Objective Classification of Atlantic Hurricanes

J. B. ELSNER, G. S. LEHMILLER, AND T. B. KIMBERLAIN

Department of Meteorology, The Florida State University, Tallahassee, Florida

(Manuscript received 29 January 1996, in final form 28 April 1996)

ABSTRACT

Previous groupings of Atlantic tropical cyclone activity into baroclinically influenced and tropical-only hurricanes have required subjective evaluations. In this paper, a set of statistically significant and valid rules are introduced that objectify this previously subjective evaluation. This is done with the aid of classification trees. The tree classifications are better than 90% accurate with respect to an earlier subjective discrimination. Objective classification rules are the basis for a climatology of Atlantic hurricanes. The average latitude of origin for tropical-only hurricanes is 18.8°N, compared to 29.1°N for baroclinically influenced storms. The baroclinically influenced hurricane season extends from-mid May to December, while the tropical-only season is largely confined to the months of August through October. There is a fairly abrupt shift to fewer numbers of tropical-only hurricanes around 1960.

1. Introduction

The annual number of hurricanes in the Atlantic varies from year to year. Long-lead forecasts of the number of storms are routinely issued (e.g., Elsner et al. 1994; Gray 1994). Forecast skill comes from variables of the large-scale tropical environment, including the stratospheric winds, the state of the El Niño–Southern Oscillation, and meteorological anomalies over western Africa and over the Caribbean (Gray et al. 1992, 1993, and 1994). This is to be expected since many Atlantic hurricanes have origins in tropical easterly waves.

Hurricanes can form from other sources as well; it is sometimes observed, for example, that frontal intrusions from latitudes north of the Tropics provide organization for initiation of tropical depressions. In addition, baroclinic disturbances sometimes serve to intensify an otherwise benign tropical depression. It thus seems natural to consider the hurricane season as the sum of tropical-only and baroclinically influenced storms.

In fact, Hess et al. (1995) show a significant increase in hindcast skill when using prediction models that incorporate both types of storms separately. To predict the annual number of hurricanes, an ordinary least squares (OLS) linear regression is used to estimate the number of tropical-only hurricanes ($\hat{H}_T$), and a seasonal average number of baroclinically influenced hurricanes ($\bar{H}_B$) is added to this. The model can be expressed as

$$\hat{H} = \hat{H}_T + \bar{H}_B.$$

To develop their forecast model, Hess et al. (1995) classified hurricanes over the period 1950–93. Stratification was done by examining the summaries of past tropical cyclone seasons published in *Monthly Weather Review*. These reviews describe the life cycle of each tropical storm individually for each season. Typically the descriptions are adequate to decide whether baroclinic influences were a factor. Other sources included daily synoptic charts and consultation with N. LaSeur, who flew on many of the pre–satellite era reconnaissance missions.

Figure 1 shows the results of Hess et al.'s (1995) stratification, where the ○ is the initial location of a tropical only hurricane and + is the initial location of a baroclinically influenced storm. A line of latitude near 20°–23°N offers a good first guess at objectively dividing the two groups, particularly over the eastern Atlantic. Note, however, that this simple division does not work over the Caribbean or the Gulf of Mexico, where it appears that more rules are needed.

The limitation of this work is that the separation of storms was done subjectively. The method of Hess et al. (1995) was subjective in that it was done by an individual (in consultation with others) looking at each storm separately. This does not mean that there are no underlying physical mechanisms supporting the classification. On the contrary, as shown in Hess et al. (1995), the subjective classification is consistent with some underlying physical processes. Further, it does not mean that someone else looking at the same resources would produce an identical classification. That is, there will always be debate about the classification of some storms.

Corresponding author address: Dr. James B. Elsner, Department of Meteorology, The Florida State University, Tallahassee, FL 32306-3034.

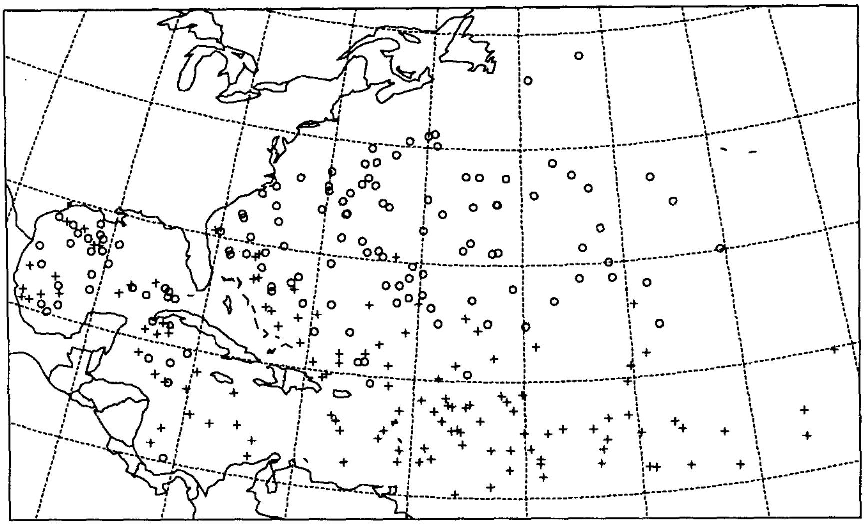

FIG. 1. Location where each of the 255 tropical cyclones over the period 1950–93 reached hurricane strength for the two groups of baroclinically influenced (O) and tropical-only (+) hurricanes, after Hess et al. (1995).

What the aforementioned limitation implies is that without some kind of objective classification other researchers will be reluctant to completely embrace this kind of storm separation because they lack confidence in the uniqueness or repeatability of the classification. Thus, a set of rules is developed here that will classify hurricanes into the two groups in an objective way. The rules classify at better than 90% accuracy and allow for a climatology over the period 1886–1994. The rules will provide guidance for classifying future Atlantic hurricanes, but the best classification will result from consideration of all data. We emphasize that this work provides an objective classification for approximating a particular subjective grouping, which is different from constructing an objective discrimination from true classifications.

The paper is outlined as follows. The data used in the present study are described in section 2. Section 3 provides first a rationale for using a tree-based regression on this problem and then a description of the classification methodology. Statistics and results of the procedure are given in section 4. A comparison climatology of tropical-only versus baroclinically influenced hurricanes is given in section 5.

2. Data

Data for this study were obtained from Atlantic tropical cyclone "best" track and intensity records managed by the Tropical Prediction Center (formerly the National Hurricane Center, Jarvinen et al. 1984), where best refers to an accurate assessment of storm location based on a postanalysis of available data. The dataset extends back to 1886 and includes all tropical cyclones that reached tropical storm strength. Each storm has latitude and longitude coordinates and maximum sustained winds every 6 hours during the storm's existence. Data are most reliable after 1944 when the U.S. Air Force began aircraft reconnaissance missions to investigate individual storms over the Atlantic.

Atlantic tropical cyclones are grouped into four stages. A depression is defined as a closed low-level wind circulation with sustained speeds of generally between 10 and 18 m s^{-1}. A tropical storm is a tropical cyclone with maximum sustained low-level winds between 18 and 32 m s^{-1}. A tropical cyclone is designated as a hurricane when sustained winds exceed 32 m s^{-1}, and it is called an intense (or major) hurricane when the winds exceed 50 m s^{-1}.

The idea is to develop a set of classification rules for the hurricanes originally grouped subjectively by Hess et al. (1995). The classification will be most reliable for storms with data over their complete morphology, including the depression stage. As such, we exclude storms in the 1950–93 period for which there are no records of a tropical depression. This leaves a total of 209 storms out of the 255 originally grouped. Of the 46 storms not included, 28 were tropical only and 18 were baroclinically influenced. These numbers reflect the fact that missing depression-stage data is more likely for earlier storms in the record. As an example, Hurricane Hazel in 1954 was not included because the best track data does not contain information on this storm until it had maximum sustained winds of 60 knots.

Table 1 is a random sample of the hurricanes used in the initial classification. Each storm contains the Julian day, the latitude, and the longitude for initial depression and initial hurricane stages, that is, the day and position for which the storm was first reported as a

TABLE 1. A small sample of the hurricanes and data used in the objective classification study, where day D and day H are the Julian days on which the storm first reached depression and hurricane strengths, respectively; long D and long H are the initial depression and hurricane longitudes (°W); respectively; lat D and lat H are the initial depression and hurricane latitudes (°N); and H_T and H_B are tropical-only and baroclinically influenced hurricanes, respectively, according to Hess et al. (1995).

Year	Name	Day D	Long D	Lat D	Day H	Long H	Lat H	Category
1951	Charlie	224	45.7	12.2	228	62.5	15.4	H_T
1952	Fox	294	77.6	11.8	296	82.2	16.8	H_T
1956	Greta	304	75.5	17.8	309	69.6	23.3	H_B
1957	Carrie	245	21.7	13.0	248	32.6	14.5	H_T
1959	Flora	252	45.8	16.8	254	41.3	28.7	H_B
1963	Beulah	232	49.5	13.7	234	56.9	17.9	H_T
1964	Gladys	257	44.3	14.7	258	52.1	18.8	H_T
1966	Lois	308	50.0	26.5	312	49.8	24.8	H_B
1969	Blanche	223	71.7	28.1	223	69.9	35.5	H_B
1971	Ginger	249	71.5	25.5	254	63.3	27.9	H_B
1974	Fifi	257	65.0	15.3	260	80.2	16.6	H_T
1977	Dorothy	269	71.5	28.5	271	59.7	35.5	H_B
1979	Gloria	247	21.0	15.5	250	37.2	24.4	H_B
1980	Allen	213	30.0	11.0	216	51.4	12.4	H_T
1985	Elena	240	74.0	19.8	241	85.0	25.0	H_T
1988	Gilbert	252	54.0	12.0	255	66.8	15.9	H_T
1989	Hugo	253	20.0	13.2	256	43.5	12.8	H_T
1993	Harvey	261	61.8	26.7	263	55.2	35.6	H_B

tropical depression and a hurricane, respectively. Thus, there are six *independent* variables for each hurricane. The Hess et al. (1995) classification is also listed, where H_T is a tropical only hurricane and H_B is a baroclinically influenced storm. This grouping defines the *dependent* variable used to build a classification tree.

Figure 2 shows the prehurricane track of the baroclinically influenced Hurricane Arlene of 1967 and of the tropical-only Hurricane Hugo of 1989. In general, tropical-only hurricanes originate from Cape Verde waves and maintain a westward motion, as was the case with Hugo. Baroclinically influenced hurricanes, on the other hand, tend to have a northward component and often do not reach hurricane strength until after curvature to the north (recurvature), like Arlene.

3. Classification method

To develop an effective set of classification rules for our purposes, the method should have several charac-

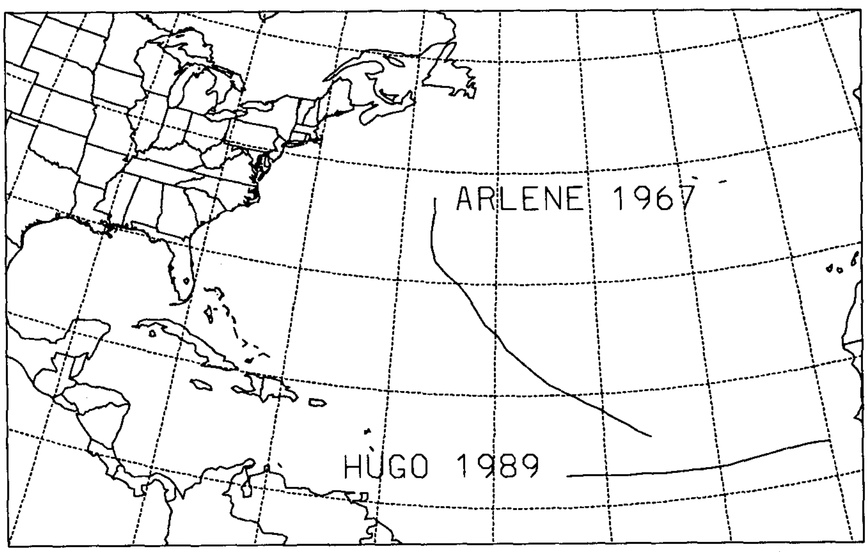

FIG. 2. Tracks of Hurricanes Hugo (1989) and Arlene (1967) during their tropical depression and tropical storm stages. Hugo is classified as a tropical-only and Arlene as a baroclinically influenced storm, according to the subjective classification of Hess et al. (1995).

teristics to ensure its validity. Among the most important of these considerations is that the methodology should allow for statistical significance testing by way of cross validation, allow for nonfunctional relationships between predictor variables and the classification rules (algorithm), and provide useful and easily interpretable results. Purely empirical classification techniques, such as linear programming methods, do not allow for statistical validation of the results, while purely statistical methods do not easily allow for nonfunctional relationships between the predictors and the prediction groups.

Hence, to create a set of objective classification rules for the subjective classification developed by Hess et al. (1995), we employ a statistical classification algorithm known as partially adaptive classification trees (PACT, Shih 1993). PACT unifies the multivariate statistical methodology of linear discriminant analysis (LDA, Mardia et al. 1979) and tree-structured classification methods (CART, Breiman et al. 1984). As will be discussed, PACT combines the advantages of both methodologies and meets the desired criteria specified above. We note that the classification method algorithm chosen here is not unique; however, it is quite simple to implement and yields satisfactory results for our purposes. For instance, the PACT algorithm halved the LDA error rate for our classification problem, as discussed in the results section. Readers wishing to investigate other classification methods are encouraged to refer to Brieman et al. (1984) and Hand (1981).

A brief review of LDA and another statistical technique known as analysis of variance (ANOVA, Casella and Berger 1990) is needed in order to understand how the PACT algorithm creates its classification rules. Linear discriminant analysis is a multivariate statistical technique that seeks to classify an observation into a group or category according to the observed values of several associated predictor variables. The choice of a linear discriminant function (LDF) depends upon the nature of the data involved. The most commonly used LDF assigns group classifications by using a generalized distance function (the Mahalanobis distance) that measures the distance of the values of the predictor variables, corresponding to an observation, to the means of those predictor variables for each classification group (Mardia et al. 1979). An observation is then assigned to that group for which its distance measure to the group mean (the centroid) is the smallest.

ANOVA is a technique to determine how much a measured response variable, or variables, changes according to different group classifications and to ascertain the corresponding statistical significance (Casella and Berger 1990). This methodology uses least squares techniques to estimate the sources of variances, so that a single test statistic (the F statistic) can indicate the statistical significance of the variance caused by the group classifications. PACT also employs another statistical technique that measures how much variance is caused by the group classifications. Known as Levine's test, it uses techniques that are based on ANOVA methods. Levine's test is quite robust and formally tests for equality of group variances in continuously valued data. Like ANOVA, Levine's test also creates the F statistic as its single test statistic.

PACT itself functions by emulating the decision trees created by CART. A decision tree is a set of sequential rules that one follows in order to classify an observation. The name itself comes from the appearance of the rules as written on a sheet of paper, which is somewhat similar in appearance to a flow chart. Within a decision tree, each time a decision (or classification rule) is to be performed, we are at what is called a decision node. The result of the decision, true or false, shunts the decision into a choice of two other nodes, which themselves may be either more decision nodes or what are known as terminal nodes. A group classification is assigned for each terminal node. Following this procedure, we begin at the first decision node (the top of the tree) and ultimately finish in a terminal node at some part of the tree.

For an example of this, and to illustrate the major advantage of PACT over LDA, refer to Fig. 3. Here, an artificially created dataset shows a separation of H_T and H_B by longitude and latitude. An optimal set of classification rules would stratify the variable space (here, just the regions) in the simplest manner possible, so that we could accurately classify every single observation. For the case here, the stratification by category is not functional, that is, there is no linear discriminant function that can divide this region into the proper subregions for H_T and H_B. In other words, since we have two categories (H_T and H_B) and two predictor

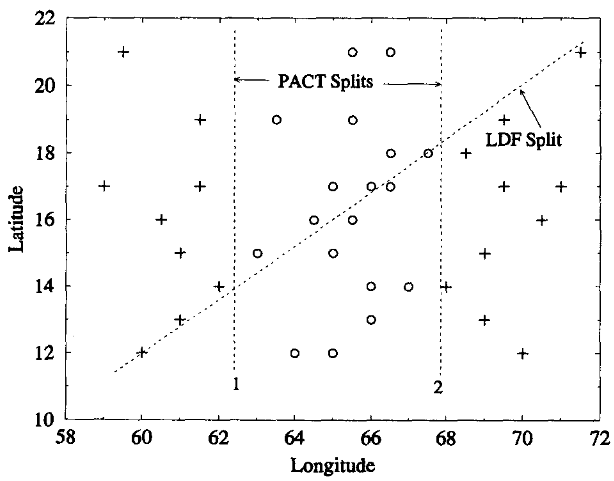

FIG. 3. A contrived example showing the difference in results between a PACT and an LDA classification scheme. Locations of baroclinically influenced hurricanes are denoted by circles and locations of tropical-only hurricanes by pluses. The splits from the PACT algorithm are represented by two vertical lines, and the split from the LDA algorithm is represented by the LDF line.

variables (latitude and longitude), LDA is limited to separating the regions by the best straight line that can be drawn in the plane. Here, that line is the dashed diagonal line; the classification accuracy achieved by the LDA method is only 0.487.

In contrast, PACT is not limited by the nonfunctional relationship. The PACT algorithm here produces a decision tree with two decision nodes and three terminal nodes and achieves a classification that is 100% accurate. The first decision node is represented by the dashed line labeled split 1. At this decision node, the question is asked, Does the observation have a longitude of less than 62.3? If yes, than we are shunted to a terminal node with a tropical-only label. Otherwise, we are shunted to the second decision node, represented by the dashed line labeled split 2, which determines whether or not longitude is greater than 67.5. If yes, then the observation is shunted to a terminal node and classified as tropical only; otherwise it is shunted to the other terminal node and classified as baroclinically influenced. Note that the PACT algorithm ignores latitude entirely in its decision process as it contains no useful information, while the LDA spuriously uses latitude in its region separation, separating the regions with a line that is constructed as a linear combination of the two predictor variables. If latitude had contained useful information, then PACT would have also used this at a decision node, but in a univariate fashion. That is, PACT separates the regions in a univariate fashion so that we do not have to evaluate linear combinations. This is particularly useful in high dimensional datasets.

PACT creates its classification rules by using a hybrid of several statistical methods. The procedure commences by creating an initial decision node and then adding further nodes as constrained by the tree growth parameters. Since it is possible to always create a 100% classification accuracy by completely partitioning the predictor space, a criterion is needed to determine the optimal tree size. Here, this was achieved by using a direct stopping rule and then maximizing the cross-validated classification accuracy as a function of the stopping rule. A direct stopping rule stops the tree growth process once the number of observations remaining within a terminal node falls below a certain percentage threshold of the total number of observations. In other words, suppose that a direct stopping rule of 6% was chosen. Then, if the number of observations in a particular node is less than 6% of the initial total, the growth process is stopped for that node and it is assigned as terminal.

The algorithm functions as follows. First, if the initial or any subsequent node has a sufficient number of observations, the algorithm performs an ANOVA on each potential predictor variable and selects the variable that has the most significant F statistic. To avoid ignoring variables that have a large degree of nonfunctional group separation, Levine's test is also conducted to identify which variable has the largest inequality of variances caused by group classification. The F statistics for this test are also obtained. PACT selects the splitting variable for the decision node based on the variable having the largest F value over both test procedures.

Next, the algorithm performs a one-dimensional linear discriminant analysis, using the variable selected above. The decision rule for the decision node in question is created from the LDF, which partitions this node into two new (sub)nodes. Finally, each of these nodes is checked to see if it has a sufficient number of observations, and the process is repeated until all of the remaining nodes become terminal nodes, thus completing the tree.

The classification tree, once completed, allows for rather straightforward group classifications. While the rules strictly create a yes/no classification assignment, probabilities of assignments may also be estimated by one of two ways. One method is simply to note the observed classification error for the corresponding terminal node and calculate the group assignment probability as one minus the node misclassification error. Since this method ignores the actual values of the predictor variables outside of the classification cutoffs, another method consists of obtaining the corresponding group classification probabilities for the LDF used at each involved decision node and then using conditional probabilities to estimate the group assignment probabilities. Note that since LDA is technically a Bayesian classifier (Mardia et al. 1979), Bayesian prior probabilities may be utilized in the LDFs for each decision node. The PACT algorithm allows for this; however, we have not made use of prior probabilities in this study.

4. Results

The above procedure is applied to the dataset consisting of six independent variables for each of the 209 hurricanes to obtain a classification tree with eight decision rules and 9 terminal nodes. The PACT algorithm gives a relative ranking of importance for independent variables based on the reduction of variance when the variable appears either in decision rules or as a surrogate. The three most important variables in order of importance are initial depression longitude, initial depression latitude, and initial hurricane latitude.

The decision rules are listed in Table 2. The rules are used on the 209 hurricanes to get an in-sample (resubstitution) error of 0.077, which means there are 16 storms classified incorrectly by the classification tree. The cross-validated error is 0.083, implying that out-of-sample classification of hurricanes using the above rules will have an accuracy exceeding 90%.

As suggested previously, there are different ways to construct classification rules. For example, LDA, which is based on linear regression, is often used in classification problems. So for a comparison we apply

TABLE 2. Objective classification rules for determining hurricane type (either tropical only H_T or baroclinically influenced H_B) based on initial depression and initial hurricane data, where lat H is the latitude at which the storm reached hurricane strength, and lat D and long D are the latitude and longitude at which the disturbance reached depression strength, respectively.

Rule number	Rule	Action
1	Is lat H < 23.5°N?	If yes, rule 2, else, rule 3.
2	Is long D < 68.7°W?	If yes, $H = H_T$, else, rule 4.
3	Is lat D < 19.85°N?	If yes, rule 5, else, $H = H_B$.
4	Is lat D < 20.4°N?	If yes, rule 6, else, $H = H_T$.
5	Is lat D < 17.35°N?	If yes, rule 7, else, $H = H_T$.
6	Is lat D < 12.6°N?	If yes, $H = H_T$, else, $H = H_B$.
7	Is lat H < 26.4°N?	If yes, rule 8, else, $H = H_B$.
8	Is long D < 72.4°W?	If yes, $H = H_T$, else, $H = H_B$.

a LDA (with equal prior probabilities) on the same dataset. In this case, the in-sample error is 0.158, which is more than double the error of the classification tree. This result should not be generalized to mean that tree-based regressions are always the best way to proceed. In this particular case, where the possibility exists for nonfunctional relationships between the independent variables and the groups, discriminant analysis is not the best choice. Indeed, we stress that the particular choice of classification methodology is not critical to the overall approach of an objective classification of Atlantic basin hurricanes.

To assess the statistical significance of this result, we employ a normal approximation on the cross-validated accuracy proportion, as opposed to the best accuracy that could be obtained by climatology alone. In this case, there are 112 baroclinically influenced and 97 tropical-only storms, so a climatological classification accuracy of 112/209 = 0.536 is the best that could be achieved blindly. We then imploy the normal approximation

$$z = \frac{a_p - a_c}{\sqrt{a_c(1 - a_c)/n}},$$

where a_p and a_c are the accuracies of the algorithm and climatology, respectively, and $n = 209$ is the number of cases. Using this, a value $z = 11.22$ is obtained. Since the algorithm iterated over 20 potential tree sizes, we multiply the corresponding p value by 20 to correct for selection bias. Doing so yields a p value of less than 10^{-4}, indicating significant results. This lends credence not only to the objective classification algorithm, but it also supports the initial subjective stratification, since largely independent datasets were used for the stratification and for the algorithm.

Further, we find that 66 (or 32%) of the hurricanes are classified as tropical only by simply applying rules 1 and 2 (type 1 hurricane), while 75 (or 36%) are classified by rules 1 and 3 (type 2 hurricane). Thus, nearly 70% of all storms can be classified as either a type 1 or a type 2 hurricane. Hurricanes Hugo of 1989 and Arlene of 1967 (with their development tracks shown in Fig. 2) serve as examples of type 1 and type 2 storms, respectively. Additional rules are needed to classify the remaining 30% of the hurricanes.

Table 3 is a list of the hurricanes that were misclassified by PACT. We find that some of these storms posed problems for the original classification. We apply the above rules to the 1994 hurricane season and determine that, of the three Atlantic hurricanes, only Chris reached hurricane strength devoid of any baroclinic influences. This is consistent with the subjective interpretation of the 1994 season by Elsner et al. (1994), but may not represent a consensus among the researchers in this area. For example, it might be argued that Hurricane Chris was aided in its early development by the proximity of an upper-level low.

It is interesting to speculate on the physical importance of these results. The fact that a judicious screening of storms by an individual can be reduced to a relatively few objective rules, with relatively high degree of accuracy, suggests some underlying simplicity to the apparent complexity of hurricane development. We speculate that perhaps the role of midlatitude westerlies may be more important in differentiating various physical mechanisms of hurricane development than has been previously considered. For example, the late season hurricane activity is probably a consequence of the seasonal return of high-amplitude midlatitude baroclinic disturbances superimposed over the still warm waters of the Atlantic.

The rules generated above will guide the classification of future hurricanes but may not be useful for grouping storms prior to 1951, since in general, depression data are not available for these earlier storms. Therefore, we develop another classification tree using

TABLE 3. Hurricanes misclassified by the eight-rule PACT.

Year	Name	Hess et al. (1995)	Objective
1951	How	H_B	H_T
1954	Carol	H_T	H_B
1954	Alice2	H_B	H_T
1955	Gladys	H_T	H_B
1962	Daisy	H_B	H_T
1965	Betsy	H_B	H_T
1966	Celia	H_B	H_T
1968	Gladys	H_T	H_B
1970	Alma	H_B	H_T
1976	Holly	H_B	H_T
1979	Gloria	H_B	H_T
1982	Debby	H_B	H_T
1985	Bob	H_T	H_B
1988	Debby	H_T	H_B
1989	Chantal	H_T	H_B
1992	Andrew	H_B	H_T

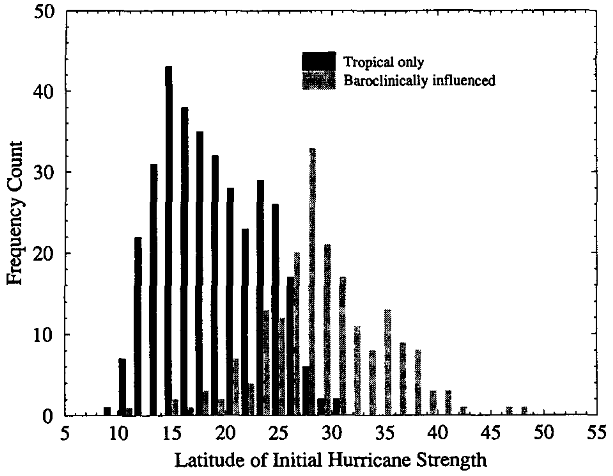

FIG. 4. Frequency of hurricanes as a function of latitude for tropical-only and baroclinically influenced storms compiled over the period 1886–1994. The years from 1950 onward were classified subjectively, and the earlier years were classified using the rules similar to those of Table 2.

only data after the cyclone reached tropical storm strength ($18 \mathrm{~m~s^{-1}}$).

We use the same six independent variables on 257 hurricanes over the period 1950–94, with the exception that initial depression data are replaced by initial tropical storm data. The results are similar to the original classification in that the PACT algorithm gives a classification tree with nine decision rules and 10 terminal nodes. The resubstitution error is 0.074 (or 19 misclassified storms), and the cross-validated error is 0.097. Unlike the classification tree built from the depression data, this tree contains a slight bias toward tropical-only hurricanes. There does not appear to be any time-dependent trend in the bias, however. The most important independent (or predictor) variables in order of importance are initial hurricane day, initial hurricane latitude, and initial tropical storm latitude. These rules are used to classify hurricanes prior to 1950.

5. Climatology

We apply the decision rules to the historical best track data over the period 1886–1949. There were a total of 279 hurricanes in this period. Because a few of these storms were not detected until hurricane strength, the initial storm and initial hurricane data are identical (e.g., the hurricane of 1890). While colinearity (correlation among the independent variables) is not a problem with classification trees as it is with other statistical models, the missing data will add a bit of uncertainty to the decision of grouping these particular hurricanes.

Over this period we find 217 tropical-only hurricanes, or 78% of the total. The percentage of tropical-only hurricanes is considerably greater in this period than in the 1950–94 period. Combining the classifications from the two periods we have a total of 342 tropical-only hurricanes, or 64% of the total number.

The average latitude at which a tropical-only storm initially becomes a hurricane is 18.8°N, while the average longitude is 65.9°W. This compares to an average latitude of 29.1°N and an average longitude of 69.1°W for baroclinically influenced storms. Figure 4 shows the distributions of the initial hurricane latitude for the two groups of storms. Tropical-only hurricanes not only tend to form at lower latitudes than baroclinically influenced storms, but also the distribution as a function of latitude is skewed toward lower latitudes. This is in contrast to baroclinically influenced hurricanes, which have a symmetric distribution with respect to their latitude of origin.

The seasonal variability of tropical-only and baroclinically influenced hurricanes is shown in Fig. 5. The greatest concentration of hurricane activity occurs from the middle of August through September. The baroclinically influenced season is longer, extending from June to November, while the tropical-only season is generally from August to October. In November, midlatitude disturbances become even more frequent over the still relatively warm tropical Atlantic waters. Occasionally, these disturbances initiate hurricane development.

The secondary, postmaximum peak in total Atlantic hurricane activity during the middle of October has been suggested by others (e.g., Cry and Haggard 1962; Neumann et al. 1987; Landsea 1993), but there has yet to be an explanation. Cry and Haggard (1962) state, "An increase of both tropical storms and hurricane

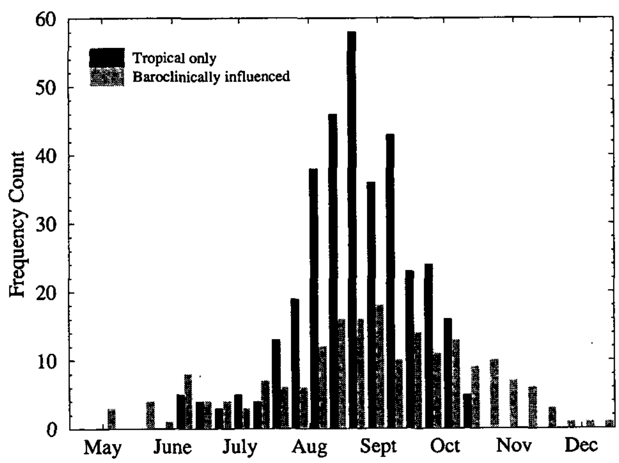

FIG. 5. Frequency of hurricanes as a function of day of the year for tropical-only and baroclinically influenced storms compiled over the period 1886–1994. The years from 1950 onward were classified subjectively, and the earlier years were classified using the rules similar to those of Table 2. Tick marks represent the middle of each month.

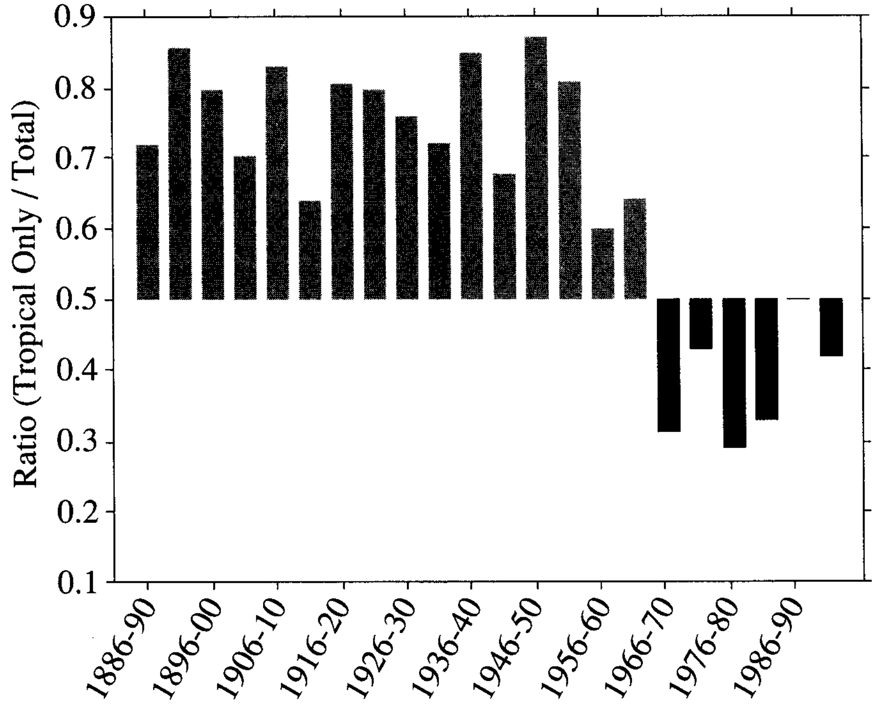

FIG. 6. Ratio of the number of tropical-only hurricanes to the total number in 5-year intervals beginning with 1886 for the Atlantic basin (1886–1995). Tropical-only hurricanes were determined subjectively from 1950 onward and objectively using rules similar to those in Table 2 for years over the period 1886–1949. Tic marks correspond to every second pentad. Ratios greater than 0.5, indicating more tropical-only hurricanes, are shown in ( ) shading, and ratios less than 0.5, indicating more baroclinically influenced hurricanes, are shown in ( ) shading.

days in mid-October is followed by a sharp decline to a low level of activity in the last part of October and in November.'' Neumann et al. (1987) state, ''A somewhat irregular decline in frequency occurs thereafter [after mid-September], interrupted by a slight increase in mid-October.'' Here, we suggest that this secondary peak is due to the relative maximum in baroclinically influenced activity at this time of year.

Multiyear variations in tropical-only hurricane activity are examined by considering the ratio of the number of tropical-only hurricanes to the total number of hurricanes (H_T/H). This idea is not without precedent. Simpson et al. (1968) were the first to catalog all synoptic-scale tropical disturbances each season. Herbert and Frank (1973) added to these efforts by creating a ratio of the baroclinic depressions to the total number of depressions. A ratio of less than 0.5 suggests a season with a ''tropical'' character, that is, more developments over the tropical belts. Ratios closer to one signal increased baroclinic activity. Avila and Clark (1989) revised this ratio by taking the total number of tropical storms originating in Africa to the total number to get a better assessment of seasonal tropical cyclone activity. They subjectively separated ''African'' years from ''non-African'' years in this fashion. If the ratio is greater than or equal to 0.7, then it is an African year. If it is less than or equal to 0.5, then it is a non-African year.

For our study, it is believed that the ratio H_T/H is more robust than the actual number of tropical-only hurricanes against changes in hurricane-detection methods over the past century of observations (see Landsea 1993). This is because some early hurricanes may have gone undetected, a situation not likely in the most recent years. We recognize that a potential bias may exist toward one type of storm or the other over this time period. While the ratio may prove more robust against inconsistent biases, it is impossible to determine whether or not the pre-1950 hurricane record contains a bias toward storm type. However, no evidence exists to indicate that it is in fact biased. The link between low-frequency variability of tropical-only hurricanes and intense hurricane activity described below gives some support to the conjecture that the ratio is not biased over this period.

Figure 6 shows the time series of this ratio using 5-yr intervals beginning in 1886. The last interval includes hurricanes through 1995. The striking feature of this graph is the abrupt change in ratios around 1960. Between 1886 and 1960, tropical-only hurricanes dominated annual totals, averaging nearly three-quarters of all storms. In stark contrast, over the past 30 years or so baroclinically influenced hurricanes have accounted for nearly 70% of the annual total. It is noted that this change occurred during the reliable portion of the hurricane record.

The above result is similar to the decrease in the number of intense hurricanes noted by Gray and Landsea (1992) during the 1970s and 1980s compared with the decades of the 1940s and 1950s. In fact, the findings are probably related, since the empirically derived conditional probability of designating a hurricane as tropical only given that it reached intense hurricane strength is 0.78. In addition, since 66% of all tropical-only hurricanes become intense hurricanes, the decrease in intense hurricane activity associated with recent prolonged drought conditions in western Africa (Landsea 1993) might be used to partly explain the decline in tropical-only activity. There may be other factors, however, since the linear correlation between the number of tropical-only hurricanes and the number of baroclinically influenced hurricanes in these 5-year intervals over the period is -0.47, which is statistically significant at the 5% level.

We note that the rather sudden drop in the relative number of tropical-only hurricanes occurred during the period in which satellite information was beginning to be routinely consulted. Thus, there is the possibility that the change in ratio resulted from more and/or better information about higher-latitude baroclinically influenced development. For example, Avila and Clark (1989) mention that organized clusters of convection over the open ocean and some midlatitude frontal lows may have been classified as tropical depressions during the late 1960s. This indicates that there is the possibility for data bias due to changes in the interpretation of the available information.

6. Summary and conclusions

A useful way to consider the Atlantic hurricane season it to separate tropical-only from baroclinically influenced storms (Hess et al. 1995). Here, we develop an objective classification procedure for grouping hurricanes based on initial depression and initial hurricane positions.

The grouping is done using a partially adaptive classification tree that provides a series of decision rules. For this problem, a useful classification tree is found with eight decision rules. We obtain a greater than 90% accuracy in a cross-validation exercise. The results are significant at $\alpha < 0.0001$. In order to group storms before 1950, a similar classification tree was considered by replacing depression data with initial tropical storm data. Applying these rules to storms over the period 1886–1949 and combining the grouping over the 1950–94 period provides useful climatological information.

The two groups of hurricanes differ most in their latitude of development. On average, tropical-only hurricanes form near 19°N latitude, compared with 29°N for baroclinically influenced storms. Seasonal variability is also different with tropical-only storms confined primarily to the months of August through October, compared with June through November for baroclinically influenced storms. The secondary maximum observed in the seasonal number of hurricanes near mid-October is partly a consequence of baroclinically influenced hurricane activity. There has been a marked decrease in the relative number of tropical-only hurricanes since 1960. This is partly explained by recent droughts in western Africa, resulting in fewer hurricanes originating from waves passing over or near Cape Verde.

This study can be extended by considering the above dichotomy of storm type to include tropical storms. Further, it might be possible to subdivide the baroclinically influenced group into the categories of baroclinically initiated and hybrid storms. Work in these directions is currently in progress. In closing, we restate the utility of considering hurricane activity in the Atlantic as the sum of tropical-only and baroclinically influenced storms to better understand seasonal and interannual variability of the Atlantic tropical cyclone climate.

Acknowledgments. We thank the anonymous reviewers for their many helpful comments on an earlier draft. Some support for this work came from the Risk Prediction Initiative, from the National Oceanic and Atmospheric Administration through the Cooperative Institute on Tropical Meteorology, and the National Science Foundation ATM 94-17528.

REFERENCES

Avila, L. A., and G. B. Clark, 1989: Atlantic tropical systems of 1988. *Mon. Wea. Rev.*, **117**, 2260–2265.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984: *Classification and Regression Trees.* Wadsworth, 358 pp.

Burrows, W. R., 1991: Objective guidance for 0–24-hour and 24–48-hour mesoscale forecasts of lake-effect snow using CART. *Wea. Forecasting,* **6**, 357–378.

——, M. Benjamin, S. Beauchamp, E. R. Lord, D. McCollor, and B. Thomsom, 1995: CART decision tree statistical analysis and prediction of summer season maximum surface ozone for the Vancouver, Montreal, and Atlantic regions of Canada. *J. Appl. Meteor.*, **34**, 1848–1862.

Casella, G., and R. L. Berger, 1990: *Statistical Inference.* Wadsworth, 650 pp.

Cry, G. W., and W. H. Haggard, 1962: North Atlantic tropical cyclone activity, 1901–1960. *Mon. Wea. Rev.*, **90**, 341–349.

Elsner, J. B., and C. P. Schmertmann, 1993: Improving extended-range seasonal predictions of intense Atlantic hurricane activity. *Wea. Forecasting,* **8**, 345–351.

——, T. B. Kimberlain, and C. P. Schmertmann, 1994: Poisson model (with maximum likelihood criterion) forecasts of Atlantic tropical storm activity for 1995. *Exp. Long-Lead Forecast Bull.*, **3**, 14–15.

Gray, W. M., 1994: LAD multiple linear regression forecasts of Atlantic tropical storm activity for 1995. *Exp. Long-Lead Forecast Bull.*, **3**, 12–13.

——, and C. W. Landsea, 1992: African rainfall as a precursor of hurricane-related destruction on the U.S. East Coast. *Bull. Amer. Meteor. Soc.,* **73,** 152–1364.

——, ——, P. W. Mielke Jr., and K. J. Berry, 1992: Predicting Atlantic seasonal hurricane activity 6–11 months in advance. *Wea. Forecasting,* **7,** 440–455.

——, ——, ——, and ——, 1993: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. Forecasting,* **8,** 73–86.

——, ——, ——, and ——, 1994: Predicting Atlantic basin seasonal tropical cyclone activity by 1 June. *Wea. Forecasting,* **9,** 103–115.

Hand, D. J., 1981: *Discrimination and Classification.* John Wiley & Sons, 218 pp.

Herbert, P. J., and N. L. Frank, 1973: Atlantic season summary, 1972. *Mon. Wea. Rev.,* **102,** 456–465.

Hess, J. C., and J. B. Elsner, 1994: Historical developments leading to current forecast models of annual Atlantic hurricane activity. *Bull. Amer. Meteor. Soc.,* **75,** 1611–1621.

——, ——, and N. E. LaSeur, 1995: Improving seasonal predictions for the Atlantic basin. *Wea. Forecasting,* **10,** 425–432.

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. Rep. NWS NHC 22, 21 pp.

Landsea, C. W., 1993: A climatology of intense (or major) Atlantic hurricanes. *Mon. Wea. Rev.,* **121,** 1703–1713.

Mardia, K. V., J. T. Kent, and J. M. Bibby, 1979: *Multivariate Analysis.* Harcourt Brace & Company, 521 pp.

Neumann, C. J., B. R. Jarvinen, A. C. Pike, and J. D. Elms, 1987: *Tropical Cyclones of the North Atlantic Ocean: 1871–1986.* National Climatic Data Center and National Hurricane Center, 186 pp.

Shih, Y.-S., 1993: Tree-structured classification. Ph.D. dissertation, University of Wisconsin-Madison.

Simpson, R. H., N. L. Frank, D. Shideler, and H. M. Johnson, 1969: Atlantic tropical disturbances, 1967. *Mon. Wea. Rev.,* **97,** 251–264.