

Ocean Sci. Discuss., referee comment RC1  
<https://doi.org/10.5194/os-2021-2-RC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on os-2021-2**

Anonymous Referee #1

---

Referee comment on "An EMD-PSO-LSSVM hybrid model for significant wave height prediction" by Gang Tang et al., Ocean Sci. Discuss.,  
<https://doi.org/10.5194/os-2021-2-RC1>, 2021

---

This manuscript presents the use of a hybrid machine learning model to predict significant wave height. While the results of this study can be of importance, the paper in its current form needs to be significantly improved. The language does not meet the quality standards of scientific publications. There are several grammatical errors, incorrect use of words, vague statements, etc. In addition, I have some concerns about the methodology (see below). My recommendation is major revision.

Some examples of grammatical errors or format issues: (this is not meant to be exhaustive, please revise carefully the whole text)

Lines 21-23. "Furthermore.." please rephrase this sentence as it does not read well.

Line 103: "Ocean wave time series is a kind of complicated nonlinear..." "kind of" is usually not used in scientific writing

Line 117 "kind of natural oscillatory mode"

Line 159: "can be got by" this is incorrect

Line 239 "human experts have performed this task" human experts??

Line 245 "As can be seen"

Other technical issues:

Line 73: "Significant wave height is a complicated, nonlinear, dynamic system" This is not accurate. Significant wave height is not a system but a variable or parameter to characterize ocean waves, among others. Please correct.

There are two Table 3 (page 7 and 9)

How many hidden layers are used in the configurations of the three single models?

Line 170: "The lag is a type of prediction error that can also be found in other work on wave forecasting using single models" Please add corresponding reference.

Figure captions need to be improved. E.g. Figure 5: No need to repeat 41025 twice. Specify what one hour and three hours means. No need to discuss results in figure caption

Line 196: "The time series of ocean waves is" This is vague. Time series of what? Need to specify the parameter used to describe ocean waves. If significant wave height is used often, the authors might want to use  $H_s$  or SWH to not have to repeat it all the time.

Figure 7: time axis label missing

In terms of the methodology I am concerned about applying EMD to significant wave height time series, as it is not an oscillatory signal around zero like sea surface height would be. Significant wave height is a parameter obtained from a sea surface height time series by averaging the 1/3 largest waves or calculating the integral over the wave spectra. I wonder if it is appropriate to decompose significant wave height in IMF, which are oscillatory signals with zero-crossings, while significant wave height has no zero-

crossing and it is, by definition, non-negative.

My understanding is that the model only considers 5 h of data to predict the following 1 to 3 h. physically speaking the evolution of waves very much depend on surface winds. Shouldn't the model also consider information about surface winds as an input parameter?