Sensitivity analysis for wind-driven significant wave height model in SWAN: A Sunda Strait case

A Wurjanto1* and J A Mukhti2

1-2 Ocean Engineering Program, Faculty of Civil Engineering and Environment, Institut Teknologi Bandung, Indonesia

E-mail: andojowurjanto@gmail.com

Abstract. Sunda Strait is one of the most important straits in Indonesia since it connects two major islands of Indonesia as well as two oceans. Due to its high importance, it is essential to have a well understanding of significant wave height within the strait. However, there are currently no long-term significant wave data from buoy records in Indonesia. The available fine-resolution meteorology-oceanography (metocean) data around the Sunda Strait is also not as fine as other locations in the world. Therefore, hydrodynamic modelling is essential to obtain significant wave data that completely covers the strait. The aim of this study was to perform a sensitivity analysis based on Sunda Strait SWAN significant wave height model. In the research, we performed a sensitivity analysis on a significant wave height SWAN model using various spatial resolution of wind forcing from the ECMWF ERA5 database. The analysis is essential to obtain the coarsest acceptable metocean data for modelling. Our result shows that higher spatial resolution of wind forcing gives better agreement with the ERA 5 wave data. For the Sunda Strait case, we recommend using no coarser than 0.5 degrees of spatial wind forcing resolution. However, we also found that the modelling result, along with models from past studies, still underestimates the reported wave height data in extreme conditions.

Keywords: sensitivity analysis, wind-driven wave height, SWAN model, Sunda Strait

1. Introduction
Sunda Strait is one of the busiest straits in Indonesia. The strait is the pathway between the Pacific and the Indian Ocean that also connects the Java and Sumatera Island. In terms of meteorology-oceanography (later referred to as ‘metocean’), the location of the Sunda Strait can be challenging especially for maritime operations. Extreme weather occurs almost every year in Sunda Strait and affects the ferry operations in the area [1] [2] [3]. It was also mentioned in a report that the waves can reach up to 5 meters in the strait [4].

This frequent occurrence of extreme weather leads to the importance of metocean study in the Sunda Strait. Significant wave height $H_s$ and wave period $T_p$ are one of the most important ocean parameters. $H_s$ is often used for engineering purposes such as the design of coastal structures [5] and offshore platforms [6] [7]. A higher significant wave period $T_p$ implies the higher wavelength that may occur in a region [8]. Wavelength information is critical to the safety of navigation. According to a past study [9], most ship accidents occur when the wave length is more than half of the ship’s length.
The most common method to acquire wind-wave information in an area is by using hydrodynamic models. Hydrodynamic models such as SWAN are based on finite difference method which is presented as grids in a certain modelling domain [10]. Source conditions or ‘forcings’ are applied in the edge and within the domain. For wind-wave modelling, the forcings for SWAN can be wave spectra or wind time series.

However, due to time constrain, lack of computational power, or limited amount of input data, it is common to optimize hydrodynamic models. Aspects which are often be optimized including grid size and time step resolution. The common practice in hydrodynamic modelling is to use a coarser grid on the larger part of the ocean and finer grid around the area of interest. This method is known as nesting and has been done in previous studies [11] [12]. Another method to optimize the model runtime is by increasing the time step resolution until the desired balance between accuracy and runtime is obtained [13].

The less preferred optimization method is done by reducing the number of forcing points. While it is a common norm in wind-wave modelling to include as many input data as possible, data with fine spatial resolution such as 6 km is rare. For example, the SEAFINE data by Oceanweather, Inc has that spatial resolution only in Offshore Madura, North of Kalimantan Island, and the eastern coast of South East Asia Countries and China [14].

The aim of this study is to perform a sensitivity analysis based on Sunda Strait SWAN significant wave height model. Three levels of wind forcing spatial resolution from publicly available wind data are used as the variable for the analysis. The modelling result is then compared to the ECMWF ERA 5 wave data as the baseline, which wave data covers globally.

2. Methodology

2.1. SWAN

SWAN is a hydrodynamic model based on the finite-difference method. It is capable to simulate wave transformation in the nearshore area. The main inputs for SWAN including wind, energy dissipation, wave-wave interaction. The principle of SWAN modelling calculation is by integrating the wave action equilibrium by using the finite difference method. Action balance is the core equation of the SWAN wind-wave model. The action density $A$ is described as follows [15]:

$$\frac{\partial A}{\partial t} + \frac{\partial}{\partial x} c_x A + \frac{\partial}{\partial y} c_y A + \frac{\partial}{\partial r} c_r A + \frac{\partial}{\partial \theta} c_\theta A = \frac{S_{tot}}{\sigma}$$ (1)

Where $S_{tot}$ or total $S(\sigma, \theta)$ is the source of energy density spectral, $A(x, y, t, r, \theta)$ is the wave action density spectrum, $t$ is time, $c_x$, $c_y$, $c_r$, $c_\theta$ represents wave propagation velocity in their respective space, $r$ is relative frequency, and $\theta$ is wave direction. Each term on the left side of the equation represents the following factors from left to right: (1) the density change over time, (2) the change of action density propagation in $x$ and $y$ space, (4) the relative frequency change due to water depth and current variation, and (5) wave refraction change due to changing water depth and current.

Those factors are balanced by the right side of Eq. 1. The $S_{tot}$ is an accumulation of the following factors: (1) wind-induced wave generation (2) dissipation caused by whitecapping (3) nonlinear quadruplet and triad interaction of waves (4) sea bottom friction, and (5) wave-breaking in shallow water.

The transfer of energy $E$ from wind to waves $S_{in}$ is formulated as follows [15]:

$$S_{in}(\sigma, \theta) = J + KE(\sigma, \theta)$$ (2)

Where $J$ and $K$ depends on the wave frequency and direction as well as wind speed and direction.
2.2. General analysis process
The research is conducted in the following process. The first step is to acquire the location, wind, and bathymetry data. Secondly, we set up the modelling domain, grid, and bathymetry on the SWAN. Thirdly, we analyse in three scenarios with each using its own set of wind forcing data points.

Fourthly, after the analysis for each simulation is completed, we extract the data from two-point of interests. These points are chosen as a representation of significant wave height before it is transformed as it approaches shallower water. The points are in the northeast and southwest part of the Sunda Strait respectively. Both locations are located on points deep water with approximately the same depth to avoid the wave shoaling effect.

The fifth step is to compare the model result from each scenario with ECMWF wave data as the baseline. The model accuracy to the wave data is measured using root-mean-square error (RMSE) and standard deviation. Finally, the result from each scenario is compared with each other, the ECMWF wave data, and past studies.

2.3. Modeling setup and input
The bathymetry used in this model is from GEBCO (General Bathymetric Chart of the Oceans). GEBCO has been used in metocean studies such as by [16] and [17]. The GEBCO bathymetry has 30 arc-sec of spatial resolution or approximately 1 km. Modelling is performed in a uniform 0.01 x 0.01 degree grid within 6.25 S 105.25 E, 6.25 S 106.50 E, 5.25 S 106.50 E, and 5.25 S 105.25 E.

The modelling boundary along with the bathymetry and wind forcing location for each scenario is shown in Figure 1. The first scenario uses all available wind forcing from ECMWF within the modelling domain with 0.25 degrees spatial resolution. The second scenario use nine points of wind forcing with each separated 0.50 degrees from each other. Finally, the third scenario use only four points of wind forcing with 1.00 degree of spatial resolution.

Figure 1. Modelling Domain, Grid, Bathymetry, Wind Forcing Locations, and Point of Interests (Loc).
2.4. Selecting wind data
In its current state, the SEAFINE data only covers the wind data on the northern part of the strait with coarser 0.5 x 0.5 degrees of the spatial grid, while the ideal modelling requires wind input from the southern part of the strait as well. The wind data input from the within the strait and southern part are important, especially since most larger waves are coming from the Indian Ocean according to past study [17]. The only freely accessible alternative to the SEAFINE is ERA 5 by ECMWF. ERA5 is a reanalysis database with data ranging from 1750 to two or five days before the present day [18].

The wind data used from the ERA 5 is the wind speed at 10 m height in transversal and longitudinal direction relative to the equator. The ERA 5 wind data has 0.25° x 0.25° degrees of resolution for the wind data. The analysis was performed using ERA5 2019 data for one year. The example of wind data time series used in the model is shown in Figure 2. The computational time step used is the same as the wind data interval which is 1 hour. The bottom friction coefficient used in the model is JONSWAP Type which is 0.038 m²/s³. Water density is set to 1,025 kg/m³.

![Figure 2. Example of ECMWF Wind Data Used in the Sunda Strait SWAN Wave Model](image)

2.5. Error calculation
Actual wave data from buoy recording is virtually non-existent in Indonesian waters. The BPPT (Indonesian Agency for the Assessment and Application of Technology) reported that all Indonesian
buoys for tsunami detection have been out of service since 2012 [19]. Therefore, the ECMWF wave data is used as the baseline to be compared with the simulation result.

The comparative study of significant wave height between an established data such as the ECMWF data or wave buoys to model simulation are often compared only visually [20] or by certain statistical parameters [21][22].

The following are the RMSE, standard deviation $D$, scatter index $SI$, and correlation coefficient $r$ formula to measure how the model compares to ECMWF wave data:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (e_i - m_i)^2} \quad (3)$$

$$D = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (m_i - \bar{m})^2} \quad (4)$$

$$SI = \frac{RMSE}{\bar{e} \bar{m}} \quad (5)$$

$$r = \frac{\sum_{i=1}^{N} (e_i - \bar{e})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^{N} (e_i - \bar{e})^2} \sqrt{\sum_{i=1}^{N} (m_i - \bar{m})^2}} \quad (6)$$

Where $N$ is the number of wind data, $e_i$ is the significant wave height number-$i$ from ERA 5, $m_i$ is the significant wave height number-$i$ from the SWAN model simulation, $\bar{e}$ is the average significant wave height from the ECMWF data, and $\bar{m}$ is the average significant wave height from the SWAN simulation.

The mean percentage of error, often expressed as:

$$Error = \frac{1}{N} \sum_{i=1}^{N} (e_i - m_i)^2 \quad (7)$$

is usually not used for error calculation of significant wave height modelling since either $e$ and $m$ can be very close to zero, which can contribute a high amount of error to the mean of error.

### 3. Results

The modelling result of the 6 x 5 wind forcing grid shows a good agreement with the ECMWF $H_s$ data as shown in Figure 3 and Figure 4. The quantitative comparison between the ECMWF ERA 5 and SWAN model results is shown in Table 1. Overall accuracy is higher in Location 1 where no data showing significant wave height difference of more than 0.5 m.

The scenario with 6 x 5 wind forcing input has the least amount of RMSE in both locations. Both Location 1 and 2 show similar lowest RMSE at 0.12 and 0.13 m respectively. The reduction to 3 x 3 wind forcing slightly increases the RMSE to 0.18 and 0.14 m respectively. However, further reduction to 2 x 2 drastically increased the RMSE. With the RMSE of up to 0.58 m, the 2 x 2 is unusable for modelling.

A similar pattern is also found on the deviation and scatter index. The reduction of wind forcing resolution increases both deviation and scatter index which means the data spread gets larger. However, there is a slight anomaly on the 3 x 3 wind forcing scenario on Location 2 where it is better than the 6 x 5 scenario. It should also be noted that the lowest scatter index is still high at 0.6 where the maximum 0.5 is preferable.

In terms of the correlation coefficient, the 6 x 5 forcing scenario topped at 0.93. There is a considerable amount of correlation reduction to 0.48 when the forcing is lowered to 3 x 3. The 1 x 1 scenario, however, is unusable with 0.01 correlation.

The modelling result itself also shows a relatively low significant wave height in both locations. Over one year of modelling, the highest significant wave height to be found is 1.06 m in Location 1. The highest significant wave height from ECMWF wave data, however, is only slightly higher at 1.33 m.
Figure 3. Significant Wave Height Based on ECMWF Data and Hydrodynamic Modelling on Location 1
**Figure 4.** Significant Wave Height Based on ECMWF Data and Hydrodynamic Modelling on Location 2
Table 1. Standard Deviation $\sigma$, RMSE, and % of Data with Large Error in Location 1 and 2

| Wind Forcing | Location 1 | Location 2 |
|--------------|------------|------------|
|              | RMSE       | $D$        | $SI$  | $r$   | RMSE       | $D$        | $SI$  | $r$   |
| 6 x 5        | 0.12 m     | 0.185      | 0.6   | 0.93  | 0.13 m     | 0.074      | 1.5   | 0.75  |
| 3 x 3        | 0.18 m     | 0.221      | 0.7   | 0.85  | 0.14 m     | 0.094      | 1.1   | 0.48  |
| 2 x 2        | 0.58 m     | 0.435      | 2.0   | 0.83  | 0.38 m     | 0.141      | 2.6   | 0.01  |

4. Discussion

Based on the sensitivity analysis with three different wind forcing resolutions, it is found that incorporating more data points improves the modelling accuracy. This is determined by the 6 x 5 wind forcing showing overall lowest error and better correlation compared to the baseline data. The insignificant increment of error by reducing the input points from 6 x 5 to 3 x 3 is likely to be caused by the model only lost four dominant forcings at 6 $^\circ$ 105.25 E, 6.25 $^\circ$ 105.50 E, 5.5 $^\circ$ 106.25 E, and 5.25 $^\circ$ 106.25 E. The rest of wind forcings are either located within the domain or inland. Points in 106.50 E, however, hypothetically can be replaced by points in 106.25 E.

Reduction of data input from 3 x 3 to 2 x 2 leaves only two points of inputs in the sea, which are 6.25 $^\circ$ 105.25 E and 6.25 $^\circ$ 106.25 E. The lack of other inputs from the sea caused a large deviation in the result with only two dominant inputs.

Past studies also show similar behavior. A study by Rusu et al. shows that wind data spatial resolution is very important in increasing the modelling accuracy [22]. The Tagus estuary significant wave model generated by SWAN shows increasing the wind data spatial resolution from 12.33 km to 4.10 reduces the RMSE from 0.602 to 0.505 and the correlation coefficient from 0.856 to 0.847 compared to wave buoy data. However, further resolution increment to 1.37 km does not give any improvement to both parameters.

On a semi-enclosed sea, however, the wind data accuracy is presumably more important than the spatial resolution as presented by León and Guedes Soares [23]. In their study, the HIPOCAS wind dataset with 0.25 degrees resolution and ECMWF ERA40 with 2.5 degrees are chosen to drive the significant wave height model in two different scenarios in the western Mediterranean Sea. The switch ECMWF ERA40 to HIPOCAS only resulted in the change of data bias from constantly negative to mostly positive. However, the positive bias does not translate to a more accurate model as the model driven by each dataset. On seven different points of interest, the RMS difference of significant wave height from the HIPOCAS-driven model varies between 0.62 m to 1.26 m, while the ERA40-driven model varies between 0.50 m to 1.36 m.

In terms of significant wave height seasonal variation, the modelling result from the research agrees with the past studies. A study by Khairunnisa et al. [24] shows that the significant wave height within Sunda Strait is mostly below 1 m through the entire year of 2016. Rachmayani et al. [20] uses the wave spectrum as the data input and also show that the highest average waves are found in January and November. The research also shows that the wave height is below 1 m within the same domain of this study.

Generally, the research is along with the mentioned two studies seems to underestimate the actual extreme wave height. As reported in [2], the wave height can be as high as 5 m. The extreme significant wave height extracted from the ERA 5 wave data also shows that the wave can only reach as high as 2.80 m as shown in Table 2.
5. Conclusion
The sensitivity analysis has been performed in the SWAN significant wave height model of the Sunda Strait. It is found that more data points lead to a more accurate result and less deviated dataset. We find that using 0.25 degrees of wind forcing spatial resolution results in good agreement with the ERA5 dataset. We acquire the best case of 0.12 m RMSE and 0.93 correlation coefficient rating. Reducing the resolution to 0.50 degrees shows a slight increment of error with 0.14 m RMSE and 0.85 correlation coefficient. Using a 1.00 degree of spatial resolution makes the model unreliable with a maximum of 0.58 m RMSE and 0.01 correlation coefficient. In terms of the scatter index, we find that the best scenario only shows 0.6 value which still needs improvement.

In a more general hydrodynamic modelling context, we recommend using at least 0.5 degrees of wind forcing spatial resolution for open seas or straits with geographical features similar to the Sunda Strait. Spatial resolution coarser than 0.5 degrees shows the data to be unusable since the error and is too high with a very low correlation to the baseline. The model also agrees with past studies where the significant wave height within the Sunda Strait is mostly below 1 m.

However, we also recommend using wind forcing spatial with the highest possible resolution to ensure modelling accuracy. The use of more wind forcings also does not impact the simulation time significantly. Using our system, all scenarios can be finished under 24 hours. The research can still be further improved by using a finer resolution of wind forcing spatial resolution on the Sunda Strait model to observe the point of diminishing return in terms of modelling accuracy.

References
[1] Kabar Banten. Cuaca Ekstrem Bayangi Penyeberangan Merak Bakauheni https://www.kabar-banten.com/cuaca-ekstrem-bayangi-penyeberangan-merak-bakauheni/(Accessed on July 22 2020) (in Bahasa Indonesia).
[2] Okezone. Gelombang Tinggi hingga 5 Meter Hambat Arus Penyeberangan Kapal Ferry di Selat Sunda https://news.okezone.com/view/2017/11/29/1/45049/gelombang-tinggi-hingga-5-meter-hambat-arus-penyeberangan-kapal-ferry-di-selat-sunda(Accessed on July 22 2020) (in Bahasa Indonesia).
[3] Detik News. Penyebrangan Merak-Bakauheni Terhambat Cuaca Buruk https://news.detik.com/berita/d-1475727/penyebrangan-merak-bakauheni-terhambat-cuaca-buruk(Accessed on July 22 2020) (in Bahasa Indonesia).
[4] Banten Raya. Cuaca Buruk, Kapal Terlambat Bersandar http://bantenraya.com/berita/2019/11/02/10831/cuaca-buruk-kapal-terlambat-sandar(Accessed on July 22 2020) (in Bahasa Indonesia).
[5] Thoresen C A.2014. PORT DESIGNER’S HANDBOOK 3rd Ed. (London: Thomas Telford).
[6] American Petroleum Institute. 2014. API RECOMMENDED PRACTICE 2A-WSD 22nd Edition: Planning, Designing, and Constructing Fixed Offshore Platforms – Working Stress Design. (Washington, DC: API Publishing Services).
[7] Det Norske Veritas. 2015. Offshore Standard DNV-OS-C101: Design of offshore steel structures, general - LRFD Method (Hovik: DNG GL).
[8] Goda Y. 2010. RANDOM SEAS AND DESIGN OF MARITIME STRUCTURES 3rd Ed. (London: World Scientific).

Table 2. Extreme Significant Wave Height $H_s$ and Period $T_p$ Based on ERA 5 Data with Gumbel Distribution Method

| Return Period (year) | Point A (6.25 S 105.50 E) | Point B (5.25 S 106.00 E) |
|----------------------|---------------------------|---------------------------|
|                      | $H_s$ | $T_p$     | $H_s$ | $T_p$     |
| 1                    | 0.88  | 9.87      | 1.41  | 14.03     |
| 10                   | 1.17  | 11.25     | 2.19  | 13.15     |
| 50                   | 1.32  | 11.96     | 2.62  | 12.95     |
| 100                  | 1.38  | 12.25     | 2.80  | 12.85     |
[9] Toffolia A, Lefevre J M, Bitner-Gregersen E, and Monbaliu J. 2005. *App. Oc. Res.* **27** 281.
[10] Booij N, Ris R C, and Holthuijsen L. 1999. *J. Geophys. Res.* **104** C4-7649.
[11] Muliati Y, Tawekal R L, Wurjanto A, Kelvin J, and Pranowo WS. 2018., *Int. J. Geomate* **15** 114.
[12] Akipnar A, van Vledder G, Kömürçu M, and Özger M. 2012. *Contl. Shelf Res.* **50-51** 80.
[13] Muliati Y, Tawekal R L, Wurjanto A, Kelvin J, and Pranowo WS. 2018. *Int. J. Geomate* **16** 176.
[14] Oceanweather, Inc. SEAFINE: Southern South China Sea, Makassar Strait, Java Sea https://www.oceanweather.com/metocean/seafine/index.html (Accessed on July 22 2020) (in Bahasa Indonesia).
[15] Hoque MA, Perrie W, and Solomon SM. 2019. *J. of Oc. Eng. Sci.* **5** 19.
[16] Hamza W, Lusito L, Ligorio F, Tomasicchio GR, and D’Alessandro F. 2018. *Water* **10** 1.
[17] Rachmayani R, Ningsih NS, Adiprabowo SR, and Nurfitri S. 2018., *IOP Conf. Ser.: Earth Environ. Sci.* p 139.
[18] Hersbach H, Bell B, Berrisford P, Horányi A, Sabater J M, Noclas J, Radu R, Schepers D, Simmons A, Soci C, and Dee D. 2019. *ECMWF Newsletter* **159** 17.
[19] CNN Indonesia. BNPB: Seluruh Buoy Deteksi Tsunami di Indonesia Rusak https://www.cnnindonesia.com/nasional/20180930160115-20-334439/bnpb-seluruh-buoy-deteksi- tsunami-di-indonesia-rusak (Accessed on July 22 2020) (in Bahasa Indonesia).
[20] Gonçalves M, Rusu E, and Guedes Soares C. 2012. *MARITIME ENG.AND TECH.* (London: England) p 481.
[21] Gonçalves M, Rusu E, and Guedes Soares C. 2013. *J. of Coastal Research* **31** 326.
[22] Rusu L, Bernardino M, and Guedes Soares C. 2009. *J. Coastal Res.* **56** 1419.
[23] Leon S, and Guedes Soares C. 2008. *Coastal Eng.* **55** 920.
[24] Khoriiunnisa H, Lubis MZ, and Anurogo W. 2017. *J. App. Geospatial Inf.* **1** 19.