

An Operational Statistical Typhoon Intensity Prediction Scheme for the Western North Pacific

JOHN A. KNAFF

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

CHARLES R. SAMPSON

Naval Research Laboratory, Monterey, California

MARK DEMARIA

NOAA/NESDIS, Fort Collins, Colorado

(Manuscript received 24 June 2004, in final form 26 January 2005)

ABSTRACT

The current version of the Statistical Typhoon Intensity Prediction Scheme (STIPS) used operationally at the Joint Typhoon Warning Center (JTWC) to provide 12-hourly tropical cyclone intensity guidance through day 5 is documented. STIPS is a multiple linear regression model. It was developed using a “perfect prog” assumption and has a statistical–dynamical framework, which utilizes environmental information obtained from Navy Operational Global Analysis and Prediction System (NOGAPS) analyses and the JTWC historical best track for development. NOGAPS forecast fields are used in real time. A separate version of the model (decay-STIPS) is produced that accounts for the effects of landfall by using an empirical inland decay model. Despite their simplicity, STIPS and decay-STIPS produce skillful intensity forecasts through 4 days, based on a 48-storm verification (July 2003–October 2004). Details of this model’s development and operational performance are presented.

1. Introduction

The prediction of tropical cyclone “intensity” or maximum sustained surface winds remains a difficult task in all tropical cyclone basins. On average, forecasts of intensity are slightly better than those produced based upon climatology and persistence (CLIPER). While there has been relatively steady improvement over the years in track forecasting skill with ever-improving numerical models, the skill (relative to CLIPER) of intensity forecasts still lags that of track forecasts (AMS 2000). In recent years, however, the rate of intensity forecast skill improvement in the Atlantic tropical cyclone basin has exceeded that of track prediction.

This has not been the case in the western North Pacific (Furze and Preble 2004).

The relatively low skill of intensity forecasts is primarily due to the complexity of the tropical cyclone intensification process, which involves scale interactions between the environment, the storm, and convection. Intensity forecasting in the western North Pacific is also made more difficult by the limited number of guidance models specifically designed for that task and the inability for current numerical models to be run at the resolutions needed to explicitly resolve convection in a real-time and operational manner.

With the complexities involved in tropical cyclone intensity change, there exists an underlying need for alternative intensity forecast methods to overcome the shortcomings of existing numerical and purely statistical intensity forecast models. Alternative methods would ideally employ the strengths of both statistical models and numerical models. Such an approach would

Corresponding author address: John Knaff, CIRA, Colorado State University, Foothills Campus, Fort Collins, CO 80521-1375.
E-mail: knaff@cira.colostate.edu

combine the statistical methodology with environmental predictors derived from numerical weather forecasts. This methodology is commonly called the statistical–dynamical approach. The Statistical Hurricane Prediction Scheme (SHIPS; DeMaria and Kaplan 1999) has been developed for use in the North Atlantic and east Pacific and is a good example of a statistical model developed using this approach. The operational use of the SHIPS model has been successful as it produces skillful (i.e., smaller errors relative to CLIPER) intensity forecast guidance in the Atlantic and east Pacific tropical cyclone basins.

In the western North Pacific, the Joint Typhoon Warning Center (JTWC) routinely produces intensity forecasts through 72 h and in 2003 began issuing forecasts through 5 days. These forecasts tend to be, as in other tropical cyclone (TC) basins, slightly better than forecasts produced using a combination of climatology and persistence.¹ For instance in 2003, JTWC's intensity forecast errors were 10.7, 16.1, 20.0, 22.4, and 19.0 kt at 24, 48, 72, 96, and 120 h, respectively, based upon 560, 464, 365, 223, and 156 cases. Forecasts made by the 5-day version of the Statistical Typhoon Intensity Forecast (ST5D) (Knaff et al. 2003) produced errors of 12.4, 19.7, 24.7, 26.2, and 26.7 kt for the same cases.

Prior to 2002, operational statistical–dynamic intensity guidance models, like SHIPS, were unavailable to the forecasters at JTWC.² Intensity guidance available at JTWC prior to 2002 included analog techniques (Sampson et al. 1990), a statistical typhoon intensity forecast model (STIFOR; Chu 1994), the Coupled Hurricane Intensity Prediction Scheme (CHIPS; Emanuel 2004), the Innovative Coastal–Ocean Network model (ICON; Grandau and Engel 2004), ST5D (Knaff et al. 2003), and a few numerical models such as the Japanese typhoon model (Kuma 1996), and the Geophysical Fluid Dynamics Laboratory–Navy model (GFDN), which is the U.S. Navy's version of the

Geophysical Fluid Dynamics Laboratory model (Renick 1999). None of the pre-2002 techniques listed above have demonstrated consistent, long-term skill. As a result, the intensity forecasting at JTWC remains rather subjective, as was described in Guard et al. (1992).

In July of 2002, as part of a one-year Office of Naval Research sponsored project, the first version of the Statistical Typhoon Intensity Forecasting Scheme (STIPS) model was implemented in operations at JTWC. On average, forecasts from this version of STIPS were only marginally skillful. A postseason evaluation revealed a few shortcomings in the model that could be linked to individual predictors. Based on this evaluation, STIPS was redeveloped for use in 2003. This paper discusses the most recent version of STIPS, which was made operational on 16 June 2003 and continues to run and provide skillful intensity guidance to forecasters at the JTWC. The datasets used in the development of STIPS are discussed in section 2. The model development, including potential predictors, statistical methodology, predictor selection, and model performance, is discussed in section 3. The operational implementation of this model and its independent performance are discussed in section 4. This will be followed by a discussion and some concluding remarks.

2. Datasets

Five and a half years of Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan et al. 2002; Hogan and Rosmond 1991) analyses were used in the most recent development of STIPS. Specifically, temperature, wind, water vapor pressure, and geopotential height data were collected twice daily for the period 21 July 1997–31 December 2002 at 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, and 1000 hPa. Surface skin temperature fields were also collected for the same period, which are used as sea surface temperature (SST). Surface type (i.e., land or ocean) is determined from a digitized land file that contains the continental areas and large islands in the western North Pacific. For operational and developmental purposes a climatological SST (Levitus 1982) is used when the NOGAPS skin temperature field is unavailable.

The tropical cyclone position and intensity information used in this study came from the JTWC's "best track," which is a postseason reanalysis using additional information not available in the operational setting (JTWC 2004). These files contain the date, time, latitude, longitude, and intensity every 6 h for all storms designated by JTWC as being tropical depres-

¹ It is a common practice to compare tropical cyclone intensity forecasts with a baseline model based upon climatology and persistence. In the western North Pacific there are two such models: the Statistical Typhoon Intensity Forecast (STIFOR) described in Chu (1994) and a similar model that produces forecasts through 5 days (ST5D), which is discussed in Knaff et al. (2003). This CLIPER combination represents a skill threshold. For instance, if an individual model produces a series of intensity forecasts that are significantly better in a statistical sense than STIFOR or ST5D, this model is considered to have skill. For the purposes of this paper the climatology and persistence intensity forecasts for the western North Pacific basin will come from ST5D.

² Models were developed for this purpose (Fitzpatrick 1997; Baik and Paek 2000), but neither model was implemented operationally.

sion³ strength or greater. Because routine aircraft reconnaissance is unavailable in this basin, one caveat concerning the best track is that the intensity is often determined solely from satellite-based methods. The actual errors associated with the use of satellite intensity estimation methods, however, are difficult to determine as the aircraft measures central pressure and flight-level winds and the most commonly used technique, the Dvorak technique, estimates maximum 1-min-averaged surface wind speed (Dvorak 1984). The intensity archived in these historical datasets, as well as operational intensity forecasts, are estimated to the nearest 5 kt (2.58 m s^{-1}) at 6-h intervals. For this reason, model formulation as well as any discussion of intensity in this paper will be in terms of knots instead of meters per second.

3. Model development

The development of the STIPS model closely follows the development of the SHIPS model in the Atlantic and east Pacific tropical cyclone basins as described in DeMaria and Kaplan (1999). As a result, STIPS is a multiple linear regression model. The dependent variables (predictands) are the intensity change from the initial forecast time (DELV) at 12-h intervals of all storms not making landfall. Potential predictors (independent variables), or more precisely parameters that have been documented in the literature to be associated with tropical cyclone intensity change, are created. Then the potential model predictors are evaluated for their combined statistical ability to predict tropical cyclone intensity change. This results in ten predictive equations that are used to make forecasts at each of the ten 12-hourly time periods, 12–120 h.

The resulting equations can predict the intensity changes associated with environmental and climatological tendencies, but not the intensity changes caused by landfall. It is known, however, that the intensity change of some tropical cyclones is strongly influenced by rapid weakening associated with landfall. To account for landfall effects on intensity along the forecast track, the empirical inland decay models discussed in Kaplan and DeMaria (1995, 2001) are used south of 36°N , and north of 40°N , respectively. Between these two latitudes a linear weighting of these two methods is applied for the decay.

³ Tropical depression: A weak tropical cyclone with a definite closed surface circulation and highest sustained wind speeds (averaged over 1 min or longer period) of less than 17 m s^{-1} (34 kt) (Elsberry 1987).

The details of the STIPS model are contained in the following subsections. Section 3a outlines the predictors used in the model development. Section 3b describes the statistical methodology. Section 3c discusses how the final model predictors were selected along with a discussion of their relative importance. It was also found during model development that many of the individual predictors are best described when combined with others. For this reason, section 3d discusses the interpretation of some of the predictor combinations in STIPS. Finally, section 3e discusses the dependent statistics associated with the STIPS model.

a. Potential predictors

The potential predictors used in STIPS development can be divided into two categories: 1) those related to climatology, persistence, and trends of intensity—“static predictors” and 2) those related to current and future environmental and SST conditions—“time dependent predictors.” All of these are derived along the tropical cyclone track. Predictors are developed using a “perfect prog” methodology (Kalnay 2003) where the analyses and actual tropical cyclone best track are used to create the statistical model. However, when STIPS is used operationally, the NOGAPS model forecasts are used to create the predictors along the JTWC tropical cyclone track forecast. Therefore, errors in both the NOGAPS forecast fields and the JTWC track forecast represent additional sources of STIPS intensity forecast errors not accounted for in the developmental data.

The potential static predictors are derived at the initialization time from the date, current intensity, and the 12-h change in intensity and motion, and location. Predictors determined during the recent development of ST5D (Knaff et al. 2003) were included as potential static predictors in STIPS. However, since SST is used in the development of this model, predictors related to location (a proxy for SST in ST5D) were not included in the static predictor pool. This left only terms related to date, intensity, and motion as static predictors (Table 1).

The potential time-dependent predictors are more

TABLE 1. The potential static predictors used in STIPS development.

Predictor	Description
VMAX	Initial intensity
VMAX ²	Initial intensity squared
DVMX	12-h change in intensity
JDAY	Absolute value of yearday minus 248
SPD	Storm translational speed

numerous and require more explanation. These predictors are divided into three basic categories, namely those related to the SST, those related to the moisture fields, and those related to the wind fields. SST values are determined at the storm center by interpolating from oceanic NOGAPS skin temperature values, while moisture and wind-related predictors are area averaged around the center. Time-dependent predictors are also averaged with respect to time along the track from the initial time to the forecast time, providing the mean conditions the storm will experience along its track.

The primary use of the SST is to determine the upper bound of tropical cyclone intensity as a function of SST. This upper bound is commonly referred to as the maximum potential intensity (MPI) and can be estimated theoretically (i.e., Miller 1958; Malkus and Riehl 1960; Emanuel 1988 and Holland 1997) or empirically (i.e., Merrill 1987; DeMaria and Kaplan 1994; Whitney and Hobgood 1997). For the purposes of developing STIPS, the empirical approach is chosen following the methodology employed in DeMaria and Kaplan (1994), who fit an exponential function to the maximum observed tropical cyclone intensity with respect to SST. The SST used to develop this MPI function for the western North Pacific is derived from a climatological analysis described in Levitus (1982), which contained data from the period 1962–88. This 1° latitude \times 1° longitude resolution SST climatology is then interpolated to the storm center following the best track for a 31-yr period (1962–92) to find the SST corresponding to the storm intensity minus the storm speed. The maximum values in each 0.5°C temperature interval are then used to determine the coefficients in the MPI function described below [in Eq. (1)]. Using this procedure, the coefficients are given by $A=38.21$ kt, $B = 170.72$ kt, $C = 0.1909^\circ\text{C}^{-1}$, and $T_0 = 30.0^\circ\text{C}$:

$$\text{MPI} = A + Be^{C(T-T_0)}. \quad (1)$$

In the formulation of STIPS the maximum value of MPI allowed is 185 kt. Figure 1 shows this MPI function along with the data used for its development.

Variations of environmental relative humidity (RH) will affect convective buoyancy through entrainment of unsaturated air. In tropical cyclones, convective available potential energy is relatively small and decreases to almost zero near the center (Bogner et al. 2000). Therefore, relative humidity values in the middle atmosphere should be relatively large, which reduces the entrainment of dry air into cumulus convection. Since convection is the direct source of the tropical cyclone's energy, variations of mid- and upper-level RH should affect tropical cyclone intensification rates. To examine

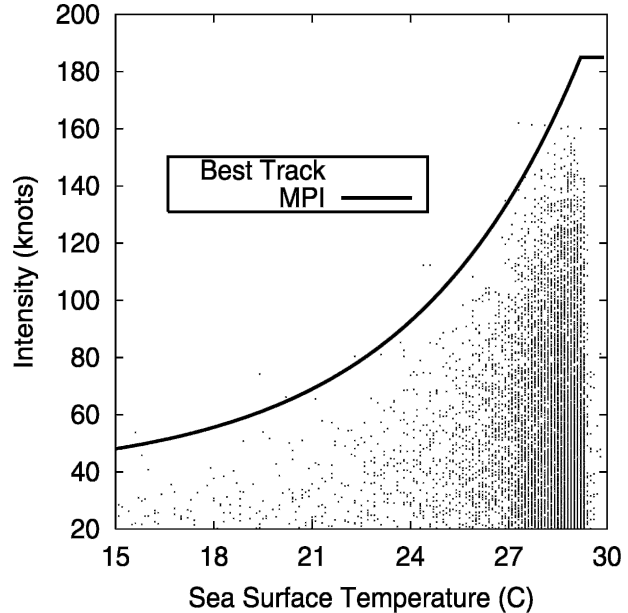


FIG. 1. Plot of the empirical relationship between MPI (kt) and SST ($^\circ\text{C}$) that is used in the development of STIPS. The relationship was derived from 31 yr (1963–93) of data and is shown along with the individual data points (i.e., intensity minus storm speed) used for its development.

the potential effects of environmental mid- and upper-level moisture on tropical cyclone intensity, average RH was calculated in atmospheric layers, 850–700 hPa (RHLO) and 500–300 hPa (RHHI), within an annulus of 200–800 km from the center of the cyclone. This annular average is used to estimate environmental parameters. The 200-km radius is used to remove the innermost regions of the tropical cyclone from the analysis where the NOGAPS model uses synthetic observations to initialize the tropical cyclone (Goerss and Jefferies 1994).

At 200 hPa, the zonal wind (U_{200}), the temperature (T_{200}), the divergence (δ_{200}), and the relative eddy flux convergence (REFC) are examined. The zonal wind and the temperature are again averaged in the same 200–800-km annulus as the relative humidity, and the divergence is averaged within a slightly larger 1000-km circle. The REFC is calculated within 600 km:

$$\text{REFC} = -r^{-2} \frac{\partial}{\partial r} (r^2 \overline{U'_L V'_L}), \quad (2)$$

where U is the radial wind, V is the tangential wind, r is radius, the overbar represents an azimuthal mean, the primes indicate a departure from the azimuthal mean, and the subscript L indicates that the calculation is done following the storm motion.

In addition to these potential predictors at 200 hPa,

the 850-hPa vorticity (ζ_{850}) is averaged within a radius of 1000 km and several measurements of vertical wind shear are calculated within the 200–800-km annulus. As was the case with relative humidity, the core region of the storm is removed for the measurement of environmental vertical wind shear. A traditional approach for calculating vertical wind shear is to simply use the magnitude of the vector difference between layers. Using this approach, two time-dependent predictors were created: the 200 – 850 hPa wind difference (SHRD) and the 500 – 850 hPa wind difference (SHRS). In addition to the scalar value of shear (i.e., SHRD and SHRS), the zonal components of the shear in these layers were also created (USHRD, USHRS). As an alternative to the traditional measures of vertical wind shear, a generalized vertical wind shear can be calculated and tested. The generalized shear at a point is calculated from the mass-weighted root-mean-square deviations of the winds from the mass-weighted deep-layer mean winds times a factor of 4 to make the values equivalent to the more conventional measure of 200–850 hPa for the case when the shear is linear with respect to pressure:

$$\text{SHRG} = 4.0 \times \sum_{p=850}^{p=200} w_p \sqrt{(u_p - \bar{u})^2 + (v_p - \bar{v})^2}, \quad (3)$$

where

$$\bar{u} = \sum_{p=850}^{p=200} w_p u \text{ is the deep layer zonal wind,}$$

$$\bar{v} = \sum_{p=850}^{p=200} w_p v \text{ is the deep layer meridional wind,}$$

and w_p are mass weights.

Potential predictors in STIPS also include several quadratic terms. The MPI squared as well as the MPI times VMAX were added following the notion that these terms may account for some inherent nonlinearity in the same way they do in SHIPS (DeMaria and Kaplan 1999) and in ST5D (Knaff et al. 2003). The terms SHRG times the sine of the latitude (along the storm track) and SHRD times the sine of the latitude (along the storm track) were also tested because storms at higher latitudes tended to be less sensitive to vertical wind shear (DeMaria 1996). Adding these predictors results in 15 synoptic-scale predictors available for testing in STIPS as listed in Table 2.

b. Statistical methodology

When developing a multiple regression model, one commonly uses a method to select predictors based

TABLE 2. Potential synoptic predictors available for inclusion into STIPS.

Predictor	Description
MPI	Maximum potential intensity based upon Eq. (1)
MPI ²	MPI squared
MPI * VMAX	MPI times the initial intensity
RHLO	Area-averaged (200–800 km) relative humidity at 850–700 hPa
RHHI	Area-averaged (200–800 km) relative humidity at 500–300 hPa
U200	Area-averaged (200–800 km) zonal wind at 200 hPa
T200	Area-averaged (200–800 km) temperature at 200 hPa
δ_{200}	Area-averaged (0–1000 km) 200-hPa divergence
REFC	Relative eddy flux convergence within 600 km [see Eq. (2)]
SHRG	Generalized 200–850-hPa vertical wind shear [see Eq. (3)]
SHRS	Area-averaged (200–800 km) 500–850-hPa wind shear
SHRD	Area-averaged (200–800 km) 200–850-hPa wind shear
USHRS	Area-averaged (200–800 km) 500–850-hPa zonal wind shear
USHRD	Area-averaged (200–800 km) 200–850-hPa zonal wind shear
SHRD * SIN(LAT)	200–850-hPa wind shear times the sine of the latitude
SHRG * SIN(LAT)	Generalized wind shear times the sine of the latitude
ζ_{850}	Area-averaged (0–1000 km) 850-hPa relative vorticity

upon their combined ability to predict the dependent variable or predictand. For this model a stepwise procedure is used to select variables from the predictor pool at each forecast time (see IMSL 1987). The significance of individual predictors is based on a standard *F* test (Panofsky and Brier 1968). A 99% statistical significance level is required for an individual predictor to be included initially in the model. Once in the model, a predictor can only be removed if its significance level becomes less than 98% by the addition/removal of another predictor. Because the model development uses two different ways of measuring vertical wind shear, namely the SHRG term, and two-layer scalar differences SHRD and SHRS, two pools of predictors were created. These pools were identical except for the treatment of vertical wind shear predictors shown in Tables 1 and 2. The stepwise procedure identifies important predictors at each forecast time, which sometimes results in erratic forecasts. To avoid this problem, all of the predictors chosen for any forecast period by the stepwise selection procedure are included in the final

group of predictors. Using the predictors in this final group, a single multiple regression model is created for each forecast time. In the next subsection the results of this regression procedure, including the predictors and their relative importance through 120 h, are discussed.

c. STIPS model formulation

The stepwise predictor selection procedure was performed on the two-predictor pools and resulted in 11 predictors being selected for use in model formulation (Table 3). There were 1921 cases available at 12 h and 608 cases at 120 h in the developmental dataset. The 11 predictors chosen came from the predictor pool containing the SHRD and SHRS terms. It was found that the regression results were slightly better using this vertical shear parameterization (i.e., SHRD, SHRS, USHRD, USHRS) than the generalized shear parameterization (SHRG). The predictor selection procedure also found that the REFC term is not significant at any time period, in agreement with Fitzpatrick (1997). The potential predictors ζ_{850} and RHLO were also found to be statistically unimportant at all forecast times. Interesting is the inclusion of RHHI, but not RHLO, which is identical to the relationship used in the 2002 version of the SHIPS model (DeMaria et al. 2003).

Table 3 also lists the forecast time at which the 11 predictors are most important (statistically significant) to the model's forecast. Not surprisingly the predictors related to current conditions, namely the static predictors, were most important to the model at the 12-h period, with the exception of storm speed. The factors related to vertical wind shear, T_{200} , RHHI, and MPI become most important for the 24- and 36-h forecasts.

The relative contribution of each predictor for each forecast period is illustrated by the values associated with the normalized regression coefficient. A simple way to interpret these coefficients is as follows: the

larger the normalized coefficient, the greater its contribution to the individual forecast equation. To form normalized coefficients, all of the predictors, as well as the predictand, are normalized before they are incorporated into the regression equation. Subtracting the population mean and dividing this result by the population standard deviation accomplishes the normalization. Table 4 lists the normalized coefficients associated with each predictor for each forecast equation through 120 h. The number of cases used to develop the regression equations are shown in parentheses at the top of the table with the forecast times, and the 99% statistical significance of each normalized coefficient is indicated by boldface italics.

d. Interpretation of predictor combinations

The information presented in Table 4 requires some explanation. The static term VMAX changes sign with time indicating that intensity tends to persist for the first 72 h and then storms in the sample decay with time. The coefficients associated with DVMX or persistence indicate that the trend in intensity persists through 60 h. As a result, a storm that has recently had increasing sustained winds will tend to intensify through 60 h. Similarly, a storm that is moving faster than the mean (~ 9 kt) will tend to intensify. One possible explanation for this finding is that when storms move slower, their intensity may be limited by the upwelling of cooler water (e.g., Schade and Emanuel 1999; Schade 2000). The interpretation of the SHRD is consistent with our understanding of tropical cyclones in that strong shear is related to weakening. The USHRD term indicates that westerly shear favors intensification. This term replaced two terms, U_{200} and LSHRD, during the predictor selection process—reducing the predictor pool by one member. The finding that westerly shear (and westerly 200-hPa winds) is favorable for intensification is different than relationships found in the Atlantic and east Pacific. One can only speculate about this relationship. There are a number of possible causes including, but not limited to, the following: 1) Storms at higher latitude can withstand greater amounts of vertical wind shear as described in DeMaria (1996). 2) The influence of westerly winds approaching the cyclone on the poleward side of the tropical cyclone [e.g., tropical upper tropospheric troughs (TUTTs; Sadler 1976, 1978) or midlatitude troughs], which acts to weaken the commonly observed easterly shear over the cyclone center, but creates weak westerlies in the 200–800-km annulus where the vertical wind shear is measured. This mechanism is likely operating in this basin where most of the storms that form are associated with the monsoon trough (Zehr 1992; Briegel and Frank 1997), which of-

TABLE 3. List of the final predictors used in STIPS along with the forecast hour when they are most statistically significant.

Predictor	Most important forecast hour
1) DVMX	12
2) SPD	60
3) VMAX	12
4) VMAX ²	12
5) MPI	24
6) MPI ²	24
7) MPI * VMAX	12
8) SHRD	12
9) USHRD	60
10) T_{200}	36
11) RHHI	24

TABLE 4. A list of normalized regression coefficients used in the STIPS model. The predictors are listed on the left side of the table and the forecast times are listed at the top with the number of dependent cases used to develop the equation displayed in parentheses. The 99% statistical significance level from an F test is indicated by bold italics.

Predictor	12 h (1921)	24 h (1716)	36 h (1530)	48 h (1353)	60 h (1188)	72 h (1040)	84 h (911)	96 h (800)	108 h (700)	120 h (608)
1) DVMAX	0.35	0.27	0.19	0.13	0.07	0.05	0.04	0.02	0.01	0.00
2) SPD	0.05	0.07	0.07	0.08	0.09	0.09	0.09	0.08	0.06	0.07
3) VMAX	-0.57	-0.54	-0.44	-0.27	-0.13	-0.08	0.24	0.38	0.48	0.48
4) VMAX ²	-0.66	-0.97	-1.14	-1.18	-1.14	-1.05	-0.94	-0.82	-0.73	-0.61
5) MPI	-0.73	-0.89	-0.94	-0.81	-0.71	-0.59	-0.52	-0.42	-0.37	-0.37
6) MPI ²	0.66	0.81	0.90	0.85	0.81	0.77	0.75	0.71	0.69	0.70
7) MPI * VMAX	1.04	1.23	1.23	1.03	0.80	0.48	0.19	-0.08	-0.28	-0.41
8) SHRD	-0.15	-0.19	-0.21	-0.21	-0.21	-0.22	-0.21	-0.21	-0.22	-0.23
9) USHRD	0.08	0.08	0.09	0.11	0.13	0.13	0.13	0.12	0.10	0.08
10) T200	-0.07	-0.08	-0.09	-0.08	-0.08	-0.08	-0.07	-0.08	-0.08	-0.08
11) RHHI	0.10	0.14	0.16	0.18	0.20	0.20	0.21	0.21	0.20	0.18

ten has strong upper-level easterlies present that inhibit intensification. 3) An empirical observation that typhoon maximum intensity is often associated with recurvature (Riehl 1972). Regardless of the exact cause, the combined statistical effect of these two wind shear factors is shown as a contour plot of 24-h intensity change as a function of SHRD and USHRD in Fig. 2. The 200-hPa temperature is a negative factor and indicates that colder 200-hPa temperatures were a favorable condition for storm intensification and is similar to

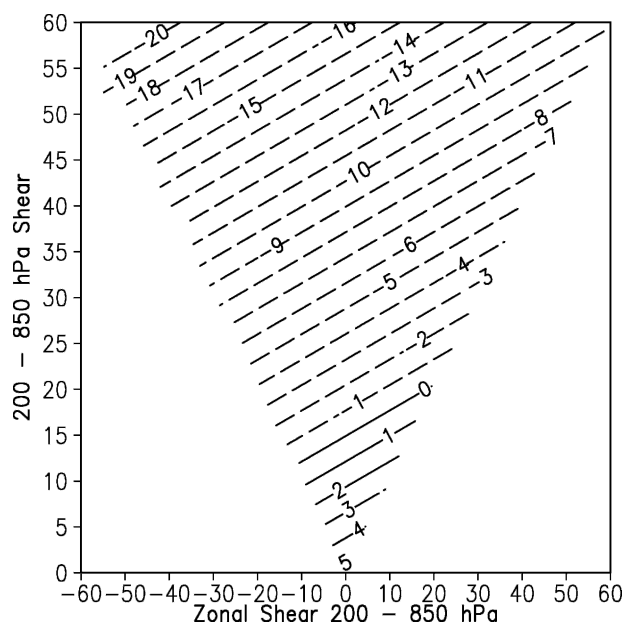


FIG. 2. Contour plot of 24-h tropical cyclone intensity change as a function of the magnitude of vector difference of the 200 - 850 hPa wind (SHRD) and the magnitude of the 200 - 850 hPa zonal wind difference (USHRD). Negative values are indicated by dashed contours.

findings of DeMaria and Kaplan (1999). Not surprisingly, greater upper-level relative humidity is also a favorable for intensification.

The most difficult terms to interpret involve the MPI, MPI², VMAX, VMAX², and MPI * VMAX terms, which describe the complicated relationships between SST and current intensity. A good way of illustrating the effect of these terms is to show the intensity change as a function of current intensity and of current SST averaged along the storm track. This type of plot is created using the STIPS regression coefficients applied to the MPI, MPI², VMAX, VMAX², and MPI * VMAX terms for the 12- and 24-h forecast times (Figs. 3a and 3b). These contours (Fig. 3) represent the mean forecast intensity change as a function of current intensity and SST alone, and should be thought of as intensification potential. The combined result of these predictors for the 12-h forecast time (Fig. 3a) shows that there is a preferred range of initial intensities and SSTs that are favored for intensification. A similar shape is observed for the 24-h forecast time (Fig. 3b). In general, storms with SSTs greater than 28°C and initial intensity ranging from 30 to 110 kt are most likely to intensify. An interesting result comes from the ranges where these factors predict large 12- and 24-h intensity changes (~15 kt day⁻¹) have SSTs greater than 29.25°C and current intensities between 50 and 75 kt. These same intensities correspond well with eye formation/existence as implied by the aircraft-based results in Weatherford and Gray (1988) and satellite-based results in Dvorak (1984). The maximum intensity change for both of these forecast time periods occurs when the SST is greater than 29.25°C and the current intensity is approximately 60 kt.

Sometimes in the forecasting procedure, it is useful to know if a particular condition is favorable or unfa-

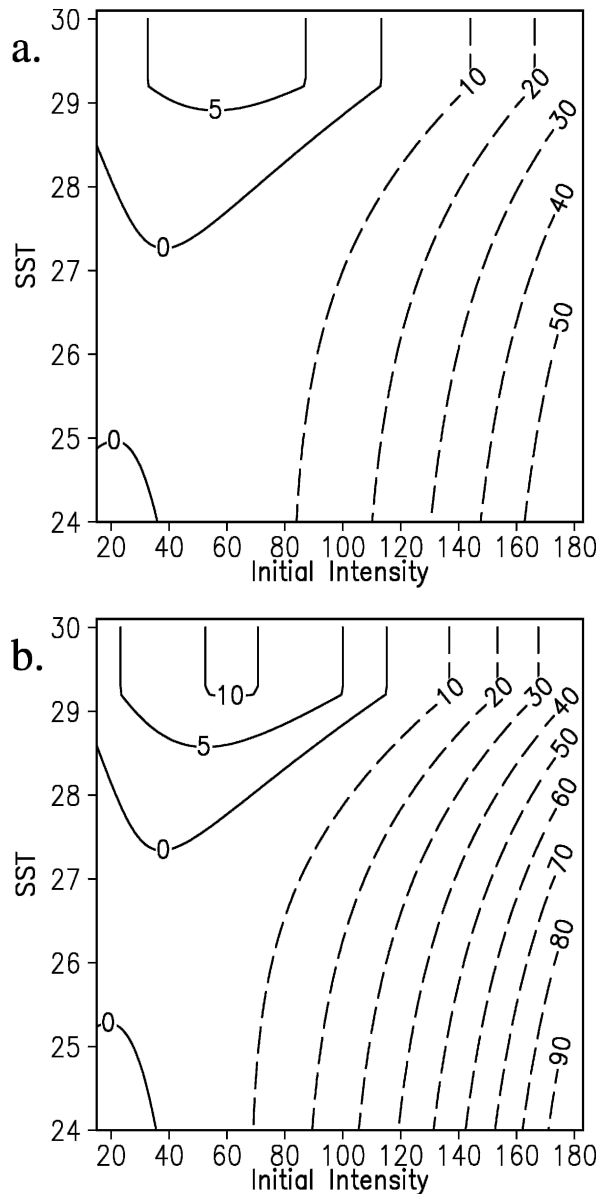


FIG. 3. The combined effect of the MPI, MPI^2 , VMAX, $VMAX^2$, and $MPI * VMAX$ predictors in the STIPS model—or intensification potential as a function of initial intensity (VMAX) and of SST. Shown are (a) the 12-h potential intensity changes and (b) the 24-h potential intensity changes. Note that MPI is solely a function of SST in this calculation [e.g., Eq. (1)] and negative values are indicated by dashed contours.

avorable. The multiple regression procedure also provides threshold values of the individual predictors as they relate to intensity change in the STIPS-dependent data. These threshold values can determine when a predictor has a positive versus negative affect on intensity change. Table 5 lists the threshold values for the synoptic predictors at the 12–120-h forecast times, excluding MPI and quadratic terms. SHRD and T200 have

negative coefficients in Table 4 while USHRD, RHHI, and SPD have positive coefficients. As a result, values of USHRD, RHHI, and SPD greater than the thresholds lead to intensification, and values of SHRD and T200 smaller than the thresholds would also lead to intensification. Regression coefficients calculated from this sample indicate that the vertical shear (i.e., SHRD) becomes a negative influence on intensification of the system in a range of 14.1–15.6 kt. At the same time the zonal wind shear (USHRD) is most favorable when it is greater than about 1 kt (also see the discussion of Fig. 2). The 500–300-hPa relative humidity (RHHI) is favorable when greater than ~59%, which is a more favorable environment for deep convection (e.g., less dry-air entrainment). Since tropical cyclones are fueled by deep convection, it follows that lower temperatures at 200 hPa are also favorable. In this sample, the temperature at 200 hPa (T_{200}) is favorable when it is colder than ~ -51°C . Storm speeds (SPD) greater than 8–9 kt also favor intensification.

For completeness it must be mentioned that the predictand (DELV) used in the STIPS model has a non-zero constant at each time, which increases from about 2 kt at 12 h to 25 kt at 120 h. These mean DELV values are shown in Table 5. This result simply indicates that storms generally intensify with time. This means intensification is primarily caused by the fact that the final intensity values in the best track tends to have higher intensities than the initial intensity values. For example, the best tracks for tropical storms that undergo extratropical transition are considered complete when the extratropical transition is complete. This can happen at intensities of 50 kt or greater.

e. Dependent model statistics

The potential forecast capabilities (best-case scenario) of this model in terms of percent variance explained (R^2) and mean absolute error (MAE) can be estimated from the dependent data (Table 6). MAE is shown to increase from a value of 5.6 kt at 12 h to a value of 21.8 kt at 120 h. The percent variance explained of DELV starts with a relatively large value of 40% at 12 h and increases to only 67.8% at 120 h, keeping in mind that most (increasing 19.6%) of the variance of this variable is explained during the 12–60-h forecast time versus the 72–120-h forecast time (increasing 8.2%). In independent predictions, these statistics are expected to degrade due to the influences of artificial statistical skill (see Knaff and Landsea 1997) and the errors associated with the perfect prog assumption—particularly when track deviations are larger than 200 km.

TABLE 5. Threshold values for the synoptic predictors and average values for the predictand found during the development of STIPS.

Predictor	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
SHRD	15.5	15.2	14.9	14.6	14.5	14.3	14.2	14.1	14.1	14.1
T200	-51.0	-51.0	-51.0	-51.0	-51.0	-50.9	-50.9	-50.9	-50.9	-50.9
USHRD	1.2	1.0	0.9	0.7	0.8	1.0	1.0	1.1	1.1	1.1
RHHI	58.8	58.9	59.0	59.0	59.0	58.9	59.0	59.0	59.1	59.3
SPD	9.2	8.9	8.7	8.5	8.4	8.4	8.3	8.4	8.4	8.3
DELV	1.6	4.1	6.8	9.8	12.3	15.1	18.0	20.6	22.9	25.0

4. Operational model performance

The STIPS model has been integrated into the Automated Tropical Cyclone Forecast (ATCF) system (Sampson and Schrader 2000). The NOGAPS analyses and forecasts are supplied via a data feed with the Fleet Numerical Meteorological Operations Center. In the ATCF system, the Typhoon Duty Officer usually completes the track forecast before starting the intensity forecast. It is at the point when the track forecast is complete that STIPS is run. In this manner the STIPS model makes use of the current track forecast. Output from the STIPS forecast can be displayed graphically (as a graph of intensity versus forecast period), or as a one-page summary detailing forecast changes of the predictors as well as their individual contributions to the intensity forecast.

The current version of STIPS as described in the previous sections was implemented on the ATCF at JTWC on 16 June 2003. This section will present the verification of the STIPS model performance since that time. Verification is based on the homogeneous comparison of STIPS forecasts and ST5D (Knaff et al. 2003), with the final 2003 and preliminary 2004 best tracks from JTWC. To ensure enough cases, forecasts made in 2003 and 2004 are verified together. Also note that other models are not verified as they reduce the sample size. However, for the record, no other model outperformed STIPS during this time period.

There were 20 tropical cyclones (8° – 27° W) in 2003 that began after the current version of STIPS was implemented at JTWC and 28 tropical cyclones (1° – 28° W) that had occurred in 2004 by the time of this evaluation. Of the 48 storms verified here, 8 became supertyphoons (>129 kt), 32 became typhoons (>64 kt), and 42 became tropical storms (>34 kt). Additional

details of each of the storms used in the verification can be found at the Joint Typhoon Warning Center Web site (<http://www.npmoc.navy.mil/jtwc.html>).

A homogeneous comparison of performance for the STIPS (STIP), decay-STIPS (STID), no change (NCHG), and 5-day STIFOR (ST5D) forecasts is shown in Fig. 4. Mean absolute intensity errors are shown in Fig. 4a and the associated biases are shown in Fig. 4b. A common way to show the results of this comparison is in terms of the reduction in error relative to the control forecast (in this case ST5D) or CLIPER diagram. The CLIPER diagram associated with these forecasts is shown in Fig. 4c. Mean absolute errors of the verification compare well with those of the developmental dataset, with observed MAE being approximately 2–3 kt larger than the developmental MAE through the 120-h forecast period. Correcting for the effects of serial correlation (30 h assumed), the reduction in errors relative to ST5D is significant at the 95% level at 24, 36, 48, and 72 h for STID and at 36, 48, and 72 h for STIP. It is interesting to note that the use of the inland decay has a profound effect on the short-term forecasting skill, but at longer leads actually degrades the forecasts. This skill performance is similar to that found with the SHIPS model in the Atlantic (DeMaria et al. 2003) and suggests that beyond 48 h errors in the track forecast can begin to adversely affect intensity forecasts.

While it is desirable to examine the performance of STIP stratified by intensity, intensity trend, location, and time of year, this is difficult to do with 1.5 yr of independent forecasts. However, it is clear that the STIPS model is unable to predict the extreme intensification rates associated with storms that intensify at rates greater than 25 kt day^{-1} . The reasons are twofold:

TABLE 6. Developmental statistics associated with the STIPS model. Shown are percent variance explained (R^2) and mean absolute error of the model estimate (MAE).

Statistic	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
R^2 (%)	40.0	49.4	54.6	57.7	59.6	61.2	63.0	64.8	66.6	67.8
MAE (kt)	5.6	9.3	12.1	14.7	17.0	18.6	19.7	20.7	21.2	21.8

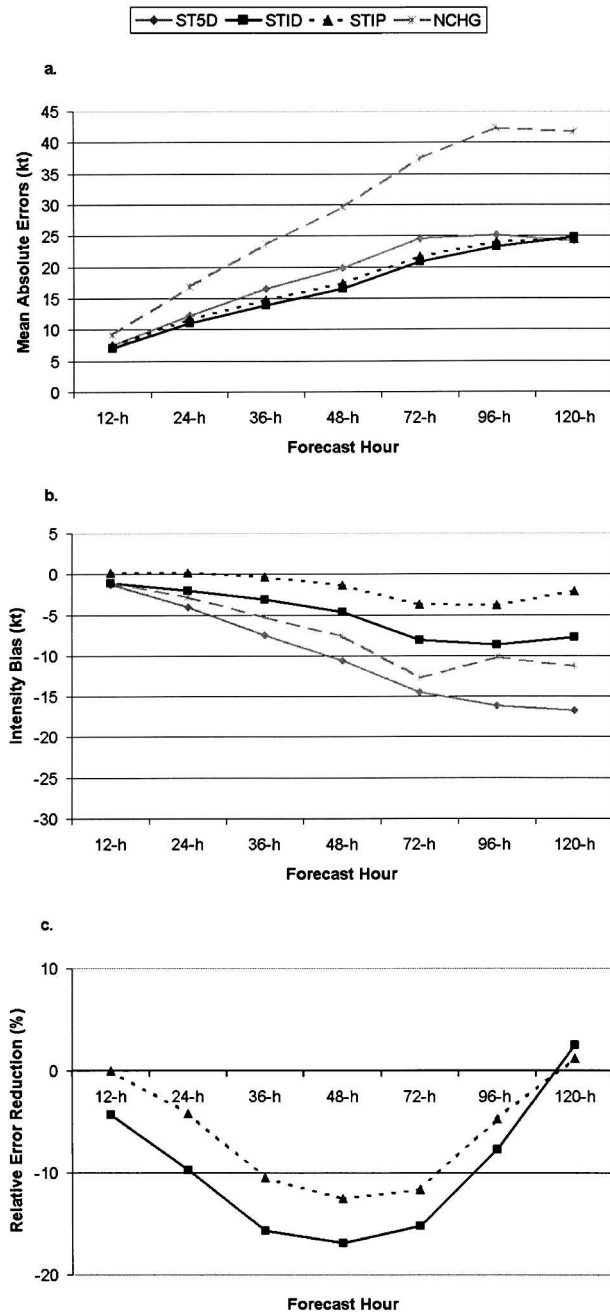


FIG. 4. Results of the independent and homogeneous verification of 48 tropical cyclones (8° – 27° W in 2003 and 1° – 28° W in 2004) for three models, namely STIPS (STIP), decay-STIPS (STID), and the baseline ST5D. Shown are (a) the mean absolute intensity errors and (b) the associated model biases, and (c) the percent reduction in error relative to ST5D. Also shown for reference in (a) and (b) are the errors and biases associated with a no-change forecast (NCHG). The number of cases is 985, 913, 837, 764, 618, 380, and 276 at 12, 24, 36, 48, 72, 96, and 120 h, respectively.

- 1) linear regression models do not predict extremes and
- 2) STIPS uses environmental conditions surrounding the storm rather than conditions associated with the near-core convective region to make its predictions.

5. Discussion of ongoing and future work

Despite the skill demonstrated by STIPS, the model possesses several shortcomings due to design and formulation shortcomings that may be overcome with future work. Among these is the use of the perfect prog assumption in combination with the use of annular area averaging of environmental predictors. STIPS could be further improved by determining the optimal averaging area for environmental predictors. As part of this work, the relative errors introduced by using the track JTWC forecast versus the use of the NOGAPS track forecast would also be determined.

Another shortcoming of the STIPS model is that it does not include any satellite-based predictors. There are several datasets that could be exploited to provide key information concerning the tropical cyclone and its environment. Examples include ocean heat content from satellite-based altimetry and the measurement of convective activity/trends associated with tropical cyclones via infrared and microwave sensors. Efforts are beginning to add satellite-based ocean heat content to the STIPS developmental data.

Potential improvement could also come from the use of consensus intensity forecasts, since track forecasting has benefited from this approach (Goerss et al. 2004). STIPS currently provides intensity forecasts based upon NOGAPS model fields and the JTWC track forecasts. Could other NWP models be used to create a consensus intensity forecast? Would the consensus forecast provide better guidance? Experimental efforts are under way at the Naval Research Laboratory, Monterey, California (NRLMRY), to evaluate the potential intensity prediction improvements that are possible using a consensus of several global models.

Potentially the most important improvement could be on the development of tools to help predict rapid intensity changes. The intensity changes associated with rapidly intensifying tropical cyclones are nearly impossible to predict using the current methodology. Techniques that include environmental along with storm-scale predictors should be developed to help with these situations. It is anticipated that these efforts will begin soon.

The relative success of the STIPS model in the western North Pacific tropical cyclone basin has also led to the development of similar capabilities in the north Indian Ocean (IO) and Southern Hemisphere (SH).

These basins have an even more limited number of tropical cyclone intensity guidance methods than in the western North Pacific. As part of this project, 5-day CLIPER models have also been developed for both of these regions, so that intensity forecast skill of all available methods can be assessed. The CLIPER models are currently running in JTWC operations and the SH and IO versions of STIPS are to be evaluated by JTWC and NRLMRY during 2005.

6. Summary

Prediction of tropical cyclone “intensity” or maximum sustained surface winds remains a difficult task in all tropical cyclone basins. While there has been relatively steady improvement over the years in track forecasting skill with ever-improving numerical models, the skill of intensity forecasts still lags that of track prediction. In recent years however, the rate of improvement in intensity forecast skill has exceeded that of track prediction in the Atlantic tropical cyclone basin. This has not been the case in the western North Pacific TC basin where, until 2002, there were no statistical–dynamical intensity forecast techniques available.

To help overcome the shortcomings associated with operational tropical cyclone intensity forecasting in the western North Pacific TC basin, a statistical–dynamical approach was employed to develop an intensity prediction scheme. This technique, the Statistical Typhoon Intensity Prediction Scheme or STIPS, makes use of 5.5 yr of NOGAPS analyses and tropical cyclone best tracks in a “perfect prog” developmental framework. STIPS has been implemented into operations at JTWC where it uses forecast tracks and forecast fields from NOGAPS to make intensity predictions. Despite errors introduced by using the perfect prog assumption, STIPS has been shown to produce skillful forecasts for most forecast times through 4 days.

The uses of statistical–dynamical methodologies in the Atlantic and east Pacific has corresponded to a trend of increasing skill associated with intensity forecasting in these basins. For the time being, it is likely that the continued use of STIPS, which is a skillful intensity prediction tool, will result in similar trends in the intensity forecasts issued by JTWC. However, even greater intensity forecast improvements could be realized in this basin if the shortcomings associated with STIPS were rectified.

Acknowledgments. The authors thank the Office of Naval Research for funding this project with Grant NA17RJ1228, Mr. Ed Fukada at the Joint Typhoon Warning Center for being supportive of this initiative,

and the two anonymous reviewers for their useful and constructive comments concerning this manuscript. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. government position, policy, or decision.

REFERENCES

- AMS, 2000: Policy statement—Hurricane research and forecasting (adopted by AMS council on 14 February 2000). *Bull. Amer. Meteor. Soc.*, **81**, 1341–1347.
- Baik, J.-J., and J.-S. Paek, 2000: A neural network model for prediction typhoon intensity. *J. Meteor. Soc. Japan*, **78**, 857–869.
- Bogner, P. B., G. M. Barnes, and J. L. Franklin, 2000: Conditional instability and shear for six hurricanes over the Atlantic Ocean. *Wea. Forecasting*, **15**, 192–207.
- Briegel, L. M., and W. M. Frank, 1997: Large-scale influences on tropical cyclogenesis in the western North Pacific. *Mon. Wea. Rev.*, **125**, 1397–1413.
- Chu, J.-H., 1994: A regression model for the western North Pacific tropical cyclone intensity forecasts. NRL Memo. Rep. 7541-94-7215, Naval Research Laboratory, 33 pp. [Available from Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.]
- DeMaria, M., 1996: The effect of vertical wind shear on tropical cyclone intensity change. *J. Atmos. Sci.*, **53**, 2076–2087.
- , and J. Kaplan, 1994: Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate*, **7**, 1324–1334.
- , and —, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337.
- , M. Mainelli, L. K. Shay, J. A. Knaff, and J. P. Kossin, 2003: Improvements in real-time statistical tropical cyclone intensity forecasts using satellite data. Preprints, *12th Conf. on Satellite Meteorology and Oceanography*, Long Beach, CA, Amer. Meteor. Soc., CD-ROM, JP1.4.
- Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, 47 pp. [Available from NOAA/NESDIS, 5200 Auth Rd., Washington, DC 20233.]
- Elsberry, R. L., 1987: Observation and analysis of tropical cyclones. *A Global View of Tropical Cyclones*, R. L. Elsberry, Ed., Naval Postgraduate School, 1–12.
- Emanuel, K. A., 1988: The maximum intensity of hurricanes. *J. Atmos. Sci.*, **45**, 1143–1155.
- , cited 2004: The Coupled Hurricane Intensity Prediction Scheme (CHIPS). [Available online at <ftp://texmex.mit.edu/pub/emanuel/Powerpoint/CHIPS.pdf>.]
- Fitzpatrick, P. J., 1997: Understanding and forecasting tropical cyclone intensity change with the Typhoon Intensity Prediction Scheme (TIPS). *Wea. Forecasting*, **12**, 826–845.
- Furze, P., and A. Preble, cited 2004: 2003 annual tropical cyclone report. [Available online at <https://metoc.npmoc.navy.mil/jtwc/atcr/2003atcr/>.]
- Goerss, J. S., and R. A. Jefferies, 1994: Assimilation of synthetic tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting*, **9**, 557–576.

- , C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638.
- Grandau, F., and G. Engel, cited 2004: 2001 annual tropical cyclone report. [Available online at <https://metoc.npmoc.navy.mil/jtwc/atcr/2001atcr/>.]
- Guard, C. P., L. E. Carr, F. H. Wells, R. A. Jefferies, N. D. Gural, and D. K. Edson, 1992: Joint Typhoon Warning Center and the challenges of multibasin tropical cyclone forecasting. *Wea. Forecasting*, **7**, 328–352.
- Hogan, T., and T. Rosmond, 1991: The description of the Navy Operational Global Atmospheric Predictions System's spectral forecast model. *Mon. Wea. Rev.*, **119**, 1786–1815.
- , M. S. Peng, J. A. Ridout, and W. M. Clune, 2002: A description of the impacts of changes to NOGAPS convection parameterization and the increase in resolution to T239L20. NRL Memo. Rep. 7530-02-52, 10 pp. [Available from Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.]
- Holland, G. J., 1997: The maximum potential intensity of tropical cyclones. *J. Atmos. Sci.*, **54**, 2519–2541.
- IMSL, 1987: *FORTRAN Subroutines for Statistical Analysis*. International Mathematical and Statistical FORTRAN Library, 1232 pp.
- JTWC, cited 2004, Tropical cyclone best track data site. [Available online at https://www.npmoc.navy.mil/products/jtwc/best_tracks/.]
- Kalnay, E. 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.
- 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.
- , and —, 2001: On the decay of tropical cyclone winds after landfall in the New England area. *J. Appl. Meteor.*, **40**, 1–12.
- Knaff, J. A., and C. W. Landsea, 1997: An El Niño–Southern Oscillation climatology and persistence (CLIPER) forecasting scheme. *Wea. Forecasting*, **12**, 633–652.
- , M. DeMaria, C. R. Sampson, and J. M. Gross, 2003: Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting*, **18**, 80–92.
- Kuma, K., 1996: NWP activities at Japan Meteorological Agency. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., J15–J16.
- Levitus, S., 1982: *Climatological Atlas of the World Oceans*. NOAA Prof. Paper 13, U.S. Government Printing Office, 173 pp.
- Malkus, J. S., and H. Riehl, 1960: On the dynamics and energy transformations in steady-state hurricanes. *Tellus*, **12**, 1–20.
- Merrill, R. T., 1987: An experiment in statistical prediction of tropical cyclone intensity change. NOAA Tech. Memo. NWX NHC-34, 33 pp.
- Miller, B. L., 1958: On the maximum intensity of hurricanes. *J. Meteor.*, **15**, 184–195.
- Panofsky, H. A., and G. W. Brier, 1968: *Some Applications of Statistics to Meteorology*. Earth and Mineral Sciences Continuing Education, College of Earth and Mineral Sciences, The Pennsylvania State University, 224 pp.
- Rennick, M. A., 1999: Performance of the navy's tropical cyclone prediction model in the western North Pacific basin during 1996. *Wea. Forecasting*, **14**, 297–305.
- Riehl, H., 1972: Intensity of recurved typhoons. *J. Appl. Meteor.*, **11**, 613–615.
- Sadler, J. C., 1976: A role of the tropical upper tropospheric trough in early season typhoon development. *Mon. Wea. Rev.*, **104**, 1266–1278.
- , 1978: Mid-season typhoon development and intensity changes and the tropical upper tropospheric trough. *Mon. Wea. Rev.*, **106**, 1137–1152.
- Sampson, C. R., and A. J. Schrader, 2000: The automated tropical cyclone forecasting system (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240.
- , R. J. Miller, R. A. Kreitner, and T. L. Tsui, 1990: Tropical cyclone track objective aids for the microcomputer: PCLM, XTRP, PCHP. Naval Oceanographic and Atmospheric Research Laboratory Tech. Note 61, 15 pp. [Available from Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.]
- Schade, L. R., 2000: Tropical cyclone intensity and sea surface temperature. *J. Atmos. Sci.*, **57**, 3122–3130.
- , and K. A. Emanuel, 1999: The ocean's effect on the intensity of tropical cyclones: Results from a simple coupled atmosphere–ocean model. *J. Atmos. Sci.*, **56**, 642–651.
- Weatherford, C. L., and W. M. Gray, 1988: Typhoon structure revealed by aircraft reconnaissance. Part II: Structural variability. *Mon. Wea. Rev.*, **116**, 1044–1056.
- Whitney, L. D., and J. S. Hobgood, 1997: The relationship between sea surface temperatures and maximum intensities of tropical cyclones in the eastern North Pacific Ocean. *J. Climate*, **10**, 2921–2930.
- Zehr, R. M., 1992: Tropical cyclogenesis in the western North Pacific. NOAA Tech. Rep. NESDIS 61, 181 pp.