Hochschule Neubrandenburg
University of Applied Sciences

Computation of Precipitable Water Vapour Using GPS and Sparse Publicity Available Meteorological Data

—

Case Studies in New Zealand and Europe

by

Andreas Krietemeyer

A Thesis Submitted
In Partial Fulfillment of the
Requirements for the Degree of

„Master of Engineering“ (M.Eng.)

in Geoinformatics and Geodesy

October 2014

Advisor: Prof. Dr.-Ing. Andreas Wehrenpfennig
Co-Advisor: Dr. Paul Denys

URN: urn:nbn:de:gbv:519-thesis2014-0607-7

Contents

List of Figures

List of Tables

List of Abbreviations

AVHRR	Advanced Very High Resolution Radiometer
AWS	Automatic Weather Station
CDMA	Code Division Multiple Access
CHAMP	CHAllenging Minisatellite Payload
EGM96	Earth Gravity Model 1996
GFZ	German Research Centre for Geosciences
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRACE	Gravity Recovery and Climate Experiment
GTRF	Galileo Terrestrial Reference Frame
IGS	International GNSS Service
ITRS	International Terrestrial Reference System
ITU	International Telecommunication Union
JGS	Japanese Geodetic System
MAPS	Mesoscale Analysis and Prediction System
NIWA	National Institute of Water and Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
PPP	Precise Point Positioning
PWV	Precipitable Water Vapour
QZSS	Quasi-Zenith Satellite System
RINEX	Receiver Independent Exchange Format
RMS	Root Mean Square
TEC	Total Electron Content
UTC	Universal Time Coordinated
WVR	Water Vapour Radiometer
ZHD	Zenith Hydrostatic Delay
ZTD	Zenith Tropospheric Delay

Abstract

Precipitable Water Vapour (PWV) data is used in meteorology studies and is not always easy to obtain. This project combines GPS derived Zenith Tropospheric Delay (ZTD) together with meteorological data for the estimation of PWV. Stations with both, meteorological and GPS on-site measurements available are used to produce reference data (verification stations). The work shows a possible way to take existing GPS stations without on-site meteorological observations for the PWV estimation by interpolating temperature and pressure measurements from sparse, publicity available, meteorological sites. First, the GPS stations network from the University of Otago in Dunedin, New Zealand, is processed. In order to have more reference stations available, the implemented approach has been adopted for a region in Central Europe, where data from the International GNSS Service (IGS) and the German Research Centre for Geosciences (GFZ) are available.

Using the interpolated meteorological data, PWV is estimated at the GPS sites and PWV time series maps are generated. In addition to the maps, the PWV calculations computed with sparse meteorological data are compared to the reference data. A PWV Root Mean Square (RMS) of smaller than ± 0.5mm is achieved for the three available stations in New Zealand. The same accuracy is observed for 10 out of 12 IGS stations, whilst only 130 out of approximately 200 of the GFZ stations have an PWV RMS of smaller than ± 0.5mm.

1. Introduction

Atmospheric processes are influenced by Precipitable Water Vapour (PWV). A better knowledge of the hydrological cycle allows a better understanding of the earth's climate system. In this context, PWV is applied in many ways such as weather forecasting and climate models. Using the Zenith Tropospheric Delay (ZTD) derived from GPS signals together with meteorological temperature and pressure observations is one opportunity to estimate PWV. This concept has been known since the early 1990s and has been described by e.g. Bevis *et al.* (1992) and Rocken *et al.* (1993). This project demonstrates a method to use available GPS stations for the PWV estimation by interpolating sparse, publicity available, meteorological data. To quantify the precision of the interpolated values, stations with both, GPS and meteorological measurements available (verification stations), are used to produce reference data that will be compared to the sparse (interpolated) estimations. Unlike optical survey methods, GPS measurements are independent of weather conditions with high observational quality and sampling characteristics in different geographic regions. Hence, GPS based water vapour estimation systems are already integrated in daily weather models, tested in realistic experiments, economically attractive and can be used for long period observations.

The University of Otago in Dunedin, New Zealand, manages GPS stations throughout the country. However, only a few of them can be used to estimate PWV with on-site meteorological data - so-called reference stations. By interpolating temperature and pressure measurements from sparse meteorological sites, GPS stations without on-site meteorological observations will be used to estimate PWV. Since there are only a few reference stations available in the chosen region in New Zealand, the implemented approach has been further developed and may be used for any geographic region. A region in Central Europe has been used that has the highest density of freely available International GNSS Service (IGS) stations. In addition, many verification datasets have been provided by the German Research Centre for Geosciences (GFZ) in this area. For each region two weeks of data, one week in summer and one in winter, has been processed. This results in regional PWV time series maps and a comparison of the interpolated PWV estimates with those of the reference data.

Chapter 2 characterizes the theoretical background of this work from the basics of Global Navigation Satellite System (GNSS) to explaining the structure of the atmosphere and concludes by describing the Global Positioning System (GPS) based estimation of PWV. Chapter 3 outlines the approach, realization and working progress, whilst the results of this project are represented in Chapter 4.

2. Background and Basic Principles

This chapter provides an overview of the basic terminology and principles for the estimation of Precipitable Water Vapour (PWV). In Sections 2.1-2.3, the fundamentals of GNSS, the water vapour in its atmospheric context and the structure of the atmosphere will be described briefly. Section 2.4 describes the estimation of PWV.

2.1. Basics of GNSS

A Global Navigation Satellite System (GNSS) is a space based system which provides global coverage of geo-spatial positioning and navigation by receiving signals transmitted along the line of sight from satellites. GNSS is also a generic term for existing and future satellite systems such as GPS, GLONASS, Galileo or BeiDou-2. This space based positioning methodology calculates precise time of the signal path and thus the GNSS receiver. To derive this data, signals to at least four satellites must be received at a time. With these four observations, the receiver calculates the pseudoranges and determines its position and the clock bias between the receiver and satellite clocks from it. The distance from the satellite to the receiver results from the elapsed time of the signal between sender and receiver. The calculated time and coordinates relate to the coordinate system of the corresponding navigation system. (Hofmann-Wellenhof *et al.*, 2008)

Currently, only the U.S. GPS and the Russian GLONASS, are fully operational GNSS. GPS uses the WGS84 as the geodetic reference system. Each satellite transmits signals on a specific carrier frequency. GPS uses two frequencies in the L band. The first carrier frequency, L1, is at 1575.42MHz and the second, L2, is at 1227.60MHz. The GPS satellites transmit their signals on these frequencies and are separated from each other by Code Division Multiple Access (CDMA). Modulated onto the carrier frequencies is information such as satellite time, ephemeris data used to calculate the satellite position in the orbit and auxilary data like the satellite clock correction. The receiver compares the satellite transmitted signal with one generated by the receiver to determine the transmission time of the signal from satellite to receiver. The satellite to receiver range (pseudorange PR) is determined by scaling the transmission time Δt by the speed of light c:

$$PR = c\Delta t$$

The propagation velocity c in a vacuum (e.g. space) is approximately $3 \times 10^{8 km/s}$. However, the signal is refracted and slowed down when travelling through the earth's atmosphere (see also Section 2.3). The three dimensional position of the

receiver is computed by using the principle of a three dimensional resection of three different pseudoranges. An additional fourth satellite is required in order to determine the error between the receiver and satellite clocks. Further satellite ranges are used as an over-determination to improve the receiver position. For this purpose, it is an advantage to use as many satellites as possible to determine the receiver position. Furthermore, removing satellites with low elevation angles typically improves the positioning, since low elevation satellites have greater distortion through the atmosphere. (Hofmann-Wellenhof *et al.*, 2008)

Since 2011, the Russian navigation system GLONASS also provides global coverage and has a similar structure to the US GPS. In contrast to GPS, GLONASS uses the PE-90 as the geodetic reference system. All Russian satellites transmit the same code (pseudo-random noise), but each on a different frequency. Whilst GPS uses CDMA, GLONASS uses frequency multiplexing. However, in future, GLONASS satellites will transmit CDMA codes on a fixed frequency to simplify the combination of multiple GNSS systems. The 24 satellites[1] orbit the earth on three orbital planes (GPS six orbital planes) with an altitude of approximately 19,100km (GPS 20,200km) and an orbit inclination of around 64.8° (GPS 55°). The higher inclination angle results in having better coverage in higher latitude regions (polar regions). (Hofmann-Wellenhof *et al.*, 2008)

The planned European navigation system Galileo aims to be fully operational in 2019-2020 (ESA-Galileo, 2014). The structure is similar to GPS and GLONASS. However, whilst GPS and GLONASS are under military control with civilian access, Galileo is entirely under civilian control. It is planned to launch 27 (+3 reserve) satellites with an orbit inclination of 56° in three orbital planes at an altitude of about 23260km. It will use the Galileo Terrestrial Reference Frame (GTRF) as the geodetic reference system. Like GPS, it will establish an independent realization for the International Terrestrial Reference System (ITRS) (ESA, 2014).

The Chinese satellite navigation system BeiDou-2 (formerly COMPASS) is also currently under construction. It is the global expansion of the regional BeiDou-1 system, which has operated since the end of 2011. It is expected that by 2020, 35 satellites will be available at a height of approximately 22,000km. BeiDou-2 will also include five geostationary orbiting satellites to enable backward compatibility with BeiDou-1 and 30 non-geostationary satellites that will be configured for global coverage. The ranging signals are based on the CDMA principle similar to that used in Galileo or GPS. The signal frequencies overlap with Galileo. However, the International Telecommunication Union (ITU) policies stipulate, that the first nation starts broadcasting in a specific frequency will have priority to

[1] status as of 28^{th} June 2014

that frequency, which has lead to a race between Galileo and BeiDou-2 for these frequencies. (Hofmann-Wellenhof *et al.*, 2008)

The Japanese Quasi-Zenith Satellite System (QZSS) provides a navigation service in East Asia and Oceania. The satellites orbit between 32000km and 40000km height with an orbital inclination of 43°. At least one satellite will always be in quasi-zenith direction over Japan. This design provides positioning services in urban canyons and mountainous environments. QZSS is primarily designed as an complementary system to GPS, but it can also operate as an independent regional service. Thus, QZSS has adopted CDMA and the same carrier frequencies to be compatible and interoperable with the GPS signals. It is proposed that until 2018, four satellites will be launched of which three will be visible at all times from locations in this region. QZSS uses the Japanese Geodetic System (JGS) as coordination system, which uses the fundamentals of the GRS-80 defined ellipsoid. (Hofmann-Wellenhof *et al.*, 2008)

GNSS estimated positions are limited in accuracy due to several measurement errors. One such error is a result of the effect of the atmosphere. When travelling through the atmosphere, the signal is refracted and slowed down. The upper layer of the earth's atmosphere (above ∼80km), the ionosphere, the free electrons causes an positioning error. This effect is eliminated by using signals on two frequencies (see also Section 2.4.2). Due to this reason, each GNSS uses at least two frequencies.

The lower layer of the atmosphere, the troposphere also causes a delay and refraction of the signal. The tropospheric effect is predominantly influenced by the amount of water vapour and air molecules in the air. For a zenith-measurement the tropospheric delay is approximately 2.3 metres. For lower signal elevations, this effect rises because of the longer propagation path through the troposphere. Section 2.4 provides an overview of how the tropospheric delay is used for the estimation of PWV.

2.2. Water Vapour

In the atmosphere, water occurs in all three states: liquid as water, solid as ice or in its gaseous state as water vapour. The amount of ice and liquid water is not measurable by GPS surveying methods. Observing these must be done with other technologies like rain gauges or weather radars. This project concentrates on water in its gaseous state, which can be determined using ground-based GPS atmosphere soundings.

Water vapour occurs either naturally or as a result of industrial activity and is the earth's most important greenhouse gas in the atmosphere. While constantly

cycling through the atmosphere and transferring heat energy around the earth and the atmosphere, condensing water vapour forms clouds that absorb and reflect the long-wave thermal radiation of the earth. This effect is called the natural greenhouse effect and keeps the earth sufficiently warm to support life. (Kraus, 2009)

Water vapour accounts for the largest percentage of the greenhouse effect. Increasing the concentration of greenhouse gases, and particularly water vapour, heavily impacts on the earth's climate. As a consequence, the earth would heat up and rainfall and climatic zones would change (Metz *et al.*, 2007). Evaluating the feedback of water vapour in global warming is therefore of eminent importance. Hence, comprehensive observations of the water vapour distribution and their variations are essential.

The distribution of water vapour varies strongly vertically as well as horizontally. Basically, water vapour concentration can be reduced by decreasing temperature and therefore by increasing height. Although water vapour contributes only about 0.25 percent of the total sea level pressure of gases, it is essential to the earth's climate. "If all the water vapour in the air at a particular time were to condense and fall as rain, it would amount to a depth of only about 2.5 cm" (Mockler, 1995). It can be measured, for instance, by radiosonde balloons that afford good vertical and regional measurements, but at high costs. It can also be estimated by remote sensing satellite systems from infra-red radiation measurements with a good horizontal resolution. However, these measurements are only reliable in cloud-free conditions (Gutman and Benjamin, 2001). Another method of acquiring atmospheric information is to use water vapour radiometers (WVR), which measure the background microwave radiation of the water vapour. There are two types of WVRs: ground-based and space-based. A ground-based, upward looking, WVR measures the radiation in a given line of sight against the cold background of space. They are not affected by light cloud cover and have a good temporal but poor spatial resolution. Downward-looking, space based, WVRs measure the absorption lines from the earth. Because this technique depends on temperature, it is more useful over oceans than over land and provides good spatial and poor temporal coverage. A disadvantage of WVRs is the poor performance in heavy cloud or during raining conditions (Bevis *et al.*, 1992).

Water vapour is important for short-term weather patterns and is already integrated in current day weather models. Cost-efficient measurements such as GPS based estimation of water vapour in the atmosphere is one possibility to obtain observations at all weather conditions for near real-time weather predictions.

2.3. Structure of the Atmosphere

The earth's atmosphere is a gaseous mixture that surrounds the earth. It is prevented by gravitation from leaking into space. From a thermal perspective, the atmosphere is divided into layers known as the troposphere, stratosphere, mesosphere and thermosphere (see Figure 1). The troposphere is the lowest layer and

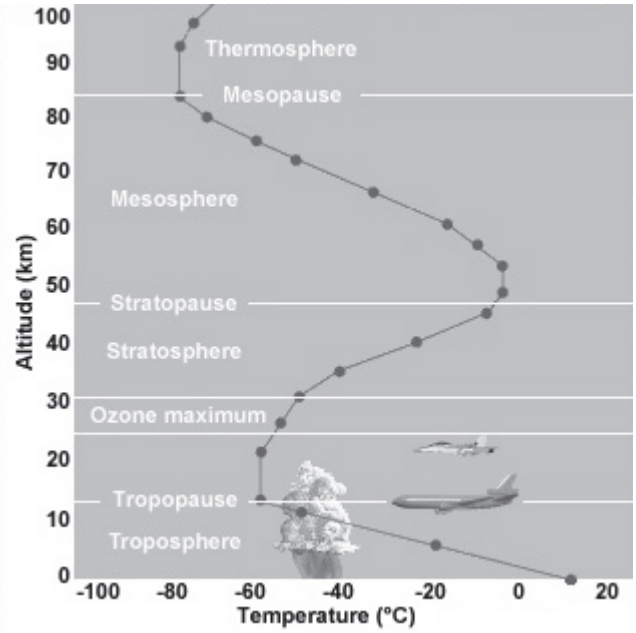


Figure 1: Structure of the atmosphere (Source: nc-climate.ncsu.edu, 2013)

contains approximately 80% of the mass of the atmosphere. The height of the troposphere ranges between 9km at the poles and 16km at the equator. In the mid-latitude regions, the troposphere ends at an elevation of about 10km. The temperature in this layer decreases approximately by -0.5 °C up to -0.6°C per 100 metres. Accordingly, at the transition to the tropopause, the temperature is about -60 °C. In the tropopause, the temperature is almost constant but then starts to increase at 20km to the stratosphere boundary (at 30km). At the boundary between the stratosphere and stratospause (50km), the temperature is about 0 °C. In the following layer, the mesosphere, continues to a height of 85km and the temperature has decreased to around -70°C. In the mesopause, the temperature remains approximately constant. In the next layer, the thermosphere, the temperature increases continuously again. (nc-climate.ncsu.edu, 2013)

Besides these thermal aspects, the atmosphere can also be divided into ionization levels of its gaseous mixtures. The resulting variation to the signal path is one reason for the refraction of the GPS signals. Regarding this property, the atmosphere is structured in an ionospheric and an tropospheric component. The ionosphere ranges from over 80km up to 1000km height and consists of electrically charged particles, which refract the GPS signal and increases the GPS signal

transmission time. The ionosphere is dispersive at 1.5GHz. Dispersive mediums are dependant upon "the phase velocity and, thus, of refraction on the frequency" (Hofmann-Wellenhof *et al.*, 2008). The effect of the ionosphere refraction can be eliminated using a linear combination of the phase measurements from two frequencies, e.g. GPS L1 and L2 frequencies. This is also known as the ionosphere free linear combination.

The troposphere and stratosphere are grouped as the tropospheric layers of the earth's atmosphere. The molecules and atoms in this layer occur mostly in a neutral state and are therefore not dispersive. Hence, an elimination of this refraction by dual-frequency methods is not possible. When passing through this layer, the neutral gaseous particles are polarized by the electromagnetic field of the signal. As a consequence, the signal covers a longer path with lower speed and therefore reaches the receiver delayed. The tropospheric refractive index is dependant of temperature, pressure and partial water vapour pressure. (Rocken *et al.*, 1993)

The tropospheric delay is divided into a dry and a wet delay. To mitigate these components, models are used. Models for the dry and wet refractivity have been known for some time (e.g. Essen and Froome (1951)). The dry delay occurs due to the different components of atmospheric gases like the dry gases oxygen, carbon dioxide and nitrogen. About 90% of the total tropospheric delay is caused by the dry component. The wet delay is due to the water vapour in the atmosphere and is dependent upon the atmospheric conditions along the path of the signal. In a homogeneous medium like a vacuum, the signal would arrive in linearly to the receiver. Though, the atmosphere is a gaseous mixture with a vertical pressure flow. Therefore, the true path deviates from the geometric one. Section 2.4 deals with the propagation of this geometric path through the atmosphere and demonstrates the estimation of PWV in this context. (Bevis *et al.*, 1992)

2.4. GPS based Estimation of PWV

2.4.1. Overview

The Global Positioning System (GPS), originally designed by the US Air Force for military navigational purpose, consists of 31 satellites[2] at approximately 20,000km height in six orbital planes with an orbital inclination of 55° (see Figure 2). Due to

Figure 2: Satellite orbits (Source: satsig.net, 2006)

the comparable short distance between the receiver and the satellite with regard to the speed of light, a high temporal resolution can be achieved. The satellites are transmitting L-band radio signals with 19 and 22cm wavelengths, which are modulated with the C/A and P/Y ranging codes. After deactivating selective availability on 2nd May 2000 on the publically available C/A-Code, civilian use has grown with time. (Navigation Center of Excellence, 2001) For example, the International GNSS Service (IGS) network has been established, which provides more than 300 globally distributed stations in support of geodetic and geophysical research activities. To achieve the accuracy required by these applications, it is necessary that all possible error sources are accounted for. Besides the satellite orbits, clock and antenna errors, another major error source lies in the propagation of the GPS signals when passing through the atmosphere.

For heights greater than 90km, pressure and water vapour are negligible (ionospheric component). As a consequence, the refractivity is directly proportional to the electron density of the ionosphere. This electron content along a straight signal path between the satellite and the receiver is called the total electron content (TEC) and depends upon sunspot activity as well as elevation and azimuth of the satellite. Even though some models have been developed to compute the TEC, the

[2] status as of 22^{nd} June 2014

most efficient method is to eliminate the ionospheric effect by using two signals with different frequencies. By taking a linear combination of the two frequencies, the ionospheric effect is eliminated (see Section 2.4.2).

In the lower troposphere, water vapour influences the refractivity by as much as 30%. The tropospheric effect can be used to derive accurate water vapour profiles together with given pressure and temperature data.

Rocken *et al.* (1993) describe a number of ways to estimate water vapour in the atmosphere by GPS measurements. One possibility is to obtain atmospheric soundings through a radio occultation technique using low-earth orbits (LEO) satellites. Using radio frequencies, the bending can not be measured directly. Instead, the bending is calculated by using the Doppler-Shift of the signal given the geometry of the satellite and receiver. The bending can be related to the refractivity, which is a function of the electron density in the ionosphere and temperature, pressure and water vapour in the neutral atmosphere (below ionosphere). With this method, useful informations about the structure of the ionosphere, stratosphere and troposphere can be obtained. (Melbourne *et al.*, 1994) In practice, the satellite missions CHAMP, GRACE, COSMIC or SAC-C were using this technique. The CHAllenging Minisatellite Payload (CHAMP) is a German satellite mission for geoscientific and atmospheric research. The mission started on 15$^{\text{th}}$ July 2000 and ended in September 2009. It has been the first satellite mission which obtained highly precise gravity and magnetic field measurements. The Gravity Recovery and Climate Experiment (GRACE) was launched in March 2002 and is expected to end in 2015. It consists of two satellites separated by a distance of 220km. Besides atmospheric research, its main objective is to achieve accurate mapping variations in the earth's gravity field. FORMOSAT-3/COSMIC is a Taiwanese-U.S. six-satellite mission. It started in April 2006 and is mainly designed for atmospheric research. It has resulted in huge improvements in global weather forecasting. A follow-on mission (COSMIC-2) is planned for 2015. SAC-C (Satelite de Aplicaciones Cientificas - C) is an international mission. Its main contributors are Argentina and the US. Its mission is to study the structure and dynamics of the earth's surface by using multi spectral images as well as researching the atmosphere, ionosphere and the geomagnetic field.

Another approach to determine water vapour in the atmosphere is to collect dual-frequency signals at ground-based receivers. These are used to obtain the signal delay and thus the water vapour along the signal path from the satellite to the receiver (Bevis *et al.*, 1992) (see also Section 2.4.2). This method was recognized during error investigations of geodetic measurements (Davis *et al.*, 1985). Kuo *et al.* (1996) described with an data assimilation system the improvements of short range numerical weather models with GPS based precipitable water va-

pour estimations. Wolfe and Gutman (2000) stated the results of a GPS, ground based, water vapour observing system for NOAA. At that moment, 35 continuously operating stations are managed by the NOAA network. Using GPS zenith water vapour estimations, accuracies up to a maximum of 1.5cm (at PWV values up to 5cm) compared to radiosondes, WVRs and the Forecast Research Division's Mesoscale Analysis and Prediction System (MAPS) model were achieved. Boccolari *et al.* (2002) also compared PWV and ZTD estimations from ground based GPS stations with radiosoundings and other meteorological measurements in Italy. Besides a PWV standard deviation of ±2.7mm, their results indicate a high correlation between precipitation, forthcoming increase and decrease of PWV. Motell *et al.* (2002) compared GPS, radiosondes and the satellite based Advanced Very High Resolution Radiometer (AVHRR) based split-window techniques over Hawaii. "The split-window technique uses the differential water vapour absorption at two adjacent thermal channels under clear sky conditions to produce PWV estimates" (Motell *et al.*, 2002). Using this method, errors of about 3.8mm (14%) of the typical PWV values were obtained. Walpersdorf *et al.* (2007) assesses the availability and quality of data (especially ZTD values) from the IGS network in Africa and examines the major error sources for the GPS based PWV estimations with studies performed in the year 2001. Wonnacott (2005) provides a good overview of the background and basic principles of water vapour and its estimation. He also compared GPS, radiosonde and numerical weather model PWV estimations. He stated a mean difference of -4.08mm between radiosonde and GPS based PWV estimations, as well as -1.86mm between NWM and GPS and 1.22mm between NWM and radiosondes. Zhengdong and Yanming (2003) used interpolated surface meteorological data from the Australian Automatic Weather Station (AWS) network to estimate PWV. They compared the PWV estimations based on GPS and interpolated meteorological observations against PWV radiosonde results. Their comparisons showed that the estimates agreed with an RMS of ±1.74mm for the four tested stations. GPS data, whether obtained in situ or occultation measurements, are potentially important for climate and Numerical Weather Prediction (NWP) applications.

2.4.2. Signal Refraction and PWV Estimation

The electromagnetic signal transmitted by a GPS satellite travels through several layers of the earth's atmosphere (see Section 2.3). Each layer slows and refracts the signal before arriving at the receiver. In fact, the ionosphere and troposphere refract the propagation of GPS signals through the atmosphere (see Figure 3).

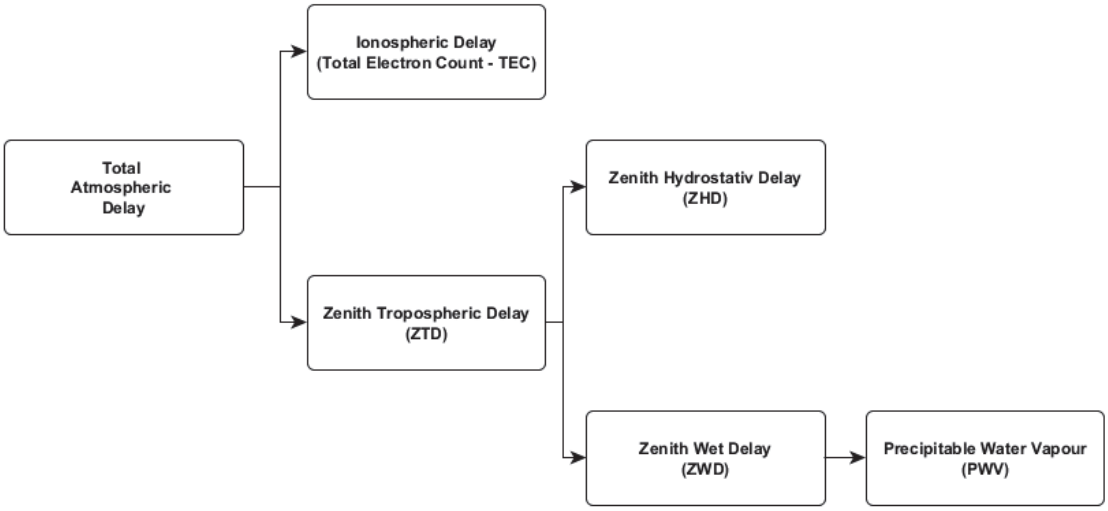


Figure 3: Break down of the atmospheric delay according to Bevis *et al.* (1992)

An appropriately derived linear combination of two different frequencies will eliminate the ionospheric delay. Using the phase range with respect to the different signal frequencies gives:

$$\begin{aligned} \lambda_1 \Phi_1 &= \varrho + c\Delta\delta + \lambda_1 N_1 - \Delta_1^{Iono} \\ \lambda_2 \Phi_2 &= \varrho + c\Delta\delta + \lambda_2 N_2 - \Delta_2^{Iono} \end{aligned} \tag{1}$$

Where λ_1 and λ_2 are the wavelengths of the different frequencies ($c = \lambda f$), Φ_1 and Φ_2 the phase pseudoranges, ϱ the distance between satellite and receiver, c the speed of light, $\Delta\delta$ the receiver clock bias, N_1 and N_2 the cycle ambiguity and Δ_1^{Iono} and Δ_2^{Iono} frequency-dependent ionospheric influences on the measured pseudoranges. After dividing the corresponding wavelengths and rearranging, the phase pseudoranges can be written as

$$\begin{aligned} \Phi_1 &= af_1 + N_1 - \frac{b_1}{f_1} \\ \Phi_2 &= af_2 + N_2 - \frac{b_2}{f_2} \end{aligned} \tag{2}$$

Where a represents the geometry term, defined as

$$a = \frac{\varrho}{c} + \Delta\delta$$

and b the ionospheric term, defined as

$$b_i = \frac{f_i^2}{c}\Delta^{Iono} \qquad i = 1, 2, ...$$

Both terms are frequency independent. The ionospheric term is eliminated by multiplying equation 2 by f_1 and f_2 respectively and differencing to give

$$\Phi_1 f_1 - \Phi_2 f_2 = a(f_1^2 - f_2^2) + N_1 f_1 - N_2 f_2 \tag{3}$$

Multiplying equation 3 by $f_1/(f_1^2 - f_2^2)$ and rearranging, the ionosphere-free combination is achieved as

$$\left[\Phi_1 - \frac{f_2}{f_1}\Phi_2\right]\frac{f_1^2}{f_1^2 - f_2^2} = \frac{f_1}{c}\varrho + f_1\Delta\delta + \left[N_1 - \frac{f_2}{f_1}N_2\right]\frac{f_1^2}{f_1^2 - f_2^2} \tag{4}$$

Equation 4 describes the ionosphere-free combination for phase measurements. An important issue of this combination is that the ambiguity term $\left[N_1 - \frac{f_2}{f_1}N_2\right]$ is now not an integer number for current GNSS. To avoid this, the derivation for the code pseudoranges is used:

$$\begin{aligned} R_1 &= \varrho + c\Delta\delta + \Delta_1^{Iono} \\ R_2 &= \varrho + c\Delta\delta + \Delta_2^{Iono} \end{aligned} \tag{5}$$

Δ^{Iono} is inversely proportional to the respective squared carrier frequency, multiplying equation 5 by f_1^2 and f_2^2 respectively, the difference forms

$$R_1 f_1^2 - R_2 f_2^2 = (f_1^2 - f_2^2)(\varrho + c\Delta\delta)$$

where the ionospheric effect is eliminated. Dividing this equation by $(f_1^2 - f_2^2)$ gives the ionosphere free combination for code pseudoranges:

$$\left[R_1 - \frac{f_2^2}{f_1^2}R_2\right]\frac{f_1^2}{f_1^2 - f_2^2} = \varrho + c\Delta\delta \tag{6}$$

To reduce the effects of the ionosphere, GPS transmits two carrier phase measurements, the L1 and L2 frequencies. By using this technique, the combination for L1 and L2 GPS phase measurements results in the ionosphere-free observable, namely L3:

$$\begin{aligned} L3 &= \frac{1}{f_1^2 - f_2^2}(f_1^2\Phi_1 - f_2^2\Phi_2) \\ L3 &= 2.546\lambda_1\Phi_1 - 1.546\lambda_2\Phi_2 \end{aligned}$$

Thus, the ionospheric effect is eliminated. However, the tropospheric delay is an electrically neutral medium and independent of frequency. Due to this, it can not

be eliminated by a linear combination like the ionospheric delay. Instead, a model is required. The tropospheric delay can be calculated as the difference between the measured range and the geometric distance using the following expression after Hofmann-Wellenhof *et al.* (2001):

$$\Delta^{Trop} = \int (n-1)ds = 10^{-6} \int N ds \tag{7}$$

where n is the refractive index of air and N the refractivity. The tropospheric delay can be divided in a dry (also known as hydrostatic) and a wet component and can also be related to the atmospheric temperature and partial pressures of water vapour and dry gases using equation 8 after Thayer (1974):

$$N = N_{dry} + N_{wet} = \underbrace{k_1 \frac{p_d}{T}}_{dry} + \underbrace{k_2 \frac{e}{T} + k_3 \frac{e}{T^2}}_{wet} \tag{8}$$

The empirical constants $k_1 = (77.604 \pm 0.014)^K/_{mbar}$, $k_2 = (64.79 \pm 0.08)^K/_{mbar}$ as well as $k_3 = (3.776 \pm 0.004)^{10^5 K^2}/_{mbar}$ have been defined in Thayer (1974). The variable p_d is the combined partial pressure of the dry gases in mbar, T the temperature in degree Kelvin and e the partial pressure of water vapour in mbar. Using the dry term, the Zenith Hydrostatic Delay (ZHD) can be obtained using the Saatamoinen or Hopfield model. According to Saastamoinen (1972) the ZHD follows as:

$$ZHD = \frac{0.0022768 * P_0}{1 - 0.00266 cos(2\Phi) - 0.00028H} \tag{9}$$

where P_0 is the total atmospheric pressure in mbar, the latitude in radians and H the height in kilometres. They are expected to be measured at the observation-point (receiver position).

The Zenith Total Delay (ZTD) is defined as the sum of the zenith dry (hydrostatic) and the zenith wet component, projected onto the vertical axis of the propagation path:

$$ZTD = ZHD + ZWD \tag{10}$$

The maximum delay at zenith is approximately 2.3 metres. 90% of the tropospheric delay is attributed to the dry delay and 10% to the wet delay. Given high quality surface pressure measurements (±1mbar), the ZHD can be estimated with an accuracy of a few millimetres (Gutman and Benjamin, 2001). The wet delay is always much smaller but more variable than the hydrostatic delay. Even though models have been developed for the estimation of the ZWD, its predictive value is poor in comparison to the ZHD estimation (Bevis *et al.*, 1992). As a consequence,

a model with a combination of GPS and atmospheric observation is used for the estimation of the ZHD component (equation 9). To calculate the Zenith Wet Delay, the difference between ZTD and ZHD is used:

$$ZWD = ZTD - ZHD \tag{11}$$

The range of ZWD is between close to zero in arid regions and almost 35 centimetres in humid regions (Wonnacott, 2005). Once the ZWD has been estimated, the PWV can be estimated using the following set of equations 12–15:

$$PWV = \Pi * ZWD \tag{12}$$

and

$$\Pi = \frac{10^6}{\rho R_v \left[(\frac{k_3}{T_m}) + k_2' \right]} \tag{13}$$

where ρ is the density of liquid water, R_v the specific gas constant for water vapour and T_m the weighted mean atmospheric temperature along the propagation path. The variable Π is non-dimensional and typically around 0.16. It can vary by up to 20% depending on the location, height and weather (Bevis *et al.*, 1994). The calculation of T_m and k_2' is described in Bevis *et al.* (1992) with:

$$T_m = 70.2 + 0.72 T_s \tag{14}$$

$$k_2' = k_2 - m k_1 \tag{15}$$

where T_s is the surface temperature in Kelvin. k_2' is composed of the earlier stated constants k_1 and k_2 and the the ratio of the molar masses of water vapour and dry air m, which leads to the units Kelvin per millibar ($\frac{\mathrm{K}}{\mathrm{mbar}}$). The above equations 9–15 assume a propagation path coming from the zenith direction. This means, the signal arrives from the vertical direction opposite to the gravitational force at the receiving antenna. Usually the signals travels to the GPS receiver at different elevation angles from the line of sight and is not received from the zenith direction. To reduce these signals to their corresponding zenith path, a mapping function is used. The most basic form of such a function is the reciprocal of the sine of the elevation angle. Marini (1972) proposed a mapping function in a parametrized form as:

$$m(\varepsilon) = \frac{1}{sin(\varepsilon) + \frac{a}{sin(\varepsilon) + \frac{b}{sin(\varepsilon) + c...}}} \tag{16}$$

where ε is the elevation angle of the incoming signal and the variables $a, b, c, ...$ are mapping constants dependent on the refractivity of the atmosphere, the defined height of the troposphere, height of the receiver and the radius of the spherical earth. This formula has been further developed e.g. by Saastamoinen (1972), Davis *et al.* (1985) or Niell (1996).

As mentioned before, GPS is not the only available technique for meteorologists to estimate atmospheric water vapour. Other methods, such as radiosondes, water vapour radiometers or traditional surface measurements have their own advantages and disadvantages. A comparison of these techniques that considers their relative accuracy and expenses would be of interest. A large number of experiments have already been carried out that compares GPS derived water vapour estimations with radiosonde, water vapour radiometers and surface meteorological measurement e.g. Coster *et al.* (1996), Motell *et al.* (2002), Boccolari *et al.* (2002), Walpersdorf *et al.* (2007) compared GPS and water vapour radiometers with radiosondes and surface measurements from eight stations in a 25km radius with 11 GPS receivers and up to three daily radiosonde launches in a 15 day period. They asserted that the GPS derived ZWD estimates agreed within 6–12mm to the WVR and radiosonde estimations. The corresponding PWV calculations agreed within 1–2mm to these techniques. Similar results have been found in different regions (Spain, Sweden, Germany) by Kruse (2001) and Reigber *et al.* (2002). Wonnacott (2005) provides a good summary of these techniques and comparisons. According to these results, Wonnacott (2005) concluded that GPS is a good method for water vapour estimations and is comparable to traditional techniques such as radiosondes and water vapour radiometers.

3. Data Preparation and Methodology

In this project, PWV values at one hour intervals for one week are calculated using sparse, publicity available, meteorological data. Due to different data sources, the input data interval varies from seconds up to several hours. This projects aims to implement a solution for the GPS sites on the South Island of New Zealand managed by the University of Otago as well as a region in Central Europe. There are only a few stations with both, meteorological and GPS data, available. Hence, the objective is to use the GPS station observations without on-site meteorological data to calculate PWV values by interpolating meteorological data recorded at meteorological sites in the region. Those stations with both ZTD and meteorological data available are used to compare the PWV estimations with interpolated meteorological data. For the purpose of this project, these sites with meteorological as well as GPS data available will be called verification or reference sites.

Based on the sparse available meteorological observations and ZTD recordings, the PWV will be estimated at each of the available GPS stations. Using the point based PWV estimations, a two dimensional spatial interpolation creates PWV maps. The calculated PWV at non-verification sites will also be compared to the available verification sites. Given that in the South Island there are only a few verification stations (typically up to three) available, the same methodology will also be adopted to a region in Central Europe. The area covered by the meteorological stations is approximately 150,000km^2 in New Zealand and 490,000km^2 in Europe. Accordingly, the European region is more than three times larger than the South Island area, whereas the amount of available meteorological stations differs only negligibly.

Sections 3.1 and 3.2 describe the general approach of this project. Section 3.3 characterizes the methodology implemented for New Zealand. The last Section concentrates on the adaptation of the implementation for New Zealand in the European region. The computation has been carried out using MATLAB. The tools are provided on the enclosed CD-ROM.

3.1. General Approach

According to Bevis *et al.* (1992), the receiver position and the ZTD together with temperature and pressure measurements must be available at the same height as the GPS receiver in order to estimate the PWV. Figure 4 provides a general overview of the single working steps. Each illustrated activity corresponds to a working step and will be explained in further detail later in this Section.

Figure 4 shows the basic work flow diagram of the PWV estimation process. Each presented activity has a specified procedure, which is defined below. In

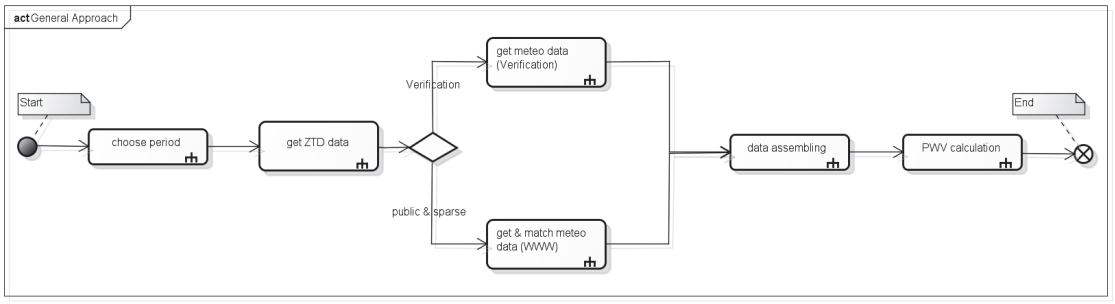


Figure 4: General approach of PWV estimation with verification or sparse meteorological data

general, to obtain PWV values, the desired period must be specified. This project used two one week periods, one during the winter and one during the summer season. The next step is to collect ZTD data. The source of this data is not critical and can be publicity available or from a secure server. The diamond symbol (◊) indicates an if-then decision in this activity diagram. The decision depends upon whether data at the verification station should be computed and thus meteorological data at the GPS site is used, or if publicity available meteorological data is used. From this point, it can be either decided whether verification with available meteorological data at the GPS sites should be calculated, or sparse available meteorological data apart of the GPS stations should be used. However, in each case, the meteorological data must be linked with the ZTD values before the PWV can be estimated. After completing all these activities, PWV values are applicable at the ZTD sites in the covered geographic region.

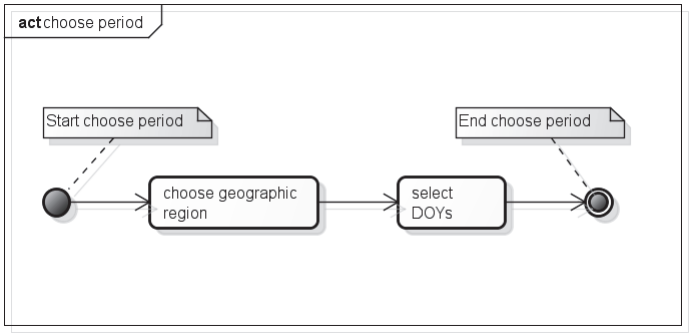


Figure 5: Activity diagram to pick an epoch for the PWV calculation

Figure 5 illustrates the very beginning of the processing work flow. The whole process starts by choosing a geographic region. The user must know where and which GPS stations with tropospheric delay data are to be used. Subsequently the epoch must be defined. Typically the tropospheric delay data, when available at that date, will be defined as either day of the year (DOY) or per GPS-Week in the Universal Time Coordinated (UTC) time zone. After defining these fundamental parameters, the first data collection is started.

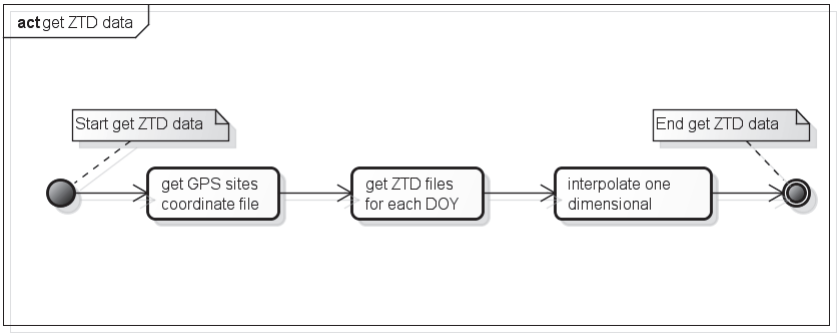


Figure 6: Obtaining ZTD values in the defined interval

The function to read in the ZTD data is shown in Figure 6. Regardless of whether verification data or estimated data with sparse meteorological data should be calculated, the tropospheric delay measurements must be gathered. A mandatory requirement is also the coordinates of all the GPS sites. Due to the fact that the collected ZTD data interval normally ranges between five minutes up to one or two hour intervals, the tropospheric delay must be interpolated over time (one dimensional) for the whole period in a self-defined interval. Throughout this project, the interpolation interval will be one hour.

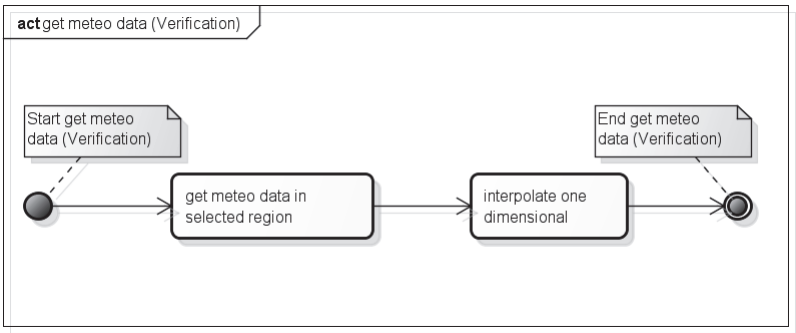


Figure 7: Getting meteorological data at the GPS site positions (verification) in the specified interval

Figure 7 displays the process to obtain meteorological data for the verification proceeding. In this case, the meteorological data must be gathered for the available GPS sites in the selected region. Once this has been acquired, the data must also be interpolated over time, similar to the tropospheric delay measurements. Figure 8 illustrates the approach when using sparse meteorological data.

In this case, arbitrary temperature and pressure measurement sites are used at sparse positions in the selected region. Similar to the verification approach, the measurements must be converted to a uniform interval that corresponds to the ZTD interval. After this, the measurements are interpolated in two dimensions over the chosen geographic region according to latitude and longitude of the meteorological site positions. As a result, a two dimensional grid with pressure and

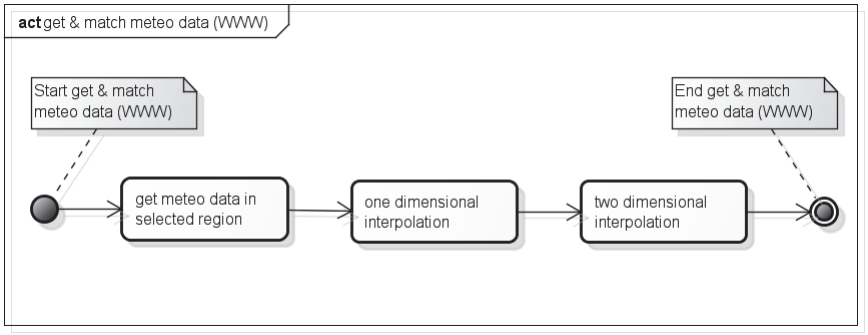


Figure 8: Approach to acquire sparse meteorological data

temperature values are available. After this, it is necessary to allocate the meteorological data with the tropospheric data. This is realized with the common interpolated time (same intervals) of the meteorological and delay data. With the meteorological measurements temperature and pressure and the available ZTD, the PWV values at these stations can be estimated by implementing the described formulae in Section 2.4.2.

The following Sections are structured as the working progress of this project. They describe the approach to estimate PWV values for the available GPS stations on the South Island of New Zealand and adopting these methods for a chosen region in Central Europe. Section 3.2 deals with implementing and verification of the described PWV formulae. For this propose, data from two DOYs of two stations with available meteorological data (verification sites) are used. To verify the implemented PWV formulae, this implementation is applied for one DOY on two sites in Hong Kong in 2013. Its output will be compared to the processed results in Denys (2010) for these two sites.

3.2. Data Preparation and Verification of the Implemented Formulae

This Section deals with the data preparation for the verification sites. Furthermore, the working process and the input data characteristics will be described. Finally, the implemented processes will be applied to two verification sites in Hong Kong, to compare the results with the already processed data from Denys (2010).

One of the outputs of the GPS Precise Point Positioning (PPP) are ZTD estimates. The GPS sites managed by the University of Otago are processed using the Bernese GPS Software Version 5.2. The ZTD estimates are stored in .TRP files. They are available on a server managed by the University of Otago in a Bernese file format similar to the Receiver Independent Exchange Format (RINEX). Each file contains header information that includes, for example, the used mapping function, or the maximum allowable satellite elevation. In addition, the data interval is defined and typically in full minutes. The .TRP files used in this project are

provided in a one hour interval. Table 1 shows some selected data columns from a .TRP file at DOY 212:

Station Name	Year	Month	Day	Hour	Min	Sec	σ	ZTD
OUS2	2013	07	31	00	00	00	0.00116	2.59033
OUS2	2013	07	31	01	00	00	0.00093	2.59185
...								
QUAR	2013	07	31	00	00	00	0.00111	2.42950
QUAR	2013	07	31	01	00	00	0.00084	2.41936
...								

Table 1: Structure .TRP file for one DOY

The .TRP file contains the estimated tropospheric data of all available GPS sites managed by the University of Otago at the DOY 212 in 2013. Table 1 shows the necessary data columns of the GPS stations OUS2 and QUAR. The required data includes the year, month, day, hour, minute and second. From this, the decimal DOY of the used data rows is calculated. The column heading "sigma" (σ) indicates the standard deviation of the ZTD data. The files also include several auxiliary data from the tropospheric GPS processing, which are not necessary in the PWV calculation. The dots indicate ongoing data rows of additional GPS sites.

station	σ	ZTD	DOY
OUS2	0.00116	2.59033	213.0000
OUS2	0.00093	2.59185	213.0416
...			
QUAR	0.00111	2.42950	213.0000
QUAR	0.00084	2.41936	213.0416
...			

Table 2: Interface read-in tropospheric delay

Table 2 shows the interface of the read-in process of the tropospheric delay data. Each data row consists of the station name, the standard deviation (σ), ZTD and the converted decimal DOY of the data.

Meteorological data managed by the University of Otago is accessible for up to three stations on the South Island of New Zealand. Compared to the .TRP file, meteorological data of each meteorological site is stored in a single data file. They are stored in the RINEX format. The header consists of information about the type of sensor used and types of observations recorded. Table 3 shows the structure of the stations OUS2 and QUAR. The file name and its extension depends on the station name, DOY and year (YY).

Like the .TRP file, the date is stored in six columns as year, month, day, hour, minute and second. On the basis of this information, the decimal DOY is also

File	Year	Month	Day	Hour	Min	Sec	P (mbar)	T (°C)	RH (%)
	13	07	31	00	00	00	1022.0	11.0	59.6
OUS2*DOY.YYM*	13	07	31	00	10	00	1021.9	11.5	58.2
	...								
	13	07	31	00	00	02	1023.5	11.0	99.6
QUAR*DOY.YYM*	13	07	31	00	05	02	1023.4	11.0	98.7
	...								

Table 3: Structure .YYM file for two verification sites

calculated for each data row. P(mbar) indicates the pressure at the station measured in millibar. T(°C) the temperature in degree Celsius and RH(%) the relative humidity in percent. An interesting feature of the meteorological data is the recording interval, which is variable and does not necessarily coincide with a full minute. Table 3 also shows this, whilst the OUS2 data is stored in a ten minute interval at full minutes, the QUAR measurements are saved in a five minute interval with a two second offset to a full minute. Hence, the meteorological information is much denser than the desired one hour interval. It is therefore necessary to interpolate the meteorological data to match the one hour interval of the corresponding ZTD data. Given that the observations are available before and after the desired point of time, the natural curve progression of the temperature, pressure and ZTD seems to be a good approximation of the reality. As a consequence, the interpolation was executed using the MATLAB cubic spline function.

station	T	P	RH	DOY
OUS2	11.0	1022.0	59.6	213.0000
OUS2	12.3	1021.3	54.8	213.0416
...				
QUAR	11.0	1023.5	99.6	213.0000
QUAR	11.9	1023.0	97.5	213.0416
...				

Table 4: Interpolated meteorological data

Table 4 shows the interface of the meteorological data. The data is interpolated in a one hour interval to full minutes to match the ZTD interval. Besides the tropospheric and meteorological data, the positions of each station is required. The coordinates are stored in a .CRD file, which contains the coordinates of every managed station. The position is stored as X, Y, and Z cartesian coordinates in metres. As the PWV estimation requires longitude, latitude and ellipsoidal height, the associated X, Y and Z coordinates of each station are transformed by using the WGS-84 ellipsoidal parameters.

Table 5 shows the imported coordinates X, Y, Z and the transformed latitude,

station	X	Y	Z	latitude	longitude	height
OUS2	-4387890.83	733420.44	-4555176.37	-0.8006	2.9760	26.1215
QUAR	-4558539.07	818911.42	-4370554.02	-0.7598	2.9638	58.0312
...						

Table 5: Interface and transformed read-in coordinates

longitude (radians) and height (metres). At this stage, the read-in process is partially automated. The data is imported for each DOY. After implementing the PWV calculation as described in Section 2.4, the meteorological and ZTD data is linked by using the joint time-stamps. After associating the geodetic coordinates latitude, longitude and height with the stations, the PWV can be estimated for the defined interval. To test the implementation, the PWV is calculated using the datasets of the DOY 213 in 2013 for the stations OUS2 and QUAR. Figure 9 shows the ZWD and PWV estimations of the OUS2 station on the DOY 213 in 2013:

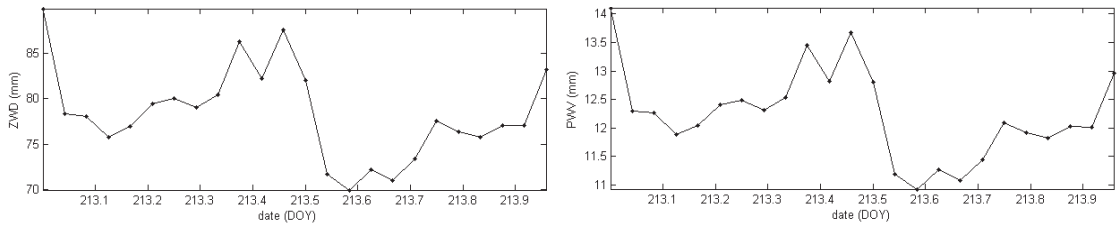


Figure 9: ZWD and PWV estimations OUS2 on DOY 213 in 2013

In order to verify the implemented formulae and reading functions, the same methodology was used with the raw data of a single day of two sites in Hong Kong in 2006 (HKWS and HKST). In contrast to the former processed ZTD data, the available tropospheric delay files are recorded in a two hour interval. The meteorological file equals the RINEX format as defined before and is stored in a one minute interval with an offset of some seconds to each full minute. Each site consists of one ZTD file and one meteorological file. In view of these amendments, the read-in process was slightly modified. In fact, the desired interval is changed to 30 minutes to match the interval used in Denys (2010). The ZTD data stored in a two hour interval is now also interpolated one dimensional. Furthermore, the read-in now handles the import from multiple files. The PWV estimation was not changed. The resulting ZWD and PWV values were then differenced from the estimations of Denys (2010). The differences between the calculation of the actual implementation and the estimations from Denys (2010) are almost zero. Even though the results in Denys (2010) has not been independently checked, it can be concluded that the PWV estimation has been implemented correctly.

3.3. Realization for the South Island of New Zealand

In the last Section, PWV and ZWD values for the two University sites OUS2 and QUAR were computed for one day in 2013 in a one hour interval. Those implemented functions are now used as fundamentals to estimate PWV values for two GPS weeks in 2009 in the South Island.

3.3.1. Concept and Input Data Description

For each week, two datasets will be computed. The first one contains verification data of three GPS sites. The second will include all available GPS sites by using sparse, publicity available, meteorological data. Figures 10 and 13 demonstrate the input for the verification and sparse PWV estimations.

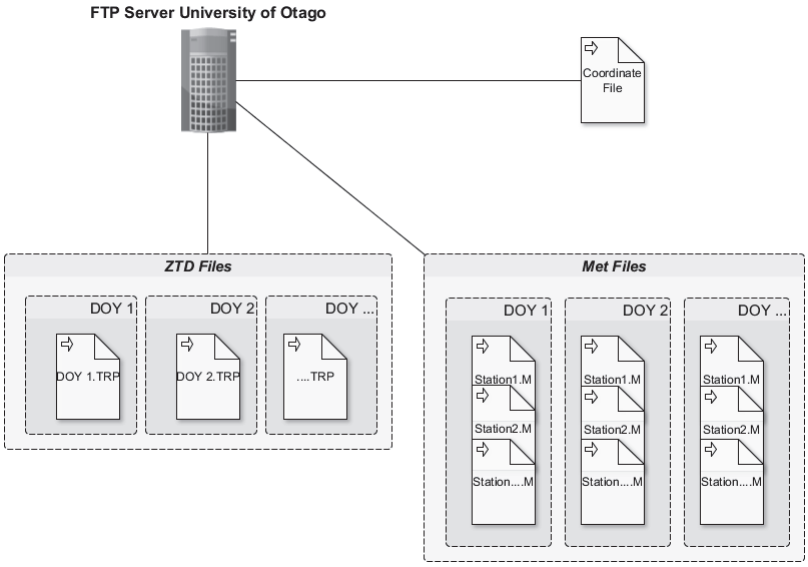


Figure 10: Use of the .TRP and .M files to obtain verification data from the FTP server managed by the University of Otago

Figure 10 shows the files from the FTP server of the University of Otago. For each DOY, one file with tropospheric delay data is available. These files include all processed GPS stations in New Zealand. For each DOY, a single meteorological datafile is provided for each available station. For all GPS stations, a single coordinate file is also supplied. All files used the same structure as defined in Section 3.2. These input files are used for the computation of PWV verification data. There are three sites for the GPS weeks 1517 (1^{st} to 7^{th} February 2009) and 1543 (2^{nd} to 8^{th} August 2009) on the South Island available. The stations are OUS2, QUAR and MQZG. The location of these stations is shown in Figure 11.

As described previously, the processing followed the methodology described in Section 3.2 but has been expanded to handle multiple DOYs. The output is used as interface for additional calculations such as statistics or the creation of PWV maps.

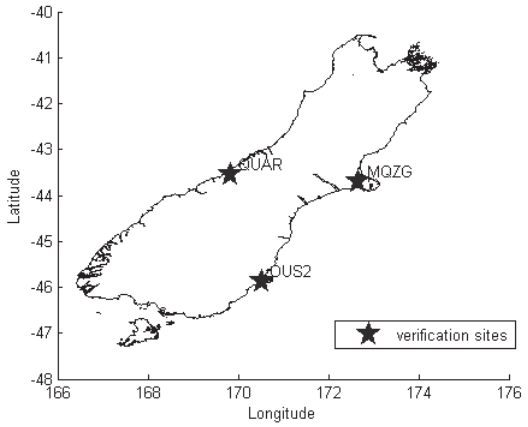


Figure 11: Position of the verification sites on the South Island

At this stage, the process is only partially automated. The data is not accessed directly via the FTP server. Instead the required files for each DOY and station are selected manually and copied to the local hard disk drive. Consequently, the DOYs and verification sites are chosen after a manual pre-processing. Figure A.1 shows the basic flowchart of the processing for one verification site.

The process illustrated in Figure A.1 uses the station names as well as the start and end DOYs of each GPS week as input. With this information, the program selects the associated meteorological and ZTD files from the data directory. This data is imported, transformed, interpolated and structured as described in Section 3.2. In the specified one hour interval, the ZWD and PWV is estimated. Each station and DOY results in an output file. The general view of this whole process is illustrated in Figure 12.

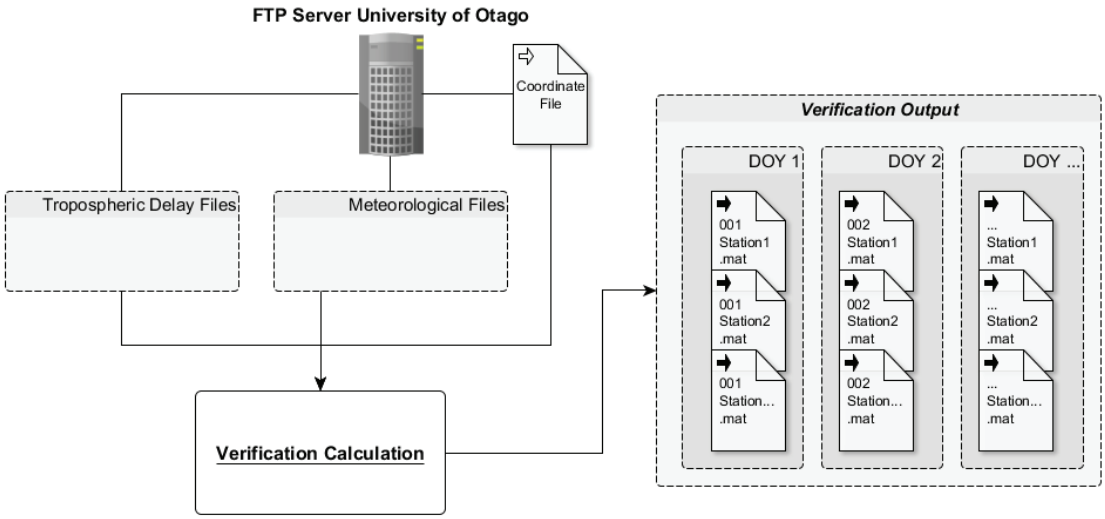


Figure 12: General view creating verification files for the South Island

Table 6 shows a selection of a sample output format of a file:

station	DOY	T (°C)	P (mbar)	ZTD (m)	ZWD (mm)	PWV (mm)
MQZG	32	27.3358	985.991	2.4004	154.25	25.19
MQZG	32.0417	24	986.1	2.4065	160.09	25.93
...						

Table 6: Sample interface/output of verification site MQZG on DOY 32 in 2009

With this method, verification data for three stations in two GPS weeks were computed. The same period is now computed with sparse, publicity available, meteorological data. The general input data files to estimate PWV values at GPS stations with sparse meteorological data are shown in Figure 13.

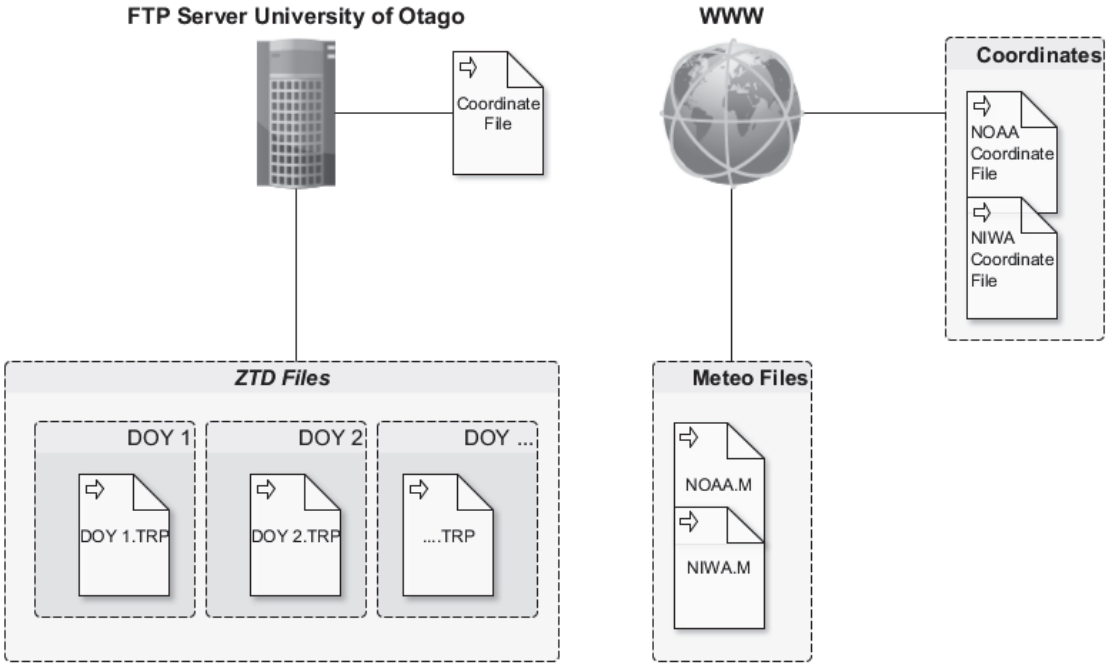


Figure 13: Input data with sparse meteorological data

The objective is to compute ZWD and PWV estimations at each verification site as well as the GPS stations without meteorological data available. The general approach to handle sparse meteorological data is illustrated in Figure 14.

The sparse, publicity available, meteorological data is imported together with the tropospheric delay files from the FTP server of the University of Otago. The meteorological observations in New Zealand were obtained from the publicity available at the "National Oceanic and Atmospheric Administration" (NOAA) and the "National Institute of Water and Atmospheric Research" (NIWA) archives. NOAA manages a world wide meteorological observation network and is a scientific agency within the Department of Commerce of the United States. NIWA offers a national climate database that has access to approximately 600 climate stations in New

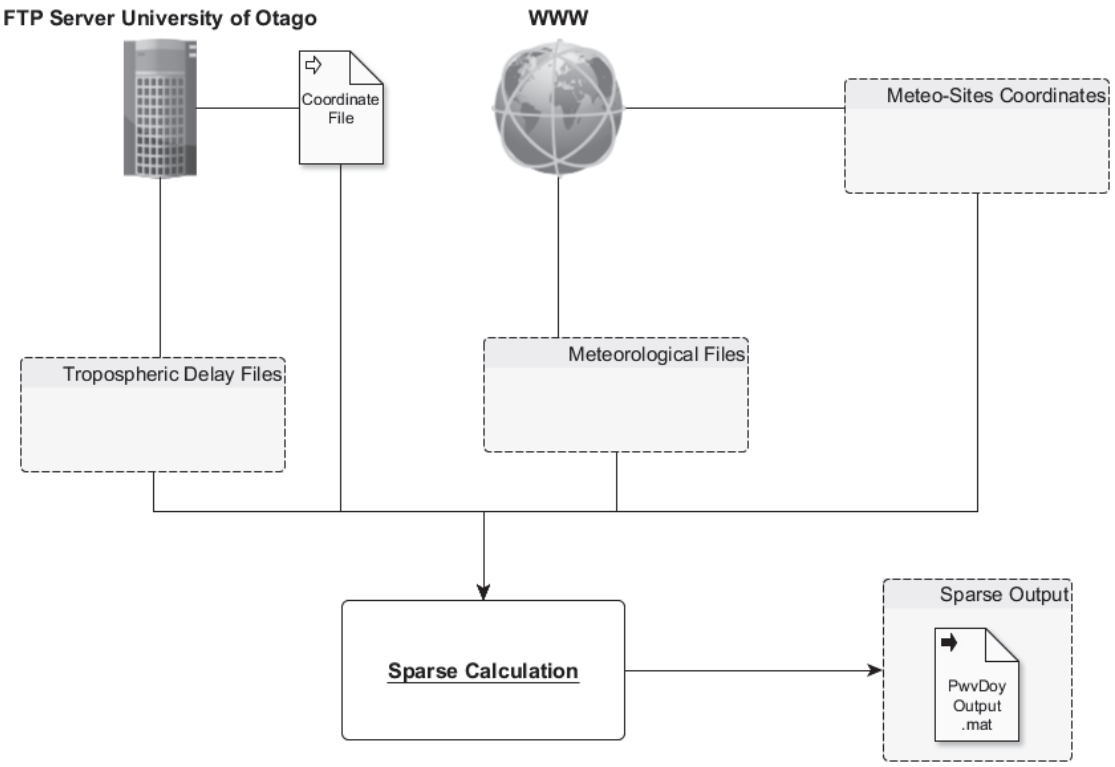


Figure 14: General approach for sparse PWV estimation for the South Island of New Zealand

Zealand. However, for the NIWA archive, the amount of downloadable data rows is restricted to 2,000,000. Hence, for New Zealand, this work concentrates on all available NOAA stations on the South Island of New Zealand and imports two NIWA stations to demonstrate the successful operation with NIWA and NOAA meteorological datasets. Figure 15 displays the NOAA, NIWA and University meteorological sites (dots) used together with the available GPS sites (crosses) processed by the University of Otago. The grey envelope is the extrapolation boundary (convex envelope).

Next, scripts to read the temperature and pressure from the NOAA and NIWA files were implemented. Table 7 and 8 show the general structure of the NOAA and NIWA files:

NOAA Meteorological File					
Name	Date	HrMn	Temp	P_{Sea}	RH
Westport Aerodome	20130729	2100	7.5	1031.7	94
Westport Aerodome	20130730	0000	13.9	1030.8	65
...					
Farewell Spit AWS	20130729	2100	8.7	1027.4	91
Farewell Spit AWS	20130730	0000	13.1	1031.3	93
...					

Table 7: Structure meteorological NOOA file

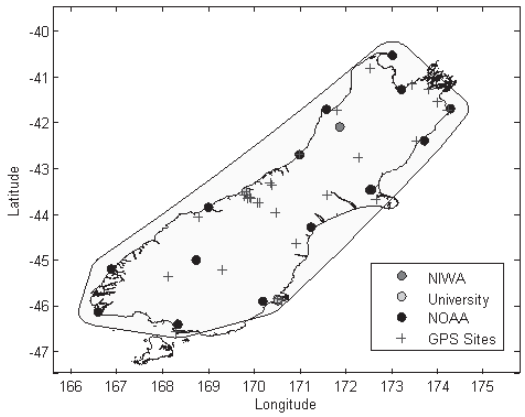


Figure 15: Positions of available meteorological stations (dots) and the GPS sites managed by the University of Otago with available tropospheric delay data (crosses) on the South Island

NIWA Meteorological File			
Name	Date		P_{Sea}
Dunedin Aerodome	20130801:00		1027.7
Reefton Ews	20130801:00		1031.5
...			
Name	Date	T	RH
Dunedin Aerodome	20130801:00	8.0	89.0
Reefton Ews	20130801:00	2.4	93.0
...			

Table 8: Structure meteorological NIWA file

Table 7 (NOAA data) is ordered by the station name and the date. One file contains the whole date-range downloaded from the NOAA website. Thus, it is possible to use only one file for a defined time span. The temperature, sea level pressure (P_{Sea}) and relative humidity are available in a single data row. In contrast, the NIWA data (Table 8) is ordered by the timestamp. As a consequence, the pressure and temperature / relative humidity are not combined in a single data row and have to be assembled during the read-in process.

After importing and structuring all the selected sparse meteorological data that corresponds to the selected troposphere data time period, both the temperature and pressure are interpolated at a one hour interval. Since the NOAA data observations are provided at 3 hour intervals, it is important to ensure that the correct one hour interval, beginning at midnight is selected. In many cases, the midnight observations are only available with the previous data days. Due to this, the interpolation generally begins at 3.00 a.m. to avoid imprecise extrapolations. This issue only occurs on the first day of the data set and for this reason, if available, data from the last observation of the previous day are included (see Table 7), to ensure that data is available at midnight.

3.3.2. Height Extrapolation and Error Analysis

After the observed temperature and pressure values are achieved on joint dates, the spatial (two dimensional) interpolation according to the latitude and longitude positions of the meteorological sites is performed. Certainly, the obtained pressure and temperature data of the NOAA and NIWA sites are measured at their station height (temperature) or are already reduced to sea level (in this case the pressure measurements). For the calculation of the PWV or ZWD values, these meteorological measurements must be known at the height of the GPS site. Given the fact that the pressure and temperature decrease with increasing height, the temperature values are reduced to sea level. Following the spatial interpolation for the PWV and ZWD calculations, the associated values are extrapolated to the elevation of the GPS station. For this step, the following standard lapse rate formulae are used:

$$\frac{\partial P}{\partial H} = -0.0342(\frac{P}{T+273.2})mbar/metre \tag{17}$$

$$\frac{\partial T}{\partial H} = -0.0055°\mathrm{C}/metre \tag{18}$$

Where P is the pressure in mbar, T the temperature in °C and H the elevation in metres. Using these formulae, it is possible to convert the temperature and pressure from sea level to the station height. The received height of the NOAA data is associated with the WGS84 and are ellipsoidal heights. Given the fact that the pressure is already reduced to mean sea level, the geoid undulation also has to be considered. Otherwise the height difference between ellipsoidal and orthometric height (geoid undulation) would incur a bias in the pressure and temperature calculation. For this calculation the MATLAB integrated Earth Gravity Model 1996 (EGM96) is used with an estimated accuracy of 1-2 meters at its maximum (Land Information New Zealand, 2014). By using another model, like the EGM2008 or the nationwide gravity model of New Zealand, this accuracy can be improved. In New Zealand, the geoid undulation range is between 0–40 metres in contrast to the ellipsoidal height of the WGS84. The influence of incorrect heights has been analysed. The investigation will be briefly described by the calculation of systematic errors for the lapse pressure rate and the ZWD formula:

$$P_H = \frac{0.0342 P_{SLN} * H}{T+273.2} + P_{SLN} \tag{19}$$

$$\frac{\partial P_H}{\partial H} = \frac{0.0342 P_{SLN}}{T+273.2} \tag{20}$$

$$ZWD = \frac{-0.0022768 P_H}{1 - 0.00266 cos(2\Phi) - 0.00028h} + ZTD \tag{21}$$

$$\frac{\partial ZWD}{\partial P} = \frac{-0.0022768}{1 - 0.00266 cos(2\Phi) - 0.00028h} \tag{22}$$

Using varying pressure and temperature between 700 mbar, 30 °C and 1025 mbar, -15°C as sample data, a height error of one meter incurs an error in the pressure of between 0.08 and 0.13 mbar. According to the ZWD error propagation, a pressure variance of one mbar leads to an error of around 2.28 mm. With an geoid undulation between approximately -5 and 40 metres in New Zealand and between 15 and 50 metres in Europe, this effect could result in a systematic error in pressure of up to approximately five mbar. This would result in an ZWD error of about 11.4 mm. Therefore it's important to use the orthometric height instead of the ellipsoidal height derived from the GPS coordinates. Figure 16 shows the differences between the verification data of QUAR and the calculated values with ellipsoidal height (dark line) and orthometric height (light line). The geoid undulation is approximately 7.9 metres at this site.

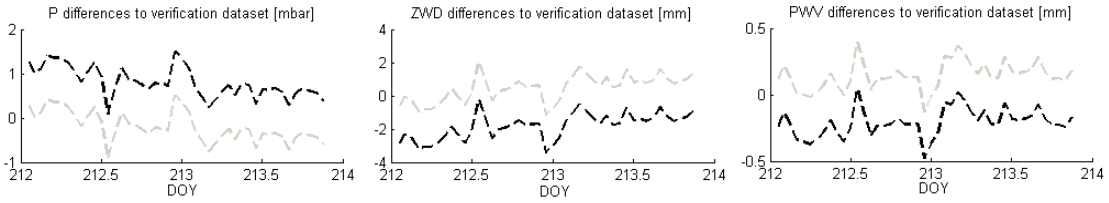


Figure 16: Differences between verification data and calculations with ellipsoidal (black) and orthometric (light) height

The differences shown in Figure 16 demonstrate, that the light line (orthometric height) leads to an smaller error compared to the ellipsoidal height (black line). Hence, the orthometric heights are necessary to come to an accurate result. Note that a negative ZWD difference results in a positive PWV difference.

3.3.3. Spatial Interpolation and PWV Estimation

After reducing the temperature and pressure to sea level, the two dimensional interpolation according to latitude and longitude with a specified grid distance of around 33 kilometres is calculated. To define the borders of the grid, the convex envelope of all used meteorological stations was calculated and a buffer of about 30 kilometres around this envelope is added (see grey boundary in Figure 15). The interpolation generates a grid for temperature and pressure at each interval. Not all meteorological stations provide data during the required period. Due to the fact that the spatial interpolation is computed for each joint date, the convex envelope

varies for each point of time and possibly uncovered GPS stations are handled. The interpolation is performed with the `scatteredInterpolant` class of MATLAB. The two methods used are the natural neighbour (based on voronoi tesselation) for interpolation and nearest neighbour for extrapolation. Though, the extrapolation must be handled separately, the `griddata` (cubic) MATLAB function was also considered. Table 9 shows the Root Mean Square (RMS) values of the temperature, pressure, ZWD and PWV with the interpolated data and the verification stations OUS2 and QUAR.

Station	Method	T(°C)	P(mbar)	ZWD(mm)	PWV(mm)
OUS2	natural-nearest	0.90	0.94	2.15	0.30
	cubic	0.96	0.91	2.09	0.29
QUAR	natural-nearest	2.91	0.39	0.89	0.17
	cubic	3.22	0.37	0.84	0.19

Table 9: RMS of two different interpolation methods

Due to the small number of available verification stations and the short distances of these sites to the closest meteorological station, the results do not differ much. However, in this project, the natural-nearest interpolation method based on the `scatteredInterpolant` MATLAB class is used to interpolate the scattered data. In this case the values within the grid are calculated using the natural neighbour interpolation, which is considered to match the temperature and pressure curves best. Based on the minor extrapolation region of around 30 kilometres, the nearest neighbour method is considered to be the best method for this process. In contrast to the cubic interpolation method, the natural-nearest extrapolation values are automatically calculated. However, to provide evidence on the optimal interpolation method for this type of project, future work on this with more verification stations would need to be done.

The next step involves locating each GPS site within the grid. For this step, the GPS stations were detected by using a inpolygon query. By searching the nearest neighbours of these stations to the closest grid points, the meteorological data is associated with the GPS stations. The last step requires the temperature and pressure at sea level to be extrapolated to the elevation of the GPS sites (equations 17 and 18). Thereby, the data preparation is completed. As a result, the temperature, pressure and ZTD values are available in the specified interval at the GPS sites. Now, the ZWD and PWV are calculated for each site at the time interval required. The process generates one output file, which contains the estimated ZWD and PWV values in the area covered by the meteorological stations. Table 10 shows the output format of each processed GPS week.

PwvDoyOutput.mat			
DOY	Met Cell	Calc Cell	Flag
32.0417	15x11 Cell	66x14 Cell	1
32.0833	15x11 Cell	66x14 Cell	1
...			
38.9583	1x11 Cell	70x8 Cell	0

Table 10: Output format GPS week 1751, sparse PWV estimation, South Island of New Zealand. The Cells represent arrays that contain data of varying types and sizes. They are described in Tables 11 and 12.

The PwvDoyOutput.mat interface consists of four data columns: the DOY, Met Cell, Calc Cell and a Flag attribute. The first column characterise the decimal DOY of the whole GPS week in the specified interval (in this case: one hour interval of GPS week 1517 - DOYs 32-38). No meteorological data was available at midnight of the DOY 32. Due to this, the interval starts at 1.00 a.m. Each interval step is linked with the corresponding Met and Calc Cells. The Flag column indicates if the spatial interpolation and accordingly the PWV has been estimated (1-yes, 0-no). If the interpolation is performed depends on the amount of available meteorological stations. In this project, the required number was simply set to 10, because the spatial interpolation with just a few meteorological stations is not reasonable. On DOY 38.9583 (11.00 p.m.) there is just a single meteorological station available. Hence follows, no spatial interpolation and thus, no PWV estimation has been performed and the Flag is set to 0. The Met Cell consists of 11 data columns. Table 11 shows the structure of the cell.

Met Cell						
station	DOY	T (°C)	P (mbar)	Lat,Lon(rad,deg), H_{ell} (m)	T Flag	P Flag
Westport	32.0417	16.9	1011.9	...	2	1
Farewell	32.0417	22.7	1007.2	...	2	1
...						

Table 11: Met Cell interface of sparse PwvDoyOutput

Every Met Cell describes the one dimensional interpolated meteorological station at the corresponding interval step. It handles the station name, the actual DOY, the temperature in °C, the pressure in mbar and the latitude, longitude and ellipsoidal height of the stations. The last two data columns are flags for the T and P rows. The number 1 indicates that the value is reduced to sea level, whilst the number 2 indicates a measurement at station height. The DOY is stored in the Met Cell interface, to keep the program simple and to use the cell itself as an interface for further purposes. The Met Cell also represents the input data for the spatial interpolation. Using the flags, the sparse procedure transforms the

measurements on station heights to sea level followed by the spatial interpolation. The Cell provides the one dimensional interpolated data from the input files. The second cell within the PwvDoyOutput interface is named Calc Cell.

Calc Cell									
station	DOY	ZTD (m)	pos	T_{Sea} (°C)	P_{Sea} (mbar)	T_{St} (°C)	P_{St} (mbar)	PWV (mm)	ZWD (mm)
AUCK	32.0417	2.3723	...	-	-	-	-	-	-
BLUF	32.0417	2.3206	...	15.1	1009.6	14.43	995.0	8.6	54.7
...									

Table 12: Calc Cell interface of sparse PwvDoyOutput

Its column length depends whether the site has been processed or not. If the site has not been processed, the cell consists of eight columns (station name, DOY, ZTD, [Latitude (radians), Longitude (radians), Latitude (degrees), Longitude (degrees)] and height). The square bracketed term corresponds the summarized pos column. A processed cell is illustrated in Table 12. Its columns are expanded with the temperature and pressure at sea level as well as on station height and the estimated PWV and ZWD values. Because the cell consists of all available GPS stations (both North and South Island), not all rows are filled with actual data (see AUCK-Auckland, North Island, in Table 12). The reason to store all available GPS sites is also to keep the program simple and to use this cell independently as an interface for future usage.

3.4. Adopting the Approach for a Region in Central Europe

As described in Section 3.3, New Zealand provides only a few available verification stations. To increase the amount of verification sites and to gain information about the accuracy of the sparse calculation, the implemented approach from Section 3.3 is now adapted for a region in Central Europe. It extends the implemented functions for any geographic region. This includes a more automatic processing in the matter of selecting stations within the desired geographic region, as well as finding verification sites. For this purpose, ZTD data of the International GNSS Service (IGS) as well as data in Germany(GFZ)[3] are used. To estimate PWV for these stations, the sparse, publicity available, NOAA recordings are used. IGS and GFZ stations, which provide besides GPS measurements also meteorological data, are used to produce verification data. The IGS manages 453 stations world wide of which approximately 389 are active. In the chosen Central European region there are up to 18 stations available. Eleven of the IGS sites can be used as verification sites. A benefit of the IGS data is the publicity availability. The produced IGS

[3] kindly provided by Galina Dick from the GeoForschungsZentrum (GFZ) Potsdam

PWV maps and comparisons to verification sites are solely computed with publicity available data. The GFZ data is not publicity available. However, it includes up to 224 stations in the region from which up to 167 can be used as verification sites. Spatially, the region is approximately limited to 49°–55° North and 6°–22° East. To process different input sources (IGS, GFZ data formats), changes to the read-in process were performed. The output of any read-in process now matches the unprocessed PwvDoyOutput interface as described in Table 10.

Some of the IGS and GFZ stations available on the corresponding FTP servers are outside the defined region. As a consequence, functions to select the GPS and verification sites in the defined region are implemented. During this process, obviously wrong observations (e.g. temperatures above 100 °C or pressure values greater than 1500mbar) are deleted automatically. The major change to the approach described in Section 3.3 is the decoupling of the spatial interpolation from the selecting and formatting segments. For this purpose, the interface described in Table 10 will be used. Consequently, the input sources are of no relevance. The only condition is to fit the defined interface. Furthermore, modifications can be easily made (e.g. interpolation model). The output of the spatial interpolation is the processed PwvDoyOutput.mat as described in Table 10. By making minimal modifications, the approach described in 3.3 can be adopted for any geographic region. The output is used to produce two dimensional PWV time series maps (see Section 4.1) or to compare with verification data (see Section 4.2).

4. Results

The implemented tools produce PWV time series maps and are able to compute statistics (e.g. differences, boxplots and RMS) of the GPS sites with sparse meteorological data at the corresponding verification sites. The decoupled routines are using the PwvDoyOutput interface. For this project, two regions with two GPS weeks each, were chosen. In order to cover a summer as well as a winter period, the observation weeks are separated by a 26 weeks time period. Two GPS weeks in 2009 were processed for the South Island of New Zealand using the ZTD data of up to 28 GPS stations managed by the University of Otago. The year 2009 has been chosen because this period provides the most verification sites (three) in comparison to other periods. In addition, the available GPS stations cover almost the whole area of the South Island. This allows good spatial coverage for the produced PWV maps. The sites provide a transverse profile crossing the Southern Alps (West Coast) with small inter-site distances and with height differences of up to 2,250 metres. The characteristics of the South Island are also very diverse. Generally, the region has a mountainous character with valleys. The weather varies not only regionally but also during the day. It is close to the cold air reservoir Antarctica and allows short-term coldness advances. The East and West Coast of the South Island differ significantly due to the predominantly South-West air flow that crosses the Southern Alps. Westerly air is forced to rise over the Southern Alps, which creates heavy rainfall on the West Coast. This in turn creates hot, drying winds on the East Coast.

With the IGS and GFZ data, another two GPS weeks in 2013 were processed. The year 2013 has been chosen to cover the latest possible week in summer as well as a week in winter. The region in Central Europe, spatially limited by 49°–55° North and 6°–22° East, has been selected due to the highest density of IGS verification sites in the world as well as the availability of the GFZ stations. Climatically the region is characterized by a mild climate with no significant mountainous regions in the north of the project area.

The PWV time series maps and temperature, pressure, ZWD and PWV estimations for both regions are included on the attached CD-ROM.

4.1. PWV Maps

A summary of the meteorological sites and distances for the two weeks of the South Island and the two weeks of the Central Europe data is given in Table 13.

Region	GPS data origin	GPS week	Max GPS sites	Max meteo sites	Avg distance GPS to meteo sites (km)
NZ	UO	1517	28	15	45.8
		1543	27	15	49.3
EU	IGS	1725	19	18	166.6
		1751	18	18	144.4
	GFZ	1725	239	18	97.6
		1751	224	18	204.4

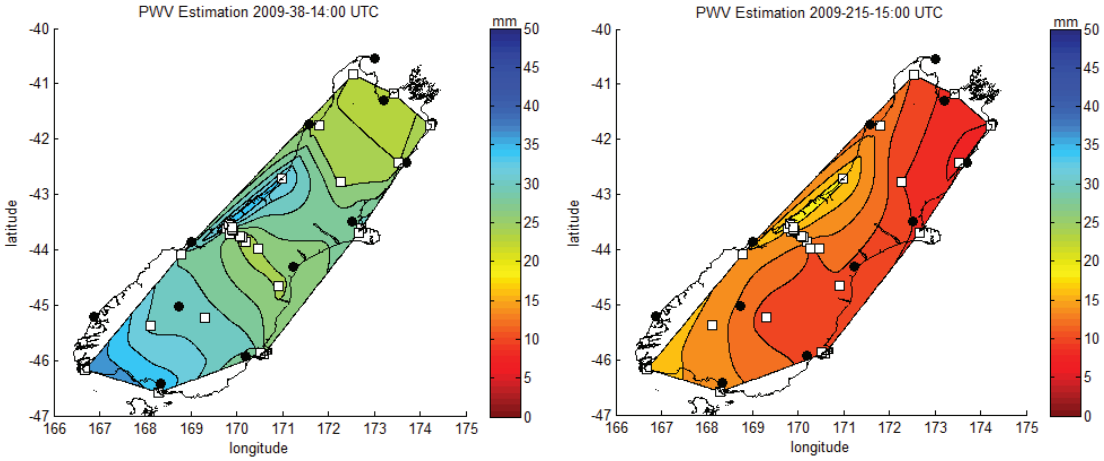
Table 13: Characteristics of input datasets with sparse meteorological data

For the South Island, the GPS weeks 1517 and 1543 include the dates 1^{st} to 7^{th} February and 2^{nd} to 8^{th} August 2009 respectively. The GPS data source is the University of Otago (UO). During these periods, a maximum of 28 GPS and 15 meteorological sites were available. The average distance from a GPS station to the nearest meteorological station is approximately 45 to 49km. The European datasets for the GPS weeks 1725 and 1751 include the dates from 27^{th} January to 2^{nd} February and 28^{th} July to 3^{rd} August 2013 respectively. Each week is processed twice with either IGS or GFZ ZTD data. IGS provides a maximum of 19 GPS sites, whereas the GFZ dataset provides 239 sites in this region. The maximum available meteorological stations is 18. The quantity does not differ for either datasets since the same meteorological input data is used for the IGS and GFZ PWV estimation. However, the average distance from IGS GPS stations to a meteorological site varies between approximately 166 and 144km, whereas the average distance for the GFZ sites to the closest meteorological station is approximately 97 and 204km. Notably, the distance increases dramatically in GPS week 1751 for the GFZ dataset. This is attributed to the removal of six meteorological stations for the last three processed DOYs and the unfavourable geographical distribution of those sites. The region with the most GPS stations (Germany) is only covered by five meteorological sites (see Figure 19). This results in the high average distance (Table 13).

The computation of PWV maps will be implemented with the Calc Cell interface of the PwvDoyOutput data as described in Table 12. A spatial interpolation of all GPS sites with available PWV data produces PWV time series maps.

4.1.1. New Zealand

The two GPS weeks 1517 and 1543 as described in Table 13 were processed to compute two dimensional PWV time series maps for the South Island of New Zealand. Figure 17 shows a sample of the calculated PWV maps in New Zealand:



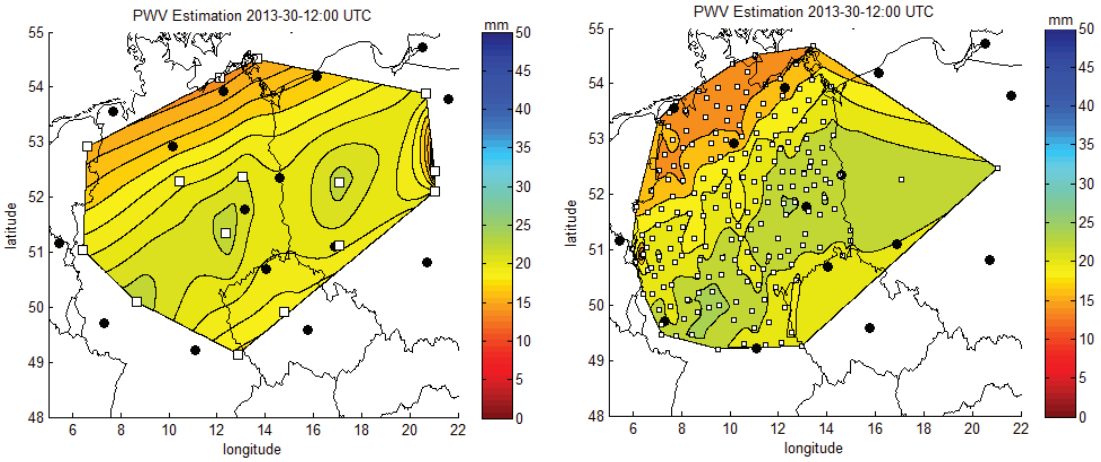
(a) Summer period (b) Winter period

Figure 17: PWV maps South Island, GPS weeks 1517 (DOY 38, 14:00UTC) and 1543 (DOY 215, 15:00UTC). The white squares represent GPS and the black dots meteorological sites.

The area covered by GPS stations accounts for approximately 140,000km^2 with a mean GPS station density of one station per 5,000km^2. Figure 17a shows a sample of the summer period (GPS week 1517). Figure 17b illustrates a sample of the winter period (GPS week 1543). Figure 17 shows that higher PWV values are estimated in summer than during the winter period. The maps also show a higher amount of PWV at higher elevations (transverse profile at West Coast in the Southern Alps). The transverse profile also shows less PWV in the atmosphere East of the Southern Alps. This is assumed to be caused by the formation of clouds that encounter the mountains. Precipitation is then released as rain in order to loose weight and hence rise to cross the Southern Alps. The atmosphere therefore contains less water vapour than on the West and the PWV values have fallen noticeable. Regarding the whole time series for each week, a wetter period (higher PWV values) usually begins either in the West/South-West or in the North. This is mostly caused by westerly winds with associated clouds. Although the summer period begins with PWV estimations between 0 and 10mm, the summer period is at most of the time wetter (PWV >40mm) than the winter period (PWV between 0 and 15mm). The implemented tools create evaluable PWV maps with a spatial coverage of almost the whole South Island.

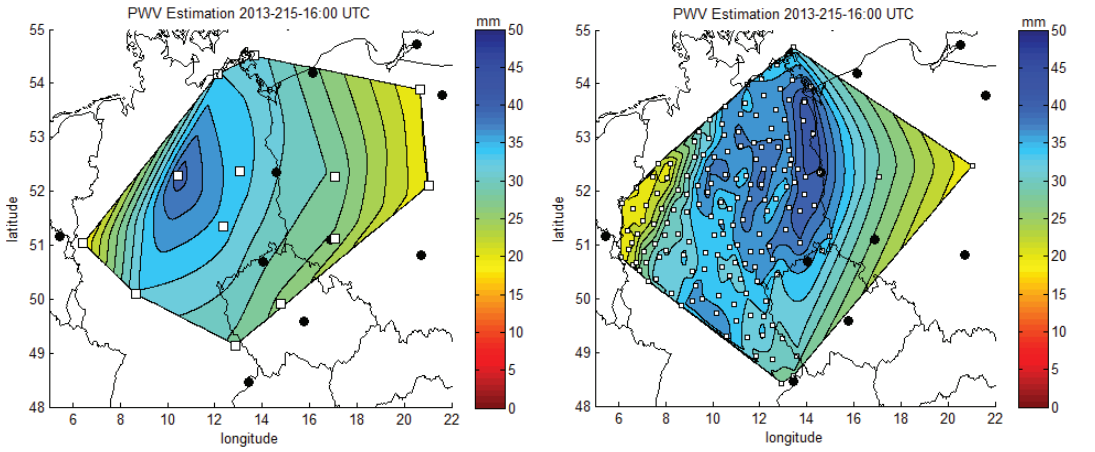
4.1.2. Europe

This Section shows a sample of the created PWV maps of the two GPS weeks with either IGS or GFZ data used. Figures 18 and 19 show the processed PWV maps for GPS weeks 1725 and 1751 using the IGS (Figures 18a and 19a) and GFZ (Figures 18b and 19b) GPS data respectively.



(a) IGS stations (b) GFZ stations

Figure 18: PWV maps winter period IGS and GFZ datasets Central Europe, GPS week 1725 (DOY 30, 12:00UTC). The white squares represent GPS and the black dots the meteorological sites.



(a) IGS stations (b) GFZ stations

Figure 19: PWV maps summer period IGS and GFZ datasets Central Europe, GPS week 1751 (DOY 215, 16:00UTC). The white squares represent GPS and the black dots the meteorological sites.

The average area covered by the IGS GPS stations for both time periods (GPS weeks 1725 and 1751) is approximately 390,000km^2. The mean density of these stations is almost one station per 20,000km^2. In contrast, the average area covered

by GFZ GPS stations is approximately 510,000km^2 (GPS week 1725). This results in a density of generally one station per 2,100km^2. Although missing meteorological stations induce a smaller covered region for the last three DOYs in GPS week 1751, the area covered by GPS stations is approximately 530,000km^2. This occurred because of the discontinuation of the most easterly GPS site in GPS week 1725. However, on the basis of less available GPS stations, the corresponding mean GPS density is roughly one station per 2,300km^2.

Figures 18a and 19a contain a sample from the IGS PWV maps. Figures 18b and 19b show samples of the GFZ PWV maps. Figure 18 demonstrates the winter period (GPS week 1725), whilst Figure 19 shows the summer period (GPS week 1751). The Figures illustrate the same time of the day (12.00 UTC and 16.00 UTC). Similar to Section 4.1.1, higher PWV values were estimated in the summer than in the winter period (note that the time of the summer and winter seasons are reversed in the Northern (Europe) compared to the Southern (New Zealand) hemisphere. Regarding the whole time series, the summer period is continuously wetter than the winter period. The time series also show that water vapour (expected as rain) comes from the West. This corresponds with the low-pressure cyclones coming from Iceland over Central Europe. As the clouds pass by, they are raining down and the amount of PWV decreases over time. The significantly higher station density of GFZ stations in contrast to the IGS density provides better accuracy of the PWV maps and detailed regional characteristics become recognizable. Figure 19 also demonstrate the discontinuation of the meteorological stations in GPS week 1751. The resulting high distances of the GPS stations to the closest meteorological station is obvious. It is logical to assume that the higher distance affects the accuracy of the PWV estimations. In order to check this assumption and to obtain an understanding of the precision of the estimated PWV values, Section 4.2 compares the results produced with aid of the sparse meteorological stations with the verification sites.

4.2. Comparing the Estimated Data to Verification Stations

This Section compares the computed PWV estimations using sparse, publicity available meteorological data to the verification stations. For this purpose, the described PwvDoyOutput interface in Section 3.3.3 is used. Differences for each available station in each processed GPS week are calculated using equation 23.

$$x_{Diff} = x_{Spar} - x_{Ver} \tag{23}$$

Where x_{Spar} is the sparse calculated value and x_{Ver} the verification data. To denote how stretched the distribution of the differences between verification and estimated PWV values are, the corresponding boxplots are computed. This includes the median, quartiles, whiskers and outliers. The median $\tilde{x}$ is calculated with the ranked set of data values by using the following equation:

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{n even number} \\ \frac{1}{2}(x_{n/2} + x_{n/2+1}) & \text{n uneven number} \end{cases}$$

The lower and upper quartiles can be calculated from the ranked data set by using the $(n+1)/4^{th}$ and $3(n+1)/4^{th}$ values. If the data is normally distributed, the whiskers correspond to approximately $\pm 2.7\sigma$ or 99.3% coverage. A whisker extends to the most extreme data value that is not an outlier. To provide summarized information between the estimated and verification PWV sites, the Root Mean Square (RMS) will be computed for each available station in each week. The RMS is calculated using equation 24.

$$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x_t)^2} \tag{24}$$

Where n is the number of measurements, x_i is the estimated value and x_t the reference data at the verification site. The variables x_i and x_t are equivalent to x_{Spar} and x_{Ver} from equation 23.

Because the meteorological data is not measured at the GPS station, the estimated data will be imprecise. A high distance to the closest meteorological stations will lead to inaccuracies of the interpolated temperature and pressure. Also, a higher GPS station elevation induces more imprecise temperature and pressure extrapolations because of the lapse rate equations used (equations 17 and 18).

4.2.1. New Zealand

During this project, PWV has been estimated for the South Island of New Zealand. For the GPS weeks 1517 and 1543, the verification stations OUS2, QUAR and MQZG are available. Hence, only these three stations with available reference data can be compared to the estimated values. The meteorological verification data is provided with measurements at the same position and height as the GPS site. Table 14 tabulates the distance to the closest meteorological station that is used for the interpolated surface meteorological data and the height of the GPS station.

GPS station	Distance closest Met site (km)	Ellipsoidal height (m)
OUS2	3.54	26.12
QUAR	75.44	58.03
MQZG	26.83	154.68

Table 14: Characteristics verification sites New Zealand

With a distance of 3.54km to the closest meteorological site and a station height of 26.12 metres, the GPS site OUS2 has the fewest disadvantages with respect to the verification data. As explained in Section 3.3.2, the ZWD and PWV strongly depend on the pressure and, of course, the ZTD observations. Due to the fact that the ZTD is the same for both, verification and estimation data, the most important indicator for accurate estimations are precise pressure values, which in turn depend on the station height.

The Figures 20 and 22 illustrate the temperature, pressure, ZWD and PWV differences between the reference and estimation data (from equation 23). In order to be able to state assumptions about the station precisions, these differences are used to compute the corresponding boxplots for the verification stations OUS2, QUAR and MQZG (Figures 21 and 23). The illustrated boxplots are containing the median, quartiles, whiskers and outliers (see Section 4.2). The rectangular box for each station represents the upper and lower quartiles. This box contains the central 50% of the data. The vertical (dashed) lines with another horizontal line at its ends are the whiskers. The crosses represent data outliers. The horizontal line within the quartile rectangle is the median. It divides the diagram in two halves from which each halve (including the whiskers and outliers) consists of 50% of the whole data.

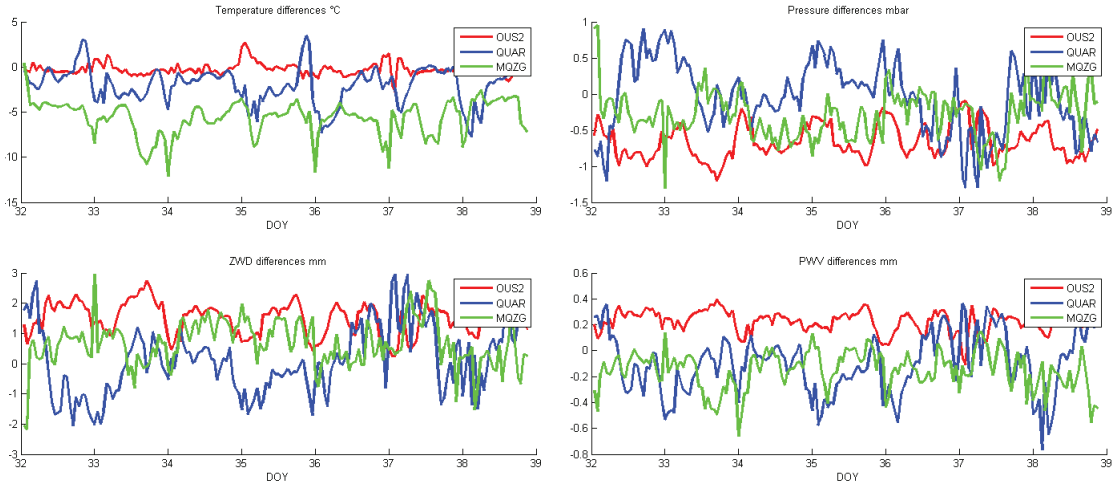


Figure 20: Differences estimated and reference data, South Island GPS week 1517

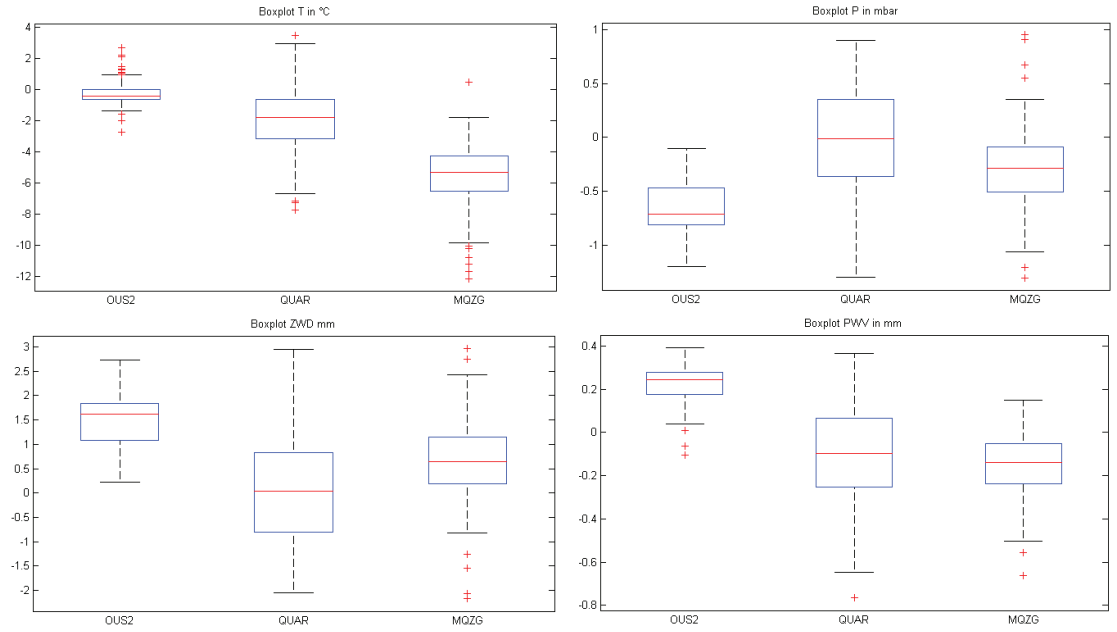


Figure 21: Temperature, pressure, ZWD and PWV boxplots, GPS week 1517

The temperature differences in Figure 20 (top left corner) show that the differences range approximately between +4 and -12°C. It becomes evident that OUS2 has the most accurate temperature interpolation (close to zero for all DOYs) and the lowest fluctuations. This is most likely caused by the small distance to the closest meteorological station. The temperature lows that can be easily seen on DOYs 34, 36 and 37 at the MQZG site are caused by the meteorological observations of the verification data. At these days, measurements are missing some hours before and after midnight. This results in a more inexact temperature interpolation at this site. Taking Table 14 into account, the figure also indicate that the interpolated data may be less reliable as the station height increases. However, this hypothesis must be verified with a larger sample size and further research.

This is also reflected in the top left corner (temperature boxes) of Figure 21. OUS2 has a very small box in comparison to the sites QUAR and MQZG. This means that the average 50% values are not subject to huge variances. Even though some outliers exist, the data range is smaller than those of the other stations. Visually, this effect is recognizable with bigger boxes as well as longer whiskers.

Regarding the pressure differences and boxplots in Figures 20 and 21, the interpolated values at the GPS sites using sparse meteorological data, are more reliable than the temperature interpolations. The differences range between approximately +1 and -1.4mbar. Even though OUS2 has the smallest distance to the closest meteorological station, the pressure differences show similar characteristics to the QUAR and MQZG sites. One possibility would be that the NOAA pressure observations are extensively interpolated and are no actual measurements at the station. Another possibility is, of course, wrong measurements of either the verification or the closest sparse meteorological stations. However, the pressure precision of OUS2 is the smallest, but the median offset is greatest. It is notable that there are fluctuations in the pressure differences at the QUAR site around the DOY 37. This effect can generally be traced back to unsteady pressure observations of the NOAA stations.

The ZWD and PWV differences indicate that a negative pressure difference results in a positive ZWD and PWV difference. Hence, the higher the pressure differences, the higher the ZWD and PWV differences. According to equation 23, a positive difference leads to a higher interpolated value than the verification data. Similarly, a lower estimated pressure observation at the GPS station results in higher ZWD and consequently higher PWV estimations. Regarding the PWV differences and boxplots, it is evident that the PWV values are only minimally affected by the temperature observations (equation 12). This can be seen at the OUS2 and MQZG boxes (Figure 21). Although the ZWD boxes are approximately the same size (depending on the pressure), the PWV box (depending on ZWD and temperature) of OUS2 is slightly smaller than the MQZG one, because of the small temperature differences of OUS2. As a consequence, only high temperature differences have a significant impact to the estimated PWV.

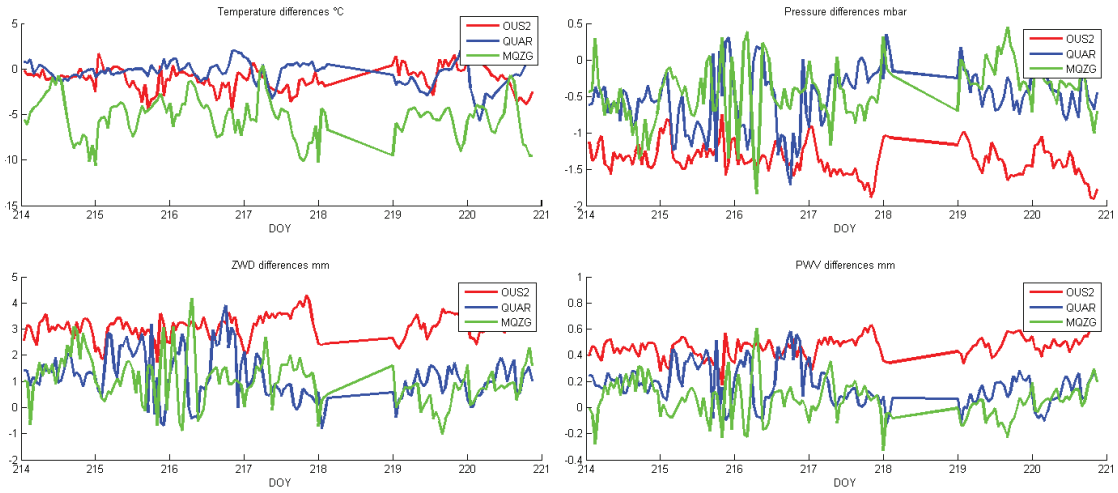


Figure 22: Differences estimated and reference data, South Island GPS week 1543

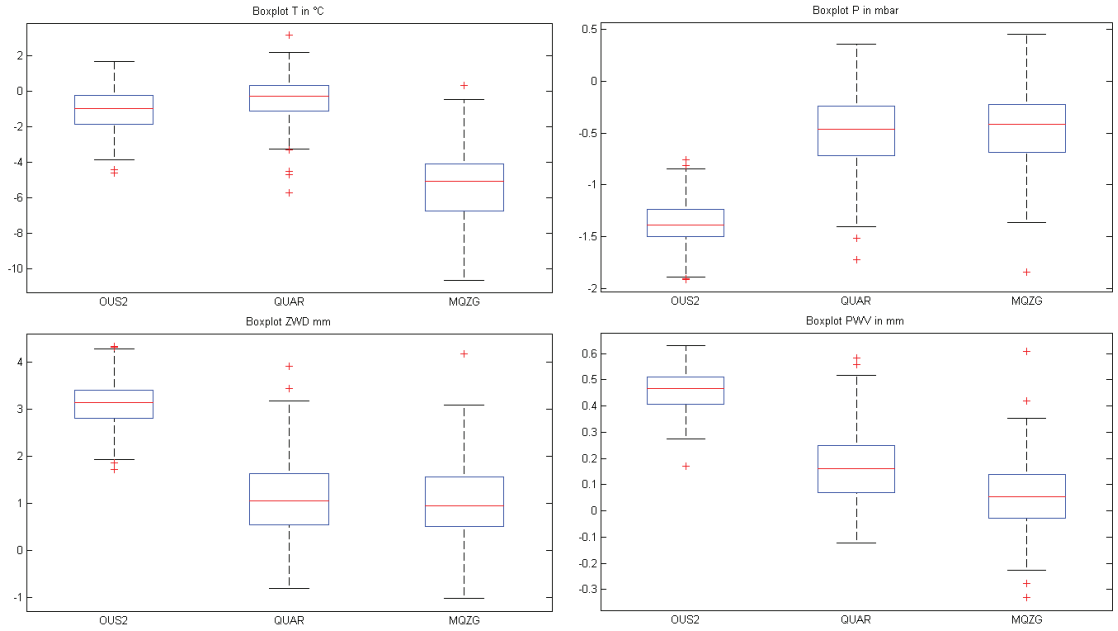


Figure 23: Temperature, pressure, ZWD and PWV boxplots, GPS week 1543

Regarding the differences in GPS week 1543 (Figure 22), the effect of missing data becomes evident for DOY 218. This is caused by insufficient ZTD observations. In fact, DOY 218 consists of only the first three hours of the day. The software tools are designed to process the delay data by day. Hence, the whole day is skipped if too few observations are available throughout the day. In general, the differences and boxplots (Figures 22 and 23) show the same characteristics like the processed GPS week 1517. The temperature differences ranges from approximately +3 to -11°C and the pressure from +0.5 to -2mbar. However, it becomes evident that the OUS2 temperature precision has decreased in contrast to the previous week. It is notable that there are strong fluctuations in the pressure differences at the MQZG site around the DOY 216. This effect becomes also apparent at

the QUAR site and partially even at the OUS2 site and can be traced back to unsteady pressure observations of the NOAA stations.

In general, the differences are also caused by the three hour observation intervals of the NOAA stations. This leads to imprecise interpolated data in contrast to the verification data that is usually available in a five minute interval. Another error source is the pressure and temperature lapse rate extrapolation (equations 17 and 18). The usage of these equations leads to a more and more inaccurate extrapolation with increasing heights, because the formulae only implement a simple model and do not represent the real values at the station height.

The pressure differences in Figures 21 and 23 show a greater distribution of QUAR and MQZG, even though the median is closer to zero than for OUS2. The bigger boxes and longer whiskers may be explained by the greater distances to the closest meteorological station. Despite the very low distance to the closest meteorological station, OUS2 has a constantly higher offset. This all lends weight to the argument that the estimation or verification pressure data at OUS2 is affected by a systematic error.

Table 15 tabulates the RMS of the difference between the estimated and verification data for each station (equation 24).

Station	Distance (km)	GPS week	RMS T (°C)	P (mbar)	ZWD (mm)	PWV (mm)
OUS2	3.54	1517	0.74	0.69	1.59	0.23
		1543	1.69	1.38	3.15	0.46
QUAR	75.44	1517	2.83	0.49	1.13	0.25
		1543	1.43	0.64	1.45	0.22
MQZG	26.83	1517	5.99	0.44	1.01	0.21
		1543	5.77	0.58	1.34	0.15

Table 15: RMS of the difference data at OUS2, QUAR and MQZG

The RMS values allow an assessment of the accuracy of the estimated data using sparse available data. For the two GPS weeks 1517 and 1543, a reasonable PWV precision of better than ±0.5mm is achieved for the three available verification stations on the South Island of New Zealand.

4.2.2. Europe

Estimates of PWV have been made in Central Europe using either IGS or GFZ stations. The data described in Section 4.1.2 is now compared to reference data at each of the verification stations. In contrast to the number of available verification stations in New Zealand, the European datasets provide significantly more verification data. For the GPS weeks 1725 and 1751, four datasets have been produced. One for each processed GPS week by using the publicity available IGS datasets and another one per week by using the provided GFZ stations. Table 16 tabulates the characteristics of the verification stations.

GPS data origin	GPS week	Max verification stations	⊘ Distance to closest met station (km)
IGS	1725	10	85.50
	1751	11	88.60
GFZ	1725	210	95.08
	1751	197	139.76

Table 16: Characteristics of IGS and GFZ verification stations

Contrary to the maximum available GPS stations from Table 13, only ten or eleven of the 18 and 19 IGS stations and 210 or 197 of the 239 or 224 GFZ stations can be used as verification stations. The ⊘ column (Table 16) specifies the mean distance to the closest meteorological station of the GPS stations. Even though the same meteorological data is used for sparse IGS as well as GFZ PWV estimations, the distance of the GPS stations to the closest meteorological site in GPS week 1751 is higher than those of the IGS stations for the same week. Again, this effect is attributable to the missing meteorological stations as described in Section 4.1.

Figures B.1-B.4 in Appendix B plot the temperature, pressure, ZWD and PWV differences between the interpolated and reference data for both IGS and GFZ verification stations. In view of the large number of stations, these figures can not show the precision of single stations. Instead, they provide a general overview of all GPS sites.

Regarding the IGS stations, Figure 24 shows the boxplots diagrammatically for GPS week 1725. They give a good overview of the accuracy of the available GPS sites for the temperature, pressure, ZWD and PWV values for each verification station (Table 16).

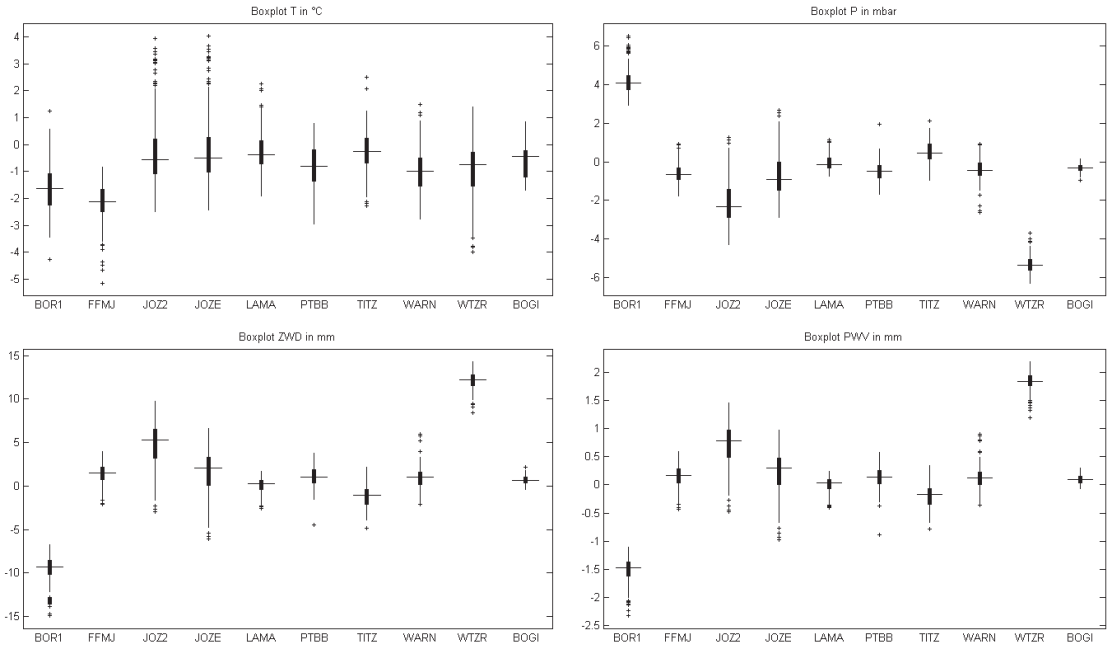


Figure 24: Boxplots IGS stations, GPS week 1725

For the temperatures, the distribution is generally higher in contrast to the pressure, ZWD or PWV boxplots. The stations JOZE and JOZ2 are very close together, which causes the almost identical boxes. However, this effect is not seen in the pressure chart, even though both stations are also almost on the same elevation. It is possible that one of the pressure gauges is measuring inaccurately. Another possibility is that there is an height offset between the pressure gauges, which leads to the offset seen in the Figure 24. Regarding the ZWD and PWV charts, it becomes apparent that the values of most of the stations are approximately zero. This implies that most of the estimates are reasonable accurate. The stations BOR1 and WTZR are the only ones with high differences. However, this may be caused by incorrect meteorological measurements of either the verification or sparse estimated data.

The visualized PWV chart peak on the DOY 210 of WTZR (line at the very bottom) in Figure B.2 (Appendix B) is caused by the abrupt temperature change which can be seen in the temperature chart. While steps have been taken to avoid wrong input data, the temperature peak demonstrates the effect of incorrect measurements at the reference data for this DOY.

Station	T (°C)	P (mbar)	ZWD (mm)	PWV (mm)	Distance closest met site (km)	Ell. height (m)
BOGI	0.93	0.41	0.93	0.13	86.20	138.90
LAMA	0.82	0.38	0.86	0.14	64.97	187.04
WARN	1.21	0.71	1.63	0.24	31.62	50.76
FFMJ	2.27	0.77	1.76	0.24	105.73	178.23
PTBB	1.11	0.71	1.61	0.25	72.77	130.25
TITZ	0.80	0.71	1.75	0.30	68.93	156.18
JOZE	1.44	0.77	2.91	0.44	89.34	141.43
JOZ2	1.44	2.36	5.37	0.81	89.28	152.53
BOR1	1.87	4.27	9.73	1.54	133.23	124.36
WTZR	1.53	5.32	12.13	1.83	112.95	666.03
Mean RMS	1.34	1.64	3.87	0.59		

Table 17: IGS station RMS values in winter period (GPS week 1725)

With respect to the boxplots, BOGI shows the smallest deviations. This is reflected by the fact that the site has the smallest RMS (Table 17). Table 17 also shows that most of the stations accorded well with the reference data. A mean PWV RMS of ±0.59mm has been achieved for the 10 stations for GPS week 1725. Figure 25 shows the boxplots for GPS week 1751.

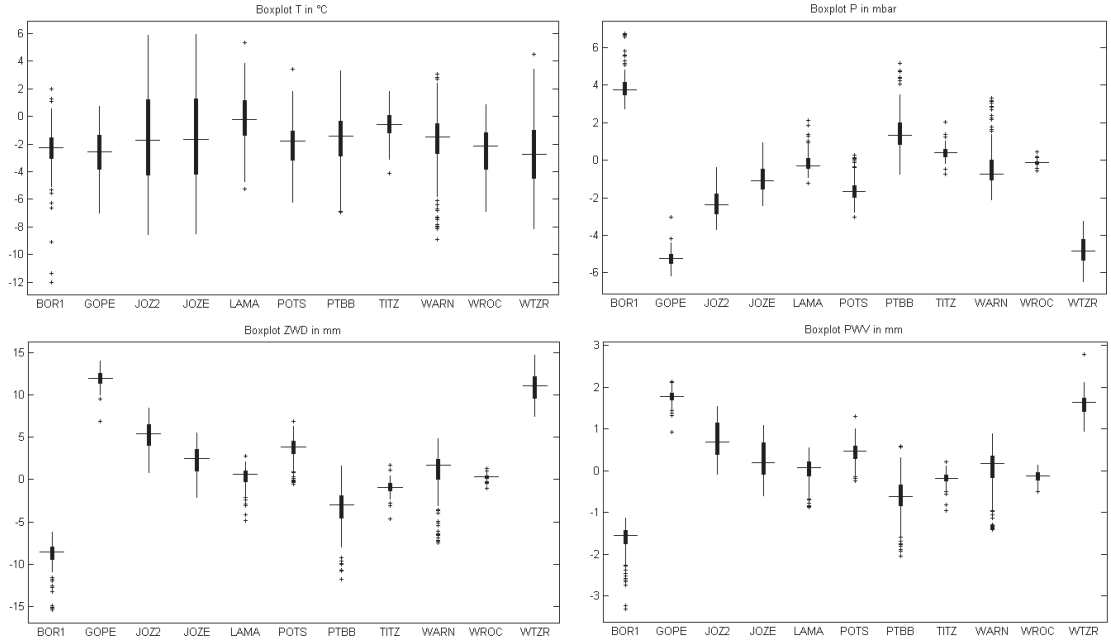


Figure 25: Boxplots IGS stations, GPS week 1751

The data consistently shows that the sites in both weeks have similar characteristics. The sites with larger differences and only a few outliers like GOPE (Figure 25) or WTZR (Figures 24 and 25), indicate a systematic error of the input data.

Station	T (°C)	P (mbar)	ZWD (mm)	PWV (mm)	Distance closest met site (km)	Ell. height (m)
GOPE	3.14	5.24	11.95	1.78	79.27	592.61
POTS	2.60	1.69	3.84	0.50	84.07	144.43
WROC	2.87	0.18	0.42	0.18	12.59	180.82
LAMA	1.72	0.51	1.16	0.28	61.13	187.04
WARN	2.87	1.26	2.87	0.54	117.46	50.76
PTBB	2.59	1.82	4.13	0.80	163.27	130.25
TITZ	1.13	0.51	1.16	0.24	68.93	156.18
JOZE	3.83	1.25	2.84	0.51	84.81	141.43
JOZ2	3.86	2.44	5.55	0.85	84.74	152.53
BOR1	2.96	3.99	9.08	1.69	131.53	124.36
WTZR	3.76	4.80	10.96	1.61	86.79	666.03
Mean RMS	2.85	2.15	4.91	0.82		

Table 18: IGS station RMS values in summer period (GPS week 1751)

The results demonstrate that the BOGI (Figure 24) and WROC (Figure 25) PWV estimations fit best for the corresponding GPS weeks with the verification data. It is also shown that BOR1, GOPE and WTZR are outliers with an PWV difference of more than 1.5mm (see RMS Tables 17 and 18). Regarding WARN, PTBB and even WTZR, Tables 17 and 18 demonstrate that increasing distances to the closest meteorological station of the same GPS station generally results in less precise PWV estimations. It can also be noted that stations at a high elevation (500m or more) have the highest differences. This highlights possible problems with the lapse extrapolation equations. It is also shown that the mean RMS values are greater in summer than in the winter period. However, ten out of twelve IGS stations have reasonable results where the PWV RMS of ±1mm or less has been achieved.

The GFZ data consists of more than 200 verification stations (Table 16). Consequently, an individual analysis is only beneficial for a few noticeable stations. Instead, the whole dataset is regarded by boxplots, differences, RMS and the visualized positions of inaccurate stations. The corresponding difference charts can be found in Appendix B. The Figures B.3 and B.4 show similar results compared to the IGS stations. Most of the meteorological interpolated temperature and pressure differences as well as the estimated ZWD and PWV values vary around zero. However, the pressure differences (Figures B.3 and B.4) and boxplots (Figures 26 and 27) show some noticeable outliers.

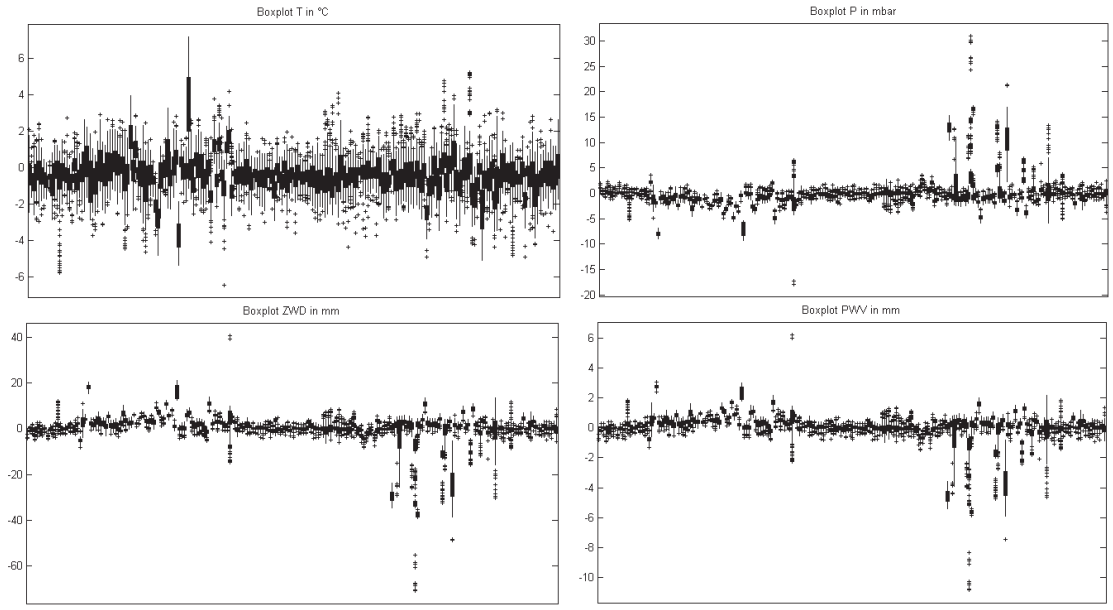


Figure 26: Boxplots GFZ stations, GPS week 1725

There are some stations with a constant pressure offset and peaks during the GPS weeks. Some stations have very high pressure differences at the beginnings of the GPS weeks (reflected in the corresponding boxplot charts as outliers). This might be caused by wrong time-related pressure interpolations. An inspection of the input data and the fact that this effect occurs frequently within the chosen period leads to the conclusion of incorrect input data of the meteorological stations. An inspection of the sites with the highest pressure differences showed that most of the stations are near to one meteorological station concentrated in the west (see 1mm maps below). Another station has a zigzag curve in the pressure, ZWD and PWV difference charts. This occurs for both GPS weeks for the same station. The amplitude is between $\pm$2mbar and changes almost each interval step (one hour). It is unlikely caused by an incorrect spatial interpolation since the effect only appears at this particular station. Instead, it is assumed that the time related interpolation or, more likely, the meteorological data is wrong. However, the processed GFZ dataset achieves a reasonable accuracy. Given the large number of

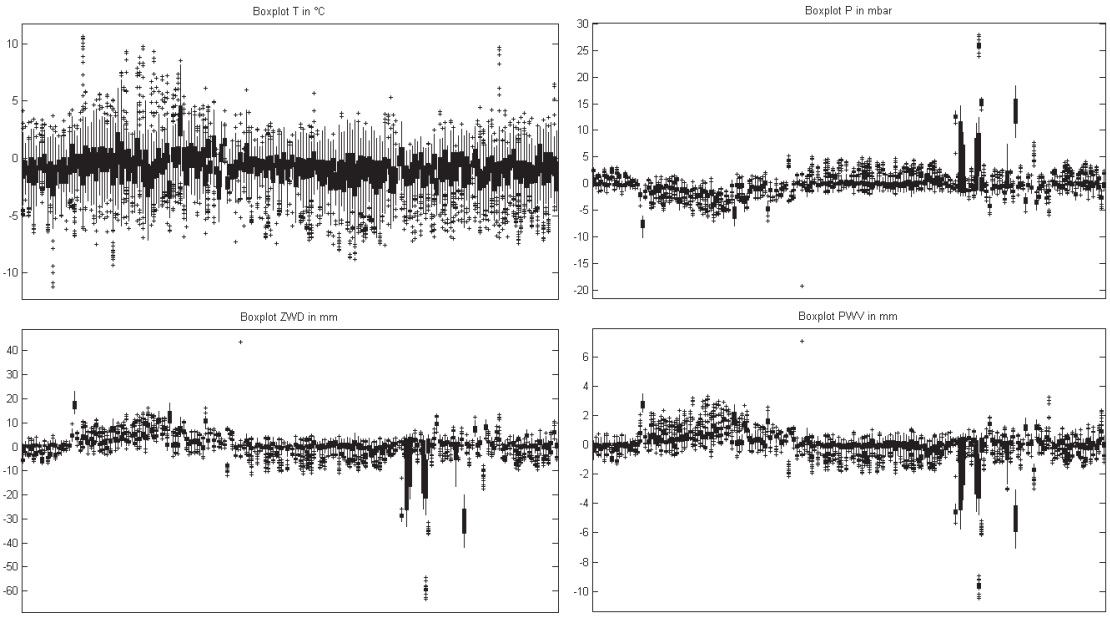


Figure 27: Boxplots GFZ stations, GPS week 1751

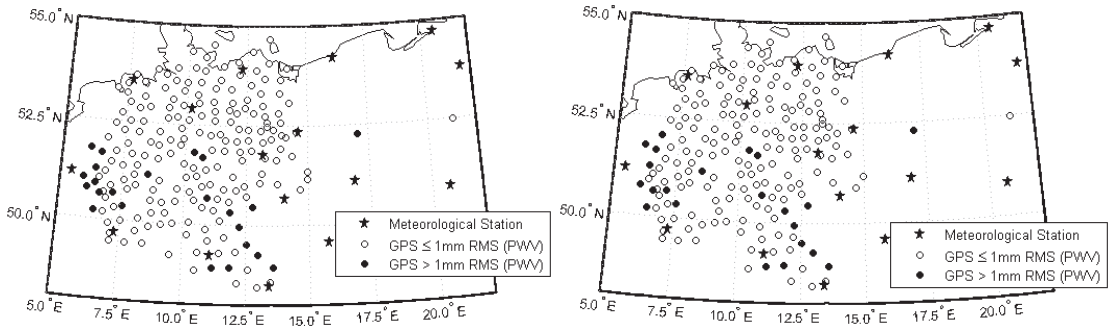
GFZ verification stations, a table of RMS values is not sensible. Instead, Table 19 summarizes those sites with a PWV RMS smaller than 0.25mm, between 0.25mm and 0.5mm, 0.5mm and 1mm as well as 1mm and above.

GPS week	PWV RMS (mm)	Number of stations	% Distribution
1725	0≤0.25	99	47.2
	0.25≤0.5	57	27.1
	0.5≤1	31	14.8
	>1	23	10.1
1751	0≤0.25	22	11.1
	0.25≤0.5	105	53.3
	0.5≤1	48	24.4
	>1	22	11.2

Table 19: Summarized PWV RMS GPS weeks 1725, 1751 of GFZ stations

Table 19 shows that most of the stations for GPS week 1725 have a PWV RMS between 0 and 0.25mm (47.2%), whilst for GPS week 1751 the RMS of the majority is between 0.25 and 0.5mm (53.3%). This is reflected by the mean PWV RMS of ±0.45mm for GPS week 1725 and ±0.59mm for GPS week 1751. Considering the percentage distribution, it is noticeable that the 0 to 0.25mm and the 0.25 to 0.5mm RMS distribution for GPS week 1725 shifted to the next classification interval for GPS week 1751. This is assumed to be primarily caused by the higher mean distance to the closest meteorological station (see Table 16). A higher average distance affects the accuracy of the meteorological interpolations. However, this could also be impacted by the summer/ winter period with generally higher or lower PWV values. Apart from this, it is noticeable that the amount of

stations with a PWV RMS greater than ±1mm only differs by one single station. Figure 28 shows more detailed information about the spatial distribution of the stations with a PWV RMS greater than ±1mm where positions of the GPS and meteorological stations are plotted.

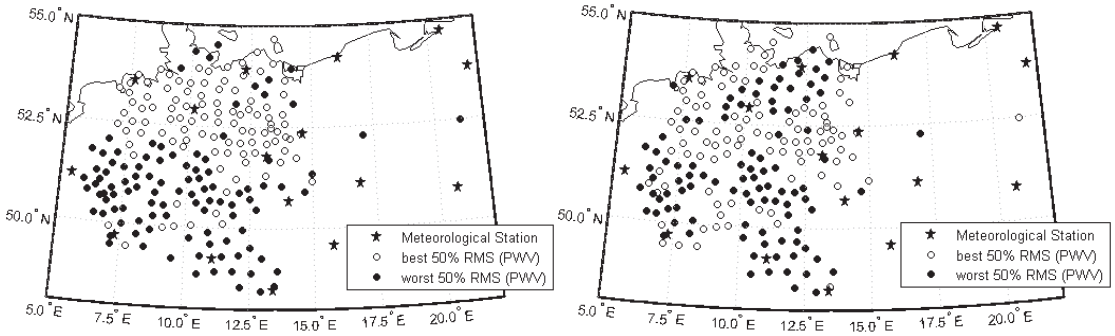


(a) GPS week 1725 (b) GPS week 1751

Figure 28: 1mm PWV RMS maps of GFZ stations. The black dots indicate stations with an PWV RMS greater than 1mm and the white dots are the other GFZ verification sites. Stars indicate the meteorological stations used for the meteorological data interpolation.

Figure 28 verifies that the less accurate stations for both GPS weeks are largely the same. Spatially, most stations are in the west and south of the region. This justifies the assumption of incorrect meteorological verification data. However, the spatial distribution of those stations has some specific features which could cause the comparatively high RMS values. First, most of the PWV RMS stations greater than ±1mm are close to one meteorological station in West Germany. Second, most stations have a high elevation (greater than 1000 metres). This could lead to an inaccurate temperature and pressure extrapolation by the lapse rate formulae and agrees with the findings of the IGS results described above.

The last ice age in Europe is responsible for some landscape characteristics in Germany. Whilst the north of Central Europe is flat, the centre is more mountainous and ends with the Alps in the South. In order to investigate the effect of errors in the temperature and pressure height extrapolations using the lapse rate formulae (equations 17 and 18), the best and worse 50% of the GFZ stations are shown in Figure 29.



(a) GPS week 1725 (b) GPS week 1751

Figure 29: 50% PWV RMS maps of GFZ stations. The black dots indicate stations of the best PWV RMS and the white dots are the other GFZ verification sites (less precise stations). Stars indicate the meteorological stations used for the meteorological data interpolation.

The 50% limit of GPS week 1725 stands for an PWV RMS of approximately ±0.27mm. The PWV RMS limit of GPS week 1751 is approximately ±0.40mm. However, especially for GPS week 1725, a pattern regarding North and South Germany is discernible. This supports the influence of the measurement error caused by the lapse rate formulae. The Figure 29 also shows a noticeable group of less precise stations in the North for GPS week 1751. This effect may be caused by summer thunderstorms which affects significantly the PWV on a regional scale. By using the sparse meteorological coverage and the available three hour intervals, the temperature and pressure can not be interpolated accurately. However, the inaccuracies in this week may also occur due to the described elimination of meteorological stations and the resulting higher average distance to the closest meteorological site. For the last three DOYs in GPS week 1751, Germany is only covered by five meteorological stations. The position of these sites is either in the east (four stations) or in the west (one station). From this perspective, Figure 29 shows white points (best 50% PWV RMS stations) in the northwest which are no more covered for those DOYs. They have been interpolated with closer meteorological stations for the DOYs before and are not subject to the high meteorological station distance of the last three DOYs. In summary, the mass of available GFZ stations allows to make better conclusions about the results than the IGS data. Regarding the density of available stations, the reasons can be better investigated with the GFZ data. However, a more detailed analysis is required to make precise statements of the results.

5. Conclusion

Water vapour is the dominating greenhouse gas, plays a fundamental role in the hydrological cycle and is highly variable in space and time across the earth's atmosphere. A better understanding of this cycle allows a better understanding of the earth's climate system. In this context, PWV is applied in many ways such as weather predictions and climate models. It can be measured, for instance, by radiosonde balloons, remote sensing satellite systems from infra-red radiation measurements or water vapour radiometers. However, this project uses GPS derived delays together with meteorological data to estimate PWV. The GPS signals are refracted and slowed down when travelling through the earth's atmosphere. In view of electromagnetic wave propagation, the atmosphere is divided into the ionosphere and troposphere. The effect of the ionosphere is eliminated using signals on two frequencies. The tropospheric effect is predominantly influenced by the amount of water vapour and air molecules in the air. The tropospheric delay is applied with a mapping function to obtain the ZTD. Bevis *et al.* (1992) describes a way to use the ZTD together with meteorological temperature and pressure measurements for the PWV estimation.

This project briefly describes the basic correlations between water vapour, the atmosphere, GNSS and using GPS, explains the signal refraction through the atmosphere. A possible way to take existing GPS stations without on-site meteorological observations for the PWV estimation by interpolating meteorological data from sparse, publicity available, meteorological sites is shown. The University of Otago in Dunedin, New Zealand, manages GPS stations throughout the country. However, only a few of them can be used to estimate PWV with on-site meteorological measurements - so-called verification stations. By interpolating temperature and pressure observations from sparse meteorological sites, GPS stations without on-site meteorological data have been used to estimate PWV. Spatially, the chosen region is limited by the South Island of New Zealand and consists of two GPS weeks of data with up to 28 available GPS stations in 2009. Since there are only a few reference stations available in that region, the implemented approach has been further developed and may be used for any geographic region. A second region in Central Europe has been considered to contrast the results from New Zealand. This region consists of two GPS weeks in Central Europe that has the highest density of freely available IGS stations in 2013. Spatially, the region is limited by approximately 49°–55° North and 6°–22° East. In addition, many verification datasets have been provided by the GFZ in this area. IGS provides a maximum of 19, whereas the GFZ datasets provide up to 239 GPS stations. For each region, one week in summer and one in winter has been processed. As meteorological

input, the sparse, publicity available, NOAA recordings have been used to interpolate temperature and pressure measurements at the GPS sites. The maximum of available meteorological stations is 15 in the South Island of New Zealand and 18 for the region in Central Europe. This results in an average distance from the GPS stations to the closest meteorological site of approximately 45 to 49km (South Island) and approximately 97 to 204km (Europe).

Several tools were developed to import data from the GPS and meteorological input files. The interfaces used in this work have been kept as generic as possible to enable independent evaluations. Analysis tools were developed and implemented to estimate PWV at one hour intervals from GPS derived ZTD measurements together with either sparse or at the GPS station available, temperature and pressure observations in the chosen regions. The implemented European approach can be adopted for any other geographic region with arbitrary time spans. Stations with erroneous measurements heavily affect their environments during the spatial interpolation. Obviously wrong observations have been eliminated by filtering temperature or pressure measurements above a defined limit (e.g. 100 °C or 1500mbar). However, an individual examination of the input meteorological and GPS datasets is necessary.

Using the sparse available meteorological data, PWV time series maps have been produced for these periods. Samples of the produced PWV time series maps are shown and characteristics are highlighted. The processed one week periods in two different regions around the world demonstrate the meteorological impacts on the PWV. In general, summer periods show a higher amount of PWV in the atmosphere than the winter periods. Effects of incoming atmospheric depressions and anticyclones with corresponding rain, rainfall in front of mountains and higher PWV values in higher elevations were observed. The PWV time series maps produced with the IGS and University of Otago GPS data allow a general perspective, whilst the GFZ data shows small-scaled regional characteristics.

Ideally, it would be best to choose as many sparse meteorological stations as possible with a reasonable spatial distribution and small distance to the GPS stations. The created PWV maps may be strongly influenced by regional characteristics. In order to generate precise PWV maps, a high GPS station density is required, e.g. by combining the IGS and GFZ data. In addition, specific regions such as mountain ranges need a good station coverage. Moreover, the meteorological and GPS stations should minimize height differences because of the inaccuracies caused by the temperature and pressure extrapolation.

GPS stations with both, tropospheric delay and meteorological measurements available, are used to produce PWV reference data. The reference values are compared to PWV estimations using sparse, publicity available, meteorological data

which allow statements about the precision of the sparse estimations. A PWV RMS of smaller than ±0.5mm is achieved for both periods at the three available stations in New Zealand. A mean PWV RMS of ±0.59mm is observed for the IGS stations in the winter period, whilst the summer shows a mean of ±0.82mm. Regarding the GFZ stations in the winter period, 156 out of 210 verification stations have an PWV RMS of smaller than ±0.5mm from which 90 are below ±0.25mm. During the summer period, 127 out of 197 stations have an PWV RMS smaller than ±0.5mm from which only 22 are below ±0.25mm. This is reflected by the mean PWV RMS of ±0.45mm for GPS week 1725 and ±0.59mm for GPS week 1751. The fluctuating differences between summer and winter RMS can be impacted by the generally higher or lower PWV values during these periods. Another error source are the regional characteristics at the GPS stations which can not be covered with the interpolated sparse meteorological data. The impreciseness might also be caused by different positions and heights of the GPS and meteorological sensors.

In the past, field studies have been performed to compare GPS based PWV estimations with radiosondes or other techniques. However, the produced PWV time series maps should be compared with e.g. AVHRR based PWV estimations to obtain an insight of the accuracy of the interpolated maps. Processing additional regions and longer time periods would enable incorrect meteorological measurements to be identified at the verification sites. Extending the area of application over many and/or larger regions throughout the world, the PWV maps could be combined into a global model. Since the temperature and pressure estimations depend on orthometric heights, the underlying gravitational model (EGM96) could be replaced by a more accurate model to increase the precision of the station heights. Regarding the interpolations, further studies are necessary to ensure a good interpolation model.

Many GPS and meteorological stations are available worldwide. The data from some of these sites are freely applicable. In summary, this project demonstrates the GPS based PWV estimation by using freely and not freely available GPS and meteorological data. By applying minimal optimizations to the implemented tools, any geographic region and time periods can be processed. Regions with no or very sparse available GPS stations can be covered by installing additional GPS stations. This enables cost-efficient PWV estimations with a precision that is possible by using sparse available meteorological data.

References

Bevis, M., Businger, S., and Chiswell, S. (1994). GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. *Journal of Applied Meteorology*, 33:379–386.

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H. (1992). GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. *Journal of Geophysical Research*, 97(D14):15,787–15,801.

Boccolari, M., Fazlagic, S., Frontero, P., Lombroso, L., Puhnaghi, S., Santangelo, R., Corradini, S., and Teggi, S. (2002). GPS Zenith Total Delays and Precipitable Water in comparison with special meteorological observations in Verona (Italy) during MAP-SOP. *Annals of Geophysics*, 45(5):599–608.

Coster, A. J., Niell, A. E., Solheim, F., Mendes, V. B., Toor, P. C., Langley, R. B., and Ruggles, C. A. (1996). The Westford Water Vapor Experiment: Use of GPS to Determine Total Precipitable Water Vapor. Presented at the ION 52nd Annual Meeting, Cambridge, MA 19-21 June 1996.

Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E., and G., E. (1985). Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio Science*, 20(6):1593–1607.

Denys, P. (2010). Water vapour estimation using Bernese 5.0. Technical report, School of Surveying, Otago University.

ESA (2014). Reference Frames in GNSS. `http://www.navipedia.net/index.php/Reference_Frames_in_GNSS` - viewed 10th September 2014.

ESA-Galileo (2014). Galileo Future and Evolutions. `http://www.navipedia.net/index.php/GALILEO_Future_and_Evolutions` - viewed 10th September 2014.

Essen, L. and Froome, K. D. (1951). The Refractive Indices and Dielectric Constants of Air and its Principal Constituents at 24 GHz. *Proceedings of the Physical Society (London)*, (64):862–875.

Gutman, S. and Benjamin, S. (2001). The Role of Ground-Based GPS Meteorological Observations in Numerical Weather Predictions. *GPS Solutions*, 4(4).

Hofmann-Wellenhof, B., Kienast, G., and Lichtenegger, H. (2001). *Global Positioning System: Theroy and Practice*. Springer-Verlag Wien New York, 5 edition.

Hofmann-Wellenhof, B., Lichtenegger, H., and Wasle, E. (2008). *GNSS - Global Navigation Satellite Systems*. Springer WienNewYork.

Kraus, H. (2009). *Die Atmosphäre der Erde*. Springer Verlag Berlin Heidelberg New York, 3 edition.

Kruse, L. P. (2001). Spatial and Temporal Distribution of Atmosperic Water Vapor using Space Geodetic Techniques. *Swiss Geodetic Commission*, 61.

Kuo, Y. H., Zou, X., and Guo, Y. R. (1996). Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model. Part I: Moisture retrieval and sensitivity experiments. *Monthly Weather Review*, 124:122–147.

Marini, J. W. (1972). Correction of satellite tracking data for an abritrary tropospheric profile. *Radio Science*, 7(22):223–231.

Melbourne, W., E., D., Duncan, C., Hajj, G., Hardy, K., Krsinsksi, E., Meehan, T., Young, L., and Yunck, T. (1994). *The Application of Space-Borne GPS to Atmospheric Limb Sounding and Global Change Monitoring*. National Aeronautics and Space Administration, Jet Propulsion Laboratory.

Metz, B., Davidson, O., P.R., B., Dave, R., and L.A., M. (2007). Climate Change 2007: Mitigation of Climate Change. `http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg3_report_mitigation_of_climate_change.htm` - viewed 6th June 2014.

Mockler, S. (1995). Water Vapor in the Climate System: Special Report. `https://www.eso.org/gen-fac/pubs/astclim/espas/pwv/mockler.html` - viewed 19th June 2014.

Motell, C., Porter, J., Foster, J., Bevis, M., and Businger, S. (2002). Comparison of precipitable water over Hawaii using AVHRR-based split-window techniques, GPS and radiosondes. *International Journal of Remote Sensing*, 23(11):2335–2339.

Navigation Center of Excellence (2001). Global Positioning System Standard Positioning Service Performance Standard. `http://www.navcen.uscg.`

gov/pdf/gps/geninfo/2001SPSPerformanceStandardFINAL.pdf - viewed 22nd June 2014.

nc-climate.ncsu.edu (2013). Structure of the Atmosphere. https://www.nc-climate.ncsu.edu/edu/k12/.AtmStructure - viewed 22nd June 2014.

Niell, A. E. (1996). Global mapping functions for the atmosphere delay at radio wavelengths. *Journal of Geophysical Research*, 101(B2):3227–3246.

Reigber, C., Gendt, G., Dick, G., and Tomassini, M. (2002). Near real time water vapor monitoring for weather forecasts. *GPS World*.

Rocken, C., Ware, Randolph, H., Hove, T. V., Solheim, F., Alber, C., Johnson, J., Bevis, M., and Businger, S. (1993). Sensing atmospheric water vapor with the global positioning system. *Geophysical Research Letters*, 20(23):2631–2634.

Saastamoinen, J. (1972). Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. *The Use of Artificial Satellites for Geodesy*, 15:247–251.

satsig.net (2006). Satellite Signals - Orbits. http://www.satsig.net/maps/finding-lat-long-with-gps.htm - viewed 22nd June 2014.

Thayer, G. D. (1974). An improved equation for the radio refractive index of air. *Radio Science*, 9(10):803–807.

Walpersdorf, A., Bouin, M.-N., Bock, O., and Doerflinger, E. (2007). Assessment of GPS data for meteorological aapplication over Africa: Study of error sources and analysis of positioning accuracy. *Journal of Atmospheric and Solar-Terrestrial Physics*, 67:1312–1330.

Wolfe, D. and Gutman, S. (2000). Developing an Operational, Surface-Based, GPS, Water Vapor Observing System for NOAA: Network Design and Results. *Journal of Atmospheric and Oceanic Technology*, 17(4):426–440.

Wonnacott, R. (2005). *The Estimation of Precipitable Water Vapour from GPS Measurements in South Africa*. PhD thesis, School of Architecture, Planning and Geomatics, University of Cape Town, South Africa.

Zealand, L. I. N. (2014). Other geoid models. website. http://www.linz.govt.nz/geodetic/datums-projections-heights/vertical-datums/other-geoid-models - viewed 28th May 2014.

Zhengdong, B. and Yanming, F. (2003). GPS Water Vapor Estimation Using Interpolated Surface Meteorological Data from Australian Automatic Weather Stations. *Journal of Global Positioning Systems*, 2(2):83–89. http://www.scirp.org/journal/PaperDownload.aspx?paperID=227&fileName=nav20030200002_25554713.pdf - viewed 26th September 2014.

A. Flow Chart Verification Process South Island

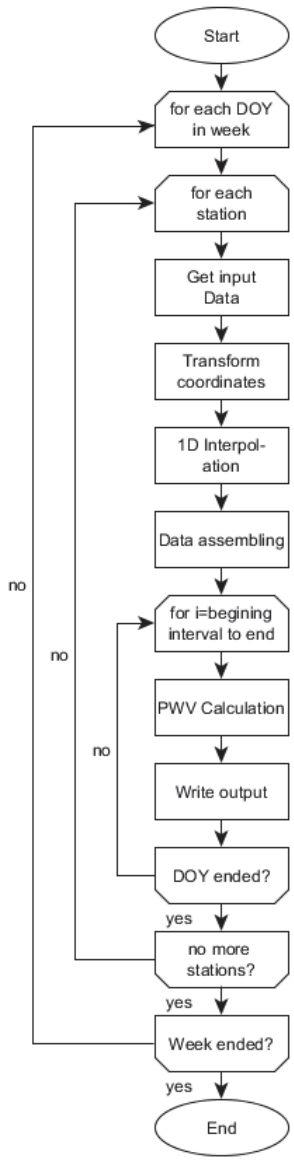


Figure A.1: Flow chart to process verification data for one station and one week

B. Difference Charts Europe

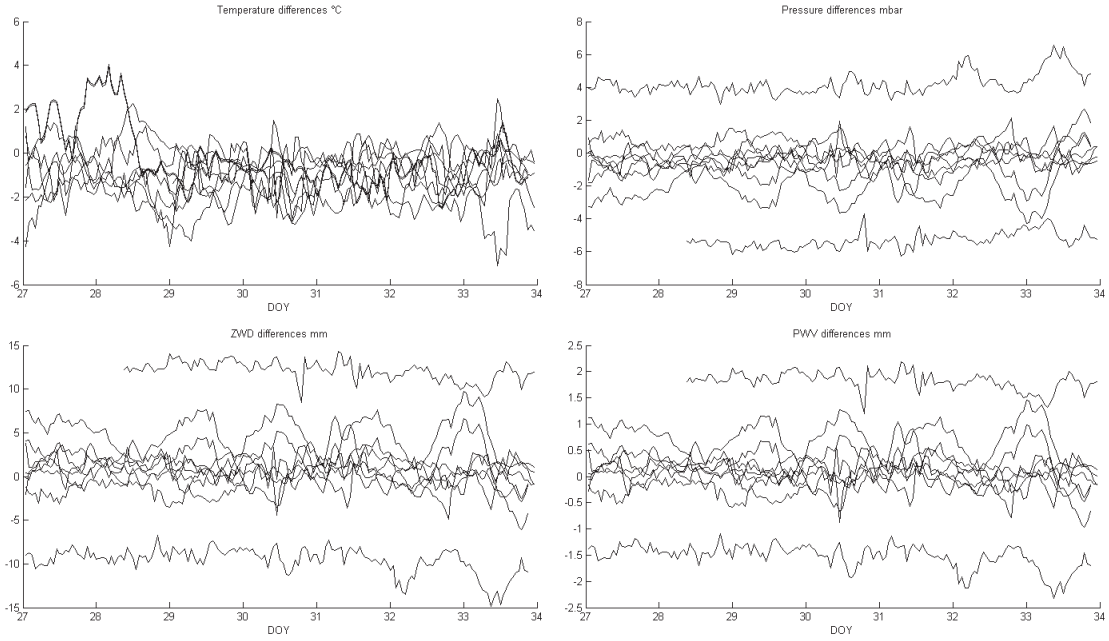


Figure B.1: IGS differences GPS week 1725

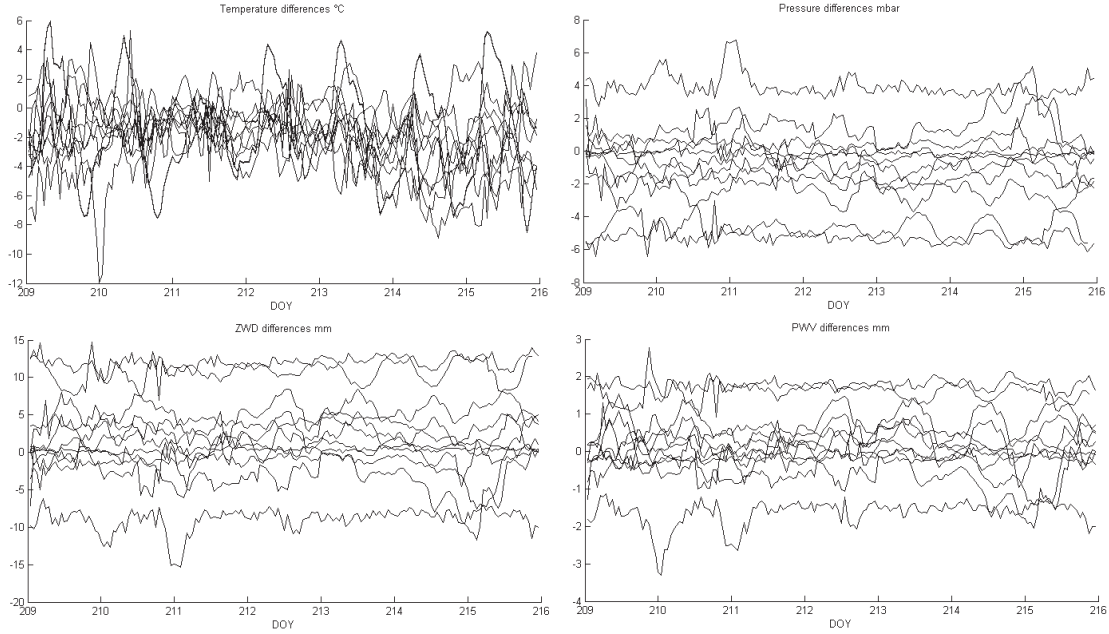


Figure B.2: IGS differences GPS week 1751

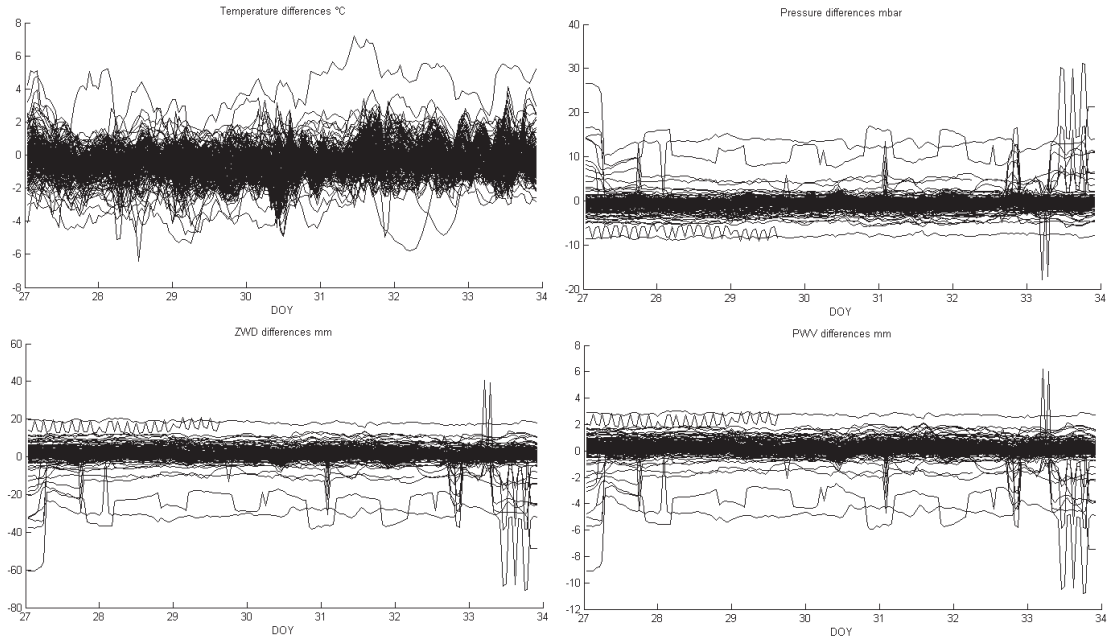


Figure B.3: GFZ differences GPS week 1725

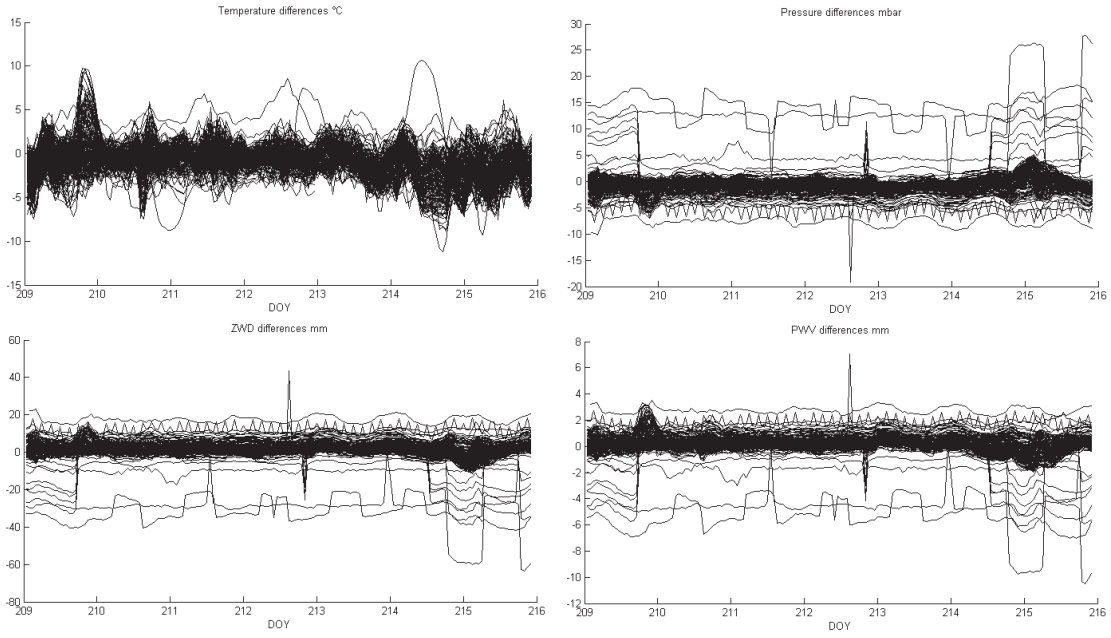


Figure B.4: GFZ differences GPS week 1751

Statutory Declaration

I declare that I have developed and written the enclosed Master Thesis completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. The Master Thesis was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere.

Place, Date

Signature

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Dr. Paul Denys from the University of Otago in Dunedin, New Zealand, for his continuous support, patience and motivation. His guidance helped me in all the time of research and writing of this thesis. I would also like to thank Prof. Dr.-Ing. Andreas Wehrenpfennig from the University of Applied Sciences in Neubrandenburg, Germany, for his insightful comments and questions.

My sincere thanks also goes to Dr. Jens Wickert from the German Research Centre for Geosciences (GFZ) Potsdam who had arranged the contact to Dr. Paul Denys. Special thanks also to Dr. Galina Dick (GFZ) for providing the large number of the European GFZ datasets used in this project. I am particularly grateful for the support of my parents who made my studying in Neubrandenburg and the stay in New Zealand possible. Last but not least, I would like to thank my friends for listening, offering me advice and supporting me through this entire process.