Ending Storm Version of the 7-day Weighted Analog Intensity Prediction Technique for Western North Pacific Tropical Cyclones

Hsiao-Chung Tsai

Department of Water Resources and Environmental Engineering, Tamkang University, New Taipei City, Taiwan

Russell L. Elsberry

Trauma, Health, Hazards Center, University of Colorado Colorado Springs, Colorado Springs, Colorado, and Department of Meteorology, Naval Postgraduate School, Monterey, California

(Manuscript received 12 October 2017, in final form 22 November 2017)

ABSTRACT

The weighted analog intensity prediction technique for western North Pacific (WAIP) tropical cyclones (TCs) was the first guidance product for 7-day intensity forecasts, which is skillful in the sense that the 7-day errors are about the same as the 5-day errors. Independent tests of this WAIP version revealed an increasingly large intensity overforecast bias as the forecast interval was extended from 5 to 7 days, which was associated with ''ending storms'' due to landfall, extratropical transition, or to delayed development. Thus, the 7-day WAIP has been modified to separately forecast ending and nonending storms within the 7-day forecast interval. The additional ending storm constraint in the selection of the 10 best historical analogs is that the intensity at the last matching point with the target TC track cannot exceed 50 kt (where 1 kt = 0.51 m s^{-1}). A separate intensity bias correction calculated for the ending storm training set reduces the mean biases to near-zero values and thereby improves the mean absolute errors in the 5–7-day forecast interval for the independent set. A separate calibration of the intensity spreads for the training set to ensure that 68% of the verifying intensities will be within the 12-h WAIP intensity spread values results in smaller spreads (or higher confidence) for ending storms in the 5–7-day forecast intervals. Thus, some extra effort by the forecasters to identify ending storm events within 7 days will allow improved intensity and intensity spread forecast guidance.

1. Introduction

Tsai and Elsberry (2014) developed a 5-day weighted analog intensity technique for western North Pacific tropical cyclones (TCs) called WAIP. Analog techniques are developed by selecting from the best-track files those cases (analogs) that have similar characteristics (in this case, tracks and initial intensities) as the target storm. After ranking the 10 best analog intensity evolutions according to how closely the track and initial intensities matched those of the target storm, Tsai and Elsberry (2014) gave a higher weight for those analogs that better matched the 3–5-day tracks, because they hypothesized that the track was a primary determinant of the intensity changes in that time interval. Tsai and Elsberry (2016) demonstrated that this simple analog technique, which can be calculated in a few minutes on a desktop computer, was more accurate than the regional numerical model intensity guidance in the 3–5-day forecast intervals. Tsai and Elsberry (2015) then developed a 7-day WAIP, but independent testing revealed an increasingly large intensity overforecast bias in the 5–7-day interval that they attributed to ''ending storms'' due to landfall, extratropical transition, or to delayed development.

The objective of this study is to develop an ''ending storm'' version of the 7-day WAIP that is similar to an ending storm version of the weighted analog intensity technique for Atlantic TCs (WAIA; Tsai and Elsberry 2017). The reason an ending storm WAIP version is so important for the western North Pacific is that of the almost 5000 forecasts in the 2000–15 Joint Typhoon Warning Center (JTWC) training set (Fig. 1a), approximately 3300 of these forecasts have an ending in the 3–7-day interval (Fig. 1c). Thus, only ~1600 of the remaining forecasts do not involve an ending storm situation

Corresponding author: R. L. Elsberry, elsberrylr@comcast.net

DOI: 10.1175/WAF-D-17-0151.1

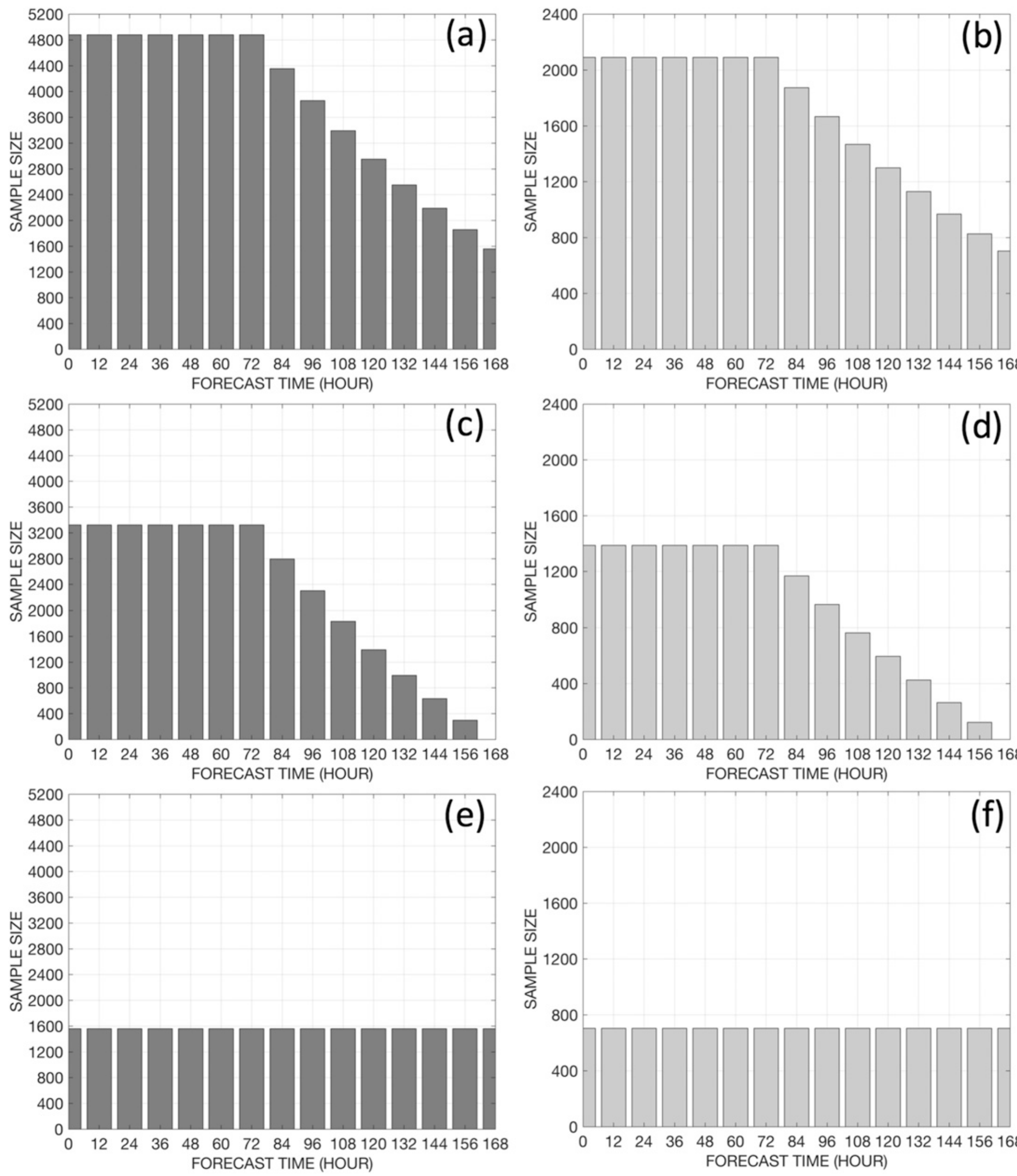

FIG. 1. Histograms of the numbers of western North Pacific TC events during 2000–15 in the JTWC best-track files utilized for WAIP 70% training sets for (a) all sample, (c) ending storms, and (e) nonending storms. (b),(d),(f) As in (a), (c), and (e), but for the 30% independent sets.

(Fig. 1e). One of the reasons for the large number of ending storm situations in the western North Pacific is the fraction of TCs that form in the South China Sea and have early landfalls, or form in the southwest monsoon trough and within 7 days make landfall along the East Asia coast or adjacent islands.

The training and independent testing sets in the original 7-day WAIP (Tsai and Elsberry 2015) were taken from the 2000–09 and 2010–14 seasons, respectively. Here, following the WAIA development, a new version of the 7-day WAIP (designated as all sample) will also be developed and tested, with 70% of the TCs during 2000–15 being randomly selected to be the training set (Fig. 1a). Separate WAIP versions will be developed and tested for those ending storms in Fig. 1c and for the nonending storms in Fig. 1e. For the 30% independent set, ~2100 events are in the all sample subset (Fig. 1b), ~1400 events are in the ending storm subsample (Fig. 1d), and ~700 events are in the nonending storm subsample (Fig. 1f). The improved intensity and intensity spread predictions from the separate ending storm and nonending storm WAIP versions will be demonstrated by comparisons with the all sample WAIP predictions.

For this development and testing of the ending storm version of WAIP, the ending time is known from the JTWC best-track file. In an operational application, the forecaster will need to provide this ending time. After the forecaster has made the 7-day TC track forecast, the forecaster will evaluate if the TC will end at any time along this 7-day track forecast as a result of a landfall, an extratropical transition, or that the TC is not likely to exceed an intensity of 35 kt (1 kt = 0.51 m s^{-1}). The extratropical transition time might be evaluated from products based on the Hart (2003) phase diagram from the forecast fields of the global model that has the track forecast most closely matching the official track forecast. Finally, the likelihood of a nondevelopment or delayed development to 35 kt along that 7-day track forecast might be evaluated based on the trends in the satellite imagery and the previous 6- and 12-h intensity forecasts.

Since the methodology for developing and testing the original 7-day WAIP is given in Tsai and Elsberry (2015), only a brief description of the ending storm modification will be given in section 2. The performance of the separate ending storm and nonending storm WAIP versions will be evaluated in section 3. A summary and discussion are provided in section 4.

2. Modification for an ending storm situation

As indicated above, the JTWC best tracks during the 2000–15 seasons are used for the new all sample, ending storm (hereafter designated as track < 168 h), and nonending storm (hereafter designated as track = 168 h) versions. Each best track is divided into several 3–7-day tracks with different starting dates. Ten candidate historical analogs within ±30 days are selected based on the average track differences that must be within 15° latitude/longitude and on the initial intensity difference that must be ≤20 kt. Separate rankings are made of the track differences and initial intensity differences between the candidate analogs and the target storm, and then the final rankings of the 10-best candidate analogs are weighted by 0.8 for the track difference ranking and 0.2 for the initial intensity ranking.

As in the ending storm version of the WAIA (Tsai and Elsberry 2017), the additional constraint in the analog selection is that the intensity at the last matching point with the end of the target TC track cannot exceed 50 kt. Note that 95% of the western North Pacific TCs have ending storm intensities ≤ 50 kt. That is, if the target TC track is shorter than 7 days, then an analog will not be selected if the intensity of that analog is 50 kt or higher at the forecast time interval corresponding to the ending time of the target TC track. To ensure good matching of the track differences, the analog track time period must overlap by at least 80% the total target track time period.

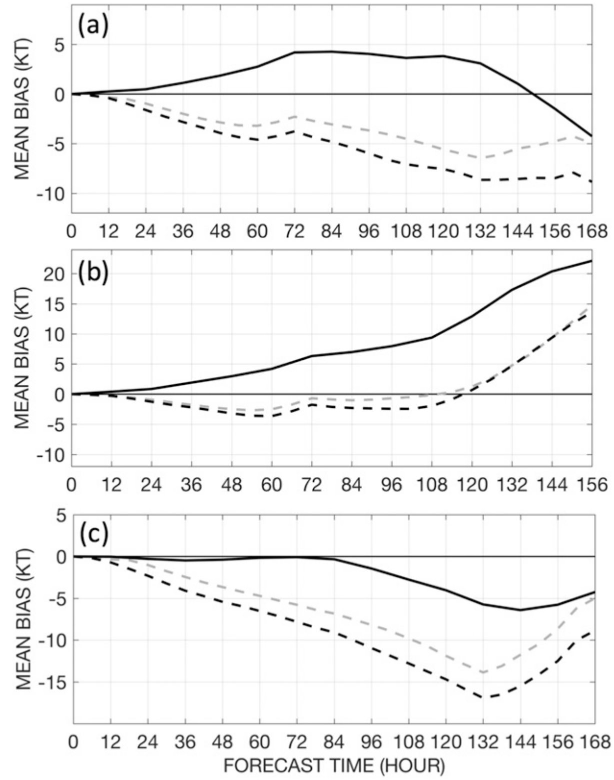

FIG. 2. Mean error biases (kt) as a function of forecast interval for the training set (gray dashed line) and the independent set prior to (dashed line) and after (solid line) application of the all sample bias correction for the (a) all sample, (b) ending storm subsample but without the ending storm constraint applied, and (c) nonending storm subsample.

3. Performance evaluation for WAIP with ending storm conditions

a. Bias corrections

The second step in the WAIP development with each of the three training sets in Figs. 1a, 1c, and 1e is a bias correction each 12 h, which is designed to reduce the mean biases in these training sets to zero. If the characteristics of the independent set TCs are similar to those of the training sets, application of these individual bias corrections to the corresponding independent sets in Figs. 1b, 1d, and 1f are expected to reduce the mean biases at each 12-h forecast interval to near-zero values.

The mean biases for the all sample training set increase steadily at a rate of about −1 kt day^{-1} (Fig. 2a).

The mean biases for the all sample independent set are similar with a slightly larger decrease to −9 kt at 168 h (Fig. 2a). Consequently, the bias correction developed for the training set and applied to this independent set does not result in near-zero mean biases with +4 kt between 72 and 120 h and −4 kt at 168 h (Fig. 2a).

The mean biases for the subsample of the ending storms training set without the ending storm constraint being applied (Fig. 2b) are very different from the all sample training set in Fig. 2a. While the mean biases do have a −1 kt day^{-1} rate during the first 2 days, between 72 and 120 h the mean biases are near zero, and then the mean biases increase to +15 kt at 156 h (Fig. 2b). The independent subsample of ending storms without the ending storm constraint being applied has a similar distribution of mean biases with a value of +17 kt at 162 h (Fig. 2b). These large positive biases at ~168 h occur because the analogs have been chosen with the all sample selection criteria rather than the ending storm selection criterion, and thus the WAIP predictions would have large overforecasts of those storms that have actually ended. Applying the all sample bias correction to this independent subset of ending storms results in WAIP predictions with positive mean biases throughout the entire forecast interval (Fig. 2b), which is the motivation for developing a separate bias correction for these ending storms.

For just the subsample of nonending storms, the mean biases for the training set (Fig. 2c) are again somewhat different from the all sample training set in Fig. 2a. For this subsample, the mean biases become progressively more negative to −14 kt at 132 h, and then become smaller to only −5 kt at 156 h. The independent set of the subsample of nonending storms has a similar distribution of mean biases with a minimum value of −17 kt at 132 h and −9 kt at 168 h (Fig. 2c). Applying the all sample bias correction to the independent subsample of nonending storms is quite successful in reducing the mean biases to near-zero values from 12 to 84 h, but the WAIP predictions would have negative biases of −6 kt at 144 h (Fig. 2c).

In summary, these nonzero mean biases for the independent sets in Figs. 2b and 2c when the all sample bias corrections have been applied demonstrate that separate bias corrections are required for the ending storm and nonending storm training sets in Figs. 2b and 2c. For the subsample of ending storm cases that now has the ending storm constraint applied (Fig. 3a), the mean biases for the training set and the independent set are rather similar in that both have increasingly large biases (−14.5 and −17 kt) at around 108 h and then have smaller biases (−3 and −6 kt) at 156 h. Consequently, applying the separate bias correction results in mean biases for the independent set that are relatively small, with the largest mean bias of −3 kt at 108 h (Fig. 3a). For the subsample of nonending storm cases (Fig. 3b), the mean biases for the training set and the independent set are quite similar in that both have increasingly large biases (−14 and −17 kt) at 132 h and then have smaller biases (−5 and −9 kt) at 162 h. Thus, the separate bias corrections for the nonending cases are also successful in that the largest independent set bias is −5 kt at 144 h (Fig. 3b).

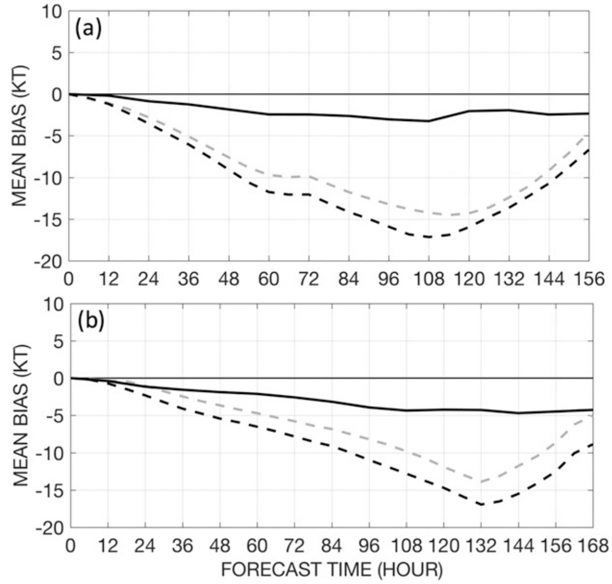

FIG. 3. Mean error biases (kt) as in Fig. 2b and 2c for the (a) ending storm subsample with the ending storm constraint applied and (b) nonending storm subsample, except separate bias corrections are derived and applied to each of these subsamples rather than applying the all sample bias corrections to these subsamples as in Figs. 2b and 2c.

b. MAEs and correlation coefficients

Impacts in terms of the independent set mean absolute errors (MAEs) of the separate bias corrections for the ending storms are shown in Fig. 4a. Although these MAEs increase rapidly to 15 kt at 36 h, the MAEs are then about 18 kt from 60 to 96 h before rapidly decreasing to only 10 kt at 156 h. Thus, the ending storm constraint of only selecting analogs with intensities < 50 kt is particularly effective if the TC is known to end in the 4–7-day forecast interval. By contrast, this independent set of ending storms has much larger MAEs if the all sample bias correction is utilized, because rather than leveling off after 84 h the MAEs continue to increase to ~23 kt at 144 and 156 h (Fig. 4a).

Another metric for assessing the impact of the separate bias correction for the ending storm cases is the

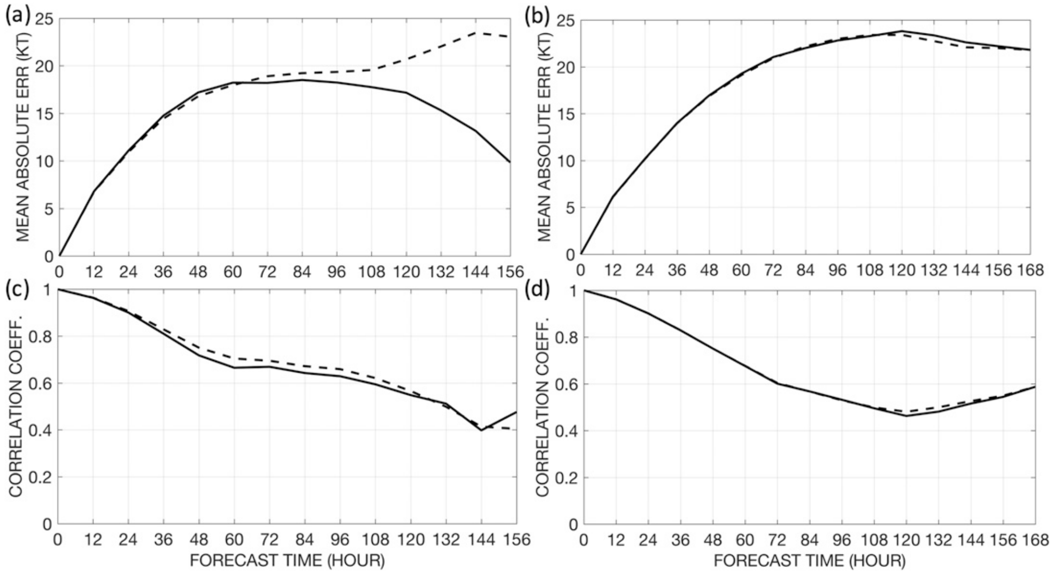

FIG. 4. MAEs (kt) for independent sets of the (a) ending storm subsample and (b) nonending storm subsample when the separate bias corrections are applied (solid line) or the all sample bias corrections are applied (dashed line). (c),(d) As in (a) and (b), but for correlation coefficients of the WAIP intensity predictions with the verifying intensities.

correlation coefficient of the WAIP intensity predictions with the verifying intensities. For the independent set of ending storm WAIP predictions with the separate bias correction, the correlation coefficients rapidly decrease to 0.66 by 60 h and then decrease more slowly to 0.6 around 120 h before decreasing to 0.45 at 156 h (Fig. 4c). One contribution to these smaller correlations beyond 120 h (even though the MAEs in Fig. 4a are smaller) is that the independent sample sizes are smaller at these longer forecast intervals and even a small bias of −5 kt (Fig. 4a) could lead to smaller correlation coefficients. Although there is no significant difference, the independent ending storm subsample that includes the all sample bias correction has correlation coefficients that decrease more rapidly after 108 h to a value of 0.40 at 144 and 156 h.

The corresponding nonending storm independent set MAEs are provided in Fig. 4b. While these MAEs are similar to the ending storm MAEs (Fig. 4a) through 36 h, these nonending storm MAEs continue to steadily increase to 23 kt at 120 h and then slightly decrease to 22 kt at 168 h. Because the all sample bias correction was so successful for the independent subsample of nonending storms (Fig. 3b), the corresponding MAEs are just as accurate as with the separate bias correction (Fig. 4b). Nevertheless, this favorable result that the 7-day intensity forecast accuracy is as good as at 5 days is similar to the original 7-day WAIP (Tsai and Elsberry 2015). Both sets of correlation coefficients with the verifying intensities of these nonending storms (Fig. 4d) continue to steadily decrease to a value of 0.5 at 120 h, and thus are not as high as for the ending storm correlation coefficients in Fig. 4c. However, the nonending storm correlation coefficients then increase to 0.6 at 168 h, which is better than for the ending storms (0.45 at 156 h).

c. PoD and mean intensity spreads

A necessary feature of the original 7-day WAIP (Tsai and Elsberry 2015) was a calibration of the ''raw intensity spread'' among the 10 analog intensities each 12 h to ensure an intensity spread that would include 68% of the verifying intensities. Separate calibrations of the intensity spreads of the ending storm and nonending storm training sets were carried out and applied to the intensity spreads of the independent sets. An example of the application of the calibration for the ending storm independent set is given in Fig. 5a. Note that the probability of detection (PoD) values prior to the calibration are overdispersive (raw intensity spreads of the 10 analog intensities are too large) during the first 30 h (Fig. 5c), and then are very underdispersive out to 156 h with a PoD of 40% rather than the desired value of 68%. The success of the calibration for this ending storm independent set is evident from the PoDs that are within 5% of the 68% value over most of the 7-day forecast interval (Fig. 5a).

The impact of the calibration on the intensity spreads for the ending storm independent set is displayed in

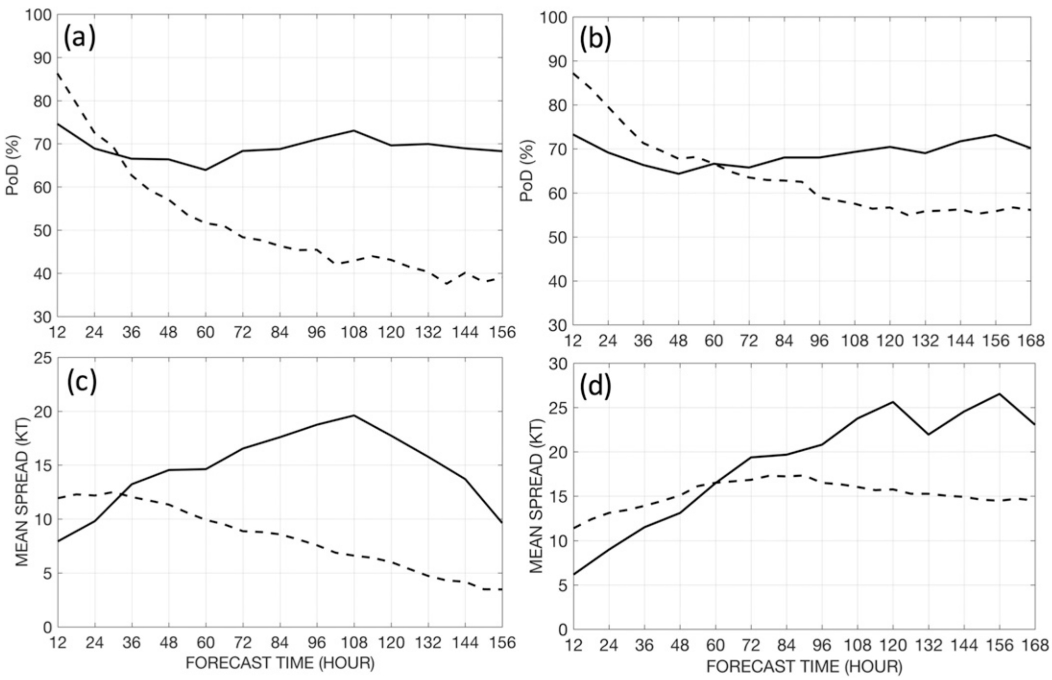

FIG. 5. (a),(b) As in Figs. 4a and 4b, but for PoD before (dashed line) or after (solid line) applying a calibration of the raw intensity spreads to the independent sets. (c),(d) As in (a) and (b), but for the mean intensity spreads (kt) and before (dashed line) or after (solid line) applying a calibration of the raw intensity spreads to these independent sets.

Fig. 5c. Because progressively smaller numbers of storms are extending to 156 h before ending (Fig. 1d), the mean intensity spreads become smaller with increasing forecast interval (and thus are underdispersive). After the calibration of the intensity spreads in Fig. 5a, the mean intensity spreads are made smaller from 12 to 24 h, but need to be much larger to 108 h to ensure 68% of the verifying WAIP intensity forecasts would be within the intensity spreads. However, the calibrated intensity spread at 156 h is only ±10 kt (Fig. 5c) because all of these storms must by definition end prior to 168 h and, thus, are more homogeneous in intensities so an intensity spread of only 10 kt still encompasses 68% of the WAIP ending storm intensities.

The PoDs for the nonending storm independent set are also overdispersive from 12 to 48 h, but are not nearly so underdispersive at larger forecast intervals (Fig. 5b) because the "raw" intensity spreads are at least 15 kt from 48 to 168 h (Fig. 5d). The calibration of the intensity spreads for this independent set is again quite successful as most of the PoDs are within 5% of the desired 68% (Fig. 5b). After this calibration, the mean intensity spreads are smaller from 12 to 54 h, but have to be larger at longer forecast intervals to ensure that the WAIP intensity spreads encompass the desired 68% of the verifying intensities (Fig. 5d).

4. Summary and discussion

The WAIP technique has been modified to separately forecast ending and nonending storms within the 7-day forecast interval. The additional ending storm constraint on the selection of the 10-best historical analogs is simply that the intensity of the last matching point with the end of the target TC track cannot exceed 50 kt. A separate bias correction for the ending storms results in improved MAEs in the 5–7-day forecast interval, and a calibration of the WAIP intensity spread values results in smaller spreads (or higher confidence). It is noted that this statistical approach, which is designed to minimize errors over a large sample, will not forecast extremely rapid intensification as a typhoon is moving northwestward, or an extreme decay rate following recurvature, and these will be the WAIP intensity forecast errors that will fall outside the WAIP intensity spread.

These new ending storm WAIP forecasts will be based on the 10-best historical analogs that match the official track forecast and the current storm intensity, and they can be produced in a few minutes on a desktop computer. Thus, the forecaster could easily vary the landfall time to take into account likely along-track forecast errors, vary the extratropical transition time, or change the nondevelopment versus the development decision to see the impacts on the WAIP forecast. This study demonstrates that this extra effort on the part of the forecaster will result in improved 7-day intensity and intensity spread forecast guidance.

Acknowledgments. RLE has been supported by the Office of Naval Research Marine Meteorology section. HCT is supported by the Taiwan Ministry of Science and Technology (105-2625-M-032-003). Mrs. Penny Jones provided excellent assistance in the manuscript preparation.

REFERENCES

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Wea. Rev.*, **131**, 585–616, https://doi.org/10.1175/1520-0493(2003)131<0585:ACPSDF>2.0.CO;2.

Tsai, H.-C., and R. L. Elsberry, 2014: Applications of situation-dependent intensity and intensity spread predictions based on a weighted analog technique. *Asia-Pac. J. Atmos. Sci.*, **50**, 507–518, https://doi.org/10.1007/s13143-014-0040-7.

——, and ——, 2015: Seven-day intensity and intensity spread predictions for western North Pacific tropical cyclones. *Asia-Pac. J. Atmos. Sci.*, **51**, 331–342, https://doi.org/10.1007/s13143-015-0082-5.

——, and ——, 2016: Skill of western North Pacific tropical cyclone intensity forecast guidance relative to weighted-analog technique. *Asia-Pac. J. Atmos. Sci.*, **52**, 281–290, https://doi.org/10.1007/s13143-016-0001-4.

——, and ——, 2017: Seven-day intensity and intensity spread predictions for Atlantic tropical cyclones. *Wea. Forecasting*, **32**, 141–147, https://doi.org/10.1175/WAF-D-16-0165.1.