Spatial-temporal Characterization of Hurricane Path using GNSS-derived Precipitable Water Vapor: Case Study of Hurricane Matthew in 2016

Hoda Tahami *tahamih@oregonstate.edu*
Civil Engineering Department
Oregon State University
Corvallis Oregon, 97331, USA

Jihye Park *Jihye.Park@oregonstate.edu*
Civil Engineering Department
Oregon State University
Corvallis Oregon, 97331, USA

Abstract

Global Navigation Satellite System (GNSS) precise point positioning (PPP) technique is capable of monitoring Precipitable Water Vapor (PWV) in high accuracy with low cost. As PWV is related to the initiation and development of a severe weather convective system, this study analyzed the characteristics of PWV variations over time and space to monitor and predict the path and the intensity of a severe rainfall during a hurricane. The PWV measurements are obtained by processing ground based GNSS data. The spatial and temporal variation of PWV and other meteorological variables are characterized for the time frames of before, during, and after the severe precipitation. The correlation effect between meteorological variables were mitigated by adapting a principle component analysis (PCA) and multivariate regression analysis. The method allows determining the expected movement of the rainfall up to 24 hours in advance. The proposed method was validated by analyzing the distribution pattern of the predicted PWV residual, its magnitude, and the actual observed PWV in the region. As a case study, we adopted one of the destructive and long-lived hurricane along the Florida, Georgia, North Carolina and South Carolina coast, namely, Hurricane Matthew, occurred in October 2016. From the experiment, we identified the areas closely fitting the prediction model by computing the residuals between the GNSS derived PWV measurements at each station in the test site. The residual of the predicted model is used for determining the track of extreme hurricane precipitation and potentially applied to evaluate its intensity. This study proved the effectiveness of the statistical model for forecasting the hurricane rainfall path that is potentially applied to a hazard early warning system.

Keywords: GNSS Meteorology, Precipitable Water Vapor (PWV), Principle Component Regression (PCR), Hurricane.

1. INTRODUCTION

Precipitable Water vapor (PWV) is defined as the total amount of water vapor in a column of air above the surface [1]. PWV varies quickly with short temporal- and spatial-scale during severe meteorological phenomena such as tropical storms and hurricanes. These weather events accompany intensive precipitations that cause significant variation of the water vapor content and other meteorological parameters in the lower atmosphere [2],[3],[4]. Conventionally the PWV is detected by direct measurement sensors such as radio sounding sensors and microwave radiometers. However, these direct measurements have the limitations to distinguish the rapid variations of PWV due to the low spatial and temporal resolution. Also, these sensors require high-cost calibration, operation, and maintenance [4]. Satellite imageries monitor PWV over a wide area. Although a few multispectral images can measure even the vertical distribution of high

resolution PWV, the reliability of imagery based PWV decreases with the presence of clouds. Space-born microwave radiometers can acquire measurements for cloudy regions, but they are reliable only over the ocean in general [5]. One of the widely used space-based observation methods is the Moderate Resolution Imaging Spectroradiometers (MODIS). The MODIS provides water vapor images using near infrared channels twice a day in the world [6]. The infrared MODIS observes water vapor with 1 km spatial resolution over cloud-free areas, oceanic areas with Sun glint, and/or above clouds over both land and ocean [7],[8]. Envisat satellites, including the Medium-Resolution Imaging Spectrometer (MERIS), measure the solar radiation reflected from the Earth's surface and clouds in the visible and near-IR spectral range during the daytime. MERIS provides near-IR water vapor products above land or ocean surfaces under cloud-free conditions at two nominal spatial resolutions: 0.3 km for full-resolution mode and 1.2 km for reduced-resolution model [9],[10]. However, neither the MODIS nor MERIS instruments provide reliable PWV values above land or ocean surfaces in the presence of clouds. Similarly, the Atmospheric Infrared Sounder (AIRS) provides global coverage with a vertical resolution of 2 km and a spatial resolution at nadir of 45 km, but only operates in clear skies or partly cloudy conditions [8], [11].

PWV can also be obtained from forecasting models. Numerical Weather Prediction (NWP) models quantitatively forecast the weather based on the prediction models of temperature, wind, precipitation, etc. However, the NWP model involves a critical challenge associated with its temporal latency. In case of the global scale NWP, the latency is approximately 4 - 6 hours. Although the local scale NWP models can be generated with less latency because of the smaller data sets, they are hardly applicable due to the limitation of the covered area so that the prediction only works when the weather event approaches in close proximity of the area to be covered by the NWP model. The global scale models are certainly applicable for monitoring the global scale weather events, but they tend to include high levels of uncertainty due to the lack of detail which is another limiting factor [12]. Additional limitations of the NWP model come from the inconsistent types of terrestrial and space-based observations. Since different types of data have different characteristics, the data integration and the analysis cannot be done in a unique way to provide precise forecasting results [12].

Global Navigation Satellite System (GNSS) has been used to detect water vapor content in the atmosphere from a tropospheric delay of a GNSS signal that has been shown in a number of researches (e.g., [3],[13],[14],[15]). One of the most valuable attributes of GNSS derived PWV is its ability of providing precise and seamless measurements under all weather conditions regardless of the presence of clouds, which derive a challenge to aforementioned traditional weather monitoring technologies [5]. Another benefit of GNSS is the data availability and accessibility. Currently, the densely distributed Continuously Operating Reference Station (CORS) network provides seamless observations with high spatial and temporal resolution. In addition, GNSS data has a good potential for improving NWP forecasts by investigating initial conditions in the NWP based on the GNSS derived tropospheric measurements [16],[17],[18]. Various studies demonstrated the capability of GNSS to monitor small-scale water vapor structures associated with the initiation and development of convective systems [14],[19],[20],[21]. From the study, the PWV rate of change (PWV ROC) over time with other atmospheric parameters were considered as a precursor to heavy precipitation. Previous studies verified that when PWV reached a certain threshold, the probability of rainfall would quickly increase [15],[22],[23],[24],[25],[26],[27]. The PWV threshold empirically derived by real-time GNSS PPP is used for forecasting rainfall [28]. Although the close relationship between the PWV and a rainfall is revealed in numerous researches, fewer studies have focused on predicting the path of severe precipitation events such as storms or hurricanes using the PWV. In this study, a new method for forecasting the path and intensity of extreme rainfall during a hurricane is proposed based on the prediction model of GNSS-derived PWV.

2. SEVERE WEATHER MONITORING USING GNSS

Several atmospheric factors contribute to the initiation and development of extreme precipitation. Among them, the instability and convergence/divergence conditions of water vapor content are key factors to forecast heavy precipitation [29],[30],[31]. The PWV as a moisture related parameter is used to forecast the precipitation and assess its severity [3],[8],[16],[21]. Therefore, the relationship between the PWV and other atmospheric parameters are crucial for characterizing the causes of the severe precipitations. The precipitation is highly associated with pressure systems as well as temperature [31],[32],[33]. This study investigates the occurrences of precipitation by analyzing the atmospheric pressure, temperature and the PWV. To secure the high spatial and temporal resolution of PWV, GNSS data are processed to extract the PWV component. The atmospheric constituents in the troposphere cause a delay on GNSS signals during transmission from a satellite to a receiver on the ground. The delay is considered as the neutral atmospheric delay, or conventionally referred to as the tropospheric delay (TD) as it is mostly affected by the troposphere [3], [34]. The total tropospheric delay along the slant ray path can be mapped onto the zenith direction, yielding the Total Zenith Delay (TZD), by a mapping function that is the function of satellite's elevation angle. The TZD is defined as the sum of a Zenith Hydrostatic Delay (ZHD) and a Zenith Wet Delay (ZWD), which are the delays caused by the hydrostatic component and non-hydrostatic (or wet) component of the atmosphere, respectively. The ZWD depends mostly on the distribution of water vapor content along the signal propagation path [35]. Due to the high variability of the atmospheric water vapor, many studies rely on the relatively accurate ZHD model to calculate the ZWD by subtracting the ZHD value computed by a ZHD model from the TZD measured at a GNSS station. The PWV can be computed using ZWD measurements from GNSS signals. The retrieving process of the PWV from the ZWD requires meteorological parameters, such as pressure and temperature. Eq (1) represents the relation between the PWV and the ZWD [36],[37]:

$$PWV = ZWD \times \Pi \tag{1}$$

and conversion factor Π is expressed in Eq (2).

$$\Pi^{-1} = 10^{-6} \times \rho \times R_v \left(\frac{k_3}{T_m} + k_2\right) \tag{2}$$

where ρ is the water vapor density (1000 kg m^{-3}) and R_v denotes the specific gas constant of water vapor (461.5 J $kg^{-1}K^{-1}$). k_2 and k_3 are set to $(22.1 \pm 2.2)\, K \cdot mbar^{-1}$ and $(3.739 \pm 0.012)10^5\ K^2 \cdot mbar^{-1}$ respectively [38]. It is worth to mention that k_2 can be represented as (17 ± 10) $K \cdot mbar^{-1}$ for a larger uncertainty [39]. T_m is the modeled weighted mean temperature of troposphere (K), which is computed based on the regression relation between T_s (surface temperature) and T_m from a linear function, $T_m = 70.2 + 0.72\, T_s$ proposed by [36]. Although the ZHD and therefore PWV computations involve the meteorological parameters, they are relatively stable in time and space and can be obtained by an interpolation method.

2.1 GNSS-based Hurricane Monitoring

The variations of PWV provide the transition pattern of the water vapor content. When GNSS based estimated PWV are analyzed with the local atmospheric elements, the formation process of the sever precipitation can be observed by assessing the correlation between atmospheric pressure, temperature and PWV. Because the PWV variation is an indication of the precipitation showing the interrelation with other atmospheric parameters, the PWV ROC can be used to monitor the atmospheric changes over time. Several studies showed that the PWV greatly increases a few hours prior to the most intense rainfall then sharply decreases when precipitation begins to weaken and finally ceases [26], [27], [38], [40], [41], [42]. Referring to [15], [27] and [42] 1) PWV peaks occur at a few hours before the precipitation onset; 2) the PWV ROC sharply varies before the start of the rain in the most rainfall events. Thus, the significant spatial-temporal variation of PWV and PWV ROC corresponds to the rain event in time and can be identified as important features for short-term precipitation forecasting purposes.

2.2 Case Study: Hurricane Matthew in October 2016

Hurricane Matthew is one of the most destructive and long-lived hurricanes in the United States occurred in October 2016. To investigate the potential of GNSS data for retrieving the perturbation of GNSS derived PWV, seven days (5 – 11 October 2016) observations of Continuously Operating Reference Station (CORS) were used to track the path of this hurricane. For this study, we formed a network, which consists of thirty-eight CORSs that are regionally distributed along the coasts of Florida, Georgia, South Carolina, and North Carolina. For all GNSS CORSs, the GNSS-derived TZD and ZWD were estimated in 300 seconds intervals using the automatic precise positioning service (APPS) developed by NASA's Jet Propulsion Laboratory (JPL). The satellite orbits and clock corrections are conducted by using the JPL final products for the orbit and clock corrections. The elevation cutoff angle of 15 degrees were applied for processing. From the GNSS derived ZWD, the PWV were computed based on Eq (1) and (2). In Eq (2), T_m is derived based on the question $T_m = 70.2 + 0.72\,T_s$ proposed by [36] considering that our sites are located in the United States with a latitude range of 27° to 65° and a height range of 0 to 1.6 km that is valid for the area of the investigation according to the literature. The surface meteorological data is provided by real-time online weather information resources. The meteorological data were collected from the weather underground website (https://www.wunderground.com). The fourteen weather stations close to the CORS stations were utilized for this experiment providing three meteorological parameters: the pressure, temperature and relative humidity). The average distance between stations is 120 km. The temporal resolution of the GNSS-derived PWV and the meteorological parameters are 5 and 60 minutes respectively. To integrate the GNSS data and the meteorological data for the statistical analysis in Section 3, the meteorological data are temporally interpolated for every 5 minutes. Figure 1 shows the geographical distribution of the GNSS CORS and meteorological stations over the study area.

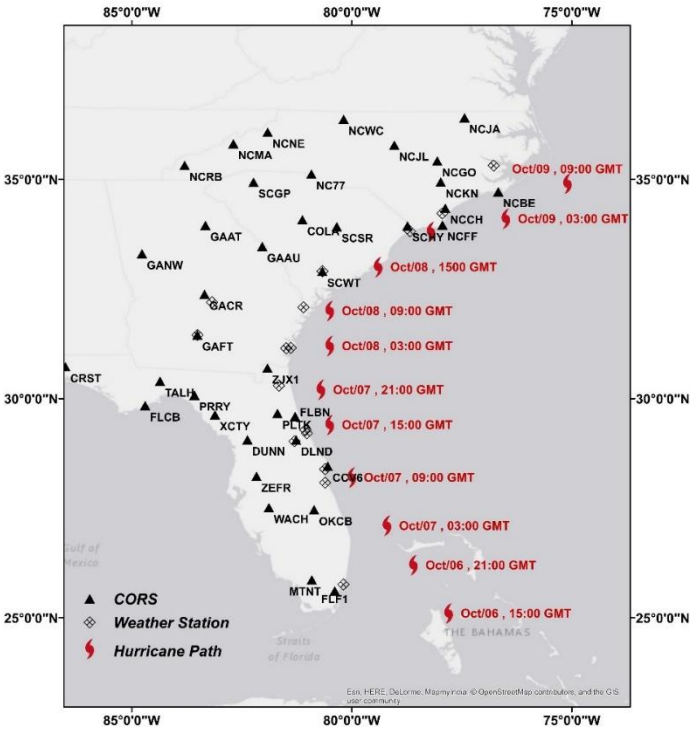

FIGURE 1: Geographical distribution of the GNSS CORS, meteorological stations and the reported hurricane path over the study area. (Produced by ESRI ArcGIS).

.

3. STATISTICAL ANALYSIS OF METEOROLOGICAL FACTORS FOR HURRICANE MATTHEW

The PWV measurements, which is the primary data in this study, are obtained from GNSS observations. From the PWV, the PWV ROC are computed by taking a time difference of PWV and it can be utilized to predict the PWV together with other meteorological parameters as shown in Eq (3). In Eq.(3), the PWV at epoch t+1 is predicted based on the weighted mean of PWV at epoch t ($\overline{PWV_t}$), and difference of PWV, temperature (T), pressure (P), and relative humidity (RH) from their corresponding mean values at epoch t. The derivative of PWV at t+1 is expressed as ΔPWV_{t+1} that is referred to as PWV derivatives. In this study, we defined the window size of 1 hour for computing the weighted mean of PWV. This approach linearizes the non-linear dynamics of PWV by using the derivative of PWV instead of PWV itself which is also supported by [45] where it claims that the forecasts of rainfall along with linear regression analysis on meteorological observations such as pressure and temperature can be used to obtain an estimate of the anticipated rainfall. Due to the nature of the meteorological parameters that are related each other, we applied a multivariate linear regression model to decorrelate the parameters. In the proposed multivariate linear regression model, the PWV derivative as a dependent variable is described with respect to the independent explanatory variables of pressure, temperature, and relative humidity. Eq (3) can be simplified to Eq (4) where the next value of dependent output (ΔPWV_{t+1}) is regressed on the basis of the previous values of the input regressor. ($\Delta PWV_t, \Delta P_t, \Delta T_t, \Delta RH_t$).

$$(PWV_{t+1}) = \overline{PWV_t} + c_1\left(PWV_t - \overline{PWV_t}\right) + c_2\left(P_t - \overline{P_t}\right) + c_3\left(T_t - \overline{T_t}\right) + c_4\left(RH_t - \overline{RH_t}\right) + e \quad (3)$$

$$(PWV_{t+1}) - \overline{PWV_t} = (\Delta PWV_{t+1}) = c_i[\Delta PWV_t, \Delta P_t, \Delta T_t, \Delta RH_t] + e \quad (4)$$

where ΔPWV is PWV derivative, P is pressure, T is temperature, RH is relative humidity, c consists of regression coefficients, t is time and e is the error term. For each prediction time window, the model estimates the prediction of PWV using the hourly PWV ROC and the derivatives of P, T, and RH. In this paper, a model with 6h, 12h and 24h lead time is presented to forecast PWV that is applied to predict the hurricane path for 6h, 12h, and 24h forecast.

3.1 Classification of Hurricane Matthew Development Stages

During a hurricane event, the PWV in the affected region fluctuates with the different intensity and duration. The rapid fluctuation of GNSS-derived PWV is highly correlated to the severe rainfall and can be modeled in the dynamic statistical multivariate regression model. However, the PWV ROC and other atmospheric parameters behave differently at the different stage of hurricane, so that the principle regression model introduced in this section should be tailored to each stage in the hurricane lifetime. The time series of hurricane observations are divided into multiple classes of “before”, “right before”, “during”, “right after”, and “after” the hurricane and for each class a unique regression model is developed. To characterize each model, the meteorological parameters and the PWVs from CORSs in the test site were analyzed during the whole period of hurricane Matthew. The analysis results show the consistent patterns in those meteorological parameters in all sites for each model that were shown in the previous research in [27]. From the findings, the PWV threshold is set to 0.055 m where most of the severe precipitation events occurred when PWV exceeded that threshold. Selection of this threshold is also supported by other studies conducted in different regions and different times that claim the highest amount of water vapor is observed typically a few hours before the heavy precipitation when the PWV exceeded that threshold [15], [42]. From the findings in the test site, the numerical criteria of the classes for this event were determined. Each class is defined by applying the threshold to the GNSS-derived PWV fluctuations and deviation of pressure. Figure 2 is an example (CCV6 station) showing the PWV, pressure, and rainfall with respect to the different classes of hurricane stage - before, right before, during, right after, and after- based on the hurricane occurrence time.

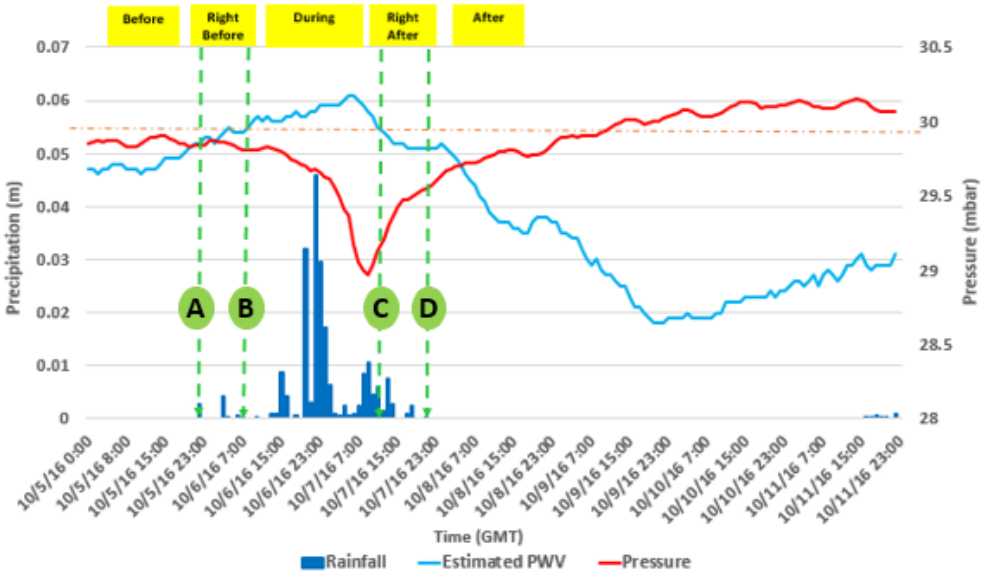

FIGURE 2: Classification of Model based on PWV Threshold.

In the figure, the GNSS-derived PWV stayed below 0.055m during the "before" class, which is a normal condition, and no severe rainfall was observed during that time. The estimated PWV continuously increases and reaches the threshold in the "right before" class when rainfall starts and increase in PWV and decrease in pressure are observed (Green symbol indicated as A in Figure2). Therefore, the first epoch of the "right before" class coincides with the beginning of the rainfall. The PWV time series show an active, increasing pattern in the "right before" class. The last epoch of "right before" class is the time when PWV exceed the threshold (Green symbol indicated as B in Figure 2). The end of the "right before" class is set as the first epoch of "during" class. In the "during class" the rainfall amounts reach to its peak. Then PWV time series dramatically decrease to reach the threshold where the "right after" class is defined (Green symbol indicated as C in Figure2). It then decreased to below the threshold while the "after" class is formed where rainfall begins to cease (Green symbol indicated as D in Figure 2). Obviously, the rainfall indicator bars in the bottom of Figure 2 are significantly high when the PWV is fluctuating above the 0.055 m. The same strategy is applied to the PWV time series of other stations to define the classes. After defining the classes, a unique regression model for predicting PWV were built in each class. It should be emphasized that this study specially focuses on the "right before" model, which is the most useful model to analyze and predict the movement of hurricane.

3.2 Training Dataset

To derive the coefficients,c_i, of each regression model of PWV in Eq (4), the representative stations are selected to generate a training dataset for each class. For deriving the coefficient of "right before" model at specific time in the study area six stations including MTNT, OKCB, ZJX1, GAAU, SCRS, and NCKN are selected. All of the selected stations in the right before model training dataset should reach the threshold in the "right before" class at the start point of the forecast period. For the purpose of the models characterization, we observed all six station's data for all the classes, which serve a priori information, and then generalized the models for testing the capability of the models for detecting the path of hurricane precipitation. It should be noted that these stations were excluded for the rest of experiment. The meteorological and GNSS measurements at a few sites near a hurricane landfall in the area (typically near the coast) are required for serving as the initial information and setting the prediction model that makes this method valid for the real world applications. This approach is validated by applying the model to other remaining CORSs in the study area as the independent test dataset.

3.3 Principle Component Regression

The PWV regression model requires preliminary analysis identifying the relationships between independent variables, PWV, and the dependent variables, P, T, and RH, to avoid the multicollinearity problem caused by the correlation between the variables in a model [43]. The multicollinearity can be evaluated by examining a correlation matrix and a variance inflation factor (VIF). The VIF identifies the correlation between independent variables and provides an index for the strength of correlation, i.e., severity of existence of multicollinearity. VIFs equal or less than 5 indicates the moderate correlation and VIFs greater than 5 imply the critical levels of multicollinearity [43], [44]. Table1 shows an example of the correlation coefficient among PWV, P, T, and RH over the stations in the training dataset station and for the right before classes. The correlation coefficients between PWV, P and RH are larger than the other coefficients. Since the correlation coefficient cannot sufficiently provide an estimate of the degree of multicollinearity, other statistics are added to detect the probable existence of collinearity and its severity.

Variables	**ΔPWV**	**ΔP**	**ΔT**	**ΔRH**
ΔPWV	1	-0.760	0.297	0.754
ΔP	-0.760	1	-0.046	-0.719
ΔT	0.297	-0.046	1	-0.243
ΔRH	0.754	-0.719	-0.243	1

TABLE 1: Correlation Matrix of Variables.

Table2 provides the multicollinearity statistics for each variable in the model. The statistics include R-squared value (R^2), the tolerance factor, which is (1-R^2), and VIF, which equals to the inverse of the tolerance factor. A high value for R^2 and the VIF value greater than 5 may cause non-significant parameter estimates in the regression due to multicollinearity problem.

Statistic	**PWV**	**P**	**T**	**RH**
$\boldsymbol{R^2}$	0.840	0.628	0.595	0.815
Tolerance	0.160	0.372	0.405	0.185
VIF	4.235	2.686	2.470	4.416

TABLE 2: Multicolinearity Statistic.

As shown in Tables 1 and 2, the moderate to significant correlations exist between the tested meteorological parameters. To avoid the multicollinearity from the suggested statistical model, we applied Principal Component Regression (PCR), which is based on the Principle Component Analysis (PCA). The PCA extracts the maximum variance of a data set as factors in an iterative process. The resulting factor represents the vector that maximizes the dispersion of the observations. The first principal component accounts for the largest variability in the data, and each succeeding component accounts for the next biggest variability of the remaining property in the data. Principal components are uncorrelated orthogonal linear functions of the original variables obtained by a transformation of the form [44], [45]:

$$F = \mathrm{ZA} \tag{5}$$

where F is the set of transformed variables as principal components, Z is the set of original variables, and A is the set of eigenvectors associated with significant eigenvalues obtained from the correlation matrix of the original variables [44]. In PCR, the independent variables are standardized to avoid the problem of non- commensurate units of the variables so that:

$$Z = [Z_{tj}] \tag{6}$$

where t = 1,2, …,n and j = 1 ,2 ,…,p and Z_{tj} is described as

$$Z_{tj} = \frac{X_{tj} - mean(X_{tj})}{std\ (X_{tj})} \quad (7)$$

where X_{tj} is $\Delta PWV, \Delta P$, ΔT and ΔRH. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity are used to evaluate the suitability of data for factor analysis [44], [46]. The Chi-square value of 5496.3 is significant at 95% confidence interval and KMO statistic of greater than 0.50 for all variables confirms the appropriateness of PCA for further analysis of data. In the next step, the eigenvalues and the eigenvectors are computed from the correlation matrix of the standardized independent variables. The eigenvalues are 2.491, 1.157, 0.266 and 0.085 and the eigenvectors are as follows:

Variables	F_1	F_2	F_3	F_4
ΔPWV	0.588	0.236	0.354	0.088
ΔP	-0.575	0.004	0.815	0.071
ΔT	0.342	0.921	0.030	-0.082
ΔRH	0.568	-0.309	0.235	-0.013

TABLE 3: Eigenvectors Resulted from PCA.

Table 3 shows the correlation between the meteorological variables in Table 1 and their principle components (F1- F4). Clearly, the first principle component is substantially correlated with PWV, pressure and relative humidity. The second and third component is correlated with temperature and pressure respectively. However, the fourth principal component is moderately not correlated with the explanatory variables. The contribution of each factor to the total variability of the dataset is shown in Table 4.

	F_1	F_2	F_3	F_4
Eigenvalue	2.491	1.157	0.266	0.085
Variability (%)	62.282	28.935	6.662	2.121
Cumulative %	62.282	91.217	97.879	100.000

TABLE 4: Eigenvalues explain most of the variability in data.

It demonstrates that the first principal component (F_1) accounts for 62.28 % of the system "variance" while the second (F_2), third (F_3) and fourth (F_4) components account for 28.93%, 6.66% and 2.12% respectively. Based on the contribution of each component in the variability explanation in the dataset and factor loadings of principal components, the first three principal components were considered in the PCR. The selected principal components explain more than 97% of the variance. The regression model is then described as:

$$y = Fb + e \quad (8)$$

where y is a dependent variable, F represents principal components, b is regression coefficients to be estimated, and e indicates an error term.

3.4 Model Parametrization

The PCR model in Eq (8) generates the set of coefficients for the principle components. The set of regression coefficients shown as b vector in Eq (8) and the 't' statistic for each coefficient are provided in Table 5. A t-test and the corresponding p-values are performed to test a null hypothesis and evaluate the significance of each component in the regression model. The null hypothesis is that the regression equation does not explain a significant variation of the dependent variable; i.e., the regression coefficients are equal to zero. The p-values are compared

to a significance level (0.05 in this study) to make the decisions whether to retain or reject the null hypotheses. If the p-value for a variable is less than the significance level, the null hypothesis will be rejected.

	VALUE	STANDARD ERROR	T-TEST	P-value> |T|
$\mathbf{b_1}$	0.928	0.004	207.867	< 0.0001
$\mathbf{b_2}$	0.254	0.003	56.882	< 0.0001
$\mathbf{b_3}$	0.183	0.004	40.866	< 0.0001

TABLE 5: Principle Component Regression Coefficients.

The t-test values and their corresponding p-values at 95% confidence interval shows that the regression coefficients are statistically significant. When the non-zero regression coefficients are statistically significant, the alternate hypothesis is accepted, i.e., the calculated coefficients are used to form the PCR model. By applying the derived coefficients to each factor in Eq (8), the final model is as follows:

$$\Delta PWV_{t+1} = b_1(F1) + b_2(F2) + b_3(F3) + e \quad (9)$$

While the normal distribution of independent variables is not essential to perform PCR, the normality of PCR residuals is a critical requirement for the ability to use the standard critical value for statistic tests such as t- and F-test. Moreover, auto-correlated PWV residuals lead regression coefficients to be statistically insignificant and make statistic tests unreliable. Therefore, the normality test and Autocorrelation Function are used to evaluate the conditions of normal distribution and detection of autocorrelation respectively. The proposed statistical model quantifies the relationship between the predicted PWV ROC and meteorological variables. The model is further extended to predict the path of the hurricane by utilizing the behavior of the PWV ROC as a significant indicator of predicting different intensity and duration of the event.

4. HURRICANE PATH PREDICTION MODEL

In Section 3, a statistical model for predicting PWV was demonstrated. Figure 3 shows the flowchart describing all the phases in the modeling of the rainfall path prediction during the hurricane in the previous sections.

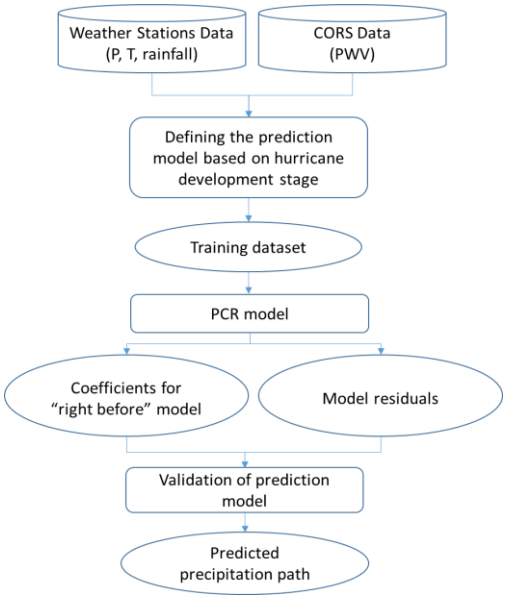

FIGURE 3: Flowchart describing the phases in prediction of hurricane rainfall path.

From the model, the temporal variation of PWV at a particular site can be obtained. By observing the temporal variation of PWV at multiple sites simultaneously, the spatial variation of PWV at a certain epoch can be monitored and analyzed that can be applied to forecast the spatial movement of severe weather event, which is a hurricane in this study.

4.1 Model Parametrization

For each class in the hurricane development phases represented in Figure 2, the predicted PWV calculated form the methodology described in Section 3, was compared with the actual PWV. The reliability of the prediction model is validated by calculating the residuals of the PWV measurements for each model. The residuals are computed by applying the derived coefficients from the training dataset to all other CORSs in the independent test dataset. As an example, the coefficients derived for each class from the training dataset, are applied to CCV6, to predict PWV for each class for CCV6. Figure 4 depicts the predicted PWV at CCV6 over the period of 5-12 October 2016 and the corresponding residuals using one-hour prediction model for all classes of event. The bottom panel in Figure 4 depicts the residuals of the predicted PWV from the estimated PWV. The amplitude of the residuals is about 5 % - 10% to the amplitudes of corresponding PWV almost all period except the “after” class where a significant rise was shown. It is possibly due to the fact that the prediction models are derived based upon the PWV fluctuation during the rainfall so that it is not sensitive to non-rain scenarios such as the “after rain” event. Also, the maximum residuals are within the first or last epochs of the corresponding class in each model except “after” class. This is because of the transition from one model to another. However, the range of residuals for each model shows how the model matches to the actual PWV time series.

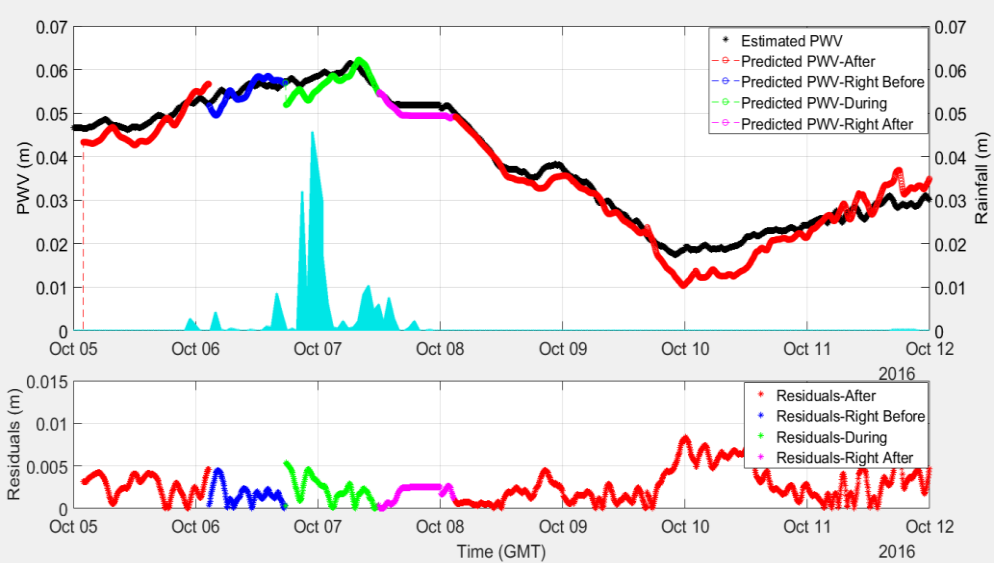

FIGURE 4: Prediction models and the corresponding residuals for CCV6 during 5-12 Oct, 2016; The top panel shows the models in different colors that are mapped over the actual PWV time series (Black line) and the blue bars represent the recorded actual rainfall on the station for the mentioned period.; the bottom panel shows the residuals of observations with respect to each model.

4.2 Spatial-temporal Analysis of Predicted PWV Residuals and Prediction of Hurricane Path

To investigate the PWV responses to the storm front, the PWV residuals of right before and right after models are shows in Fig 5. Fig 5 (a) shows the residuals of “right before” model when the rain event approaches near the stations in the south region (near Florida) and Fig 5 (b) shows the residuals of the “right after” model when the rain event passed Florida. It shows that the location of the minimum residuals along the hurricane days agree with the observed rain pattern, which is used for prediction of the hurricane path. For example, the smaller residual takes place a few

hours before the hurricane passage demonstrating a higher concentration of the rainfall in the area of the right before model. In this model, the larger residuals corresponded to either 1) the area not affected by the hurricane yet or 2) the area where rain ceases; i.e., the area with smaller residuals can be considered as a probable area to be affected by a hurricane.

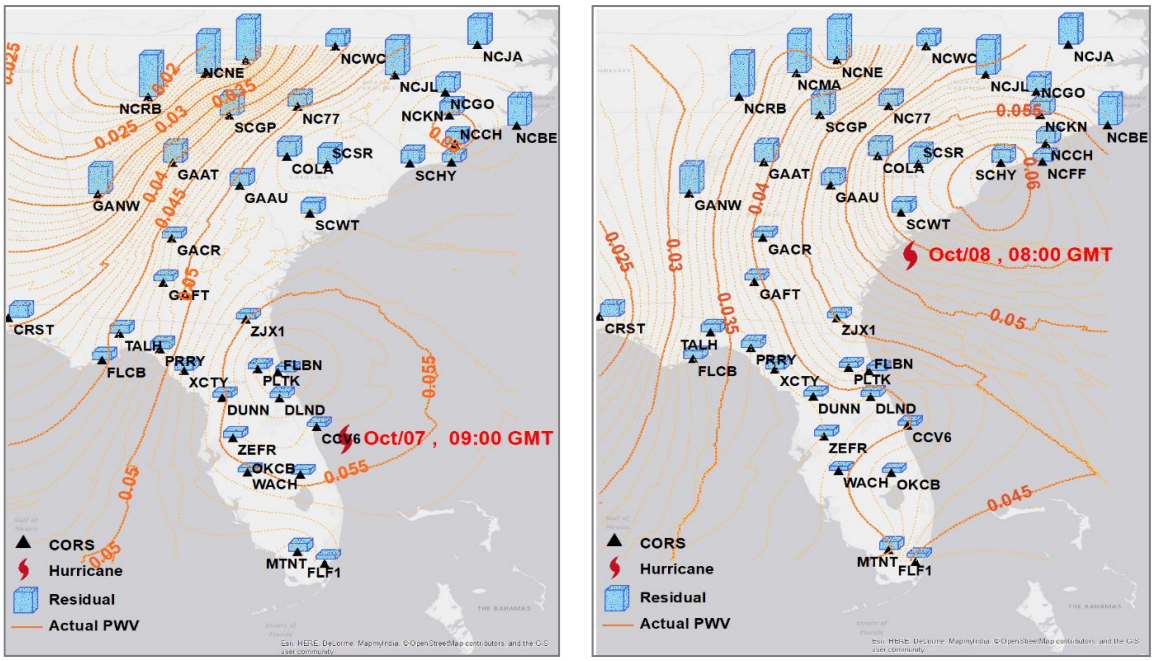

FIGURE 5: Distribution of actual PWV and the residuals for the "right before" and "right after" models over the area corresponding to 7 October at 9:00 am (Figure5 (a) : left panel) and 8 October at 12:00 pm GMT (Figure5 (b) : right panel) respectively. The red hurricane mark shows the actual reported time and location of the hurricane center provided by National Weather Service (Figures are created by ArcGIS).

Table 6 summarizes the result of a prediction model at CCV6. In this example, the "right before" model is examined over different time lags and computed the residual of the observed PWV from the prediction model. Table 6 provides the statistics of the residuals, which are mean and standard deviation for the "right before" model in different prediction time lags.

	1h	6h	12h	24h
Mean	0.004	0.003	0.003	0.003
Standard deviation	0.000	0.001	0.000	0.002

TABLE 6: Residuals for the prediction model (right before the hurricane) in prediction time lags at CCV6 (unit:meter).

The station within with the minimum residuals of the "right before" model can be considered as an indicator for the most probable region to be affected by hurricane. The hurricane path can be determined by observing the predictions of PWV derived at multiple sites in different directions. The residuals of the "right before" model at those sites are compared for the prediction time lags of 1 – 24 hours over the study area. In each model, the minimum residuals over the stations show where and when the model is the best fit to the actual observations on the specific time window. For more robust analysis considering the uncertainty due to the noise and other error sources in the input data, we evaluated the residual PWV at multiple sites in a local area by forming grids of 200×200 km^2.The grid size was determined with the consideration of the test site and the distribution of available GNSS stations. As shown in the map in Figure 5, each cell is wide

enough to cover multiple CORSs that it reduces the risk of failure or error of a particular station. Also, the cell is small enough to distinguish the directions of potential paths of hurricanes. Unlike the conventional method of NWS forecast, which provides the different size of probability circles for the location of hurricane centers, our proposed approach indicates the most probable cell for the hurricane path with the mean and the standard deviation of the model residuals. Figure 6 presents the residuals to the right before models at the stations in the study area for 1 hour and 24 hour forecast in panels (a), (b) respectively. It should be noted that the storm front passed over the region around CCV6 from 9:00 am GMT on October 7. The right before model is applied to CCV6 station and the residuals for different prediction time are shown in the figures. The hurricane path is determined based on the direction of grids containing the sites with smaller residuals than another grids. The red line in Fig 6 is the reported path during hurricane time from National Weather Service (NWS). The reported time of hurricane centers along this path was presented in Fig 1.

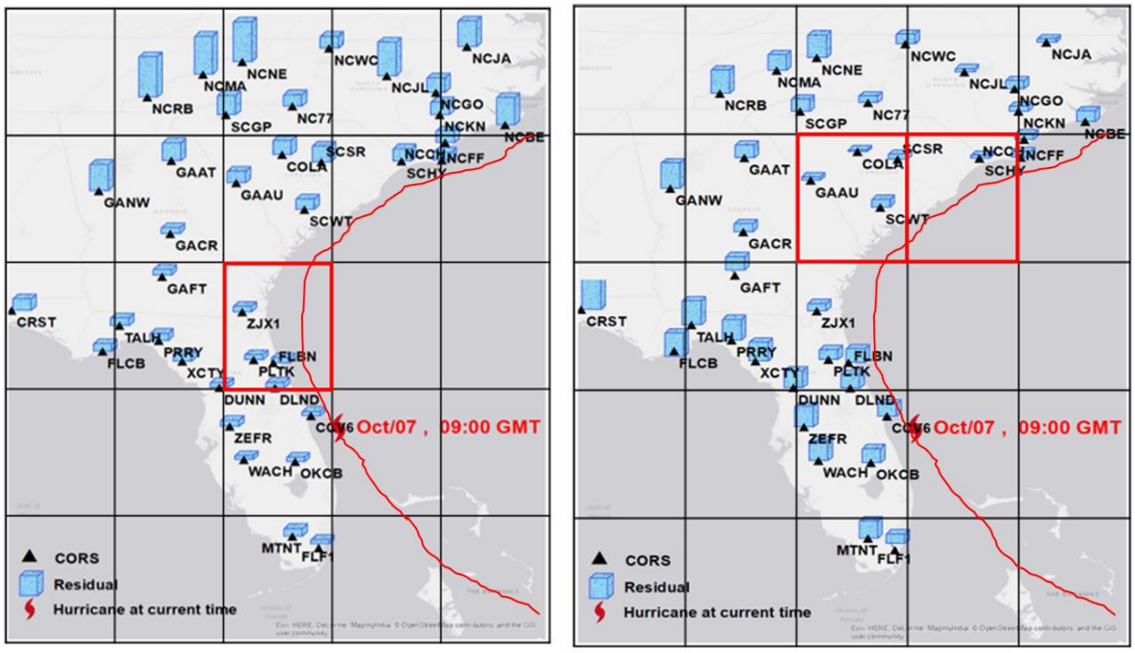

FIGURE 6: Predicted residuals over the study area by applying prediction model to CCV6 on 7 Oct, 2016. The hurricane hit the area close to CCV6 on 7 October at 9:00 am GMT. The red hurricane mark shows the reported time and location of the hurricane provided by NWS. The red line also shows the reported hurricane path. In each image, red grid shows the stations fall into the hurricane and black girds show the grids correspond the stations located outside the hurricane center. The residuals are corresponding to the "right before" model in 1h and 24h time prediction lead time in the left panel (Figure6 (a)) and right panel (Figure6 (b)) respectively (Figures are created by ArcGIS).

By comparing the average residuals of stations within each grid, the hurricane path can be predicted to move eastward. Moreover, in order to investigate the predicted rainfall intensity in each grid in the hurricane path, the residuals within the corresponding grid are compared to each other. For instance, among the stations in Figure 6 (a) ZJX1 in the red grid represents the lowest residual (Figure 7), which means the area near ZJX1 station is highly likely hit by the hurricane around 10:00 am GMT on 7 October. The same strategy is applied for 24 hour prediction time window in Fig 6 (b). The numerical comparisons of prediction residuals for the grids located inside and outside of the hurricane are provided in the Figure 7 and Figure8 respectively.

	0.034 (0.001)	0.021 (0.001)	0.018 (0.001)	0.018 (0.003)
0.010 (0.000)	0.010 (0.001)	0.011 (0.002)	0.009 (0.000)	
0.008 (0.002)	0.006 (0.000)	0.003 (0.001)		
		0.004 (0.000)		
		0.005 (0.002)		

FIGURE 7: Residuals and its standard deviation in parenthesis (in meter) for the grids inside the hurricane (red) and outside the hurricane (black). The results are attributed to the "right before" prediction model (1h) at 9:00 am GMT on 7 Oct to predict 1h later.

	0.016 (0.000)	0.012 (0.003)	0.005 (0.000)	0.007 (0.000)
0.021 (0.000)	0.009 (0.001)	0.003 (0.001)	0.003 (0.000)	
0.027 (0.001)	0.016 (0.003)	0.010 (0.003)		
		0.014 (0.001)		
		0.011 (0.000)		

FIGURE 8: Residuals and its standard deviation in parenthesis (in meter) for the grids inside the hurricane (red) and outside the hurricane (black). The results are attributed to the "right before" prediction model (1h) at 9:00 am GMT on 7 Oct to predict 24h later.

As shown in Figure 7 and 8, the characteristics PWV residual in each grid plays a role as a forecast indicator, to predict the path of the hurricane and its relative intensity. The path of a hurricane can be forecasted by analyzing the locations of lower residual grids based on the prediction models with different prediction time window. Also, the relative intensity can be estimated by relative comparison of residuals within each grid in the hurricane path. The predicted PWV behavior preceding the rain event varies from a station to another station in terms of the magnitude of residuals in each grid, which can be due to the variation of the storm intensity in each location. And, the stations with the lower residuals within that grid shows the more severe parts of hurricane. The experimental results shown in Figs 6, 7, and 8, indicate a sequence of rapid predicted PWV residual decrements for each station, starting from Florida at 9:00 am GMT October 7 and reaching in South Carolina and North Carolina after 24 hours at 8:00am GMT on October 8, our experimental result is confirmed by the reported hurricane path in Fig 5.

5. SUMMARY AND CONCLUSION

In this experiment, the relationship between GNSS derived PWV and atmospheric parameters have been analyzed for Hurricane Matthew in 2016. The analysis of the GNSS-derived PWV time series agree that the occurrence of rain events in the study area is related to atmospheric low pressures and water vapor entries. During the event, the patterns of the GNSS-derived PWV fluctuations were observed showing sudden and sharp increments in the PWV followed by sharp descending trends, a few hours prior to the onset of precipitation. This supports a connection between large PWV increments and convective precipitation. Through the case study, the trace of the hurricane was characterized based on the relationship between predicted GNSS-derived PWV and heavy rainfall by monitoring the troposphere water vapor distribution trend before,

during, and after the severe precipitation due to the hurricane. Based on the relationship between GNSS-derived PWV rate of change and actual atmospheric parameters, a short-term forecasting method is established to track the rainfall during the hurricane. The PCA and multivariate regression analysis were performed in order to derive the prediction model. The method can forecast the severe hurricane precipitation track 1-24 hours in advance. The result of this study provides evidence that the level of predicted PWV fluctuations can be precisely matched to the meteorological and geographic conditions. Moreover, this method can be applied to different regions, and different durations of rainfall events because the model is defined by using the measurements of actual test site to define the initial stage of hurricane right after it lands to the ground stations By testing the prediction models of before, during and after hurricane, it has been verified that the precipitation frequency increased rapidly when the PWV reached a certain threshold. The definition of each class (before, during and after) is associated with the threshold defined by the variations of GNSS-derived PWV. Also, our results show that the predicted PWV ROC significantly increased prior to the severe precipitation that are more extreme change than those predicted for other weather conditions. This is verified by representing the distribution pattern of the predicted PWV residual, its magnitude and the actual observed PWV in the region. Our results illustrate that tracking the location of the minimum residual during the hurricane agree with the observed rain pattern and can be used for rainfall path prediction during the hurricane. For the model used for hurricane onset prediction, the smaller residuals occur a few hours before the hurricane passage, demonstrating that a bigger concentration of the hurricane rainfalls in the area. From this model, the area with the larger residuals is considered either not to be affected by rainfall induced by the hurricane or agreed with the rain's stop observed for the corresponding time. Therefore, the magnitude of the predicted model residuals is used for hurricane rainfall tracking and potentially applies to evaluate the intensity prediction respectively.

This study shows the feasibility of a short term (up to 24 hours) forecast for hurricane path by applying high resolution GNSS-derived PWV. In addition, the availability and widespread spatial distribution of GNSS stations provide continuous PWV estimates in near real time. Although the other metrological measurements were also the part of the prediction model, the interpolated values from relatively sparse meteorological data still generated reliable solution in our case study that was considerably thanks to their characteristics that are relative stable over time and space. Since various hurricanes have their own characteristics, this study will be applied to other hurricane events to validate or refine the method.

6. REFERENCES

[1] G. Li, D. Huang, B. Liu, and J. Chen. "Experiment on driving precipitable water vaport from ground-based GPS network in Chengdu Plain," Geo-Spatial Information Science. vol. 10, pp. 181–185, 2007.

[2] A. Akilan, K.K. AbdulAzeez, S.Balajj, H.Schuh, and Y.Srinivas. "GPS derived Zenith Total Delay (ZTD) observed at tropical locations in South India during atmospheric storms and depressions," Journal of Atmospheric and Solar-Terrestrial Physics, vol.125-126, pp.1-17, 2015.

[3] P. Benevides, J. Catalao, and P.Miranda, "On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall," Natural Hazards and Earth System Sciences. vol.15, pp. 2605–2616, 2015.

[4] X. Dongdong, Z. Jie, X. Jingming, L. Zhenyu, and H. Hai. "The Study on Detection Method of Water Vapor on Boundary Layer Based on Multiagent System," Mathematical Problems in Engineering, vol.6, pp.1-9, 2015.

[5] S.Gutman and S. Benjamin. "The Role of Ground-Based GPS Meteorological Observations in Numerical Weather Prediction," GPS Solution. vol. 4, pp.16–24, 2001.

[6] Z. Liu , M. Wong, J. Nichol, and P. W. Chan. “A multi-sensor study of water vapour from radiosonde, MODIS and AERONET: A case study of Hong Kong," International Journal of Climatology. vol. 33, pp.109–120 (2013).

[7] B. Gao and C. Kaufman, ”Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels," Journal of Geophysical Research Atmosphere, vol.108 , 2003.

[8] C. Yu, N.T. Penna, and Z. Li. “Generation of real-time mode high-resolution water vapor fields from GPS observation,” Journal of Geophysical Research Atmosphere Research, vol.22, pp. 2008-2025, 2017.

[9] P. Albert, R.Bennartz, and J.Fischer. “Remote sensing of atmospheric water vapor from backscattered sunlight in cloudy atmospheres,” Journal of Atmospheric and Oceanic Technology, vol.18, pp. 865–874, 2001.

[10] R. Bennartz and J. Fischer. “Retrieval of columnar water vapour over land from backscattered solar radiation using the medium resolution imaging spectrometer,” Remote Sensing Environment, vol. 78, pp. 274–283, 2001.

[11] H.H. Aumann, M.T. Chahine, C. Gautier, M.D. Goldberg, E. Kalnay , L.M. McMillin , H. Revercomb, P.W.Rosenkranz, W.L.Smith, D.H.Staelin, L.L.Strow, and J.Susskind. “AIRS/AMSU/HSB on the aqua mission: Design, science objectives, data products, and processing systems,” IEEE Transactions on Geoscience and Remote Sensing, vol.41, pp. 253–264 2003.

[12] N. Žagar, “A global perspective of the limits of prediction skill of NWP models,” Tellus: A Dynamic Meteorological Oceanography, vol. 69, pp. 1–15, 2017.

[13] A. Seco, F. Ramírez, E.Serna, E.Prieto, R.García, A.M. Juan, C. Cantera, L.,Jose, and E.Priego. “Rain pattern analysis and forecast model based on GPS estimated atmospheric water vapor content," Atmosric Environvironmet, vol. 49, pp. 85–93, 2012.

[14] A.Caloria, J.R.Santosb, M.Blancob, H.Pessanocd, P.Llamedoe, P.Alexanderf, and A.de la Torree. “Ground-based GNSS network and integrated water vapor mapping during the development of severe storms at the Cuyo region (Argentina),” Atmospheric Research, vol. 176–177, pp. 267–275, 2016.

[15] Y. Yao, L.Shan, and Q. Zhao, “ Establishing a method of short-term rainfall forecasting based on GNSS-derived PWV and its application,” Scientific Report, vol. 7, pp. 1–11, 2017.

[16] K. Boniface, V. Ducrocq, G. Jaubert, X. Yan, P. Brousseau, F. Masson, C. Champollion, J. Chéry, and E. Doerflinger. “Impact of high-resolution data assimilation of GPS zenith delay on Mediterranean heavy rainfall forecasting,” Annual Geophysics, vol. 27, pp. 2739–2753, 2009.

[17] M. Lindskog, M. Ridal, and S.Thorsteinsson, T. Ning. “Data assimilation of GNSS zenith total delays from a Nordic processing centre,” Atmospheric Chemistry and Physics, vol. 17, pp. 13983–13998, 2017.

[18] G. Guerova, J. Jones, J. Douša, G. Dick, S. de Haan, E. Pottiaux, O. Bock, R. Pacione, G. Elgered, H. Vedel, and M. Bender. “Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe,” Atmospheric Measurement Technique, vol.9, pp. 5385–5406, 2016.

[19] Y. Cao, H. Guo, R. Liao, and M. Uradzinski . “Analysis of water vapor characteristics of regional rainfall around Poyang Lake using ground-based GPS observations,” Acta Geodesia et Geophysica, vol. 51, pp.467–479, 2016.

[20] L.F. Sapucci, L. A. T. Machado, E. M de Souza, and T.B. Campos. “GPS-PWV jumps before intense rain events,” Atmospheric Measurement Technique, pp.1–27, 2016.

[21] W. Suparta and R.Rahman ,”Spatial interpolation of GPS PWV and meteorological variables over the west coast of Peninsular Malaysia during 2013 Klang Valley Flash Flood,” Atmospheric Research, vol. 168, pp. 205–219, 2016.

[22] T. Hadas, J. Kaplon, J. Bosy, J. Sierny and K. Wilgan. “Near-real-time regional troposphere models for the GNSS precise point positioning technique,” Measurement Science and Technology, vol.24, 2013.

[23] K. Sato, E. Realini, T. Tsuda, and H. Seko. “A High-Resolution , Precipitable Water Vapor Monitoring System Using a Dense Network of GNSS Receivers,” Journal of Disaster Research, vol. 8, pp. 37–47, 2013.

[24] L.Yang, C.Hill, and T. Moore,” Numerical weather modeling-based slant tropospheric delay estimation and its enhancement by GNSS data,” Geo-spatial Information Science, vol. 16, pp. 186–200, 2013.

[25] Y. Shoji, H. Yamauchi, W. Mashiko, and E. Sato. “Estimation of Local-scale Precipitable Water Vapor Distribution around Each GNSS Station Using Slant Path Delay," Sola, vol.10, pp. pp. 29–33, 2014.

[26] E. Priego, J. Jones, M.J. Porres & A. Seco. “Monitoring water vapour with GNSS during a heavy rainfall event in the Spanish Mediterranean area,” Geomatics, Natural Hazards Risk, vol. 8, pp. 282–294, 2017.

[27] H. Tahami, J. Park, and Y. Choi. “The Preliminary study on the prediction of a Hurricane path by GNSS derived PWV Analysis,” Proc. ION 2017 Pacific PNT Meeting, pp. 500–513, 2009.

[28] J. Shi, Ch. Xu, J. Guo, and Y. Gao. “Real-Time GPS precise point positioning-based precipitable water vapor estimation for rainfall monitoring and forecasting,” IEEE Transaction on Geoscience and Remote Sensing, vol. 53, pp. 3452–3459, 2015.

[29] G. Kiely, J.D. Albertson, M.B. Parlange, and R.W. Katz. “Conditioning stochastic properties of daily precipitation on indices of atmospheric circulation,” Meteorological Application, vol. 5, pp. 75–87, 1998.

[30] U. Dayan, K. Nissen, and U.Ulbrich, “Review Article: Atmospheric conditions inducing extreme precipitation over the eastern and western Mediterranean,” Natural Hazards Earth System Science, vol. 15, pp. 2525–2544, 2015.

[31] F. Tian , Y. Zheng ,T. Zhang , X. Zhang , D. Mao, J. Sun, and S. Zhao. “Statistical characteristics of environmental parameters for warm season short-duration heavy rainfall over central and eastern China,” Journal of Meteorological Research, vol. 29, pp. 370–384, 2015.

[32] S. Wernli and P.HeinI. “Quantifying the Relevance of Cyclones for Precipitation Extremes,” Journal of Climate, vol. 25, pp. 6770–6780, 2012.

[33] I. Prayoga and E.L.Siadari, “Atmospheric conditions analysis of the heavy rain phenomenon in Biak (case study 4-5 December 2014),” Journal of Physics, Solo, Indonesia,2017.

[34] J. Boehm, B.Werl, and H. Schuh. “ Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data," Journal of Geophysics Research. Solid Earth, vol. 111, pp. 1–9, 2006.

[35] B. Hofmann-Wellenhof, H. Lichtenegger, and E.Wasle. ”GNSS- Global Navigation Satellite Systems GPS,GLONASS, GALILEO, and more,” Springer, pp. 105- 160, 2006.

[36] M. Bevis, S. Businger, T. A. Herring, Ch. Rocken, A. Anthes, and R.H. Ware. “GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system,” Journal of Geophysical Research, vol. 97, issue.1578,1992.

[37] C. Rocken, R.Ware, T. Van Hove, F. Solheim, C. Alber, J. Johnson, M. Bevis, S. Businger. “Sensing of Atmospheric Water Vapor with the Global Positioning System,” Geophysical Research Letters, vol.20, issue.125,1993.

[38] E. Realini, K. Sato, T. Tsuda, Susilo, and T. Manik. “An observation campaign of precipitable water vapor with multiple GPS receivers in western Java, Indonesia,” Progress in Earth and Planetary Science, vol.1, pp. 1–15, 2014.

[39] P. Jiang, S. Ye, D. Chen,Y. Liu, and P. Xia. “Retrieving precipitablewater vapor data using GPS zenith delays and global reanalysis data in China,” Remote Sensing, vol.8, 2016.

[40] W.Zhao and M.A. Khalil, “The relationship between precipitation and temperature over the contiguous United States,” Journal of Climate, vol. 11, pp. 1520-0442, 1993.

[41] R. Madden and J.Williams, "The correlation between temperature and precipitation in the United State and Europe,” Monthly Weather Review, vol. 106, pp. 142–147, 1978.

[42] S.Manandhar and Y.H. Lee, “GPS-Derived PWV for Rainfall Now casting in Tropical Region,” IEEE Transaction in Geoscience. Remote Sensing, vol. 56, pp. 4835–4844, 2018.

[43] M. Abadi and M. Kermanshah. “Exploring Spatial Variation in Socioeconomic Determinants of Private Car Ownership,” Transportation Research Board (TRB) 93rd Annual Meeting, Washington D.C, 2013.

[44] A.F. Zuur, E.N. Ieno, and C.S. Elphick, “A protocol for data exploration to avoid common statistical problems," Methods Ecology and Evolution, vol. 1, pp. 3–14, 2010.

[45] S. Wasimi. “A Hybrid Model for Forecasting Daily Rainfall," Water Resource Research, vol. 26, pp. 2741–2746, 1990.

[46] F. Karimi, S. Sultana, A. Shirzadi Babakan, and S. Suthaharan. “An enhanced support vector machine model for urban expansion prediction”. Computers Environment Urban System, vol. 75, pp. 61-75, 2019.