

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

Forecasting the Distribution of Water Vapour Observed By GPS Radio Occultations and Comparison with the Results Obtained From Radiosonde Data Over Indian Regions (8° - 38° N and 66° - 98° E) Using Data Mining Applications

Venkata Sheshanna Kongara¹, Dr. D.Punyaseshudu²

Research Scholar, Dept. of Computer Science & Technology, Sri Krishnadevaraya University, Anantapuramu, Andhra Pradesh, India¹

Professor, Dept. of Physics, Rayalaseema University, Kurnool, Andhra Pradesh, India²

ABSTRACT: Atmospheric water vapour plays an important role in radio communications for both terrestrial and earth space communication systems as well as for global climate change studies. Hence the water vapour data and its information plays vital role for scientists and researchers to evaluate the hidden patterns, trends for analysis and forecasting. In general the balloon-borne radiosonde observations are the primary source of upper air water vapour data and are being used for estimation of different atmospheric parameters over the last few decades by using conventional methods. However Global Positioning System Radio Occultation (GPS RO) is a comparatively new technique for retrieving the properties of the atmospheric conditions that radio signals transmitted by the Global Positioning System (GPS) for obtaining the various atmospheric parameters like vertical profiles of bending angles, refractivity, electron density, temperature, pressure, and water vapour in the atmosphere for monitoring the weather conditions and climate change studies. Hence as part of study considered the most recent datasets of RO water vapour parameter for Indian region (latitude of 8° to 38° N, and longitude of 66° to 98° E) by FORMOSAT-3/COSMIC (Formosa Satellite mission-3/Constellation Observing System for Meteorology, Ionosphere, and Climate) data distribution measurements data. And also obtained the most recent version of the meteorological radiosonde data from British Atmospheric Data Centre (BDAC) with various radiosonde observations are integrated and analyzed from the different met parameters of air temperature, dew point temperature and pressure levels for water vapour concentrations estimation at different heights. As part of the study we have assessed and forecasting the distribution of water vapour observed by COSMIC GPS Radio Occultations and comparisons with the results obtained from radiosonde data distribution. The detailed comparisons have been accomplished and found that the radiosonde data accessibility is a big gap over oceans, however the GPS RO data available for most of the land and ocean occurrences, consequently it illustrates an advantage of on cyclone forecast and other oceanic based exploration. In conclusion our studies gain that with the COSMIC water vapour retrievals obtained from GPS RO are providing reliable data and good quality for detailed studies and investigation of the atmosphere than the radiosonde data for advanced analysis and decision supporting systems.

KEYWORDS: Atmospheric Water Vapour, Radiosonde, Radio Communications, Climate Change, Hidden Patterns, Global Positioning System, Radio occultation, Decision Supporting Systems.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

I. INTRODUCTION

1.1. Importance of the GPS Radio Occultation and radiosonde water vapour and objectives of the present work

The Water vapor (WV) is the most significant greenhouse gas and plays an important role in radio communications for both terrestrial and earth space communication systems as well as for global climate change studies and weather forecasting. The Water Vapour and Oxygen needs to be considered in the effective estimation performance of earth space communication connections. In the clear air the Water Vapour and oxygen absorption causes frequency dependant signal attenuation, propagation delay, ray bending and medium noise. Therefore it is an essential area of practice and research about Atmospheric water vapour [1] [2]; In general the balloon-borne radiosonde observations are the primary source of upper air water vapour data and are being used for estimation of different atmospheric parameters over the last few decades by using conventional methods. However Global Positioning System Radio Occultation (GPS RO) is a comparatively new technique for retrieving the properties of the atmospheric conditions that radio signals transmitted by the Global Positioning System (GPS) for obtaining the various atmospheric parameters like vertical profiles of bending angles, refractivity, electron density, temperature, pressure, and water vapor in the atmosphere for monitoring the weather conditions and climate change studies. Hence as part of study considered the datasets for the year 2014 with GPS RO water vapour parameter for Indian region (latitude of 8° to 38°N, and longitude of 66° to 98°E) by FORMOSAT-3/COSMIC (Formosa Satellite mission-3/Constellation Observing System for Meteorology, Ionosphere, and Climate) data distribution measurements data based on the availability. And also obtained the meteorological radiosonde data for year 2014 from British Atmospheric Data Centre (BDAC) with various radiosonde observations are integrated and analysed from the different met parameters of air temperature, dew point temperature and pressure levels for water vapour concentrations estimation at different Mean Sea Level geometric height in kilo meters using Artificial Neural network (ANN) approaches in MATLAB for better analysis and review. The detailed comparisons have been accomplished and the radiosonde data accessibility is a big gap in oceans, however the GPS RO data available for most of the land and ocean occurrences, hence it shows an advantage of on cyclone forecast. In conclusion our studies gain that with the COSMIC water vapour retrievals obtained from GPS RO are providing reliable data and good quality for detailed studies and investigation of the atmosphere than the radiosonde data for advanced analysis and decision supporting systems.

The rest of the paper is organized as follows. The literature review and related work for GPS RO and Radiosonde water vapour studies are discussed in section II. The details of the GPS RO and Radiosonde water vapour data evaluation, processing, integration and classification are described in Section III. The data analysis and results are presented in section IV and concluded the paper with accomplishment in the study and future work in Section V.

II. RELATED WORK

Folorunsho Olaiya et al [3] investigated the use of data mining techniques in forecasting maximum temperature, rainfall, evaporation and wind speed. This was carried out using Artificial Neural Network and Decision Tree algorithms and meteorological data. A data model for the meteorological data was developed and this was used to train the classifier algorithms. The performances of these algorithms were compared using standard performance metrics, and the algorithm which gave the best results used to generate classification rules for the mean weather variables. A predictive Neural Network model was also developed for the weather prediction program and the results compared with actual weather data for the predicted periods

Abhineet Shyam et al [4] Artificial neural network (ANN) technique have been used to derive water vapour pressure profiles in the troposphere from radio occultation data over India and adjoining region. Fully connected 3 layer networks with 1 hidden layer have been constructed and standard back propagation algorithm is used to train the network. While month, latitude and vertical profiles of refractivity/bending angle constitute the input vector. The water vapour partial pressure profiles forms the output vector. Only the moisture-laden summer monsoon months of June, July, Aug, Sept of 2010 have been consider for developing the retrieval algorithm. There are 2120 input and

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

output pairs, out of which 1696 pairs from the training set and remaining pairs constitute the validation set. The retrieved profiles of water vapour pressure in the validation set are compared with the corresponding COSMIC operational products of water vapour pressure profiles. The effectiveness of the algorithm is apparent from this comparison and also from the vertical profiles of bias and the root mean square error (RMSE). The statistics show that the better performance of the algorithm and the retrieved data has found to be highly consistent.

Stefania Bonafoni et al [5] conducted study on different approaches based on multilayer perceptron neural networks and proposed a new retrieval algorithm to derive profiles of atmospheric parameters from RO refractivity profiles overcoming the requirement for temperature profile availability at each GPS occultation. They have considered dataset of 445 occultations covering the land surface within the Tropics, Split into desert and vegetation zone. The neural networks were trained with refractivity profiles as input computed from geometrical occultation parameters provided by the FORMOSAT-3/COSMIC Satellites, while the targets were the dry and wet refractivity profiles and the dry pressure profiles obtained from the contemporary European Centre from Medium-Range Weather Forecast data.

P. Kishore et al [6] studied about global (50 S–50 N) distribution of water vapor and investigated using COSMIC GPS RO measurements for the period of September 2006–December 2009, As a first step towards building the COSMIC water vapor climatology, reasonable global maps of seasonal mean of specific humidity at different pressure levels. The detailed comparisons have been made between COSMIC and high resolution GPS radiosonde measurements across 13 tropical stations and model outputs (ERA-Interim, NCEP, and JRA-25 reanalyses datasets). In comparison within dependent techniques like radiosonde (Vaisala) and it is found that COSMIC GPS RO wet profiles are accurate up to 7–8km (assuming radiosonde as standard technique).

R.Norman et al [7] post processed radio occultation data from CDAAC at UCAR and provided by the COSMIC/FORMOSAT-3 were compared to co-located radiosonde data from Antarctica for the period from day 194 in 2006 to 2010. The robustness and accuracy of the COSMIC GPS radio occultation were examined and GPS RO temperature and Pressure profiles were compared with a spatial and temporal buffer of 200 km and 3hrs. The results presented clearly demonstrate that the GPS RO technique provides a very useful meteorological data type for the data sparse Antarctic region.

Fabrizio Pelliccia et al [8] discussed about a procedure for retrieval of tropospheric profiles (temperature, pressure, humidity) using only refractivity profiles coming from Global Positioning Systems' Low-Earth-Orbit radio occultations, without the constraint of independent knowledge of atmospheric parameters at each GPS occultation. In order to achieve this goal, they have used an approach based on neural networks (NN), exploiting a data set of 1106 occultations collected over the Arctic region during the winter season of 2007 and 2008. The total refractivity (N) profiles from FORMOSAT-3/COSMIC satellites have been used as input for training the Network and the targets were derived as dry and wet components based on the information provided by the analysis of the European Centre for Medium Range Weather Forecast (ECMWF). Once trained the NN they have computed the other parameters (temperature, pressure and vapor); and then identified relatively co-located ECMWF data. Finally comparative analysis carried out with the radiosonde observations (RAOB); the comparison shows good agreement of both NN-derived and CDAAC profiles.

F.Ladstadter et al [9] conducted studies on Climate comparison of GPS radio occultation and vaisala radiosonde, They have used 12 years (from 2002 to 2013) of GPSRO measurements, co-located radiosonde soundings from vaisala RS90/92 sensors, and 5 years of RS92 GRUAN data products to rigorously intercompare temperature and humidity data for these 3 upper air observational datasets. They have presented various analytical results in detailed and expressed that GPS RO provides valuable long term stable reference observations with well defined error characteristics for climate applications and for anchoring other upper air measurements. However they have concluded that overall very good agreement is demonstrated by this multi-year inter-comparison of GPS RO and Radiosonde reference data sets. This strongly encourages further development and advancement of both systems for the benefit of future climate monitoring and research.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

XU Xiaohua et al [10] conducted studies on the RO refractivity profiles obtained from the COSMIC Mission and statistically compared with the observations of 38 radiosonde stations provided by the Australian Bureau of Meteorology during the period from 15 July 2006 through 15 July 2007. Initially they have compared different collection criteria from Low/middle and latitude regions; and then the COSMIC RO soundings that occur within 3 hours and 300 KM of radiosonde measurements are used for their studies and the final statistical comparisons. After the review the overall results shows that the agreements between the COSMIC refractivity profiles and the radiosonde soundings from the 38 stations are very good at 0-30 KM altitude, The comparison of COSMIC RO refractivity profiles and radiosonde observations for 3 polar stations in four different seasons indicate that the accuracy of GPS RO profiles is better in the Australian summer and autumn in the Austral spring and winter during the year from Sept 2006 to Aug 2007.

Shu-peng Ho et al [11] compared specific humidity profiles derived from COSMIC RO from August to November 2006 with those from different types of radiosonde and ECMWF global analysis. A comparison shows that COSMIC specific humidity data agree well with ECMWF analysis over different regions of the world for both day and night times. On the Country, evaluation against COSMIC specific humidity shows a distinct dry bias of Shang-E radiosonde (China) and an obvious wet bias of VIZ-type (USA). No Obvious specific humidity biases are found for MRZ (Russia) and MEISEI (Japan) radiosondes. Those results demonstrate the usefulness of COSMIC water vapour for quantifying the dry/wet biases among different sensor types.

S. K. A. V. Prasad Rao Anisetty et al [12] present an encouraging example of prediction of super cyclone Gonu over the northern Indian Ocean in 2007. A series of experiments are conducted using the advanced Weather Research and Forecasting model with three-dimensional variational method to assimilate GPS RO refractivity from FORMOSAT-3/COSMIC (hereafter referred as GPS) and radiosonde sounding (GTS) to highlight the relative impact of GPS RO data on model prediction.

III. METHODOLOGY

3.1. Source of the data:

Radio Occultation Data:

FORMOSAT-3/COSMIC uses a constellation of six remote sensing micro-satellites to collect atmospheric data to demonstrate their uses and applications in weather prediction, ionosphere, climate and gravity research. The mission sponsors encourage the use of these data in a wide variety of applications and research projects. The datasets are managed and produced by CDAAC (COSMIC Data Analysis and Archival Center). As part of the study we have downloaded the GPS Radio Occultations level 2 product profiles from the CDAAC website [13]. The datasets are made available in multidimensional data formats of NetCDF (Network Common Data Form) covering the period from 1996-2015 and the wet atmospheric profiles (wetprf) of refractivity, temperature and water vapor pressure with altitude range is 0-40 km at 100 meter vertical resolution, these profiles are based on 1D variational analysis using ECMWF low resolution analysis data. We have considered the recent radio occultation datasets for various positions of Indian region (latitudes between 8° to 38°N, and longitudes 66° to 98°E) at surface level for 2014 with different mean Sea Level geometric height in Kilo meters and Pressure levels in Mb, the COSMIC data availability for Indian region latitude, 8° to 38° north and longitude, 66° to 98° east from 1 to 365 Days of 2014. The Figure.1 represents the radio occultation data availability for the specified period and the spots highlighted in red are the different radio occultation's atmospheric profiles for different latitude and longitudes over India.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

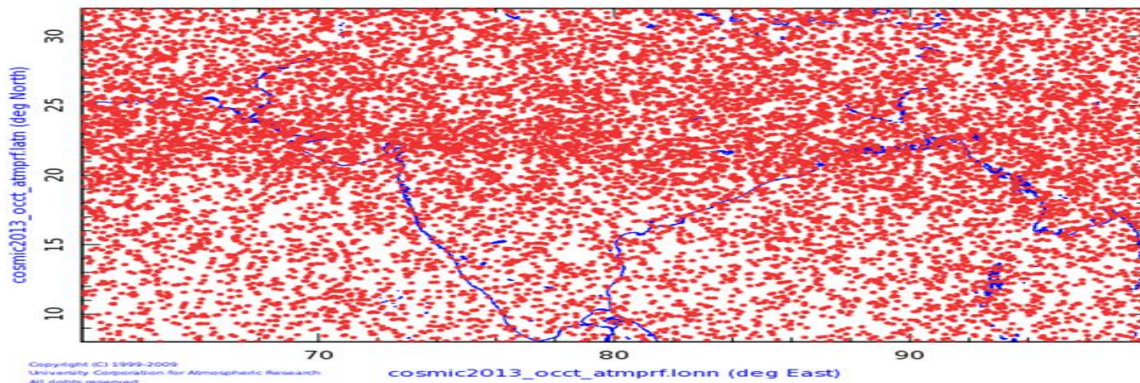


Figure.1: COSMIC data availability in the year 2014 for Indian region (latitude 8° to 38°N and longitude 66° to 98° E) and the highlighted red spots are the radio occultation's atmospheric profiles.

Radiosonde data:

The global atmospheric radiosonde data has millions of soundings from worldwide upper air stations over the period 1996 to present. Each sounding is composed of records of pressure, geo-potential height, temperature, dew-point temperature, wind-speed and wind-direction at standard and significant pressure levels proposed by World meteorological Organization (WMO) from the surface to approximately 20-30 km. This data was extracted from the daily occurrences of atmospheric soundings compiled by British Atmospheric Data Centre (BDAC) and is made available for individuals and organizations for atmospheric studies[14]. The radiosonde takes measurements at intervals of approximately 2 seconds. The high resolution data files contain all such data. The standard resolution data files contain measurements taken at particular levels of the atmosphere. We obtained different stations details, air pressure, air temperature, dew point temperature, wind speed, wind direction and various other parameters for the time period for 2014 over India. Further based on the data availability we have selected the necessary stations for various geographical locations and data cleansing activity is conducted to focus on the target data. The Pressure levels are considered at the surface levels while validating the altitude in Meters to identify Surface Pressure for various stations based on the station height. Extracted the rules and map them into different dimensions and measures to produce understandable and useful knowledge for analysis. Figure.2 represents the global radiosonde stations which are highlighted in yellow, however in this study we have considered set of radiosonde stations over India based on the data availability.

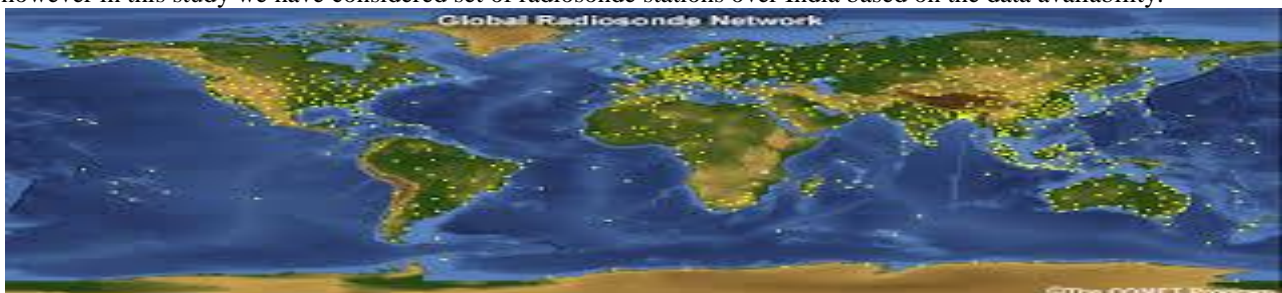


Figure.2: The Global radiosonde network by "The COMET Program": The highlighted yellow spots are the different radiosonde stations.

3.2 Radio Occultation, Radiosonde data assessment and Integrated Water Vapour Concentration:

The GPS Radio Occultation (RO) and Radiosonde (RS) data assessment and computation of integrated water vapour have been carried by using different functions, B.-R. Wang et al [15] discussed on the atmospheric refractivity of radio occultation was calculated using function (1)

$$N = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{e}{T^2} \quad (1)$$

Where N Refractivity Index, T is temperature in Kelvin and P is total air pressure and e is the water vapour pressure hpa, respectively. Hence as part of the study we have assessed the obtained radio occultation data; and considered water vapour pressure and temperature parameters to compute the Integrated Water Vapour Concentration (M) for further

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

analysis. Khamphoui et al [16] discussed on the impact of water vapor in atmosphere and clear sky attenuation is mainly due to the absorption caused by water vapour and oxygen molecules. It increases with the relative humidity as well as the temperature and alleged that the water vapour concentration is strong function temperature and humidity. In their studies they have used most revised version of the function (2) to calculate the water vapour concentration (M) as specified below.

Water Vapour Concentration (M) in $g/m^3 = 216.7 \times \frac{e}{T}$ (2) where e is water vapour pressure hPa and the T is temperature in Kelvin.

Hence we are using this function our calculation and the calculated water vapour concentration datasets are arranged in the dimensional tables of the data warehouse for our analyses, which are may be reusable for the further studies. We have analyzed all the datasets with various analytics, the detailed results are presented in section IV.

3.3. GPS RO Water Vapour Data Integration and Classification:

Data pre-processing is a fundamental stage of data analysis for better performance and good quality of results. In general missing data are unintended and uncontrolled by the researchers, but the overall result is that the observed data cannot be analysed because of the incompleteness of the data sets. Missing Values and its problems are very common in the data cleaning process. Several methods have been proposed [17] [18] so as to process missing data in datasets and avoid problems caused by it. In the Radiosonde and COSMIC GPS RO source data also there are some missing values for different parameters for different MSL Altitude, Pressure, Latitude and longitudes. Both are filled those missing data sets with 999 in their data, Hence we have handled those missing values filled with précised one by using cluster based missing values algorithm to process the water vapour data for better results.

Building effective classification systems is one of the central tasks of data mining. Many different types of classification techniques are available that includes Decision Trees, Naive-Bayesian methods, Neural Networks, Logistic Regression. We conducted a comparative review for various methods of classification and measured the accuracy of the results. It show that using Neural Networks obtains the best result among the other methods therefore in the water vapour studies we have used improved version of neural networks algorithm to classify the water vapour data for better classification and prediction.

India is a country of great geographical extent with a diversity and rich cultural heritage. The physical features of India are not same across the country. They are varying from place to place with significant features which are highest mountains in the north, wide plains in the centre, desert region in the northwest, narrowing peninsula to the south, mountain ranges to the northeast and a lower range running along the west coast of the peninsula. The vast expanse of the Indian Ocean lies to the south of the land mass with the Bay of Bengal and Arabian Sea pushing north engulfing the peninsula [19], Following 1.The Kashmir Region, 2.The Assam Region, 3.The Northern Plains, 4.The Central Highlands, 5.The Desert Region, 6.The Peninsular Plateau, 7.The Eastern Coastal Plains, 8.The Western Coastal Plains and 9.The Island are the different Physiographic divisions and classification zones are being considered for our studies.

3.4. Artificial Neural Network Models:

As acknowledged in the Introduction of this paper data mining Artificial Neural Network applications are the best suitable for atmospheric water vapour studies. The nervous system as a network of cells specialized for the reception, integration and transmission of information. It comprises the brain and spinal cord, which is the central nervous system (CNS) and sensory and motor nerve fibres that enter and leave the CNS or are wholly outside the CNS. The Fundamental unit of the nervous system is the neuron [22]. The neurons are specialized to receive information from the internal and external environment and capable to transmit signals to other neurons, in Figure.3 presented the biological aspects of the ANN model. The dendrites of the cell body receive the input from the other neurons and the signals are transmitted down the Axon, then the terminal branches again send the output to the other neurons which is a recurring process in the nervous system, with that biological inspiration [McCulloch and Pitts] proposed a artificial neural network (ANN) mathematical model based on biological neural networks, in other words, is an emulation of biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

external or internal information that flows through the network during the learning phase. In the Figure.4: represented the components of simple neuron with set of input nodes ($x_1, x_2, x_3, \dots, x_n$) which are having the data to be evaluated and ($w_1, w_2, w_3, \dots, w_n$) are its associated weights and set of output nodes (y_1, y_2, \dots, y_n); to process each input value is multiplied by associated weights and apply the summation to get the single value, then pass through the nonlinear activation function called sigmoid activation for $\text{logsig}(n) = 1 / (1 + \exp(-n))$ for data normalization and control the output values.

The ANN architecture with different layers shown in Figure.5 is the most commonly used structure. The ANN is presented in three layers, namely the input layer, hidden layer, and the output layer. Every layer consists of one or more nodes represented in the diagram by the small circles. The lines starting from input nodes indicate that flow of information from one node to the next node. In this particular type of ANN, the information flows only from the input layers to the output layer. However other types of ANN have more complex connections with feedback paths between outputs to the input layers. In general the back propagation algorithm with various training parameters are used in layered feed forward ANNs, hence that the artificial neurons are organized in layers, and send their signals “forward”, and then the errors are propagated backwards to calculate the difference between actual and expected results.

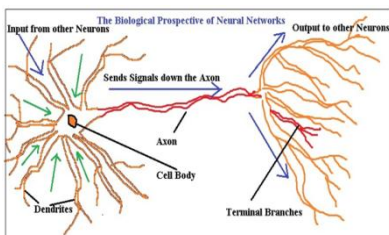


Figure.3:

Figure.3: The Biological aspects of the ANN model: The dendrites receive inputs from other neurons, axon transmits the signals and the terminal branches send the output to other neurons.

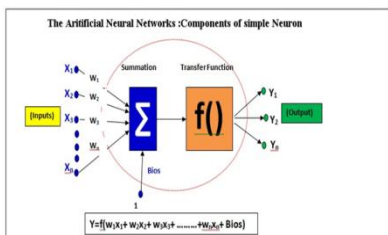


Figure.4

Figure.4: Components of simple neuron in the Artificial Neural Network: with set of input nodes and associated weights pass through the nonlinear sigmoid activation function for data normalization and control the output values.

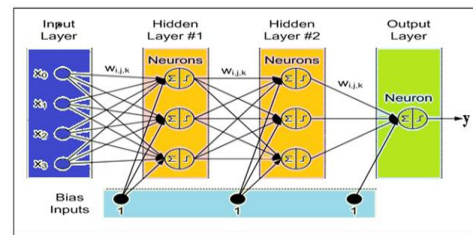


Figure.5

Figure.5: Different Layers of the ANN architecture: The ANN is presented in three layers, namely the input layer, hidden layer, and the output layer.

Building a Water Vapour Distribution Forecasting Neural Network Model (WVDFNN) using Gradient Descent Training:

In our earlier studies we have establish that improved version of neural networks Back Propagation algorithm of gradient descent with momentum and adaptive learning rate back propagation method [20]. Hence in the same way the neural network training was done using MATLAB R2015a applications with effective forecasting of the GPS RO water vapour distribution for different physiographic divisions and variations in prospect trends, which are presented in section IV Results and Analysis.

Build the Network with different parameters and training:

We have build the neural network model using MATLAB with different parameters and training, validation and testing of the GPS RO water vapour data. While training a network we have used the feed forward back propagation method with the log-sigmoid transfer function for data normalization and control the output values and provided the set of seasonal input vectors as showing in Figure.6 for the output seasonal forecasting, preferred the training function gradient descent with momentum and adaptive learning rate back propagation, and used the mean squared normalized error performance function for error checking between output and the input vectors . The training error performance is validated for different seasons as shown in Figure.7 for 1000 epochs the changing curve of error in different fitting options of gradients, validation checks and the learning rates at epoch 1000. Training performance as different hidden neurons chose for training network for the best validation out of 1000 epochs, while training the epochs progress with in 1000 iterations and validated the best fit for the stopping criteria; the model has a good training performance then all the plots have a high value correlation with better than 0.98 to 0.99 where a regression value is 1 as presented in

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

Figure.8 for training, validation and testing with the target values. In this scenario the outputs of the training network are quite close to the targets. Hence the detailed analytical results with the actual radio occultation data and the ANN based forecasted values are presented in the section IV.

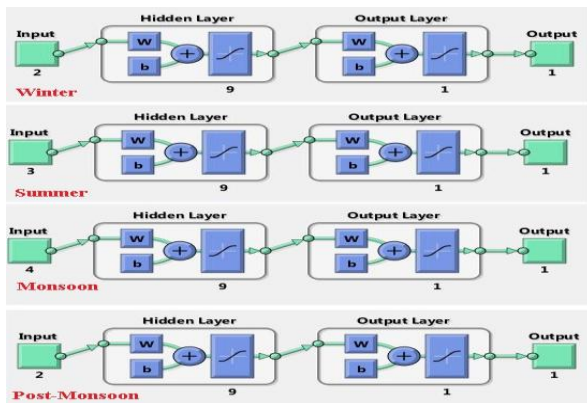


Figure.6

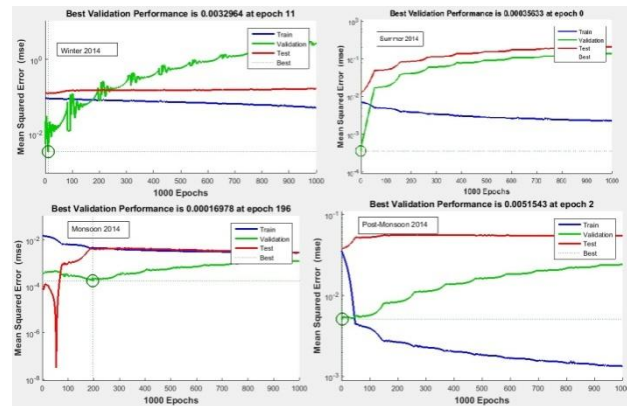


Figure.7

Figure.6: Different seasonal input and output Vectors for neural networks build and training.

Figure.7: The training error performance as different hidden neurons chosen for training network for the best validation out of 1000 epochs

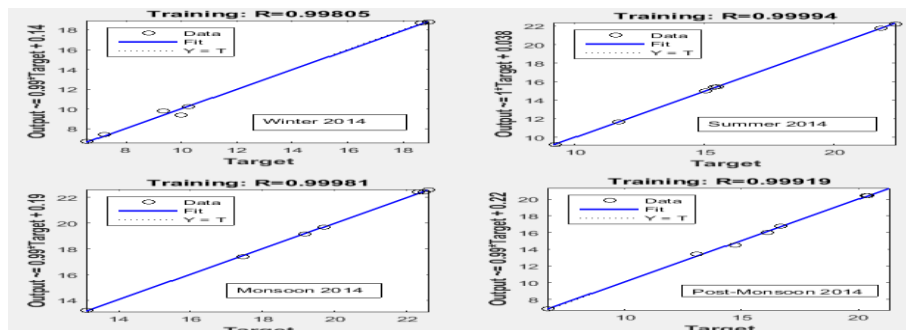


Figure 8: shows that network training regression R Values measure the correlation between Input and targets of GPS RO Data Distribution over India

IV. ANALYSIS AND RESULTS

The Concentration of water vapour, the third most abundant atmospheric gas, is highly variable, mainly because the maximum possible vapour pressure is a strong function of temperature. Once the air is saturated with water vapour, any reduction in temperature will cause condensation and No other constituent gas changes phase under normal atmospheric conditions [21]. Hence the water vapor is the most significant greenhouse gas and plays an important role in global climate change studies and weather forecasting, though increasing of water vapour leads to warmer temperature, which cause more water vapour to be absorbed into the air, the water vapour feedback can also increase the effective monitoring on the impact of other atmospheric parameters. Understanding the distribution of water vapour is of central importance to the representation of seasonal variability with different observation techniques. The balloon-borne radiosonde observations are the primary source of upper air water vapour data and are being used for estimation of different atmospheric parameters varying with temperature and various pressure levels. The Global Positioning System Radio Occultation (GPS RO) is a comparatively new technique for retrieving the properties of the atmospheric conditions and water vapour studies. These two observational systems are obtaining large volume of data. However atmospheric parameter's data assessment and usage is critical and challenging for different studies. The data

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

warehousing and data mining applications are the most emerging technologies with powerful data managing features for better data processing and analysis for effective decision supporting systems.

The water vapour data analytics provide the details on the seasonal mean water vapour distributions for radiosonde technique as shown in Figure.13 and the radio occultation techniques as shown in Figure.14 from different locations and physiographical regions are illustrates the new precise observational results which are agree with existing assessments of water vapour's impact. And also the forecasting the magnitude of water vapour seasonal distribution from radiosonde and radio occultations technique's feedback for various physiographic divisions as shown in the Figure.9 illustrates the monsoon seasonal water vapour forecasting feedback , Figure.10 illustrates the post monsoon seasonal water vapour forecasting feedback, Figure.11 illustrates the summer seasonal water vapour forecasting feedback, Figure.12 illustrates the summer seasonal water vapour forecasting feedback and the Figure.15 illustrates the annual mean distribution of water vapour feedback for radiosonde, radio occultation techniques along with the forecasting results. The research discloses that the water vapour concentrations at the surface level for various physiographic divisions are strongly dependant on altitude of location from the mean sea geometric height. The comparative results of radiosonde and radio occultation techniques are presented in Figure.15.The Annual mean water vapour distribution for the 1st group Kashmir Region are varying with (~ 8 g/m³),The Annual mean water vapour distribution for the 2nd group, Assam Region is varying with (~ 16 g/m³),The Annual mean water vapour distribution for the 3rd group, Northern Plains are varying with (~ 14 g/m³),The Annual mean water vapour distribution for the 4th group, Central Highlands are varying with (~ 12 g/m³),The Annual mean water vapour distribution for the 5th group, Desert Region are varying with (~ 10 g/m³),The Annual mean water vapour distribution for the 6th group, Peninsular Plateau are varying with (~ 14 g/m³),The Annual mean water vapour distribution for the 7th group, Eastern Coastal Plains are varying with (~ 20 g/m³),The Annual mean water vapour distribution for the 8th group, Western Coastal Plains are varying with (~ 19 g/m³), and The Annual mean water vapour distribution for the 9th group, Island are varying with (~ 18 g/m³) all these results are derived based on the different climate conditions during the period from Jan to Dec in 2014.

The comparisons between seasonal radio occultation and radiosonde water vapour distribution presented in Figure.16: illustrate the monsoon seasonal feedback during the period monsoon months (Jun, Jul, Aug, Sept) in 2014 and the maximum water vapour distribution varies for the Assam Region (~ 22 g/m³ with the variance of 2.72%) and the minimum water vapour distribution varies for the Kashmir Region (~ 12 g/m³ with the variance of 2.91%).

Figure.17: illustrate the post monsoon seasonal feedback during the period post monsoon months (Oct, Nov) in 2014 and the maximum water vapour distribution varies for the Island Region (~ 20 g/m³ with the variance of 9.45%) and the minimum water vapour distribution varies for the Kashmir Region (~ 7 g/m³ with the variance of -18.12%).

Figure.18: illustrate the summer seasonal feedback during the period summer months (Mar, Apr, May) in 2014, and the maximum water vapour distribution varies for the Eastern Coastal Plains Region (~ 21 g/m³ with the variance of 13.81%) and the minimum water vapour distribution varies for the Kashmir Region (~ 7 g/m³ with the variance of -18.99%) .

Figure.19 illustrate the winter seasonal feedback during the period winter months (Jan, Feb) in 2014, and the maximum water vapour distribution varies for the Western Coastal Plains Region (~ 17 g/m³ with the variance of -0.33%) and the minimum water vapour distribution varies for the Kashmir Region (~ 5 g/m³ with the variance of -18.09%) respectively for various physiographic divisions over India. The detailed results shown in the below sections A and B as follows

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

A). Comparisons between Seasonal GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014- CDAAC data Vs. 2014-Forecasting data using ANN over India for different Physiographic Divisions

Monsoon

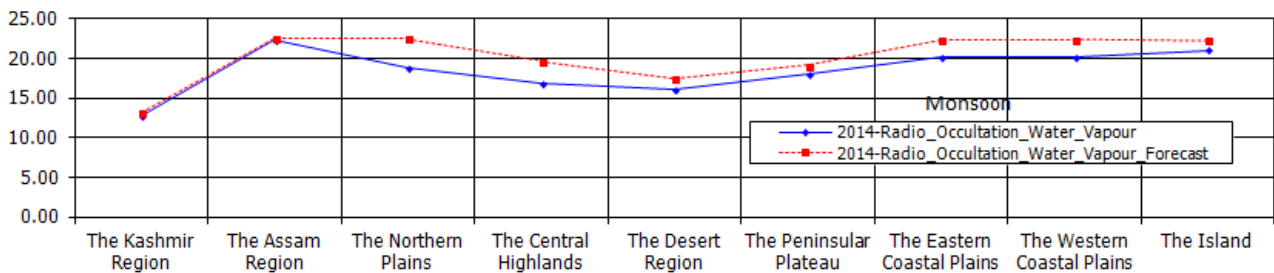


Figure.9: Monsoon Seasonal GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014- Actual Vs. 2014-Forecasting for different Physiographic Divisions Over India

Post Monsoon

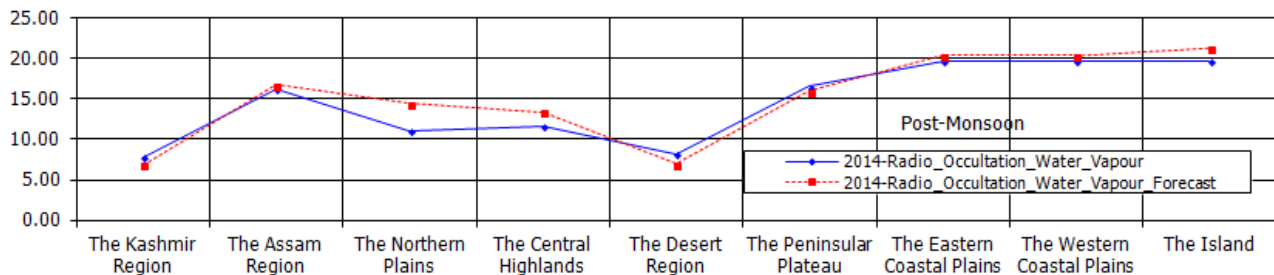


Figure.10: Post-Monsoon Seasonal GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014- Actual Vs. 2014-Forecasting for different Physiographic Divisions Over India

Summer

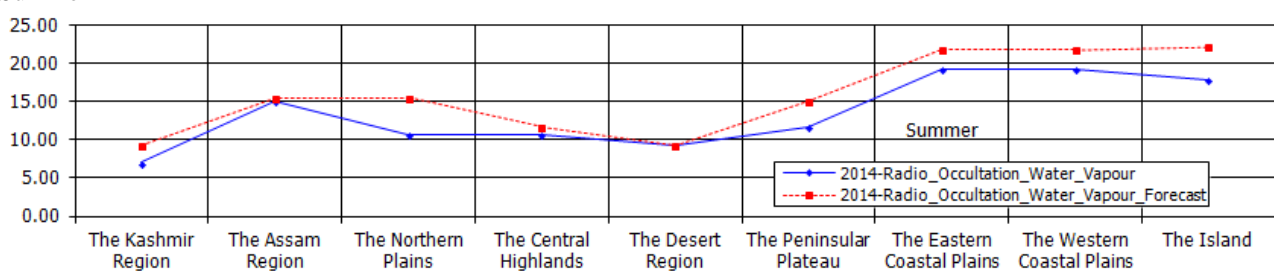


Figure.11: Summer Seasonal GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014- Actual Vs. 2014-Forecasting for different Physiographic Divisions Over India

Winter

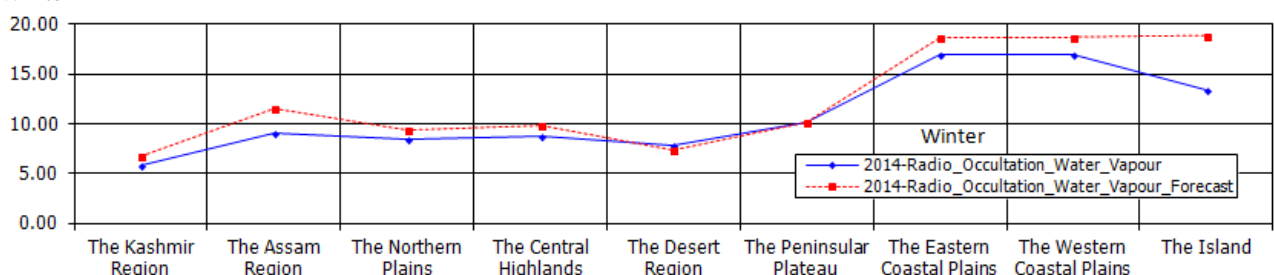


Figure.12: Winter Seasonal GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014- Actual Vs. 2014-Forecasting for different Physiographic Divisions Over India

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

Radiosonde Water Vapour Seasonal Distribution in 2014

Radiosonde Water Vapour Distribution for various Physiographic Divisions over India for 2014

Radiosonde Data for 2014	Monsoon	Post Monsoon	Summer	Winter
The Kashmir Region	13.08	6.65	8.67	4.32
The Assam Region	22.31	17.32	17.42	11.84
The Northern Plains	22.4	15.11	15.44	10.35
The Central Highlands	13.73	13.04	10.31	8.83
The Desert Region	17.43	3.62	3.76	7.22
The Peninsular Plateau	19.21	15.32	14.27	10.83
The Eastern Coastal Plains	22.62	20.4	22.38	17.23
The Western Coastal Plains	22.02	19.3	20.86	16.32
The Island	22.55	21.86	22.2	20.02

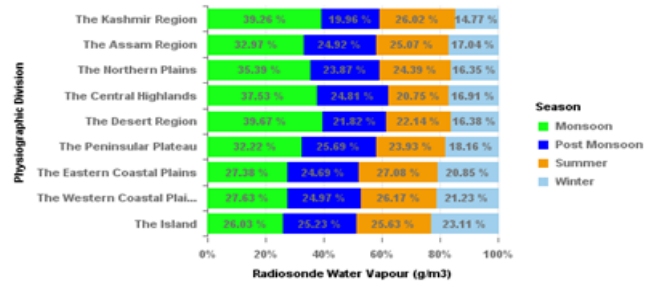


Figure.13: The Seasonal Radiosonde Water Vapour (g/m3) distribution % trend for different Physiographic Divisions over India in 2014

GPS Radio Occultation Water Vapour Seasonal Distribution in 2014

GPS Radio Occultation Water Vapour Distribution for various Physiographic Divisions over India for 2014

Radio Occultation Data for 2014	Monsoon	Post Monsoon	Summer	Winter
The Kashmir Region	12.70	7.86	7.02	5.81
The Assam Region	22.23	16.26	15.05	3.11
The Northern Plains	18.80	11.06	10.76	8.41
The Central Highlands	16.84	11.65	10.68	8.82
The Desert Region	16.08	8.18	3.36	7.73
The Peninsular Plateau	18.06	16.57	11.76	10.17
The Eastern Coastal Plains	20.20	19.72	19.23	16.38
The Western Coastal Plains	20.20	19.72	19.23	16.38
The Island	21.05	19.80	17.33	13.40

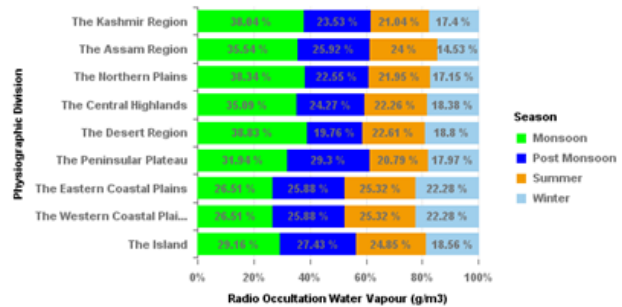


Figure.14: The Seasonal Radio Occultation Water Vapour (g/m3) distribution % trend for different Physiographic Divisions over India in 2014

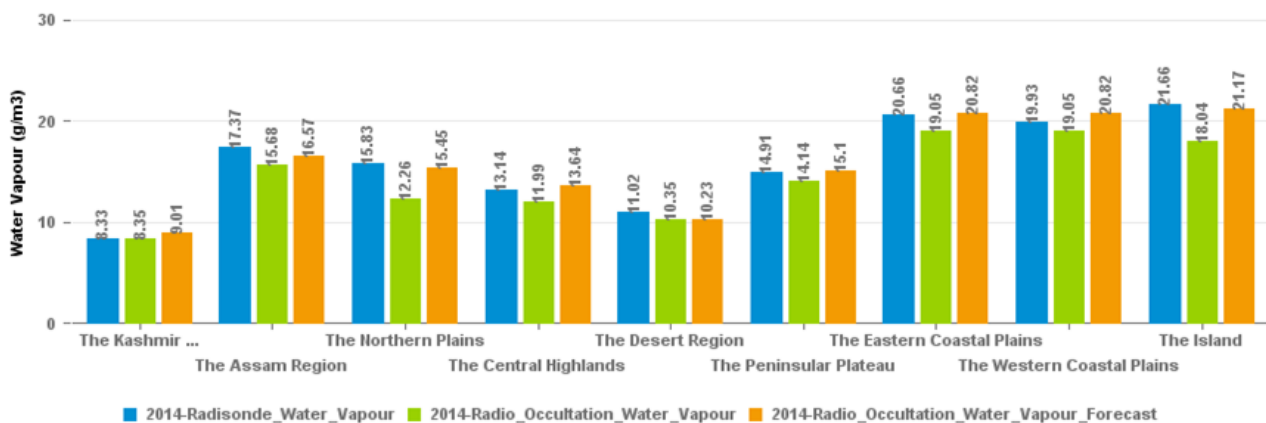


Figure.15: The Annual Radio Occultation Water Vapour (g/m3) distribution trend for different Physiographic Divisions over India in 2014

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

B). Comparisons between GPS Radio Occultations Vs. Radiosonde Water Vapour (g/m³) Distribution Trends for 2014 over India for different Physiographic Divisions.

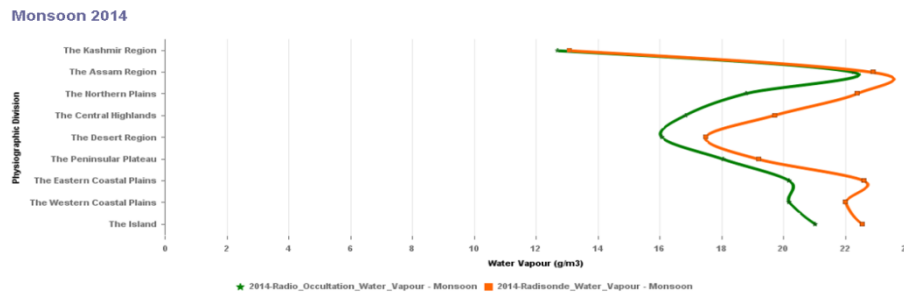


Figure.16: The Monsoon Seasonal Radiosonde and GPS Radio Occultation Water Vapour (g/m³) distribution trend for different Physiographic Divisions over India in 2014

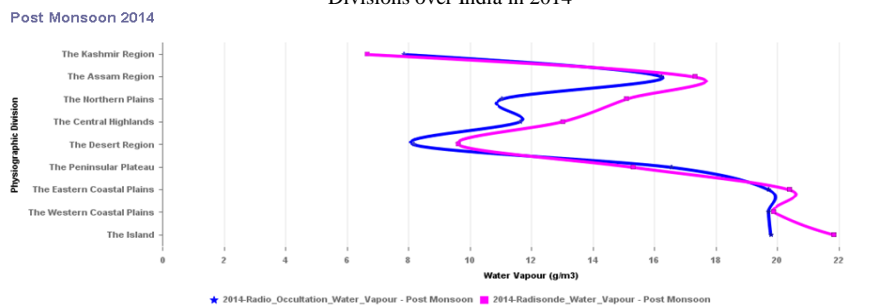


Figure.17: The Post Monsoon Seasonal Radiosonde and GPS Radio Occultation Water Vapour (g/m³) distribution trend for different Physiographic Divisions over India in 2014

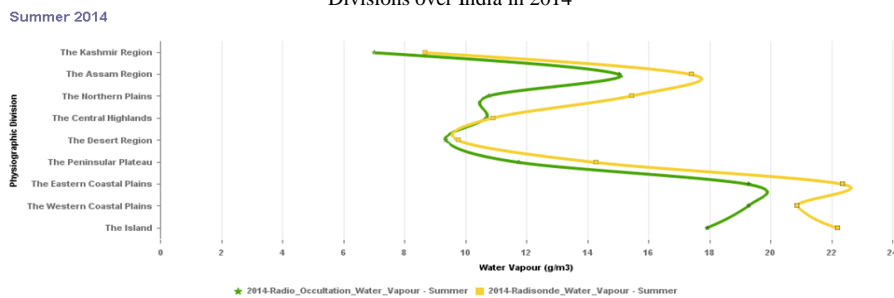


Figure.18: The Summer Seasonal Radiosonde and GPS Radio Occultation Water Vapour (g/m³) distribution trend for different Physiographic Divisions over India in 2014

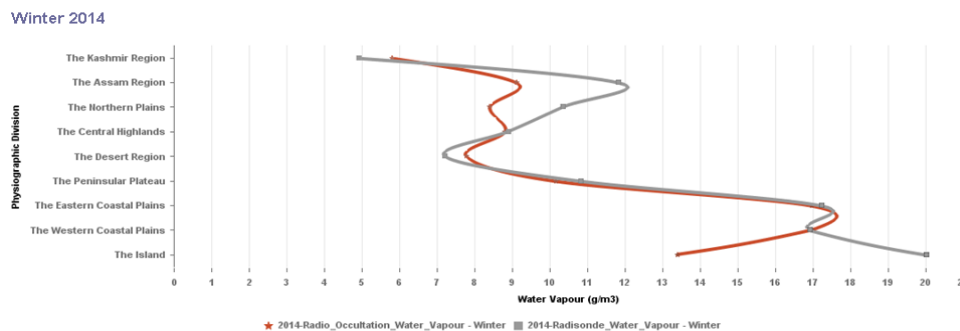


Figure.19: The Winter Seasonal Radiosonde and GPS Radio Occultation Water Vapour (g/m³) distribution trend for different Physiographic Divisions over India in 2014

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

V. CONCLUSION AND FUTURE WORK

As part of the study datasets assimilation and classification techniques are carried out by using data warehousing and data mining applications and also applied better data analytics for precise results on GPS Radio Occultation and Radiosonde water vapour distribution studies over India. The Pre-processed datasets are reusable for further water vapour studies.

Following are the accomplishments.

- Build the GPS Radio Occultations water vapour data based dimensional model for ease of use and analysis, the evaluated data is maintained in the data warehouse for further studies.
- Forecast the GPS Radio Occultation Water Vapour (g/m³) Distribution Trends for 2014 over India for different Physiographic Divisions
- Comparison between GPS Radio Occultations and Radiosonde Water Vapour (g/m³) Distribution Trends for 2014 over India for different Physiographic Divisions.

The detailed comparisons have been accomplished between Radio Occultation and Radiosonde water Vapour Data and found that the radiosonde data accessibility is a big gap over oceans, however the GPS RO data available for most of the land and ocean occurrences, consequently it illustrates an advantage of on cyclone forecast and other oceanic based exploration. In conclusion our studies gain that with the COSMIC water vapour retrievals obtained from GPS RO are providing reliable data and good quality for detailed studies and investigation of the atmosphere than the radiosonde data. In future work we would like to explore various ANN methods for different interdisciplinary research areas for better decision supporting systems

ACKNOWLEDGMENT

- ❖ We wish to thank the British Atmospheric Data Centre (BADC), which is part of the NERC National Centre for Atmospheric Science (NCAS), for providing the radiosonde data and the supporting documentation to understand the data and usage for our water vapour distribution studies over India.
- ❖ We wish to thank COSMIC Data Analysis and Archive Center (CDAAC) for the availability of the radio occultation data and the supporting documentation to understand the data and usage for our water vapour distribution studies over India.

REFERENCES

- [1] IM Held., and BJ Soden., "Water vapor feedback and global warming 1", Annual Review of Energy and the Environment 25 (1), pp.441-475, 2000.
- [2] S.K. Sarkar., H.N. Dutta., M.V.S.N. Prasad., and B.M. Reddy., "TROPOSPHERIC WATER VAPOUR OVER INDIA", Antennas and Propagation, ICAP 89., Sixth International Conference on (Conf. Publ. No.301),vol.2, PP.302-306,1989.
- [3] Folorunsho Olaiya., Adesesan Barnabas Adeyemo., "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies", I. J. Information Engineering and Electronic Business, 1, pp.51-59,2012
- [4] Abhineet Shyam., B.S. Gohil., and Sujit Basu., "Retrieval of water vapour profiles from radio occultation refractivity using artificial neural network", Indian Journal of Radio & Space Physics Vol.42,pp 411-419,2013.
- [5] Stefania Bonafoni., Vinia Mattioli., Patrizia Basili., Piero Ciotti., and Nazzareno Pierdicca., "Satellite-Based Retrieval of Precipitable Water Vapor Over Land by Using a Neural Network Approach", IEEE,TGARS, Vol.49, No 9, pp.3236-3248,2011.
- [6] P. Kishore., M.VenkatRatnam., S.P.Namboothiri., Isabella Velicogna., GhouseBasha., J.H.Jiang., K. Igarashi., S.V.B.Rao., and V.Sivakumar., "Global (501S–501N) distribution of water vapor observed by COSMIC GPS RO Comparison with GPS radiosonde, NCEP, ERA-Interim, and JRA-25 reanalysis datasets", Journal of Atmospheric and Solar-Terrestrial Physics,pp.1849–1860,2011.
- [7] R. Norman., J.Le Marshall., B.A. Carter., K. Zhang., G.Kirchengast., S.Alexender., C.S.Wang., and Y.Li., "Comparing GPS radio occultation observations with radiosonde measurements over Antarctica", International Global Navigation Satellite Systems Society IGNS Symposium, pp.1-10,2013.
- [8] Fabrizio Pelliccia., Fabio Pacifici., Stefania Bonafoni., Patrizia Basili., Nazzareno Pierdicca., Piero Ciotti., and William J.Emery., "Neural Networks for Arctic Atmosphere Sounding From Radio Occultation Data", Geoscience and Remote sensing, IEEE transactions on Vol.49,No.12,pp.4846-4855,2011.
- [9] F. Ladstadter., A.K. Steiner., M. Schwarz., and G. Kirchengast., "Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013", Atmospheric Measurement Techniques,8, pp.1819-1834,2015.
- [10] XU Xiaohua., LUO Jia., and SHI Chuang., "Comparison of COSMIC Radio Occultation Refractivity Profiles with Radiosonde Measurements", Advances in Atmospheric Sciences, Vol.26, No.6, pp.1137-1145, 2009.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

- [11] Shu-peng Ho., Xinjia Zhou., Ying-Hwa Kuo., Hunt, Douglas., and Jun-hong Wang., "Global Evaluation of Radiosonde Water Vapor Systematic Biases using GPS Radio Occultation from COSMIC and ECMWF Analysis", Remote Sensing, Vol. 2 Issue 5, pp.1320-1330,2010.
- [12] S. K. A. V. Prasad Rao Anisetty., Ching-Yuang Huang., and Shu-Ya Chen., "Impact of FORMOSAT-3/COSMIC radio occultation data on the prediction of super cyclone Gonu (2007): a case study", Natural Hazards, Vol. 70, pp.1209-1230, 2014.
- [13] University Corporation for Atmospheric Research (UCAR), UCAR-COSMIC Data Analysis and Archive Center (CDAAC), 2053, <http://cdaaac-www.cosmic.ucar.edu/cdaac/>.
- [14] Met Office (2006): Met Office Global Radiosonde Data. NCAS British Atmospheric Data Centre (BDAC), date of citation. <http://catalogue.ceda.ac.uk/uuid/f2afaf808b61394b78bd342ff068c8cd>
- [15] B.-R.Wang., X.-Y.Lui., and J.-K.Wang., "Assessment of COSMIC radio occultation retrieval product using global radiosonde data", Atmospheric Measurement Techniques, 6, pp.1073-1083, 2013.
- [16] Khamphoui Southisombath., Toshio Wakabayashi., and Yoshiaki Moriya., "The impact of water vapor in atmosphere on Ku band satellite DTV reception in a sub-tropical region", Proceedings of ISAP'04, Sendai, JAPAN, IEICE , pp.1381-1384, 2004.
- [17] Ding Youqing., Yumei Fu., Zhu Fang., and Zan Xinwu., " Comparison of Missing Data Filling Methods in Bridge Health Monitoring System", Proceedings. 12th IEEE International Conference on Cognitive Informatics & Cognitive Computing, IEEE, 978-1-4799-0783-2013.
- [18] B. Mehala., P. Ranjit Jeba Thangaiah., and K. Vivekanandan., "Selecting Scalable Algorithms to Deal With Missing Values" International Journal of Recent Trends in Engineering, Vol. 1, No. 2, pp.80-83,2009.
- [19] <http://www.yourarticlelibrary.com/geography/the-main-physiographic-divisions-of-india-geography/5461/>
- [20] Venkata Sheshanna Kongara., and Dr. D. Punyasesudu., "Forecasting the Water Vapour Distribution Over India Using Artificial Neural Networks", "International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 3, Issue 4, pp.1-12,2015.
- [21] Thomas W. Schlatter., "Atmospheric Composition and Vertical Structure", Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA, Environmental Impact and Manufacturing, ArticleID: eae31MS.2009.
- [22] O.S.Eluyode., and Dipo Theophilus Akomolafe., "Comparative Study of biological and artificial neural networks", European Journal of Applied Engineering and Scientific Research, 2(1), pp.36-46, 2013.

BIOGRAPHY



Venkata Sheshanna Kongara is a Research Scholar, Department of Computer Science & Technology, Sri Krishnadevaraya University, Anantapuramu, Andhra Pradesh, India. His research interests are in the field of Data Warehousing, Data Mining, and Atmospheric Sciences and looking forward to fruitful collaborations with researchers in interdisciplinary fields.



Dr. D. Punyaseshudu is a Professor & Dean of Sciences, Department of Physics, Rayalaseema University, Kurnool, 518 003, Andhra Pradesh, India. He is specialized in Microwave Communications and Atmospheric Physics.