Ocean Science Discussions

Ocean Sci. Discuss., author comment AC2
https://doi.org/10.5194/os-2021-84-AC2, 2021

Reply on RC2

Brandon J. Bethel et al.

Author comment on "Forecasting hurricane-forced significant wave heights using a long short-term memory network in the Caribbean Sea" by Brandon J. Bethel et al., Ocean Sci. Discuss., https://doi.org/10.5194/os-2021-84-AC2, 2021

Reviewer Comment 1: The data is provided by 10 buoys and the number of scenarios (as shown in table 2) is not big which may represent a challenge for the generality of the model. The choice of merging multiple observations in one time series (lines 75-76) is particularly important for preventing overfitting but the separation between training and testing is not motivated. Probably a better description of Figure 1 may be used help with this.

Response: We thank the reviewer for their close inspection of the manuscript. In response to another reviewer, we have updated Figure 1 to better demarcate the training and test cases from each other. Here, we have added the following text so that readers can have a better understanding of the motivation for the cases chosen for our test cases:

"Training sets were chosen to represent hurricanes that passed through the study area (Figure 1) over the 2010 – 2019, ten-year period to ensure that the model is sufficiently general to be used, regardless of prevailing environmental conditions/phenomena which may influence hurricanes of a particular year (e.g., El Niño/La Niña). Test sets, by contrast, were chosen to be extremely powerful systems as it is highly expected that TCs, due to anthropogenic climate change, would gradually grow in intensity and frequency over time. Hurricane Dorian, in particular, was chosen given the widescale destruction it caused in The Bahamas and will be used in a comparison with numerical model output in Section 3.4."

Reviewer Comment 2: The model proposed is a standard LSTM. This recurrent neural network is well known to be good in training time series. The choice of the hyperparameters is not described or motivated. The only information is about the number of epochs and the batch size but there is not mentioned any reason for these choices and there are not comparisons with others.

Response: We have updated the description of the LSTM used and corrected an error that was present in the original manuscript. We have chosen those parameters following experiments but not save the results for display within the article. We have provided references of our recent using similar settings to justify the applicability of those parameters. That section was updated as follows:

"LSTM is set up with four layers that correspond to a time step of four. The recursive linear unit (ReLu) was used as the activation function to maximize the model's ability to capture nonlinearities. The Adaptive Moment Estimation (Adam) optimizer is used to compute adaptive learning rates. The number of epochs was set to 100 and the batch size set to 3. Throughout each experiment, the operating parameters were held constant. These settings were chosen after experiments (not shown) as they produced the best results while avoiding overfitting. Similar settings can be found in Bethel et al. (2021a) and Zhou et al. (2021a, 2021b). The data was partitioned along a 70/30 split into training and validation datasets. For clarification, here, and only here, the word 'dataset' should be interpreted as a given test hurricane (the test set hurricanes of Table 2). A general model is trained using the training set hurricanes of Table 2, but the model is specified to a given test set hurricane using 70% of its time series, and the remaining 30% is used to validate the forecast. "

Reviewer Comment 3: The results in Figures 3,4,5 show some discrepancies between the forecasting and the observation. In case the authors are willing to, this may be solved by implementing an adversarially-trained LSTM.

Response: We thank the reviewer for close attention. Results in Figures 3, 4, and 5 were incorrectly plotted and were replotted which, excluding the LSTM-inherent phase shifting phenomenon issue, resolved this issue without needed to use an adversarially-trained LSTM. The phase shifting, most notable in Hurricane Igor, is a problem with LSTM itself and remains an open problem in the mathematics underlying this network. That being said, we would be happy to try the usage of an adversarially-trained LSTM in a future problem addressing the phase shifting in hurricane predictions and thank the reviewer for the suggestion.

Reviewer Comment 4: The authors provide results of the accuracy. However, it would be interesting to check the efficiency, at least in terms of execution times for both training-testing and for the forecasting.

Response: Following a few minutes of model training, forecasting occurred in a very small fraction of a second. We have added this feature of the model in Section 3.4 when we discuss the model in comparison with a numerical model (SWAN). We thank the reviewer for the suggestion.

Reviewer Comment 5: As the authors state in the introduction (from lines 29), there are other methods for nowcasting which are actually used for these purposes. The paper may provide comparisons in terms of accuracy or/and efficiency with these methods at least in terms of order of magnitude.

Response: We thank the reviewer for the suggestion. Another reviewer has also made the same suggestion and thus, we introduce Section 3.4 to compare LSTM output with a SWAN numerical model hindcast from the perspective of accuracy, computational requirements and required level of technical expertise.