FY2017 Joint Hurricane Testbed (JHT) NOAA-OAR-OWAQ-2017-2005004
NOAA grant NA17OAR4590137

Evolutionary Programming for Probabilistic Tropical Cyclone Intensity Forecasts

PI: Paul J. Roebber, University of Wisconsin – Milwaukee (roebber@uwm.edu; 414-229-3950)
Co-PI: Clark Evans, University of Wisconsin – Milwaukee (evans36@uwm.edu; 414-229-4469)

January 30, 2018

Board of Regents of the University of Wisconsin System
for the University of Wisconsin at Milwaukee
P.O. Box 340
Milwaukee, WI 53201-0340

Project/Grant Period: 7/1/2017–6/30/2019

Reporting Period End Date: 1/30/2018

Report Term or Frequency (semi-annual)

Final Annual Report? Select Yes or **No**

1. ACCOMPLISHMENTS

Table 1 provides planned versus actual accomplishments as identified in the funded proposal. As shown in the Table, the project is on-track and on-schedule.

Action	Planned completion	Actual completion	Status (if not completed)
Finish development, cross-validation, and retrospective testing of EP and BMC intensity, RI, and RW forecasts.	11-1-2017	12-15-2017	Complete.
With JHT collaborators, finalize year one testbed evaluation plan.	1-1-2018	1-22-2018	Complete.
Port EP and BMC code to JHT computational infrastructure.	3-1-2018		Pending
Interdepartmental Hurricane Conference in Orlando, FL. Semiannual project meeting between PIs and NHC/JHT facilitators.	3-15-2018		Both PIs and grad student registered and travel plans completed.
2018 AMS Conference on Hurricanes & Tropical Meteorology, Ponte Vedra, FL.	4-20-2018		Both PIs and grad student registered and travel plans completed. Abstract submitted.
Year one JHT evaluation.	Sept 2018		Pending
One-week PI visit to NHC/JHT for semiannual project meeting and EP method training activities.	Sept 2018		Pending
Complete evaluation of 2018 real-time JHT forecasts.	1-1-2019		Pending
With JHT collaborators, finalize year two testbed evaluation plan.	3-1-2019		Pending
Interdepartmental Hurricane Conference, location TBD, and semiannual project meeting between PIs and NHC/JHT facilitators.	Spring 2019		Pending
Year two JHT evaluation.	Sept 2019		Pending
One-week PI visit to NHC/JHT for semiannual project meeting.	Sept 2019		Pending
Complete evaluation of 2019 real-time JHT forecasts. Final report	1-1-2020		Pending

Table 1: Planned versus actual project accomplishments

With funding from the grant, we are supporting one masters-level research assistant, who will be presenting some of the initial findings (from the training) at the 2018 AMS conference on Hurricanes and Tropical Meteorology in April 2018. In addition to this presentation, we will be reporting on our activities at the Interdepartmental Hurricane Conference in Orlando, FL in March 2018. Further dissemination will occur as summarized in Table 1.

As noted, the project is on-schedule and the activities as laid out in the proposal and listed in Table 1 will impose a clear and robust timeline for continued progress towards the project goals and objectives.

2. PRODUCTS

As shown in Table 1, the project test plan has been produced. Additionally, a research-to-operations transition plan and this report have been produced. Further deliverables are pending as per the schedule listed in Table 1.

3. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

At the University of Wisconsin at Milwaukee, the following individuals have worked on this project: PI Paul J. Roebber, Distinguished Professor of Atmospheric Sciences; co-PI Clark Evans, Associate Professor of Atmospheric Sciences; Jesse Schaffer, M.S. candidate, Atmospheric Sciences. No changes have been needed or made to the personnel identified in the proposal. Testbed point of contacts at the National Hurricane Center include Michael Brennan, John Cangialosi, and Christopher Landsea, and additional NOAA input has been provided by Mark DeMaria, Jose Salazar, and Matt Onderlinde.

4. IMPACT

It is too early in the project to provide information on this point at this stage.

5. CHANGES/PROBLEMS

There have been no significant problems or changes to the proposed research at this stage.

6. SPECIAL REPORTING REQUIREMENTS

For the Joint Hurricane Testbed, we have the following special reporting requirements:

- Assessment of the project's Readiness Level (current and at the start of project): at project start we were at RL5, we have established RL6 (demonstration of prototype system, subsystem, process, product, service or tool in relevant or test); and we are in the process of moving toward the primary test goal RL7: successful implementation of the models' codes into the JHT/NHC infrastructure for real-time model demonstration.
- Test Plan: we have been in discussion of a draft with NOAA POCs and are submitting the Test plan for the 1-30-2018 deadline.
- Transition to operations activities in the last six months: we are submitting the Research-to-Operations Transition plan for the 1-30-2018 deadline.
- Summary of testbed-related collaborations, activities, and outcomes: these are pending (Spring 2018)
- Has the project been approved for testbed testing yet (if it's a testbed project)? Yes.
- What was transitioned to NOAA? In progress over the course of the project.

7. BUDGETARY INFORMATION

There are not budgetary anomalies or deviations from the original proposal.

8. PROJECT OUTCOMES

It is too early in the project to provide information on this point at this stage.

Test Plan for NOAA/OAR/OWAQ Testbed Project
NOAA grant NA17OAR4590137

Evolutionary Programming for Probabilistic Tropical Cyclone Intensity Forecasts

PRINCIPAL INVESTIGATORS: Dr. Paul J. Roebber, University of Wisconsin - Milwaukee
Dr. Clark Evans, University of Wisconsin - Milwaukee

Submitted: January 2018

I. What major **concepts/techniques** will be tested? What is the scope of testing (what will be tested, what won't be tested)?

Roebber (2010, 2013, 2015abc) has developed a form of "evolutionary programming" (EP; Fogel 1999), suitable for producing large member ensemble weather forecasts. The collections of algorithms so-produced, together with Bayesian model combination (BMC; Monteith et al. 2011; Roebber 2015a), result in forecast probability distribution functions (PDFs) superior in probabilistic and deterministic skill to traditional numerical weather prediction (NWP) model ensembles, particularly at the tails of the distribution (Roebber 2015ab). For tropical cyclone intensity, these include rapid intensification and rapid weakening, which are particular foci of the JHT program (e.g., Rappaport et al. 2012).

Briefly, the model is developed using SHIPS (e.g., DeMaria et al. 2005) developmental data between 2000-2016 for the Atlantic and East/Central Pacific basins (using separate models for each basin). Model outputs to be tested include (a) consensus deterministic TC intensity (every 12 h out to 120 h) and (b) probabilistic RI forecasts at the Kaplan et al. (2015) thresholds. We expect that testbed evaluation and forecaster interaction will make clear the most appropriate forecast quantities and their formats. For RI, our focus will be on the formal threshold of 30 knots at 24 hours, but we will also explore other thresholds at other lead times: 20 knots at 12 hours, 45 knots at 36 hours, 55 knots at 48 hours, and 65 knots at 72 hours. Real-time evaluation for 2018-2019 will primarily occur in the JHT environment, with a backup version running in real-time at UWM.

II. **How** will they be tested? What **tasks** (processes and procedures) and activities will be performed, what preparatory work has to happen to make it ready for NOAA testing, and what will occur during the experimental testing in the testbed?

The newly developed models described above will be tested in real-time during the 2018 and 2019 East Pacific and North Atlantic hurricane seasons. The models rely on SHIPS diagnostic data as input; consequently, task I is that the models need to be interfaced with the real-time SHIPS diagnostic data feed for testing purposes. This may require minor revisions to the model codes, which used development rather than real-time data in their formulation.

During testing, we expect that the models will run in parallel with the operational SHIPS model. Output will be provided in text and ATCF formats. The models' code will need minor revisions to accommodate these specific output formats. Verification and evaluation will occur after each season (December-May) outside of the testbed confines. Archiving the outputs described above will be necessary to permit these activities.

III. **When** will it be tested in coordination with the NOAA testbed? What are **schedules and milestones** for all tasks described in section II that need to occur leading up to testing, during testing, and after testing?

We expect the code revisions described above as being necessary for testbed evaluation to be completed by 1 April 2018. We propose to complete testing during the 2018 and 2019 East Pacific and Atlantic hurricane seasons, roughly comprising the 15 May-30 November period of each year. Verification and evaluation activities will occur during the December-May period following each tropical season.

IV. **Where** will it be tested? Will it be done at the PI location or at a NOAA testbed location?

We propose to complete testing on the NHC/JHT hardware infrastructure.

V. Who are the key **stakeholders** involved in testbed testing (PIs, testbed support staff, testbed manager, forecasters, etc.)? Briefly what are their **roles and responsibilities**?

There are several key stakeholders involved in testbed testing:

- PIs + M.S. student: prepare model codes for implementation on JHT infrastructure; revise model codes between years one and two of testbed evaluation if necessary; complete post-season verification and evaluation activities; provide routine reports to JHT and NHC personnel regarding model performance characteristics; provide training as needed on model formulation.
- JHT support staff: facilitate implementation and execution of model codes on JHT infrastructure.
- NHC forecasters: subjectively evaluate model results (as time and resources permit); provide feedback on forecast quality and format(s).

VI. What **testing resources** will be needed from each of the above participants (hardware, software, data flow, internet connectivity, office space, video teleconferencing, etc.), and who will provide them?

The JHT IT facilitator will be provided with our Fortran code to be implemented on the JHT infrastructure. The JHT IT staff will use the existing JHT workstation and NHC data flow for the real-time testing at NHC. We will also run the EP and BMC locally at UW-Milwaukee as a backup measure – resources sufficient for this purpose are already in place

VII. What are the **test goals, performance measures, and success criteria** that will need to be achieved at the end of testing to measure and demonstrate success to advance to higher Readiness Levels and to proceed to full transition to NOAA operations (Readiness Level 8)?

The primary test goal is that associated with NOAA R2O Readiness Level 7: successful implementation of the models' codes into the JHT/NHC infrastructure for real-time model demonstration. To measure and demonstrate success, we propose the following performance measure and accompanying success criterion:

- **Performance Measure**: Deterministic intensity forecast skill relative to Decay-SHIFOR5 (as a no-skill model) and the simple IVCN consensus model.
- **Success Criteria**: Forecast improvement of at least 10-25% (Atlantic) or 10-20% (East Pacific) relative to Decay-SHIFOR5, with the higher thresholds applicable at later forecast times in each basin except 120 h for the East Pacific, and forecast skill equal to or greater than that of the IVCN model at all lead times for each basin.

The Decay-SHIFOR5-based performance measure is chosen as it is the standard measure by which NHC expresses forecast skill for official and model-based intensity forecasts. The Decay-SHIFOR5-based success criterion is chosen relative to recent seasons' intensity forecast model performance (e.g., Figs. 11 and 23 of the NHC 2016 Hurricane Season Verification Report; http://www.nhc.noaa.gov/verification/pdfs/Verification_2016.pdf), such that the newly developed models must demonstrate at least comparable skill to that of existing intensity guidance in order to be deemed successful (or skillful). The IVCN-based performance measure are success criterion are chosen following JHT facilitator feedback and represent a stricter evaluation with the same aims as the Decay-SHIFOR5-based measure.

The metric to be used for probabilistic performance will be the Brier Skill Score (BSS) as per Kaplan et al. (2015) and Rozoff et al. (2015). We will document the BSS performance relative to the levels outlined in those studies.

VIII. How will testing **results** be documented? Describe what information will be included in the **test results final report**.

Testing results will be documented in a post-2018-season report, a post-project report, and a manuscript to be submitted for peer-review and eventual publication in either the *Monthly Weather Review* or *Weather and Forecasting* journals of the American Meteorological Society. In the test results final report, the same sort of information that is found in annual NHC verification reports will be conveyed: mean absolute intensity error, intensity bias, and skill

relative to Decay-SHIFOR5 and IVCN, both aggregated over all storms (separate for each basin) as well as for each individual storm.