Australian Meteorological and Oceanographic Journal 58 (2009) 9-18

Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part II: statistical – dynamical forecasts

John A. Knaff
NOAA/NESDIS, Center for Satellite Research and Applications, Fort Collins, Colorado, USA
and
Charles R. Sampson
Naval Research Laboratory, Monterey, California, USA

(Manuscript recieved June 2008;
Revised October 2008)

The development and performance of a statistical - dynamical tropical cyclone intensity forecast model, which was developed for the United States of America's Joint Typhoon Warning Center (JTWC), is described. This model, called the Southern Hemisphere Statistical Typhoon Intensity Prediction Scheme (SH STIPS), mirrors similar capabilities created for use in the western North Pacific and North Indian Ocean tropical cyclone basins. The model is created by fitting an optimal combination of factors related to climatology and persistence, intensification potential, vertical wind shear, dynamic size/intensity forecasts and atmospheric stability. All of these factors except the climatology and persistence information are derived from global forecast model analyses and forecasts. In July 2005 the SH STIPS model began a real-time evaluation period. The forecasts from the SH STIPS model have outperformed the combined climatology and persistence based forecast and thus are skillful in independent testing since that time. Since October 2006, SH STIPS has been the primary member in an operational consensus forecast of tropical cyclone intensity change provided to the JTWC. Documentation is provided for potential users of forecasts based on this methodology and for researchers interested in developing similar capabilities in the future.

Introduction

The United States of America's Joint Typhoon Warning Center (JTWC) makes tropical cyclone (TC) forecasts in the southern hemisphere. These forecasts are typically produced every 12 h, extend through 48 h, and consist of position, intensity and significant (e.g. gale-force, etc.) wind radii. These tactical forecasts support the United States of America's military and civilian operations in this part of the world. Until recently, few objective forecast aids for TC intensity existed and operational intensity forecasts heavily depended on trends in the satellite analysis. For more information about objective intensity guidance techniques used at JTWC and when they became available, a comprehensive list is provided in Table 1 of the Part I companion paper, Knaff and Sampson (2009).

Corresponding author address:
John Knaff, NOAA/NESDIS, CIRA, Colorado State University,
Campus Delivery 1375, Fort Collins, CO 80523-1375, USA
Email: John.Knaff@noaa.gov

In 2005, a TC intensity model based on a statistical – dynamical approach was developed for use in the southern hemisphere. The model was called the Southern Hemisphere Statistical Typhoon Intensity Prediction Scheme (or SH STIPS) after its counterpart used in the western North Pacific (STIPS; Knaff et al. 2005) and is the subject of this paper. The SH STIPS model was designed to make statistical forecasts of intensity using environmental forecast information from a global forecast model (the dynamical component) along the official JTWC tracks and static predictors provided by the operational best track. In addition to the actual forecast of intensity, the model output provides information about environmental conditions along the forecast track and potential influences of land by employing an inland decay model. The development of the SH STIPS model is nicely complemented by the purely statistical model based on climatology and persistence referred to as SH ST5D and described in Knaff and Sampson (2009). For this discussion, it is important to note that due to changes in JTWC operational capabilities, the SH STIPS output along the JTWC forecast track, which has been created in real-time since June 2005[1], was never available to JTWC for consideration in making their operational intensity forecasts. The SH STIPS model however is an important part of a consensus-based method to predict TC intensity using track and analysis/forecast fields from a number of models discussed in the Part III companion paper Sampson and Knaff (2009); again creating capabilities similar to those available in the western North Pacific (Sampson et al. 2008).

The SH STIPS model, because of its continued use in the consensus intensity forecast methods, needs to be formally documented. The following sections provide such documentation by giving details about the datasets and techniques used to develop the SH STIPS model. In addition, the dependent or expected performance is compared to an independent sample of forecasts made from July 2005 to July 2007. Documentation is provided for potential users[2] of forecasts based on this methodology as well as those seeking to create similar models.

Datasets

Seven years of Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan et al. 2002; Hogan and Rosmond 1991) analyses were used in the development of SH STIPS. Specifically, temperature, wind, water vapor pressure and geopotential height data were collected twice daily for the period 21 July 1997 through 30 July 2004 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 southern hemisphere. For operational and developmental purposes a climatological SST derived from the Reynolds and Smith (1995) SST is used when the NOGAPS skin temperature field is unavailable.

The tropical cyclone position and intensity information used for the development of this model came from the JTWC's "best track", which is a post-season reanalysis using additional information not available in the operational time frame (JTWC, cited 2008). These files contain the date, time, latitude, longitude and intensity every six hours for all storms designated by JTWC as being tropical depression[3] strength or greater. Because routine aircraft reconnaissance has never been available in this region, one caveat concerning the best track intensities is that they are determined solely from satellite-based methods (e.g., Dvorak 1984; Demuth et al. 2006, 2004; Olander et al. 2007; Velden et al. 1998) the majority of the time. The actual errors associated with the use of satellite intensity estimation methods have been quantified in Olander et al. (2007), Velden et al. (1998) and Demuth et al. (2006). Those results, which show all of the satellite techniques are capable of capturing intensification trends, give some confidence in the operational and best track intensity estimates. The intensity archived in these historical datasets, as well as operational intensity forecasts, are estimated to the nearest 5 knots (kn – nautical miles per hour) at 6 h intervals. For this reason, model formulation as well as any discussion of intensity in this paper will be in terms of kn instead of ms^{-1} (1 ms^{-1} = 1.94 kn).

Model development

The development of the SH STIPS model closely follows the development of the STIPS model in the western North Pacific tropical cyclone basins (Knaff et al. 2005), but incorporates some additional low-level thermodynamic predictors similar to those used by DeMaria et al. (2005). SH STIPS is a multiple linear regression model. The dependent variables (predictands) are the intensity changes from the initial forecast time (DELV) at 12 hour 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. The potential model predictors are then evaluated for their combined statistical ability to predict tropical cyclone intensity change. This process yields ten predictive equations that are used to make forecasts at each

[1]The SH STIPS model based on JTWC forecast tracks was never provided to JTWC due to latency issues, but SH STIPS has been provided as a STIPS ensemble and as part of a multi-model consensus as documented in Sampson and Knaff (2009).

[2]Forecasts from the operational intensity consensus (Sampson and Knaff 2009) are being provided to the Australian Bureau of Meteorology.

[3]Tropical Depression: A weak tropical cyclone with a definite closed surface circulation and highest sustained wind speeds (averaged over one minute or longer period) of less than 17 m s^{-1} (34 knots) (Elsberry 1987)

of the ten 12-hourly time periods, 12 h through 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 model discussed in DeMaria et al. (2006) is used. The coefficients used for inland decay come from Kaplan and DeMaria (1995) and Kaplan and DeMaria (2001) where they are used north of 36°S, and south of 40°S, respectively. Between these two latitudes, a linear weighting of two sets of coefficients is used.

The details concerning the development of the SH STIPS model are contained in the following subsections which, firstly, outline the predictors used in the model development, secondly, describe the statistical methodology and finally discuss how the final model predictors were selected along with the discussion of their relative importance.

Potential predictors

The potential predictors used in model development can be divided into two categories: (1) those related to climatology, persistence, current/past motion and trends of intensity – "static predictors" and (2) those related to current and future environmental and SST conditions – "time dependent predictors". All of the time dependent predictors are derived along the tropical cyclone track. The various predictors are developed using a "perfect prognosis" methodology (Kalnay 2003) where the analyses and actual tropical cyclone best tracks are used to create the statistical model. However, when SH STIPS is run in real-time, 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 SH STIPS intensity forecast errors not accounted for in the developmental data.

The potential static predictors are derived from the current date and intensity, and the 12 h change in intensity, motion, and location. Predictors determined during the development of climatology and persistence based models (Knaff et al. 2003; Knaff and Sampson 2009) were included as potential static predictors in STIPS. However, since SST information is used in the development of this model, predictors related to location (a proxy for SST in the SH ST5D CLIPER model) were not included in the static predictor pool. In addition the pressure corresponding to the steering level (PSLV) is also examined as a static predictor. The PSLV was estimated from the steering flow at all the analysis levels at 0 h. It has been found that a lower PSLV is a favourable condition for intensification in other basins (Knaff et al. 2005, DeMaria et al. 2005). A lower value of PSLV is likely to be associated with wind environments that are generally more uniform in the vertical, and thus more conducive to intensification. Static predictors used for model development are listed in Table 1.

Table 1. The potential static predictors used in SH STIPS development.

predictor	description
VMAX	Initial intensity
VMAX^2	Initial intensity squared
DVMX	12-hour change in intensity
JDAY	Absolute number of day before or after the 45th day of the year
SPD	Storm translational speed
PSLV	The pressure level of the layer mean flow that most closely approximates the initial steering motion

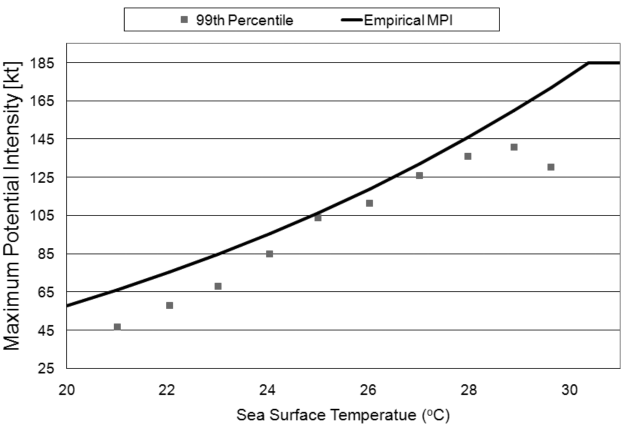
The potential time dependent predictors are more numerous and require more explanation. These predictors are divided into three basic categories, namely those related to the SST, those related to atmospheric stability and those related to the wind fields. SST values are determined at the storm centre by interpolating from oceanic NOGAPS skin temperature values, while atmospheric stability and wind related predictors are area averaged around the centre. 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 (e.g. Miller, 1958; Malkus and Reihl 1960; Emanuel 1988; Holland 1997) or empirically (e.g. Merrill 1987; DeMaria and Kaplan 1994; Whitney and Hobgood 1997). For the purposes of developing SH STIPS, the empirical approach is chosen following the methodology employed in DeMaria and Kaplan (1994), which fits an exponential function to the maximum observed tropical cyclone intensity with respect to SST. The SST used to develop this MPI function for the southern hemisphere is derived from a 21-year climatology of Reynolds SST (Reynolds and Smith 1995), which contained data from the period 1982-2002. This 1° latitude by 1° longitude resolution SST climatology is then interpolated to the storm centre following the best track for a 22-year period (1980-2002) to find the SST corresponding to the storm intensity. The storm intensity used is actually the best track intensity minus the storm speed following the convention used by DeMaria and Kaplan (1994), though a potentially more accurate method to remove the effect of storm translation on intensity is provided by Schwerdt et al. (1979). The maximum values in each half-degree temperature interval are then used to determine the coefficients in the MPI function described in Eqn 1. Using this procedure, the coefficients are A=–42.1 kn, B=220.58 kn, C=0.0792°C^{-1} and T_0=30.0°C.

$$MPI = A + Be^{C(T-T_0)} \qquad \ldots(1)$$

In the formulation of STIPS the maximum value of MPI allowed is 185 kn. Figure 1 shows this MPI function and its fit to the 99th percentile of the data.

Fig. 1 The empirical maximum potential intensity (MPI) function used in SH STIPS given by the solid line along with the binned 99th percentile of intensity observed in the JTWC best track dataset (1980-2000).



Atmospheric stability is known to affect tropical cyclone intensification and development. The effect of middle-level moisture (e.g. moist entropy) is subtle and is fundamentally related to the ventilation (as defined by Simpson and Reihl (1958)) of the storm (Emanuel et al. 2004). Variations of environmental relative humidity (RH) will affect convective buoyancy through entrainment of subsaturated air. In tropical cyclones, convective available potential energy is relatively small and decreases to almost zero near the centre (Bogner et al. 2000). Therefore, relative humidity values and thus moist entropy 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 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 centre of the cyclone. To further examine the combined effects of low-level moisture and temperature on the atmospheric stability the equivalent potential temperature at 925 hPa (E925) is also used as a potential predictor and is calculated in the same annulus. This annular average is used to estimate environmental parameters. The 200 km radius is used to remove the inner-most regions of the tropical cyclone from the analysis where the NOGAPS model uses synthetic observations to initialize the tropical cyclone (Goerss and Jefferies 1994). The size of the annulus was chosen to maximize the predictive ability of the vertical wind shear, and thus provides a good estimate for the average-sized tropical cyclone.

At 200 hPa, the zonal wind (U200), temperature (T200), divergence (D200) and 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 using Eqn 2,

$$REFC = -\frac{1}{r^2}\frac{\delta}{\delta r}\left(r^2\,\overline{U_L'V_L'}\right), \qquad \ldots(2)$$

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. The T200 is thought to help correct any shortcomings of using a climatological MPI that is solely related to SST conditions. U200 is an indicator of the direction of vertical wind shear, which has been shown to be important to tropical cyclone structure and intensification (Frank and Ritchie 2001). REFC is a measure of atmospheric torque and has been shown to be related to TC intensity change (DeMaria et al. 1993).

In addition to these potential predictors at 200 hPa, the 850 hPa vorticity (Z850) 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 hPa minus 850 hPa wind difference (SHRD) and the 500 hPa minus 850 hPa wind difference (SHRS). In addition to the scalar values of shear (i.e. SHRD and SHRS), the zonal components (USHRD, USHRS) of the shear in these layers were also created. 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 (SHRG) 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 four 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, as shown in Eqn 3.

$$SHRG = 4.0 * \sum_{p=850}^{p=200} w_p \sqrt{(u_p-\bar{u})^2 + (v_p-\bar{v})^2} \qquad \ldots(3)$$

where

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

$\overline{v} = \sum_{p=850}^{p=200} w_p v$ 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 the initial intensity VMAX, were added following the notion that these terms may account for some inherent non-linearity in the same way they do in STIPS (Knaff et al. 2005), SHIPS (DeMaria and Kaplan 1999) and in ST5D (Knaff et al. 2003; Knaff and Sampson 2009). The terms SHRG times the cosine of the latitude (along the storm track) and SHRD times the cosine of the latitude (along the storm track) were also tested because storms at higher latitude tended to be less sensitive to vertical wind shear (DeMaria 1996). Since the potential effect of vertical wind shear is somewhat determined by the current storm intensity, a quadratic term combining these effects is also tested (VXSH). This results in 13 synoptic predictors being available for testing in SH STIPS, as listed in Table 2.

Statistical methodology

When developing a multiple regression model one commonly uses a method to select predictors based 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 and Wilks 2006). Significance of individual predictors is based on a standard F-test (Panofsky and Brier 1968). A 99 per cent 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 per cent 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 that are listed in Table 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 hours, are discussed.

Table 2. Potential synoptic predictors available for predictor selection for the SH STIPS model.

Predictor	Description
MPI	Maximum potential intensity based upon Eqn 1
MPI²	MPI squared
MPI*VMAX	MPI times the initial intensity
RHLO	Area-averaged (200 km to 800 km) relative humidity 850 – 700 hPa
RHHI	Area-averaged (200 km to 800 km) relative humidity 500 – 300 hPa
E925	Area-averaged (200 km to 800 km) equivalent potential temperature at 925 hPa
U200	Area-averaged (200 km to 800 km) zonal wind at 200 hPa
T200	Area-averaged (200 km to 800 km) temperature at 200 hPa
D200	Area-averaged (0 km to 1000 km) 200 hPa divergence
REFC	Relative eddy flux convergence within 600 km (see Eq. 2)
SHRG	Generalized 200 to 850 hPa vertical wind shear (see Eq. 3)
SHRS	Area-averaged (200 km to 800 km) 500 hPa to 850 hPa wind shear
SHRD	Area-averaged (200 km to 800 km) 200 hPa to 850 hPa wind shear

STIPS model formulation

The stepwise predictor selection procedure was performed on the two predictor pools and resulted in thirteen predictors being selected for use in model formulation (Tables 3 and 4). There were 2181 cases available at 12 h and 860 cases at 120 h in the developmental dataset. The thirteen 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 using the generalized shear parameterization (SHRG), as was also the case in the western North Pacific STIPS (Knaff et al. 2005). The predictor selection procedure also found that the REFC term is not significant at any time period, in agreement with Fitzpatrick (1997) and Knaff et al. (2005). Storm translation speed (SPD) also was found to be an unimportant factor, but the pressure of the steering level (PSLV) was found to be important. The potential predictors D200, RHLO and all the 500 - 850 hPa shear measures were also found to be statistically unimportant at all forecast times. Interesting is the inclusion of RHHI (i.e. relative humidity in the 500-300 hPa layer), but not RHLO (i.e. relative humidity in the 850 – 700 hPa layer), which is identical to the relationships used in the 2002 version of the SHIPS model (DeMaria et al. 2003) and the western North

Pacific STIPS model (Knaff et al. 2005).

Table 3 lists the forecast time at which the thirteen 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 the pressure of the steering level (PSLV). Most of the factors related to vertical wind shear (except VXSH), T200, RHHI and MPI, become most important at lead times beyond 24 hours.

Table 3. The final predictors used in STIPS along with the forecast hour when they are most statistically significant. The first three predictors are static while the others (4-13) are time dependent.

Predictor	Most important forecast hour
1. DVMAX	12
2. PSLV	72
3. VMAX^2	12
4. MPI	36
5. MPI^2	48
6. MPI*VMAX	12
7. SHRD	72
8. USHRD	24
9. T200	120
10. RHHI	36
11. Z850	36
12. E925	108
13. VXSH	12

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: 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 (what is being predicted), are normalized before they are incorporated in 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 hours. 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 per cent statistical significance of each normalized coefficient is indicated by bold face italics.

The thirteen predictors in Table 4 can be grouped by effect into those related to persistence (i.e. DVMX), the dynamic prediction of intensification and growth (i.e. Z850), the vertical wind shear, intensification potential and thermodynamic effects. The predictors related to vertical wind shear are PSLV, SHRD, USHRD and VXSH, which when grouped together show that SHRD and PSLV are inversely related to intensification, whereas USHRD is preferred for intensification. This final result, that increasing USHRD is related to in-

Table 4. Normalized regression coefficients used in the STIPS model. The predictors are listed at 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 per cent statistical significance level from an F-test is indicated by bold italic print.

	12 h (2181)	24 h (1995)	36 h (1825)	48 h (1661)	60 h (1502)	72 h (1354)	84 h (1215)	96 h (1088)	108 h (970)	120 h (860)
1. DVMAX	***0.35***	***0.28***	***0.20***	***0.15***	***0.10***	***0.07***	0.04	0.02	0.01	0.00
2. PSLV	***–0.06***	***–0.07***	***–0.08***	***–0.09***	***–0.10***	***–0.11***	***–0.11***	***–0.09***	***–0.10***	***–0.09***
3. VMAX^2	***–0.67***	***–0.81***	***–0.79***	***–0.76***	***–0.71***	***–0.69***	***–0.67***	***–0.65***	***–0.60***	***–0.57***
4. MPI	–0.21	***–0.35***	***–0.41***	–0.45	–0.34	–0.19	–0.19	–0.04	0.29	0.40
5. MPI^2	0.17	0.33	***0.43***	***0.54***	0.48	0.38	0.43	0.30	0.00	–0.08
6. MPI*VMAX	***0.72***	***0.79***	***0.67***	***0.51***	***0.33***	0.18	0.08	–0.02	–0.12	–0.17
7. SHRD	***0.11***	0.08	0.00	***–0.09***	***–0.16***	***–0.24***	***–0.27***	***–0.30***	***–0.31***	***–0.29***
8. USHRD	***0.10***	***0.14***	***0.17***	***0.16***	***0.14***	***0.12***	***0.09***	0.05	0.01	–0.02
9. T200	–0.03	–0.03	–0.03	–0.03	–0.03	–0.04	–0.04	–0.05	***–0.06***	***–0.08***
10. RHHI	***0.11***	***0.10***	***0.11***	***0.10***	***0.10***	***0.08***	***0.07***	0.06	0.04	0.04
11. Z850	***0.07***	***0.10***	***0.13***	***0.14***	***0.14***	***0.15***	***0.16***	***0.15***	***0.14***	***0.13***
12. E925	0.04	***0.07***	***0.08***	***0.09***	***0.10***	***0.11***	***0.12***	***0.12***	***0.13***	***0.14***
13. VXSH	***–0.46***	***–0.52***	***–0.46***	***–0.34***	***–0.19***	–0.01	***0.11***	***0.22***	***0.32***	***0.36***

tensification, may be related to intensification trends during extratropical transition or the propensity for South Pacific TCs to intensify post recurvature (Knaff 2009). The predictors VMAX2, MPI, MPI2, and MPI*VMAX can be thought of as a potential for intensification and when grouped together show that weaker storms with larger MPI values can intensify at greater rates, which is similar to results presented in Knaff et al. (2005). Finally, T200, E925, and RHHI form the basis of a correction to the MPI based on climatological SST, where T200 provides corrections related to variations in atmospheric temperature profiles, and E925 and RHHI give corrections related to the variability of moisture.

Model performance

Model performance can be demonstrated from both the developmental dataset - which can be thought of as the model's predictability limit - and by verification against an independent dataset. Table 5 shows the statistics related to the model developments. As is the case for other statistical models that predict the intensity change from the initial time or DELV (e.g. DeMaria et al. 2005; Knaff et al. 2005), the variance explained increases with increasing lead time and the model errors saturate near five days. Nonetheless, a substantial amount of the variance is explained even at shorter lead times and the dependent errors are smaller than those produced by the baseline model SH ST5D (Knaff and Sampson 2009). Both of these measures indicate that this model should be skillful in independent testing.

To verify this assertion, the real-time SH STIPS forecasts were verified for the period July 2005 to June 2008 using final best track intensities, noting that preliminary best track intensities are used for verification of storms occurring after July 2007[4]. This results in 321, 282, 243, and 210 cases for the 12, 24, 36, and 48 h forecast times, respectively. Figure 2(a) shows the forecast verification results in terms of mean absolute error (MAE) through to 48 h. Real-time forecasts beyond 48 h are rare because SH STIPS is run off of the JTWC forecast track, which is typically run through to 48 h. Only twelve forecasts exist for the 72 h forecast period. For comparison, forecasts from persistence (PER), a constant 65 kn forecast (CLIM) and from the SH ST5D (Knaff and Sampson 2009) model are also shown. Figure 2(b) shows the biases associated with these forecasts and Fig. 2(c) shows the percent improvement of SH STIPS achieved with respect to SH ST5D – all results are significant at the 95 per cent level. When a 30 h serial correlation reduces the number of degrees of freedom, these improvements are statistically significant at the 99, 97, 90, and 87 percent level for the 12, 24, 36, and 48 h forecast times, respectively. Biases for this sample suggest SH STIPS had a tendency to over forecast intensification which mirrors the biases associated with SH ST5D, suggesting that the independent data may deviate from climatology. However, in this analysis the model is skillful through the 48 h forecast.

While Fig. 2 gives the independent statistical performance of the SH STIPS model, it is important to supplement those findings with subjective, but potentially more important information for potential users. From the statistical analysis of errors it is clear that, even with this model, intensity forecasts remain challenging. There are a few important observations that have been made pertaining to this and similar statistical-dynamical models. First, rapid intensification events will not be predicted by this type of model, which by design predicts the most likely outcome –not the outliers. Secondly, because of the time averaging employed in the SH STIPS model, large fluctuations in forecast intensity are less likely as the forecast lead time increases. The resulting inability of the model to forecast large changes at longer lead times can lead to particularly poor forecasts when storms are forecast to rapidly weaken at short lead times and then encounter more favorable conditions at longer lead times. For instance, this sequence of events can occur when TCs make landfall on a large island (e.g. Madagascar) or peninsula (e.g. Cape York Peninsula) in the short term and then re-emerge into favourable marine conditions at longer lead times. Finally, these types of models, even with perfect developmental datasets, are ultimately dependent on reliable initial conditions and quality forecasts of environmental conditions and future track.

The 12-hourly intensity change (DELV) and current intensity (VMAX) are relatively important to the SH STIPS model as shown in Table 4, so errors in intensity estimate are important – and intensity estimates are currently all satellite based in the southern hemisphere. Forecasts of environmental conditions are likely to continue to improve, but are still problematic in intensity forecasting at longer lead times. Since the forecast track determines the timing of encounters with land, cooler SSTs, vertical wind shear etc., poor forecast tracks can also significantly degrade model performance. Some of the track based errors can and are being mitigated by the use of multi-track STIPS-based consensus forecasts (Sampson et al. 2008; Sampson and Knaff 2009). Ultimately, while the statistical-dynamic technique can be improved by higher quality forecast tracks, initial conditions, forecasts of environmental conditions and by the addition of predictors related to initial convective organization and oceanic heat content, it is likely that such methods are ultimately limited by the lack of detailed information about the tropical cyclone wind and thermal structure and the vigour and organization of convection. These later factors are more likely to be captured by the use of advanced data assimilation techniques and numerical modelling. For now, however, statistical-dynamical models remain viable tools for making operational

[4]Preliminary best tracks are an intermediate stage of the best tracking procedures. In some cases changes to the track and intensity estimates in the best track are made after the operational forecast is released (e.g. based on updated information), but before the best tracks are finalised.

Table 5. Developmental statistics associated with the SH STIPS model. Shown are percent variance explained (R^2), and mean absolute error of the model estimate (MAE).

	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
R^2 (%)	41.1	49.5	53.3	55.0	56.5	57.8	59.2	59.0	59.4	58.4
MAE (kn)	5.3	9.0	12.1	15.0	17.3	18.9	20.0	20.8	21.1	21.7

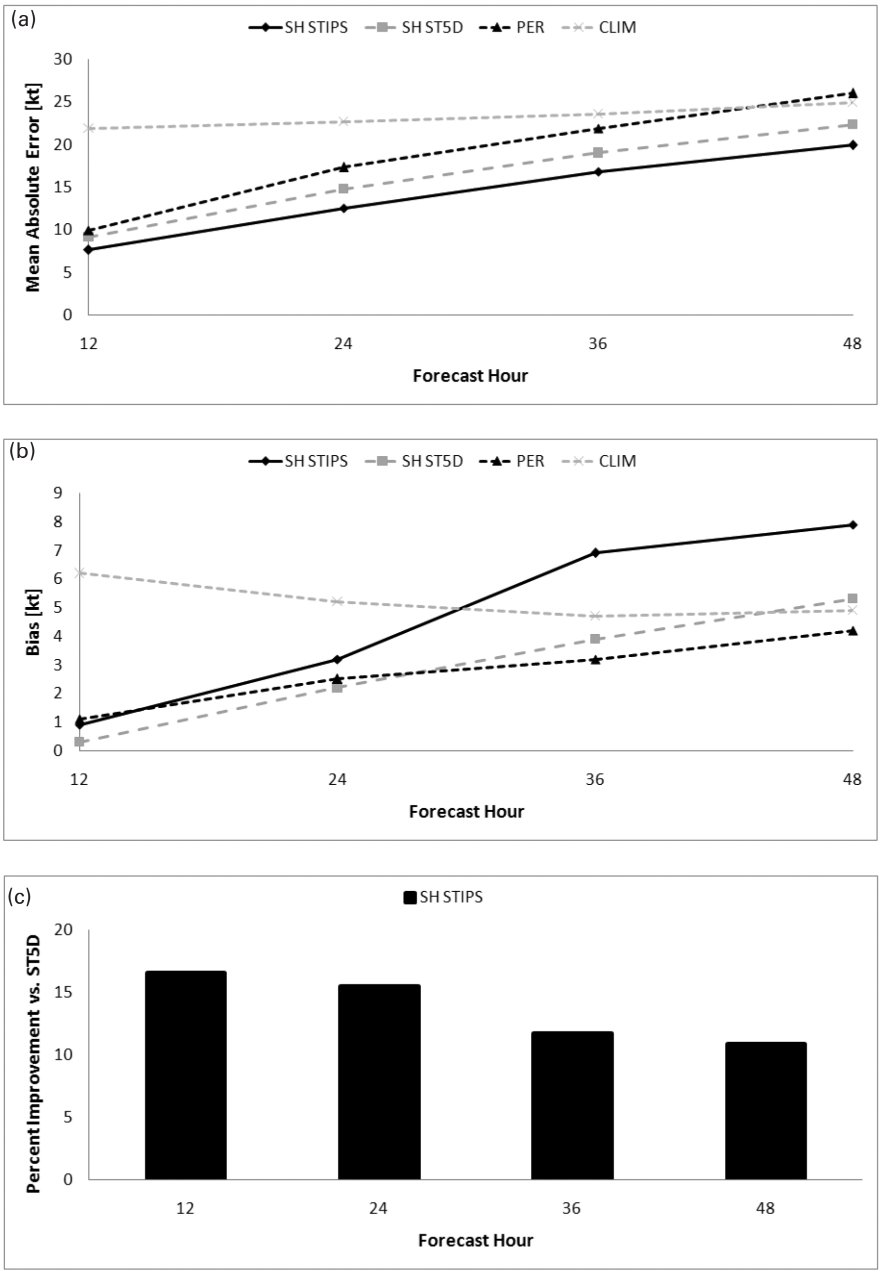


Fig. 2 The verification statistics for SH STIPS from July 2005 to the present. Mean absolute errors are shown along with similar statistics for the SH ST5D model, persistence (PER), and climatology (CLIM) in the top panel (a) and biases associated with SH STIPS, SH ST5D, PER and CLIM are shown in the central panel (b). The bottom panel (c) shows the percent improvement in the SH STIPS model when compared to climatology and persistence (i.e. SH ST5D).

tropical cyclone intensity forecasts in the southern hemisphere and other basins as shown here and in DeMaria et al. (2007, see their Table 3).

Summary and future plans

The development of a statistical-dynamical model for forecasting TC intensity change through five days in the southern hemisphere (SH STIPS) for use at the Joint Typhoon Warning Center has been documented. The model makes use of an optimal combination of factors related to climatology and persistence, intensification potential, vertical wind shear, dynamic size/intensity forecasts and atmospheric stability. SH STIPS is based on a multiple linear regression equation for each forecast time and forecasts the change in intensity from the initial value. The model was developed to mirror similar capabilities available to JTWC forecasters in the western North Pacific and Indian Ocean. SH STIPS is an improvement over other individual model intensity guidance methods in this basin. The statistics from both the dependent developmental data and from independent verification during July 2005 to June 2008 indicate that the model provides forecasts superior to combined climatology and persistence (SH ST5D). These performance statistics suggest that the SH STIPS model is one of the better models available for making tropical cyclone intensity forecasts in the southern hemisphere.

More important to the operational meteorologist, this model has become a pivotal part of an intensity consensus forecasting method described in a companion paper (Sampson and Knaff 2009), which has been able to improve the reliability, forecast length and skill of southern hemisphere intensity forecasts at the JTWC. Consensus-based intensity forecasts are also available to the Australian Bureau of Meteorology in real-time. In the near future, there are also plans to incorporate oceanic heat content information into the SH STIPS model via the method described in Goni et al. (1996) and created from the fields available from the Navy coupled ocean data assimilation (Cummings 2005). Testing of this formulation of the model in a consensus-based approach is planned to start during the 2009 tropical cyclone season.

Acknowledgments

This research was supported by ONR and NOAA (NA17RJ1228) grants. The authors would like to thank the reviewers Andrew Burton and Noel Davidson for their constructive comments. 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

Bogner, P. B., Barnes, G. M. and Franklin, J. L. 2000. Conditional instability and shear for six hurricanes over the Atlantic Ocean. *Wea. Forecasting, 15,* 192-207.

Cummings, J. A. 2005. Operational multivariate ocean data assimilation. *Quart. J. Roy. Meteor. Soc., 131,* 3583–604.

DeMaria, M. 1996. The effect of vertical wind shear on tropical cyclone intensity change. *J. Atmos. Sci., 53,* 2076-87.

DeMaria, M. and Kaplan, J. 1994. Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate, 7,* 1324-34.

DeMaria, M. and Kaplan, J. 1999. An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting, 14,* 326-37.

DeMaria, M., Knaff, J. A. and Sampson, C. R. 2007. Evaluation of long-term trends in tropical cyclone intensity forecasts. *Meteor. Atmos Phys., 97,* 19-28.

DeMaria, M., Knaff, J. A. and Kaplan, J. 2006. On the decay of tropical cyclone winds crossing narrow landmasses. *J. App. Meteor. and Clim., 45,* 491–9.

DeMaria, M., Baik, J. J. and Kaplan, J. 1993. Upper level eddy angular momentum fluxes and tropical cyclone intensity change. *J. Atmos. Sci., 50,* 1133-47.

DeMaria, M., Mainelli, M., Shay, L.K., Knaff, J.A. and Kossin, J.P. 2003. Improvements in real-time statistical tropical cyclone intensity forecasts using satellite data. *12th Conference on Satellite Meteorology and Oceanography,* 10-14 February.

DeMaria, M., Mainelli, M., Shay, L. K., Knaff, J. A. and Kaplan, J. 2005. Further Improvement to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. and Forecasting, 20,* 531-43.

Demuth, J., DeMaria, M. and Knaff, J. A. 2006. Improvement of advanced microwave sounding unit tropical cyclone intensity and size estimation algorithms, *J. App. Meteor. and Clim., 45,* 1573–81.

Demuth, J., DeMaria, M., Knaff, J. A. and Vonder Haar, T. H. 2004. Validation of an advanced microwave sounder unit (AMSU) tropical cyclone intensity and size estimation algorithm. *J. App. Meteor. and Clim., 43,* 282-96.

Dvorak, V. 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.]

Emanuel, K. A. 1988. The maximum intensity of hurricanes. *J. Atmos. Sci., 45,* 1143-55.

Emanuel, K. A., DesAutels, C., Holloway, C. and Korty R. 2004. Environmental control of tropical cyclone intensity. *J. Atmos. Sci., 61,* 843-58.

Elsberry, R. L. 1987. Observation and analysis of tropical cyclones, *A Global View of Tropical Cyclones,* R. L. Elsberry, Ed., Naval Postgraduate School, pp 1-12.

Fitzpatrick, P. J. 1997. Understanding and forecasting tropical cyclone intensity change with the Typhoon Intensity Prediction Scheme (TIPS). *Wea. Forecasting, 12,* 826-45.

Frank, W. M. and Ritchie, E. A. 2001. Effects of vertical wind shear on the intensity and structure of numerically simulated hurricanes. *Mon. Wea. Rev., 129,* 2249–69.

Goerss, J. S. and Jefferies, R. A. 1994. Assimilation of synthetic tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting, 9,* 557-76.

Goni, G., Kamholz, S. Garzoli, S. and Olson, D. 1996. Dynamics of the Brazil-Malvinas Confluence based on inverted echo sounders and altimetry, *J. Geophys. Res., 101,* 16273-89.

Hogan, T. and Rosmond, T. 1991. The description of the Navy Operational Global Atmospheric Predictions System's spectral forecast model. *Mon. Wea. Rev., 119,* 1786-815.

Hogan, T., Peng, M. S., Ridout, J. A. and Clune, W. M. 2002. A description of the impacts of changes to NOGAPS convection parameterization and the increase in resolution to T239L20. *NRL Memorandum Report 7530-02-52,* pp 10. [Available from Naval Research Laboratory, 7 Grace Hopper Avenue, Monterey, CA 93943-5502]

Holland, G. J. 1997. The maximum potential intensity of tropical cyclones. *J. Atmos. Sci., 54,* 2519-41.

IMSL 1987. FORTRAN subroutines for statistical analysis. International Mathematical and Statistical FORTRAN library, 1232 pp.

JTWC cited 2008. Tropical Cyclone Best Track Data Site. [Available online from https://metocph.nmci.navy.mil/jtwc/best_tracks/]

Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, ISBN 0-521-79629-6, 341 pp.

Kaplan, J. and DeMaria, M. 1995. A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *J. App. Met., 34,* 2499-512.

Kaplan, J. and DeMaria, M. 2001. On the decay of tropical cyclone winds after landfall in the New England area. *J. App. Met., 40,* 1-12.

Knaff, J.A. 2009. Revisiting the maximum intensity of recurving tropical cyclones. *Int. J. Climatology, 29,* 827-837

Knaff, J. A. and Sampson, C. R. 2009. Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part I: control forecasts based on climatology and persistence. *Aust. Met. Oceanogr. J., 58,* 1-7

Knaff, J. A., DeMaria, M., Sampson, C. R. and Gross, J. M. 2003. Statistical, five-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting, 18,* 80-92.

Knaff, J. A., Sampson, C. R. and DeMaria, M. 2005. An operational statistical typhoon intensity prediction scheme for the Western North Pacific. *Wea. Forecasting, 20,* 688-99.

Malkus, J. S. and Riehl, H. 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. I. 1958. On the maximum intensity of hurricanes. *J. Meteor., 15,* 184-95.

Olander, T.L. and Velden, C. S. 2007. The Advanced Dvorak Technique: continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. *Wea. Forecasting, 22,* 287-298.

Panofsky, H. A. and Brier, G. W. 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.

Reynolds, R.W. and Smith, T. M. 1995. A high-resolution global sea surface temperature climatology. *J. Climate, 8,* 1571-83.

Sampson, C. R. and Knaff, J. A. 2009. Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning

Center, Part III: forecasts based on a multi-model consensus approach. *Aust. Met. Oceanogr. J., 58,* 19-27

Sampson, C. R., Franklin, J. L., Knaff, J. A. and DeMaria, M. 2008. Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting, 23,* 304–12.

Schwerdt, R. W., Ho, F. P. and Watkins, R. R. 1979. Meteorological criteria for standard project hurricane and probable maximum hurricane wind fields, Gulf and East Coasts of the United States. *NOAA Tech. Rep. NWS 23,* 317pp. [Available from National Hurricane Center, Library, 11691 SW 117 St, Miami 33165-2149]

Simpson, R. H. and Riehl, H. 1958. Mid-tropospheric ventilation as a constraint on hurricane development and maintenance. *Proc. Tech. Conf. on Hurricanes, Miami Beach, FL, Amer. Meteor. Soc.,* D4-1–D4-10.

Velden, C.S., Olander, T. L. and Zehr, R. M. 1998. Development of an objective scheme to estimate tropical cyclone intensity from digital geostationary satellite infrared imagery. *Wea. Forecasting, 13,* 172-86.

Whitney, L., D. and Hobgood, J. S. 1997. The relationship between sea surface temperatures and maximum intensities of tropical cyclones in the eastern North Pacific Ocean. *J. Climate, 10,* 2921-30.

Wilks, D.S. 2006. *Statistical Methods in the Atmospheric Sciences,* 2nd Ed. International Geophysics Series, Vol. 59, Academic Press, 627 pp.