Weighted Mean Temperature Models for estimating Global Navigation Satellite System Precipitable Water Vapor over East Africa

By

PETER MBAYO

BU/GS19/MSP/3

A research dissertation submitted to the Directorate of Graduate Studies, Research and Innovations in partial fulfillment of the requirements for the award of Master of Science in Physics of Busitema University

MARCH, 2024

DECLARATION

I Mbayo Peter, declare that this dissertation is my original work and that it has never been submitted to any other institution for an academic award. I affirm that I consistently acknowledge the sources when I quote from the works of other authors. Aside from these quotations, this dissertation represents my original work.

Name: MBAYO PETER
Registration Number: BU/GS19/MSP/3
Signature:
Date 6/03/2024

Approval

This study was conducted under guidance of the undersigned supervisors and certify that they have read and hereby recommend for examination by Busitema University a dissertation titled: **Weighted Mean Temperature Models for estimating Global Navigation Satellite System Precipitable Water Vapor over East Africa**, in (Partial) fulfillment of the requirements for the award of Master of Science in Physics of Busitema University.

1. Dr. Richard Cliffe Ssenyunzi

Department of Physics, Faculty of Science and Education, Busitema University

Signature...

Date....6/03/2024....

2. Assoc Prof. Emirant Bertillas Amabayo

Department of Physics, Faculty of Science and Education, Busitema University

Signature...

Date 06/03/2024

ACKNOWLEDGMENT

I would like to thank the All-Powerful God, through Apostle Kayiwa David and Pastor Rodrick for providing me with guidance, courage, wisdom, safety, skills, and the opportunity to live a long and healthy life.

I would also like to thank my supervisors from Busitema University, Dr. Richard Cliffe Ssenyunzi and Assoc Prof. Amabayo Emirant Bertillas, for the guidance and contributions throughout this research.

Special thanks also goes to Dr. Geoffrey Andima for introduction to MATLAB programming. His insights and constructive suggestions were immensely beneficial to this research.

I thank my beloved Parents, Mr. Takunya Peter and Ms.Nakirya Edith, for raising me well and providing for me in numerous ways. To my wife Kisakye Faith, thank you very much for encouraging me unceasingly.

I thank Busitema University for granting me the opportunity to conduct this research. I must appreciate too, all the staff of the Department of Physics for their academic, administrative and moral support.

I also thank everyone who has given me advice and motivation to finish my studies, including my family, sisters, cousins, mentors, friends, coworkers, classmates and lecturers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACROYNMS AND ABBREVIATIONS

AIRS	Atmospheric infrared Sounder
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF Reanalysis 5
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
IGS	International GNSS Service
IWV	Integrated Water Vapor
NGL	Nevada Geodetic Laboratory
NWM	Numerical Weather Model
NWP	Numerical weather prediction
PWV	Precipitable Water Vapor
RH	Relative Humidity
RINEX	Receiver Independent Exchange Format Data
RS	Radiosonde
SHADOZ	Southern Hemisphere ADditional OZonesonde Network
T_m	Weighted mean temperature
UT	Universal Time
VMF	Vienna Mapping Function
WMO	World Meteorological Organisation
ZHD	Zenith Hydrostatic Delay
ZTD	Zenith Tropospheric Delay
ZWD	Zenith Wet Delay

ABSTRACT

The weighted mean temperature (T_m) is an important parameter in deriving accurate and reliable Global Navigation Satellite System (GNSS) Precipitable Water Vapor (PWV) values especially in tropical regions. However, the accuracy of GNSS - PWV in the East African region is limited by the availability of precise and reliable meteorological data for computing accurate T_m values. To address this limitation, this study developed site-specific and regional T_m models for East Africa (E.A) using radiosonde and Ozonesonde profiles from 1971 to 2019. The models were validated using radiosonde profiles from two stations 63741 (Dagoretti) and 63894 (Dares-salaam) based on Root Mean Square Error (RMSE) and Mean Bias Error (MBE) values of T_m. The T_m values from the established models were compared with the two global (Bevis and Yao) and one Tropical regional (Raju) T_m models in order to assess their performance. At station 63741, the RMSE of the E.A model, the site-specific T_m model, the Bevis, Yao, and Raju T_m models are 2.01 K, 1.76 K, 1.79 K, 2.54 K, 2.74 K, and 2.47 K, respectively. At station 63894, the RMSE of the E.A model, the site-specific T_m model, the Bevis, Yao, and Raju T_m models are 1.66 K, 1.63 K, 1.70 K, 2.52 K and 2.64 K, respectively. At both stations, the site-specific T_m models have higher accuracy followed by the E.A model as compared to Raju T_m model and the two global T_m models (Bevis and Yao). The study further evaluated the accuracy of T_m models in the retrieval of PWV by computing PWV values at three GNSS sites (ARSH, DODM, and RCMN). The GNSS - PWV values computed from different T_m models were compared with PWV values derived using T_m from ERA5 reanalysis dataset from 2013 to 2020. The use of GNSS-PWV from ERA5-Tm as the reference values was due to the limited Tm values from Radiosonde during the same periods of available GNSS data required for PWV estimation. In terms of RMSpwv, the site-specific models exhibit values ranging from 0.107 mm to 0.11 mm, with a mean value of 0.109 mm, which is the lowest among the five models. This is followed by, the E.A regional model with mean value of 0.127 mm, Yao T_m with a mean value of 0.129 mm, Bevis with a mean value of 0.130 mm, and Raju with a mean value of 0.148 mm.The results show that the E.A regional T_m model and site-specific models perform better than the global and other tropical regional T_m models. Therefore, the established T_m models are the preferred choices for estimating T_m for GNSS - PWV retrieval in the East African region. Overall, the study provides improved T_m models for accurate estimation of GNSS - PWV in the East African region, which could improve weather and climate forecasts.

Chapter 1

Introduction

1.1 Background

Atmospheric water vapor is a vital component of the atmosphere that has a significant impact on weather and climate. Information regarding the amount of water vapor present in the atmosphere can be obtained by analysing the impact of the troposphere on the Global Navigation Satellite System (GNSS) signal propagation (Bevis et al., 1992; Hu et al., 2021). When a GNSS signal travels through the troposphere, the signal is refracted and therefore delayed. The troposphere is not dispersive in nature. Consequently, the primary cause of this delay and preventing its elimination is atmospheric water vapor. The delay of signals from the satellite received at different angles can be computed as slant tropospheric delay (s) (STD). The total slant delay when mapped in zenith direction is defined as zenith tropospheric delay (ZTD) (Mendes, Prates, Santos, & Langley, 2000). Using ground-based GNSS receivers, ZTD may be obtained with high accuracy (Jiang, Ye, Chen, Liu, & Xia, 2016). The ZTD consists of two components which include the zenith hydrostatic delay (ZHD) and the non-hydrostatic (wet) delay (ZWD) (Boutiouta & Lahcene, 2013). The ZHD is dependent on air pressure and can be estimated using empirical models like, (Saastamoinen, 1972) model and (Hopfield, 1969) among others. According to Seeber (2003), ZHD contributes about 90 % and ZWD component only makes up to about 10 % of the total tropospheric delay. Conversely, ZWD depends on the amount of water vapor in the atmosphere and therefore a crucial parameter in determining the atmospheric water vapor content (Davis, Herring, Shapiro, Rogers, & Elgered, 1985). Due to the great temporal and spatial fluctuation in the water vapor quantity in the atmosphere, ZWD is a very difficult component to estimate. As a result, reliable modeling of non-hydrostatic delay remains difficult. However, ZWD can be obtained accurately as the difference between ZTD and ZHD.

Due to crucial impact of atmospheric water vapor on weather and climate, its observations are of utmost relevance. These have a wide range of applications, including; monitoring climate change and improving numerical weather prediction models for better weather forecasting accuracy (Yuan et al., 1993). The formation of critical weather elements such as clouds, rain, snow, and other precipitation occurs when water vapor condenses into liquid or ice (Chahine, 1992; X. Li et al., 2014). The condensation results into releasing latent heat that redistributes through the atmosphere. This latent heat is a crucial factor in the atmospheric energy budget at both local and global levels and plays a significant role in most significant weather phenomena, such as tropical cyclones and severe thunderstorms (W. Zhang, Lou, Huang, & Liu, 2018). According to Emmanuel et al. (2018), an accurate assessment of Precipitable Water Vapor (PWV) is crucial for predicting severe weather and precipitation. Additionally, research by Bernet et al. (2020) shows that a precise assessment of PWV is necessary for the prediction of precipitation and severe weather conditions. Therefore, understanding and monitoring the amount and distribution of atmospheric water vapor, as well as its effects on atmospheric radiation and circulation, are essential when analyzing and predicting short- and long-term climate changes. The PWV refers to the total amount of atmospheric water vapor present in a vertical column of air of unit area, rising from the Earth's surface to the upper atmosphere (Bevis et al., 1992). The PWV is a reliable indicator of atmospheric water vapor content and is important in precipitation forecasting (J. Wang, Zhang, & Dai, 2005; Suparta & Iskandar, 2013). Consequently, a quantitative understanding of atmospheric water vapor distribution is vital for weather forecasting and climate research. As per the study by (Mircheva, Tsekov, Meyer, & Guerova, 2017), comprehending changes in the hydrological cycle, including atmospheric water vapor, is crucial in predicting heat waves.

The parameter known as weighted mean temperature ($\mathrm{T_m}$) is a critical component in computation of PWV present in the atmosphere through the use of zenith wet delay (ZWD) data from the GNSS signal propagation (Bevis et al., 1992). The accuracy of $\mathrm{T_m}$ is crucial for producing highly precise PWV values from GNSS data, as it is the conversion factor used to convert ZWD to PWV (Davis et al., 1985). The value of $\mathrm{T_m}$ varies significantly based on location, height, season, and weather conditions, making it highly correlated with climate conditions (Suwantong, Satirapod, Srestasathiern, & Kitpracha, 2016; Jiang et al., 2019). Accurate calculation of $\mathrm{T_m}$ requires numerical integration, which can only be achieved if atmospheric vertical profiles of water vapor partial pressure and temperature are available in the area of interest. Traditionally, Radiosonde (RS) data and the Atmospheric infrared Sounder (AIRS) data have been used to generate these in - situ profiles. But the limited number of these sounders and their high operating costs make it difficult to achieve the necessary spatial-temporal coverage. Moreover, many GNSS stations in East Africa lack in-situ meteorological data. This creates significant gaps in estimated $\mathrm{T_m}$ values, limiting their use in PWV estimation for weather forecasting and prediction. To overcome these limitations, numerical weather prediction (NWP) and other empirical models are being used as alternatives to compute precise $\mathrm{T_m}$ values (Yang et al., 2019). Some popular NWP models used for this purpose include the Vienna Mapping Functions 1 (VMF1), the fifth-generation European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA5), and the National Centers for Environmental Prediction (NCEP). However, NWP data is released in a big volumes usually every 6 hours except for

ERA5 which is updated hourly. Since collecting data can take a while, certain significant real-time variations in weather which could be particularly strong during some extreme weather events may go unnoticed. Consequently, the modeled $\mathrm{T_m}$ estimates from these NWP data are inadequate for high-precision meteorological applications, such as providing real - time GNSS-PWV estimates for weather forecast.

Empirical models that relate $\mathrm{T_m}$ to atmospheric profiles are crucial for computing $\mathrm{T_m}$ especially in regions with limited access to in-situ measurements or when large amounts of data or real-time performance are needed. The Bevis et al. (1992) model, which relates surface temperature ($\mathrm{T_s}$) to $\mathrm{T_m}$ using a linear connection, has been widely used in PWV retrieval, but its accuracy may be limited in tropical areas. Several researchers like, Mendes et al. (2000), Schueler, Pósfay, Hein, and Biberger (2001), Y. Yao, Zhu, and Yue (2012), Y. B. Yao, Zhang, Yue, Xu, and Peng (2013), Y. Yao, Zhang, Xu, and Yan (2014), Sun, Zhang, and Yao (2019), and Yang et al. (2019) have developed and enhanced linear empirical models between $\mathrm{T_m}$ and $\mathrm{T_s}$ with aim at advancing the accuracy of $\mathrm{T_m}$.

While there have been many studies conducted to estimate $\mathrm{T_m}$ in various regions, the tropical East African region remains largely unexplored due to limited radiosonde networks and significant data gaps. The parameters used in studies conducted in other regions are not representative of the climate and weather conditions in the East African tropical region. Therefore, there is a need for a comprehensive investigation into $\mathrm{T_m}$ estimation in this area. This is crucial because $\mathrm{T_m}$ varies greatly based on factors such as latitude, surface altitude, and time of the year. This current research established empirical models of $\mathrm{T_m}$ for East Africa based on radiosonde and Ozonesonde profiles from 1971 to 2019.

1.2 Problem Statement

The East African region, like any other tropical zone is characterized by large amounts of water vapor with high spatial and temporal variability. As a result, there are variations in the region's rainfall distribution, with certain locations experiencing more rain than others. The rainfall is vital for agricultural production, which is the primary source of livelihood for many communities in the region. However, excessive rainfall has led to flooding, posing threats to the safety of people and property. Due to this, the region is extremely vulnerable to the effects of climate change and global warming. Because GNSS technology is essential for tracking these changes, it is vital to embrace it through the concept of GNSS meteorology in order to improve the region's capacity to predict weather and climate changes. The GNSS technology offers the capacity to identify and monitor extreme weather events including hurricanes, tropical cyclones, and severe thunderstorms through measurements of accurate atmospheric water vapor. However, deriving high-accuracy PWV values from GNSS data requires precision in a crucial parameter, known as weighted mean temperature, T_m. Although the number of GNSS ground receiver sites in the East African region that can be used to estimate the PWV has increased, getting accurate and reliable results remains a difficulty. This is because there are limited precise and reliable devices available to provide the surface and vertical profiles of meteorological data required for computing accurate T_m. As a result, the use of GNSS-PWV in weather and climate forecasts across the East African region has been significantly constrained. This study therefore, developed weighted mean temperature models for the region using data from six radiosonde and two ozonesonde stations to accurately estimate T_m which is important in precise retrieval of GNSS-derived water vapor and enhancing weather forecasting in the region.

1.3 Aim of the study

To develop weighted mean temperature models for estimating Global Navigation Satellite System precipitable water vapor over East Africa.

1.4 Objectives

The objectives include;

1. Estimate the weighted mean temperature over East Africa.
2. Develop the weighted mean temperature models for the East African region.
3. Validate the weighted mean temperature models.
4. Determine the impact of the developed weighted Mean temperature models on GNSS - derived PWV.

1.5 Scope of study

The study was conducted over East Africa which lies within the latitude range 7°S and 4°N, longitude range 28°E to 44°E. It is composed of several countries, such as Kenya, parts of Tanzania, Uganda, Rwanda, and Burundi, and is part of the African tropical zone. The study focused on development and validation of T_m models for GNSS-PWV calculation. This study establishes regional T_m-model and site - specific empirical T_m models for East Africa using the radiosonde and ozonesonde datasets. Then, the performances of the developed T_m models are compared with the two global models (Bevis et al., 1992) and

(Y. Yao et al., 2014) and one tropical model (Suresh Raju, Saha, Thampi, & Parameswaran, 2007) to verify their reliability. The study further evaluated the impact of developed models on GNSS-PWV retrieval.

1.6 Significance

This study has developed regional and site specific T_m models from ozonesonde and radiosonde datasets. The developed models will generate T_m data at any time at these sites and any other locations with same weather conditions. The generated T_m data from the models will always help in calibrating or validating other T_m sources with high accuracy. The T_m models will also lead to the estimation of accurate GNSS - PWV data useful for many environmental applications, such as meteorology, hydrology and climate monitoring. Forecasts of the weather that are both dependable and accurate will help communities plan more effectively for sectors like agriculture and disaster preparedness, as well as for activities that are sensitive to our changing climate.

Chapter 2

Literature Review

2.1 Water Vapor in the Atmosphere

The water vapor content in atmosphere is the most important parameter in establishing the Earth's climate and its short term changes (Bevis et al., 1994). Because of the higher rate of evaporation, it rises when air temperature rises. Moreover, water vapor affects cloud formation, energy transmission, the Earth's radiation budget, and the distribution of precipitation that results (Long, Hu, Dong, & Wang, 2021; Hu et al., 2021). The main source of this water vapor is evaporation from water bodies when heated by solar energy and supplemented by transpiration from plants. The study by Shao et al. (2023), indicates that there is an asymmetry in the latitudinal distribution of water vapor between the northern and southern hemispheres, with the polar areas having the smallest amount and the equator zone (tropics) having the greatest quantity (Jiang et al., 2016). The exchange and movement of water between the earth and atmosphere occurs in the troposphere and begin as the sun evaporates large amounts of water from the earth's surface and the moisture is transported to other regions by the wind. As air rises, it expands and cools, water vapor condenses and clouds develop. The sun warms the air at the equator more than the air at the poles causing convection currents, large-scale patterns of winds that move heat and moisture around the globe.

The biggest percentage of water vapor is at low altitudes, which increases the absorption of solar radiation and long wave terrestrial absorption, hence reducing the amount of outgoing long-wave terrestrial radiation. The energy released when water vapor condenses causes disturbances in the atmosphere's vertical stability influencing global weather systems and their associated precipitation patterns. In this way, evaporation - condensation cycle is a significant mechanism for moving heat energy from the earth's surface to its atmosphere and in moving heat around the earth (Boutiouta & Lahcene, 2009).

Whereas atmospheric water vapor constitutes less than 5 % of the atmosphere's mass, it has a tremendous influence on the spatial and temporal evolution of the Earth's atmospheric properties (Huang, Mo, Liu, et al., 2021). It is one of the key factors in monitoring climate change, the earth's greenhouse effect, and in precipitation forecasting making it important component of the atmosphere (Elgered & Jarlemark, 1998; S. Wang et al., 2020). It influences various processes in the ecosystem; such as hydrological cycle, atmospheric radiations, weather patterns, and climate changes (Elhaty, Abdelfatah, Mousa, & El-Fiky, 2019; Sun et al., 2019). The hydrological cycle describes the transfer of water through various phases (e.g. solid, liquid and gaseous form), along with the physical movement of water within the ecosystem of the earth, and between its atmosphere, oceans and continents (Huang, Mo, Liu, et al., 2021). Being the most highly variable component of the atmosphere, atmospheric water vapor is unevenly distributed in time and space, therefore its changeability in the atmosphere affects the short-term and long-term changes in weather and climate (Suresh Raju et al., 2007; Yang et al., 2019). It is therefore, vital to frequently monitor the content of water vapor in the atmosphere for reliable weather forecasts (Liu, Yao, & Sang, 2018; Sun et al., 2019). Despite its importance, water vapor is one of the elements of the atmosphere that is very difficult to measure accurately because of its temporal and spatial variability.

Over the past decades, various techniques have been developed to measure the amount of water vapor in the atmosphere, and these include; using meteorological data from Light Detection and Ranging (LIDAR) systems and the Atmospheric infrared Sounder (AIRS) data (Maghrabi, Clay, Wild, & Dawson, 2009; Zheng, Lou, Gu, Gong, & Shi, 2018). Other techniques include; radiosondes, ground-based sun photometers and microwave radiometers. These, however, proved to have poor spatial and temporal resolutions (J. Wang et al., 2005). Among these, determination of PWV from radiosondes is the prime in-situ measurement method due to its high accuracy. However, due to its limitations in terms of cost and spatiotemporal resolution, this approach often yields errors in PWV estimates (Maghrabi et al., 2009; Elhaty et al., 2019). A lower utilization of microwave radiometers for PWV measurement in East Africa is caused by the scarcity of microwave radiometer sites and their irregular measurements in the region (Satirapod, Anonglekha, Choi, & Lee, 2011). Because of their various resolution problems and high cost, none of these approaches can thus meet the growing demands of the developing meteorological applications.

In light of all these problems faced by the existing PWV determining methods, the initiation of the GNSS meteorology; a technique for retrieving PWV from GNSS signals has gained attention because of its high accuracy, low cost, all-weather usability, real-time and continuous information accessibility (Satirapod et al., 2011). The GNSS water vapor data can help researchers in mesoscale modelling and data assimilation which provides accurate information for prediction of severe weather, precipitation, cloud dynamics, climate and hydrology at a smaller cost (Manandhar, Lee, Meng, & Ong, 2017). Various studies have proved that GNSS is capable of providing accurate water vapor estimates with an accuracy comparable to the measurements from meteorological sensors in both post-processing and near-real-time modes (Bernet et al., 2020). Studies by Bevis et al. (1992), Rocken, Van Hove, and Ware (1997), De Haan (2013), Lindskog, Ridal, Thorsteinsson, and Ning (2017) have also showed that GNSS improves short-term weather forecasts justifying its applicability in estimation of PWV. As a result, the increasing usage of GNSS applications for PWV estimate has led to the incorporation of these observations into weather forecasting models, hence improving weather forecasting (Businger et al., 1996; Jiang et al., 2019).

2.2 Global Navigation Satellite System

The Global Navigation Satellite System (GNSS) refers to all constellations of satellites that provide autonomous Positioning, Navigation and Timing (PNT) data information to user receivers round the globe. The GNSS comprises three segments namely; space segment, control segment and user segment. The space segment consists of satellites flying in selected orbits around the earth. The control segment comprises of a network monitors, control and upload stations to ensure that the GNSS satellite orbits and clocks remain within acceptable limits. The GNSS user segment is the equipment used for processing the received GNSS signals to derive location and time information. It consists of L-band radio receiver/processors and antennas which receive GNSS signals, determine pseudoranges (and other observables), and solve the navigation equations in order to obtain their coordinates and provide a very accurate time (Misra & Enge, 2006). The GNSS satellite systems are described in the following paragraphs.

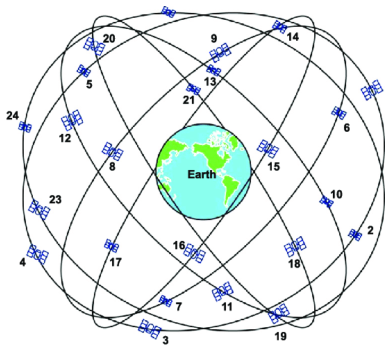


Figure 2.1: GNSS Constellation
Adapted from https://www.vectornav.com/resources/inertial-navigation-primer/theory-of-operation/theory-gnss

The United States Global Positioning System (GPS) is the oldest complete and prevalent GNSS network. GPS is the most accurate navigation system in the world with its space segment currently consists of 32 satellites that transmit radio signals from six almost circular equally spaced orbital planes, with altitudes of about 22,000 kilometers above the Earth's surface and orbits the earth about every 12 hours (Kumar et al., 2021).

The Russian Global Navigation Satellite System (GLONASS). This is developed and operated by the Russian Federation and currently consists of 24 satellites in operation. The satellites are placed into nominally circular orbits with target inclinations of 64.8° and an orbital radius of 19,140 kilometres (11,893 miles), about 1,060 km (659 miles) lower than GPS satellites. Each orbital plane contains eight equally spaced satellites (?, ?)

The Chinese global Navigation Satellite System (BeiDou) consists of two separate satellite constellations. The first BeiDou system also known as BeiDou-1 consisted of three satellites by 2000. This offered limited coverage and navigation services and was decommissioned at the end of 2012. The second generation, officially called BeiDou Navigation Satellite System (BDS) and also known as COMPASS or BeiDou-2 became operational in December 2011 designed for covering the whole Asia-Pacific region. In 2015, China launched the third generation called BeiDou-3 for global coverage. Since then, the BeiDou system has developed and improved continuously into a global navigation system. The BeiDou-3 system consists of satellites in three different orbits, including 24 satellites in medium-circle orbits (covering the world), 3 satellites in inclined geosynchronous orbits (covering the Asia-Pacific region), and 3 satellites in geostationary orbits (covering China). All these orbit the earth about every 12 hours.

The European Union Navigation Satellite System (Galileo). Developed by European Union through the European Space Agency (ESA) and started offering services globally in 2016. Its space segment consists of at least 30 satellites in orbit (24 active, 6 spare) with altitudes of about 23,200 kilometers above the Earth's surface and orbits the earth every about 12 hours.

However, two regional satellite navigation systems have also emerged over a decade. These include the; Indian Regional Navigation Satellite System (IRNSS) which is owned and operated by the Government of India. It is as autonomous system designed to consist of seven satellites in its constellation. The IRNSS covers the Indian region and 1500 kilometers around Indian mainland (Leick, Rapoport, & Tatarnikov, 2015). The other is Japanese Quasi-Zenith Satellite System (QZSS) owned and operated by Japan government. It provides regional communication services and positioning information for the navigation purposes. The focus of this system is for the Japan region, but it provides service to the entire Asia-Oceania region. Japan declared the official start of QZSS services in 2018 with 4 operational satellites and plans to expand the constellation to 7 satellites for autonomous capability (J. Chen, Wang, Wang, Ding, & Zhang, 2020).

The GNSS is used to track the dynamic processes of the Earth's crust, tracking changes in the absolute sea level, and tracking plate tectonics movements. In addition, GNSS is utilized in atmospheric science, glaciology, geology, geophysics, natural hazards, and ocean surface monitoring with GNSS reflectometry (Bai, 2005; Long et al., 2021).

2.2.1 The GPS signal

The GPS signal contains both ranging codes and navigation messages using binary phase-shift keying (BPSK). The GPS uses code-division multiple access (CDMA), which is the use of distinct range codes to distinguish satellites operating on the same frequency, but with a restricted number of central frequencies (Hofmann-Wellenhof, Lichtenegger, & Collins, 2012). The ranging codes are called chipping codes with regard to CDMA. They are also known as pseudorandom noise or pseudorandom binary sequences because they are predictable but statistically similar to noise. The ranging signals are used to determine the distance between the GPS receivers and the satellites (Nicholas, 2016). The navigation messages contain information about the time and status of the entire satellite constellation, known as the almanac, as well as ephemeris data, which is used in trilateration to determine each satellite's position in orbit (Misra & Enge, 2006).

For the ranging codes and navigation message to travel from the satellite to the receiver, they must be modulated onto a carrier wave. The GPS was originally designed, to operate in frequency band referred to as L - band which is a portion of the radio spectrum between 1 GHz and 2 GHz (M. Wang, Shan, Zhang, & Huan, 2021). The carrier signal utilizes two frequencies; 1575.42 MHz, called $L1$ and 1227.60 MHz, called $L2$ (Hofmann-Wellenhof et al., 2012). The GPS satellite system uses a broad spectrum, which means that the information signal is sent across the whole carrier bandwidth. A high bit-rate pseudorandom sequence that is specific to each satellite is utilized to encode the low bit-rate message data using the Code Division Multiple Access approach (Spilker Jr, 1978). Two pseudorandom codes are modulated onto the carrier frequencies. The Coarse Acquisition Code (C/A-code) for civilians is modulated onto the L1 carrier only, while the Precise Code (P-code), for U.S. military use, is modulated onto both $L1$ and $L2$ carriers (Spilker Jr, 1978).

In order to ascertain position, velocity, and/or timing, a GPS receiver examines the GPS signals that it receives via its antenna. A variety of chaining operations may be used to achieve different results. At the antenna, the signal is amplified, down converted to base-band or intermediate frequency, filtered (to eliminate frequencies outside the specified frequency range for the digital signal that might alias into it), and digitalized. As long as the satellite is in use, the receiver for each employed satellite must first acquire the signal and then track it. In the many receivers, if not all of them, these processes are carried out in the digital domain presented in Figure 2.2. Tracking is a phase-locked loop because it involves constantly modifying the estimated frequency and phase to as closely resemble the received signal as feasible. Tracking is done as long as the target satellite is in use,

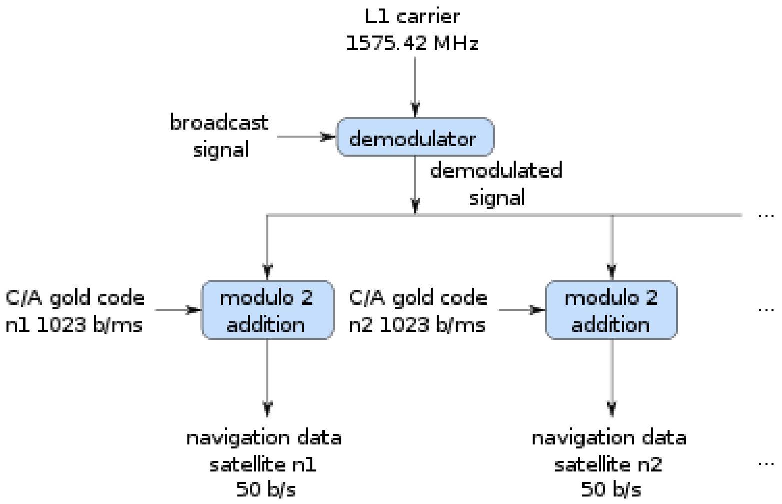


Figure 2.2: Demodulating and Decoding GPS Satellite Signals using the Coarse Gold code. (Gold, 1967)

whereas acquisition is done to begin using a certain satellite.

2.3 Sources of Errors in GPS and their corrections

There are many GPS errors some of which are given in subsections that follow.

2.3.1 Satellite Orbital Variations (Ephemeris)

Although the satellite orbits are extremely stable and predictable, some inaccuracies (perturbations) in the satellite orbit do exist. These are caused by gravitational effects of the Earth and Moon and the pressure of solar radiation. The orbit errors are reduced through a technique known as double differencing of ranges between pairs of satellites and receivers (Misra & Enge, 2006).

2.3.2 Satellite Clock Error

Timing errors due to inaccuracies in both the satellite and receiver clocks, as well as relativity effects, can result in position errors of up to two metres (Nicholas, 2016). To keep the satellite clock error as small as possible, every GPS satellite carries an atomic clock. Those atomic clocks still experience noise and clock drift errors and because of this, a correction need to be made. The clock errors can also be reduced through a technique known as double differencing of ranges between pairs of satellites and receivers (Misra & Enge, 2006).

2.3.3 Multipath

The signal transmitted by a GPS satellite can reach the receiving antenna in a number of ways other than the direct path between satellite and receiver. It takes place when the signal from the satellite reaches the receiver after one or more reflections or scattering from the ground or an object. This causes an error in the pseudo-range measurement. We call this Multipath error. Because of interference of the reflected signals, multipath slightly changes the amplitudes and phases of the direct path signals. Multipath can be minimized by: using choke ring antennas and by strategically placing antennas (Misra & Enge, 2006).

2.3.4 Dilution of Precision

Because of the quantity and relative geometry of satellites in view at the moment of calculation, a phenomenon known as Geometric Dilution of Precision (GDOP) reduces the precision of a position fix. The specified number represents the factor that is used to multiply the system range errors in order to get the overall system error. A subset of GDOP called Position Dilution of Precision (PDOP) affects altitude, longitude, and latitude. A lot of GPS receivers can give an approximation of PDOP (Nicholas, 2016).

2.3.5 Atmospheric Errors

The earth is surrounded by an atmosphere which is composed of distinct layers as presented in Figure 2.3. When GNSS signal propagates through this atmosphere from the satellites to receiver on the Earth, the signal is refracted and therefore delayed (Andrei & Chen, 2009). The atmospheric delay mainly occurs in the neutral atmosphere and in the ionosphere. These two layers cause two different kinds of delays. The ionosphere is that part of the upper atmosphere where free electrons occur in sufficient density to have an appreciable influence on the propagation of radio frequency electromagnetic waves. It extends from about 50 km above Earth's surface to altitudes of about 1000 km atmosphere (Bai, 2005). On the other hand, neutral atmosphere is the non-ionized part of the atmosphere. This region delays GNSS signals travelling through the same path equally, irrespective of their frequencies. It comprises the stratosphere and troposphere. The stratosphere extends approximately 20 - 60 kilometers above the Earth's surface (El-Mowafy & Lo, 2014). The neutral atmosphere is non dispersive in nature for radio waves up to frequencies of about 15 GHz, which is far higher than the 1.6 GHz and 1.2 GHz frequencies, the $L1$ and $L2$ frequencies transmitted by the GPS satellites, respectively (Hofmann-Wellenhof et al., 2012). The GNSS signal delay through the neutral atmosphere is, therefore, mainly caused by atmospheric water vapor, which is mostly concentrated in the troposphere.

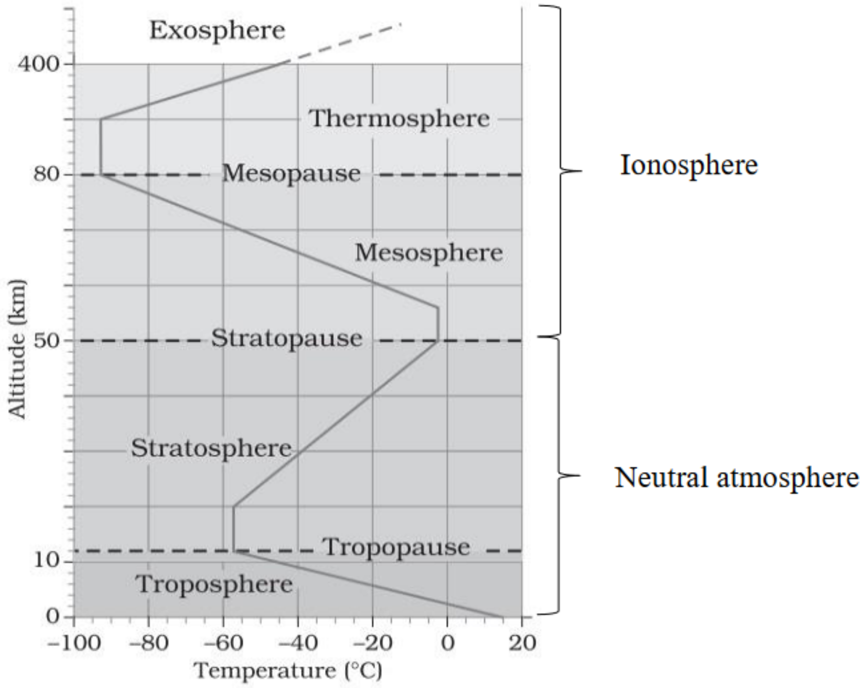


Figure 2.3: Layers of Earth's Atmosphere
Adapted from https://byjus.com/free-ias-prep/ncert-notes-structure-of-atmosphere.

2.3.5.1 Ionospheric delay

This occurs in the ionosphere. The ionospheric delay is a function of the Total Electron Content (TEC) and signal frequency, therefore, a dispersive medium for electromagnetic waves (Hofmann-Wellenhof et al., 2012). The sun's radiation intensity (insolation), determine the density of ions and electrons in this region. Gases in this region are heated by the Sun until they lose one or two electrons, producing a pool of electrically charged particles. Among these, the free electrons affect the propagation of the GNSS signals, delaying signal propagation to receivers on earth's surface. Therefore, the ionosphere plays a role in our everyday communications and navigation systems. The ionospheric delays can be predicted and being frequency dependent, can be effectively eliminated by using dual frequency receivers in a process called ionospheric-free combination (Hofmann-Wellenhof et al., 2012). Ionosphere-free linear phase combinations are linear combinations of the carrier phase data that are not affected by ionospheric delays. The estimate of the ionospheric group delay at $L1$ frequency and $L2$ frequency using the code measurements are:

$$I_{L1} = \frac{f_{L2}^2}{f_{L1}^2 - f_{L2}^2} (\rho_{L2} - \rho_{L1}) \tag{2.1}$$

$$I_{L2} = \frac{f_{L1}^2}{f_{L1}^2 - f_{L2}^2} (\rho_{L2} - \rho_{L1}), \tag{2.2}$$

where I_{L1} and I_{L2} are the ionospheric delays on the $L1$ and $L2$ frequency bands respectively, f_{L1} and f_{L2} are the frequencies of the $L1$ and $L2$ frequency bands, the ρ_{L1} and ρ_{L1} are the code measurements on band $L1$ and $L2$, respectively (Sunehra, 2013).

2.3.5.2 Tropospheric Delay

This occurs in the troposphere. The troposphere is the lower part of the earth's atmosphere where temperature decreases with an increase in altitude. The height of the troposphere extends to about 9 km over the poles and upto about 16 km near the equator (Sunehra, 2013). The troposphere is composed of both moist (wet) and dry (hydrostatic) components, which contribute differently to the overall tropospheric delay. Study by Langley, Wells, and de Brito Mendes (1995) shows that propagation of GPS signals is affected by the presence of neutral atoms and molecules in the troposphere.The troposphere causes a delay in both the code phase and carrier phase observations. Delay caused by the troposphere originates from the sum of the propagation delay and signal bending which induces an excess propagation path length on the GNSS signals (Mendes et al., 2000). Unlike the ionosphere, the troposphere is non-dispersive at GPS frequencies. Since the tropospheric delay is not frequency dependent, it cannot be canceled out by using dual frequency measurements. However, it can be modeled accurately using several empirical models including Hopfield (1969) and Saastamoinen (1972) models, among others. These models use meteorological data including surface temperature, pressure and relative humidity, and satellite elevation angle to compute the tropospheric delay. The total tropospheric delay can be split into two parts: the dry (hydrostatic) component (ZHD), which constitutes about 90 % of the total refraction, and the wet component (ZWD), which constitutes the remaining 10 % (Seeber, 2003). The former component depends mostly on dry gasses and is computed using the surface pressure value obtained from the station while wet component mostly dependents on moisture content. The variation in the tropospheric delay is much larger in the wet delay, as this component varies more both spatially and temporally. The tropospheric delay depends on the refractive index of the troposphere. The refractive index, n to a medium is defined as;

$$n = \frac{c}{v} \tag{2.3}$$

where c and v , are the speeds of GNSS signal in vacuum and in the medium interacting with the signal (Mendes et al., 2000).

The increase in time required to cover the excess path length is called time delay. Figure 2.4 shows the true path, G along which the signal propagates and the ideal path, S along which the signals would travel if the troposphere is a vacuum. Time delay in figure 2.4 is given by

$$\Delta T = T_G - T_S \tag{2.4}$$

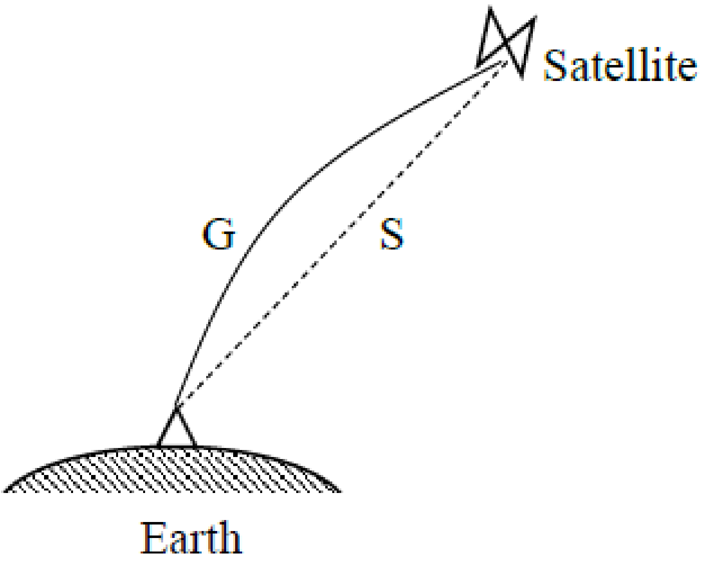


Figure 2.4: GPS signal paths

where T_G and T_S , are the times taken by the GNSS signal to propagate along the true path and the ideal path respectively.

Therefore, integration along the signal path gives:

$$\Delta T = \int_{TP} \frac{1}{v} dS_{TP} - \int_{IP} \frac{1}{c} dS_{IP} \tag{2.5}$$

The excess path length in figure 2.4 is given as follows:

$$\Delta L = c\Delta T \tag{2.6}$$

According to Mendes et al. (2000), the excess path length, ΔL is computed through integration along the signal path through the troposphere as follows:

$$\Delta L = \int_{TP} n dS_{TP} - \int_{IP} dS_{IP} \tag{2.7}$$

where dS_{TP} and dS_{IP} are the distances along the true path and ideal path, respectively and n is the refractive index of the medium interacting with the signal.

Therefore;

$$\Delta L = \int_{TP} n dS_{TP} - \int_{TP} dS_{TP} + [\int_{TP} dS_{TP} - \int_{IP} dS_{IP}] \tag{2.8}$$

The bending term (second term on the right-hand side of equation) is much smaller, about 1cm or less, for paths with elevation greater than 10° (Leick et al., 2015). Assuming the atmosphere is spherically symmetric (in absence of horizontal gradients in n), and the propagation is in the zenith, the signal path is a straight line and the bending term varnishes (Leick et al., 2015).

$$\Delta L^o = \int_{TP} n dS_{TP} - \int_{TP} dS_{TP} \tag{2.9}$$

$$\Rightarrow \Delta L^o = \int_l (n-1)\, dl \tag{2.10}$$

where l is the zenith path, dl is the distance along the signal path in the zenith direction. The refractivity, N, of the medium and its effect on propagation of GPS signals is vital to GNSS measurements. The tropospheric delay is the integral of refractivity index N along the entire path taken by the GPS signals through the atmosphere (from antenna height to the top of the neutral atmosphere). So, $(n-1)$ can be expressed in terms of refractivity (N) of the troposphere as shown in equation;

$$N = 10^6(n-1) \tag{2.11}$$

$$\Rightarrow n-1 = 10^6 N \tag{2.12}$$

Now integrating from the GNSS antenna height to point at infinity (tropopause) where the effects of the atmosphere on the signal can be assumed to be zero, we obtain;

$$\Delta L^o = 10^{-6} \int_l N\, dl. \tag{2.13}$$

The signal refractivity of the neutral atmosphere is a function of pressure, temperature, and water vapor pressure. According to Rózsa, Weidinger, Gyöngyösi, and Kenyeres (2012) and Schueler et al. (2001), the total refractivity, N, for radio frequencies can be expressed as;

$$N = k_1 \frac{P_d}{T} Z_h^{-1} + \left(k_2 \frac{e}{T} + k_3 \frac{e}{T^2}\right) Z_w^{-1}, \tag{2.14}$$

where P_d is the partial pressure of dry gases in hectopascals (hPa), T is absolute temperature in Kelvin (K), e is the partial pressure of water vapor, Z_h^{-1} and Z_w^{-1} are the inverse compressibilities for dry air and water vapor respectively. The two terms Z_h^{-1} and Z_w^{-1} are almost constant in value. These terms differ from unity by small parts in thousands and can all be ignored (Schueler et al., 2001; Suresh Raju et al., 2007; Y. Wang et al., 2017). The constants; k_1, k_2 and k_3 are the atmospheric refractivity constants. Different authors have computed values of refractivity constants and their associated uncertainties (Smith & Weintraub, 1953; Davies, Vincent, & Beresford, 1985; Bevis et al., 1994; Rüeger, 2002). The choice taken for this study is by Rüeger (2002) which gives the weighted mean

Table 2.1: Estimates of refraction constants

Constant	Estimate
k_1	77.6848 ± 0.0094 K/hPa
k_2	71.2152 ± 1.3 K/hPa
k_3	375463 ± 0.000007694 K^2/hPa

values of the refraction constants and their uncertainties as outlined in table 2.1; The first term of N represents the sum of the hydrostatic component and the second term represents the wet component due to effects of water vapor. Therefore, N can be expressed as the sum of hydrostatic (N_h) and wet refractivity (N_w).

$$N = N_h + N_w. \tag{2.15}$$

Schueler et al. (2001), further showed that the atmospheric refractivity can be expressed as

$$N = N_h + N_w = k_1.R_h.\rho + k_2'\left(\frac{e}{T}\right) + k_3\left(\frac{e}{T^2}\right) \tag{2.16}$$

where ρ is total mass density of air in Kgm^{-3}, $R_h = 287.06 \pm 0.01 Jkg^{-1}K^{-1}$ is the specific gas constant for dry air, T is absolute temperature, Z_w^{-1} are the inverse compressibility for dry air. The constant k_2' is given by

$$k_2' = k_2 - k_1\left(\frac{M_w}{M_h}\right) \tag{2.17}$$

where $k_2' = 22.97 \pm 2.2$ K/hPa (Rüeger, 2002), M$_w$ and M$_h$ are molar masses of the water vapor and dry air, with values of 0.00289644 kg/mol and 0.018016 kg/mol, respectively; while e is partial pressure of water vapor given by;

$$e = \rho_w\,.\,R_w\,.T \tag{2.18}$$

where ρ_w and R_w are the total mass density and specific gas constant for water vapor, respectively. Therefore,

$$N_h = k_1\left(\frac{P_d}{T}\right) \tag{2.19}$$

$$N_w = k_2\left(\frac{e}{T}\right) + k_3\left(\frac{e}{T^2}\right) \tag{2.20}$$

Also, the tropospheric delay

$$\Delta L^o = L_h^o + L_w^o \tag{2.21}$$

$$\Delta L^o = 10^{-6}\int_l N_h\,dl + 10^{-6}\int_l N_w\,dl \tag{2.22}$$

The magnitude of total tropospheric delay depends on the signal path through the neutral atmosphere. It is lowest in the zenith direction and increases as the satellite elevation angle is reduced. The slant delay is converted to zenithal path delay called the Zenith Tropospheric Delay (ZTD) using a mapping function, such as the Vienna Mapping Function (VMF), depending on the elevation angle of the satellite as follows; (Tuka & El-Mowafy, 2013).

$$\Delta L^{o} = m_h(\theta) \times ZHD + m_w(\theta) \times ZWD, \tag{2.23}$$

where θ is the satellite elevation angle, m_h and m_w are the hydrostatic and wet mapping functions and ZHD, and ZWD are the hydrostatic and wet delay components in zenith direction, respectively. The Zenith Tropospheric Delay (ZTD) is therefore, a function of zenith hydrostatic delay (ZHD) due to dry gases, and zenith wet delay (ZWD) caused by water vapor in the troposphere (Bevis et al., 1994; He, Shen, Wan, & Li, 2020). Therefore,

$$ZTD = ZHD + ZWD. \tag{2.24}$$

2.4 Mapping Functions

The mapping functions are aimed at attaining precise and accurate mapping for radio space geodetic applications such as Global Navigation Satellite System(s) (GNSS). The GNSS measurement is not conducted in the zenith direction; instead, it is taken along the slant direction, which is the line of sight from the ground station to the satellite as presented in Figure 2.5. It is necessary to map functions for neutral atmospheric delays in order to take satellite elevation dependence into account. The mapping function maps the slant delays to zenith direction and increases as the satellite elevation angle is reduced (Troller, 2004). So, hydrostatic and wet mapping functions are generally functions of elevation defined as:

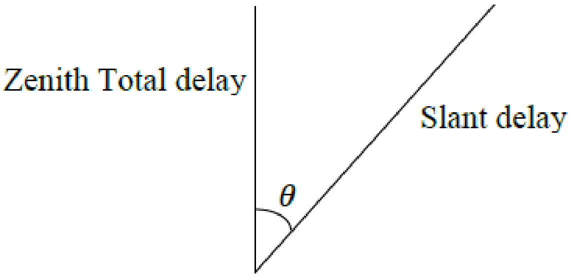


Figure 2.5: Relationship between the slant delay and the zenith total delay.

$$mf(\theta) = \frac{\Delta L(\theta)}{\Delta L^{z}} \tag{2.25}$$

where ΔL^z is the delay in zenith direction, and $\Delta L(\theta)$ is the delay for any given elevation angle θ, for both hydrostatic and wet delays.

Previous research shows that mapping functions that adopt information from ray - tracing through numerical weather prediction models such as ; Vienna mapping function 1 (VMF1) (Böhm, Niell, Tregoning, & Schuh, 2006), Herring mapping function (Herring, 1992), Niell Mapping Functions (NMF) (Niell, 1996) and Vienna mapping function VMF3 released by Landskron and Böhm (Landskron & Böhm, 2018) among others. The VMF1 mapping functions are still considered the most accurate over the past years and their use is recommended by the International Earth Rotation and Reference Systems Service for high accuracy geodesy applications (Petit & Luzum, 2010; Feng, Li, Yan, Zhang, & Barriot, 2020). Mapping functions were first formally adopted for Very Long Baseline Interferometry (VLBI) data analysis in the form of a continued fraction (Marini, 1972). This is given as;

$$m(\theta) = \frac{1 + \frac{a}{1+\frac{b}{1+c}}}{sin\theta + \frac{a}{sin\theta+\frac{b}{sin\theta+c}}} \tag{2.26}$$

where θ represents the elevation angle, while, a, b, c, ... are the profile dependent coefficients determined from atmospheric conditions and specialized for the hydrostatic and wet mapping functions.

This format has been adopted by Herring (1992), and Böhm et al. (2006) by extending the form of the continued fraction to three terms. Niell (1996) modified Herrings continued fraction mapping function coefficients by putting into consideration the temporal changes and geographical location rather than the surface meteorology parameters. The coefficients are given by the mean and seasonal varying values as;

$$\alpha(\varphi, DOY) = \alpha(\varphi)_{average} - \alpha(\varphi)_{amplitude} . \cos\left(\frac{DOY - DOY_w}{365.25}\right)2\pi, \tag{2.27}$$

where α is the mapping function coefficients either a, b, or c, separated into average value and amplitude, φ is site latitude in radians, DOY is day of year, DOY_w is day of year for maximum winter, 28 for Northern Hemisphere and 211 for Southern Hemisphere.

2.5 Tropospheric Delay Models

The total tropospheric delay can be divided into the hydrostatic (ZHD) and wet components (ZWD) in the zenith direction (Saastamoinen, 1972). Since ZHD and ZWD are affected differently by the troposphere, they are modeled differently. Tropospheric models fall into two primary categories: The first type of models are the traditional ones, which require real-time surface meteorological observations such as temperature and pressure at the user's location to compute the tropospheric delays. These traditional empirical methods have been used for the estimation of the ZHD with an accuracy of greater than 90 %, (Tuka & El-Mowafy, 2013). The ZHD models include:the Saastamoinen model (Saastamoinen, 1972), Hopfield model (Hopfield, 1969), Baby model (Berrada Baby, Gole, & Lavergnat, 1988), Davis hydrostatic model (Davis et al., 1985), Askne model (Askne & Nordius, 1987). Even though the fluctuation of the water vapor makes it challenging to compute ZWD accurately, some models have been developed to approximate this delay. Examples of ZWD models include:the Saastamoinen model (Saastamoinen, 1972), Hopfield model (Hopfield, 1969), Baby model (Berrada Baby et al., 1988), Berman model (Berman, 1976), Ifadis model (Ifadis, 1986) and Askne model (Askne & Nordius, 1987).

The second type of models are blind models. The reliance of traditional models on real-time surface meteorological data limits their real-time application as real-time meteorological data is difficult to get (Liu, Chen, Sun, & Liu, 2017). The blind tropospheric models were therefore developed as a solution to overcome these limitations. Blind models estimate meteorological characteristics derived from numerical weather prediction (NWP) models by using the user's location and time information (Tuka & El-Mowafy, 2013). Examples include: the UNB3m model (Leandro, Langley, & Santos, 2008), the EGNOS model (Penna, Dodson, & Chen, 2001), GPT2w model (Böhm, Möller, Schindelegger, Pain, & Weber, 2015) and its updated version, the GPT3 model (Landskron & Böhm, 2018), space Agency GAL-TROPO (Martellucci, 2012), TropGrid (Krueger, Schueler, Hein, Martellucci, & Blarzino, 2004) and its update TopGrid2 (Schüler, 2014). However, the atmospheric constituents in a given area vary according to different climatic conditions, therefore, tropospheric effect on GNSS signals also varies, which affects how well these empirical prediction models (both traditional and blind) perform at user locations (Tuka & El-Mowafy, 2013).

2.5.1 Saastamoinen Tropospheric Models

The ZHD model is a function of surface pressure P_s and geographical location of GNSS receiver at a specific site. This model is dependent on the choice of refractivity constants and on the modeling of the height and latitude dependence of the acceleration of gravity (Mendes et al., 2000). According to Saastamoinen (1972), the refractivity constants of Essen and Froome (1951) were used to get the following expression for the zenith hydrostatic delay.

$$ZHD = \frac{0.0022768 P_s}{1 - 0.00266 \cos 2\phi - 0.00000028 h}, \tag{2.28}$$

where h is the height of the surface above the geoid and ϕ is the latitude of the station. The P_s is the corresponding air pressure at the station in hPa. The pressure can be computed from an empirical model, a numerical weather prediction model (NWP) or with in-situ measurements at the observing station.

(Saastamoinen, 1972) also established an empirical model to calculate ZWD and is expressed as;

$$ZWD = 0.0027768 \times (1255 + 0.05T)\,\frac{e}{T}, \tag{2.29}$$

where T is absolute temperature and e is the partial pressure of water vapor which can be calculated from the relative humidity (RH) measurements as:

$$e = 6.108\,\frac{RH}{100}\Big(\frac{17.15T - 4684}{T - 38.45}\Big) \tag{2.30}$$

2.5.2 Hopfield Tropospheric Models

The Hopfield model uses a quartic model for the determination of the ZHD which is an empirical representation of the dry refractivity as a function of the height of the observing station and the assumption that the “dry atmosphere is approximated by a single polytropic layer extending from the surface to an altitude of approximately 40 km” (Janes et al, 1989). The zenith hydrostatic model by Hopfield is expressed as (Mendes et al., 2000).

$$ZHD = 10^{-6}\,N_{ds}\,\frac{H_d^e - H_s}{5} \tag{2.31}$$

where H_s is the height of the receiver, N_{ds} is the dry refractivity factor at the surface, H_d^e is the equivalent height which can be expressed as:

$$H_d^e = 40136 + 148.72(T_s - 273.16) \tag{2.32}$$

Mendes et al. (2000) and Seeber (2003) reported that the ZHD model by Hopfield can also be expressed as:

$$ZHD = \frac{10^{-6}}{5}\Big[40136 + 148.72(T_s - 273.16)\Big] \times 77.64\Big(\frac{P_s}{T_s}\Big), \tag{2.33}$$

where P_s is surface air pressure in hectopascal (hPa) at the station and T_s is surface temperature in kelvin (K). Hopfield (1969) developed a wet delay model expressed as:

$$ZWD = \frac{10^{6}}{5} \times \Big[-12.96T_s + 3.718 \times 10^{5}\Big] \times 1100 \times \Big(\frac{P_v}{T_s^2}\Big), \tag{2.34}$$

where P_v is partial water pressure in hectopascal (hPa) at the station, computed from surface temperature T_s and relative humidity (RH).

2.5.3 Baby Tropospheric Models

Another model dependent on the choice of refractivity constants and on the modeling of the height and latitude dependence of the acceleration of gravity is the baby et al. model (Berrada Baby et al., 1988). Bean and Dutton (1966) refractivity constant, k_1, was used to developed the . Baby ZHD model presented as:

$$ZHD = \frac{0.0222277\, P_s}{g_s}\left[1 + \frac{2}{r_s\,\sigma\,(\mu\ +\ 1)}\right], \tag{2.35}$$

where g_s is the gravity at the surface, r_s is the mean geocentric radius of the station in meters (m). The parameters, σ and μ are given by the following expression respectively;

$$\sigma = \frac{\alpha}{T_s}\,; \qquad \mu\ =\ \frac{g_s}{R_d\,\alpha}\left[1 - \frac{2}{r_s\,\sigma}\right]. \tag{2.36}$$

In which α is the lapse rate, T_s is the surface temperature and R_d is the specific gas constant for dry air.

2.5.4 Ifadis Tropospheric Models

Ifadis (1986) established a ZWD model expressed as;

$$ZWD = 0.55\times 10^{-2} - 0.88\times 10^{-4}\left(P_s - 1000\right) + 0.27\times 10^{-4} \times P_v + 2.771\left(\frac{P_v}{T_s}\right), \tag{2.37}$$

where P_s is surface air pressure in hectopascal (hPa) at the station, P_v is partial water vapor pressure and T_s is surface temperature in kelvin (K).

2.5.5 Askne Tropospheric Model

Askne and Nordius (1987) established an empirical model for estimating ZWD expressed as:

$$ZWD\ =\ 10^{-6}\left[\left(k_2' + \frac{k_3}{\mathrm{T\,m}}\right)\left(\frac{R_d . e}{(\lambda\ +\ 1)g_m}\right)\right], \tag{2.38}$$

where k_2' and k_3 are empirically determined coefficients, R_d denotes the specific gas constant for the dry constituents and g_m is the gravity constant g_m is set as 9.784ms^{-2}. The variable e is the water vapor pressure at the site, λ is the water vapor lapse rate, $\mathrm{T_m}$ is the weighted mean temperature of the water vapor in kelvin, (Böhm et al., 2015).

Study by Mendes et al. (2000) shows that Saastamoinen hydrostatic model is superior in accuracy compared to other ZHD models. This makes it the best performing and widely used model for ZHD computation and therefore the best choice for this research.

2.5.6 EGNOS and UNB3m Tropospheric Models

The EGNOS and UNB3m models are based on the estimation of five meteorological parameters such as: pressure (P), temperature (T), water vapour partial pressure (e) , temperature lapse rate (β) and water vapor pressure height factor (λ) derived from the U.S. Standard Atmosphere Supplements, 1966 using user location (latitude and height) and day of year (Penna et al., 2001). These meteorological parameters are stored in a look-up table and account for the annual average and amplitude/seasonal variations (Leandro et al., 2008; Penna et al., 2001). The annual average and amplitude of each parameter (P_o ,T_o , RH_o , e_o , β_o and λ_o) can be computed as follows:

$$Avg_{\phi} = \begin{cases} \varepsilon_{15}, & \phi \leq 15 \\ \varepsilon_i + \frac{\varepsilon_{i+1}}{15}.(\phi - lat_i), & 15 < \phi < 75 \\ \varepsilon_{75}, & \phi \geq 75 \end{cases} \tag{2.39}$$

where ϕ is the latitude of interest in degrees, ε is the computed amplitude of the meteorological parameter, i is the index of the nearest lower tabled latitude and lat represents the latitude from the look-up table (Leandro et al., 2008). Each meteorological parameter value (P, T, e, β, λ) is then estimated for the desired day of year (DOY) as follows:

$$r(\phi, DOY) = Avg_{\phi} - Amp_{\phi}.\cos(DOY - D_{min})\frac{2\pi}{365.25}, \tag{2.40}$$

where $r(\phi, DOY)$ represents the calculated parameter value for latitude ϕ and DOY, Amp_{ϕ} represents amplitude ε of the meteorological parameter, D_{min} is 28 for northern latitudes and 211 for southern latitudes.

After computing the meteorological parameters, the tropospheric delays ZHD and ZWD are then computed based on Saastamoinen model as:

$$ZHD = 10^{-6}.k_1.\frac{R_d.P}{g_m}.\big(1 - \frac{\beta.H}{T}\big)^{\frac{g}{R_d.\beta}}, \tag{2.41}$$

$$ZWD = \frac{10^{-6}.k.R_d}{g_m.(\lambda + 1) - \alpha.R_d}.\frac{e}{T}.\big(1 - \frac{\beta.H}{T}\big)^{\frac{(\lambda+1)g}{R_d.\beta} - 1}, \tag{2.42}$$

where k is an arbitrary constant for refractivity.

For UNB3m model, k is computed from:

$$k = \mathrm{T_m}.k_2' + k_3 \tag{2.43}$$

where $\mathrm{T_m}$ is the weighted mean temperature of water vapor in (kelvin) expressed as:

$$\mathrm{T_m} = T\left[1 - \frac{\beta.R_d}{g_m.(\lambda + 1)}\right]. \tag{2.44}$$

For the EGNOS model, $k = k_2$ with value 382000, R_d is the specific gas constant of dry air, k_1 , k_2' and k_3 are refractivity constants with values 77.60 K/mbar, 16.6 K/mbar and 377600 K^2/mbar, respectively, H is the orthometric height in meters above sea level, g is the surface acceleration due to gravity (9.80665ms^{-2}), g_m is the gravity acceleration at the centroid of the atmospheric column above the GNSS antenna (in ms^{-2}).

For the EGNOS model, g_m is set as constant 9.784ms^{-2}. For the UNB3m model g_m is computed as:

$$g_m = 9.784(1 - 0.00266\cos\phi - 2.8 \times 10^{-7}H). \tag{2.45}$$

For UNB3m model, e values are computed from RH as:

$$e = \frac{\mathrm{RH}}{100} \times e_s \times f_w, \tag{2.46}$$

where e_s is the saturation vapor pressure given by:

$$e_s = 0.01{\times}exp\left[1.2378847{\times}10^{-5}T^2{-}1.9121316{\times}10^{-2}T{+}33.93711047{-}6.3431645{\times}10^3T^{-1}\right]. \tag{2.47}$$

While f_w is the enhancement factor which can be determined from:

$$f_w = 1.00062 + 3.14 \times 10^{-6}\mathrm{P} + 5.6 \times 10^{-7}(T - 273.15)^2 \tag{2.48}$$

2.5.7 GPT Tropospheric Models

The Global Pressure Temperature (GPT) empirical model (Böhm, Heinkelmann, & Schuh, 2007), and its revisions GPT2 (Lagler, Schindelegger, Böhm, Krásná, & Nilsson, 2013) followed by GPT2w (Böhm et al., 2015) and the GPT3 model recently developed by (Landskron & Böhm, 2018) provide internally the mean values together with annual and semi-annual amplitudes of various surface meteorological parameters for calculating the tropospheric delays. These parameters include: pressure (P), temperature (T) and its lapse rate, water vapour pressure (e) and its decrease factor (λ), weighted mean temperature

($\mathrm{T_m}$), as well as hydrostatic (a_h) and wet(a_w) mapping function coefficients of the VMF1 and VMF3 respectively (Böhm et al., 2006).

These parameters are computed based on the analysis of 37 monthly mean pressure level profiles of the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis for a period of 10 years (2001-2010). Though both GPT2w and GPT3 models are the advanced versions of the GPT series, GPT3 is the most recent version and refinement of the GPT2w model (Landskron & Böhm, 2018). These models are both provided on a regular 1° S x 1° S and 5° S x 5° S global grids. Each meteorological parameter is computed from equation (2.49) using the date of observation and station coordinates latitude (ϕ), longitude, ellipsoidal height (h)) as input:

$$r(t) = A_o + A_1 cos\Big(\frac{DOY}{365.25}2\pi\Big) + B_1 sin\Big(\frac{DOY}{365.25}2\pi\Big) + A_2 cos\Big(\frac{DOY}{365.25}4\pi\Big) + B_2 sin\Big(\frac{DOY}{365.25}4\pi\Big), \tag{2.49}$$

where, A_o is the mean value, (A_1 , B_1) and (A_1 , B_2) are the annual and semi-annual amplitudes of the selected parameters, r.

The zenith delays ZHD and ZWD are computed based on Saastamoinen model as revised by Davis et al. (1985) and Askne and Nordius (1987) model expressed as:

$$ZHD = \frac{0.0022768P_s}{1 - 0.00266\cos 2\phi - 0.00000028h} \tag{2.50}$$

$$ZWD = 10^{-6}(k_2' + \Big(\frac{k_3}{\mathrm{T\,m}}\Big)\Big(\frac{R_d}{(\lambda + 1)g_m}\Big)e. \tag{2.51}$$

2.6 Obtaining PWV from GNSS Observations

The GNSS techniques can be used to compute ZTD with high accuracy. The zenith hydrostatic delay (ZHD) and the zenith wet delay (ZWD) are the two elements that make up ZTD. While the second component is dependent on water vapor, the hydrostatic delay is dependent on air pressure . The ZHD can be accurately calculated with the empirical models such as the Saastamoinen model (Saastamoinen, 1972). After retrieving the ZTD, the ZWD is calculated by subtracting the Zenith Hydrostatic Delay (ZHD) from the ZTD;

$$\mathrm{ZWD} = \mathrm{ZTD} - \mathrm{ZHD}. \tag{2.52}$$

Following (Bevis et al., 1994), PWV is retrieved from GNSS observations as follows;

$$PWV = \Pi \cdot ZWD \tag{2.53}$$

where ZWD is the Zenith Wet Delay and Π is the water vapor conversion factor. Following (Davis et al., 1985), the conversion factor was determined by:

$$\Pi = \frac{10^6}{\rho_v R_v \left(\frac{k_3}{T_m} + k_2'\right)}, \tag{2.54}$$

where ρ_v is the density of liquid water (1000 kg^{-3}), R_v is the specific gas constant of water vapour (461.5 $Jkg^{-1}K^{-1}$), $k_3 = 3.739 \pm 0.004 \times 10^5 K^2 hPa^{-1}$ and $k_2' = 22.1 \pm 2.2K\ hPa^{-1}$ are refractivity constants while T_m is the weighted mean temperature. The weighted mean temperature (T_m) of the troposphere at the GNSS station serves as the formulation's primary parameter.

2.7 The Weighted Mean Temperature (T_m)

The GNSS-PWV retrieval from ZWD requires a conversion factor which is a function of T_m (Lu, Trenberth, Qin, Yang, & Yao, 2015; Suwantong et al., 2016). The accuracy of the conversion factor depends mainly on the precision of the T_m. Therefore, the T_m is a vital parameter in the ground-based GNSS atmospheric research. The T_m is an average temperature of atmosphere weighted by the pressure of water vapor (Bevis et al., 1992). It varies significantly based on location, height, season, and weather conditions, making it highly correlated with climate conditions (Jiang et al., 2019). It is possible to determine the T_m in a variety of methods, each with advantages and disadvantages. The most accurate approach is by numerical integration of vertical temperature and water vapor pressure profiles as expressed in (Davis et al., 1985).

$$T_m = \frac{\int \frac{e}{T}\, dh}{\int \frac{e^2}{T^2}\, dh} \tag{2.55}$$

where e is the partial pressure of water vapor (in hPa), T is the atmospheric temperature (in K) and dh is the the height difference between the discrete layers of air. Radiosondes and NWP datasets can be used to extract these data profiles but they are not suitable for real-time calculations (X. Wang, Zhang, Wu, Fan, & Cheng, 2016).

Secondly, a number of empirical models based on linear relationship between some surface meteorological parameters such as surface air temperature T_s, pressure (P) have been established to estimate T_m (Bevis et al., 1992; Jiang et al., 2019; B. Chen, Yu, Wang, Zhang, & Dai, 2021). Using 8718 radiosonde profiles across the United States for the region between Florida and Alaska in North America, Bevis established the first and simplest linear relationship between $\mathrm{T_s}$ and $\mathrm{T_m}$ (Bevis et al., 1992). This is expressed as:

$$\mathrm{T_m} = a + b \cdot \mathrm{T_s}. \tag{2.56}$$

Bevis et al. (1992) derived the linear relationship $\mathrm{T_m}$ = 70.2 + 0.72 $\mathrm{T_s}$ commonly called the Bevis model. Studies by Schueler et al. (2001), Ross and Rosenfeld (1997), Mendes et al. (2000), J. Wang et al. (2005) and Y. Yao et al. (2012) reported that Bevis model can be applied globally. Schueler et al. (2001) analyzed global-scale numerical weather fields collected by the National Center of Environmental Prediction (NCEP/ NOAA) from July 1999 to July 2000 to derive a linear global relation of $\mathrm{T_m}$ = 86.9 + 0.657$\mathrm{T_s}$. Mendes et al. (2000) used 32,467 radiosonde profiles from 50 unevenly spread stations around the world covering a latitude range of 62°S to 83°N for the year 1992 to deduce a relationship $\mathrm{T_m}$ = 50.4 + 0.789$\mathrm{T_s}$. Y. Yao et al. (2014) developed a global $\mathrm{T_m}$ model $\mathrm{T_m} = 0.8116\mathrm{T_s} + 43.69$ by modifying the Bevis model with data from 135 radiosonde stations over a ten-year period, which can improved the precision of $\mathrm{T_m}$ on a global and local scale.

However, studies by Ross and Rosenfeld (1997), Boutiouta and Lahcene (2009) Y. Yao et al. (2014), Sapucci (2014) and Jiang et al. (2019) among others showed that the coefficients of a and b are season and location dependent. Therefore, Bevis model that was developed using data from North America hardly represents climate globally and therefore inadequate when used to calculate $\mathrm{T_m}$ world wide especially in tropics. For this reason, some regional linear functions of $\mathrm{T_m}$ and $\mathrm{T_s}$ have been established to compute $\mathrm{T_m}$ precisely. Study by Sapucci (2014) over Brazil using dataset of approximately 78000 Radiosonde profiles collected at 22 stations during a 12- year period (1999-2010) shows that regional $\mathrm{T_m}$ model were more accurate when used than the global models. After conducting a comprehensive study on $\mathrm{T_m}$ utilizing meteorological data from 53 worldwide stations, Ross and Rosenfeld (1997) came to the conclusion that a site-specific model would be preferable to the region and global models.

Many other linear functions of $\mathrm{T_m}$ and $\mathrm{T_s}$ have been developed to meet the demands of different regions. Li et al. (1999) developed a linear regression equation of $\mathrm{T_m}$ and $\mathrm{T_s}$ in eastern China by using the radiosonde data of Eastern China to deduce that $\mathrm{T_m}$ = 44.05 + 0.81$\mathrm{T_s}$, which has a good performance in that region. Suresh Raju et al. (2007) derived an empirical model $\mathrm{T_m}$ = 0.75$\mathrm{T_s}$ + 62.6 for Indian zone using about 4104 RS measurements extending up to or above 10 km for a period of 1995-1997 from 8 stations spread over Indian sub-continent(8.5°N-32.6°N). Liou obtained the linear model coefficients as $\mathrm{T_m}$ = 1.07$\mathrm{T_s}$ + 31.5 to compare the PWV results using GPS, microwave radiometers, and radiosondes Taipei (Liou, Teng, Van Hove, & Liljegren, 2001). Suparta and Iskandar (2013) used RS from 15 radiosonde stations within latitude and longitude range 20°N-20°S and 95°E-156°E respectively for the year 2011 to developed a T_m model $T_m = 0.84T_s + 48$ for Western Pacific region and the results showed improved accuracy in estimation of GPS-derived PWV while using this model compared Global models. Using radiosonde data from 2005 to 2007, Boutiouta and Lahcene (2009) also discovered these coefficients as $T_m = 0.96T_s + 14.79$ and developed a T_m model for Algeria; Suwantong et al. (2016) used two years of AIRS and AMSU data to determine a local relationship describing Bangkok, Thailand as $T_m = 50.4 + 0.789T_s$.

Using data from three radiosonde sites in Hunan, China, L. Li et al. (2017) calculated the coefficients for the linear model between T_s and T_m as $T_m = 0.65T_s + 87.08$. Table 2.2 presents the commonly used global and regional empirical models. Elhaty et al. (2019) developed T_m model for Egypt using 3600 radiosonde profiles from four stations in Egypt between 2015 and 2016 which outperformed the Bevis global model. Using information from a radiosonde station in Guilin, China, Huang, Mo, Xie, et al. (2021) also established a linear model between T_s and T_m as $T_m = 0.62T_s + 101.01$. Several other regional and local models were developed for Europe, Johansson (1998); Australia Tregoning, Boers, O'Brien, and Hendy (1998); Germany, Solbrig (2000); Taiwan, Liou et al. (2001) among others.

Table 2.2: Commonly used existing T_m models and their Authors (the units of T_m and T_s are Kelvin).

T_m model	Equation	Author
Global 1	$Tm = 0.72T_s + 70.2$	Bevis et al. (1992)
Global 2	$T_m = 0.64T_s + 86.9$	Schueler et at. (2001)
Global 3	$T_m = 0.8116T_s + 43.69$	Yao et al. (2014)
Tropic 1	$T_m = 1.07T_s - 31$	Liou et al. (2001)
Tropic 2	$T_m = 0.75T_s + 62.6$	Suresh Raju et al. (2007)
Tropic 3	$T_m = 0.84T_s + 48$	Wayan et al. (2013)

Since the majority of GNSS sites lack meteorological sensors, in situ T_s readings are not accessible. Therefore, the calculations based on linear functions of T_m and T_s become invalid. As a result, certain NWP models also called blind models for T_m computation using only the site's coordinates and the time of observation have been established as another method (Y. Yao et al., 2012; Pany, Pesec, & Stangl, 2001). In the development of these models, the Earth's surface is divided into a series of grids according to latitude and longitude, then a trigonometric function is used to simulate the periodic variation characteristics of T_m over each single grid (Lagler et al., 2013; Y. Yao et al., 2014). These T_m models include the; Global Weighted Mean Temperature (GWMT) model, Global

Tropospheric Model (GTrop) model, UNB3m model, Global Weighted Mean Temperature-Diurnal (GWMT-D) model, Global Pressure Temperature (GPT) empirical models, among others (Böhm et al., 2007). Y. Yao et al. (2012), took seasonal and geographic variations into account and established the empirical Global Weighted Mean Temperature-Diurnal (GWMT) based on spherical harmonics, and solved well the problem of calculating $\mathrm{T_m}$ independent of measured meteorological parameters. They developed one meteorological factor model (GTm) and a multi - meteorological factor model (PTm) using radiosonde data of 135 global stations from 2005 to 2009. The results showed that PTm model performed better than the GTm model in computation of $\mathrm{T_m}$. Later, Y. B. Yao et al. (2013) took the seasonal and geographic variation characteristics of the $\mathrm{T_m}$ - $\mathrm{T_s}$ and $\mathrm{T_m}$ - $\mathrm{T_s}$, e_s relations into account to develop the GTm-I model and PTm-I model. The $\mathrm{T_m}$ accuracy greatly improved compared with the GTm model and PTm models, more so, in sea areas. Y. Yao et al. (2014), used spherical harmonics, considering seasonal and geographic variations and considering the semi-annual and diurnal variations of $\mathrm{T_m}$ recalculated the coefficients for an updated and very accurate one-/multi-parameter-based model-GTm-II using radiosonde data from 135 globally distributed stations over a period of 10 years.

The Global Pressure Temperature (GPT) empirical model (Böhm et al., 2007) and its revisions GPT2 (Lagler et al., 2013) followed by GPT2w (Böhm et al., 2015) may be used to compute $\mathrm{T_m}$ precisely. Recently, (Landskron & Böhm, 2018) established the GPT3 model which can generate $\mathrm{T_m}$ and $\mathrm{T_s}$ values on global grids of $5^\circ \times 5^\circ$ and $1^\circ \times 1^\circ$ resolutions which proves to be more accurate than the previous versions of GPT models (Landskron & Böhm, 2018). The variation characteristics of meteorological parameters over time for the GPT3 model are characterized by equation 2.49.

The applications and accuracy of the $\mathrm{T_m}$ blind models would be restricted without a vertical adjustment as the height differences between the station and the reference level do exist. The products from numerical weather prediction (NWP) models generally have higher spatiotemporal resolutions, such as reanalysis data from the European Center for Medium-Range Weather Forecasts (ECMWF) and the America National Centers for Environmental Prediction (NCEP), but they are typically updated with some time delay on the web and thus can't be applied to real-time or even near-real-time $\mathrm{T_m}$ estimation (Pany et al., 2001).

Chapter 3

Data and Methods

3.1 Area of Study and Datasets

The East African region being studied lies between latitudes 7° S and 4° N and longitudes 28° E and 44° E. This region is composed of several countries, such as Kenya, Tanzania, Uganda, Rwanda, and Burundi, and is part of the African tropical zone. The region is characterized by a warm and humid tropical climate, with average temperatures ranging between 25°C and 30°C throughout the year, due to its proximity to the equator (Camberlin, 2018). The high temperatures remain consistent throughout the year, with little seasonal variation of up to 5°C (Camberlin, 2018). The region is subject to diverse topography with a complex diversity of mountains, including steep, snow-capped mountains like Mount Rwenzori and significant low-lying plains. Several huge inland lakes, including Lake Victoria, rift valleys, and tropical forests, river basins like the River Congo Basin

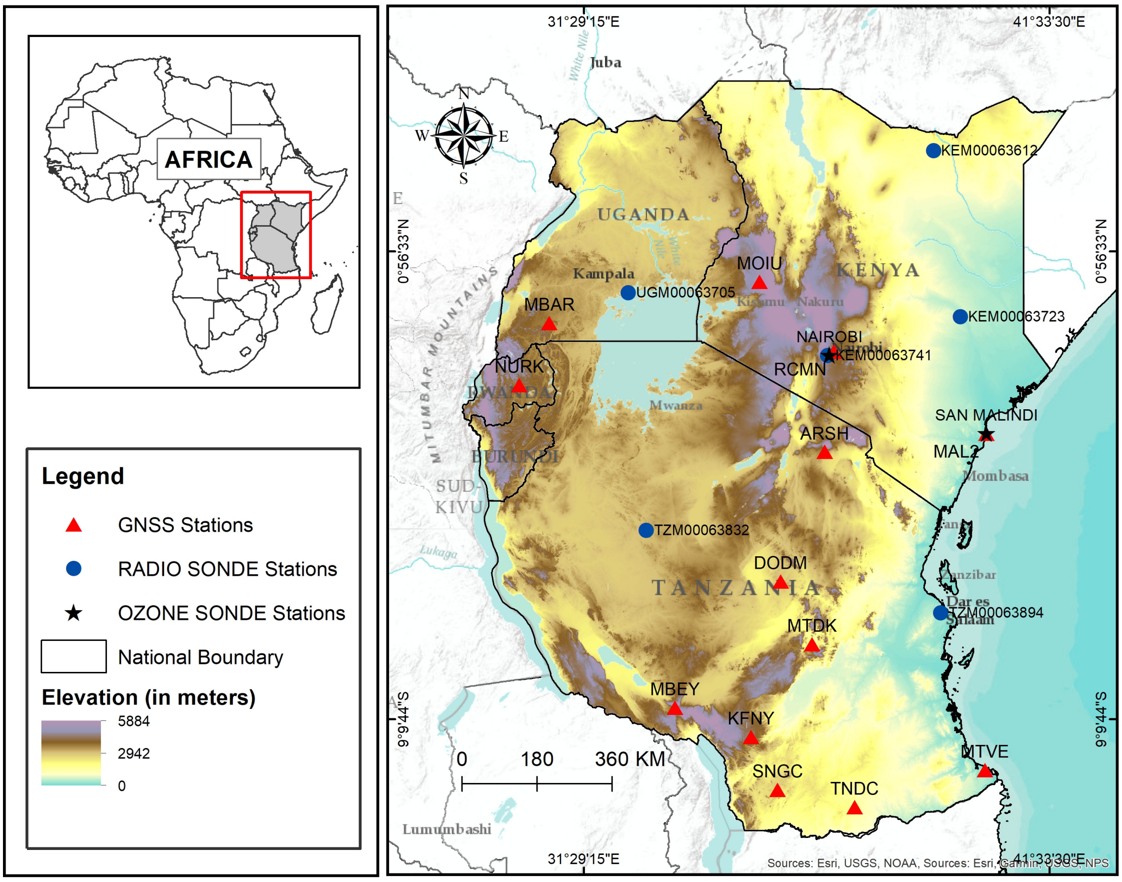


Figure 3.1: Distribution of GNSS , radiosonde and Ozonesonde stations in East Africa.

and the East African Highlands. All these do influence the climate and weather in this region.

The East African region experiences two rainy seasons, namely the long rains from March to May and the short rains from October to December. However, southern Tanzania experiences a single rainy season from November to April. During these rainy seasons, the region receives high amounts of rainfall ranging from 500 mm to 2000 mm per year, depending on the location (Camberlin, 2018). The distribution of rainfall varies across the region, with some areas receiving higher amounts of rainfall than others. The rainfall is vital for agricultural production, which is the primary source of livelihood for many communities in the region. However, excessive rainfall has led to flooding, posing threats to the safety of people and property. In this study, data from the 6 Radiosonde stations, 2 ozonesonde stations, 3 GNSS stations, and the ECMWF-Reanalysis 5 (ERA5) dataset are used. The corresponding stations are presented in figure 3.1. The details of ozonesonde and radiosonde stations, including the station name, geodetic coordinates, elevations, and available data are presented in table 3.1.

Table 3.1: Details of radiosonde, and ozonesonde stations used in this study.

Station	**Station Id & City**	**Country**	**Available Data**	**Data Type**	**Lon(°) E**	**Lat(°) N**	**Alt (m)**
63741	DAGO (Dagoretti)	Kenya	1971-2006	Radiosonde	36.760	-1.30	1798.3
63894	DARE(Daressalaam)	Tanzania	1971-2013	Radiosonde	39.200	-6.87	55.000
63705	ENTB (Entebbe)	Uganda	1971-2018	Radiosonde	32.450	0.05	1155.0
63723	GARI (GARISSA)	Kenya	1975-1999	Radiosonde	39.633	-0.47	138.00
63612	LOWD (Lodwar)	Kenya	1998	Radiosonde	35.640	3.12	506.00
-	MALI (Malindi)	Kenya	1999-2006	ozonesonde	40.190	-2.99	-6.000
-	NAIR (Nairobi)	Kenya	1998-2019	ozonesonde	36.800	-1.27	1795.0
63832	TOBR (Tabora)	Tanzania	1977-1988	Radiosonde	32.833	-5.08	1182.0

3.1.1 Radiosonde Data

Radiosondes (RS) are balloons equipped with a collection of sensors that are sent into the high atmosphere to measure the atmospheric variables such as absolute temperature (T), relative humidity (RH) and geopotential height (H) at every height level in its path. It also provides surface parameters, such as surface temperature (T_S) and pressure (Ps). The GNSS Meteorological Ensemble Tools (GMET) online service, available at http://gmet.users.sgg.whu.edu.cn/, provides air temperature, pressure, T_m, ZHD, ZWD, ZTD, and PWV data for all the Integrated Global Radiosonde Archive (IGRA) radiosonde stations (H. Zhang, Yuan, Li, & Zhang, 2019). The radiosonde data used in this study can also be obtained from the IGRA radiosonde stations, which can be accessed at https://www.ncdc.noaa.gov/data-access/. These observations are available at 00:00 and

12:00 Coordinated Universal Time (UTC). In this study, radiosonde observations from 1971 to 2018 were used to compute T_m values for establishing the empirical T_m models for East African tropical region. Table 3.1 presents the details of the RS data used in this study.

3.1.2 Ozonesonde Data

The Southern Hemisphere Additional Ozonesonde Network (SHADOZ) is a global collaboration established in 1998 with the aim of increasing the number of tropical ozone soundings in the troposphere and stratosphere (Thompson et al., 2017). The SHADOZ uses Electrochemical Concentration Cell (ECC)-type sensors released with a conventional radiosonde to collect the ozonesonde data. The sondes are launched two to four times per month between 8:00 and 14:00 local time (Thompson et al., 2017). Ozone profiles, temperature, pressure, and relative humidity records are available in the SHADOZ archives at http://tropo.gsfc.nasa.gov/. Data profiles from 1998 to 2019 from the two Ozonesonde stations was used to compute the T_m values. Table 3.1 presents the details of the ozonesonde data used in this study.

3.1.3 ERA5 Reanalysis Data

The ERA5 is the latest fifth-generation reanalysis dataset version of European Centre for Medium-Range Weather Forecasts (ECMWF), featuring a horizontal resolution of 31 km grid spacing and several innovative capabilities that enhance the quality of its estimates compared to other NWP models (Hersbach et al., 2020). The study conducted by Ssenyunzi et al. (2020) found that the meteorological values obtained from ERA5 could be used as an alternative to sensor meteorological datasets in estimating T_m and other meteorological parameters. The ERA5 data set is produced by using the 4D-Var data assimilation scheme in the CY41R2 model of the ECMWF's Integrated Forecast System (IFS).This study used eight years ERA5 data (2013 to 2020) of specific humidity and surface temperature with vertical dimensions at 137 model levels and a horizontal resolution of $0.25° \times 0.25°$ and with temporal resolution of 6 h (00, 06, 12, 18 UTC).

3.1.3.1 Obtaining ERA5 Parameters

The set of meteorological parameters applied in this study consists of relative humidity, surface geopotential, temperature and pressure. All these parameters are obtained from the ECMWF-ERA5 dataset using a Web API. The reconstruction of the full level pressure and the geopotential height associated with each model level uses the log of the surface pressure and the surface geopotential data. The ERA5 data is Archived at 137 model levels arranged from top to bottom and these model levels are defined by pressures at interfaces between them. Pressures was defined as:

$$p_{(k+\frac{1}{2})} = a_{(k+\frac{1}{2})} + b_{(k+\frac{1}{2})} p_s, \tag{3.1}$$

where k is the model level number with possible values 0-60. The $a_{(k+\frac{1}{2})}$ and $b_{(k+\frac{1}{2})}$ are constants whose values commendably define the vertical coordinate and p_s is the surface pressure data. The values of $a_{(k+\frac{1}{2})}$ and $b_{(k+\frac{1}{2})}$ for all model levels in the ECMWF files of all fields archived on model levels to allow the reconstruction of full-level p_k associated with each model level given by:

$$p_k = \frac{1}{2}(p_{(k-\frac{1}{2})} + p_{(k+\frac{1}{2})}). \tag{3.2}$$

The height used in ERA5 is the geopotential height (Z) and is zero at sea level. The geopotential heights of the surface and levels were obtained by dividing the geopotential by constant gravitational acceleration, G_o = 9.80665 ms^{-2}. The computation of $\mathrm{T_m}$ using integration method involves integration of vertical profiles of partial pressures of water vapor and temperatures through the entire atmospheric column. Therefore, geopotential height was converted to the geometric height. The geopotential height, Z (unit: m) is first converted to the geoid / orthometric height, H (unit: m) using the following equations (Yang et al., 2019);

$$H = \frac{R_e(\varphi) \cdot Z}{\frac{G(\varphi)}{G_o} \cdot R_e - Z} \tag{3.3}$$

where

$$R_e(\varphi) = \frac{6378.137}{1.006803 - 0.0069435 \sin^2(\varphi)} \tag{3.4}$$

$$G(\varphi) = \frac{9.780325[1 + 1.93185 \times 10^{-3} \sin^2(\varphi)]}{1 - 6.69435 \times 10^{-3} \sin^2(\varphi)} \tag{3.5}$$

where G_φ and G_o are the normal gravity (unit: ms^{-2}) of latitude φ and 45^o, respectively; R_e is the effective earth radius (unit: m) at latitude, φ.

After calculating the orthometric height, the geodetic/ geometric height h at the model levels was computed from equation (3.6).

$$h = H + \xi, \tag{3.6}$$

where the height anomaly (ξ) was computed from the official earth Gravitational model 1996 (EGM96) and H is ellipsoid/geoid height.

The specific humidity and the pressure data at the 137 model levels recorded in ERA5 were used to calculate the partial pressure of water vapor. Specific humidity refers to the amount of water vapor in a unit of humid air. The partial pressure of water vapor e_i in hpa was calculated from (Mateus, Catalão, Mendes, & Nico, 2020):

$$e_i = \frac{p_i q_i}{(q_i - (q_i - 1))\,\varepsilon} \tag{3.7}$$

where q_i is the specific humidity (unit: kg/kg), p_i is the air pressure in hPa, and ε is calculated from equation (3.8) and is equal to 0.622:

$$\varepsilon = \frac{R_d}{R_v}, \tag{3.8}$$

where R_d is the molecular weight of water; and R_v is the apparent molecular weight of dry air. Both q_i and p_i can be taken from each grid point at the model level. T_s and T_i were recorded as the surface temperature data and the temperature data at each model level, respectively. At this point, in each model level, the values of temperature, water vapor pressure, and altitude have been obtained from respective grids.

Spatial adjustments of the reanalysis data to the stations are also necessary because the grid points of the ERA5 data sets are not strictly collocated with the stations. Interpolating the grid point meteorological parameters to the station of interest entails making both vertical and horizontal spatial modifications. First, the meteorological data from the four grid points surrounding the station's location are subjected to bilinear interpolation for the horizontal adjustment basing on longitudes and latitudes. Details of Bilinear interpolation have been described in Section 3.1.3.2. Lastly, for those stations that are higher than the lowest level of ECMWF-ERA5, linear interpolation is done to bring the meteorological data to the same height as the station. If the station altitude is lower than the height of the lowest model level of ECMWF-ERA5, the parameters are linearly extrapolated.

3.1.3.2 Bilinear Interpolation of ERA5 Grid Parameters

For spatial alignment, all the comparisons were made at the locations of interest (or GNSS) stations. Four grid cells adjacent to a station were used to perform the bilinear interpolation. Figure 3.2 presents a schematic diagram of bilinear interpolation, where point M(x_2, y_2) denotes the meteorological station or point of interest while different background colors represent different grids cells. A(x_1, y_3), B(x_3, y_3), C(x_1, y_1), and D(x_3, y_1) are the grid points of the four adjacent grid cells. The different colors of the grid represent different values of ERA5 grid meteorological parameters, namely, P_A, P_B, P_C and P_D. The station value of the meteorological parameter, P_M, can be calculated using the bilinear

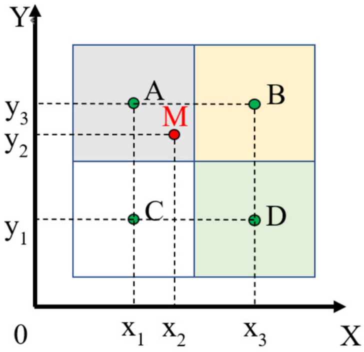


Figure 3.2: Schematic diagram of bilinear interpolation.

interpolation as follows:

$$P_M = P_A \frac{x_3 - x_2}{x_3 - x_1}\frac{y_2 - y_1}{y_3 - y_1} + P_B \frac{x_2 - x_1}{x_3 - x_1}\frac{y_2 - y_1}{y_3 - y_1} + P_C \frac{x_3 - x_2}{x_3 - x_1}\frac{y_3 - y_2}{y_3 - y_1} + P_D \frac{x_2 - x_1}{x_3 - x_1}\frac{y_3 - y_2}{y_3 - y_1}. \tag{3.9}$$

In this interpolation, the elevation of the points was not taken into account.

3.1.4 NGL/GNSS data

This study utilized Zenith Total Delay (ZTD) data from the Nevada Geodetic Laboratory (NGL) troposphere products (Blewitt, Hammond, et al., 2018) to analyze GNSS data. Table 3.2 presents the details of the GNSS stations whose data was used in this study. Table 3.2 presents the details of the GNSS data used in this study. The NGL, based at the University of Nevada, Reno, processes ZTD data using GipsyX version 1.0 software with JPL's Repro 3.0 orbits and clocks, VMF1 gridded data, and mapping function parameters as inputs (Ding & Chen, 2020). NGL provides highly accurate 5-minute ZTD time series, with an estimated accuracy of 4 mm, which are available at: http://geodesy.unr.edu from 1996 onwards. NGL continually updates its data and, as of November 2019, it has improved its modeling and data processing to include integrated water vapor and weighted mean tropospheric temperature. A Matlab script was used to extract the 5-minute resolution data,

Table 3.2: Details of GNSS stations used in this study.

Station Id & City	**Country**	**Available Data**	**Lon(°) E**	**Lat(°) N**	**Alt (m)**
ARSH (Arusha)	Tanzania	2013-2020	36.698	-3.387	1345.161
DODM (Dodoma)	Tanzania	2013-2020	35.478	-6.187	1122.595
RCMN (Nairobi)	Kenya	2013-2020	36.894	-1.221	1591.967

and then hourly and daily averages were calculated for each epoch, using equation 3.10.

$$\overline{ZTD}_{NGL} = \frac{1}{N}\sum_{i=1}^{N} ZTD_i, \tag{3.10}$$

where N is the total number of epochs and ZTD_i is the ZTD at each epoch.

3.2 Estimating the weighted mean temperature

The numerical integration method is considered to be the most accurate for computation of $\mathrm{T_m}$ (Davis et al., 1985). The $\mathrm{T_m}$ values were calculated by using (equation 3.11) (Davis et al., 1985):

$$\mathrm{T_m} = \frac{\int_h^{\infty} \frac{e}{T}\, dh}{\int_h^{\infty} \frac{e^2}{T^2}\, dh} \tag{3.11}$$

$$e = \frac{\mathrm{RH}}{e_s} \times 100\% \tag{3.12}$$

$$e_s = 6.112\mathrm{exp}\left(\frac{17.62\mathrm{T}}{\mathrm{T} + 243.5}\right), \tag{3.13}$$

where RH, is the Relative humidity, e is the partial pressure of water vapor, e_s is saturation vapour pressure, T is the temperature and dh represents the height difference between the layers. These parameters are obtained from ozonesonde, Radiosonde and ERA5 datasets.

In practice, a series of discrete sampling data of the profiles is collected and then equation 3.11 is discretized into;

$$\mathrm{T_m}, = \frac{\sum_{i=1}^{n} \frac{1}{2}\left(\frac{e_i}{T_i} + \frac{e_{i+1}}{T_{i+1}}\right)(h_{i+1} - h_i)}{\sum_{i=1}^{n} \frac{1}{2}\left(\frac{e_i}{T_i^2} + \frac{e_{i+1}}{T_{i+1}^2}\right)(h_{i+1} - h_i)} \tag{3.14}$$

where e_i, T_i, h_i and e_{i+1}, T_{i+1} and h_{i+1} are the water vapor pressure, absolute temperature and height of two adjacent levels of the atmospheric profile, respectively.

The ERA5-$\mathrm{T_m}$ values were computed six hourly; at 00, 06, 12, 18 and 00 of the next day. These were then interpolated using a piece-wise cubic Hermite interpolating polynomial to generate one hourly, $\mathrm{T_m}$ values for 24 h per day as explained in section 3.2.1.

3.2.1 Piece-wise Cubic Hermite Interpolating

For any pair of epochs $[t_i, t_{(i+1)}]$, the cubic Hermite polynomial (CHP) interpolant, $\hat{\mathrm{T}}_{\mathrm{m}}\ (t_k)$ between the given points can be estimated using equation (3.15) (Rabbath & Corriveau, 2019);

$$\hat{\mathrm{T}}_{\mathrm{m}}\,(t_k) = \mathrm{T}_{\mathrm{m}}(t_i) + f_1(t_k - t_i) + f_2(t_k - t_i)^2 + f_3(t_k - t_i)^2(t_k - t_{i+1}), \tag{3.15}$$

where

$$f_1 = \frac{\mathrm{T}_{\mathrm{m}}(t_i) - \mathrm{T}_{\mathrm{m}}(t_{i+1})}{t_i - t_{i+1}} \tag{3.16}$$

$$f_2 = \frac{f_4 - f_1}{t_{i+1} - t_i} \tag{3.17}$$

$$f_3 = \frac{f_6 - f_2}{t_{i+1} - t_i} \tag{3.18}$$

$$f_4 = \frac{\mathrm{T}_{\mathrm{m}}(t_{i+1}) - \mathrm{T}_{\mathrm{m}}(t_i)}{t_{i+1} - t_i} \tag{3.19}$$

$$f_5 = \frac{\mathrm{T}_{\mathrm{m}}(t_{i+2}) - \mathrm{T}_{\mathrm{m}}(t_{i+1})}{t_{i+2} - t_{i+1}} \tag{3.20}$$

$$f_6 = \frac{f_5 - f_4}{t_{i+1} - t_i} \tag{3.21}$$

The hourly T_{m} values were used to calculate the daily (24 h) average values using equation (3.22);

$$\bar{\mathrm{T}}_{\mathrm{m}} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathrm{T}}_{\mathrm{m}}, \tag{3.22}$$

where N is the total number of observations, $\bar{\mathrm{T}}_{\mathrm{m}}$ is the daily (24 h) average values and $\hat{\mathrm{T}}_{\mathrm{m}}$ is the hourly T_{m}.

3.3 Establishing the T_{m} Models

To establish the T_{m} models, a regression analysis method is used, which relies on the linear correlation between T_{m} and T_{s} as shown in equation 3.23. The T_{m} values were accurately determined using equation 3.11 as discussed in subsection 3.2.1. This method is frequently used in studies because of the linear relationship between T_{m} and T_{s} (Y. Yao, Shan, & Zhao, 2017).

$$\mathrm{T}_{\mathrm{m}} = a + b \cdot \mathrm{T}_{\mathrm{s}} \tag{3.23}$$

where the coefficients of the regression equation, a and b, are estimated using a least squares adjustment.

$$b = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}\left(\mathrm{T}_{\mathrm{m_i}} - \overline{\mathrm{T}}_{\mathrm{m}}\right)\left(\mathrm{T}_{\mathrm{s_i}} - \overline{\mathrm{T}}_{\mathrm{s}}\right)}{\sum_{\mathrm{i=1}}^{\mathrm{n}}\left(\mathrm{T}_{\mathrm{s_i}} - \overline{\mathrm{T}}_{\mathrm{s}}\right)} \tag{3.24}$$

$$a = \frac{\sum_{i=1}^{n} Tm_i - b\sum_{i=1}^{n} Ts_i}{n} = \overline{T}_m - b\overline{T}_s \tag{3.25}$$

where $\overline{T}m = \frac{1}{n}\sum_{i=1}^{n} Tm_i$, and $\overline{T}s = \frac{1}{n}\sum_{i=1}^{n} Ts_i$, and N is number of observations.

3.4 Performance Evaluation of the Developed T_m Model

To assess the effectiveness of the developed T_m model, radiosonde T_m data from two stations that were not used to establish the models as references is employed. The performance of the developed models was compared with the widely used Bevis model (Bevis et al., 1994) and the (Y. Yao et al., 2014) and (Suresh Raju et al., 2007) models. The T_s data from radiosonde was used as input for the T_m models. The performance of the models was assessed by computing several metrics, including the root mean square error (RMSE), mean bias (MnB), and correlation coefficient R. These metrics allowed measurement of the models' accuracy, efficiency, and linear covariation as shown in the following equations:

$$\text{MnB} = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i) \tag{3.26}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2} \tag{3.27}$$

$$R = \frac{\sum_{i=1}^{N}\left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)}{\sqrt{\left(P_i - \overline{P}\right)^2\left(O_i - \overline{O}\right)^2}} \tag{3.28}$$

$$R^2 = \frac{\left[\sum_{i=1}^{N}\left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)\right]^2}{\left[\sum_{i=1}^{N}\left(P_i - \overline{P}\right)^2\right]\left[\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2\right]}, \tag{3.29}$$

where $\overline{P}$ and $\overline{O}$ are the respective means of the predicted and observed T_m, while P_i and O_i are their i^{th} values respectively.

3.5 Retrieval of PWV from GNSS observations

To retrieve PWV from GNSS data, ZTD has to be determined first. In this study, the ZTD was obtained from NGL as explained in section 3.1.4. The ZTD consists of the zenith hydrostatic delay (ZHD) and the zenith wet delay (ZWD). The ZHD is estimated using the

Saastamoinen model given by equation 3.30 (Saastamoinen, 1972):

$$\mathrm{ZHD} = \frac{0.0022768\mathrm{P_S}}{1 - 0.00266\cos 2\phi - 0.00000028\mathrm{h}}, \tag{3.30}$$

where h is the height of the surface above the geoid, ϕ is the latitude of the station and $\mathrm{P_S}$ is the corresponding air pressure at the station in hPa. The air pressure $\mathrm{P_S}$ at the GNSS site in this study is obtained from ERA5 data as explained earlier in section 3.1.3. After retrieving the ZTD, the ZWD is calculated by subtracting the Zenith Hydrostatic Delay (ZHD) from the ZTD using equation 3.31.

$$\mathrm{ZWD} = \mathrm{ZTD} - \mathrm{ZHD}. \tag{3.31}$$

Following (Bevis et al., 1994), PWV is retrieved from GNSS observations using equation 3.32;

$$PWV = \Pi \cdot ZWD \tag{3.32}$$

where Π is the water vapor conversion factor. Following Askne and Nordius (1987), the conversion factor was determined by:

$$\Pi = \frac{10^6}{\rho_v \mathrm{R}_v \left(\frac{\mathrm{k_3}}{\mathrm{T_m}} + \mathrm{k}_2'\right)}, \tag{3.33}$$

where ρ_v is the density of liquid water (1000 $\mathrm{kgm^{-3}}$), R_v is the specific gas constant of water vapour (461.5 $\mathrm{Jkg^{-1}K^{-1}}$), $\mathrm{k_3} = 3.739 \pm 0.004 \times 10^5 \mathrm{K^2hPa^{-1}}$ and $\mathrm{k}_2' = 22.1 \pm 2.2\mathrm{K\ hPa^{-1}}$ are refractivity constants. Figure 3.3 presents the procedure for retrieval of GNSS-PWV, where P and T are Pressure and Temperature respectively. The $\mathrm{T_m}$ values were obtained from the ERA5 data and linear empirical relationships of $\mathrm{T_m}$ and $\mathrm{T_S}$ developed in this study and from (Bevis et al., 1994), (Y. Yao et al., 2014) and (Suresh Raju et al., 2007) models.

3.6 The Impact of the Developed $\mathrm{T_m}$ Models on GNSS-PWV

The validity of the developed $\mathrm{T_m}$ models was assessed using GNSS-PWV calculations based on GNSS observations from three stations: ARSH, DODM, and RCMN. The GNSS-PWV values were computed using Equation 3.32. The $\mathrm{T_m}$ values were obtained from various models, including our developed models, (Bevis et al., 1994), (Y. Yao et al., 2014), (Suresh Raju et al., 2007) and ERA5 $\mathrm{T_m}$ models. For each model, $\mathrm{T_m}$ values were calculated using ERA5 $\mathrm{T_S}$ data. GNSS-PWV obtained using ERA5 - $\mathrm{T_m}$ was used

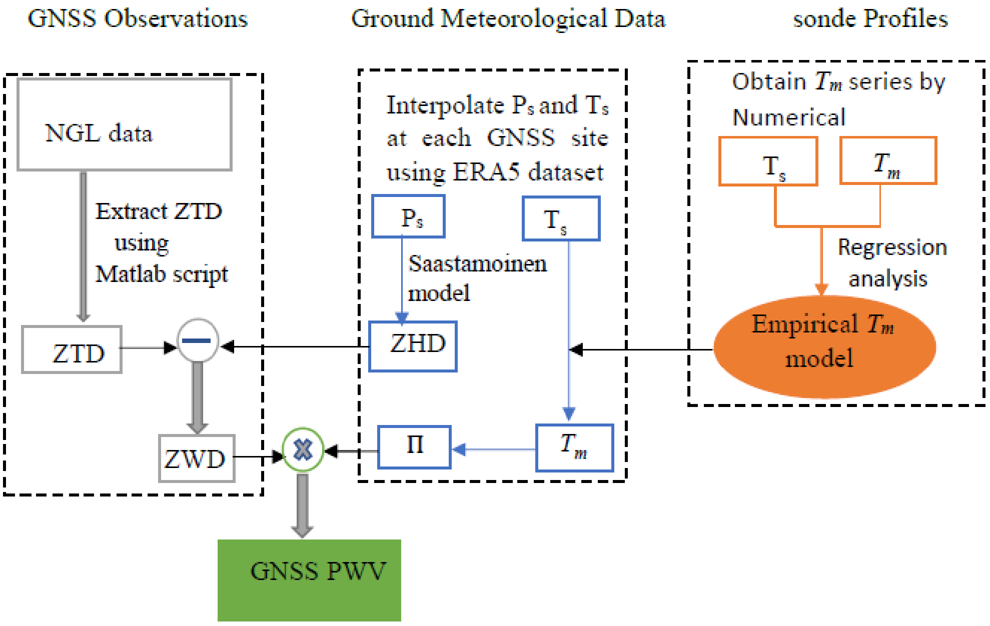


Figure 3.3: Procedure for GNSS PWV retrieval at a GNSS station (Huang, Mo, Liu, et al., 2021).

as reference values to assess the other PWV values. The RMSE of PWV values from different models at selected stations were calculated and later used for assessing the impact of developed $\mathrm{T_m}$ models on GNSS-PWV retrieval. To investigate the impact of $\mathrm{T_m}$ values on GNSS-PWV, the relationship between the root mean square (RMS) of $\mathrm{T_m}$ and PWV was determined using the equation 3.34 proposed by (Huang, Mo, Liu, et al., 2021).

$$\frac{\mathrm{RMSE_{PWV}}}{\mathrm{PWV}} = \frac{\mathrm{RMSE}_\pi}{\pi} = \frac{\mathrm{k_3} \cdot \mathrm{RMSE_{Tm}}}{\left(\frac{\mathrm{k_3}}{\mathrm{Tm}} + \mathrm{k}_2'\right) \mathrm{T_m^2}} = \frac{\mathrm{k_3}}{\left(\frac{\mathrm{k_3}}{\mathrm{Tm}} + \mathrm{k}_2'\right) \mathrm{Tm}} \cdot \frac{\mathrm{RMSE_{Tm}}}{\mathrm{Tm}} \tag{3.34}$$

In the context of $\mathrm{T_m}$ models and GNSS PWV results, the impact of $\mathrm{T_m}$ errors were assessed by calculating the relative error of PWV, defined as $\frac{\mathrm{RMSE_{pwv}}}{\mathrm{PWV}}$. Where $\mathrm{RMSE_{pwv}}$ is the root mean square of PWV, while $\mathrm{RMSE_{T_m}}$ is the root mean square of $\mathrm{T_m}$. Both of these values are obtained from the $\mathrm{T_m}$ models, with PWV from ERA5 $\mathrm{T_m}$. By comparing $\mathrm{RMSE_{pwv}}$ and $\frac{\mathrm{RMSE_{pwv}}}{\mathrm{PWV}}$, the impact of the errors in $\mathrm{T_m}$ on its resultant GNSS PWV results were obtained.

Chapter 4

Results and Discussion

4.1 Spatial Variation of T_m and T_s

This study investigated the spatial distribution of T_m data across the East African region, utilizing data from Radiosonde, ozonesonde stations, and ERA5 as shown in Figures 4.1(a) and 4.1(b). Figure 4.1(a) shows that the lowest average T_m values below 280 K are observed at stations Dagoretti (DAGO) and Nairobi (NAIR), respectively, which have the highest altitude of all available stations. Conversely, the Malindi station (MALI), which has the lowest altitude, exhibits the highest average T_m values above 289 K. The figures demonstrate that high T_m values are concentrated near the coast and in Northern Kenya. The high T_m values near the coast or in part of Northern Kenya may be attributed to low altitudes and high T_s values, among other factors. It is also observed in Figure 4.1(b) that ERA5 has captured the spatial distribution of T_m very well in comparison to Radiosonde T_m. Figure 4.2 presents the mean values of T_s from radiosonde. It is observed that areas with high T_m values also present high T_s values, with mean T_s values above 300 K. Stations at Dagoretti (DAGO) and Nairobi (NAIR) present the lowest mean T_s values below 290 K, respectively, which are also the stations with the lowest T_m values (see Figure 4.1). The altitude is found to have a significant impact on T_m and T_s values. This study aligns with the findings of the study by S. Wang et al. (2020) who also reported that, temperature decreases with the increase in height. So, generally, as altitude increases, temperatures decrease. This is called lapse rate, where the temperature decreases by about

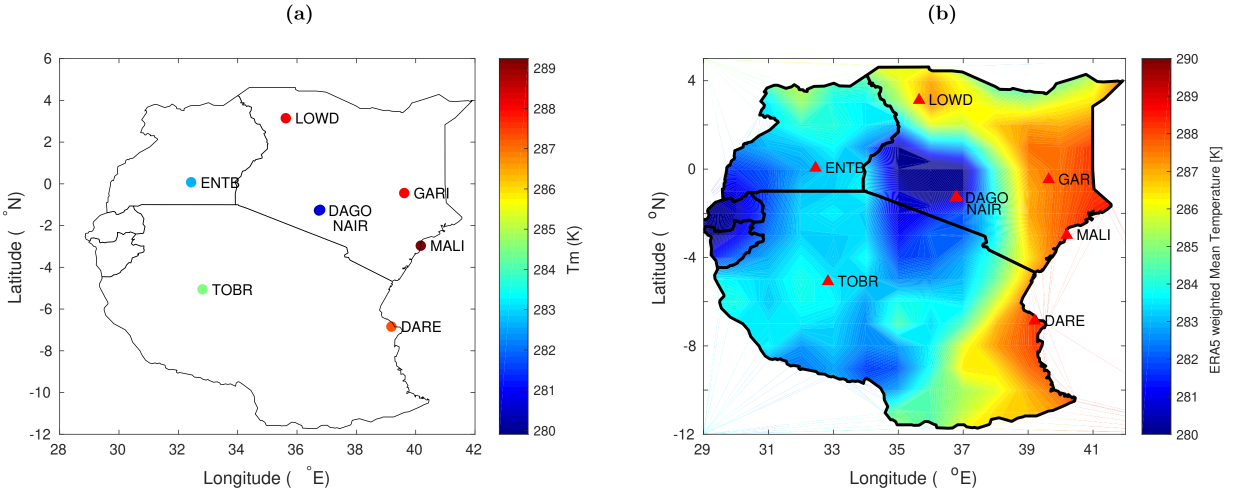


Figure 4.1: Spatial variation of T_m from (a) Radiosonde and (b) ERA5.

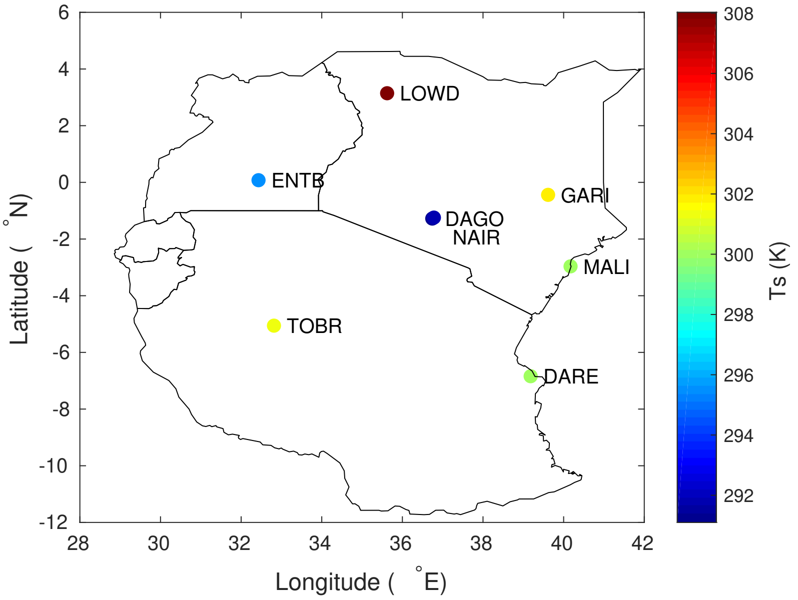
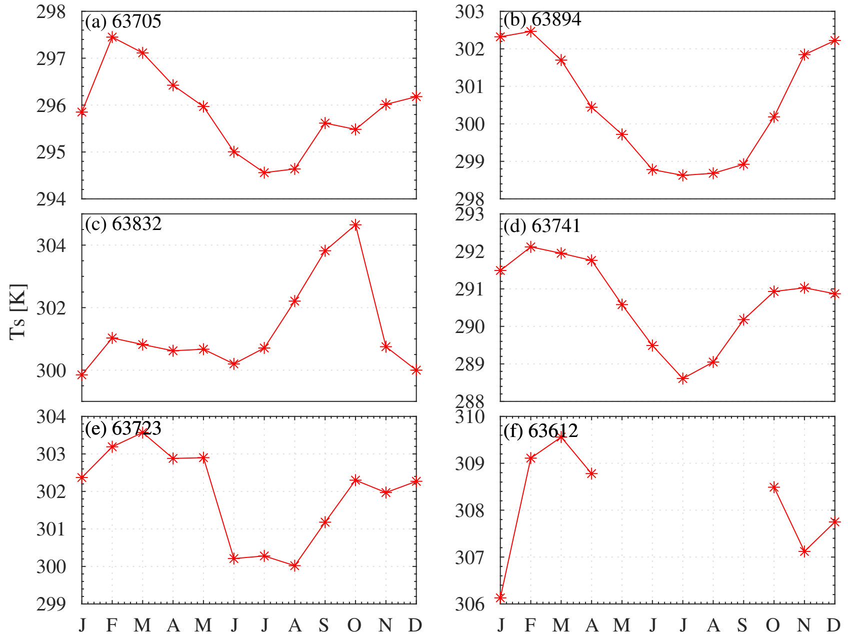


Figure 4.2: Spatial variation of T_s from Radiosonde.

6.5 degrees Celsius per kilometer increase in altitude (Huang, Mo, Liu, et al., 2021). This means that highland areas in East Africa, such as the Kenyan Highlands, tend to have cooler temperatures compared to lowland areas. However, T_m in the vertical direction varies nonlinearly due to the irregular variations in both water vapor pressure and temperature in this direction. Secondly, Urban areas in East Africa, tend to be warmer than rural areas due to the urban heat island effect, where urban areas retain more heat due to the concentration of buildings and human activities. Additionally, the proximity to large water bodies, such as Lake Victoria and the Indian Ocean, also influences temperature variation in East Africa.

4.2 Temporal Variation of T_m and T_s

Figures 4.3 and 4.4 illustrate the temporal variation of Radiosonde T_s and T_m respectively. The T_s in the East African region varies between 280 K and 311 K, while the range of T_m is smaller than T_s, ranging from 271 K to 295 K. From Figures 4.3 and 4.4, it is observed that T_s and T_m exhibit unique seasonal patterns across various radiosonde stations. In Figures 4.3(a) and 4.3(b), we observe an upward trend in average T_s values at stations 63705 (Entebbe) and 63894 (Dares-salaam), starting from a minimum value in January to a maximum value in February. Subsequently, there is a decline leading to the lowest values between June and August. The average T_s values then exhibit a gradual rise, reaching another maximum value in December. Notably, the maximum average T_s value in February exceeds that observed in December. From Figure 4.3(c), the average T_s values at station 63832 (Tabora) also show an increase, starting from a minimum value in January



.

Figure 4.3: Temporal variation of T_s from the East African Radiosonde stations.

and reaching a maximum value in February. Following this maximum value, there is a slight decline leading to a minimum value in June. Subsequently, the average T_s values gradually increase to a maximum value in October before decreasing to another minimum value in December. It is observed that the maximum average T_s value in February is lower than that observed in October. Figure 4.3(d) presents the average T_s values at station 63741 (Dagoretti). Here, we observe an increase from a minimum value in January to a maximum value in February, followed by a slight decline leading to a minimum value in July. Subsequently, the average T_s values gradually increase, reaching another maximum value between October and November, before declining once more to a minimum in December. Finally, in Figure 4.3(e), the average T_s values at station 63723 (GARISSA) also exhibit a rise from a minimum value in January to a maximum value in March. Afterward, a slight decrease, resulting in minimum values between June and August is observed. The average T_s values then gradually increase again, reaching another maximum value in October before declining.

Figures 4.4(a), 4.4(b), 4.4(d), and 4.4(e), clearly indicate an upward trajectory in average T_m values at stations 63705 (Entebbe), 63894 (Daressalaam), 63741 (Dagoretti), and 63723 (GARISSA) from January, reaching a peak in February. Subsequently, a decline in T_m values is noticeable, reaching the lowest values between June and August at these stations. Following this, there is again an increase in T_m values, reaching maximum values between September and December. However, station 63832 (Tabora) exhibits a distinct

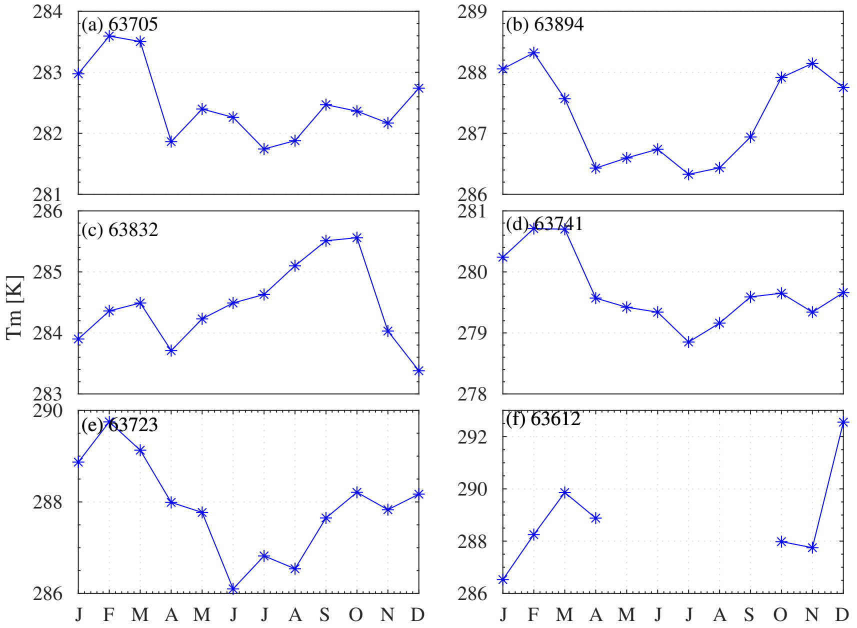


Figure 4.4: Temporal variation of $\mathrm{T_m}$ from the East African Radiosonde station.

pattern. The rise in $\mathrm{T_m}$ values is observed from January, culminating in a peak in March. A decline in $\mathrm{T_m}$ is then noted, reaching its lowest in April, followed by a subsequent increase to a maximum value in October.

It is worth noting, as illustrated in Figures 4.3 and 4.4, that the trends of $\mathrm{T_m}$ and $\mathrm{T_s}$ closely align over time. The variations in temperatures may be caused by the monsoon climate and terrain, and on the other hand, the water vapor content along coastal areas that exceeds one of other places. The variability of T_s is influenced by an influx of dry air driven by north eastern monsoon winds leading to dry seasons in December, January and February. The study sites therefore experience cloudless conditions during this time of the year, resulting into overheating of the earth and consequently hot season. On the contrary, the inflow of cold air from southern hemisphere attributes to the low temperatures in June, July and August (JJA). The dominant cloud cover during these months (JJA), block the short wave solar radiations from reaching the earth surface, resulting into low temperatures and consequently cold conditions.

4.3 The Established Site Specific T_m Models

In order to develop site-specific T_m models for East Africa, T_m and T_s values were calculated for each station using data from six radiosonde and two ozonesonde stations spanning from 1971 to 2019. The relationship between T_m and T_s was established using coefficients in Equation 3.23, determined by the linear least-squares fitting method, as illustrated in Figure 4.5. The resulting T_m models, their correlation coefficients of determination and the number of data points applied in the establishment of the models are presented in Table 4.1. Table 4.1 reveals that the constant, b, values range from 0.25 to

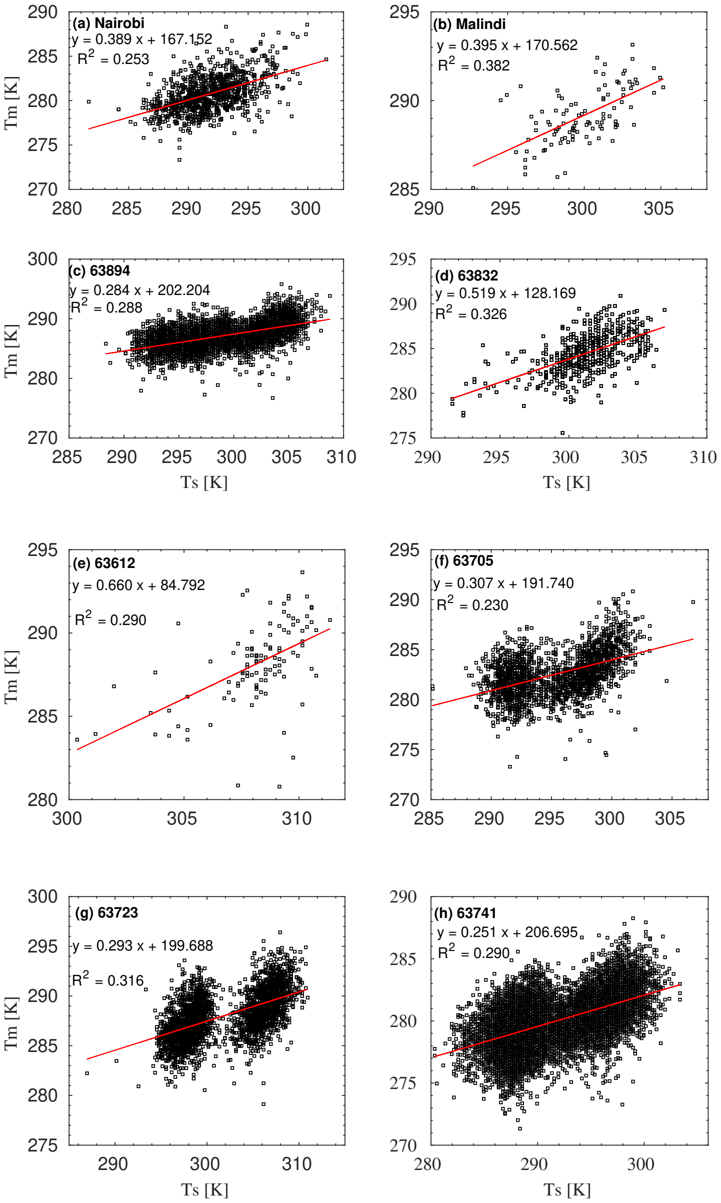


Figure 4.5: T_m-T_s scatter plots at different locations.

0.53, and constant, a, is between 84.8 and 206.7. In contrast, Wang et al. (2020) reported that constant, b, values in the middle to high latitudes over the Northern Hemisphere vary from 0.6 to 0.8, with constant, a, approximately 100.5. This suggests that the constant, b, values in the region of this study are smaller than those in the middle and high latitudes regions. However, Wang et al. (2020) also observed smaller values of the constant, b, at low latitudes, which is attributed to the low variations of T_s and T_m in this region. Additionally, Figure 4.5 and Table 4.1 show that the correlation coefficients of determination for the linear regression T_m models, range from 0.23 to 0.38. This suggests weak correlations between T_m -T_s in this region. (Ross & Rosenfeld, 1997) also reported weak correlations at tropical radiosonde stations, where the correlations were found to be less than 0.5. Similarly, Wang et al. (2020) reported a weak T_m -T_s relationship in the areas near the equator due to the entire variation ranges of T_s and T_m both being within 10 K.

Table 4.1: The East African site specific and regional T_m -T_s regression equations.

Station	**City Name**	**Regression equations**	**Data points modeled**	**R^2**
63741	Dagoretti	$T_m = 0.251T_s + 206.695$	11,628	0.290
63894	Daressalaam	$T_m = 0.284T_s + 202.204$	4,011	0.288
63705	Entebbe	$T_m = 0.307T_s + 191.740$	1,989	0.230
63723	Garissa	$T_m = 0.293T_s + 199.688$	2,684	0.316
63612	Lodwar	$T_m = 0.660Ts + 84.792$	108	0.290
-	Malindi	$Tm = 0.395T_s + 170.562$	102	0.382
-	Nairobi	$T_m = 0.389T_s + 167.152$	937	0.253
63832	Tabora	$T_m = 0.519T_s + 128.169$	522	0.326
All	**Regional**	$\mathbf{T_m} = 0.526\mathbf{T_s} + 127.691$	**21,981**	**0.603**

4.4 The Established East African Regional T_m Model

To establish the East African regional T_m model, we utilized a dataset comprising 21,981 data points of T_m and T_s from eight radiosonde and ozonesonde profiles spanning from 1971 to 2019 in East Africa. The coefficients governing the relationship between T_m and T_s for the regional model were determined using the linear least-squares fitting method as outlined in Equation 3.23. The linear fitting of these datasets is illustrated in Figure 4.6, and the resulting regional T_m model is summarized in Table 4.1. Analyzing Figure 4.6 and Table 4.1, it is evident that the regression model between T_m and T_s exhibit a correlation coefficient of determination of 0.603 which is better when compared with one for site specific models. This is because East Africa has diverse topography and climate conditions. So, the E.A regional model can capture this variability better than site-specific models, which might not account for larger-scale climate patterns affecting the region. Secondly, the E.A regional model can leverage a larger dataset, which may result in more robust and

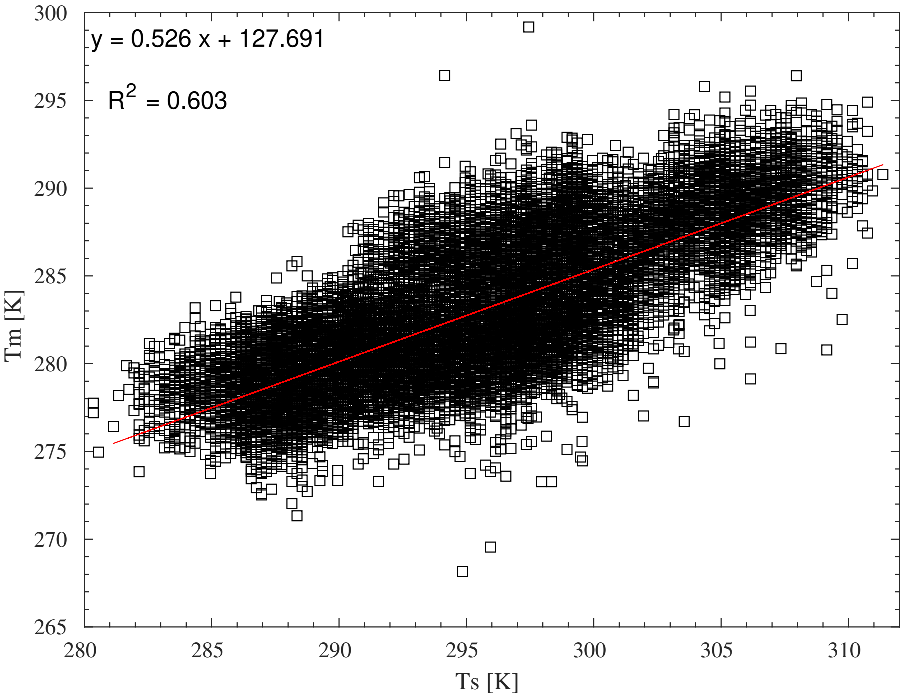


Figure 4.6: Tm-Ts scatter plot for East African region.

accurate estimates compared to site-specific models that might have limited data points or data quality issues.

4.5 Assessments of the Developed T_m Models

To assess the performance of the developed models, we used T_m data from two radiosonde stations: 63741 in Dagoretti (DAGO), Kenya, and 63894 in Dares-salaam (DARE), Tanzania. Data utilised is different from one used in model establishment. For station 63741, out of 15,361 profiles, 11,628 were employed in modeling and 3,733 in validation. For station 63894, 4,011 out of 4,999 profiles were used in modeling and 988 in validation. To assess the effectiveness of the site-specific and regional T_m models, two global T_m models (Bevis and Yao) and one tropical regional model (Raju T_m model) were used. The RMSE, MnB, and correlation coefficients were used as metrics for model evaluation. Figures 4.7 and 4.8 depict results derived from data at stations 63741 and 63894. Figure 4.7 presents RMSE values of 2.70 K, 1.76 K, 2.01 K, 2.47 K, 1.79 K, and 2.54 K for Yao, site-specific, regional, Raju, Nairobi, and Bevis, respectively. Corresponding MnB values were 0.81 K, 0.40 K, -0.36 K, -0.20 K, -0.08 K, and 0.87 K, while R values were consistent across all models at 0.446. Notably, Yao demonstrated the highest RMSE, followed by Bevis and Raju. Based

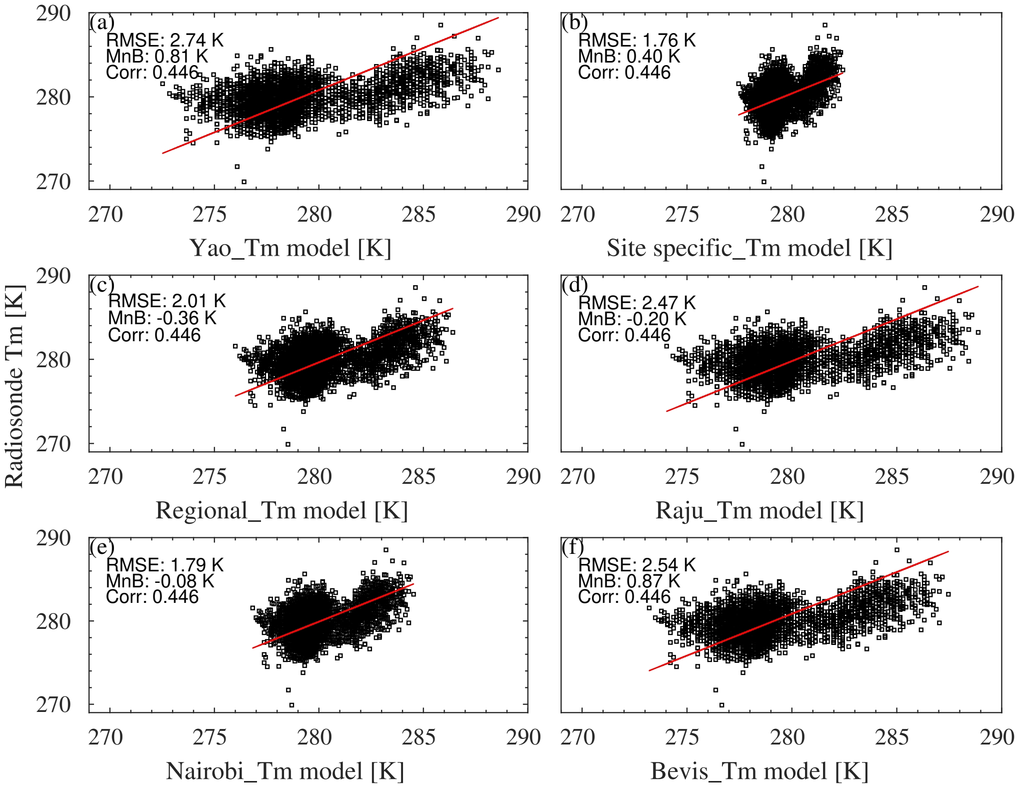


Figure 4.7: Scatter plots showing the statistics results of developed models against other models at Dagoretti (63741), Kenya

on RMSE values, the site-specific model exhibited superior performance, boasting the lowest RMSE of 1.76 K, followed by Nairobi at 1.79 K, and then the regional model at 2.01 K. Additionally, it is evident that the global models and the tropical regional model (Yao, Bevis, and Raju) delivered less favorable outcomes. In Figure 4.8, we extracted RMSE values of 2.52 K, 1.63 K, 1.66 K, 2.64 K, and 1.70 K for Yao, site-specific, regional, Raju, and Bevis models, respectively. The corresponding MnB values were -1.91 K, -0.43 K, 0.70 K, -2.10 K, and -0.64 K, with R values consistently hovering between 0.587 and 0.588 for all models. Notably, Yao and Raju exhibited the highest RMSE and MnB values. In terms of RMSE, the site-specific model demonstrated superior performance with the lowest value of 1.63 K, followed closely by the regional model at 1.66 K, and then the Bevis model at 1.70 K. It's worth noting that at this station, the differences in RMSE values between the site-specific, E.A regional models are relatively small. This suggests that established T_m models have superiority over global models and therefore will provide more accurate and relevant temperature predictions for East Africa, taking into account local climate characteristics and trends.

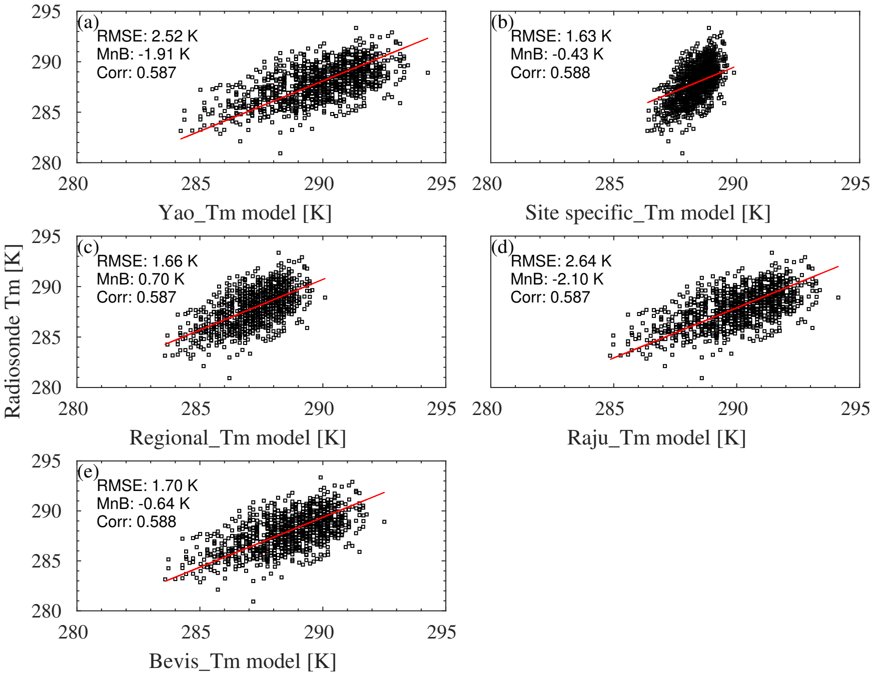


Figure 4.8: Scatter plots showing the statistics results of developed models against other models at Daressalaam (63894), Tanzania.

4.6 Seasonal Variation of MnB and RMSE

This study analyses the seasonal variations in T_m models including Bevis et al. (1994), Y. Yao et al. (2014), Suresh Raju et al. (2007), Nairobi, as well as site-specific and East African regional models. The seasonal fluctuations were assessed using Radiosonde T_m and T_s values collected from station 63741 Dagoretti (DAGO) spanning from 1995 to 2006. The data used here was not utilized in the modeling process for both the site-specific and E.A regional models. Figure 4.9 illustrates the seasonal trends in MnB and RMSE for the various T_m models under study. Figure 4.9(a) reveals that during the December - January - February (DJF) season, Yao, site-specific, Nairobi, and Bevis T_m models exhibit positive MnB values of 0.89 mm, 0.91 mm, 0.33 mm and 1.02 mm respectively. Conversely, the E.A regional and Raju T_m models display negative MnB values of -0.05 mm and -0.07 mm respectively. In the June - July - August (JJA) season ,Yao, site-specific, Raju, and Bevis T_m models show positive values of 1.38 mm, 0.18 mm, 0.29 mm and 1.31 mm respectively, whereas the regional and Nairobi T_m models reveal negative MnB values of -0.19 mm, -0.11 mm respectively. During this season, the site-specific, regional, and Nairobi T_m models exhibit the lowest MnB values, while Yao and Bevis present the highest values.

Figure 4.9(a) further depicts that in the March - April - May (MAM) season, Yao, site-specific, and Bevis T_m models record positive MnB values of 0.09 mm, 0.30 mm

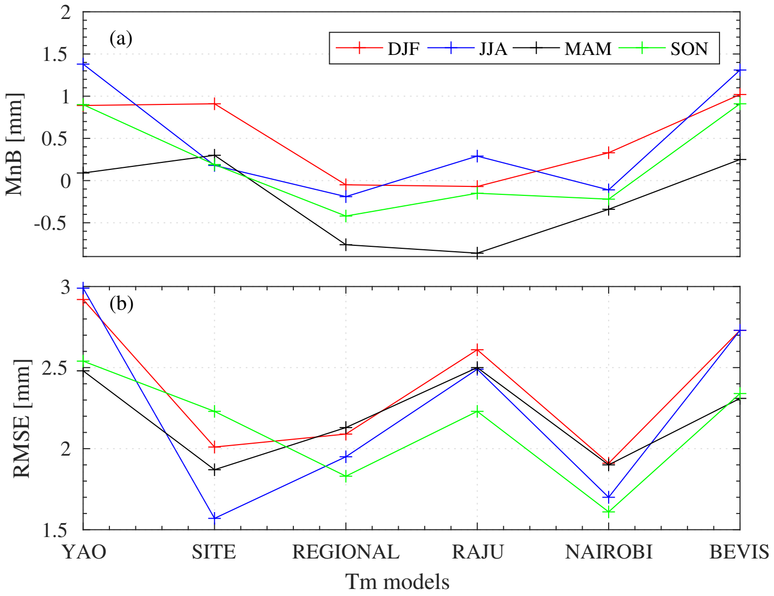


Figure 4.9: Seasonal variation of MnB and RMSE from radiosonde measurements at Dagoretti (63741), Kenya.

and 0.25 mm respectively . In contrast, T_m models including the E.A regional, Raju, and Nairobi T_m models display negative MnB values of -0.76 mm, -0.86 mm,-0.34 mm respectively. Within this season, Yao T_m model presents the lowest MnB value, whereas Raju T_m model exhibits the highest values. During the September - October - November (SON) season, Yao, site-specific, and Bevis T_m models demonstrate positive MnB values of 0.9 mm, 0.19 mm and 0.91 mm respectively. Conversely, the regional, Raju, and Nairobi T_m models exhibit negative values of -0.42 mm, -0.15 mm, -0.22 mm respectively. The Bevis and Yao T_m models present the highest MnB values, while Raju displays the lowest MnB values, followed by site-specific and Nairobi. Overall, Figure 4.9(a) shows that Yao, site-specific, and Bevis T_m models consistently yield positive MnB values across all seasons, indicating an underestimation of Radiosonde T_m values. Conversely, the regional T_m model consistently yields negative MnB values in all seasons, indicating an overestimation of Radiosonde T_m values.

Figure 4.9(b) illustrates the Root Mean Square Error (RMSE) values for the six T_m models throughout the four seasons. During the DJF season, Yao exhibits the highest RMSE value at 2.99 mm, followed by Bevis and Raju with RMSE values of 2.73 mm and 2.61 mm, respectively. The site-specific exhibits RMSE value of 2.01 mm while and E.A regional model exhibits RMSE value of 2.09 mm. These two T_m models show a very slight difference of only 0.08 mm. In contrast, the Nairobi T_m model displays the lowest RMSE value at 1.91 mm for this season. Based on the RMSE values, it is evident that the Nairobi

T_m model performs the best, followed by the site-specific and regional models during the DJF season. According to the JJA season, the Yao T_m model displays the highest RMSE value of 2.99 mm, followed by Bevis and Raju with RMSE values of 2.73 mm and 2.49 mm, respectively. The site-specific and regional T_m models present RMSE values of 1.57 mm and 1.95 mm, respectively. The Nairobi T_m model again displays the lowest RMSE value at 1.7 mm. Therefore, considering the RMSE values, site-specific and E.A regional models demonstrate better performance, respectively.

In the MAM season, the Raju T_m model shows the highest RMSE value at 2.50 mm, followed closely by Yao and Bevis models with RMSE values of 2.48 mm and 2.31 mm, respectively. The regional T_m model displays the RMSE value of 2.13 mm. The site-specific and regional T_m models exhibit the lowest RMSE values of 1.87 mm and 1.90 mm respectively, with a negligible difference of only 0.03 mm. As for the September-October - November (SON) season, the Yao T_m model demonstrates the highest RMSE value at 2.54 mm, followed by Bevis with RMSE value of 2.34 mm. Both Raju and the site-specific models display the same RMSE values of 2.23 mm. The Nairobi T_m model registers the lowest RMSE value at 1.61 mm, followed by the regional model with RMSE value of 1.83 mm. Figure 4.9(b) generally indicates that, across all four seasons with the exception of MAM, the Yao T_m model consistently exhibits the highest RMSE values.

Additionally, it is noteworthy that the established models outperform Y. Yao et al. (2014), Bevis et al. (1994), and Suresh Raju et al. (2007) models for all the four seasons. This is because the developed models were established using higher resolution data which allows for better representation of small scale climate features. So incorporate local trends and patterns that may not be captured by global models.

4.7 The Impact of the Developed T_m Models on GNSS-PWV

The accuracy of T_m plays a crucial role in determining the precision of PWV. In order to assess the impact of T_m on GNSS-PWV, we derived five different sets of PWV data, using T_m values from various sources, for three different GNSS stations: ARSH, DODM, and RCMN. These sets are labeled as ERA5 (T_m) PWV , Bevis (T_m) PWV , Yao (T_m) PWV, regional (T_m) PWV , and site-specific (T_m) PWV , corresponding to different T_m sources namely; ERA5 T_m, Bevis T_m, Yao T_m, regional T_m, and site-specific T_m, respectively. ERA5 (T_m) served as the reference to evaluate the errors in the other methods. The use of ERA5-Tm was due to the limited or lack of Tm values from Radiosonde during the same periods of available GNSS data required for PWV estimation. The absolute errors in

GNSS-PWV were then computed for the Bevis T_m, Yao T_m, regional T_m, and site-specific T_m at all the three stations. The mean biases (MnB) and Root Mean Square Error (RMSE) values of the GPS-PWV from the three GNSS stations are presented in Figures 4.10, 4.11 and 4.12. Table 4.2 displays the mean values of RMSpwv and RMSpwv/PWV for the PWV (Bevis T_m), PWV (Yao T_m), PWV (regional T_m), and PWV (site-specific T_m).

Table 4.2: Statistical results of theoretical RMS and relative errors in PWV resulting from the five models validated by using ERA5 profiles from 2013- 2020

	RMS_{PWV} (mm)			RMS_{PWV}/PWV (%)		
Models	Max	Min	Mean	Max	Min	Mean
Bevis	0.15	0.109	0.130	2.21	0.30	1.26
Yao	0.15	0.107	0.129	2.57	0.14	1.36
Raju	0.17	0.126	0.148	2.71	0.50	1.61
E.A-Model	0.14	0.113	0.127	1.93	0.30	1.12
Site-Model	0.11	0.107	0.109	0.75	0.07	0.41

From these figures, it is observed that the MnB values in the estimated GPS-PWVs derived from Bevis, Raju, regional, and Yao models range between 0.087 mm and -0.056 mm, indicating relatively small differences. The figures also display RMSE values at the stations, ranging between 0.170 mm and 0.107mm. In Figure 4.10 (ARSH station), the GPS-PWV from the regional T_m model exhibited the smallest RMSE value of 0.113 mm, followed by Raju (0.133 mm), Yao (0.141 mm), and then Bevis (0.144 mm). Based on the

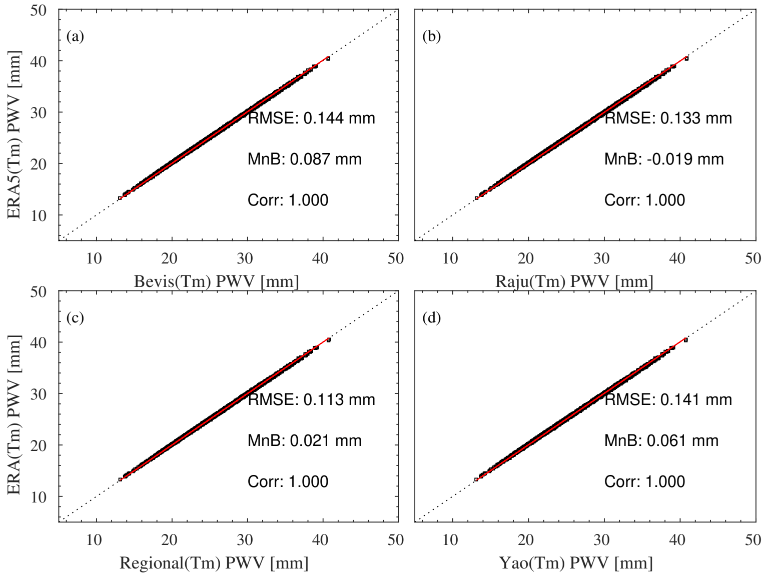


Figure 4.10: Scatter plot for retrieved PWV at ARSH-GNSS station

RMSE values, the regional T_m model out performs the other models for deriving PWV values at this station. In Figure 4.11 (DODM station), the GPS-PWV from the regional T_m model also exhibited the smallest RMSE value of 0.142 mm, followed by Bevis and Yao (0.150 mm), and then Raju (0.170 mm). In Figure 4.12 (RCMN station), the GPS-PWV

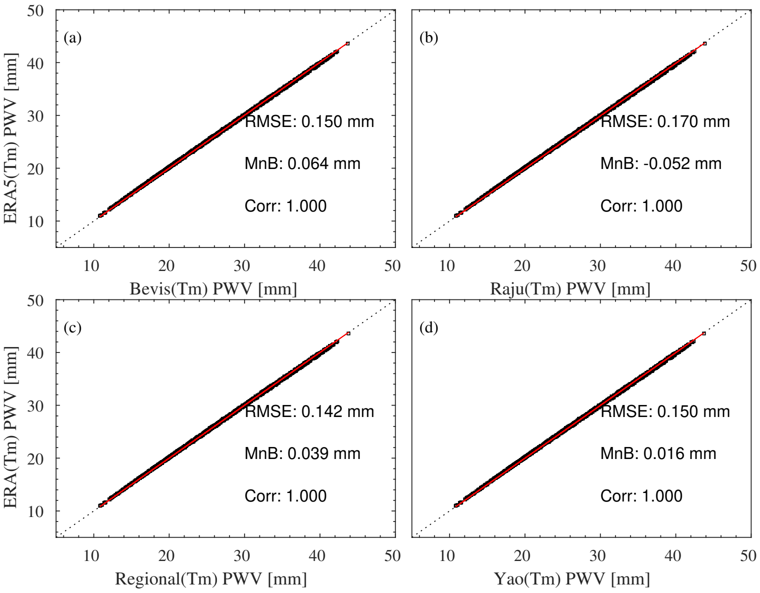


Figure 4.11: Scatter plot for retrieved PWV at DODM-GNSS station

from the site-specific and Yao T_m models exhibited the smallest RMSE value of 0.107 mm, followed by Bevis (0.109 mm), regional (0.116 mm), and then Raju (0.126 mm). Based on the RMSE values, the site-specific and Yao T_m models can be considered accurate for deriving PWV values at this station. Again, based on the RMSE values, the regional T_m model is suitable for accurately deriving PWV values at this station. Overall, based on the RMSE values, it is evident that both the site-specific and regional T_m models can be used effectively to derive PWV values in the study region.

In terms of RMSpwv, the site-specific models exhibit values ranging from 0.107 mm to 0.11 mm, with a mean value of 0.109 mm, which is the lowest among the five models. This is followed by, the E.A regional model with mean value of 0.127 mm, Yao T_m with a mean value of 0.129 mm, Bevis with a mean value of 0.130 mm, and Raju with a mean value of 0.148 mm. Regarding RMSpwv/PWV, the site-specific models report values between 0.07% and 0.75%, with a mean value of 0.41%. The results suggest that the impact of T_m values derived from site-specific models on GNSS PWV is smaller compared to the other four models. The regional model records values ranging from 0.30% to 1.93%, with a mean of 1.12%. Comparing these results to the other three models, the models based on local data from the East African region demonstrate smaller errors and broader applicability.

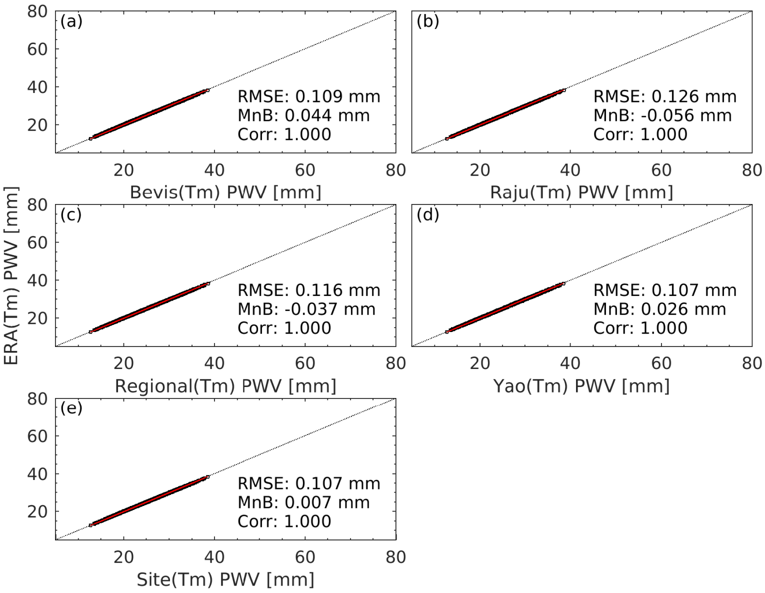


Figure 4.12: Scatter plot for retrieved PWV at RCMN-GNSS station

Therefore, the site-specific and regional models are the preferred choices for estimating T_m in GNSS PWV retrieval in the East African region.

Chapter 5

Conclusions and Recommendations

5.1 Conclusions

This study provides a thorough examination of weighted mean temperature (T_m) variations in the East African region, introducing new T_m models derived from Ozonesonde and Radiosonde profiles from 1971 to 2019. The spatial analysis reveals distinct patterns, with higher altitudes exhibiting lower T_m values (e.g., Dagoretti and Nairobi), and lower altitudes showing higher values (e.g., Malindi). Coastal and northern areas consistently exhibit high T_m values. Additionally, altitude emerges as a significant factor influencing both T_m and surface temperature (T_s), with lower altitudes corresponding to higher T_m and T_s values.

The study further examines the seasonal T_s and T_m variations, exhibiting unique patterns across different stations. The analysis of MnB and RMSE highlights the performance of various T_m models. Notably, the site-specific and regional models consistently outperform globally used models such as Bevis, Yao, and Raju. This suggests that the newly developed T_m models provide more accurate estimations for the East African region.

Moreover, assessing the impact of T_m errors on GNSS PWV, the study highlights the superiority of site-specific and regional models over global models. These models demonstrate smaller errors and broader applicability, making them the preferred choices for GNSS-PWV retrieval in the dynamic atmospheric conditions of East Africa. Overall, the accuracy of T_m proves pivotal in determining the precision of PWV calculations, emphasizing the significance of region-specific models in atmospheric research.

5.2 Recommendations

The study established the T_m models for GNSS PWV retrieval in East Africa based on a single meteorological parameter (T_s). Future studies, should focus on establishing a more accurate multi-parameter input T_m model by considering other factors that affect precipitation, such as pressure, geography (longitude and latitude), elevation, time and temperature.

The findings indicate that the generated models outperform the global models in terms of GNSS PWV accuracy. As a result, it is advised to use established models to determine the T_m in the East African region.

References

Andrei, C.-O., & Chen, R. (2009). Assessment of time-series of troposphere zenith delays derived from the global data assimilation system numerical weather model. *GPS solutions*, *13*, 109–117.

Askne, J., & Nordius, H. (1987). Estimation of tropospheric delay for microwaves from surface weather data. *Radio science*, *22*(03), 379–386.

Bai, Z. (2005). Near-real-time gps sensing of atmospheric water vapour.

Bean, B. R., & Dutton, E. (1966). *Radio meteorology* (Vol. 92). Superintendentof Documents, US GovernmentPrint. Office.

Berman, A. L. (1976). *The prediction of zenith range refraction from surface measurements of meteorological parameters* (Tech. Rep.).

Bernet, L., Brockmann, E., von Clarmann, T., Kämpfer, N., Mahieu, E., Mätzler, C., ... Hocke, K. (2020). Trends of atmospheric water vapour in switzerland from ground-based radiometry, ftir and gnss data. *Atmospheric chemistry and physics*, *20*(19), 11223–11244.

Berrada Baby, H., Gole, P., & Lavergnat, J. (1988). A model for the tropospheric excess path length of radio waves from surface meteorological measurements. *Radio science*, *23*(6), 1023–1038.

Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C., & Ware, R. H. (1994). Gps meteorology: Mapping zenith wet delays onto precipitable water. *Journal of Applied Meteorology (1988-2005)*, 379–386.

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., & Ware, R. H. (1992). Gps meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *Journal of Geophysical Research: Atmospheres*, *97*(D14), 15787–15801.

Blewitt, G., Hammond, W., et al. (2018). Harnessing the gps data explosion for interdisciplinary science. *Eos*, *99*.

Böhm, J., Heinkelmann, R., & Schuh, H. (2007). Short note: a global model of pressure and temperature for geodetic applications. *Journal of Geodesy*, *81*, 679–683.

Böhm, J., Möller, G., Schindelegger, M., Pain, G., & Weber, R. (2015). Development of an improved empirical model for slant delays in the troposphere (gpt2w). *GPS solutions*, *19*, 433–441.

Böhm, J., Niell, A., Tregoning, P., & Schuh, H. (2006). Global mapping function (gmf): A new empirical mapping function based on numerical weather model data. *Geophysical research letters*, *33*(7).

Boutiouta, S., & Lahcene, A. (2009). Algerian weighted mean temperature equation (awmte) and gnss meteorology technique application. In *Proceeding of 1st international symposium on modeling and implementation of complex systems* (Vol. 4, pp. 208–

218).

Boutiouta, S., & Lahcene, A. (2013). Preliminary study of gnss meteorology techniques in algeria. *International Journal of Remote Sensing*, *34*(14), 5105–5118.

Businger, S., Chiswell, S. R., Bevis, M., Duan, J., Anthes, R. A., Rocken, C., . . . Solheim, F. S. (1996). The promise of gps in atmospheric monitoring. *Bulletin of the American Meteorological Society*, *77*(1), 5–18.

Camberlin, P. (2018). Climate of eastern africa.

Chahine, M. T. (1992). The hydrological cycle and its influence on climate. *Nature*, *359*(6394), 373–380.

Chen, B., Yu, W., Wang, W., Zhang, Z., & Dai, W. (2021). A global assessment of precipitable water vapor derived from gnss zenith tropospheric delays with era5, ncep fnl, and ncep gfs products. *Earth and Space Science*, *8*(8), e2021EA001796.

Chen, J., Wang, J., Wang, A., Ding, J., & Zhang, Y. (2020). Shatrope—a regional gridded ztd model for china and the surrounding areas. *Remote Sensing*, *12*(1), 165.

Davies, T., Vincent, C., & Beresford, A. (1985). July-august rainfall in west-central kenya. *Journal of Climatology*, *5*(1), 17–33.

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

De Haan, S. (2013). Assimilation of gnss ztd and radar radial velocity for the benefit of very-short-range regional weather forecasts, qj roy. meteor. soc., 139, 2097–2107.

Ding, J., & Chen, J. (2020). Assessment of empirical troposphere model gpt3 based on ngl's global troposphere products. *Sensors*, *20*(13), 3631.

Elgered, G., & Jarlemark, P. O. (1998). Ground-based microwave radiometry and long-term observations of atmospheric water vapor. *Radio Science*, *33*(3), 707–717.

Elhaty, N. M., Abdelfatah, M. A., Mousa, A. E., & El-Fiky, G. S. (2019). Gnss meteorology in egypt: Modeling weighted mean temperature from radiosonde data. *Alexandria Engineering Journal*, *58*(2), 443–450.

El-Mowafy, A., & Lo, J. (2014). Dynamic modeling of gnss troposphere wet delay for estimation of precipitable water vapour. *Journal of Applied Geodesy*, *8*(1), 31–42.

Emmanuel, M., Sunilkumar, S., Ratnam, M. V., Muhsin, M., Parameswaran, K., & Murthy, B. K. (2018). Diurnal variation of the tropospheric water vapour over a coastal and an inland station in southern indian peninsula. *Journal of Atmospheric and Solar-Terrestrial Physics*, *179*, 11–21.

Essen, L., & Froome, K. (1951). The refractive indices and dielectric constants of air and its principal constituents at 24,000 mc/s. *Proceedings of the Physical Society. Section B*, *64*(10), 862.

Feng, P., Li, F., Yan, J., Zhang, F., & Barriot, J.-P. (2020). Assessment of the accuracy of the saastamoinen model and vmf1/vmf3 mapping functions with respect to ray-tracing from radiosonde data in the framework of gnss meteorology. *Remote Sensing*, *12*(20), 3337.

Gold, R. (1967). Optimal binary sequences for spread spectrum multiplexing (corresp.). *IEEE Transactions on information theory*, *13*(4), 619–621.

He, Q., Shen, Z., Wan, M., & Li, L. (2020). Precipitable water vapor converted from gnss-ztd and era5 datasets for the monitoring of tropical cyclones. *IEEE Access*, *8*, 87275–87290.

Herring, T. (1992). Modeling atmospheric delays in the analysis of space geodetic data. *Proceedirws of Refraction of Transatmospheric simals in Geodesy, eds. JC De Munck and TA Spoelstra, Netherlands Geodetic Commission Publications on Geodesy*, *36*(4).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... others (2020). *The era5 global reanalysis, qj roy. meteor. soc., 146, 1999–2049.*

Hofmann-Wellenhof, B., Lichtenegger, H., & Collins, J. (2012). *Global positioning system: theory and practice*. Springer Science & Business Media.

Hopfield, H. (1969). Two-quartic tropospheric refractivity profile for correcting satellite data. *Journal of Geophysical research*, *74*(18), 4487–4499.

Hu, J., Bao, Y., Liu, J., Liu, H., Petropoulos, G. P., Katsafados, P., ... Cai, X. (2021). Temperature and relative humidity profile retrieval from fengyun-3d/hiras in the arctic region. *Remote Sensing*, *13*(10), 1884.

Huang, L., Mo, Z., Liu, L., Zeng, Z., Chen, J., Xiong, S., & He, H. (2021). Evaluation of hourly pwv products derived from era5 and merra-2 over the tibetan plateau using ground-based gnss observations by two enhanced models. *Earth and Space Science*, *8*(5), e2020EA001516.

Huang, L., Mo, Z., Xie, S., Liu, L., Chen, J., Kang, C., & Wang, S. (2021). Spatiotemporal characteristics of gnss-derived precipitable water vapor during heavy rainfall events in guilin, china. *Satellite Navigation*, *2*(1), 1–17.

Ifadis, I. (1986). The atmospheric delay to radio waves: modeling the elevation dependence on a global scale. *Technical Report*, *381*.

Jiang, P., Ye, S., Chen, D., Liu, Y., & Xia, P. (2016). Retrieving precipitable water vapor data using gps zenith delays and global reanalysis data in china. *Remote Sensing*, *8*(5), 389.

Jiang, P., Ye, S., Lu, Y., Liu, Y., Chen, D., & Wu, Y. (2019). Development of time-varying global gridded t s–t m model for precise gps–pwv retrieval. *Atmospheric Measurement Techniques*, *12*(2), 1233–1249.

Johansson, G. E. J. M. (1998). Three months of continuous monitoring of atmospheric water vapor with a network of gps. *joURNAL OF GEOPHYSICAL RESEARCH*, *103*(D2), 1807–1820.

Krueger, E., Schueler, T., Hein, G. W., Martellucci, A., & Blarzino, G. (2004). Galileo tropospheric correction approaches developed within gstb-v1. In *Proceedings of enc-gnss* (Vol. 2004, pp. 16–19).

Kumar, A., Kumar, S., Lal, P., Saikia, P., Srivastava, P. K., & Petropoulos, G. P. (2021). Introduction to gps/gnss technology. In *Gps and gnss technology in geosciences* (pp. 3–20). Elsevier.

Lagler, K., Schindelegger, M., Böhm, J., Krásná, H., & Nilsson, T. (2013). Gpt2: Empirical slant delay model for radio space geodetic techniques. *Geophysical research letters*, *40*(6), 1069–1073.

Landskron, D., & Böhm, J. (2018). Vmf3/gpt3: refined discrete and empirical troposphere mapping functions. *Journal of geodesy*, *92*, 349–360.

Langley, R. B., Wells, W., & de Brito Mendes, V. (1995). Tropospheric propagation delay: A bibliography. *Remote Sensing*, *28*, 811–816.

Leandro, R. F., Langley, R. B., & Santos, M. C. (2008). Unb3m_pack: a neutral atmosphere delay package for radiometric space techniques. *GPS solutions*, *12*, 65–70.

Leick, A., Rapoport, L., & Tatarnikov, D. (2015). Gps satellite surveying.

Li, L., Wu, S., Wang, X., Tian, Y., He, C., Zhang, K., et al. (2017). Seasonal multifactor modelling of weighted-mean temperature for ground-based gnss meteorology in hunan, china. *Advances in Meteorology*, *2017*.

Li, X., Dick, G., Ge, M., Heise, S., Wickert, J., & Bender, M. (2014). Real-time gps sensing of atmospheric water vapor: Precise point positioning with orbit, clock, and phase delay corrections. *Geophysical Research Letters*, *41*(10), 3615–3621.

Lindskog, M., Ridal, M., Thorsteinsson, S., & Ning, T. (2017). Data assimilation of gnss zenith total delays from a nordic processing centre. *Atmospheric Chemistry and Physics*, *17*(22), 13983–13998.

Liou, Y.-A., Teng, Y.-T., Van Hove, T., & Liljegren, J. C. (2001). Comparison of precipitable water observations in the near tropics by gps, microwave radiometer, and radiosondes. *Journal of Applied Meteorology and Climatology*, *40*(1), 5–15.

Liu, J., Chen, X., Sun, J., & Liu, Q. (2017). An analysis of gpt2/gpt2w+ saastamoinen models for estimating zenith tropospheric delay over asian area. *Advances in Space Research*, *59*(3), 824–832.

Liu, J., Yao, Y., & Sang, J. (2018). A new weighted mean temperature model in china. *Advances in Space Research*, *61*(1), 402–412.

Long, F., Hu, W., Dong, Y., & Wang, J. (2021). Neural network-based models for estimating weighted mean temperature in china and adjacent areas. *Atmosphere*, *12*(2), 169.

Lu, N., Trenberth, K. E., Qin, J., Yang, K., & Yao, L. (2015). Detecting long-term trends in precipitable water over the tibetan plateau by synthesis of station and modis observations. *Journal of Climate*, *28*(4), 1707–1722.

Maghrabi, A., Clay, R., Wild, N., & Dawson, B. (2009). Design and development of a simple infrared monitor for cloud detection. *Energy Conversion and Management*, *50*(11), 2732–2737.

Manandhar, S., Lee, Y. H., Meng, Y. S., & Ong, J. T. (2017). A simplified model for the retrieval of precipitable water vapor from gps signal. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11), 6245–6253.

Marini, J. W. (1972). Correction of satellite tracking data for an arbitrary tropospheric profile. *Radio Science*, *7*(2), 223–231.

Martellucci, A. (2012). Galileo reference troposphere model for the user receiver. *ESA-APPNG-ReF/00621-AM v2*, *7*.

Mateus, P., Catalão, J., Mendes, V. B., & Nico, G. (2020). An era5-based hourly global pressure and temperature (hgpt) model. *Remote Sensing*, *12*(7), 1098.

Mendes, V., Prates, G., Santos, L., & Langley, R. (2000). An evaluation of the accuracy of models for the determination of the weighted mean temperature of the atmosphere. , 433–438.

Mircheva, B., Tsekov, M., Meyer, U., & Guerova, G. (2017). Anomalies of hydrological cycle components during the 2007 heat wave in bulgaria. *Journal of atmospheric and solar-terrestrial physics*, *165*, 1–9.

Misra, P., & Enge, P. (2006). Gps measurements and error sources. *Global Positioning System-Signals, Measurements, and Performance, 2nd ed.; Ganga-Jamuna Press: Lincoln, MA, USA*.

Nicholas, Z. (2016). Satellite and receiver clock errors. *Retrieved from Tekmon Geomatics: http://www. tekmon. eu/1-3-2-satellite-and-receiver-clock-errors*.

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

Pany, T., Pesec, P., & Stangl, G. (2001). Elimination of tropospheric path delays in gps observations with the ecmwf numerical weather model. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, *26*(6-8), 487–492.

Penna, N., Dodson, A., & Chen, W. (2001). Assessment of egnos tropospheric correction model. *The Journal of Navigation*, *54*(1), 37–55.

Petit, G., & Luzum, B. (2010). Chapter 9 models for atmospheric propagation delays. In *Iers convention* (pp. 132–150).

Rabbath, C., & Corriveau, D. (2019). A comparison of piecewise cubic hermite interpolating polynomials, cubic splines and piecewise linear functions for the approximation of projectile aerodynamics. *Defence Technology*, *15*(5), 741–757.

Rocken, C., Van Hove, T., & Ware, R. (1997). Near real-time gps sensing of atmospheric water vapor. *Geophysical research letters*, *24*(24), 3221–3224.

Ross, R. J., & Rosenfeld, S. (1997). Estimating mean weighted temperature of the atmosphere for global positioning system applications. *Journal of Geophysical Research: Atmospheres*, *102*(D18), 21719–21730.

Rózsa, S., Weidinger, T., Gyöngyösi, A. Z., & Kenyeres, A. (2012). The role of gnss infrastructure in the monitoring of atmospheric water vapor. *Application of remote sensing and geoinformatics in environmental sciences and agriculture*, *116*(1), 1–20.

Rüeger, J. M. (2002). Refractive indices of light, infrared and radio waves in the atmosphere.

Saastamoinen, J. (1972). Atmospheric corrections for the troposphere and stratosphere in radio ranging of satellites. *The Use of Artificial Satellites for Geodesy, Geophysics Monograph Service*, *15*, 274–251.

Sapucci, L. F. (2014). Evaluation of modeling water-vapor-weighted mean tropospheric temperature for gnss-integrated water vapor estimates in brazil. *Journal of Applied Meteorology and Climatology*, *53*(3), 715–730.

Satirapod, C., Anonglekha, S., Choi, Y.-S., & Lee, H.-K. (2011). Performance assessment of gps-sensed precipitable water vapor using igs ultra-rapid orbits: a preliminary study in thailand. *Engineering journal*, *15*(1), 1–8.

Schueler, T., Pósfay, A., Hein, G. W., & Biberger, R. (2001). A global analysis of the mean atmospheric temperature for gps water vapor estimation. , 2476–2489.

Schüler, T. (2014). The tropgrid2 standard tropospheric correction model. *GPS solutions*, *18*(1), 123–131.

Seeber, G. (2003). Satellite geodesy: foundations, methods and applications. *INTERNATIONAL HYDROGRAPHIC REVIEW*, *4*(3), 92–93.

Shao, X., Ho, S.-P., Jing, X., Zhou, X., Chen, Y., Liu, T.-C., ... Dong, J. (2023). Characterizing the tropospheric water vapor variation using cosmic radio occultation and ecmwf reanalysis data. *Atmospheric Chemistry & Physics Discussions*.

Smith, E. K., & Weintraub, S. (1953). The constants in the equation for atmospheric refractive index at radio frequencies. *Proceedings of the IRE*, *41*(8), 1035–1037.

Solbrig, P. (2000). Untersuchungen ber die nutzung numerischer wettermodelle zur wasserdampfbestimmuna mit hilfe des global positioning systems, diploma. *Thesis, Institute of Geodesy and Navigation, University of FAF Munich*.

Spilker Jr, J. J. (1978). Gps signal structure and performance characteristics. *Navigation*, *25*(2), 121–146.

Ssenyunzi, R. C., Oruru, B., D'ujanga, F. M., Realini, E., Barindelli, S., Tagliaferro, G., ... van de Giesen, N. (2020). Performance of era5 data in retrieving precipitable water vapour over east african tropical region. *Advances in Space Research*, *65*(8), 1877–1893.

Sun, Z., Zhang, B., & Yao, Y. (2019). An era5-based model for estimating tropospheric delay and weighted mean temperature over china with improved spatiotemporal resolutions. *Earth and Space Science*, *6*(10), 1926–1941.

Sunehra, D. (2013). Estimation of prominent global positioning system measurement errors for gagan applications. *European Scientific Journal*, *9*(15).

Suparta, W., & Iskandar, A. (2013). Modeling of weighted mean temperature over the western pacific region to estimate gps pwv. In *2013 ieee international conference on space science and communication (iconspace)* (pp. 190–193).

Suresh Raju, C., Saha, K., Thampi, B. V., & Parameswaran, K. (2007). Empirical model for mean temperature for indian zone and estimation of precipitable water vapor from ground based gps measurements. In *Annales geophysicae* (Vol. 25, pp. 1935–1948).

Suwantong, R., Satirapod, C., Srestasathiern, P., & Kitpracha, C. (2016). Deriving the mean tropospheric temperature model using airs and amsu for gnss precipitable water vapour estimation. , 1642–1648.

Thompson, A. M., Witte, J. C., Sterling, C., Jordan, A., Johnson, B. J., Oltmans, S. J., ... others (2017). First reprocessing of southern hemisphere additional ozonesondes (shadoz) ozone profiles (1998–2016): 2. comparisons with satellites and ground-based instruments. *Journal of Geophysical Research: Atmospheres*, *122*(23), 13–000.

Tregoning, P., Boers, R., O'Brien, D., & Hendy, M. (1998). Accuracy of absolute precipitable water vapor estimates from gps observations. *Journal of geophysical research: atmospheres*, *103*(D22), 28701–28710.

Troller, M. R. (2004). Gps based determination of the integrated and spatially distributed water vapor in the troposphere.

Tuka, A., & El-Mowafy, A. (2013). Performance evaluation of different troposphere delay models and mapping functions. *Measurement*, *46*(2), 928–937.

Wang, J., Zhang, L., & Dai, A. (2005). Global estimates of water-vapor-weighted mean temperature of the atmosphere for gps applications. *Journal of Geophysical Research: Atmospheres*, *110*(D21).

Wang, M., Shan, T., Zhang, W., & Huan, H. (2021). Analysis of bds/gps signals' characteristics and navigation accuracy for a geostationary satellite. *Remote Sensing*, *13*(10), 1967.

Wang, S., Xu, T., Nie, W., Jiang, C., Yang, Y., Fang, Z., ... Zhang, Z. (2020). Evaluation of precipitable water vapor from five reanalysis products with ground-based gnss observations. *Remote Sensing*, *12*(11), 1817.

Wang, X., Zhang, K., Wu, S., Fan, S., & Cheng, Y. (2016). Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend. *Journal of Geophysical Research: Atmospheres*, *121*(2), 833–852.

Wang, Y., Yang, K., Pan, Z., Qin, J., Chen, D., Lin, C., ... others (2017). Evaluation of precipitable water vapor from four satellite products and four reanalysis datasets against gps measurements on the southern tibetan plateau. *Journal of Climate*, *30*(15), 5699–5713.

Yang, F., Guo, J., Meng, X., Shi, J., Xu, Y., & Zhang, D. (2019). Determination of weighted mean temperature (tm) lapse rate and assessment of its impact on tm calculation. *IEEE Access*, *7*, 155028–155037.

Yao, Y., Shan, L., & Zhao, Q. (2017). Establishing a method of short-term rainfall forecasting based on gnss-derived pwv and its application. *Scientific reports*, *7*(1), 12465.

Yao, Y., Zhang, B., Xu, C., & Yan, F. (2014). Improved one/multi-parameter models that consider seasonal and geographic variations for estimating weighted mean temperature in ground-based gps meteorology. *Journal of Geodesy*, *88*, 273–282.

Yao, Y., Zhu, S., & Yue, S. (2012). A globally applicable, season-specific model for estimating the weighted mean temperature of the atmosphere. *Journal of Geodesy*, *86*, 1125–1135.

Yao, Y. B., Zhang, B., Yue, S. Q., Xu, C. Q., & Peng, W. F. (2013). Global empirical model for mapping zenith wet delays onto precipitable water. *Journal of Geodesy*, *87*(5), 439–448.

Yuan, L. L., Anthes, R. A., Ware, R. H., Rocken, C., Bonner, W. D., Bevis, M. G., & Businger, S. (1993). Sensing climate change using the global positioning system. *Journal of Geophysical Research: Atmospheres*, *98*(D8), 14925–14937.

Zhang, H., Yuan, Y., Li, W., & Zhang, B. (2019). A real-time precipitable water vapor monitoring system using the national gnss network of china: Method and preliminary results. *IEEE journal of selected topics in applied earth observations and remote sensing*, *12*(5), 1587–1598.

Zhang, W., Lou, Y., Huang, J., & Liu, W. (2018). A refined regional empirical pressure and temperature model over china. *Advances in Space Research*, *62*(5), 1065–1074.

Zheng, F., Lou, Y., Gu, S., Gong, X., & Shi, C. (2018). Modeling tropospheric wet delays with national gnss reference network in china for beidou precise point positioning. *Journal of Geodesy*, *92*, 545–560.