UNIVERSITY OF NAIROBI

A NEURAL NETWORK IMPLEMENTATION FOR NEAR REAL TIME TROPOSPHERIC WATER VAPOUR PROFILING OVER NAIROBI USING GROUND-BASED GPS RECEIVER

BY

ONYANGO MICHAEL OTIENO

I56/76281/2009

A Thesis Submitted for Examination in Partial Fulfillment of the Requirements for Award of the Degree of Master of Science in Physics of the University of Nairobi.

2015

I declare that this thesis is my original work and has not been submitted elsewhere for examination, award of a degree or publication. Where other people's work or my own work has been used, this has properly been acknowledged and referenced in accordance with the University of Nairobi's requirements.

Signature: ……. Date: ……

Onyango Michael Otieno

I56/76281/2009

Department of Physics, University of Nairobi

This thesis is submitted with our approval as Research Supervisors:

Dr. Collins Omulo Mito

Department of Physics Signature: …….

University of Nairobi Date: …….

P. O. Box 30197 – 00100, Nairobi, Kenya

collins@uonbi.ac.ke

Prof. Paul Baki

Department of Technical and Applied Physics Signature: …….

Technical University of Kenya Date: …….

P. O. Box 52428 – 00200, Nairobi, Kenya

paulbaki@gmail.com

Dr. Gilbert Ouma

Department of Meteorology Signature: ……

University of Nairobi Date: ……

P. O. Box 30197 – 00100, Nairobi, Kenya

gouma@uonbi.ac.ke

DEDICATION

To my Parents, C. J. Onyango and Hellen A. Onyango

ACKNOWLEDGEMENTS

I begin by acknowledging and conveying utmost thanks to my principal supervisor Dr. Collins O. Mito for the introduction to the field of Remote Sensing and for the supervision, advice, guidance and encouragement throughout this research. I appreciate also the very useful input by my supervisors Prof. Paul Baki and Dr. Gilbert Ouma. Their insights and constructive suggestions were immensely beneficial to this project.

I thank the University of Nairobi for granting me the opportunity to conduct this research. I acknowledge too, my MSc classmates for their consultative and sharing spirit which was enriching in numerous ways. I must appreciate too, all the staff of the Department of Physics for their academic, administrative and moral support.

I cannot fail to express gratitude to the Kenya Meteorological Department and the Regional Centre for Mapping of Resources for Development for collecting and making available through the International Global Radiosonde Archive and the International Global Navigation Satellite Systems Service respectively, the radiosonde and GPS data that was used in this study.

Special thanks go to the Massachusetts Institute of Technology, the Harvard-Smithsonian Centre for Astrophysics and the Scripps Institution of Oceanography for developing and making available for free the GAMIT Software which was used to process the GPS data.

This work would never have been completed without the support, love, encouragement and prayers of my family and friends.

Most importantly, I acknowledge God's love and guidance throughout the duration of this project.

ABSTRACT

A remote sensing tool employing an Artificial Neural Networks algorithm was proposed for near real time determination of the relative humidity profile using Global Positioning System (GPS) data recorded by a ground-based GPS receiver. The GPS data was processed to obtain the Integrated Water Vapour. This Integrated Water Vapour in conjunction with ground level information for temperature, pressure and relative humidity were fed as inputs to the developed neural network which in turn generated the instantaneous relative humidity profile as output.

GPS and radiosonde data for the years 2009 and 2010 were used to train the system while the same data for 2011 were used to validate the system. The RH profile results for 2011 generated using GPS data and the neural network, upon comparison with recorded *in situ* radiosonde relative humidity profile measurements for the same days and times in the year 2011, had Root Mean Square Error of less than 4%, which fell within the margin of error of the Vaisala RS92 Radiosonde's humidity measurement regime.

TABLE OF CONTENTS

ABBREVIATIONS

AGU – American Geophysical Union

ANN – Artificial Neural Networks

EUMETSAT – European Organisation for the Exploitation of Meteorological Satellites

GAMIT – GPS Analysis at the Massachusetts Institute of Technology

GNSS – Global Navigation Satellite Systems

GPS – Global Positioning System

GPSRO – Global Positioning System Radio Occultation

IGRA – Integrated Global Radiosonde Archive

IGS – International GNSS Service

IP – Internet Protocol for Telecommunication Networks

ITU – International Telecommunications Union

IWV – Integrated Water Vapour

KMD – Kenya Meteorological Department

LEO – Low Earth Orbit Satellite

MEO – Medium Earth Orbit Satellite

NCDC – National Climatic Data Centre

NOAA – National Oceanic and Atmospheric Administration

PWV – Precipitable Water Vapour

RCMRD – Regional Centre for Mapping of Resources for Development

RH – Relative Humidity

RINEX – Receiver Independent Exchange Format Data

RMSE **–** Root Mean Square Error

UTC – Coordinated Universal Time

WMO – World Meteorological Organisation

ZHD – Zenith Hydrostatic Delay

ZTD – Zenith Tropospheric Delay

ZWD – Zenith Wet Delay

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 The Troposphere

The Troposphere is the lowest of the earth's atmospheric layers and holds approximately 80% of the atmospheric mass and 99% of its water vapour content. Most of the parameters associated with the temporal evolution of the weather are generated in this layer. The average height of the troposphere is 7 km in the Polar Regions and up to 20 km at the low latitudes (Danielson *et al*., 2003). The Troposphere height has been shown to increase with increase in mean temperature: the air in cold regions experiences much less convection than in the warmer regions, which explains the lower heights at the poles and higher heights at the equator (Geerts and Linacre, 1997).

Pressure levels in the troposphere and indeed the atmosphere are highest at sea level and decrease exponentially with increase in altitude. Temperature too, is higher at low altitudes and generally decreases linearly with increase in altitude until the end of the troposphere where there is a discontinuity called the Tropopause, at which there is a temperature inversion and temperature starts to increase with altitude into the Stratosphere. Figure 1-1 represents this.

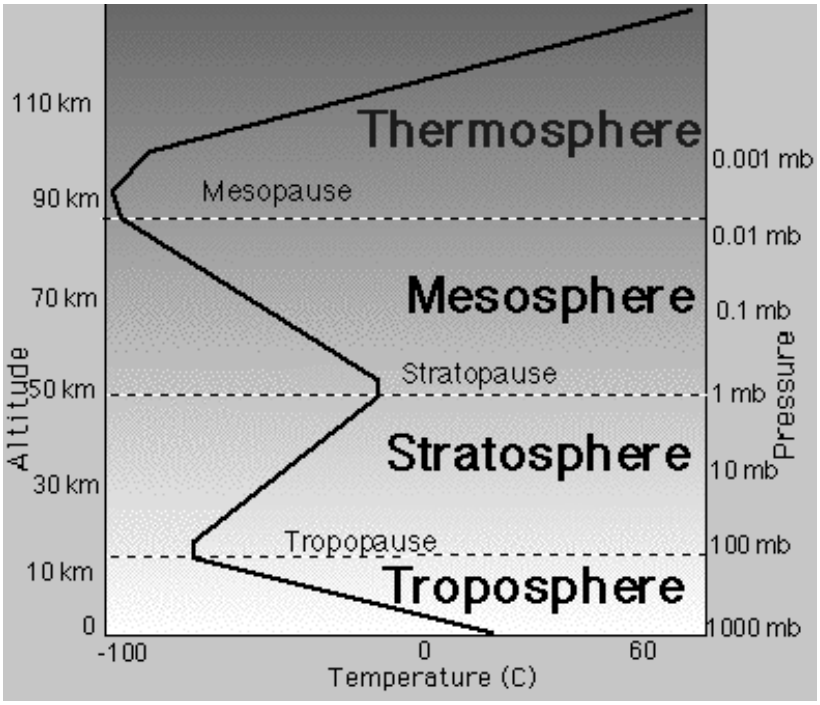

Figure 1-1: Graphical representation of Earth's atmospheric layers (Artinaid, 2014)

The chemical composition of the troposphere, as shown in Table 1-1, is generally uniform, components having same distribution at all levels.

Table 1-1: Composition of the Troposphere (NOAA, 1993)

Constituent	**Tropospheric Mixing Ratio**
Nitrogen	78 %
Oxygen	21 %
Water Vapour	≤ 4.00 %
Argon	1 %
Carbon Dioxide	0.0345 %
Neon	0.0008 %
Helium	0.0005 %
Methane	0.00017 %
Nitrous Oxide	0.00003 %
Ozone	0.00001 %
CFC	0.00000006 %

The exception to this rule is water vapour whose quantity in the troposphere and distribution in the spatial and temporal dimensions is very variable. It reduces due to precipitation and increases majorly due to evaporation of surface water. Other processes like plant transpiration, combustion, volcanism etc. are also contributors of atmospheric water vapour.

1.2 Water Vapour in the Troposphere

Even though it constitutes less than 5 % of the atmosphere's mass, water vapour has a tremendous influence on the spatial and temporal evolution of the Earth's atmospheric properties (Mockler, 1995). To begin with, water vapour is the most abundant of the atmospheric greenhouse gases and also the one that exhibits the most temporal and spatial variation (NCDC, 2013). As explained in Braun *et al.* (2004), it influences the Earth's radiation budget, energy transfer, cloud formation and the resultant precipitation distribution. Both the solar radiation and the longwave terrestrial radiation are absorbed by atmospheric water vapour. The energy absorbed as latent heat of vaporisation at the equator is transferred from the equator towards the poles by atmospheric circulation via Hadley cells and released as latent heat of condensation when the water vapour condenses at the higher latitudes. The energy released when water vapour condenses causes disturbances in the Troposphere's vertical stability influencing global weather systems and their associated precipitation patterns.

According to the Clausius-Clapeyron relation (Lawrence, 2005), the air's capacity to contain water vapour increases with increase in air temperature. With rising atmospheric temperature therefore, the atmosphere is bound to contain more water vapour. An increase in water vapour at low altitudes will reduce the amount of outgoing long-wave terrestrial radiation, as more radiation will be absorbed, and the atmospheric equilibrium temperature, corresponding to the long-wave radiation will have to rise to balance the incoming solar radiation. This water vapour feedback has the effect of doubling the extent of global warming due to greenhouse gases (Dessler *et al.*, 2008).

1.3 Water Vapour Measurement

Humidity is the measure of the amount of water vapour in the air. Tropospheric humidity profiling therefore refers to the representation of the variation of the tropospheric water vapour in the vertical dimension. Accuracy in weather prediction is greatly improved when there is accuracy in tropospheric profiling (Marchuk, 1974).

Quantification of the amount of water vapour in the air may be accomplished in terms of: water vapour mixing ratio, specific humidity and relative humidity. The water vapour mixing

ratio is the ratio between the mass of water vapour and the mass of dry air contained in a volume of air. Specific humidity is the amount of water vapour measured in grams contained in a volume of air with mass of one kilogram. The most commonly used parameter for specifying humidity measurement is the Relative Humidity (RH), which is the ratio of the water vapour pressure to the saturation vapour pressure, the saturation vapour pressure being the pressure at which the water vapour contained in the air starts to condense (Mockler, 1995). At 100% RH, the air can be said to be saturated with water vapour. As mentioned by Peixoto and Oort (1996), Relative Humidity is the most commonly used humidity definition for general applications such as in agriculture, hydrology, management of water resources and human comfort among others. Integrated Water Vapour (IWV) is the total amount of water vapour in the atmospheric column overlying a given point per unit area, the SI unit being in kilograms per square metre. Upper air water vapour can be measured directly using radiosondes or using humidity sensors fitted onto aircrafts. It can also be measured indirectly via remote sensing techniques.

The classical tool for tropospheric water vapour profiling is the radiosonde whose setup is a meteorological probe unit attached to a balloon. In the course of its flight, the radiosonde measures the pressure, altitude, position, temperature, humidity and wind (speed and direction) and transmits this information via radio to a ground-based receiver connected to a recording computer (Du Bois *et al.* 2002). While the instantaneous data obtained from radiosonde is quite accurate (Nash *et al.* 2011), the high cost of the equipment means it can only be launched a limited number of times in a day and even then only at a limited number of stations. Significant variations in the weather parameters that occur between the radiosonde launches go undetected (Elliot, 1995). The deployment of radiosondes is sparse in the tropics especially as compared to the northern hemisphere (WMO, 2005). Figure 1-2 shows the global distribution of radiosonde stations that have shared data with the International Global Radiosonde Archive (IGRA) which is a common database storing data collected from radiosonde stations all over the world, including those that record data only a few times in a year (Durre *et al.*, 2006).

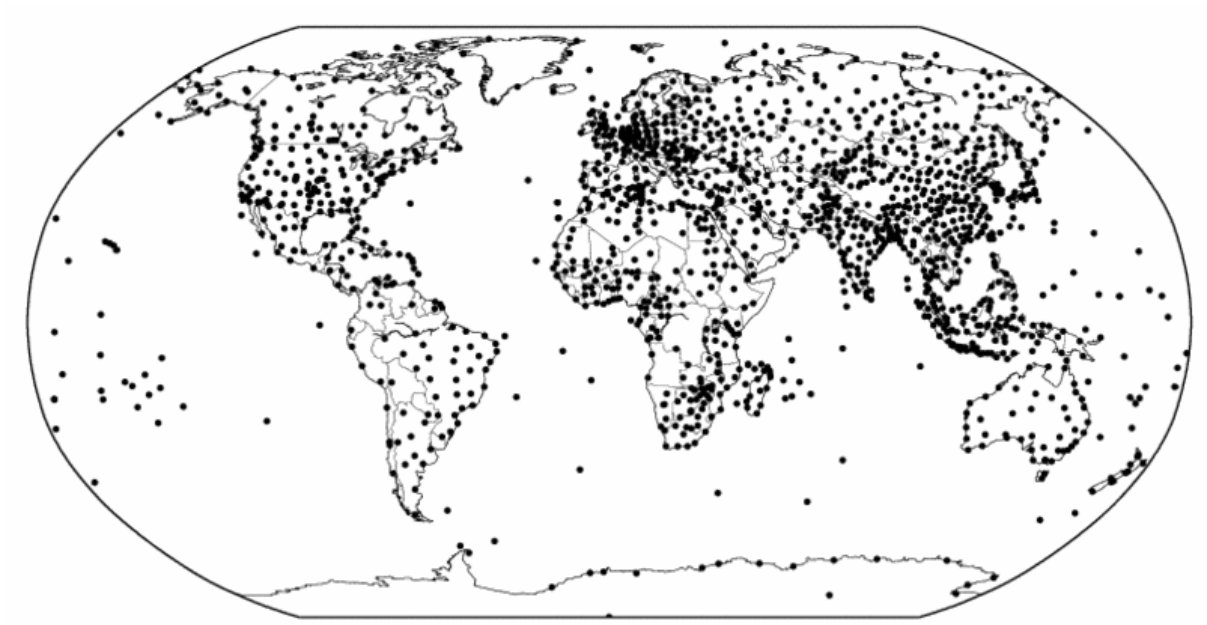

Figure 1-2: Locations of global radiosonde stations that have shared data with IGRA (Durre *et al.*, 2006)

The scarcity of radiosonde data in the tropics and especially over the oceans makes the reliability of the predictions from this data insufficient over the large scale. It is necessary that techniques that enhance the availability of meteorological data be employed for such regions. Remote sensing enables greater detail in the observation of meteorological parameters in regions where *in situ* observational tools would provide inadequate resolution due to low distribution of in situ observation stations. Remote Sensing is therefore a very important method for enabling greater resolution in regular meteorological observation for many tropical regions.

Imaging via satellite makes it possible to conduct qualitative analysis of the horizontal distribution of water vapour. The Meteosat satellite system for instance evaluates the thermal emission from tropospheric water vapour in the 5.7 μm to 7.1μm band (EUMETSAT, 2013). During cloud-free periods, acceptably accurate profile information on water vapour distribution can be determined from these observations. This is because in clear weather, the Troposphere is transparent to the thermal emissions from water vapour and these can be observed by satellite sensors; but in cloudy weather, the Troposphere is opaque to these thermal water vapour thermal emissions. The advantage of the observations via Meteosat is the high update frequency meaning good temporal resolution, the disadvantage being its degraded performance during cloudy periods which makes it unfavourable for use in the

tropical regions because many parts of the tropics are very cloudy especially during the rain seasons. Figure 1-3 is an example of a Meteosat image representing the relative humidity at the middle and upper troposphere.

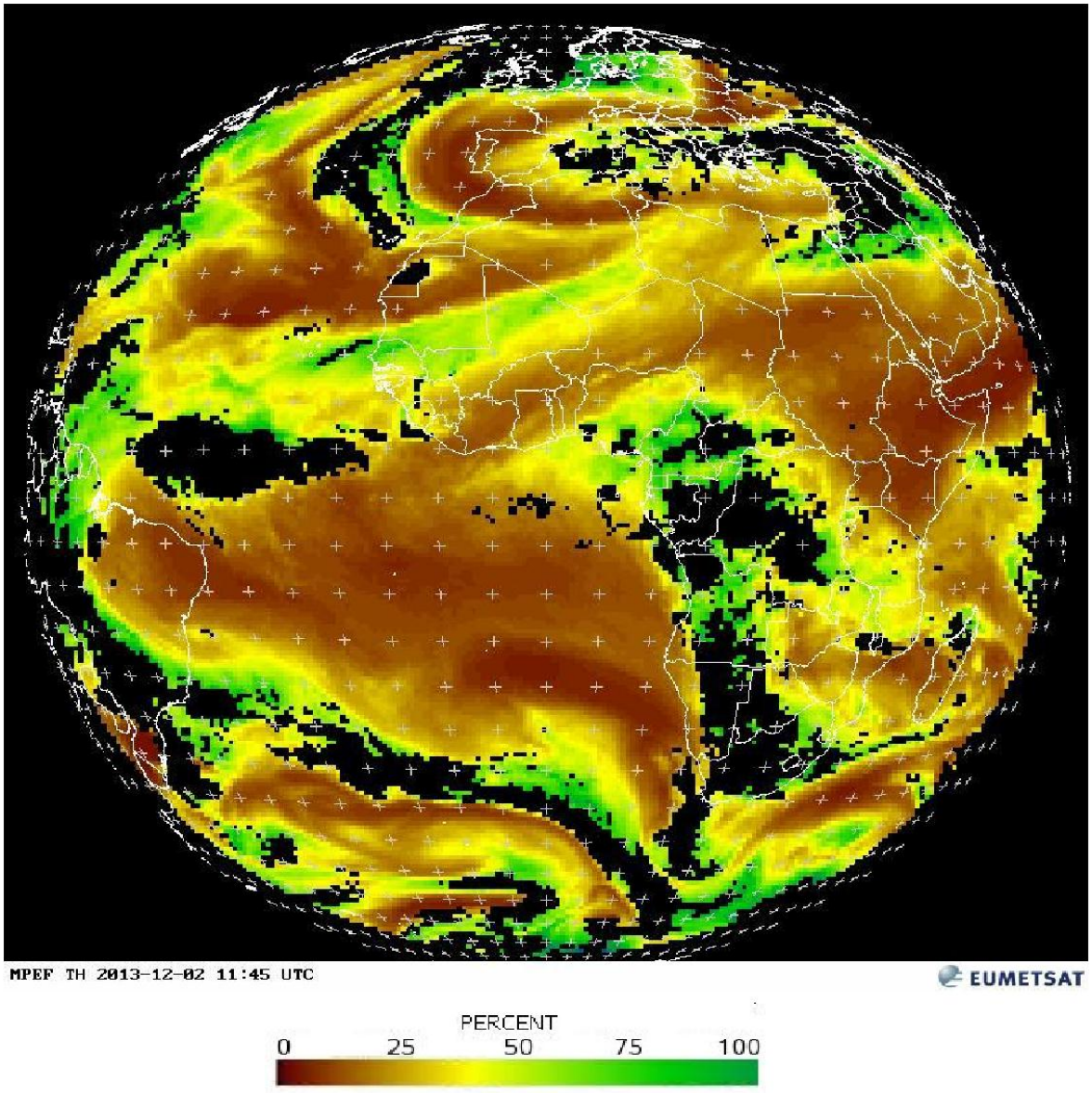

Figure 1-3: Meteosat image representing relative humidity at mid and upper layers of the Troposphere (EUMETSAT, 2013)

Water vapour radiometers, which measure the background microwave radiation produced by the atmospheric water vapour can estimate IWV with very good accuracy (Resch, 1984); however as pointed out by Bevis *et al.* (1992), the radiometers do experience some degradation in performance under heavy rainfall. A Hong Kong study by Chan (2009) on the performance of ground based radiometers under intense convective weather reported rapid

fluctuations in the IWV data generated by radiometers operating under conditions with rain rates of 30 mm/hour and above that made radiometer data inconsistent with evaluations from radiosonde and GPS data under similar conditions.

Evaluation of atmospheric water vapour using the Global Positioning System helps to overcome some of these challenges. Studies have shown that the evaluation of GPS signal phase delay due to water vapour can be performed to great degree of accuracy even under very rainy or cloudy weather (Solheim *et al.*, 1999). This suggests that an atmospheric water vapour monitoring system employing GPS would be reliable under diverse weather conditions including those under which radiometers and satellite imaging typically experience performance degradation.

A ground based GPS Receiver can be used to determine the flight time of a time-stamped GPS signal. On processing the raw data obtained from it, an estimate of the signal phase delay due to the atmospheric water vapour can be evaluated, from which the Integrated Water Vapour can be computed. The density of GPS receivers deployed globally continues to grow year by year (IGS, 2013) making the spatial resolution of the IWV data obtained by GPS potentially very high. Figure 1-4 shows the global distribution of stations equipped with GPS Receivers capable of IWV evaluation that share their data with the International Global Navigation Satellite Systems (GNSS) Service (IGS).

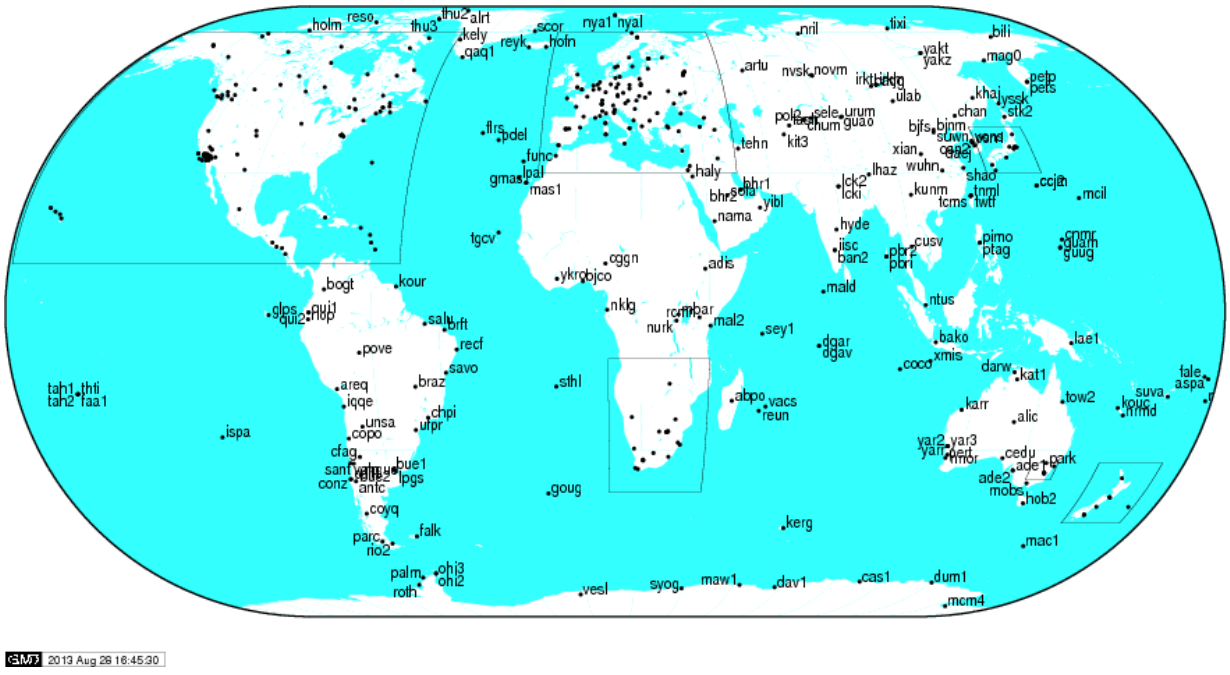

Figure 1-4: Map of Global GNSS Stations that share data with the IGS (IGS, 2013)

Rather than restrict the water vapour observations to the zenith IWV properties, it would be more beneficial to meteorologists if the water vapour's vertical profile were evaluated too. A modelling technique to relate the IWV to the RH profile can be employed to execute this function. This adds a second dimension to the water vapour data. With increase in density of the GPS receivers, a three-dimensional evaluation of atmospheric RH is possible too. While having high horizontal resolution is not very important for upper air observation, it is very useful in the boundary layer to detect small scale variations in water vapour distribution.

In this study, a remote sensing technique employing a ground-based GPS receiver for tropospheric water vapour profiling was proposed. The instantaneous vertical profile of RH was determined by processing of the GPS signal phase delay data as observed by a ground-based GPS receiver. This technique will enable a much more widespread deployment of tropospheric relative humidity profiling stations because of the low cost of GPS receivers and also the ease in deploying them. In addition, taking advantage of the GPS constellation's 24-hr availability and worldwide coverage, this model enables near real-time tropospheric humidity profiling at any point on the Earth's surface.

1.4 Importance of Water Vapour Profiling in Meteorological Forecasting

Knowledge of the atmospheric water vapour profile and its trends are useful in the prediction of precipitation. Accuracy in precipitation forecasting improves with increase in frequency of water vapour sampling (Schuman, 1978). With real time knowledge of the water vapour profile therefore, the short term prediction of cloud formation for example, would be much more accurate, hence a more accurate prediction of degree of cloud cover and precipitation.

Accuracy in prediction of storm severity for instance depends on knowledge of tropospheric water vapour distribution. Studies by Park and Droegemeier (2000) indicated that even a 1% variation in the water vapour content in the environment of a storm had a statistically considerable impact on the severity of the storm. Researchers at the Georgia Tech Research Institute (2010) underscore how important water vapour information is in weather forecasting, pointing out that a small increase in water vapour can mean the difference between no precipitation and a major storm.

1.5 Study Area

Kenya's capital city, Nairobi, at latitude approximately 1°17′S and longitude approximately 36°49′E, lies at averagely 1,700 m above sea level. At this altitude, Nairobi's climate is moderate, classified as Subtropical Highland Climate (McKnight and Darrel, 2000).

The data used in this study was collected at two stations, both within Nairobi. The first station, which was the source of all the radiosonde data, was the Kenya Meteorological Department (KMD) Upper Air Observatory, located in Dagoretti Corner (1°20′S, 36°46′E, altitude 1795 m). The second station was the Regional Centre for Mapping of Resources and Development (RCMRD). All the GPS Data was recorded at this station located in Kasarani, on the north eastern side of Nairobi (1°13′S, 36°54′E, altitude 1605 m).

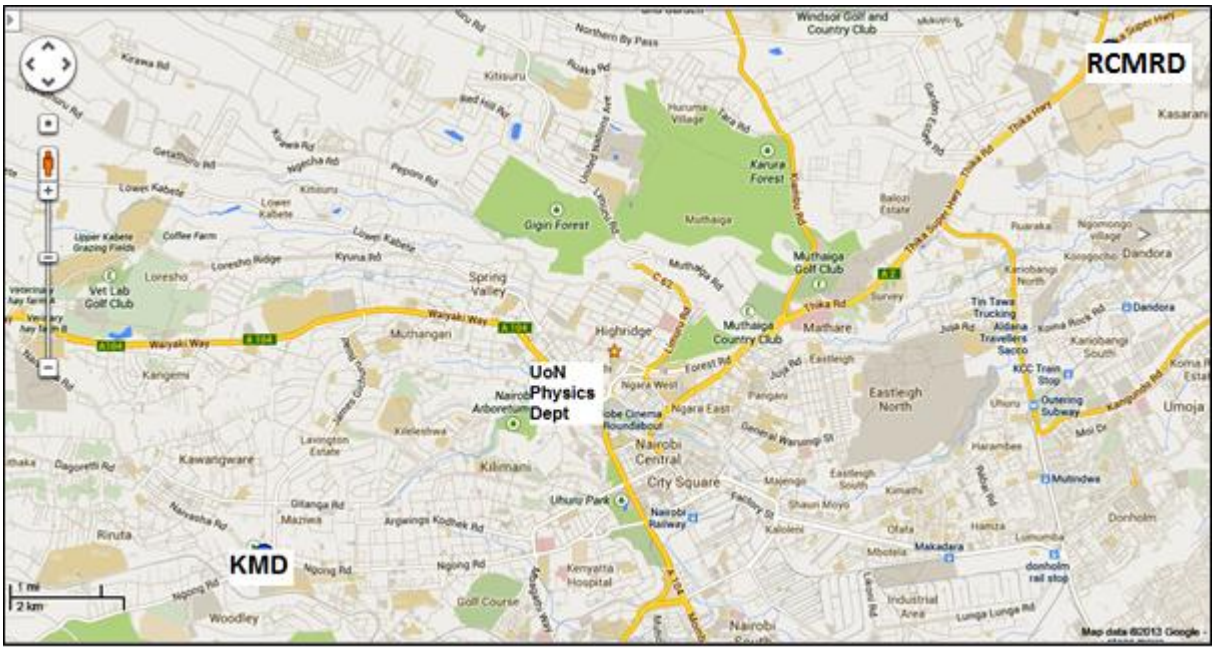

Figure 1-5: Map of Nairobi showing the locations of the stations (Google Maps, 2013)

The two stations are 17.53 km apart and the height difference between them is 190 m. The topography of the region separating the two stations is a gentle slope from Dagoretti Corner to Kasarani with no intervening mountains or hills as shown in the profile in Figure 1-6. At either site, observing from a height of 50 m above ground level, there is a clear line of sight between the two stations. For this reason it was expected that it would be possible to determine by extrapolation the ground level temperature and pressure at one station if the ground level temperature and pressure at the other station were known.

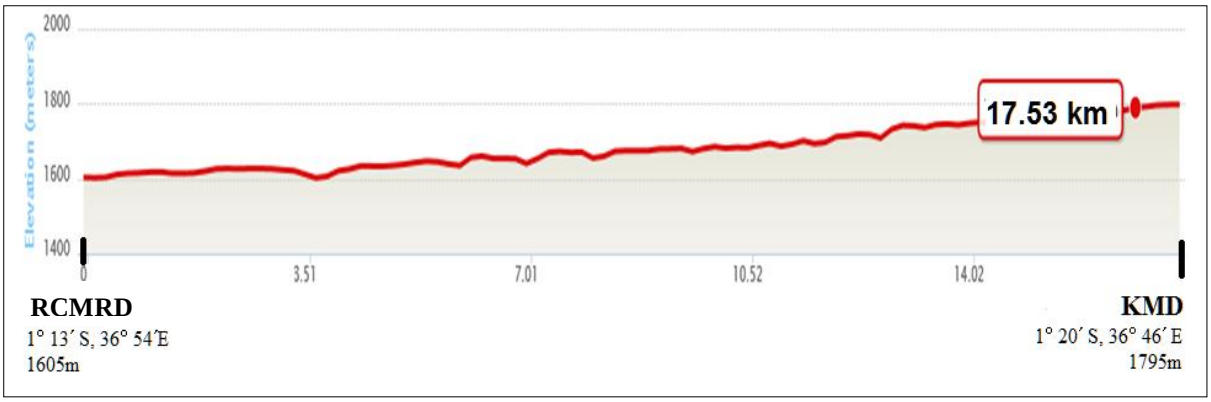

Figure 1-6: Topographical cross-section between RCMRD and KMD (Google Maps, 2013).

Naturally, it was expected that the IWV evaluated from GPS data recorded at RCMRD would be greater than the IWV calculated using radiosonde data collected at KMD because RCMRD lies at lower altitude than KMD. If the IWV at KMD and IWV at RCMRD exhibited a known relationship, then this relationship would be used to evaluate the IWV at one station if the IWV at the other station was known.

1.6 Problem Statement

Improvements in the resolution of atmospheric water vapour measurement have been demonstrated to lead to increased accuracy in the prediction of precipitation levels (Schuman, 1978). Weather forecasting would be more accurate if the water vapour profiling tools had the highest temporal, vertical and horizontal resolution possible. If the atmospheric water vapour profile could be determined in real time, then the accuracy of short term weather prediction would also greatly ameliorate.

Radiosondes and water vapour radiometers, the equipment most commonly employed currently in tropospheric water vapour vertical profiling have temporal resolution limitations which limit their usefulness for short term weather forecasting. These equipment are also costly to deploy, meaning their data can be obtained from only a limited number of stations, lowering also the horizontal resolution of this data.

A tropospheric water vapour profiling tool that has high temporal resolution and also that operates under all weather conditions would be very helpful. Making the tool low in cost

would enable its widespread and dense deployment, increasing also the horizontal resolution of its data.

1.7 Research Questions

The accuracy of GPS in determination of IWV over a site is already proven by many studies (Bevis *et al.*, 1992; Braun *et al.*, 2004; De Haan, 2008). To what extent does the IWV above a site influence the RH profile above the same site? Can this relationship between IWV and RH profile be accurately evaluated? And can this relation be used to determine the RH profile if IWV is known? What is the influence of ground level conditions on RH profile? If the ground level temperature, pressure and RH are known, can this information be useful in evaluating RH profile? Is the method of Artificial Neural Networks (ANN) a good method to use in predicting these relations? How accurate would the predictions of RH profile from IWV using ANN be?

1.8 Hypothesis

As a general trend, the RH profile for sites in tropical regions takes a C-shape, with high RH values at the low altitudes, low RH values at the mid-troposphere and high RH values as we approach the Tropopause (Folkins *et al.*, 2002). There ought to exist a relationship between the IWV above a site and the water vapour profile above that site – high IWV should lead to higher humidity values at each altitude level and vice versa. The technique of ANN can be used to model this relationship. Taking advantage of the accuracy of GPS in determining the IWV over a site, it is possible to develop a water vapour profiling algorithm that combines GPS and neural networks in evaluating the water vapour profile accurately, GPS determining the IWV and the neural network, taking the IWV as input, determining the Relative Humidity profile, factoring in the ground level temperature, pressure and relative humidity.

Because of the high accuracy of GPS receivers and the highly developed predictive capability of ANN, the technique of water vapour profiling using a ground-based GPS receiver and ANN should be as accurate as existing water vapour profiling techniques e.g. radiosonde.

1.9 Objectives

1.9.1 Main Objective

The main objective was to develop and validate a tropospheric water vapour profiling tool that employed the technique of Artificial Neural Networks to determine the instantaneous values of the vertical profile of atmospheric relative humidity using the IWV evaluated from GPS data collected by a ground based GPS receiver.

1.9.2 Specific Objectives

The specific objectives were to:

i. Use GAMIT Software to process GPS data collected for selected days of the years 2009 to 2011 using the GPS receiver at RCMRD – Kasarani in order to determine the IWV over the GPS site.
ii. Develop an Artificial Neural Network (ANN) model that would determine the atmospheric RH profile as output using ground level measurements (temperature, pressure, RH and IWV) as inputs.
iii. Train the developed system using IWV as evaluated from GPS data for 2009 and 2010 as inputs and RH profile observed using radiosonde for 2009 and 2010 as the targets, the GPS-derived IWV data corresponding to the same days and time of day as the radiosonde data.
iv. Validate the developed ANN algorithm, using GPS data for randomly selected days of a different year, 2011; evaluating its accuracy by comparing its output RH profile values with actual observed and recorded RH profile values from radiosonde data for the same randomly selected days of 2011.

1.10 Justification and Significance of the Study

The very survival of the human race is partly predicated on the ability of human societies to reduce their vulnerability to weather and climate changes. Weather catastrophe mitigation and even timely evacuations depend greatly on the ability to predict the weather. Success in

events planning also requires good weather forecasting. Knowledge of the spatial and temporal distribution of precipitation is useful in the management of agriculture, water usage, public health and renewable energy (Katz et al. 1997). A nationwide study in the United States of America indicated that the American public made use of the weather forecast service provided by the American government's meteorological bureau, generating an average of US$ 31.5 billion per year in benefits compared to costs of US$ 5.1 billion (Lazo et al., 2009). Weather and climatological information would typically be obtained by observational analysis and predictive modelling. This demands that the instruments for weather parameters information acquisition be accurate and that the tools for predictive modelling be reliable.

In order to accurately predict and model the weather, it is important that information on the vertical profile of the weather parameters be accurately and frequently collected. The accuracy of short term weather forecasting improves with increase in frequency of weather parameters sampling and measurement. Schuman (1978) reported evidence that improving the horizontal resolution of weather models improved the wind, temperature and precipitation forecasts. Weather forecasting would therefore be more accurate if the weather parameters information were gathered with the highest temporal and horizontal resolution possible. To enable good horizontal resolution of water vapour profiles, then the water vapour profiling tools ought to be low in cost in order that their deployments are widespread.

The types of equipment currently employed in tropospheric water vapour profiling have certain limitations which this study seeks to help alleviate. A radiosonde system has good vertical accuracy but due to its high cost, can only be launched a limited number of times in a day and even then only at a limited number of stations. Changes occurring between radiosonde launches go undetected and unrecorded. Consequently, radiosonde data has poor temporal resolution. Water Vapour Radiometers have good accuracy in water vapour measurement, but experience some performance degradation under very heavy rainfall (Chan, 2009). A reliable meteorological observation tool ought to operate and collect accurate data under all weather conditions.

In this work, a model that generated accurate data on the vertical profile of atmospheric humidity using a ground based GPS receiver was proposed. If the accuracy of existing water vapour profiling tools could be matched by the proposed tool, then because this proposed tool will be low in cost, it will be possible for it to be deployed widely and with high density.

Tropospheric humidity profile data will now be available for many more locations than currently possible; meaning the horizontal resolution of humidity profile data will be much higher. Future studies could focus on improving the vertical resolution of the proposed tool.

This will improve the accuracy of both short term and long term weather forecasting. With the cost of GPS receivers getting lower and lower each year, it is expected that these GPS-based tropospheric profiling stations be deployed with even greater proliferation. One exciting possibility is the convergence of mobile telephony and GPS meteorology. Virtually all smartphones are nowadays equipped with GPS receivers as a standard feature (Malm, 2009). This gives rise to the possibility of the development of mobile phone applications which could turn every smartphone into a mobile tropospheric profiling station.

1.11 Report Outline

Following the introduction in this chapter in which the objectives for the study are outlined, Chapter 2 presents the literature review. Chapter 3 then explores the theoretical background of the study and describes the basics of the Atmosphere's influence on radio signals, the basics of GPS Meteorology, Radiosonde and Artificial Neural Networks.

The methodology used to achieve the specific objectives is expressed in Chapter 4, where the data acquisition and selection criteria are explained. The techniques for processing the meteorological and GPS data are also described. Also described is the procedure of design and implementation of the neural network that was used for RH profiling. Chapter 5 presents the results of the study and their accuracy assessment. The details of the analysis and discussion into the results are also included in this chapter.

Chapter 6 presents the conclusions from this study and recommendations for improvement of the ANN technique.

CHAPTER 2. LITERATURE REVIEW

Meteorology, being an observational science, is highly dependent on the accuracy of the observation instruments. Meteorological knowledge is enhanced by improvements in the instruments of the meteorological observation network. Atmospheric scientists have done a great deal of work towards the development and improvement of techniques for measuring the horizontal and vertical distribution of weather parameters (Schuman, 1978). The most important tool for observing and recording the vertical profile data for weather parameters is the radiosonde. Radiosonde collects weather data with good vertical resolution and also good accuracy. The radiosonde equipment deployed at the Kenya Meteorological Department for instance, measures temperature with accuracy of 0.5 °C and relative humidity with accuracy of 5% (Vaisala, 2011). Because of the low temporal resolution of radiosonde data however, a lot of the changes that occur between launches are missed. Limitations due to the unavailability of data on temporal variation of weather parameters are the major source of errors in short term (0-24 hr.) weather forecasting (Hooke, 2000). In addition, radiosonde observation stations are very sparsely deployed in the tropical regions (WMO, 2005) hence low horizontal resolution of tropical radiosonde data.

Meteorological observation via remote sensing is one of the solutions to the problems mentioned above because of its inherently much higher spatial and temporal availability. This enables weather data collection for many more regions of the earth than currently possible using *in situ* data acquisition techniques. The assimilation of remote sensing data into numerical weather prediction models has greatly improved the performance of the models (Marshall *et al*., 2005). Remote sensing via microwave radiometers has better temporal resolution than radiosonde but experiences degraded performance under heavy rain, as demonstrated in the study by Chan (2009).

Bevis *et al*. (1992) introduced the concept of GPS meteorology outlining methods in which atmospheric effects on GPS signals could be used for meteorological instrumentation. The propagation speed of radio waves through the Earth's neutral atmosphere is influenced by the atmosphere's refractive index, which in turn depends on the atmospheric temperature, pressure and humidity. There are currently three methods of GPS meteorology:

i. Determination of Integrated Water Vapour (IWV) from GPS signal phase delay using ground-based GPS receiver.

ii. Vertical profiling of temperature and water vapour using radio occultation techniques.
iii. Tropospheric water vapour tomography using networks of ground-based GPS receivers.

IWV is determined by evaluating the Wet Delay, which is the delay due to the contribution of the dipole component of water vapour to atmospheric refraction, and which can be calculated from the atmospheric wet refractivity. This wet delay is nearly proportional to the amount of Precipitable Water Vapour (PWV) above the GPS receiver (Askne and Nordius, 1987). The IWV is defined as the product of PWV and the density of water.

While this method determines the total column IWV over a site quite accurately, it is not helpful when the vertical profile of water vapour is required. This is because the technique attributes the total phase delay to the total water vapour content of the overlying atmospheric column rather than its distribution along the ray transmission path.

The technique of radio occultation was originally developed and applied to the study of the atmospheres of several solar planets and their moons (Lindal, 1987). Bevis *et al*. (1992) proposed that the technique of radio occultation using GPS signals could be used to study the Earth's atmosphere. GPS Radio Occultation (GPSRO) technique enables the vertical profiling of not only water vapour, but also temperature and pressure. In this method, a GPS receiver is attached to a Low Earth Orbit (LEO) satellite located on the opposite side of the Earth's planetary limb from the GPS transmitter. As illustrated in the Figure 2-1 below (Hardy *et al.*, 1992), for sixty seconds prior to the occultation of the GPS satellite from the LEO satellite by the Earth, the signals travelling from the GPS transmitter to the LEO satellite receiver are refracted by the Earth's atmosphere. The extent of refraction can then be evaluated.

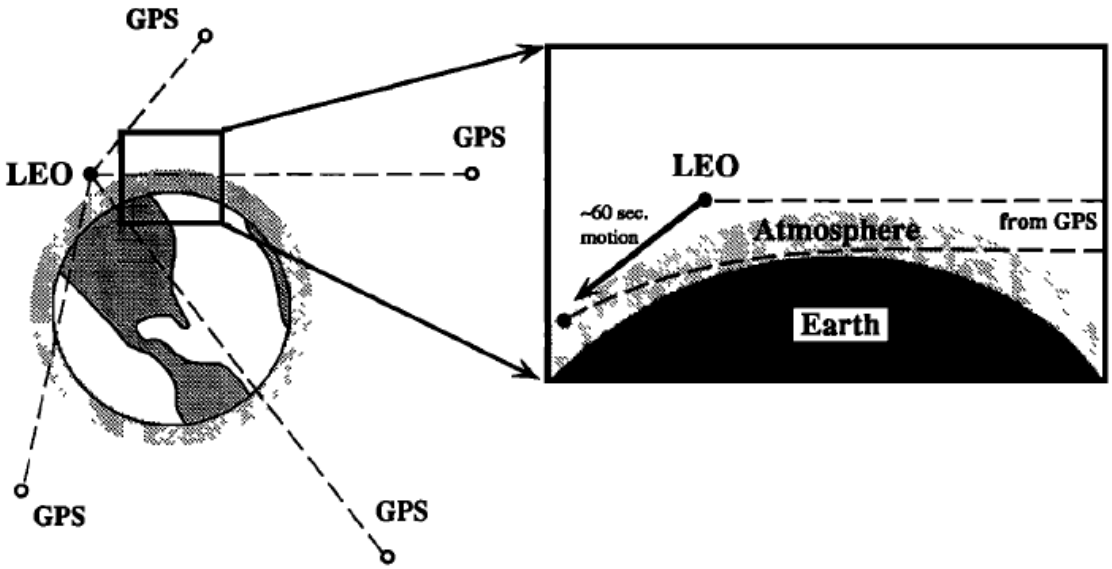

Figure 2-1: Diagrammatic representation of GPS Radio Occultation (Hardy *et al.*, 1992)

Using the GPS-obtained refractivity profile of the atmosphere, it is possible to determine these meteorological parameters (water vapour, temperature and pressure) from heights of ~60km up to the surface with vertical resolution of 0.5 km in the lower troposphere (Kursinski, 1997; Kursinski *et al*., 1997).

Accurate atmospheric water vapour profiling using GPSRO requires the knowledge of temperature profiles obtained from independent observations e.g. radiosondes. Where radiosonde data is unavailable, GPSRO suffers inaccuracies due to errors in estimation of the temperature profiles. At lower altitudes, the temperature and water vapour profiles obtained from GPSRO suffer inaccuracies due to the strong interaction between temperature and water vapour via the Clausius-Clapeyron relation which makes it difficult to isolate the refractivity effects of temperature from those of water vapour (Kursinski *et al*., 1997). Moreover, in the case whereby the time variation of atmospheric refractivity at a targeted location was required, the atmospheric sampling by occultation for that location would be too sparse temporally to enable adequate resolution.

Basili *et al*. (2007) proposed a neural network model to be used in conjunction with GPSRO which determined the refractivity profile hence weather parameters profile without the need for independent knowledge of the temperature profile at each GPS occultation. The use of neural networks eliminated the need for externally sourced temperature profile data. The

resultant vertically averaged RMSE for the water vapour pressure was found to be 0.21 hPa which was within the climatological standard deviation of 0.24 hPa.

Ground-based techniques for evaluating tropospheric profiles have the potential to provide more information than space-based measurements about horizontal variations of refractivity profiles because of the possibility to deploy numerous GPS receivers enabling high density and wide area coverage. There is therefore potential for greater accuracy in tropospheric profiling if ground-based GPS receivers are used rather than GPSRO. Lowry *et al*. (2002) developed a model which estimated the refractivity profile using ray propagation models to fit GPS tropospheric delays measured at ground level using a least squares metric. In order to improve the performance of such a system, it was recommended to employ a more robust method of data parameterisation. Applications of ground-based GPS receivers for refractivity profiling are few because for ground-based observations where the receiver altitude is near that of the ground horizon, the ray elevation is positive and the inversion to determine the refractivity profile is described by an ill-posed Fredholm integral equation of the first kind whose solution is unstable in the presence of observational errors (Lowry *et al*., 2002).

Leke *et al.* (2009) showed that for a single ground-based GPS receiver, an algorithm using neural networks and Slant Phase Delay predicted the profile with better accuracy than existing meteorological models, yielding RMS average error of 2.97 for wet refractivity and 3.06 for hydrostatic refractivity, an improvement on the general meteorological model for evaluating refractivity profile based on the Hopfield Model which yields average RMS errors of 4.77 for wet refractivity and 3.17 for hydrostatic refractivity (Hopfield, 1969). This refractivity profile, when evaluated in conjunction with the temperature, and pressure profile obtained from independent sources, could then yield the RH profiles.

A simulation on the expected performance of the algorithms to retrieve RH profile values from Microwave Radiometers and Imagers aboard the Megha Tropiques, a French-Indian satellite launched in October 2011, stated that a neural network, with the satellite observables as input and RH profile as output would have high accuracy at the low and medium altitudes, while accuracy would reduce at the high altitudes. However, at the very low altitudes, where pressure is above 1000 hPa, the accuracy would also greatly reduce (Sivira *et al*., 2012).

In this study, a method that employed the method of Artificial Neural Networks was used to determine the vertical profile of atmospheric RH using as the predictor the IWV calculated

from the measured phase delay of GPS signals as observed by a ground based GPS receiver. The use of ANN enabled the model to adapt easily to the changes in the factors that determine the RH profile due to the more effective model-training regime of the ANN technique (Beale *et al*, 2009). The ease in adaptability of ANN also gave more robustness to the system in the face of diverse possibilities of weather parameters configuration (Anderson and Rosenfeld, 1998).

CHAPTER 3. THEORETICAL BACKGROUND

3.1 The Global Positioning System

The Global Positioning System (GPS) is an assembly of radio transmitters attached to twenty-eight Medium Earth Orbit (MEO) satellites at an altitude of approximately 20,200 km transmitting to specialised radio receivers that employ the processing of these signals in navigation and relative time and position determination (Lieck, 1990). The satellites are distributed in six orbital planes as shown in figure 3-1.

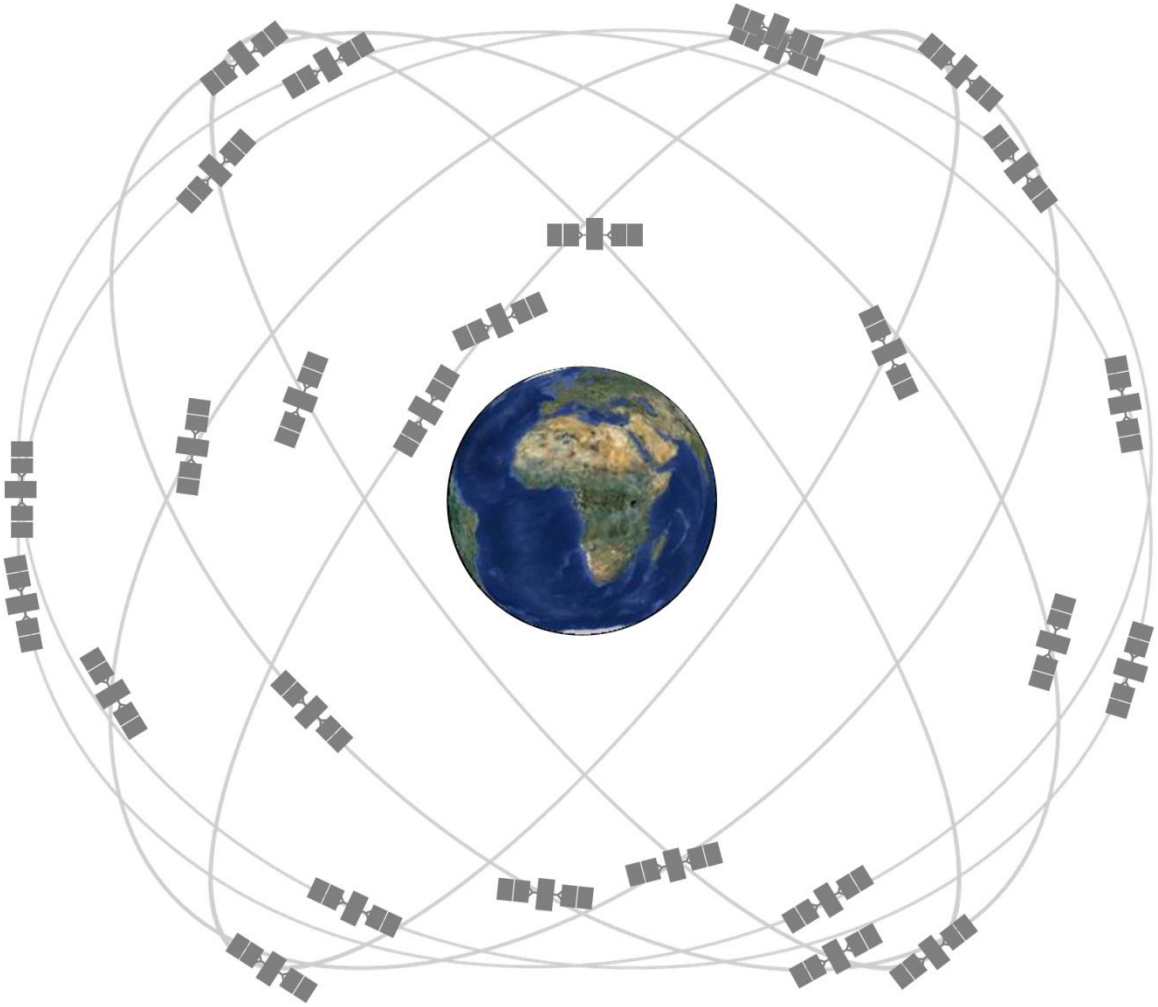

Figure 3-1: Constellation of GPS Satellites orbiting the Earth (GPS.gov, 2013).

The design of the GPS constellation is such that at any one time; any spot on the earth's surface would have at least four of the satellites visible allowing for triangulation using the satellites to determine position and time. This is represented in figure 3-2 (De Haan *et al.*, 2008).

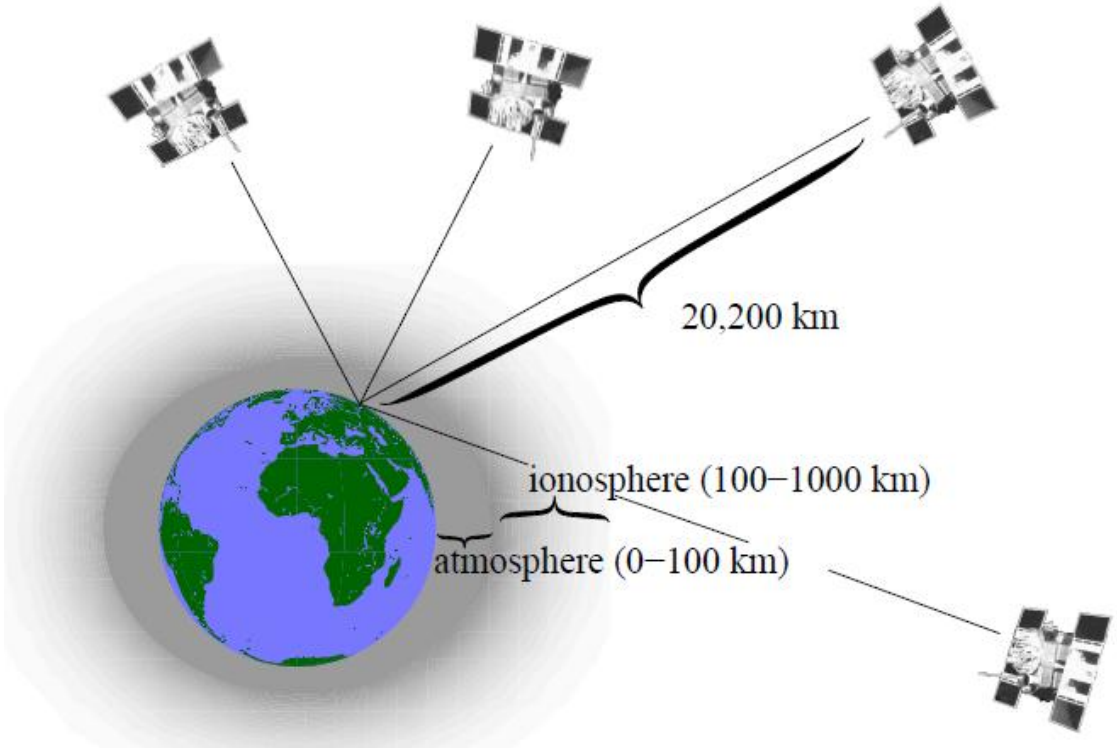

Figure 3-2: GPS Satellites orbiting the Earth. At least four satellites are visible at every point of the globe at any time (De Haan *et al.*, 2008).

Some of the characteristics of the GPS assembly are summarised in Table 3-1 below:

Table 3-1: Characteristics of the GPS Assembly (Braun *et al.*, 2004)

Number of Satellites	28
Number of Orbital Planes	6
Orbital Period	~12 Hrs
Orbital Radius	~26,400 km
Inclination	~55°
Eccentricity	~0

3.1.1 GPS Signals

The GPS signals, as they propagate through the atmosphere from the satellite transmitter to the ground-based receiver are affected by among other things: geomagnetic storms, ionospheric scintillations as well as meteorological factors such as water vapour, temperature and atmospheric pressure. The effect of these is to slow down the phase velocity of the signal due to their refractive and scattering properties. Refraction causes the GPS signal to travel along a curved path rather than a straight line to reach the GPS receiver meaning the signal takes longer to reach the receiver than it would if it were travelling along a straight line. The extent to which these signals are delayed can be evaluated by taking into account the prevailing weather conditions (Bevis *et al*., 1992).

The GPS signals contain both ranging codes and navigation messages. The ranging signals are used to measure the distance between the GPS receivers and the satellites. The navigation signals contain the data used to evaluate the orbital position of each satellite and also contain information about the time and status of the entire satellite constellation.

The navigation message is modulated on top of the ranging codes at 50 bits per second. This navigation message consists of three components, each providing information on:

- the GPS date and time
- orbital information, also called ephemeris data; it enables calculation of satellite position
- the status of all the satellites in the GPS constellation also called the almanac

The almanac guides the GPS receiver in choosing the satellite to search for. Once the receiver picks up the relevant satellites, the ephemeris data is downloaded directly from the satellite to the receiver.

3.1.1.1 Navigation Message Components

The navigation message is a 1,500 bit frame, divided into five subframes each 300 bits long and transmitted at 50 bits per second. Subframe 1 contains the GPS date and information to match the satellite's GPS time to the GPS receiver's time. Subframes 2 and 3 contain the transmitting satellite's ephemeris data while subframes 4 and 5 contain the almanac's components.

Each subframe contains ten words. The first word is the Telemetry Word (TLM) which the receiver uses to detect the start of a new subframe and determine the receiver clock time at which the subframe begins. The following word is the Handover Word (HOW), which the receiver uses to determine the time at which the first bit of the next subframe is to be transmitted and identifies the individual Subframe from within the frame.

3.1.1.2 GPS Signal Frequencies and Modulation Techniques

The ranging codes and navigation messages are modulated onto a carrier signal. The carrier signal, as GPS was originally designed, utilises two frequencies. The frequencies are 1575.42 MHz, called L1 and 1227.60 MHz, called L2. At these frequencies, there is no significant attenuation of GPS signals due to atmospheric conditions, making GPS an all-weather solution. Both carrier frequencies are derived from the fundamental frequency f_0 which is the frequency of the GPS Satellite's internal atomic clock, 10.23 MHz. All other GPS signals are also derived from f_0. The GPS satellite network employs a spread spectrum (whereby the information signal is transmitted over the entire carrier bandwidth) Code Division Multiple Access technique to encode the low bit-rate message data with a high bit-rate pseudorandom sequence that is unique for each satellite (Spilker, 1978).

Two pseudorandom codes are modulated onto the carrier frequencies. The Coarse Acquisition Code (C/A-code) is modulated onto the L1 carrier only, while the Precise Code (P-code) is modulated onto both L1 and L2 carriers. P-code signals can be encrypted so as to allow access to the US military only, and on encryption, the P-code becomes known as the Y-code, though advanced civilian receivers are able to read Y-code too, albeit at reduced accuracy. The C/A-code is accessible to civilians. The characteristics of GPS signals are listed in Table 3-2.

Table 3-2: GPS Signal Characteristics (Braun *et al.*, 2004)

Observation Name	Frequency Multiplier	Frequency	Wavelength	Precision
C/A	$f_0/10$	1.23 MHz	300 m	~10 m
P1 (Y1), P2 (Y2)	1* f_0	10.23 MHz	30 m	~3 m
L1	154* f_0	1575.42 MHz	19.0 cm	~1-2 mm
L2	120* f_0	1227.60 MHz	24.4 cm	~1-2 mm

The results of the measurement of the pseudorandom codes (C/A, P1 and P2) yield the measure of the time of travel for the radio signal from satellite and the receiver. The receiver in turn generates a copy of the pseudorandom code in order to compare it side by side with the pseudorandom code received from the satellite. A function introduced into the receiver correlates automatically the time difference between the pseudorandom code received from the satellite and the one generated by the receiver. This time difference combines the signal travel time from satellite to receiver and the error in synchronisation between the satellite clock and the receiver clock. This measurement is called pseudorange measurement and has precision of 7.8 m in signal travel distance for up to 95% of global users (GPS.GOV).

The GPS receiver can also conduct phase delay measurement, using the GPS carrier frequencies, L1 and L2 by first removing the pseudorandom codes from the carrier phase and combining the received signal with an internally generated signal. Due to the motion of the satellite relative to the receiver, the carrier frequency from satellite transmitter is Doppler shifted as it arrives at the receiver antenna. By integrating the signal output, the distance between the satellite and the receiver are obtained, the accuracy being between 1 to 2 mm. There is an ambiguity in this measurement because the receiver does not have the ability to determine the initial distance between it and the satellite, only being able to track the evolution of the range from the moment it begins tracking the satellite. It is possible to estimate this ambiguity by processing the observation equations to yield the station position, atmospheric delay and other parameters. The estimation of the phase delay due to the atmospheric parameters is the basis for GPS Meteorology.

3.1.2 Physics of Atmospheric Phase Delay of Radio Signals and its Application to GPS Meteorology

The theory behind GPS Meteorology is well explained in the literature of Bevis *et al.* (1992), from where the material in this section is lifted. Electromagnetic waves propagating through the atmosphere travel slower than they would in a vacuum. In addition, the waves travel along a curvilinear path rather than in a straight line. These effects are due to the variable refractive index of the medium of propagation, i.e. the atmosphere. The radio signal time delay, which is the difference between the travel time of the signal if it were travelling along a straight line in a vacuum and the actual travel time, can be expressed in terms of the excess travel path length as given in equation (3-1) below, which is valid for rays along the zenith path, and approximated to within 1 cm for paths with elevations greater than 15°.

$$\Delta L = \int_L (n(s) - 1) ds \tag{3-1}$$

where $n(s)$ is the refractive index as a function of position s along the ray path L through the atmosphere. The atmospheric refractive index is related to the atmospheric refractivity N, as

$$N = 10^6 (n - 1) \tag{3-2}$$

For measurements on the earth's surface along a given zenith angle z, Equation (3-1) approximately becomes

$$\Delta L = \frac{10^{-6} N}{cosz} \tag{3-3}$$

The atmospheric refractivity is not vertically homogenous, rather it varies with height. If the vertical refractivity profile is known, then the path delay due to any arbitrary section of the refractive atmosphere can be calculated as

$$\Delta L = \int_{r_0}^{r} \frac{10^{-6} N(r)}{cosz} dr \tag{3-4}$$

where r is the geocentric distance to a point in the atmosphere and r_0 is the geocentric position of the observer. This refractivity is a function of the atmospheric temperature, pressure and water vapour and can also be approximately evaluated using Equation (3-5) (Smith and Weintraub, 1953).

$$N = 77.6\left(\frac{P}{T}\right) + 3.732 \times 10^5 \left(\frac{P_w}{T^2}\right) \tag{3-5}$$

where P is the total atmospheric pressure in hPa, P_w is the partial pressure of water vapour and T is the atmospheric temperature in Kelvin. The first term in Equation (3-5) is the hydrostatic refractivity while the second term is the wet refractivity.

Thayer (1974) provides a more accurate formula for refractivity, given in equation (3-6).

$$N = k_1\left(\frac{P_d}{T}\right)Z_d^{-1} + k_2\left(\frac{P_w}{T}\right)Z_v^{-1} + k_3\left(\frac{P_w}{T^2}\right)Z_w^{-1} \tag{3-6}$$

where $k_1 = 77.604 \pm 0.014\text{K } hPa^{-1}$;

$k_2 = 64.79 \pm 0.08\ hPa^{-1}$

$k_3 = 3.7776 \pm 0.004\ K^2 hPa^{-1}$

P_d is the partial pressure of dry air in hPa.

Z_d^{-1} and Z_w^{-1} are the inverse compressibility factors for dry air and water vapour respectively correct for non ideal gas behaviour, but are almost always equal to 1. Due to uncertainty in the values of the constants k_1, k_2 and k_3, the accuracy of the computed refractivity is limited to 0.02%.

The total atmospheric delay comprises two components, one part which depends on surface pressure only (known as Hydrostatic Delay) and a lesser part which is a function of atmospheric water vapour distribution (Saastamoinen, 1972). This lesser part is known as the Wet Delay, because it is produced by Water Vapour alone. For a signal that is approaching from vertically upwards, the signal delay due to the troposphere is on the zenith path and is therefore called the Zenith Tropospheric Delay (ZTD). ZTD is a sum of the hydrostatic and

wet delays along the zenith angle, referred to as the Zenith Hydrostatic Delay (ZHD) and the Zenith Wet Delay (ZWD) respectively as shown in equation (3-7).

$$ZTD = ZHD + ZWD \tag{3-7}$$

ZHD and ZTD are related to the hydrostatic and wet delays for the GPS signal along any elevation angle by a mapping function which is a mathematical model for the elevation dependence of the signal delay such that for a signal of elevation angle θ:

$$\begin{aligned} Hydrostatic\ delay &= ZHD \times M_h(\theta) \\ Wet\ Delay &= ZWD \times M_w(\theta) \end{aligned} \tag{3-8}$$

where $M_h(\theta)$ is the hydrostatic mapping function and $M_w(\theta)$ is the wet mapping function. $M_h(\theta)$ and $M_w(\theta)$ are generally similar at elevation angles above 10° to 15° (Tralli and Lichten, 1990).

ZHD can usually be evaluated using ground level conditions according to the relation below derived by Saastamoinen (1973).

$$ZHD = \frac{0.0022767 \times P_s}{1 - 0.00266 \times cos2\varphi - 0.00028 \times H} \tag{3-9}$$

where P_s is the ground level atmospheric pressure in hPa;

φ is the ellipsoidal latitude of the station;

H is the station altitude above the ellipsoid in km.

The ZWD can then be easily estimated by simply subtracting ZHD from ZTD. More accurately though, ZWD can also be evaluated using the radiosonde-acquired values of *T* and P_w as below (Davis *et al*., 1985):

$$ZWD = 10^{-6}\left[k'_2 \int \left(\frac{P_w}{T}\right) dz + k_3 \left(\frac{P_w}{T^2}\right) dz\right] \tag{3-10}$$

where $k'_2 = 17 \pm 10\ K^2 hPa^{-1}$and the integral is along the zenith path.

The ZWD is related to the Integrated Water Vapour along a zenith path by Equation (3-11) (Askne and Nordius, 1987):

$$IWV = \kappa \cdot ZWD \tag{3-11}$$

where $\frac{1}{\kappa} = 10^{-6}\left(\frac{k_3}{T_m} + k'_2\right) R_w$

where $T_m = \frac{\int\left(\frac{P_w}{T}\right)dz}{\int\left(\frac{P_w}{T^2}\right)dz}$ and is the weighted mean temperature of the atmosphere;

R_w is the specific gas constant for water vapour.

The ionosphere extends from approximately 50 km to 1000 km altitude. At frequencies above 30 MHz, radio waves passing through the ionosphere undergo dispersion due to the free electrons in the ionosphere. Consequently, the phase velocities of the carrier waves are increased while the group velocities of the modulations superposed on the carrier are reduced. The magnitude of the phase advance is equal but opposite in sign to the group delay. This magnitude is approximately proportional to the total electron content (TEC) along the ray path and inversely proportional to square of the carrier frequency. If measurements are made at two widely spaced frequencies, it is possible to model the ionospheric effects and isolate them from the neutral atmospheric phase delay (Vorobev and Krassil'nikova, 1994).

Taking *L1* and *L2* to be the two transmission frequencies, a simple linear correction for the time delay for *L1* can be expressed as

$$Ionospheric\ TD_{L1} = 1.5336\ \times \Delta T \tag{3-12}$$

where $Ionospheric\ TD_{L1}$ is the ionospheric delay on *L1* and ΔT is the measured difference in the total time delay between *L1* and *L2*. These computations are discussed in the literature of Bevis *et al.* (1992).

3.2 Upper-Air Sounding Systems (Radiosonde)

The upper-air sounding system consists of a flight subsystem (comprising an inflated balloon, flight-train and radiosonde systems) and a ground-based subsystem (comprising the tracking,

receiving and signal-processing and data-processing equipment). The balloon carries the radiosonde to desired altitudes at a desired ascension rate. The flight-train (which may include a combination of parachute, train regulator, lights or radar reflector) interfaces the radiosonde and the balloon and is designed to help in the launching, flight and descent of the radiosonde. The radiosonde equipment is composed of the meteorological sensors, the data encoding electronics and the telemetry transmitter as shown in Figure 3-3 (Federal Meteorological Handbook 3, 1997). This radiosonde takes in-flight measurements and transmits these to the ground-based subsystem. The measurements taken by radiosonde are the atmospheric state properties of pressure, temperature and relative humidity of the atmospheric constituents.

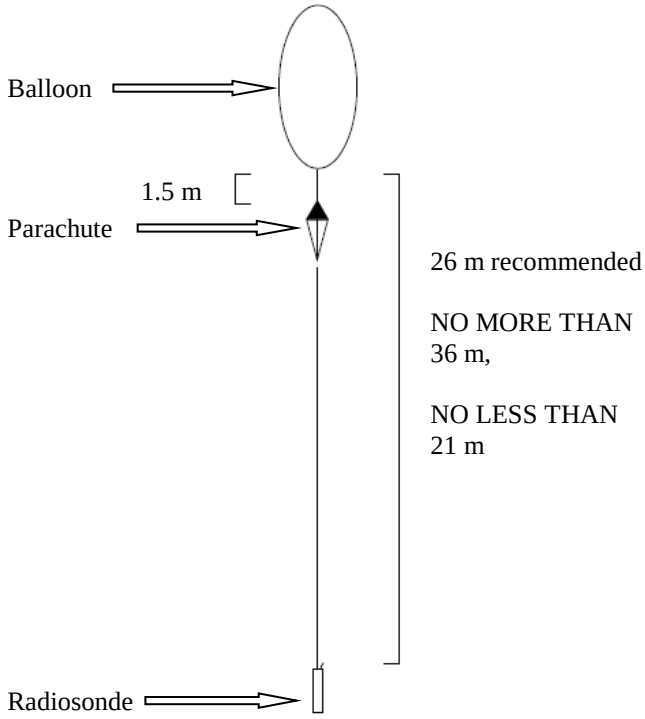

Figure 3-3: Radiosonde Flight Subsystem Assembly (FMH 3, 1997)

The ground subsystem performs the tracking, receiving, signal processing and data processing functions. The various signals transmitted from the flight subsystem are received, filtered and demodulated by the receiver. The signal processing and data processing units convert these demodulated signals into values of pressure, temperature and humidity at the various altitudes (Dubois *et al*., 2002).

3.3 Artificial Neural Networks (ANN)

ANN is a supervised learning technique originally developed to simulate the human neural system. In analogy to the biological brain, an ANN, which is a network of simple processing elements (called neurons) that exhibit complex collective behaviour determined by the connections between the processing elements, can be trained to find solutions, recognise patterns and predict a system's outputs with respect to its inputs. A general overview of the technique of neural networks can be found in the publications by Bose and Liang (1998) and Hagan and Demuth (1996).

Figure 3-4 is a simplified representation of the processing element of a neural network with a single input *p*, weight *w*, bias *b* and transfer function *f* which could be linear or nonlinear:

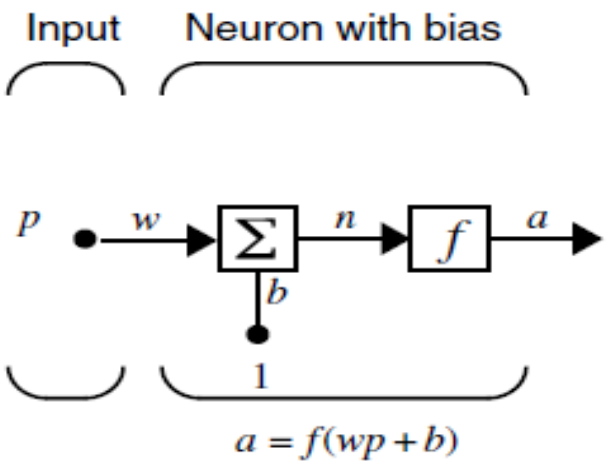

Figure 3-4: Single-Input Neuron (Hagan and Demuth, 1996)

The values of *w* and *b* are adjustable and scalar while the function *f* is chosen by the neural network designer. The scalar input *p* is multiplied by the scalar weight *w* to form *wp* and the product sent to the collator. The other input into the collator is the bias *b*. The collator's output n is then sent to a transfer function which operates on it to yield the neuron output *a*. This is calculated as:

$$a = f(wp + b) \tag{3-13}$$

The neural network output depends on the chosen transfer function and the weight and bias whose values are adjusted by a learning rule so that the neuron's input/output relationships

meets a specific goal. Neural networks are trained in order that an input leads to a specific targeted output. This concept is illustrated in Figure 3-5 (Beale *et al.*, 2009):

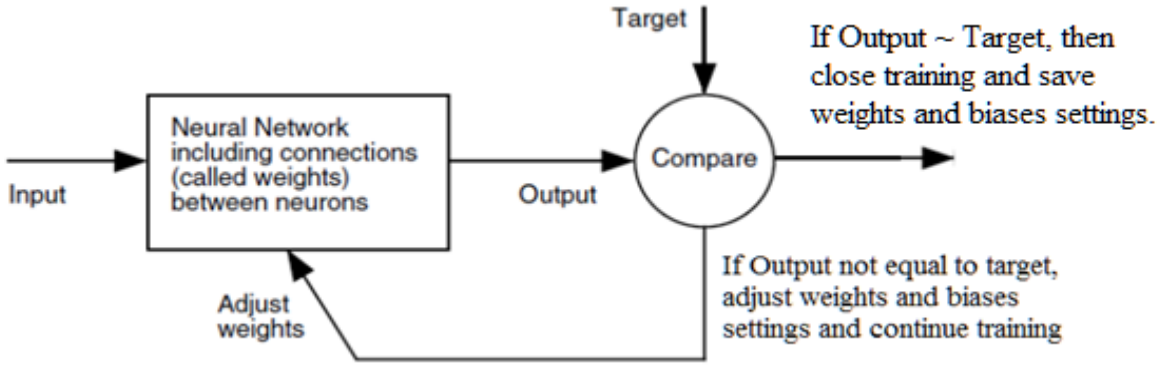

Figure 3-5: Illustration of Neural Network Training (Beale *et al.*, 2009)

During training for supervised learning, the input is presented to the input layer, the output being as yet undefined. The output of the system is then compared with the result from experimental observation. The observed difference between the network's output and the expected result target is used to determine an error function which is propagated back to update the weights and biases using an optimization technique which strives to minimize the error. The propagation of the error function back to the input of the neuron is called backpropagation. The entire procedure is repeated for a number of epochs until the desired output accuracy is achieved.

The training setup for the neural network used in this study was as illustrated in Figure 3-6:

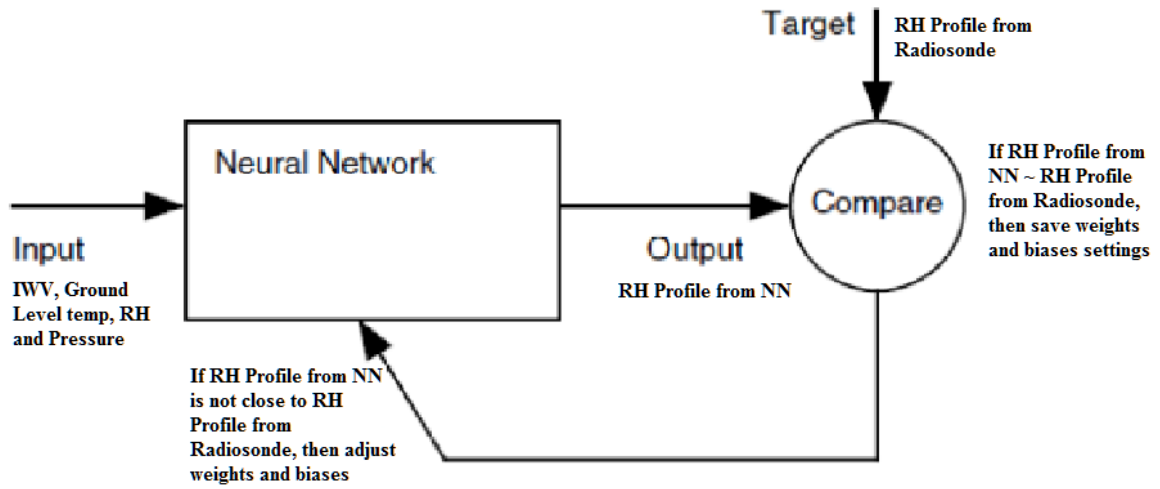

Figure 3-6: Neural Network Training setup for the study

During validation, the model output variables are directly calculated, without iteration, from the input variables and the trained weights and biases.

In typical neural network applications, it is usually necessary to use more than one processing element layer. The output of one processing element layer serves as the input to the subsequent layer. A neural network should comprise an input layer (with as many inputs as there are variables) an output layer and at least one hidden layer with full connections between neighbouring layers.

For this study, two transfer functions were employed in the neural networks, the linear transfer function and the hyperbolic tangent sigmoid transfer function. These two functions were chosen because previous research had proven that networks with sigmoid transfer functions in the hidden layers and linear transfer functions in the output layer could be used to accurately approximate any function or relation between inputs and desired outputs (Hornick *et al.*, 1989). For the linear transfer function, the output *a* of the function is equal to the input *n*. as shown in Equation (3-14). Neurons with linear transfer functions are very useful in applications for linear approximation (Beale *et al.*, 2009).

$$a = n \tag{3-14}$$

The hyperbolic tangent sigmoid transfer function takes an input of any value between minus infinity and positive infinity and compresses the output into the range of between -1 to 1 according to the expression given in Equation (3-15) below (Hagan and Demuth, 1996).

$$a = \frac{e^{n} - e^{-n}}{e^{n} + e^{-n}} \tag{3-15}$$

The sigmoid transfer functions are useful in multilayer backpropagation networks because they are differentiable.

The technique of ANN has powerful capabilities in approximating nonlinear relationships and correlations with arbitrary complexity and can be a very useful tool for modelling complex relationships between inputs and outputs (Hammerstrom, 1993). In this work, ANN is to be used to correlate the IWV evaluated using GPS phase delay with the corresponding RH profiles.

CHAPTER 4. METHODOLOGY

4.1 Experimental Setup

4.1.1 GPS Setup at RCMRD

A GPS receiver at the Regional Centre for Mapping of Resources for Development (RCMRD), (1° 13′ S, 36° 54′E, altitude 1605 m) receives signals from the GPS constellation. The receiver model is LEICA GRX1200GGPRO with elevation cut off setting of 5 degrees. The antenna type is LEIAT504GG. Figure 4-1 and Figure 4-2 are pictures of the equipment setup at RCMRD. The observable characteristics of the signals received and their time evolution from the satellites are routinely recorded and uploaded to the IGS global archive (Dow *et al.*, 2009).

Figure 4-1: GPS Antenna at RCMRD, the GPS Receiver is housed in RCMRD Office in the background.

Figure 4-2: GPS Receiver Setup at RCMRD including receiver on the left and computer for data processing.

4.1.2 Radiosonde Setup at KMD

Radiosondes are launched on a daily basis, usually around 45 minutes before 0000 Coordinated Universal Time (UTC) and 45 minutes before 1200 UTC. Radiosonde launches at all the upper air sounding stations in the world are synchronised for simultaneous launching. This synchronised launching enables a snapshot of the global state of the atmosphere at 0000 UTC and 1200 UTC which are the official observation times. At the Kenya Meteorological Department (KMD) Upper Air Observatory in Dagoretti Corner (1°20′S, 36°46′ E, altitude 1795 m), radiosondes are launched on a daily basis for the 0000 observation time and occasionally for the 1200 observation time. The radiosondes used in this study for temperature, pressure and RH profiling are the Vaisala Radiosonde RS92 model with accuracy of 0.5 °C for temperature, 5% for relative humidity and 1 hPa for pressure (Nash *et al.*, 2011). Figure 4-3 is a picture of the Vaisala radiosonde RS92.

Figure 4-3: Picture of the Vaisala Radiosonde RS92

The setup for the radiosonde equipment as configured for this project was as in Figure 4-4.

After radiosonde launch the sensors attached to the balloon send the sensed data via radio to the Telemetry Receiving System which uses the signal processing system to interpret the signals and translate them into readable data which is onward-transmitted to the computer for archiving.

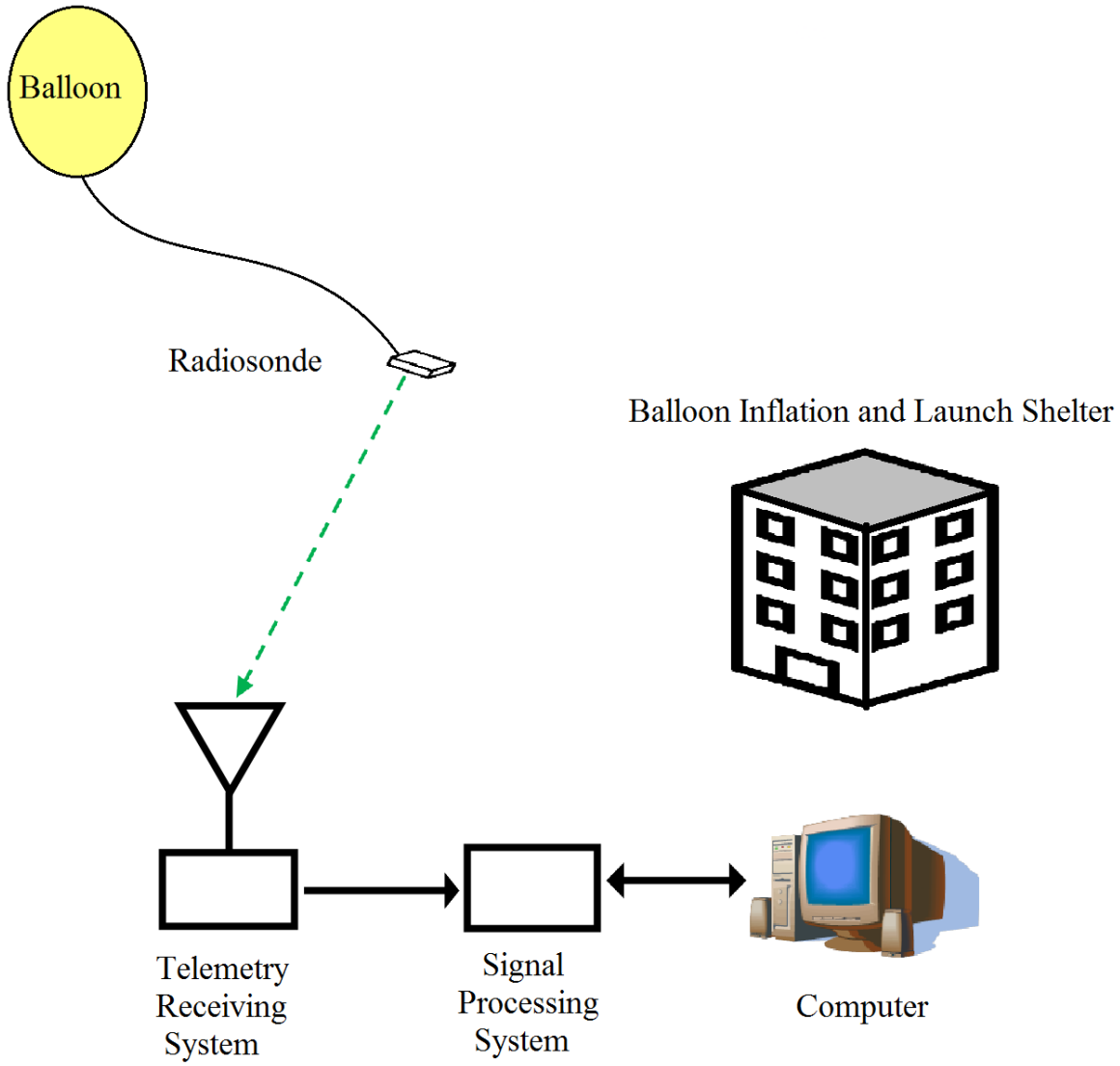

Figure 4-4: Radiosonde Setup at KMD (NOAA, 2013)

4.2 Data Acquisition

The radiosonde and GPS data, collected over three years (1 Jan 2009 to 31 Dec 2011) using the equipment setup as outlined in Section 4.1 above, were used in this work. The days from which data was selected for both training and validation of the system had to satisfy the following conditions:

1. The radiosonde data had to be available for that day at 2200 UTC to 0059 UTC.
2. The radiosonde had to have recorded soundings for temperature, pressure and RH from ground level up to at least 20,000 m.
3. GPS observables data had to be available for the same day, 2200 UTC to 0059 UTC.

4. GPS navigation data had to be available for the same day, 2200 UTC to 0059 UTC.

These conditions were set in order to ensure that for every radiosonde data set used in the study, there was a corresponding GPS data set. This was important as both training and validation of the system required time-synchronised GPS and radiosonde data sets.

Most of the days of the months of August to December for all the three years did not satisfy the conditions set above and therefore data sets for these months in the years 2009 and 2010 were not used in training the system. However, the available data sets for these months for 2011 were separately used in validating the system to check the system performance on data sets from the months not included in the system calibration.

4.2.1 Meteorological Data Acquisition and Pre-processing

The radiosonde data recorded from 1 Jan 2009 to 31 Dec 2011 was acquired from the Kenya Meteorological Department. This data contains for each radiosonde launch the information on:

- Time of radiosonde launch
- Elapsed time since radiosonde launch in minutes and seconds (miss)
- Atmospheric pressure at the respective altitudes in hundredths of hPa (0.01 hPa)
- Corresponding geopotential height in whole geopotential meters.
- Temperature (dry bulb) to the nearest 0.1 Celsius degree.
- Relative humidity to the nearest 0.1 %
- Dew point depression to the nearest 0.1 Celsius degree.
- Wind direction to the nearest whole degree
- Wind speed to the nearest 0.1 m/s

The data of interest was the atmospheric pressure, temperature and relative humidity at the respective geopotential heights.

4.2.1.1 IWV Evaluation from Radiosonde Data

The radiosonde data included the temperature, pressure and RH profiles. From the RH profile, the water vapour pressure profile was approximated using Equation (4-1) (ITU-R, 2012):

$$P_w(h) = H(h) \times \frac{E \cdot a \cdot \exp\left[\frac{\left(b-\frac{t}{d}\right)\cdot t}{t+c}\right]}{100} \tag{4-1}$$

where P_w *(h)* is the water vapour pressure in hPa at altitude *h*, *t* is the Celsius temperature, *H(h)* is the relative humidity at altitude *h* and the coefficients *a*, *b*, *c* and *d* have the values:

a = 6.1121; *b* = 18.678; *c* = 257.14and d = 234.5 for water (valid between 0°C and +50°C)

and

a = 6.1115; *b* = 23.036; *c* = 279.82 and d = 333.7 for ice (valid between -80°C and 0°C).

with the coefficient *E* defined for both water and ice as:

$$E_{water} = 1 + 10^{-4}[7.2 + P_d \cdot (0.00320 + 5.9 \cdot 10^{-7} \cdot t^2)]$$

$$E_{ice} = 1 + 10^{-4}[2.2 + P_d \cdot (0.00320 + 6.4 \cdot 10^{-7} \cdot t^2)]$$

With the P_w profile now determined, the IWV above the radiosonde site was determined using Equation (4-2) below with 1795 m being the altitude of the radiosonde launch station and 20,000 m being the presumed maximum height of the Tropopause.

$$IWV_{KMD} = \frac{100}{R_w} \int_{1795\ m}^{20{,}000\ m} \frac{P_w(h)}{T(h)} \tag{4-2}$$

where $P_w(h)$ is the water vapour pressure at altitude *h*,

$T(h)$ is the temperature in Kelvin at altitude *h* and

R_w is the specific gas constant for water vapour.

This IWV was evaluated for each day for which Radiosonde and GPS data were available from which a monthly average of monthly IWV values are presented on Figure 5.1.

4.2.1.2 Relative Humidity Profile from Radiosonde Data

The radiosonde system's output was the meteorological data which included the RH profile from ground level up to radiosonde burst height which for the three years was on average between 33 km and 34 km. Also included in the RH profile data was the RH values at the standard pressure levels. For this study, six standard pressure levels were considered: 700 hPa, 500 hPa, 300 hPa, 200 hPa, 150 hPa and 100 hPa. The RH at these standard pressure levels was recorded for each day that both Radiosonde and GPS data were available.

4.2.2 GPS Data Acquisition and Processing

GPS data recorded from 1 Jan 2009 to 31 Dec 2011 using the GPS receiver at RCMRD, Kasarani was used in this study. The GPS data, in RINEX format contained information on:

- Positions of the GPS transmitting and receiving antennas
- Time of signal transmission from satellite-born GPS transmitter and reception at ground based GPS receiver.

Using GAMIT software, the GPS signal phase delay which is the difference between the actual time taken by GPS signals to travel from satellite transmitter to the ground-based receiver and the time the signals would have taken to travel based on the straight-line distance between transmitter and receiver was determined. GAMIT software also determined and eliminated the delay due to the ionosphere using the dual frequency technique described in Chapter 3 (Dach *et al*., 2011). The GAMIT output was the Precipitable Water Vapour (PWV) which was converted into IWV by multiplying PWV with the density of water.

The procedure for determining IWV from GPS data using GAMIT was as in Figure 4-5:

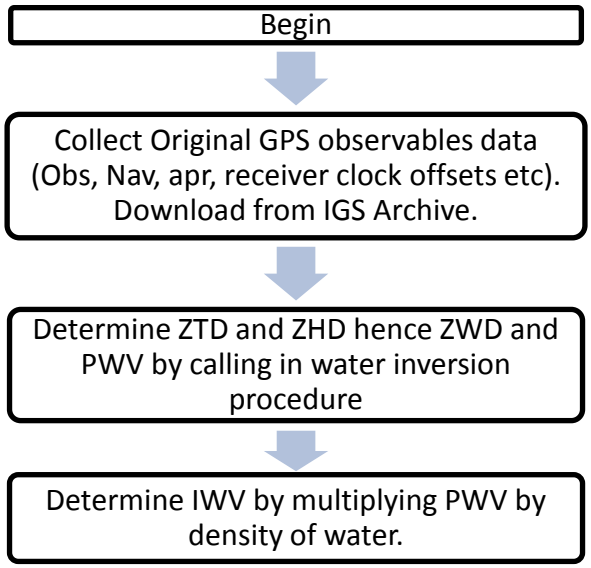

Figure 4-5: Procedure for Evaluating IWV from GPS data.

4.3 Relating IWV Measurements from Radiosonde and GPS

The next step was to compare side by side the IWV derived from radiosonde data and the IWV obtained from GPS data. Since the GPS site (RCMRD) and the radiosonde station (KMD) were a different locations, with the GPS site being 190 m lower in altitude than the radiosonde launch site, it was necessary to find a way to relate the readings from these two sites. One method was to extrapolate the meteorological data at the radiosonde site altitude so as to yield the temperature, pressure and humidity at the GPS site altitude. With these, then the IWV could be evaluated and compared to the IWV obtained from GPS. The alternative method was to simply do a comparison between the IWV at KMD derived from radiosonde and the IWV at RCMRD obtained from GPS. Once this was done, the next step was to investigate if the relation between the IWV at KMD and IWV at RCMRD exhibited any pattern or trend and if this trend was reasonably consistent through the period of the experiment. If the trend showed consistency, then it would be reasonable to assume that there was a direct correlation between the temporal evolution in the value for IWV at RCMRD as evaluated form GPS data and the temporal evolution of RH profile as observed via radiosonde at KMD. This was to say that the values for RH at the six standard pressure levels starting from 700 hPa up to 100 hPa would be similar at KMD and RCMRD, which is not unexpected seeing that both stations fall within the same climatic region and are within 20 km of each other. The IWV data evaluated from GPS data recorded at RCMRD could

therefore be extrapolated in order to estimate the IWV at KMD, and this value, in conjunction with the ground level temperature, pressure and RH and also the RH profiles be used to train the neural network system and also to validate the system.

4.4 Artificial Neural Network (ANN) Development and Configuration

A neural network system was created, the predictors being the temperature, pressure and RH at ground level, and also the IWV. The output of this system was the RH profile, values for RH specified at six standard pressure levels: 700 hPa, 500 hPa, 300 hPa, 200 hPa, 150 hPa and 100 hPa. The manner in which the predictors influence the output RH profile pattern is analysed in section 5.3.

The neural network to retrieve the relative humidity profile was designed using the Matlab Neural Network Toolbox. This neural network was of the feed-forward type, having an input layer with four neurons, one intermediate layer of six neurons with tan-sigmoid transfer functions, a further two intermediate layers of 24 neurons each with tan-sigmoid transfer functions, a fourth intermediate layer of six neurons with linear transfer functions and an output layer of six neurons with linear transfer functions. The linear transfer functions were used in the input and output layers for linear function fitting and the tan-sigmoid transfer functions were used in the intermediate neurons for pattern recognition and to compute the non-linear relations the inputs and the outputs. The inputs for the neural network were the values for atmospheric temperature, pressure, RH and IWV as measured at ground level. The output of this neural network was the RH profile specified at six standard pressure levels as mentioned in Section 4.4 above.

The number of neurons at the input layer was set to be same as number of input variables and number of neurons at output layer to be equal to number of pressure levels, hence four input neurons and six output neurons. The network was initially set to have one intermediate layer with one neuron and then the modelling process was carried out and error analysis performed. The number of neurons in the intermediate layer was increased by one, and the modelling process carried out again, and error analysis also computed to check if there was improvement. At the low number of intermediate layer neurons, the network's accuracy was found to be extremely low with almost zero correlation detectable between the input weather parameters and the output RH profiles.

One neuron after another was added to the single intermediate layer without much improvement in network performance getting observed until the number neurons in the intermediate layer reached fifty. The decision was then made to include a second intermediate layer and the number of neurons in the two intermediate layers incremented step by step from one neuron each until 24, where some improvement in the network performance was observed. The best results were realised upon adding a four-neuron intermediate layer just after the input layer and another six-neuron intermediate layer just before the output layer. The introduction of the second 24-neuron intermediate layer and the other four-neuron and six-neuron layers next to the input and output layers respectively is supported in the literature of Bengio and LeCun (2007) where they propose to solve the problem of representing a highly-varying function without having to use an extremely large number of neurons, as would have been necessary in a single intermediate layer network, by using multiple non-linearities in the network. This kind of network is called a deep neural network.

The advantage of a deep neural network architecture like the one realised here is its ability to model complex relations by representing a much larger number of functions and relations by making use of the additional layers (Szegedy *et al.*, 2013). The main disadvantage of a deep neural network is its propensity for overfitting and exposure to noise because the additional layers could model undesirable input-output relations in the training data sets. This was overcome using ideas borrowed from the greedy layer-wise training method (Bengio et al., 2007). This technique was proposed by Hinton *et al.* (2006) to help minimise overfitting in a deep neural network by training the network one layer at a time. For the application in this study, once the final network architecture had been implemented, the network was trained in steps with layers incremented after each training round. Firstly, the network comprising only the first intermediate layer was trained, then the second intermediate layer was added and training done again, then with three and finally four intermediate layers. The resultant weights from the step-incremental training of the layers were then set to be the weights of the final deep network with all the layers. A final training was then done with all four intermediate layers which served the purpose of fine tuning the network.

The eventual architecture for this system was as in Figure 4-6:

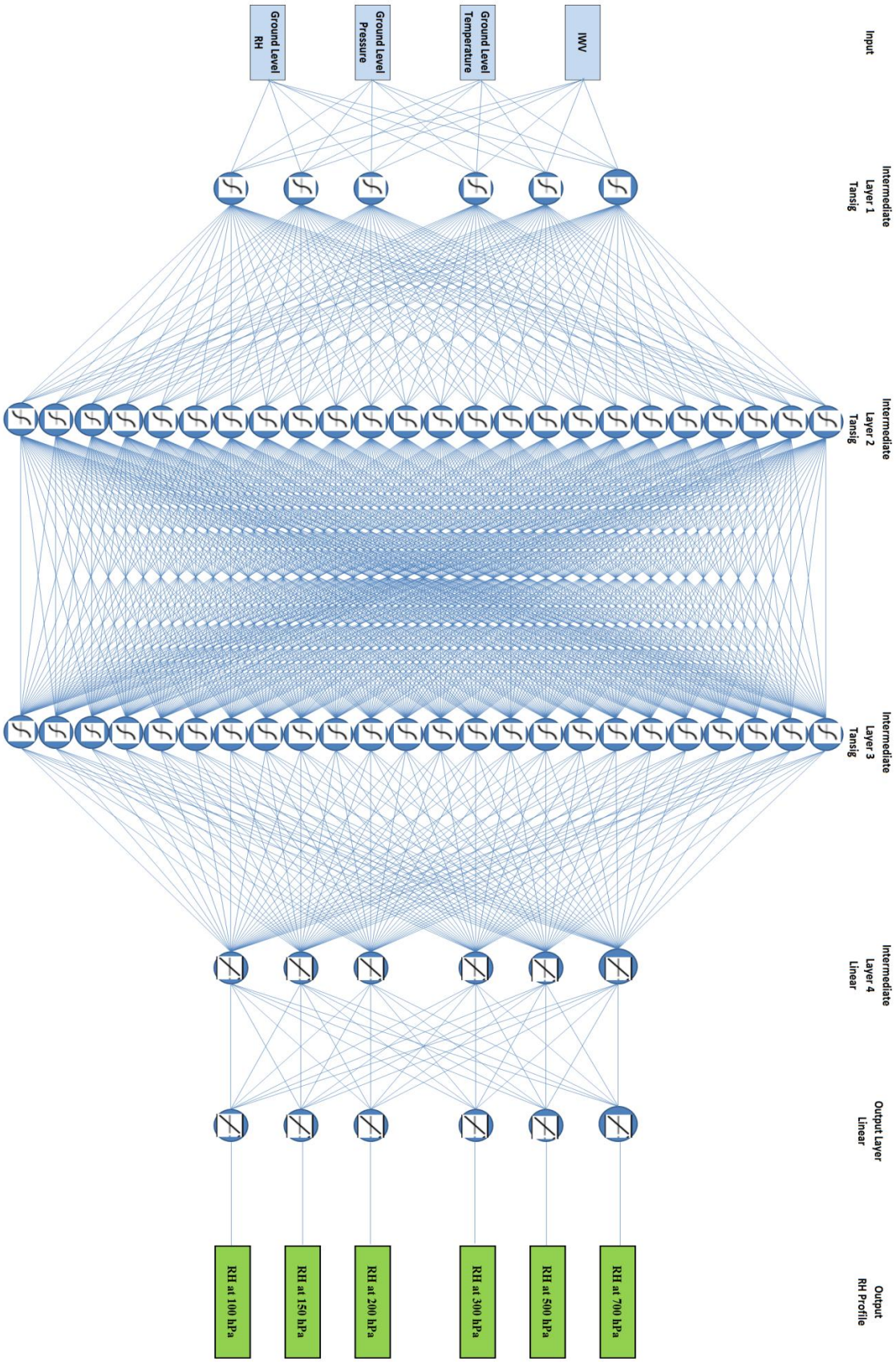

Figure 4-6: Architecture of Neural Network to determine RH Profile

The neural network was trained using the Levenberg-Marquardt optimisation algorithm, which was preferred rather than the standard gradient descent backpropagation algorithm in order to save on training time. The ground level temperature, pressure, RH and IWV for the years 2009 and 2010 and the RH profile data for the same years were used to train the system, the ground level temperature, pressure, RH and IWV being the inputs and the RH profile being the target. One training pattern representing each day for which GPS and radiosonde RH profile data was available was used in the training. In total 235 patterns were used in the training.

The neural network was expected to recognise two things:

i. that the tropospheric RH profile should take a C-shape as predicted in the literature of Folkins *et al.* (2002) with highest RH values at the low altitudes. Lowest RH values at mid altitudes and higher RH values towards the Tropopause.
ii. that ground level conditions should have an influence on form of the C-shape mentioned in (i) above, and the influence should be quantifiable.

The neural network was then expected to take advantage of these relations and model itself so that its output mimics the relationship between the RH profile and the ground level conditions, taking into account the expected C-shaped profile.

Generalisation was achieved by employing the technique of 'early stopping': having a high number of set training iterations but small number of maximum allowed epochs without improvement in performance. To avoid memorisation, a value for *max_fail epochs in a row* was set to 5. The training algorithm automatically selects part of the training sample and uses them for validation and testing during training. If for five consecutive iterations, there is no improvement in performance, then training is automatically stopped at the fifth consecutive non-improving iteration. The initial configuration contemplated 1000 iterations. But by the end of 1000 iterations, it was observed that for the final architecture during the fine tuning stage, 4730 iterations achieved generalisation.

4.5 Determination of Relative Humidity Profiles using Neural Network

The system was then validated using the GPS and radiosonde data for the year 2011. The input into the system was the ground level temperature, pressure, RH and IWV for the year 2011. Simulation of the network was then done to see what the output RH profile for the six standard pressure levels would be.

The system's output values for RH profile for the year 2011were compared with the directly observed values for RH profile from radiosonde for the year 2011 in order to determine the accuracy of the system's output as compared to radiosonde data. This would guide as to whether the model provided satisfactory predictions of RH profile if ground level conditions (temperature, pressure, RH and IWV) were known. This would determine whether this model could be considered a valid RH profiling tool and whether it ought to be adopted for weather monitoring.

4.6 Algorithm for RH Profile Determination using GPS Data

The algorithm that was used in this work to determine the RH profile from GPS data and ground level weather conditions is described in this section.

Running the *sh_gamit* script on GAMIT instructs the GAMIT software to fetch the GPS data from the GPS observables data stored in RINEX format in the processing computer's local system memory for the specified days and process the GPS observables to yield the GPS baseline information and also the time delay information for each day using the Equations (3-1) to (3-4) (Herring *et al.*, 2010). The ZTD information for each day is then stored in a ZTD file generated as part of the output of each *sh_gamit* run. Then upon running the *sh_metutil* script on GAMIT, the GAMIT software calls the ZTD file and from it generates the ZHD using the previously mentioned Equation (3-8). Ground level atmospheric pressure information is needed for this step. In the absence of available ground level atmospheric pressure information, GAMIT makes use of the Global Pressure and Temperature (GPT) model developed by Boehm and Schuh (2006) which generates surface pressure and temperature values as a function of latitude and season based on a spherical harmonic fit taking into account twenty years of global meteorological data (Herring *et al.*, 2010). The script, *sh_metutil* generates also the ZWD and consequently the PWV data for each day

which are then stored in an output *met_[site]* file making use of the Global Mapping Functions (GMF) derived by Boehm *et al.* (2006) which include longitude dependence in addition to the latitude and time-of-year dependence used in the older mapping functions. The PWV received as output are converted to IWV using the equation (3-11) and transmitted onwards as input to the neural network.

The neural network will then take as input the IWV, and also the ground level temperature, pressure and RH, and operate on these to yield as output the RH profile. A flow chart algorithm for the procedure to obtain RH profile from GPS observables as input via GAMIT and ANN is outlined in Figure 4-7:

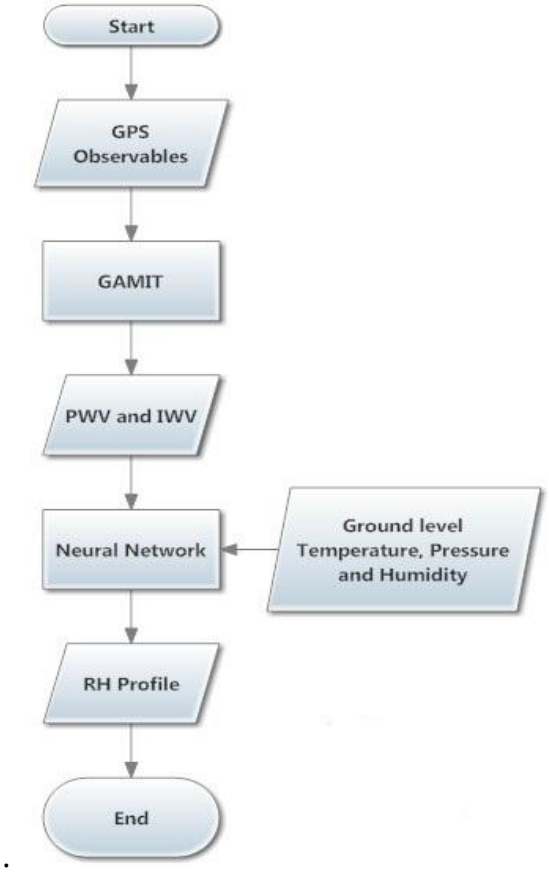

Figure 4-7: Flow Chart Algorithm for RH profile determination from GPS data

Before the neural network was used to determine RH profile using the inputs (IWV and ground level temperature, pressure and RH), it needed to be trained. The training was conducted using the procedure described in Section 4.4.2. Figure 4-8 the flow chart algorithm describing the training and operation of the ANN for prediction of RH profile using as input the ground level IWV, temperature, pressure and RH.

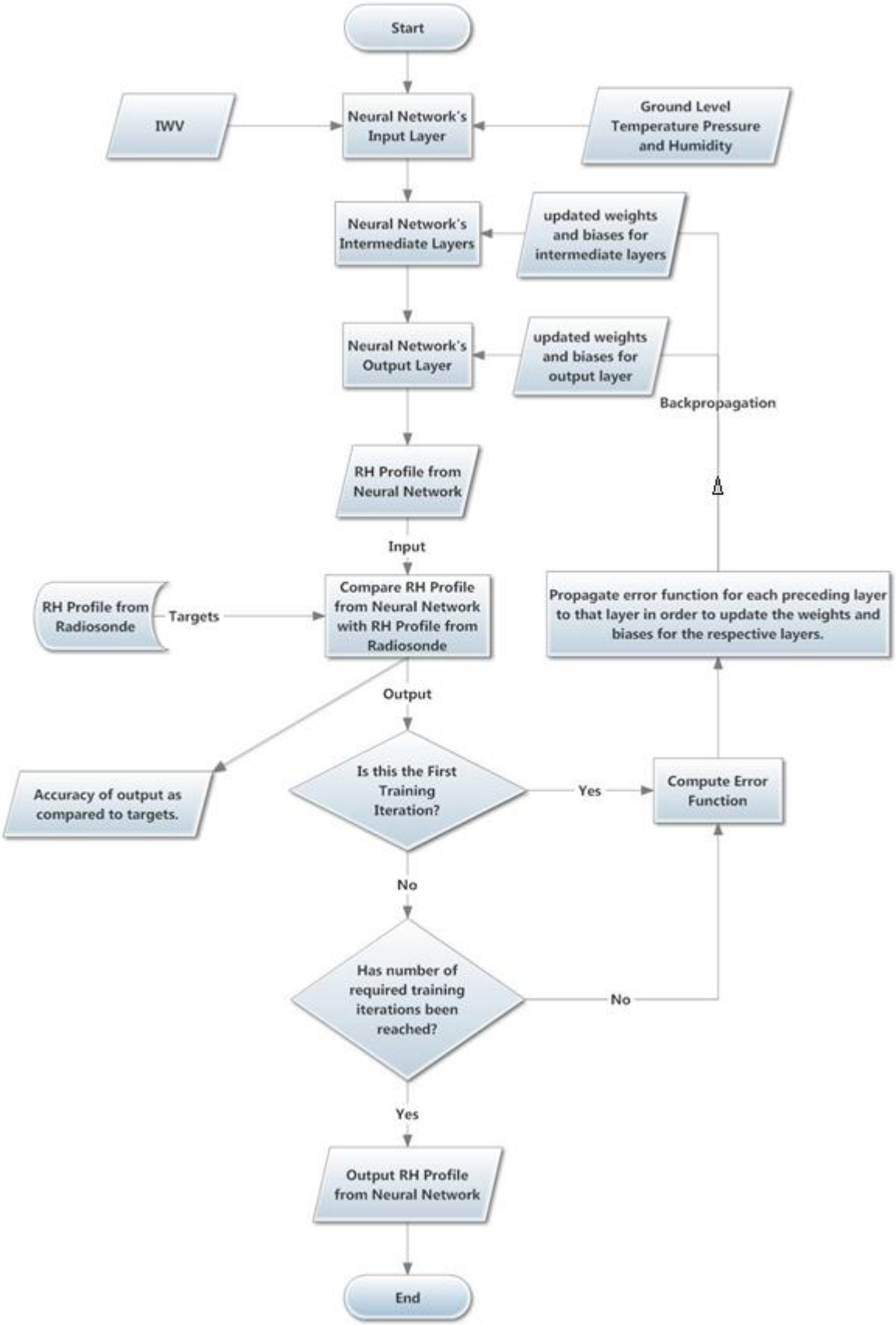

Figure 4-8: Flow Chart Algorithm for training and operation of neural network for determination of RH Profile from IWV

The created ANN consisted of an input layer of four neurons with linear transfer functions, an intermediate layer of six neurons with tan-sigmoid transfer functions, two intermediate layers of 24 neurons each with tan-sigmoid transfer functions, an intermediate layer of six neurons with linear transfer functions and finally an output layer of six neurons with linear transfer functions. Each of the four neurons at the input layer denoted one of the four input parameters: IWV and ground level temperature, pressure and RH; while each of the six neurons at the output layer denoted one of the six standard pressure levels at which the RH profile is specified.

Each layer has an input vector **p**, a transfer function **f**, a weight matrix **W**, a bias vector **b**, and an output vector **a**. When denoting the ANN in equation form, for each of the layers, the number of the layer is given as a superscript to the variable whether **f**, **a**, **W** or **b**. The weight matrices are specified whether **IW** for the input layer or **LW** for subsequent layers. In addition, the weight matrix is labelled with a two-index superscript with the first index denoting its destination and the second index denoting its source. Using the standard equation for a neuron in equation (3-3), for the six layers therefore the equation notation was as below:

$$\begin{aligned}
\mathbf{a^1} &= \mathbf{f^1}(\mathbf{IW^{1,1}p} + \boldsymbol{b^1}) \\
\mathbf{a^2} &= \mathbf{f^2}(\mathbf{LW^{2,1}}\boldsymbol{a^1} + \boldsymbol{b^2}) \\
\mathbf{a^3} &= \mathbf{f^3}(\mathbf{LW^{3,2}}\boldsymbol{a^2} + \boldsymbol{b^3}) \\
\mathbf{a^4} &= \mathbf{f^4}(\mathbf{LW^{4,3}}\boldsymbol{a^3} + \boldsymbol{b^4}) \\
\mathbf{a^5} &= \mathbf{f^5}\left(\mathbf{LW^{5,4}}\boldsymbol{a^4} + \boldsymbol{b^5}\right) \\
\mathbf{a^6} &= \mathbf{f^6}\left(\mathbf{LW^{6,5}}\boldsymbol{a^5} + \boldsymbol{b^6}\right)
\end{aligned} \tag{4-3}$$

p is the input vector denoting the four input parameters while $\mathbf{a^6}$ is the output vector denoting the RH profile at the six standard pressure levels. Being a multilayer network, the output vector of one layer is the input vector for the subsequent layer. Equation (4-3) is rewritten as:

$$\mathbf{a^6} = \mathbf{f^6}\left(\mathbf{LW^{6,5}}\left(\mathbf{f^5}\left(\mathbf{LW^{5,4}}\left(\mathbf{f^4}\left(\mathbf{LW^{4,3}}\left(\mathbf{f^3}\left(\mathbf{LW^{3,2}}\left(\mathbf{f^2}\left(\mathbf{LW^{2,1}}\left(\mathbf{f^1}(\mathbf{IW^{1,1}p} + \boldsymbol{b^1})\right) + \boldsymbol{b^2}\right)\right) + \boldsymbol{b^3}\right)\right) + \boldsymbol{b^4}\right)\right) + \boldsymbol{b^5}\right)\right) + \boldsymbol{b^6}\right) \quad (4\text{-}4)$$

As specified in the design of the ANN, $\mathbf{f^1}$, $\mathbf{f^5}$ and $\mathbf{f^6}$ are linear transfer functions while $\mathbf{f^2}$, $\mathbf{f^3}$ and $\mathbf{f^4}$ are tan-sigmoid transfer functions. During training, the weights and the biases for all the neurons at each layer are adjusted after every iteration until the number of specified iterations is reached. The number of required iterations is set to be higher than the number of iterations after which the performance of the neural network does not improve with further iterations. The values of the weights and biases for all the neurons at each layer are then fixed at the figures reached after the final training iteration. Finally, during the operation of the ANN to determine RH profile, the input parameters are fed into the network and processed at each layer without iteration and the output received at the output layer.

CHAPTER 5. RESULTS, ANALYSIS AND DISCUSSION

5.1 IWV Evaluated at KMD using Radiosonde Data and at RCMRD using GPS Data

The first step was to determine the IWV for both stations, the IWV at KMD being determined using radiosonde data, and the IWV at RCMRD using GPS phase delay data. The average monthly IWV values evaluated for both stations is represented in Figure 5-1 below:

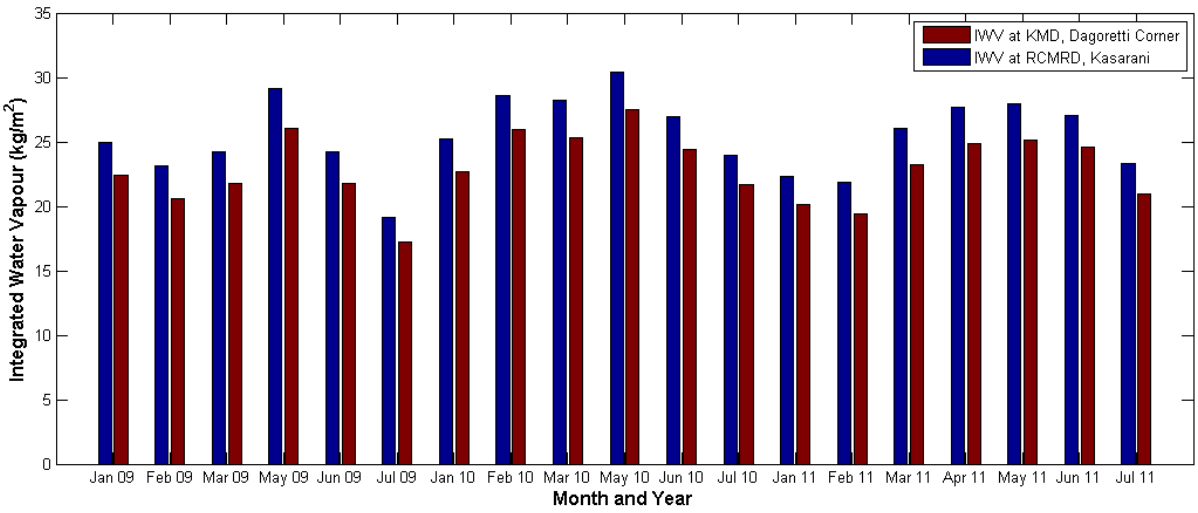

Figure 5-1: Monthly Average IWV Values Evaluated for KMD and RCMRD Stations

Studying the IWV data for both stations, it was observed that the IWV evaluated from GPS data at RCMRD was in all cases higher than the IWV evaluated from radiosonde data at KMD as represented in the monthly averaged values in Figure 5-1 above. This was expected because RCMRD site is 190 m lower in altitude than KMD site.

Over the three years, the IWV at KMD was observed to be on average 89.96% of the IWV at RCMRD. This ratio did not show any particular response to season, but rather ranged randomly between 86.2% and 94.7%. More accurately, the relationship between the IWV at KMD and the IWV at RCMRD could be represented by Figure 5-2.

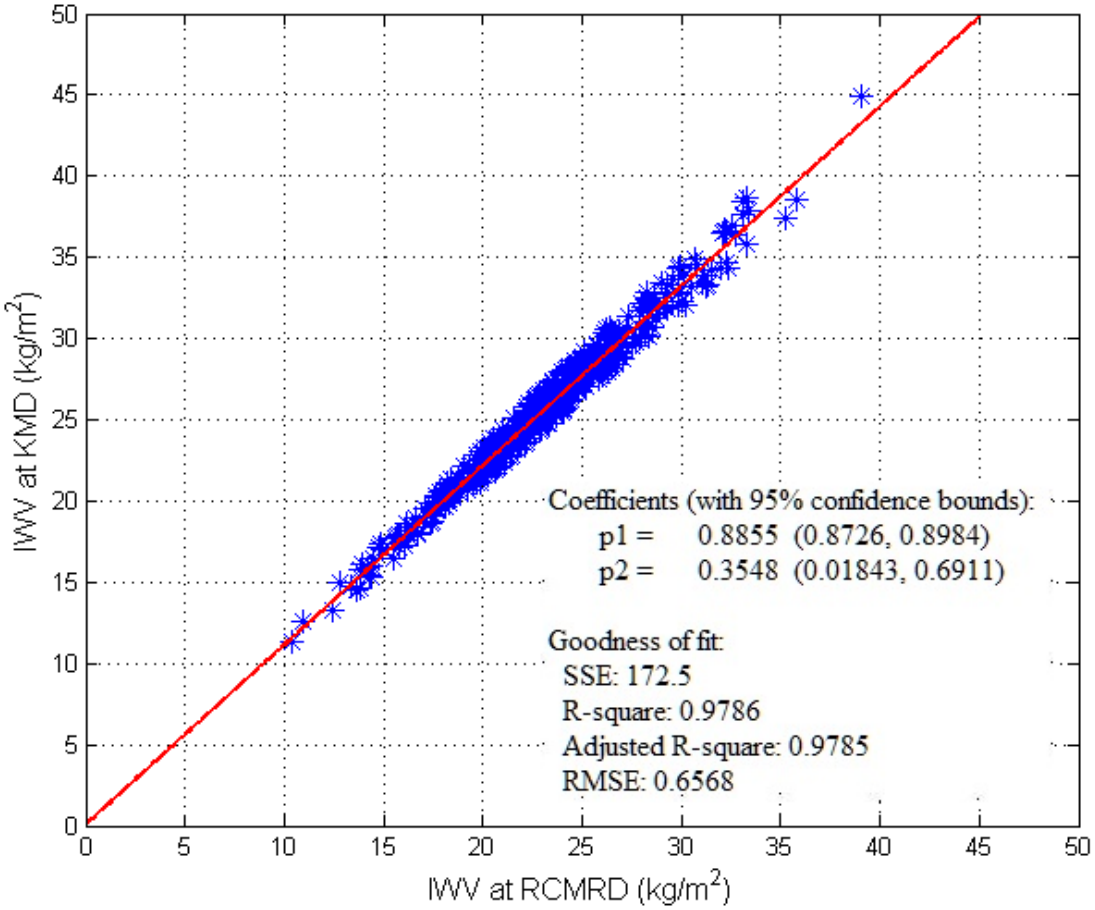

Figure 5-2: Comparison between IWV values at KMD and RCMRD over three years (2009-2011)

The relationship between the IWV at KMD and the IWV at RCMRD over the three years could therefore be approximated by Equation (5-1) below.

$$IWV_{KMD} = 0.8855 \times (IWV_{RCMRD}) + 0.3548 \quad (5\text{-}1)$$

The estimation of IWV at RCMRD by extrapolation of the temperature and pressure values from KMD to RCMRD (using the known atmospheric temperature lapse rate and atmospheric pressure scale height), assuming same ground level RH values at RCMRD as at KMD, and then calculating using Equation (4-2) yielded the same relationship over the three years.

5.2 Relative Humidity Profile evaluated from Radiosonde Data

The RH profile at KMD was observed by radiosonde for the three years in study. The RH profile values for 2009 and 2010 were used for training the neural network while the RH profile values for 2011 were used for validation of the system by comparing with the RH profile values obtained as output of the neural network.

Figure 5-3 below presents the RH Profiles for all the days of the years 2009 to 2011 for which radiosonde RH Profile data was collected.

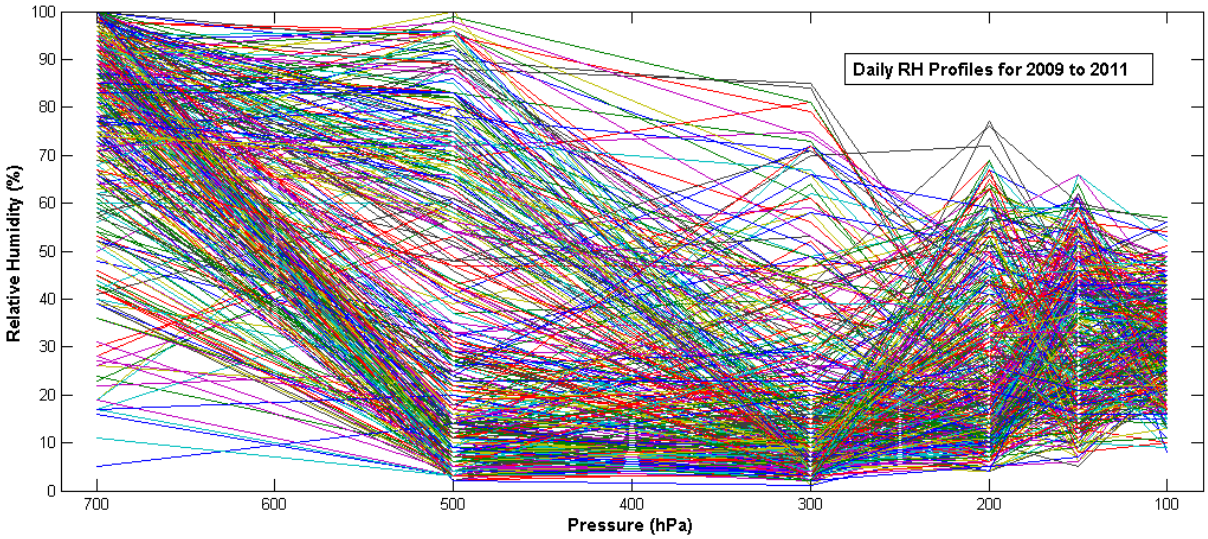

Figure 5-3: Daily RH Profiles recorded at KMD for 2009 to 2011

Majority of the RH profiles exhibited an approximate C-shape with highest RH values at the low altitudes and the lowest RH values at the mid altitudes. The RH values in the higher Troposphere regions tended to be high too. In this study, it was observed that the highest RH values were at the altitude corresponding to 700 hPa. At 500 hPa, the RH values were lower than those observed at 700 hPa for more than 85% of the profiles and the lowest RH values were at 300 hPa. The observed RH value at 300 hPa was below 50% for 94.6% of the soundings. Between 300 hPa and 100 hPa, there was an increase in the RH values with RH values at 100 hPa being higher than at 300 hPa for 82.6% of the soundings.

This observation is consistent with the literature in the report by Folkins *et al.* (2002) indicating that the mean RH profile for tropical regions takes a C-shape with highest RH values in the boundary layer, low RH values in the mid troposphere and high RH values in

the upper troposphere. This is observable also in the monthly average RH profiles in Figure 5-4 and also the yearly average RH profiles in Figure 5-5.

In some cases, the expected C-shape of the RH profile was not observed. For the days that IWV was high (over 28 kgm^{-2}), the RH observed at 500 hPa was over 50% for 88% of these cases, the average RH being 72.8% as opposed to average RH of 33.4% for all the measurements over the three years. For cases of high IWV therefore, the RH measurement at 500 hPa was not expected to drop significantly from the value observed for 700 hPa, and could even be higher in many of the cases.

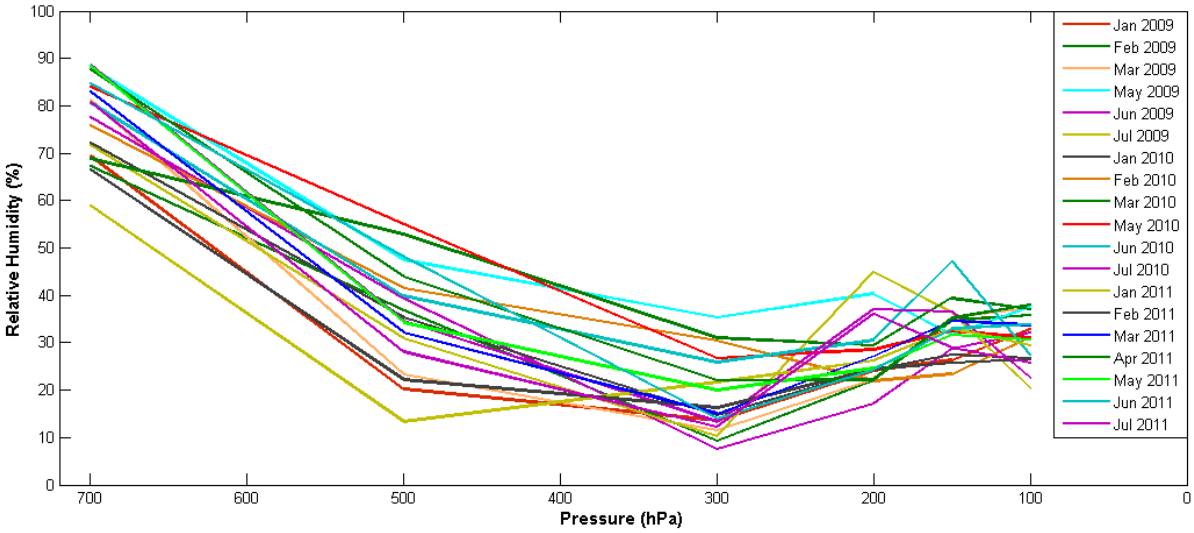

Figure 5-4: Monthly Average RH Profile recorded at KMD for 2009 to 2011

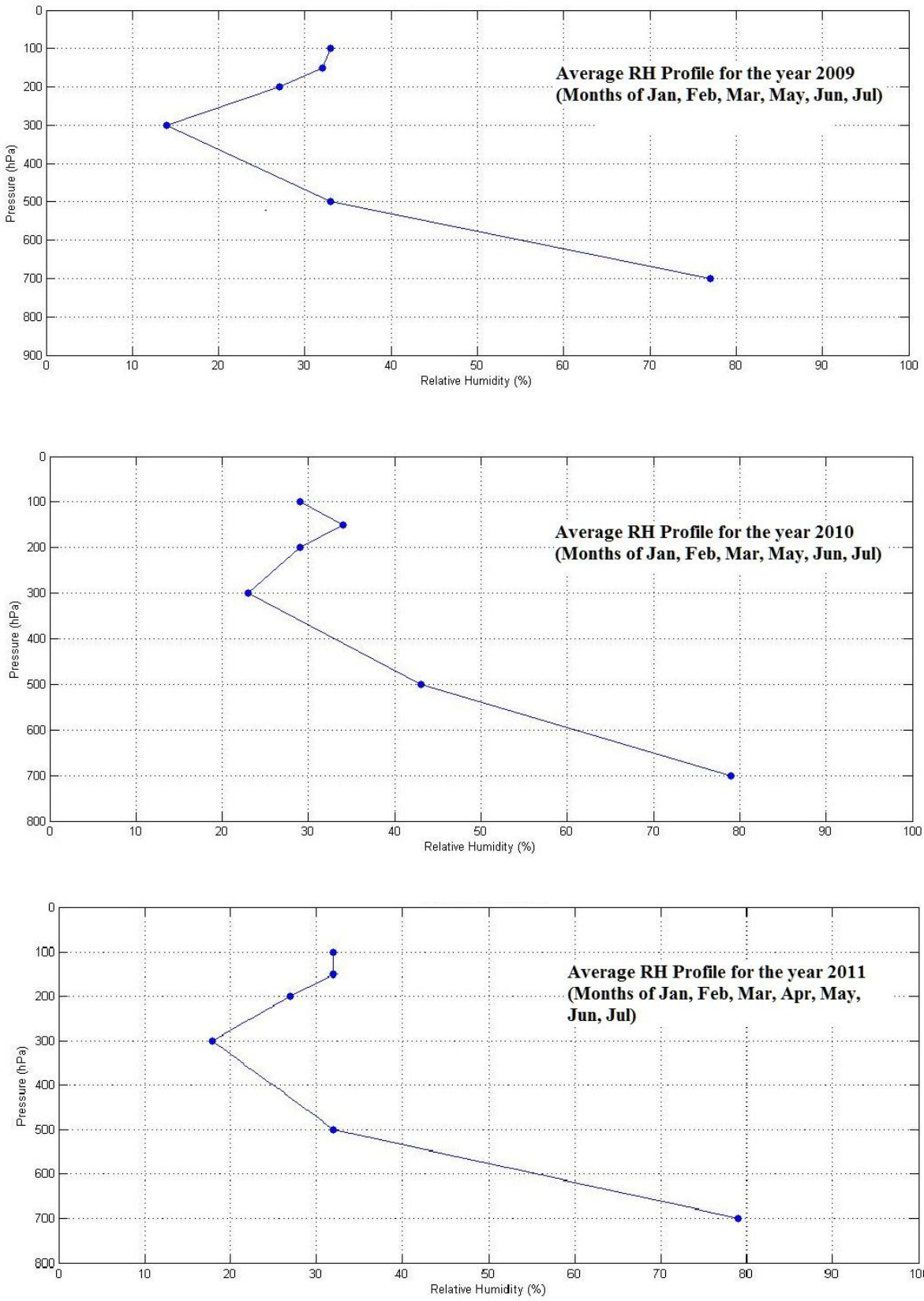

Figure 5-5: Average RH Profile for the years 2009 to 2011

5.3 Influence of IWV and Ground Level Temperature, Pressure and RH on RH Profile

The average IWV calculated from the available radiosonde data using equations (4-1) and (4-2) for the three years of study was found to be as presented in Table 5-1.

Table 5-1: Average Annual IWV for 2009 to 2011.

Year	IWV (Average)	Ground Level Temp (Average)	Ground Level P (Average)	Ground Level RH (Average)
2009 (Jan, Feb, Mar, May, Jun, Jul)	21.657	15.97 °C	822.28 hPa	82.84 %
2010 (Jan, Feb, Mar, May, Jun, Jul)	24.550	15.92 °C	822.74 hPa	88.29%
2011 (Jan, Feb, Mar, Apr, May, Jun, Jul)	22.802	15.97 °C	822.46 hPa	82.80%

Relating the values in Table 5-1 above with the RH profiles in Figure 5-5, it could be suggested that the IWV does have an influence on the RH profile. A higher value of IWV tends to make the C-shape of the RH profile shallower, so that the RH at the mid altitudes where the pressure is ~500 hPa to ~200 hPa will have higher values than would have been the case if the IWV value was lower. This is represented in Figure 5-6 and Figure 5-7. The Units for Pressure are in hPa, Relative Humidity in %, IWV in kg/m^2 and temperature in °C.

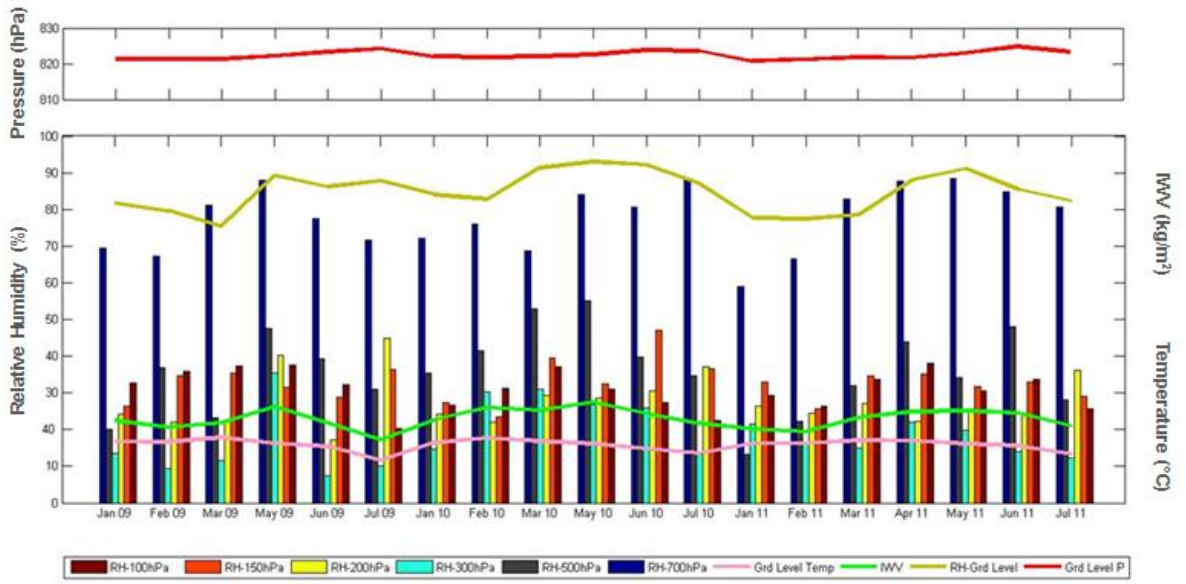

Figure 5-6: Representation of Influence of Ground Level Conditions on RH Profile

This influence was not easily recognisable because the temperature profile also had an influence on the RH profile. It was observed that higher ground level temperatures coincided with higher IWV values. This result is consistent with the Clausius-Clapeyron Relation (Lawrence, 2005) which states that higher atmospheric temperatures result in higher ability by the atmosphere to contain water vapour. The increase in IWV readings due to temperature increase did not necessarily lead to higher RH values though. At same ground level temperature, higher IWV values generally corresponded to high RH values at each pressure level. But if IWV remained constant, increase in temperature would lead to reduction in RH readings because of the expected rise in the saturation vapour pressure. The relationship between the temperature, IWV and the RH readings at the pressure levels was therefore complex, but it was expected that the neural network would be able to recognise this complex relationship and model the output RH values from the input IWV and temperature combinations. The atmospheric temperature profile can be determined by extrapolation from ground level temperature, making use of the atmospheric temperature lapse rate, which is well characterised for most regions on earth.

Some of the RH profiles for random days of the year 2009, accompanied by their respective ground level temperature, pressure, RH and IWV are shown in Figure 5-7.

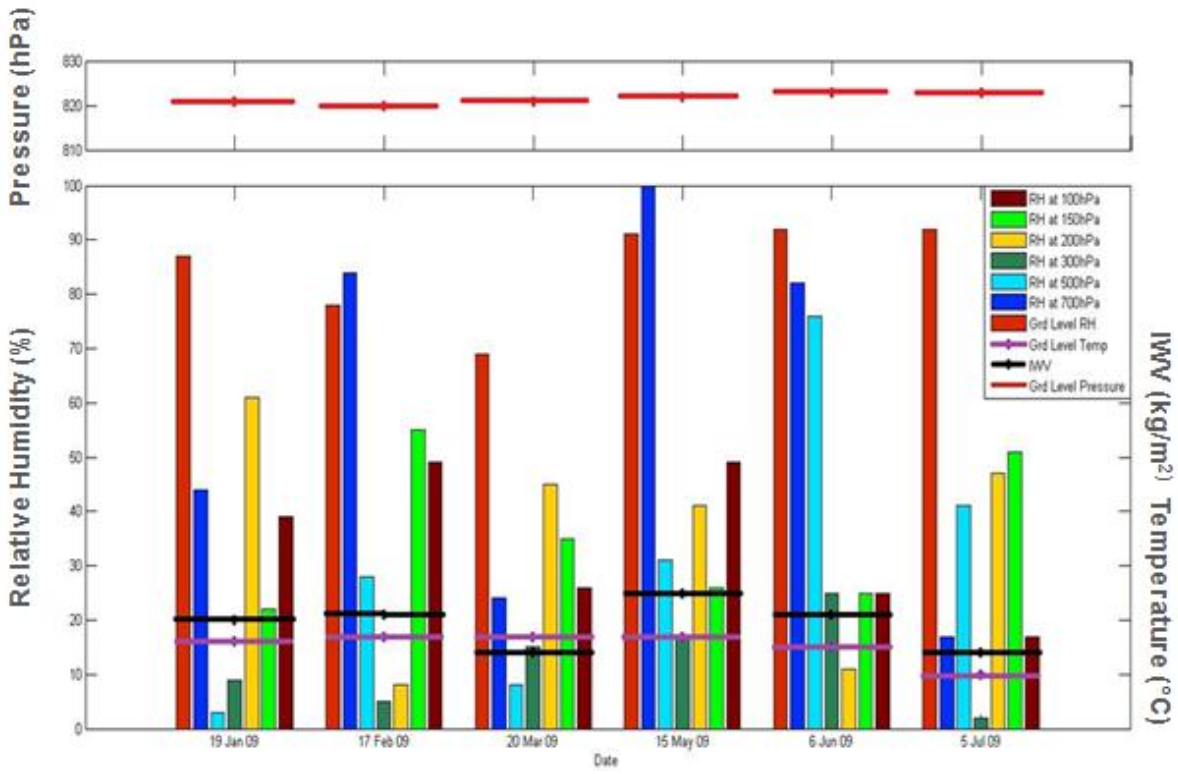

Figure 5-7: RH Profiles for Randomly Selected Days in 2009

On examining the daily RH profiles and also their respective ground level conditions, the relation between IWV and RH profile as earlier mentioned was generally maintained with few exceptions. It was noted too that low ground level temperature led to lower values for RH at all the levels with the C-shape of the RH profile generally maintained. Higher RH values at ground level tended to coincide with higher RH values at the other levels and vice versa, especially when high ground level RH was accompanied by high IWV or low ground level RH accompanied by low IWV as seen in the monthly averages in Figure 5-6. No clear relationship between ground level pressure and the RH profile could be recognised. This was expected though because the RH profile values are recorded at altitudes based on the standard pressure levels rather than geopotential height, meaning the effect of pressure variations on the RH profiles would already have been factored into the RH measurements. The ground level pressure exhibited very little variation over the three years with lowest value being 819 hPa and highest value being 827 hPa.

Even though a relation could be inferred to exist between IWV, ground level conditions and the RH profile as mentioned above with higher IWV and RH tending to lead to higher values for RH in the mid and higher altitudes, this relation could not be accurately evaluated directly from the data available. It was hoped however, that a neural network algorithm would be able

to take advantage of this inherent relation between RH profile and IWV together with ground level conditions to accurately predict the RH profile using IWV and ground level conditions as input, also taking advantage of the expected C-shape of the RH profile.

5.4 Relative Humidity Profile obtained from the Neural Network

After the training of the network was complete, the network was then validated. The available GPS phase delay data recorded for the days of the year 2011 was processed to obtain the IWV at RCMRD for the same days. This IWV at RCMRD was operated on to obtain the IWV at KMD using Equation (5-1). The calculated value for IWV-KMD was then used as input into the neural network system, along with three other input parameters: Ground level temperature, pressure and RH. The output of the system was the RH profile at RCMRD as determined using the neural network system and represented graphically in Figure 5-8.

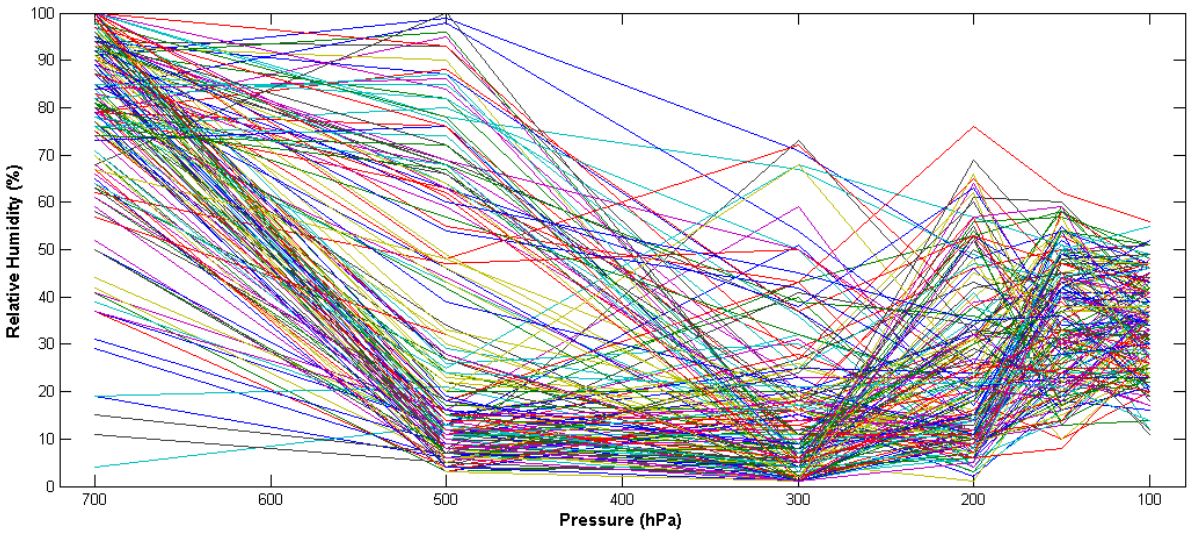

Figure 5-8: Daily RH Profiles for Jan 2011 to Jul 2011 as Determined by Neural Network

The monthly average values for RH Profile were then found to be as shown in Figure 5-9 below for the months in question for the year 2011:

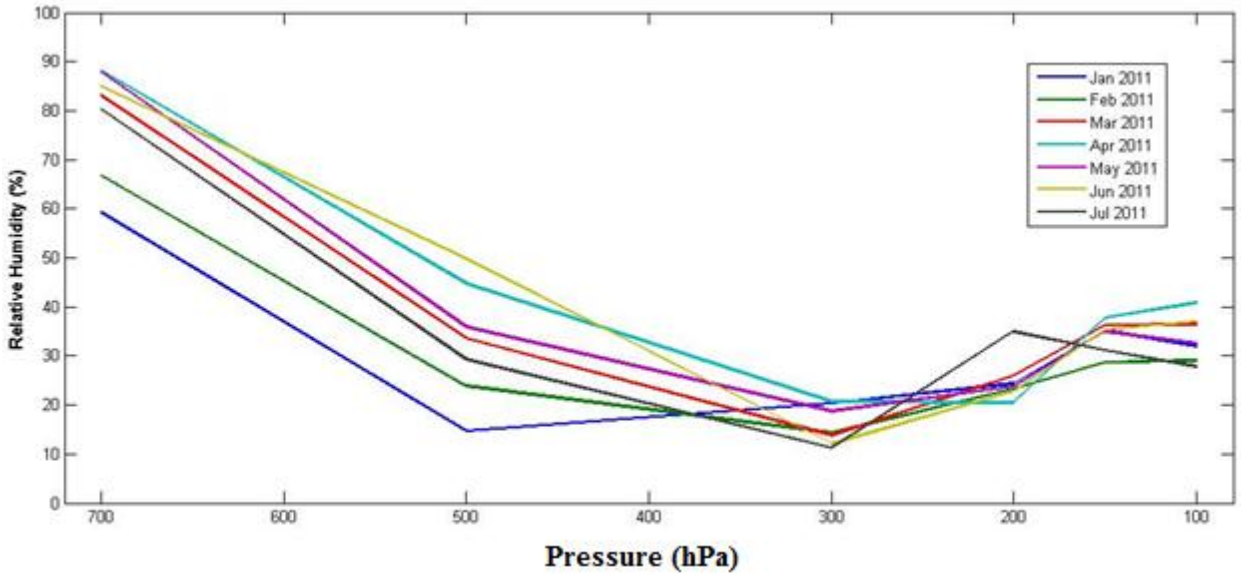

Figure 5-9: Monthly Average Profile values for Jan to Jul 2011 as determined by Neural Network.

The next step was to determine the accuracy of these values for RH profile obtained as outputs of the neural network.

5.5 Comparison between Neural Network and Radiosonde RH Values

A side by side comparison between the RH profile values obtained from radiosonde data and the RH profile values obtained as output of the GPS and neural network system showed that the neural network, taking as input the IWV obtained from GPS analysis and the ground level weather conditions, yielded the RH profile fairly accurately. Some RH profile values generated from the neural network for random days in 2011 are shown in Table 5-2, compared with radiosonde-recorded RH profile values the same days.

Table 5-2: Side by Side Comparison between Radiosonde and Neural Network RH Values

Day	RH at 700 hPa (%)		RH at 500 hPa (%)		RH at 300 hPa (%)		RH at 200 hPa (%)		RH at 150 hPa (%)		RH at 100 hPa (%)	
	Rad	NN	Rad	NN	Rad	NN	Rad	NN	Rad	NN	Rad	NN
4-Jan-11	87	89	13	14	4	2	25	24	22	20	38	36
22-Jan-11	64	63	23	27	9	5	29	30	59	59	35	39
27-Feb-11	17	19	20	21	5	5	22	24	19	24	19	24
22-Mar-11	100	100	4	8	14	13	28	25	28	26	34	32
8-Apr-11	96	94	66	68	23	23	23	23	18	22	19	19
22-Apr-11	95	97	11	11	9	11	35	35	44	51	39	50
3-May-11	85	84	96	98	53	54	29	26	50	55	37	35
19-May-11	77	76	24	25	66	68	59	57	48	51	25	27
28-May-11	77	77	21	22	15	16	6	7	29	33	17	21
6-Jun-11	72	73	74	76	5	7	14	15	37	40	35	39
17-Jun-11	78	78	65	68	12	7	4	5	43	42	43	46
7-Jul-11	74	74	6	8	11	10	36	32	18	21	26	33
17-Jul-11	76	77	8	9	2	4	68	69	47	48	48	49
23-Jul-11	92	91	17	20	72	68	28	24	59	57	17	23
30-Jul-11	85	83	32	32	3	2	43	45	39	38	35	38

The first and most direct measure of accuracy was the difference between the radiosonde RH value and the corresponding RH value predicted by the GPS and neural network system at a particular pressure level. At 700 hPa, the largest difference was 2 % and the mean difference 1.006 %; at 500 hPa, it was 4 % and 1.695 % respectively; at 300 hPa: 5 % and 2.257 % respectively; at 200 hPa: 5 % and 2.251 % respectively; at 150 hPa: 7 % and 3.048 % respectively and lastly at 100 hPa: 12 % and 3.281 % respectively.

Further accuracy analysis for each of the six considered standard pressure levels is done in Figure 5-10 to Figure 5-16 along with the respective scatter plots included too:

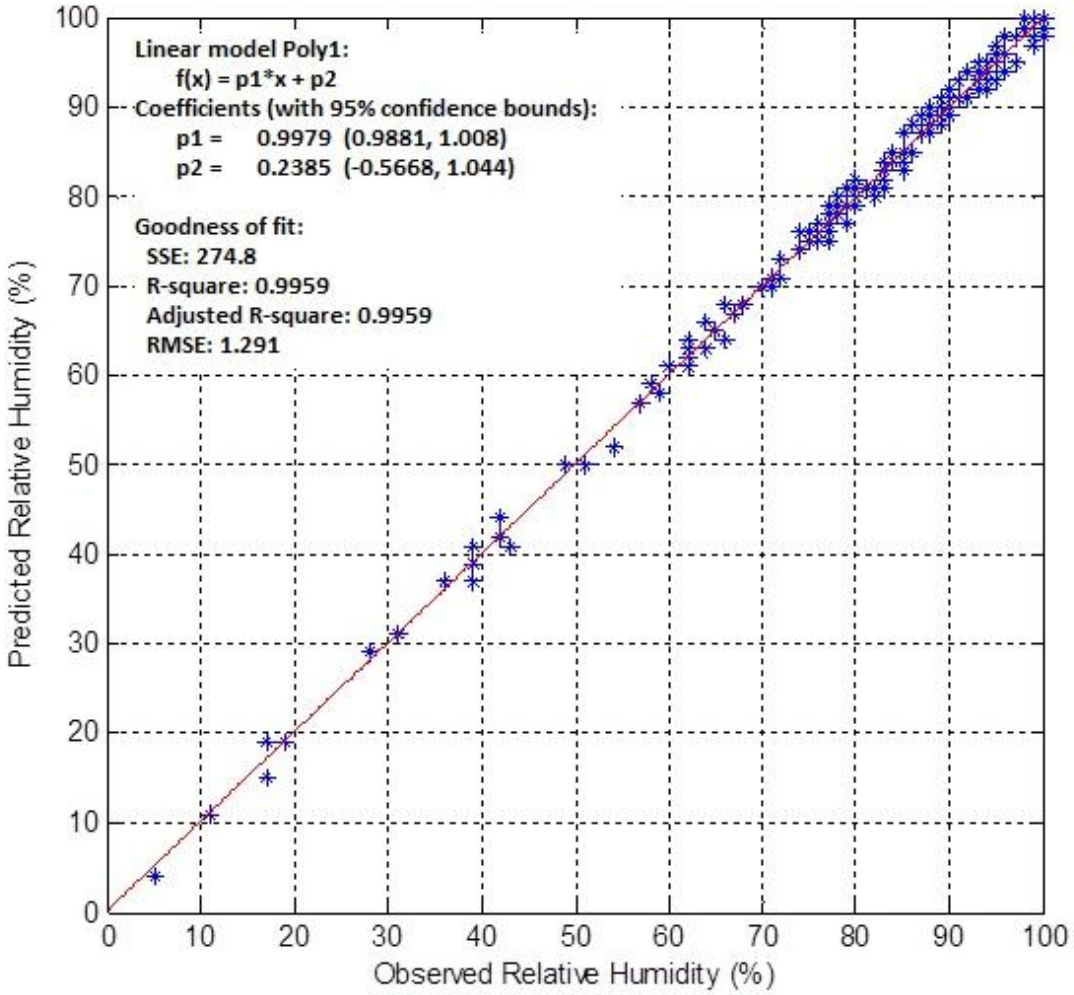

Figure 5-10: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 700 hPa pressure.

At 700 hPa, the Root Mean Square Error (RMSE) in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 1.291 %.

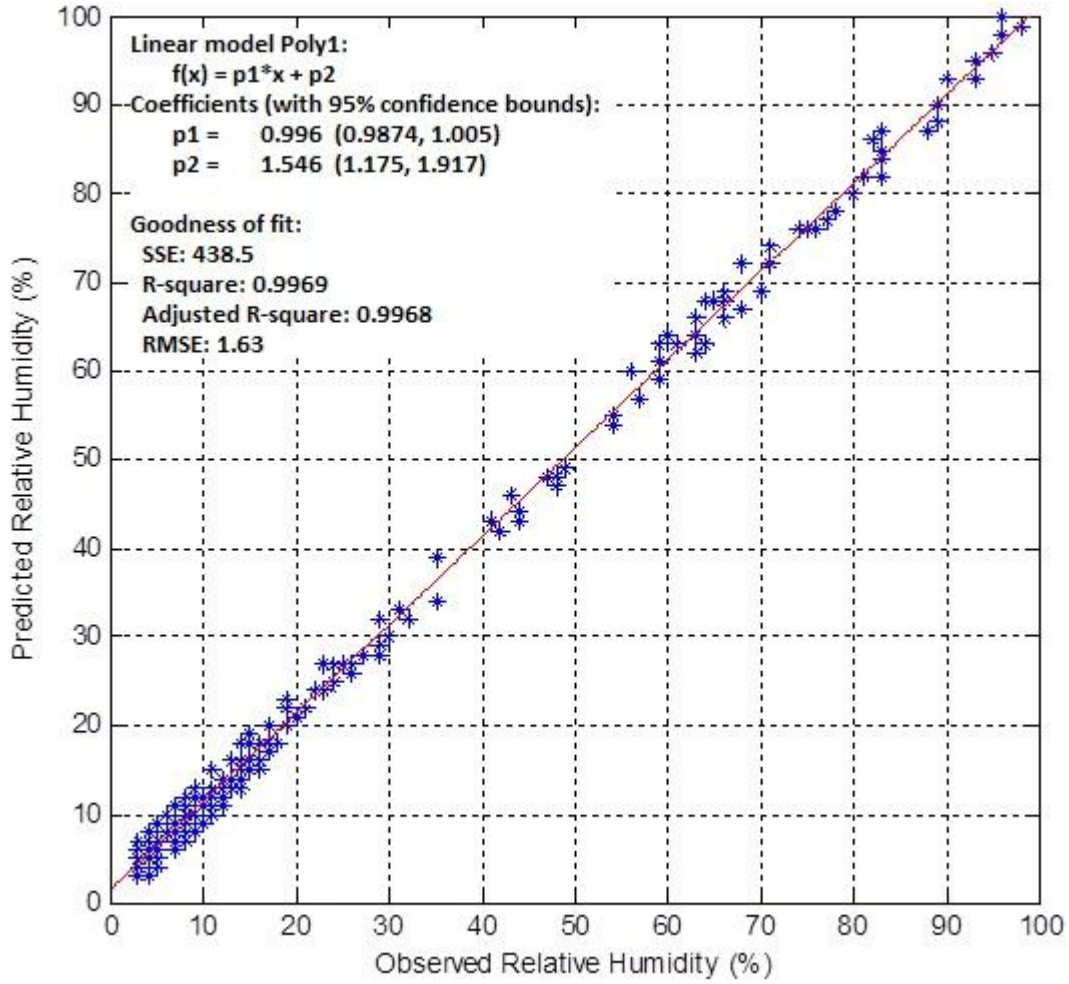

Figure 5-11: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 500 hPa pressure.

At 500 hPa, the RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 1.63 %.

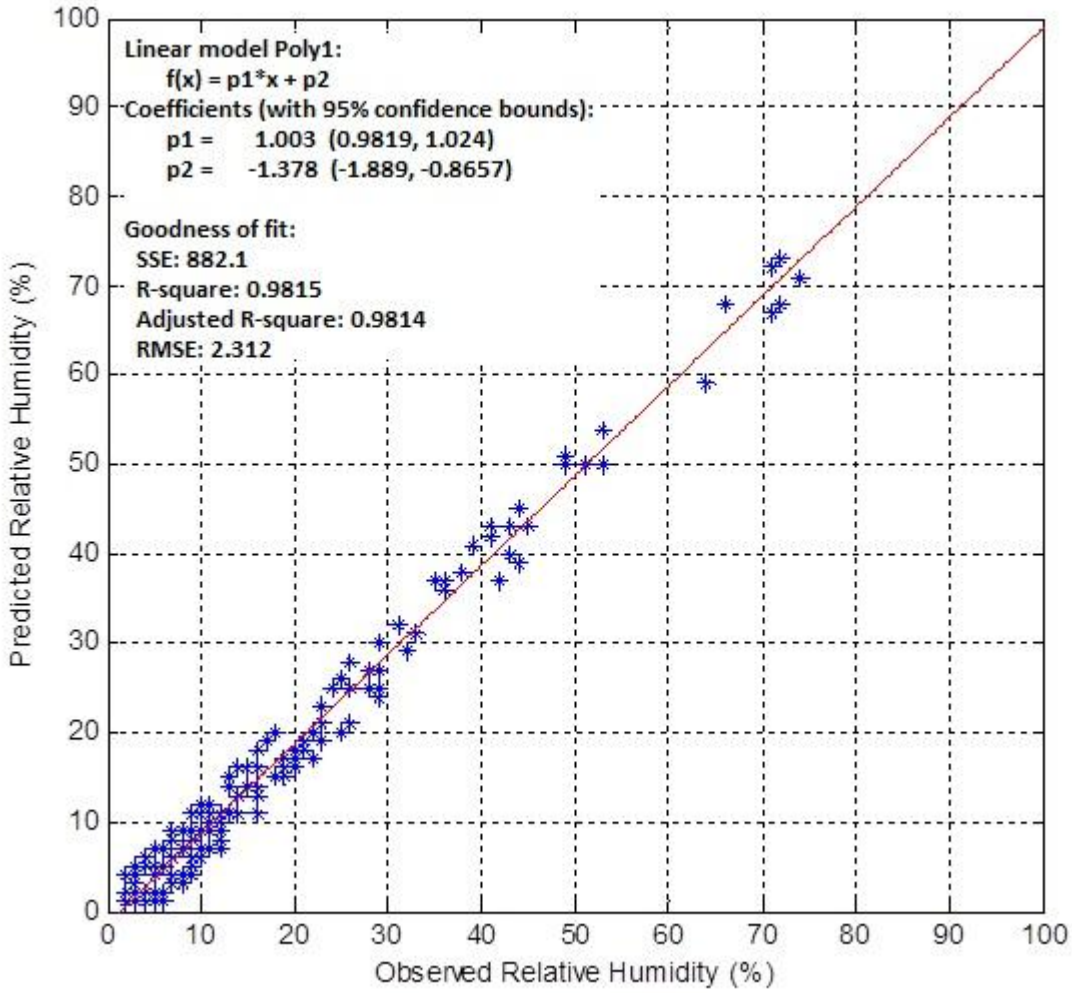

Figure 5-12: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 300 hPa pressure.

At 300 hPa, the RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 2.312 %.

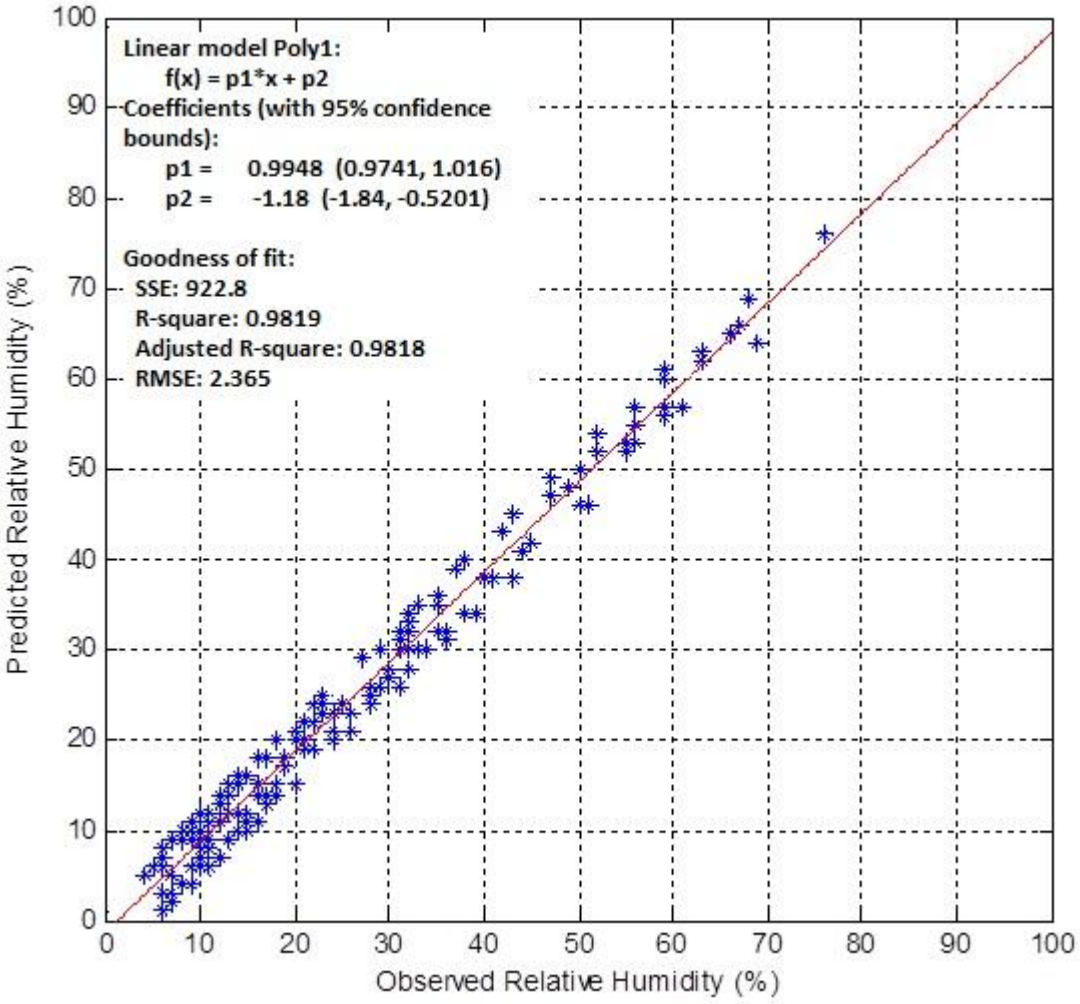

Figure 5-13: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 200 hPa pressure.

At 200 hPa, the RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 2.365 %.

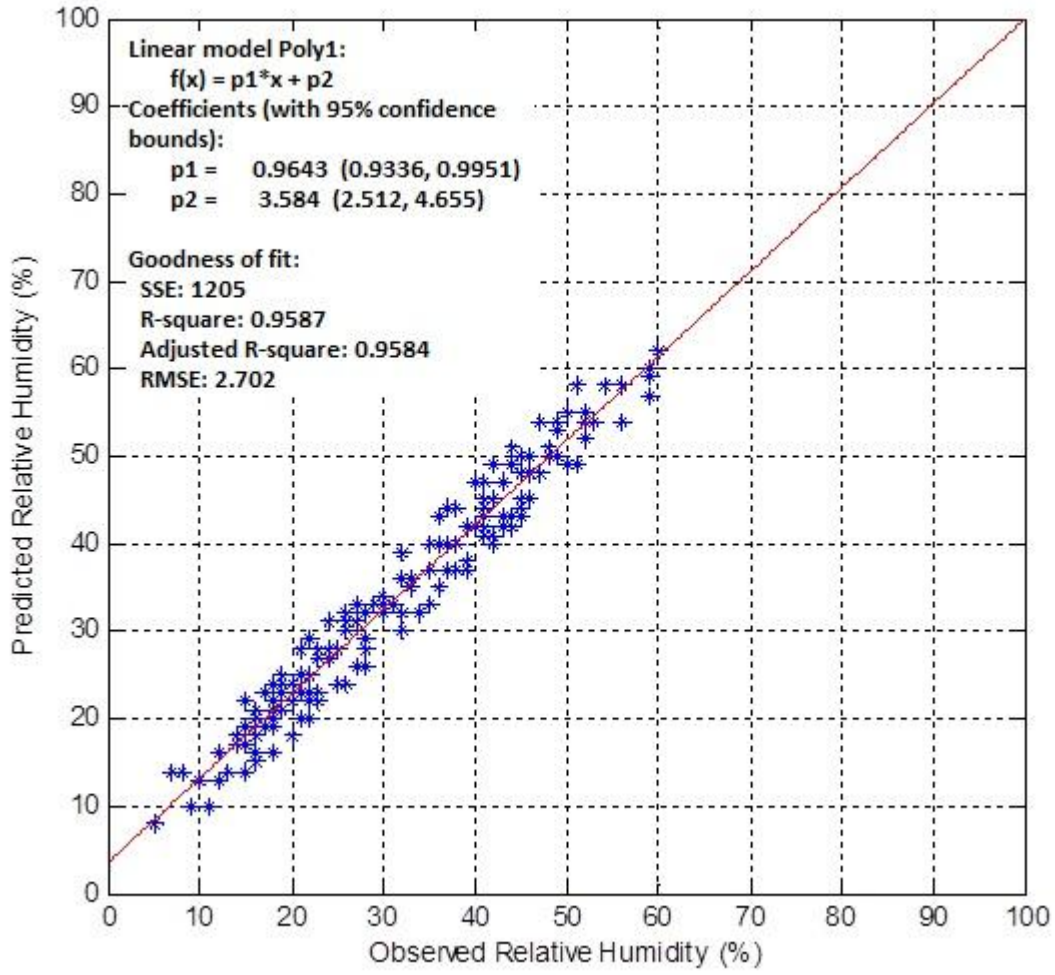

Figure 5-14: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 150 hPa pressure.

At 150 hPa, the RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 2.702 %.

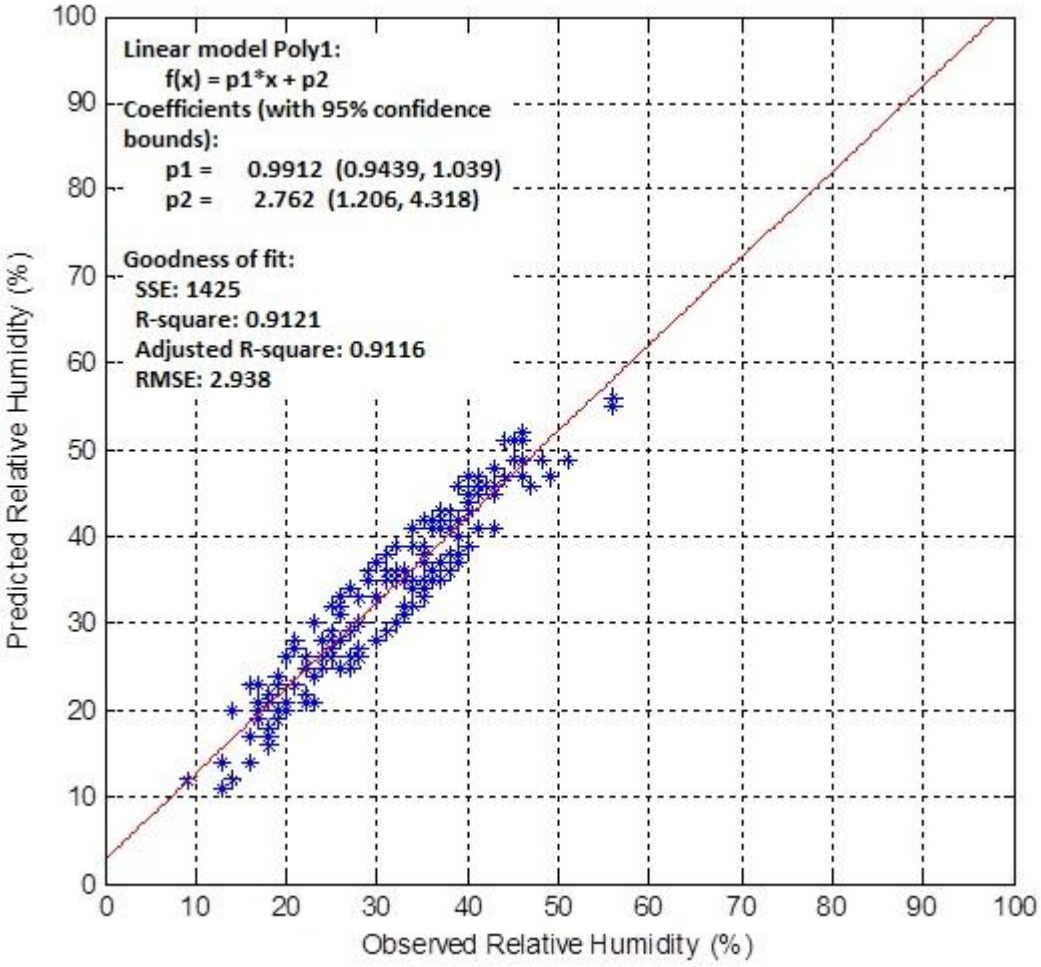

Figure 5-15: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at altitude corresponding to 100 hPa pressure.

At 100 hPa, the RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 2.938 %.

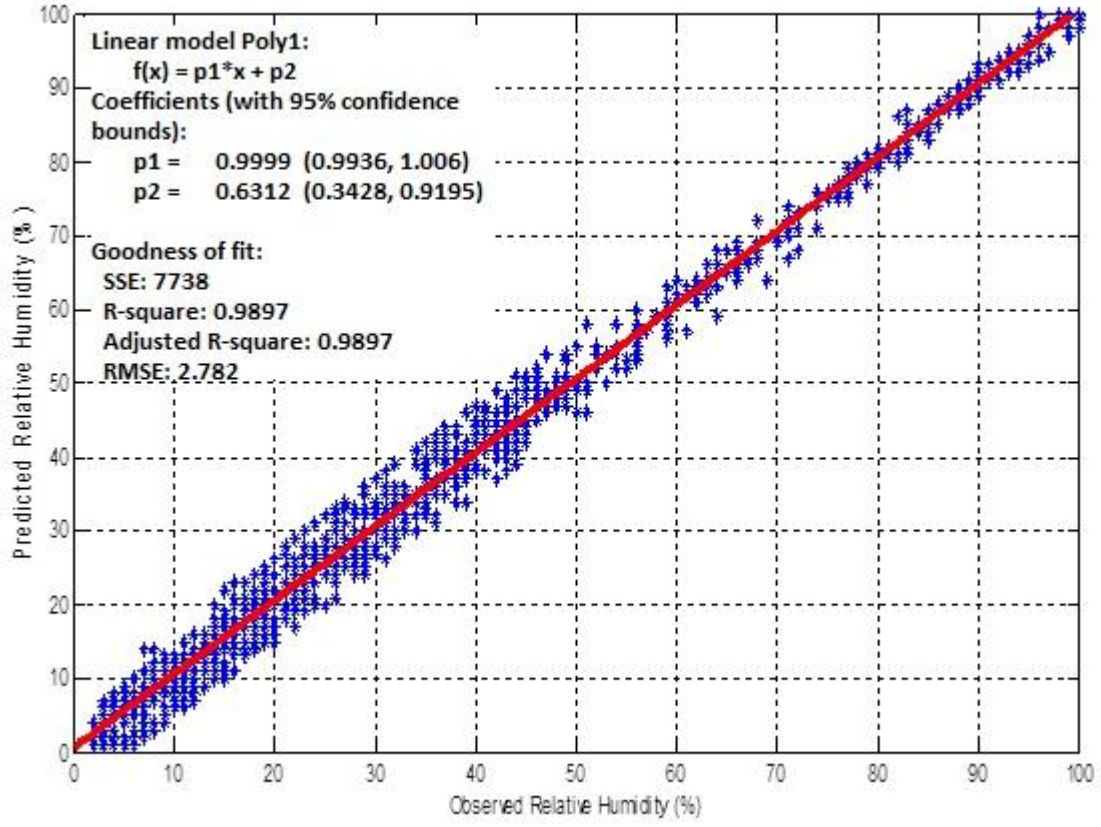

Figure 5-16: Scatter plot for RH obtained from GPS and neural network system vs. RH evaluated from radiosonde data at all six standard pressure levels.

In this case, the Vertically Averaged RMSE in the RH values obtained from GPS and neural network system when compared to radiosonde RH values was 2.782 % which is within the 5 % margin of error of the Vaisala RS 92DRH radiosonde equipment which was used for the radiosonde measurements.

Table 5-2 summarises the accuracy of the neural network system in predicting the RH profile using as input the IWV derived from GPS data and the ground level temperature, pressure and RH:

Table 5-3: Summary of the Accuracy Analysis

Pressure Level	R	RMSE	R^2	Adjusted R^2
700 hPa	0.9979	1.291 %	0.9959	0.9959
500 hPa	0.9960	1.630 %	0.9969	0.9968
300 hPa	1.0000	2.312 %	0.9815	0.9814
200 hPa	0.9948	2.365 %	0.9819	0.9818
150 hPa	0.9643	2.702 %	0.9587	0.9584
100 hPa	0.9912	2.938 %	0.9121	0.9116
All	0.9999	2.782 %	0.9897	0.9897

5.6 System Validation using Data from Months not used in Calibration

Further validation of the system's performance was done using GPS and radiosonde data from the months of August 2011 to December 2011. These correspond to the months of August 2009 to December 2009 and August 2010 to December 2010, months for which GPS and radiosonde data had earlier not been included in the system's training data as mentioned in section 4.2. This additional validation was done in order to check the robustness of the training system, if it would result in predictions as accurate for August to December as it was for January to July, hence reliable for all-year-round prediction even if only part-of-the-year data is used for calibration. The neural network-generated RH profile values for the random days in August 2011 to December 2011 is shown in Figure 5-17.

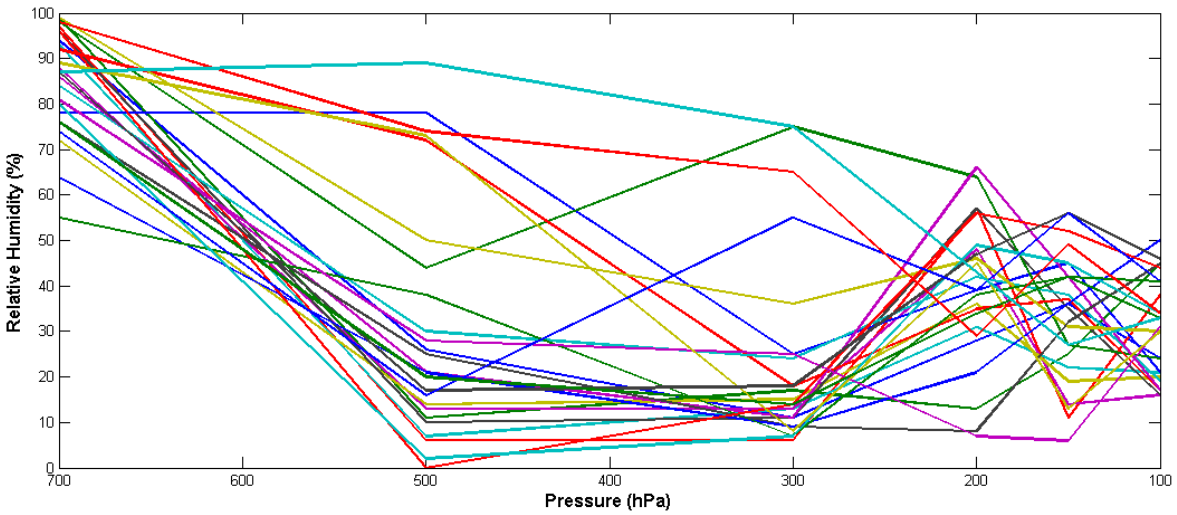

Figure 5-17: RH profile for random days in Aug to Dec 2011

The comparison between the neural network-generated RH profiles for the randomly selected days of the months of August 2011 to December 2011 and the actual radiosonde-recorded RH profiles for the same months is given in the scatter plot in Figure 5-18.

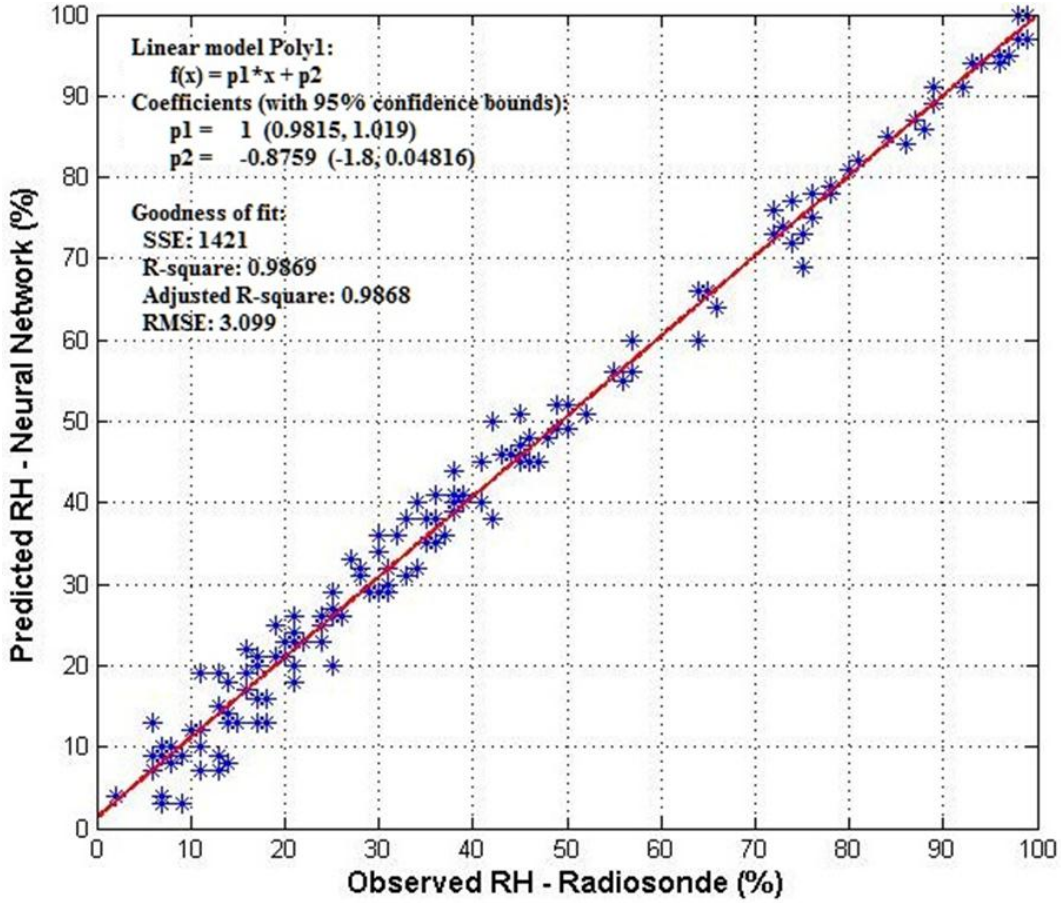

Figure 5-18: Scatter Plot for RH Profile for Aug to Dec 2011

The vertically averaged RMSE for August 2011 to December 2011 was 3.099 %, still within the 5 % margin of error of the Vaisala RS 92DRH radiosonde equipment, but noticeably a little higher than the 2.78 % RMSE for the months of January 2011 to July 2011. This suggests that the system was still accurate at modelling for months not used in calibrating it, but accuracy of the model is slightly reduced.

CHAPTER 6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

The objective of the study, which was to develop a tool that would use a ground based GPS receiver in conjunction with the technique of neural networks to determine the RH profile, was achieved.

The study relied entirely on data collected by the RCMRD and the KMD. Even though the aim of both organisations is to have the GPS and radiosonde data recorded daily, the reality is that this data is not always recorded for all the days of the year. One requirement of the study, for both the training and validation data sets, was that both GPS and radiosonde data needed to be available for the same day and time instance in order that it is used in the study. Concerning the period of this study, 2009 to 2011, it happened that for many of the days and sometimes entire months of the three years, either GPS or radiosonde data was not recorded and was therefore unavailable. It is felt that if the radiosonde and GPS data had been available for all the days, or at least a few days of all the months of the three years, the training of the neural network would have had better results because data representing the atmospheric characteristics of all the seasons of the year would be available enabling the neural network to learn the underlying structures and diverse configurations.

Training of the neural network using the Levenberg-Marquadt Optimisation Algorithm resulted in great savings in training time. When training of the network was done with the study's data set using the standard gradient descent backpropagation algorithm, on average at least half a day was required for generalisation to be achieved for each training episode. Using the Levenberg-Marquardt Optimisation algorithm on the other hand, the same degree of generalisation could be achieved in approximately 30 minutes. Neural network architecture development is essentially a trial and error process, and the final architecture had 60 neurons in the four intermediate layers. Each time a new neuron was added to the network architecture or removed, the training exercise needed to be freshly performed a few times. In the initial stages of the study, the standard backpropagation algorithm was used in training the network, and it soon became clear that over a year would be required to arrive at the final architecture. The decision was made to switch to the Levenberg-Marquadt Optimisation algorithm, the trade-off being that this method was much more demanding of the computing

resources and even necessitated a computer processor and memory upgrade. With this new algorithm, architecture development was achieved in slightly over three months.

Only four variables, IWV, ground level temperature, pressure and relative humidity, were used as inputs to the neural network. But these are not the only influencers on the RH profile. There are many other factors such as wind speed and direction, ground surface temperature, presence of aerosols, etc. that also have an influence on the water vapour profile. These other variables could possibly have improved the performance of the system if they were included too as inputs to the neural network.

Even though the neural network was successful in evaluating the water vapour profile from the input parameters, it must be admitted that a neural network being inherently a black box, makes it difficult to study the underpinning relations between the input variables and how they interact to give rise to the outputs. This meant that troubleshooting the network to identify the sources of overfitting for instance was difficult, with the only solution to detected inaccuracies being training the entire network afresh each time, which was quite tedious and time-consuming.

The accuracy of the proposed system was higher at the lower altitudes and reduced with increase in altitude. This conforms to the trend observed by previous studies including the report by Sivira *et al.* (2012). From the specifications of the radiosonde equipment, and also based on the results of a WMO study in Yangjiang, China in 2010, the accuracy of the Radiosonde model, Vaisala RS 92DRH, which is used for the taking the RH measurements is 5% (Nash *et al.*, 2011). For the days within the months whose data was used in training the neural network, the system yielded RH values whose difference with the radiosonde RH values was on average, 2.3 %. The vertically averaged RMSE of the RH data obtained using the GPS and neural network system was 2.782 %. This is within the range of the radiosonde equipment's specified accuracy which is 5 %. The accuracy of the system in measuring RH at the six standard pressure levels specified in this study was therefore within the range of the radiosonde's RH measurement margin of error.

Even for months whose data was not used in training the system, the RMSE of the RH data obtained at the neural network output was 3.099 %, still within the 5 % margin of error of the radiosonde equipment. It was concluded that despite the system being trained using data from

only part of the year, it was accurate enough at determining RH profiles for the whole year, even though accuracy is slightly lower for the months not used in system calibration.

The Vertical Resolution of RH measurements using the proposed tool is quite low (200 hPa at lower altitudes and 50 hPa at the higher altitudes). This was the highest resolution possible to achieve while maintaining reasonable degree of accuracy in the RH profile results. Attempts to improve the resolution by having more neurons at the output layer were unsuccessful. This was a difficult problem to solve because the design parameters for multilayer feed-forward backpropagation networks are generally determined through trial and error; this is due to the absence of firm theoretical guidelines on selection of these parameters (Jain and Mao, 1996).

6.2 Recommendations

There is a 190 m difference in altitude between the radiosonde station and the GPS site. There is also a 17.53 km horizontal separation between the two sites. Statistically significant errors could have been introduced into the data during the extrapolation of temperature, pressure and RH data from KMD to RCMRD and also of IWV data from RCMRD to KMD which served to distort the evaluation. It is recommended that a GPS receiver of at least similar capability to the one at RCMRD be installed at KMD – Dagoretti Corner in order to enable collocation between the GPS receiver and the radiosonde. With collocation of GPS and Radiosonde measurements, researchers will enjoy more suitable and better-correlated radiosonde and GPS data which could improve the accuracy of their studies.

Having observed the accuracy of this technique that evaluates the RH profile over a site using the phase delay data obtained using a ground-based GPS receiver at the site coupled with a neural network to process the data, it is recommended therefore that the technique be promoted as a method for tropospheric RH profiling. More work should be done into the neural network architecture proposed in this work in order to improve its predictive accuracy and also the vertical resolution of the data it yields.

In order to improve the characterisation of the temporal and spatial evolution of water vapour, it is recommended that a network of stations equipped with the GPS receivers, the software to process the GPS data into IWV, and the neural networks to generate the RH profiles be

considered. Figure 6-1 outlines a suggested prototype of how such a network could be implemented.

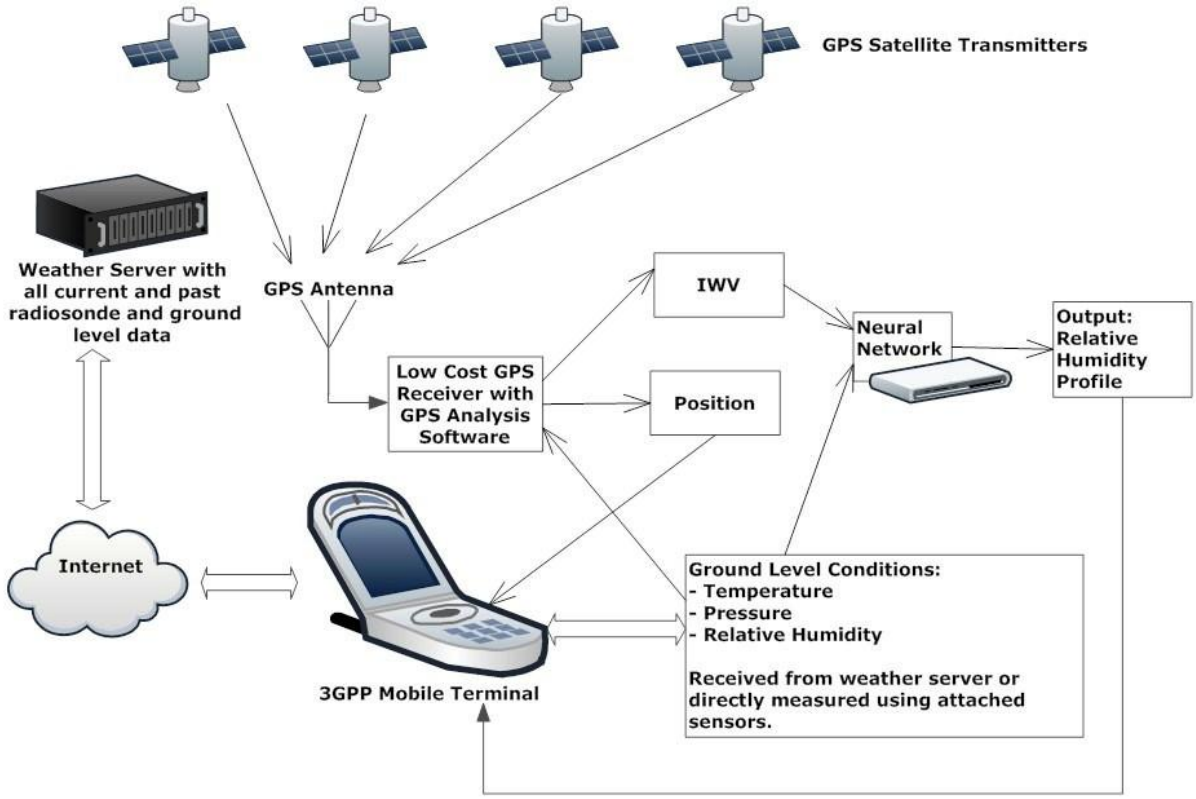

Figure 6-1: Proposed Network Architecture for Real Time Monitoring of Spatial and Temporal Evolution of Water Vapour.

The weather server will be connected to the KMD's network and all radiosonde results will be stored there. The weather server will also store a record of all ground level weather parameters information which could be sent over an (Internet Protocol) IP network to any GPS station needing them. The GPS receivers could be fixed or mobile; the ground level conditions necessary for evaluating RH profile in conjunction with IWV can either be sent to the GPS station over the internet or can be captured directly at the local GPS site via sensors attached to the site assembly. A 3GPP Mobile Terminal, which is just any standard mobile phone with packet data connectivity, is included in the setup to provide connectivity to the internet, enabling sharing of information with the central weather server and also other GPS sites through an agreed data sharing protocol. If the GPS site has connectivity to broadband internet via DSL or FTTx, then the mobile terminal would not be necessary.

Majority of smartphones manufactured today have GPS receivers embedded in them (Malm, 2009). It is hoped that the GPS receivers embedded in smartphones will improve in performance in the years to come, so that it will be possible to develop smartphone applications that evaluate accurately the IWV. If a neural network is embedded into the smartphone too, then the smartphone can act as a mobile standalone tropospheric RH profiling station. The RH profile generated can then be forwarded to the weather server which will be combined with the profiles generated from other smartphones to be employed in water vapour modelling and weather forecasting activities.

REFERENCES

Anderson, J. A. and Rosenfeld, E. (1988) *Neurocomputing: Foundations of Research*. MIT Press, Boston.

Anderson, L. J. (1958) Tropospheric Bending of Radio Waves. *Transactions of the American Geophysical Union* **39,** 208-212.

Askne, J. and Nordius, H. (1987) Estimation of Tropospheric Delay for Microwaves from Surface Weather Data. *Radio Science* **23,** 379-386.

Basili, P., Ciotti, P., Bonafoni, F., Mattioli, V. and Peliccia, F. (2007) Neural Networks for Tropospheric Profiling from GPS-LEO Radio Occultation. *IEEE/1-4244/1212-9/07*, 1693-1696.

Beale, M. H., Hagan, M. T. and Demuth, H. B. (2009) *Neural Network Toolbox™ 6 User's Guide*. The Mathworks Inc, Natick, Massachusetts.

Bengio, Y. and LeCun, Y. (2007) Scaling Learning Algorithms towards AI. In: Bottou, L., Chapelle, O., DeCoste, D., and Weston, J. Eds. *Large-Scale Kernel Machines*, MIT Press, Cambridge, MA.

Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H. (2007) Greedy Layer-Wise Training of Deep Networks. In: Scholkopf, B., Platt, J., and Hoffman, T. Eds. *Advances in Neural Information Processing Systems **19***. MIT Press, Cambridge, MA, 153-160.

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A. and Ware, R. H. (1992) GPS Meteorology: Remote Sensing of Atmospheric Water Vapour Using The Global Positioning System. *Journal of Geophysical Research* **97**, 15787-15801.

Boehm, J. and Schuh, H. (2006) Global Pressure and Temperature (GPT): A Spherical Harmonic Expansion of Annual Pressure and Temperature Variations for Geodetic Applications. *Journal of Geodesy* **2006**.

Boehm, J., Niell, A., Tregoning, P. and Schuh, H. (2006) Global Mapping Function (GMF): A New Empirical Mapping Function based on Numerical Weather Model Data. *Geophysical Research Letters* **33**.

Bose, N. and Liang, P. (1998) *Neural Network Fundamentals with Graphs, Algorithms and Applications*. Tata McGraw-Hill Publications, New Delhi.

Braun, J. J., Curry, J. and Nerem, S. (2004) Remote Sensing of Atmospheric Water Vapour with the Global Positioning System. PhD Thesis, University of Colorado.

Chan, P. W., (2009) Performance and Application of a Multi-wavelength, Ground-based Microwave Radiometer in Intense Convective Weather. *Meteorologische Zeitschrift*, **Vol. 18**, **No. 3**, 253-265.

Dach, R., Hugentobler, U. and Walser, P. (2011) *GAMIT GPS Software Version 5.0 Tutorial*. Astronomical Institute, University of Bern.

Danielson, E. W., Levin, J. and Abrams, E. (2003) *Meteorology*. McGraw Hill, Dubuque, Iowa.

De Haan, S. (2008) Meteorological Applications of a Surface Network of Global Positioning System Receivers. PhD Thesis, University of Wageningen.

Dessler, A. E., Zhang, Z. and Yang, P. (2008) Water Vapor Climate Feedback Inferred from Climate Fluctuations, 2003–2008. *Geophysical Research Letters* **35,** L20704.

Dow, J. M., Neilan, R. E., and Rizos, C., (2009) The International GNSS Service in a Changing Landscape of Global Navigation Satellite Systems. *Journal of Geodesy* **83,** 191–198

DuBois, J. L., Multhauf, R. P. and Ziegler, C. A. (2002) *The Invention and Development of the Radiosonde*. Smithsonian Institution Press, Washington, DC.

Durre, I., Vose, R. S. and Wuertz, D. B. (2006) *Overview of the Integrated Global Radiosonde Archive*. National Climatic Data Centre, Asheville, North Carolina.

Elliot, W. P. (1996) On Detecting Long Term Changes in Atmospheric Moisture. *Climate Change* **31**, 349-367.

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

Folkins, I., Kelly, K. K. and Weinstock, E. M. (2002) A Simple Explanation for the Increase in Relative Humidity between 11 and 14 km in the Tropics. *Journal of Geophysical Research* **107 (D23):** 26.1-26.7.

Geerts, B. and Linacre, E. (1997) The Height of the Tropopause, *Atmospheric Science **12/97***, University of Wyoming.

Hagan, M. T. and Demuth, H. B. (1996) Neural Network Design. PWS Publishing Company, Boston, MA.

Hammerstrom, D. (1993) Working with Neural Networks. *IEEE Spectrum* **July 93**, 46-53.

Hardy, K. R., Hinson, D. P., Tyler, G. L. and Kursinski, E. R. (1992) Atmospheric Profiles from Active Space-based Radio Measurements. *6th Conference on Satellite Meteorology and Oceanography,* Atlanta, GA, Jan 5-10.

Herring, T. A., King, R. W. and McClursky, S. C. (2010) *GAMIT Reference Manual Release 10.4.* Massachusetts Institute of Technology.

Hinton, G. E., Osindero, S., and Teh, Y. (2006) A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation **18***, 1527–1554.

Hooke, W. (2000) Short-term Weather Prediction: An Orchestra in Need of a Conductor. In: *Prediction: Science, Decision-Making and the Future of Nature*. Island Press, 61-83.

Hornick, K. M., Stinchcombe, M. and White, H. (1981) Multilayer Feedforward Networks are Universal Approximators. *Neural Networks* **2, No. 5**, 359-366.

Hopfield, H. S. (1969) Two-quartic Tropospheric Refractivity Profile for Correcting Satellite Data. *Journal of Geophysical Research* **74 (18)**, 4487-4499.

ITU-R (2012) Recommendation ITU-R P.453-10, 2012: The Radio Refractive Index: its Formula and Refractivity Data, International Telecommunications Union.

Jain, A. K. and Mao, J. (1996) Artificial Neural Networks: A Tutorial. *IEEE 0018-9162/96.*

Katz, R. W., and Murphy, A. H. (1997) *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, 222.

Kidder, S. Q. and VonderHaar, T. H. (1995) *Satellite Meteorology, An Introduction*. Academic Press, San Diego, CA, 466.

Kursinski, E. R. (1997) The GPS Radio Occultation Concept: Theoretical Performance and Initial Results, PhD Thesis. California Institute of Technology.

Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P. and Hardy, K. R. (1997) Observing The Earth's Atmosphere with Radio Occultation Measurements Using the Global Positioning System. *Journal of Geophysical Research* **102,** 23429-23465.

Lawrence, M. G. (2005) The Relationship between Relative Humidity and the Dew Point Temperature in Moist Air: A Simple Conversion and Applications. *Bulletin of The American Meteorological Society* **86**, 225-233.

Lazo, J. K., Moss, R. E. and Demuth, J. L. (2009) 300 Billion Served: Sources, Perceptions, Uses and Values of Weather Forecasts. *Bulletin of the American Meteorological Society* **90**, 785-798.

Leke, L., Shifeng, K., Yerong, Z. and Zhenwei, Z. (2008) Tropospheric Refractivity Profiling Based on Single Ground-based GPS, ICMMT 2008 Proceedings.

Lieck, A. (1990) *GPS Satellite Surveying*. John Wiley and Sons, New York.

Lindal, G. F. (1987) The Atmosphere of Jupiter: An Analysis of the Voyager Radio Occultation Measurements. *Journal of Geophysics Research* **86,** 8721-8727.

Lowry, A. R., Rocken, C., Sokolovskiy, S. V. and Anderson, K. D. (2002) Vertical Profiling of Atmospheric Refractivity from Ground-based GPS. *Radio Science* **37(3),** 13.1-13.10.

Malm A. (2009) *GPS and Mobile Handsets*. Berg Insight AB, Gothenburg.

Marchuk, G. I. (1974) *Numerical Methods in Weather Prediction*. Academic Press, London.

Marshall, J. Le, Jung, J., Derber, J., Chahine, M., Treadon, R.,Lord, S. J., Goldberg, M., Wolf, W., Liu, H. C., Joiner, J., Woollen, J., Todling, R., Van Delst, P. and Tahara, Y. (2005) Improving Global Analysis and Forecasting with AIRS. *Bulletin of the American Meteorological Society* **87,** 891-894.

McKnight, T. L. and Darrel, H. (2000) Climate Zones and Types. In: *Physical Geography: A Landscape Appreciation*, Prentice Hall, Upper Saddle River, NJ.

Mockler, S. B. (1995) *Water Vapor in the Climate System Special Report*. American Geophysical Union. Washington, DC.

Nash, J., Oakley T., Vömel H. and Wei L. (2011) *WMO Intercomparison of High Quality Radiosonde Systems,* Yangjiang, China,

Park, S. K. and Droegemeier, K. (2000) Sensitivity Analysis of a 3D Convective Storm: Implications for a Variational Data Assimilation and Forecast Error. *Monthly Weather Review,* **12,** 140-159.

Peixoto, J. P. and Oort, A. H. (1996) The Climatology of Relative Humidity in the Atmosphere. *Journal of Climate* **9,** 3444.

Saastamoinen, J. (1972) Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging of Satellites. In: *The Use of Artificial Satellites for Geodesy*, Geophysical Monograph Series **15**, 247-251.

Schuman, F. G. (1978) Numerical Weather Prediction. *Bulletin of the American Meteorological Society* **59**, p5-17.

Sivira, R. G. F., Brogniez, H., Mallet, C., and Oussar, Y. (2012) Tropospheric Relative Humidity Profile Statistical Retrievals and their Confidence Interval from Megha-Tropiques Measurements. *9th International Symposium on Tropospheric Profiling (ISTP),* L'Aquila, Italy.

Smith, E. K. and Weintraub, S. (1953) The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies. *Proceedings of the IRE* **41**, 1035-1037.

Solheim, F. S., Rocken, C., Vivekanandan, J. and Ware, R. H. (1999) Propagation Delays Induced in GPS Signals by Dry Air, Water Vapour, Hydrometeors and Other Particulates, *Journal of Geophysical Research* **104,** 9663-9670.

Spilker, J. (1978) GPS Signal Structure and Performance Characteristics. *Journal of the Institute of Navigation* **25,** 121-146.

Szegedy, C., Toshev, A. and Erhan, D. (2013) Deep Neural Networks for Object Detection. *Advances in Neural Information Processing Systems* ***2013***, 2553-2561.

Thayer, D. (1974) An Improved Equation for the Radio Refractive Index of Air, *Radio Science* **9**, p803-807.

Tralli, D. M. and Lichten, S. M. (1990) Stochastic Estimation of Tropospheric Path Delays in Global Positioning System Geodetic Measurements, *Bull. Geod* **64,** 127-159.

Ulaby, F. T., Moore, R. K. and Fung, A. K. (1981) Microwave Remote Sensing. Addison-Wesley, Reading, MA, 344-412.

National Oceanic and Atmospheric Administration (1997) *Federal Meteorological Handbook No. 3, Rawinsonde and Pibal Observations*. Office of the Federal Coordinator for Meteorological Services and Supporting Research.

Vaisala Radiosonde RS92-D Datasheet (2010) Vaisala.

Vorobev, V. V. and Krassil'nikova, T. G. (1994) Estimation of the Accuracy of the Atmospheric Refractive Index Recovery from Doppler Shift Measurements as Frequencies used in the NAVSTAR System. *Izvestiya Atmospheric and Oceanic Physics* **29,** 602-609.

World Meteorological Organisation (2005) The Global Observing Systems: 20 Second Status Report on the Implementation of the World Weather Watch. *WMO-No 986*.

Internet References

Artinaid. Earth's Atmosphere: The Tropopause and the Troposphere. Accessed on 6 Jan 2014. http://www.artinaid.com/2013/04/earths-atmosphere-the-tropopause-and-the-troposphere/

Geerts, B and Linacre, E. (1997) The Height of the Tropopause. Accessed on 11 Feb 2015. http://www-das.uwyo.edu/~geerts/cwx/notes/chap01/tropo.html

Georgia Tech Research Institute Case Study (2010) Better Weather: Water Vapour Profiling System Could Result in More Accurate Less Costly Forecasting. Accessed on 17 Dec 2013. http://www.gtri.gatech.edu/casestudy/weather-water-vapor-profiling-lidar

Google Maps. Accessed on 15 Dec 2013. https://www.google.com/maps/@-1.2640159,36.8297848,13z

GPS.GOV. GPS Accuracy. Accessed on 11 Jul 2014. http://www.gps.gov/systems/gps/performance/accuracy/

European Organisation for the Exploitation of Meteorological Satellites. Meteosat 0 degree Tropospheric Humidity. Accessed on 2 Nov 2013. http://oiswww.eumetsat.org/IPPS/html/MSG/PRODUCTS/TH/

European Organisation for the Exploitation of Meteorological Satellites. Meteosat Design. Accessed on 2 Dec 2013. http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Meteosat/MeteosatDesign/index.html

Official US Government Information about the Global Positioning System. GPS Constellation, Accessed on 16 Sep 2013. http://www.gps.gov/multimedia/images/constellation.jpg

International GNSS Service. The IGS Tracking Network, Accessed on 30 Aug 2013. http://igscb.jpl.nasa.gov/network/complete.html

National Climatic Data Centre, Greenhouse Gases, Accessed on 2 Dec 2013, http://www.ncdc.noaa.gov/cmb-faq/greenhouse-gases.php

National Oceanic and Atmospheric Administration. Composition of the Atmosphere, Accessed on 20 Dec 2013, http://esrl.noaa.gov/gsd/outreach/education/climgraph/docs/cg_8.pdf

National Oceanic and Atmospheric Administration. Radiosonde System Design, Accessed on 3 Sep 2013, http://www.nws.noaa.gov/rrs/system.htm

US National Research Laboratory, Monterrey, Meteo7 WV Indian-Southern, Accessed on 30 Aug 2013, http://www.nrlmry.navy.mil/tropics-bin/tropics.cgi?REGION=indian&SECTOR=southern&DISPLAY=Single&AGE=Archive&SIZE=Full&PRODUCT=vapor&PATH=/atlantic/tropics/vapor&SHOW=Image&ARROW=next&CURRENT=20130829.2115.goes13.vapor.x.atl_tropics.x.jpg&TITLE=Region/Sector