7B.5

USING EVOLUTIONARY PROGRAMMING TO GENERATE IMPROVED TROPICAL CYCLONE INTENSITY FORECASTS

Jesse Schaffer, Paul J. Roebber, and Clark Evans *
University of Wisconsin-Milwaukee, Milwaukee, Wisconsin

1. INTRODUCTION

The difficulty of forecasting tropical cyclone (TC) intensity has led to the development of many different TC forecasting models, which can be dynamical, statistical, or a combination thereof. While these models have made notable advances in forecasting track, improvements in forecasting intensity have been comparatively lacking (DeMaria et al. 2014). A significant contribution to intensity errors is the difficulty of predicting times of rapid intensification and rapid weakening. The goal of this research is to use evolutionary programming (EP; Fogel 1999; Roebber 2015) to construct improved forecasts for TC intensity through a lead-time of 120 h for both the Atlantic and East Pacific basins. Previously, EP generated ensembles have been tested on 156-h minimum temperature forecasts as well as 500-hPa height forecasts and been shown to outperform traditional dynamical model ensembles and multiple linear regression ensembles (Roebber 2013). This improvement is a result of an EP architecture that allows for both flexible and adaptive forecasts.

2. METHODS

a, Overview

Data for training, cross-validation, and testing came from the Statistical Hurricane Intensity Prediction Model (SHIPS; DeMaria and Kaplan 1994) developmental database for all TCs from the respective basin for the years 2000-2016. TCs from this period were separated into three categories based on the storm's maximum achieved intensity: tropical storms, weak hurricanes (category 1 or 2), and major hurricanes (category 3, 4, or 5). These three categories were then evenly divided into training, cross-validation, and testing data so that no category was biased toward storms of a certain intensity.

Corresponding author address: Jesse Schaffer, Department of Mathematical Sciences, University of Wisconsin-Milwaukee, 3200 North Cramer Ave., Milwaukee, WI 53211
Email: schaff26@uwm.edu

The basic structure of EP employed consisted of a population of 10,000 algorithms. The 10,000 algorithms were then divided into five tribes where each tribe used a random set of thirteen out of a total of 46 input variables. Additionally, all tribes had available a constant value of 10 to use. Each variable was converted into a standardized anomaly so they could be functionally compared and thus 10 provided a value outside the range of anomalies. The algorithms were initialized with random coefficients and operators as well as random selections from the thirteen assigned variables and the constant. Then the 100 best algorithms based on cross-validation performance were saved.

Then began the process of producing the next generation of algorithms. The 400 worst performing algorithms in each tribe (based on training performance) were removed. These equations were replaced with clones of the best performing algorithms in the tribe. However, during cloning a single mutation occurs and a variable, operator, or coefficient is altered in one of the lines. This gives a new population of algorithms, which were again sorted based on their performance on the cross-validation data. Any algorithm that outperformed a member on the list of 100 best equations joined the list and the worst algorithm was removed. This process was repeated for 300 iterations. After this point more iterations could be done, but little skill is added (not shown). Thus instead the whole process is run through five more times with the entire population being re-initialized, though the list of 100 best algorithms is retained. After this, Bayesian Model Combination (BMC; Monteith et al. 2011) is performed on five algorithms to obtain further refinement. These five algorithms were chosen based on relatively low root-mean square error and relatively high root-mean square difference (relative to other algorithms) so that the five algorithms are both skillful and diverse.

b. Algorithm Structure

Figure 1 shows an example of an algorithm. Each line features a conditional followed by an adjustment to be made to a persistence forecast.

If no conditional is met then no adjustment is made and it becomes a persistence forecast. The first line is red to indicate that the condition will always be met, as RHCN will always equal RHCN. Thus intensity change over the previous 6-h (INCV), radially averaged 200-hPa zonal wind (U20C), and the maximum potential intensity (VMPI; Emanuel 1988) are all used to calculate an adjustment to persistence. While looking at the second recall that each variable inputted into the EP algorithm was converted to a standardized anomaly and 10 was specifically chosen to lie outside this range. In this way T150 will always be less than 10 and the condition will always be met.

Contrasting the red lines the conditionals in black won't always be true. Looking at the third line, the adjustment only occurs if the depth of the 26˚C thermocline (RD26) is greater than the radially averaged 850-200 hPa vortex removed shear (SHDC). Again, the variables are converted to standardized anomaly, thus the conditional is more accurately read as, if the anomaly in RD26 is greater than the anomaly in SHDC, then perform the following adjustment. The adjustments from each line are added together to provide a forecast given by that algorithm.

3. RESULTS

Figure 2 shows the skill of the model on the testing data for the Atlantic Basin. Over all lead times the five best algorithms combined using BMC (BMC5) see a 13-21% improvement over the decay version of the Statistical Hurricane Intensity Forecast Model (Decay-SHIFOR) with the single-best algorithm being equally as skillful as the decay-SHIFOR model. However, the official forecast is still best with anywhere from 1-26% improvement over BMC5. In the East Pacific the same general pattern from the Atlantic holds (Figure 3). The best individual algorithm is as skillful as decay-SHIFOR. BMC5 is 5-17% better than the decay-SHIFOR and official forecasts still performs best, with 10-30% better forecasts over BMC5.

4. CONCLUSION/FUTURE WORK

The EP model has shown considerable improvement in skill over the Decay-SHIFOR model. While it doesn't outperform official forecasts the implimentation of this improved model would likely lead to improved official forecasts. Future work involves examining testing data to identify cases where the model succesfully forecasted TC intensity and where it struggled which might highlight ways to further refine the model. Additionally, real-time forecasting and testing can be performed during the 2018 TC season. Lastly, the model will be developed to give probabilistic forecasts for rapid intensification and rapid weakning.

REFERENCES

Emanuel, K. E., 1998: The maximum intensity of hurricanes. *J. Atmos. Sci.* **45,** 1143-1155.

DeMaria, and M., J. Kaplan, 1994: A Statistical Hurricane Intensity Predictions Scheme (SHIPS) for the Atlantic Basin. *Wea. Forecasting,* **9**, 209-220

—, M., C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398

Fogel, L.J., 1999: Intelligence through Simulated Evolution : Forty Years of Evolutionary Programming, John Wiley. 162 pp.

Monteith, K., J. Carroll, K. Seppi, and T. Martinez, 2011: Turning Bayesian model averaging into Bayesian model combination. *Proc. Int. Joint Conf. on Neural Networks (IJCNN'11)* San Jose, CA, IEEE, 2657-2663

Roebber, P. J., 2013: Using evolutionary programming to generate skillful extreme value probabilistic forecasts. *Mon. Wea. Rev.*, **141,** 3170-3185

—, 2015: Evolving Ensembles. *Mon. Wea. Rev.*, **143**, 471-490

FIGURES

1	IF	RHMD <= RHMD	THEN	0.29 * INCV	+	-0.51 * LON	*	0.18 * RSST
2	IF	T150 <= 10	THEN	0.29 * MSLP	+	-0.95 * 10	*	-0.01 * V850
3	IF	RD26 > SHDC	THEN	-0.13 * Z000	*	0.71 * RHCN	+	-0.3 * SHDC
4	IF	D200 > IR00_2	THEN	0.11 * RD20	+	0.02 * 10	*	-0.61 * MTPW_2
5	IF	RHMD > IR00_2	THEN	0.14 * SDDC	+	-0.12 * PSLV_5	*	0.39 * T200

Figure 1: Algorithm structure with five if-then statements followed by an adjustment. Red highlights lines where the if-statement is always true and thus the following adjustment will always be performed.

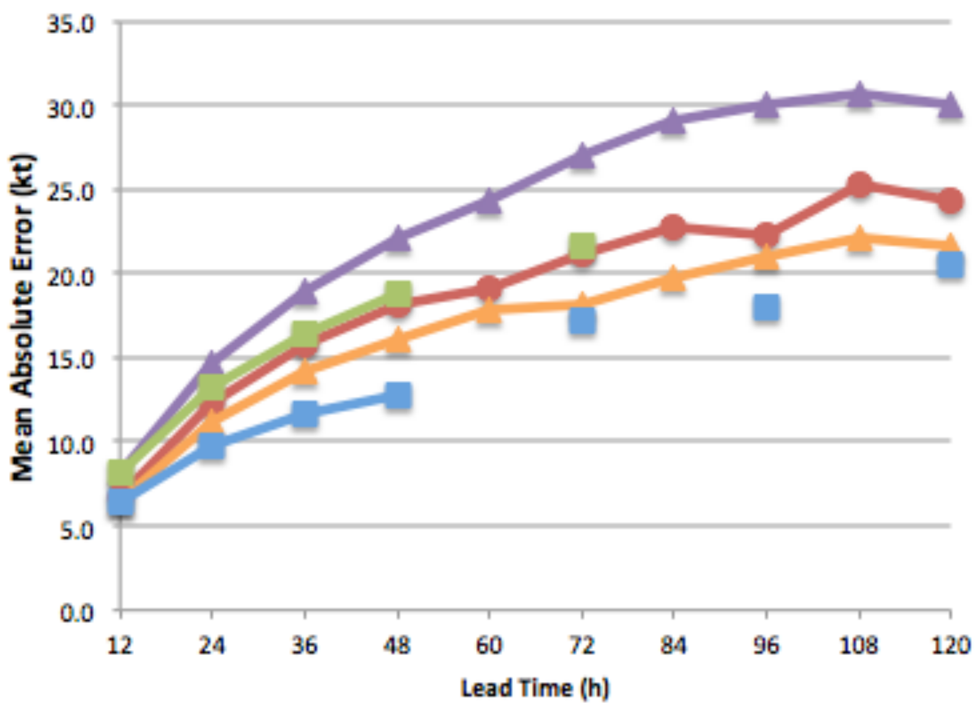

Figure 2: Skill (as determined by mean absolute error) of the Official Forecast (blue), Statistical Hurricane Intensity Forecast Model (green), persistence forecast (purple), Bayesian Model Combination of the best 5 algorithms (orange), and best individual algorithm (Red) based on testing data for the Atlantic Basin.

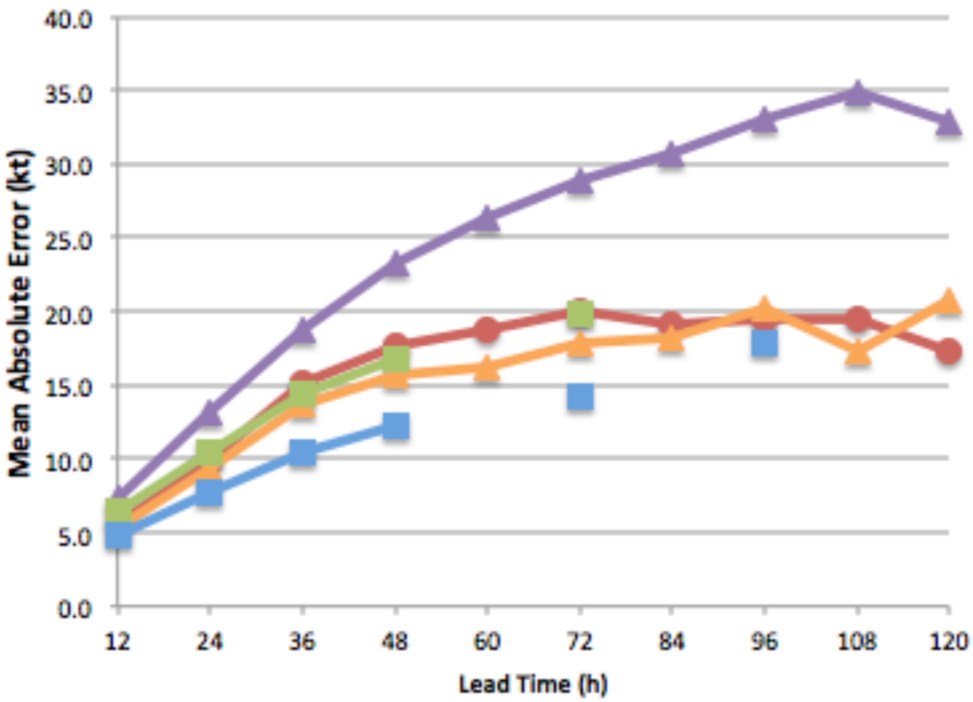

Figure 3: Same as for Figure 2 except for the East Pacific Basin.