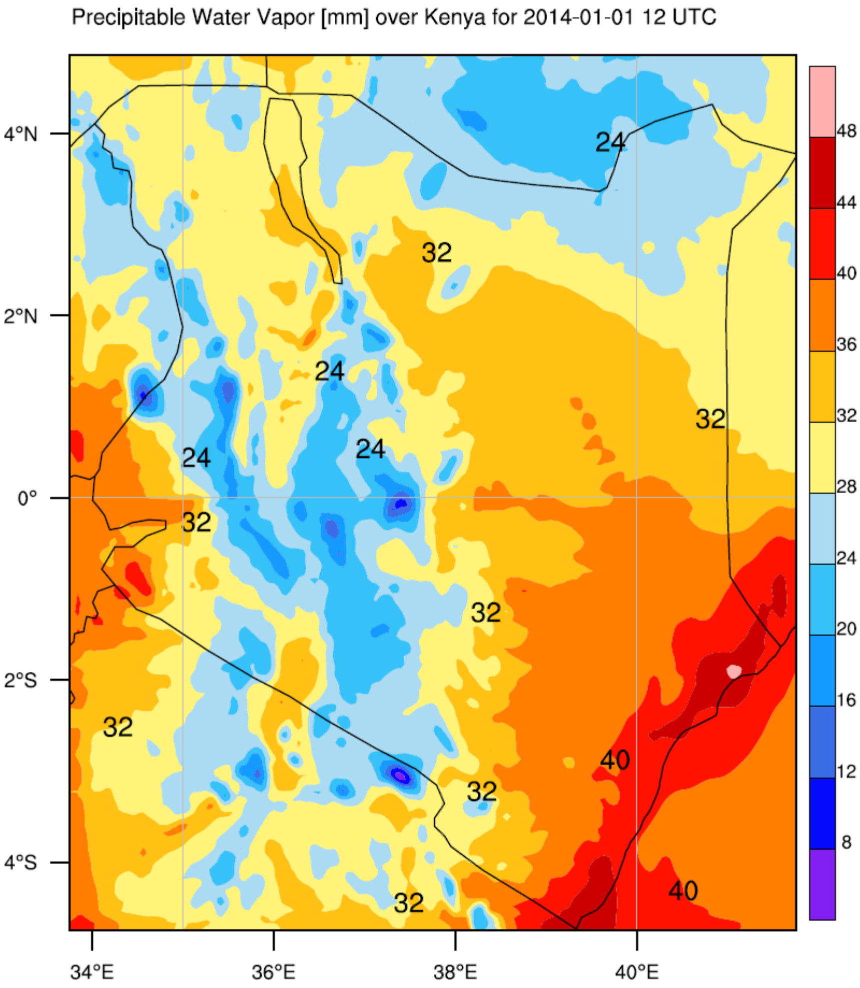

Precipitable Water Vapor [mm] over Kenya for 2014-01-01 12 UTC

Figure 8.13: NWP simulations for precipitable water vapor across Kenya for 1st January 2014 12 UTC.

The resulting trends in PWV derived from GPS observations (using NWP T_m) and NWP model data after every six hours in the month of January 2014 at the IGS stations MAL2, MOIU and RCMN are shown in Figure 8.14. The mean GPS and NWP PWV estimates were found to be 39 mm and 43 mm respectively at MAL2. Those at MOIU and RCMN compared quite well with mean values of 15 mm at MOIU and 22 mm at RCMN. The largest PWV amounts were, therefore, observed at MAL2 while the least amounts were observed at MOIU possibly owing to it being on a high altitude region.

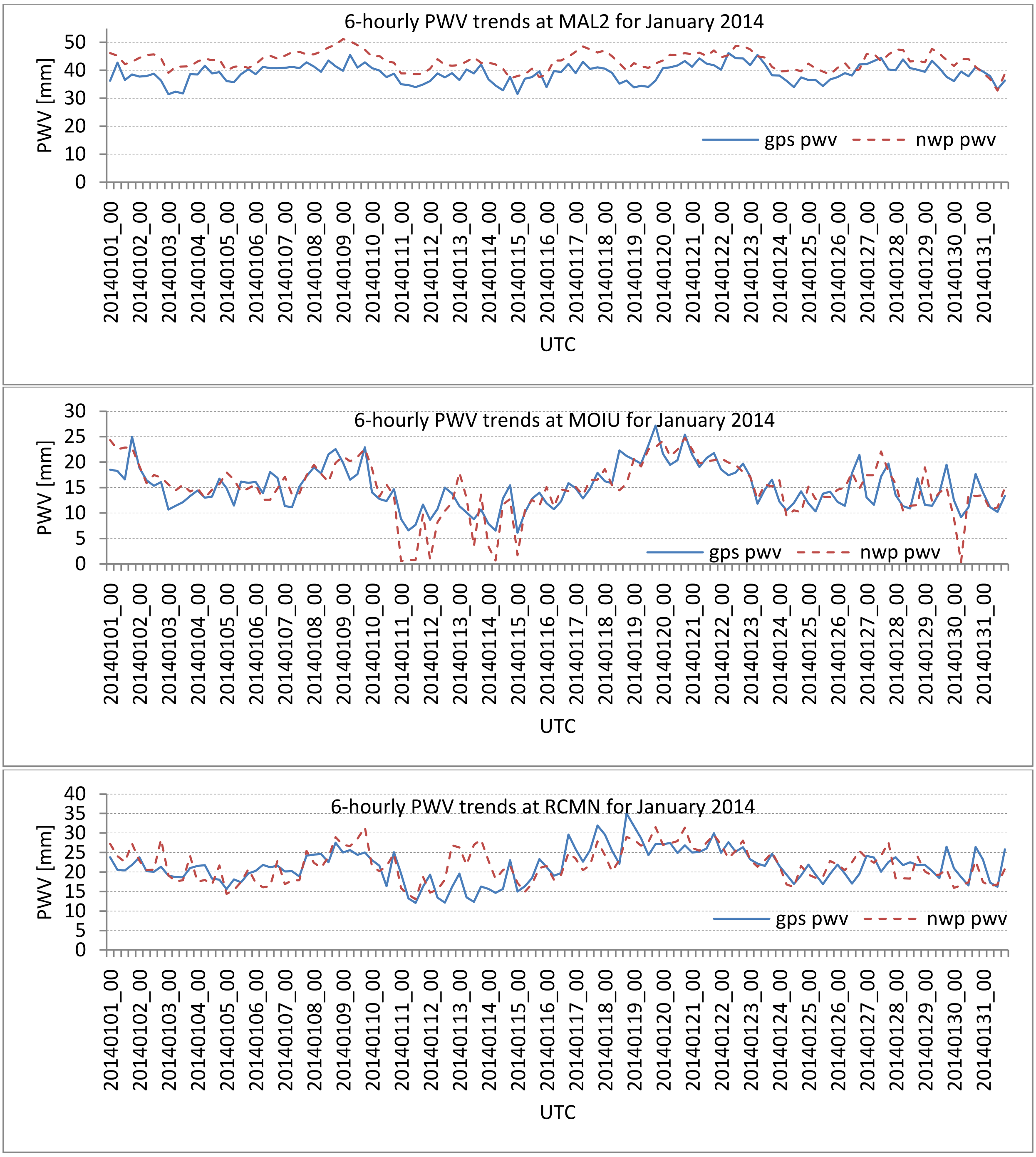


Figure 8.14: GPS and NWP PWV trends at MAL2 (top pannel), MOIU (middle pannel) and RCMN (bottom pannel) for January 2014.

The accuracy and precision in PWV derived using numerical integration of data in the NWP model and using GPS observations (with the Bevis T_m) was computed in terms of bias and standard deviation errors with GPS-derived PWV (with NWP T_m) as the reference. The results are presented in Table 8.4. GPS-derived PWV with Bevis T_m (PWV$_2$ in Table 8.4) was found to be more accurate and precise than the NWP-derived PWV (PWV$_1$ in Table 8.4) at all the sites. The largest bias errors are observed at the

MAL2 site with the least being observed at MOIU. Standard deviation errors are however varied with the least errors being observed at MAL2.

The different bias errors in PWV at the sites considered were thought to be related to the sites' altitude as it also relates to the PWV amounts in the neighbouring regions as seen in Figure 8.13 and discussed in a previous paragraph in this section. With respect to the errors at the other sites, the bias error at MAL2, of approximately 4 mm, was considered significant. In addition to it being a low elevation region, the error was also thought to originate from the site's proximity to seashore (Indian Ocean).

Table 8.4: Mean accuracies in PWV at IGS sites in Kenya

Mean accuracies in PWV [mm]				
	PWV_1		**PWV_2**	
IGS site	**bias**	**std**	**bias**	**std**
MAL2	4.0	2.3	-1.6	0.1
MOIU	0.5	2.7	0.0	1.6
RCMN	-0.4	3.7	-0.9	0.7

8.9. Summary

The NWP-derived ZHD and ZWD estimates across Kenya were found to average approximately 2.3 m and 0.2 m respectively. However, altitude variations across the country were found to divide it into roughly, an eastern and western part, which then becomes the main determining factor of the magnitude of troposphere delays across the country with the average magnitudes varying depending on the part (eastern or western) being referred to.

The western portion, characterized by elevations averaging about 1600 m, was found to have ZTD estimates averaging about 2.2 m, while those on the eastern portion, characterized by elevations averaging less than 400 m, were found to be about 2.4 m with the greatest amount of approximately 2.5 m seen along the coastal strip. Generally, however, the NWP-derived ZTD estimates over Kenya were found to have a mean value of approximately 2.3 m.

A study of the NWP-derived ZTD estimates at the IGS sites in Kenya resulted into mean ZTD values of approximately 2.5 m, 1.9 m and 2.0 m at MAL2, MOIU and RCMN respectively. Accuracy assessment of these estimates involved comparing them with IGS ZTD observations.

MAL2 showed a negative mean bias error of approximately 24 mm and a standard deviation error of approximately 10.7 mm, MOIU showed a negative mean bias error of approximately 2.7 mm and a standard deviation error of approximately 14.1 mm, while RCMN showed a mean bias error of approximately 6.7 mm and a standard deviation error of approximately 14.4 mm. The large bias error at MAL2 was attributed to a number of factors including proximity to the Indian Ocean and failure of the NWP model to reproduce the terrain close to the site.

A weighted mean temperature model derived using one month NWP data over Kenya was found to compare quite well with the Bevis T_m model. The Bevis T_m global model was therefore found sufficient for application in Kenya and east Africa at large. However, one month NWP data may not be sufficient to confidently make this conclusion. It is therefore recommended that more data, over a longer duration and covering different seasons, by used in order to make a better comparison.

Additionally, despite the advantage in spatial and temporal resolution, the NWP model data is simulated data and more not accurately reflect the actual observations as may be measured by sensors in radiosonde equipment, for instance. It is therefore recommended to have additional radiosonde sites in Kenya.

The study also found that T_m models for a region covering Kenya are mostly site-specific or regional in nature. While some of the ideas presented here might not be new to the scientific community, the site-specific T_m models developed in this work have not been studied before. It may therefore act as a basis for other related studies in future.

Precipitable water vapor over Kenya has been found to vary with altitude with the low elevation regions having the greatest amounts of water vapor compared to the high elevation regions. GPS-derived PWV computed using the Bevis T_m has been to be more accurate than that computed by numerical integration in an NWP model. However, due to the lack of surface observations in Kenya, the surface temperature required for input into the Bevis T_m model can be sourced from NWP model data.

Information on precipitable water vapor distribution is important for climate studies and this study therefore served to provide an insight into its distribution across Kenya. GPS water vapor observations have become an alternative source for water vapor data, a vital component in climate modeling. The accuracy in numerical weather predictions depends on among other factors the accuracy in the assimilated data that includes GPS water vapor observations.

There is, therefore, need to enhance not only the number of GPS reference sites in the Kenya, but also the number of radiosonde/ground-based microwave radiometers.

9. Conclusion and recommendations

9.1. Conclusion

This section mainly seeks to assess whether the research questions stated in Section 1.5 have been addressed. Based on this assessment we will be able to know whether the research objectives have been achieved.

The primary aim of this research was to explore the accuracy of troposphere delays derived from a numerical weather prediction (NWP) model. The study, therefore, sought to use an optimized method of retrieving data from the 3D grid of an NWP model whose accuracies form the basis upon which the first research question is answered, “What will determine the accuracy of NWP-derived troposphere delays?”

Part of this study was also to analyze precipitable water vapor observations derived using several approaches including GPS, NWP and radiosonde. A vital part of this study included an assessment of the accuracy of the weighted mean temperature parameter results of which were used to answer the second research question: Which weighted mean temperature model is most suitable for the derivation of water vapor observations using GPS and how does GPS-derived precipitable water vapor compare with NWP- and radiosonde-derived precipitable water vapor observations?

This study was also set out to assimilate GPS troposphere delay observations into an NWP model and assess the impact that this would have on short-range NWP forecasts of selected variables. Through an analysis of the accuracy of selected NWP-derived variables, the study made it possible to attempt to answer the third research question: What will determine meaningful impact of the assimilation of GPS total zenith path delays into an NWP model?

An NWP model can be used to estimate troposphere delay at practically every point of a given area. This makes it suitable for application in regions that lack enough ground sites to characterize the delays, such as the African equatorial region. Furthermore, given that GPS-derived water vapor observations have become an important element of GNSS meteorology, it would be good to know the weighted mean temperature model suitable for the said region which is an important parameter in the accuracy of the GPS-derived water vapor observations. In this respect, a study was conducted covering Kenya, the results of which helps us attempt the last question: What could be the expected troposphere delays over Kenya and what is the weighted mean temperature model suitable for the region?

All the findings relating to the various studies conducted in order to attempt to answer the stated research questions are discussed and summarized under studies presented in Chapters 4 to 8. The final conclusion, therefore, focuses on the research questions and is given in the sections that follow.

9.1.1. The accuracy of NWP-derived troposphere delays.

The accuracy of troposphere delays derived from an NWP model is linked to the accuracy of the parameters that are used to compute them, which is in turn linked to the accuracy in the method used to extract them from the NWP model grid. The study presented in Chapter 4 used interpolation methods and approximations using the standard atmosphere model to extract near-surface variables and variables within the 3D grid of an NWP model.

Surface variables of the target parameters needed to compute troposphere delays have been derived from the NWP model and compared to similar observations coming from meteorological sensors and radiosonde data.

When compared with meteorological sensor data at selected EPN sites, the surface pressure variables were found to be negatively biased with a mean value of approximately 0.2 hPa and a mean standard deviation of approximately 0.7 hPa. Surface temperature observations were found to agree with a mean bias and standard deviation of approximately 0.5 °C and 2.0 °C respectively, while near-surface relative humidity accuracy results yielded mean values of approximately 6.8% and 8.8% for the bias and standard deviation errors respectively.

Comparison with radiosonde data at selected radiosonde sites in central Europe, yielded accuracy results that were almost similar to those obtained from comparison with meteorological sensor data at EPN sites. The mean bias errors were, however, positive for the surface pressure observations and significantly small for the near-surface relative humidity observations with a value of approximately 2.5% error.

The accuracy assessment also helped identify some erroneous results coming from some of the ground-based meteorological sensors (like in the case of the EPN site BOR1) and some locations where the NWP model may have been unable to properly model the local terrain probably due to limitations in model resolution configurations.

The study also established that height differences in NWP model orography and meteorological sensor or GNSS antenna heights were likely to cause errors in NWP-derived variables particularly in situations where large differences were involved and the sensor or antenna is below model orography. This, however, was seen to be a potential cause of error in some situations in addition to the NWP model's ability to mimic the local terrain.

The accuracy levels resulting from a comparison of NWP-derived meteorological variables with similar variables sourced from meteorological sensors and radiosonde data were found sufficient for onward application in the computation of troposphere delays using NWP data.

NWP-derived troposphere delays were computed by numerically integrating the hydrostatic and wet refractivities for the hydrostatic and wet delays respectively. A study to assess the accuracy of the resulting delays was presented in Chapter 5. The study established that ignoring the portion of the atmosphere above the NWP model top led into smaller zenith hydrostatic delays by approximately 0.1 m while the zenith wet delays remained unaffected.

A comparison of the NWP-derived zenith hydrostatic delays with similar observations computed using the Saastamoinen model and radiosonde data resulted into larger mean bias and standard deviation errors from the comparison with radiosonde delays with bias and standard deviation errors of approximately 0.01 mm and 1.4 mm respectively for NWP and bias and standard deviation errors of approximately 0.2 mm and 4.0 mm respectively for radiosonde.

The magnitude of the mean bias error obtained when NWP-derived zenith wet delays were compared with similar observations obtained from the difference in GPS total path delays and Saastamoinen hydrostatic delays was found to be significantly large, approximately 8.0 mm, when compared to the 2.0 mm bias error obtained from the comparison with radiosonde-derived wet delays. On the other hand, the precision errors seemed to be of an almost equal magnitude of approximately 18.0 mm on average.

The total zenith path delays obtained from summing up the NWP-derived zenith hydrostatic and wet delays were compared to similar observations sourced from EPN GNSS sites and also computed from radiosonde data. A comparison with the former resulted into a bias error of approximately 8.0 mm in magnitude and a 15.0 mm standard deviation error, while a comparison with the latter resulted into a 2.0 mm bias error and a standard deviation error of approximately 18.0 mm.

9.1.2. The weighted mean temperature and the accuracy of GPS-derived water vapor observations

The important role that GPS water vapor observations continue to play in the mapping of atmospheric water vapor and in meteorology applications cannot be over emphasized. Additionally, knowing that the accuracy of these observations is mainly dependent upon the accuracy of the weighted mean temperature parameter makes us want to model the parameter using any available data.

The study presented in Chapter 6 found the NWP model to be better than the Bevis T_m-T_s in estimating the weighted mean temperature. A comparison of the two methods with radiosonde-derived estimates resulted into a negative systematic error of approximately 0.4 K for the NWP model estimates and a positive systematic error of approximately 0.7 K for the Bevis T_m-T_s relation.

The resulting accuracies in the computed precipitable water vapor estimates showed negative systematic errors of approximately 0.8 mm for the NWP-derived water vapor estimates and approximately 1.0 mm for the water vapor estimates derived using surface observations of GPS ZTD, pressure and temperature.

While it is not a big issue deriving precipitable water vapor using the NWP model (numerical integration of the water vapor density), its accuracy would mainly be determined by the model vertical resolution. On the other hand, the GPS-derived precipitable water vapor makes use of several inputs including the weighted mean temperature, surface pressure and surface temperature, all of which play a critical role in determining the accuracy of the water vapor observations. Consequently, based on the accuracies obtained, the GPS water vapor estimates can be said to have agreed quite well with the NWP estimates. This means that the GPS water vapor estimates can sometimes be used to validate NWP water vapor observations particularly due to the limitation in radiosonde data availability.

9.1.3. The impact of assimilating GPS troposphere delay observations into an NWP model

This was investigated in the study presented in Chapter 7. The impact was generally found to be neutral on short-range forecasts of surface pressure and 2 m temperature with a modest positive impact (a small reduction in forecast errors) on short-range forecasts of relative humidity.

The results obtained are seen to be nearly consistent with those found by, among others, Guerova (2003) in which ZTD data assimilation impacts on 2 m temperature were found to be small and TOUGH Project (2006) in which impacts on near surface relative humidity were found to vary from neutral to positive.

It was established that this activity can possibly result into meaningful impact if a greater number of GNSS sites with their corresponding ZTD observations are used. Additionally, since the accuracy of numerical weather prediction depends a lot on the quality of the input observations, subjecting the ZTD data into some kind of quality assurance process can possibly result into better results.

Nevertheless, the results obtained from this study serve to show that the use of ground-based GNSS ZTD observations can add value to NWP efforts and hence contribute positively to climate monitoring and research.

9.1.4. Troposphere delays and weighted mean temperature over Kenya

The study of Chapter 8 has established that Kenya is roughly divided into an eastern and a western part with respect to terrain variations which has had a direct effect on the troposphere delays experienced over the country. Generally, the eastern part was seen to have the largest amounts of troposphere delays with mean values of approximately 2.4 m compared to 2.2 m in the western part. Similarly, the eastern part was also found to have the largest amounts of precipitable water vapor particularly along the coastal strip which averaged about 40 mm.

A formula defining the relationship between the weighted mean temperature and surface temperature across Kenya, the so-called T_m model, was demonstrated using a month's NWP model data and several years (varies from one radiosonde site to another) of radiosonde data. The NWP T_m model was found to compare quite well with the Bevis T_m model, while radiosonde T_m models were found to be more of site-specific. The frequency of the available radiosonde data also limits the accuracy of the T_m models formulated using radiosonde data.

9.2. Recommendations

NWP models are undoubtedly capable of giving highly accurate tropospheric delay estimates. High resolution regional models would be capable of modeling the local terrain more accurately compared to global models hence proving highly accurate tropospheric delay corrections on a regional scale. These delay corrections will enhance the accuracy and precision of GNSS positioning that is essential for the monitoring of earth deformation processes and hence aid in the realization of geodetic reference systems.

The use of high resolution NWP models can also be used to provide highly accurate weighted mean temperature data that will help in improving the accuracy of GPS-derived water vapor.

Setting up additional GNSS reference sites will help a lot in providing troposphere delay observations that can be assimilated into NWP models and help in improving weather forecasts, particularly with the current climate change that is characterized by severe changes in expected weather patterns.

Though the NWP T_m model for Kenya covers the entire country, it still needs to be validated or probably derived using data covering a longer period. With the dwindling of radiosonde launch sites in the country (probably due to costs, used to be eight, now only two are active), developing a reliable T_m model would go a long way to ensure that GPS-derived precipitable water vapor estimates are accurate which shall in turn help in the monitoring of water vapor.

There is need to put up more GNSS reference sites in Africa. This should be a collective effort by research institutions, weather forecasting authorities and universities. It will help in enhancing the effort of unifying and modernizing the African geodetic reference frame (AFREF).

It is hoped that this study will form a basis on which similar studies will continue to be carried out in future even as additional GNSS reference stations are installed in Kenya and Africa at large.

References

AGARD, 1970. *Tropospheric Radio Wave Propagation.* Dusseldorf,Germany, North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

Aggarwal, S., 2008. Principles of Remote Sensing. In: *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology.* s.l.:s.n., pp. 22 - 38.

Agnew, J., 2002. *Lidar Measurement of Tropospheric Radio Refractivity, Radiocommunications Agency Project 41, Final report,* Harwell Oxford, United Kingdom: Communications Research Unit, CLRC Rutherford Appleton Laboratory.

Alizadeh, M. *et al.*, 2013. *Ionospheric Effects on Microwave Signals - In: Böhm, J., Schuh, H. (Eds.), Atmospheric Effects in Space Geodesy, (Springer Atmospheric Sciences),* Potsdam.

Askne, J. and Nordius, H., 1987. Estimation of tropospheric delay for microwaves from surface weather data. *Radio Science,* 22(3), pp. 379 - 386.

Baek S., *et al*, 2009. Correcting for Surface Pressure Background Bias in Ensemble-Based Analyses. *American Meteorological Society, DOI: 10.1175/2008MWR2787.1,* pp. 2349 - 2364.

Betts, A., 1996. *The parameterization of deep convection: a review,* Reading, UK: ECMWF.

Bevis, M., Businger, S. and Chiswell, S., 1994. GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. *Journal of Applied Meteorology,* 33(3), pp. 379 - 386.

Bevis, M. *et al.*, 1992. GPS Meteorology: Remote Sensing of Atmospheric Water Vapor using the Global Positioning System. *Journal of Geophysical Research,* 97(D14), pp. 15,787 - 15,801.

Boniface, K. *et al.*, 2009. Impact of high-resolution data assimilation of GPS zenith delay on Mediterranean heavy rainfall forecasting. *Annales Geophysicae,* Band 27, pp. 2739 - 2753.

Boone, A., 2013. *Vertical Hybrid-pressure Coordinate.* Available at: http:/ /www.aaron.boone.free.fr/ aspdoc/node7.html. [Accessed on 25 June 2014].

Bosy, J. and Rohm, W., 2007. Verification of the Meteorological Observations on the EPN Stations. Institute of Geodesy and Geoinformatics, Wroclaw University of Environmental and Life Sciences, EUREF 2007 Symposium, London.

Boudouris, G., 1963. On the Index of Refraction of Air, the Absorption and Dispersion of Centimeter Waves by Gases. *Journal of Research of the National Bureau of Standards - D. Radio Propagation,* 67D(6), pp. 631 - 684.

Boumis, M. *et al.*, 2002. *Cost Action 255 Final report - Propagation Effects due to Atmospheric Gases and Clouds,* Noordwijk, Netherlands: European Space Agency, Publications Division.

Calais, E., 2012. GPS signal Propagation. Available at: http:/ /web.ics.purdue.edu/~ecalais/teaching/ geodesy/GPS_signal_propagation.pdf. [Accessed on 03.05.2012].

Carter, J. *et al.*, 2012. *Lidar 101: An Introduction to Lidar Technology, Data, and Applications. Revised.,* Charleston, SC: NOAA Coastal Services Center: National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center.

CASA, 2006. *Global Navigation Satellite Systems Overview,* Canberra: Civil Aviation Safety Authority Australia.

Cassano, J., 2010. *Introduction to Atmospheric Science, Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder.* Available at: http://atoc.colorado.edu/. [Accessed on 13.06.2012].

Cavcar, M., 2013. *Anadolu University, Flight Mechanics and Airplane Performance, The International Standard Atmosphere (ISA).* Available at: http://home.anadolu.edu.tr/~mcavcar/common/ISAweb.pdf. [Accessed on 12.9.2013].

Charney, J. G. and Phillips, N. A., 1953. Numerical Integration of the Quasi-geostrophic Equations for Barotrophic and Simple Baroclinic Flows. *Journal of meteorology,* 10(2), pp. 71-99.

Christian, R. *et al*., 1995. GPS/STORM - GPS Sensing of Atmospheric Water Vapor for Meteorology. *Journal of Atmosheric and Oceanic Technology*, p. 468–478.

Collins, S. N. *et al*., 2013. Grids in Numerical Weather and Climate Models. In: Y. Z. a. P. Ray, Hrsg. *Climate Change and Regional/Local Responses.* s.l.:s.n., pp. 111-128.

Cove, K., 2002. *Improvements in GPS Tropospheric Delay Estimation With Numerical Weather Prediction, MSc Thesis,* Fredericton, New Brunswick, Canada, 98 pp.: Department of Geodesy and Geomatics Engineering Technical Report No. 230, The University of New Brunswick Library.

Cucurull, L. *et al*., 2004. Three-Dimensional Variational Data Assimilation of Ground-Based GPS ZTD and Meteorological during the December 14th Storm Event Over the Western Mediterranean Sea, Proceedings of the 4th EGS Plinius Conference. *Monthly Weather Review,* Band 132, p. 749–763.

Dach, R. and Jean, Y., 2013. *International GNSS Service, Technical Report 2012,* Pasadena: IGS Central Bureau.

Davis, J. *et al*., 1985. Geodesy of Radio Interferometry: Effect of atmospheric modelling errors on estimates of baseline length. *Radio Science,* 20(6), pp. 1593 - 1607.

Derks, H. *et al*., 1997. *GPS Water Vapor Meteorology, Scientific report: WR 97-04,* De Bilt, Netherlands: Royal Netherlands Meteorological Institute (KNMI).

Dudhia, J. *et al*., 2005. *PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and Users' Guide (MM5 Modeling System Version 3).* Available at: http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html. [Accessed on 8.03.2014].

Durran, D., 1998. Runge–Kutta (Multistage) Methods. In: J. Marsden, L. Sirowich and S. Antman, Hrsg. *Numerical Methods for Fluid Dynamics, With Applications to Geophysics.* New York: Springer, pp. 49-78.

Durre, I., Vose, R. S. and Wuertz, D. B., 2006. Overview of the Integrated Global Radiosonde Archive. *Journal of Climate,* 19(1), pp. 53 - 68.

Ehret, U., *et al*., 2012. HESS Opinions “Should we apply bias correction to global and regional climate model data?”. *Hydrology and Earth System Sciences Discussion,* p. 5355–5387.

Elgered, G., Davis, J., Herring, A. and Shapiro, I., 1991. Geodesy of Radio Interferometry: Water Vapor Radiometry for Estimation of the Wet Delay. *Journal of Geophysical Research,* 96(B4), pp. 6541 - 6555.

Elgered, G. and Jarlemark, O., 1998. Ground-based microwave radiometery and long-term observations of atmospheric water vapor. *Radio Science,* 33(3), pp. 707 - 717.

Elgered, G. *et al*., 2005. *Exploitation of ground-based GPS for operational NWP and climate applications, COST Action 716, Final Report,* Luxembourg: Office for Official Publications of the European Communities.

Emardson, R. and Derks, H., 2000. On the relation between the wet delay and the integrated precipitable water vapour in the European atmosphere. *Meteorology Applications,* Band 7, p. 61 68.

EPN, 2013. *EUREF Permanent Network.* Available at: http://www.epncb.oma.be/_networkdata/ stationmaps.php. [Accessed on 16.09.2013].

Erasmaa, R., 2007. *Exploiting ground-based measurements of the global positioning system for numerical weather prediction,* Helsinki: Finnish Meteorological Institute.

Essen, L. and Froome, K., 1951. *The Refractive Indices and Dielectric Constants of Air and its Principal Constituents at 24,000 Mc/s.* Available at: http://iopscience.iop.org/0370-1301/64/10/ 303. [Accessed on 15.06.2012].

Eumetcal, 2014. *Training Resource Library: Eumetcal.* Available at: http:// www.eumetcal.org. [Accessed on 23 June 2014].

Feng, Y., 2003. Combined Galileo and GPS: A Technical Perspective. *Journal of Global Positioning Systems,* 2(1), pp. 67 - 72.

Fox-Rabinovitz, M. S., 1994. Computational Dispersion oProperties of Vertically Staggered Grids for Atmospheric Models. *Monthly Weather Review,* Band 122, pp. 377-392.

Gaffen, D., Barnett, T. and Elliott, W., 1991. Space and Time Scales of Global Tropospheric Moisture. *American Meteorological Society, Journal of Climate,* Band 4, pp. 989 - 1008.

Gauthier, P. *et al.*, 1999. Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmosphere-Ocean,* Band 37 , pp. 103-156.

Ghoddousi-Fard, R., 2009. *Modelling Tropospheric Gradients and Parameters from NWP Models: Effects on GPS Estimates, PhD Thesis,* Fredericton, Canada: Department of Geodesy and Geomatics Engineering, University of New Brunswick.

GNSS Scientific Advisory Commitee, ESA, 2010. *Galileo Scientific Opportunity Document,* Toulouse: European Space Agency.

Gorokhovich, Y. and Voustianiouk, A., 2006. Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics. *Elsevier, Remote Sensing of Environment,* Band 104, pp. 409 - 415.

Guerova, G., 2003. *Application of GPS derived water vapour for Numerical Weather Prediction in Switzerland. PhD Thesis,* Switzerland: University of Bern.

Guerova, G., 2003. *GPS observing system experiment with the Local Model. Report from the COST 716 Short Term Scientific Mission. Research Report No. 2003-16,* Switzerland: Institute of Applied Physics, Department of Microwave Physics, University of Bern.

Guo, Y. *et al.*, 2004. *Assimilation of ground-based GPS PWV with a 3DVAR system for an IHOP case.* Boulder, USA, 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction.

Haan, S., 2008. *Meteorological applications of a surface network of Global Positioning System Receivers, Ph.D. Thesis,* s.l.: s.n.

Haase, J., Ge, M., Vedel, H. and Calais, E., 2003. Accuracy and Variability of GPS Tropospheric Delay Measurements of Water Vapor in the Western Mediterranean. *American Meteorological Society,* Band 42, pp. 1547 - 1568.

Hagemann, S., Bengtsson, L. and Gendt, G., 2003. On the determination of atmospheric water vapor from GPS measurements. *Journal of Geophysical Research,* Band 108.

Hall, A. and Manabe, S., 1999. The Role of Water Vapor Feedback in Unperturbed Climate Variability and Global Warming. *J. Climate,* pp. 2327 - 2346.

Han, Y., Snider, B. and Westwater, E., 1994. Observations of Water Vapor by Ground-based Radiometers and Raman Lidar. *Journal of Geophysical Research,* 99(D9), pp. 18,695 - 18,702.

Hasegawa, S. and Stokeberry, D., 1975. Automatic Digital Microwave Hygrometer. *Rev. Sci. Instruments, doi: 10.1063/1.1134331,* 46(7), pp. 867 - 873.

Hefty, J., 2001. The Permanent Modra-Piesok GPS Station Stability Modra-Piesok and its Long-term and Short-term. *Slovak Journal of Civil Engineering,* IX(1 - 2), pp. 31 - 37.

Hein, G. *et al*., 2007. Envisioning a Future GNSS System of Systems, Part 3, A role for C-band?. *Inside GNSS*, pp. 64 - 73.

Huang, X. and Vedel, H., 2000. *Introduction to Data Assimilation.* Copenhagen, Denmark: Danish Meteorological Institute.

IGRA, 2008. *NOAA Satellite and Information Service, Integrated Global Radiosonde Archive, Data Access - ASCII Files.* Available at: ftp://ftp.ncdc.noaa.gov/pub/data/igra/igra-stations.txt. [Accessed on 5.03.2014].

IGS Central Bureau, 2014. *International GNSS Service, IGS Tracking Network.* Available at: http://www.igs.org/network/netindex.html. [Accessed on 28.02.2014].

Jensen, A., 2002. *Numerical Weather Predictions for GPS Positioning,* Munich, Germany: International Association of Geodesy.

Jensen, A., Tscherning, C. and Madsen, F., 2002. *Integrating Numerical Weather Predictions in GPS Positioning.* ENC GNSS-2002, Copenhagen, s.n.

Jones, J., 2010. *An Assessment of the Quality of GPS Water Vapor Estimates and their Use in Operational Meteorology and Climate Monitoring, Ph.D. Thesis*, Nottingham: Institute of Engineering Surveying and Space Geodesy (IESSG), University of Nottingham.

Karbou, F., Aires, F., Prigent, C. and Eymard, L., 2005. Potential of Advanced Microwave Sounding Unit-A (AMSU-A) and AMSU-B measurements for atmospheric temperature and humidity profiling over land. *Journal of Geophysical Research*, 110(D07109), pp. 1 - 16.

Kasahara, A., 1974. Various Vertical Coordinate Systems Used for Numerical Weather Prediction. *Monthly Weather Review*, 102(7), pp. 509-522.

Klobuchar, J. A., 1987. Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users. *IEEE Transactions on Aerosapce and Electronic Systems,* Vol. AES-23(3), pp. 325 - 331.

Köpken, C., 2001. Validation of Integrated Water Vapor from Numerical Models Using Ground-Based GPS, SSM/I, and Water Vapor Radiometer Measurements. *Journal of Applied Meteorology,* Band 40, p. 1105–1117.

KOS, T. e. a., 2009. Mitigating GNSS Positioning Errors Due to Atmospheric Signal Delays. *Pomorstvo (Journal of Maritime Studies),* 23(2), pp. 495-513.

Kruizinga, G., 1997. *Validation and Applications of Satellite Radar Altimetry,* Austin: University of Texas.

Laing, A. and Evans, J., 2011. *Introduction to Tropical Meteorology, 2nd edition, A Comprehensive Online and Print Textbook, Version 2.0.* Available at: http://www.meted.ucar.edu/tropical/textbook_2nd_edition/index.htm. [Accessed on 2.02.2014].

Langley, R., 1996. Propagation of GPS signals. *SpringerLink, Earth and Environmental Science, GPS for Geodesy Lecture Notes in Earth Sciences,* Volume 60 (DOI: 10.1007/BFb0117680), pp. 103-140.

Lemoine, F. *et al*., 1998. *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96,* Greenbelt, Maryland, USA: National Aeronautics and Space Administration (NASA), Goddard Space Flight Center.

Léna, P. *et al*., 2012. *The Earth Atmosphere and Space, Chapter 2 in Observational Astrophysics.* 3rd Hrsg. s.l.:Springer, http://www.springer.com/978-3-642-21814-9.

Lide, D. R., 2005. "Acceleration Due to Gravity". In: D. R. Lide, Hrsg. *CRC Handbook of Chemistry and Physics, Internet Version, <http://www.hbcpnetbase.com>.* Boca Raton, Florida: CRC Press, p. 2367.

Lorenz, E. N., 1960. Energy and Numerical Weather Prediction. *Tellus XII,* Band 4, pp. 364-373.

Lynch, A. H. and Cassano, J. J., 2006. Composition of the atmosphere. In: *Applied Atmospheric Dynamics.* Chichester, England: John Wiley and Sons, Ltd, pp. 43 - 45.

Macpherson, S., Deblonde, G., Aparicio, J. and Casati, B., 2008. Impact of NOAA Ground-Based GPS Observations on the Canadian Regional Analysis and Forecast System. *Monthly Weather Review,* Band 136, pp. 2727 - 2746.

Mendes, V. B., 1999. *Modeling the neutral-atmospheric propagation delay in radiometric space techniques, Ph.D. dissertation, Technical Report No. 199,* Fredericton: Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Mesinger, F. and Arakawa, A., 1976. *Numerical Methods Used in Atmospheric Models,* Geneva, Switzerland: Global Atmospheric Research Program, WMO-ICSU Joint Organizing Committee.

MetEd, 2014. *Impact of Model Structure and Dynamics, NWP Equations, How Mesoscale Models Work.* Available at: https://www.meted.ucar.edu/nwp. [Accessed on 5.03.2014].

Misra, P. and Enge, P., 2006. GPS Measurements and Error Sources. In: *Global Positioning System, Signals, Measurement and Performance, Second Edition.* Massachusetts: Ganga-Jamuna Press, p. 148.

Morton, Y. and Graas, F., 2009. *High Precision GPS Measurements,* Oxford: Miami University.

Musa, T., Lim, S. and Rizos, C., 2005. *Low Latitude Troposphere: A Preliminary Study Using GPS CORS Data in South East Asia.* San Diego, California, USA, The Institute of Navigation, pp. 685 - 693 .

Muvvala, P., 2011. *Feasibility of Troposphere Propagation Delay Modeling of GPS Signals using Three-Dimensional Weather Radar Reflectivity Returns, MSc Thesis,* Ohio: Russ College of Engineering and Technology of Ohio University.

NASA, 2014. *NASA Earth Observatory, Global Maps, Water vapor.* Available at: http://earthobservatory.nasa.gov/GlobalMaps/view.php?d1=MYDAL2_M_SKY_WV. [Accessed on 15.02.2014].

National Climatic Data Center, 2008. *NOAA Satellite and Information Service.* Available at: http://www.ncdc.noaa.gov/oa/climate/igra/. [Accessed on 28.02.2014].

NCAR, 2014. *WRF ARW, Version 3 Modeling System User's Guide,* Boulder, Colorado: Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.

Neely, R. and Thayer, J., 2011. Raman Lidar Profiling of Tropospheric Water Vapor over Kangerlussuaq, Greenland. *Journal of Atmospheric Oceanic Technology,* Band 28, p. 1141–1148.

Nesbitt, S., 2008. *ATMS410 Radar Meteorology.* Available at: http://www.atmos.illinois.edu/~snesbitt/ATMS410/files/attenuation.ppt. [Accessed on 12.03.2012].

Nichols, N., 2003. Data Assimilation: Aims and Basic Concepts. *Data Assimilation for the Earth System, NATO Science Series,* Band 26, pp. 9 - 20.

OFCM, 1997. *Federal Meteorological Handbook No. 3, Rawinsonde and Pibal Observations,* Washington, DC: Office of the Federal Coordinator for Meteorology, US Department of Commerce/ National Oceanic and Atmospheric Administration.

Owens, J., 1967. Optical refractive index of air: Dependence on Pressure, Temperature and Composition. *Journal of Applied Optics,* 6(1), pp. 51 - 59.

Parrish, D. F. and Derbe, J. C., 1992. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review,* Band 120, p. 1747–1763.

Plana-Fattori, A. *et al*., 1998. Estimating the Atmospheric Water Vapor Content from Sun Photometer Measurements. *Journal of Applied Meteorology,* Band 37, p. 790–804.

Popovic, J., 2006. *Eta Model in Weather Forecast: MSc Thesis,* Stockholm: Royal Institute of Technology.

Pósfay, A., Krueger, E., Schueler, T. and Hein, G., 2003. *Tropospheric Delay Modeling for the European Space Agency's Galileo Testbed: Methods of Improvement and First Results.* Anaheim, CA, USA, The Institute of Navigation.

Rama Varma Raja, M., Gutman, S., Yoe, J. and McMillin, L., 2006. *Comparison of Column Integrated Water Vapor Measurements from Atmospheric InfraRed Sounder (AIRS) and Surface-based Global Positioning System Receivers.* s.l., 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS).

Randall, D. A., 2004. *Department of Atmospheric Science, Colorado State University, An Introduction to Atmospheric Modeling.* Available at: http://kiwi.atmos.colostate.edu/group/dave/at604pdf/Chapter_1.pdf. [Accessed on 15.02.2014].

Rocken, C., Sokolovskiy, S., Johnson, J. and Hunt, D., 2001. Improved Mapping of Tropospheric Delays. *Journal of Atmospheric and Oceanic Technolology,* Band 18, p. 1205–1213.

Ross, R. and Rosenfeld, S., 1997. Estimating mean weighted temperature of the atmosphere for Global Positioning System applications. *Journal of geophysical research,* 102(D18), pp. 21,719-21,730.

Rüeger, J., 2002. *Refractive Indices of Light, Infrared and Radio Waves in the Atmosphere,* Sydney, Australia: School of Surveying and Spatial Information Systems, University of New South Wales.

Saastamoinen, J., 1972. Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. *Geophysical Monogram Series, American Geophysical Union,* Band 15, pp. 247-251.

Sánchez-Arriola, J., Navascués, B. and García-Moya, J., kein Datum *INM Assimilation Results, Deliverable 46 and 32 for the TOUGH project,* s.l.: Instituto Nacional de Meteorología, Spain.

Schrier, G. e. a., 2013. *Climate Indicator Bulletins (CIBs)/Central European flooding 2013.* Available at: http://cib.knmi.nl/mediawiki/index.php/Central_European_flooding_2013. [Accessed on 16.09.2013].

Schueler, T., 2001. *On Ground-Based GPS Tropospheric Delay Estimation, PhD Thesis,* München: Universität der Bundeswehr.

Schueler, T., Pósfay, A., Hein, G. and Biberger, R., 2002. *A Global Analysis of the Mean Atmospheric Temperature for GPS Water Vapor Estimation,* Munich: Institute of Geodesy and Navigation, University FAF Munich.

Schueler, T., 2014. The TropGrid2 standard tropospheric correction model. *GPS Solutions,* 18(1), pp. 123-131.

Schwitalla, T., 2012. *Model evaluation and data assimilation impact studies in the framework of COPS, Ph. D. Thesis,* Hohenheim, Germany: Institute of Physics und Meteorology, Faculty of Natural Sciences, University of Hohenheim.

Sibylle, V. *et al*., 2010. 2010: Validation of Precipitable Water Vapor within the NCEP/DOE Reanalysis Using Global GPS Observations from One Decade. *Journal of Climate Research,* Band 23, pp. 1675 - 1695.

Singh, D., J., G. and Kashyap, D., 2013. Weighted mean temperature model for extra tropical region of India. *Journal of Atmospheric and Solar-Terrestrial Physics,* Band 107, pp. 48-53.

Skamarock, W. *et al*., 2008. *A Description of the Advanced Research WRF Version 3, NCAR TECHNICAL NOTE: NCAR/TN–475+STR,* Boulder, Colorado, USA: Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research (NCAR).

Smith, E. and Weintraub, S., 1953. The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies. *Proceedings of the Institute of Radio Engineers,* 41(8), pp. 1035 - 1037.

Smith, R. J., 2010. *Minimising Time-Stepping Errors in Numerical Models of the Atmosphere and Oceans: MSc. Thesis,* Reading: University of Reading, School of Mathematics, Meteorology and Physics.

Solheim, F., 1993. *Use of pointed water vapor radiometer observations to improve vertical GPS surveying accuracy, PhD Thesis,* Colorado, USA: University of Colorado.

Solheim, F. and Vivekanandan, J., 1999. Propagation delay induced in GPS signals by dry air, water vapor, hydrometeors, and other particulates. *Journal of Geophysical Research,* 104(D8), pp. 9663 - 9670.

SOO/STRC, 2013. *The National SOO Science and Training Resource Center, Introduction to the EMS.* Available at: http://strc.comet.ucar.edu/software/newrems/. [Accessed on 5.03..2014].

Thayer, G., 1961. A Formular for Radio Ray Refraction in an Exponential Atmosphere. *Journal of Research of the National Bureau of Standards - D. Radio Propagation,* 65(2), pp. 181 - 182.

Thayer, G., 1974. An improved equation for the radio refractive index of air. *Radio Science,* 9(10), p. 803–807.

Thessin, R. N., 2005. *Atmospheric Signal Delay Affecting GPS Measurements Made by Space Vehicles during Launch, Orbit and Reentry, MS Thesis,* Massachusetts: Massachusetts Institute of Technology.

Thompson, A. M., 2014. *National Aeronautics and Space Administration/Goddard Space Flight Center, SHADOZ: Southern Hemisphere ADditional OZonesondes.* Available at: http://croc.gsfc.nasa.gov/shadoz/. [Accessed on 5.03.2014].

Thompson, A. *et al*., 2007. Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles. *Journal of Geophysical Research,* 112 (D3), pp. 1 - 18.

Throner, R. H., 2001. *Engineering Geology Field Manual, Second Edition, Volume 3,* Denver: U.S. Department of the Interior, Bureau of Reclamation.

TOUGH Project, 2006. *Targeting Optimal Use of GPS Humidity Measurements in Meteorology, 6th Periodic Report.* Available at: http://web.dmi.dk/pub/tough/deliverables/d14-final-rep.pdf. [Accessed on 4.03.2014].

Tranquilla, J. and Al-Rizzo, H., 1994. Range Errors in Global Positioning System During Ice Cloud and Snowfall Periods. *IEEE Transactions on Antennas and Propagation,* 42(2), pp. 157 - 165.

Troller, M., 2004. *Dissertation: GPS based PhD Dissertation: Determination of the Integrated and Spatially Distributed Water Vapor in the Troposphere,* Zurich: Swiss Federal Institute of Technology Zurich.

UK Met Office, 2011. *National Meteorological Library and Archive Fact sheet No. 3 – Water in the Atmosphere.* Available at: http://www.metoffice.gov.uk/media/pdf/9/g/No._03-Water_in_the_Atmosphere.pdf. [Accessed on 15.06.2012].

uk.sci.weather resources, 2014. *NWP Basics.* Available at: http://weatherfaqs.org.uk/node/192. [Accessed on 15.02.2014].

UNAVCO, 2013. *Community Science, Geoid Height Calculator.* Available at: http://www.unavco.org/community_science/science-support/geoid/. [Accessed on 6.03.2014].

University of Wyoming, 2014. *College of Engineering, Department of Atmospheric Science, Weather.* Available at: http://weather.uwyo.edu/upperair/sounding.html. [Accessed on 4.03.2014].

Varner, C., 2000. *DGPS Carrier Phase Networks and Partial Derivative Algorithms, PhD Thesis, UCGE Reports No. 20129,* Calgary, Alberta: Department of Geomatics Engineering, University of Calgary.

Wang, J., Zhang, L. and Dai, A., 2005. Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications. *Journal of geophysical research,* 110(D21101).

Ware, R. H. *et al.*, 2000. Suominet: A real-time national GPS network for atmospheric research and education. *Bulletin of the American Meteorological Society,* Band 81, pp. 677 - 694.

Yan, X. *et al.*, 2008. Mesoscale GPS Zenith Delay assimilation during a Mediterranean heavy precipitation event. *Advanced Geosciences,* Volume 17, p. 71 – 77.

Yuan, L. *et al.*, 1993. Sensing Climate Change Using the Global Positioning System. *Journal of Geophysical Research,* 98(D8), pp. 14,925 - 14,937.

Yunck, T. and Hajj, G., 2003. *Global Navigation Satellite Sounding of the Atmosphere and GNSS Altimetry: Prospects for Geosciences,* Pasadena, California: Jet Propulsion Laboratory, California Institute of Technology.

Zhenhong, L., 2005. *Correction of Atmospheric Water Vapour Effects on Repeat-Pass SAR Interferometry Using GPS, MODIS and MERIS Data,* London: Department of Geomatic Engineering, University of London.

Appendices

Appendix A: Maps of common GNSS reference networks and stations

Figure A.1: IGS world stations map.

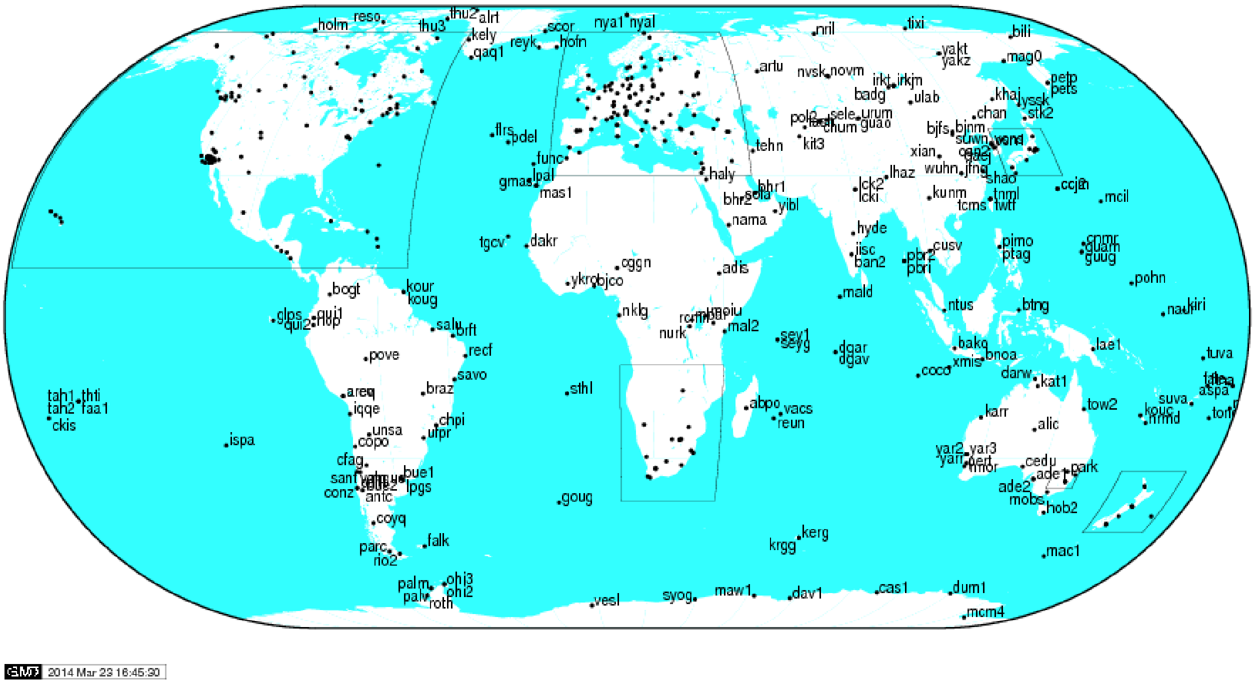
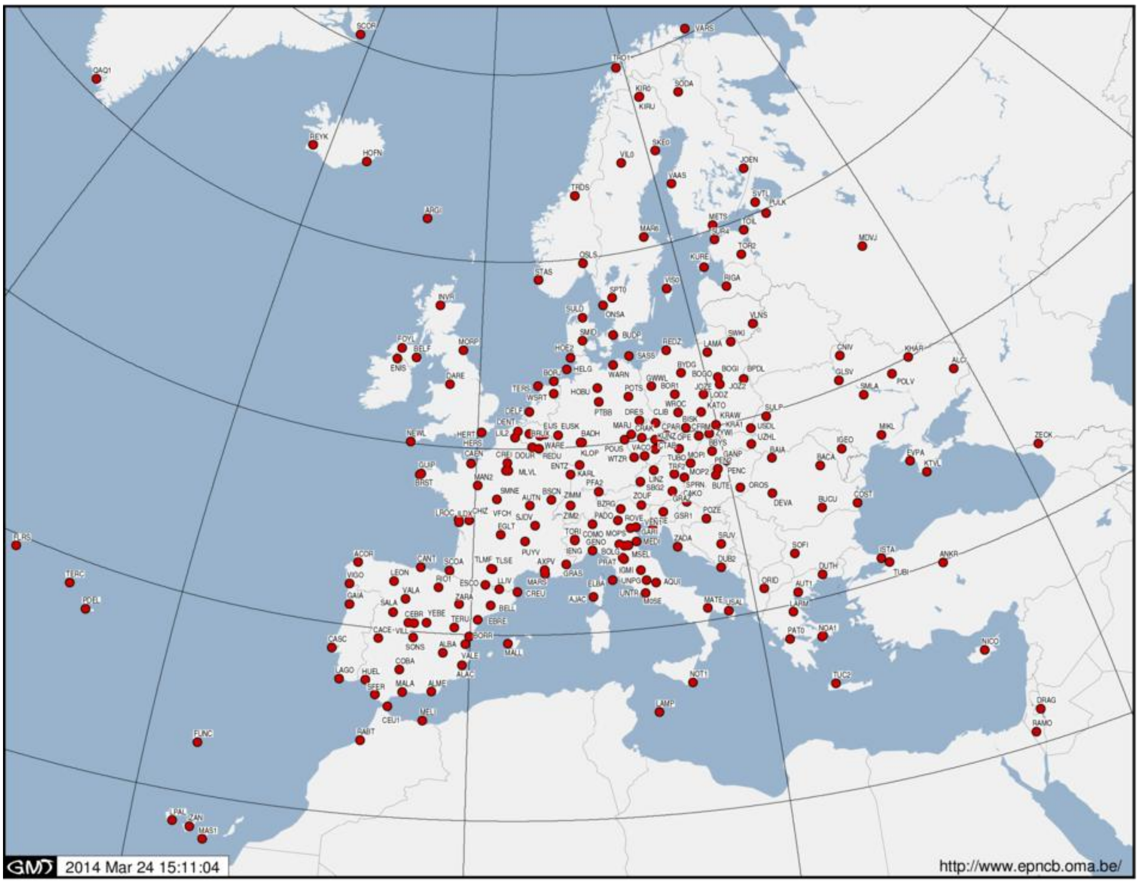


Figure A.2: EPN stations map.



Appendix B: The EMS and WRFDA

This appendix presents a brief discussion of how the EMS and WRFDA systems were set up and implemented. Most of the discussion is based on detailed overviews of these systems that can be found in e.g. Skamarock *et al.*(2008) and SOO/STRC (2013).

The EMS is essentially the WRF system in a single package that was developed in a way that is easy to install, configure and execute in order to encourage the use of WRF for private, operational and research purposes(SOO/STRC, 2013). The system incorporates two NWP models, the Advanced Research Weather Research and Forecasting (ARW) model and the Non-hydrostatic Mesoscale Model (NMM).

From a developer's point of view a number of differences could exist between the two models most of which may not be important from a user's point of view. Most importantly, both are non-hydrostatic models and NMM is mainly used for forecasts while ARW is mainly used in both forecasts and atmospheric research.

1. The EMS: set up and implementation

The EMS was acquired and installed into a 64-bit 4 core Intel processor computer operating at 3.40 GHz frequency and with a Linux operating system (openSUSE). This was done with the intention of running it solely for research (as opposed to real-time forecasting).

For the installation, a Perl script, ems_install.pl, manages the installation of the NWP model by downloading and installing all the required installation files from servers of the model developers. This means that an internet connection is required. However, it is also possible to install from DVD or a local network (SOO/STRC, 2013).

Making a successful EMS model run that would result into a weather forecast involves three required and one optional procedure. The compulsory steps include creating a domain (or geographical area to be modeled), initializing a simulation and running the actual simulation. The model output files are usually in a form to be specified by the user (by editing some configuration files prior to running a simulation). The optional step is therefore meant to prepare the model output into a form that can be visualized, if need be. All the procedures are run by executing preloaded Perl scripts from the command line. These steps are summarized as follows:

i. Domain creation

The geographical area to be modeled is specified using either a graphical user interface (domain wizard) or some line commands to run the domain creation routine, ems_domain.pl. Some of the processes involved in creating a domain include:

a. Selection of a map projection that is suitable for the selected geographical area.

The selected map projection type determines the WRF model to be implemented among the two core models, NMM and ARW. The available map projections and the corresponding model to be used are summarized in the table below:

Table B.1: EMS map projection and NWP model options.

Map projection type	**WRF model**
Rotated Latitude-Longitude (Rot Lat-Lon)	NMM
Lambert Conformal (preferred for mid-latitudes)	ARW
Polar Stereographic (preferred for high-latitudes)	ARW
Mercator (preferred for low-latitudes)	ARW
Regional Lat-Lon	ARW
Global Lat-Lon (preferred for low-latitudes)	ARW

b. Specification of the horizontal grid dimensions and spacing.

The horizontal dimensions in the x and y directions are specified which is equivalent to the number of grid points expected in the x- and y-axis. The grid spacing is equivalent to the distance between grid points in km. Consequently, the end-to-end distance in either the x- or y- horizontal dimension is the product of the horizontal component (number of grid points in x- or y-dimension) and the grid spacing.

c. Localization of a domain.

This step serves to extract terrestrial data over the area covered by the defined domain at a specified resolution (i.e. the scale of the features you can simulate with the model). This then allows for the visualization of the terrestrial data.

ii. Model configuration

After successfully creating a domain, the system creates a folder for the simulation that is about to be run. This folder contains symbolic links to all the necessary Perl routines required to run a simulation, copies of default configuration files, folders into which output files shall be placed, folder into which the log files shall be placed, a folder into which model initialization files shall be placed, and a folder containing the namelist files[1].

The copies of the configurations and namelist files can be edited to fit a simulation to a desired output. Among the configurations to be made includes the number of model vertical levels, the format of the

---

[1] A namelist file is an ASCII file that allows a program to transfer a group of variables by referencing the name of the group which they belong.

model output files (NetCDF, GRIB[1]1, GRIB2), parameterization schemes, and frequency of output files.

iii. Model initialization

Initialization of a model run mainly involves setting the initial and boundary conditions that specify the state of the atmosphere at the start of the integration process and the meteorological variables at the upper, lower and lateral boundaries of the specified domain.

The initial and boundary conditions are obtained from a data set that is to be specified. The available data sets that are available for use with the EMS are specifically tailored for operational modeling.

This procedure therefore identifies, acquires and processes the specified data set(s) through a method (e.g. http[2], ftp[3], and nfs[4]) and a source (e.g. NCEP, STRC) that is to be specified.

iv. Model simulation

During model simulation, model configurations are ascertained to be valid by, among other things, confirming the WRF core (NMM or ARW) to be used and the set initial and boundary conditions. The procedure also creates initial and boundary condition files for the user domain.

v. Post-simulation

The resulting output files from the model run contain data which is only useful if it can be read and visualized. This is achieved using software that is able to understand the format of the output files, read and give graphical visualizations as an output.

The default format of the model output files is NetCDF and can be changed by editing some configuration files before running a simulation. Otherwise, a post-simulation routine can be run to process the output files into other formats as desired.

Several tools are available for reading/visualizing the post-processed files including: NetCDF4excel, NCL (**N**CAR **C**ommand **L**anguage), and GrADS. NetCDF4excel can only read the NetCDF file and present the output in the Microsoft Excel spreadsheets, while the NCL and GrADS can both read the NetCDF file and make graphical visualizations if need be.

2. The WRFDA: set up and implementation

The WRFDA system is a collection of a number of programs including the observations pre-processor program (obsproc), the WRF variational data assimilation program (wrfvar) and the boundary conditions updater program.

---

[1] GRIdded Binary.

[2] HyperText Transfer Protocol.

[3] File Transfer Protocol.

[4] Network File System.

The obsproc program takes the observations file as input and processes them, by among other things, removing observations outside the specified temporal and spatial domains, re-ordering and merging duplicate (in time and location) data reports, assigning observation errors based on a pre-specified error file (obserr.txt), and writing out an observations file (in ASCII[1] or BUFR[2] format) to be an input for the wrfvar program.

The wrfvar program ingests the observations processed by the obsproc program into the initial conditions files created during the NWP model run.

The boundary conditions updater program updates the boundary conditions file generated during the NWP model run so that it is consistent with the modified initial conditions file.

The WRFDA[3] system was installed in the same computer as the EMS. Its implementation involved three main steps, namely:

i. Preparing the observations for input to the DA system.

Preparation includes converting the observations files into the Little_r[4] format and running the Little_r format observations file through the obsproc program.

For a 3D variational (3D-Var) DA system, only observations closest to the center of the assimilation window are chosen. The assimilation window is a time interval (typically 6 to 12 hours) that encompasses the observations to be assimilated.

The obsproc program also requires a namelist file (namelist.obsproc) as input.

ii. Assimilating the observations into the NWP initial fields.

This is the core of the DA system. It ingests observations into the NWP initial fields which will then be used by the NWP model to run the next forecast.

iii. Updating the boundary conditions to be consistent with the new initial fields file.

Since the original NWP initial conditions file has been modified, there is need to update the original boundary conditions file in order for it to be consistent with the new initial conditions file.

3. Testing the EMS/WRFDA systems on the hardware

A seemingly successful installation of the EMS and WRFDA systems was followed by an exercise that involved testing them on the computer system in which they were residing. A series of tests were done

---

[1] American Standard Code for Information Interchange.

[2] Binary Universal Form for the Representation of meteorological data.

[3] The version used is WRF 3D-Var version 3.2.1.

[4] Little_r format is an ASCII standard text file originally developed for the MM5 weather model. The format of the file is described in (Dudhia *et al.*, 2005). It is required for input into the observations pre-processor (obsproc) program that processes data for input to the data assimilation program.

all the while developing a strategy to be used in making all the other simulations including the main ones whose results were presented and discussed in this thesis.

The general strategy observed is as follows:

a. Prepare the observations

i. Study possible data sources and their corresponding coverage (spread of reference stations).

- Target data: GPS ZTD.
- Potential data sources and coverage.

  E.g. EUREF (Europe), IGS (global), SuomiNet (USA, global)

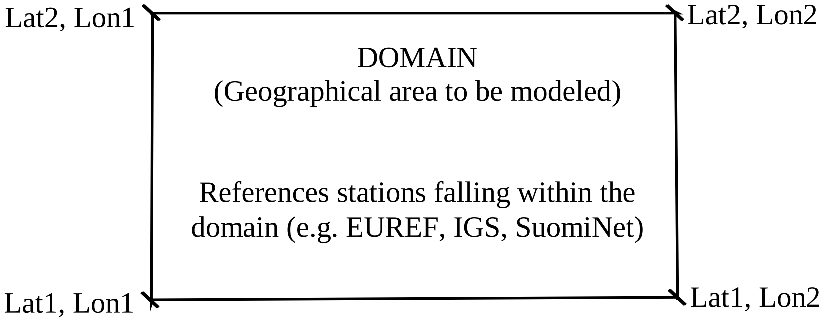
  Note: Some of the stations may be inactive.

ii. Based on the findings of the preceding step, decide on the geographical domain that will be covered.

- Define the domain more precisely by using the “Domain Wizard” of the EMS.
- A definition of the horizontal dimensions and grid spacing helps in defining the domain more precisely. The horizontal dimensions specify the number of grid points from one end of the domain to the other in both the x (longitude) and y (latitude) directions while the grid spacing (distance between grid points) defines the resolution. The end-to-end horizontal distance in x and y direction can therefore be determined by multiplying together a horizontal dimension (in x or y) by the grid spacing.
- Experience has shown that the NWP model may run without segmentation fault errors by having large grid spacing and smaller number of grid points in both the x and y dimensions. This differs from one domain to another.
- Specifying the coordinates at the edges of the defined domain may help in providing an early detection of reference stations falling outside the defined domain.

iii. Based on the selected data source and the defined domain, prepare a list of all the reference stations falling within the domain (Figure B.1).

Figure B.1: Domain specification.



iv. Decide upon the period (date) for data assimilation and check for data availability (for the period decided upon) from the data source selected in Step a. i.

This date is required for NWP model initialization. Furthermore, GPS ZTD and PWV retrievals are not available in real time/current day (date). This presents the following implications:

- Need to find out the most recent date (backdate from current) for data availability from the selected data source. For instance, GPS ZTD retrievals from the EUREF network are available two GPS weeks back from the current GPS week.
- The implication is that the NWP model can as well be initialized using a historical data set. Therefore, there is need to study the available historical data sets for NWP model initialization and the periods of availability.

For instance, among the historical data sets available for model initialization in the WRF EMS model is the ECMWF historical data set and it is currently available for the period 1989 to 2008.

Relying on the EMS historical data sets may limit our choice to only use the ECMWF data set. This may as well make it the main determining factor in deciding the date (GPS week for GPS ZTD retrieval must lie within the period 1989 to 2008). However, with proper planning and timing, an archive of historical data sets can be created from a real-time archive (such as GFS). This may be archived by running NWP simulations using a real-time archive for initialization (up to the pre-simulation stage is sufficient). The downloaded model initialization files can then be used on a later date for model initialization.

b. First NWP model run.

i. Finish the domain started in Step a. ii. Localize the domain (extracting terrestrial data over the area covered by the defined domain at the specified resolution, the model’s coordinate system is terrain-following). At the end of this process a domain directory is created.

ii. Run the NWP model.

- Execute the routines ems_prep.pl and ems_run.pl.

Note: The use of historical data sets has already been foreseen. Therefore, the routine shall be run with, among other options, one that specifies a historical data set e.g. ECMWF, the date to coincide with that of the GPS data file.

- If necessary, modify the configuration files to enable all the desired outputs.
- If necessary, run the EMS post-simulation routine ems_post.pl.
- Create a backup directory (the initial and boundary condition files, the NetCDF output files).

c. Assimilate the GPS ZTD data.

The main purpose for this activity is to ingest GPS ZTD retrievals into the initial conditions file generated during the first model run and to update the boundary conditions file.

i. Download the data files from the appropriate source based upon the findings of Step i., ii. iv. concerning the data type (GPS ZTD or PWV), the data source (EUREF, IGS, and SuomiNet), the domain specified, and data availability.
ii. Convert the downloaded data file to Little_r format using a decoder program that takes in the data file and the stations list file (as described in Step a. iii.).
iii. Process the Little_r format data file with the obsproc program.
iv. Ingest the GPS data into the initial analysis file with the wrfvar program.
v. The main output of this program is a new data file (wrfvar_output) which will now become the initial analysis file for the second NWP model run. This file is basically the initial analysis file generated during the first model run but now with the GPS data ingested into it.
vi. Update the boundary conditions file.
vii. Since the original initial conditions file has been modified, there is need to update the original boundary conditions file in order for it to be consistent with the new initial conditions file.

d. Second NWP model run.

The purpose for a second NWP model run is to make a simulation that includes the ingested GPS data. The output of this model run will then be compared to the output of the first model run. Consequently, the domain definition and the pre-simulation process are as defined in Step b.

This model run will therefore begin from the simulation stage, but this time using the new initial and boundary conditions files generated in Step c.

Backup the results of the second model run.

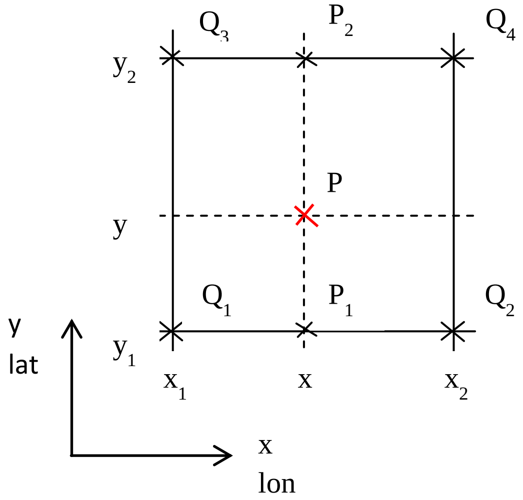
e. Assess/interpret the results before DA (as backed up in Step b) and after DA (as backed up in Step d).

This activity can be carried out by a combination of many different tools including NetCDF4Excel, NCL, GrADS, Matlab, etc.

Appendix C: Bi-linear interpolation to estimate variables at a point of interest

In Figure 3.1, longitude and latitude variable combinations of the point of interest P is (x,y). The four nearest neighboring grid points surrounding P are represented by Q_1, Q_2, Q_3, and Q_4 and defined by the longitude, latitude combinations (x1,y1), (x2,y1), (x1,y2) and (x2,y2) respectively. This situation is replicated in Figure C.1 which is also used to illustrate the bilinear interpolation process to estimate the variables at P.

Figure C.1: Illustration of the bilinear interpolation process.



The target variables at the four nearest neighboring grid points, Q_1, Q_2, Q_3, and $Q4$, are known. What need to be determined are the variables at P.

By first interpolating in the horizontal, the functions at P_1 and P_2 (i.e. $f(P_1)$ and $f(P_2)$ respectively) can be determined. Then by interpolating in the vertical, $f(P_1)$ and $f(P_2)$ can be used to solve for variables at P. Thus,

$$V_P = \frac{(y - y_2)}{(x_2 - x_1)(y_1 - y_2)} \{(x_2 - x)[Q_1 + Q_2(y_1 - y)] + (x - x_1)[Q_3 + Q_4(y_2 - y)]\}$$

Where, V_P represents a variable at P

Appendix D: Accuracy computations

The accuracy of a given observed quantity made up of N samples was made by comparison with reference observations in terms of bias, standard deviation and root mean square by using the following definitions: if x_{o_i} is the i^{th} observation value with the corresponding reference value x_{r_i}, then,

$$\text{bias} = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(x_{r_i} - x_{o_i} \right)$$

$$\text{standard deviation} = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^{N} \left(\left(x_{r_i} - x_{o_i} \right) - \text{bias} \right)^2}$$

$$\text{root mean square} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(x_{r_i} - x_{o_i} \right)^2}$$

Appendix E: EPN stations list.

Table E.1: EPN stations list.

Site ID.	City	Country	Latitude [degrees]	Longitude [degrees]	Ellipsoidal height [m]
BBYS	Banska Bystrica	Slovak Republic	48.75	19.15	487.4
BISK	Zlate Hory	Czech Republic	50.26	17.43	950.9
BOGI*	Borowa Gora	Poland	52.47	21.04	139.9
BOR1*	Borowiec	Poland	52.1	17.07	124
BPDL*	Biala Podlaska	Poland	52.04	23.13	196.8
BUTE	Budapest	Hungary	47.48	19.06	180.9
DRES*	Dresden	Germany	51.03	13.73	202.9
GOPE*	Ondrejov	Czech Republic	49.91	14.79	592.6
GWWL*	Gorzow Wielkopolski	Poland	52.74	15.21	127.6
JOZE*	Jozefoslaw	Poland	52.1	21.03	141.4
KATO*	Katowice	Poland	50.25	19.04	332.9
KRAW	Krakow	Poland	50.07	19.92	267.3
KUNZ*	Kunzak	Czech Republic	49.11	15.2	703.1
LINZ	Linz	Austria	48.31	14.28	337
LODZ	Lodz	Poland	51.78	19.46	259.9
MARJ	Jachymov	Czech Republic	50.36	12.89	904.7
MOP2	Modra-Piesok	Slovak Republic	48.37	17.27	581.1
MOPI*	Modra-Piesok	Slovak Republic	48.37	17.27	579
PENC	Penc	Hungary	47.79	19.28	291.7
POTS*	Potsdam	Germany	52.38	13.07	174
POUS	Poustka	Czech Republic	50.14	12.3	572.2
PTBB*	Braunschweig	Germany	52.3	10.46	130.2
SPRN	Sopron	Hungary	47.68	16.58	279
TUBO*	Brno	Czech Republic	49.21	16.59	324.4
VACO	Vacov	Czech Republic	49.13	13.72	799.4
WROC	Wroclaw	Poland	51.11	17.06	181
WTZR*	Bad Koetzting	Germany	49.14	12.88	666
ZIMM	Zimmerwald	Switzerland	46.88	7.47	956.7
ZYWI*	Zywiec	Poland	49.69	19.21	412.8

Note: Stations in asterisk were used for both meteorological data and GPS ZTD observations.

Appendix F: Error analysis results for surface meteorological parameters at selected EPN sites in January 2014

The NWP model simulation for the one month period was run on model configurations similar to those used and discussed in Section 3.2. Only that the period here was longer.

Table F.1: Mean accuracies in surface pressure, temperature and relative humidity at selected epn sites in January 2014.

		pressure [hPa]		**temperature [°C]**		**relative humidity [%]**	
Station id.	**Samples**	**bias**	**stdev**	**bias**	**stdev**	**bias**	**stdev**
BPDL	720	-0.1	0.9	1.2	2.0	-1.2	6.9
BYDG	720	**-7.7**	**9.5**	**9.7**	**10.0**	1.2	10.2
DRES	720	-0.4	1.2	0.8	1.9	5.3	11.4
GOPE	720	-0.3	0.8	0.1	1.4	3.0	7.6
GWWL	720	0.5	1.0	2.2	3.8	4.2	7.5
JOZE	720	0.8	1.0	0.2	1.3	7.6	9.3
KATO	720	-0.4	1.0	**5.2**	**5.6**	-6.6	11.1
KUNZ	720	1.9	2.1	0.2	1.5	-3.5	8.3
MOP2	720	1.6	2.1	-0.8	1.5	10.4	12.3
MOPI	720	1.6	2.1	-0.8	1.5	10.4	12.3
POTS	720	0.4	0.8	-0.4	1.7	1.1	7.1
TUBO	720	1.2	1.3	2.7	2.9	8.5	10.6
WROC	720	0.1	0.9	1.7	2.2	-0.9	7.9
WTZR	720	-0.1	0.7	-0.2	2.2	8.0	12.9
ZIMM	720	0.5	0.9	-2.4	2.8	-7.7	16.0
mean		**-0.03**	**1.75**	**1.29**	**2.82**	**2.65**	**10.09**

Figure F.1: Trends in sensor and NWP surface pressure observations at BYDG in January 2014.

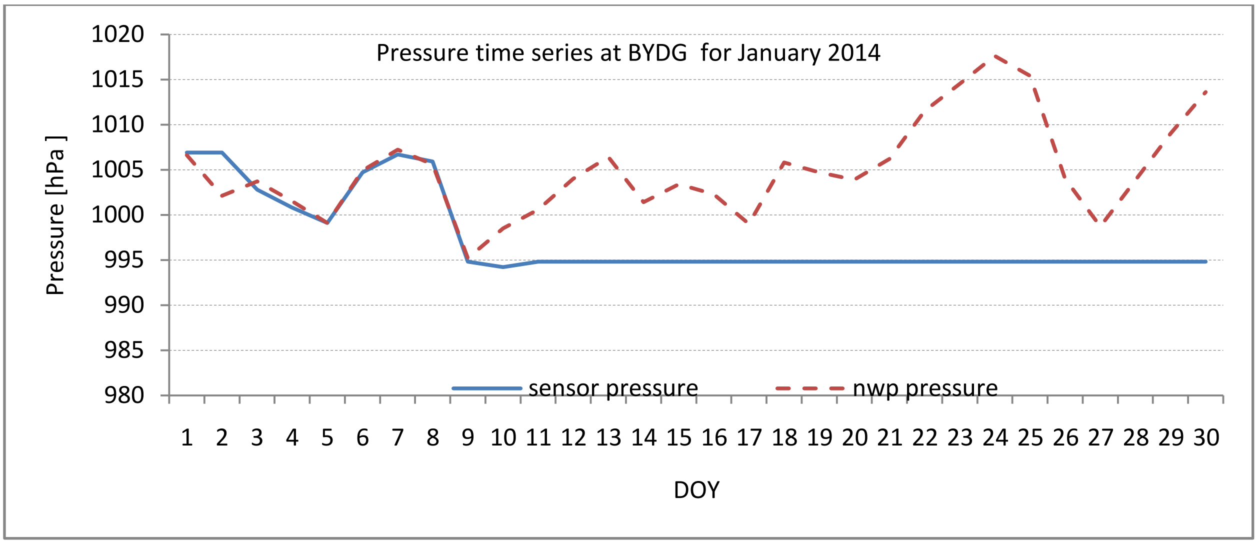
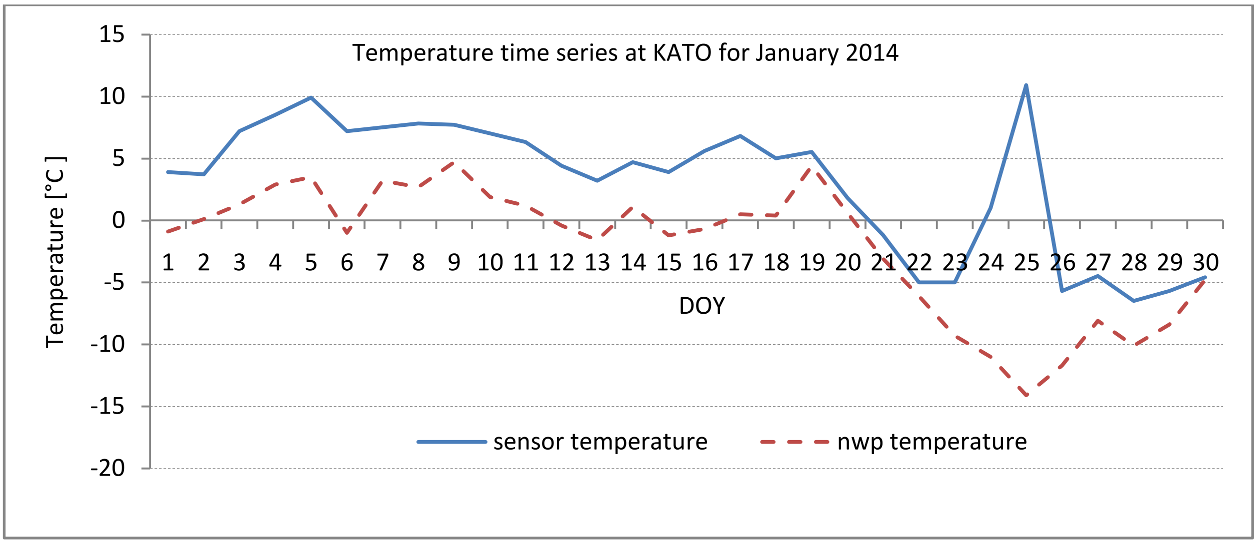


Figure F.2: Trends in sensor and NWP surface temperature observations at KATO in January 2014.



Appendix G: Zenith path total delays at IGS sites in Kenya.

The daily mean zenith path total delay estimates and their mean accuracies and precision for the three IGS sites currently operating in Kenya in the month of January 2014 are shown in Table G.1. RCMN has some missing entries due to missing data in the SINEX troposphere delay files.

Table G.1: Mean accuracies in ZTD for MAL2, MOIU and RCMN in January 2014.

	MAL2			**MOIU**			**RCMN**		
DOY	**NWP_ZTD**	**bias**	**stdev**	**NWP_ZTD**	**bias**	**stdev**	**NWP_ZTD**	**bias**	**stdev**
1	2566.6	-30.4	11.7	1928.8	-20.1	15.3	2057.8	-21.1	8.8
2	2567.9	-35.8	12.4	1890.6	-7.1	7.6	2040.1	-3.5	16.2
3	2553.7	-43.8	19.3	1877.4	-13.5	12.6	2021.1	9.7	17.0
4	2565.3	-18.2	7.4	1881.1	-3.3	7.7	2014.5	14.4	15.8
5	2551.6	-17.0	11.3	1884.6	-3.7	17.7	2009.4	10.6	15.0
6	2569.0	-16.8	5.9	1876.1	10.2	20.1	2012.9	24.9	11.8
7	2580.5	-29.8	8.2	1884.0	-5.4	15.4	2023.7	11.6	7.8
8	2590.4	-35.1	9.8	1903.4	10.7	16.2	2052.7	7.9	12.3
9	2600.8	-30.6	14.0	1917.4	-6.9	14.9	2073.9	-7.8	21.7
10	2570.3	-26.4	5.8	1875.9	-9.7	13.7	2036.3	8.8	22.4
11	2543.5	-22.9	3.7	1847.1	-2.4	12.4	2000.8	7.3	17.2
12	2565.5	-24.6	6.2	1862.4	8.0	21.3	2033.2		
13	2569.0	-19.9	11.2	1876.4	-20.9	9.3	2064.2		
14	2550.7	-29.5	16.4	1863.5	-6.5	16.7	2033.2	-14.5	17.2
15	2545.1	-12.6	12.0	1864.8	-1.5	12.5	2020.9	10.9	7.9
16	2574.4	-17.9	8.2	1879.4	0.0	10.0	2044.6	14.6	12.9
17	2594.2	-30.6	9.3	1891.1	0.2	9.9	2055.8	25.1	8.1
18	2570.3	-32.6	7.5	1892.7	22.9	21.1	2058.5	30.6	18.0
19	2559.9	-35.5	10.3	1927.8	7.7	6.9	2084.7	1.1	18.9
20	2585.3	-21.0	7.2	1933.8	-1.7	9.0	2084.0	-7.5	13.8
21	2583.6	-21.3	14.3	1917.0	2.5	9.1	2077.8	-3.5	5.3
22	2594.9	-15.5	15.2	1907.0	-2.1	12.1	2062.6	7.6	13.9
23	2571.2	-11.9	12.0	1883.8	-6.3	9.6	2050.6	-1.8	10.5
24	2546.6	-25.1	9.4	1860.6	-0.1	19.4	2022.2	3.6	12.9
25	2546.3	-26.5	6.9	1870.3	-7.3	11.8	2029.8		
26	2554.2	-14.9	11.7	1886.2	-3.6	19.9	2040.7		
27	2578.3	-21.3	13.8	1900.9	-16.0	20.2	2051.0	-9.4	23.4
28	2575.7	-26.4	12.3	1867.1	-1.7	17.8	2026.0	14.0	16.1
29	2575.9	-25.6	8.2	1875.3	-2.8	23.7	2025.8	13.3	14.3
30	2564.1	-21.3	16.9	1860.5	1.4	17.7	2017.6	17.7	14.4
31	2529.6	-7.0	13.6	1863.6	-3.2	6.8	2016.9	17.3	14.3

Acknowledgement

I am grateful to my sponsors, Germany's *Deutsche Akademischer Austauschdienst* (DAAD) and Kenya's National Commission for Science, Technology and Innovation (NACOSTI), for availing the funds that have made it possible for me to have uninterrupted studies.

I am greatly indebted to my supervisor, Prof. Dr.-Ing Torben Schüler, for his guidance, valuable advice and support throughout the research period.

Thanks also to the Institute of Space Technology and Space Applications (ISTA) at the University of the Federal Armed Forces Munich (*Universität der Bundeswehr München*), through Prof. Dr.-Ing Bernd Eissfeller, for hosting me throughout the entire study period.

Without prior experience in numerical weather modeling, the Environmental Monitoring System (EMS) made it very easy to learn the art and finally achieve the main objectives of this research. I therefore wish to thank the developers of the EMS (and all the systems therein), notably Robert Rozumalski, and the affiliated institutions, the NOAA/NWS Science and Training Resource Center (STRC), the National Center for Atmospheric Research (NCAR) and the National Center for Environmental Predictions (NCEP).

I am grateful to all the organizations/institutions/projects that availed data that was used in this work. They are all mentioned under the respective chapters covered in this thesis.

There are many I would like to mention, who have in one way or another assisted me in the course of carrying out my research work. Kindly feel appreciated. I have also tried as much as possible to acknowledge cited works of other researchers in this thesis. This can be found in the references.

Finally, thanks to my employer, the Jomo Kenyatta University of Agriculture and Technology, for giving me some time-off to concentrate on my studies.