Improving Tropical Cyclone Intensity Guidance in the Eastern North Pacific

KEVIN R. PETTY
National Center for Atmospheric Research, Boulder, Colorado

JAY S. HOBGOOD
Atmospheric Sciences Program, The Ohio State University, Columbus, Ohio

(Manuscript received 16 September 1998, in final form 15 November 1999)

ABSTRACT

The primary objective of this research is the development of a statistical model that will provide tropical cyclone (TC) intensity guidance for the eastern North Pacific (ENP) superior to that provided by climatology and persistence models. Toward this goal, the authors investigate the use of shear and thermal variables in statistical forecasts of TC intensity change. European Centre for Medium-Range Weather Forecasts (ECMWF) global analyses are used to develop an Eastern Pacific Intensity Change (EPIC) model, forecasting tropical cyclone intensity changes at 12-h intervals, out to 72 h. The dataset consists of ENP tropical cyclones during the years 1989–96 along with ECMWF analyses for those years.

The synoptic predictors examined in this study consist of shear and thermal variables, including zonal and meridional components of shear between 200 and 850 mb, the difference in temperature between 200 and 850 mb, and equivalent potential temperature at 700 mb. The time tendency of these variables is also explored, and multiple linear regression analysis is used to detect which variables best explain the variance in tropical cyclone intensity change in the ENP.

The EPIC model forecasts are compared to those from a climatology–persistence model developed from the 1989–96 dataset, and to the current operational statistical models, SHIFOR (a model based on climatology and persistence) and SHIPS (Statistical Hurricane Intensity Prediction Scheme—a model based on climatology, persistence, and synoptic variables), in the ENP basin for the 1997 and 1998 hurricane seasons. Results for these seasons reveal that EPIC may provide better intensity guidance than statistical models based on climatology and persistence, and confirm that the inclusion of synoptic predictors in a statistical intensity prediction scheme improves intensity change forecasts and that, overall, EPIC may better forecast intensity change in the ENP.

1. Introduction

Progress on improved guidance for the prediction of intensities of tropical cyclones has lagged behind improvements in the accuracy of track forecasts. As stated by Avila (1998), a hurricane specialist at the Tropical Prediction Center’s National Hurricane Center (NHC) in Miami, Florida, “The problem of forecasting tropical cyclone intensity change continues to be a real challenge for meteorologists despite recent advances in numerical weather prediction” and “very little guidance is available to forecast intensity changes.” Forecast intensity errors increase substantially in cases when a tropical cyclone undergoes rapid changes of intensity. Avila cited the case of Hurricane Linda (1997) in the eastern Pacific and noted that “the 72 hour intensity forecast valid for 1800 UTC 12 September was underestimated by 100 kts.” As with other basins, prediction of intensity change in the eastern Pacific is sometimes made more difficult by the absence of aircraft reconnaissance and the relative sparsity of conventional observations.

Intensity is defined here as the maximum 1-min sustained surface wind. An understanding of the atmospheric variables that contribute to intensity change is an essential component in developing improved intensity forecast models. In this regard, previous research resulted in the development of objective intensity prediction schemes for the Atlantic basin, the western Pacific, and Australia (Pike 1985; Elsberry et al. 1988; DeMaria and Kaplan 1994b; Leslie and Holland 1991; Chu 1994). Although these studies have attempted to further our understanding of the factors that affect intensity change, improvements in forecasts are still lacking.

Using satellite imagery, forecasters rely on the Dvorak technique to produce initial intensity estimates, and it can also be used to produce intensity forecasts. In this pattern recognition technique, tropical cyclone intensity change is deduced from successive estimates of intensity
fitted to climatological deepening and filling models (Dvorak 1975, 1984). However, the use of this technique in operational intensity forecasts has occasionally resulted in substantial errors (Sheets and McAdie 1989).

The Dvorak technique is also limited to producing intensity forecasts out to 24 h. SHIFOR, a statistical model based on climatology and persistence, provides guidance for extended forecasts out to 72 h (Pike 1987). The model uses latitude, longitude, translation speed, Julian date, maximum wind speed, and 12-h change in the maximum wind speed to forecast intensity change. A later statistical model (Statistical Hurricane Intensity Prediction Scheme or SHIPS), which includes some synoptic predictors in addition to climatology and persistence, also provides guidance to forecasters (DeMaria and Kaplan 1994b, 1999).

The objective of this research is the development of a statistical model that will provide improved tropical cyclone (TC) intensity guidance for the eastern North Pacific Ocean (ENP) basin over that provided by climatology and persistence models. Toward this objective, we investigate the use of shear and thermal variables in statistical forecasts of TC intensity change. This goal is achieved using a methodology analogous to that of DeMaria and Kaplan (1994b, 1999). They used multiple linear regression analysis to produce SHIPS for the Atlantic basin and later for the ENP.

2. Data

Best-track data from NHC are used in this study as the source of nonsynoptic variables. This database includes all observed ENP tropical cyclones from 1949 through 1998. Observations from the best-track data include tropical depression, tropical storm, and hurricane stages. The latitude and longitude at the center of each tropical cyclone are provided at 6-h intervals (0000, 0600, 1200, 1800 UTC), as well as the cyclones’ maximum wind speeds rounded to the nearest 5 kt. Since aircraft reconnaissance is rarely available and there is relatively little ship traffic in this area, maximum winds are usually estimated from satellites (Dvorak 1984).

The model is designed specifically for NHC’s region of responsibility in the ENP: the area east of 140°W. This research uses data for the 8-yr period 1989–96, but excludes all observations in the best track database that occur west of 140°W. Observations that take place when cyclones are over land are also removed from the best-track file. These restrictions ensure that a homogeneous set of cases is used in the development of the model. The guidelines and the usual decreasing sample size with time yield 1260 and 707 cases at 12 and 72 h, respectively. This is more than adequate for the development of the regression equations (Neter et al. 1989).

European Centre for Medium-Range Weather Forecasts (ECMWF) global analyses are used to compute all of the synoptic variables in this study. These data are provided on a 2.5° latitude–longitude grid produced twice daily (00 and 1200 UTC) at 14 (15 level starting January 1992) standard pressure levels. Synoptic variables used in this investigation are computed from the 850-, 700-, and 200-mb pressure levels. Between 1980 and 1988, the ECMWF model underwent major changes in the data assimilation and analysis processes. These changes were implemented in an effort to improve the model’s performance. Trenberth and Olson (1988) outlined these changes in their study that compared the analysis–forecast system at ECMWF to that at the National Meteorological Center (now the National Centers for Environmental Prediction). They indicated that the most substantial changes related to the ECMWF system included physical parameterization, increased resolution, and the introduction of diabatic effects. The changes administered during the 8-yr period have further enhanced the model’s ability to provide quality forecasts in the Tropics. Since our dataset begins in 1989, we take full advantage of these improvements to the ECMWF model. Molinari et al. (1992) suggested that the use of ECMWF analyses in data-poor regions is likely to be meaningful and that future studies should investigate the use of these analyses in predicting TC intensity change.

3. Analysis

a. Climatology and persistence variables

Climatology and persistence have been used to predict TC intensity changes (Jarvinen and Neumann 1979; Hobgood 1998), but research has also shown that the addition of synoptic predictors leads to reduced forecast errors of TC intensity change (Elsberry et al. 1988; Leslie and Holland 1991; DeMaria and Kaplan 1994b, 1999). The research presented here centers on synoptic predictors but also emphasizes the importance of climatology and persistence predictors. Table 1 presents the climatology and persistence variables used in this study. These variables parallel those in DeMaria and

<table>
<thead>
<tr>
<th>TABLE 1. Climatology and persistence variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMAX</td>
</tr>
<tr>
<td>PVMAX12</td>
</tr>
<tr>
<td>TCLAT</td>
</tr>
<tr>
<td>TCLONG</td>
</tr>
<tr>
<td>TS</td>
</tr>
<tr>
<td>UTS</td>
</tr>
<tr>
<td>VTS</td>
</tr>
<tr>
<td>ABSIDAY</td>
</tr>
</tbody>
</table>
Kaplan (1994b). Methods used to compute these variables are discussed in the following paragraphs.

For each tropical cyclone, the best-track data provide an estimate of the maximum wind speed every 6 h. A TC intensity change variable is calculated based on the best-track wind speed given 12 h prior to the start of the forecast period. The difference between the intensity at the start of the forecast period (VMAX) and the intensity 12 h earlier gives the change in intensity over the last 12 h (PVMAX12), a persistence predictor that indicates strengthening or weakening at the start of the forecasts. It will be shown later that this predictor performs well in short- to medium-range forecasts, but it becomes less reliable in long-term prediction of intensity change. VMAX is preserved as a possible predictor in the model. A small magnitude at the start of the forecast period implies that the system is in the early stages of development and the possibility of strengthening is greater.

Latitude (TCLAT) and longitude (TCLONG) from each 6-hourly time period are explored as possible predictors, because storm position has been found to be correlated with TC intensity (DeMaria and Kaplan 1994b; Hobgood 1998). The average motion of ENP TCs is west-northwest from warm to cooler water, and Miller (1958) was one of the first to propose that the upper bound intensity a TC can achieve is a function of the SST. Moreover, formation and intensification of a TC is dependent upon the temperature of the underlying ocean. When ENP cyclones move to higher latitudes (cooler water) the amount of energy transferred to the system is reduced and the storms weaken. Vertical shear of the horizontal wind is also a possible explanation for weakening systems traveling to higher latitudes. At high latitudes, strong westerly flow can result in increased vertical shear within the environment of a TC. Elsberry et al. (1987) noted that the combination of vertical shear and lower SSTs contributed to reducing the intensity of recurving TCs in the ENP.

An increase in the translation speed of a TC is correlated with vertical shear, weakening of the system, and with cooler SSTs. The zonal (UTS) and meridional (VTS) components of translation speed are investigated as possible TC intensity predictors. The translation speed and its components are found using the position of the storm at the beginning of the forecast and 12 h earlier.

The method used in this study to determine the absolute Julian day predictor (ABSJDAY) is comparable to that used in SHIPS (DeMaria and Kaplan 1994b, 1999). Julian date 237 (25 August) has been found to be the peak of the hurricane season in the ENP (Neumann 1993). ABSJDAY is the absolute value of the difference between the observed date and the peak season date (the effect of leap years is not considered in the calculation). ABSJDAY suggests that a TC closer to the peak season date may have a higher intensity bound. In other words, the TC may be more likely to intensify as compared to a TC observed farther away from the peak season date. Higher SSTs and/or favorable environmental conditions near the peak season date may be responsible.

### b. Synoptic variables

Understanding the environment of TCs is an important step in predicting the future intensity of these systems. The research presented here uses ECMWF data to derive the synoptic-scale environment in an attempt to improve intensity change guidance in the ENP. This is accomplished through examination of shear and thermal variables. Previous research (DeMaria and Kaplan 1994a,b, 1999; DeMaria et al. 1993; Dunnavan 1981; Elsberry et al. 1988; Molinari and Vollaro 1989) has shown the importance of these synoptic variables in predicting TC intensity change. An effort is made here to survey specific components of each group and identify which are more significantly related to intensity change.

The development of the Eastern Pacific Intensity Change model (EPIC) uses a perfect-prognosis (prefect-prog) approach in which observed fields are used to simulate forecast fields. In other words, the future position of the storm at each time interval is simulated from best-track positions, while ECMWF initial analyses are used to simulate forecast fields. Operationally, it would be necessary to obtain future TC positions from a track forecast model and ECMWF gridded forecasts to calculate the synoptic variables at each forecast interval. Best-track positions and ECMWF initial analyses are also used to derive the synoptic predictors during the evaluation of the EPIC model (section 5). The following paragraphs explain what specific synoptic variables are investigated and how each variable is computed (the synoptic variables investigated in this study are listed in Table 2).

Vertical shear variables are determined from the zonal and meridional components of the wind at 200 and 850 mb. If a data point is less than 139 km from the center of the TC, it is excluded from the calculations (a point this close to the TC's center is not always representative of the large-scale flow). In Frank's (1977) study of tropical cyclone structure, he found that the scale of circulation was expansive, with considerable mean radial flow being present beyond 14° of latitude radius and

<table>
<thead>
<tr>
<th>Table 2. Synoptic variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td>U2-8</td>
</tr>
<tr>
<td>DU2-8</td>
</tr>
<tr>
<td>V2-8</td>
</tr>
<tr>
<td>DV2-8</td>
</tr>
<tr>
<td>T2-8</td>
</tr>
<tr>
<td>DT2-8</td>
</tr>
<tr>
<td>TLE700</td>
</tr>
<tr>
<td>DTE700</td>
</tr>
<tr>
<td>DTL</td>
</tr>
</tbody>
</table>
tangential circulation being positively correlated with central pressure. Large-scale circulation is captured in this study by using data points out to 1400 km. McBride and Zehr (1981) found distinct differences in the zonal and meridional vertical shear fields of developing and nondeveloping tropical clusters. Most importantly, a region of zero shear was located over developing systems. The time tendencies of the zonal and meridional shear variables are examined to account for the continually evolving environmental flow. For example, during a 48-h period the zonal and meridional components of shear between 200 and 850 mb are computed at the beginning of the time period. Best-track data are used to simulate the position of the TC 48 h from the initial time, and ECMWF initial analyses at that location and time are then used to compute the zonal and meridional components of the shear. Finally, the time tendency variables are computed by determining the difference between the shear magnitudes at the end of the period and those at the beginning.

Frank (1977) found that maximum temperature anomalies occur at 250 mb in steady-state typhoons. The 200-mb level is used to investigate the upper-level warming associated with TCs in the study presented here, in order to be consistent with the levels used for the shear calculations. All ECMWF data points within a 700-km radius are used to determine the 200- and 850-mb environmental averaged temperature. As with vertical shear computations, points within 139 km of the TC’s center are not used in the calculations. The thermal structure of the environment is explored using the absolute difference of environmental average temperature between the 200- and 850-mb levels. Like shear, the evolving thermal structure is also examined by calculating the temporal changes in the absolute difference of the temperature between 200 and 850 mb.

Equivalent potential temperature ($\theta_e$) can be viewed as a measure of convective available potential energy in the lower levels of the atmosphere. Earlier research has revealed the importance of $\theta_e$ in predicting changes in tropical cyclone intensity (Dunnavan 1981; Sikora 1976). In this study, $\theta_e$ at 700 mb is considered as a possible predictor. Errors of several degrees kelvin can occur when $\theta_e$ is calculated by conventional methods in the Tropics (Holland 1997). Therefore, the equivalent potential temperature in this investigation is obtained through a method suggested by Bolton (1980) and Holland (1997). The equivalent potential temperature at 700 mb is derived from ECMWF data points within 700-km radius of the TC’s center, while points within 139 km are disregarded. Time tendencies of the equivalent potential temperature variable are also analyzed. The computation of the equivalent potential temperature time tendency is identical to the method by which the vertical shear time tendency variables are derived.

Variations in TC intensity have been observed when a storm approaches a large landmass. For this reason, distance to land (DTL) is analyzed to see what role it may play in the prediction of tropical cyclone intensity. The positions of the centers of TCs in the best-track data are designated by latitude–longitude coordinates. If the location of the TC is determined to be inland, the observation is excluded from the study. A TC that enters the Gulf of California is considered to be inland in this analysis, because once in the Gulf of California, TC intensity would become heavily influenced by Baja California and the Sierra Madre Occidental of Mexico. Therefore, removal of these observations is reasonable.

c. Regression

An EPIC model is developed through multiple linear regression analysis. As noted by DeMaria and Kaplan (1994b), the distribution of intensity change for a given time period is approximately normal. Thus, multiple linear regression is appropriate and is relied on to ascertain which variables best explain TC intensity change in the ENP. The dependent variable in this study is intensity change at 12, 24, . . . , 72 h, with the independent variables being made up of climatology, persistence, and synoptic predictors given in Tables 1 and 2. Six forecast equations (one for each forecast period) are constructed through regression analysis, with each equation attempting to utilize the smallest possible subset of independent variables to explain the largest amount of variance in the dependent variable.

One of the most crucial steps in the development of a regression model is the selection of the independent variables. Although an effective method of developing a regression model would be to examine all the possible models that can be developed from the initial pool of variables, this approach becomes computationally infeasible when the initial pool of variables becomes too large. Instead, a stepwise selection approach is adopted to identify the best subset of independent variables. In the first step, regression is run and the variable that is the most significant is entered into the model. A second variable is entered based on the highest partial correlation. The first variable is then examined for removal. After each step, the variables not in the model are examined for entry. If the most significant variable not in the model meets the entry criterion, it is entered into the model and the equation is recalculated. This process continues until no variables meet the entry and removal criteria (Norusis 1992).

4. Model descriptions

a. Climatology and persistence

The regression analysis is first run on the climatology and persistence variables listed in Table 1. Normalized regression coefficients for the climatology–persistence model, hereafter CLIMO model, are displayed in Table 3. The coefficients statistically significant at the 95% level are in bold face. Coefficients not in bold are the
normalized regression coefficients that the predictor would have if it were entered into the model at the next step in the stepwise regression analysis.

The PVMAX12 is significant at all forecast times excluding the 72-h forecast and the sign of the regression coefficients conform to what would be expected in the physical environment (Table 3). PVMAX12 is a persistence predictor that measures the intensity trend of a TC showing a greater ability to explain future intensity change during the short-term forecasts, but losing skill in long-term forecasts. Intensity trends in TCs are likely to persist in the short term, but over time, it is more probable that the intensity trend of a system will be disrupted.

Predictor VMAX is significant at all forecast intervals, with a negative correlation between future intensity change and the initial intensity of the storm. More intense TCs are closer to their maximum intensities and are more likely to weaken. The magnitude of the normalized regression coefficients is a minimum at the 12-h forecast and a maximum at 72 h.

The ABSJDAY predictor is negatively correlated with intensity change because farther away from the peak frequency date the probability becomes greater that a TC will weaken. ABSJDAY can be viewed as a climatological predictor of the synoptic-scale environment, with the early and late season large-scale environment being less favorable for the development of intense storms. Like VMAX, ABSJDAY is significant at all forecast periods and the minimum and maximum magnitude of the coefficients are found at 12 and 60 h, respectively.

The only other predictors that are significant at all forecasts periods are TCLAT and TCLONG. An increase in a TC’s latitude places it in an environment that is less conducive to intensification. Decreasing SSTs at higher latitudes further decrease the TC’s available energy while increased vertical shear acts to disrupt the storm’s circulation. In the research presented here, longitude in the Western Hemisphere is represented by a negative value (e.g., 100°W = -100). The predictor suggests that the farther west a system is located the higher the likelihood that it will weaken. Higher magnitudes of vertical shear and lower SST values beyond 120°W make this inference physically plausible.

The meridional component of the translation speed is significant at five forecast periods, 24–72 h. Negative normalized regression coefficients are in agreement with the theory that an increase in the translation speed would lead to higher magnitudes of vertical shear. Changes in intensity are also associated with the fact that northward moving systems encounter cooler SSTs at higher latitudes.

The coefficient, $r^2$, is a measure of how well the model is able to predict the dependent variable. In the case of the climatology model, $r^2$ is relatively large at all forecast times, being a minimum at 24 h and a maximum at 72 h. The $r^2$ increases at all subsequent forecast intervals except the 24-h forecast. This would suggest that the model’s ability to forecast intensity changes gets better at longer forecast periods, but the final analysis does not support this indication. A similar pattern is seen in the synoptic model produced in this investigation. These results are consistent with DeMaria and Kaplan’s (1994b, 1999) research, attributing this pattern to problems in the accuracy of the intensity estimates and the practice of rounding best track intensity estimates to the nearest 5 kt. Because the average intensity change is larger at longer forecast periods, the accuracy of intensity estimates poses less of a problem as compared to shorter intervals. This increase in $r^2$ over time may also be the result of the standard deviation increasing at a different rate than the standard error.

b. Synoptic model

Five synoptic variables are used as predictors in the EPIC model: the zonal and meridional components of shear, the time tendency of the 200–850-mb temperature difference, 700-mb equivalent potential temperature time tendency, and the distance to land (Table 4). Along with these synoptic predictors, four climatology–persistence predictors are used. These include the 12-h intensity change, initial intensity, absolute Julian date, and initial TC latitude, with normalized regression coefficients at a different rate than the standard error.

The change in PVMAX12 is significant at all the forecast periods, with normalized regression coefficients revealing that this predictor dominates all other predic-
tors at 12 h. VMAX is also significant at all forecast periods, with the magnitude of the coefficients increasing at each subsequent forecast period. Predictor ABSJDAY gradually becomes more important over time, reaching its largest magnitude at 60 and 72 h. Predictor TCLAT is significant at all forecast periods in the EPIC model, most significant at later periods.

The meridional component of the vertical shear from 200 to 850 mb (V2-8) is significant at every forecast period, the magnitude of the coefficients being smallest at 12 and 72 h. The zonal component of vertical shear is only significant at the 36-, 48-, and 60-h periods. The sign of the regression coefficients is physically reasonable, since a negative correlation exists between TC intensity change and shear (McBride and Zehr 1981).

Two thermal variables are included as predictors in the EPIC model, DT2-8 and DT E700. DT2-8 is significant at the last three forecasts periods with its largest coefficient at 72 h. The variable DT E700 is representative of the time tendency of \( \theta_e \) at 700 mb, being significant at the 95% level from 24 to 60 h. Inclusion of DT E700 in the final regression may be particularly important in cases where drier air entrains into the circulation around the tropical cyclone.

Predictor DTL is significant at the last three forecast periods. Predictor TCLONG is strongly correlated with DTL due to the orientation of the coast of Mexico. Therefore, TCLONG was removed from the regression analysis during the development of the EPIC model.

The values of \( r^2 \) associated with the EPIC model are very similar to CLIMO. Beyond 24 h, \( r^2 \) continues to increase reaching a maximum at 72 h, with the 24-h forecast period displaying the smallest value of \( r^2 \). The addition of synoptic predictors only results in minor increases in \( r^2 \) at each interval. It is probable that the lack of a substantial increase in \( r^2 \) is the result of utilizing best-track data in the developmental process. Had we used operational data to derive the climatology and persistence predictors, the value of \( r^2 \) would have been lower in Table 3. There is a marked difference in the quality of best-track data and operational data, with best-track data being superior. The best-track dataset has the advantage of assimilating information that is not known operationally, with initial conditions of climatology and persistence predictors being more realistically represented. Consequently, the climatology and persistence predictors exhibit higher regression coefficients and \( r^2 \) than would be observed if operational data had been applied during model development. Although, the addition of synoptic predictors does not produce a significant increase in \( r^2 \), the evaluation of the models clearly shows that EPIC is capable of producing smaller forecast errors than CLIMO.

5. Evaluation

The performance of the EPIC model is discussed in this section. The model was run on independent datasets from two ENP seasons, 1997 and 1998. The 1997 season produced 17 named TCs, just above the long-term average, with Guillermo and Linda being two of the most intense TCs ever observed in the ENP. These two seasons should provide an excellent test for the EPIC model. Performance is evaluated in terms of mean and mean absolute wind speed forecast errors (kt). Mean forecast errors (forecast minus observed) are used to investigate possible biases in EPIC, CLIMO, SHIFOR, and SHIPS forecasts, while mean absolute forecast errors measure the actual skill of the forecasting schemes. Although past research has used SHIFOR as a baseline for skill, direct comparisons are used herein. The intensity change forecasts from the statistical models are compared to those derived from best-track wind speed data.

The statistical models produced in this research were evaluated using operational data files from the NHC. These files provide the current storm position, storm position 12 h ago, current maximum wind speed, the maximum wind speed 12 h ago, and intensity forecasts from SHIPS and SHIFOR. As discussed previously, the data contained in these files consist of the operational data that forecasters have available during the time of the storm and are inferior to best-track data, used to develop the EPIC and CLIMO models. The climatology and persistence predictors derived from the operational files are utilized for input in EPIC and CLIMO. Because an archive of ECMWF forecast fields was not available,

### Table 4. Normalized regression coefficients for predictors used in EPIC. Coefficients statistically significant at the 95% level are in bold; \( N \) is the sample size.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>12 h</th>
<th>24 h</th>
<th>36 h</th>
<th>48 h</th>
<th>60 h</th>
<th>72 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVMAX12</td>
<td>0.58</td>
<td>0.42</td>
<td>0.27</td>
<td>0.17</td>
<td>0.09</td>
<td>0.03</td>
</tr>
<tr>
<td>VMAX</td>
<td>-0.26</td>
<td>-0.34</td>
<td>-0.38</td>
<td>-0.41</td>
<td>-0.43</td>
<td>-0.45</td>
</tr>
<tr>
<td>TCLAT</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.30</td>
<td>-0.34</td>
<td>-0.37</td>
<td>-0.38</td>
</tr>
<tr>
<td>ABSJDAY</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-0.09</td>
<td>-0.11</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td>V2-8</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.12</td>
<td>-0.09</td>
</tr>
<tr>
<td>DT2-8</td>
<td>0.02</td>
<td>0.03</td>
<td>0.07</td>
<td>0.09</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td>U2-8</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.06</td>
</tr>
<tr>
<td>DT E700</td>
<td>-0.02</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.04</td>
</tr>
<tr>
<td>DTL</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.12</td>
</tr>
<tr>
<td>( r^2 ) (%)</td>
<td>52.5</td>
<td>49.5</td>
<td>51.5</td>
<td>54.7</td>
<td>57.0</td>
<td>58.1</td>
</tr>
<tr>
<td>( N )</td>
<td>1260</td>
<td>1139</td>
<td>1023</td>
<td>908</td>
<td>801</td>
<td>707</td>
</tr>
</tbody>
</table>
EPIC's synoptic predictors are determined from ECMWF initial analyses and best-track positions. Operational use of EPIC would incorporate ECMWF forecasted fields and track model forecasted positions, so we would expect that the EPIC forecasts presented in the following sections would experience some degradation if forecast fields and positions were applied. It also should be noted that, unlike EPIC, the synoptic variables in SHIPS are computed from forecast analyses and storm positions. Although an effort was made to compile a homogenous sample, the authors are aware that the comparisons to SHIPS and SHIFOR are not entirely equitable since EPIC uses some information not available operationally. Nevertheless, the comparisons are presented in an attempt to provide some understanding of EPIC's performance during the two seasons.

A brief overview of each hurricane season is given, and model forecast errors for each year are computed and analyzed. Because NHC does not issue a 60-h forecast, the models produced here were not compared to SHIFOR and SHIPS forecasts during that period. Performance of the models was further analyzed using individual storms. This was conducted to investigate how the models handle intense TCs.

a. Seasonal forecasts errors

The 1997 hurricane season produced 17 tropical storms and nine hurricanes. Although this was a near-normal year, it included one of the strongest hurricanes on record, Linda, whose maximum wind speed reached an astounding 160 kt (Lawrence 1998). Also during this year there were four landfalling TCs. Pauline, one of the storms that made landfall, was reportedly responsible for more than 230 deaths near Acapulco (Lawrence 1998). Guillermo was the second strongest TC during the season and lasted for over 15 days.

Figure 1a displays the mean forecasts errors for the 1997 ENP hurricane season, showing that EPIC maintains a very small bias during all forecast periods. In addition, EPIC has the smallest bias of all the statistical models displayed, whereas SHIPS displays a consistently small negative bias during all forecast periods. Beyond the 36-h forecast period, CLIMO and SHIFOR have considerably larger biases than EPIC and SHIPS indicating that the models tend to overforecast and underforecast intensity change, respectively.

The mean absolute forecasts errors for 1997 are given in Fig. 1b, with EPIC showing the most skill. The skill of the EPIC, CLIMO, SHIPS, and SHIFOR model forecasts are similar at 12 and 24 h, but by 36 h, the EPIC model begins to demonstrate more skill than the others. By 72 h, EPIC's forecast errors are over 4 kt smaller than the models based on climatology and persistence. SHIPS outperforms CLIMO and SHIFOR at the last two forecast times, but continues to produce larger forecast errors than EPIC. The EPIC forecasts for 1997 demonstrate the model's ability to provide valuable guidance during a very difficult forecasting season. These results suggest that the EPIC model may be capable of providing operational forecasters with improved guidance at extended forecast times. Paired t tests reveal that the differences between the EPIC and CLIMO models were significant at 12, 36 and 48 h at the 99% level and significant at the 95% level during the 24-h forecast period. As previously discussed, the EPIC comparisons to SHIPS and SHIFOR are not entirely fair. Therefore, significance tests between EPIC and these models were not computed.

The following year 13 tropical storms occurred and 9 of these resulted in hurricanes, being below the long-term average of 16 tropical storms and 10 hurricanes. Two tropical cyclones made landfall resulting in eight deaths. Isis lasted three days after forming south of Baja California and made landfall as a weak category one hurricane, and Javier made landfall as a tropical depression near Cabo Corrientes, Mexico (Avila and Giuney 1999). The strongest hurricane during the season was Howard, reaching wind speeds in excess of 130 kt,
forming in late August and maintaining a commonly observed west-northwest track before weakening on 30 August.

Figure 2a contains the mean forecast errors for the 1998 ENP hurricane season. Like 1997, EPIC features similar forecast trends during the 1998 season. The model displays the smallest bias at all the forecast periods except the 72-h forecast where it has a large positive bias. SHIPS has a positive bias at the first two forecast periods, negative at later periods. SHIFOR has a negative bias, and CLIMO exhibits a positive bias at all forecast times.

The absolute errors of the forecasts for the 1998 season are shown in Fig. 2b. During the first 36 h, all forecasts have relatively the same forecast errors. At 48 h, CLIMO demonstrates the most skill, with EPIC and SHIPS displaying the most skill at 72 h in 1998. Again, EPIC and SHIPS begin to show more skill than climatology and persistence at later forecast periods. There is a notable difference between SHIFOR and EPIC at the 72-h period, with less division between CLIMO and EPIC being observed during the same period. Although EPIC shows some skill at the last forecast period, the skill displayed is not as substantial in the 1998 season as in 1997. The differences between the EPIC and CLIMO models are significant at the 95% level at all forecast periods excluding the 72-h forecast.

In summary, the addition of synoptic predictors adds skill to intensity forecasts over that provided by operationally determined climatology and persistence alone. The 1997 and 1998 evaluation of the forecast models (Figs. 1 and 2) shows that at almost all periods EPIC exhibits smaller biases than climatology and persistence and at most periods equal or smaller absolute errors. The addition of thermal predictors to the synoptic variables may account for the improvement over SHIPS observed mainly in 1997. In the following section, case studies investigate how EPIC, CLIMO, SHIFOR, and SHIPS handle intense storms. It is possible that the inclusion of synoptic variables into statistical forecasts results in higher skill during rapid intensification or weakening situations.

b. Case studies

Linda was the strongest storm, not just during the 1997 season, but on record in the ENP, reaching a maximum intensity on 12 September, with winds speeds estimated at 160 kt (82 m s\(^{-1}\)) (Lawrence 1998). Linda experienced a 120-kt change in wind speed during a 48-h period. Another powerful storm, which occurred in 1998, was Hurricane Howard, with winds of 130 kt (67 m s\(^{-1}\)) on 23 August, making Howard the strongest storm during the season. Together, these two storms are studied to investigate how robust the forecast schemes are during the intensification, transitional, and weakening stages of intense storms in the ENP.

Mean forecast errors for Hurricane Linda reveal that all the forecasts schemes have a negative bias (Fig. 3a), not unexpected since Linda was the most intense TC on record in the ENP. Because of Linda’s extremely high wind speeds, the forecast schemes consistently underestimate intensification. The CLIMO model has the smallest bias through all forecasts periods followed by EPIC, with SHIFOR experiencing the largest bias at all forecast intervals. Figure 3b shows the mean absolute forecast errors for Linda, with the difference in absolute errors being negligible at the 12- and 24-h forecasts. By 36 h, the statistical models begin to diverge with EPIC showing moderate performance and SHIPS and SHIFOR displaying the most skill at 72 h.

The mean forecast errors for Howard are given in Fig. 4a. Bias trends for Howard closely resemble those for the entire 1998 season. During the first four forecast periods, EPIC has a bias close to zero. Large negative biases are exhibited by SHIFOR and SHIPS at the final two forecasts times while CLIMO displays a positive bias. The mean absolute forecasts errors shown in Fig. 4b indicate that EPIC has better skill than the other forecast schemes from 36 to 72 h. By 72 h, EPIC’s absolute errors are smaller than SHIPS, CLIMO, and SHIFOR by 4.6, 5.0, and 8.6 kt, respectively.
Figures 3 and 4 demonstrate EPIC’s ability to perform skillfully during the lifetime of two very intense TCs. EPIC, as well as the other models, showed a negative forecast bias for Hurricane Linda, and did not display the smallest absolute errors at any period for Linda, but did have more skill than CLIMO at 48 and 72 h and SHIFOR at 36 and 48 h. The forecasts produced by EPIC for Hurricane Howard clearly show more skill, with the mean absolute errors being smaller than all other models beyond 24 h. Linda and Howard test EPIC’s ability to perform in adverse situations. Although the results for these TCs are not statistically significant, there is still an indication the EPIC has more skill than climatology and persistence when forecasting intense tropical cyclones.

A further analysis of the robustness of the forecasts schemes with respect to Linda and Howard is provided in Figs. 5, 6, and 7. In these figures, actual forecasts from specified times during the existence of Linda and Howard are given, the forecast times being representative of stages during which the TCs were intensifying, in a state of transition, or weakening.

Figure 5a shows the forecast for Linda at 1200 UTC 10 September. Linda was beginning to go through a rapid intensification, reaching its maximum intensity 36 h from the initial time of this forecast. EPIC, CLIMO, and SHIPS have similar intensity change forecasts during the first two periods, outperforming SHIFOR. At 36 and 48 h, SHIPS displays the best performance followed by EPIC and CLIMO, whereas EPIC performs slightly better than CLIMO and SHIFOR at 72 h. A different pattern is seen during the Howard intensification period (Fig. 5b). Here CLIMO’s performance exceeds all other forecast schemes from 24 to 72 h followed by EPIC, and SHIPS and SHIFOR forecast significantly less intensification than EPIC and CLIMO. Again, EPIC displays moderate skill during the intensification of Howard.

The transitional phase of Linda and Howard can be seen in Figs. 6a and 6b, respectively. Figure 6a shows that Linda reached its maximum intensity and was beginning to weaken. The SHIFOR forecasts are close to the actual change at 12 h, but EPIC, SHIPS, and CLIMO continue to forecast an intensification of Linda at this time. By 24 h, SHIFOR again supplies the best guidance, but EPIC and CLIMO begin to pick up the weak-
FIG. 5. (a) Actual forecast of intensity change for Hurricane Linda (1997) during intensification stage from EPIC (line with filled circular markers), CLIMO (line with open circular markers), SHIPS (line with square markers), and SHIFOR (line with triangle markers). Observed intensity change denoted by dashed line. (b) Forecast for Howard (1998) during intensification stage.

FIG. 6. As in Fig. 5 except for Linda (1997) and Howard (1998) during transitional stage.

ning by forecasting a slightly smaller magnitude of intensification. At 36- and 48-h periods, EPIC provides improved guidance over that for SHIFOR, SHIPS, and CLIMO. SHIFOR is somewhat closer to the actual change of Linda at 72 h, but is forecasting too much weakening of the storm. During the transition stage of Howard, the statistical models have a difficult time forecasting the mild weakening (Fig. 6b). SHIPS and SHIFOR do well at 36 h, but continue to forecast substantial weakening at 48 and 72 h, whereas EPIC and CLIMO forecast a slight intensification through 36 h before diagnosing weakening. At 72 h, the actual change in intensity is very close to the change forecasted by EPIC.

Finally, forecasts are given for the weakening stages of each storm (Figs. 7a and 7b). EPIC and SHIFOR do an excellent job of forecasting the intensity change in Linda at this time. SHIPS does well through 36 h, but at later times forecast too much weakening. CLIMO also exhibits less skill at the last two periods, but in contrast to SHIPS, CLIMO does not forecast enough weakening.

Figure 7b shows the forecasts for Howard’s weakening stage, with SHIPS showing the most skill beyond 24 h, EPIC doing a poor job of detecting the substantial weakening of Howard at longer forecast times, and CLIMO exhibiting the least skill overall for the forecasts made at 0000 UTC 26 August.

The actual forecasts display model performance at different times during a TC’s lifetime and in explosive types of situations. From these analyses, it is difficult to postulate about EPIC’s ability to provide improved guidance over that provided by CLIMO, SHIFOR, or SHIPS, but Figs. 5, 6, and 7 indicate that the statistical models presented here display an interesting and common characteristic. If a model does well during one portion of a TC lifetime, it tends to do poorly in the other. For example, when a statistical model performs well during the intensification stage of a TC it will typically do poorest during the weakening stage of the same storm. (This can be seen in Figs. 5 and 7.) SHIPS shows the most skill beyond 36 h when Linda is intensifying (Fig. 5a). During the 48- and 72-h periods corresponding to Linda’s weakening stage, SHIPS shows the least amount of skill compared to the other statistical models.
Fig. 7. As in Fig. 5 except for Linda (1997) and Howard (1998) during weakening stage.

(Fig. 7a). Figure 5b indicates that CLIMO has the least forecast errors when Howard goes through intensification on 21 August. On 26 August when Howard is weakening, CLIMO displays the largest absolute errors (Fig. 7b). This attribute has been seen in intensity change forecasts for other TCs (not shown). In addition, all statistical models investigated here show difficulties in forecasting the transition stage of TCs. This may be the result of utilizing climatology and persistence in the forecasting schemes. There are indications that EPIC may feature more stable forecasts during the combined intensification and weakening states of TCs, but more investigation is needed.

6. Summary and conclusions

A statistical model has been developed to provide improved tropical cyclone intensity change guidance for the eastern North Pacific basin at 12-h intervals out to 72 h over climatology and persistence. This study is based on the hypothesis that adding synoptic predictors into a statistical model would improve intensity change forecasts over that provided by climatology and persistence predictors, with multiple linear regression being relied on to determine which variables are the best predictors.

First, the Eastern Pacific Intensity Change (EPIC) model was developed using an 8-yr period of observations, 1989–96, ECMWF initial analyses were used to calculate synoptic variables. The statistical development of the synoptic model began with nine synoptic variables and eight climatology–persistence variables, and five synoptic variables were retained as predictors in the final model. Normalized regression coefficients for the synoptic model indicated that the meridional shear between 200 and 850 mb was the most important synoptic predictor, with the change in $\theta$, at 700 mb being the next most significant synoptic predictor, and the change in the temperature difference between 200 and 850 being important at later forecast times. Four climate variables were also included in the EPIC model. Overall, the initial wind speed of the TC was the most significant climatology/persistence predictor included in the model, with the 12-h change in the wind speed being more significant at early forecast periods.

Next, the EPIC model forecasts were compared against CLIMO (climatology and persistence model), SHIFOR (operational climatology and persistence), and SHIPS (operational synoptic model) to see if the addition of synoptic predictors increased the skill of intensity change forecasts in the ENP. For a 2-yr period, 1997 and 1998, the EPIC model developed in this research consistently maintains a very small bias through 48 h and display a positive bias at 72 h, while the model based on climatology and persistence displayed a positive bias during these seasons. In terms of mean absolute errors, the EPIC model provides better intensity guidance than CLIMO and SHIFOR, specifically at later forecast periods. EPIC noticeably outperforms CLIMO and SHIFOR at 36-, 48-, and 72-h periods during 1997 and at 72 h in 1998, indicating that the model may be able to improve guidance at extended times. EPIC also exhibits smaller errors than SHIPS in 1997 and has comparable forecast skill in 1998.

It is believed that the EPIC model may be more robust than climatology and persistence models. In order to investigate this hypothesis, forecasts for individual cases of strong deepening and decay were analyzed, as well as the transitional state of intense TCs. The TC cases used to analyze SHIFOR, SHIPS, EPIC, and CLIMO’s abilities to forecast intensity change during adverse situations consisted of Hurricane Linda (1997) and Howard (1998). These cases were chosen because of the intense nature of each TC. Inspection of the forecasts from the statistical models did not establish EPIC’s ability to produce more robust forecasts than the other models and further investigation is needed. What was evident is that statistical models investigated in this study have some difficulty producing good forecasts during the complete lifetime of a TC. Preliminary indications suggest that when a model performs well during one
stage of a TC (intensification or weakening), its skill decreases during the opposite stage.

In conclusion, it is apparent that the EPIC model is capable of providing improved intensity change guidance over that supplied by operationally determined climatology and persistence. Evidence also suggests that EPIC may yield better guidance in more difficult forecasting situations, with an improvement in forecast skill over SHIFOR and CLIMO being evident during very intense storms. This study presents evidence reaffirming the use of synoptic predictors in statistical forecasts of intensity change. Although on average EPIC does improve intensity change forecasts, the results are not as good as we had hoped, suggesting the need for additional work on methodology and the examination of additional predictive information. Continued investigations of statistical–dynamical procedures may not only improve overall forecast guidance, but also lead to more robust forecasts.

Acknowledgments. This research was supported by the Advanced Study Program at the National Center for Atmospheric Research and the Atmospheric Science Program at The Ohio State University. We thank Dr. Mark DeMaria for valuable comments and suggestions during the course of this investigation and Dr. Richard Katz for editorial support. The National Center for Atmospheric Research, funded by the National Science Foundation, maintains the ECMWF gridded analyses that were used in this study. We also would like to thank the reviewers for their comments.

REFERENCES


