The Utilization GPS Radio Occultation Data to Improve Numerical Weather Prediction Skill through Assimilation Data Procedure Using WRF3DVAR Technique over Jakarta Region

Wasfi Qordowi[1], Adi Mulsandi[2]
School of Meteorology, Climatology and Geophysics Agency
STMKG
Jl. Perhubungan I No. 5, Pondok Betung, Pondok Aren, Tangerang Selatan, 15221 INDONESIA
E-mail : wasfi.qordowi@stmkg.ac.id

***Abstract*—The model of Weather Research and Forecasting-Advanced Research WRF (WRF-ARW) is one of Numerical Weather Prediction (NWP) model which often used to study and to predict weather phenomenon in the atmosphere. Initial condition and boundary condition are two essential elements which needed by WRF model in order to produce forecast. Initial condition is a part of WRF that needs to be corrected in order to make the prediction more accurate. Various methods have been developed to improve the initial condition one of them through data assimilation. There are several methods of assimilation data process which combines NWP products with information from different types of observation, one of them is Three Dimensional Variational (3D-VAR). The purpose of this research is to analyze and compare the accuracy of Weather Research Forecasting (WRF) prediction before and after assimilation the Global Positioning System Radio Occultation (GPS RO) Refractivity data, where the GPS data will be assimilated into the WRF-ARW model by 3D-VAR technique to simulate rain event in Jakarta area on 14 until 16 February 2018. Verification technique to quantify the accuracy of the assimilation model was conducted towards 24 hours accumulated rainfall. The result of this research shows that by applying the data assimilation procedure of the GPS RO Refractivity which goes into WRF-ARW model can increase the accuracy predictions level of heavy rainfall phenomenon which is occurred at that time where able to predict the occurrence for the first category correctly through the percentage of average POD reaching 66% with a prediction error rate of average rainfall (POFD) of 26.1%. Furthermore, for the light rain category, on average only around 59.2% of events can be predicted correctly and with an average percentage of 12.5% prediction errors.**

***Keywords*—Weather Prediction, WRF-ARW, GPS RO Refractivity**

I. INTRODUCTION

One basic reference in making weather forecasting is the output of various types of numerical models starting from short-term to long-term weather forecasting models such as the WRF meso scale (Weather Research and Forecasting) to climate models. Numerical weather forecasts have been carried out more than 150 years ago since the United Kingdom Meteorological Office, UKMO, issued weather forecasts for the public by using weather maps in 1854 [10]. Numerical weather prediction as a form of modern weather forecasting is done based on computer model applications that are able to describe changes in the atmosphere using mathematical equations.[2]

The numerical model requires two important elements in producing forecasts, namely the initial condition and boundary conditions. The target of numerical prediction simulation is to calculate atmospheric conditions based on time [8] which means that the simulation aims to calculate speed, density, pressure, temperature, and humidity (atmospheric parameters) from each point in the air. The accuracy of the initial condition data as the initial input of the model is one part of the problem that still requires improvement and improvement in the NWP model system. Until now, many researches and developments have been carried out to improve the initial conditions of model initialization. One such technique is assimilation of data.

Data assimilation is a technique of combining observation data with NWP products (first guess or initial forecast) to improve estimation and analysis of atmospheric conditions. The data assimilation system will combine all available atmospheric information at a certain time to produce an estimate of its state at a predetermined time analysis [3]. In its application, assimilation is done to improve the initial conditions and perfect the predictions produced by the model.

This study focuses on improving the value of the initial condition of the model as the material for making estimates through assimilating the model data with Global positioning system (GPS) radio occultation (RO), and Satellite rainfall estimation data (GSMAP). Improvements were made using the 3DVAR technique as one of the assimilation techniques into WRF (Weather Research and Forecasting) that is quite easy to do and has been widely used in several previous studies such as by [6] and [7].

Global positioning system (GPS) radio occultation (RO) is an active satellite-to-satellite limb sounding technique in which a transmitter and its corresponding receiver are placed on a GPS satellite and on a low-Earth-orbiting (LEO) satellite. GPS RO sounding data have many unique characteristics, including high accuracy, high vertical resolution, no need for calibration, all-weather sounding capability, and global coverage. Owing to these characteristics, GPS RO data are ideally suited for climate monitoring and weather prediction, and hence they have played an increasingly important role in cli-mate and meteorology studies. Depending on the level of data processing, a variety of data retrieved from GPS RO can be used for weather prediction (e.g., phase and amplitude of GPS signals, bending angle, refractivity, and retrieved temperature and moisture).

In this study, we investigated the impact of GPS RO refractivity on simulations of a heavy rainfall event that occurred over Jakarta region on 14-16 February 2018. The results of the study showed that assimilation carried out had a good influence on the results of the forecast, namely the forecast of the outcome of the model had a better level of accuracy than the non-assimilated model. The purpose of this study was to determine how the comparison of the accuracy and output performance of the WRF model in predicting rain before and after assimilation was carried out using Satellite rainfall estimation data/GSMAP data. The selection of GSMAP data as assimilation data is based on one of them is because precipitation estimation is very important for the NWP model because the model shows a greater level of sensitivity to the data when compared with other data types in the assimilation.

GPS Radio Occultation FORMOSAT-3/COSMIC

The term occultation is a term often used in the field of astronomy to define the event of the closure of a celestial object or object by a larger object or other celestial body [12]. Meanwhile, radio occultation is a technique that uses the frequency of radio signals to obtain information about the vertical gradient of the Refractivity atmosphere, which is related to gradient density (pressure and temperature) and moisture density [18].

The basic measurement used in radio occultation techniques is when radio signals that move from a GPS transmitter to a recipient of Low Earth Orbit (LEO) through the atmosphere will experience bending or curvature caused by varying density differences in each layer of the atmosphere. To measure the atmospheric refraction index (n), the magnitude of the curvature angle (α) that occurs when measuring the atmospheric refraction index (n) can be used to obtain an atmospheric Refractivity value (N) which can be formulated in the following equation :

$$\alpha(a)= 2\int_{r_t}^{\infty} d\alpha = -2a\int_{r_t}^{\infty} \frac{1}{\sqrt{r^2n^2-a^2}}\frac{d\ln(n)}{dr}dr \quad (1)$$

$$n(r) = Exp\left[\frac{1}{\pi}\int_{a_t}^{\infty}\frac{\alpha}{\sqrt{a^2-{a_t}^2}}da\right] \quad (2)$$

$$\mathrm{n} = 1+(\mathrm{N} \times 10^6) \quad (3)$$

$$\mathrm{N} = 77.6\frac{\mathrm{P}}{\mathrm{T}}+3.73 \times 10^5\frac{\mathrm{P_w}}{\mathrm{T}^2}+4.03 \times 10^7\frac{\mathrm{n_e}}{\mathrm{f}^2}+1.4\,\mathrm{W} \quad (4)$$

α = bending angle
a = impact parameter
r = ray trajectory to the center of the arch
R = tangent point
N = refractivity index
N = refractivity
P = atmospheric pressure (hPa)
T = atmospheric temperature (K)
Pw= atmospheric water vapor pressure (hPa)
ne = electron density (Number of electrons / m^3)
W = liquid water content (gr / m^3)
f = transmitter frequency (Hz)

Radio Occultation (RO) data obtained from FORMOSAT-3 /COSMIC has much better data quality when compared to previous missions [12]. The satellite circuit was first launched from the Vadenberg air force center in California, U.S. on April 15, 2006.

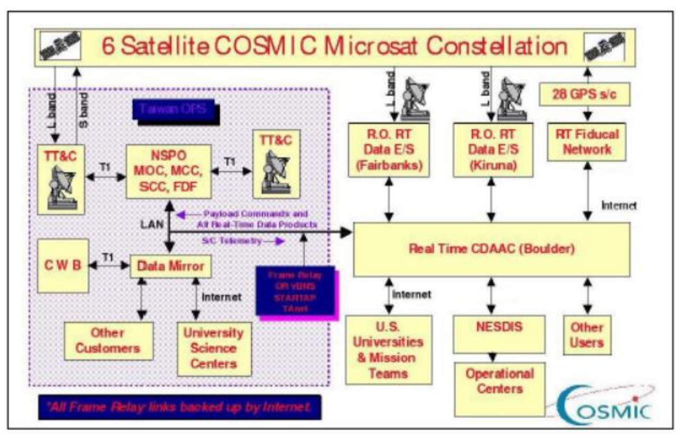

Fig. 1. COSMIC land and satellite components

The results of research using COSMIC data have also proven theoretically that RO sounding has a fairly high accuracy. Accuracy in the upper troposphere or lower stratosphere reaches up to 0.05 K or higher. For example, the comparison between COSMIC data and AMSU-A data obtained from the National Oceanic and Atmospheric Administration (NOAA) satellite shows a high correlation (~ 0.99 or higher) on brightness temperature with a standard deviation ranging from 0.95 K and 0.97 K. This shows that the observation of RO has a significant positive impact on global and regional weather forecasts, as well as on the estimation of important weather phenomena such as tropical cyclones (tropical cyclone) [16].

II. WRF 3DVAR TECHNIQUE ASSIMILATION

Data assimilation is a technique where data is observed combined with a numerical weather model (NWP) good for the first guess and background forecast as well as statistical errors covering it to predict better atmosphere. Assimilation of data Varitional (Var) achieves this through a minimum iteration of the value function (cost function) that has been determined [22]. The data assimilation technique has the purpose of using all information available todetermine the state of the atmosphere as accurately as possible.[5]

3DVAR assimilation is one part of the data assimilation system varitional which aims to produce optimal estimates of the situation the actual atmosphere in a time analysis through an iterative solution fromcost function [11] as follows:

$$J(x)=J_b(x)+J_o(x)=\frac{1}{2}(x-x^b)^T B^{-1}(x-x^b)+\frac{1}{2}(y-y^o)^T(E-F)^{-1}(y-y^o) \quad (4)$$

$Jb\ (x)$ = first guess cost function (background)

$Jo\ (x)$ = cost function observation

x = expected analysis data

xb = first guess data (background)

y = observation / observation data

Yo = observation in the grid model

B = background error covariance matrices

E = observation error covariance matrices

F = representivity error covariance matrices

In 3DVAR assimilation, Verification of predicted results has an important role in the quality of the analysis produced. Second it determines to what extent the background will be corrected so according to observation.Verification of the predicted results of spatial rainfall distribution was also carried out using the BMKG daily rainfall category. The threshold used in accordance with the BMKG predetermined category is as follows: rain (> 1mm / day), light rain (> 20mm / day), heavy rain (> 50mm / day), and very heavy rain (> 100mm / day) [21]. Comparison of performance values and the ability of both models in predicting rainfall distribution is analyzed through two indices, namely the POD and POFD values where the value of the prediction error average rainfall.

III. RESEARCH METHOD

a. Data

The data used in the study consisted of :
1. GFS data (Global Forecast System) resolution of 0.250 x 0.250 with the initiation of February 14, 2018 at 00 until February 16, 2018 12 UTC
2. GPS RO Refractivity data from February 14, 2018 at 00 to February 16, 2018 12 UTC which can be obtained from cosmic.ucar.edu as assimilation data
3. Background Error (BE) data as model correction data on predictions that have been provided in the WRFDA application on the default BE option (CV3).
4. Observation data from the Meteorological Station and rainfall posts representing the Jabodetabek area as verifier data per point.
5. Satellite rainfall estimation data (GSMAP) with a resolution of 0.250 x 0.250 as a spatial verifier that occurs.
6. Final Data Analysis (FNL) with a resolution of 1 ° x 1 ° on the date from February 14, 2018 at 00 UTC to February 16, 2018 12 UTC obtained from rda.ucar.edu

b. Research site

The research location is between 104.59^0 - 109.08^0 BT and 4.02^0 - 8.48^0 LS with the main focus of the research being the Jakarta area.

c. Research mind

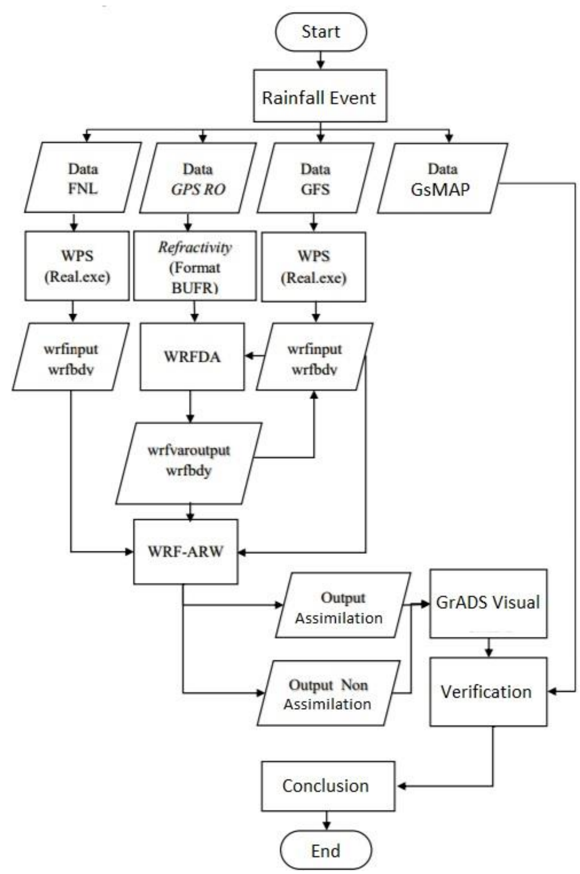

Fig. 2. Research Flow Diagram

IV. SIMULATION AND RESULT

a. Comparison and Verification of 24hour rain event

Spatial verification is done by comparing the rainfall observation data of daily GSMAP accumulation with the results of the model resolution of 0.25^o x 0.25^o. Verification of spatial model performance tests is carried out, namely the forecast of rain dichotomy (yes / no rain) and forecast of rain using the threshold according to the daily rainfall category of BMKG.

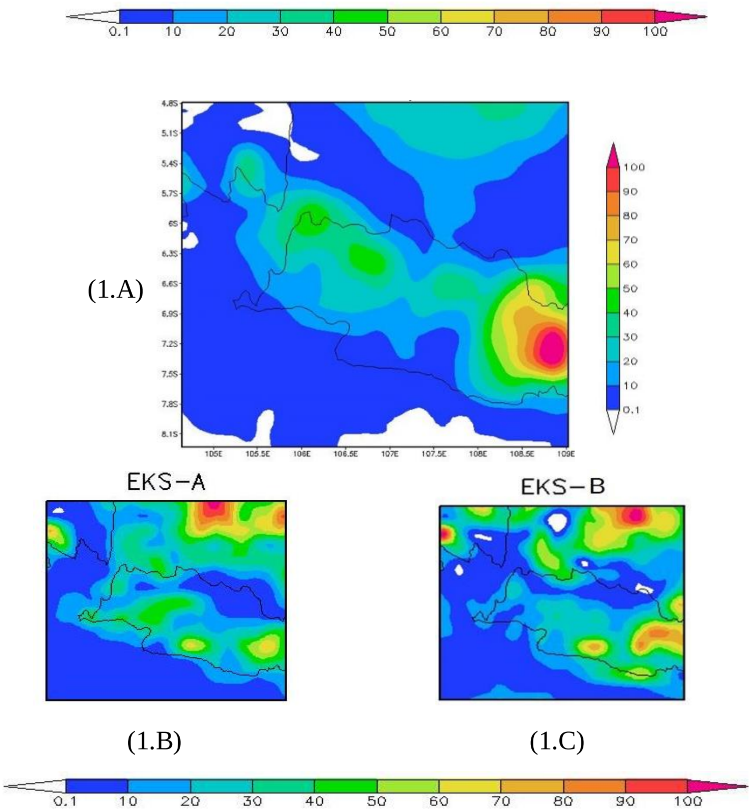

Figure 3. (1.A) GsMAP data 14 February 2018 ; (1.B) No Assimilation ; (1.C) GPS RO Assimilation

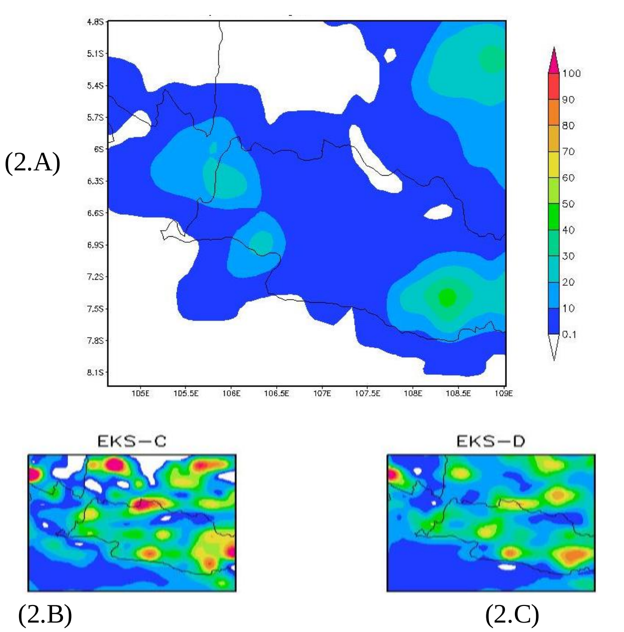

Fig. 4. (2.A) GsMAP data 15 February 2018 ; (2.B) No Assimilation ; (2.C) GPS RO Assimilation

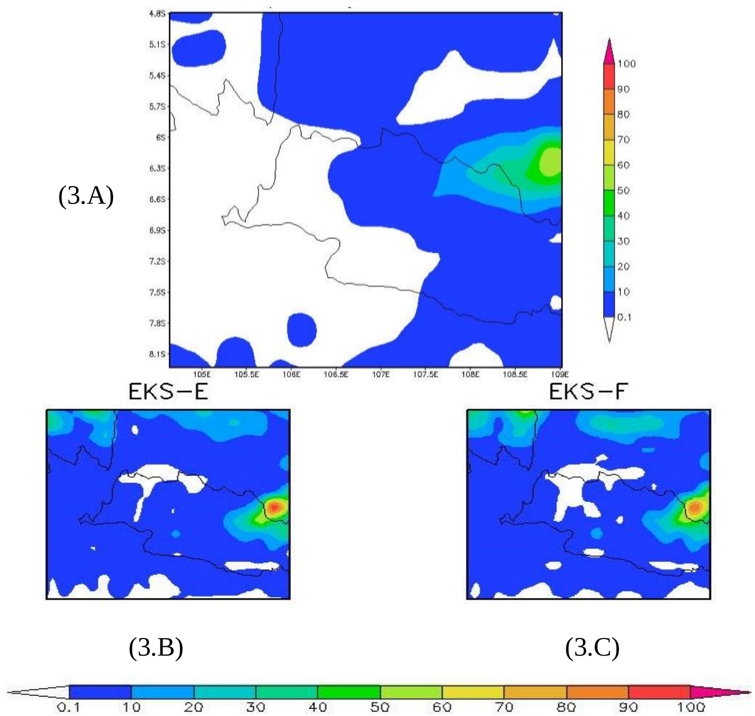

Fig. 5. (3.A) GsMAP data 16 February 2018 ; (3.B) No Assimilation ; (3.C) GPS RO Assimilation

Based on the subjective analysis of the outputs of the two models above, it is shown that there was a good improvement after assimilation. This can be seen from the distribution of patterns and similarities formed in the assimilation model with a fairly good degree of similarity when compared with GSMAP data.

Furthermore, in order for the analysis to be more objective, some statistical indexes are calculated, namely POD, FAR, TS, BIAS, and POFD as a result of comparison between GsMAP observation data and model predictions. Based on the calculation of several indices, it is shown that the two models have different values and different fluctuations in showing the spatial distribution of rainfall.

Based on index calculations on both models and trends for three days, the assimilation model was able to show consistency to have a POD and TS index value that was greater than the non-assimilation model. In addition to the BIAS, FAR, and POFD indices, the assimilation model has an index value smaller than the non-assimilation model.

From all the results of the analysis of the distribution pattern of rain distribution that occurred during the four days of research subjectively and objectively through the calculation of several statistical indices, assimilation models have outputs that are closer to GSMAP observation data, although a little more overestimate for the amount of rainfall especially in Jakarta.

b. Verification of spatial rainfall accumulation using the rainfall classification

Verification of the predicted results of spatial rainfall distribution was also carried out using the daily rainfall category. Comparison of performance values and the ability of both models in predicting rainfall distribution is analyzed through two indices, namely the POD and POFD values.

POD (Probability of Detection) values is a calculation to measure the number of "yes" observations that are predicted correctly. This index shows the probability of events that can be detected by the model. High values are used for overforecasting events. The index value is between 0 and 1 with the best value when POD is 1.

POFD (Probability of False Detection) value is the possibility of many "yes" occurrences according to predictions but does not occur. The range of POFD values ranges from 0 to 1. A value of 0 indicates a perfect value.

The POD and POFD values for each rain category for the four days of the study are shown in Table 1.

Category	Index	Non-Assimilation	Assimilation
Rainfall >1mm/day	POD	0.661	0.662
	POFD	0.265	0.257
Rainfall >20mm/day	POD	0.594	0.596
	POFD	0.120	0.119
Rainfall >50mm/day	POD	0.132	0.157
	POFD	0.036	0.036
Rainfall >100mm/day	POD	0.025	0.025
	POFD	0.005	0.004

Table 1. The accumulation of the average POD and POFD indexes of daily rainfall distribution according to the daily rainfall category of BMKG

Both models are able to predict the occurrence for the first category correctly through the percentage of average POD reaching 66% with a prediction error rate of average rainfall (POFD) of 26.1%. Furthermore, for the light rain category, on average only around 59.2% of events can be predicted correctly and with an average percentage of 12.5% prediction errors.

The ability to predict heavy rain correctly (hits) as a whole is only about 55 events and very thick only one incident. For heavy rain only around 13.5% and only 2.2% for very heavy rain. Overall, there was a decline in the quality of predictions quite dramatically, especially in the category of heavy rain and very heavy.

But based on the consistency of the model's performance for four days for all of these categories, the assimilation model has a better level of consistency in predicting rain based on the above categories with the POD and POFD values better than the non-assimilation models.

c. Verification of Model Performance at Observation Points

Verification to see the ability of the model to produce the next prediction is through verification using observational data from the observation points in the Jakarta area which is the research area and its surroundings which cover Depok, Bogor, Tangerang, and Bekasi (Jabodetabek) areas. The analysis was carried out objectively through the standard deviation, correlation, and RMSE index plotted into the Taylor diagram.

The results of the comparison between model output data and station observation data and rain observation post for 24 hours daily rainfall show overall at all points for four days, from the values of the three indices, the non assimilation model has better performance than the assimilation model. This is indicated by the position of point A (as a non assimilation model) which has a distance closer to point O (observation) than the assimilation model (point B) on the Taylor diagram.

Based on further analysis at each observation point, it was found that the non-assimilation model performed better, especially for the Rorotan region on the first day, Kemayoran and Curug on the second and third days.

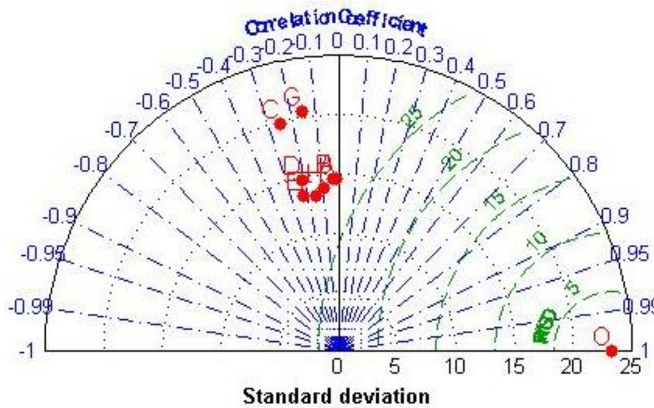

Fig. 6. Taylor's statistical index diagram model experiment and third-day rainfall observation of all observation points

V. CONCLUSIONS

From the results of the analysis carried out by comparing the results of the non-assimilation model with the assimilation model of the observation data in the case of rain events on February 14-16, 2018 in the Greater Jakarta area (Jabodetabek), the following conclusions were produced:

1. In general, there is an increase in forecast accuracy after assimilating data from Jason. There was an improvement in the predictive results of the WRF model until the third day of the initial initiation time despite very significant improvements.

2. Improving the prediction of the assimilation model for rain parameters is for spatial rainfall (using the BMKG rain category). Whereas for the rain at the observation point, based on the correlation index value and assimilation RMSE carried out did not give a positive influence on the results of the model.

REFERENCES

[1] Bao, Y., Huang X., Min J., Zhang Xin., dan Xu D., 2010, Application Research of Radiance Data Assimilation in Precipitation Predicton Based on WRFDA[online],http://www2.mmm.ucar.edu/wrf/users/workshops /WS2010/abstracts/P-07.pdf, diakses Access 1st december 2018

[2] Golding, B., 2013, Numerical Weather Prediction NWP [online],*Royal Meteorology Society* [online], https://www.rmets.org/weather-andclimateweather numericalweather-prediction-nwp, 1st december 2018

[3] Barker, D. M., Huang W., Guo Y R., dan A. Bourgeois., 2003, *A three-dimensional variational (3DVAR) data assimilation system for use with MM5*. NCAR Tech. Note. NCAR/TN-453 1 STR, 68 pp.

[4] Ferdousi, N., Debsarma S K., dan Mannan Abdul M.D., 2015, *Impact of Data Assimilation in Simulation of Thunderstorm Event over Bangladesh Using WRF Model*, K. Ray et al. (eds.), *High-Impact Weather Events over the SAARC Region*, Springer, New Delhi.

[5] Talagrand, O., 1997: Assimilation of Observations, An Introduction. J. Met Japan Special Issue, 1B, 191-209.

[6] Gustari, I., 2014. Perbaikan Prediksi Cuaca Numerik Kejadian Hujan Sangat Lebat Terkait dengan Sistem Awan di Jabodetabek Menggunakan Asimilasi Data Radar C-Band. *Disertasi*, ITB, Bandung.

[7] Paski, J., 2016, Perbandingan Hasil Model Numerik WRF dengan Model Asimilasi Data Radar dalam Prediksi Cuaca (Studi Kasus di Lampung). *Skripsi*, Meteorologi, Diploma IV STMKG, Jakarta.

[8] Wiegand, B, 2015, *Introduction to Numerical Weather Prediction*. htw saar – Hochschule für Technik und Wirtschaft des Saarlandes (University of Applied Sciences). Fakultät für Ingenieurwissenschaften.

[9] Athumani, C., 2012, Implementation of WRF-3DVAR data Assimilation over East Africa, A Case Study of Tanzania[online], http://www.geus.dk/DK/int_dev el_projects/Documents/Chuki_DataAssmPresentation.pdf, Access 1st december 2018

[10] Gustari, I., Hadi T., Hadi, S., dan Renggono, F., 2012, Akurasi Prediksi Curah Hujan Harian Operasional di Jabodetabek: Perbandingan dengan Model WRF, Jurnal Meteorologi dan Geofisika, Vol.1, no.2, 119-130, Badan Meteorologi Klimatologi dan Geofisika.

[11] Ide, K., Courtier,P., Ghil,M., dan Lorenc, A.C., 1997. Unified notation for data assimilation: Operational,sequential and variational. Journal of Meteorology Society. Vol. 75, pp 181-189.

[12] Yen, N.L., Fong, C.-J., Chu, C.-H., Miau, J.-J., Liou, Y.-A., Kuo, Y.-H., 2010, Global GNSS Radio Occultation Mission for Meteorology, Ionosphere & Climate, Arif,T.T,Aerospace technology Advancements,1st ed , Intech, India.

[13] Kalnay, E., 2003. Atmospheric Modelling, Data Assimilation and Predictability. Cambridge, UK: Cambridge University Press.

[14] Kizhner, L.I., A.A. Bart., D.P Nahtigalova., 2013, Using the Numerical WRF Model for the Prediction of Weather Parameters in Tomsk Region, BioClimLand, No. 1, pp 29-35.

[15] Liu, J., M. Bray., dan D.Han., 2013, A Study on WRF Radar Data Assimilation for Hydrological Rainfall Prediction, Hydrological Earth System, Vol.17 pp 3095-3110.

[16] Anthes, R.A., 2011, Exploring Earth's Atmosphere With Radio Occultation: Contributions to Weather, Climate and Space Weather,Atmospheric Measurement Techniques, No. 1, 135-212.

[17] Maiello, I., R. Ferretti, S. Gentile, M. Montopoli, E. Picciotti, F. S. Marzano, dan C. Faccani,2013, Impact of radar data assimilation using WRF threedimensional variational system, Atmospheric Measurement Techniques, Vol. 6, pp 7315-7363.

[18] Melbourne, W.G., Davis, E.S., Duncan, C.B., Hajj, G.A., Hardy, K.R., Kursinski, E.R., Meehan, T.K., Young, L.E., Yunck, T.P., 1994, The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring, Jet Propulsion Laboratory Report 94-18.

[19] Rinehart, R. E., 2010. Radar for Meteorologist Fifth Edition, Nevada, Missouri, Rinehart Publications. Satrya, L. I., 2012, Asimilasi Data Radar dalam Penerapan Prediksi Cuaca Numerik di Indonesia (Studi Kasus di Jawa Barat). Meteorologi ITB: Bandung.

[20] Schulze, G.C. 2007, Atmospheric Observation and Numerical Weather Prediction. South African Journal of Science, Vol. 103, pp 318-323.

[21] BMKG, 2010, Peraturan Kepala Badan Meteorologi Klimatologi dan Geofisika Nomor: KEP.009 Tahun 2010 tentang Prosedur Standar Operasional Pelaksanaan Peringatan Dini, Pelaporan, dan Desiminasi Informasi Cuaca Ekstrim, BMKG, Jakarta.

[22] Wang, H., Bruyère, C., Duda, M., Dudhia, J., Gill, D., Lin, H. C., Michalakes, J., Rizvi, S., dan Zhang, X., 2016, WRF-ARW Version 3 Modeling System User's Guide, National Center for Atmospheric Research, Amerika Serikat.