September 1, 2011 – November 30, 2013

Improvements to the SHIPS Rapid Intensification Index

Principle Investigator: John Kaplan (NOAA/AOML)

Co-PIs: Christopher M. Rozoff (CIMSS), Charles R. Sampson (NRL), James P. Kossin (NOAA/NCDC), Christopher S. Velden (CIMSS), Mark DeMaria (NOAA/NESDIS)

Collaborator: Mark DeMaria (NESDIS)

NHC Points of Contact: Stacy R. Stewart, Eric S. Blake, Christopher W. Landsea

Computer scientist support: Paul A. Leighton NOAA/AOML

1. Project overview

Predicting episodes of tropical cyclone (TC) rapid intensification (RI) remains one of the highest operational forecasting priorities of the National Hurricane Center (NHC). Thus, in recent years, a statistically based rapid intensification index (RII) that employs predictors from the SHIPS model to estimate the probability of RI over the succeeding 24-h has been developed utilizing linear discriminant analysis for both the Atlantic and eastern North Pacific basins (Kaplan et al 2010). Although the SHIPS-RII is currently used as an operational forecasting tool by the NHC, its utility has been somewhat restricted since the original version was developed exclusively for a single (24-h) lead-time and its skill has tended to be rather limited particularly for the Atlantic basin.

Thus in an effort to improve the overall forecasting usefulness of the current operational SHIPS-RII, a number of model enhancements have been developed as part of our current Joint Hurricane Testbed (JHT) project. First, new consensus-based versions of the RII that employ both the current SHIPS-discriminant RII as well as newly developed Bayesian and logistic regression versions (Rozoff and Kossin 2011) have been derived for the current 24-h operational forecast lead-time as well as the added lead times of 12-h, 36-h, and 48-h. Secondly, new versions of the rapid intensity aid that provides deterministic intensity forecasts utilizing both existing operational intensity forecast models and the probabilistic RII have been developed utilizing the new multi lead-time consensus RII. Lastly, microwave imagery-based versions of the RII that have been shown to be capable of providing a more accurate measure of the overall inner-core tropical cyclone structure have been derived. A summary of the above efforts to improve the utility of the operational SHIPS-RII is provided below.

2. Project Accomplishments

a. Multi-lead time consensus-based RI model

As noted above, the original version of the operational SHIPS-RII was developed exclusively for the 24-h lead-time. Thus, in an effort to provide added RI guidance to NHC forecasters for the critical watch/warning period that has recently been extended to 48 h, additional versions of the SHIPS-RII were developed for the 12-h, 36-h and 48-h lead times

utilizing the 1995-2012 SHIPS developmental database. Following the methodology that was used previously to develop the current operational 24-h lead-time version, new versions of the SHIPS-RII were developed for RI thresholds that represented approximately the 95th percentile of over-water intensity change of all subtropical and tropical cyclones that formed during the developmental data period. More specifically, new versions of the SHIPS-RII were developed for the 20-kt RI threshold at the 12-h lead-time, 45-kt at the 36-h lead-time, and 55-kt at the 48-h lead-time. Thus, versions of the SHIPS RII were developed for a total of seven RI thresholds since versions of that model had been previously developed for the 24-h RI lead-time for four other RI thresholds (i.e., 25 kt, 30 kt, 35 kt, and 40 kt). It should be noted that the new multi-lead time versions of the SHIPS-RII employ the same predictors that are used in the new experimental 24-h lead-time version of the SHIPS-RII that had been previously developed by Kaplan et al. (2011) and was subsequently accepted for operational implementation by the NHC in April 2012. For a complete list of the predictors used in the new multi-lead time versions of the SHIPS-RII, consult Table 1 of the Appendix.

New versions of the Bayesian and logistic regression RI models were also developed for all four lead-times (12 h, 24 h, 36 h, and 48 h) following the methodology described in Rozoff and Kossin (2011) since that study indicated that by simply averaging the forecasted RI probabilities from each of the three individual RI models (i.e., the SHIPS-RII, Bayesian, and logistic regression), a new consensus RI model could be constructed that provided improved skill over any of three individual RI models. In the study by Rozoff and Kossin (2011), the same predictors were employed to develop the Bayesian and logistic regression versions of the RI models for all RI thresholds in a given basin even though the predictors utilized to derive those two models differed slightly between the two basins. In the current study, however, the RI predictors utilized to derive the new Bayesian and logistic multi-lead time versions of the RII were optimized so that the set of predictors that were selected for each RI model maximized the cross-validated BSS for each forecast lead-time and basin. Thus, the predictors used in the newly developed Bayesian and logistic regression models differ both as a function of lead-time and basin. More specifically, for the Atlantic basin, up to 11 (8) of the RI predictors listed in Table 2 (Table 3) of the Appendix were employed to derive the Bayesian (logistic regression) versions of the RI models. For the E. Pacific basin, up to 11 (7) of the RI predictors listed in Table 4 (Table 5) of the Appendix were used to derive the Bayesian (logistic regression) versions of those RI models.

The new multi-lead time consensus RII was run in near real-time commencing at the beginning of August 2013 at CIRA on the Colorado State University campus with the output made available to forecasters at the NHC via a web site to demonstrate the real-time capability of these models. It is important to note that all of the data required to run the aforementioned RII models in real-time are currently available from the operational WCOSS computer located at NCEP with the exception of the total precipitable water (TPW) data that is being processed at CIRA. However, accessing the TPW data via NCEP computers has been denoted as a high priority by the NHC (particularly since it is being used as part of the version of the SHIPS-RII that was accepted for operational implementation in April 2012) and it is anticipated that the capability to access TPW data via NCEP computer will be fully functional prior to the beginning of the 2014 Hurricane Season.

Although as noted above the multi-lead time RII models were run in real-time during the latter part of the 2013 Hurricane Season, the period during which the models were run was limited and both the Atlantic and E. Pacific seasons were unusually quiet. Thus, the new models

were re-run for the period 2004-2013 using the real-time GFS forecast fields and initial storm data that was archived for this 10-year period in an effort to better evaluate the model's forecasting capabilities. It is important to note that prior to performing the 2004-2013 re-run RI forecasts, cross-validated versions of each of the three versions of the RI models (Bayesian, SHIPS-RII, and logistic regression) were re-derived for all seven RI thresholds for each of the ten individual forecast years (e.g., 2004, 2005, 2006, …, 2013) after having first excluded all of the cases for each of those same forecast years. Thus, the re-run results that were obtained and that are presented below are considered independent.

Figure 1 shows the skill of each version of the RI models as well as their consensus as a function of lead-time and RI threshold for both the Atlantic and E. Pacific basin for a homogeneous sample of all of the subtropical and tropical forecasts cases during the period 2004-2013 for which forecasts were available for all three of the individual RI models. Following the methodology described in Kaplan et al. (2010), skill was computed relative to the climatological probability of RI for a given RI threshold and lead-time when evaluated based upon the Brier Skill Score (BSS) for all systems that remained over-water during the specified forecast lead-time.

It can be seen that each of the three individual RI models generally exhibits skill relative to climatology by this measure at all lead-times and RI thresholds (save for the logistic version at the 40-kt threshold at the 24-h lead time) and that their skill is noticeably larger in the E. Pacific basin than the Atlantic basin. Overall, the performance of the individual models tended to be similar at the shorter (i.e., 12-h and 24-h) lead times with the logistic regression version performing somewhat better at the longer (i.e., 36-h and 48-h) lead times. Nevertheless, the consensus of the three RI models proved to be more skillful than any of the individual RI models at each lead-time in each basin consistent with the dependent results that Rozoff and Kossin (2011) obtained for a dependent sample of forecasts at the 24-h lead-time. Interestingly, the skill of the consensus RI for the 55-kt RI threshold at the 48-h lead-time exceeded the skill that was obtained for that same model for the 20-kt RI threshold at the 12-h lead-time in both basins suggesting that RI may be more predictable at longer lead-times.

Figure 2 shows reliability diagrams (Wilks 2011) depicting the forecasted versus observed RI probabilities for both the Atlantic and E. Pacific basins that were constructed utilizing the consensus RI model forecasts since Fig. 1 indicated that these were the most skillful. When constructing the reliability diagrams, the consensus RI probability forecasts were divided into five equally spaced bins and the average predicted versus forecasted RI probability as well as the frequency of the RI forecasts were computed for each of the five bins. It is worth noting that the number of bins used to construct the reliability diagrams was determined so as to maximize the total number of bins employed while minimizing the number of bins that contained a single forecast.

It can be seen that the Atlantic consensus RI probability forecasts were fairly well calibrated except for those made for the 35-kt and 40-kt RI thresholds with the latter exhibiting substantial under confidence (the forecasted RI probabilities were appreciably lower than the observed RI probabilities) and the former showing a tendency for under confidence for lower RI probabilities and over confidence (the forecasted RI probabilities were higher than the observed) for moderate RI probability values. It is worth noting that while the ensemble RI probabilities did not exceed 50% very often in the Atlantic basin, RI generally did occur when such probabilities were forecast. Also, while the individual versions of the Atlantic RI models exhibited the capability to forecast higher RI probabilities than did the consensus (not shown), these forecasts

tended not to be nearly as well calibrated as those of the consensus with some models exhibiting a fairly marked tendency to forecast appreciably higher probabilities than were observed.

Figure 2 shows that the E. Pacific basin consensus RI forecasts were fairly well calibrated up to probabilities of around 40-50% after which they exhibited under confidence particularly for the 45-kt and 55-kt RI thresholds. Nevertheless, the E. Pacific consensus RI model demonstrated the capability of correctly forecasting much higher RI probabilities than did the Atlantic version.

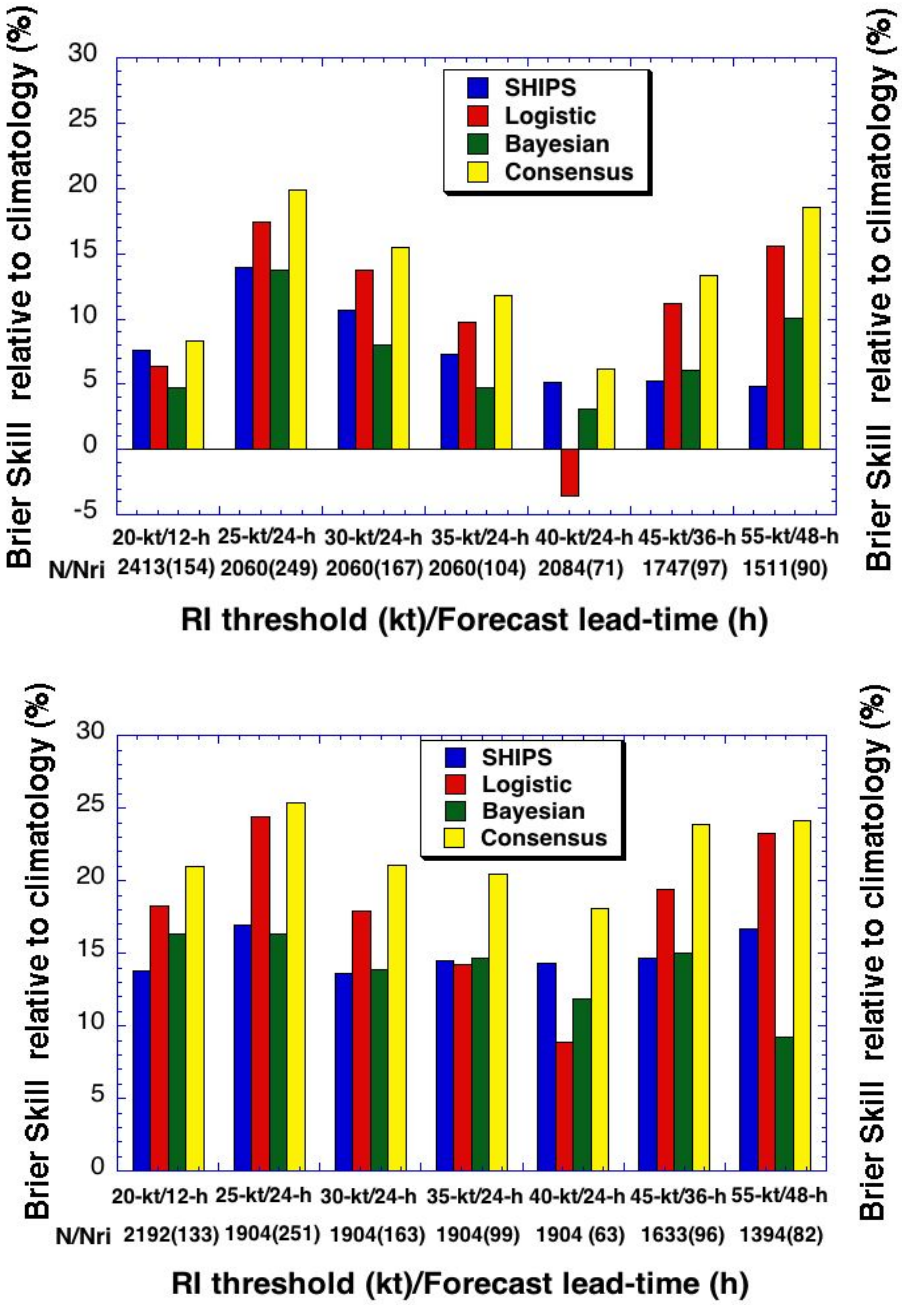

Fig. 1. Skill of the Atlantic basin (top) and E. Pacific basin (bottom) 2004-2013 RI re-run RI forecasts. Skill is shown for the SHIPS, Bayesian, logistic regression (logistic), and consensus versions of the RI model. The total number of forecasts (N) and number of RI cases (Nri) are shown for each of the seven RI thresholds.

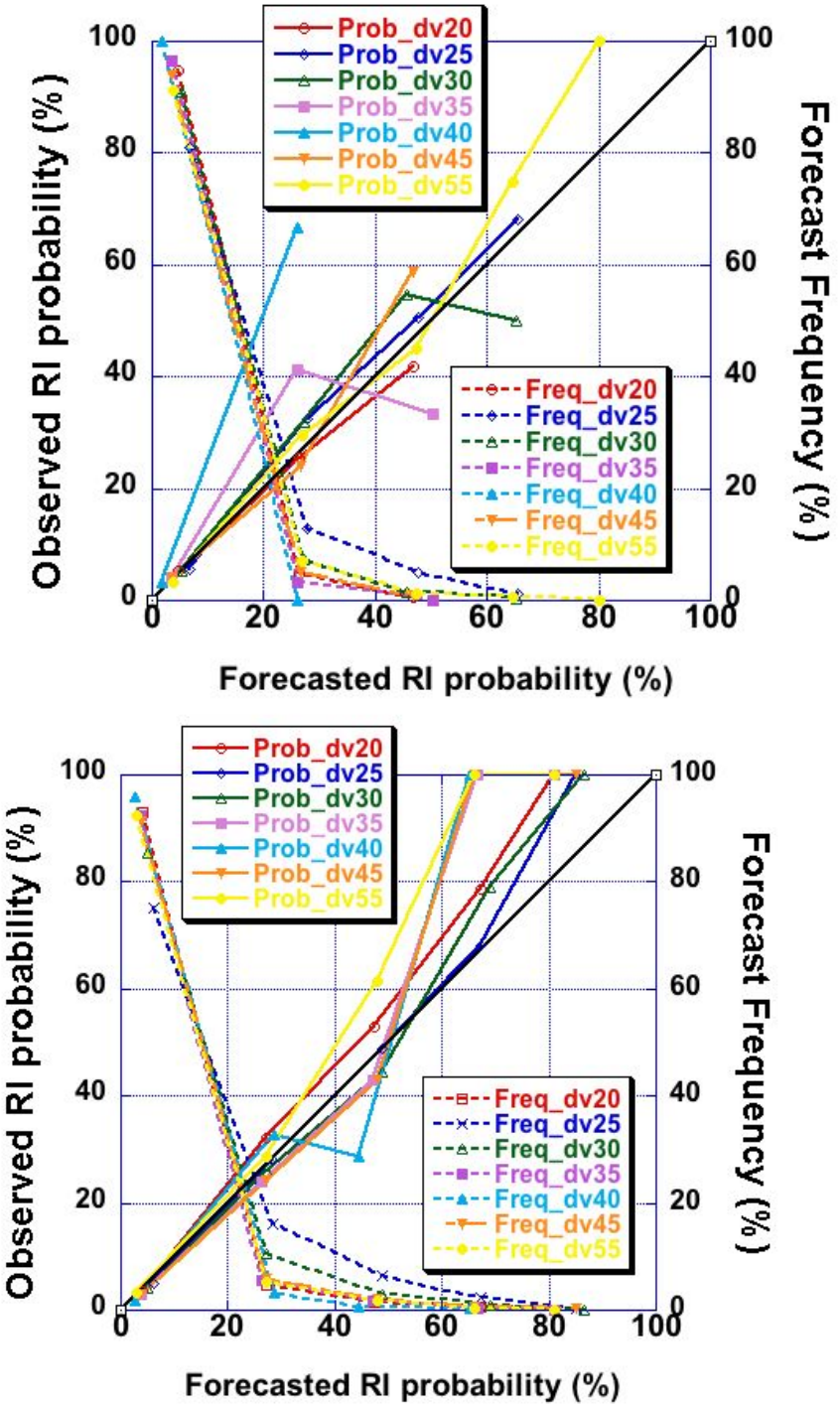

Fig. 2. Reliability diagrams for the Atlantic (top) and. E. Pacific (bottom) basins depicting the forecasted versus observed RI probability for the consensus version of the RI for each of the seven RI thresholds. The solid lines show the reliability for the 20-kt (red), 25-kt (dark blue), 30-kt (green), 35-kt (magenta), 40-kt (light blue), 45-kt (orange) and 55-kt (yellow) RI thresholds. The dashed lines depict the frequency of the forecasts as a function of the forecasted RI probability for each of the RI thresholds. The black diagonal line depicts forecasts with perfect reliability.

Figure 3 shows an example of the performance of the consensus RI model during an episode of extreme RI that occurred during Hurricane Wilma (2005). For simplicity, only the forecasted RI probabilities for the 20-kt, 30-kt, 45-kt and 55-kt RI thresholds are shown. It can be seen that the RI probabilities for the 55-kt RI reached ~80% when Wilma was still a weak tropical storm and had not yet started to undergo RI. Interestingly, the probabilities for the 48-h lead-time exceeded those predicted for any of the other RI thresholds again suggesting that RI may be more predictable at the longer lead times as was discussed previously.

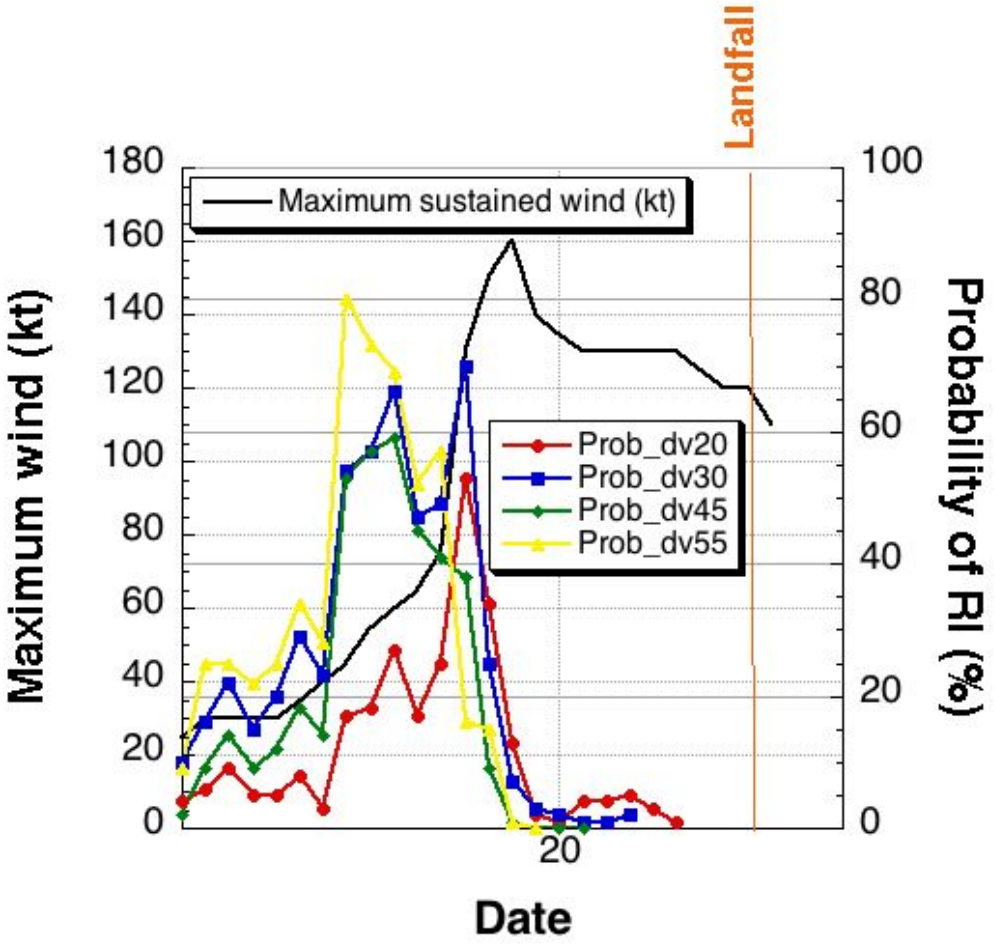

Fig. 3. Re-run consensus forecasted RI probabilities for the 20-kt (red), 30-kt (dark blue), 45-kt (green), and 55-kt (yellow) RI thresholds from the time of Hurricane Wilma's (2005) initial formation until the time of landfall along the Mexican Yucatan Peninsula. The sold black line shows the NHC best track 1-miniute maximum sustained wind estimates and the vertical orange line depicts the time of landfall along the Mexican coastline.

b. Deterministic RI Aid

A deterministic rapid intensification aid (RAPID) that provides deterministic intensity forecasts utilizing both existing operational intensity and probabilistic SHIPS-RII model forecasts has been previously developed (Sampson et al. 2011). The current operational SHIPS-RII produces RI probabilities for intensification rates (25 kt, 30 kt, 35 kt, and 40 kt) through a 24-h period. We use SHIPS-RII thresholds (currently 40%) to trigger whether the deterministic RAPID aids are produced. We implemented a temporary solution in operations and the resultant output from the IBM was then transmitted back to NHC and ingested into the ATCF in real-time for the 2011 and 2012 seasons. For 2013, we implemented a more permanent solution where the RAPID aids are computed within SHIPS.

We recently ran new versions of the RAPID aid that now have the capability to produce 12-h, 24-h, 36-h and 48-h forecasts by employing the new multi-lead time consensus RI models, for cases from the 2008-2013 Atlantic and eastern North Pacific seasons. The code implemented to perform the recomputed guidance is written to be an integral part of the new consensus RI model and it could be transitioned to operations at the same time as the newly developed consensus RI guidance suite. After the RAPID aids were recomputed utilizing the consensus RI guidance for the extended RI periods, we developed an intensity consensus to include all the new RAPID aids with the current members of IVCN (DSHP, LGEM, GHMI, and HWFI). The new intensity consensus was developed prior to the evaluation, so no tuning was done to get better performance. The new intensity consensus (IVRI) can be defined with the existing intensity consensus code (Sampson et al. 2008) so that the work to implement this is minimized.

Results from the 2008-2013 consensus RI reruns are shown in Fig. 4. Figure 4 (top panel) is an evaluation of the Atlantic cases for which at least one ensemble RII aid probability was greater than 40%. The number of cases is quite limited (an average of only 10 per year at 24 h), so these events are quite rare. Still, we can see in Fig. 4 that inclusion of the RAPID aids in the consensus does indeed reduce the mean forecast error by approximately 5% and reduce the biases by several kt. The improvement is larger in the eastern North Pacific (bottom panel), and the mean errors of IVRI are significantly lower than those of IVCN. In both basins the top performer is generally the official forecast although the IVRI forecasts represented somewhat of an improvement over the official forecast at the longer lead times (particularly for the intensity bias) in the eastern North Pacific basin. Nevertheless, the overall results shown in Fig. 4 demonstrate that the forecasters are better at RI forecasting than any of the individual guidance.

We recently experimented with a GPCE for intensity (Goerss and Sampson 2013) and have investigated cases where RAPID aids were run and their potential impact on the intensity GPCE. So far, we find that the GPCE 67th percentile is generally shifted up appropriately in RI cases (Fig. 5). Although this is not part of our original proposal, we have continued to evaluate this topic since the authors feel the RAPID aids should improve the intensity GPCE.

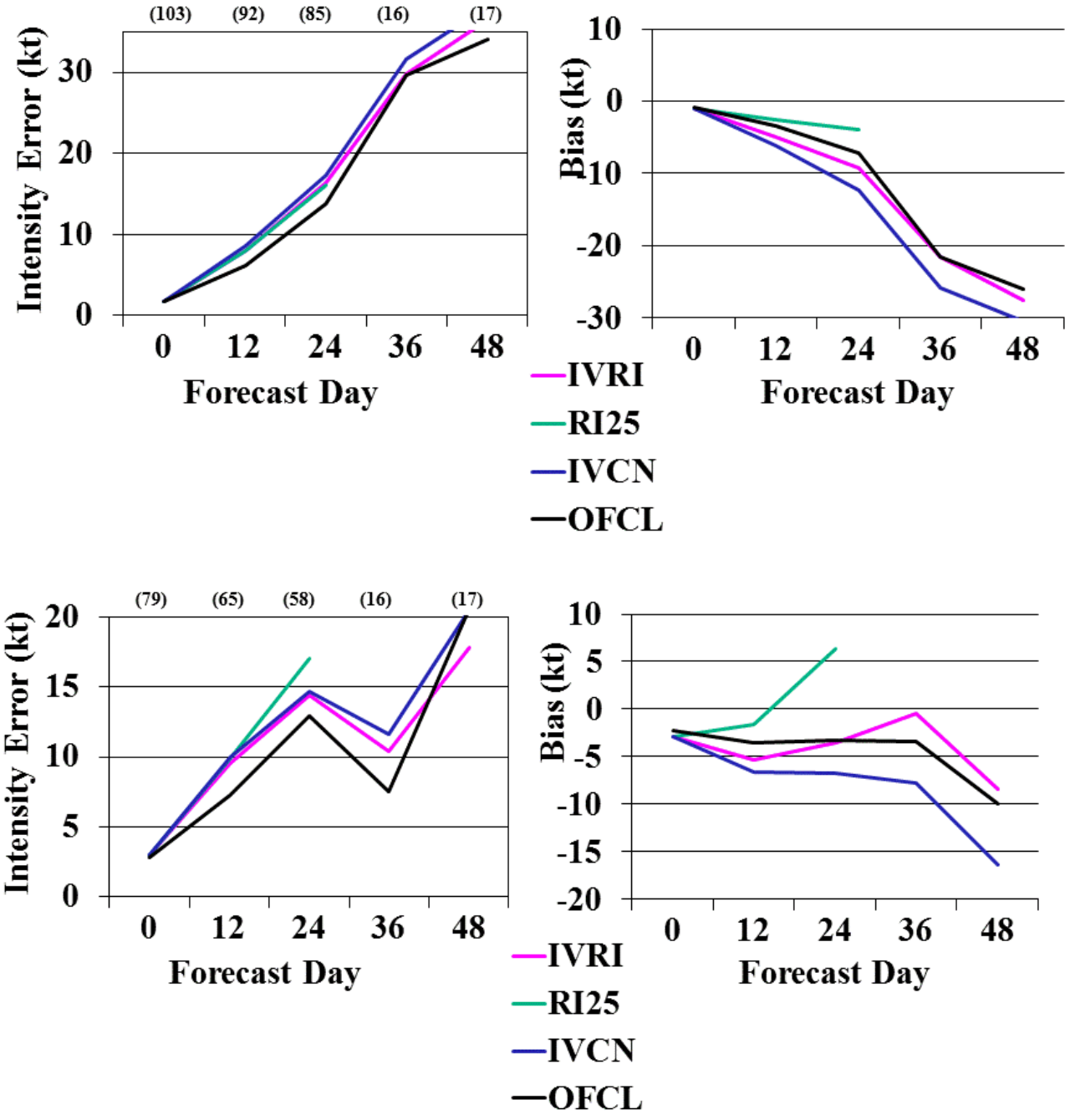

Fig. 4. Results from 2008-2013 Atlantic (top) and eastern North Pacific (bottom) re-runs of the operational NHC intensity consensus (IVCN) and the IVCN with the addition of RAPID aids computed from the consensus RI model (IVRI). Evaluation shows lower mean errors and reduced bias when RAPID aids are included. RI25 is a deterministic RI aid computed for a 25-kt RI event. OFCL is the official forecast and the number of cases is shown in parentheses.

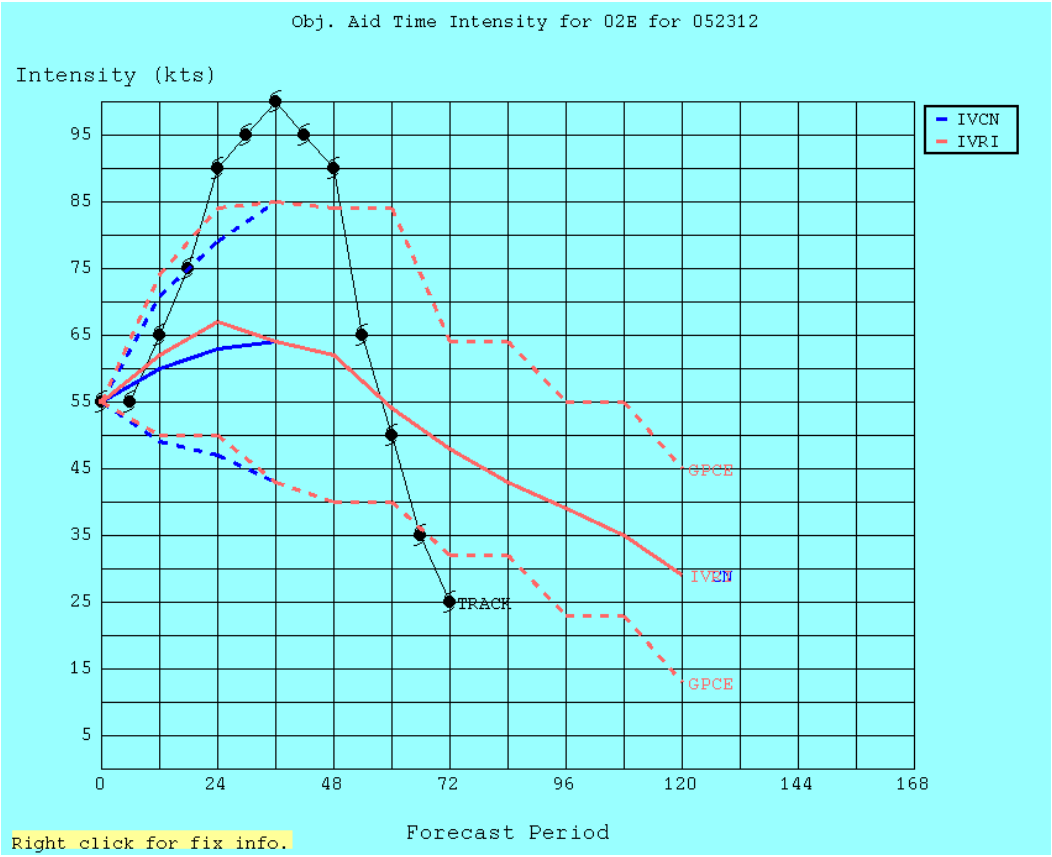

Fig. 5. Intensity vs. forecast time graph for EP022012 at 2012052312. The solid (dashed) blue line(s) represent the IVCN forecast (GPCE 67th percentile) and the solid (dashed) red line(s) represent the same for IVRI (IVCN members + the RAPID aids).

c. Microwave-based RI model

A final angle to this JHT project involves using information about a TC's structure to improve the probabilistic prediction of RI. To delineate aspects of TC structure relevant to RI, passive microwave imagery (MI) from low-earth orbiting satellites has been collected and analyzed for many Atlantic and Eastern Pacific TCs. In contrast to infrared and visible satellite imagery, MI can depict the distribution of precipitation underneath thick, overlying clouds often associated with TCs. In the current project, physically based statistical features are created from the MI to represent the distribution and intensity of precipitation, including warm rain and ice hydrometeors in the inner core of developing and mature TCs. These predictors are specifically developed to enhance existing probabilistic RI models.

In this JHT project, the logistic regression RI model described in Rozoff and Kossin (2011) has been updated to include MI-based predictors. It is important to note that, unlike in section 2a, this MI-enhanced model continues to use the same SHIPS developmental dataset predictors described in Rozoff and Kossin. The additional MI-based predictors are derived from the Special Sensor Microwave Imager (SSMI), Special Sensor Microwave Imager/Sounder (SSMIS), Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), and the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and include the 19.4-, 37.0-, and 85.5-GHz brightness temperatures, or similar frequencies depending on the specific sensor. The developmental dataset used here spans the years 1998-2012 over the Atlantic and Eastern Pacific Ocean basins.

Multiple MI channels are chosen for this study because of each channel's somewhat unique perspective on storm structure. Lower frequency channels, including the 19.4- and 37.0-GHz channels, measure the emission of microwave radiation from liquid hydrometeors. These channels often depict the precipitation structure of TCs at lower levels of the troposphere. The higher frequency 37.0-GHz and 85.5-GHz channels tend to capture ice scattering signals from precipitation at higher levels of the troposphere. Figure 6 shows an example of these different perspectives for Atlantic Hurricane Danielle (2004) at a stage in which it was undergoing RI. While the 19-GHz channel has lower spatial resolution compared to the higher frequency channels, the brightness temperatures in Fig. 6a show Danielle possesses a low-level ring structure around the storm center at 1527 UTC 14 August, a structure often identified in TC RI events (Kieper and Jiang 2012). This ring structure is also seen in the intermediate-resolution 37-GHz vertical polarization brightness temperatures (Fig. 6b). On the other hand, the polarization-corrected temperatures (PCT) at 37 GHz and 85 GHz (Figs. 6c,d) show ice scattering associated with deeper convective clouds. Collectively, these different channels show the storm structure is azimuthally asymmetric aloft despite a higher degree of azimuthal symmetry at lower levels of the storm.

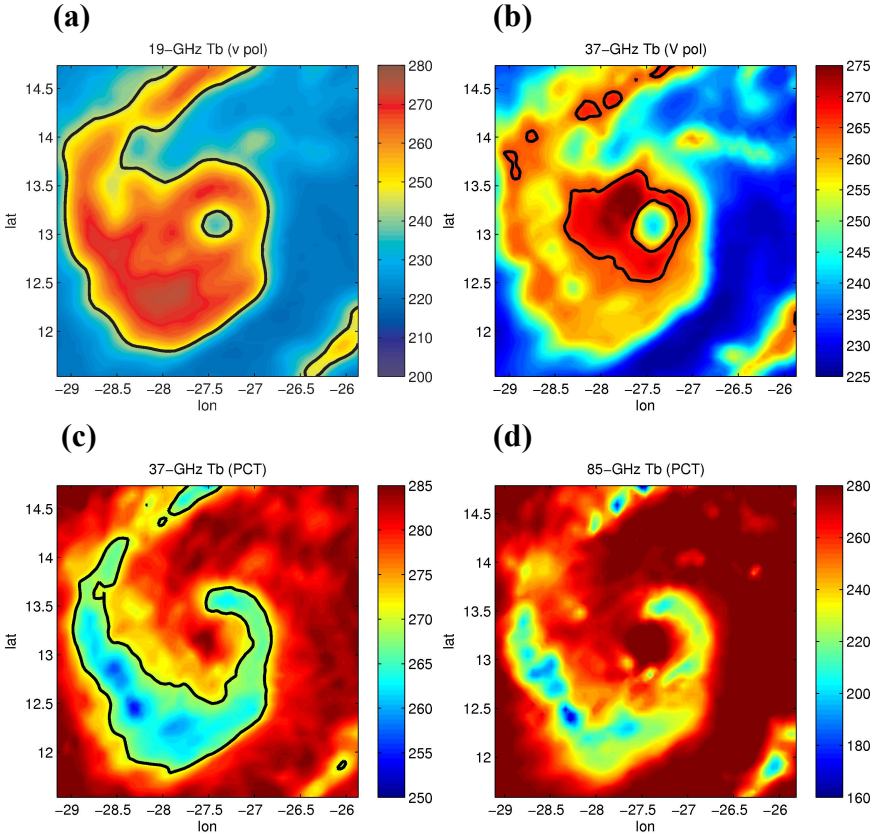

Fig. 6. Atlantic Hurricane Danielle at 1527 UTC 14 August 2004 depicted by TMI's (a) 19-GHz $T_{b,v}$, (b) 37-GHz $T_{b,v}$, (c) 37-GHz PCT, and (d) 85-GHz PCT.

A few salient details on the creation of predictors are worth noting. First, AMSR-E channels (18.7, 36.5, and 89.0 GHz) are calibrated to match TMI/SSMI channels (19.4, 37.0, and 85.5 GHz) using the histogram matching technique described in Jones and Cecil (2006). Similar histogram matching methods were applied to the somewhat different SSMIS channels – the SSMIS has a 91.7-GHz channel rather than the 85.5-GHz channel on TMI. However, the results

were not highly sensitive to these corrections in the SSMIS data. Therefore, SSMIS data are used without correction. Once AMSR-E data are calibrated, accurate center estimates are sought before creating MI-based predictors. All center fixes are based on the objective technique of Wimmers and Velden (2010). For the 19.4- and 37.0-GHz MI-based predictors, the 37.0-GHz channel is used for centering, whereas the TC centers in 85.5-GHz MI are found using the 85.5-GHz channel. Using multiple channels for center finding reduces parallax errors between the 37.0 and 85.5-GHz channels, while use of the 37.0-GHz-based center for the 19.4-GHz MI-based predictors is beneficial due to the lower spatial resolution of the 19.4-GHz channel.

Two types of MI-based predictors are defined. First, similar to the satellite infrared imagery-based predictors in the SHIPS developmental dataset, a variety of fixed-geometry predictors are developed using the MI. A second set of MI-based predictors is based on an objective estimate of the nascent or mature eye and eyewall regions and derived from parameters that result from the algorithm of Wimmers and Velden (2010). An example of an objectively determined ring (and eye enclosed within the ring region) in Hurricane Alex (2004) is shown in Fig. 7.

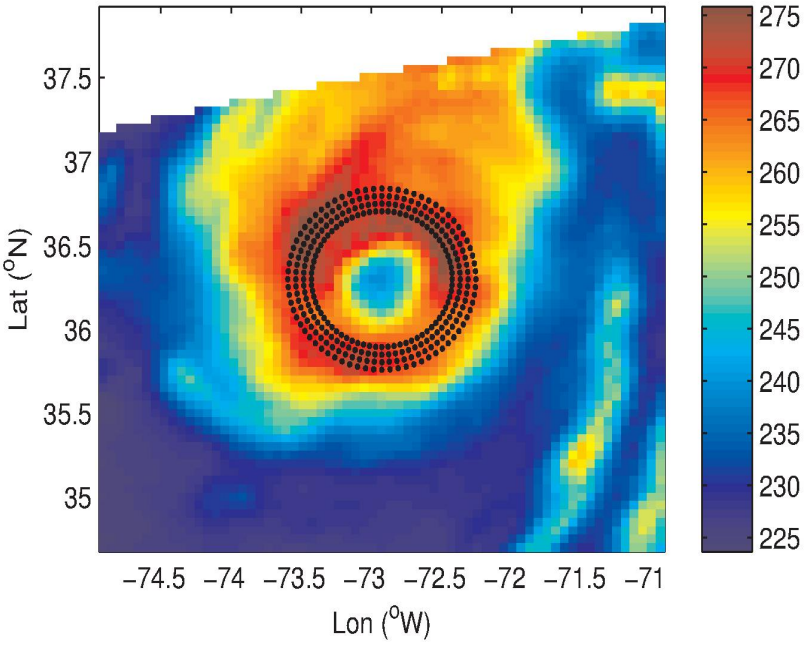

Fig. 7. An example of the objective ring found for Hurricane Alex (4 Aug. 2004) using the 37-GHz (vertical polarization) brightness temperatures (K).

Optimal storm-centered predictors are chosen such that they are statistically independent from other model predictors and so that they maximize BSS in independent leave-one-year-out cross validation. The years 1998-2012 are specifically used to train the models. Only data over the ocean are considered in the development of the logistic regression model predictors. Also, in order for an MI-based predictor to be considered, the difference in the composite means of the RI and non-RI samples must be statistically significant at the 95% level according to a two-sided student-*t* test. To properly compare the BSS of the MI-enhanced logistic regression model with the MI-free version, each model is trained and evaluated on only the forecast times in which all of the SHIPS-based and optimal MI-based predictors are available. Despite the uneven temporal coverage of satellite passes over a storm, the models are developed for making forecasts at the

synoptic times of 00, 06, 12, and 18 UTC. To deal with the irregular times in which MI of a TC exists, the MI-based schemes are only developed for forecasts in which the MI is less than 6-h old. Based on the data sets used here, microwave data are available roughly 40-60% of the time.

In both the Atlantic and eastern Pacific Ocean basins, MI-based logistic regression models have been derived from the SHIPS developmental dataset and the retrospective MI dataset for the RI thresholds of 25, 30, and 35 kt per 24 h. The optimal predictors for these models are detailed in Tables 6 and 7 of the Appendix. The BSS values derived from leave-one-year-out cross validation (1998-2012) for logistic regression models at various RI and TC intensity thresholds are shown in Fig. 8. In each case, significant forecast skill is provided by the addition of MI-based predictors. The Atlantic models experience greater relative improvement from the inclusion of MI-based predictors when compared to those tailored for the eastern Pacific Ocean basin, which tend to already hold an elevated baseline forecast skill.

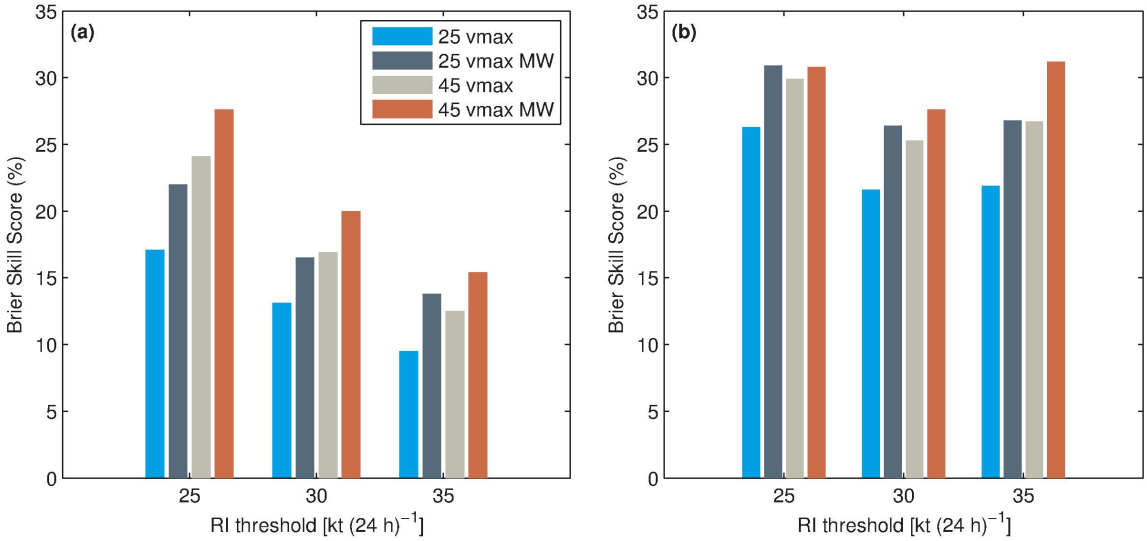

Fig. 8. Brier skill scores for (a) Atlantic and (b) eastern Pacific TCs with RI thresholds of 25, 30, and 35 kt per 24 h for the logistic regression model excluding MI-based predictors (light blue and gray for TCs of at least 25 kt and 45 kt intensity, respectively) and including MI-based predictors (dark blue and red for TCs of at least 25 kt and 45 kt intensity, respectively). These results are for the independent testing over the years of 1998-2012. In the Atlantic, the sample size for the 25-kt and 45-kt storm intensity thresholds is 1966 and 1479, respectively. The number of RI cases for the 25-, 30-, and 35-kt per 24 RI definitions is 231, 142, and 86 for the 25-kt storm intensity threshold and 174, 112, and 73 for the 45-kt storm intensity threshold. In the Eastern Pacific, the sample sizes for the two storm intensity thresholds are 1569 and 1073. The number of RI cases for the 25-, 30-, and 35-kt per 24 RI definitions is 231, 142, and 86 for the 25-kt storm intensity threshold and 174, 112, and 73 for the 45-kt storm intensity threshold.

Using Matlab scripts, the MI-based logistic regression models were run in near real-time at CIMSS late in 2013. (Fortran code to run these programs will soon be available.) Real-time streams of TMI, SSMI, and SSMIS data were obtained directly from NESDIS, which enabled the models to run in real time. It should be noted that real-time data streams like those that have already been established at NESDIS will need to be set up at NCEP as well for the MI-based

models to be run on the operational WCOSS computer and we are willing to assist with this task as needed.

As was the case for the consensus SHIPS-RII described in section 2a, to obtain a more robust evaluation of the MI-based models, the MI-based models were re-run for the period 2004-2013 using the real-time GFS forecast fields and initial storm data that was archived for this 10-year period. As before, before carrying out the re-run forecasts, cross-validated versions of each of the logistic regression-based RI models were re-derived for each of the years in the period of 2004-2012 by excluding data from a given year when evaluating the model's real-time performance over that year. This ensures that the reforecast results are independent from the training data. In evaluating the models, only forecast times in which all of the non-MI and MI-based predictors are available are included in the analysis to allow for a fair comparison of performance. Note, satellite data were available roughly 40% and 45% of the time for storms of at least 25 kt intensity in the Atlantic and eastern Pacific, respectively, and 47% and 60% of the time for storms having an intensity of at least 45 kt in the Atlantic and eastern Pacific, respectively.

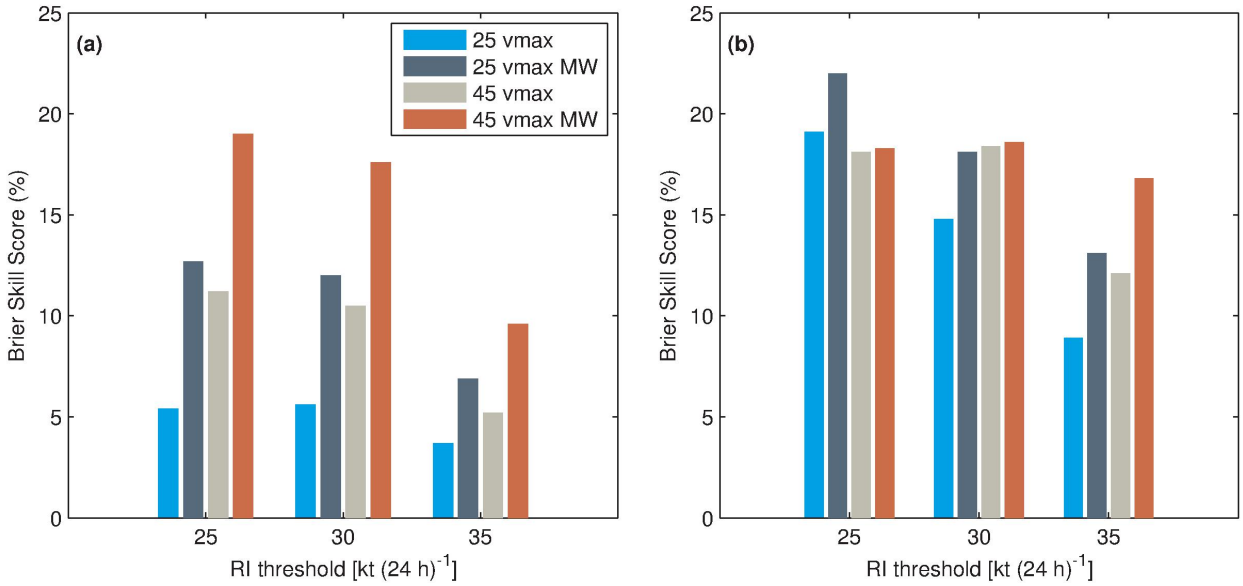

Fig. 9. Brier skill scores for (a) Atlantic and (b) eastern Pacific TCs with RI thresholds of 25, 30, and 35 kt per 24 h for the logistic regression model excluding MI-based predictors (light blue and gray for TCs of at least 25 kt and 45 kt intensity, respectively) and including MI-based predictors (dark blue and red for TCs of at least 25 kt and 45 kt intensity, respectively). These results are for the 2004-2013 re-rerun forecasts. In the Atlantic, the sample size for the 25-kt and 45-kt storm intensity thresholds is 1415 and 1021, respectively. The number of RI cases for the 25-, 30-, and 35-kt per 24 RI definitions is 153, 113, and 72 for the 25-kt storm intensity threshold and 118, 89, and 56 for the 45-kt storm intensity threshold. In the Eastern Pacific, the sample sizes for these two storm intensity thresholds are 1312 and 880. The number of RI cases for the 25-, 30-, and 35-kt per 24 RI definitions is 153, 99, and 72 for the 25-kt storm intensity threshold and 106, 73, and 54 for the 45-kt storm intensity threshold.

Figure 9 shows the performance of the real-time models (2004-2013) according to the BSS defined relative to climatology. In the Atlantic, the MI-free models suffer worse forecast skill at all RI thresholds (Fig. 9a) when compared to the results of the logistic regression model

in Fig. 1, but it is important to recall the logistic regression models in section 2a and 2c use different SHIPS developmental dataset-based predictors from one another, and furthermore, that the sample sizes are substantially smaller in the MI-based model evaluation due to only using forecasts from when MI is available. As in Fig. 8, analyzing storms of 45 kt or greater intensity improves the forecast skill for all three of the RI thresholds. The relative improvement of the MI-based models is quite striking in the Atlantic for storms having at least 25-kt and 45-kt intensity. On the other hand, the relative improvement (similar to Fig. 8) gained by using a MI-based model in the eastern Pacific is relatively modest at some of the intensity categories/RI thresholds. These skill differences that were observed between basins in the real-time evaluation is characteristic of the developmental dataset results (Fig. 8) as well. Nonetheless, there is uniform improvement in the eastern Pacific, which is significant in some cases (e.g., the 35 kt per 24 h RI threshold).

In conclusion, the addition of MI-based predictors to the logistic regression RI model was found to be quite beneficial even when using real-time data. These results boost confidence in the fidelity of MI-based RI models in the operational forecasting setting. While recent satellite data (less than 6-h old) are not always available, when they are, the improvements are considerable, particularly in the challenging Atlantic basin. It should also be noted that the loss of AMSR-E and the eventual loss of TMI, SSMI and SSMIS satellites all degrade the coverage of TCs by passive MI. In spite of this, the recent additions of the Advanced Microwave Scanning Radiometer 2 (AMSR2) and Megha-Tropiques Microwave Analysis and Detection of Rain and Atmospheric Systems (MADRAS) sensors and the upcoming SSMIS (DMSP-F19 and F20 satellites) and Global Precipitation Measurement (GPM) Microwave Imagery (GMI) missions alleviate concerns on data coverage. The continued availability of low-earth orbiting satellites offer the opportunity to exploit MI for improved statistical prediction of RI into the near future. Furthermore, there exists vast potential for further improvements in RI prediction through more sophisticated uses of MI and its adoption into other statistical models (e.g., SHIPS-RII) and numerical weather prediction (e.g., data assimilation).

Appendix

Previous 12-h intensity change
850-200 mb vertical shear from 0-500 km radius with vortex removed (time-averaged)
200 mb divergence from 0-1000 km radius (time-averaged)
Percent area with total precipitable water < 45 mm within 500-km radius 90° up-shear at $t = 0$ h
Second principle component of GOES-IR imagery within 440-km radius at $t = 0$ h
Std. dev. of 50-200 km GOES-IR brightness temperatures at $t = 0$ h
Potential intensity (Current intensity – maximum potential intensity)
Reynolds oceanic heat content (time-averaged)
Inner-core dry air predictor (time-averaged)
$t = 0$ h maximum sustained wind

Table 1. List of predictors used in the new Atlantic and E. Pacific multi lead-time versions of the SHIPS-RII. When a predictor is noted as being time-averaged then the magnitude of that particular variable was averaged every 6-h over the specified lead-time.

Previous 12-h intensity change
850-200 mb vertical shear from 0-500 km radius with vortex removed (time-averaged)
850-200 mb vertical shear from 200-800 km radius (time-averaged)
Generalized 850-200-hPa shear magnitude from 200-800-km radius that takes into account flow at all levels (time-averaged)
850-200 mb divergence from 0-1000 km radius (time-averaged)
200 mb divergence from 0-1000 km radius evaluated at 850 mb vortex location (time-averaged)
850-700 mb relative humidity (%) from 200-800 km radius
200 mb zonal wind averaged from 200-800 km radius
% area with Total precipitable water < 45 mm within 100-200 km radius
Total precipitable water averaged from 0-100 km radius
% area with Total precipitable water <45 mm from r=400 to r=600 km right quadrant
First principle component of GOES-IR imagery within 440 km radius at t = 0 h
Second principle component of GOES-IR imagery within 440 km radius at t = 0 h
Std. dev. of 50-200 km GOES-IR brightness temperatures at t = 0 h
% Area from 50-200 radius with GOES-RI brightness temperatures < -30 C at t = 0 h
% Area from 50-200 radius with GOES-RI brightness temperatures < -50 C at t = 0 h
Maximum GOES-IR brightness temperature from 0 to 30 km radius at t = 0 h
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Reynolds sea surface temperature (time-averaged)
Reynolds oceanic heat content (time-averaged)

Table 2. List of predictors used in the new Bayesian multi-lead time SHIPS-RII for the Atlantic basin. Note that only 11 of the above predictors were used at any one forecast lead-time.

Previous 12-h intensity change
850-200 mb vertical shear from 0-500 km radius with vortex removed (time-averaged)
Generalized 850-200-hPa shear magnitude from 200-800 km radius that takes into account flow at all levels (time-averaged)
200 mb divergence from 0-1000-km radius evaluated at 850 mb vortex location (time-averaged)
850-700 mb relative humidity (%) from 200-800-km radius
200 mb zonal wind averaged from 200-800 km radius
% Area with total precipitable water <45 mm from $r = 0$ to $r = 200$ km west quadrant
% Total precipitable water <45 mm, $r = 0$ to $r = 200$ km E quadrant
% Area with total precipitable water <45 mm from $r = 0$ to $r = 200$ km back quadrant
Second principle component of GOES-IR imagery within 440 km radius at $t = 0$ h
Trend in second principle component of GOES-IR imagery within 440-km radius during previous 6 h.
Std. dev. of 100-300 km GOES-IR brightness temperatures at $t = 0$ h
% Area from 50-200 radius with GOES-RI brightness temperatures < -10 C at $t = 0$ h
% Area from 50-200 radius with GOES-RI brightness temperatures < -20 C at $t = 0$ h
% Area from 50-200 radius with GOES-RI brightness temperatures < -50 C at $t = 0$ h
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Reynolds sea surface temperature (time-averaged)
Reynolds oceanic heat content (time-averaged)

Table 3. List of predictors used in the new logistic regression multi-lead time Atlantic RI models. Note that only 8 of the above predictors were used at any one forecast lead-time.

Previous 12-h intensity change
850-200 mb vertical shear from 0-500-km radius with vortex removed (time-averaged)
200 mb divergence from 0-1000-km radius evaluated at 850 mb vortex location (time-averaged)
Equivalent potential temperature excess of parcel relative to the 0-1000 km environment (time-averaged)
% Area with Total precipitable water < 45 mm from 0-200 km radius rear quadrant
Std. dev. of 100-300 km GOES-IR brightness temperatures at $t = 0$ h
% Area from 50-200 radius with GOES-RI brightness temperatures < -30 C at $t = 0$ h
% Area from 50-200 radius with GOES-RI brightness temperatures < -50 C at $t = 0$ h
Average GOES-IR brightness temperature from 100-300 km radius at $t = 0$ h
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Reynolds sea surface temperature (time-averaged)
Reynolds oceanic heat content (time-averaged)

Table 4. List of predictors used in the new Bayesian multi-lead time RI model for the E. Pacific basin. Note that only 11 of the above predictors were used at any one forecast lead-time

Previous 12-h intensity change
850-200 mb vertical shear from 0-500 km radius with vortex removed (time-averaged)
850-200-hPa shear magnitude from 200-800 km radius (time-averaged)
Vertical average of the negative differences between the equivalent potential temperature of a parcel lifted from the surface and the environmental saturation equivalent potential temperature from 200-800 km radius (time-averaged)
200 mb divergence from 0-1000 km radius evaluated at 850 mb vortex location (time-averaged)
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Std. dev. of 100-300 km GOES-IR brightness temperatures at t = 0 h
% Area from 50-200 radius with GOES-RI brightness temperatures < -20 C at t = 0 h
Average GOES-IR brightness temperature from 0–200-km radius at t = 0 h
% Area from 50-200 radius with GOES-RI brightness temperatures < -50 C at t = 0 h
Maximum GOES-IR brightness temperature from 0 to 30 km radius at t=0 h

Table 5. List of predictors used in the new logistic regression multi-lead time E. Pacific RI models. Note that only up to 7 of the above predictors were used at any one forecast lead-time.

Previous 12-h intensity change
Reynolds sea surface temperature (time averaged)
200-hPa divergence from 0–1000-km radius evaluated at 850 mb vortex location (time-averaged)
850-200-mb vertical wind shear magnitude calculated with the vortex removed and relative to the 850-hPa center (r = 0–500 km) (time-averaged)
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Std. dev. of 100–300 km infrared brightness temperatures at t = 0 h
Mean infrared cloud-top brightness temperature (r = 0–30 km) at t = 0 h
MI: Min. 37-GHz (horiz. pol.) brightness temperatures in objective ring (within t = -6 h)
MI: Max 37-GHz (horiz. pol.) brightness temperatures in objective ring (within t = -6 h)
MI: Max. 85-GHz polarization corrected temperature of the eye (within t = -6 h)
MI: Radius of max. 37-GHz (vert. pol.) bright. temp. (r = 30–130 km) (within t = -6 h)
MI: Radius of min. 85-GHz (horiz. pol.) bright. temp. (r = 30–130 km) (within t = -6 h)

Table 6. List of predictors used in the 24-h lead-time MI-enhanced logistic regression model for the Atlantic Ocean. MI-based predictors are specifically labeled.

Previous 12-h intensity change
Equivalent potential temperature excess of parcel relative to the 0–1000-km environment (time-averaged)
850-200-mb vertical wind shear magnitude calculated with the vortex removed and relative to the 850-hPa center (r = 0–500 km) (time-averaged)
Potential intensity (Current intensity – maximum potential intensity) (time-averaged)
Std. dev. of 50–200-km infrared brightness temperatures at t = 0 h
Mean infrared cloud-top brightness temperature (r = 100–300 km) at t = 0 h
Mean infrared cloud-top brightness temperature (r = 0–30 km) at t = 0 h
MI: Mean 37-GHz (vert. pol.) brightness temperatures in objective ring (within t = -6 h)
MI: Std. dev. 19-GHz (horiz. pol.) brightness temperatures of the eye (within t = -6 h)
MI: Radius of min. 85-GHz pol. corrected temp. (r = 30–130 km) (within t = -6 h)
MI: Mean 37-GHz (horiz. pol.) brightness temp. (r = 100–300 km) (within t = -6 h)
MI: Std. dev. 37-GHz polarization corrected temp. (r = 100–300 km) (within t = -6 h)

Table 7. List of predictors used in the 24-h lead-time MI-enhanced logistic 24-h lead-time Eastern Pacific RI model. MI-based predictors are specifically labeled.

References

Goerss, J. S. and C. R. Sampson, 2013: Prediction of Consensus Tropical Cyclone Intensity Forecast Error, *Wea. Forecasting*, review.

Jones, T. A., and D. J. Cecil, 2006: Histogram matching of AMSR-E and TMI brightness temperatures. Preprints, *14th Conf. on Satellite Meteorology and Oceanography*, Atlanta, GA, Amer. Meteor. Soc., CD-ROM, P1.23.

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification for the Atlantic and eastern North Pacific basins. *Wea. Forecasting,* ***25*****,** 220-241.

_____, J. Cione, M. DeMaria, J. Dostalek, J. Dunion, J. Knaff, J. Zhang, T. Lee, J. Hawkins, J. Solbrig, E. Kalina, and P. Leighton, 2011: Improvement in the rapid intensity index by incorporation of inner-core information. JHT final report. Available from http://www.nhc.noaa.gov/jht/09-11reports/final_Kaplan_JHT11.pdf.

Kieper, M., and H. Jiang, 2012: Predicting tropical cyclone rapid intensification using the 37 GHz ring pattern identified from passive microwave measurements. *Geophys. Res. Letts.*, **39**, L1804, doi:10.1029/2012GL052115.

Rozoff, C. M., and J. P. Kossin, 2011: New probabilistic forecast models for the prediction of tropical cyclone rapid intensification. *Wea. Forecasting*, **26**, 677-689.

Sampson, C. R., J. Kaplan, J. A. Knaff, M. DeMaria, and C. A. Sisko, 2011: A deterministic rapid intensification aid, *Wea. Forecasting*, **26**, 579-585.

_____, J. L. Franklin, J. A. Knaff and M. DeMaria, 2008: Experiments with a Simple Tropical Cyclone Intensity Consensus. *Wea. Forecasting,* **23**, 304-312.

Wilks, D.S., 2011: *Statistical methods in the Atmospheric Sciences*, 3d ed.. Elsevier, pp 676.

Wimmers, A. J., and C. S. Velden, 2010: Objectively determining the rotational center of tropical cyclones in passive microwave satellite imagery. *J. Appl. Meteor. Climatol.*, **49**, 2013-2034.