Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 61 (2015) 422 – 427

Procedia Computer Science

Complex Adaptive Systems, Publication 5
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2015-San Jose, CA

Diagnosing Tropical Cyclone Rapid Intensification Using Kernel Methods and Reanalysis Datasets

Andrew Mercer[a]*, Alexandria Grimes[a]

[a]*Department of Geosciences, Mississippi State University, Mississippi State, MS 39762, USA*

Abstract

Tropical cyclone rapid intensification (RI) continues to be a problem that eludes operational forecasters. Recent work in this area has revealed the value of applying machine learning techniques to classifying storms as RI or non-RI at 24-hours lead time. However, that work showed that differing reanalysis datasets represented the storms in unique ways, offering different discrimination capability and unique predictor sets that are important for RI. The scope of this research is to identify factors important for RI that are consistent among three reanalysis datasets, as these are likely the fields that will provide the greatest discrimination capability. An S-mode rotated principal component analysis was used to formulate unique patterns within RI and non-RI storms, and the resulting RPC scores were used to train a support vector machine classification algorithm that yielded binary RI occurrence output. Base-state meteorological variables (geopotential height, temperature, u and v wind components, vertical velocity, and relative humidity) at single horizontal levels were tested individually as predictors for the SVM. Base-state fields that were consistently good at discriminating RI events from non-RI events among all three reanalysis datasets were deemed most useful for RI classification and will be considered for future forecast applications.

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology

Keywords: Tropical cyclones; support vector machines; kernel principal component analysis; kernel methods; reanalysis datasets

1. Introduction

Tropical cyclone (hereafter referred to as TC) forecasting continues to be a challenge in operational meteorology. TC forecasts consist of two primary forecast modes, track forecasts and intensity forecasts. Track forecasts are improving with continued updates to operational weather forecast models and the advent of the satellite era, yet

* Corresponding author. Tel.: 1-662-325-3915; fax: 1-662-325-9423.
E-mail address: a.mercer@geosci.msstate.edu

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology
doi:10.1016/j.procs.2015.09.179

intensity forecasts continue to demonstrate limited effectiveness, primarily due to the limited understanding of the mechanisms driving TC intensification and the inability of weather models to render the important underlying thermodynamic processes. Forecasts of rapid intensification (hereafter referred to as RI), which involve a rapid strengthening of the cyclone over a short period of time (typically 24-hours), are of particular concern since most moderately strong (Category 2-3) TCs and all Category 4 and 5 storms undergo RI at some point in their life cycle [1]. Hence, RI forecasting remains a major forecasting challenge.

Initial attempts at forecasting RI in a TC's life cycle have primarily consisted of statistical approaches (i.e. multivariate linear regression [2]). These methods require prior knowledge of the variables relevant for RI, which typically consist of thermodynamic and kinematic fields. In particular, horizontal wind fields are shown to be critical in several kinematic studies on RI [3-5], while other studies reveal the importance of thermodynamic fields (mainly moisture and temperature or equivalent potential temperature) for RI [4, 6, 7]. Blends of kinematic and thermodynamic quantities are utilized as predictors in the most current operational RI classification model, the Statistical Hurricane Intensification Prediction Scheme Rapid Intensification Index (SHIPS-RII [1]). This model utilizes several kinematic variables (vertical wind shear, upper level divergence, potential intensity, and initial maximum sustained wind) and thermodynamic variables (low-level relative humidity, ocean heat content, total precipitable water, inner-core dry-air), as well as a predictor dealing with persistence (previous 12-hour intensification). It is evident from the literature on this topic that while a blend of kinematic and thermodynamic variables is useful, the exact physical mechanisms driving RI are still not well understood.

Recently, dynamic modeling of the TC environment has improved with the implementation of the Hurricane Weather and Research Forecasting (HWRF) model [8]. Recent work [9] showed vast improvements in HWRF's rendering of Hurricane Earl (2010) over previously utilized dynamic models. Eventually, the forecast skill of dynamic models such as HWRF will surpass statistical approaches, but until that time, forecast skill still remains sufficiently low to continue consideration of statistical models. Further, no machine learning techniques have been considered for the RI forecast problem, with the exception of [10], which utilized a single reanalysis dataset.

It is evident from previous work that without a dynamic rendering of the RI process, an optimal suite of predictors useful for discriminating RI/non-RI events is essential. The goal of this project is another look at the predictors useful for RI determination using kernel methods. In particular, three reanalysis datasets' [11-13] renderings of RI and non-RI storms will be formulated on individual base-state meteorological variables with the goal of identifying the variables with the greatest discrimination skill. Section 2 briefly describes the reanalysis datasets, while section 3 discusses the methods employed to discriminate RI and non-RI results. Section 4 shows results, while section 5 provides some discussion and conclusions.

2. Data

Since large-scale predictors were of interest for this work, reanalysis datasets, which consist of gridded three-dimensional large-scale representations of the atmosphere, were used. Three reanalysis datasets were tested, the NCEP/NCAR reanalysis dataset (NNRP – 11), the NCEP-DOE Reanalysis II dataset (DOE – 12) and the 20th century Regional Reanalysis dataset (20th – 13). Details of each dataset are provided in Table 1. Since each reanalysis dataset was formulated using a unique underlying dynamic modeling system, subtle but often important differences in their renderings of the atmosphere exist. As such, any "best classifying predictors" consistent among all reanalysis datasets would likely suggest important variables for discriminating RI and non-RI events. Six base-state meteorological fields for all vertical levels up to 100 mb were retained from each reanalysis dataset: geopotential height, temperature, *u* and *v* wind components, vertical velocity, and relative humidity. Each of these fields has high reliability in each reanalysis dataset, as they are primarily based upon observational data instead of model parameterizations, making them useful for the study.

Table 1. Summary of selected reanalysis datasets

Dataset name	Longitude-Latitude Resolution	Vertical Levels	Years available
NCEP/NCAR reanalysis (NNRP – 11)	2.5° x 2.5° (144x73)	17	1948-present
NCEP/DOE reanalysis II (DOE – 12)	2.5° x 2.5° (144x73)	17	1948-present
20th century reanalysis (20th – 13)	2° x 2° (180x91)	24	1871-present

In addition to the required reanalysis datasets, a database of all TCs with an associated RI definition was needed. The National Hurricane Center has multiple definitions of RI that are based on 24-hour changes in peak wind gusts [1,6]. To ensure a large sample size of RIs, the weakest of the NHC's criteria (a 25 kt increase of maximum wind speed in 24 hours) was selected, yielding a final case set of 158 RI events and 146 non-RI events. Storm locations and intensities were obtained directly from the NHC Atlantic Hurricane database for 1985 to 2009. All reanalysis fields pulled from the datasets listed in Table 1 were centered on the gridpoint nearest the NHC-defined hurricane center associated with the lowest mean sea level pressure of the grid (to ensure the fields were centered on the reanalysis solution of the storm center, not the NHC solution). All reanalysis grids were retained at 24 hours prior to RI (that is, at the beginning of the RI process). In the case of multiple occurrences of RI within a storm, the first time was used as the valid time, and for non-RI storms, the point of greatest 24 hour wind speed change was deemed the valid time (again with the first time being used in the case of ties).

3. Methodology

As stated previously, all vertical levels up to 100 mb were retained for the six base-state meteorological variables. However, the NNRP and DOE reanalysis datasets only provide humidity information up to 300 mb. In total, 68 grids were tested for the DOE and NNRP datasets, while 114 grids were tested for the 20th century reanalysis, owing to its different vertical resolution. Each grid consisted of a subset of datapoints centered on the gridpoint with the lowest mean sea level pressure. For the NNRP and DOE, each grid was 11 x 11 (121 total points), while for the 20th century reanalysis, each grid was 15 x 15 (225 gridpoints), owing to the different horizontal spatial resolution of the datasets (see Table 1).

Since each dataset had multiple gridpoints per variable and per vertical level, each of which were tested individually, a data reduction method was implemented to reduce the high correlation among the gridpoints within each horizontal field. The utility of rotated principal component analysis (RPCA) for data reduction in statistical modeling has been seen in many studies [10, 14-16], and as such, this method was selected for reduction of each gridded field. The RPCA methodology was completed by:

1. Extraction of the horizontal grid of interest, and scaling of the grid through the computation of standard anomalies. Gridpoints were scaled to a mean of 0 and standard deviation of 1 for all times at that point.
2. Formulation of the correlation matrix on the gridpoint dimension (S-mode [17]), and eigenanalysis of the correlation matrix
3. Truncation of the eigenvectors through North's test [18]
4. Calculation of the Varimax-rotated RPC loading matrix [17] from the reduced eigenvector matrix
5. Calculation of the RPC score matrix (same dimensionality as the number of cases).

The resulting RPC score values represent a relative match of the actual spatial field observed for that particular storm to each of the RPC loadings derived in step 4 of the RPCA methodology described above. These RPCs were used as predictors in the RI/non-RI classification model. Figures 1 and 2 below show the number of RPCs retained for each of the grids for the three reanalysis datasets (Fig. 1) and the variance explained by each configuration (Fig. 2). As has been demonstrated in numerous previous studies [10, 14-16], support vector machine (SVM [19]) classification schemes provide superior classification performance to more traditional logistic or linear regression approaches. As such, a SVM was used as the classification scheme for this study. RPC scores were used as inputs into the SVM, with a single binary RI/non-RI output selected as the final outcome. However, SVMs have multiple configuration options (see the appendix in [15] for more details on the SVM algorithm), including a cost function, which determines the impacts of extreme classifiers, the kernel function, which is responsible for mapping data into a nonlinear hyperspace on which the decision hyperplane is constructed, and all associated parameters within the kernel functions (polynomial degrees and values of γ within the radial basis kernel function). Theoretically, there are an infinite number of possible configurations, but a subset of these configurations was tested in this study. In particular, polynomial kernels with degrees of 2, 3, 4, 5, and ½ were considered, and radial basis functions with values of $\gamma = 0.05$, 0.01, and 0.1 were tested (8 total kernel function configurations). Additionally, four cost function values between 1 and 1000 in $\log_{10}$ units were tested, yielding a total of 32 possible SVM configurations considered.

To establish the "best" configuration, the cases were pairwise bootstrap-resampled 500 times with 90% of the cases retained for training and the remaining 10% for independent testing. This approach allowed for the formulation of bootstrap confidence intervals on Heidke skill scores (HSS [15]) for each of the 32 possible kernel/cost configurations for each grid. The kernel/cost configuration that yielded the highest median HSS was deemed the best of those tested and was retained for that diagnostic variable and vertical level. A logistic regression was also conducted for each grid as a baseline. Generally speaking, most temperature and vertical velocity grids required high costs to get the best results, while the geopotential height fields (which were often the best classifying fields as seen below) frequently showed that simple logistic regression outperformed the SVM. Polynomial kernels of varying degrees tended to be the most common "best" configuration in the *u* and *v* wind classification fields, and no real patterns in the relative humidity SVM configurations emerged. The resulting best kernel/cost configurations for each diagnostic variable grid were retained to compare the performance of each variable. These results were further compared among the three reanalysis datasets to find consistently good classifying diagnostic variables.

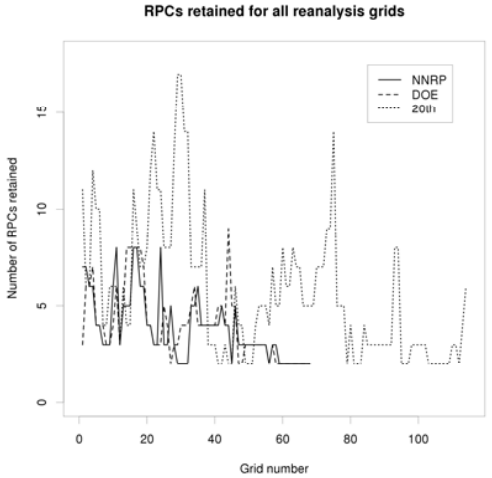

Fig. 1. Number of RPCs retained for all grids

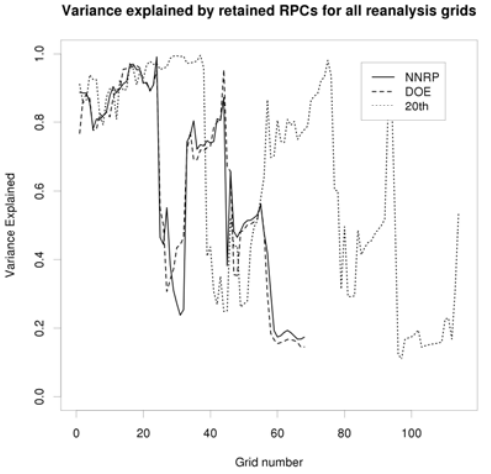

Fig. 2. Variance explained by RPCs for each grid

4. Results

After completion of the SVM step, the resulting best kernel/cost combination was retained for each of the grids for each reanalysis dataset. The quality of the results was assessed using 95% bootstrap confidence intervals on the 500 resampled training/testing sets. The NNRP and DOE (Figs. 3 and 4) yielded surprisingly consistent results, suggesting that mid and upper level temperature fields, near surface and upper level geopotential height, and near surface relative humidity were the best classifiers of RI/non-RI. Interestingly, no consistent result between these two reanalysis datasets with regards to the kinematic fields (either *u* or *v* wind components or vertical velocity) was observed. These results support the work of others [4,6,7] that suggests that a good rendering of the low and upper level thermodynamics is critical for identifying RI/non-RI. Additionally, among the top 10 ranked predictors for these two reanalysis datasets, no consistent winner is able to be identified, and skill scores remain quite poor, with medians near 0.2 for all top 10 ranked sets. Of particular concern are the lower confidence levels in all 68 predictors for both reanalysis datasets, which overlap the 0 skill line. This is particularly problematic since 0 skill suggests that simple climatological forecasts may be better classifiers of RI/non-RI than the diagnostic variable in some instances.

The 20th century reanalysis (Fig. 5) results showed somewhat inconsistent results with the NNRP and DOE, as five of the top 10 performing grids were kinematic fields. The zonal (east-west) component of the winds were good classifiers in the 20th century reanalysis, possibly due to the increased horizontal resolution in the dataset. Additionally, 200 mb vertical velocity yielded a lower confidence limit that was greater than 0, an improvement over the NNRP and DOE results. The major differences between the reanalysis datasets arise more due to the horizontal resolution than the model underlying the formulation of each dataset, owing to the consistency of the results between the DOE and the NNRP.

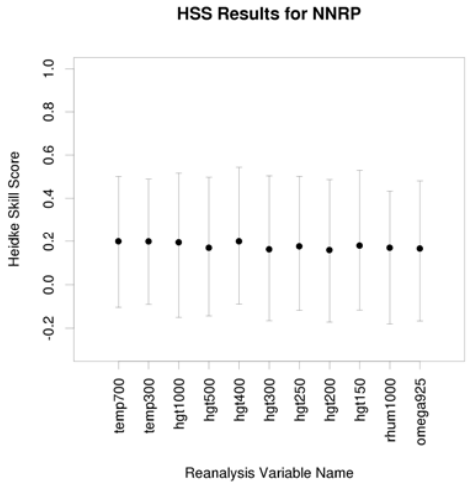

Fig. 3. All SVM results ranked in top 10 of median HSS for the NNRP. The variable name and vertical level are indicated under each plot.

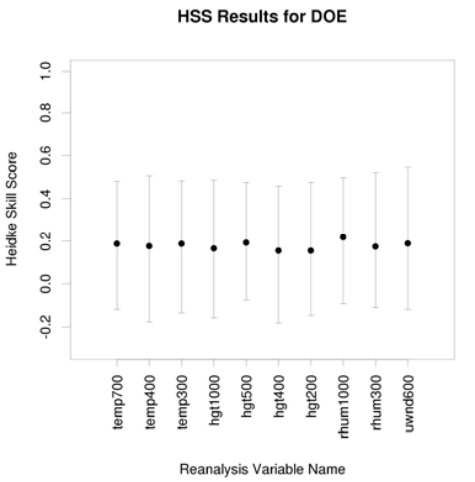

Fig. 4. Same as Fig. 3, but for DOE.

In comparing the top ranked predictor for each of the three datasets (Fig. 6), it is clear that the 20th century reanalysis is yielding improved results over the DOE or NNRP. The median HSS for the 20th century reanalysis is roughly 30% improved over the other two reanalysis datasets (0.29 for 20th, 0.22 for NNRP and DOE). The spread of the HSS results is consistent throughout the three best performing datasets as well (standard deviation of the HSS bootstrap replicates of 0.161 for the NNRP, 0.163 for the DOE, and 0.164 for the 20th century reanalysis). While none of the observed differences are statistically significantly different, the upward shift in the replicate distribution for the 20th century reanalysis, and the consistency in the replicate spread, suggests better SVM classification performance when using the 20th century reanalysis as a baseline. Interestingly, all three datasets had the highest median HSS with a lower-level thermodynamic variable (700 mb temperature for the NNRP, 1000 mb relative humidity for the DOE, and 800 mb geopotential height for the 20th century reanalysis).

5. Discussion/Conclusions

It is clear based on the plethora of previous work on the topic that forecasting TC RI remains a major challenge. Since dynamic model renderings of TC intensification processes remain inadequate, further refinement of statistical approaches is necessary to ensure the best possible forecast quality. In this study, kernel methods were utilized to formulate a SVM classification scheme on individual base-state diagnostic variables, including winds, temperature, moisture, and geopotential height. The NNRP/DOE results clearly demonstrated that thermodynamic variables are more important than kinematic variables in diagnosing RI, and that more effort should be focused on improved thermodynamic renderings of RI processes. However, the 20th century reanalysis results were quite different, suggesting the kinematic fields are equally important. Regardless of this discrepancy, the best performing predictor for each dataset was a thermodynamic quantity. Unfortunately, the classification skill of the SVMs trained with these variables remained inadequate (roughly 0.22 for the best NNRP/DOE field, 0.29 for the best 20th century reanalysis), even demonstrating negative skill in some instances. The contribution of the RPCA to these differences is likely minimal due to the inconsistent results in the variance explained for each best field (85% for NNRP, 54% for DOE, and 97% for the 20th century). Regardless, this study was able to define important predictors more conclusively than previous work, which should improve predictions of RI onset in future studies.

Future work on this topic will involve implementing nonlinear kernel PCA as a data reduction technique, which is more complex due to the numerous configurations of the kernel matrix. Derived fields, such as vertical wind shear, lapse rates, divergence, and equivalent potential temperature will be considered as classifiers as well. Additionally, this work will need to be implemented using forecast data as opposed to reanalysis fields in order for it to become useful for forecast applications.

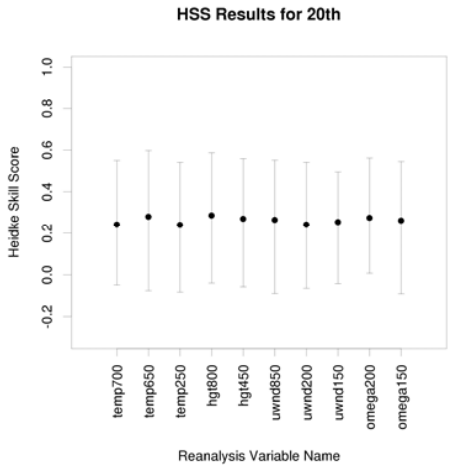

Fig. 5. Same as Fig. 3, but for 20th century reanalysis

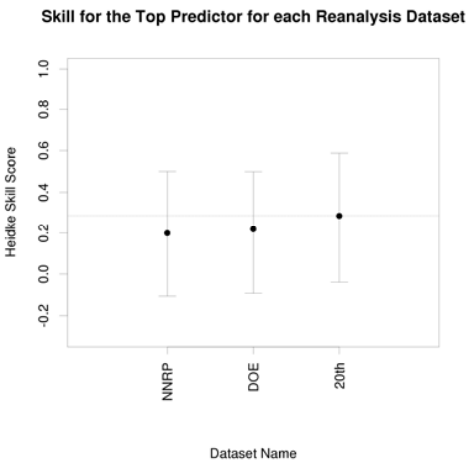

Fig. 6. Top performing predictor for each reanalysis dataset. The solid line is passing through the highest median.

Acknowledgements: This work was partially funded by NOAA-OAR #NA11OAR4320199.

References

1. Kaplan, J., and Coauthors. Improvements to the SHIPS Rapid Intensification Index. JHT Final Rep., 2013, 19 pp. [Available online at http://www.nhc.noaa.gov/jht/11-13reports/Final_Kaplan_ JHT13.pdf]
2. Kaplan, J., and DeMaria, M. Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting* 2003; **18**:1093-1108.
3. Sitkowski, M., and Barnes, G. Low-level thermodynamic, kinematic, and reflectivity fields of Hurricane Guillermo (1997) during rapid intensification. *Mon. Wea. Rev.* 2009; **137**:645-663.
4. Vigh, J., and Schubert , W. Rapid development of the tropical cyclone warm core. *J. Atmos. Sci.* 2009; **66**: 3335-3350.
5. Sitkowski, M., Kossin, J., Rozoff, C., and Knaff, J. Hurricane eyewall replacement cycle thermodynamics and the relict inner eyewall circulation. *Mon. Wea. Rev.* 2009; **140**: 4035-4045.
6. DeMaria, M., and Kaplan, J. A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting* 1993; **9**:209-220.
7. Gao, S., and Chiu, L. Surface latent heat flux and rainfall associated with rapidly intensifying tropical cyclones over the western North Pacific. *Int. J. Remote Sens.* 2010; **31**: 4699-4710.
8. Tallapragada, V., Kieu, C., Kwon, Y., Trahan, S., Liu, Q., Zhang, Z., and Kwon, I. Evaluation of storm structure from the operational HWRF during 2012 implementation. *Mon. Wea. Rev.* 2014; **142**: 4308-4325.
9. Chen, H., and Golpalakrishnan, S. A study on the asymmetric rapid intensification of Hurricane Earl (2010) using the HWRF system. *J. Atmos. Sci.* 2015; **72**: 531-550.
10. Grimes, A., and A. Mercer, Synoptic-scale precursors to tropical cyclone rapid intensification in the Atlantic Basin. *Adv. Meteorology* 2015; 814043, 16 pp.
11. Kalnay, E., and co-authors. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Soc.* 1996; **77**:437-471.
12. Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S-K, Hnilo, J.J., Fiorino, M., and Potter, G.L. NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Amer. Soc.* 2002, **83**:631-1643.
13. Compo, G. P., and co-authors. The twentieth century reanalysis project. *Quart. J. Roy. Meteor. Soc.* 2011; **137**:1-28.
14. Mercer, A. E., Richman, M., Bluestein, H., and Brown, J. Statistical modelling of downslope windstorms in Boulder, Colorado. *Wea. Forecasting* 2008; **23**:1176-1194.
15. Mercer, A. E., Shafer, C., Doswell, C., Leslie, L., and Richman, M. Objective classification of tornadic and nontornadic severe weather outbreaks. *Mon. Wea. Rev.* 2009; **137**: 4355-4368.
16. Shafer, C., Mercer, A., Leslie, L., Richman, M., and Doswell, C. Evaluation of WRF model simulations of tornadic and nontornadic outbreaks occurring in the spring and fall. *Mon. Wea. Rev.* 2010; **138**: 4098-4119.
17. Richman, M. Rotation of principal components. *J. Climatology* 1986; **6**: 293-335.
18. North, G. R., Bell, T. L., Galahan, R. F., Moeng, F. J. Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.* 1982; **17**: 699-706.
19. Cristianini, N., and Shawe-Taylor, J. *An Introduction to Support Vector Machines and other Kernel-Based Learning Methods*. Cambridge University Press, Cambridge, England, 2000, 189 pp.