Ensemble Physics of the Weather Research and Forecasting (WRF) Model for Predicting Heavy Rainfall in the Bandung area, West Java

Abdul Hamid Al Habib*,[1,2] Fathan Ilham Ramadhani,[2] and Nurjanna J. Trilaksono[3]

[1]*Department of Meteorology, Faculty of Earth Sciences and Technology, Bandung Institute of Technology, Bandung, Indonesia*
[2]*Gusti Syamsir Alam Meteorological Station (Kotabaru), Meteorology Climatology and Geophysics Agency, Indonesia*
[3]*Atmospheric Science Research Group, Faculty of Earth Sciences and Technology, Bandung Institute of Technology, Bandung, Indonesia*

Abstract: The complex topography of the Bandung region, with the presence of mountains and valleys, can affect air flow patterns and rainfall distribution. Accurate weather predictions and spatial precision are crucial for anticipating the impacts of heavy rainfall. This study aims to evaluate the capability of the WRF physics ensemble prediction system in forecasting heavy rainfall events in the Bandung region. The use of an ensemble prediction system is a viable approach to quantifying uncertainty in numerical weather prediction and provide more reliable information. The case study used is the heavy rainfall event that caused flooding on October 4, 2022, in the Pagarsih area. Global Forecasting System (GFS) data with a spatial resolution of 0.25 x 0.25 and a temporal resolution of three hours were used as input for downscaling in the WRF-ARW model. This study used 9 configuration schemes of the WRF-ARW model parameterization as ensemble members. The results of the study indicate that the WRF model (a combination of the Purdue Lin, Yonsei University Scheme, and Betts-Miller-Janjic Scheme) provided the most accurate heavy rainfall prediction, with an RMSE value of 2.13. The probability maps of rainfall products can effectively identify peak heavy rainfall between 1:00 PM - 4:00 PM. This is indicated by the large area with a greater than 90% probability of rainfall exceeding 10 mm. The ensemble mean product of rainfall predictions tends to underestimate heavy rainfall in the Pagarsih area. The ensemble mean product of surface air temperature can effectively identify the pattern of observational fluctuations with a low RMSE value (0.77), and the ensemble mean product of surface layer air humidity can identify the pattern of observational fluctuations with a relatively high RMSE value (13.28).

Keywords: Ensemble; Parameterization; WRF-ARW; Forecast; Rainfall

*Corresponding author: jpatiani@itb.ac.id

http://dx.doi.org/10.12962/j24604682.v21i1.18883

I. INTRODUCTION

Extreme rainfall tends to trigger natural disasters such as floods, landslides, and debris flows, which can cause significant losses both in terms of human and economic aspects [1]. Numerous studies have documented a global increase in the frequency and intensity of extreme rainfall events [2-4]. Complex topographical features, variations in land use, and distance from the sea are some of the key characteristics that affect synoptic scale and local weather conditions [5-7].

Convective activity in the Bandung Basin is local and can occur due to the unique and complex topography, consisting of a series of mountains in the north, east, and south. Since the 1960s, the Pagarsih area in Bandung has frequently experienced high rainfall and recurrent flooding. The impacts of flooding in Pagarsih include loss of life, infrastructure damage, economic losses, transportation disruptions, and health risks. Therefore, evaluating the predictability of models in this area is crucial. Developing spatially and temporally responsive simulations through numerical weather prediction (NWP) models is essential to anticipate the potential consequences of disasters. According to a study by Prein *et al.* (2015), by handling initial and lateral boundary conditions and appropriate physical model configurations, convection simulations (with resolutions finer than 4 km) using NWP models show significant potential in improving rainfall forecasts [8].

Forecasting the spatial and temporal variations of extreme rainfall is a challenging issue, especially in regions with complex topography [7, 9]. As a more advanced generation of numerical weather prediction models, the WRF (Weather Research and Forecasting) model has been widely used in over 150 countries for atmospheric research and operational forecasting purposes [10, 11]. However, WRF and similar numerical models are imperfect due to measurement, analysis errors, and biases. Initial conditions rarely match reality, with errors growing over time, especially at smaller scales [12, 13]. Moreover, the model equations do not fully encompass all atmospheric processes. Therefore, ensemble methods were developed to quantify uncertainty and provide more reliable information to users.

Ensemble prediction systems can produce probabilistic forecasts that provide users with information on the likelihood of an event. Research by Joslyn and Savelli (2010) shows that with probabilistic forecasts, users can make more accurate

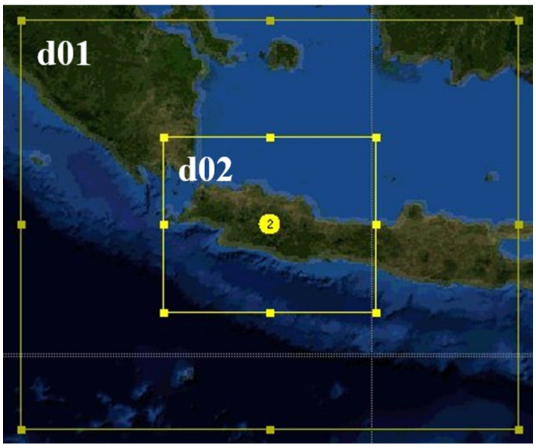


FIG. 1: The domain configuration of the WRF ARW model is centered on the research area, specifically in the Pagarsih region, Bandung.

decisions [14]. Additionally, probabilistic forecasts can better maintain user confidence in the forecast results compared to deterministic forecasts [15]. This study aims to evaluate the capability of the ensemble physics of the WRF model in forecasting heavy rainfall events in the Bandung region. The ensemble physics consists of several members, each utilizing different configurations of parameterization schemes within the WRF model.

II. METHODOLOGY

This study uses the Weather Research and Forecasting-Advanced Research WRF (WRF-ARW) model version 4.4.2 to process GFS data. The Global Forecasting System (GFS) data with a spatial resolution of 0.25° × 0.25° and a temporal resolution of three hours is processed using the WRF-ARW model. The GFS data has a grib file extension. GFS data is global model prediction data released by National Oceanic and Atmospheric Administration (NOAA), with prediction times every three hours and analysis cycles every six hours, accessible via https://nomads.ncdc.noaa.gov/data/gfs4/. The data period taken is from October 3, 2022 (00 Universal Time Coordinated, UTC) to October 6, 2022 (00 UTC).

In processing data using WRF-ARW, it is necessary to configure the model to determine the domain, input data, parameterization schemes, and grid resolution to be used (Table I). WRF is run to downscale GFS data at 00:00 UTC on October 3, 2022, with a prediction time of up to three days ahead.

This study aims to simulate heavy rainfall predictions that could trigger flooding in the Pagarsih area of Bandung. Pagarsih is located at the coordinates 6°55'22"S 107°35'39"E (Fig. 1). In this study, heavy rainfall is categorized as occurring when the precipitation exceeds 10 mm over a 3-hour period. The Pagarsih area of Bandung falls within two districts: Astanaanyar District and Bojongloa Kaler District. In this study, one parent domain and a two-domain setup (one-time nesting) are used (Fig. 1).

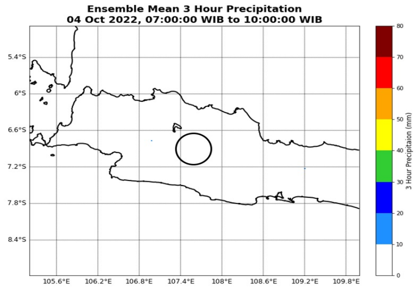
TABLE I: WRF ARW model configuration in two domains.

WRF Settings	Information	
	Domain 1 (9 km)	Domain 2 (3 km)
history_interval	180 seconds	60 seconds
e_we	140	178
e_sn	120	154
e_vert	28	28
ra_sw_physics	Dudhia Shortwave Scheme	Dudhia Shortwave Scheme
ra_lw_physics	RRTM Longwave Scheme	RRTM Longwave Scheme
bl_pbl_physics	YSU scheme	YSU scheme
cu_physics	Kain Fritsch Betts Miller Janjic GrellDevenyi	0
mp_physics	WSM 6 class Kessler Purdue Lin	WSM 6 class Kessler Purdue Lin

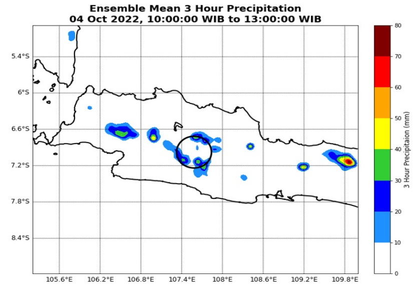
TABLE II: Configuration of the WRF ARW model parameterization scheme used as an ensemble member.

Name	Parameterization Scheme		
Ensemble member	Microphysics	Planetary Boundary Layer	Cumulus
Member 1 Member 2 Member 3	WRF Single-Moment 6-Class (WSM6)	Yonsei University Scheme (YSU Scheme)	Kain-Fritcsh Betts Miller Janjic Grell Devenyi
Member 4 Member 5 Member 6	Kessler	Yonsei University Scheme (YSU Scheme)	Kain-Fritcsh Betts Miller Janjic Grell Devenyi
Member 7 Member 8 Member 9	Perduelin	Yonsei University Scheme (YSU Scheme)	Kain-Fritcsh Betts Miller Janjic Grell Devenyi

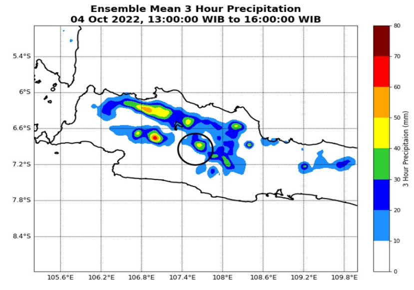
Nine configuration schemes of the WRF-ARW model (Table II) are utilized to obtain initial data as ensemble members. The selection of parameterization schemes in Table II is based on literature studies identifying the best parameterization schemes for simulating heavy rainfall events. After configuring the model, the next step is to perform computations with the model. This computation process will generate wrfinput_d0x, wrfbdy_d0x, and wrfout data. The wrfout data is the main output file containing the results of the atmospheric simulation. This file stores numerical data for various meteorological variables such as temperature, pressure, humidity, wind, and other parameters predicted by the WRF model. The wrfout data output from WRF in domain 2 is then visualized and extracted to produce ensemble members. These ensemble members are used as inputs for various ensemble products to predict extreme heavy rainfall events.



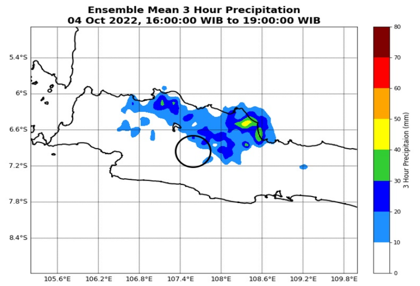
(a)October 4, 2022, from 7:00 AM to 10:00 AM



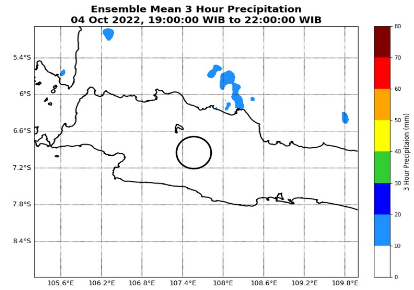
(b)October 4, 2022, from 10:00 AM to 13:00 AM



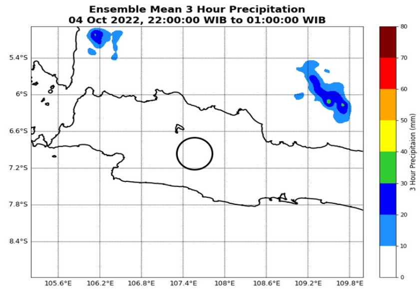
(c)October 4, 2022, from 1:00 PM to 4:00 PM



(d)October 4, 2022, from 4:00 PM to 7:00 PM



(e)October 4, 2022, from 7:00 PM to 10:00 PM



(f)10:00 PM October 4, 2022, to 01:00 AM October 5, 2022

FIG. 2: Ensemble mean rainfall (3 hours) on October 4 2022 at 07.00 AM to October 5 2022 at 01.00 PM.

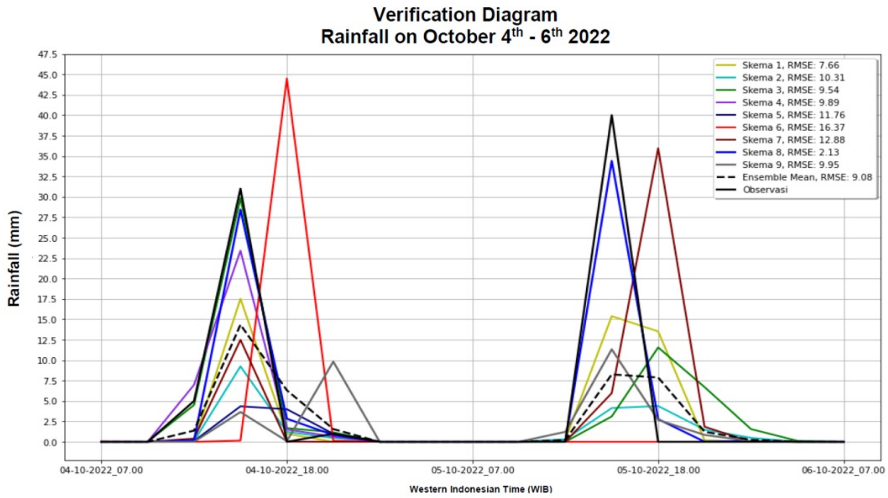


FIG. 3: Comparison diagram of rainfall from each WRF parameterization scheme (ensemble member) and ensemble mean against observed rainfall at Husein Sastranegara, Bandung (7:00 AM on October 4, 2022, to 7:00 AM on October 6, 2022).

There are several types of ensemble methods according to WMO No. 1091, including the following:

1. Ensemble Mean
   The Ensemble Mean method involves averaging all parameter values across all ensemble members. The average result from the model runs can be used as a prediction result;

2. Ensemble Spread
   The growth of error values over time throughout the prediction period causes the model solutions to become increasingly divergent. This divergence of model solutions indicates the level of uncertainty of a condition. Ensemble spread is the standard deviation value of all ensemble members; the larger the standard deviation, the further the data deviates from the mean value. This condition also indicates a high level of uncertainty;

3. Basic Probability
   Basic probability displays the likelihood of an event or parameter from part of the ensemble members at a specific grid point or location. Probability is defined as the

simple proportion of ensemble members that predict a certain phenomenon at a certain point. This probability product is generated from the model consensus calculation for a specific value.

The data used for verification are direct weather observation data from the Husein Sastranegara Weather Observation Station in Bandung. This station is the closest observation point to the Pagarsih area. The observatory data obtained is in ASCII format with a temporal resolution of 3-hour accumulated rainfall, so the temporal resolution of the WRF prediction data used for verification is also 3 hours.

The method used to verify the model output against the observational data in this study is the Root Mean Square Error (RMSE). The Root Mean Square Error (RMSE) is a measurement method that measures the difference between the predicted values of a model and the observed values. The accuracy of the measurement error estimation method is indicated by a low RMSE value. An estimation method with a lower RMSE is considered more accurate than one with a higher RMSE.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q_\circ - Q_s)^2}{n}} \quad (1)$$

information:
$Q_\circ$ = actual data value
Q_s = forecasted value
n = number of data points
Σ = summation (total sum of values)

III. RESULT AND DISCUSSION

A. Ensemble Mean Precipitation

Based on the results of the ensemble mean precipitation every 3 hours (Fig. 2), it can be identified that rainfall of 10-20 mm started to occur between 10:00 AM and 1:00 PM. The intensity of the rainfall increased to 40-50 mm from 1:00 PM to 4:00 PM, and then decreased to 10-20 mm from 4:00 PM to 7:00 PM in the West Java region. The ensemble mean product tends to underestimate the peak heavy rainfall in the Pagarsih area, which occurred between 1:00 PM and 4:00 PM.

Fig. 3 represents the comparison of RMSE values for each member and the ensemble mean for the rainfall variable compared to observed rainfall data at the Husein Sastranegara Weather Station. The figure shows that member 8 (parameterization scheme 8) is the most accurate compared to other members and the ensemble mean product. This indicates that member 8 provides the best rainfall prediction among all members.

The test of parameterization scheme 6 (member 6) of the WRF model in the Bandung area, West Java, resulted in the worst rainfall accumulation compared to other parameterization schemes (Fig. 3). Therefore, in identifying heavy rainfall events in the Pagarsih area, Bandung, West Java, the combination of the Kessler scheme (microphysics) and the Grell-Devenyi scheme (cumulus) is less suitable for this case. The Kessler scheme is a relatively simple cloud microphysics scheme, as it does not account for several complex processes such as cloud phase transitions and tends to focus on isolated cloud microphysics processes (not considering external factors such as wind patterns, advection, or interactions with the broader atmospheric environment). Consequently, this scheme may not accurately model some more complex cloud phenomena [16]. Meanwhile, the Grell-Devenyi ensemble scheme currently cannot handle ideal convection on a small grid [17].

B. Probability Map Precipitation

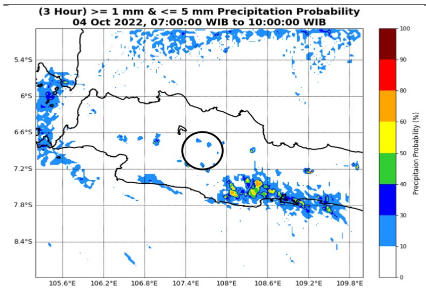
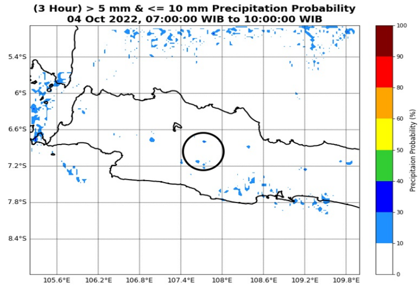
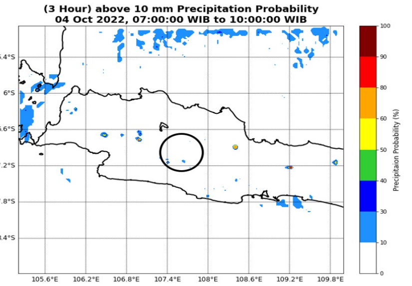
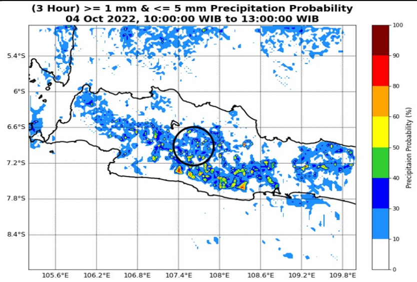
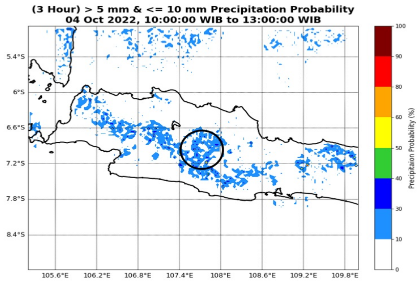
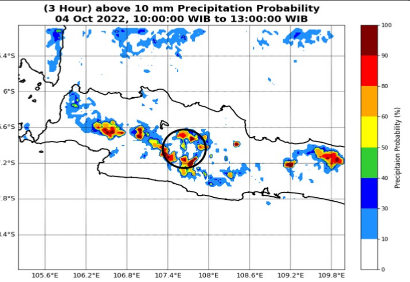
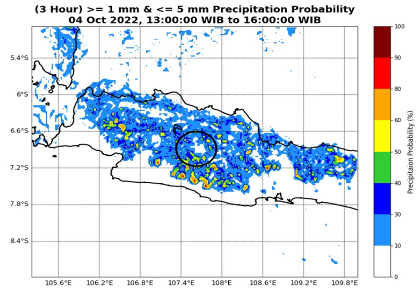
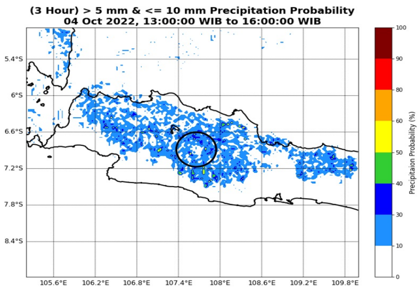
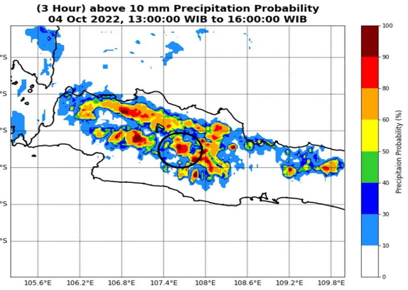
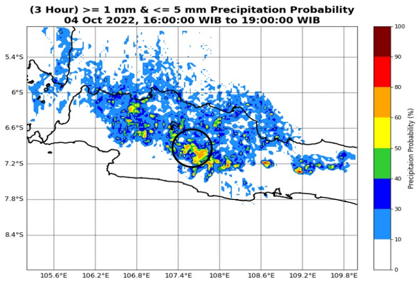
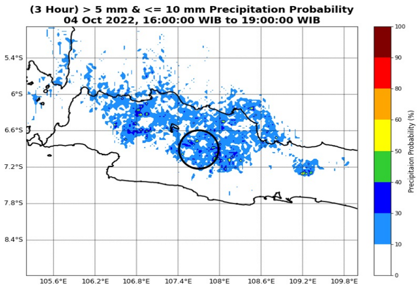
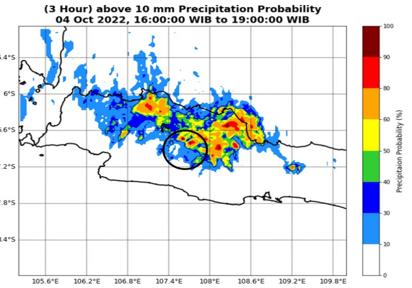
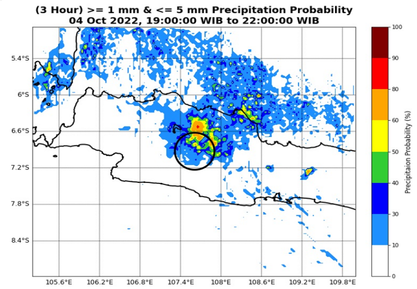
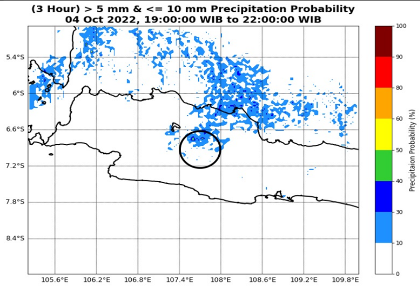
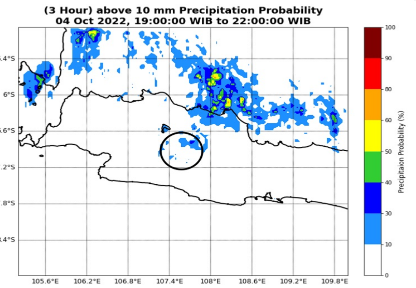
Based on the probabilistic prediction of rainfall events (Table III), there is a likelihood of heavy rainfall (threshold of more than 10 mm/3 hours) in Bandung City, West Java, on October 4, 2022, from 10:00 AM to 7:00 PM. The peak rainfall is expected between 1:00 PM and 4:00 PM, indicated by a predicted area with more than a 90% chance of experiencing rainfall greater than 10 mm/3 hours. Furthermore, the probabilistic prediction of light rain in Bandung, West Java, starts from 4:00 PM to 10:00 PM. This aligns with observed rainfall data, suggesting that the rainfall probability maps have good skill in predicting heavy rainfall events, which is one of the factors causing floods in the Pagarsih area.

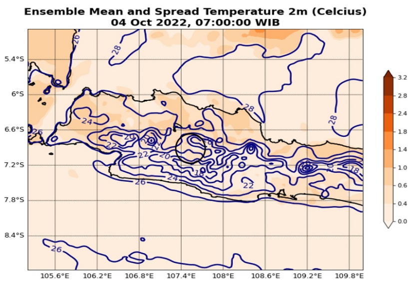
C. Ensemble Mean and Spread (Surface Air Temperature)

Fig. 4 represents the surface temperature using the ensemble mean and spread prediction method from October 4, 2022, at 7:00 AM to October 4, 2022, at 10:00 PM. During the rainfall event from 1:00 PM to 4:00 PM, the standard deviation of surface air temperature in the Pagarsih area is the highest compared to other times. This variation is attributed to the differing rainfall intensities produced by each member of the ensemble. These differences in rainfall intensity can significantly impact the variation in surface air temperature (2 meters), highlighting the role of parameterization schemes in predicting temperature fluctuations. The figure further shows that during the heavy rainfall in the Pagarsih area, Bandung, the dominant surface air temperature ranges from 20°C to 22°C, with a standard deviation between 0 and 1.8, indicating relatively low uncertainty.

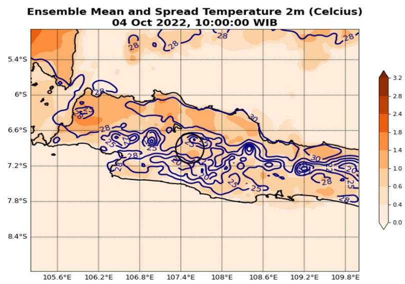
The comparison of RMSE values for each member and the ensemble mean for the 2-meter surface air temperature variable, when compared to the observed surface air temperature at the Husein Sastranegara Station, is presented in Fig. 5. According to Fig. 5, the ensemble mean with an RMSE value of 0.77, provides the most accurate prediction compared to the individual members.

TABLE III: Probabilistic prediction of 3-hour rainfall events with three categories of rainfall intensity (from 7:00 AM on October 4, 2022, to 1:00 AM on October 5, 2022).

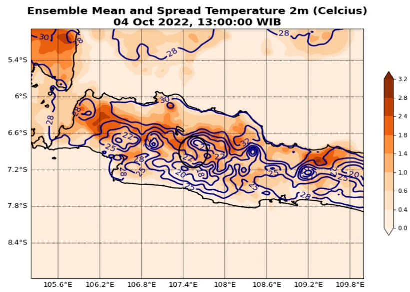
1 <= Rainfall <= 5 (mm)	5 < Rainfall <= 10 (mm)	Rainfall > 10 (mm)
 Oct 4, 07:00 to 10:00 AM	 Oct 4, 07:00 to 10:00 AM	 Oct 4, 07:00 to 10:00 AM
 Oct 4, 10:00 AM to 1:00 PM	 Oct 4, 10:00 AM to 1:00 PM	 Oct 4, 10:00 AM to 1:00 PM
 Oct 4, 1:00 to 4:00 PM	 Oct 4, 1:00 to 4:00 PM	 Oct 4, 1:00 to 4:00 PM
 Oct 4, 4:00 to 7:00 PM	 Oct 4, 4:00 to 7:00 PM	 Oct 4, 4:00 to 7:00 PM
 Oct 4, 7:00 to 10:00 PM	 Oct 4, 7:00 to 10:00 PM	 Oct 4, 7:00 to 10:00 PM



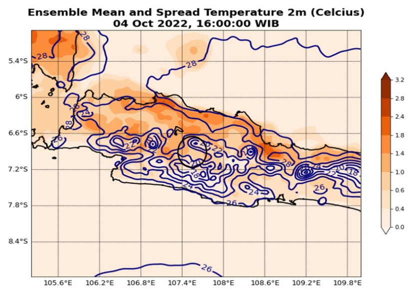
(a)October 4, 2022 at 07:00 AM



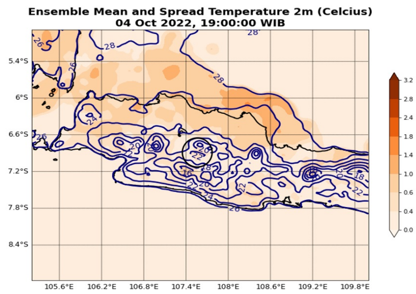
(b)October 4, 2022 at 10:00 AM



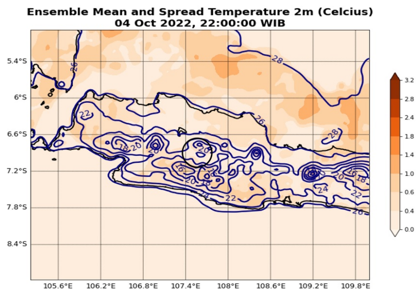
(c)OctoberOctober 4, 2022 at 1:00 PM



(d)October 4, 2022 at 4:00 PM



(e)October 4, 2022 at 7:00 PM



(f)October 4, 2022 at 10:00 PM

FIG. 4: Ensemble mean (contour) and ensemble spread (shaded) of surface air temperature (2 meters) on October 4, 2022, from 07:00 AM to 10:00 PM.

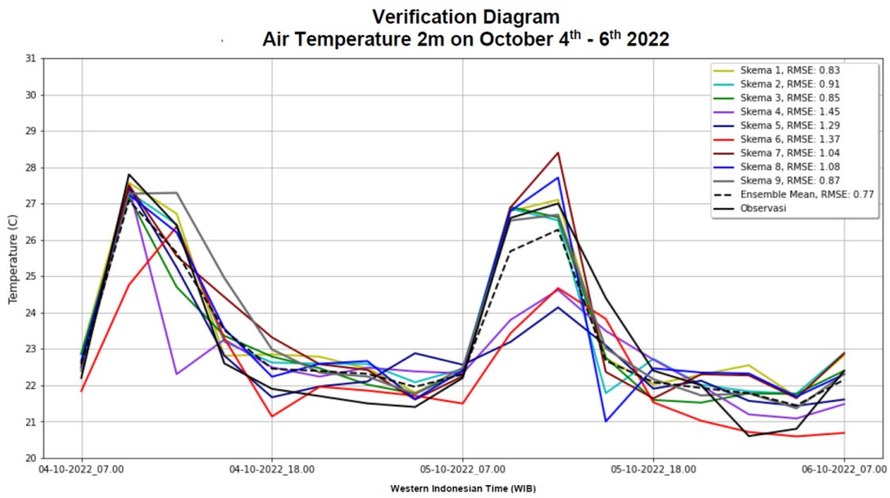
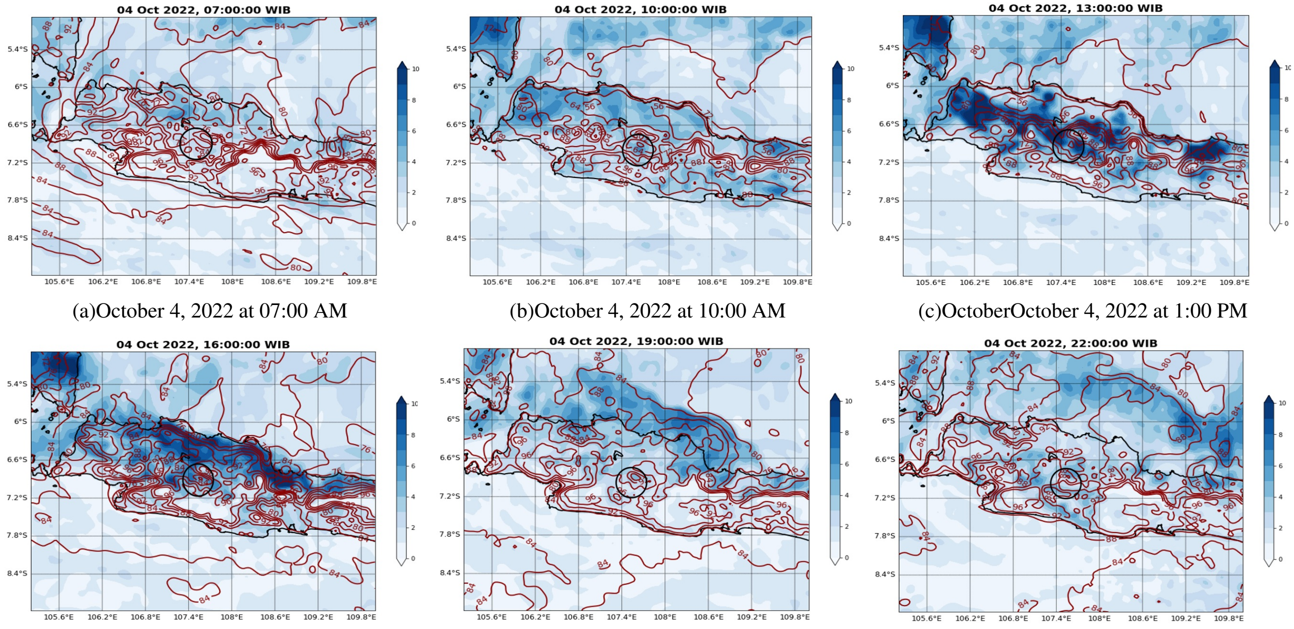


FIG. 5: Comparison diagram of 2-meter surface air temperature from each WRF parameterization scheme (ensemble member) and ensemble mean against observed surface air temperature at Husein Sastranegara Station, Bandung (7:00 AM on October 4, 2022, to 7:00 AM on October 6, 2022)

D. Ensemble Mean and Spread (Surface Relative Humidity)

Figure 6 represents surface air humidity using the ensemble mean and spread prediction method from October 4, 2022, at 7:00 AM to October 4, 2022, at 10:00 PM. During the rainfall from 1:00 PM to 4:00 PM, the standard deviation of surface air humidity in the Pagarsih area is the highest compared to other times. This is related to the varying rainfall intensities of each member (parameterization schemes) during the rain, which can affect the variation in surface air humidity (2 meters). The figure shows that during heavy rainfall in the Pagarsih area, the surface air humidity ranges from 88% to 96%, with a standard deviation reaching 9, indicating relatively high uncertainty. Finer grids produce forecasts with smaller bias and more realistic rainfall intensity distributions, but have higher standard deviations of errors [18].

Figure 7 shows the verification and comparison of RMSE values for each member and the ensemble mean for the surface air humidity variable. Based on the figure, the prediction of surface air humidity with the smallest RMSE value is mem-



(a)October 4, 2022 at 07:00 AM (b)October 4, 2022 at 10:00 AM (c)OctoberOctober 4, 2022 at 1:00 PM

(d)October 4, 2022 at 4:00 PM (e)October 4, 2022 at 7:00 PM (f)October 4, 2022 at 10:00 PM

FIG. 6: Ensemble mean (contour) and ensemble spread (shaded) of surface air humidity (2 meters) on October 4, 2022, from 07:00 AM to 10:00 PM.

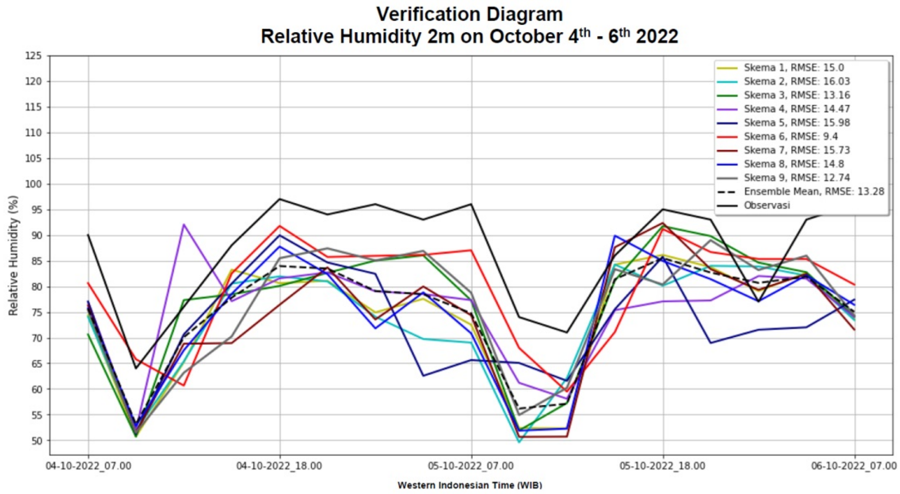


FIG. 7: Comparison diagram of 2-meter surface air humidity from each WRF parameterization scheme (ensemble member) and ensemble mean against observed surface air humidity at Husein Sastranegara Station, Bandung (7:00 AM on October 4, 2022, to 7:00 AM on October 6, 2022).

ber 6 (parameterization scheme 6), while the ensemble mean RMSE value is 13.28. In this case, member 6 performs better in predicting surface air humidity compared to other members and the ensemble mean.

IV. CONCLUSION

Based on the analysis of the application of the ensemble mean method, ensemble spread, and probability maps for predicting rainfall, surface air temperature, and surface air humidity in the case study of heavy rain on October 4, 2022, in Pagarsih, Bandung, West Java, the following conclusions were obtained. The WRF model (parameterization scheme 8 or member 8) is the best at predicting very heavy rainfall compared to other parameterization schemes. Parameterization scheme 1 is the best for predicting surface temperature, and parameterization scheme 6 is the best for predicting surface air humidity. The ensemble mean product has the lowest RMSE value for predicting surface air temperature.

The rainfall probability maps product can effectively identify the peak of heavy rainfall (October 4, 2022, from 1:00 PM to 4:00 PM), which is one of the causes of flooding in the Pagarsih area. This is indicated by the extensive area with more than a 90% chance of experiencing rainfall greater than 10 mm/3 hours. In contrast, the ensemble mean rainfall prediction product tends to underestimate heavy rainfall in the Pagarsih area.

High values of ensemble spread (for surface air temperature and surface air humidity), indicating high uncertainty, correlate with the potential for rain event in the area. This is evidenced by the coincidence of high ensemble spread values (surface air temperature and surface air humidity) with the potential for rain in the ensemble mean and probability maps (rainfall) products at the same time and location. This can be attributed to the interrelation between these weather elements and the variation in prediction values from each ensemble member.

Acknowledgments

This study was supported by the LPDP Scholarship and Faculty of Earth Sciences and Technology, Bandung Institute of Technology. I would like to thank the Earth Sciences Program (Faculty of Earth Sciences and Technology), for providing server facilities to run the WRF model in this study.

[1] D.R. Easterling, *et al.*, "Observed variability and trends in extreme climate events," A brief review. Bull. Am. Meteorol. Soc., vol. 81, p. 417-425, 2000.

[2] L.V. Alexander, *et al.*, "Global observed changes in daily climate extremes of temperature and precipitation," J. Geophys., vol. 111, D05109, p. 1-22, 2006. https://doi.org/10.1029/2005JD006290

[3] D. Coumou, S. Rahmstorf, "A decade of weather extremes," Nat. Clim. Chang., vol. 2, p. 491-496, 2012.

[4] W. Ingram, "Extreme precipitation: Increases all round," Nat. Clim. Chang., vol. 6, p. 443-444, 2016.

[5] I. Yucel, *et al.*, "Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall," Journal of Hydrology, vol. 523, p. 49-66, 2015.

[6] B. Bulut, M. T. Yilmaz, *et al.*, "Evaluation of Remotely-Sensed and Model-Based Soil Moisture Products According to Different Soil Type, Vegetation Cover and Climate Regime Using Station-Based Observations over Turkey," Remote Sensing, vol. 11, no. 16, p. 1875, 2019

[7] M. Amjad, *et al.*, "Performance evaluation of satellite- and model-based precipitation products over varying climate and complex topography," Journal of Hydrology, vol. 584, 124707, 2020.

[8] A.F. Prein, *et al.*, "A review on regional convectionpermitting climate modeling: Demonstrations, prospects, and challenges," Reviews of geophysics, vol. 53, no. 2, p. 323-361, 2015.

[9] I. Yucel, and A. Onen, "Evaluating a mesoscale atmosphere model and a satellite-based algorithm in estimating extreme rainfall events in northwestern Turkey," Natural Hazards and Earth System Sciences, vol. 14, no. 3, p. 611-624, 2014.

[10] F. Xie, H. He, and Y. Zhang, "Evaluation of simulation results for different parameterization schemes of land process," J. Anhui Agric. Sci., vol. 44, p. 210-215, 2016.

[11] Q. Yang, *et al.*, "Performance of the WRF model in simulating intense precipitation events over the Hanjiang River Basin, ChinaA multi-physics ensemble approach," Atmos. Res., vol. 248, 105206, 2021.

[12] W.A. Gallus, "Application of object-based verification techniques to ensemble precipitation forecasts," Weather Forecast., vol. 25, p. 144-158, 2010.

[13] J. Wei, *et al.*, "Role of reservoir regulation and groundwater feedback in a simulated ground-soil-vegetation continuum: A long-term regional scale analysis," Hydrol. Process., vol. 35, no. 8, 2021.

[14] S. Joslyn, and S. Savelli, "Communicating Forecast Uncertainty: Public Perception of Weather Forecast Uncertainty," Royal Meteorology Society, vol. 17, p. 180-195, 2010.

[15] S. Joslyn, and J.E. Le Clerc, "Uncertainty Forecast Improve Weather Related Decisions and Attenuate the Effects of Forecast Error," Journal of Experimental Psychology, vol. 18, no. 1, p. 126-140, 2012.

[16] E. Kessler, "On the Distribution and Continuity of Water Substance in Atmospheric Circulations. In: On the Distribution and Continuity of Water Substance in Atmospheric Circulations. Meteorological Monographs," American Meteorological Society, Boston, MA., 10, https://doi.org/10.1007/978-1-935704-36-2-1, 1969.

[17] G.A. Grell, and D. Dvnyi, "A generalized approach to parameterizing convection combining ensemble and data assimilation techniques," Geophys. Res. Lett., vol. 29, no. 14, doi:10.1029/2002GL015311, 2002.

[18] J. Jeworrek, G. West, and R. Stull, "WRF Precipitation Performance and Predictability for Systematically Varied Parameterizations over Complex Terrain," Wea. Forecasting, vol. 36, p. 893-913, https://doi.org/10.1175/WAF-D-20-0195.1, 2021.