

A New Method for Estimating Tropical Cyclone Wind Speed Probabilities

MARK DEMARIA AND JOHN A. KNAFF

NOAA/NESDIS, Fort Collins, Colorado

RICHARD KNABB

NOAA/NWS/Central Pacific Hurricane Center, Honolulu, Hawaii

CHRIS LAUER

NOAA/NCEP/National Hurricane Center, Miami, Florida

CHARLES R. SAMPSON

Naval Research Laboratory, Monterey, California

ROBERT T. DEMARIA

CIRA/Colorado State University, Fort Collins, Colorado

(Manuscript received 11 March 2009, in final form 5 June 2009)

ABSTRACT

The National Hurricane Center (NHC) Hurricane Probability Program (HPP) was implemented in 1983 to estimate the probability that the center of a tropical cyclone would pass within 60 n mi of a set of specified points out to 72 h. Other than periodic updates of the probability distributions, the HPP remained unchanged through 2005. Beginning in 2006, the HPP products were replaced by those from a new program that estimates probabilities of winds of at least 34, 50, and 64 kt, and incorporates uncertainties in the track, intensity, and wind structure forecasts. This paper describes the new probability model and a verification of the operational forecasts from the 2006–07 seasons.

The new probabilities extend to 120 h for all tropical cyclones in the Atlantic and eastern, central, and western North Pacific to 100°E. Because of the interdependence of the track, intensity, and structure forecasts, a Monte Carlo method is used to generate 1000 realizations by randomly sampling from the operational forecast center track and intensity forecast error distributions from the past 5 yr. The extents of the 34-, 50-, and 64-kt winds for the realizations are obtained from a simple wind radii model and its underlying error distributions.

Verification results show that the new probability model is relatively unbiased and skillful as measured by the Brier skill score, where the skill baseline is the deterministic forecast from the operational centers converted to a binary probabilistic forecast. The model probabilities are also well calibrated and have high confidence based on reliability diagrams.

1. Introduction

In recognition of the increasing coastal population vulnerable to tropical cyclones and the inherent uncertainty in the National Hurricane Center (NHC) track forecasts, the National Weather Service (NWS) imple-

mented a quantitative probability product beginning with the forecast of Hurricane Alicia in August of 1983 (Sheets 1985). Decision makers often relied on NHC's watches and warnings, but those do not directly correspond to the probability of an event occurring within the watch-warning areas. After consideration of a number of factors, the decision was made to supplement the deterministic NHC track, intensity, and wind structure forecasts and the watches and warnings with quantitative probabilities, rather than replace any of the existing products. It was

Corresponding author address: Mark DeMaria NOAA/NESDIS/StAR, 1375 Campus Delivery, CIRA/CSU, Fort Collins, CO 80523.
E-mail: mark.demaria@noaa.gov

felt that the familiar products were well suited to the general public, but the new quantitative probability products could be used by more sophisticated users such as government officials and other decision makers in cost-benefit analyses.

The original probability products were developed under the Hurricane Probability Program (HPP). The HPP only considered track error uncertainty, where bivariate normal distributions were fitted to the recent history of NHC track errors. A tropical cyclone “strike” was defined as occurring when the cyclone center moved within 50 n mi to the right or 75 n mi to the left of a given location, and probabilities were provided at selected coastal locations from 12 to 72 h. The products were issued for tropical cyclones above and below hurricane intensity, but only for the Atlantic basin. Except for periodic updating of the track error statistics, the HPP products changed very little from 1983 through 2005.

NHC and the Central Pacific Hurricane Center (CPHC), both part of the NWS, along with the Department of Defense’s Joint Typhoon Warning Center (JTWC), extended their track and intensity forecasts from 72 to 120 h beginning in 2003 after an experimental test period in 2001–02 (Rappaport et al. 2009). During the test period, the average 5-day NHC Atlantic track error was 350 n mi and the average intensity error was 25 kt. These relatively large errors raised concerns that too much attention would be paid to the exact forecast location and intensity in these long-range forecasts. Since extension of the existing HPP products from 72 to 120 h would only address track uncertainties and because some product users had expressed limitations in the utility of the HPP products, the program was reevaluated and an alternate approach was taken.

Beginning with the 2006 hurricane season, after an experimental period in 2005, a new probability program was implemented for the Atlantic and eastern, central, and western North Pacific basins that takes into account the uncertainties in track, intensity, and wind structure forecasts out to 5 days. The new program uses a Monte Carlo technique to estimate the probability of winds of at least 34, 50, and 64 kt at individual marine, coastal, and inland locations throughout each basin, from 6 to 120 h, based on official forecast error statistics from the previous 5 yr. This paper describes the new Monte Carlo probability (MCP) model.

The MCP model utilizes the error distributions from the official track and intensity forecasts. NHC has forecast responsibility for all tropical cyclones in the North Atlantic and the eastern North Pacific out to 140°W, CPHC has that responsibility in the central North Pacific from 140°W to the date line, and JTWC forecasts tropical cyclones (TCs) in the western North Pacific, the

North Indian Ocean, and many portions of the Southern Hemisphere. In the initial version of the MCP model, the Indian Ocean and Southern Hemisphere cyclones were not included. Three versions of the program were developed, for 1) the Atlantic, 2) the combined eastern and central North Pacific, and 3) the western North Pacific. The eastern and central Pacific cyclones were combined because the sample size is very small in the central Pacific, and the 140°W partition has no physical significance. In this paper, the term “official forecast” represents the operational tropical cyclone forecasts issued by NHC, CPHC, or JTWC, which include the storm center positions, maximum winds, and wind structure parameters described below.

The wind structure in the official forecasts is represented by the maximum radial extent of the 34-, 50-, and 64-kt winds in four quadrants relative to the storm center (NE, SE, SW, and NW). If the maximum radii were used to determine areas receiving these winds, the probabilities would be overestimated because the average radius in a quadrant can be much smaller than the maximum. Instead, the radius at the azimuthal center of each quadrant is needed, which can then be interpolated to give the radii at any azimuth. Using a large sample of surface wind analyses from the Hurricane Research Division’s H*Wind program (Powell et al. 1998) it was found that, on average, the radial extent of the winds of a given threshold in the center of each quadrant was about 85% of the maximum value in the quadrant. Another complication with estimating the storm structure is that the official wind radii are not available for all forecast periods out to 5 days. For this reason, the radii in the MCP model are obtained from the simple climatology and persistence (radii-CLIPER) model described by Knaff et al. (2007). For the wind radii discussed in the remainder of the paper, whether from the official forecasts, the best track (used for verification), or the radii-CLIPER model, the 0.85 correction factor was included.

The MCP model is described in section 2, a brief summary of operational products is provided in section 3, and the verification for the 2006 and 2007 seasons is presented in section 4. Conclusions, potential new applications, and improvements to the MCP model are described in section 5.

2. The Monte Carlo wind speed probability model

As described in the introduction, the original HPP product estimated the likelihood that the storm center would pass within a specified distance of a given location. These estimates were based on fitting bivariate normal distributions to the NHC track errors from the previous 10 yr (Sheets 1984). The fitting of a distribution

is a reasonable approach when only the track errors are considered. In principle, distributions could also be fitted to the intensity and structure errors. However, the track, intensity, and wind structure forecasts are not independent, especially when a cyclone is close to land. For example, suppose a particular cyclone is forecast to move northward just off the east coast of Florida, but without the forecasted center crossing land. The corresponding intensity forecast would assume that the cyclone remains over water, but there would be a reasonable chance that the actual track would cross land at some point. Utilizing basin-wide or ocean-only intensity error statistics would not be appropriate in this case. In principle, separate distributions could be developed for over-land and over-water forecasts. However, due to the complexity of the coastal geometry and the large number of combinations of tracks crossing land and water at various times during the forecast, a very large number of years would be needed to adequately sample the error distributions for all combinations of land and water tracks during the 120-h forecast period. A related problem occurs for the wind structure uncertainty (34-, 50-, and 64-kt wind radii). These wind radii are dependant on the intensity forecast, which also depends on the track forecast. In addition, for a given track and intensity forecast, proximity to land itself can affect the official wind radii forecast.

Due to the interdependence of the track, intensity, and structure forecasts and the interaction of the tropical cyclone with land, a Monte Carlo (MC) method was utilized for the new probability program. The MC method was originally developed to model the interaction of subatomic particles (Cashwell and Everett 1959) and is used in problems where geometric or other considerations make analytic approaches impractical. In MC methods a large number of plausible realizations of the physical process of interest are simulated directly. For example, MC methods have been used extensively to model visible light scattering in clouds with complicated geometries (e.g., Iwabuchi 2006). The paths of large numbers of photons are simulated, where the path after each scatter is determined by randomly sampling from phase functions appropriate for a given set of cloud particles.

For the wind probability model, a large number of plausible 5-day cyclone forecast tracks (realizations), each having its own 5-day intensity forecast, are determined by randomly sampling from the distributions of the official track and intensity errors, and then adding these to the official deterministic forecast of those parameters. An advantage of the MC method is that it is not necessary to assume an analytic form for the error distributions because they are sampled directly.

The official 34- and 50-kt wind radii forecasts only extend to 72 h, and the 64-kt radii forecasts only extend

to 36 h. For this reason the wind structure for each realization is determined from the radii-CLIPER model and its associated error distributions. Once these radii are determined, 0–120-h wind swaths for each wind speed threshold can be determined for each realization. The probabilities at any given location and for any time period out to 120 h can then be estimated by counting the number of realizations for which the point is inside the radii of the wind speed threshold of interest (34, 50, or 64 kt), relative to the total number of realizations. Further details of the method for constructing the realizations are provided below.

a. Track realizations

The first step in the MCP model is to generate the track realizations from the official track forecast, which includes a position estimate (latitude and longitude to the nearest 0.1°) at 0, 12, 24, 36, 48, 72, 96, and 120 h. The official positions are linearly interpolated to provide estimates at 60, 84, and 108 h, yielding a forecast position every 12 h. The tracks for the realizations are determined by randomly sampling from the previous 5-yr history of the official track errors and then adding these to the official forecast positions. These error distributions are calculated by comparing the official forecast positions to the “best track” positions, which are the poststorm best estimates of the cyclone track and intensity (Jarvinen et al. 1984). The track error is the great-circle distance from the forecast to the best-track position. The error vector is decomposed into the along-track (AT) and cross-track (CT) components, relative to the direction of the cyclone motion vector in the forecast track. The direction from the forecast track is used instead of that from the best track because the best track is not available in real time. The AT error is defined to be positive when the forecast position is ahead of the best-track position, and the CT error is positive when the forecast position is to the right of the best-track position.

Figure 1 shows the 48-h AT error distributions from the 2003–07 Atlantic sample, which was used in the 2008 operational MCP model. The mean of the distribution is near zero, indicating that the forecasts had relatively small biases. Similar distributions were calculated for the 12–120-h AT and CT errors. Although the $t = 0$ operational position estimates have some error, these are much smaller than those at later times. Therefore, the $t = 0$ h position errors are neglected and all realizations start at the same location for a given forecast.

During the initial development of the MCP model, it quickly became apparent that the serial correlation of the errors needs to be accounted for. That is, the 24-h AT and CT errors are not independent of the 12-h errors, and so on out to 120 h. As an example, Fig. 2a shows the

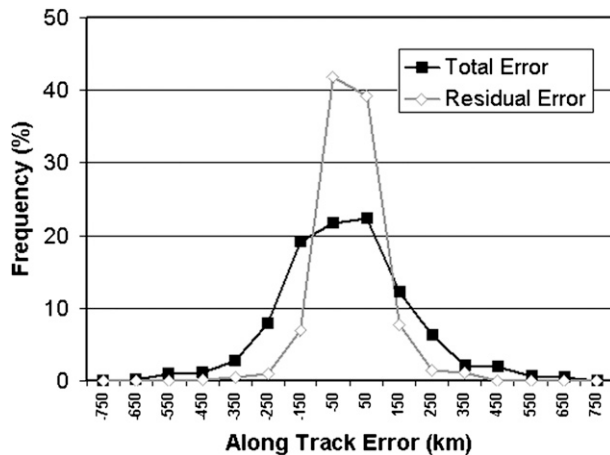


FIG. 1. The 48-h AT error distributions for the NHC Atlantic forecasts from 2003 to 2007. The distributions of the total errors and the residuals from the linear prediction are shown.

first 10 track realizations for a case from Hurricane Ike (2008) starting when the cyclone was east of Cuba. When the AT and CT errors are randomly sampled at each 12-h period, independent of the error from the previous 12-h period, the resulting tracks show unrealistic variability relative to the official track.

To account for the track error serial correlation, a simple autoregressive technique was utilized. For this purpose, the CT or AT error at each time period was assumed to be a linear function of the error at the previous time period. For example, letting AT_{t-12} and AT_t be the AT errors at time $t - 12$ h and t , respectively, and similarly for CT, then AT_t and CT_t are estimated from

$$AT_t = a_t AT_{t-12} + b_t \quad \text{and} \quad (1a)$$

$$CT_t = c_t CT_{t-12} + d_t, \quad (1b)$$

where a_t , b_t , c_t , and d_t are constants. The two constants in each equation at each time period are estimated from a least squares fit to the 5-yr sample of AT and CT errors used to generate the probability distributions. Since the $t = 0$ errors are assumed to be zero, $a_{12} = c_{12} = 0$ and the coefficients b_{12} and d_{12} are the sample mean AT and CT track error biases at 12 h. At later times, b_t and d_t are the y intercepts of the linear error prediction equations.

Table 1 shows the coefficients from the least squares fit of (1a) and (1b) to the 2003–07 NHC Atlantic track errors for all time periods from 12 to 120 h, and the variance explained by the linear model. All of the a_t and c_t values are positive, indicating a serial correlation of the AT and CT errors. The variance explained by the fit generally increases with forecast length and exceeds 90% at the longer times. This result indicates a high degree of serial correlation among the track errors.

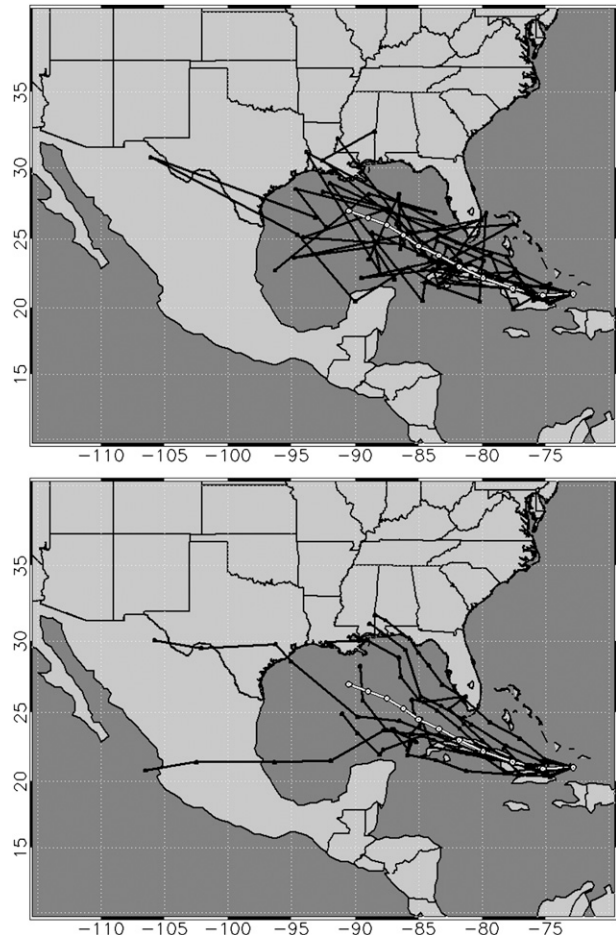


FIG. 2. The first 10 track realizations for a Hurricane Ike forecast starting at 1200 UTC 7 Sep 2008. The tracks (bottom) with and (top) without the correction for serial correlation of the track errors are shown. The white line is the NHC official track forecast.

Once the coefficients of the linear relationships in (1) are determined, the distributions of the AT and CT errors minus the linear predictions of these values (residuals) are calculated. Figure 1 shows the distributions of the residual 48-h AT errors for the Atlantic sample. The distribution of the residuals is much narrower than is that of the AT errors due to the high degree of serial correlation.

To calculate the track realizations, the CT and AT errors at 12 h are first predicted from (1a) and (1b), and then the residual distributions are randomly sampled and added to the predicted values of AT and CT. The sum of the predicted and random components of AT and CT are then added to the official forecast track. At 24 h, the AT and CT values are predicted from the 12-h values, and then a random component from the 24-h residual distribution is added, and so on out to 120 h. Figure 2b shows the first 10 track realizations for the Hurricane Ike example. Because the residual error distributions are much

TABLE 1. The slope (a and c), y intercept (b and d), and error variance explained (r^2) for the autoregression in (1a) and (1b) that account for the serial correlation of the AT and CT errors. The constants a , c , and r^2 are nondimensional and b and d have units of km.

Time (h)	Along track			Cross track		
	a	b	r^2	c	d	r^2
12	0.0	-6.5	0.00	0.0	4.9	0.00
24	1.3	3.4	0.74	1.3	3.3	0.78
36	1.3	-2.8	0.86	1.2	-0.2	0.84
48	1.2	-2.8	0.89	1.2	1.2	0.89
60	1.2	9.1	0.90	1.2	-7.1	0.90
72	1.2	16.8	0.94	1.2	21.7	0.94
84	1.1	13.2	0.88	1.0	-5.4	0.88
96	1.1	9.1	0.95	1.2	19.5	0.95
108	1.2	-11.6	0.91	1.0	-13.1	0.92
120	1.1	-12.2	0.97	1.2	19.1	0.96

TABLE 2. The coefficients in (2) for the estimation of the intensity error needed to account for the serial correlation and bias corrections that are a function of the forecasted intensity and the distance to land. The constants e and f are nondimensional, g has units of kt km^{-1} , and h has units of kt . The correlation coefficient of the prediction equation for intensity error (r) is also shown.

Time (h)	e	f	g	h	r^2
12	0.0	-0.048	0.0061	0.62	0.21
24	0.93	-0.031	0.0016	0.56	0.66
36	0.90	-0.022	0.0005	0.63	0.74
48	0.92	-0.034	0.0011	1.20	0.80
60	0.95	-0.009	0.0027	-0.62	0.85
72	0.92	-0.047	0.0022	1.23	0.86
84	0.91	0.005	0.0015	-1.16	0.88
96	0.88	-0.041	0.0007	1.29	0.88
108	0.93	-0.012	0.0008	-0.38	0.91
120	0.93	-0.060	0.0016	1.78	0.92

narrower than the total error distributions the track realizations are much less likely to jump back and forth between a position behind and in front of the forecast track, or between the right and left of the forecast track. Using this procedure, 1000 track realizations are generated for each forecast case. This number was chosen as a compromise between accuracy and run time. The convergence of the algorithm as a function of the number of realizations will be discussed in more detail later in this section.

b. Intensity realizations

For each of the 1000 track realizations, the maximum wind (intensity) at each 12-h interval is determined using a random sampling approach similar to that for the track, but also taking into account land interaction. The starting point is the 0–120-h official forecast of intensity that is linearly interpolated to include values at 60, 84, and 108 h. Because of the interdependence of the track and intensity forecasts, the track of each realization is checked for cases where the official forecast is over land but the realization position is over water, and vice versa. If the official forecast is over land but the realization is over water at a given time, the official intensity for that realization is replaced by a simple persistence forecast from the last time period where the official track position was over water. If the official forecast is over water but the realization position is over land, the official intensity for that realization is replaced by a forecast from the empirical inland wind decay model of Kaplan and DeMaria (1995) starting from the point where the realization track first moved over land.

Once the official intensity forecast for each realization is modified to account for land–water differences, a random component is added using a method similar to that for the track. The serial correlation of the intensity

error is accounted for by developing equations analogous to (1a) and (1b), but with additional terms. Letting VE_t represent the error in the forecast maximum wind (kt) at time t , V_t the forecast maximum wind at time t , and D_t the distance of the cyclone center from land (in km, where negative values indicate the cyclone is inland) at time t , then the intensity errors at each time period are estimated from

$$VE_t = e_t VE_{t-12} + f_t V_t + g_t D_t + h_t, \tag{2}$$

where e_t , f_t , g_t , and h_t are constants at each forecast time. The first term on the right side of (2) accounts for the autocorrelation of the intensity errors in a similar way as for track, and the last term on the right is the constant part of the bias correction. The second and third terms on the right correct the forecast bias as a function of intensity and distance inland, respectively. When a cyclone is sufficiently far from land, it should not make much difference how far away from land it is. For this reason, D_t is set equal to 500 km when it is greater than that distance.

The coefficients in (2) are determined from a least squares fit to the most recent 5-yr official forecast error sample and are shown in Table 2. The $t = 0$ intensity error is neglected, so $e_{12} = 0$. All other e values are positive, indicating a serial correlation of the intensity errors. Nearly all of the f values are negative, indicating that the error bias is inversely correlated with the forecast intensity. When high-intensity values are forecast, they tend to be too high, and vice versa when low-intensity values are forecast. All of the g values are small but positive, indicating that the intensity forecasts have a slight low bias for inland cyclones.

Equation (2) is used to calculate the expected intensity error at each forecast time, and then the probability distributions of the residuals from this prediction

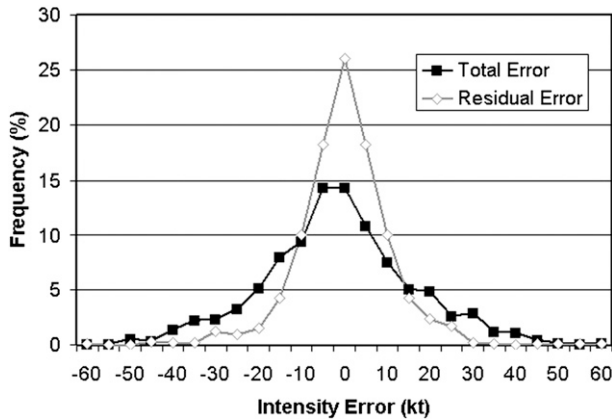


FIG. 3. The histograms of the 48-h intensity error distributions for the NHC Atlantic forecasts from 2003 to 2007. The original distributions and the residuals from the linear prediction of the errors are shown.

are determined. Figure 3 shows the 48-h intensity error distributions for the 2003–07 Atlantic sample before and after the removal of the linear error prediction. The intensity of each realization at 12 h is determined by first estimating the 12-h error from (2) and then randomly sampling from the residual intensity error distribution, and then both are added to the official intensity forecast. The 12-h error is then used as input to predict the 24-h intensity, and so on out to 120 h.

The above procedure works well except for cases when the official forecast track is over water, but the realization track moves inland. In this case, the intensity is adjusted by the inland wind decay model, but the addition of the residual can still sometimes make the intensity unrealistically large for an inland cyclone. To correct this problem, one final adjustment is made for realizations that are inland. The maximum intensity of any Atlantic tropical cyclone from 1967 to 2007 was calculated as a function of the distance inland, as shown in Fig. 4. An empirical curve was fit to the maximum intensity as a function of the distance to land, which is given by

$$V_i = 20 + 120e^{(0.0035D)}, \quad (3)$$

where V_i is the maximum wind (kt) and D is the distance to land (km), where D is negative for inland cyclones. If the intensity in any of the inland realizations exceeds this value at any forecast time, its intensity is set to this value, and the intensity errors are recalculated based on the adjusted intensity for use in the serial correlation for the following 12 h. Also, if the intensity drops below 15 kt at any time for an inland realization, the intensity of the cyclone is set to zero for all later times. This prevents cyclones from unrealistically reintensifying over

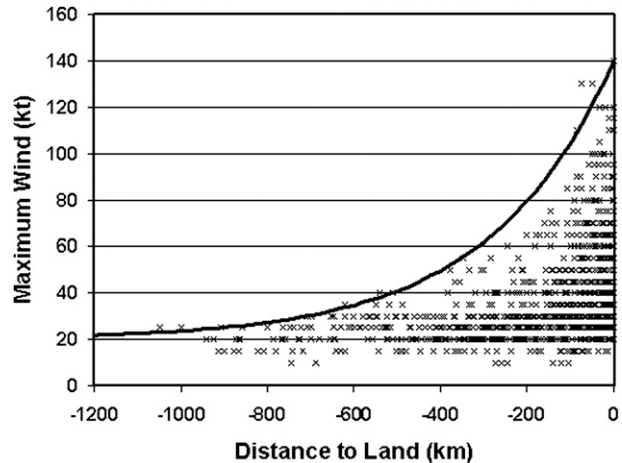


FIG. 4. A scatterplot of the maximum wind (kt) of any Atlantic tropical cyclone that made landfall in the continental United States vs the distance (km) from land (inland distances are negative) from the 1967–2007 NHC best track. An empirical function that represents the upper-bound intensity as a function of distance inland is shown by the thick black line.

land. This correction was implemented beginning with the 2008 hurricane season.

c. Wind structure realizations

After the procedures described in sections 2a and 2b are applied, each of the 1000 realizations have position and intensity estimates out to 120 h. To determine whether or not a specific point will experience the wind thresholds of interest (34, 50, and 64 kt), the radial extents of these wind speeds from the cyclone center as a function of azimuth are needed. As described previously, the official radii forecasts are not available out to 120 h. Even if the official forecast included all radii out to 120 h, some would still be missing for some of the realizations if their intensities were above a particular wind threshold but the official intensity was below it.

Because the wind structure needs to be estimated from the very limited information available for each realization (position and maximum wind out to 120 h and the $t = 0$ value of the wind radii), the radii-CLIPER model was used. This model uses a wind speed field that is the sum of an axisymmetric modified Rankine vortex profile and a wavenumber 1 asymmetry, given by

$$V(r, \theta) = (V_m - a) \left(\frac{r}{r_m} \right) + a \cos(\theta - \theta_0) \quad r < r_m \quad \text{and} \quad (4a)$$

$$V(r, \theta) = (V_m - a) \left(\frac{r}{r_m} \right)^x + a \cos(\theta - \theta_0) \quad r \geq r_m, \quad (4b)$$

where V is the wind speed, r is the radius from the cyclone center, θ is the azimuth measured counterclockwise relative to the direction 90° to the right of the cyclone direction of motion, V_m is the maximum wind speed, r_m the radius of maximum wind, a is an asymmetry factor, x is the cyclone size parameter, and θ_0 is a constant that allows the maximum wind speed to be located at an azimuth other than 90° to the right of the direction of motion. The complete cyclone wind speed field in the rotated cyclone-centered coordinate system can be determined once the five parameters V_m , r_m , a , x , and θ_0 are specified. For each realization, the coordinate system center and rotation angle are known from the track, and V_m is the maximum wind. The other four parameters are determined by climatological relationships with the cyclone maximum wind, latitude, and translational speed. The parameters r_m and x are functions of maximum wind and latitude, the asymmetry factor a depends on the translational speed and latitude, and the wind speed maximum rotation factor θ_0 depends on the latitude and translational speed. Generally speaking, the cyclone becomes larger with increasing latitude and intensity and more asymmetric with increasing translational speed and latitude. The azimuthal location of the maximum wind tends to rotate from the right to the front of the cyclone with increasing translational speed and latitude. Separate statistical relationships for the wind field parameters were developed for the Atlantic, the combined eastern-central North Pacific, and the western North Pacific.

The wind model above represents the climatological component. Persistence is included in two ways. First, the value of the size parameter x at $t = 0$ is adjusted to best fit the $t = 0$ values of the 34-, 50-, and 64-k winds from the official forecast. Second, the residuals from the radii predicted by the fitted wind model and the official $t = 0$ radii are calculated and added back to the model radii. This ensures that the wind radii exactly match those of the official forecast at $t = 0$. These residuals are also added during the forecast period, with a weight that exponentially decays with an e -folding time of 32 h. The e -folding time was determined from an analysis of the errors of the radii-CLIPER model.

Perturbations to the radii-CLIPER model are introduced through the size parameter x . In the development of the MCP model, the distribution of the difference in the x value between that from the climatological value and the value that provides the best fit to the observed radii was calculated. Beginning at 12 h, these differences are randomly added to the climatological estimate of x . The serial correlation of the x values is included in a similar way as for track and intensity, to account for the serial correlation in the deviations in cyclone size from the climatological value.

Once the wind model parameters are determined, the radii of 34-, 50-, and 64-kt winds are calculated in four quadrants relative to the cyclone center at 12-h intervals out to 120 h. The radii-CLIPER model provides inner and outer radii by solving (4a) and (4b) for r , respectively, for each wind threshold. The wind speed is zero at the cyclone center, so the inner radius is where the wind speed first reaches that of the specific threshold for increasing r . The outer radius is where the wind speed drops below the specified threshold. Because (4a) and (4b) are continuous functions of r with a single maximum for $r \geq 0$, the radii will always be physically consistent in each quadrant (the outer 50-kt wind radius can never exceed the 34-kt radius, etc.).

d. Probability calculation

Once the calculations described in sections 2a–2c are completed, the cyclone track, maximum wind, and radii of 34-, 50-, and 64-kt winds are available at 12-h intervals for each of the 1000 realizations. These parameters are linearly interpolated to a calculation time step, currently set to 2 h. The calculation time step must be small enough so that the cyclone does not move very far between time steps in relation to the size of the outer wind radii. The 2-h value was chosen to accommodate a cyclone with typical 64-kt wind radii (40 n mi) moving at a fairly fast speed of 20 kt. At each time step, the cyclone is repositioned and a determination is made of whether any given spatial grid point where the probabilities are being calculated is contained within the inner and outer radii of the wind speed of interest. For this determination, the wind radii in the four earth-relative quadrants are azimuthally interpolated to the angle between the cyclone center and the spatial grid point. The probabilities are determined by counting the number of realizations for which the grid point came within the radii of interest during a specified time period, and then dividing the result by 1000.

The model outputs two basic types of probabilities at 6-h intervals: cumulative and incremental. The cumulative values at each grid point indicate the probabilities that winds of at least 34, 50, and 64 kt will occur sometime during the entire period from $t = 0$ to a given forecast time (0–6, 0–12, . . . , 0–120 h). The incremental values are the probabilities that winds of at least 34, 50, and 64 kt will occur sometime within each 6-h time interval (0–6, 6–12, . . . , 114–120 h). The 6-h interval was chosen for consistency with other NWS products. For reference, fields of the 0-h probability values are also created. These values are 100% for points within the initial wind radii and 0% for points outside of these radii.

The probabilities are needed at 12-h time intervals for some applications. For the cumulative probabilities,

these are obtained simply by using values from every other 6-h cumulative period (e.g., skip 0–6 h, use 0–12 h). For the incremental probabilities, the 12-h values (0–12, 12–24, . . . , 108–120 h) are not equal to the sum of two consecutive 6-h values because realizations at the end of one time interval and the start of the next contribute to both intervals. However, the 12-h incremental values can be obtained from the 6-h incremental and cumulative values as follows. Letting $I_{t,t+n}$ represent the incremental probabilities from $t = t$ to $t = t + n$ and similarly for the cumulative probabilities C , then incremental probabilities over the 12-h interval can be determined from the 6-h values using

$$I_{t,t+12} = I_{t,t+6} + C_{t+6,t+12} - C_{t,t+6}. \quad (5)$$

e. Convergence tests

As described in Cashwell and Everett (1959), the accuracy of the MC method increases with increasing numbers of realizations (N). For some simplified cases the convergence rate can be estimated and is usually slower than $1/N$. This slow rate is one of the limitations of the MC method, and for some applications, methods have been developed to accelerate the rate of convergence (Iwabuchi 2006).

To gain some insight into the convergence rate and the error introduced by using 1000 realizations in the MCP model, the Hurricane Ike case starting at 1200 UTC 7 September 2008 is used as an example. The tracks of the first 10 realizations for this case are shown in Fig. 2, and the 0–120-h cumulative probability of 64-kt winds from one of NHC's official products (to be discussed in the next section) is shown later (see Fig. 6). This is a challenging case because many of the realizations interact with land. For the convergence tests, this case was run with 10, 25, 50, 100, 250, 500, . . . , 100 000 realizations. The 34-, 50-, and 64-kt probabilities were compared to runs with 500 000 realizations, which were considered the converged solutions.

As will be described in the next section, the operational version of the MCP model is run on a 0.5° latitude–longitude grid over a very large domain. For the convergence tests, the grid spacing was halved (0.25° latitude–longitude grid) on a domain centered on the forecast track (10° – 40° N, 60° – 100° W). For each value of N for each wind speed threshold, the average and maximum probability error on the domain were calculated relative to the runs with $N = 500\,000$. The average error is the mean absolute probability difference between that from a run with a given N to the converged solution. Because the probability is near zero over a large fraction of the domain, the average errors were calculated only for those

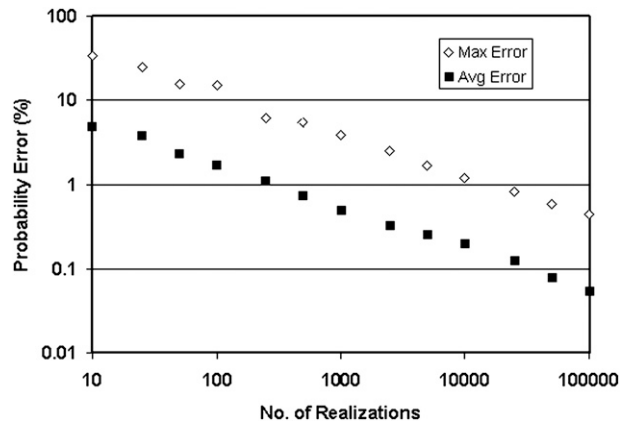


FIG. 5. The maximum and average error of the 64-kt wind probabilities for Hurricane Ike starting at 1200 UTC 7 Sep 2008 as a function of the number of realizations. Note that both axes have log scales.

points where the probability from the run being evaluated, or the converged run, was at least 1%. All points were used to find the maximum error.

Figure 5 shows the 64-kt probability error as a function of N on a log–log plot. For $N = 1000$, the average error is 0.49% and the maximum error is 3.8%. In the log–log diagram, the errors are nearly linear functions of N , which indicates that the error E can be accurately estimated from

$$E = \frac{C}{N^z}, \quad (6)$$

where C and z are constants. Taking the natural log of both sides of (6) gives

$$y = mx + b, \quad (7)$$

where $y = \ln(E)$, $x = \ln(N)$, $b = \ln(C)$, and $m = -z$. Fitting (7) to the data in Fig. 5 gives m and b , which then determine C and z in (6). The least squares fit of (7) explained more than 99% of the variance, and resulted in C and z values of 109.2% and 0.485 for the maximum error and 15.8% and 0.490 for the average error. The z values indicate that, to a good approximation, the maximum and average errors are both inversely proportional to the square root of N .

The above results indicate that the 64-kt wind probabilities with $N = 1000$ have converged to less than 0.5% in terms of the average error, and less than 4% in terms of the maximum error anywhere in the domain. Equation (6) can also be used to determine the number of realizations required for a given level of convergence. For example, to reduce the maximum error to 1% would require about 16 000 realizations. On the other hand, if a 1% average error was acceptable, then only about 280 realizations would be required.

The convergence of the 34- and 50-kt probabilities was also analyzed and the results were very similar to those for the 64-kt probabilities. For $N = 1000$, the average and maximum errors for the 50-kt probabilities were 0.54% and 3.4%, and 0.60% and 2.8% for the 34-kt probabilities. The convergence of the 50- and 34-kt probabilities can also be accurately estimated with equations in the form of (6). Similar to the 64-kt errors, the maximum and average errors were very close to being inversely proportional to the square root of N .

In summary, the MCP model converges at a fairly slow rate, but the average error due to the use of 1000 realizations is less than 1%.

3. Probability products

The original HPP information was disseminated by NHC as an Atlantic basin text product with probabilities for specified points, primarily at coastal locations (Sheets 1985). Shortly after NHC began making products available via their Web page in 1995, a graphical version of the HPP product was provided, where the probabilities were calculated on a latitude–longitude grid and then contoured. The output from the new MCP model has been provided via a suite of gridded, graphical, and text formats since replacing the HPP products in 2006, with coverage expanded to include not only the Atlantic but also the eastern, central, and western North Pacific basins.

The model is run on an evenly spaced (0.5°) latitude–longitude grid, in the domain 1° – 60° N, 100° E– 1° W, to generate 6-h cumulative and incremental probabilities resulting from all active cyclones in the domain forecasted by NHC, CPHC, and JTWC. The probabilities from all cyclones are combined at each grid point through the recursive use of the formula for two cyclones given by

$$P_t = P_1 + P_2 - P_1P_2, \quad (8)$$

where P_t is the probability of receiving winds of a given threshold from either cyclone and P_1 and P_2 are the probabilities for each individual cyclone.

The gridded cumulative and incremental probabilities with the 0.5° resolution are disseminated through the Advanced Weather Interactive Processing System (AWIPS) on a Northern Hemisphere domain, where the regions outside of the MCP model computational domain are filled with zeros. The probabilities are also interpolated to a 5-km grid for dissemination through the National Digital Forecast Database (NDFD) over the conterminous United States (CONUS) and Puerto Rico. They are also interpolated onto a 10-km NDFD grid over a much larger area that includes Hawaii and Guam.

The cumulative, combined probabilities on the 0.5° grid serve as input to graphical products issued by NHC and CPHC. An example of the 64-kt, 0–120-h cumulative probabilities for Hurricane Ike (2008) from the NHC Web page is shown in Fig. 6. Figure 7 shows an example of the 34-kt probabilities for a time with four tropical cyclones within the computational domain.

Due to the time required to generate these combined probabilities for all active cyclones, NHC and CPHC also issue a preliminary graphic for each cyclone that displays only the probabilities produced by that cyclone. These are generated using the Automated Tropical Cyclone Forecasting System (ATCF; Sampson and Schrader 2000). At the JTWC, single-storm cumulative probabilities are generated in a similar fashion for western North Pacific tropical cyclones and posted to a Web page along with the standard suite of tropical cyclone warning products, and are considered the final product.

One text product per active cyclone in the Atlantic and eastern and central North Pacific basins is also generated by NHC and CPHC by running the model at a selected set of coastal and inland points, and then listing the cumulative probabilities that exceed certain minimum values for each of the three wind speed thresholds. These text products list the cumulative probabilities through the official NHC and CPHC forecast periods (12, 24, 36, 48, 72, 96, and 120 h). The 120-h cumulative values in the text product correspond to the cumulative values displayed in the preliminary graphical products. Differences between the cumulative time periods are also shown in the text product, in order to convey the probabilities of onset of the various wind speeds, since event timing is important for decision making by many users such as emergency managers. A complete description of how the text and graphical products are interpreted and utilized by users will be discussed in a future paper.

4. Evaluation of the Monte Carlo probability model

In this section, the MCP model forecasts are evaluated through a comparison with observed occurrences of each wind threshold. The “ground truth” is determined by constructing binary grids indicating whether or not the wind threshold occurred during the period of interest. These grids are determined from the NHC best-track positions and wind radii. The MCP model is then evaluated using a number of standard metrics for probabilistic forecasts described later in this section. The skill of the MCP model is evaluated by comparison with the official deterministic forecast, which is converted to a probability binary grid using the official forecast positions and wind radii. As described in section 2, the official

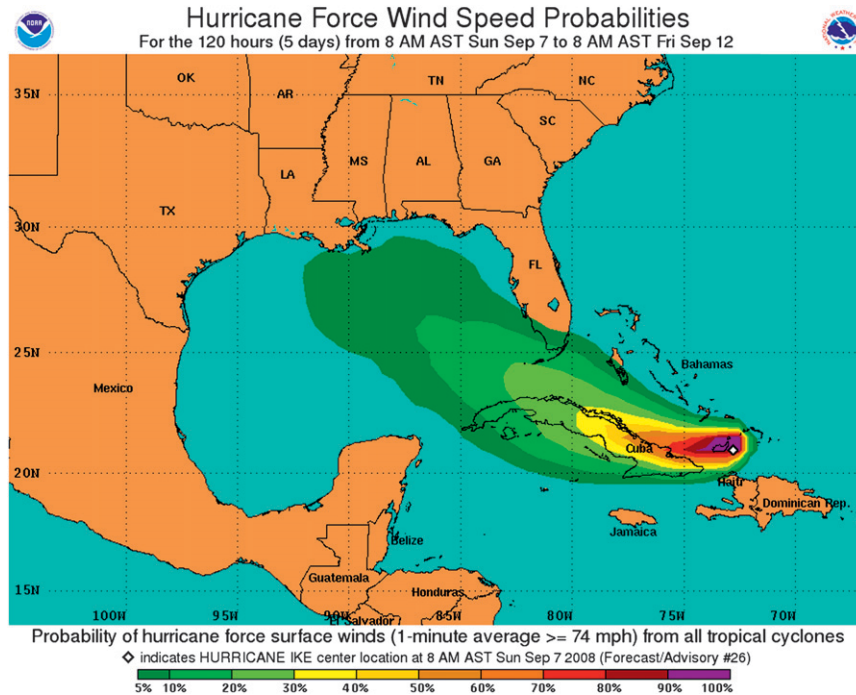


FIG. 6. The 0–120-h cumulative 64-kt wind probabilities for Hurricane Ike starting at 1200 UTC 7 Sep 2008 obtained from the NHC Web page.

radii are not available at all forecast times through 120 h, so the same radii-CLIPER model used in the MCP model is used to obtain the missing radii. A discussion of the results of the 2006–07 verification¹ then follows. The verification is performed for the incremental and cumulative probabilities with the 6-h time interval.

a. Verification–evaluation methodology

The input for the verification includes the official forecasts from NHC, CPHC, and JTWC; the radii-CLIPER model forecasts; the MCP model output on the 0.5° gridded multibasin domain; and the best-track files. All inputs except the best track were created in real time. The best-track data include estimates of the position, maximum wind, and radii of 34-, 50-, and 64-kt winds at 6-h intervals from a poststorm analysis of all available information. To compare the deterministic and the MCP model forecasts with the best-track observations, a set of common grids is constructed. This construction requires the following steps: 1) determining the number of 6-hourly times when there were storms

active in the verification area (active times), 2) determining how many storms had forecasts made on those dates (active storms), 3) constructing the deterministic forecasts by combining the official forecasts (position, intensity, and wind radii) and the radii-CLIPER forecasts of wind radii for times when the official forecast does not contain wind radii forecasts, and 4) matching deterministic forecasts and best-track verification times (i.e., verification will occur only for cases that have official forecasts). The first two steps and the last one listed above are simple accounting exercises. Combining the official and radii-CLIPER forecasts, however, requires more explanation.

The radii-CLIPER forecasts are created in real time using the official intensity and track forecasts as input. The merging is accomplished by first determining the last wind radii forecast time from the official forecasts. Once these times are known, the wind radii forecasts from radii-CLIPER are substituted for all of the remaining times when the official forecast exists. The resulting merged forecasts are then linearly interpolated to a 2-hourly temporal resolution. This is followed by a consistency check between intensity and the corresponding wind radii (where radii are set to zero when the interpolated intensity falls below the wind threshold), ultimately resulting in a 2-hourly deterministic forecast consisting of position, intensity, and wind radii.

¹ The verification contains all storms in the three basins beginning 1200 UTC 11 May 2006 and continuing through the end of 2007. This date corresponds to the operational implementation of the MCP model at the National Centers for Environmental Prediction (NCEP).

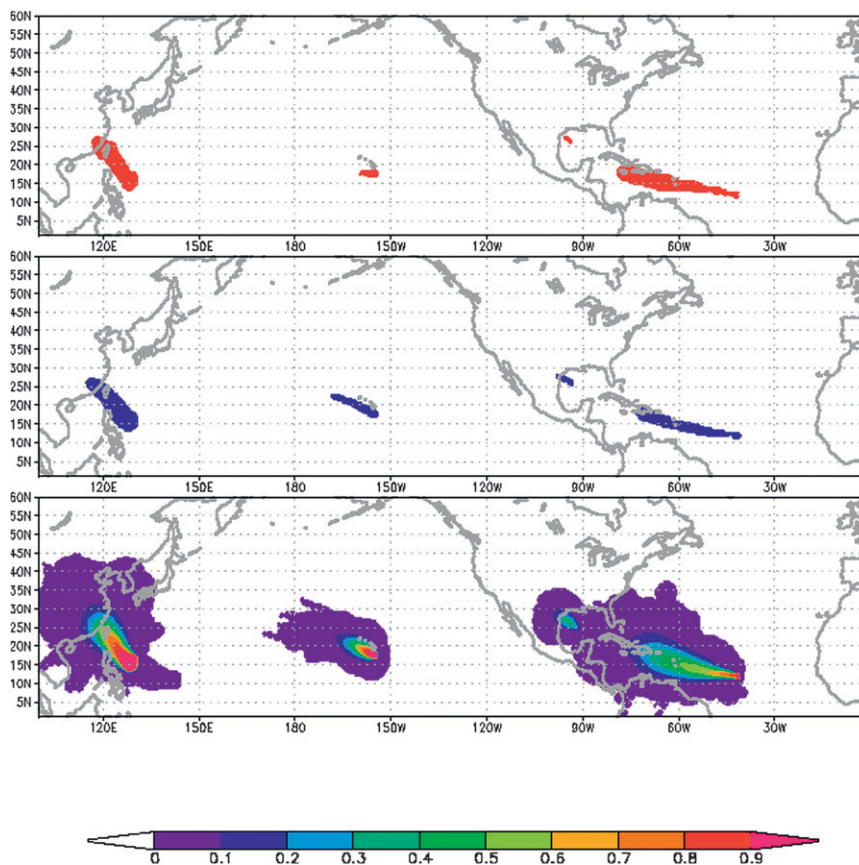


FIG. 7. Examples of the fields used for the verification starting 0000 UTC 15 Aug 2007 and extending through 5 days (120 h). (top) The observed occurrence of 34-kt winds from the best-track files (red), (middle) the forecast occurrence of 34-kt winds based on the deterministic forecast (blue), and (bottom) the 120-h cumulative 34-kt MCP model forecast (colors correspond to the color bar). During this time Tropical Storm Dean and Tropical Depression 5 (Erin) were active in the Atlantic, and Hurricane Flossie and Typhoon Seput were active in the central Pacific and western North Pacific, respectively.

Using the best-track positions, intensities, and wind radii, an identical interpolation and consistency-checking procedure is used to create 2-hourly best tracks of positions, intensities, and wind radii. Since the best tracks often contain periods of the storm histories during which no forecasts were made (e.g., extratropical stages), it is necessary to truncate “the verification best tracks” to include only the times during which there were corresponding official forecasts. The inclusion of only those times when an official forecast was available is consistent with the annual verification procedures of NHC’s track and intensity forecasts (Franklin 2008).

Using the 2-hourly interpolated best-track and deterministic forecast data, and corresponding active storm forecasts, for each active time, verification frequency grids and deterministic probability grids are constructed for each 6-h time interval. The verification and deterministic forecast binary grids contain a series of ones and

zeros corresponding to regions where the wind threshold is reached, or is not reached, respectively, over each 6-h interval. Examples of each of these gridded fields are shown in Fig. 7 for the multibasin domain at 0000 UTC 15 August 2007 when there were four active storms in the various basins. There were 359, 365, and 761 six-hour periods when there was at least one active storm in the Atlantic, eastern–central North Pacific, and western North Pacific basins, respectively. Typically, there are about 1000 grid points at each time period with nonzero probabilities for a given storm, so the verification sample sizes are very large.

A number of verification measures have been developed for probabilistic forecasts (Wilks 2006). For this study, the multiplicative biases, Brier skill scores, reliability diagrams, and various statistics calculated from 2×2 contingency tables found by assigning conditional probability thresholds are utilized. Specifically, the contingency table

metrics include the relative operating characteristics (and related skill score) (Mason and Graham 1999) and threat score. These statistics answer the following questions, respectively: How does the average forecast magnitude compare to the average observed magnitude? What is the relative skill of the probabilistic forecast over that of a reference forecast in terms of predicting whether or not an event occurred? How well do the predicted probabilities of an event correspond to their observed frequencies? What is the ability of the forecast to discriminate between events and nonevents? How well did the forecast “yes” events correspond to the observed yes events? Further details of each of these metrics are described by Wilks (2006).

b. Verification–evaluation results

The first aspect of the MCP model to be evaluated is its gross calibration by examining the multiplicative bias (Bias) defined by

$$\text{Bias} = \frac{\frac{1}{N} \sum_{i=1}^N F_i}{\frac{1}{N} \sum_{i=1}^N O_i}, \quad (9)$$

where F_i are forecasted probabilities and O_i are observed frequencies. These are summed over the entire domain for each forecast time. If the Bias $<$ ($>$) 1, then the average probability forecasts are too small (large) for that forecast period.

Figure 8 shows the multiplicative biases associated with the cumulative and incremental probability forecasts from the MCP model as a function of time for the Atlantic Basin (1° – 50° N, 110° – 1° W), the combined eastern and central Pacific basins (1° – 40° N, 180° – 75° W), the western North Pacific (1° – 50° N, 100° E– 180°), and the entire domain (1° – 60° N, 100° E– 1° W). Figure 8 shows that the cumulative probabilities have relatively small biases in the Atlantic and western North Pacific, with high biases most evident in the combined eastern–central Pacific. The biases are larger for the incremental probabilities (dashed lines). However, the entire domain shows very little multiplicative bias for 34- and 64-kt probabilities with a low bias evident for 50-kt probabilities. It is noteworthy that the cumulative probability biases are within $\pm 15\%$, save the eastern Pacific, where storms tend to be approximately 20%–30% smaller (Knaff et al. 2007, their Table 2). The deterministic forecasts exhibited similar biases (not shown) in all basin areas. This suggests that the intensity and radii biases in the official forecasts might be responsible for the biases in the MCP model. In addition, since the sample only includes 2 yr, it is possible

that the storms during that year were smaller than the long-term average in the eastern Pacific.

The above results indicate that the MCP model provides only slightly biased estimates of the wind probabilities, but are these forecasts more skillful than the deterministic forecasts produced by the various operational centers? To examine this question, Brier skill scores (BSSs) of the MCP model were computed using the deterministic forecast as the reference. This analysis provides the percent improvement or degradation a forecast has relative to the reference forecast. Results of this comparison (Fig. 9) depict a favorable interpretation of the MCP model. The MCP model forecasts are superior (BSS $>$ 0) to the deterministic forecasts beyond 12 h in all regions and for all wind thresholds. The relatively poor performance of the MCP model in the early periods is most likely caused by 1) the rapid relaxation of the wind radii to that of the climatology and persistence (i.e., e -folding time of 32 h) and 2) the linear interpolation between the $t = 0$ h observed wind radii and the first perturbations that are assigned at $t = 12$ h. Nonetheless, these statistics clearly show that the MCP model improves the mean square error (Brier score) associated with the frequencies of occurrence of winds at the 34-, 50-, and 64-kt wind thresholds. This error reduction has implications for the possibility of improving hurricane watches–warnings issued by NHC and CPHC, as well as tropical cyclone conditions of readiness (TC-CORs; NRL 2009) issued by JTWC, which is one of the applications to be discussed in a future paper.

To address the model calibration, reliability diagrams were prepared that display the forecast probabilities as a function of observed frequency as well as presenting information about the frequency of various probability forecasts. In this representation, perfect calibration would be represented as a 45° line. For brevity, the calibration results are shown only for the cumulative probabilities on the total MCP model domain for selected time periods. The results for the individual basins are similar.

Figure 10 shows the reliability diagrams for the 36-, 72-, and 120-h cumulative probabilities. These diagrams consist of two parts: the calibration function, which is the line plot, and the refinement distributions, which is the smaller bar plot. Good calibration is indicated by a near 1:1 correspondence in the calibration function and high confidence is indicated by the relatively frequent forecasts of the extremes (i.e., probability forecasts near 0 and 1.0). The good calibration and high confidence hold at all time periods shown in Fig. 10. Similar results were found for the individual basins, although there are some differences in the biases consistent with those shown in Fig. 8 (i.e., high biased in the eastern–central Pacific, slightly low biased in the Atlantic and western Pacific).

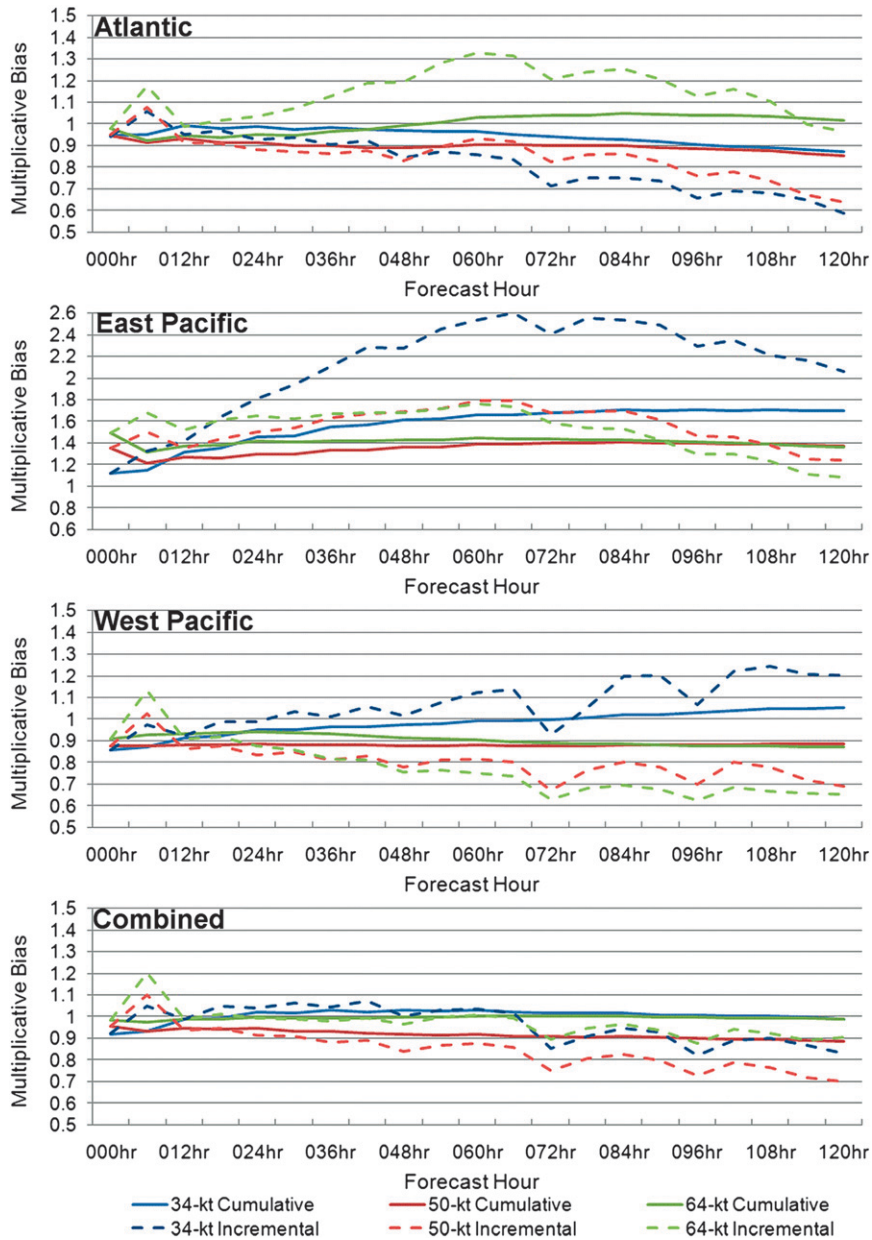


FIG. 8. The multiplicative biases associated with the 2006–07 MCP model verification in the North Atlantic (1° – 50° N, 110° – 1° W), eastern North Pacific (1° – 40° N, 180° – 75° W), western North Pacific (1° – 50° N, 100° E– 180°), and the multibasin domain (1° – 60° N, 100° E– 1° W) are shown in the panels starting from the top, respectively. Biases for the cumulative probabilities are given by solid lines and for the incremental probabilities they are given by dashed lines. Line colors blue, red, and green correspond to biases associated with 34-, 50-, and 64-kt wind probabilities. The scale is identical for all basins except the eastern North Pacific, where the scale is doubled.

The final statistics examined in this section are those constructed from the 2×2 contingency table, including the relative operating characteristics (ROC) and related skill scores. For this purpose, a probability threshold is specified, and then each event (at each grid point for each

forecast case for each 6-h interval) is classified into one of four categories in the contingency table: *a* if the event was predicted (the probably exceeded the specified threshold) and was observed, *b* if the event was predicted but did not occur, *c* if the event was not predicted but did occur, and

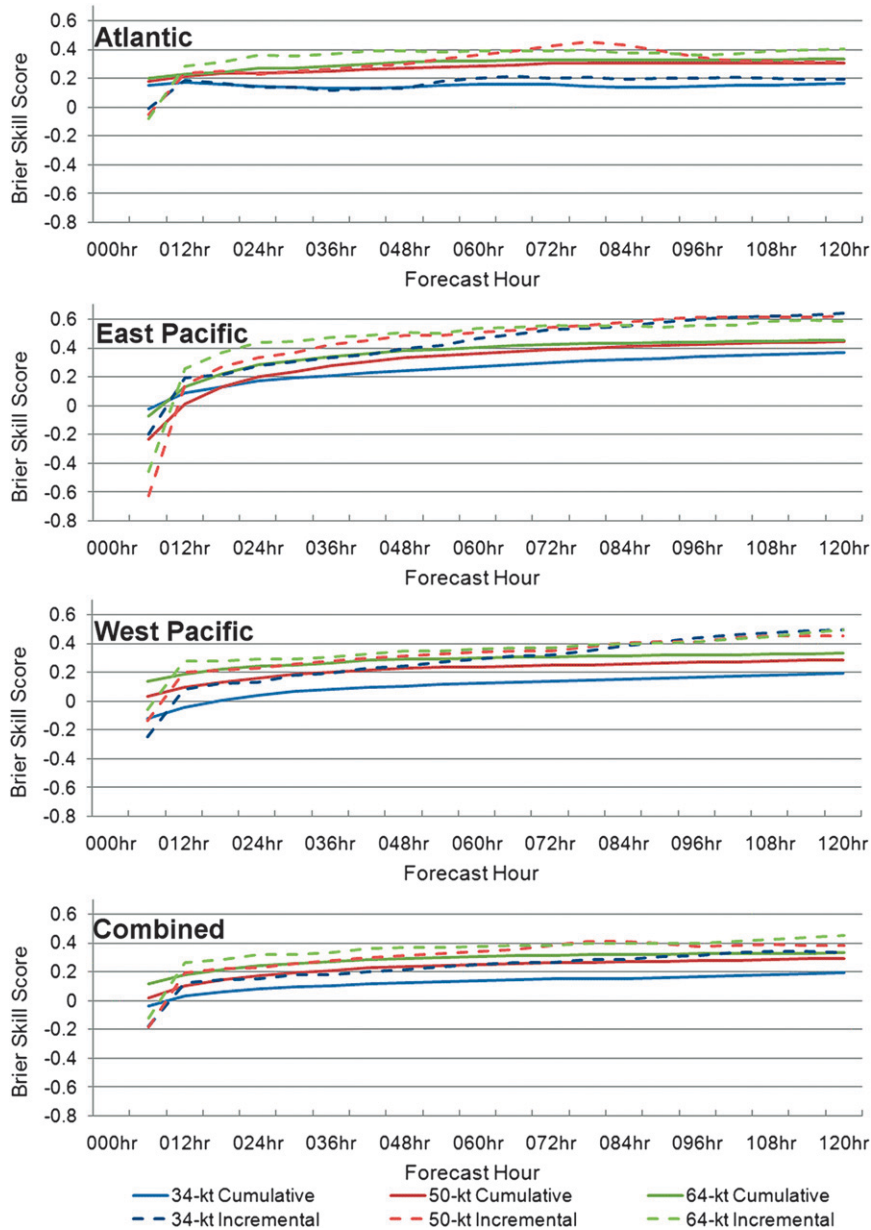


FIG. 9. The BSSs associated with the 2006–07 MCP model verification in which the deterministic forecast is used as the reference for the North Atlantic (1° – 50° N, 110° – 1° W), eastern North Pacific (1° – 40° N, 180° – 75° W), western North Pacific (1° – 50° N, 100° E– 180°), and the multibasin domain (1° – 60° N, 100° E– 1° W) are shown in the panels starting from the top, respectively. Results for the cumulative probabilities are given by solid lines and those from incremental probabilities are given by dashed lines. Line colors blue, red, and green correspond to biases associated with 34-, 50-, and 64-kt wind probabilities. The scale is identical for all basins.

d if the event was not predicted and did not occur. The values of a , b , c , and d are calculated for probability thresholds ranging from 0 to 1 with an increment of 0.01.

The ROC diagram is created by plotting the false alarm rate $c/(c + d)$ along the x axis and the hit rate $a/(a + b)$ along the y axis for each of the probability

thresholds. The ROC skill score is the area under this curve minus the area under the line where the false alarm rate equals the hit rate, with the result multiplied by two. The ROC skill score determines the ability of a method to discriminate between events and nonevents. A perfect skill score is equal to 1.0 (where the hit rate is 1

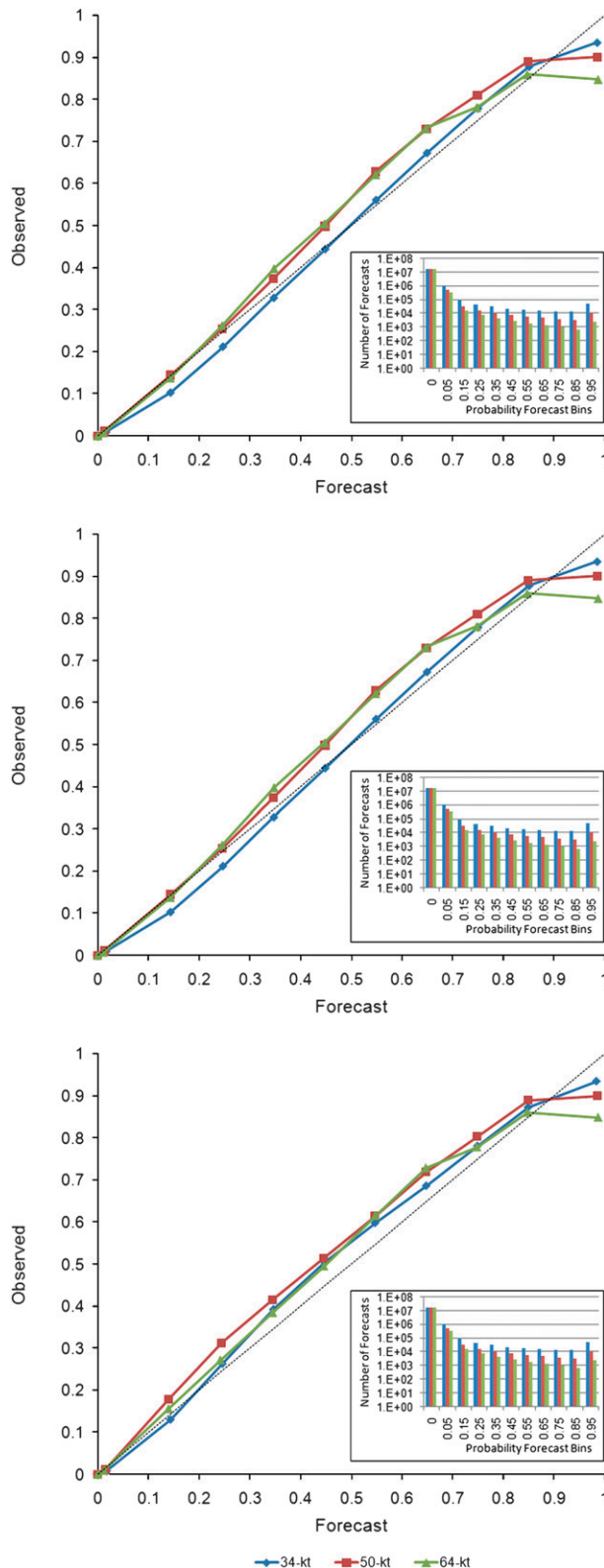


FIG. 10. Reliability diagrams (calibration functions) for (top) 36 h, (middle) 72 h, and (bottom) 120 h for 34-kt (blue), 50-kt (red), and 64-kt (green) probability forecasts over the entire MCP model domain. The insets show the refinement distributions for each case.

TABLE 3. The ROC skill scores ($\times 100$) for the MCP model forecasts of cumulative and incremental 34-, 50-, and 64-kt wind speed probabilities.

Cumulative probability forecasts					
Wind speed (kt)	24 h	48 h	72 h	96 h	120 h
34	92.6	92.7	92.4	92.0	91.5
50	91.5	92.1	91.9	91.4	90.7
64	91.9	91.9	90.9	89.8	88.1
Incremental probability forecasts					
Wind speed (kt)	24 h	48 h	72 h	96 h	120 h
34	90.7	89.4	83.6	83.4	85.3
50	89.1	87.7	80.7	72.5	50.2
64	91.8	83.9	65.4	43.2	16.8

and the false alarm rate is 0 for all probability thresholds except for a threshold of 0, where both equal 1) and the worst possible score is -1.0 (i.e., completely incorrect discrimination). Any score greater than zero is considered skillful. The MCP model was found to have high ROC skill scores (Table 3). However, because the verification is performed on the large grid, which is composed of mostly nonevents, the ROC statistics are rather difficult to interpret beyond the information provided by the skill score, which indicates that the MCP model has the ability to discriminate events from nonevents. The ROC statistics are more meaningful for specific forecasts when the events and nonevents are of more equal magnitudes, such as those associated with landfalling TC events for restricted sections of the coast.

The ROC diagrams can also be used as guidance for choosing a probability threshold for a yes–no decision. For example, the probability that corresponds to the point closest to the upper-left corner of the diagram is sometimes used because it has the best balance between a high hit rate and low false alarm rate. However, because of the very large number of grid points with zero probabilities, this ROC diagram is less useful for this application and the threat scores based on conditional probability thresholds were examined. The threat score is calculated by dividing the number of correct forecasts (hits) by the sum of the hits, misses, and false alarms $a/(a + b + c)$ and has values that range from 0 to 1 (perfect). It can also be interpreted as the ratio of the intersection of the area where an event was predicted to occur and the area where it did occur to the union of those two areas. An advantage of the threat score is that it does not include element d of the contingency table, which is the large number of grid points away from the storm where the probability is at or near zero.

Since the contingency table values were calculated for a large number of probability thresholds, the probability

threshold that maximizes the threat score was determined. Figure 11 shows the maximum threat score associated with the MCP model forecasts versus forecast time, and Fig. 12 shows the corresponding probability threshold associated with that maximum threat score for each basin and the full MCP model domain, respectively. The probability values that maximize the threat scores can be used as guidelines for optimal yes–no decisions. For example, from Fig. 12 the optimal probability for the 24-h cumulative probability of 64-kt wind for the Atlantic was about 35% (solid green line in top panel at the 24-h point). In real time, when the probability exceeds this value at 24 h, a yes decision would be made. The corresponding threat score for this point from Fig. 11 is about 45%, which indicates that a little $<1/2$ of the predicted and observed areas will overlap.

The probabilities that optimize the threat score provide a reasonable basis for a yes–no decision since they maximize the overlap between the predicted and observed areas of occurrence. However, these threshold values are highly application dependent, and could also include additional factors such as cost and risk. For example, a decision with a very high cost of inaction (such as coastal evacuation) would likely require a lower probability threshold for the yes–no decision.

To summarize this section, the MCP model forecasts were shown to be more skillful than the deterministic forecasts in determining the probability of the occurrence of 34-, 50-, and 64-kt winds (Fig. 9) with relatively small overall biases (Fig. 8). The model produces well-calibrated and high-confidence probabilistic forecasts of those same wind thresholds (Fig. 10) and shows skill in discriminating events from nonevents on the large-scale verification grids examined here (Table 3). Furthermore, threshold probabilities (Fig. 12) that maximize the threat score (Fig. 11) provide general guidance for the use of the MCP model output for yes–no decisions. Thus, the MCP model provides useful probabilistic forecasts of 34-, 50-, and 64-kt wind occurrences that further enhance the information contained within official 5-day forecasts of TC track, structure, and intensity.

5. Concluding remarks

This paper described the new wind probability model that became operational at NHC, CPHC, and JTWC beginning in 2006. This model replaced the older Hurricane Probability Program (HPP) product that estimated the probability of the center of a tropical cyclone coming within 60 n mi of a given point out to 72 h, and which had been utilized at NHC since 1983. The new model accounts for the uncertainties in the track, intensity, and wind structure forecasts, and estimates the probabilities

of 34-, 50-, and 64-kt winds out to 120 h for all tropical cyclones in the Northern Hemisphere from the Greenwich meridian to 100°E. Because of the interdependence of the track, intensity, and structure forecasts, especially when cyclones interact with land, a Monte Carlo method is used where 1000 realizations are generated by randomly sampling from the operational track and intensity forecast error distributions from the past 5 yr. The horizontal extents of the 34-, 50-, and 64-kt winds for the realizations are obtained from a climatology and persistence wind radii model and its underlying error distributions. Serial correlations of the track, intensity, and wind radii forecast errors are accounted for in the random sampling technique, and special procedures are implemented to account for cases where the official forecast is over land, but the track in a realization is over water, and vice versa.

The convergence of the MCP model was evaluated by running cases with different numbers of realizations. Results showed that with 1000 realizations, the average probability error was $<0.6\%$ and the maximum error anywhere in the domain was $<4\%$ for all three wind speed thresholds. To a good approximation, the error of the MCP model is inversely proportional to the square root of N , where N is the number of realizations.

The operational MCP model forecasts from 2006 to 2007 were evaluated using a number of metrics commonly used for probabilistic forecasts. Results show that over the combined Atlantic and Pacific domains, the model is relatively unbiased and the forecasts are skillful using a Brier skill score. The baseline for the Brier skill score is the deterministic forecast from the operational centers converted into a binary probabilistic forecast. The model is also skillful based on the relative operating characteristic skill score, and the results are well calibrated and have high confidence based on reliability diagrams. Probability thresholds that optimize the threat score were also shown for rough guidance in utilizing the MCP model products for yes–no decisions.

The output from the MCP model is disseminated in a number of forms from the operational centers, including text, graphical, and gridded products. A separate paper is in preparation describing these products in greater detail. A number of applications that utilize the new MCP model output are also being developed. For example, the new products are being incorporated into tropical cyclone weather support at the Kennedy Space Center in Florida (Winters et al. 2007). An application for automating NWS Weather Forecast Office (WFO) products during tropical cyclone landfalls, based in part on the MCP model, is currently under development (Santos et al. 2009). The MCP products have also shown promise in providing guidance for the issuance of tropical

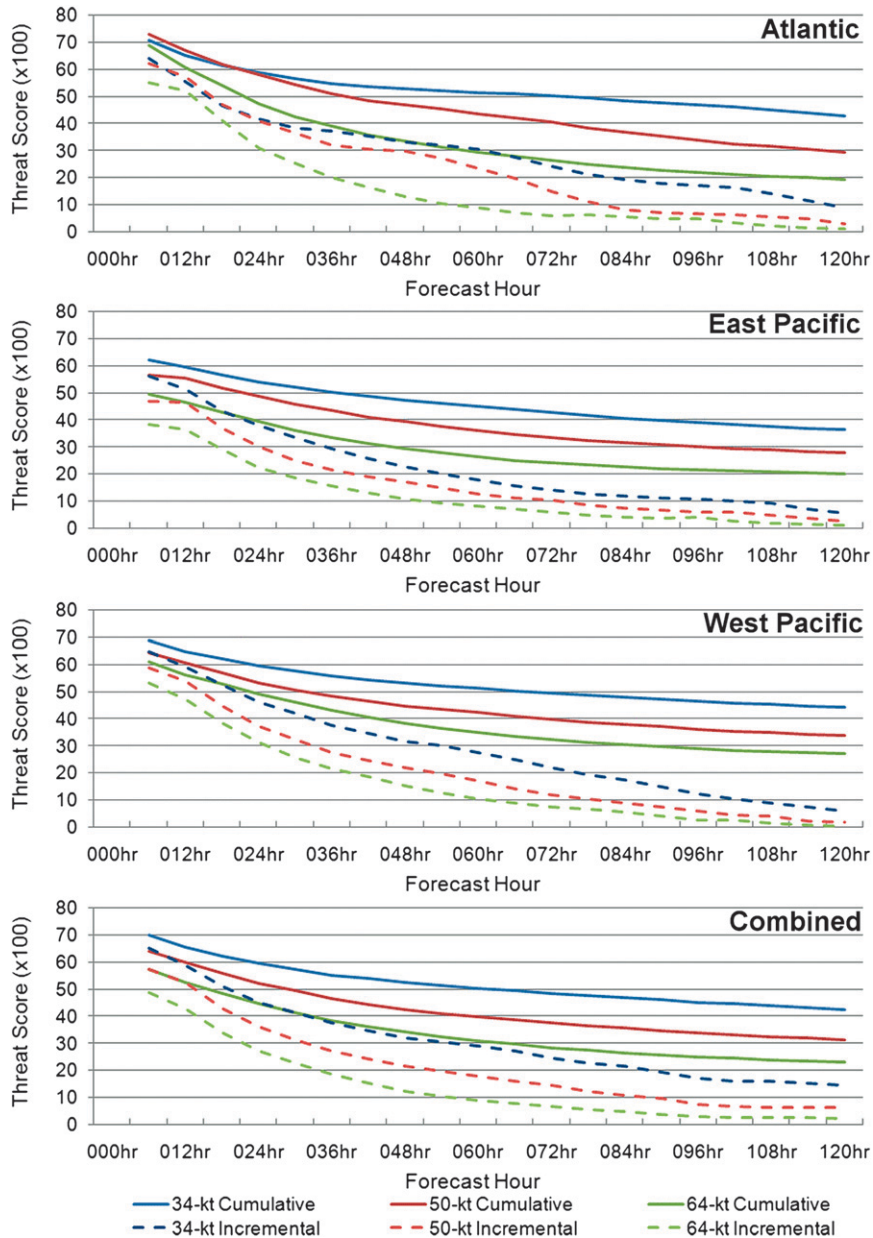


FIG. 11. The maximum conditional threat scores ($\times 100$) associated with the 2006–07 MCP model verification for the North Atlantic (1° – 50° N, 110° – 1° W), eastern North Pacific (1° – 40° N, 180° – 75° W), western North Pacific (1° – 50° N, 100° E– 180°), and the multibasin domain (1° – 60° N, 100° E– 1° W) are shown in the panels starting from the top, respectively. Results for the cumulative probabilities are given by solid lines and those from incremental probabilities are given by dashed lines. Line colors blue, red, and green correspond to biases associated with 34-, 50-, and 64-kt wind probabilities. The scale is identical for all basins.

cyclone watches and warnings (Mainelli et al. 2008). It is expected that new applications will continue to be developed as users gain experience with the new probability products.

The official track and intensity error probability distributions utilized in the MCP model will continue to

be updated annually using the previous 5 yr of forecast errors. As the official forecasts improve, the MCP model probabilities will exhibit less spread around the official forecasts, which increases the probabilities near the official track positions. A current limitation of the model is that basin-wide error statistics are utilized. Goerss

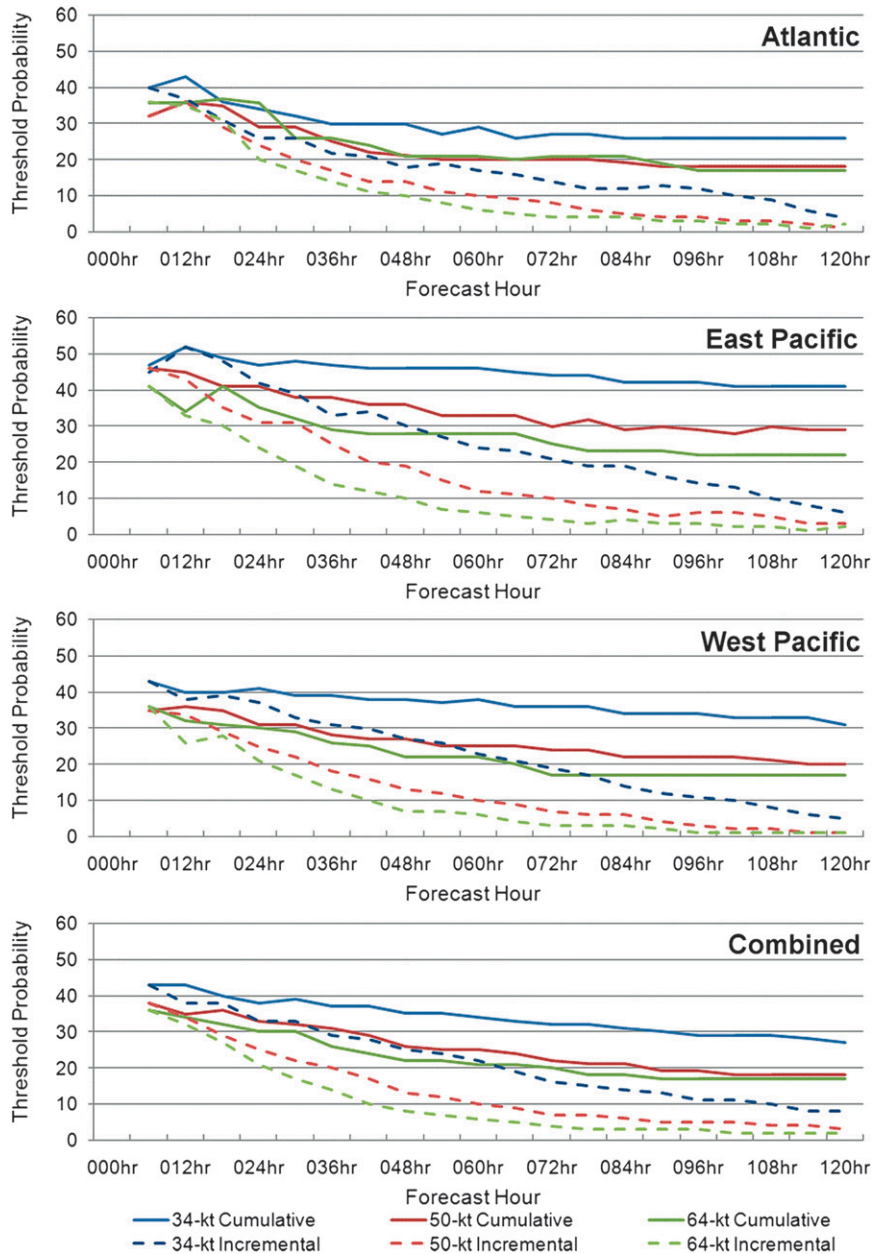


FIG. 12. The probability thresholds associated with the maximum conditional threat scores shown in Fig. 11 and based on the 2006–07 MCP model verification for the North Atlantic (1°–50°N, 110°–1°W), eastern North Pacific (1°–40°N, 180°–75°W), western North Pacific (1°–50°N, 100°E–180°), and the multibasin domain (1°–60°N, 100°E–1°W) are shown in the panels starting from the top, respectively. Results for the cumulative probabilities are given by solid lines and those from incremental probabilities are given by dashed lines. Line colors blue, red, and green correspond to biases associated with the 34-, 50-, and 64-kt wind probabilities. The scale is identical for all basins.

(2007) showed that it is possible to estimate the error of a given track forecast based upon the spread of an ensemble of track models, and other information known at the time the official forecast is made. Work is under way to incorporate this information into the MCP model, so

that the probability distributions will depend on the current forecast situation, rather than just on the error statistics from the past 5 yr. The probabilities will have a wider spread when the uncertainty is large and a narrower spread for cases with a higher-confidence forecast.

This new version of the model will be evaluated during the 2009 hurricane season and be implemented in the 2010 operational MCP model if approved by NHC.

Acknowledgments. This research was partially supported by the NOAA Joint Hurricane Testbed and the Insurance Friends of the National Hurricane Center. C. Sampson was also partially supported by the Office of Naval Research. The authors thank Ed Rappaport, Max Mayfield, Jim Gross, Chris Landsea, Chris Sisko, James Franklin, Alison Krautkramer, and Christopher Juckins for their support of and valuable input to this research. Two anonymous reviews also provided valuable suggestions. The views, opinions, and findings in this report are those of the authors and should not be construed as an official NOAA or U.S. government position, policy, or decision.

REFERENCES

- Cashwell, E. D., and C. J. Everett, 1959: *A Practical Manual on the Monte Carlo Method for Random Walk Problems*. Pergamon Press, 153 pp.
- Franklin, J. L., 2008: National Hurricane Center forecast verification. Preprints, *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc., 13A.1. [Available online at <http://ams.confex.com/ams/pdfpapers/137870.pdf>.]
- Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985–1993.
- Iwabuchi, H., 2006: Efficient Monte Carlo methods for radiative transfer modeling. *J. Atmos. Sci.*, **63**, 2324–2339.
- Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC 22, Coral Gables, FL, 21 pp. [Available online at www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.pdf.]
- Kaplan, J., and M. DeMaria, 1995: A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *J. Appl. Meteor.*, **34**, 2499–2512.
- Knaff, J. A., C. R. Sampson, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence. *Wea. Forecasting*, **22**, 781–791.
- Mainelli, M., R. D. Knabb, M. DeMaria, and J. A. Knaff, 2008: Tropical cyclone wind speed probabilities and their relationships with coastal watches and warnings issued by the National Hurricane Center. Preprints, *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc., 13A.3. [Available online at <http://ams.confex.com/ams/pdfpapers/137827.pdf>.]
- Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting*, **14**, 713–725.
- NRL, cited 2009: The typhoon havens handbook. [Available online at <http://www.nrlmry.navy.mil/~cannon/thh-nc/0start.htm>.]
- Powell, M. D., S. H. Houston, L. R. Amat, and N. Morisseau-Leroy, 1998: The HRD real-time hurricane wind analysis system. *J. Wind Eng. Indust. Aerodyn.*, **77&78**, 53–64.
- Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419.
- Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240.
- Santos, P., M. DeMaria, D. Sharp, and S. Kiser, 2009: The determination of optimal thresholds of tropical cyclone incremental wind speed probabilities to support expressions of uncertainty in text forecasts. Preprints, *Eighth Symp. on the Urban Environment*, Phoenix, AZ, Amer. Meteor. Soc., P1.5. [Available online from <http://ams.confex.com/ams/pdfpapers/147019.pdf>.]
- Sheets, R. C., 1984: The National Weather Service Hurricane Probability Program. NOAA Tech. Rep. NWS 37, Silver Spring, MD, 70 pp.
- , 1985: The National Weather Service Hurricane Probability Program. *Bull. Amer. Meteor. Soc.*, **66**, 4–13.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 59, Academic Press, 627 pp.
- Winters, K. A., J. W. Weems, F. C. Flinn, G. B. Kubat, and S. B. Cocks, 2007: Providing tropical cyclone weather support to space launch operations. Preprints, *27th Conf. on Hurricanes and Tropical Meteorology*, Monterey, CA, Amer. Meteor. Soc., 9A.5. [Available online from <http://ams.confex.com/ams/pdfpapers/108570.pdf>.]