



ESTABLISHMENT AND PERFORMANCE OF THE OCEAN WAVE ENSEMBLE FORECAST SYSTEM AT CWB

Heng-Wen Chang

Green Energy and Environment Research Laboratories, Industrial Technology Research Institute, Hsinchu County, Taiwan, R.O.C. Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei County, Taiwan, R.O.C.

Chieh-Cheng Yen

Green Energy and Environment Research Laboratories, Industrial Technology Research Institute, Hsinchu County, Taiwan, R.O.C.

Ming-Chung Lin

Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei County, Taiwan, R.O.C.

Chi-Hao Chu

Central Weather Bureau, Ministry of Transportation and Communication, Taipei County, Taiwan, R.O.C.

Follow this and additional works at: <https://jmstt.ntou.edu.tw/journal>

Recommended Citation

Chang, Heng-Wen; Yen, Chieh-Cheng; Lin, Ming-Chung; and Chu, Chi-Hao (2017) "ESTABLISHMENT AND PERFORMANCE OF THE OCEAN WAVE ENSEMBLE FORECAST SYSTEM AT CWB," *Journal of Marine Science and Technology*: Vol. 25: Iss. 5, Article 11.

DOI: 10.6119/JMST-017-0622-1

Available at: <https://jmstt.ntou.edu.tw/journal/vol25/iss5/11>

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

ESTABLISHMENT AND PERFORMANCE OF THE OCEAN WAVE ENSEMBLE FORECAST SYSTEM AT CWB

Acknowledgements

This research is funded by the CWB through the Establishment of Wave Ensemble forecast System project. We thank Dr. Hong Ching-Shan, Dr. Chang Hui-Ling and Dr. Terng Chuen-Teyr for support and valuable discussions. We also acknowledge very useful source of buoy information available from the Marine Meteorology Center of CWB.

ESTABLISHMENT AND PERFORMANCE OF THE OCEAN WAVE ENSEMBLE FORECAST SYSTEM AT CWB

Heng-Wen Chang^{1,2}, Chieh-Cheng Yen¹, Ming-Chung Lin², and Chi-Hao Chu³

Key words: wave ensemble forecast system, reliability diagram, relative operating characteristic, ensemble spread.

increase SPRD in the near future.

ABSTRACT

A wave ensemble forecast system is being developed based on the NOAA WAVEWATCH III (NWW3) two nesting multi-grid model over Taiwan area. The ensemble system consisted of 20 ensemble members and was set with spatial resolutions of 0.25° and 0.1° . The wind forcing is coming from the WRF-based ensemble forecast system (WEPS) 10 m wind field of Central Weather Bureau (CWB). The cycle initial condition of each wave ensemble member from the previous run of the same ensemble member is applied to generate a history perturbation of swell. The objectives of this work are to verify the impact of different wind forcing formulas, to find the better composition of ensemble members, and to evaluate the forecast capacity of resulting ensemble forecast system. We proposed the combination of using two built-in wind forcing formulas to form the ensemble members, which can reserve the advantages of different formulas under various wind fields (monsoon and typhoon period), increase the average ensemble spread (SPRD) and decrease the difference between the root mean square error (RMSE) and average SPRD based on the truth value at open seas. With Reliability diagram, Brier Skill Score and Relative Operating Characteristic analyses, the ensemble system has better forecast skill than the operational deterministic forecast and can discriminate between the events and non-events. Nevertheless the overestimation near the coast could be improved by increasing the grid resolution and resolving nearshore wave simulation to reduce RMSE. For the underestimation of SPRD we intend to add perturbation at low frequency swell as initial condition to

I. INTRODUCTION

In recent years, due to the improvement of numerical forecast capacity and the advancement of computing power, numerical forecast has been implemented in applications of wave forecast products and served as the main guideline of wave forecast. Nevertheless, the numerical forecast is subject to the restricted understanding of physical phenomena, physical parameterizations and numerical methods. Moreover, without sufficient observation, errors and uncertainties still exist in the numerical wave forecast and increase with the forecast hour.

From the perspective of users or data providers, the deterministic forecast could not meet their requirements because they always ignore the forecast errors that existed. In order to make up the deficiency of deterministic forecast and meanwhile expect that a forecast can estimate model uncertainties, ensemble forecast was developed which could quantify forecast uncertainties and provide the probability forecast by building different ensemble members. Besides it can also facilitate forecasters and decision makers in respect of prediction analysis. In general, a forecast explicitly cast in probability terms is better not only because it provides the user with an estimate of the error, but because it is more realistic and truthful. So a probability forecast conveys a message which explicitly reminds the user that there is always a forecast uncertainty which should be considered, computed and taken into account when making any practical decision from the forecast. In fact even deterministic forecasts are in reality probability forecasts in disguise, since an error can and should always be associated with it (Massel, 2013).

The propagation of wind waves belongs to a process with weakly nonlinearity and high dissipative. However, since the interaction of wind and wave and the nonlinearity of wind fields are significant factors for energy into wave, uncertainties of such two factors make it more difficult to grasp the forecast accuracy of wind waves. Therefore, the deterministic forecast could not simulate real wave fields precisely. Chen (2006) who evaluated two cases of the storm within three months after the operation of NCEP ensemble forecast system, indicated that the ensemble system is more realistic and a better tool for decision making

Paper submitted 11/28/16; revised 05/09/17; accepted 06/22/17. Author for correspondence: Heng-Wen Chang (e-mail: BHWChang@itri.org.tw).

¹Green Energy and Environment Research Laboratories, Industrial Technology Research Institute, Hsinchu County, Taiwan, R.O.C.

²Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei County, Taiwan, R.O.C.

³Central Weather Bureau, Ministry of Transportation and Communication, Taipei County, Taiwan, R.O.C.

than the deterministic system.

Reviewing the development of the wave ensemble forecast system, the National Centers for Environmental Prediction (NCEP) has developed and implemented the Global Ensemble Ocean Wave Forecast System (GEOWaFS) operationally. GEOWaFS contains one control run and ten ensemble members by using (NWW3) version 2.22 as the wave model to generate global wave data. Its resolution was 1.0° in latitude and 1.25° in longitude from 78°S to 78°N . The wind forcing of the ensemble members were obtained from NOAA/NCEP Global Ensemble Forecast System (GEFS) 10 meter high wind fields (U_{10}), which was generated using Global Forecast System (GFS) by breeding method. Its initial condition of each member started from the same initial conditions. It made 126 hours forecast four times a day since 2006.

New NWW3 model version 3.14 possess the multi-grid approach which features two way nesting where grids with various resolution in a single wave model. GEOWaFS has updated to this new model. More important modifications to the ensemble system include $1.0^\circ \times 1.0^\circ$ spatial resolution of grids, ensemble member increased to 20 members, adding a control run forced by the GFS at the ensemble resolution, cycle initial conditions of each member from the previous cycle run of the same ensemble member, use of bias-corrected GEFS wind fields instead of raw GEFS winds, extending to the forecast to 10 days. The new GEOWaFS was in operation since June 1, 2008.

Additionally, NCEP started to have cooperation with the Fleet Numerical Meteorology and Oceanography Center (FNMOC) of US Navy since November 1, 2011. The combined NCEP/FNMOC wave ENSEMBLES (NFCENS) with 40 ensemble members (each with 20) were established. A detailed description is provided in Alves et al. (2013). Two centers used the same wave model and settings, with the differences regarding input sources of ensemble wind fields.

In 1998, the coupling between wind (IFS) and waves operationally had been completed by European Center for Medium-range Weather Forecasting (ECMWF). The wave model in the Wave Ensemble Prediction System meteorogram (Wave EPSgram) is the ECMWF version of WAM cycle 4. The system consisted of 51 ensemble members including one control run. All ensemble members use the unperturbed analysis as the initial condition. The divergence between the wave ensemble members is therefore due only to different wind forcing when the coupled atmospheric ensemble members are subject to different evolutions. The model wind input is obtained from EPS by singular vectors method. Meanwhile, spectrum resolutions of Wave EPSgram were increased to 24 directions and 30 frequencies. In 2010, the spatial resolution became $0.5^\circ \times 0.5^\circ$ and the forecast hour extended to 10 days for every 6 hours (http://www.ecmwf.int/products/forecasts/guide/Wave_EPSgrams.html).

The establishment of the wave ensemble forecast system at CWB with 20 ensemble members had been accomplished in June 2014. NWW3 version 3.14 was used with the multi-grid calculation (0.25° , 0.1°). Computation domains and water depth of the wave model is shown in Fig. 1. Wind forcing was ob-

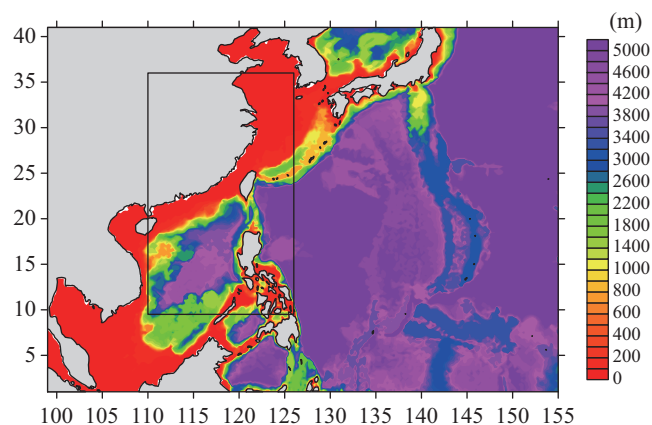


Fig. 1. Computation domains.

tained from WEPS of CWB with two spatial grid resolution (45 km and 15 km). The forecast time is expanded to 72 hours for every 6 hours.

The main outputs of ensemble system include point output and gridded output. The former utilizes boxplots to show ensembles (include significant wave height H_s , wind speed U_{10} , mean wave direction T_{m02}), ensemble mean and observed data in previous 48 hours. The latter contains 20 ensembles, ensemble mean and spread, 10% exceeding probabilities, probability and spaghetti diagrams at different thresholds every three hours. Those outputs are posted on the following website: <http://61.56.11.156/ens/viewernewd.htm>

The main goal of this work are to investigate the influence of different wind forcing input formulas, to find the better compositions of ensemble members, and to evaluate the performance of resulting operational ensemble forecast system at CWB to find to which extent the ensemble system resolves the uncertainties.

This paper is organized in the following manner: The buoy data, model parameters and methodology of analysis indexes are explained in section 2; in section 3, the ensemble members determination procedure and resulting performance of ensemble system for wind and wave are discussed separately; some discussions and conclusions are drawn in section 4.

II. BUOY DATA FOR VERIFICATION AND METHODOLOGY OF ANALYSIS

1. Buoy Data for Verification

When conducting the analysis, truth value must be presented. The study took the hourly (sometimes 2 hours) and quality controlled observed buoy data from CWB as the truth value. Fig. 2 shows locations of buoys. Among that, buoys No. 14 and No. 15 are located in open seas and called sea stations. The others which called coast stations are located a distance less than 5 km to the coast. Buoys No. 16 and stations No. 17 are not available in 2012. Buoys No. 7 stopped working after 2012.

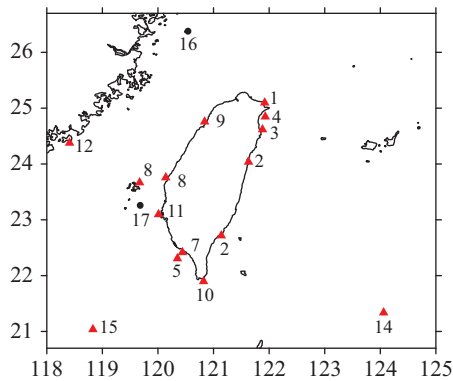


Fig. 2. Buoy locations used in the ensemble wave forecast verification.

2. Model Options and Parameters

Important parameters of numerical models contain 25 frequencies from lowest frequency 0.04178 Hz, frequency increment factor 1.1 and 15° directional resolution. Cycle initial conditions of each member refers to Cao (2009) are taken from the previous cycle run of the same ensemble member to generate a natural history of perturbations of swell. The method could avoid initial zero spread that each member started from the same (deterministic) initial conditions.

3. Methodology of Analysis

The spread enough forecast system could capture the forecast uncertainties and provide the probability products. Small spread means low predictability uncertainty, while large spread means high predictability uncertainty. The spread mainly comes from physical perturbations, including initial conditions, boundary conditions and numerical models. The wave forecast is an issue of forcing problem and is slightly relevant to initial conditions. Therefore, boundary conditions of wind fields and numerical model perturbations become primary factors for adjustment. Verification technique were applied to evaluate the ensemble spread (ensemble members and ensemble system).

In Taiwan, the wave climate appears two obvious categories, wave in northeast monsoon and typhoon periods. In typhoon season, southwest monsoon wave is prevailing but smaller compared with typhoon wave and northeast monsoon wave. Therefore, The study covers the periods to determine ensemble members when the higher wave often happened, northeast monsoon period during Jan.-Feb. 2012 and the typhoon period during Jul.-Aug. 2012 (forecast every 12 hours). Ignoring the warming-up period, 118 records of data were available during Jan. and Feb. (each forecast). As to the period from Jul. 30 to Aug., there were 5 typhoons went through this area. To evaluate the ensemble system forecast capacity which need more longer computations, data for the period from Jun. 2014 to Apr. 19, 2016 was collected to conduct analysis (forecast every 6 hours).

The study used the following conventional verification technique qualitatively and quantitatively to assess ensemble members combinations, including Talagrand Rank Histogram, ensemble spread and member equal-likelihood. In addition, it utilized re-

liability diagrams, Brier skill scores (BSS) and Relative Operating Characteristic (ROC) curves to measure the quality of ensemble system. The following states briefly the definitions of various statistical metrics (Wilks, 2006; Li and Hong, 2011). It also can be found in many mathematical books.

(1) Talagrand Rank Histograms (TRH):

In terms of known truth-value, TRH are used to assess bias and dispersion characteristics of the ensemble system. It is constructed from a concept that an ideal ensemble system will correspond to a verification analysis that is equally distributed between any two ordered adjacent ensemble members, including the cases when the analysis will be beyond the ensemble range on either sides of distribution. If diagram of statistical results presents a U-shape distribution, it indicates that the system doesn't spread out sufficiently. The truth-value always falls within intervals of larger value or smaller value. Forecasts could not contain all probabilities of occurrence. In reality the distribution is slightly U-shape. If the diagram presents a A-shape distribution, it indicates that the spread of system is too much and uncertainties possessed by the system are greater than actual conditions. If the diagram is flat, it indicates reasonable dispersion degree of the system.

(2) Ensemble Spread (SPRD):

Evaluation made by TRH regarding the dispersion degree tends to be qualitative information only. Therefore, the use of ensemble spread can offer a quantitative information as a measure of uncertainty defined in Eq. (1). However SPRD only could not be judged the correctness of ensemble because without comparison with truth value. Therefore, RMSE between ensemble mean and truth-value can be thought of as a typical magnitude for forecast errors. If RMSE is equal to SPRD, it indicates that the dispersion degree of ensemble spread is reasonable and capture all the uncertainties. When RMSE is higher than SPRD, it indicates insufficient dispersion. When RMSE is lower than SPRD, it indicates excessive dispersion.

$$SPRD = \sqrt{\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})^2} \quad (1)$$

Where \bar{f} the ensemble mean, f_n the ensemble member, N is the ensemble number.

(3) The Member Equal-Likelihood (MEL):

In this method, a bin is set for each ensemble member, and then the result is checked to see which member's forecast is closest to the observed data. In general, good ensemble forecast system expects that all members should have equal ability to capture the observations. That means the observed data set should uniformly distributed among the ensemble members. Therefore, when the diagram presents flat shape, it indicates the best condition and all of the members have similar ability to capture the observation.

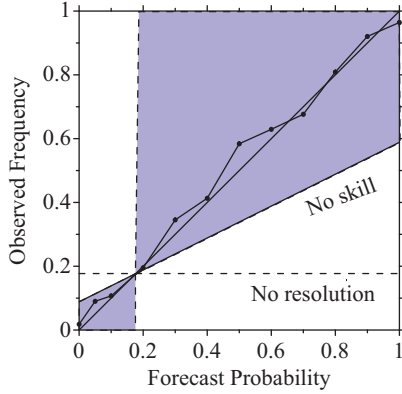


Fig. 3. Reliability diagram. Subsamples in light blue stippled region will contribute positively to forecast skill.

(4) Brier score (BS) and Brier skill score (BSS):

BS mainly estimates the mean squared error between the probability that forecast is higher than the threshold (e.g., $H_s > 2$ m or $U_{10} > 10$ m/s) and the probability that truth-value is higher than the threshold as the definition in Eq. (2), which. The best value for BS is 0 for perfect forecast system ($BS_{perfect}$), and the worst BS is 1. However BS alone without a base line could not get the forecast capacity. The BSS is the conventional skill-score form using the Brier score as the underlying accuracy measure and is defined as the percentage improvement over the reference forecasts, see Eq. (3). Usually the reference forecasts BS_{ref} could be the relevant climatological relative frequencies or other deterministic forecasts.

$BSS > 0$ represents more forecast capacity with respect to climate mean or other deterministic model. $BSS < 0$ represents the absence of forecast capacity or poorer forecast capacity.

$$BS = \frac{1}{n} \sum_{j=1}^n (P_j - o_j)^2 \quad (2)$$

where n is total event number, P_j is the forecast probability, o_j is the observed data (1 for the event happening, 0 for the event not happening).

$$BSS = \frac{BS - BS_{ref}}{BS_{perfect} - BS_{ref}} = 1 - \frac{BS}{BS_{ref}} \quad (3)$$

(5) Reliability diagram:

The reliability diagram plots the observed frequency against the forecast probability for different probability under different thresholds. It mainly checks the consistency between the forecast probability and observed frequency, and could answer how well the forecasts corresponding to its observations.

The diagonal line represents perfect reliability (Fig. 3). The

		Observed		
		Yes	No	
Forecast	Yes	a (hits)	b (false alarm)	a+b
	No	c (misses)	d (correct rejection)	c+d
		a+c	b+d	

Fig. 4. Relationship between counts of forecast/event pairs for the dichotomous non probabilistic verification situation as displayed in a 2×2 contingency table.

closer the curve to the diagonal line, the smaller the probabilistic forecast bias and the higher the reliability. No resolution line refers to climatological relative frequency, which is the ratio of the observation which exceeds thresholds (ex. $H_s > 2$ m) to all the observations. Forecasts not defining event subsets with different relative frequencies of the forecast event would exhibit all points on the dashed no-resolution line. In other words, the closer the probability to climatological relative frequency, the less we need forecasting system, because using climatological mean is enough. No skill line means $BSS = 0$, which is midway between the perfect reliability and no-resolution line, and delimits the stippled region, in which subsamples contribute positively to forecast skill.

(6) Relative Operating Characteristic (ROC):

It mainly focuses on measure the ability of forecast which discriminates between the event and non-event. ROC curve is plotted using false alarm rate (Fi) with hit rate (Hi) against a set of varying probability thresholds, which defined in Eq. (4) and Fig. 4. Generally, the horizontal axis is false alarm rate or error rate while the vertical axis is detection rate or hit rate. When the curve is close to the upper left corner, it indicates good ROC with high hit rate and low false alarm rate.

$$\text{Hit rate} = \text{hit} / (\text{hit} + \text{miss}) = a / (a + c)$$

$$\text{False alarm rate} = \text{false} / (\text{false} + \text{correct rejection}) = b / (b + d) \quad (4)$$

ROC area is an index for measuring whether the system has forecast skill. When the index is 1, it indicates perfect forecast which only takes place in a perfect deterministic forecast. The index 0.5 indicates that there is no forecast skill, namely, observed probabilities and forecast probabilities are equivalent. In general, only ROC area bigger than 0.7 can be referred to as forecast skill that could discriminate between the events and non-events.

III. RESULTS

1. Ensemble Members Determination

The analysis results respectively explain as follows in terms of wind fields and wave fields:

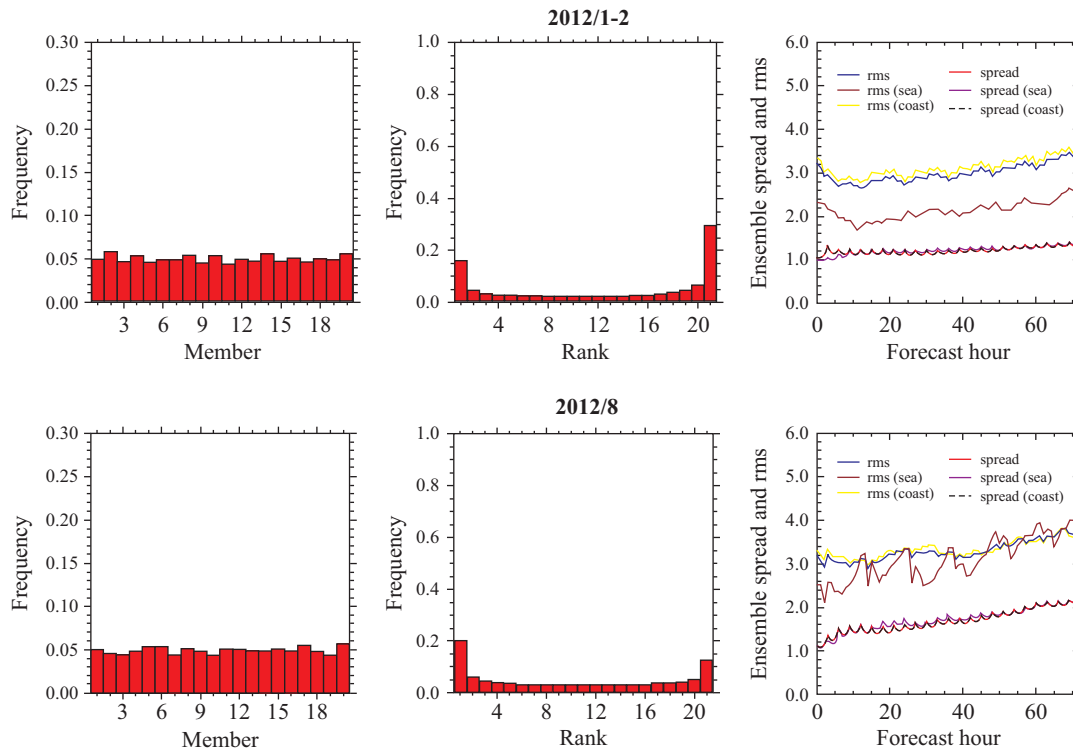


Fig. 5. MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) for U_{10} for Jan.-Feb. (left panels) and Aug. (right panels).

1) Wind Field

Fig. 5 is the MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) of wind speed at buoy stations from Jan.-Feb. (left panels) and Aug. (right panels). Among that, the RMSE and SPRD diagram additionally contain values of stations number 14 and 15 at open seas, and the rest stations near the coast. The MEL diagram shows that the equality among the 20 members is generally the same. No member gets particular high percentage or particular low percentage which indicates the best condition and all of the members have similar ability to capture the observation. As to the TRH diagram, the distribution is slightly U-shape. U-shape stands the ensemble doesn't spread out sufficiently.

The RMSE and SPRD diagrams of Jan. to Feb. presents that the ensemble spread was smaller than model error, but both grew similarly as a function of forecast hour. For the stations at open seas, RMSE was obviously smaller than that at near coast stations, while the SPRD unlike the RMSE are nearly the same. This might be the fact that wind at open sea is less interfered by lands, therefore it has better simulation results.

In the typhoon period, RMSE and SPRD showed the similar trend as value during the period from Jan.-Feb. but became larger. Nevertheless no obvious difference between RMSE and SPRD was found at coast stations. The RMSE at open seas stations became more fluctuation up and down and smaller than the RMSE at coast stations.

2) Wave Field

Wind fields have regional characteristics. Its relations with waves growth usually utilize field experiments to regress empirical formula (Tolman, 2008; Kuznetsova, 2016). Hence, using different forecast wind field under specific areas shall choose proper empirical formula as wind input source term in numerical model to generate wave. We are using two built-in wind input source term (Tolman and Chalikov, WAM4) in NWW3 (Tolman, 2008) for computations to find the suitable formula under WEPS.

(1) Tolman and Chalikov (1996)

Tolman and Chalikov (Tolman, 2008) presented a approach to parameterization of the terms of energy balance equation as input source term and two dissipation constitutes for low frequency and high frequency regimes. In the low frequency regimes, the dissipation can be described using an analogy with the dissipation of wave energy due to oceanic turbulence. For the high frequency range, the dissipation term has been derived from the other source term to result in a consistent source term balance for infinity high frequency, which is proportion to square of wind friction velocity. The formula has been applied in the past studies at Taiwan area. (ITRI, 2012; 2013)

Fig. 6 is the MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) of significant wave height at buoy stations from Jan.-Feb. (left panels) and Aug. (right panels). The RMSE and SPRD diagram additionally contain information from the stations at open sea and the rest near the coast.

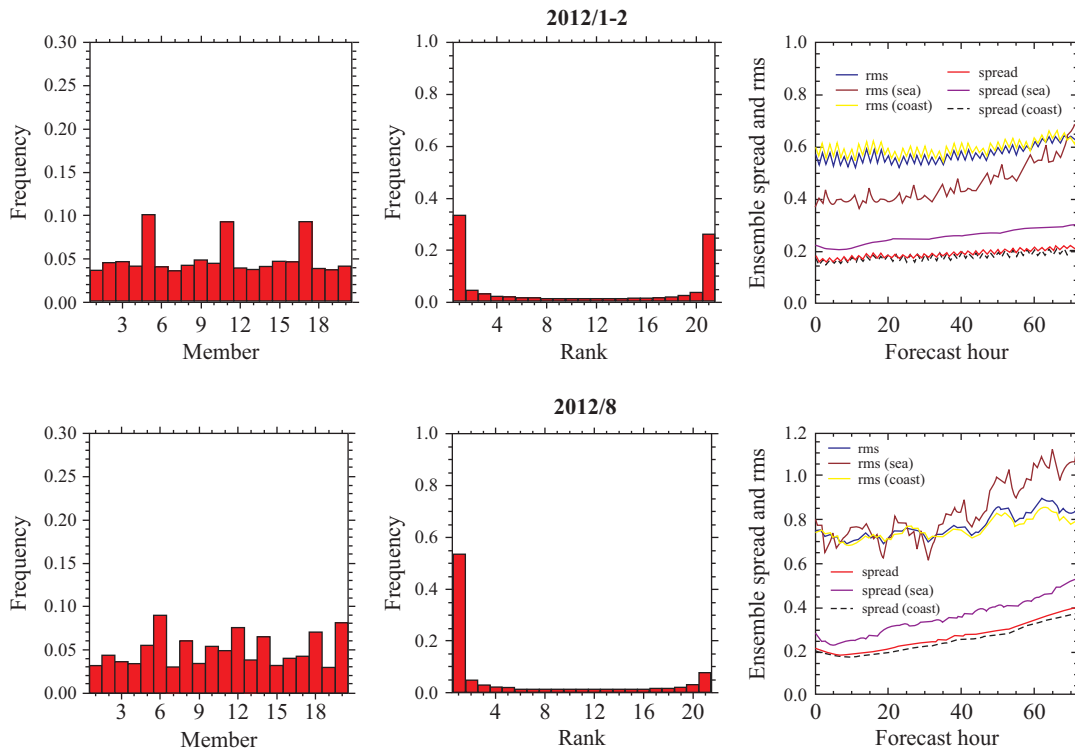


Fig. 6. MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) for significant wave height for Jan.-Feb. (left panels) and Aug. (right panels) (Tolman).

In Jan. and Feb., the MEL showed flat except for the 5th, 11th and 17th ensemble members. The TRH diagram present U-shape and the truth-value falls within the interval with higher or lesser value. This indicated that the average ensemble spread was smaller than the actual model uncertainty. The RMSE and SPRD increased with the forecast hour, and RMSE was larger than SPRD, consistent with TRH. Compare the means of stations at open seas and the rest near coast stations, it showed that RMSE of stations at open seas before 68 forecast hours was obviously smaller and SPRD was also larger. Therefore the difference between RMSE and SPRD became smaller, it indicated better forecast capacity of uncertainty at open seas. This had the same conclusion with wind fields.

The MEL diagram for Aug. doesn't present ensemble members with extremely high frequency like those in Jan. and Feb., yet it is not very flat. As to the TRH diagram, it presented L-shape which indicates excessive higher forecast, and very small part of forecast tended to be lower. The RMSE and SPRD diagram showed the same tendency as in the Jan.-Feb., except that the RMSE at open sea became close to the RMSE at near coast stations before 36 forecast hour and became larger after 36 forecast hour, and the difference between RMSE and SPRD became larger. It indicated that forecast at open seas was better than that at near-shore stations in Jan. and Feb., while both became worse in August.

(2) WAM4 (2013)

The development basis of WAM4 is presented by the adjusted Miles formula (1957) and the correction made by Janssen (1982). It mainly considers the boundary layer effect of atmosphere and the roughness of water surfaces. This roughness is related to the shearing force of waves. Additionally, a linear damping of swells is added, either linear viscous decay or nonlinear turbulence decay depends on the air-sea boundary layer significant Reynolds number. As to the dissipation term, it integrates WAM-3 proposed the white-capping dissipation (modified into WAM4), the saturation-based dissipation (Phillips,1984) and wave-turbulence dissipation.

Rascle and Ardhuin (2013) concluded that the Climate Forecast System Reanalysis (CFSR) and NCEP analysis have systematically higher values than ECMWF analysis of the wind speed. Smaller wind growth parameter should be used for NCEP winds. The boundary condition of CWB WEPS is coming from NCEP winds, therefore the wind growth parameter setting by Rascle and Ardhuin (2013) is used in the study.

MEL diagram in Jan. and Feb. (Fig. 7) showed that the higher frequency of three ensemble members has decreased. TRH diagram indicated the forecast became smaller on Aug. As to the RMSE and SPRD diagram, both of them decrease compared to using Tolman formula except for the RMSE at open sea in Jan. and Feb. The discrepancy between RMSE and SPRD also became slightly smaller than using Tolman at open seas on Aug. This is the desirable feature as it shows

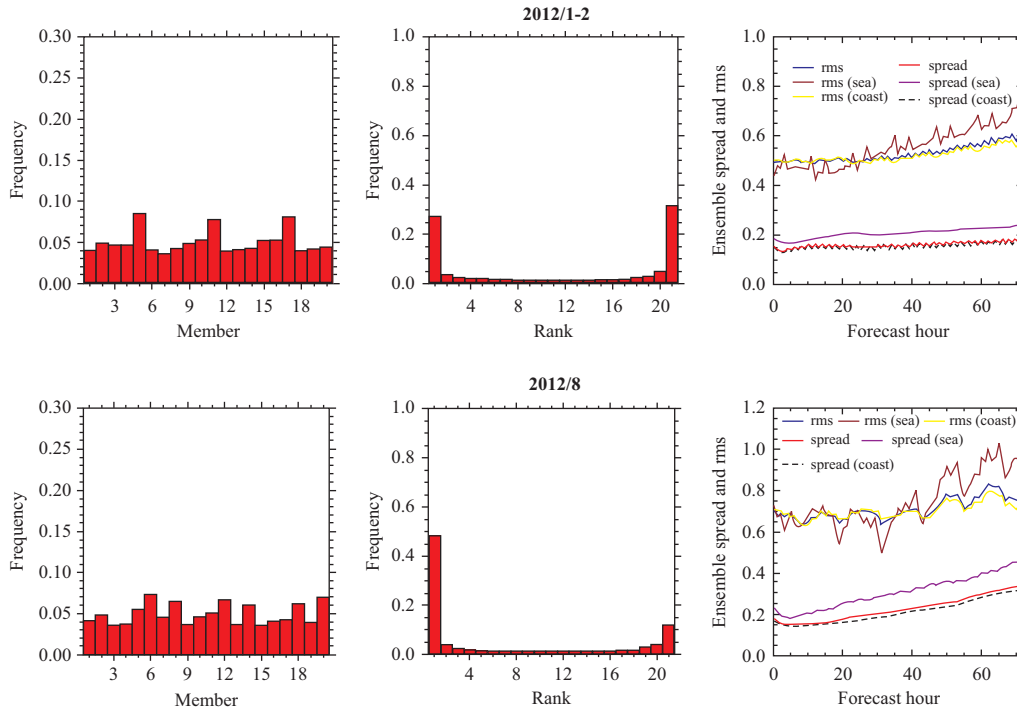


Fig. 7. MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) for significant wave height for Jan.-Feb. (left panels) and Aug. (right panels) (WAM4).

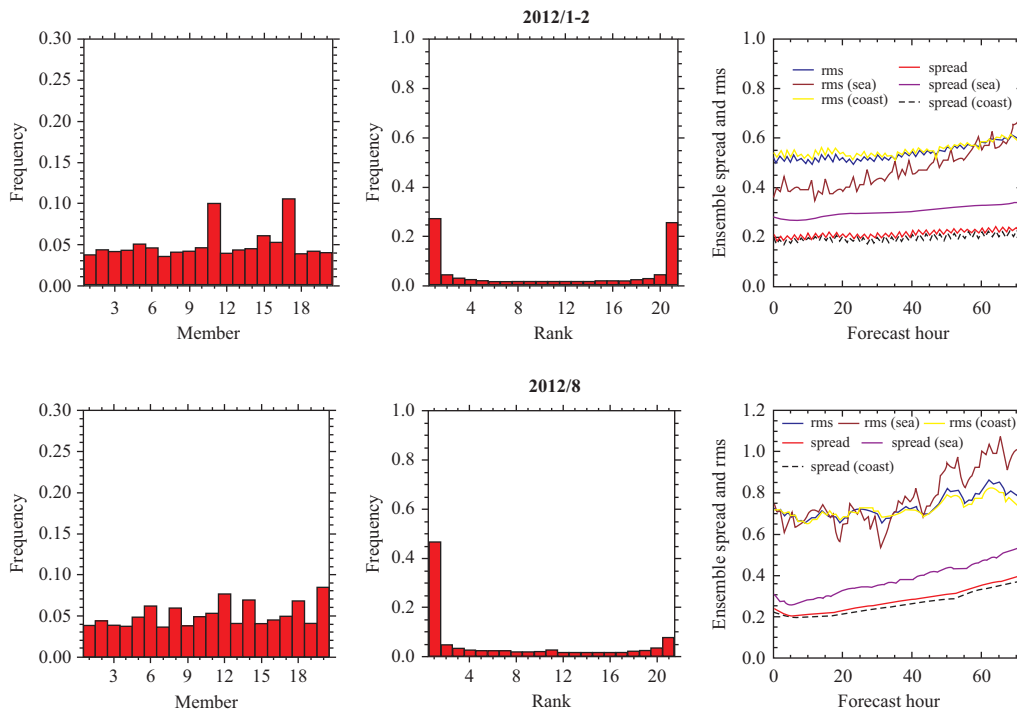


Fig. 8. MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) for significant wave height for Jan.-Feb. (left panels) and Aug. (right panels) (10T10W).

that the model has improved at the open seas during typhoon period. It seems that using different formulas presented their own

advantages during different months. It indicates that using only one formula could not capture whole sea states. However, it's hard to modify the parameters in formulas because

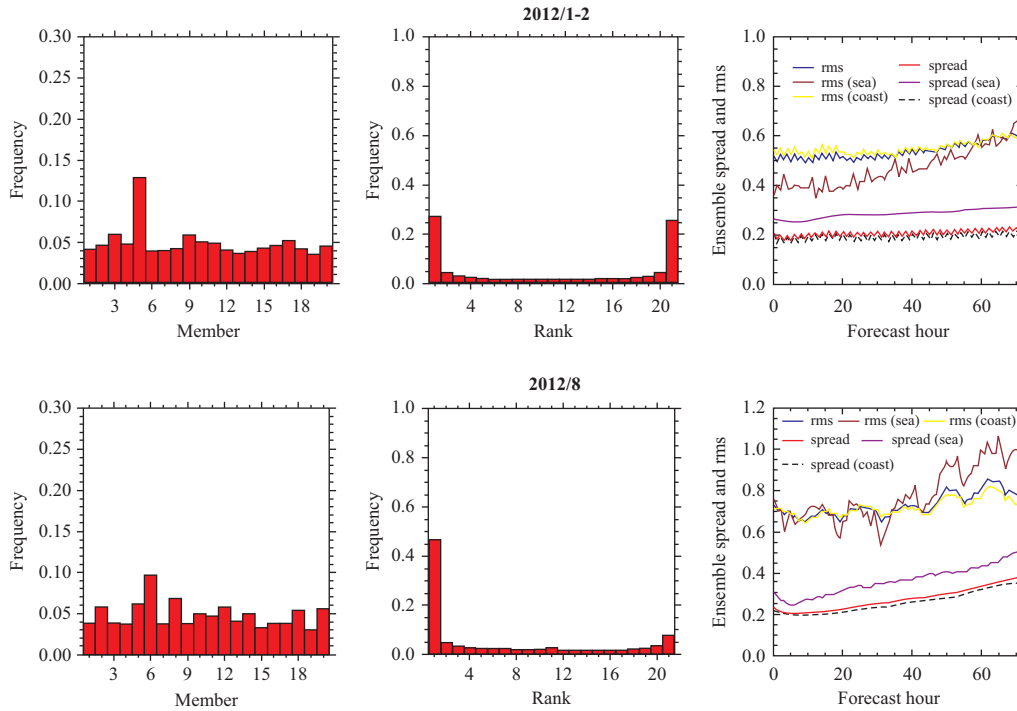


Fig. 9. MEL (upper panels), TRH (center panels), RMSE and SPRD (bottom panels) for significant wave height for Jan.-Feb. (left panels) and Aug. (right panels) (10W10T).

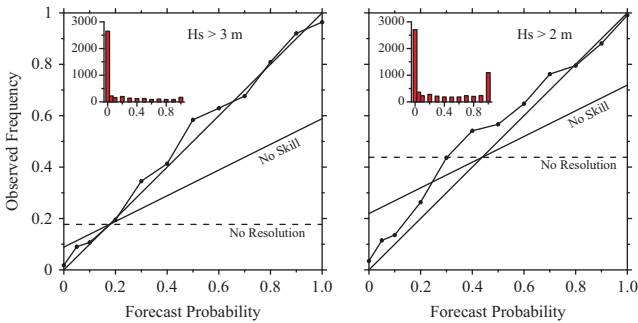


Fig. 10. Day 3 reliability diagram for wave filed (stations 14 & 15). Left $H_s > 3$ m, right $H_s > 2$ m.

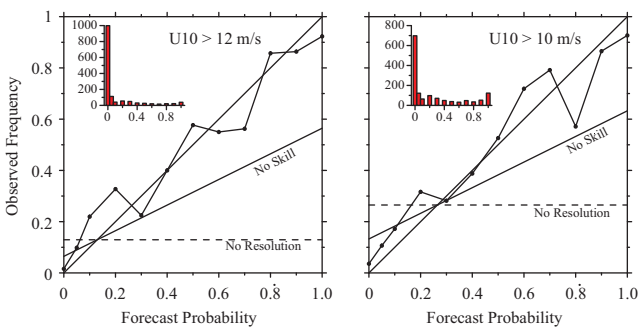


Fig. 11. Day 3 reliability diagram for wind filed (stations 14 & 15). Left $U_{10} > 12$ m/s, right $U_{10} > 10$ m/s.

them were tuned on the basis of the observation and forecast winds at specific locations. Therefore, we proposed com-

bination of two built-in formulas as model perturbation for further analysis. It was selected that the first 10 ensemble members using Tolman formula and the last 10 ensemble members using WAM4 formula. hereafter named as 10T10W. While the first 10 members using WAM4 formula and last 10 members using Tolman formula was named as 10W10T.

(3) 10T10W and 10W10T

Fig. 8 shows that the composition of ensemble members with two built-in formulas obviously increased SPRD and retained advantages of using different formulas for different sea states. It decreased the difference between RMSE and SPRD and improved the forecast capacity of ensemble system for grasping model uncertainties. Computation results as shown in Fig. 9 was similar to the results of 10T10W. Based on the above analysis, we utilized the compositions of ensemble members (10W10T) to establish and implement an operational ensemble system on June, 2014. The forecast skill analysis of ensemble system needs to collect long period of computation to prevent samples not enough problems.

2. Ensemble System Forecast Skill

1) Reliability Diagram

Figs. 10 and 11 are the reliability diagrams for 72-hour forecast of wave height ($H_s > 2$ m, 3 m) and wind speed ($U_{10} > 10$ m/s, 12 m/s) at open seas stations. The forecast probabilities are divided into 12 ranges from 0 to 1.0. This study covers the period from June 2014 to April 19, 2016 for wave and the period from June 2015 to April 19, 2016 for wind speed. The less

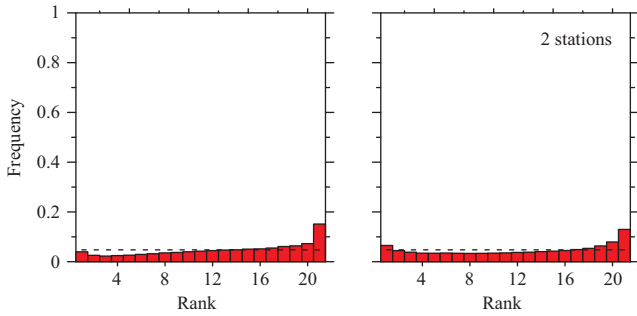


Fig. 12. Talagrand rank histogram for wave (left) and wind (right).

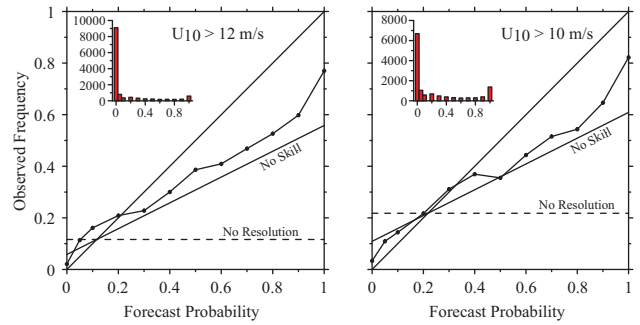


Fig. 14. Day 3 reliability diagram for wind filed (all stations). Left $U_{10} > 12$ m/s, right $U_{10} > 10$ m/s.

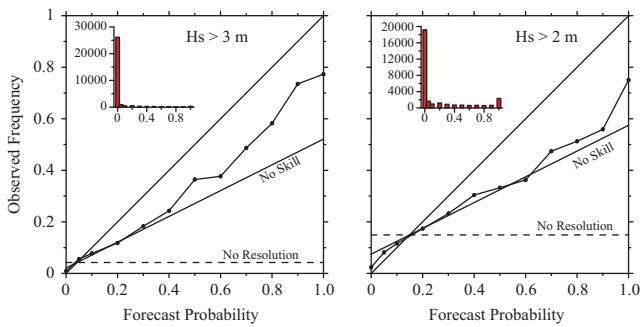


Fig. 13. Day 3 reliability diagram for wave filed (all stations). Left $H_s > 3$ m, right $H_s > 2$ m.

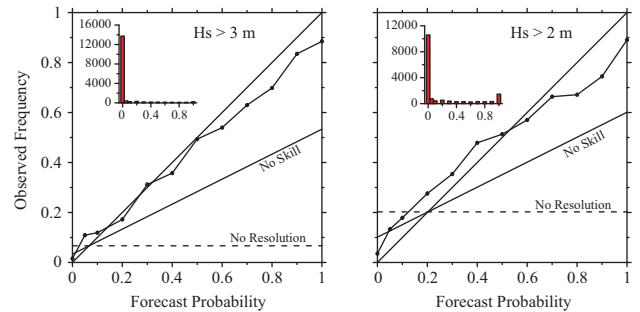


Fig. 15. Day 3 reliability diagram for wave (8 stations). Left $H_s > 3$ m, right $H_s > 2$ m.

wind data is due to the data management errors.

The figures show that the observation probability is approximately equal to forecast probability at open seas when $H_s > 3$ m. When $H_s > 2$ m, the model underestimates slightly at low probability and the subsamples for forecast probability 0.3 and 0.4 failed to contribute positively to the overall forecast skill. It can be also seen from TRH for 3 days forecast (Fig. 12) that the distribution is slightly U-shape and forecast tends to be slightly underestimated at the open seas. As to wind filed, observation probability is approximately equal to forecast probability but fluctuate up and down, which are mainly caused by a lack of sufficient observation data. From the distribution of TRH, it appears a slightly U-shape.

As far as all the observation stations are concerned (Fig. 13 and Fig. 14), the model overestimates and it has higher forecast capacity when $H_s > 3$ m than when $H_s > 2$ m. The diagram of wind speed shows the similar overestimate distribution as wave height.

It seems that many near coast stations have been greatly influenced by terrain and 10 km spatial resolution of model could not be able to resolve the terrain changes and nearshore wave propagations. Review the locations of nearshore stations which are not influenced by terrain seriously such as buoy Xiao Liuqiu (No. 5), Penghu (No. 6), Kinmen (No. 12), station Dongjidao (No. 17) and buoy Qigu (No. 11), and plus two open-seas stations and buoy Matsu (No. 16) a total of 8 stations for further analysis. It shows better result (Fig. 15) which the distribution is closer to diagonal line. It is expected that further improvement is necessary at the other stations including Northeast, East

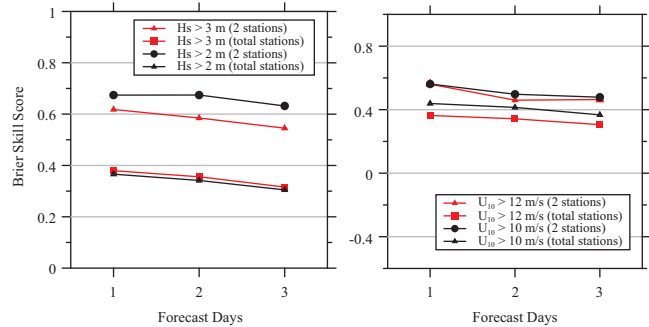


Fig. 16. Brier Skill Score using the climate data as the reference at 2 open sea stations and total stations.

and South Cape ones in the future.

2) Brier Skill Score

In Fig. 16, BSS are plotted for three-days forecast of wave height and wind speed at the open seas stations (2 stations in legend of figure) and total stations separately. The reference is the climate filed.

The skill scores are rather good for wave and wind. The BSS in the three days forecast are higher than 0 which means forecast probability error of ensemble system is less than the error of climate value and has forecast ability in terms of climate field. Higher BSS values at open seas stations for wind and wave shows its better prediction ability. In addition, compared with NCEP

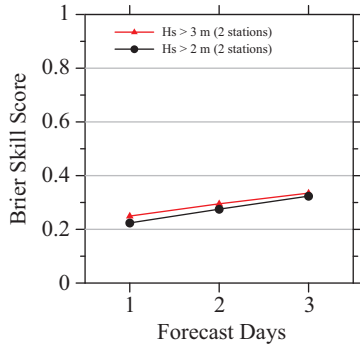


Fig. 17. Brier Skill Score using the ITRI deterministic run as the reference.

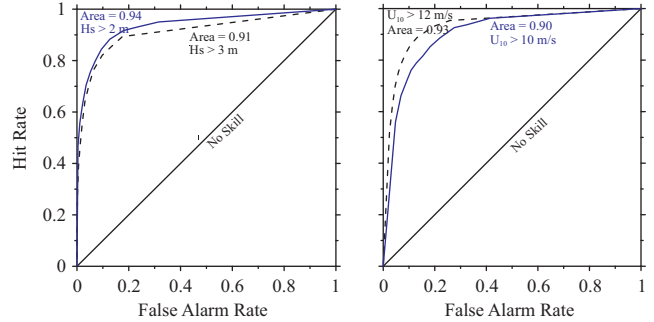


Fig. 19. day 3 ROC curve for wave ($H_s > 3$ m and $H_s > 2$ m) and wind ($U_{10} > 12$ m/s and $U_{10} > 10$ m/s) for all stations.

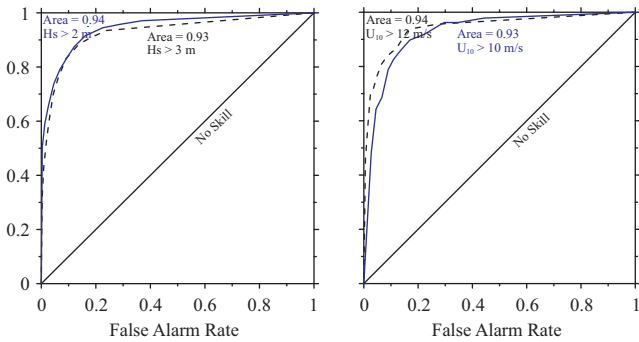


Fig. 18. day 3 ROC curve for wave ($H_s > 3$ m and $H_s > 2$ m) and wind ($U_{10} > 12$ m/s and $U_{10} > 10$ m/s) at open seas stations.

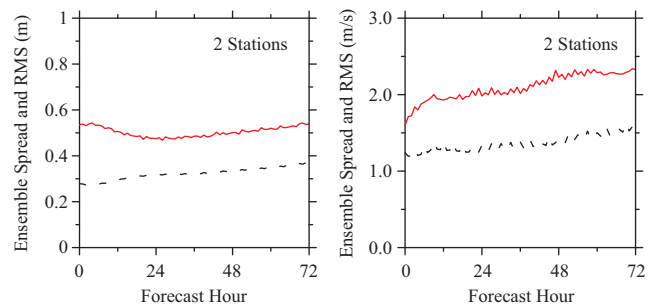


Fig. 20. RMS error and ensemble spread for wave (left) and wind (right), red line is RMSE, the black dash line is SPRD.

ensemble system (Cao, 2007), its BSS for $H_s > 2$ in the three days are less than 0.45, which means the forecast ability of our system at the open seas station is better than NCEP. Fig. 17 is the BSS plots for wave, but the reference is the operational ITRI deterministic forecast model which uses the NCEP GFS wind field, Tolman formula and four layers nesting multi-grids. All BSS increase over the 3 days forecast. This means that the ensemble forecast has higher forecast skill than the deterministic operational forecast.

3) ROC

ROC analysis for different thresholds could be used to perceive the performance of ensemble forecast system and ROC area is a useful summery measure of a forecast skill. ROC curves in Figs. 18 and 19. are close to the upper left corner and ROC areas are over 0.9 for wind and wave which means that the ensemble system has the ability of discriminating between the events and non-events.

4) RMSE & SPRD

Fig. 20 is RMSE and SPRD for wind and wave at open seas stations. SPRD is smaller than RMSE. This means the situation of ensemble spread not enough and matches the analysis of TRH diagram (Fig. 12). Compared with the NCEP ensemble system (Cao, 2009), the distribution of wave height RMSE at beginning of forecast is approximately between 0.32 to 0.55 m,

while in 72 hours forecast is 0.4-0.65 m roughly (estimated from the figure). The distribution of U_{10} RMSE is approximately 1.3-2 m/s at first, and 1.8-2.3 m/s in 72 hours. It does show similar result as our study. The literature also mentioned that ECMWF and NCEP ensemble prediction system both have appeared the tendency of smaller SPRD. They supposed the reasons that the disturbance of wind field may only generate the disturbance of wind wave which do not have impact on the swell in a short time. The solution may be directly adding system disturbance as the initial condition on the low frequency swell which might be able to solve above mentioned problem. However this is just the speculation that further study is necessary.

IV. DISCUSSION AND CONCLUSION

This study sets up an operational wave ensemble forecast system in the range of Taiwan area. Due to the limitation of computer resources, 20 ensemble members with only two layers of multi-grid can be applied. The finest resolution of spatial grid is 10 km and has poor capacity to analysis the wave near the coast which caused larger RMSE at part of near coast stations. Therefore the ensemble members are mainly selected on the basis of two stations at open seas. We proposed a combination of two built-in wind forcing input formulas to form the ensemble members (10 ensembles use WAM4 formula and 10 use Tolman formula) as model perturbation. This combination retains advantages of using different formulas for various wind fields (monsoon and

typhoon period), increases the ensemble spread and decreases the difference between the RMSE and average SPRD. At present, the wave ensemble forecast system at CWB with 20 ensemble members composition of 10W10T had been accomplished and in operation since June 2014.

We uses Reliability diagram, BSS and ROC curves to measure the quality of ensemble forecasts. The result shows that the ensemble system has good forecast capacity and discriminate between events and non-events. It also has better skill than the operational deterministic forecast, and can be comparable with NCEP ensemble system.

Nevertheless, the average spread is not enough which caused the difference between RMSE and SPRD. In order to decrease the difference between the RMSE and SPRD, it could be solved either from decreasing RMSE or increasing SPRD or together. For wind, to decrease RMSE could be used by using the sub-model on land to modify the roughness of boundary layer and coupling wind and wave computation through its roughness at seas. In terms of wave, high resolution of grid can be applied to resolve nearshore simulation to reduce RMSE, and perturbation of the low frequency can be added to increase SPRD.

ACKNOWLEDGEMENTS

This research is funded by the CWB through the Establishment of Wave Ensemble forecast System project. We thank Dr. Hong Ching-Shan, Dr. Chang Hui-Ling and Dr. Terng Chuen-Teyr for support and valuable discussions. We also acknowledge very useful source of buoy information available from the Marine Meteorology Center of CWB.

REFERENCES

Alves J.-H. G. M., P. Wittmann, M. Sestak, J. Schauer, S. Stripling, N. B. Bernier,

- J. McLean, Y. Chao, A. Chawla, H. Tolman, G. Nelson and S. Klotz (2013). The NCEP-FNMOC combined wave ensemble product: expanding benefits of interagency probabilistic forecasts to the oceanic environment. American Meteorological Society, 1893-1905.
- Cao, D., H. S. Chen and H. L. Tolman (2007). Verification of ocean wave ensemble forecast at NCEP. The 10th international workshop on wave hindcasting and forecasting & coastal hazards symposium, Turtle Bay, Oahu, paper G1.
- Cao, D., Tolman, H. L., Chen, H. S., Chawla, A., & Gerald, V. M. (2009). Performance of the ocean wave ensemble forecast system at NCEP. The 11th International Workshop on Wave Hindcasting & Forecasting and 2nd Coastal Hazards Symposium.
- Chen, H. S. (2006). Ensemble prediction of ocean waves at NCEP. Proceedings of 28th Ocean Engineering Conference in Taiwan, NSYSU. http://www.ecmwf.int/products/forecasts/guide/Wave_EPSgrams.html. <http://dx.doi.org/10.1016/j.ocemod.2012.12.001>.
- Industrial Technology Research Institute (2012). Establishment of a wave ensemble forecasting system (2/4), Central Weather Bureau 2012 Project Report. (in Chinese)
- Industrial Technology Research Institute (2013). Establishment of a wave ensemble forecasting system (3/4), Central Weather Bureau 2013 Project Report. (in Chinese)
- Kuznetsova, A., G. Baydakov, V. Papko, A. Kandaurov, M. Vdovin, D. Sergeev and Y. Troitskaya (2016). Adjusting of wind input source term in WAVEWATCH III model for the middle-sized water body on the basis of the field experiment. *Advances in Meteorology*, vol. 2016, Article ID 8539127, 13 pages, doi:10.1155/2016/8539127.
- Li, J. S. and J. S. Hong (2011). The study of regional ensemble forecast: physical perturbations. *Atmospheric Sciences*, 39, 95-115. (in Chinese)
- Massel, S. R. (2013). *Ocean surface waves, their physics and prediction*, 2nd Ed. Advanced series on ocean engineering (Vol. 36).
- Raschle, N. and F. Ardhuin (2013). A global wave parameter database for geophysical applications. Part 2: model validation with improved source term parameterization. *Ocean modeling*.
- Tolman, H. L. (2008). User manual and system documentation of WAVEWATCH III version 3.14. NOAA/NWS/NCEP MMAB Tech. Note 268.
- Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences*, 2nd Ed. (Vol. 91). Academic press.