J8.3 **The Oceanic Convection Diagnosis and Nowcasting System**

Cathy Kessinger*, Huaqing Cai, Nancy Rehak, Daniel Megenhardt, Matthias Steiner,
National Center for Atmospheric Research
Boulder, CO

Richard Bankert, Jeffrey Hawkins
Naval Research Laboratory
Monterey, CA

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

1. Introduction

The Oceanic Convection Diagnosis and Nowcasting system uses geostationary satellite-based methodologies to identify deep convection over oceanic regions and produces short-term nowcasts of its future location. Satellite-derived environmental parameters from polar-orbiting satellites in low earth orbit and a global numerical weather prediction model identify favorable conditions for oceanic storm initiation and continuance. Within the system, convection is identified through a fuzzy-logic combination of three satellite-based algorithms to form the Convective Diagnosis Oceanic (CDO) product. As a first step, the CDO-identified convection is extrapolated using an object-tracking methodology to form the Convective Nowcasting Oceanic (CNO) product for 1-hr and 2-hr nowcasts. The methodology of both the diagnosis and the nowcasting systems are described.

Independent validation of the convective diagnosis product is accomplished with the National Aeronautical and Space Administration (NASA) Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (PR) and Lightning Imaging System (LIS) with the validation methodology and results described in Donovan et al. (2009). Investigation into using Random Forest methodology to improve short-term nowcasts is underway with preliminary results described in Cai et al. (2009).

*Corresponding author address: Cathy Kessinger, NCAR/RAL, P.O. Box 3000, Boulder, CO 80301; email: kessinge@ucar.edu

2. Aviation Weather Center Products

The National Weather Service (NWS) Aviation Weather Center (AWC) produces current and forecast weather products in oceanic regions that are available through facsimile transmission or through the World Wide Web within their “International Flight Folder” (http://aviationweather.gov). The NWS AWC presents international significant meteorological information (SIGMETs) for the Atlantic and Pacific oceans in a textual format and plots that with geophysical references at 4-hr intervals (Fig. 1a). The SIGMETs indicate the presence of hazardous convection, turbulence and icing. Each event is described by its horizontal and vertical location, its intensity and movement as well as assigning an expiration time. For hurricanes and tropical storms, SIGMETs are issued at 6-hr intervals. The Significant Weather (SIGWX) prognostic chart (Fig. 1b) approximates a human-drawn weather chart and depicts the type, the location, movement and intensity of weather features such as fronts, cumulonimbus clouds and regions of turbulence; it is updated at 6-hr intervals. In summary, these products are issued infrequently and cover large domains. For instance, the area enclosed by a SIGMET is typically so large that aircraft have little option but to traverse through it.

The goal of our research is to provide higher resolution (both spatially and temporally) convection products for use by the oceanic aviation community. The convective diagnosis and

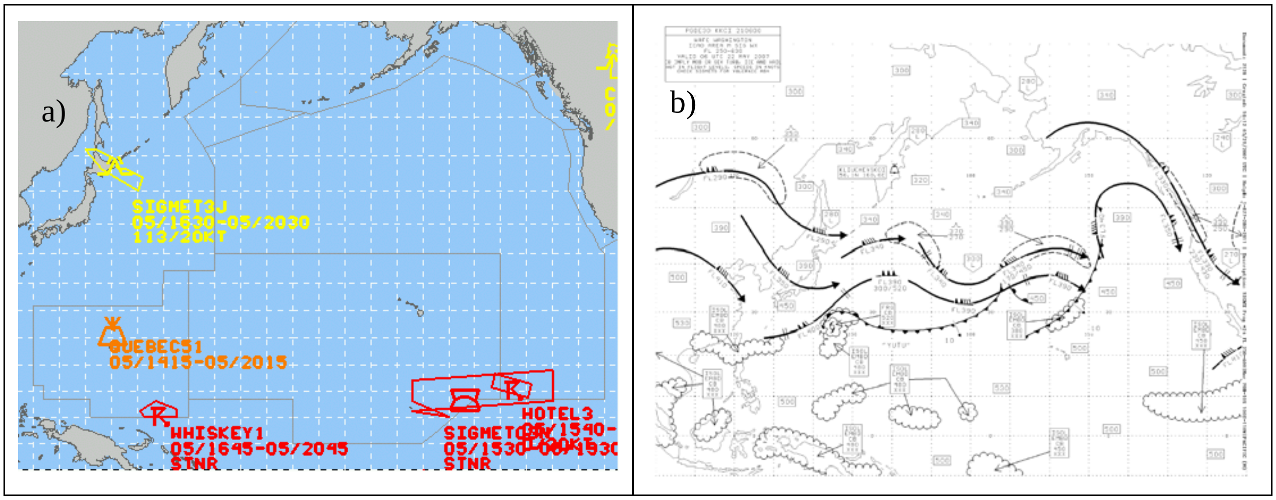

Figure 1: AWC products are shown that are available through the International Flight Folder on the WWW and include a) international SIGMETs, updated every 4-hrs, and the b) SIGWX facsimile chart, updated every 6-hrs.

nowcasting products are geared toward oceanic pilots and dispatchers of transoceanic routes where aircraft fly at en-route altitudes. Satellite remote sensing provides the best means to monitor convective attributes near the desired temporal and spatial resolution. These products are geared toward fulfillment of the Federal Aviation Administration's Next Generation Air Transportation System (FAA NextGen) goal of a global convection product. With expected increases in the number of oceanic flights over the next few years (FAA, 2008), convective products with higher spatial and temporal resolution will serve to improve safety and efficiency of airline operations.

3. Hurricane Dean Case Study

To illustrate the methodology used within the convective diagnosis and nowcasting products, the Hurricane Dean case from 12-23 August 2007 was selected because the hurricane traversed the entire domain from east to west, providing a long time period of convective activity. Hurricane Dean was the first land-falling category 5 hurricane in the Atlantic basin since Hurricane Andrew in 1992 and was responsible for 32 fatalities. Hurricane Dean traversed the domain of interest that includes the greater Gulf of Mexico region (Fig. 2). On 11 August, Hurricane Dean formed from a tropical wave off the west coast of Africa, became a tropical depression around 06 UTC on 13 August, and reached hurricane status early on 16 August (Fig. 3; Franklin 2008). By 19 August, Hurricane Dean was a category 4 hurricane with a well-defined eyewall as seen by microwave and visible imagery (Fig. 4). At landfall on 21 August, its central pressure was estimated at 905 mb with maximum sustained winds of 150 kt.

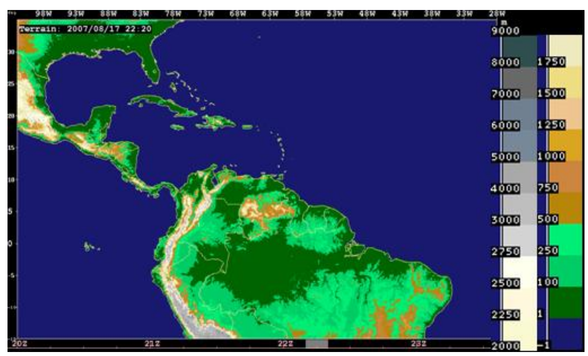

Fig. 2. The domain covered by the Oceanic Convection Diagnosis and Nowcasting effort is shown. Terrain heights (m) are plotted and color-coded via the scale to the right.

GOES-East infrared imagery on 17 August 2007 at 2245 UTC (Fig. 5) shows the position of Hurricane Dean as well as significant amounts of convection elsewhere. Purple-shaded regions define cloud top brightness temperatures (BT) of -35°C or less.

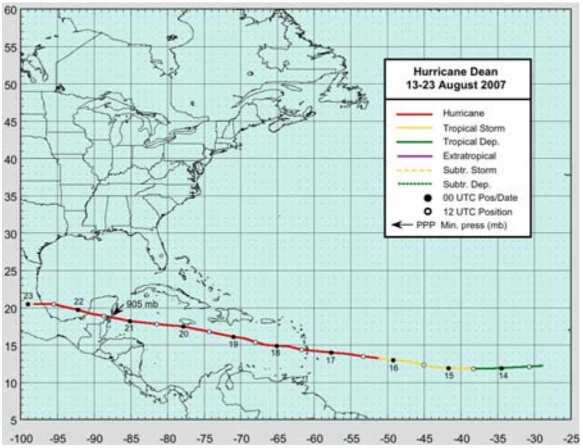

Fig.3. The best track positions of Hurricane Dean from 13-23 August 2007. Figure courtesy of Franklin (2008).

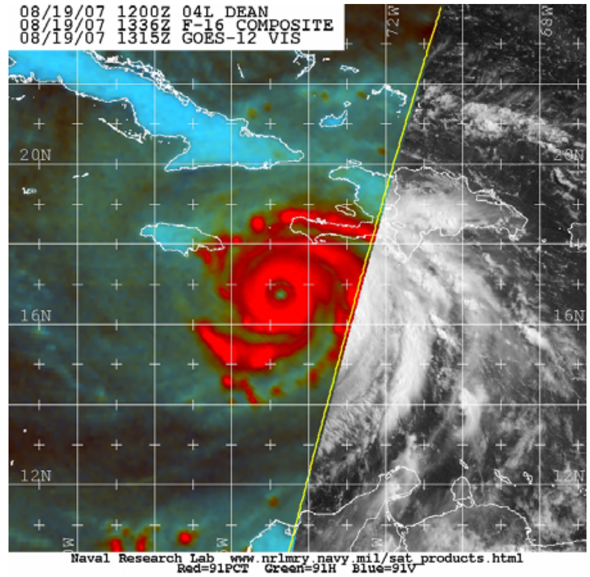

Fig. 4. Microwave imagery of Hurricane Dean (1336 UTC) superimposed over GOES-E visible imagery (1315 UTC) on 19 August 2007. Image from NRL Tropical Cyclone webpage at http://www.nrlmry.navy.mil/tc_pages/tc_home.html.

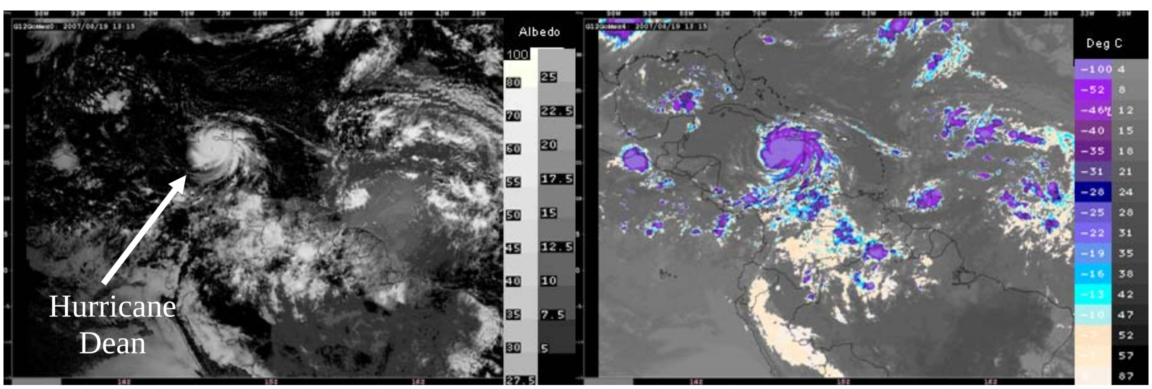

Fig. 5. Satellite imagery is shown from GOES-East on 19 August 2007 at 1315 UTC and includes the a) visible and b) the longwave infrared channels. Hurricane Dean is indicated.

4. Methodology for Diagnosis

The National Oceanic and Atmospheric Administration (NOAA) Geostationary Operational Environmental Satellite-East (GOES-East) longwave infrared (IR) and visible (VIS) imagery are used as input into the three convection detection algorithms, described below. The National Center for Environmental Prediction (NCEP) Global Forecasting System (GFS) numerical model provides global coverage on a 0.5 degree latitude/longitude grid and is input into the cloud top height (CTOP) algorithm.

4.1. Component algorithms of the CDO

Convective clouds are identified via a fuzzy logic combination of three satellite-based algorithms (Cloud Classification, Cloud Top Height, and

Global Convective Diagnosis) to form the Convective Diagnosis Oceanic (CDO) product and are briefly described here (Kessinger et al. 2008). The data fusion methodology used within the CDO is described in the following section.

4.1.1. CLOUD CLASSIFICATION (CCLASS)

Using a supervised learning methodology that was first applied to AVHRR data (Tag 2000), a cloud classifier has been developed and further refined for application to GOES data (Bankert and Wade 2007; Bankert et al. 2008). A training data set is established through independent expert agreement of thousands of labeled 16*x*16 pixel samples. The classes used by the experts (and of relevance to this research) include cumulonimbus (Cb) and cirrostratus anvil (CsAn) for daytime classifications and a deep convection (DC) class at night. CsAn represents relatively deep cirrostratus (Cs) near turrets in thunderstorms and is more closely related to deep convection than "garden variety" Cs. These four categories are inputs into the CDO product.

Each training set sample is represented by a vector of characteristic features computed or extracted from each spectral channel in the GOES imager. Various training sets were established, differentiated by satellite (GOES-East or GOES-West), sea or land, and day or night. A 1-nearest neighbor algorithm is used within the classifier. The minimum distance in feature space between an unclassified sample presented to the classifier and the training data samples is found and the class label of the nearest-neighbor training sample is subsequently assigned to each pixel in the unclassified sample.

Classifications of overlapping boxes (moving 16x16 pixel window) within each image are performed such that each image pixel is classified four times with the majority class assigned (ties broken randomly). Since each box is assigned a specific class, no "multiple", "overlapping", or "unknown" classes are used.

4.1.2. CLOUD TOP HEIGHT (CTOP)

The CTOP algorithm (Miller et al. 2005) combines geostationary IR data with the temperature profile data from the GFS to estimate the heights of convective cloud tops over ocean and land surfaces during day- and night-time hours. For a given pixel location, the algorithm converts the satellite 11-μm IR BT (approximate cloud top temperature) to a cloud top height (pressure level) using the GFS vertical profile. The estimated pressure level is converted to height above sea level using the pressure vs. height relationship given by the standard atmosphere convention, which has been widely adopted for aviation use. Note that this algorithm is intended for use over deep cloud systems, not for cloud tops lower than 15K ft.

4.1.3. GLOBAL CONVECTIVE DIAGNOSIS (GCD)

The GCD algorithm (Mosher 2002), for a given pixel location, computes the BT difference between the water vapor channel (6.7-μm) and the longwave IR channel (11-μm). Deep, convective (i.e., optically thick) clouds that reach the tropopause are overlaid by dry, stratospheric air such that the BT of these two channels will be nearly equal at storm top. Within the GCD, near-zero differences are associated with deep convection. The GCD, as originally devised by Mosher, used the GFS 4-layer lifted index to remove thermodynamically stable regions. This step is not utilized here, as undesirable discontinuities are created within the CDO due to the large grid spacing (0.5 degrees) of the GFS model.

4.2. Methodology of the CDO product

The CDO product is computed using a fuzzy logic, data fusion procedure (Fig. 6a) that ingests output from the three algorithms described above. Output from each of the three algorithms is scaled by a stepwise linear "membership function" such that values that positively indicate the desired feature (i.e., convective clouds) are scaled to unity while values that do not indicate the desired feature are scaled to zero (see Figs. 6b-6d). The output from the membership function scaling is termed an "interest field". The interest outputs are weighted (GCD and CTOP use a weight of 1 while CClass has a weight of 2) and summed to form the initial CDO product with a maximum value of four. The final CDO product is formed after the application of a threshold of 2.5 thus creating a binary

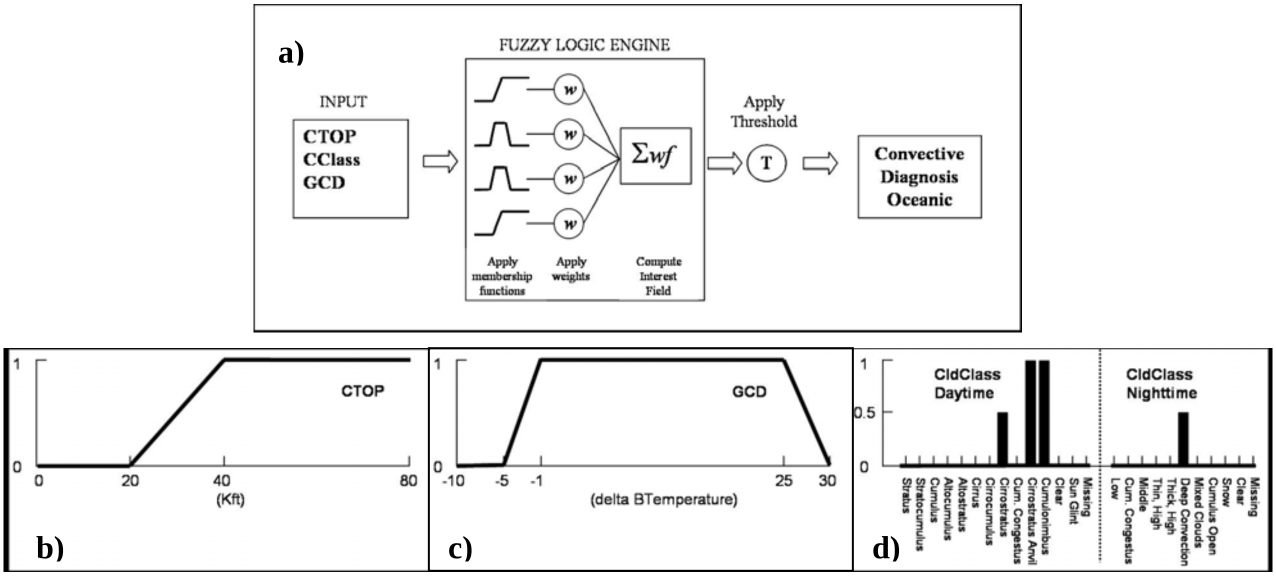

Fig. 6. In a), a schematic shows the fuzzy logic, data fusion process used to calculate the Convective Diagnosis Oceanic (CDO) product. The membership functions for b) CTOP, for c) GCD and for d) CClass are shown.

indicator for the presence (=1) or absence (=0) of convection. The threshold value ensures positive contributions from at least two algorithms.

The target audience for the CDO/CNO product suite is transoceanic, commercial aircraft that are flying at altitudes between 30-40 kft. Membership functions for the CDO component algorithms reflect this emphasis by the selection of categories for CClass (Fig. 6d), the scaling of higher cloud top levels in CTOP (Fig. 6b) and the emphasis on deep convection by the GCD (Fig. 6c).

Figure 7 shows an example of the three component algorithms that are input into the CDO for the Hurricane Dean case study, taken from 19 August 2007 at 1315 UTC (Fig. 5). In the top row, Figs. 7a-7b shows the cloud top height algorithm output with its associated interest field after application of the membership function (Fig. 6b). Notice in Fig. 7b, how the highest cloud tops approach an interest value of unity. Likewise, Figs. 7c-7d show the CClass algorithm for daytime only conditions. The CClass algorithm has performance differences between day and nighttime conditions due to the loss of the high resolution visible imagery at night. Considerably more daytime, small-scale structure is realized when compared to the nighttime, as expected. Because extrapolation of convective cell positions is done for all hours, consistent storm cell area is needed between the day and the night to ensure good tracking performance is attained. For this reason, the area enclosed by the daytime categories of Cb, CsAn and Cs are balanced against the area of the nighttime DC category. Because the CClass produces a cloud type rather than a range of values, the interest values are constant within a given category. The GCD algorithm output and interest field are shown in Fig. 7e-7f. Regions of deep convection, particularly the eyewall portion of Hurricane Dean, that have GCD values near zero, have interest values near unity and indicate mature updrafts.

After application of the data fusion procedure, the final CDO interest field is calculated. CDO interest values range between 0-4 during the day and between 0-3 during the night due to the weight applied to the CClass algorithm output. Application of a threshold converts the CDO interest field into the CDO product (Fig. 8) that indicates the convective regions. Hurricane Dean is clearly resolved.

4.3. Validation of the CDO product

To validate the CDO product, the NASA TRMM PR, LIS and mission products are utilized to evaluate performance for over 1800 hazardous storms that were identified and validated through a manual process. The methodology and results are contained within Donovan et al. (2009). The TRMM data sets provide an independent validation source for the CDO. Results show that the CDO does detect hazardous deep convection, but can miss warm rain clouds, and can be overly generous in the detected area compared to the location of convective precipitation, updrafts and lightning.

For all storms, a Probability of Detection (POD; Wilks 1995) of 0.72 was attained for the CDO, with a False Alarm Ratio (FAR; Wilks 1995) of 0.26 and a Critical Success Index (CSI; Donaldson et al. 1975) of 0.58. When the storms are stratified between daytime and nighttime conditions, the CDO attains a CSI score of 0.64 during the day and drops to 0.40 at night; however, the ratio of the number of daytime storms to nighttime storms is about 2:1. See Donovan et al. (2009) for complete details.

5. Methodology for Nowcasting

With the goal of providing high resolution, tactical decision aids to oceanic pilots and dispatchers, short-term nowcasts of the location of convection, as identified by the CDO product, are produced for 1-hr and 2-hr intervals. The extrapolation is accomplished via a cell-tracking technique, called the Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN; Dixon and Wiener, 1993). The TITAN was developed for tracking 2- or 3-dimensional storms as identified by radar reflectivity, but for our purposes, the software performs similarly when used to track the 2-dimensional CDO product. A threshold of 2.5 interest value is used to define storms. TITAN extrapolates the storm cell position and anticipates its growth and dissipation from past trends. A minimum storm size of 300 km^2 is a criterion that must be met before a storm is tracked.

For validation, a statistical comparison is done between the area enclosed by the TITAN shape at forecast time and the CDO product ($\geq$2.5 interest) at verification time for all forecast grids produced between 12-22 August 2007. Standard statistical indicators are computed with results shown in Table 1 for the CSI and bias. While this analysis does not provide a fully independent comparison such as was possible for the TRMM-CDO validation, this process does validate the extrapolation of CDO storm positions and is consistent with methodologies used for validating forecast skill over the CONUS (Pinto et al. 2006). The TRMM validation provides an estimate of the quality of the CDO product while this analysis provides an estimate of the quality of the CNO extrapolation process.

As expected, the best CSI performance is realized at the 1-hr nowcast with declining performance at the next hour. The CNO CSI and bias scores produced for these 11 days compare favorably to those produced by the National Convective Weather Forecast - 6hr (NCWF-6) system (Pinto et al. 2006) for one day. The NCWF-6 is primarily a radar-based nowcasting system developed with FAA support to extend convective nowcasts to 6-hr using a blended observation- and NWP-based methodology. In the NCWF-6 analysis (Pinto et al. 2006), the CSI scores are plotted hourly over the diurnal cycle for a Great Plains squall line initiation case to illustrate performance differences related to convection initiation, extrapolation and dissipation and varied from 0.2-0.4 for 1-hr nowcast and from 0.05-0.35 for the 2-hr nowcast with maximum scores realized several hours after

Table. 1. Statistical indicators are summarized for 1- and 2-hr intervals for the CNO for the period from 12-22 August 2007.

		Nowcast Period	
		1-hr	2-hr
Indicators	Critical Success Index (CSI)	0.45	0.35
	Bias	1.23	1.20
	Number of Forecast Grids	319	315

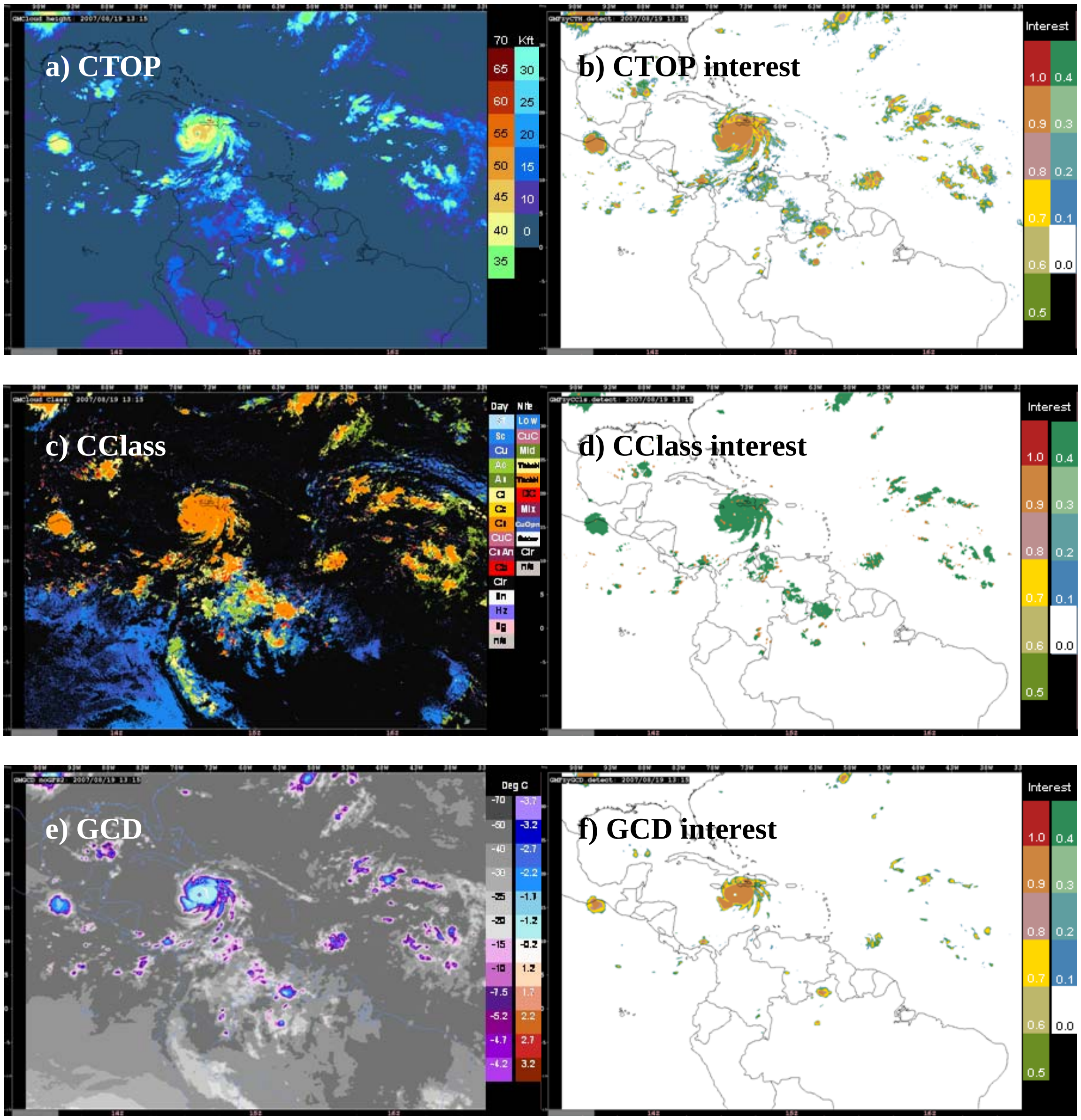

Fig. 7. For the same data shown in Fig. 5, the output from the CTOP algorithm and its interest field are shown in a) and b). Likewise the CClass algorithm output and interest field are shown in c) and d) and the GCD algorithm output and the interest field are shown in e) and f).

the squall line formed. Further evaluation in the same vein is planned for the CNO.

Figure 9 compares a 1-hr and 2-hr CNO nowcasts to the CDO product, both having the same verification time of 1315 UTC. The 1-hr nowcasts (red polygons in Fig. 9a) enclose the CDO validation product fairly well. The polygons tend to be generous in size compared to the area of the CDO with occasional location displacements. In Fig. 9b, the 2-hr nowcast polygons show similar results with some reduction in performance. For both, the position predictions for Hurricane Dean validated very well.

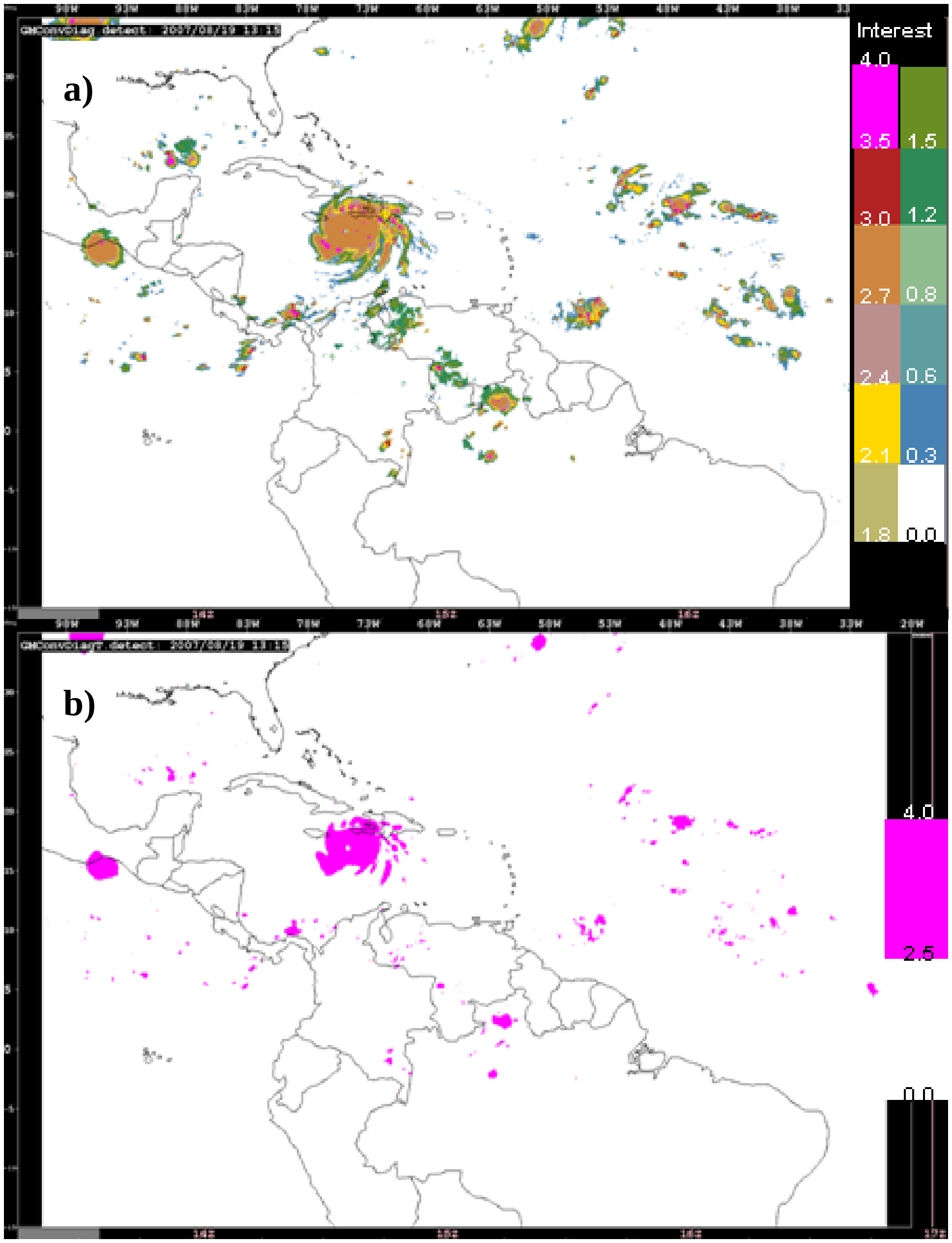

Fig. 8. The CDO output for the data shown in Figs. 5 and 7. The CDO interest field with values ranging from 0-4 is shown in a) and the CDO product is shown in b) after a threshold of 2.5 is applied.

6. Random Forest for Nowcasting

The CNO system as currently configured can only extrapolate existing storm positions and apply growth and dissipation adjustments based on past CDO trends. Inclusion of oceanic environment characterization and GFS model-derived quantities into the CNO system will add greater complexity yet should improve our understanding of where new convection may form, given the presence of a triggering mechanism, or where mature convection should dissipate (Cai et al., 2008).

To accomplish the inclusion of environmental and model-derived data sets into the CNO, a data fusion methodology called "random forest" is explored with preliminary results presented in Cai

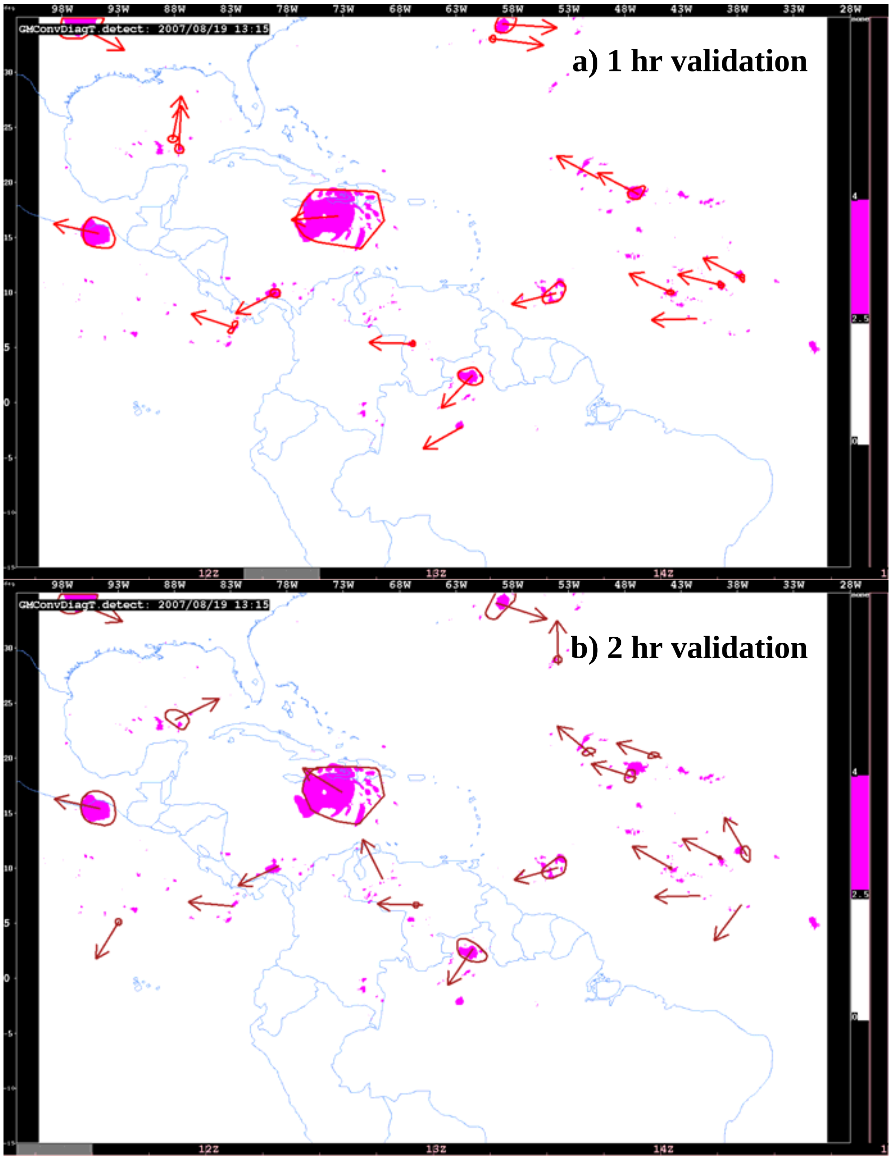

Fig. 9. For 19 August 2007, the CDO product (magenta shapes where CDO$\geq$2.5) is shown at the validation time of 1315 UTC for the a) 1 hr nowcast made at 1215 UTC and for the b) 2 hr nowcast made at 1115 UTC. The 1-hr nowcast is indicated in a) with the red polygons and the 2-hr nowcast is in b) with brown polygons. Vectors (arrows) indicate storm motion direction but not speed.

et al. (2009). Random forest is a powerful, non-linear statistical analysis technique that consists of a collection of independent decision trees. These decision trees are produced from a "training set" of predictor variables (i.e., SST, convergence, etc.) that are paired with their corresponding set of "truth" values (the CDO). Each decision tree's forecast logic is based on a random subset of data and predictor variables, making it independent from all others. A trained random forest functions as an "ensemble of experts" and uses a consensus vote to classify each new data point.

7. Summary

In this paper, we have shown that geostationary visible and infrared imagery can be used to detect hazardous convection via the CDO product over remote, oceanic regions. The CNO product provides 1-hr and 2-hr nowcasts of convection location and is shown to have good performance at extrapolating existing storm positions. Additional oceanic environmental data sets should expand its capabilities to give better indication of convection initiation and evolution.

Acknowledgements

This research is funded by the NASA under grants NNS06AA22G Cooperative Agreement Notice (CAN) and NNA07CN14A Research Opportunities in Space and Earth Sciences (ROSES). The TRMM vis, IR, and PR data are provided by the Goddard Distributed Active Archive Center (DAAC). TRMM LIS data are provided in near real-time from the Global Hydrology Research Center (GHRC), NASA Marshall Space Flight Center (MSFC). Near real-time digital data sets for AMSU, AIRS, AMSR-E SSTs are provided to NRL via collaborations with NOAA, NASA Near Real Time Processing Effort (NRTPE) and Remote Sensing Systems (RSS).

References

Bankert, R.L. and R.H. Wade, 2007: Optimization of an instance-based GOES cloud classification algorithm, *J. Appl. Meteor. Clim.*, **46**, 36-49.

Bankert, R.L., C. Mitrescu, S.D. Miller and R.H. Wade, 2008: Comparison of GOES cloud classification algorithms employing explicit and implicit physics, *Proceedings, 5th GOES Users' Conf.*, Amer. Meteor. Soc., New Orleans, LA, 20-24 Jan 2008.

Cai, H., C. Kessinger, N. Rehak and D. Megenhardt, 2008: Investigation into environmental conditions for storm initiation over the ocean using satellite data, *13th Conf. Aviation, Range, and Aerospace Meteor.*, Amer. Meteor. Soc., New Orleans, LA, 20-24 Jan 2008.

Cai, H., C. Kessinger, D. Ahijevich, J. Williams, N. Rehak, D. Megenhardt, R. Bankert, J. Hawkins, M. Donovan, and E. Williams, 2009: Nowcasting oceanic convection using random forest classification, *16th Conf. Satellite Meteor. Ocean.*, Amer. Meteor. Soc., Phoenix, AZ, 12-15 Jan 2009.

Dixon, M., and G. Wiener, 1995: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting - A radar-based methodology, *J. Atmos. Ocean. Tech.*, 10, 785-797.

Donaldson, R. J., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluation of techniques for predicting severe weather events, *Preprints*, *Ninth Conf. on Severe Local Storms*, Norman, OK, Amer. Meteor. Soc., 321–326.

Donovan, M., E. Williams, C. Kessinger, G. Blackburn, P. H. Herzegh, R. L. Bankert, S. D. Miller, and F. R. Mosher, 2008: The identification and verification of hazardous convective cells over oceans using visible and infrared satellite observations, *J. Appl. Meteor. Clim.*, **47**, 164-184.

Donovan, M., E. Williams, C. Kessinger, N. Rehak, H. Cai, D. Megenhardt, R. Bankert, and J. Hawkins, 2009: An evaluation of a convection diagnosis algorithm over the Gulf of Mexico using NASA TRMM Observations, *16th Conf. Satellite Meteor. Ocean.*, Amer. Meteor. Soc., Phoenix, AZ, 12-15 Jan 2009.

Federal Aviation Administration, 2008: FAA Aerospace Forecast Fiscal Years 2008-2025. Available at http://www.faa.gov.

Franklin, J. L., 2008: Tropical cyclone report, Hurricane Dean (AL042007). National Hurricane Center, http://www.nhc.noaa.gov/2007atlan.shtml.

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, J. Hawkins, and E. Williams, 2008: Oceanic Convection Diagnosis and Nowcasting, *13th Conf. Aviation, Range, and Aerospace Meteor.*, Amer. Meteor. Soc., New Orleans, LA, 20-24 Jan 2008.

Miller, S., T. Tsui, G. Blackburn, C. Kessinger and P. Herzegh, 2005: Technical description of the cloud top height (CTOP) product, the first component of the Convective Diagnosis Oceanic (CDO) product. Unpublished manuscript available from the author at NRL.

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.

Pinto, J., C. Mueller, S. Weygandt, D. Ahijevych, N. Rehak, and D. Megenhardt, 2006: Fusing observation- and model-based probability forecasts for the short term prediction of convection, *12th Conf. Aviation, Range, and Aerospace Meteor.*, Amer. Meteor. Soc., Atlanta, GA, 30 Jan-2 Feb 2006.

Tag, P. M., R. L. Bankert, and L. R. Brody, 2000: An AVHRR multiple cloud-type classification package, *J. Appl. Meteor.*, **39**, 125–134.

Wilks, Daniel S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, San Diego, 467 pp.