An Operational Rapid Intensification Prediction Aid for the Western North Pacific

JOHN A. KNAFF

NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado

CHARLES R. SAMPSON

Naval Research Laboratory, Monterey, California

KATE D. MUSGRAVE

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

ABSTRACT

This work describes tropical cyclone rapid intensification forecast aids designed for the western North Pacific tropical cyclone basin and for use at the Joint Typhoon Warning Center. Two statistical methods, linear discriminant analysis and logistic regression, are used to create probabilistic forecasts for seven intensification thresholds including 25-, 30-, 35-, and 40-kt changes in 24 h, 45- and 55-kt in 36 h, and 70-kt in 48 h ($1 \, \mathrm{kt} = 0.514 \, \mathrm{m \, s}^{-1}$). These forecast probabilities are further used to create an equally weighted probability consensus that is then used to trigger deterministic forecasts equal to the intensification thresholds once the probability in the consensus reaches 40%. These deterministic forecasts are incorporated into an operational intensity consensus forecast as additional members, resulting in an improved intensity consensus for these important and difficult to predict cases. Development of these methods is based on the 2000–15 typhoon seasons, and independent performance is assessed using the 2016 and 2017 typhoon seasons. In many cases, the probabilities have skill relative to climatology and adding the rapid intensification deterministic aids to the operational intensity consensus significantly reduces the negative forecast biases.

1. Introduction

Forecasters at the Joint Typhoon Warning Center (JTWC) have to predict the future movement, intensity, and the extent of the wind field for tropical cyclones (TCs). To aid in the forecast process, progress has been made in developing forecast guidance tools. As guidance tools improve, so do the TC forecasts produced by operational centers. Illustrating this point, JTWC's operational track forecasts have steadily improved since the 1990s due primarily to the improvement in global (and regional) numerical weather prediction (NWP) models (e.g., Elliot et al. 2014). More recently, operational intensity forecasts have shown slight improvements due to increasingly skillful guidance and to the use of simple consensus forecasts (Sampson et al. 2008; DeMaria et al. 2007; DeMaria et al. 2014). Based on Atlantic basin validations, gale-force—a wind exceeding

34 kt ($1 \, \mathrm{kt} = 0.514 \, \mathrm{m \, s}^{-1}$)—wind field guidance now seems to be able to provide skillful information (Sampson and Knaff 2015) leading to improved operational forecasts that are currently generated out to day 5 in the western North Pacific. While track, intensity, and wind-field forecasts have improved, operational guidance for unanticipated rapid TC intensification is essentially nonexistent in the western North Pacific TC basin.

Many definitions of rapid intensification (RI), also referred to as rapid development/deepening, exist for the western North Pacific. Brand (1973) defined RI as a 50-kt change in intensity in 24 h. Holliday (1977) and Holliday and Thompson (1979) define RI as a 42-hPa drop in central pressure in a day. Looking at these definitions from the perspective of the Atkinson and Holliday (1977) wind–pressure relationship (WPR) and the Dvorak (1984) intensity estimation method, these techniques share similarities. A central pressure of 985 hPa—the higher range of onset central pressures

Corresponding author: John Knaff, john.knaff@noaa.gov

corresponds to the 19th percentile in Holliday (1977)—corresponds to ~55 kt in the Dvorak method (Dvorak 1984). Subtracting 42 hPa from that central pressure value (985 hPa) roughly corresponds to 45 kt of intensity change with the Atkinson and Holliday WPR. These studies focused on the most rapidly intensifying systems, or about 25% of the storms that initially had central pressures ranging from 995 to 950 hPa (Holliday 1977) that correspond to ~35 kt of change or an intensity of 115 kt, respectively, using the operational WPR employed at the time.

While central pressures were considered more accurate in 1977, maximum wind speed estimates—the primary intensity metric in today's historical records of TC intensity—have uncertainties around 8 kt (Torn and Snyder 2012), and the Dvorak method has been shown to have RMSEs ranging from 8 to 12 kt (Knaff et al. 2010). Since wind speeds are the basis for intensity verification, more recent RI work uses definitions based on wind speed changes (Kaplan and DeMaria 2003; Kaplan et al. 2010, 2015; Shu et al. 2012). We will use the change in maximum wind speeds for our RI definitions, as has been done in more recent papers.

This study develops intensification guidance designed to anticipate TC rapid intensification in the western North Pacific, a tool recently transitioned into operational use at JTWC. The work uses a statistical–dynamical methodology, wherein application prognostic predictors from dynamic forecast models are used to make statistical forecasts, and uses a "perfect prog" development approach, where analyses of the prognostic data, considered perfect, are used for statistical inference (Neumann and Lawrence 1975). Section 2 describes the data and methods used to create the guidance, and section 3 describes the details of the statistical–dynamical models used and how those models provide the information to make deterministic forecasts. Section 4 shows results based on two years of independent data, and section 5 summarizes and provides comments on the utility of our algorithm and some ideas for improvement.

2. Data and methods

a. Independent and dependent variables

The intensity records from JTWC's best tracks provide the TC location and intensity based on a postseason reanalysis of all the data available and current operational practices. These data are in the native Automated Tropical Cyclone Forecast (ATCF) system format (Sampson and Schrader 2000; available at http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/). The files contain 6-hourly position, intensity, and wind

radii information for each storm reaching tropical depression status in JTWC's warnings. In these files, the native units—operational units—are expressed in knots for intensities (i.e., 1-min maximum sustained winds). For consistency with operational practices, this study uses knots throughout.

Table 1 contains the forecast parameters that are considered during the development of the rapid intensity change guidance. The parameters are grouped into three categories: 1) a subset of the environmental condition parameters available in the RAMMB (2017) developmental dataset, 2) storm-centered infrared (IR) imagery–based initial conditions, and 3) real-time, best-track parameters. This limited set of predictors was chosen based on past research on RI forecasting, the authors' experience, and the many years of forecaster-derived insight on this topic.

Environmental condition parameters (top section in Table 1) used in this application come from the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1999) developmental dataset (see RAMMB 2017). The developmental data consisted of western North Pacific TCs during the years 2000–15. Most of the environmental predictors [e.g., relative humidity in a 200–800-km annulus around the TC (RHMD), 200-hPa divergence (DIVC), and oceanic heat content between the surface and the depth of the 26°C isotherm (OHC)] require little explanation. However, a brief description and/or justification of some of the less straightforward environmental predictors is provided. The 850–200-hPa layer shear (the vector difference between 200 and 850 hPa) is the traditional measure of vertical wind shear, but it is probably less reliable in complicated vertical wind profiles. Furthermore, shear at lower levels is better correlated with intensification/weakening (e.g., Velden and Sears 2014; Wang et al. 2015), and the depth and height of wind shear can vary with individual situations (Finocchio and Majumdar 2017), so we want to capture as much of the predictive signal of shear as possible in a single parameter. We satisfy this objective by relying on the generalized vertical wind shear parameter (GSHR). The GSHR is the mass-weighted, root-mean-square deviations of the winds from 4 times the mass-weighted, deep-layer mean winds. The factor of 4 is used to make the values comparable to the more conventional measure of 200–850 hPa and is equal to that scalar difference for the case when shear is constant with respect to pressure. We also want to investigate a measure of upper-level eddy fluxes to ascertain TC interactions with troughs. This is done using the relative eddy flux convergence (REFC; see Knaff et al. 2005). Finally, low-level temperature advection (TADV) can potentially be

TABLE 1. Potential predictors for algorithms used to predict the probabilities of RI at various intensification rate thresholds. Predictors include forecast parameters (environmental predictors) and initial conditions (IR predictors and best-track/advisory-based predictors). Static predictors (i.e., those available only at $t = 0$) are italicized.

Acronym	Description
	Environmental predictors (time averaged from $t = 0$ to time of the forecast)
GSHR	850–200 hPa generalized wind shear calculated as the mass-weighted root-mean-square deviations of the winds from the mass-weighted deep-layer mean winds times a factor of 4 calculated in a 200–800-km annulus (Knaff et al. 2005)
OHC	Oceanic heat content between the surface and the depth of the 26°C isotherm (Shay et al. 2000, and references within)
RHMD	700–500-hPa relative humidity averaged within a 200–800-km annulus
DIVC	200-hPa divergence following the storm calculated in a 500-km circle centered on the TC
POT	Potential intensification calculated from the potential intensity as a function of SST at the storm center and the current intensity (at $t = 0$)
REFC	Average relative eddy momentum flux convergence ($\mathrm{m\,s^{-1}\,day^{-1}}$) calculated in a 100–600-km annulus vs time
TADV	The temperature advection between 850 and 700 hPa averaged from 0 to 500 km calculated from the geostrophic thermal wind
	IR predictors
PC50	Percentage of IR pixels colder than −50°C within a 50–200-km annulus
PC60	Percentage of IR pixels colder than −60°C within a 50–200-km annulus
SDO	Standard deviation of IR brightness temperature at 100–300 km
RMNT	Radius of minimum brightness temperature (0–150 km)
FR5	The deviation of IR-based TC size (R5) from the climatological population as a function of TC intensity
	Best-track/advisory-based predictors
VMAX	Current TC intensity ($t = 0$)
DV	12-h change in TC intensity, which is limited by the following function: $DV = \min[\min(VMAX \times 0.33, 17.5), DV]$

important to the TC intensification process (see Callaghan and Tory 2014) and is included for inspection.

IR imagery is a subjective analysis tool for predicting the onset of RI, and it is available almost instantaneously at operational centers, thus making the imagery ideal for an application such as ours. With this premise in mind, we use IR imagery to capture two characteristics related to intensification (see the middle section of parameters in Table 1). The first is convective vigor and symmetry. Previous work has shown that convective vigor can be discerned through cold-temperature-based pixel counts and that convective symmetry can be computed with standard deviations of the azimuthal temperatures surrounding the storm in the 100–300-km annulus (SDO; see Mundell 1990; Fitzpatrick 1997; Kaplan et al. 2010). The convective vigor predictors used here, PC50 and PC60, were chosen based on sensitivity work shown in Knaff et al. (2014a), where −50°C and colder pixel counts improved discrimination of RI events in the Atlantic and eastern Pacific. The storm size and structure is the second characteristic provided by the IR imagery. Knaff et al. (2014a, 2016) developed a normalized IR-based TC size (FR5), where the TC size parameter (R5 from Knaff et al. 2014b) is divided by its intensity-based climatological values. FR5, in essence, determines whether a TC at a specific intensity is large or small relative to the global climatology. We include this parameter because studies (e.g., Xu and Wang 2015) find that small TCs have a tendency to intensify more quickly than larger TCs. In addition to the overall TC size, we also make use of the radius of minimum azimuthally averaged IR temperature (RMNT)—a crude measure of inner-core and eye size, the latter being inversely correlated to intensification (Musgrave 2011; Carrasco et al. 2014).

Because of the nature of operations, nonphysical values can arise in the real-time, best-track parameters (bottom section in Table 1). Since large changes in operational intensity estimates are sometimes related to the inspection of time-late information (e.g., microwave imagery and scatterometry) and/or restrictions mandated by diagnostic techniques (e.g., Dvorak intensity estimate), we limit our 12-h intensification rate (DV) to physically realistic values using the following relationship (VMAX is the current TC intensity):

$$DV = \min[\min(VMAX \times 0.33, 17.5), DV]. \quad (1)$$

This effectively limits DV to constraints used in Dvorak (1984) and provides a more stable predictor for the linear statistical techniques used in this work. If our DV development dataset has large positive outliers, those will adversely influence the statistics. If DV is large and positive in real time (e.g., the forecaster received new

data that dramatically increased the current intensity estimate, but did not have data to modify the past intensity estimates), those will adversely influence our RI algorithm forecast. Including this parameterized DV value improved the dependent model fit, and sensitivity tests (not shown) indicated that capping DV in this physically consistent way resulted in better modeling of RI and answered the more basic question of whether the storm has been intensifying.

Construction of the RI algorithm follows the same process as past RI algorithm development efforts (e.g., Kaplan and DeMaria 2003; Kaplan et al. 2010, 2015), using a combination of current TC conditions, environmental conditions, and information about the current IR structure to forecast the probability of various intensification rates. In development, we use analyses (i.e., perfect prog) of environmental conditions. In applications, environmental conditions are based on forecasts. We also use two statistical methods to create forecasts from which we construct a two-member consensus forecast. The two methods are a linear discriminant analysis and logistic regression.

b. Linear discriminant analysis

Linear discriminant analysis (LDA) is a classification method originally developed by Fisher (1936). In LDA, a linear combination of variables that best separates two or more groups is developed. We define just two groups for the LDA: group 1, for when the intensification threshold is reached or exceeded, and group 2, for when the intensification threshold is not reached. In the two-class LDA, the goal is to find the n-dimension vector of observations that best assigns a case to belonging to either group 1 or 2. In our application, we assume both groups have the same covariance structure, so the vector has a direction in n-dimensional space that maximizes the distance between the means of groups 1 and 2 in standardized units. This is formalized as the discriminant function δ, which is the scalar projection of the data vector $\mathbf{x}$ in the direction of maximum separation (i.e., Mahalanobis distance), which is called the discriminant vector $\mathbf{a}$. The equation for δ is

$$\delta = \mathbf{a}^{T}\mathbf{x}, \tag{2}$$

where $\mathbf{x}$ is the data vector (our predictors) and $\mathbf{a}$ is the discriminant vector.

We have used the International Mathematical and Statistical Libraries (IMSL 2017) to make the calculations [see Wilks (2006) for a discussion of LDA]. Forward variable selection is used with a 95% confidence (F test), with an occasional backward step to remove variables that are no longer statistically significant. To estimate probabilities from the discriminant function provided by LDA, a windowing procedure relates prior probabilities (i.e., dependent data) to discriminant function values. This windowing procedure is a one-dimensional single-pass Barnes (1964) analysis, sampling the discriminant function at discrete intervals, 0.1 in this case, where all the discriminant function values within a radius of influence of 0.3 of each interval are weighted by the square distance from the interval. In application, a cubic spline provides a probability given the discriminant function value.

c. Logistic regression

Logistic regression (LRE) is a model where the dependent variable is a defined category. In our case, "1" is used for reaching the intensification threshold and "0" for not having met the intensification threshold. LRE is a special case of the generalized linear model, where the natural log of the odds ratio or logit based on categorical data is fit to a linear combination of independent predictors $(x_1, \ldots, x_n)$ with intercept b_o and weights $(b_1, \ldots, b_n)$ that are determined via the method of maximum likelihood:

$$\ln\left(\frac{p_e}{1-p_n}\right) = b_o + b_1 x_1 + \cdots + b_n x_n. \tag{3}$$

To perform variable selection and the model fit, we use FORTRAN 90 code from CSIRO (2017) that produces linear logistic models by iteratively reweighted least squares. Model fit is based on maximum likelihood criteria. Forward variable selection was used with the knowledge of the LDA model formulations. An occasional backward step was performed to remove predictors that had lost their statistical significance (99%, chi squared). The logistic regression model has different assumptions about the relationship between dependent and independent variables when compared to linear regression. The two primary differences are 1) since the dependent variable is binary, the conditional distribution is a Bernoulli distribution rather than a Gaussian distribution, and 2) the predicted values are probabilities of a particular outcome. Once fitted, the probability of exceeding the intensification threshold takes the form

$$p_{\mathrm{RI}} = \frac{1}{(1 + e^{-(b_o + b_1 x_1 + \cdots + b_n x_n)})}. \tag{4}$$

The measure of the quality of fit for logistic regression is in terms of deviance: a generalization of the idea of using the sum of squares of residuals in ordinary least squares to cases, but where the model is fit using a maximum likelihood criterion. Deviance is defined as −2 times the log-likelihood ratio of the fitted model

TABLE 2. Statistics associated with the LDA models for various intensity change thresholds.

Intensification threshold	Climatological frequency (%)	BS	BSS	No. of predictors	No. of cases
RI25	18.8	1162.8	23.8	8	5447
RI30	13.0	890.4	21.1	8	5396
RI35	9.3	681.0	19.4	8	5301
RI40	6.5	519.7	15.0	8	5211
RI45	12.3	871.4	19.5	9	4432
RI55	7.4	576.4	15.8	9	4166
RI70	7.4	594.7	13.0	9	3290

compared to the full (i.e., perfect) model. One can also define the percent deviance explained as 1 minus the ratio of the fitted model deviance to the deviance of a model containing only the intercept b_o (Knaff and DeMaria 2017).

3. Statistical–dynamical model formulations

For this work, we will examine several intensification thresholds. These include 25-, 30-, 35-, and 40-kt changes in 24 h; 45- and 55-kt changes in 36 h; and 70-kt changes in 48 h. These changes in the western North Pacific correspond to the 81.2, 87.0, 90.7, and 93.7 percentile values of 24-h intensity change; the 87.7 and 92.6 percentile values for the 36-h intensity change; and the 92.6 percentile value of the 48-h intensity change. We will refer to the thresholds as RI25, RI30, RI35, RI40, RI45, RI55, and RI70, respectively. We also tried to predict the 85-kt increase of intensity in 72 h, but dependent fits were, in our opinion, not good enough to pursue real-time prediction. With these thresholds, we now describe the statistical–dynamical models for the LDA and LRE methods for each of the intensity change thresholds.

a. LDA-based models

As mentioned in section 2, LDA is primarily a categorization method and probabilistic forecasts based on discriminant function values come from prior/dependent frequencies. For our intensification thresholds, we seek the model with the best developmental Brier scores (BSs; i.e., the mean square distance in probabilistic space) and Brier skill scores (BSSs). The calculation of BSS is provided in (5), where BS_f is the BS of the forecasts and BS_r is the BS of the reference (Wilks 2006), In this case, the reference is climatology:

$$BSS = 1 - \frac{BS_f}{BS_r}. \tag{5}$$

This also corresponds to the discriminant vector that provides the best categorization of cases that meet or exceed the intensification threshold and cases that do not. For fitting LDA models, we calculate time-averaged values of nonstatic predictors up until each forecast lead time. In application, we use forecast values of these quantities for the calculation of averaged predictor values. The previous section provided the climatological rate of RI for each threshold, and Table 2 provides statistics including climatological frequency of RI for each threshold, BS and BSS values, the number of predictors used, and the number of cases for each intensification threshold. BSS values decrease with increasing rates of intensity change and forecast difficulty.

The following set of eight predictors is used to make forecasts for RI25, RI30, and RI35 intensification rates: VMAX, DV, GSHR, OHC, PC50, SDO, potential intensification (POT), and DIVC (see Table 2). Forecasts for RI40 make use of the same predictors, save the PC50 predictor. In this case, PC60, the colder pixel count predictor, replaces PC50. The remaining intensification thresholds, RI45, RI55, and RI70, use the same predictors as RI40 with the addition of the inner-core predictor RMNT.

Figure 1 shows the normalized (by their standard deviations) discriminant function weights [i.e., vector **a** in (2)] used for each intensification threshold; 24-, 36-, and 48-h lead times in blue, green, and red hues, respectively. The first observation with respect to the individual predictors is the similarity between the different intensification threshold forecasts. Static predictors tend to become less important for the longer lead times, especially the pixel counts, PC50 and PC60, and the initial intensity. Forecast conditions for GSHR and DIVC appear to become more important as lead time increases, while OHC contributions are remarkably consistent among the thresholds. It is also important to note that VMAX and POT predictors appear to be inversely related to each other. In the absence of VMAX, one would expect POT to be positively correlated with RI, but in Fig. 1, we see the opposite. Upon further inspection, we found a much better discrimination when we used both of these variables despite their collinearity. Other that the POT coefficients, the others make sense physically with lower initial intensity (VMAX), vertical
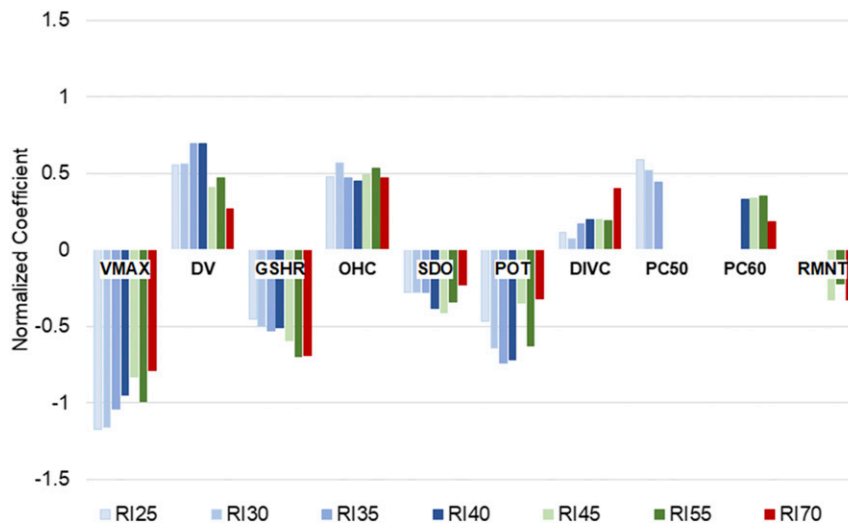
FIG. 1. Normalized discriminant function coefficients for LDA models of RI25, RI30, RI35, and RI40 (blue hues); RI45 and RI55 (green hues); and RI70 (red). Predictors, which are listed just under the $y = 0$ line, are described in Table 1.

wind shear (GSHR), convective variability (SDO), and smaller inner-core conditions (RMNT) and recent intensification (DV), higher OHC, DIVC, and more deep convection (PC50, PC60) being favorable for rapid intensification events. It is noteworthy that when VMAX is not used as a predictor, the POT coefficient has a positive sign, further suggesting that the negative POT coefficient is due to collinearity.

b. Logistic regression-based models

As described in section 2, we used logistic regression to create probabilistic forecast modes for the intensity change thresholds. For the development of these models, we used the same potential predictors (Table 1), but here we seek to minimize the deviance (i.e., maximize the deviance explained) for each intensity change threshold. Table 3 presents the number of predictors, the deviance explained, and the BSS for each intensity change threshold. These statistics show that these models consistently explain about 24%–29% of the deviance using between 8 and 10 predictors and dependent BSSs that are just a little bit larger than the LDA-based models. Table 4 shows the predictors used in each model. It is interesting to note that the static predictors DV, PC50, and RMNT, and time-varying predictors GSHR and OHC, are selected for every lead time. In the LRE-based models, the importance of the SDO predictor is reduced compared to the LDA model. Also, the LRE predictor selection included FR5 (TC size) as a predictor for many of the lead times while it was not for the LDA model. RHMD and TADV become important specifically for the RI70 forecasts in the LRE model, and

the reasons for this and other predictive relationships are now explored.

Since the coefficients of the logistic regression have essentially the same convention and meaning as linear regression coefficients, we show the normalized coefficients for these models in Fig. 2. The positively correlated RHMD and negatively correlated TADV that are important specifically for the 48-h RI70 forecasts appear on the right in Fig. 2. These are likely related to extratropical transition cases. Positive temperature advection and decreased humidity could both be effects from an approaching midlatitude trough, and so should suppress RI. As is the case for LDA, the LRE predictors VMAX and POT are covariant. And as is the case with the LDA, removing either predictor significantly increased the deviance and decreased the percent deviance explained. This covariance between VMAX and

TABLE 3. Number of predictors and percentage deviance explained by dependent logistic regressions for the intensification thresholds used in this study. BSSs are also provided for comparison with LDA-based models in the column labeled BSS.

Intensity change threshold	No. of predictors	Percentage deviance explained	BSS
RI25	10	27	26.5
RI30	8	28	24.4
RI35	9	29	21.4
RI40	9	28	22.2
RI45	10	26	21.5
RI55	9	28	18.0
RI70	9	24	12.4

TABLE 4. Listing of the predictors used in LRE models for each intensification threshold and each model. The symbol O is for LDA and X is for LRE models.

Predictor	RI25	RI30	RI35	RI40	RI45	RI55	RI70
VMAX	O, X	O, X	O, X	O, X	O, X	O, X	O
DV	O, X	O, X	O, X	O, X	O, X	O, X	O, X
GSHR	O, X	O, X	O, X	O, X	O, X	O, X	O, X
OHC	O, X	O, X	O, X	O, X	O, X	O, X	O, X
PC50	O, X	O, X	O, X	X	X	X	X
PC60				O	O	O	O
SDO	O, X	O, X	O	O	O, X	O	O
POT	O, X	O, X	O, X	O, X	O X	O, X	O, X
DIVC	O, X	O		O, X	O, X	O, X	O, X
FR5	X		X	X	X	X	
RMNT	X	X	X	O, X	X	O, X	O, X
RHMD							X
TADV							X

POT causes a nonphysical coefficient sign on the VMAX predictor. Similarly, if POT is removed from the model, the sign of the VMAX coefficient becomes negative and physically consistent. The remaining predictors all have physically consistent sign conventions. Similar to the LDA results, the time-averaged environmental predictors generally become more important for longer forecasts (e.g., RI70) and the static predictors exert less influence. Both TC size (FR5) and inner-core size (RMNT) make significant contributions in most models, suggesting that smaller storms and those with the coldest brightness temperatures near the center are more likely to be rapid intensifiers. The result that smaller TC sizes are favorable for rapid deepening is consistent with the work of Weatherford (1989), who found that rapid deepeners started with higher central pressures [smaller TCs generally have higher central pressures; see Knaff and Zehr (2007), Courtney and Knaff (2009), and Chavas et al. (2017)]. It is also consistent with the inner/outer convection ratio predictor described in Mundell (1990) that measures the concentration of convection in the 0°–2° range (pixels colder than −70°C) relative to the 2°–6° range (pixels colder than −60°C). The finding that RI favors smaller inner-core convective rings also agrees with results obtained by Xu and Wang (2015) and Carrasco et al. (2014). It is also worth mentioning that there is no advantage to using PC60 in any of the equations. It is also clear that convective vigor is more important as the intensification threshold/rate increases in the LRE models. This is different from coefficient weights found in the LDA models. Also unlike the LDA models, there are the relatively small or zero weights for the SDO predictor.

The differences in weights and predictors in the LDA and LRE models suggest that the results of these two methods may be different. This difference implies some independence in the methods, making them ideal for use in a consensus forecast. Consensus forecasts, averages of forecasts from more than one method, have been shown to yield improvements over individual forecasts in a variety of fields varying from economics (Bates and Granger 1969) to investment decisions (Jones 2014) to sensible weather (Sanders 1973) to political elections (Graefe et al. 2014). Furthermore, these studies all indicate that the degree of independence among consensus members is an important factor when combining forecasts, contributing to forecast improvements (e.g., Sampson et al. 2008, their appendix B). With the goal of creating the most skillful forecasts and useful guidance for rapid intensification, the next section discusses the preprocessing and combining of the LDA- and LRE-based models and the creation of deterministic intensity forecasts.
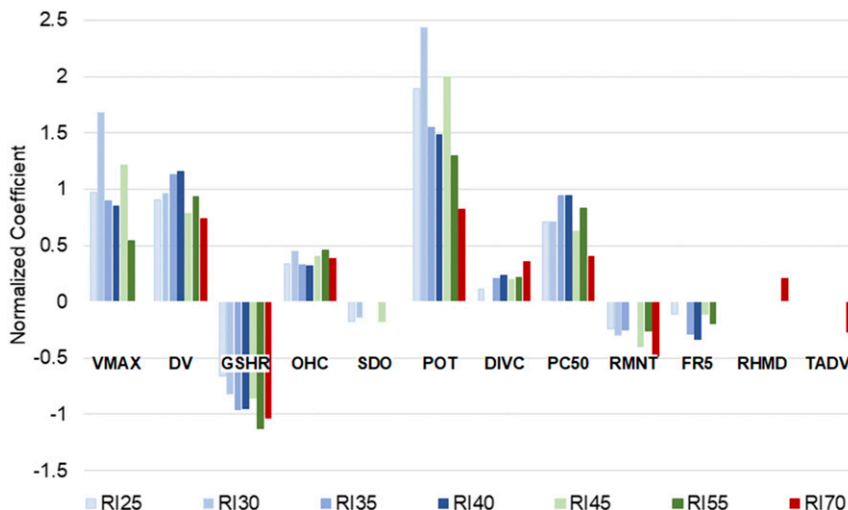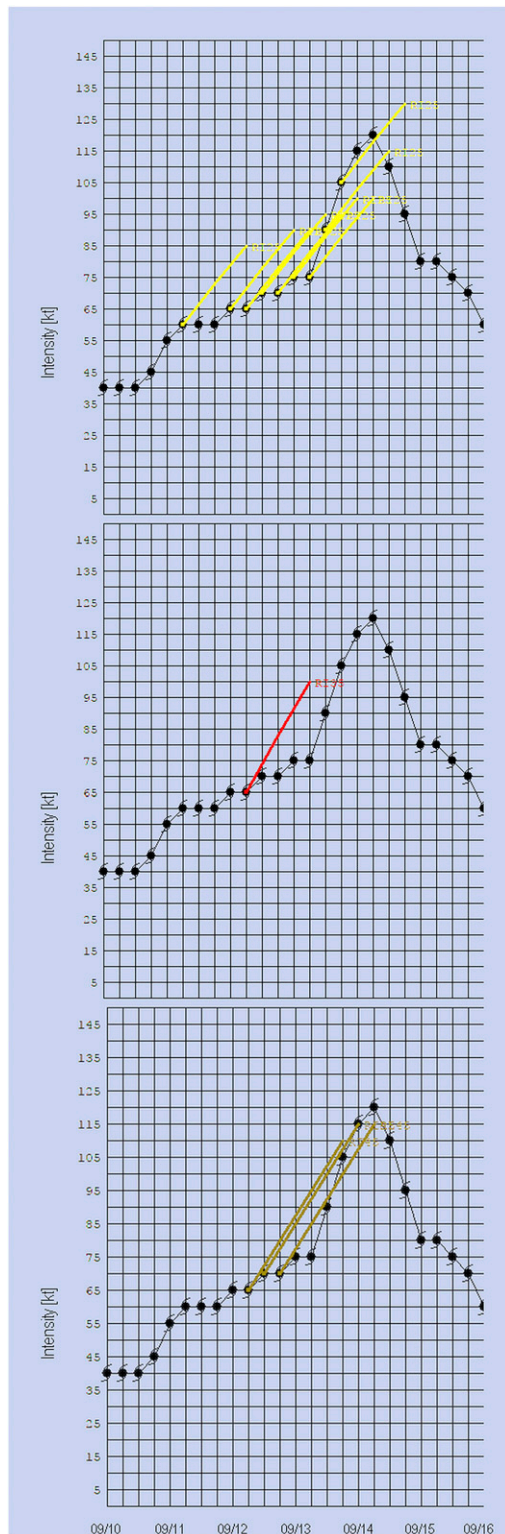

FIG. 2. As in Fig. 1, but for LRE.

FIG. 3. Examples of deterministic RI forecasts triggered during Typhoon Talim (wp202017) vs the working best-track intensities. Deterministic forecasts are shown for (top) RI25, (middle) RI35, and (bottom) RI45. These forecasts were produced in real time at JTWC, and graphics represent how these are displayed in ATCF.

c. Preprocessing and combining forecasts and deterministic forecasts

The LDA- and LRE-based models described above provide probabilistic RI forecasts for distinct intensification thresholds and lead times. The models use slightly different predictors and therefore exhibit some degree of independence. As a result, the RI model forecasts can occasionally produce probabilities that are inconsistent; that is, a higher RI threshold (e.g., RI35) may have a higher probability than the lower RI thresholds for that lead time (RI25 and RI30). In such cases, probabilities of the lower RI thresholds are assigned the probability of the higher RI threshold. In our example above, the probabilities associated with RI25 and RI30 are assigned the RI35 probability. This consistency check among RI thresholds is performed for both 24- and 36-h forecasts, and for each forecast methodology (LDA and LRE) independently. Following the consistency check, we use probabilities from the two forecast methodologies for each intensification threshold to create an equally weighted average (CON).

Traditionally the rapid TC intensification forecast problem led to categorical/binary (Mundell 1990) or probabilistic (i.e., Kaplan and DeMaria 2003) forecasts. Despite the reasoning for this decision, forecasters are still required to provide a deterministic forecast of intensity. Intensity guidance has been improving (DeMaria et al. 2007, 2014) in terms of mean absolute error (MAE), but intensity forecast verification shows that most models have low biases, especially in RI cases (Kaplan et al. 2015). To address the negative biases in RI forecasts, Sampson et al. (2011) demonstrated a method of providing deterministic intensity forecasts based on RI forecast probabilities. In this method, threshold values of the probabilistic forecast trigger deterministic forecasts for the valid forecast lead time. For instance, when the RI35 forecast exceeds the threshold probability, the algorithm generates a 12-hourly deterministic forecast of 35 kt in 24 h starting with the observed intensity at $t = 0$ and ending 24 h later with an intensity of the initial intensity plus 35 kt. These deterministic forecasts are then added to the operational intensity consensus forecast at JTWC, which is the most skillful intensity guidance (DeMaria et al. 2014). Both mean errors and biases are smaller when the intensity forecast aid contains deterministic RI aids (Sampson et al. 2011).

The threshold probability for triggering rapid aids is determined using past forecasts. For our purposes, and based on independent CON forecasts, we found the threshold of ~40% ± 10% for all of the intensification thresholds. For this reason, the 40% probability triggers deterministic rapid intensification forecast aids for RI25,

RI30, RI35, RI40, RI45, RI55, and RI70 based on the CON forecast probabilities. Only the deterministic member with the highest intensification rate for each lead time is triggered. So conceivably, for one consensus forecast, three deterministic RI members could be added, one each for 24-, 36-, and 48-h leads. Figure 3 shows an example for Typhoon Talim (wp202017) for the thresholds RI25, RI35, and RI45 at the times they were triggered, capturing the periods of time when more rapid intensity changes were occurring. Note this was the first TC captured after these aids became available in operations at JTWC, and they represent an operational forecast.

Above, we described probabilistic rapid intensification aids using LDA and LRE methods for seven intensification thresholds (RI25, RI30, RI35, RI40, RI45, RI55, and RI70), how a simple consensus (CON) is formed, and how CON is used to trigger deterministic forecast members that are used by JTWC's operational intensity consensus guidance tool. In the next section, we discuss the independent verification of LDA, LRE, and CON forecasts.

4. Independent verification

To evaluate the performance of the models described in section 3, we now present verification statistics based on almost two years of independent forecasts produced as part of the vetting process. Verification data include all forecasts from the 2016 season and forecasts through 27 October 2017. We present BSSs with the climatology of the individual intensification thresholds, reliability, and the verification of the impact of the deterministic rapid intensification members triggered by the probabilistic models.

Before presenting statistics, we review the methods. The BS is simply the mean squared probability error for a given probabilistic forecast (Murphy 1973). One can form a skill score referred to as the BSS using (5), where BS_f and BS_r are the BS for the forecast method and for a reference forecast, respectively. In our case, the reference forecast is the climatological frequency of the event occurring. The BSS in (5) answers the following question: What is the relative skill of the probabilistic forecast over that of climatology, in terms of predicting whether an event occurred? The reliability diagrams (Fig. 4) show observed frequency against the forecast probability, where the range of forecast probabilities is divided into 10 bins (e.g., 0%–10%, 10%–20%, 20%–30%, etc.). The sample size in each bin is also included as a histogram inside the reliability diagram. Reliability diagrams answer three questions graphically including 1) "reliability," or the agreement between forecast probability and mean observed frequency; 2) "sharpness," measuring the tendency to forecast probabilities near 0 or 1 versus values clustered around the mean; and 3) "resolution," or the ability of a forecast to resolve events into subsets with characteristically different outcomes (CAWCR 2017).

Table 5 shows the independent BSSs (%) for LDA, LRE, and CON forecasts for each intensification threshold. This is admittedly a limited sample with less than two years of verification. Brier skill scores show that the LRE methods are skillful for all the intensification threshold forecast models developed and that the LDA method produces skillful forecasts for only the RI25 intensification threshold model. Consensus forecasts provided skill for all but the RI55 cases. These results are discouraging for the LDA-based methods, but suggest that the LRE and CON have delivered skillful guidance. Results are subject to change due to changes in the final 2017 best tracks, particularly because modification of best-track intensities is likely when rapid intensity changes occur. Nonetheless, results are encouraging. Experience also suggests that at least three independent typhoon seasons are typically needed to make solid verification inferences.

Figure 4 shows reliability diagrams associated with our sample of independent forecasts. The LDA methods generally produce low-biased reliability (i.e., below the 1:1 line), whereas for the more rare intensification thresholds LRE results are indicative of high bias. We wish we could report that the CON is the best method in this sample, but it appears that the LREs have better reliability for RI25, RI30, RI35, and RI45. The CON forecasts generally produce reliabilities in between those of the LDA and LRE, but this is not always the case, suggesting that there may be a fair bit of independence between the LDA and LRE methods (e.g., RI45 for higher probabilities and frequencies). In general, these performance results are similar to rapid intensification aids in other basins—overconfident and low biased (e.g., Kaplan et al. 2015). It is also interesting to note that the 40% CON forecasts in general would trigger deterministic forecasts for about 15%–25% of the cases (i.e., overconfidence). We now move on to the discussion of the impact of using the rapid intensification deterministic forecasts in the intensity consensus at JTWC.

The addition of deterministic forecasts of rapid intensification based on the probabilistic models developed here should help with this issue by reducing the biases and possibly reducing the MAEs. Results based on our independent forecast sample used here have found this to be the case. Using the ideas presented in Sampson et al. (2011) and a 40% triggering probability
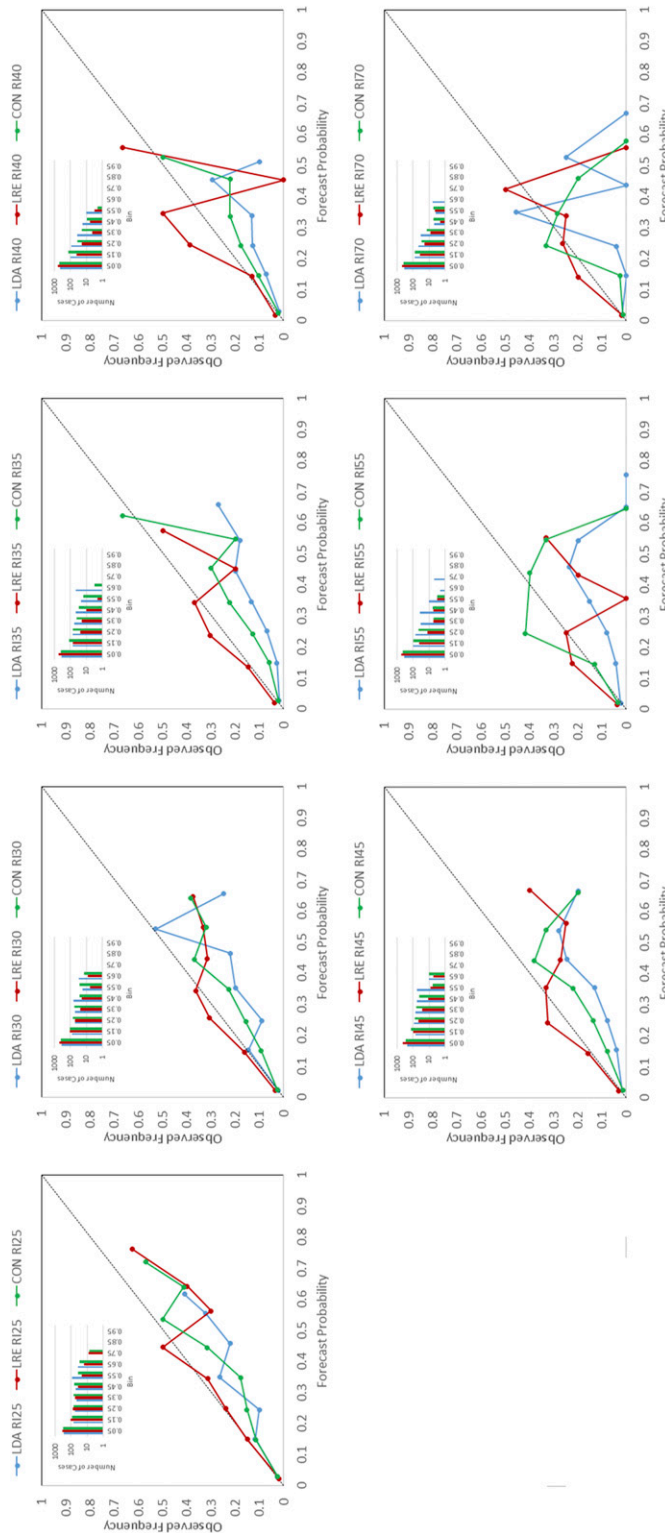
FIG. 4. Reliability diagrams showing LDA (blue), LRE (red), and CON (green), based on independent forecasts. Shown are reliability diagrams for (top) RI25, RI30, RI35, and RI45, and (bottom) RI45, RI55, and RI70. Note the distributions of cases are shown in the top left of each panel and have a log scale.

TABLE 5. Probabilistic RI algorithm evaluation. Dataset is independent and includes western North Pacific cases from 1 Jan 2016 through 27 Oct 2017.

	RI25	RI30	RI35	RI40	RI45	RI55	RI70
No. of forecasts	678	678	678	678	601	601	531
Percentage of climate	12	13	9	6	12	6	4
Percentage of observed	14	10	7	5	10	6	5
LDA BSS (%)	2	−4	−32	−10	−19	−27	−5
LRE BSS (%)	18	13	10	9	11	4	9
CON BSS (%)	13	8	2	5	4	−2	5

results in a dramatic decrease in the biases, which shows reductions in MAEs for 24-, 36-, and 48-h forecasts. These results represent a significant improvement of season intensity given the slow rate of change in intensity forecast performance. Table 6 shows the intensity verifications of the intensity consensus with the deterministic rapid intensification member versus those without. The 12- and 24-h periods have enough cases to provide a sense of the performance. The improvement in MAE at 24 h is 0.4 kt, and the expected low bias for RI cases is reduced by 1.9 kt. The improvements in MAE are very small or negative, but the reduction in bias is relatively large. The performance at the 36- and 48-h periods is promising, especially upon inspection of the individual cases, but there are too few cases to make any firm conclusions other than that the biases are reduced.

The improvements in MAE are small for these time periods, but this is partly a construct of availability. The RI aids at 24 h are available in 15% of all forecasts. As discussed in Sampson et al. (2011), raising the threshold above 40% improves the performance in terms of MAE and lessens the bias correction, but also reduces the availability. One can think of the 40% threshold as a tuning knob: turn it up and you get better MAEs, less bias correction, and fewer RI forecasts nudging the consensus; turn it down and you typically increase the MAE, improve the bias correction, and more often nudge the consensus. In this work, we turned the knob up to 50% and noticed the same behavior as discussed in detail in Sampson et al. (2011, their Fig. 2), and decided that 40% would also work in this basin. We also felt that overtuning of this parameter was possible given the short sample of developmental and independent data. Tuning this factor is something that can be revisited once the independent sample increases in size.

5. Summary

This article describes the development of TC rapid intensification models for the western North Pacific basin. We chose seven intensification thresholds including 25-, 30-, 35-, and 40-kt intensity changes in 24 h; 45- and 55-kt intensity changes in 36 h; and 70-kt intensity changes in 48 h, also referred to as RI25, RI30, RI35, RI40, RI45, RI55, and RI70, respectively. The models were developed as probabilistic algorithms following work in other basins and using two methodologies: linear discriminant analysis (LDA) and logistic regression (LRE). The years 2000–15 were used in the development, and then independent testing was performed with the years 2016 and 2017. Equally weighted averages of the LDA and LRE probabilities are computed (CON), and those are then used to trigger deterministic forecasts for each of the seven intensification thresholds when the probabilities reach 40%.

Dependent LDA models had Brier skill scores ranging from 13% to 24% that indicated they could produce skillful (relative to climatology) probabilistic forecasts. Dependent LRE models, on the other hand, explained a little less than 30% of the deviance and had slightly better BSSs than LDA models. In independent tests, LDA models with the exception of RI25 failed to produce skillful forecasts, but both LRE and CON produced skillful probabilistic forecasts. We also examined

TABLE 6. Deterministic intensity consensus evaluation. ICNW is the JTWC operational consensus that includes the deterministic RI aids while ICNC is the same consensus without RI aids. Cases include only those where at least one of the deterministic RI aids for a given forecast time was available. The dataset is independent and includes western North Pacific cases from 1 Jan 2016 through 27 Oct 2017.

Quantity/forecasts	Statistic	12 h	24 h	36 h	48 h
No. of forecasts		127	125	58	14
ICNW	MAE (kt)/bias (kt)	8.9/−3.7	13.9/−7.2	18.8/−12.8	16.5/−4.8
ICNC	MAE (kt)/bias (kt)	9.0/−4.6	14.3/−9.1	18.7/−14.2	14.4/−6.9
Percentage improvement	MAE (%)/bias (%)	1.0/19.6	2.7/21.0	−0.5/9.9	−14.6/30.4

reliability diagrams for independent forecasts of the LDA, LRE, and CON probabilities. The reliability results are typical for the rapid intensification problem, resulting in forecasts that are overconfident (e.g., Kaplan et al. 2015). The overconfidence suggests that the 40% CON forecast threshold would ultimately trigger deterministic forecasts for about 15% of the cases.

In independent verification with the 2016 and 2017 data, the JTWC intensity consensus that includes deterministic RI guidance clearly shows reduced negative biases and somewhat improved MAEs (i.e., reduction in RMSE), indicating that some independence and skill are garnered by including them in the consensus. The deterministic RI aids should also give forecasters improved intensity guidance spread, though this is yet to be shown.

We are also happy to report that experiments running these aids in JTWC's operations in the Indian Ocean and Southern Hemisphere have been rather successful, anticipating the rapid intensification associated with TCs they were forecasting and capturing the attention of forecasters at JTWC and at the Australian Bureau of Meteorology. Forecasters, and the authors, are still trying to formulate how to best use these guidance tools at these centers. Feedback thus far has concerned overprediction and thus poor performance in cases when the TC is relatively weak and the creation of deterministic forecasts when the TC is making or is expected to make landfall (J. Courtney and B. Strahl 2018, personal communications). The exchange of ideas of how to best use these aids and how to best nudge the intensity consensus is critical and still ongoing.

Future work will revolve around improving the use of these relatively new RI aids by providing examples of how these aids work and their shortcomings to those making forecasts at JTWC and elsewhere. We will also continue experimenting with using these aids in the other TC basins where best tracks are less reliable and the numbers of cases are relatively few, respectively. It is likely that a larger independent dataset and more forecaster feedback will further elucidate benefits and detriments of the new RI models, prompting changes and improvements in how these aids are used and presented to forecasters, as well as training on best practices

REFERENCES

Atkinson, G. D., and C. R. Holliday, 1977: Tropical cyclone minimum sea level pressure/maximum sustained wind relationship for the western North Pacific. *Mon. Wea. Rev.*, **105**, 421–427, https://doi.org/10.1175/1520-0493(1977)105<0421:TCMSLP>2.0.CO;2.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409, https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2.

Bates, J., and C. Granger, 1969: The combination of forecasts. *Oper. Res. Quart.*, **20**, 451–468, https://doi.org/10.1057/jors.1969.103.

Brand, S., 1973: Rapid intensification and low-latitude weakening of tropical cyclones of the western North Pacific Ocean. *J. Appl. Meteor.*, **12**, 94–103, https://doi.org/10.1175/1520-0450(1973)012<0094:RIALLW>2.0.CO;2.

Callaghan, J., and K. Tory, 2014: On the use of a system-scale ascent/descent diagnostic for short-term forecasting of tropical cyclone development, intensification and decay. *Trop. Cyclone Res. Rev.*, **3**, 78–90.

Carrasco, C. A., C. W. Landsea, and Y. Lin, 2014: The influence of tropical cyclone size on its intensification. *Wea. Forecasting*, **29**, 582–590, https://doi.org/10.1175/WAF-D-13-00092.1.

CAWCR, 2017: Forecast verification - Issues, methods and FAQ. Collaboration for Australian Weather and Climate Research, http://www.cawcr.gov.au/projects/verification/verif_web_page.html.

Chavas, D. R., K. A. Reed, and J. A. Knaff, 2017: Physical understanding of the tropical cyclone wind–pressure relationship. *Nat. Commun.*, **8**, 1360, https://doi.org/10.1038/s41467-017-01546-9.

Courtney, J., and J. A. Knaff, 2009: Adapting the Knaff and Zehr wind–pressure relationship for operational use in tropical cyclone warning centres. *Aust. Meteor. Oceanogr. J.*, **58**, 167–179, https://doi.org/10.22499/2.5803.002.

CSIRO, 2017: Software from Alan J. Miller. Commonwealth Scientific and Industrial Research Organisation, http://wp.csiro.au/alanmiller/.

DeMaria, M., and J. Kaplan, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337, https://doi.org/10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2.

——, J. A. Knaff, and C. R. Sampson, 2007: Evaluation of long-term trend in tropical cyclone intensity forecasts. *Meteor. Atmos. Phys.*, **97**, 19–28, https://doi.org/10.1007/s00703-006-0241-4.

——, C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, https://doi.org/10.1175/BAMS-D-12-00240.1.

Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. 11, 47 pp., http://severe.worldweather.wmo.int/TCFW/RAI_Training/Dvorak_1984.pdf.

Elliot, G., and Coauthors, 2014: Topic 1, Motion - Recent advances. Rep. on the Eighth International Workshop on Tropical Cyclones, WMO, http://www.wmo.int/pages/prog/arep/wwrp/new/documents/Topic1_AdvancesinForecastingMotion.pdf.

Finocchio, P. M., and S. J. Majumdar, 2017: A statistical perspective on wind profiles and vertical wind shear in tropical cyclone environments of the Northern Hemisphere. *Mon. Wea. Rev.*, **145**, 361–378, https://doi.org/10.1175/MWR-D-16-0221.1.

Fisher, R. A., 1936: The use of multiple measurements in taxonomic problems. *Ann. Eugen.*, **7**, 179–188, https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

Fitzpatrick, P. J., 1997: Understanding and forecasting tropical cyclone intensity change with the Typhoon Intensity Prediction Scheme (TIPS). *Wea. Forecasting*, **12**, 826–846, https://doi.org/10.1175/1520-0434(1997)012<0826:UAFTCI>2.0.CO;2.

Graefe, A., J. S. Armstrong, R. J. Jones, and A. G. Cuzán, 2014: Combining forecasts: An application to elections. *Int. J. Forecasting*, **30**, 43–54, https://doi.org/10.1016/j.ijforecast.2013.02.005.

Holliday, C. R., 1977: On the rapid intensification of typhoons. M.S. thesis, Department of Meteorology, Texas A&M University, 87 pp., http://hdl.handle.net/1969.1/ETD-TAMU-1977-THESIS-H739.

——, and A. H. Thompson, 1979: Climatological characteristics of rapidly intensifying typhoons. *Mon. Wea. Rev.*, **107**, 1022–1034, https://doi.org/10.1175/1520-0493(1979)107<1022:CCORIT>2.0.CO;2.

IMSL, 2017: IMSL Fortran Numerical Stat Library. International Mathematics and Statistics Library, http://docs.roguewave.com/imsl/fortran/7.0/stat/stat.htm.

Jones, R. C., 2014: Making better (investment) decisions. *J. Portfolio Manage.*, **40**, 128–143, https://doi.org/10.3905/jpm.2014.40.2.128.

Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting*, **18**, 1093–1108, https://doi.org/10.1175/1520-0434(2003)018<1093:LCORIT>2.0.CO;2.

——, ——, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and east Pacific basins. *Wea. Forecasting*, **25**, 220–241, https://doi.org/10.1175/2009WAF2222280.1.

——, and Coauthors, 2015: Evaluating environmental impacts on tropical cyclone rapid intensification predictability utilizing statistical models. *Wea. Forecasting*, **30**, 1374–1396, https://doi.org/10.1175/WAF-D-15-0032.1.

Knaff, J. A., and R. M. Zehr, 2007: Reexamination of tropical cyclone wind–pressure relationships. *Wea. Forecasting*, **22**, 71–88, https://doi.org/10.1175/WAF965.1.

——, and R. T. DeMaria, 2017: Forecasting tropical cyclone eye formation and dissipation in infrared imagery. *Wea. Forecasting*, **32**, 2103–2116, https://doi.org/10.1175/WAF-D-17-0037.1.

——, C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688–699, https://doi.org/10.1175/WAF863.1.

——, D. P. Brown, J. Courtney, G. M. Gallina, and J. L. Beven II, 2010: An evaluation of Dvorak technique–based tropical cyclone intensity estimates. *Wea. Forecasting*, **25**, 1362–1379, https://doi.org/10.1175/2010WAF2222375.1.

——, M. DeMaria, S. P. Longmore, and R. T. DeMaria, 2014a: Improving tropical cyclone guidance tools by accounting for variations in size. *31st Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 51, https://ams.confex.com/ams/31Hurr/webprogram/Paper244165.html.

——, S. P. Longmore, and D. A. Molenar, 2014b: An objective satellite-based tropical cyclone size climatology. *J. Climate*, **27**, 455–476, https://doi.org/10.1175/JCLI-D-13-00096.1.

——, C. J. Slocum, K. D. Musgrave, C. R. Sampson, and B. R. Strahl, 2016: Using routinely available information to estimate tropical cyclone wind structure. *Mon. Wea. Rev.*, **144**, 1233–1247, https://doi.org/10.1175/MWR-D-15-0267.1.

Mundell, D. B., 1990: Prediction of tropical cyclone rapid intensification events. M.S. thesis, Dept. of Atmospheric Science, Colorado State University, 186 pp. [Available from Dept. of Atmospheric Science, Colorado State University, 3915 W. Laporte Ave., Fort Collins, CO 80523.]

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600, https://doi.org/10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

Musgrave, K. D., 2011: Tropical cyclone inner core structure and intensity change. Ph.D. dissertation, Colorado State University, 103 pp. [Available from Dept. of Atmospheric Science, Colorado State University, 3915 W. Laporte Ave., Fort Collins, CO 80523.]

Neumann, C. J., and M. B. Lawrence, 1975: An operational experiment in the statistical-dynamical prediction of tropical cyclone motion. *Mon. Wea. Rev.*, **103**, 665–673, https://doi.org/10.1175/1520-0493(1975)103<0665:AOEITS>2.0.CO;2.

RAMMB, 2017: SHIPS developmental data. Regional and Mesoscale Meteorology Branch, Cooperative Institute for Research in the Atmosphere, Colorado State University, http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting system (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, https://doi.org/10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2.

——, and J. A. Knaff, 2015: A consensus forecast for tropical cyclone gale wind radii. *Wea. Forecasting*, **30**, 1397–1403, https://doi.org/10.1175/WAF-D-15-0009.1.

——, J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting*, **23**, 304–312, https://doi.org/10.1175/2007WAF2007028.1.

——, J. Kaplan, J. A. Knaff, M. DeMaria, and C. Sisko, 2011: A deterministic rapid intensification aid. *Wea. Forecasting*, **26**, 579–585, https://doi.org/10.1175/WAF-D-10-05010.1.

Sanders, F., 1973: Skill in forecasting daily temperature and precipitation: Some experimental results. *Bull. Amer. Meteor. Soc.*, **54**, 1171–1178, https://doi.org/10.1175/1520-0477(1973)054<1171:SIFDTA>2.0.CO;2.

Shay, L. K., G. J. Goni, and P. G. Black, 2000: Effects of a warm oceanic feature on Hurricane Opal. *Mon. Wea. Rev.*, **128**, 1366–1383, https://doi.org/10.1175/1520-0493(2000)128<1366:EOAWOF>2.0.CO;2.

Shu, S., J. Ming, and P. Chi, 2012: Large-scale characteristics and probability of rapidly intensifying tropical cyclones in the western North Pacific basin. *Wea. Forecasting*, **27**, 411–423, https://doi.org/10.1175/WAF-D-11-00042.1.

Torn, R. D., and C. Snyder, 2012: Uncertainty of tropical cyclone best-track information. *Wea. Forecasting*, **27**, 715–729, https://doi.org/10.1175/WAF-D-11-00085.1.

Velden, C. S., and J. Sears, 2014: Computing deep-tropospheric vertical wind shear analyses for tropical cyclone applications: Does the methodology matter? *Wea. Forecasting*, **29**, 1169–1180, https://doi.org/10.1175/WAF-D-13-00147.1.

Wang, Y., Y. Rao, Z. Tan, and D. Schönemann, 2015: A statistical analysis of the effects of vertical wind shear on tropical cyclone intensity change over the western North Pacific. *Mon. Wea. Rev.*, **143**, 3434–3453, https://doi.org/10.1175/MWR-D-15-0049.1.

Weatherford, C., 1989: The structural evolution of typhoons. Dept. of Atmospheric Science Paper 446, Colorado State University, Fort Collins, CO, 198 pp., https://tropical.colostate.edu/media/sites/111/2016/10/446_Weatherford.pdf.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction.* 2nd ed. Academic Press, 627 pp.

Xu, J., and Y. Wang, 2015: A statistical analysis on the dependence of tropical cyclone intensification rate on the storm intensity and size in the North Atlantic. *Wea. Forecasting*, **30**, 692–701, https://doi.org/10.1175/WAF-D-14-00141.1.