Evaluating skill of BMKG wave model forecast (Wavewatch-3) with observation data in Indian Ocean (5 – 31 December 2017).

R Kurniawan¹, H Harsa¹, A Ramdhani¹, W Fitria¹, D Rahmawati¹, M N Habibie¹, and T D Hutapea¹.

¹Meteorological, Climatological, and Geophysical Agency, BMKG, Jakarta, 10610, Indonesia

e-mail: ronie_354@yahoo.co.id, hastuadi@gmail.com, andri.dhani@gmail.com, wellyfitria234@gmail.com, rahmawatidhita119@gmail.com, najib.habibie@gmail.com, hutapea.df@gmail.com

Abstract. Providing Maritime meteorological forecasts (including ocean wave information) is one of BMKG duties. Currently, BMKG employs Wavewatch-3 (WW3) model to forecast ocean waves in Indonesia. Evaluating the wave forecasts is very important to improve the forecasts skill. This paper presents the evaluation of 7-days ahead BMKG’s wave forecast. The evaluation was performed by comparing wave data observation and BMKG wave forecast. The observation data were obtained from RV Mirai 1708 cruise on December 5th to 31st 2017 at the Indian Ocean around 04°14’S and 101°31’E. Some statistical properties and Relative Operating Characteristics (ROC) curve were utilized to assess the model performance. The evaluation processes were carried out on model’s parameters: Significant Wave Height (Hs) and Wind surface for each 7-days forecast started from 00 UTC. The comparison results show that, in average, WW3 forecasts are over-estimate the wave height than that of the observation. The forecast skills determined from the correlation and ROC curves are good for the first- and second-day forecast, while the third until seventh day decrease to fair. This phenomenon is suspected to be caused by the wind data characteristics provided by the Global Forecasts System (GFS) as the input of the model. Nevertheless, although statistical correlation is good for up to 2 days forecast, the average value of Root Mean Square Error (RMSE), absolute bias, and relative error are high. In general, this verifies the overestimate results of the model output and should be taken into consideration to improve BMKG’s wave model performance and forecast accuracy.

1. Introduction

Marine meteorological information services are important and must be available continuously for public information, especially for ocean wave information [1][2][3][4][5][6][7][8][9][10][11]. Ocean wave forecast is very vital in supporting many activities in maritime sector for example, safety of voyage, coastal development, etc[12][13][14][15][16][17][18][19][20][21]. The Indonesia Meteorology Climatology and Geophysics Agency (BMKG) is an institution which is responsible for providing maritime forecast information to the public; including ocean wave forecast over Indonesian waters for 7 days ahead.

BMKG utilizes WW3 model to provide the ocean wave forecast because the model is able to provide more complete and continuous data. The data serve as complement to observation and measurement records. The other problem is the observation and measurement data in the ocean are very limited and,
rarely continuous, based on WMO-No.702, the ocean wave information can be provided from three types of wave data that are, observed, measured and hindcast [22]. BMKG used WW3 model since 2014, the model parametrization, configuration and validation based on Ramdhani 2015 [23], this model is run using 10-meter level height from GFS (Global Forecasting System), NCEP, NOAA as an input model.

Some wave forecast users informed that the results of the wave forecast were higher than reality, this condition will be an important concern to conduct a validation study on the BMKG’s wave forecast results, so the aim of this research is to evaluate BMKG’s wave model (WW3) performance for 7 days forecast, and to analyzes the forecast skill for recommendation and improvement the WW3 model.

2. Data and Methods

2.1. Data

The significant wave height (Hs) and wind surface data observation time series obtained from RV Mirai 1708 cruise, located in the Indian ocean around 04°14’S, 101°31’E (figure 1), and the measurement period is carried out from 5 to 31 December, 2017. Significant wave height and wind surface dataset was produced from surface meteorological measurement systems (SMET/MIRAI Surface Meteorological observation system).

Wavewatch3 model output obtained from BMKG Ocean Forecast System [24], the data point position collected at 04°14’S, 101°31’E (following the position of observation data), the parameters assessed covers are significant wave height (Hs) and Wind surface for 7 days forecast, start from 00 UTC in each day, the spatial resolution is 0.0625 deg (7km), and the spatial temporal is 3 hours, the time series data period from 5 to 31, December 2017, and provides 7 days forecast information in each date.

![Figure 1. Location of the observation data measurement](image)

2.2. Methods

The correlation coefficient (r) is a value that indicates the strength of the linear relationship between variables. The coefficient values from -1 to 1, and the interpretations of the values are -1 (Perfect negative correlation), 0 (No correlation), 1 (Perfect positive correlation), The closer that the absolute value of r is to one, the better that the data are described by a linear equation [25].

\[ r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2\sum(y_i - \bar{y})^2}} \]  \hspace{1cm} (1)

where:
- \( r \) = the correlation coefficient relationship between the variables x and y
- \( x \) – the values of the x-variable
- \( y \) – the values of the y-variable
To identify skill performance from the Wavwacth3 model forecast against observation data used the ROC (Relative Operating Characteristic) curve, ROC curves is highly flexible method for representing the quality of the forecasts to discriminate observations in each category [26], [27].

ROC is a means of testing the skill of categorical forecast, the ROC curve is the visualization methods used for classifying quality. ROC is a signal detection curve for binary forecasts obtained by plotting a graph of the Hit rate (y-axis) versus the False alarm rate (x-axis) over a range of different thresholds (figure 2). It shows the dependency between the HR (Hit Rate) and the FAR (False Alarm Rate). Hit Rate (HR) = hits / (hits + misses), False Alarm Rate (FAR) = false alarm / (correct negative + false alarm) (table 1). The ROC curve represents a skill forecast of a system wherein the hit rate and the false alarm rate are compared. A rough guide for classifying the accuracy of a diagnostic test is given by the traditional academic point system; 0.90-1 = excellent, 0.80-.90 = good, 0.70-.80 = fair, 0.60-.70 = poor, 0.50-.60 = fail

![Figure 2. ROC Space](image)

**Table 1. Contingency table**

| (F) Forecast | Yes | No | Total |
|--------------|-----|----|-------|
| Yes          | hits|    |       |
| No           | Misses| Correct Negatives |       |
| Total        | Observed yes | Observed No | Total |

Threshold’s categories were based on Douglas scale. The categorical method used for analyzing which wave categories were successfully predicted. Douglas Sea Scale is a scale which measures the height of the sea waves. The scale is very simple to follow and is expressed in one of 10 levels [28].
Table 2. Douglas Sea Scale

| Scale | Height   | Desc       |
|-------|----------|------------|
| 0     | 0        | No wave    |
| 1     | 0 - 0.1  | Calm       |
| 2     | 0.1 - 0.5| Smooth     |
| 3     | 0.5 - 1.25| Slight     |
| 4     | 1.25 - 2.5| Moderate   |
| 5     | 2.5 – 4  | Rough      |
| 6     | 4 – 6    | Very Rough |
| 7     | 6 – 9    | High       |
| 8     | 9 – 14   | Very High  |
| 9     | > 14     | Phenomenal |

3. Results and discussion

Previous research, evaluation of WW3 model was carried out by Dedi (Presented in Blue Earth Symposium, March, 8-9 2016, Tokyo), the study conducted an evaluation of the WW3 model by comparing with wave data observation from R/V Mirai cruise (MR1504), and the correlation obtained is 0.789. The other WW3 model evaluation conducted by Ramdhani (2015) [23], on his study utilized altimetry data to evaluate WW3 model, and the correlation obtained is 0.89. Subsequently, on this study also evaluate WW3 model with observation data from R/V Mirai cruise (MR1708), the correlation obtained is 0.72 (figure 3). The points of this study shows that the WW3 model is accurate enough having correlation more than 0.7. However, it should be considered that the results of this evaluation use at the same time series data between observation and WW3 model (not a forecast data).

Figure 3. SWH Time series of WW3 (Analysis) and R/V Mirai 1708 observation data (left), Correlation value (right), 5-31 December, 2017

On this study compared the WW3 model output for seventh days forecast with the observation data (figure 4), the pattern of Significant Wave Height (Hs) chart line shows that the WW3 model provide provided a good forecast until 2 days forecast, even for the extreme conditions (the wave height more than 2 meters), and the forecast accuracy begun to deteriorate from 3 days up to 7 days forecast.
Figure 4. SWH Time series of WW3 and Observation data (R/V Mirai 1708)

The scatter plot indicates the relationship of Significant Wave Height (Hs) between WW3 and observation data (figure 5). The figure of 8 scatters plots represent the correlation between WW3 analysis and WW3 forecast for day-1 until day-7 respectively are given above. It is shown that the first day WW3 forecast has the best correlation with a value is 0.79, followed by WW3 analysis with a correlation value of 0.72, with a better absolute bias average value of 0.20 m and RMSE value 0.24 compared to day-1 forecast value. In contrast, the absolute bias average and RMSE value showed relatively high value namely 0.48 for abs bias average and 0.52 for RMSE value, on day-1 WW3 forecast. In general, this verifies that even though the WW3 model can capture the observation pattern but it’s still overestimated the observation value. The correlation is decreasing from day-2 forecast until 7-day forecast with correlation value 0.63 in day-2 to 0.59-0.57 in day-3 to day-7 forecast, with absolute bias average range between 0.42 m – 0.49 m and RMSE range between 0.51 – 0.47, this is should be taken into consideration for wave model forecast improvement.
Figure 5. SWH Correlation of WW3 and Observation data (R/V Mirai 1708)

The graph of WW3 temporal bias and observational data (figure 6) shows that the WW3 analysis has the smallest bias value compared to the WW3 model of day-1 until day-7. This chart line shows that the WW3 forecast from day-1 to day-7 came out with an overestimated value compared to observations. The bias from the WW3 forecast day 1 shows relatively better results than the WW3 forecast day-2 to day-7. The pattern of this temporal bias also shows that the model's behavior has decreased in performance further away from the initial time.

Figure 6. Temporal Bias of WW3 and Observation data (R/V Mirai 1708)

The ROC curve represents a skill forecast of a system wherein the hit rate and the false alarm rate are compared. We can analyze the ROC curve with 2 methods, first we can see the curve line. When the
curve line getting away from the base line, the skill forecast is good. The second method is by analyze the area under curve (AUC), wider AUC shows the better skill of forecast, but both of them are have similar value. The result of ROC between WW3 versus observation data (figure 7), shows that at the first day forecast has the best skill forecast whose AUC value is 0.87. The AUC in second day is decrease, but it is not significant, the AUC value is 0.86. Both of skill forecast in first and second day still in good level. The skill forecast is decreasing significantly starting in day-3 until day-7 forcast, presented by AUC’s value decreased from 0.86 in second day to 0.78-0.77 in day-3 until day-7 day forecast. The ROC curve determine that operated wave model in BMKG have a good skill forecast in two days, after that, it is decreasing significantly representing the skill forecast getting worst.

![ROC Curve WW3 vs Observation data (R/V Mirai 1708)](image-url)

**Figure 7.** ROC Curve WW3 vs Observation data (R/V Mirai 1708)
The graph on figure 8 gives the information about the time series of UV-Wind of WW3 model (analysis) and observation data for 5 until 31 December, 2017. The chart line shows that, the trend of UV-Wind time series data gives similar pattern between WW3 analysis and observation data from 5 until 31 December, 2017. Overall, the WW3 analysis can provide a good response to the observation data.

![UV-Wind of WW3 and Observation data](image)

**Figure 8.** Wind time series of Wavewatch3 and Observation data (R/V Mirai 1708)

The line graph shows the UV wind correlation of WW3 model with observation data (figure 9). The start of the comparison WW3 analysis with observation data obtained a good correlation, it is 0.77, and V-Wind is 0.82, and following the correlation has decreased from begun the first day forecast until the day-7 forecast.

![Wind Correlation of WW3 against Observation](image)

**Figure 9.** Wind Corelation of Wavewatch3 and Observation data (R/V Mirai 1708)
4. Conclusion

The comparison results between WW3 and observation data show that, in average, BMKG’s wave model forecasts (WW3) are over-estimate in significant wave height. The comparison of both on time series data shows that the forecast by the WW3 model until two days provided good skill forecast even for the extreme conditions, after that the forecast begun to deteriorate, the correlation value shows that after two days forecast decreased significantly. For the forecast skills determined from the ROC curves are good at the first- and second-day, while the third until seventh day decreased to fair.

BMKG’s wave forecasts provides good results for the first- and second-day forecast, this phenomenon is suspected to be caused by the wind surface data characteristics provided by the Global Forecasts System (GFS) as the input of the model. Nevertheless, although statistical correlation is good for up to 2 days forecast, the average value of Root Mean Square Error (RMSE), absolute bias, and relative error are high, therefore, these results could be taken into consideration to improve BMKG’s wave model performance and forecast accuracy.

For suggestion, prior to performing WW3 model enhancement, it is advisable to carry out preliminary research to some materials as follows; (1) the forecast accuracy can be maintained or improved by updating the forecast every day or more frequent using the improved data input, (2) verifies and analysis of Global Forecast Data (GFS) wind output data as the input utilized by WW3, (3) attempting to use the other wind data products as the input of WW3, e.g., wind data of WRF output.

References

[1] Guide to marine meteorological sciences., no. 471. 1982.
[2] D. J. Doong, S. T. Chen, Y. C. Chen, and C. H. Tsai, “Operational probabilistic forecasting of coastal freak waves by using an artificial neural network,” J. Mar. Sci. Eng., vol. 8, no. 3, 2020, doi: 10.3390/jmse8030165.
[3] T. Gruchmann, N. Pratt, J. Eiten, and A. Melkonyan, “4PL Digital Business Models in Sea Freight Logistics: The Case of FreightHub,” Logistics, vol. 4, no. 2, p. 10, 2020, doi: 10.3390/logistics4020010.
[4] S. Irawan, R. Fahmi, M. Z. Lubis, S. B. Aji, A. Roziqin, and H. Khoirunnisa, “Hydro-oceanographic condition (Tide, Sea Current, and Waves) of Nongs Batam Sea,” J. Appl. Geospatial Inf., vol. 2, no. 2, pp. 135–144, 2018, doi: 10.30871/jagi.v2i2.968.
[5] S. Katsanevakis et al., “Twelve Recommendations for Advancing Marine Conservation in European and Contiguous Seas,” Front. Mar. Sci., vol. 7, no. October, pp. 1–18, 2020, doi: 10.3389/fmars.2020.565968.
[6] J. M. Maina et al., “Aligning marine spatial conservation priorities with functional connectivity across maritime jurisdictions,” Conserv. Sci. Pract., vol. 2, no. 2, pp. 1–13, 2020, doi: 10.1111/csp2.156.
[7] J. W. Milon and S. Alvarez, “Coastal resources economics and ecosystem valuation,” Water (Switzerland), vol. 11, no. 11, pp. 2–5, 2019, doi: 10.3390/w11112206.
[8] Q. Qian, P. Liang, L. Qi, Y. Ding, and J. He, “Sub-Seasonal Variability of Meridional Activity of Western Pacific Subtropical High in Boreal Late Summer,” Front. Earth Sci., vol. 8, no. November, pp. 1–16, 2020, doi: 10.3389/feart.2020.597969.
[9] Y. Teff-Seker et al., “Do Alternative Dispute Resolution (ADR) and Track Two Processes Support Transboundary Marine Conservation? Lessons From Six Case Studies of Maritime Disputes,” Front. Mar. Sci., vol. 7, no. November, pp. 1–14, 2020, doi: 10.3389/fmars.2020.593265.
[10] E. Simulation, H. Wang, C. Dong, Y. Yang, and X. Gao, “Parameterization of Wave-Induced Mixing Using,” no. I.
[11] K. Yan et al., “Managing information uncertainty in wave height modeling for the offshore structural analysis through random set,” Complexity, vol. 2017, 2017, doi: 10.1155/2017/9546272.
[12] R. Kurniawan and M. K. Khotimah, “Ocean Wave Characteristics in Indonesian Waters for Sea Transportation Safety and Planning,” IPTEK J. Technol. Sci., vol. 26, no. 1, 2016, doi: 10.12962/j20882033.v26i1.767.

[13] Z. Yang, W. Shao, Y. Ding, J. Shi, and Q. Ji, “Wave simulation by the swan model and fvcom considering the sea-water level around the zhoushan islands,” J. Mar. Sci. Eng., vol. 8, no. 10, pp. 1–18, 2020, doi: 10.3390/jmse8100783.

[14] H. S. Eum, W. M. Jeong, Y. S. Chang, S. H. Oh, and J. J. Park, “Wave energy in korean seas from 12-year wave hindcasting,” J. Mar. Sci. Eng., vol. 8, no. 3, 2020, doi: 10.3390/jmse8030161.

[15] A. de S. de Lima et al., “Hydrodynamic and waves response during storm surges on the southern brazilian coast: A hindcast study,” Water (Switzerland), vol. 12, no. 12, 2020, doi: 10.3390/w12123538.

[16] A. Wiese, J. Staneva, J. Schultz-Stellenfleth, A. Behrens, L. Fenoglio-Marc, and J.-R. Bidlot, “Synergy between satellite observations and model simulations during extreme events,” Ocean Sci., pp. 1–25, 2018, doi: 10.5194/os-2018-87.

[17] J. Schulz-Stellenfleth and J. Staneva, “A multi-collocation method for coastal zone observations with applications to Sentinel-3A altimeter wave height data,” Ocean Sci., vol. 15, no. 2, pp. 249–268, 2019, doi: 10.5194/os-15-249-2019.

[18] A. Golshani, A. Vaselali, and S. Mazaheri, “Marine Forecasting Systems in Caspian Sea, Persian Gulf and Oman Sea,” Int. J. Coast. Offshore Eng., vol. 3, no. 2, pp. 9–12, 2016.

[19] M. Roh, H. Kim, P. Chang, S. O.-J. of M. S. and, and undefined 2021, “Numerical Simulation of Wind Wave using Ensemble Forecast Wave Model: A Case Study of Typhoon Lingling,” mdpi.com. Accessed: Jul. 01, 2021. [Online]. Available: https://www.mdpi.com/2077-1312/9/5/475.

[20] W. Shao, F. Nunziata, Y. Zhang, V. Corcione, and M. Migliaccio, “Wind speed retrieval from the Gaofen-3 synthetic aperture radar for VV- and HH-polarization using a re-tuned algorithm,” Eur. J. Remote Sens., vol. 54, no. 1, pp. 318–337, 2021, doi: 10.1080/22797254.2021.1924082.

[21] W. M. Organization, Guide to Wave Analysis and Forecasting, vol. 1998, no. 702. 1998.

[22] W. M. Organization, Guide to Wave Analysis and Forecasting, vol. 1998, no. 702. 1998.

[23] A. Ramdani, “Effect of Tropical Cyclones and Madden-Julian Oscillation (MJO) on High Waves Occurrences in The Inner Indonesian Seas,” ITB, 2015.

[24] “BMKG OFS - Interactive Maps.” https://peta-maritim.bmkg.go.id/of/ (accessed Jul. 01, 2021).

[25] “Basic Concepts of Correlation | Real Statistics Using Excel.” https://www.real-statistics.com/correlation/basic-concepts-correlation/ (accessed Jul. 01, 2021).

[26] S. J. Mason and N. E. Graham, “Conditional probabilities, relative operating characteristics, and relative operating levels,” Weather Forecast., vol. 14, no. 5, pp. 713–725, 1999, doi: 10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2.

[27] K. Kadarsah, “Aplikasi Roc Untuk Uji Kehandalan Model Hybmg,” J. Meteorol. dan Geofis., vol. 11, no. 1, pp. 33–43, 2010, doi: 10.31172/jmg.v11i1.60.

[28] “(5) (PDF) Encyclopedia of Coastal Science | Historically Digitized - Academia.edu.” https://www.academia.edu/1157484/Encyclopedia_of_Coastal_Science (accessed Jul. 01, 2021).