J8.4 NOWCASTING OCEANIC CONVECTION FOR AVIATION USING RANDOM FOREST CLASSIFICATION

Huaqing Cai*, Cathy Kessinger, David Ahijevych, John Williams, Nancy Rehak, Daniel Megenhardt and Matthias Steiner
National Center for Atmospheric Research, Boulder, CO

Richard L. Bankert, Jeffrey Hawkins
Naval Research Laboratory, Monterey, CA

Michael F. Donovan, Earle R. Williams
Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA

1. INTRODUCTION

The traditional methodology employed within any nowcasting system usually involves an extrapolation scheme. The variations between different techniques come from different methodologies of computing a motion vector field, different advection schemes, and/or different methods of representing storms (i.e., object-based or a gridded field). The Oceanic Convection Diagnosis and Nowcasting system (Kessinger et al., 2008, 2009) has been designed to detect and predict the locations of deep convection within remote, oceanic regions and to do this at higher spatial and temporal scales than is currently available from the Aviation Weather Center (http://aviationweather.gov). Pilots, dispatchers and air traffic managers are the users for whom these products are designed.

The Oceanic Convection Diagnosis and Nowcasting system has two parts: the Convective Diagnosis Oceanic (CDO) and the Convective Nowcasting Oceanic (CNO) products. Real-time, experimental results can be seen at http://www.rap.ucar.edu/projects/ocn under the "Operations" menu. The CDO detects convection through a data fusion scheme that has three satellite-based algorithms as input and produces an interest field as output whose values range from 0-4 during the day and 0-3 during the night. See Kessinger et al. (2008, 2009) for additional details. Currently, to make 1- and 2-hr nowcasts of convection location, the CNO utilizes an object-tracker called the Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN) (Dixon and Weiner, 1993) as its extrapolation scheme with the CDO as input. The CNO output is a polygon that represents the future location of the storm. Storm growth/decay is indicated by trending of previous storm history of the storm size. The advantage of the CNO technique is its computational efficiency, particularly over large domains as are used in the oceanic regions, and its ability to capture storm growth/decay. However, the polygons used to represent storm position may not produce realistic looking storms and also tend to over-forecast the amount of storm area.

To address these limitations associated with the CNO polygon forecast, another nowcasting technique called CNO-Gridded (CNO-G) has been developed at NCAR. The CNO-G technique creates a gridded forecast of CDO interest values by using a motion vector field derived from combining TITAN motion vectors with the Global Forecasting System (GFS) numerical model steering level winds. In this technique, overall storm structure is retained but the growth/decay information associated with storm shapes is lost in the process. A very similar technique has been used for nowcasting satellite brightness temperature three hours into the future for GOES-R Algorithm Working Group's hydrology team (Cai et al., 2008). The CNO-G forecast was not computed in real-time due to the heavy computational load.

Corresponding author address: Huaqing Cai, National Center for Atmospheric Research, Boulder, CO 80307; e-mail: caihq@ucar.edu

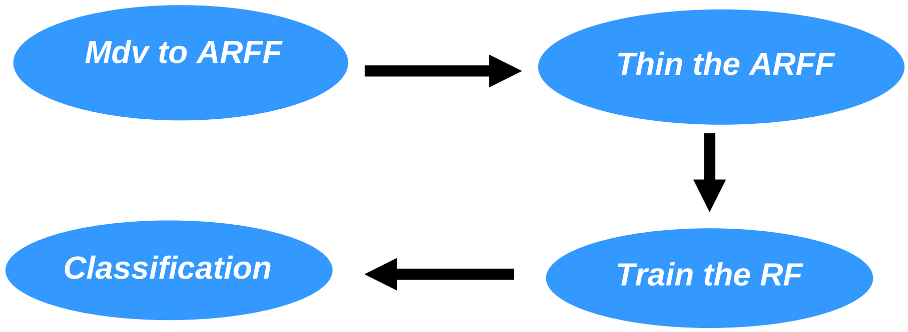


Fig.1. Flow chart for training the random forest and for classifying using the trained forest.

The third nowcasting technique, which is the primary focus of this paper, utilizes a data fusion technique called Random Forest to incorporate other data sources (e.g., NWP models, sea surface temperature, satellite soundings, and quikSCAT surface winds, etc) in addition to extrapolating the existing CDO interest field. The idea is that by incorporating information that characterizes the storm environment, something a simple extrapolation scheme can never do, storm growth/decay or even storm initiation could possibly be captured.

As a data fusion technique, Random Forest has been widely used in various scientific fields, including the development of the Federal Aviation Administration's (FAA) Consolidated Storm Prediction for Aviation (CoSPA) (Williams et al., 2008). In machine learning, a Random Forest utilizes many decision trees to ascertain the appropriate classification by taking the mode of the class as voted by each individual tree. (Breiman, 2001; Ho, 2002). The following sections will describe the procedures of running Random Forest and present some preliminary results of nowcasting oceanic convection using the Random Forest classification and compare them with the CNO and CNO-gridded forecast.

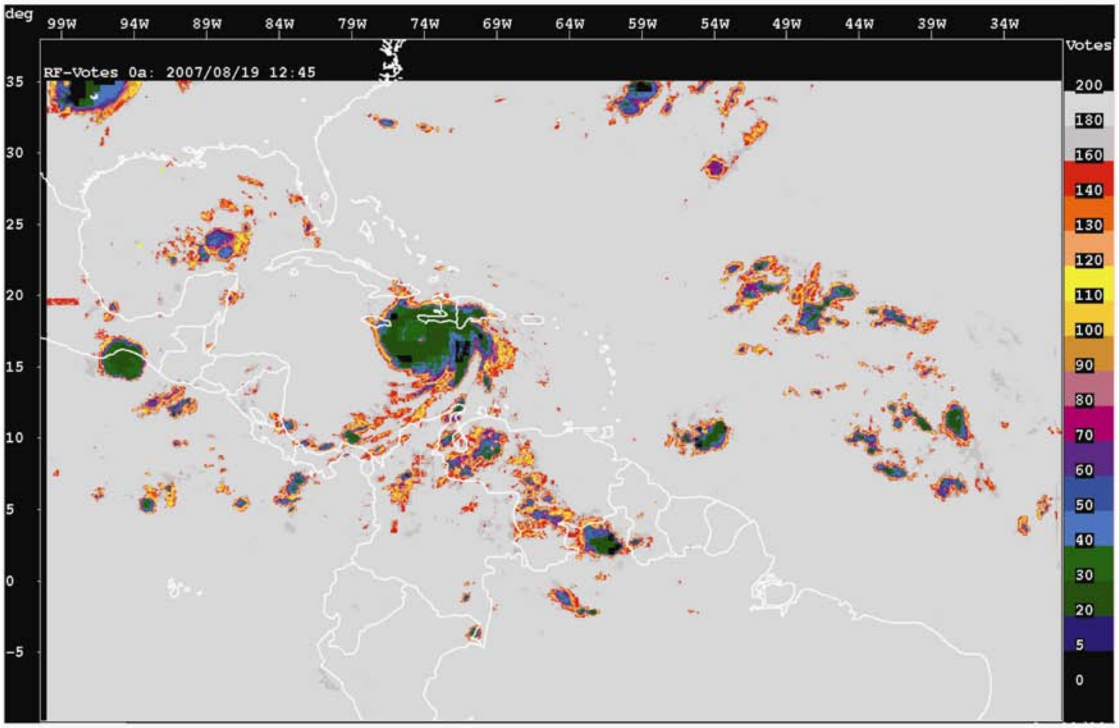
2. METHODOLOGY

The detailed methodology of creating CNO and CNO-G nowcasts of oceanic convection can be found elsewhere (Kessinger et al., 2008, 2009 and Cai et al., 2008). This section will focus on details of the Random Forest technique as applied to the CNO (hereafter referred to the CNO-RF).

A set of predictors derived from geostationary and polar-orbiting satellites and the GFS numerical model are used as input variables to the CNO-RF. The output variable (i.e., the forecast), which represents oceanic convection, contains values of the CDO interest field (Kessinger et al., 2008, 2009). The goal of nowcasting oceanic convection is thus converted to forecasting CDO intensity, which ranges in values between 0 and 4 during the day and 0 to 3 during the night, by using a set of predictors derived from satellite observations and GFS model fields.

The flow chart of training and subsequent classification using the trained forest is shown in Fig.1. As a first attempt of exploring the Random Forest technique in oceanic weather, eight days of data (12-18 August 2007 from Hurricane Dean over the Gulf of Mexico domain) are used to train a forest of 200 decision trees, while data from another four days (19-22 August 2007, also from Hurricane Dean) are used for independent classification and verification. An initial set of 17 predictors, which includes various satellite and GFS model-derived fields over the Gulf of Mexico domain, is employed in the Random Forest training/classification. All the satellite-based input predictors are advected 1-2 hr into the future using motion vectors derived by blending TITAN vectors with GFS steering level winds. Both the input predictor fields and the validation CDO interest field are converted from Meteorological Data Volume (MDV), an internal NCAR format to the Attribute-Relation File Format (ARFF) format, a format the Random Forest software can read. The ARFF

(a) Votes for CDO interest = 0



(b) Votes for CDO interest = 1

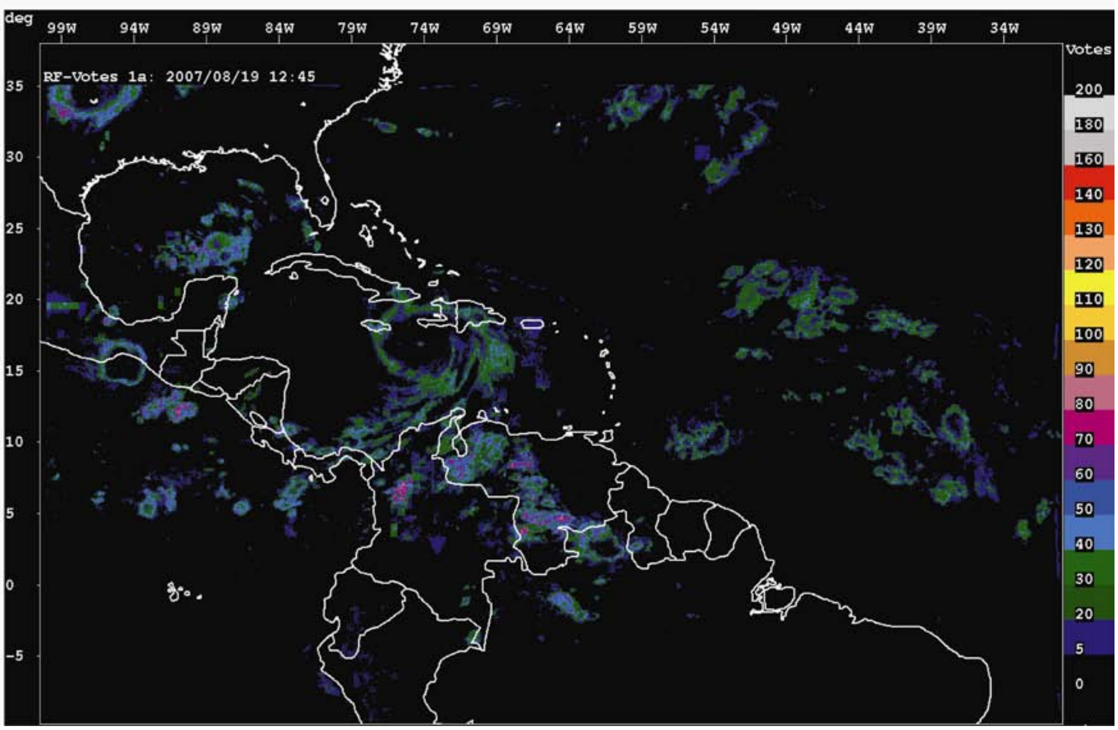
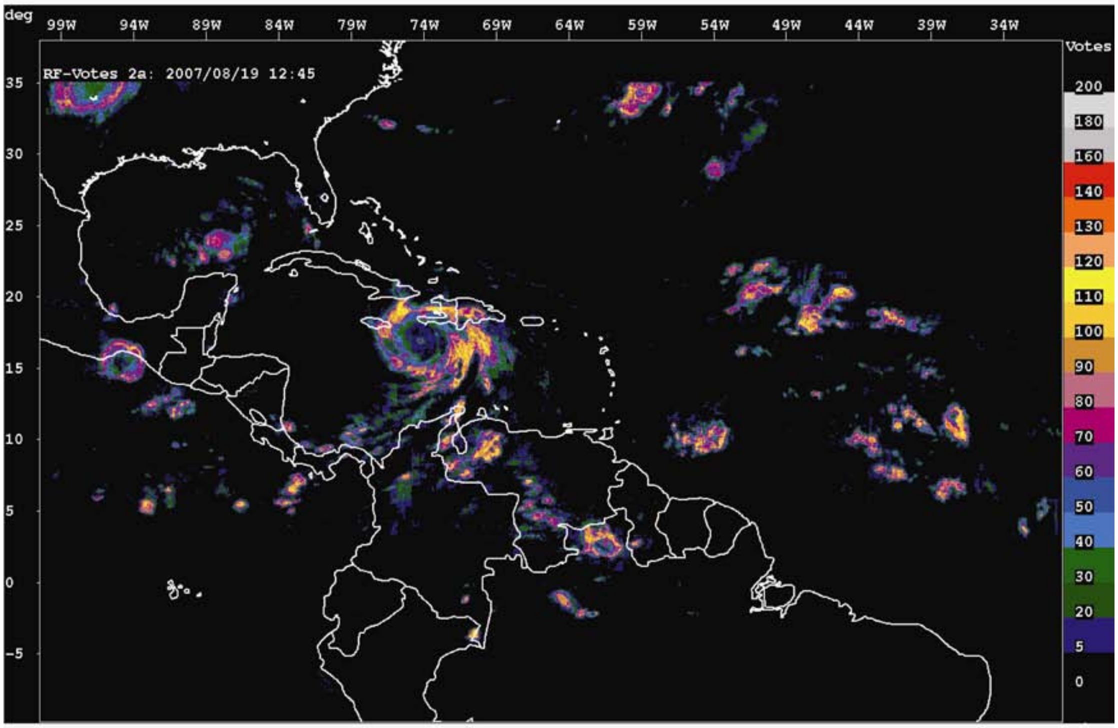


Fig. 2. An example showing the number of votes the random forest produced for various CDO interest values at 1245 UTC on August 19, 2007 over the Gulf of Mexico domain for a) CDO interest = 0 and b) CDO interest = 1. Figure continued on next page.

(c) Votes for CDO interest = 2



(d) Votes for CDO interest = 3

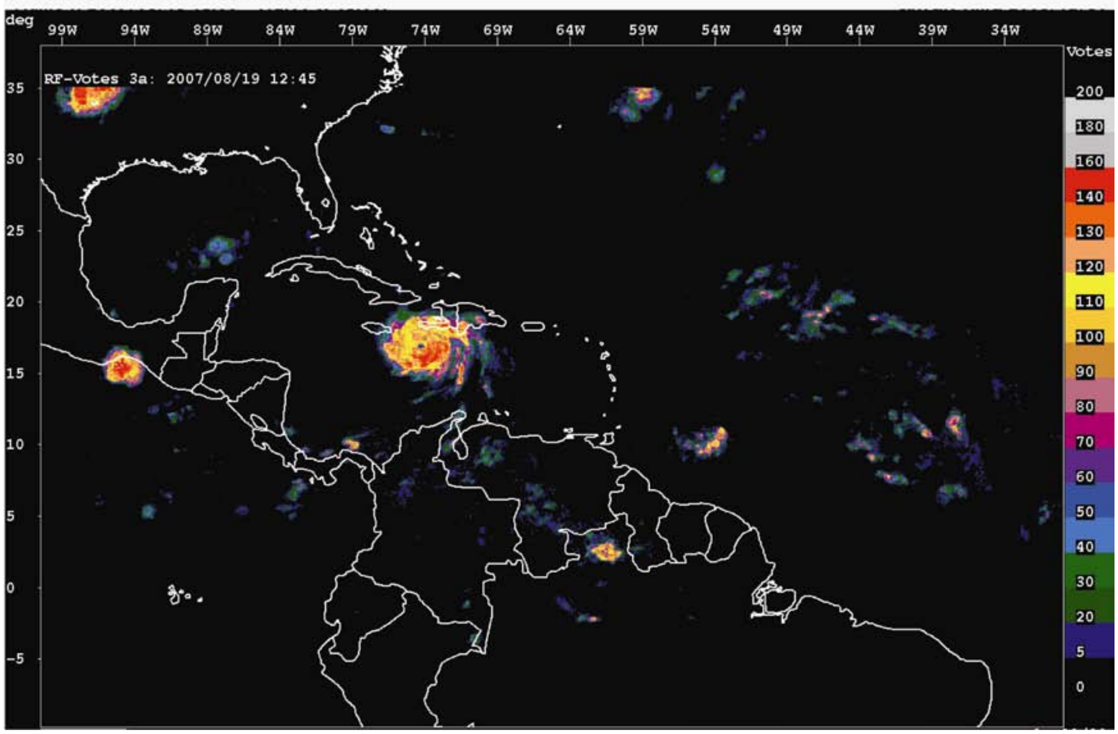


Fig. 2, con't. An example showing the number of votes the random forest produced for various CDO interest values at 1245 UTC on August 19, 2007 over the Gulf of Mexico domain for c) CDO interest =2 and d) CDO interest = 3. Figure continued next page.

(e) Votes for CDO interest = 4

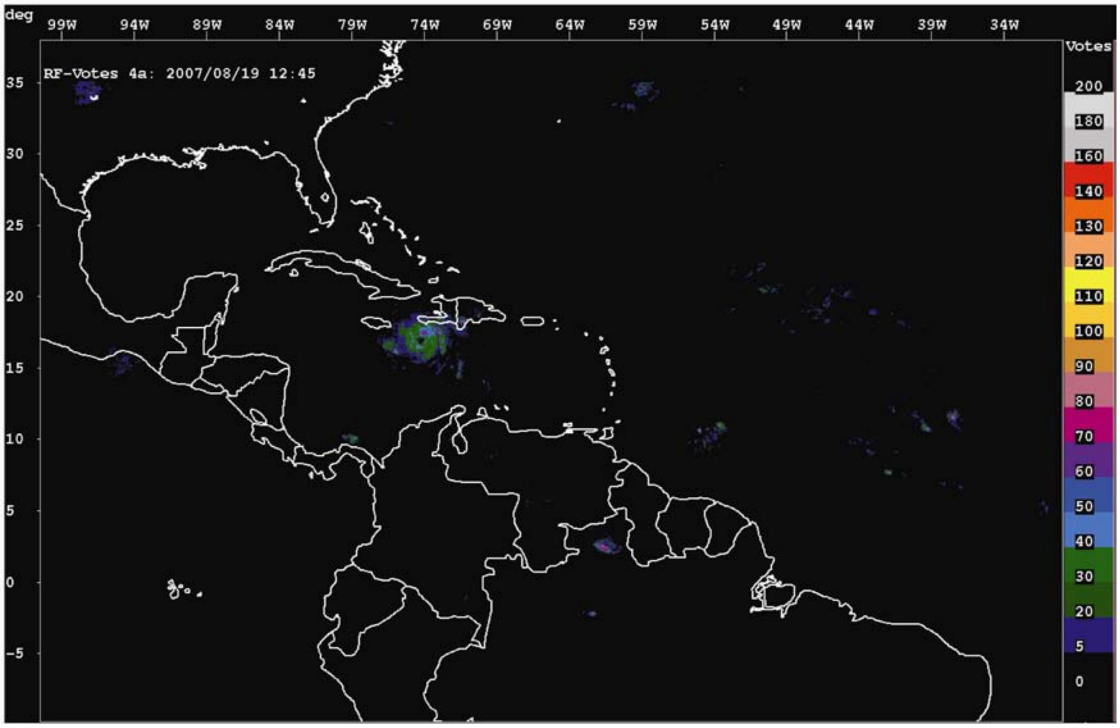


Fig.2, con't. An example showing the number of votes the random forest produced for various CDO interest values at 1245 UTC on August 19, 2007 over the Gulf of Mexico domain for e) CDO interest = 4.

files are then thinned and used for training a forest with 200 decision trees. To achieve reasonable accuracy, at least 100 decision trees are needed. The trained forest is used for independent classification/verification for the following four other days.

The Random Forest produces votes of each CDO interest category (i.e., CDO interest value equals 0, 1, 2, 3 or 4) for each set of input predictors at forecast time. When it was found that the classification process, which creates the CNO-RF forecast, could take ~1 hr to run, we decided to thin the ARFF input file by using cloud top height as a threshold. Only regions with cloud top height over 10,000 ft are classified. By reducing the number of input grid points, the classification process takes ~20 min to run. Certainly the reduction in computing time depends on the weather condition inside the domain.

One example of the votes for CDO interest equals 0, 1, 2, 3 and 4 for one hour forecasts is shown in Fig. 2. Hurricane Dean can be seen clearly in the middle of the domain. As you would expect, the majority of the domain with no convection has most decision trees voting CDO = 0 (see Fig.2a); at the same time, very few decision trees vote yes in the convection-free region for CDO values greater than zero. The strong convection associated with Hurricane Dean has the majority of trees voting CDO interest = 3 (see Fig. 2d).

An example of a 1 hr CNO-RF nowcast and its corresponding verification is shown in Fig. 3. It is interesting to notice that a CDO interest forecast purely derived through decision tree votes looks very similar to the CDO validation field, considering totally different techniques are used in calculating them. The CNO-RF forecast was able to capture Hurricane Dean as well as other relatively weak convection over the Gulf of Mexico. It should be pointed out that the CNO-RF forecast seemed unable to forecast CDO = 4 very well, therefore, some calibration might be needed for better verification results.

In addition to producing a deterministic forecast by choosing the mode of the Random Forest classification, probabilistic forecast of CDO interest can

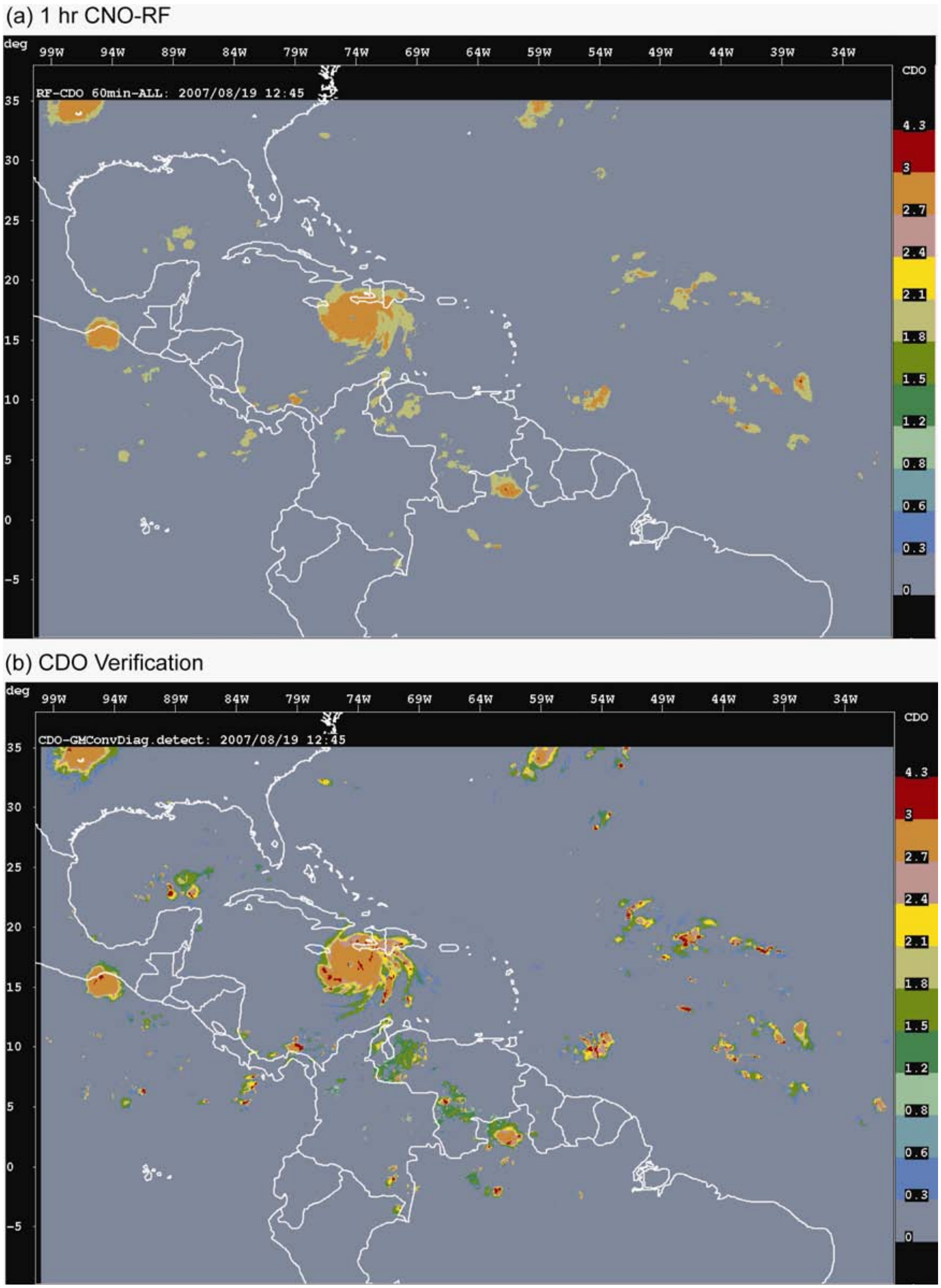


Fig.3. An example of random forest created CDO 1 hr forecast (CDO-RF) and its corresponding verification at 1245 UTC on August 19, 2007 over the Gulf of Mexico domain. a) 1 hr CDO-RF, and b) CDO verification.

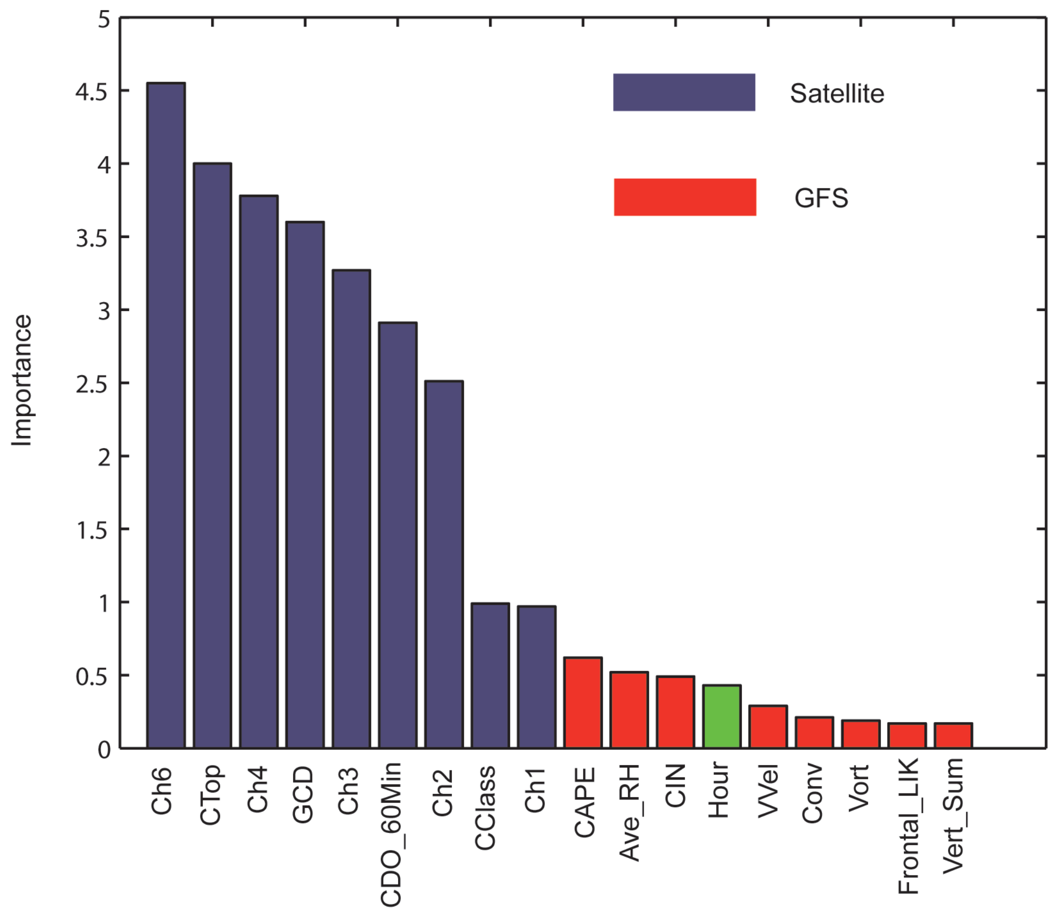


Fig.4. Bar chart showing the imporance ranking of selective predictors for 1 hr, 200 tree Random Forest forecast.

also be created using the votes information of each decision trees. This is an area of active ongoing research and results will be presented in future papers.

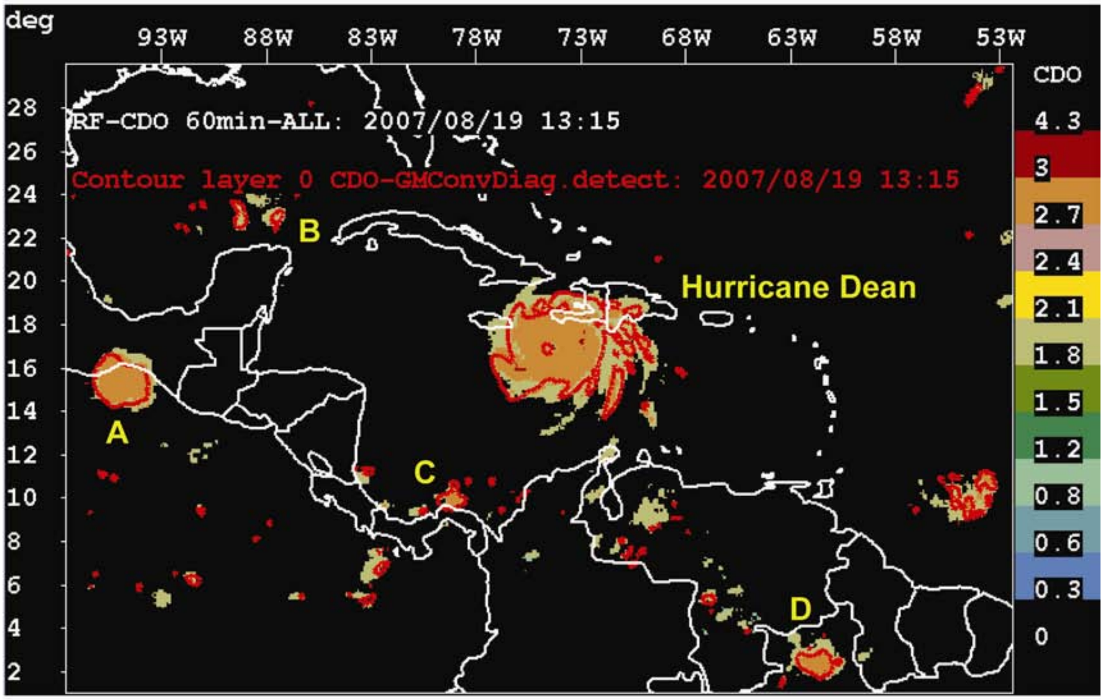
3. RESULTS

3.1. Predictor importance

One unique aspect of Random Forest technique is its capability of ranking the importance of various predictors. The relative ranking of various predictors have several implications. First, it reveals which predictor is contributing most to a correct forecast; thus, in the future training of a new forest, only the important predictors should be included if computational efficiency is an issue. Secondly, if a fuzzy logic forecast system is to be designed based on a set of predictors, the weights of each predictor can be decided by proper usage of Random Forest importance ranking. The latter application of Random Forest has profound impact on the designing/tuning of fuzzy-logic-based forecasting system, since it changes the subjective way of assigning weights for different predictors into an objective, systematic way.

The importance ranking of predictors for a 1hr, 200 tree Random Forest forecasting is shown in Fig. 4. As we would expected, the extrapolation of satellite-based observational fields such as cloud top height, original satellite channels, Global Convective Diagnosis (GCD; Mosher, 2002), and cloud types rank at the top of importance, followed by GFS model-derived environmental fields (i.e., CAPE, CIN, averaged relative humidity). Predictors related

(a) 1 hr CNO-RF



(b) 1 hr CNO

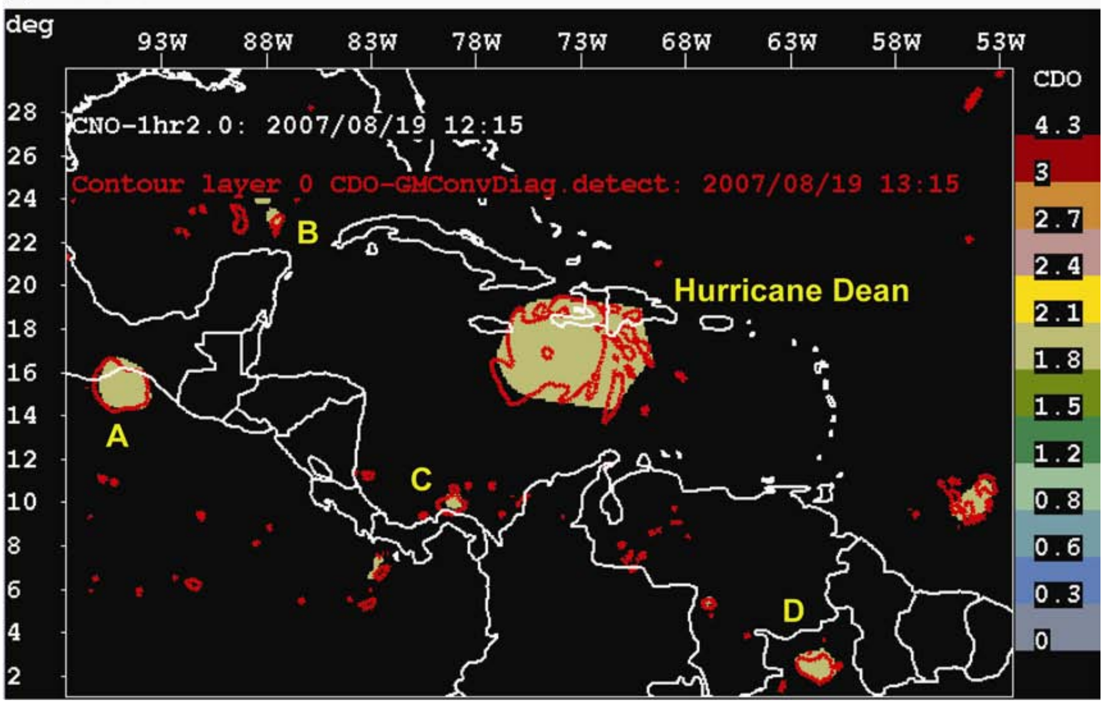


Fig.5. Comparison of 1 hr CDO interest forecasts by a) CNO-RF, and b) TITAN-based CNO technique. The forecasts are valid at 1315 UTC on 19 August 2007 over the Gulf of Mexico domain. The red lines represent the verification of CDO interest =2.5.

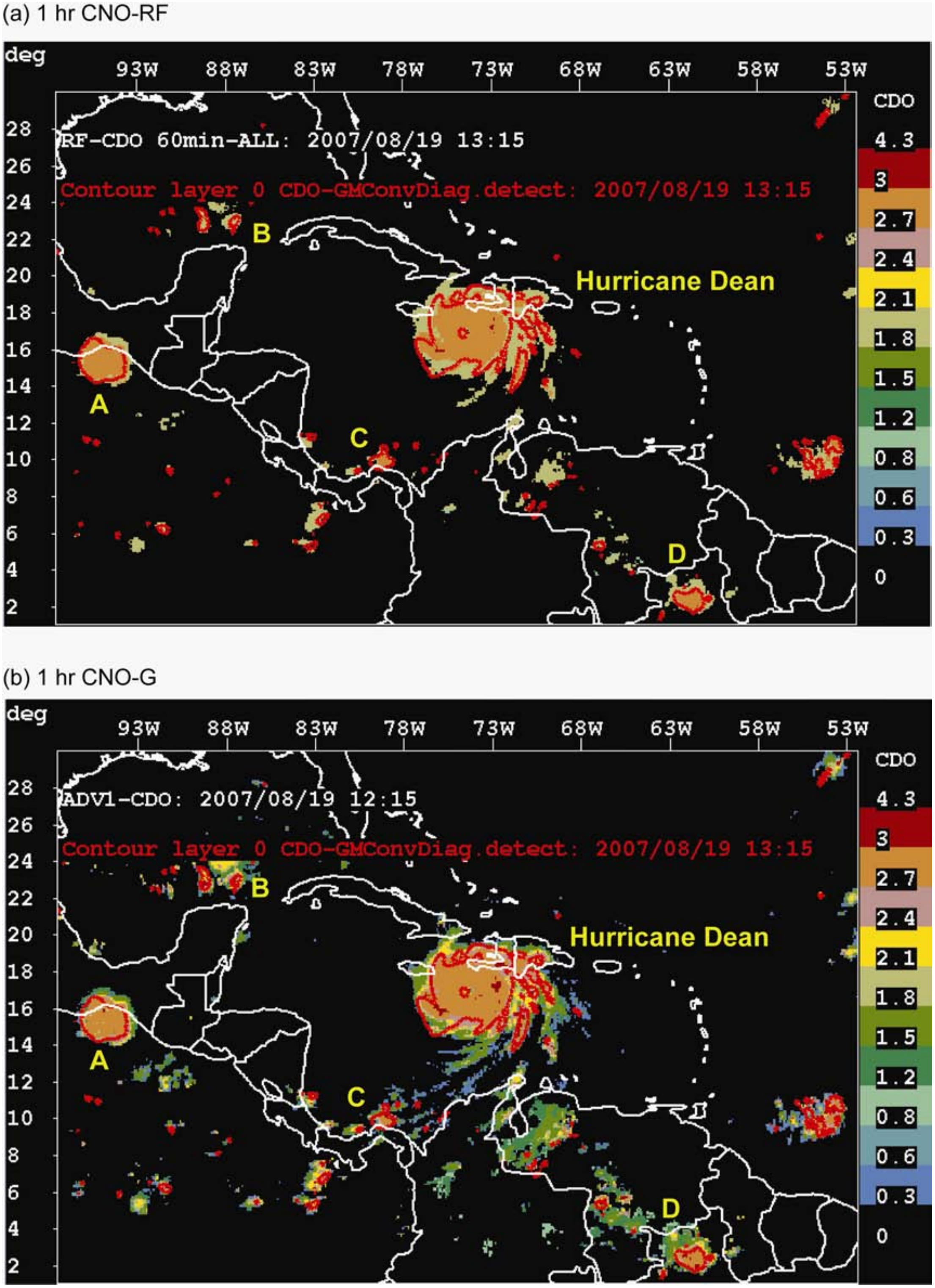


Fig.6. Comparison of 1 hr CDO interest forecasts by a) CNO-RF, and b) CNO-G technique. The forecasts are valid at 1315 UTC on 19 August 2007 over the Gulf of Mexico domain. The red lines represent the verification of CDO interest =2.5.

to new storm initiation did not show up near the top of the importance ranking probably as a result of the dominance of existing convection in the training dataset. Efforts are under way to separate existing storms from new storms such that a Random Forest can be trained to only detect new storms, rather than a combination of existing storms and new storms. In the Random Forest results shown here, existing storms seem to dominate the results, particularly within the importance ranking (Fig. 4).

3.2 Statistical Evaluations

One hour CNO-RF nowcasts for August 19-22, 2007 are produced and compared with corresponding CNO and CNO-G forecasts. An example of the CNO-RF nowcast is compared to a TITAN-based CNO nowcast valid at the same time in Fig. 5. The same CNO-RF nowcast is also compared with CNO-G forecast in Fig. 6. As we can see from Fig. 5, the CNO-RF technique did rather well on Hurricane Dean by producing a realistic looking hurricane with rainbands stretching out from the storm, while the CNO 1 hr forecast in Fig. 5b represents the hurricane by a polygon shape, which certainly caused some over-forecasting problems. Subtle differences between the two techniques also exist for storm A, B, C and D in Fig. 5. The CNO-RF did slightly better for storm A, C and D in terms of their growth/decay trends, but it totally missed forecasting storm B, if a CDO interest threshold of 2.5 is imposed (notice there is no orange color associated with storm B in Fig. 5a). As for the relative performance between CNO-RF and CNO-G, visual inspections of Fig. 6 could not tell any significant differences. Subjective evaluation of many 1 hr forecasts similar to Figs. 5 and 6 suggests that the skills of all three techniques are comparable for 1 hr forecasts, with the CNO-RF and CNO-G techniques showing the advantage of more realistic looking storms.

Standard verification scores based on CDO interest threshold of 2.5 for CNO, CNO-G and CNO-RF forecasts between 19-22 August 2007 over the Gulf of Mexico domain are shown in Table 1. Consistent with the subjective evaluations discussed earlier, all techniques have comparable Critical Success Index (CSI) performance scores for the 1 hr forecasts. It is interesting to notice that the CNO-G technique has the lowest bias, while the CNO (CNO-RF) technique tends to over (under)-forecasting the CDO interest. It is possible that the performance score of the CNO-RF technique could be improved by some proper calibrating processes. Although the three techniques show no significant difference in skills for the 1 hr forecast, it is hoped that the advantage of the CNO-RF technique could be better exploited at longer forecast lead times, owing to its capability of handling storm growth/decay through including GFS-derived storm environmental variables. Both 2 hr and 3 hr CNO-RF forecasts are being pursued now and results will be presented at the conference if they are available.

4. SUMMARY and FUTURE WORK

While this first attempt to use the Random Forest machine learning technique to nowcast oceanic convection has shown promise, plenty of improvements will be pursued in the near future. First, the input predictor list will be expanded to include more fields; secondly, the forecast should be extended to longer lead times since the Random Forest technique could potentially show more skill at longer lead time; and finally, both the training and classification datasets need to be expanded dramatically so that statistically meaningful results can be obtained regarding the performance of different techniques.

As refinements to the CNO-G and the CNO-RF methodologies are made, comparison of their results and statistical performance to the existing CNO system will be made. If statistical performance is improved by the new techniques, the current CNO will be upgraded to incorporate them, as computational load allows.

REFERENCES

Breiman, L., 2001: Random forests, *Machine Learning,* **45** (1), 5-32.

Cai, H., R. Kuligowski, G. Lee, N. Rehak, G. Cunning, D. Albo, D. Megenhardt and M. Steiner, Brightness temperature nowcasting for satellite-based short-term prediction of storms: opportunities and challenges, *Proc. SPIE*, Vol. **7088**, 708807 (2008); DOI:10.1117/12.795621.

Dixon, M., and G. Wiener, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis and Nowcasting - A radar-based methodology. *J. Atmos. Oceanic Technol.,* **10**.,785-797.

Table 1. Comparisons of standard verification scores for 1 hr CNO, CNO-G and CNO-RF forecasts

Techniques	POD	FAR	CSI	Bias	No. of Forecasts
CNO	0.69	0.44	0.45	1.23	114
CNO-G	0.64	0.35	0.48	0.98	114
CNO-RF	0.58	0.24	0.48	0.76	144

Ho, T., 2002: A data complexity analysis of comparative advantages of decision forest constructors, *Pattern Analysis and Applications*, **5**, 102-112.

Kessinger, C., M. Donovan, R. Bankert, E. Williams, J. Hawkins, H. Cai, N. Rehak, D. Megenhardt, and M. Steiner, 2008: Convection diagnosis and nowcasting for oceanic aviation applications, *Proc. SPIE*, Vol. **7088**, 708808 (2008); DOI:10.1117/12.795495.

Kessinger C., H. Cai, N. Rehak, D. Megenhardt, M. Steiner, R. Bankert, J. Hawkins, M. Donovan and E. Williams, 2009: The oceanic convection diagnosis and nowcasting system. 16th Conf. Satellite Met. Ocean., Amer. Meteor. Soc., Phoenix, AZ, 12-15 Jan. 2009.

Mosher, F., 2002: Detection of deep convection around the globe. Preprints, 10th Conf. Aviation, Range, and Aerospace Meteor., Amer. Meteor. Soc., Portland, OR, 289-292.

Williams, J. K., D. Ahijevych, S. Dettling, and M. Steiner, 2008: Combining observations and model data for short-term storm forecasting. *Proc. SPIE*, Vol. **7088**, 708805 (2008); DOI:10.1117/12.795737.

ACKNOWLEDGMENTS

This study is supported by the National Aeronautical and Space Administration (NASA) Research Opportunities in Space and Earth Sciences (ROSES) and NASA Advanced Satellite Aviation Weather Products (ASAP) program.