

Forecasting Tropical Cyclogenesis over the Atlantic Basin Using Large-Scale Data

CHRISTOPHER C. HENNON* AND JAY S. HOBGOOD

The Ohio State University, Columbus, Ohio

(Manuscript received 17 September 2002, in final form 13 June 2003)

ABSTRACT

A new dataset of tropical cloud clusters, which formed or propagated over the Atlantic basin during the 1998–2000 hurricane seasons, is used to develop a probabilistic prediction system for tropical cyclogenesis (TCG). Using data from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis (NNR), eight large-scale predictors are calculated at every 6-h interval of a cluster's life cycle. Discriminant analysis is then used to find a linear combination of the predictors that best separates the developing cloud clusters (those that became tropical depressions) and nondeveloping systems. Classification results are analyzed via composite and case study points of view.

Despite the linear nature of the classification technique, the forecast system yields useful probabilistic forecasts for the vast majority of the hurricane season. The daily genesis potential (DGP) and latitude predictors are found to be the most significant at nearly all forecast times. Composite results show that if the probability of development $P < 0.7$, TCG rarely occurs; if $P > 0.9$, genesis occurs about 40% of the time. A case study of Tropical Depression Keith (2000) illustrates the ability of the forecast system to detect the evolution of the large-scale environment from an unfavorable to favorable one. An additional case study of an early-season nondeveloping cluster demonstrates some of the shortcomings of the system and suggests possible ways of mitigating them.

1. Introduction

There is increasing evidence in the recent literature that the crucial physical processes regarding tropical cyclone formation [tropical cyclogenesis (TCG)] involve convective and dynamical interactions that occur on scales that are not currently resolved in operational forecast models (Emanuel 1989; Chen and Frank 1993; Bister and Emanuel 1997; Simpson et al. 1997; Ritchie and Holland 1997; Montgomery and Enagonio 1998). Consequently, such models have shown virtually no skill at forecasting TCG (e.g., Beven 1999), though there has been some improvement as resolution has been increased and model parameterizations improved (Pasch et al. 2002). This leaves forecasters with the dilemma of trying to forecast an event without useful objective guidance and sparse, if any, in situ observational data.

Fortunately, several promising developments suggest that skillful forecasting of TCG is an attainable goal in the short term. First, advances in computing power have led to the development of dynamical models with fine (down to 3-km horizontal) resolution. These models have

provided researchers and forecasters with tools for testing theories and refining knowledge of processes that result in TCG. Using one such model, the fifth-generation Pennsylvania State University (PSU)–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5; Grell et al. 1994), Davis and Bosart (2001) were able to simulate the genesis of Hurricane Diana (1984), beginning from a weak baroclinic disturbance through hurricane stage. This type of research was not possible only a few years ago. Second, quality atmospheric data are now available globally on a regular grid in the form of reanalysis datasets. These datasets provide surface and upper-air data derived from a merged field of model-based first-guess estimates and quality controlled in situ observations. The National Centers for Environmental Prediction (NCEP)–NCAR reanalysis (NNR) provided the atmospheric data for this research.

Third, and perhaps most relevant to this study, is the evidence that suggests that skillful TCG forecasts could be made even if the complex mesoscale features and interactions are neglected. This evidence comes from several sources. McBride and Zehr (1981), in a large global survey of tropical cloud clusters, showed that a single parameter computed from observational upper-air data provides valuable information regarding the likelihood of tropical depression development. The daily genesis potential (DGP), defined simply as the difference of the *large-scale* ($\sim 4^\circ$ – 6° of the cluster center) 900- and 200-mb relative vorticity fields, supported a

* Current affiliation: UCAR Visiting Scientist Program, NOAA/TPC/National Hurricane Center, Miami, Florida.

Corresponding author address: Dr. Christopher C. Hennon, Postdoctoral Fellow, UCAR Visiting Scientist Program, NOAA/TPC/National Hurricane Center, 11691 SW 17th St., Miami, FL 33165-2149.
E-mail: Hennon@noaa.gov

fundamental way of thinking about TCG that was first proposed by Gray (1968). He argued that in terms of large-scale factors important to TCG, the thermodynamic state of the tropical atmosphere during the hurricane season is nearly always favorable; tropical development depends primarily on dynamical factors. A numerical model sensitivity study performed by Tuleya (1991) seemed to confirm that conclusion. The DGP is used as a predictor of TCG in this study. The second source is the classification work of Perrone and Lowe (1986; Lowe and Perrone 1989). Their methodology, which has numerous similarities to this work, involved identifying candidate cloud clusters in the Pacific basin and stratifying them by their eventual development status. Using linear discriminant analysis, they obtained significant skill in separating the developing and nondeveloping cloud clusters and concluded that the large-scale environment may contain more predictive power than has been previously thought. A comparison of their work with this research is discussed in more detail in section 5. Finally, DeMaria et al. (2001) developed a genesis parameter for the Atlantic, derived from operational NCEP analysis fields and Geostationary Operational Environmental Satellite (GOES) water vapor imagery, which was shown to be correlated with active and inactive genesis periods over the Atlantic basin. They conclude by suggesting that a disturbance-centered parameter could be developed to evaluate systems individually, rather than evaluating the environmental properties of the genesis region.

This research is an attempt to develop such a parameter. This paper examines the ability to differentiate between cloud clusters that eventually form into tropical depressions [“developing” (DV)] and those that do not [“non-developing” (ND)] based on large-scale reanalysis data. Building on the studies of Perrone and Lowe (1986), Lowe and Perrone (1989), McBride and Zehr (1981), and DeMaria et al. (2001), we investigate eight predictors either extracted or calculated from the NNR for every identifiable tropical cloud cluster from the Atlantic hurricane seasons (1 June–30 November) of 1998–2000. The validity of using a large-scale dataset to forecast TCG is tested, given that the coarse resolution ($2.5^\circ \times 2.5^\circ$) of the data renders it unable to resolve any mesoscale phenomena, including tropical convection.

To accomplish this task, this study was based on the following premise. TCG is not a single event in time, but a series of events that result in a surface vortex with an upper-level warm core. Each subsequent event cannot occur without the success of the one that immediately preceded it. As the sequence progresses, the probability of tropical cyclogenesis gradually increases until the event is successful (genesis occurs) or a link in the chain fails and the cluster becomes stagnant or falls apart. *The series of mesoscale and smaller-scale phenomena that occur near the end of the sequence cannot do so without the establishment of favorable larger-scale conditions at the beginning of the sequence.* A statistical model

that analyzes the larger-scale cloud cluster features of many developing and nondeveloping systems may be able to provide users with a useful probabilistic forecast for development.

The ultimate goal of this research is to provide an objective tool for forecasters to assess the likelihood of a tropical cloud cluster developing into a tropical depression (TD) from large-scale data that can be routinely derived from operational analyses. It is meant to be used in conjunction with other valuable data and information, particularly satellite tools, which have recently been developed or enhanced. Examples of these exciting advances that are being applied to the TCG problem are atmospheric sounders such as the Advanced Microwave Sounding Unit (AMSU; Kidder et al. 2000), quick scatterometer (QuikSCAT)-derived wind fields (Sharp et al. 2002), the Tropical Rainfall Measuring Mission (TRMM) space-borne precipitation radar (e.g., Cecil et al. 2002), and geostationary satellite data that is being used to detect the existence of dry Saharan air layers in clusters embedded in easterly waves that move offshore of western Africa (J. Dunion 2002, personal communication).

The next section presents the data sources employed in this work. Section 3 describes the methodology, including the creation of the cloud cluster database, the selection of the predictors used to classify developing and nondeveloping systems, and the statistical technique used for classification. Results of the discriminant analysis classification will be presented in section 4, with subsections dedicated to a composite analysis and two individual case studies. Section 5 contains a brief discussion of the differences between this study and that of Perrone and Lowe (1986) that may explain dissimilarities in the results. Finally, the conclusions will be summarized in section 6, followed by suggestions for future work in this area.

2. Data sources

a. NCEP–NCAR reanalysis

The atmospheric data for this study were obtained from the 6-hourly global NNR dataset (Kalnay et al. 1996). The reanalysis is a global assimilation of observations from many different sources, carefully checked for quality and merged with a numerical model first-guess field. The resultant dataset includes pressure level variables on a global $2.5^\circ \times 2.5^\circ$ grid with 17-level vertical resolution. Direct observations include surface stations, aircraft, and rawinsonde and satellite data. Used for numerous applications in meteorological research, the NNR provides quality data coverage in areas that have historically had little or no observational data (i.e., ocean basins). It provides tremendous potential for application to tropical cyclone research for those who are interested in large-scale phenomena that has not existed previously, and is thought to be generally better in the Tropics than the European Centre for Medium-

Range Weather Forecasts (ECMWF) reanalysis (Trenberth et al. 2001; D. Bromwich 2002, personal communication).

For this research, the first of our knowledge to exploit a reanalysis dataset for the study of TCG, we use the following data from the NNR: air temperature at 12 pressure levels (925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, and 70 hPa), sea level pressure, precipitable water, zonal and meridional winds at the surface and at 3 pressure levels (850, 700, and 200 hPa), and specific humidity at the same pressure levels. We are confident in the quality of the temperature and wind data. Kalnay et al. (1996) label temperature, wind fields, and sea level pressure as class “A” variables, meaning that they are almost entirely derived from observations rather than the model. Precipitable water and specific humidity are classified as “B” variables, meaning that the global model has some influence on the output, but that the direct observations still significantly impact the result. Section 3 documents how the data from the NNR are used to calculate the predictors for this study.

b. Sea surface temperature

To provide a lower boundary of temperature for the computation of the maximum potential intensity predictor (discussed in section 3), we use Reynolds sea surface temperature (SST; Reynolds and Smith 1994). Real-time observations are available with a temporal resolution of 1 week, distributed on a $1^\circ \times 1^\circ$ grid. To make the dataset spatially consistent with the NNR, the grid was devolved into a $2.5^\circ \times 2.5^\circ$ grid through a simple bilinear interpolation routine. The data fields are derived from a somewhat complex optimum interpolation (OI) technique (Reynolds and Smith 1994) that improves upon the blended technique of Reynolds (Reynolds 1988).

c. Satellite imagery

In order to track the tropical cloud clusters that make up the foundation of this study, *GOES-8* infrared (IR) satellite imageries were retrieved from the archive at the Wisconsin Space Science and Engineering Center (SSEC) at a time resolution of 6 h and a wavelength of $3.9 \mu\text{m}$. To provide coverage of the eastern Atlantic, archived *Meteosat-7* IR imageries were obtained from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) at similar temporal intervals. A mosaic was produced that combined each satellite’s view into one coherent image. If either image was missing, not available, or not within 1 h of the corresponding view, the time was labeled as “missing” and not available for cloud cluster tracking. Of the 2196 potential scenes available for tracking, less than 5% were not available.

d. Genesis time and location

Determining the time of “genesis” is difficult and open to interpretation. To simplify the question and remain objective regarding the time and location of genesis, we use the National Hurricane Center (NHC) “best track” database (Jarvinen et al. 1984) with the cautionary note that there are several possible sources of uncertainty and error in the database. For example, there is usually little or no direct observational data on hand to aid analysis in determining the occurrence of genesis and the location of the center of circulation (Bracken and Bosart 2000). Nevertheless, the best-track file is a valuable resource and provides the best estimate of the time and location of TCG in the Atlantic basin. Tropical depressions that failed to develop into tropical storms do not appear in the database—those cases were obtained via the NHC Web site.

3. Methodology

a. Assumptions

In our identification and evaluation of cloud clusters we make the following assumptions. For simplicity, we assume that each system is axisymmetric. In the identification process, the position of the cloud cluster is fixed at the estimated center of the cloud mass, unless a more precise circulation center is evident from the satellite imagery. The most accurate position is not crucial to the results, however, because of the coarseness of the dataset ($2.5^\circ \times 2.5^\circ$). We assume that genesis occurs when the system first appears in the best-track database. This is the only objective method for determining genesis at this time. Cloud clusters over major land areas are not considered in this study; otherwise, we neglect any effects that proximity to land may have on the systems. The large majority of the cases thrown out because of this assumption were African easterly waves that had not left the continent.

b. Identification of cloud cluster candidates

Geostationary IR imageries were visually examined for the 1998–2000 Atlantic hurricane seasons (1 June–30 November) to find all of the tropical cloud clusters that formed within the study domain, shown in Fig. 1. To identify a cloud cluster, we had to first decide on a definition of a cloud cluster. In general, tropical cloud clusters are considered to be mesoscale convective systems (MCSs) with length scales of 250–2500 km and timescales ≥ 6 h (Maddox 1981). The main idea is that a cloud cluster should be a system that has the *potential* to develop into a tropical depression. Hence, several requirements immediately surface; namely, a cluster must have sufficient size, must persist for an extended period of time (i.e., not diurnal in nature), and must exist in a region where genesis is a genuine possibility (not in the high latitudes). To quantify these require-

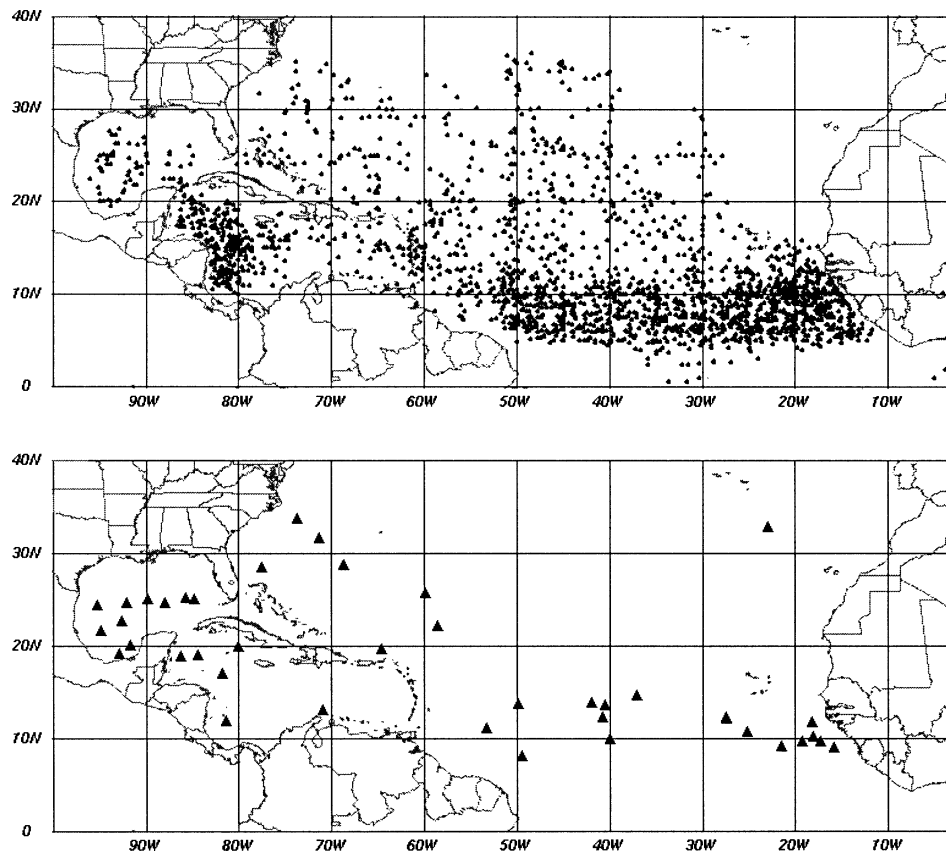


FIG. 1. (top) Location of nondeveloping cases and (bottom) developing clusters 24 h prior to genesis for the 1998–2000 Atlantic hurricane seasons.

ments, we use the following criteria initially adopted by Lee (1989) for his work in the Pacific basin:

- i) Each cluster must be an independent entity, disassociated with either a cyclone or precyclone system
- ii) Cluster must be at least 4° in diameter and not elongated in shape
- iii) Cluster must be located south of 40°N
- iv) Cluster must persist for at least 24 h

Essentially, TCG was highly unlikely if any one of these criteria was not met. These criteria provided a somewhat

objective method for identifying viable candidates. Note that the last requirement (iv) is satisfied if convection persists for any 24-h period of the cluster lifetime. This requirement primarily applies to the ND systems, since all DV clusters were tracked as long as traceable convection was apparent. We were not able to group a number of DV clusters into longer forecast lead times because convection first developed at a time close (inside of 48 h) to genesis. Although a robust automated system for objectively identifying cloud clusters is highly desirable, the clusters in this paper were identified through visual inspection of the satellite brightness temperatures. Table 1 presents a summary of the cloud cluster characteristics for the three Atlantic seasons in this study. The 2000 season was the most active in terms of both developing and nondeveloping candidates, though 1998 appeared to be a more active year for fertile African waves (those easterly waves that exhibited significant and persistent convection). Full-season summaries for all three seasons can be found in the recent literature (Pasch et al. 2001; Lawrence et al. 2001; Franklin et al. 2001).

We define a “case” as one 6-h forecast time within the entire lifetime of a cloud cluster. Therefore, a cloud cluster is typically made up of many cases. Once all

TABLE 1. Summary of documented cloud clusters for the 1998–2000 Atlantic hurricane seasons.

	1998	1999	2000
Total number of clusters	90	91	110
Longest in duration (h)*	198	258	294
Mean duration (h)*	58.9	55.1	54.8
Median duration (h)*	42	36	42
Number of African waves**	42	32	41
African wave ratio (%)	46.7	35.2	37.3
Number of TDs	14	16	18

* Nondeveloping clusters only.

** Only those easterly waves that exhibited significant and persistent convection are included.

candidate cloud clusters were identified, their cases were grouped into DV or ND bins. A case is categorized as DV if the cluster developed into a tropical depression within 48 h of the image time. The DV cases were further stratified by the number of hours before genesis. For example, Tropical Depression Bonnie (1998) achieved genesis at 1200 UTC 19 August. The pre-Bonnie cloud cluster at 0600 UTC 19 August was categorized as a “6-h developing case.” At 2 days prior to genesis, the 1200 UTC 17 August case was labeled a “48-h developing case.” If Bonnie’s convection was evident before that time, it was categorized as an ND case and included with other cases that never achieved tropical depression status. This “double counting” of DV clusters occurred 26 times over the 3-yr period of study. Thus, there were actually 317 clusters over the 3-yr period, but only 291 of them were unique.

Figure 1 illustrates the spatial distribution of all of the ND (top) cases and DV (bottom) cloud clusters at 24 h prior to genesis. There were 2265 ND cases from 269 ND clusters, meaning that each cluster existed for approximately 8.5 six-hour periods on average. Of all the cases that made up those clusters, 114 were excluded from the final analysis because of missing satellite data, 48 because convection became insignificant or absent, and 12 because the cluster progressed over a major landmass. Each individual case is assumed to be independent of the other cases in the same cloud cluster for the discriminant analysis. The two areas of greatest concentration easily stand out: the easterly wave track along 8°N from the African coast into the central North Atlantic and clusters spawned by the monsoon trough in the southern Caribbean. It can be seen that the concentration of DV clusters is spatially coherent with the ND cases, although there is a higher development ratio for the region immediately off the eastern U.S. shore and in the Gulf of Mexico. This suggests that these areas are favored TCG regions climatologically.

c. Technique for creation of predictor dataset

All of the predictors (except pressure tendency, which is dependent on storm location) were computed for each grid point in the domain at 6-h intervals for the entire Atlantic hurricane seasons of 1998–2000. Figure 2 shows an illustration of the averaging technique. Once the location of the cloud cluster is identified, a swath with a radius of 2° [6° for maximum potential intensity (MPI)] is envisioned around that point. All NNR grid points that are located within that radius are averaged to yield a single value that represents the cloud cluster conditions. Typically, 2–3 grid points fall within the averaging radius. We had also examined averaging radii of 4°, 6°, and 8°, but found much greater classification skill with a 2° radius. The lone exception was MPI—we found enhanced predictive abilities when we included MPI averaged at a 6° radius. Thus, all MPI cal-

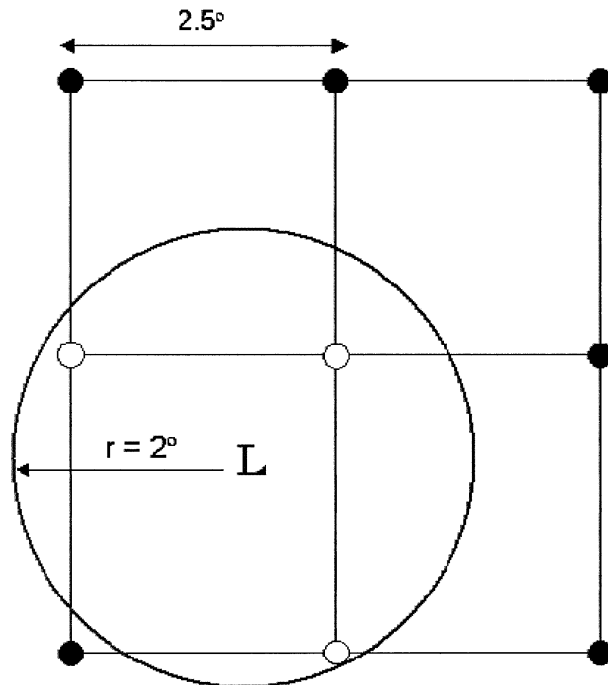


FIG. 2. Schematic of the areal averaging technique for obtaining the predictor dataset. A circle of 2° radius from the center of the cloud cluster (“L”) delimits the outer boundary for the averaging. All grid points that fall within the circle (open dots) are averaged together to yield the value for that cloud cluster and time.

culations presented here were calculated using the larger averaging area.

d. Selection of predictors

The large-scale factors that must be favorable for TCG to occur are well-established and are not necessarily mutually exclusive:

- 1) There must exist a sufficiently large area of preexisting convection (Riehl 1948)
- 2) There must be sufficient planetary vorticity, usually requiring the cluster to be at least 5° latitude away from the equator (Gray 1968)
- 3) There must exist near-zero vertical wind shear *over the center of the system*, accompanied by a high-shear gradient across the system in both the zonal and meridional directions (McBride and Zehr 1981)
- 4) The midtroposphere must be sufficiently moist (Gray 1968)
- 5) The ocean must be sufficiently warm and have a deep mixed-layer depth (Gray 1968)
- 6) The atmosphere must be in a state of conditional instability (Gray 1968)

Note that these are necessary but not sufficient conditions for genesis. The thermodynamic parameters are nearly always favorable during the Atlantic hurricane season (Gray 1968), and yet TCG is still a rare event.

Nevertheless, we selected eight large-scale predictors that we believed captured the crucial requirements of day-to-day TCG. Each predictor is briefly described below, along with the large-scale requirement (numbers 1–6 above) that it satisfies. The condition of preexisting convection (1) is satisfied by default since we are only considering tropical cloud clusters as defined in the previous section.

1) LATITUDE

Latitude satisfies condition 2. We include latitude in our set of predictors as a scaled Coriolis parameter (COR), computed as:

$$f = 2\omega \sin\phi \times 10^4, \quad (1)$$

where ω is the angular rotation of the earth ($=7.29 \times 10^{-5} \text{ s}^{-1}$), and ϕ is latitude in degrees.

2) DAILY GENESIS POTENTIAL

McBride and Zehr (1981) found that the development potential for a cloud cluster could be assimilated into one variable, which they called the “daily genesis potential.” It was formulated based on the conclusion of their study that TCG required near-zero vertical wind shear near the storm center and a strong vertical shear gradient across it. DGP can be calculated as the vertical gradient of relative vorticity, or

$$\text{DGP} = \zeta_{900} - \zeta_{200}, \quad (2)$$

where ζ_{900} and ζ_{200} are the 900-hPa and 200-hPa relative vorticities. Since the NNR does not contain a 900-hPa pressure level, the 850-hPa level was substituted. A higher (lower) DGP indicates a more (less) favorable development environment. To verify the importance of the 200-hPa level contribution to the DGP, we performed classifications using just the 850-hPa relative vorticity as a predictor. We found that the Heidke skill scores (HSSs; Wilks 1995) were 0.1–0.25 higher with the 200-hPa level included, indicating important information was gained from the consideration of both levels.

3) MAXIMUM POTENTIAL INTENSITY

MPI is defined as the theoretical upper limit on intensity that a mature tropical cyclone could attain [see Camp and Montgomery (2001) for a review]. In the Holland (1997) formulation, which is used here, the MPI can also be used as a proxy for SST and the stability of the atmospheric column, accounting for requirements 5 and 6 from above. The Holland MPI is highly sensitive to changes in SST, especially in the ranges typically seen in the tropical Atlantic during the hurricane season (26° – 29°C). The solution is obtained from a columnar temperature profile, a surface temperature (typically $\text{SST} - 1 \text{ K}$), and surface pressure by means of an iterative process. In terms of pressure, lower (higher) MPI

values are thought to correspond to more (less) favorable conditions for development. MPI was chosen in lieu of SST as a thermodynamic predictor because it captures conditions in the upper troposphere that are important for determining the capacity of the system in establishing a warm core [see Holland (1997) for a more thorough discussion].

4) 925–850-hPa MOISTURE DIVERGENCE

To examine the large-scale low-level moisture convergence into the cluster, the moisture divergence at 925 and 850 hPa was calculated, averaged together, and then scaled so that the magnitude was comparable to the COR and DGP predictors. Moisture divergence (MDIV) is calculated as

$$\text{MDIV} = r\nabla \cdot \mathbf{V} + \mathbf{V} \cdot \nabla r, \quad (3)$$

where r is the mixing ratio, and \mathbf{V} is the total horizontal wind. Lower (higher) values indicate more (less) favorable conditions for genesis.

5) PRECIPITABLE WATER

The columnar precipitable water (PWAT; taken directly from the NNR dataset) addresses requirement 4. It is important for the midlevels of the atmosphere to be moist so that the rising saturated air does not evaporate and hence cool the environment. This process will lead to a stabilization of the sounding and the suppression of the additional convection needed to ultimately form a warm core.

6) 24-H PRESSURE TENDENCY

The 24-h pressure tendency (PTEND) predictor does not specifically address a large-scale requirement for genesis; rather, it is included as a diagnostic predictor. Forecasters typically evaluate the pressure changes in cloud clusters when assessing the probability of genesis (R. Pasch 1999, personal communication). In this study we use the 24-h pressure tendency to eliminate the strong diurnal signal in the surface pressure field. Because of the nature of the predictor (information is needed 24 h prior to the analysis time), there were several instances when this predictor was not available for use. Also, it was discovered that the pressure tendency exhibited very little if any differentiation between DV and ND clusters at lead times greater than 24 h before genesis. For these two reasons, pressure tendency was only considered for DV clusters 6–24 h prior to genesis. Those cases were then compared to all of the ND cases for which pressure tendency data were available in the discriminant analysis. PTEND was computed in a Lagrangian frame of reference.

7) 6-H SURFACE AND 700-hPa RELATIVE VORTICITY TENDENCIES

Two other diagnostic predictors are included in the same spirit as pressure tendency: the 6-h surface ($VTEND_{SFC}$) and 700-hPa ($VTEND_{700}$) relative vorticity tendencies. Positive “vorticity spinup” indicates a strengthening of the circulation and thus more favorable conditions for warm-core development. For simplicity in the calculations, the vorticity tendencies, unlike pressure tendency, were computed from an Eulerian frame of reference. This simplification means that the vorticity tendency values may be partially due to the movement of the system rather than vorticity spinup or spindown. Consider a worst-case scenario of a fast-moving easterly wave (10 m s^{-1}). The system would still traverse less distance in a 6-h period (about 216 km) than the grid spacing ($\sim 277.5 \text{ km}$) of the data. Given this and the fact that additional forecast skill was harnessed from the use of the relative vorticity tendency predictors, we believe that this simplification is acceptable.

e. Statistical technique

The sorting of cases into DV and ND groups is accomplished through the use of linear discriminant analysis (DA), of which a thorough examination is given in Tabachnick and Fidell (2001). This technique has been used in several other classification studies in meteorology with success (e.g., Perrone and Lowe 1986; Elsberry and Kirchoffer 1988; DeGaetano et al. 2002). Given a set of normally distributed and independent predictors and their corresponding group membership, a function is derived that maximally separates the groups through a linear combination of the predictor variables. The resulting “discriminant function” can then be applied to an independent set of predictors that the user wishes to classify. An alternative method is to classify each case in the dataset by the “leave one out” method. That is, the discriminant function is computed with all cases except one, and then that case is classified against the function. Then the procedure repeats for all of the cases. This was the method employed in this study. We assumed equal prior probabilities of development. It should be noted that the DA requirement of multivariate normality in this case is *not* satisfied. This does not invalidate the results but may degrade the classification ability of the technique (Tabachnick and Fidell 2001).

The group classification was made in the following manner. The discriminant analysis was run eight times, one for each forecast period (6–48 h). For example, all DV cloud clusters that developed into tropical depressions 6 h after the image time were classified with *all* nondeveloping clusters, regardless of time of year. This produced a set of “6-h classifications.” Then, all clusters that developed into depressions 12 h after the image time were classified with the *same set of nondeveloping clusters*, producing a “12-h classification” set. Intui-

tively, the 12-h forecasts will be somewhat less skillful than the 6-h forecasts since the DV cases at 12 h resemble the ND cases to a greater degree. This effect becomes increasingly prominent at longer forecast times.

4. Results

a. Composite qualitative prediction

The DA output gives a probability (P) of a case belonging to the DV group given the values of its predictors. To evaluate the performance and usefulness of the predictions, we stratified all cases into five bins by this probability value. Table 2 lists each bin, the number of cases that were grouped in that bin by forecast hour, and the number of those cases that developed into a tropical depression. For example, bin 1 lists all of the cases where the DA procedure predicted DV group membership with a probability of 0.9 or higher. For the 24-h forecast period, 24 cases met this criterion. Of those 24, 10 developed into tropical depressions 24 h later, a development percentage of 41.7. For lower P forecasts at the same forecast time, the development percentage drops from 21.4 (bin 2) to 1.1 (bin 5).

Table 2 reveals an interesting aspect of the classification results. The formation rate of tropical depressions is still less than 50% even when the cluster environment is especially favorable (corresponding to probabilistic forecasts $P > 0.9$). If one were to rely on these results alone, the false alarm rate (FAR; Wilks 1995) would be far too great for a skillful forecast. If the probability of development drops below 0.7 (as it is with over 90% of all cases), TCG almost never occurs, and the forecaster could make a rather confident forecast given this situation. If one uses the value of 0.7 as a “decision boundary” between a 1 and 0 forecast, the probability of detection (POD; Wilks 1995) ranges from 64% at the 6-h forecast period to 37% at the 48-h forecast period. The FAR ranges from 4% (6 h) up to 8% for other forecast periods.¹

Given the development percentage values in Table 2, we can categorize the development likelihood given the probabilistic prediction of development by the DA procedure. This is shown in the far-right column of Table 2. If the probability of development $P \geq 0.9$, we label the cluster as having a “good” chance of developing. If $0.8 \leq P < 0.9$, the cluster has a “fair” chance of developing. If $P < 0.7$, development is unlikely to extremely unlikely. These qualitative labels will be used in the case study sections presented later.

¹ The selection of a decision boundary is usually found by calculating a series of forecast skill scores for a range of possible boundaries and then choosing the boundary where the maximum scores occur [see Wilks (1995) for more information]. Theoretically, this would minimize the FAR and maximize the POD further.

TABLE 2. Development rate by probabilistic prediction and forecast hour.

Forecast hour	No. of cases	Developed	DV percentage (%)	DV likelihood
Bin 1 ($P \geq 0.9$)				
6	36	16	44.4	Good
12	27	12	44.4	
18	30	14	46.7	
24	24	10	41.7	
30	29	10	34.5	
36	37	9	24.3	
42	26	8	30.8	
48	18	1	5.6	
Bin 2 ($0.8 \leq P < 0.9$)				
6	16	6	37.5	Fair
12	36	7	19.4	
18	26	6	23.1	
24	28	6	21.4	
30	76	7	9.2	
36	57	10	17.5	
42	44	3	6.8	
48	49	6	12.2	
Bin 3 ($0.7 \leq P < 0.8$)				
6	18	4	22.2	Unlikely
12	43	2	4.7	
18	30	1	3.3	
24	42	0	0	
30	92	6	6.5	
36	85	3	3.5	
42	124	7	5.6	
48	115	5	4.3	
Bin 4 ($0.5 \leq P < 0.7$)				
6	50	1	2.0	Very unlikely
12	106	3	2.8	
18	93	3	3.2	
24	144	2	1.4	
30	326	7	2.1	
36	284	7	2.5	
42	385	7	1.8	
48	471	9	1.9	
Bin 5 ($P < 0.5$)				
6	1108	8	0.7	Extremely unlikely
12	1015	10	1.0	
18	1045	7	0.7	
24	984	11	1.1	
30	1781	9	0.5	
36	1840	9	0.5	
42	1721	10	0.6	
48	1643	11	0.7	

b. Significance of predictors

Although it is difficult to extract significance information from DA predictors, it is possible to examine the strength of the relationship each predictor has with the classification output. Table 3 shows the correlation between each predictor and the output of the classification. Higher values for a predictor indicate a more significant role in determining group membership and hence a greater degree of separation between developing and nondeveloping groups. For all forecast hours, the DGP was by far the most significant predictor, followed

on average by COR (latitude). This was expected (though maybe to a lesser degree) and reemphasizes the importance of the large-scale dynamics to TCG. The weakest predictors overall were the VTEND₇₀₀, MPI, and PWAT. The vorticity tendency signal was relatively weak in most instances and showed little differentiation (as did PWAT) between DV and ND clusters. This result emphasizes the need to consider the vertical change of vorticity (as in the DGP) instead of simply the lower-level vorticity when assessing the potential for TCG. In fact, the correlation was negative at a couple of the forecast times, indicating that little confidence should be given to them. The significance of the MPI was low for nearly all forecast times. This seems to confirm, as shown in Gray (1968) and McBride and Zehr (1981), that for daily TCG prediction during the hurricane season the thermodynamic environment is similar for nearly all candidate cloud clusters. Therefore, proxies for the thermodynamic environment appear to be of little use within this type of prediction system, although predictions did degrade slightly when the MPI was removed as a predictor.²

c. Statistical measures of skill

There are known problems when applying traditional measures of probabilistic forecast performance to rare-event situations such as TCG (Marzban 1998). For example, a better Brier score can be obtained by blindly forecasting nondevelopment for *all* events. Nevertheless, we calculated the threat and Brier scores for the composite dataset for comparison purposes with the Perrone and Lowe (1986) study and to find differences in performance across forecast periods. The threat score, typically used to evaluate qualitative precipitation forecasting, is calculated as

$$\text{THREAT} = \frac{C}{A + B - C}, \quad (4)$$

where A is the number of DV forecasts made, B is the number of DV cases observed, and C is the number of correct DV cases. A threat score of 1 is a perfect score for a suite of forecasts; a score of 0.5 or higher is considered a highly skilled forecast. To compute threat scores, we defined a DV forecast as one in which the probability of DV was at least 0.7 (as discussed previously and similar to Perrone and Lowe's $P \geq 0.65$ boundary). Figure 3 shows the threat scores for the 6–48-h forecast periods. They range from 0.329 at 6 h to 0.059 at 48 h. These values are significantly lower than the scores computed by Perrone and Lowe (1986). This may

² The 1998–2000 seasons were similar in that each had higher than normal SSTs. The MPI predictor may be more useful for seasons in which the tropical Atlantic SSTs were cooler than normal. This may be especially true for early-season systems, when development is typically limited by the thermodynamic environment (DeMaria et al. 2001).

TABLE 3. Correlations between predictor value and classification prediction for each forecast hour. Rank is given in parentheses for each forecast hour.

Hour	6	12	18	24	30	36	42	48	Avg rank
DGP	0.864 (1)	0.815 (1)	0.813 (1)	0.781 (1)	0.830 (1)	0.859 (1)	0.860 (1)	0.880 (1)	1.0
COR	0.279 (3)	0.343 (3)	0.323 (2)	0.260 (3)	0.495 (2)	0.424 (2)	0.432 (2)	0.360 (2)	2.4
PTEND	-0.211 (5)	-0.356 (2)	-0.286 (3)	-0.197 (5)	N/A	N/A	N/A	N/A	3.7
MDIV	-0.232 (4)	-0.130 (5)	-0.281 (4)	-0.415 (2)	-0.201 (3)	-0.94 (5)	-0.100 (4)	-0.128 (5)	4.0
VTEND _{SFC}	0.307 (2)	0.129 (6)	0.194 (5)	0.231 (4)	0.030 (6)	-0.056 (6)	0.100 (4)	-0.079 (7)	5.0
VTEND ₇₀₀	-0.012 (8)	0.242 (4)	0.017 (8)	0.189 (6)	0.136 (5)	0.259 (3)	0.057 (7)	-0.194 (3)	5.5
MPI	-0.260 (7)	-0.075 (8)	-0.160 (6)	-0.097 (7)	-0.145 (4)	-0.110 (4)	-0.263 (3)	-0.116 (6)	5.6
PWAT	0.209 (6)	0.111 (7)	0.130 (7)	0.068 (8)	0.002 (7)	0.016 (7)	-0.076 (6)	-0.131 (4)	6.5

be attributed to differences in the methodologies of the studies and will be discussed further in section 5.

The Brier score (Brier 1950) is a measure of skill for a probabilistic forecasting system. It is essentially the sum of the squared probability errors:

$$BS = \frac{1}{N} \sum_{i=1}^N (f_i - O_i)^2, \quad (5)$$

where N is the number of forecasts, f_i is the forecast probability (from the DA) of the occurrence of the i th event, and O_i is the observed value of the event ($1 = DV$; $0 = ND$). The lower (higher) the Brier score, the more (less) skillful the probabilistic forecast. A Brier score of 0 indicates a perfect suite of forecasts. The Brier scores for each forecast periods are shown in Fig. 3. They are of similar magnitude to Perrone and Lowe

and indicate some value in the probabilistic forecast method derived here.

d. Tropical Depression Keith (2000) case study

We now present two case studies that highlight the strengths and weaknesses of the forecast system. Tropical Depression Keith (2000) was an easterly wave that initially left the west coast of Africa on 16 September. Figure 4 is a plot of the track of the cloud cluster that eventually developed into Keith. This plot illustrates one of the major limitations of the “cloud cluster tracking by convection” method. Convection remained strong across much of the Atlantic but development did not occur. When the convection dissipated, the cluster was declared dead in the database. But the wave remained

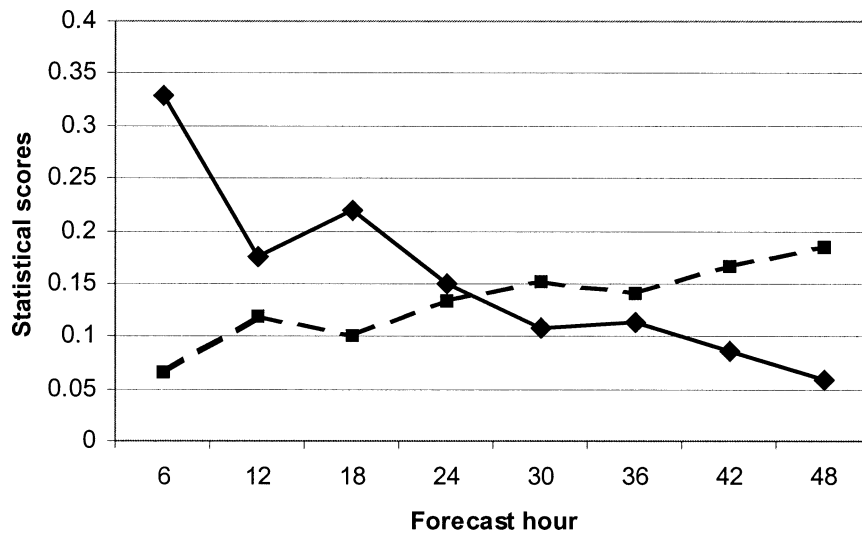


FIG. 3. Threat (solid line with diamond points) and Brier (dotted line) scores for each forecast time. These values reflect the entire dataset of cloud clusters.

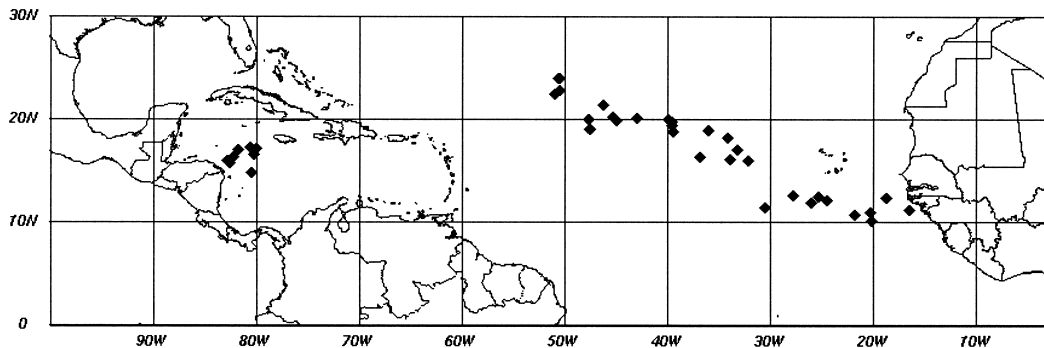


FIG. 4. Track of a tropical cloud cluster that would eventually form into TD 15 (2000), the future Hurricane Keith.

intact, unable to be tracked because of the absence of any significant convection, and by 27 September the wave began showing signs of intensification in the Caribbean Sea as a new convective phase began. This regeneration of convection and subsequent intensification led to its declaration of tropical depression status at 1800 UTC 28 September. This system was treated as two separate clusters until the tropical cyclone report was consulted on the NHC website. We believe this type of situation is a common one, but because of time constraints and the subjective nature of the problem we thought it best to consider these types of situations as separate entities except when alternative sources (such as the NHC tropical cyclone reports) clarify the situation. Note also the irregular jumps in the track of the cloud cluster—this is a typical problem with fixing centers in tropical waves, as the perceived center tends to

“jump” erratically as convection is generated in a different region.

We ran the classification program for each time plotted on Fig. 4—all of the case times in the eastern and central Atlantic were ND cases and those in the Caribbean Sea were DV cases. Figure 5 shows the series of eight 6-h forecasts generated for each image time of the cloud cluster as it progressed across the Atlantic. The shaded region corresponds to a $P(DV) < 0.7$, from which the development likelihood can be interpreted from Table 2 as “unlikely” to “extremely unlikely.” Genesis did not occur during this period, despite the impressive convection at times seen from the IR imagery, because the cluster environment was unfavorable. If we examine the overall trends, we see a general improvement of conditions from 17 September through 20 September, followed by a 2-day deterioration. This was

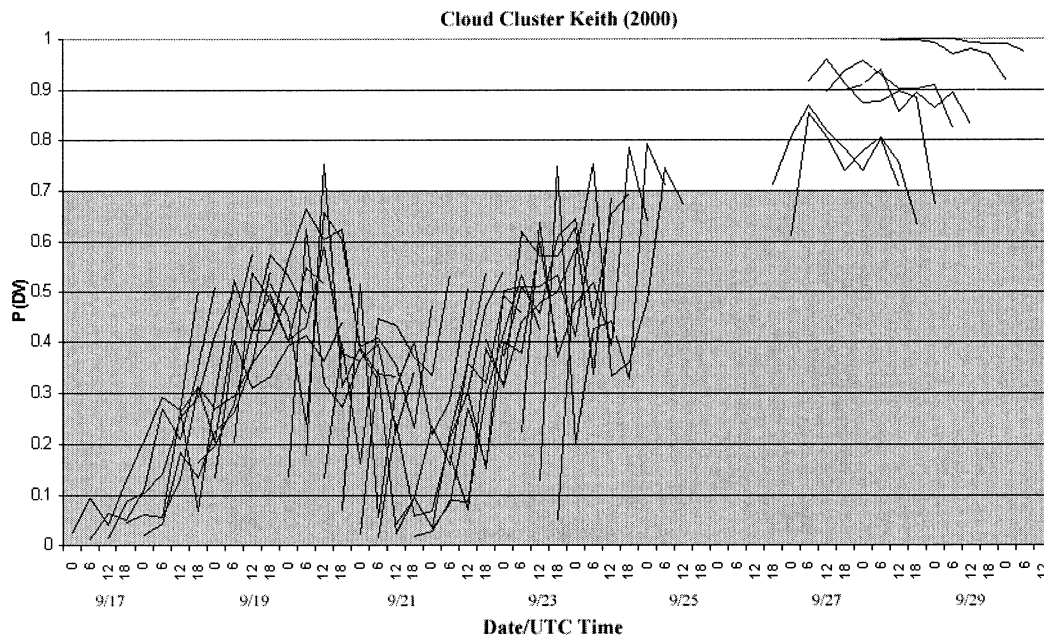


FIG. 5. Series of 6-h probabilistic development forecasts for the cloud cluster that achieved genesis on 1800 UTC 28 Sep (TD 15). The shaded region, which corresponds to development probabilities of <0.7 , delineates the area that corresponds to an “unlikely” to “extremely unlikely” chance of development.

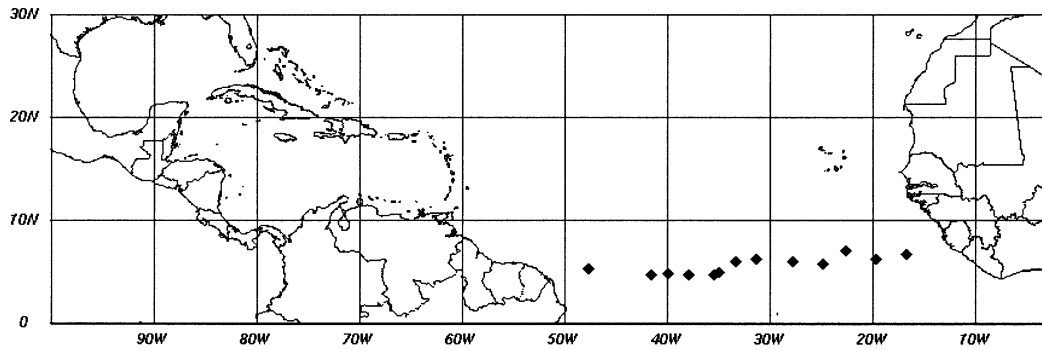


FIG. 6. Track of tropical cloud cluster ND6 (2000) from 1200 UTC 3 Jun to 0600 UTC 7 Jun.

primarily due to unfavorable trends in the vertical wind shear structure and marginal MPI. Convection became so disorganized and sparse by 25 September that the wave became untraceable via IR imagery despite the improving cluster environment. By late on 26 September, convection was regenerated within a much more favorable dynamic and thermodynamic environment, as evidenced by a large majority of the development probabilistic forecasts in the range of 0.7 to 0.99 (“fair” to “good” chance of development). Almost all predictors at this time were favorable, especially DGP, MPI, and PTEND. Genesis occurred at 1800 UTC 28 September; about 48 h after the convection associated with this long-lasting wave first reappeared in the Caribbean.

Keith was a midseason storm that was forecast very well by the system. We have found that the system in general does well with the “prime” development season (August–early October), in which the highest percentages of storms form. However, there were several instances of cloud clusters, especially those that formed early (June) or late (October–November) in the season, that were not handled with as much skill. The next case study reveals one such storm and leads to a discussion in section 6 on strategies that may be employed to improve the forecasting in these situations.

e. ND6 (2000) cloud cluster case study

The sixth cloud cluster of the 2000 season was an easterly wave that moved off Africa with a somewhat disorganized convective structure on 3 June. Figure 6 illustrates the west-southwest track of the cluster over the next few days. During the day on 4 June, convection increased in intensity, and a large ($\sim 5^\circ$ diameter) cloud shield formed. During these early stages of the cluster life cycle, most of the large-scale conditions were favorable for development. The DGP was in the 1.0–2.0 range (the mean DGP for all developing clusters during the 1998–2000 seasons was 0.8–1.4 for 48- to 6 hour forecasts), and the moisture convergence and precipitable water were very favorable (significantly above average) as well. As diagnosed by the MPI, the cluster was over SSTs sufficient for development.

Figure 7 shows that the forecast system recognized the favorable environment at this stage and forecast a high probability for development. The first forecast made at 1800 UTC 4 June shows a “good” chance of development in 24–30 h. The following two forecasts were also predicting genesis around 0000 UTC 6 June. But throughout the day on 6 June, convection decreased, and by 7 June the cluster could not be resolved on the IR imagery. The forecasts during 6 June were significantly less favorable, with most probabilities centered around 0.5, thus showing some indication that the forecast system was sensing a decline in the favorable cluster environment that had been in place.

Why did the system predict a “fair” to “good” chance of development, and yet the cluster did not undergo TCG? The probable answer illustrates the complexity of the process. First, the only large-scale parameter that was marginal/unfavorable during the entire life cycle of the cluster was COR, with values of 0.75–1.93 (average COR for DV systems was about 4.5). As mentioned previously in section 4b, latitude was on average the second most important predictor (i.e., second highest correlation with the predicted result) in the DA. However, the correlation magnitude was only on the order of 0.25–0.40 for each forecast time (compared to 0.70–0.85 for DGP). Thus, if the DGP was high, as it was in this case, it is unlikely the forecast system would predict nondevelopment unless almost all of the other predictors were unfavorable. We believe this may be a limitation of the linearity of the DA; we suspect that a quadratic or nonlinear classification scheme could do a better job in these types of situations. Another reasonable explanation is the neglect of the smaller-scale features and processes within the cloud cluster that cannot be resolved in this type of analysis. Work on incorporating new satellite tools to resolve mesoscale processes is an exciting new avenue for research in this area.

5. Discussion

The reader is encouraged to digest Perrone and Lowe’s TCG work with Pacific cloud clusters (Perrone and Lowe 1986), as there are many parallels to this

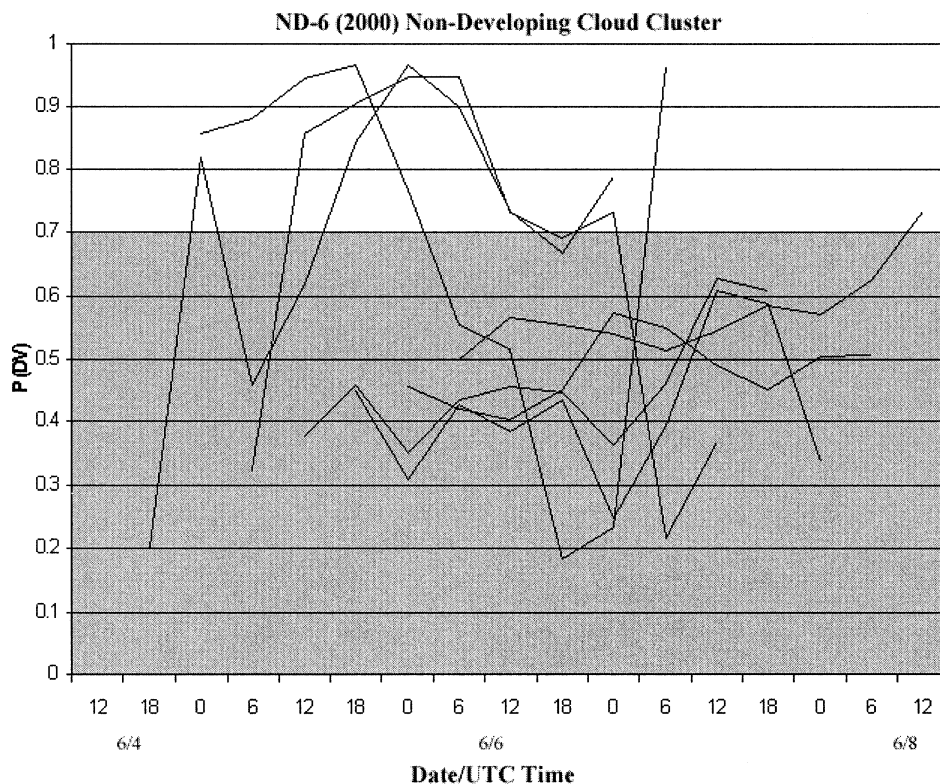


FIG. 7. As in Fig. 5, except for nondeveloping cluster ND6 (2000).

research. But there are several crucial differences that bring about significantly different results, especially in the statistical scores. First, they are predicting tropical storm formation. We think this improved their classification rate since the differences between developing and nondeveloping clusters are likely to be greater during the time the cluster is a depression. In our study, TCG is assumed to occur when a cluster makes the transition to a tropical depression—the transition from a tropical depression to a tropical storm is assumed to be the intensification of a preexisting tropical cyclone. The smaller differences between developing and nondeveloping tropical disturbances make statistical differentiation more challenging, but they form the core of the forecast problem.

Second, Perrone and Lowe used a more relaxed procedure for the selection of their ND cases. They had a size requirement of only 1° diameter, 4 times smaller than the requirement in this study. In addition, it is not clear if their clusters met a persistence requirement nor was there any mention of cloud structures that may have been associated with other systems. If we define a “cloud cluster” as a convective entity that satisfies the condition of “preexisting convection” for TCG, it is quite possible that Perrone and Lowe included clusters in their dataset that we would not consider as such. Although midget typhoons may form from small clusters in the western Pacific (Lander 1994), TCG normally occurs from larger clusters over the Atlantic. The se-

lection criteria of 4° diameters and persistence for 24 h eliminates smaller, more transient features that have little likelihood of developing into a tropical depression. Since DA works best with a high degree of separation between predictors, we think that the inclusion of smaller, transient cloud features might artificially inflate the prediction success rate.

Third, they chose some predictors that are important from a *climatological* point of view rather than those with a day-to-day focus. This may have hurt their model performance, as several of their predictors are not necessarily the best ones for daily genesis studies. For example, vertical wind shear is highly correlated with tropical cyclone formation on a seasonal scale. But as has been shown in several studies (e.g., McBride and Zehr 1981), tropical cyclones frequently form in high-shear environments. That there is near-zero shear *over the cluster itself* is the important criterion. The importance of the DGP in this study clearly validates this statement.

6. Summary, conclusions, and future work

A prediction system for tropical cyclogenesis is developed using large-scale predictors derived from the NCEP–NCAR reanalysis. Using IR satellite imagery for the Atlantic tropical cyclone seasons of 1998–2000, candidate cloud clusters that satisfy predetermined characteristics are identified and their development status determined. Eight predictors are then computed for each

6-h period of the cluster's life cycle. The entire dataset are independently classified using a linear discriminant analysis technique.

Composite results showed that genesis was a good possibility if the probabilistic forecast $P > 0.7$ and unlikely if $P < 0.5$. The discriminant function loadings revealed that the DGP was the most significant predictor at all forecast times, followed by COR. Statistical measures of skill were somewhat lower than a previous study, possibly because of critical differences in the definition of genesis and the selection of cloud clusters. A case study of the cloud cluster that eventually formed into Hurricane Keith (2000) showed that the prediction system does an excellent job at capturing the evolution of the large-scale fields important to TCG, despite the limitations of linearity of the analysis method and uncertainties in the track of the system. A separate look at a nondeveloping cluster highlighted those limitations and emphasized the need for a more detailed look at the smaller-scale features and processes within the clusters.

We believe we have shown that a large degree of predictability of genesis can be derived from a simple analysis of the cloud cluster using large-scale data at the analysis time. We choose predictors that are easy to calculate and data that are regularly available in hourly analysis fields. An interesting future study would be to evaluate TCG forecasts on numerical model *forecast* fields rather than analysis fields, similar to the latest operational version of the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1999). Several concerns would have to be addressed, specifically, how well classifications can be made on pure model data and whether the model can accurately forecast the track and intensity of the convection. We believe one of the biggest deficiencies with the current model is the representation of moisture in the reanalysis data. Making forecasts from model fields rather than analysis fields would magnify this problem.

We believe using improved reanalysis data supplemented by improved moisture fields from satellite sensors will make the biggest improvements in this model. To ascertain any improvement with a nonlinear classifier, we have applied the cloud cluster dataset to a probabilistic neural network. Results were mixed and will be reported at a later time. Ultimately, the application of the system to operational analyses will be the true test of its usefulness. If results prove beneficial, the development of tools such as a graphical user interface will give forecasters near-real-time objective guidance to tropical cyclone formation.

Acknowledgments. The authors wish to thank EUMETSAT and the Space Science and Engineering Center at the University of Wisconsin—Madison for providing the satellite imagery used in this research. We are also grateful for the forecasting perspective of tropical cyclogenesis given to us by Dr. Richard Pasch,

James Franklin, and Stacy Stewart at the Tropical Prediction Center and Commander Marge Nordman at the Joint Typhoon Warning Center. Their insights convinced us that there was a need for this type of research. This paper also benefited from discussions with Drs. Jeffrey Halverson, John Rayner, Caren Marzban, and Hugh Willoughby. We would also like to thank two anonymous reviewers who provided many beneficial comments and suggestions that greatly enhanced this paper.

REFERENCES

- Beven, J. L., 1999: The boguscan—A serious problem with the NCEP medium range forecast model in the Tropics. Preprints, *23d Conf. on Hurricanes and Tropical Meteorology*, Dallas, TX, Amer. Meteor. Soc., 845–848.
- Bister, M., and K. A. Emanuel, 1997: The genesis of Hurricane Guillermo: TEXMEX analyses and a modeling study. *Mon. Wea. Rev.*, **125**, 2662–2682.
- Bracken, W. E., and L. F. Bosart, 2000: The role of synoptic-scale flow during tropical cyclogenesis over the North Atlantic Ocean. *Mon. Wea. Rev.*, **128**, 353–376.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.
- Camp, J. P., and M. T. Montgomery, 2001: Hurricane maximum intensity: Past and present. *Mon. Wea. Rev.*, **129**, 1704–1717.
- Cecil, D. J., E. J. Zipser, and S. W. Nesbitt, 2002: Reflectivity, ice scattering, and lightning characteristics of hurricane eyewalls and rainbands. Part I: Quantitative description. *Mon. Wea. Rev.*, **130**, 769–784.
- Chen, S. S., and W. M. Frank, 1993: A numerical study of the genesis of extratropical convective mesovortices. Part I: Evolution and dynamics. *J. Atmos. Sci.*, **50**, 2401–2426.
- Davis, C. A., and L. F. Bosart, 2001: Numerical simulations of the genesis of Hurricane Diana (1984). Part I: Control simulation. *Mon. Wea. Rev.*, **129**, 1859–1881.
- DeGaetano, A. T., M. E. Hirsch, and S. J. Colucci, 2002: Statistical prediction of seasonal East Coast winter storm frequency. *J. Climate*, **15**, 1101–1117.
- DeMaria, M., and J. Kaplan, 1999: An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337.
- , J. A. Knaff, B. H. Connell, 2001: A tropical cyclone genesis parameter for the tropical Atlantic. *Wea. Forecasting*, **16**, 219–233.
- Elsberry, R. L., and P. J. Kirchoffer, 1988: Upper-level forcing of explosive cyclogenesis over the ocean based on operationally analyzed fields. *Wea. Forecasting*, **3**, 205–216.
- Emanuel, K. A., 1989: The finite-amplitude nature of tropical cyclogenesis. *J. Atmos. Sci.*, **46**, 3431–3456.
- Franklin, J. L., L. A. Avila, J. L. Beven, M. B. Lawrence, R. J. Pasch, and S. R. Stewart, 2001: Atlantic hurricane season of 2000. *Mon. Wea. Rev.*, **129**, 3037–3056.
- Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700.
- Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note 398, 121 pp.
- Holland, G. J. 1997: The maximum potential intensity of tropical cyclones. *J. Atmos. Sci.*, **54**, 2519–2541.
- Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC-22, 21 pp.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kidder, S. Q., M. D. Goldberg, R. M. Zehr, M. DeMaria, J. F. W. Purdom, C. S. Velden, N. C. Grody, and S. J. Kusselson, 2000:

- Satellite analysis of tropical cyclones using the Advanced Microwave Sounding Unit (AMSU). *Bull. Amer. Meteor. Soc.*, **81**, 1241–1260.
- Lander, M. A., 1994: Description of a monsoon gyre and its effects on the tropical cyclones in the western North Pacific. *Wea. Forecasting*, **9**, 640–654.
- Lawrence, M. B., L. A. Avila, J. L. Beven, J. L. Franklin, J. L. Guiney, and R. J. Pasch, 2001: Atlantic hurricane season of 1999. *Mon. Wea. Rev.*, **129**, 3057–3084.
- Lee, C. S., 1989: Observational analysis of tropical cyclogenesis in the western North Pacific. Part I: Structural evolution of cloud clusters. *J. Atmos. Sci.*, **46**, 2580–2598.
- Lowe, P. R., and T. J. Perrone, 1989: An operational technique for prediction tropical storm formation in the western North Pacific Ocean. *Natl. Wea. Dig.*, **14**, 5–13.
- Maddox, R. A., 1981: The structure and life-cycle of midlatitude mesoscale convective complexes. Atmospheric Science Paper 336, Dept. of Atmospheric Science, Colorado State University, 311 pp.
- Marzban, C., 1998: Scalar measures of performance in rare-event situations. *Wea. Forecasting*, **13**, 753–763.
- McBride, J. L., and R. Zehr, 1981: Observational analysis of tropical cyclone formation. Part II: Comparison of non-developing versus developing systems. *J. Atmos. Sci.*, **38**, 1132–1151.
- Montgomery, M. T., and J. Enagonio, 1998: Tropical cyclogenesis via convectively forced Rossby waves in a three-dimensional quasigeostrophic model. *J. Atmos. Sci.*, **55**, 3176–3207.
- Pasch, R. J., L. A. Avila, and J. L. Guiney, 2001: Atlantic hurricane season of 1998. *Mon. Wea. Rev.*, **129**, 3085–3123.
- , J.-G. Jiing, F. M. Horsfall, H.-L. Pan, and N. Surgi, 2002: Forecasting tropical cyclogenesis in the NCEP global model. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 178–179.
- Perrone, T. J., and P. R. Lowe, 1986: A statistically derived prediction procedure for tropical storm formation. *Mon. Wea. Rev.*, **114**, 165–177.
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.
- , and T. M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, **7**, 929–948.
- Riehl, R. J., 1948: On the formation of typhoons. *J. Meteor.*, **5**, 247–264.
- Ritchie, E. A., and G. J. Holland, 1997: Scale interactions during the formation of Typhoon Irving. *Mon. Wea. Rev.*, **125**, 1377–1396.
- Sharp, R. J., M. A. Bourassa, and J. J. O'Brien, 2002: Early detection of tropical cyclones using SeaWinds-derived vorticity. *Bull. Amer. Meteor. Soc.*, **83**, 879–889.
- Simpson, J., E. Ritchie, G. J. Holland, J. Halverson, and S. Stewart, 1997: Mesoscale interactions in tropical cyclone genesis. *Mon. Wea. Rev.*, **125**, 2643–2661.
- Tabachnick, B. G., and L. S. Fidell, 2001: Discriminant function analysis. *Using Multivariate Statistics*, R. Pascal, Ed., Allyn and Bacon, 456–516.
- Trenberth, K. E., D. P. Stepaniak, J. W. Hurrell, and M. Fiorino, 2001: Quality of reanalysis in the Tropics. *J. Climate*, **14**, 1499–1510.
- Tuleya, R. E., 1991: Sensitivity studies of tropical storm genesis using a numerical model. *Mon. Wea. Rev.*, **119**, 721–733.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.