PROJECT REPORT

NOAA/OAR Joint Hurricane Testbed

Federal Grant Number: NA15OAR4590205

Probabilistic Prediction of Tropical Cyclone Rapid Intensification Using Satellite Passive Microwave Imagery

Principal Investigators
Christopher S. Velden[1], chris.velden@ssec.wisc.edu
Christopher M. Rozoff[2], rozoff@ucar.edu

Submission Date: 30 March 2018

[1]Cooperative Institute for Satellite Meteorological Studies (CIMSS)
University of Wisconsin-Madison
1225 West Dayton Street
Madison, WI 53706

[2]National Security Applications Program
Research Applications Laboratory
National Center for Atmospheric Research
P.O. Box 3000
Boulder, CO 80307-3000

Project/Grant Period: 1 September 2017 – 30 March 2018

Report Term or Frequency: Semi-Annual

Final Annual Report? No

1. ACCOMPLISHMENTS

The primary goal of this project is to improve the probabilistic prediction of rapid intensification (RI) in tropical cyclones (TCs). The framework in which we work is probabilistic models. We specifically are innovating upon existing statistical models that use environmental and TC-centric predictors. The statistical models used in this work include the Statistical Hurricane Intensity Prediction System (SHIPS) RI Index (RII) (Kaplan et al. 2015) and the logistic regression and Bayesian models of Rozoff and Kossin (2011) and Rozoff et al. (2015).

The objectives of this project are to update the three statistical models to include a new class of predictors derived from passive microwave imagery (MI) evincing aspects of storm structure relevant to RI, using a comprehensive dataset of MI that includes all available relevant sensors, and to develop a skillful consensus model that can be tested and deployed in real-time operations.

Milestones

a. **Climatological microwave dataset**

We created a large climatological passive microwave dataset containing virtually all satellite passes over all Atlantic and Eastern Pacific TCs at the 18.7, 36.5, and 89 GHz channels. This dataset covers the years 1995-2016 and is being updated to include 2017. The satellites used include the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), the Advanced Microwave Scanning Radiometer 2 (AMSR-2), the Special Sensing Microwave Imager (SSM/I), the Special Sensing Microwave Imager/Sounder (SSMIS), the Tropical Rainfall Measuring Mission Microwave Imager (TMI), and the Global Precipitation Measurement Microwave Imager (GMI). The SSM/I, SSMIS, and TMI are calibrated to match the exact frequencies of the AMSR-E/2 and GMI instruments following the histogram matching technique described in Rozoff et al. (2015). Data are saved into native swath grids and also polar grids centered on the TC center. These data are used in this project to create MI-based predictors and are now being used by a number of research groups across the country, including SUNY-Albany, RSMAS/Miami, and Penn State University. Recently, a publication using this dataset has uncovered unique storm structures normalized by storm intensity that provide insight into RI (Fischer et al. 2018).

b. **Baseline and New Models**

Using SHIPS developmental data and new microwave developmental datasets, we developed the new Bayesian, logistic regression, and SHIPS-RII models. Conforming with the operational SHIPS-RII consensus model, we have derived models for the following RI thresholds: 20 kt / 12 h; 25, 30, 35, and 40 kt / 24 h; 45 kt / 36 h; 55 kt / 48 h; and 65 kt / 72 h. The predictors for these models were described in the previous semi-annual report.

The Brier skill score improvements to the consensus of the RI models by including microwave-based predictors are shown in Fig. 1. The Brier skill score with respect to a climatological baseline is used to evaluate the model skill. The models with and without microwave-based predictors are evaluated for the exact same forecasts over the period 1998-2016 in both the Atlantic and Eastern

Pacific using leave-one-year-out cross validation. In both basins, and for all models, skill is substantially improved by the inclusion of the microwave-based predictors listed in Tables 2 and 4, although the relative improvements become small or zero at the 65 kt / 72-h RI threshold due to the lack of MI-based predictors. We note that all consensus member models also experience enhanced skill by including MI-based predictors except at the 65 kt / 72-h threshold, where the improvements are small or zero due to few or no MI-based predictors used. The consensus produces the highest skill, consistent with the results of the non-microwave-based models in Kaplan et al. (2015).

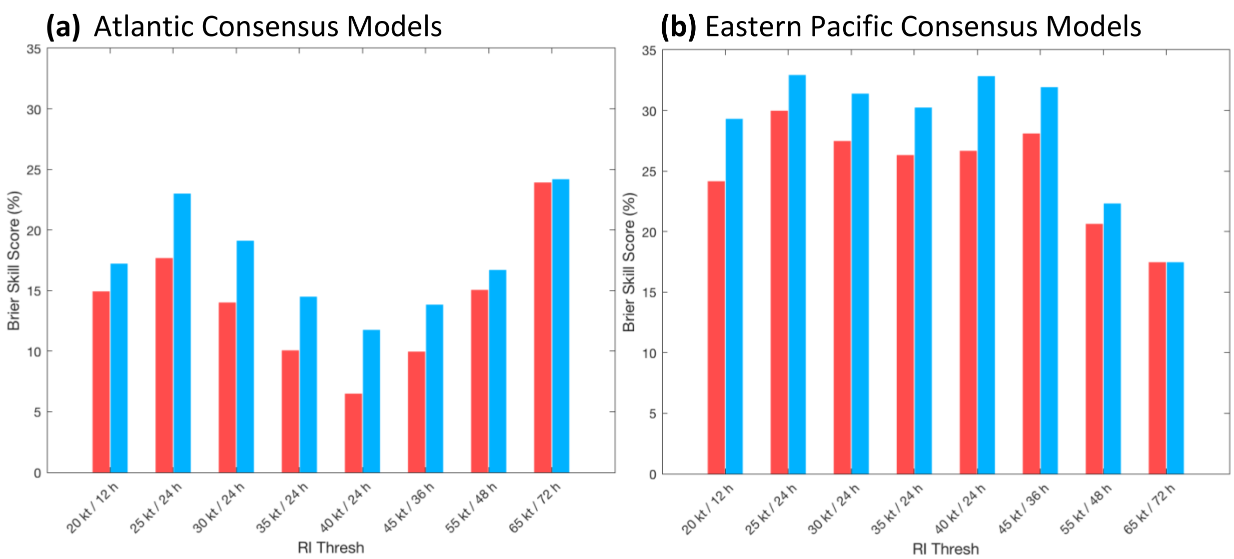

Figure 1. The Brier skill scores of the consensus RI model with (blue) and without (red) microwave-based predictors for the (a) Atlantic and (b) Eastern Pacific using leave-one-year-out cross-validation for the years 1998-2016.

c. **Real-time testing of models during 2017 Atlantic and Eastern Pacific hurricane seasons**

An experimental website was developed to demonstrate and allow a quick assessment of the probabilities of RI with the updated models and the inclusion of satellite microwave data in real-time. The site can be found at http://tropic.ssec.wisc.edu/real-time/mw-ri-prob/. Four different microwave sensors are being used in this real-time demonstration:

1) Special Sensor Microwave Imager (SSM/I) from the Defense Meteorological Satellite Program (DMSP) F15 satellite
2) Special Sensor Microwave Imager/Sounder (SSMI/S) from the DMSP F16, F17, and F18 satellites
3) Advanced Microwave Scanning Radiometer 2 (AMSR2)
4) Global Precipitation Measurement (GPM) Microwave Imager (GMI)

An example of the real-time website from 2240 UTC 24 August 2017 during Hurricane Harvey is illustrated in Fig. 2. The probabilities of RI with microwave predictors are displayed in the left table, and the concurrent operational RI model probabilities are displayed in the right table. Both probabilities are shaded based on percentages. Past probabilities can be viewed in the chart below the left table. Boxes are shaded based on consensus RI probabilities using microwave data. White boxes indicate probabilities were not available due to the lack of MW data in that analysis cycle, which has an occurrence rate of about 39% in the Atlantic Basin. In addition, no RI probabilities are calculated using microwave data if the TC center is too close to land.

Past RI probability tables for a particular storm can also be viewed by clicking on the “History” link or the TC name in the right column. An example of the RI probabilities history for TC Katia (2017) is shown in Fig. 3. The intensity of the TC and the maximum potential intensity (MPI) are listed. (In the SHIPS-RII model, the model probabilities are set to zero if the RI thresholds exceeds the MPI.) Between 0600 UTC 6 September 0600 UTC 7 September 2017, Katia rapidly intensified by 35 kts, and 30kts between 1200 UTC 6 September and 1200 UTC 7 September. In both of these cases, higher RI probabilities were predicted with the microwave-enhanced models.

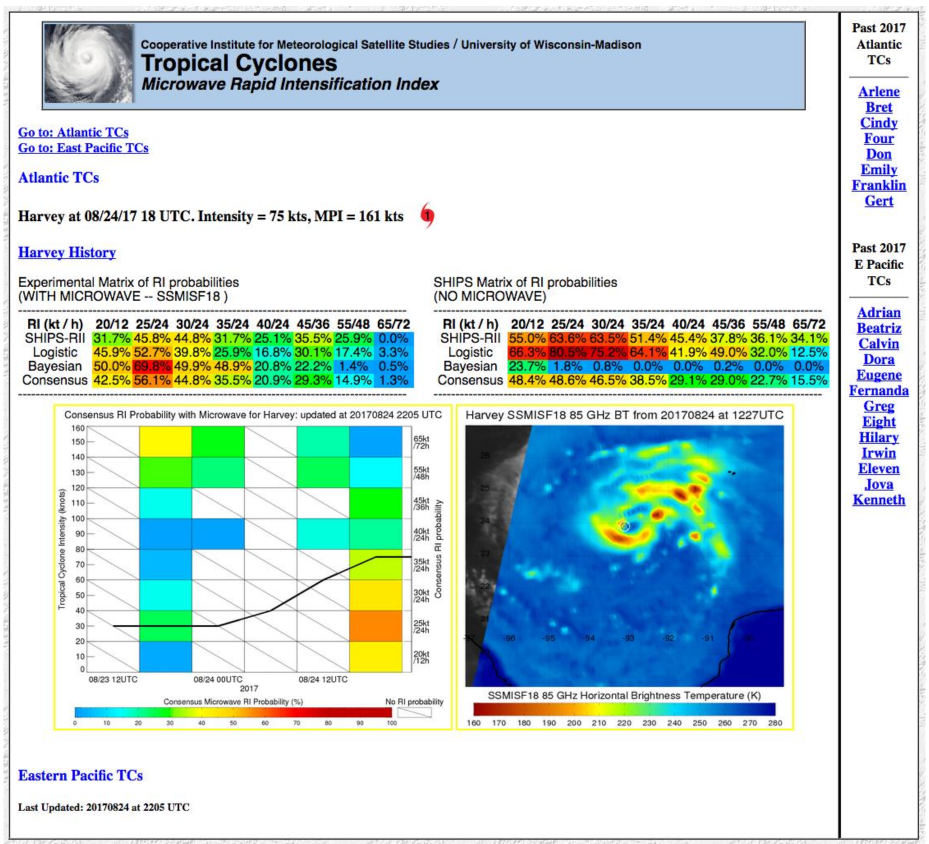

Figure 2. Example of the real-time TC RI prediction site designed to demonstrate the upgraded RI models using microwave data.

Katia at 09/06/17 18 UTC. Intensity = 65 kts, MPI = 162 kts

Experimental Matrix of RI probabilities
(WITH MICROWAVE -- SSMISF18)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	34.4%	50.0%	50.0%	31.7%	14.4%	34.0%	35.6%	57.5%
Logistic	38.4%	71.2%	64.2%	39.3%	16.7%	51.7%	55.8%	67.0%
Bayesian	30.0%	49.9%	10.7%	3.1%	0.5%	5.1%	0.6%	3.1%
Consensus	34.3%	57.0%	41.6%	24.7%	10.5%	30.3%	30.7%	42.5%

SHIPS Matrix of RI probabilities
(NO MICROWAVE)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	46.8%	64.8%	63.4%	54.3%	44.3%	51.4%	41.0%	59.5%
Logistic	55.8%	85.1%	80.1%	70.7%	42.6%	65.9%	51.8%	41.8%
Bayesian	29.8%	13.8%	13.4%	6.4%	5.4%	2.7%	0.3%	0.5%
Consensus	44.1%	54.6%	52.3%	43.8%	30.8%	40.0%	31.1%	33.9%

Katia at 09/06/17 12 UTC. Intensity = 40 kts, MPI = 170 kts

Experimental Matrix of RI probabilities
(WITH MICROWAVE -- AMSR2)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	31.8%	50.1%	50.1%	32.2%	6.4%	39.0%	28.8%	51.3%
Logistic	36.7%	82.0%	70.6%	50.4%	12.5%	52.4%	56.0%	74.9%
Bayesian	12.3%	71.6%	24.7%	4.2%	0.8%	17.1%	5.1%	23.7%
Consensus	26.9%	67.9%	48.5%	28.9%	6.6%	36.2%	30.0%	50.0%

SHIPS Matrix of RI probabilities
(NO MICROWAVE)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	18.5%	50.0%	38.3%	23.2%	12.7%	31.9%	38.7%	60.0%
Logistic	34.1%	69.9%	59.8%	43.2%	21.1%	54.1%	51.2%	51.2%
Bayesian	11.8%	41.6%	28.7%	6.3%	5.3%	5.2%	2.0%	12.1%
Consensus	21.5%	53.9%	42.3%	24.2%	13.0%	30.4%	30.6%	41.1%

Katia at 09/06/17 06 UTC. Intensity = 35 kts, MPI = 164 kts

Experimental Matrix of RI probabilities
(WITH MICROWAVE -- SSMISF18)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	22.6%	45.5%	39.1%	7.7%	4.2%	38.6%	24.4%	44.9%
Logistic	48.8%	54.7%	43.6%	23.8%	17.1%	60.8%	40.1%	64.3%
Bayesian	12.5%	11.6%	11.6%	0.9%	0.9%	15.5%	1.9%	6.6%
Consensus	28.0%	37.3%	31.4%	10.8%	7.4%	38.3%	22.1%	38.6%

SHIPS Matrix of RI probabilities
(NO MICROWAVE)

RI (kt / h)	20/12	25/24	30/24	35/24	40/24	45/36	55/48	65/72
SHIPS-RII	9.1%	37.3%	21.6%	10.2%	8.8%	15.4%	24.4%	50.3%
Logistic	20.0%	46.2%	34.1%	19.7%	13.2%	43.7%	30.6%	37.5%
Bayesian	4.8%	27.7%	18.1%	4.2%	2.6%	4.8%	1.2%	3.8%
Consensus	11.3%	37.1%	24.6%	11.4%	8.2%	21.3%	18.7%	30.5%

Figure 3. Example of RI probabilities from Atlantic TC Katia (2017).

Unfortunately, as the season ended, the results of this particular set of models were not superior to the operational RI models that do not use MI-based predictors. We briefly summarize the results here. A homogeneous comparison of the Brier skill scores of all the new MI-enhanced RI models run in real-time (with real-time quality data) vs. their non-MI-based operational counterparts is shown in Fig. 4. In the Atlantic, the BSS of the MI-based models is positive and not terribly worse than the operational models at most lead-times, although there are a couple of instances where the MI-based models outperform the operational models. In particular, both the Bayesian and logistic regression models outperformed the operational versions at the 45 kt / 24 h threshold. The MI-enhanced logistic regression model also slightly outperformed the operational version at 20 kt /12 h and 65 kt / 72 h. The MI-enhanced Bayesian model also outperformed the operational model at the 55 kt / 24 h and 65 kt / 72 h thresholds. Overall, however, the consensus is slightly worse at most RI thresholds in the Atlantic. The results in the Eastern Pacific are similar. If we stratify the results by RI and non-RI cases, we see the MI-enhanced models actually produce higher BSS than the operational models at the majority of RI thresholds for non-RI cases. Therefore the MI-based models reduce false alarms compared to the operational models. However, a primary objective is to also improve RI detection. We suspect part of the reason for the slightly worst performance of the MI-based models is that the training dataset is much smaller than the operational dataset, since satellite overpasses cover TCs only roughly half of the time and the training period is also shorter (1998-2016 vs. 1995-2016). Thus, while MI-based predictors can enhance skill when the models are trained on datasets of equal size and for equivalent dates and times, we hypothesize it would take exceptional improvements from the MI-based predictors to overcome the training size deficit conundrum when comparing the MI-based model against operational models trained on much larger datasets.

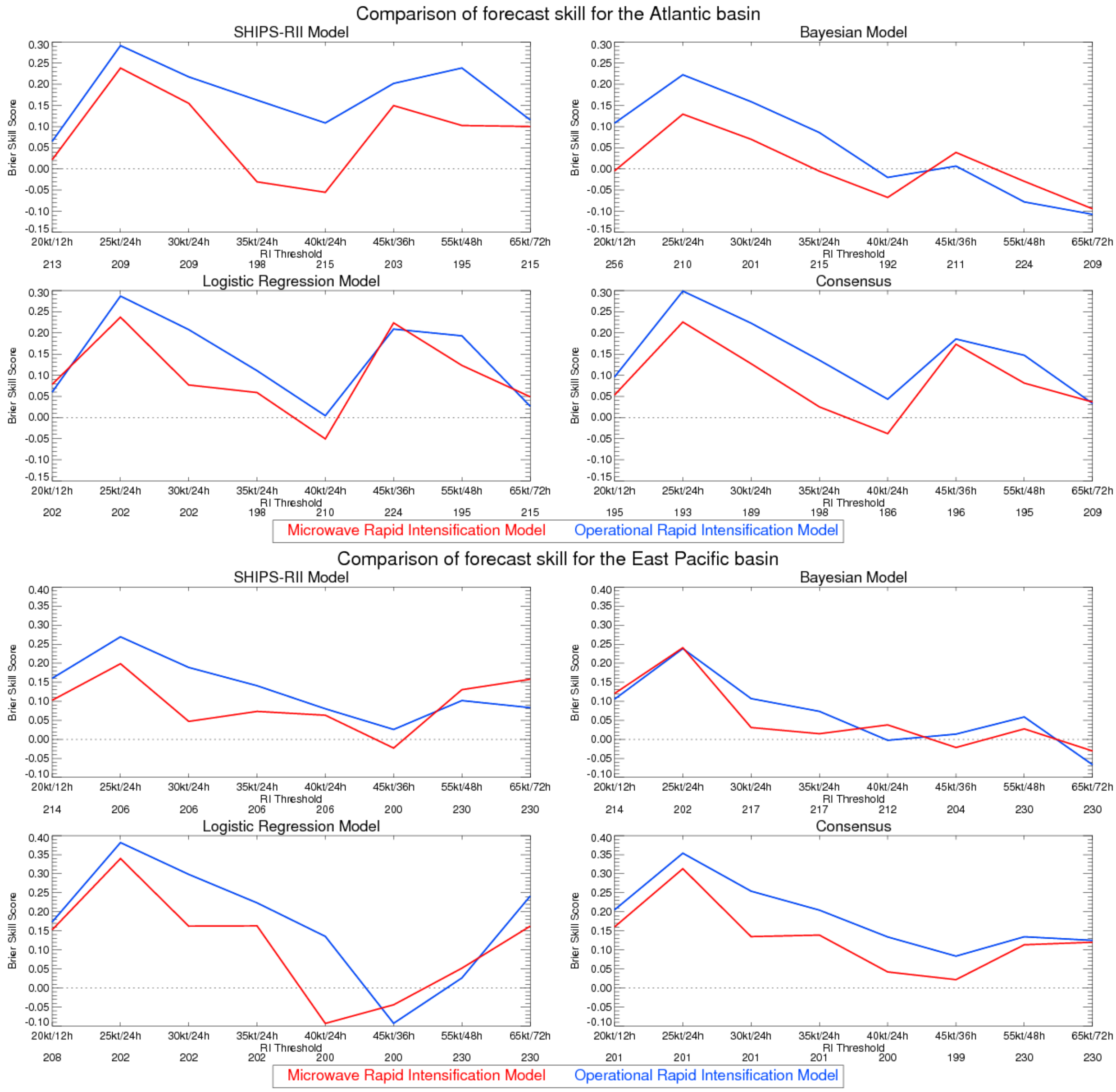

Figure 4. Brier skill scores for all RI thresholds for the MI-enhanced RI models (red) and non-MI-based operational RI models (blue) for the Atlantic (top) and Eastern Pacific (bottom).

d. **A Revised Model**

Given these negative results, we are now seeking a new focus to see if we can find exceptional MI-based predictors that overcome the training sample challenge. Using our MI dataset, Fischer et al. (2018) developed a normalized MI analysis technique that can better detect TC structure associated with RI. The procedure to create normalized fields is done with respect to a TC's current intensity, such that the climatological microwave imagery structure for a given intensity bin is subtracted from each microwave image in the same intensity bin. Each difference field is then normalized by the standard deviation of the climatological structure. This analysis is done on imagery rotated with respect to the deep-layer shear vector. The 37-GHz (horizontal polarization) and 85-GHz polarization corrected temperatures were evaluated and found to be the most useful to detect robust RI structure. Robust anomaly structures were found in the normalized brightness temperature fields for weakening, steady, slowly intensifying, and rapidly intensifying storms. Simple tests were conducted in Fischer et al. (2018) showing that simple azimuthal means of the normalized fields over some radial range can provide skillful RI predictions.

Given these promising results, we discard the previous MI-based predictors (from previous reports and results above) and simply test a small set of predictors using the normalized 37-GHz and 85-GHz fields of Fischer et al. (2018), including azimuthal mean brightness temperatures for the radial ranges of r = 0 to 50, 75, 100, and 150 km and quadrant average brightness temperatures for r = 100 km, where the quadrants are defined with respect to the deep-layer shear vector. Given time constraints, development and testing were performed only on the logistic regression and Bayesian models, but will also be extended to SHIPS-RII by the final report. It turns out, for each model and threshold in both basins, one 37 GHz-based predictor and sometimes a second 85-GHz predictor always benefit the model compared to an equivalent model not possessing a MI-based predictor. More often than not, the upshear-left and upshear-right quadrants of the normalized 37-GHz imagery provides the most useful predictors, indicating a more anomalously active convective region upshear of the storm is more conducive to RI (i.e., the storm is more symmetric).

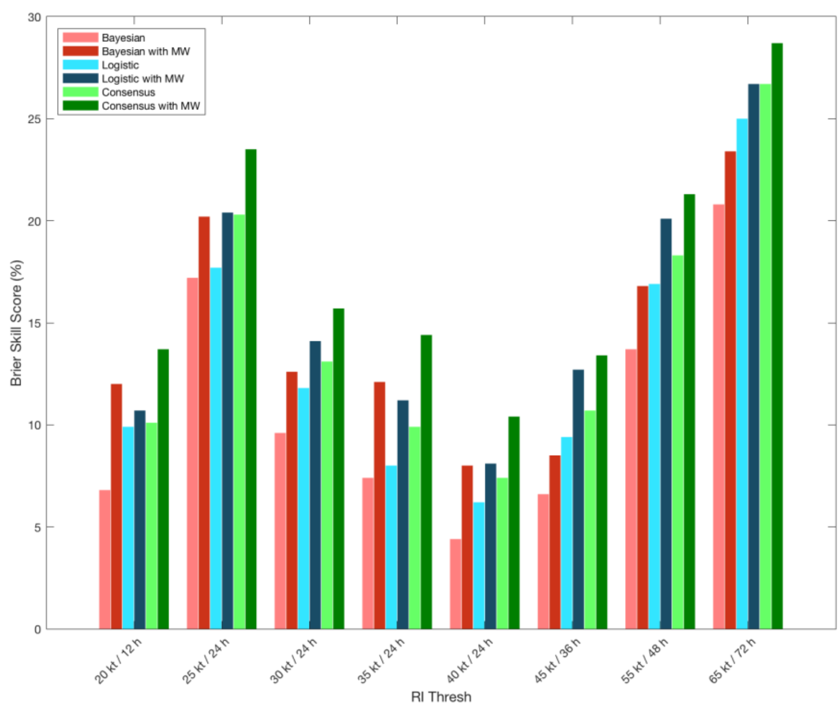

Figure 5. The Brier skill scores of the Bayesian model, logistic regression model, and their consensus with and without microwave-based predictors for the Atlantic using leave-one-year-out cross-validation for the years 1998-2016.

The Brier skill scores for the new Bayesian and logistic regression models, along with a two-model consensus, for the Atlantic and Eastern Pacific are provided in Figs. 5 and 6. At every threshold, the MI-based predictors add significant skill. As before, the Eastern Pacific offers much higher skill. Interestingly, the two-model consensus here performs competitively, if not better, than the three-model consensus with the older MI-based models in Fig. 1. A 2017 post-season analysis is currently being performed at the time of this report but has not yet been completed but should be concluded soon.

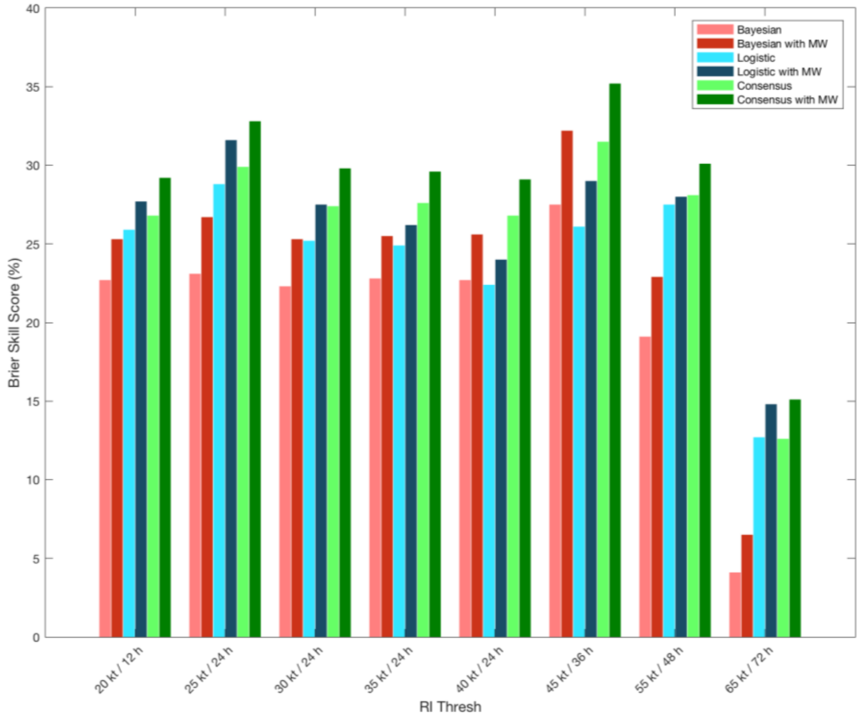

Figure 6. The Brier skill scores of the Bayesian model, logistic regression model, and their consensus with and without microwave-based predictors for the Eastern Pacific using leave-one-year-out cross-validation for the years 1998-2016.

Status of Project Tasks / Milestones

The following table summarizes the tasks originally proposed with some updated dates due to a NCE, and the status of these tasks.

Task	Proposed Activity	Status
1	Update developmental dataset to include MI of Atlantic and Eastern Pacific TCs from all available sensors (1998-2016). [September 2015 – January 2017]	Completed
2	Examine and test for significance of new MI-based predictors. [September 2015 – January 2016]	Completed
3	Update logistic regression model to incorporate improved MI predictors and evaluate on retrospective and real-time cases. [January – March 2016]	Completed
4	Enhance the Bayesian and linear discriminant analysis-based SHIPS-RII models with up-to-date MI dataset. [January – March 2016]	Completed
5	Evaluation of updated SHIPS-RII and Bayesian models on retrospective dataset. [March – May 2016]	Completed
6	Convert code from Matlab (development framework) to Fortran and C so that code is portable to NCEP operations. [April 2016 – December 2017]	In Progress pending new results

7	In-house real-time testing of models in the Atlantic and Eastern Pacific and continue reforecasts of previous seasons in simulated operational conditions with archived real-time data. [June – November 2016]	Completed
8	Evaluation of models and model updates. [January – December 2017]	Completed
9	Prepare final NCEP-ready code and documentation for running and maintaining models at the conclusion of the project. [February – December 2017]	In Progress pending new results
10	Operational demo real-time test. [June – November 2017]	Completed

We are currently testing new model revisions given the real-time testing results of 2017. We will provide real-time results to JHT/NHC points of contact for these revised models as soon as they become available. Items 6 and 9 will be completed, and will be delivered and supported at no-cost post project if the NHC is interested in the JHT results here and as found in our real-time evaluation.

References

Fischer, M. S., B. H. Tang, K. L. Corbosiero, and C. M. Rozoff, 2018: Normalized convective characteristics of tropical cyclone rapid intensification events in the North Atlantic and Eastern North Pacific. *Mon. Wea. Rev.*, in press.

Kaplan, J., C. M. Rozoff, and Co-Authors, 2015: Evaluating environmental impacts of tropical cyclone rapid intensification predictability utilizing statistical models. *Wea. Forecasting*, **30**, 1374-1396.

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.

Rozoff, C. M., C. S. Velden, J. Kaplan, J. P. Kossin, and A. J. Wimmers, 2015: Improvements in the probabilistic prediction of tropical cyclone rapid intensification with passive microwave observations. *Wea. Forecasting*, **30**, 1016-1038.

What opportunities for training and professional development has the project provided?

This project will provide training for forecasters for the use of MI-based probabilistic RI models in operations.

How were the results disseminated to communities of interest?

We provided the real-time results on a shared webpage with our points of contact at NHC. Prior results have been presented at conferences. We will provide a real-time-capable version of our algorithm code to NHC if approved.

What do you plan to do during the next reporting period to accomplish the goals and objectives?

We plan to do final real-time testing on revised RI models and finish the work on developing a real-time Fortran/C-based algorithm that can operate successfully on NOAA computers.

2. PRODUCTS

Presentations in this reporting period
None to report.

Publications
None to report. However, we will submit a paper on the results of this project pending 2017 post-season analysis with the revised RI models.

Products
We will submit a Fortran/C-based algorithm of the MI-enhanced RI models to be run on NOAA HPC systems pending the 2017 post-season analysis with the RI models using MI-based predictors derived from the normalized brightness temperatures dataset.

3. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on this project?
Christopher Velden (PI), Christopher Rozoff (Co-PI), Sarah Griffin (CIMSS/UW-Madison research assistant)

Has there been a change in the PD/PI(s) or senior/key personnel since the last reporting period?
No

What other organizations have been involved as partners? Have other collaborators or contacts been involved?
Forecasters and Program Officials (e.g., Shirley Murillo and Christopher Landsea and all other NHC points of contact) at the National Hurricane Center/Joint Hurricane Testbed have been briefed on the project progress.

4. IMPACT

What was the impact on the development of the principal discipline(s) of the project?
We have anticipated that this project will improve one of the NHC’s most reliable forecast tools for predicting RI in TCs, thereby helping NHC improve intensity prediction of TCs. While this project is highly practical, the results of this project may also contribute to increased scientific understanding of intensification processes in TCs.

What was the impact on other disciplines?
While the impact may be minimal, other disciplines often use the types of statistical models we have used in this project, and therefore researchers may find our project methodology useful.

What was the impact on the development of human resources?
None to report.

What was the impact on teaching and educational experiences?

None to report.

What was the impact on physical, institutional, and information resources that form infrastructure?
None to report.

What was the impact on technology transfer?
None to report.

What was the impact on society beyond science and technology?
Improved TC intensity prediction, especially RI, will be extremely valuable to society, particularly emergency management planning.

What percentage of the award's budget was spent in a foreign country(ies)?
0%.

5. CHANGES/PROBLEMS

The original set of MI-based models did not perform up to our expectations, so efforts were recently made to try out new innovations in the analysis of microwave imagery with respect to storm intensification.

6. SPECIAL REPORTING REQUIREMENTS

We report here on the project's Readiness Level as part of the Joint Hurricane Testbed.

Transition to operations activities
The statistical modeling framework is being developed to run in real-time and also in Fortran/C-based code (as opposed to the Matlab developmental framework) so that it will be readily able to run in an operational environment, including the WCOSS high performance computing system.

Summary of testbed-related collaborations, activities, and outcomes
We have communicated with points of contact (POC) Christopher Landsea, John Beven, Daniel Brown, and Dave Roberts at the NHC for real-time analysis and testing during the 2017 hurricane season. They also requested results from the 2018 hurricane season given our difficulties with an earlier version of our RI models.

Has the project been approved for testbed testing yet?
The 2017 real-time testing is being performed on CIMSS computing platforms. We were requested to run on CIMSS platforms again in 2018, so we will continue to do so and provide results in our final report.

What was transitioned to NOAA?
Nothing at this time.

Thus, the project's ***Readiness Level (RL) is at level 6***, and we continue working on RL stages 7-8.

7. BUDGETARY INFORMATION

A NCE was granted to extend the project in order to perform real-time testing. This project will wrap up 8/31/18.

8. PROJECT OUTCOMES

What are the outcomes of the award?
We have developed a multi-model consensus of probabilistic models that predict the likelihood or rapid intensification in tropical cyclones. In particular, we have updated these models to use new predictors from satellite passive microwave imagery. This consensus model improves forecast skill over its constituent models and over the same models not employing microwave data in leave-one-year-out cross validation, but did not outperform operational versions of the same models that do not use MI-based imagery. A simpler and more elegant set of MI-based models were more recently developed that appear promising but require further real-time testing to verify whether they will outperform existing operational models.

Are performance measures defined in the proposal being achieved and to what extent?
Besides the delay in demonstrating a new model in real-time, performance measures are being otherwise achieved.