Author: Jiang, Haoyu; Babanin, Alexander V.; Chen, Ge
Title: Event-based validation of swell arrival time
Year: 2016
Journal: Journal of Physical Oceanography
Volume: 46
Issue: 12
Pages: 3563-3569
URL: http://hdl.handle.net/1959.3/433734

The definitive version is available at: https://doi.org/10.1175/JPO-D-16-0208.1

Event-based Validation of Swell Arrival Time

Haoyu Jiang[1,2], Alexander V. Babanin[2,3], and Ge Chen[1,4]

[1] Qingdao Collaborative Innovation Center of Marine Science and Technology, College of Information Science and Engineering, Ocean University of China, Qingdao, China

[2] Centre for Ocean Engineering, Science and Technology, Swinburne University of Technology, Melbourne, Victoria, Australia.

[3] Department of Infrastructure Engineering, Melbourne School of Engineering, University of Melbourne, Victoria, Australia.

[4] Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, China

Corresponding author: Alexander V. Babanin (a.babanin@unimelb.edu.au)

ABSTRACT

The arrival time of ocean swells is an important factor for offshore and coastal engineering, naval and recreational activities which can also be used in evaluating the numerical wave model. Using the continuity and pattern of wave heights during the same swell event, a methodology is developed for identifying swell events and verifying swell arrival time in models from buoy data. The swell arrival time in a WAVEWATCH-III hindcast database is validated by in-situ measurements. The results indicate that the model has a good agreement with the observations, but usually predicts an early arrival of swell, about 4 hours on average. Histogram shows that about one quarter of swell events arrive early and three quarters late by comparison with the model. Many processes that may be responsible for the arrival time errors are discussed, but at this stage it is not possible to distinguish between them from the available data.

1. Introduction

Storms in the ocean can generate long surface gravity waves propagating away from their sources as swells. These waves can propagate over thousands of kilometers with little energy loss (e.g., *Snodgrass et al*. 1966; *Collard et al.* 2009; *Ardhuin et al*. 2009), radiating momentum and energy across ocean basins (*Munk et al*. 1963). Swells have impacts on many aspects of the human life, from industrial activities such as port operations to recreational activities such as surfing. They are also important to many physical processes of the Earth system such as momentum exchange at the air-sea boundary and sediment transport in the coastal areas.

Beginning with the needs of forecasting waves in wars, the studies of swells, as well as their forecast have been conducted since the 1940s (e.g., *Barber and Ursell* 1948; Rogers et al. 2014). Through years of development, numerical wave models nowadays can give a fairly good forecast of many swell parameters comparing with observations (e.g. *Ardhuin et al*. 2010; *Zieger et al.* 2015). To evaluate or verify a numerical model, the wave parameters from altimeters or buoys are usually collocated with the model outputs in the same time and location, and comparisons are made between two sets of data. In this process, each model-observation data pair is regarded as being independent of others, and the statistics of these data are employed as the evaluation criterion for models. For the swells coming from the same storm, their wave parameters vary continuously and are correlated, and are independent of swells from other storms. Therefore, the above comparison of wave parameters in fixed time might be still insufficient to fully demonstrate the performance of models for swells. As the swells can travel large distances, the arrival time of a swell event should be also taken into consideration when evaluating a model, which means a comparison in the dimension of time. In fact, the arrival time of swells itself is a very practical parameter as many human activities are time-sensitive, such as arranging enough time for contingency plan or forecasting the best time for surfing.

Some previous studies mentioned the idea of swell arrival time and tried to make some comparison of it between the model and observation (*Wingeart et al.* 2001; *Delpey et al.* 2010; *Ardhuin et al.* 2016). Meanwhile, anecdotal pieces of evidence are full of stories of both early and late arrival of swells by comparison with the model forecast. However, it seems there is still no clear definition of the "arrival time" of swell, which makes it hard to do the comparison quantitatively. The aim of this study is trying to present a consistent and robust method to validate the model performance for the arrival time of swells. Using this method, we also identified the problem of the swell early/late arrival which needs to be explained and solved in the future. The data used in this study and the methods to isolate swell information from the same event and to estimate the error of swell arrival time is described in Section 2. Next, some results of arrival time comparison and the discussion on swell early/late arrival are presented in Section 3, followed by the summary in Section 4.

2. Data and Methods

a. Data

To investigate the arrival time of swells, the variation of wave information against time, that is, a successive time series, is needed. Therefore, quality controlled in-situ measurements from the National Data Buoy Center (NDBC) are employed as the reference data for the model output. To compare with the global model output, 13 deep water buoys (32012, 46011, 46012, 46015, 46050, 46086, 46089, 51000, 51001, 51004, 51028, 51100, and 51101) sufficiently far from the coastlines with directional spectral information available during any periods between 2000 and 2012 are selected, and a 3-h running average is applied to smooth the temporal variation of the spectral data. The spectra are partitioned using the procedure of *Portilla et al.* (2009) without identifying wind-seas and swells.

The model hindcast data used in this study are two-dimensional spectra computed by WAVEWATCH-III® (WW3) with the physical parameterization of *Ardhuin et al.* (2010). This parameterization produces a swell wave height compare well with measurements and is widely used in many operational applications (*Stopa et al.* 2015). The model hindcast has two-dimensional spectral outputs in the buoys' locations from 2000 to 2012 with a time resolution of 1 h and a spectral resolution of 15° and 32 frequency bins exponentially space from 0.037 to 0.75 Hz. The data is downloaded from the IFREMER ftp server where more detailed information about them is also available. These wave spectra are also partitioned using the method of *Portilla et al.* (2009) without identifying wind-seas and swells.

b. Swell event identification and collocation

For comparison of arrival time, wave information at a given location from the same meteorological event (e.g. an extratropical storm or a tropical cyclone) needs to be organized into one time series which is regarded as a wave event. This event-assignment is operated using the same idea of wave tracking technology as that proposed by *Hanson and Philips* (2001). As we only focus on the arrival time of swells, three criteria are adopted to filter the noises and wind-sea events after the wave tracking: (1) The maximum of the peak periods needs to be more than 12 s. (2) The duration of the event should be more than 60 h. (3) The peak frequencies in the same event need to rise significantly (P = 0.01) with time. After the identification, the event from the buoy is collocated with the event from the model as the same event if (1) the two events have more than 45 h intersections in the range of time and (2) the difference in average swell wave height, peak frequencies, and peak wave directions between the two events are less than 1m, 0.02 Hz, and 30° respectively.

Three examples of collocated swell events are shown in Figure 1 from which it is clear that the measurement and the model outputs represent the same event. Some fundamental features of swell events are nicely shown in both data sets: the peak frequency increases nearly linearly with time due to the deep water dispersion relation while the peak direction remains almost constant during the whole event. The variation of wave height generally behaves as a convex function of time which first increases since the forerunner's arrival, and then decays after the energy peak of the wave group passed.

c. Arrival time difference

There seems to be no recognized definition of swell arrival time. *Wingeart et al.* (2001) simply use the time when the energy of a swell event is detected, an intuitive definition. The shortcoming of this method is that it depends much on the sensitivity of the measurement, as the forerunner of a swell event usually has low energy. For instance, in the right panel of Figure 1, the first detected wave height in the model data is 0.5 m while that of the buoy is only 0.3 m. *Delpey et al.* (2010) used different arrival time for different wave periods. This seems to be a practical definition, but people care more about the arrival of swell energy instead of frequency in most applications. Also, as the peak frequency rises nearly linearly with time, it might be hard to distinguish between the early/late arrival of swells and the overestimation/underestimation of peak frequency. *Ardhuin et al.* (2016) used the time when the swell energy reaches its peak as the arrival time. The peak is usually well-defined as the energy evolution in a given location is in general a convex function and its energy is large enough to be detected. Yet, due to the data noises and the complexity of the real swell in the nature, sometimes there are more than one energy peak in the event (e.g., wave height in the middle panel of Figure 1) and the peak might be not very sharp (e.g., wave height in the right panel of Figure 1).

The wave parameters vary continuously with time in each swell event, thus, it is not necessary to define the exact arrival time of a swell event. To validate the model, it is more important to know the difference of swell arrival time of each event between model and observation. Here we use wave height as the parameter to estimate this time difference, because (1) the wave height is, until now, the best-forecasted wave parameters (e.g. *Stopa et al.* 2015, Zieger et al. 2015), and (2) the variation of wave height is, in general, a convex function which has more features to match the model with the observation and is not sensitive to the systematical error on wave heights.

The method to define the time difference between two events is to find their best match by shifting one of the sequences in time. When the two events are best matched, they should have the least average error and the best correlation. Therefore, we use two schemes to define the arrival time difference by (1) the maximum of normalized cross-correlation and (2) the minimum of shift of the normalized root mean square difference (NRMSD):

$$r(\tau)=\frac{\sum_{t=1}^{n}\left[H_B(t-\tau)-\overline{H_B}\right]\left[H_M(t)-\overline{H_M}\right]}{\sqrt{\sum_{t=1}^{n}\left[H_B(t-\tau)-\overline{H_B}\right]^2}\sqrt{\sum_{t=1}^{n}\left[H_M(t)-\overline{H_M}\right]^2}} \qquad (1)$$

$$\sigma(\tau) = \sqrt{\frac{\sum_{t=1}^{n} \left[H_B(t-\tau) - H_M(t) \right]^2}{n\overline{H_B}^2}} \tag{2}$$

where $H_B(t)$ and $H_M(t)$ are the time series of swell heights from the buoy and the model respectively for a given swell event, τ is the shifting time of buoy measurements, and n is the time span when model outputs and buoy measurements overlap. The variation of cross-correlation and NRMSD are almost symmetrical with respect to each other, thus, these two definitions of arrival time difference are consistent (The 4th line of Figure 1). Using this consistency, some wrong-identified and wrong-collocated events can also be excluded from our dataset: if the time difference defined by the NRMSD corresponds to a correlation coefficient significantly smaller than the maximum (P = 0.05), the collocated data pair will be eliminated from our analysis, and vice versa.

3 Results and discussions

Using the above method, 421 collocated swell events are captured to estimate the arrival time difference between the model outputs and in situ observations. The distribution of the difference in swell arrival time of all these cases is shown in Figure 2. From the scatter plot (a), it is clear that the results derived from the two definitions are consistent. In 331 cases out of the 421 ones, the estimated time lags are within 1 hour, showing the validity of the definition of the arrival time difference. Regarding the distributions of the arrival time errors in the model, both the histograms of Figure 2 (b) and (c) show a unimodal distribution which can be regarded as normal ones with a mean value of about minus four hours, and more than three quarters of the data show the minus time differences in the model (that is, swell arrives early in the model by comparison with the observations). As there is no other systematic and quantitative validation of numerical wave models from the perspective of time, there is no specific reference for the model's performance regarding swell arrival time. However, considering the swells have usually propagated for more than 100 hours before reaching the buoys, this is a small error. This model output has a good performance on forecasting the swell arrival time in general, but for practical purposes, the difference of many hours can be significant. The result indicates that the swell arrival on average tends to delay with respect to the model forecast, and a delay of ten hours which might have some impacts on the model's application is still possible, at least at the given observational locations.

The wave height comparison of these 421 swell events between model output and buoy observation is shown in Figure 3(a) from which it is clear that the model can already give quite good results for wave height. But after shifting the wave heights sequence with the arrival time difference (here, we use the average of the two time differences to give a more robust estimation), the result can be even further improved. After the correction in Figure 3(b), the total number of collocated pairs reduce 95, but the correlation coefficient between the two series increased and the RMSD between them decreased. The time shifting should not impact the bias, but the shifting changed the data number of collocated time so that the bias also reduces a bit. Especially, the suspected outliers in Figure 3(a) are all eliminated in Figure 3(b). These features all demonstrate that the differences in the arrival time of swell between model and observation are authentic, not due to the noises or errors in the observation.

Considering that the swell can propagate over thousands of kilometers, a small difference in wave parameters related to propagation such as wave direction and the group speed can be magnified over such long distances, and will lead to a larger error in the far field. Although the propagation of the wave is believed to be understood from first principles, there are many potential possibilities to explain the differences in swell arrival time. One category of the possibilities is that the error is embedded in the generation area of the swells. The errors of wind input, resonant wave-wave interaction, and dissipation term can all lead to errors in energy distributions along different frequencies. Another category of possibilities is that the error is due to the processes in propagation, such as non-breaking dissipation, wave-current interactions, or some nonlinear effects. According to *Ardhuin et al.* (2009) or *Babanin et al.* (2012), the swells with higher frequencies will have higher dissipation rate, which may cause the downshifting of energy peak along frequency in the swell field and "accelerate" the swells. Large-scale currents can impact the absolute group speed of the swells by

Doppler Effect while mesoscale eddies could refract the swells (*Gallet and Young*, 2014) prolonging the propagation paths of swells and delay the swell arrival. Besides, the Raman-like effect of the modulational instability of wave trains, the adverse currents with gradients, and the interactions between wind and wave in the course of swell propagation can also change the swell carrier frequency.

If the difference in arrival time is due to the errors in energy distribution along frequency in generation area, or frequency shifting effects during propagation, the delayed/early arrival of a swell event means the swell energy peaks shifted to a higher/lower frequency. However, these effects will not have much impact on the swell peak frequency at the arrival point, as the peak frequency is determined by the spatiotemporal position of the swell source. In this case, there will be no underestimation or overestimation of the peak frequencies at a given time in the model. On the contrary, if the difference is due to the position error of source or the impact of currents such as refraction, the propagation time will be actually shorten or prolonged. Then the delayed/early arrival of a swell event in wave heights will also correspond to a delayed/early arrival in frequency, which will behave as the peak frequency being overestimated/underestimated in the model. In our dataset, the model on average overestimated the frequency for 1.4%, which is also corresponding to a delayed arrival of about 4 hours. This seemingly supports the conjecture of the prolonging propagation time. However, the correlation coefficient is calculated between arrival time differences and the average bias of frequency which is found not to be significant. Other parameters including the RMSD of peak frequency, the RMSD of swell direction, and the distance from the storm estimated by the slope of the peak frequency also all give little correlation with the differences in swell arrival time. Based on the data involved in this study, we cannot underpin physics of swell delay or early arrival, and much more effort is needed to answer this question. All of the above explanations are reasonable so that the arrival time error might be the superposition of many physical processes. More data and efforts are needed to split and identify them, particularly the data along swell propagation.

4. Summary

Using the pattern of wave height during a swell event, a methodology of validating the swell arrival time from numerical wave models compared with buoy data was developed. Measurements from the same swell events were organized and identified by tracking the partitioned parameters of energy spectrum in time. The difference of swell arrival time between model and buoy is defined using the maximum of cross-correlation and the minimum of shifting NRMSD, two consistent estimations. The comparison was made between WW3 hindcast data and 13 NDBC deep water buoys far from coastlines with spectral information. The result from 421 swell events indicated that the wave model has a good prediction of the arrival time of swells, but the swells are on average about 4 hours delayed to the model. The present method of comparing swell arrival time provides another perspective of evaluating and validating numerical wave model, which could be applied synergistically with other validation methods.

We identified the problem of swell early/late arrival which is an interesting question itself and deserves much effort. As the swell arrival time is related to both the generation and the propagation of swells, it is a complicated problem and many processes might be responsible for its difference between model and observation. Many possible conjectures having impacts on the swell arrival time are discussed, but we are unable to distinguish between different mechanisms from available data at this stage. Many further studies are needed in the future to split the problem and figure out the importance of each factor on the early/late arrival of swells.

Acknowledgments.

The model data are from IFRMER FTP server (ftp.ifremer.fr/ifremer/ww3/HINDCAST). The buoy data are from NDBC (www.ndbc.noaa.gov). The authors would like to thank Dr. Henrique Rapizo for finding that swells are often delayed to the forecast in his surfing life. This work is supported by the National Natural Science Foundation of China (41331172 and U1406404). AVB acknowledges the Australian Research Council Discovery grant DP130100227 and the US Office of Naval Research grant N00014-13-1-0278. HJ is also supported by China Scholarship Council.

References

Ardhuin, F. B. Chapron, and F. Collard, 2009: Observation of swell dissipation across oceans, *Geophys. Res. Lett.* **36**, L06607, doi:10.1029/2008GL037030.

Ardhuin, F. P. Sutherland, M. Doble, and P. Wadhams, 2016: Ocean waves across the Arctic: Attenuation due to dissipation dominates over scattering for periods longer than 19 s, *Geophys. Res. Lett.* **43**, 5775–5783, doi:10.1002/2016GL068204.

Ardhuin, F. et al., 2010: Semi-empirical dissipation source functions for wind-wave models: Part I, Definition, calibration and validation, *J. Phys. Oceanogr.* **40**, 1917–1941.

Babanin, A. V., 2012: Swell attenuation due to wave-induced turbulence, *paper presented at 31st International Conference On Ocean, Offshore and Arctic Engineering Division*, Riode Janeiro, Brazil, 1–6 July.

Barber, N. F. and F. Ursell, 1948: The generation and propagation of ocean waves and swell. Part I. Wave periods and velocities, *Phil. Trans. R. Soc. London A*, **240**, 527–560.

Collard, F. F. Ardhuin, and B. Chapron, 2009: Monitoring and analysis of ocean swell fields from space: New methods for routine observations, *J. Geophys. Res.* **114**, C07023, doi:10.1029/2008JC005215.

Delpey, M. T. F. Ardhuin, F. Collard, and B. Chapron, 2010: Space-time structure of long ocean swell fields, *J. Geophys. Res.* **115**, C12037, doi:10.1029/2009JC005885.

Gallet, B. and W. R. Young, 2014: Refraction of swell by surface currents, *J. Mar. Res.* **72**, 105–126.

Hanson, J. L. and O. M. Phillips, 2001: Automated analysis of ocean surface directional wave spectra, *J. Atmos. Oceanic Technol*. **18**, 277–293.

Munk, W. H. G. R. Miller, F. E. Snodgrass, and N. F. Barber, 1963: Directional recording of swell from distant storms, Philos. *Trans. R. Soc. London, Ser. A*, **255**, 505 – 584.

Portilla, J. F. J. Ocampo-Torres, and J. Monbaliu, 2009: Spectral Partitioning and Identification of Wind Sea and Swell, *J. Atmos. Oceanic Technol.* **26**, 107-122

Rogers, W. E. J. D. Dykes, and P.A. Wittmann, 2014: US Navy global and regional wave modeling, *Oceanography*, **27**, 56–67.

Snodgrass, F. E. G. W. Groves, K. Hasselmann, G. R. Miller, W. H. Munk, and W. H. Powers, 1966: Propagation of ocean swell across the Pacific, *Philos. Trans. R. Soc. London, Ser. A*, **249**, 431–497.

Stopa, J. E. F. Ardhuin, A. V. Bababin, and S. Zieger, 2015: Comparison and validation of physical wave parameterizations in spectral wave models, *Ocean Model.* doi:10.1016/j.ocemod.2015.09.003

Wingeart, K. M. T. H. C. Herbers, W. C. O'Reilly, P. A. Wittmann, R. E. Jensen, and H. L. Tolman, 2001: Validation of operational global wave prediction models with spectral buoy data, *in Proceedings of the 4th International Symposium Ocean Wave Measurement and Analysis*, pp. 590–599, ASCE, San Francisco, Calif. doi:10.1061/40604(273)61.

Zieger, S. A.V. Babanin, W. E. Rogers, and I. R. Young, 2015: Observation-based source terms in the third-generation wave model WAVEWATCH, *Ocean Model.* **96**, 2-25.

List of Figures

FIG. 1. Peak frequencies (1st line), wave heights (2nd line), peak directions (3rd line), and arrival time differences (4th line) of collocated swell events from NDBC platform 46086 in 2007 (left), 51101 in 2011 (middle), and 32012 in 2008 (right): In the first three lines, the blue dots are buoy measurements and the red dots are model outputs. In the 4th line, the red lines are the moving correlation coefficients and the blue lines are the moving NRMSDs (see Equation 1 and 2 for the definitions), and the shadow areas represent the 95% confidence intervals respectively. The respective dot vertical lines are where the maximum of moving correlation coefficients and the minimum of moving NRMSDs appear.

FIG. 2. (a) The scatter plot of the arrival time differences defined by shifting NRMSD versus those defined by cross-correlation, with respective histograms of arrival time differences (b) and (c).

FIG. 3. The scatter plot of collocated wave heights of WW3 hindcast against NDBC buoys for the 421 swell events (a) without any correction, and (b) after shifting the wave heights sequence with the arrival time difference.

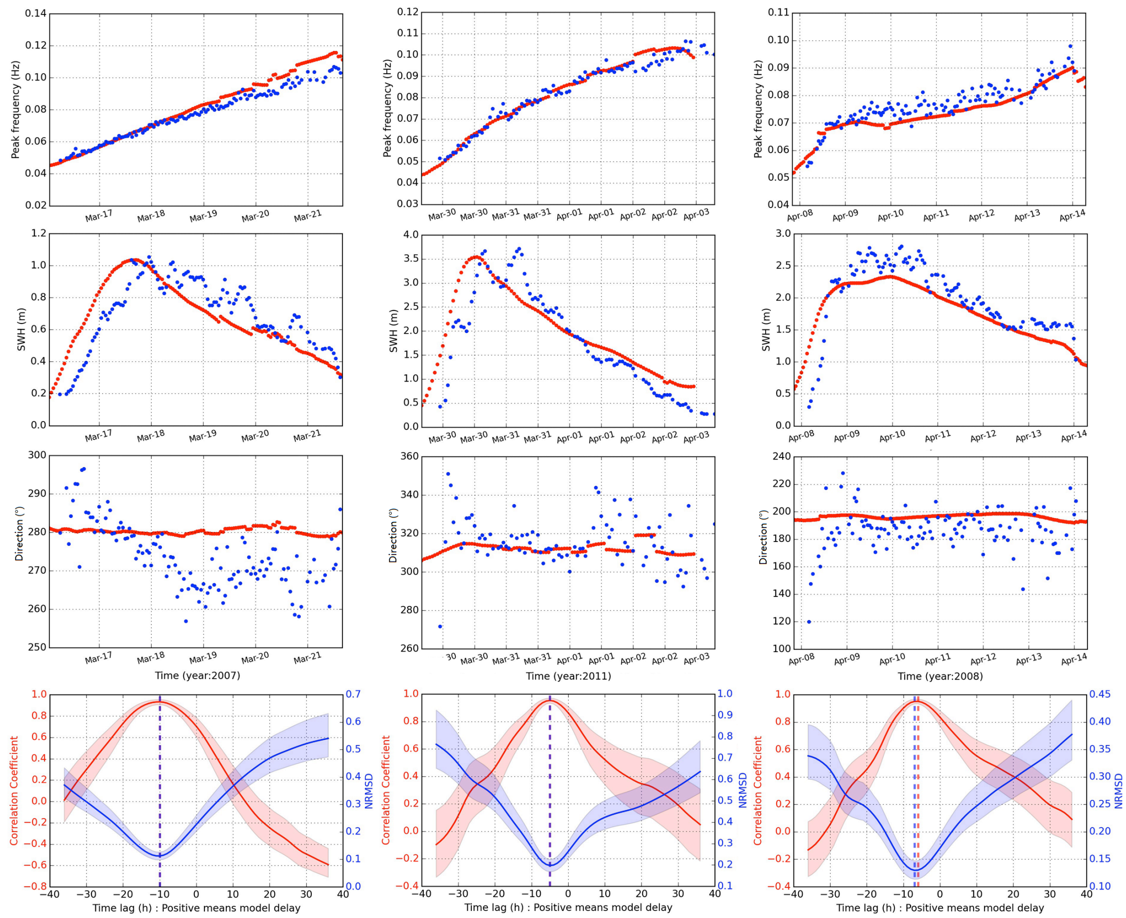

FIG. 1. Peak frequencies (1st line), wave heights (2nd line), peak directions (3rd line), and arrival time differences (4th line) of collocated swell events from NDBC platform 46086 in 2007 (left), 51101 in 2011 (middle), and 32012 in 2008 (right): In the first three lines, the blue dots are buoy measurements and the red dots are model outputs. In the 4th line, the red lines are the moving correlation coefficients and the blue lines are the moving NRMSDs (see Equation 1 and 2 for the definitions), and the shadow areas represent the 95% confidence intervals respectively. The respective dot vertical lines are where the maximum of moving correlation coefficients and the minimum of moving NRMSDs appear.

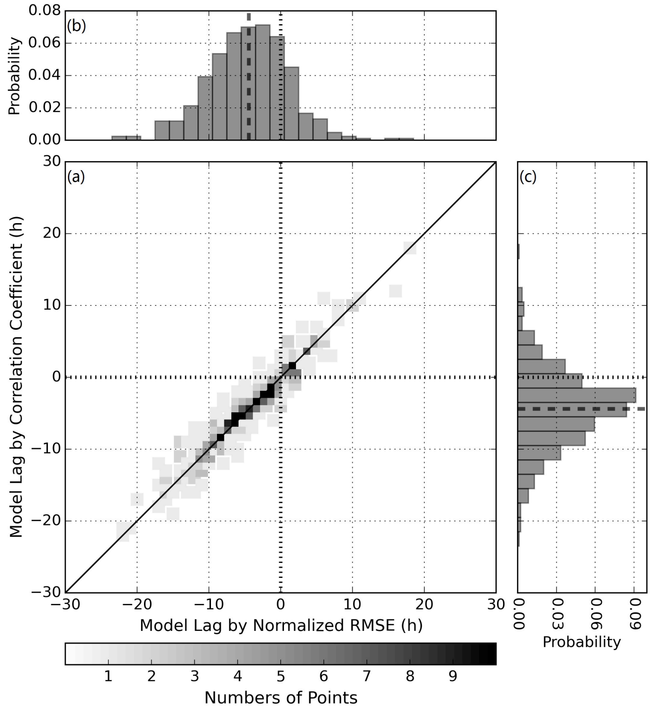

FIG. 2. (a) The scatter plot of the arrival time differences defined by shifting NRMSD versus those defined by cross-correlation, with respective histograms of arrival time differences (b) and (c).

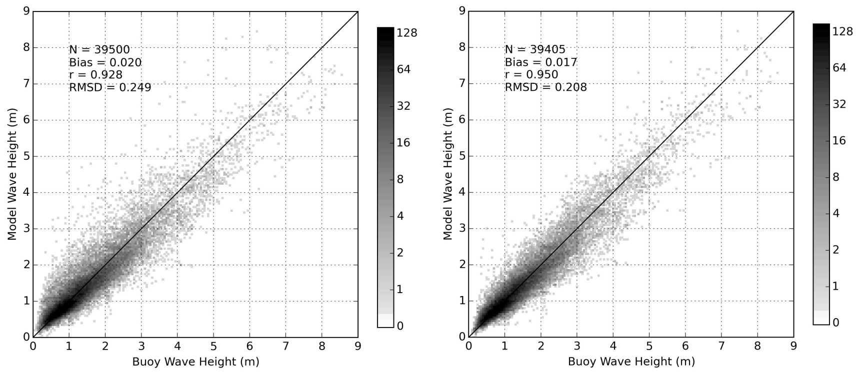

FIG. 3. The scatter plot of collocated wave heights of WW3 hindcast against NDBC buoys for the 421 swell events (a) without any correction, and (b) after shifting the wave heights sequence with the arrival time difference.