OPERATIONAL PREDICTION SYSTEM NOTES (OPS NOTES)

A Rapid Intensification Deterministic Ensemble (RIDE) for the Joint Typhoon Warning Center's Area of Responsibility

JOHN A. KNAFF,[a] CHARLES R. SAMPSON,[b] ALAN BRAMMER,[c] AND CHRISTOPHER J. SLOCUM[a]

[a] *NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado*
[b] *Naval Research Laboratory, Monterey, California*
[c] *NOAA/Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado*

(Manuscript received 19 January 2023, in final form 19 April 2023, accepted 27 April 2023)

ABSTRACT: The Rapid Intensification Deterministic Ensemble (RIDE) is an operational method used to estimate the probability of tropical cyclone rapid intensification in the Joint Typhoon Warning Center's area of responsibility. Inputs to RIDE are current intensity, storm latitude, intensity change forecasts from seven routinely available operational deterministic models of intensity change, and the number of those models exceeding their individual 90th percentile of intensity change. Deterministic model inputs come from four numerical weather prediction models, two statistical–dynamical models, and one purely statistical model. In RIDE, logistic regression combines the deterministic inputs to form a probabilistic rapid intensification forecast model. RIDE then also generates deterministic intensity forecasts from these probabilistic forecasts that serve as forecaster guidance and as input to intensity consensus aids. Results based on a year of independent verification suggest good reliability and discrimination with a general tendency to underpredict rapid intensification events, but with few false alarms.

SIGNIFICANCE STATEMENT: An operational tropical cyclone forecaster makes a forecast with deterministic and probabilistic intensity guidance tools at their disposal. These models have a varying degree of abilities for predicting both intensity change and rapid intensification. The forecaster faces a dilemma in how to combine this disparate guidance to anticipate rapid intensification events. Here, the RIDE model provides probability forecasts associated with rapid intensification at 12-, 24-, 36-, 48-, and 72-h lead times and associated deterministic forecasts. RIDE provides skillful rapid intensification forecasts and helps rectify this forecast dilemma.

KEYWORDS: Hurricanes/typhoons; Forecasting techniques; Postprocessing

1. Introduction

The U.S. Department of Defense's Joint Typhoon Warning Center (JTWC) makes tropical cyclone (TC) forecasts through five days for all TCs in its area of responsibility (AOR), which includes the western North Pacific (WP), north Indian Ocean (IO), and Southern Hemisphere (SH). Typically, JTWC's forecasts begin when storms reach an intensity of 25 kt (1 kt $\approx$ 0.51 m s^{-1}) in the WP and an intensity of 35 kt in the IO and SH in terms of the maximum 1-min sustained winds. Because intensity is reported in knots in JTWC operations, that unit is used throughout the remainder of this work.

JTWC's TC intensity forecasts have shown slow and steady improvement from 2018 to 2021 (Zhang et al. 2022). However, gains in skill for rapid intensification forecasts remain challenging. Rapid intensification is often associated with the development of a TC with intensities greater than 95 kt (Lee et al. 2016) increasing the potential hazards and impacts. And, if unanticipated and close to land can lead to disasters like Supertyphoon Omar[1] that rapidly intensified prior to making landfall in Guam in August 1992 (Mautner and Guard 1993). To address this challenge, JTWC is prioritizing research and development of basin-specific probabilistic and deterministic forecast guidance targeting intensity change, particularly the onset, duration, and magnitude of rapid intensity change events at 2–3-day lead times (Francis and Strahl 2022; JTWC 2021). To address this priority, we develop a rapid intensification deterministic ensemble (RIDE) that provides probabilistic guidance for rapid intensity change events at fixed thresholds. The thresholds chosen are as follows: 20 kt in 12 h (RI20), 25, 30, 35, and 40 kt in 24 h (RI25, RI30, RI35, and RI40), 45 and 55 kt in 36 h (RI45, RI55), 55 kt in 48 h (RI56), and 65 kt in 72 h (RI65). Collectively, these thresholds will be called rapid intensification (RI) events.

Denotes content that is immediately available upon publication as open access.

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

[1] Supertyphoon Omar (1992) resulted in $457 million (U.S. dollars) in damages, 3000 people left homeless, and because of the rapid intensification, two U.S. Navy ships were unable to sortie, broke their moorings, and went aground.

DOI: 10.1175/WAF-D-23-0012.1

The deterministic to probabilistic statistical model (DTOPS; DeMaria et al. 2021) at the National Hurricane Center provided inspiration for the RIDE model. Like RIDE, DTOPS combines several deterministic forecasts with logistic regression to generate a probability of several RI intensity change thresholds.

JTWC has a suite of RI guidance in addition to RIDE that includes statistical–dynamical aids and ensemble-based aids that estimate probabilities of RI events, and provides deterministic forecasts based on probability thresholds, and two mesoscale models. The statistical–dynamic aids are the Rapid Intensification Prediction Aid (Knaff et al. 2018, 2020), and the Forest Rapid Intensification Aid (Slocum 2021). The ensemble-based aids are based on the U.S. Navy's Coupled Ocean–Atmosphere Mesoscale Prediction System for Tropical Cyclones (COAMPS-TC, Doyle et al. 2014) ensemble system (Komaromi et al. 2021) and the Coupled Hurricane Intensity Prediction ensemble (Emanuel 2023). The mesoscale deterministic forecasts used for forecasting RI come from COAMPS-TC, run with GFS boundary conditions, and from the Hurricane Weather Research and Forecast (HWRF) Model (Biswas et al. 2018).

In this work, we employ an approach similar to DTOPS, but specifically to forecast rapid intensity changes for JTWC's AOR with forecasts from models routinely available to JTWC forecasters. Work developing RIDE started in late 2019, and RIDE was transitioned to JTWC's operations in July 2021. So this note provides documentation of this now operational capability. We describe the input models, model development, and deterministic forecast constructs in section 2. In section 3, we provide independent results based on the 2021 TC season in JTWC's AOR. Finally, we summarize and discuss future work in section 4.

2. Data and methods

a. Deterministic model descriptions

RIDE uses intensity forecast output from seven deterministic models that were routinely available (and useful) from 2018 through 2020: one statistical model, two global numerical weather prediction (NWP) models, two mesoscale NWP models, and two statistical–dynamical models.

The one statistical model, the Trajectory Climatology and Persistence Model (TCLP), blends current location and motion with trajectories of past TCs and current intensity and intensity change with a logistic growth intensity model (DeMaria 2009) based on climatological growth rates and SST-based maximum intensity estimates (DeMaria et al. 2021). TCLP provides a 7-day forecast skill baseline for both track and intensity and an independent climatological approach to forecast intensity given current intensity, location, and trends for all the JTWC basins. See Sampson et al. (2008) for a discussion on the importance of independence in a consensus aid.

The global models include the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014) and NOAA Global Forecast System (GFS 2021), both of which produce TC intensity and track forecasts via a TC vortex tracker (Marchok 2002, 2021). Because this tracker output is available late (generally available six hours after the model analysis time), it is postprocessed or "interpolated" to the current forecast cycle using current conditions (Goerss and Sampson 2014). To delineate the postprocessed interpolated model forecasts, the Automated Tropical Cyclone Forecast System (ATCF; Sampson and Schrader 2000) changes the fourth character of the "early model" technique name to "I." So the interpolated Navy Global Environmental Model is NVGI (the late model tracker output is NVGM), and the interpolated GFS model is AVNI (the late model tracker output is AVNO).

The two mesoscale model tracks used in RIDE are from COAMPS-TC (GFS boundary conditions) and from HWRF (Biswas et al. 2018). As is the case for the global models, these mesoscale models produce time late intensity forecasts via the TC vortex tracker and so their ATCF aids end with "I" (CTCX becomes CTCI and HWRF becomes HWFI).

RIDE also uses statistical–dynamical models including the Statistical Hurricane Prediction Scheme (DSHA; DeMaria et al. 2005) and the Logistic Growth Model (LGEA; DeMaria 2009)—ported to operations at JTWC in 2015. The models are similar in construction to those run at National Hurricane Center; however, the model coefficients are AOR specific and receive periodic updates (most recently in 2020). In operations, the statistical–dynamical models run prior to the issuance of a forecast. The statistical–dynamical models environmental predictors require input derived from GFS 6- to 126-h forecasts for the 0–120-h lead times along the GFS forecast track, and oceanic heat content along the GFS forecast track from the Navy Coupled Ocean Data Assimilation System (Cummings 2005) and described in Sampson et al. (2022). In addition to the environment, digital infrared brightness temperature input from the global constellation of geostationary satellites provides convective information at the initial time. Finally, the statistical–dynamical models require ATCF TC track and intensity data, namely, JTWC's official track and 12-h track and intensity histories.

The seven model forecasts used in RIDE all have their weaknesses but are typically available between 77% and 81% of the forecast periods for which JTWC issued warnings during the 2018–20 seasons for a total of 2719, 2556, 2392, 2143, and 1758 cases at 12-, 24-, 36-, 48-, and 72-h forecast leads, respectively. High availability is important for RIDE, and these models are among the most reliably available at JTWC. Figure 1a shows intensity skill (vs TCLP) through 72 h, and Fig. 1b shows Peirce skill scores (PSS) associated with the intensification rates RI20, RI30, RI45, RI56, and RI65 for the 2018–20 TC season in JTWC's AOR. Skill in Fig. 1a is in terms of percent improvement versus TCLP. PSS (Fig. 1b) answers the question "what is the accuracy of the forecast in predicting the correct category relative to that of random chance?" and is a good metric for probabilistic forecasts of rare events (Manzato 2007). TCLP provides a skill baseline, but also demonstrates a limited ability to predict RI events (i.e., positive PSSs) given that RI events likely involve cases where the climatological conditions are favorable. Global models like NVGI and AVNI generally lack the horizontal resolution to resolve a TC's inner core and its intensity; however, AVNI does demonstrate skill versus TCLP. NVGI and AVNI also have limited skill in predicting RI with respect to PSS, but skill versus TCLP of both increases through 48 h. CTCI and HWFI have considerable skill in predicting

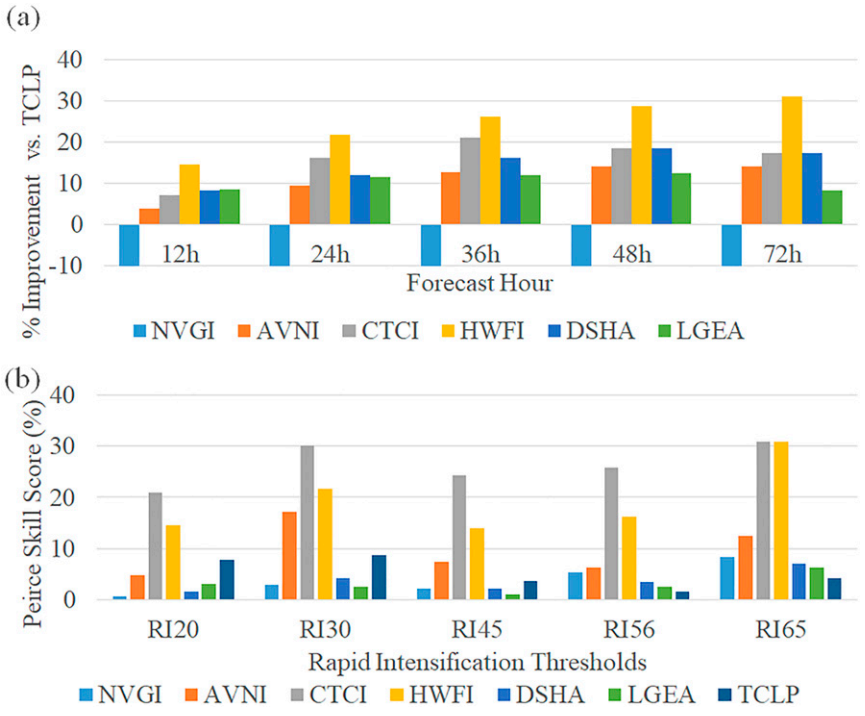

FIG. 1. (a) The skill or percent improvement versus TCLP in forecasting intensity change through 72 h for NVGI, AVNI, CTCI, HWFI, DSHA, LGEA and (b) the Peirce skill score for the rapid intensification thresholds RI20, RI30, RI45, RI56, and RI65 (see text for descriptions) for these same models plus TCLP. Models appear from left to right as they are listed in the legend. Results are based on 2018–20 seasons.

intensity and RI events, which suggests that these may be the most useful members in RIDE. Finally, the statistical–dynamic models (DSHA and LGEA) generally have skill in forecasting intensity changes but demonstrate little skill in predicting RI events.

The seven deterministic forecast models have different abilities for predicting both intensity change and rapid intensification thresholds. The forecasting dilemma is how to best use this disparate guidance to anticipate rapid intensification events. Here, we will use a method called logistic regression to develop the RIDE model that provides forecasts of probabilities associated with rapid intensification measured by several thresholds of intensity change.

b. *Model development and methodology*

As mentioned in the introduction, DTOPS inspired the development of RIDE. That being said, the input models are different, and the treatment of the input data is also different. The most common theme is the use of logistic regression to estimate probabilities. Logistic regression uses a logit function to model a binary dependent variable (p for probability) shown in (1), where β_x are constants and x_x are the predictors:

$$\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots. \tag{1}$$

The function for the odds ratio is given by (2):

$$\left(\frac{p}{1-p}\right) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots). \tag{2}$$

And, via rearrangement and substitution, p is estimated by (3):

$$p = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots)}. \tag{3}$$

For fitting RIDE, we use the FORTRAN 90 logistic regression code (Miller 2002; CSIRO 2022) that produces linear logistic models by iteratively reweighted least squares. A model using β_0 serves as a baseline for determining the amount of deviance explained, which is similar to the amount of variance explained by a linear regression model.

Instead of discussing the differences between DTOPS and RIDE, we now describe the components used to create the predictors for the logistic regression, dependent model performance, and information about predictor contributions to the probabilistic forecast. The first couple of predictors involve initial conditions. These are the cosine of latitude, denoted as $\cos(\varphi)$, and the current/initial intensity capped at 75 kt (V_{mc}). For instance, if the initial intensity is 50 kt, 50 kt is used in RIDE, but if the initial intensity is 90 kt, 75 kt is used in RIDE. The use of $\cos(\varphi)$ follows from DTOPS development and provides an estimate of the gradient of the Coriolis term, and is thought to be related to lower latitudes being characterized by

smaller TCs in JTWC's AOR (see Knaff et al. 2014), and generally more favorable and persistent environments. Knaff et al. (2020) discuss capping the initial intensity at 75 kt in statistical–dynamical models, specifically the Rapid Intensification Prediction Aid (RIPA). The initial intensity is largely a negative factor to rapid intensity changes when storms are weak (i.e., ≤45 kt), but above that intensity, TCs are more likely to undergo RI. At intensities above 75 kt, an eyewall is usually present (Vigh et al. 2012), and so long as conditions remain favorable, a TC can intensify rapidly (Malkus 1958; Yanai 1961; Mundell 1990; Willoughby 1990). The use of V_{mc} versus the initial TC intensity also produces an improved model (not shown) and without adding quadratic predictors. The remaining predictors are the deterministic forecasts coming from the seven models (i.e., NVGI, AVNI, CTCI, HWFI, DSHA, LGEA, and TCLP), and the number of these models that made a forecast exceeding 1.24 standardized deviations of their mean intensity change, which corresponds to the 90th percentile of intensity increase. In summary there are a total of ten predictors: $\cos(\varphi)$, V_{mc}, seven deterministic model intensity forecasts, and the number of intensity increase forecasts exceeding their 90th percentile.

Goodness of fit for logistic regression is measured by deviance, which is related to distance from a perfect fit. Deviance ranges from zero to infinity where zero indicates a perfect fit. Deviance from our model alone is difficult to interpret, so this deviance is compared with deviance from a baseline. The simplest baseline is the model without any independent predictors, which we prescribe by using only β_0 to solve Eq. (1). Figure 2 shows the dependent percent deviance explained for each RI threshold, which is remarkably constant near 40%, suggesting that the model reduces deviance by 40% over the baseline.

The normalized predictor coefficients in logistic regression have a similar interpretation as for multiple linear regression with positive and negative coefficients leading to higher and lower probabilities, respectively. One difference is that these coefficients are estimating the nonlinear, S-shaped logit [Eq. (1)]. Figure 3 provides the normalized logistic regression coefficients that provide a measure of how important each normalized predictor is to variations of the logit. Normalization also accounts for each predictor's sample mean and variance, meaning biases in both mean and magnitude are captured implicitly. From this analysis, it is evident that initial conditions related to V_{mc} and $\cos(\varphi)$ provide a strong positive contribution to the forecasting of RI probabilities. This suggests that intensities approaching or greater than 75 kt are favorable (note the mean is ~46 kt) as are lower latitude TCs. Predictors based on model forecasts, on the other hand, have more variable weights with some contributing heavily and others contributing little. Global models appear to contribute either trivially (AVNI) or negatively (NGVI). This result makes some sense as AVNI provides input or boundary conditions to CTCI, HWFI, DSHA, and LGEA. NVGI is largely independent in this respect and only shows nonnegative coefficients for the RI65 forecasts—a result that mirrors the PSS shown in Fig. 1. It is also shown that CTCI and HWFI provide the next most impactful and positive contributions, again mirroring information presented in Fig. 1. DSHA and LGEA also contribute positively, but somewhat surprisingly based on the verification statistics shown in Fig. 1, LGEA is more impactful at the shorter leads, which is not supported by the statistics in Fig. 1. This result suggests that the logistic regression is overcoming LGEA and DSHA negative RI biases (not shown) contributing to the errors. Also, TCLP primarily contributes to the shortest (RI20) and the longest forecasts (RI56 and RI65). For RI20, this may indicate TCLP handles the 12-h changes in intensity well or that TCLP, DSHA, and LGEA are no better than this baseline model. Noting that TCLP, DSHA, and LGEA all use the most recent 12-h intensity (0 h minus 12 h) change as a predictor. Finally, the number of models exceeding their 90% of intensity change (N90%) is a rather weak contributor, suggesting that the number of models rapidly intensifying a given TC provides limited additional information. This is contrary to consensus findings in Sampson et al. (2008) and suggests that RI timing may be an issue. Finally, we offer reliability diagrams for the dependent model fits in Fig. 4. As is often the case with statistical RI models, these suggest a general high bias for all forecasts except RI65, but rather good reliability overall.

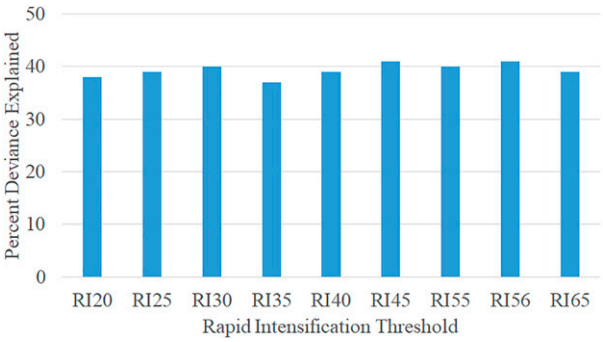

FIG. 2. The percent deviance explained in models for the RI thresholds forecast by RIDE. Results are based on the developmental data from 2018 to 2020.

c. *Creation of deterministic forecasts*

Having probabilistic forecasts of the rare RI events is useful, but deterministic forecasts are required for the operational forecasts. To help forecasters in their intensity forecast decision, RIDE transforms probabilistic forecasts into deterministic forecasts by using a 40% probability threshold trigger as discussed in Sampson et al. (2011). As discussed by Knaff et al. (2018, 2020), most statistical–dynamical probabilistic RI models use a 40% probability threshold to trigger deterministic guidance. When multiple RI thresholds are triggered, RIDE uses the maximum intensity change along the track to make a continuous albeit short deterministic forecast, which is available to add to the intensity consensus and for forecasters to view. For instance, if RI20, RI25, and RI45 all forecast probabilities of RI greater than 40%, the intensity change forecast would be +20, +30, and +45 kt at 12, 24, and 36 h, respectively.

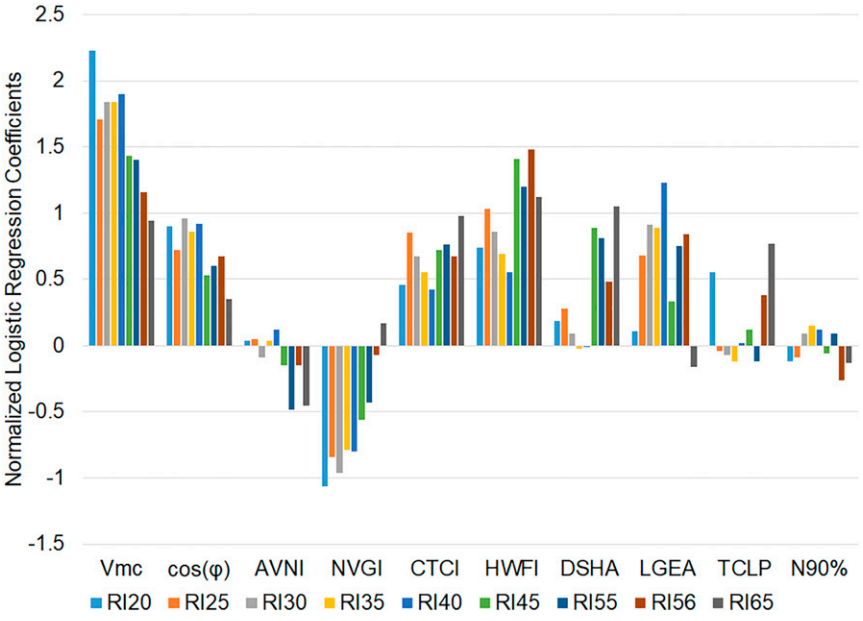

FIG. 3. Normalized weights for predictors for the logistic regression models developed to predict RI20, RI25, RI30, RI35, RI40, RI45, RI55, RI56, RI70, and RI65. Normalization is accomplished by dividing by the standard deviation. Results are based on the developmental data from 2018 to 2020.

3. Results and performance

As mentioned in the introduction, RIDE was transitioned to JTWC's operations in July 2021. Prior to July 2021 RIDE was run on the operationally available guidance to provide an independent verification dataset for the 2021 season in JTWC's AOR. We first examine the probabilistic verification of RIDE. Figure 5 shows reliability diagrams based on these independent forecasts. The relationship between the predicted probabilities and the observed frequencies are generally reliable (along the diagonal). The notable exception is RI20, which did not produce probabilities larger than 28% in our independent sample. These diagrams also show slight low biases for RI20, RI25, RI30, RI35, and RI40, and slight high biases for RI55, RI56, and RI70, which are different than shown in the developmental data.

A more complete picture of the probabilistic verifications can be shown using a performance diagram (Roebber 2009). The performance diagram shows several metrics typically used to verify probabilistic forecasts. In these diagrams, the false alarm ratio, or more specifically the inverse called the success ratio (SR), is the abscissa, and the hit rate or probability of detection (POD) is the ordinate. The multiplicative bias and critical success index (CSI), or threat score are also shown as positively tilted dashed and curved colored lines, respectively. SR, POD, bias, and CSI answer the following questions:

1) SR: What fraction of the forecast "yes" events were correctly observed?
2) POD: What fraction of the observed "yes" events were correctly forecasted?
3) Bias: How did the forecast frequency of "yes" events compare to the observed frequency of "yes" events?
4) CSI: How well did the forecast "yes" events correspond to the observed "yes" events?

An unbiased and highly skilled forecast would appear in the upper right corner of the performance diagram, a desirable trait but one not likely achievable due to observation error.

Figure 6 shows the performance diagram visualization of independent RIDE forecasts. Note that since none of the RI20 forecasts exceeded 40%, there are no forecasts for RI20 to display in the diagram. All other forecasts appear low biased with longer lead forecasts, RI65, RI56, and RI45, showing the most skill in forecasting RI events. The results are somewhat surprising in that the fraction of correct forecasts is quite good, but the frequency of observed versus forecast frequency is low. This is particularly true for the 24-h forecast lead times RI thresholds. For the largest RI thresholds RI40, RI55, and RI70, RIDE has limited correspondence with the occurrence of these rarest events. From a forecasting standpoint, RIDE forecasts tend to be correct when they are forecasted but are relatively infrequent. Finally, Brier skill scores (not shown), which are hampered by small samples of rare events, indicate all the forecasts have skill above their climatological values.

We next provide verification of the deterministic forecasts, again triggered at 40%, against the final best track estimates of intensity changes. Figure 7 shows homogeneous mean errors, biases, and PSS based on all events for RIDE and an intensity consensus without specialized RI aids (ICNC = HWFI + CTCI + DSHA + DSHN + AVNI), note DSHN is SHIPS run with NVGM-based environmental conditions. This independent verification indicates that the mean forecast errors are similar between the two, but that RIDE biases are nearer to zero. RIDE maintains positive and near zero biases while most

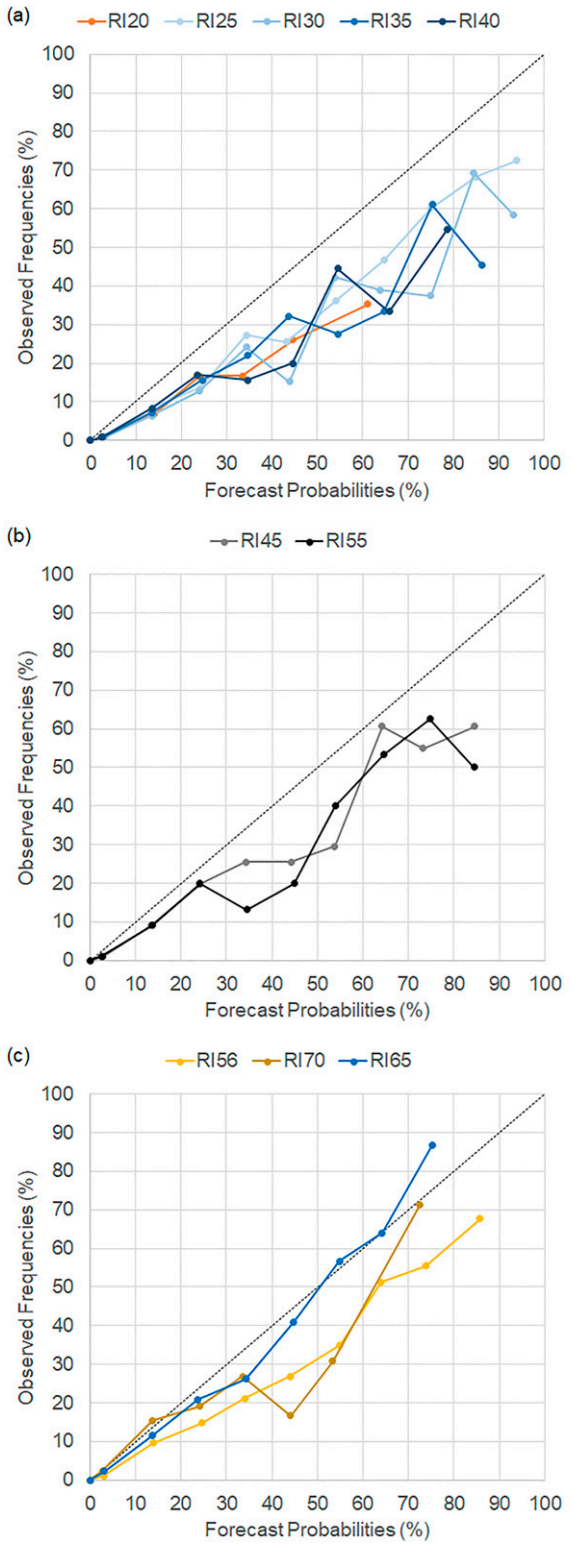

FIG. 4. Reliability diagrams for (a) RI20, RI25, RI30, RI35, and RI40; (b) RI45 and RI55; and (c) RI56, RI70, and RI65. Results are based on the developmental data from 2018 to 2020.

other models are noticeably negative. PSSs are higher for RIDE in the 24-, 36-, and 48-h lead times with both CTCI and HWFI performing better at 72-h forecasts. These statistics indicate that RIDE is providing a less negatively biased forecast for RI cases, although homogeneous errors among these methods are similar. The PSSs indicate RIDE's guidance on guidance provides superior forecasts of RI events in the 24–48-h lead times. A closer examination of contingency tables, not shown, indicates that RIDE is conservative in triggering its deterministic RI aid, which limits false alarms and is different from other RI aids such as RIPA. These results echo those of Fig. 6.

Concerning the 40% trigger probability, we conduct a simple analysis where the median intensity change is calculated for probabilistic forecasts $< 40\%$ and $\geq 40\%$ using the 2021 independent forecast cases. There are 1594, 1476, 1358, and 1241 forecasts made at 24, 36, 48, and 72 h, respectively. Note again that no probabilities were forecast larger than 40% for the RI20 threshold. These intensity change analyses are shown in Fig. 8. Figure 8a shows the 24-h forecast thresholds that confirm that 40% is a reasonable threshold for triggering simple forecasts. RI25 and RI30 appear to overforecast RI events, which is largely due to the fact that in many instances RI25, RI30, and RI35 are all triggered. When this occurs, the deterministic forecast constructed is a +17.5- and +35-kt intensity change at 12- and 24-h lead times, respectively. The few cases in Fig. 8b shows the median outcomes of RI45 and RI55 36-h deterministic forecasts to be 45 and 50 kt, respectively. Similarly, the median 48-h intensity change for RI56 and RI70 is less than 55 and 70 kt, respectively (Fig. 8c). However, for RI65 (Fig. 8c), the choice of 40% seems appropriate in this limited sample. Note mean values of intensity change, not shown, are nearly identical. Overall, and given the small number of TCs that have intensity changes greater than 55 kt within 48 h, the 40% trigger probability for deterministic forecasts seems to work reasonably well.

4. Summary and discussion of future work

We outline the method used to develop the Rapid Intensification Deterministic Ensemble model (RIDE) model. RIDE uses routinely available TC intensity guidance in the ATCF to compute and provide both probabilities and deterministic forecasts for RI events. RIDE input models include one purely statistical model (TCLP), two interpolated global models (AVNI, NVGI), two interpolated mesoscale hurricane models (HWFI, CTCI), and two statistical–dynamical models based on GFS tracks and dynamic fields (DSHA and LGEA). As inputs to RIDE, intensity changes from these aids, TC current intensity, latitude, and the number of input models exceeding their individual 90th percentile of intensity change comprise the predictors in a logistic regression equation that predict the probability of rapid intensity change. Intensity change thresholds of 20 kt in 12 h, 25, 30, 35, and 40 kt in 24 h, 45 and 55 kt in 36 h, 55 kt in 48 h, and 65 kt in 72 h are all derived using the logistic regression. Then, using the 40% or greater trigger, RIDE generates simple deterministic intensity forecasts that can be used in operational intensity consensus aids.

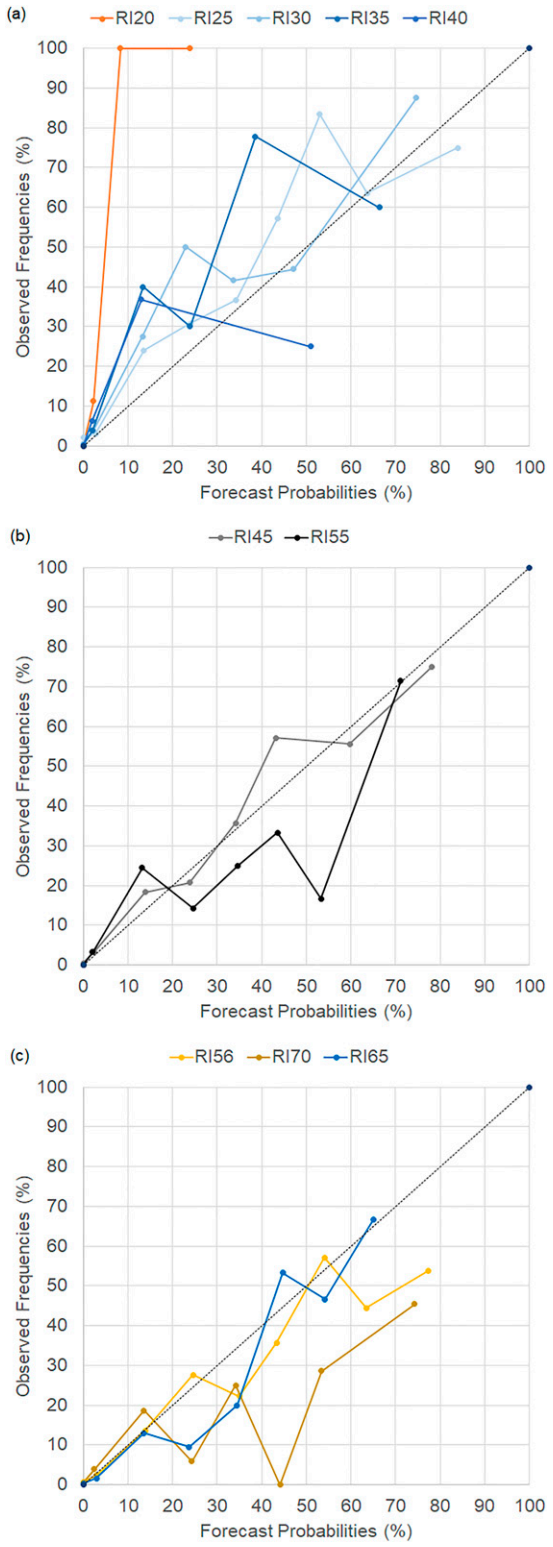

FIG. 5. Reliability diagrams associated with independent RIDE forecasts made during the 2021 season in JTWC's AOR. Shown are (a) RI20, RI25, RI30, RI35, and RI40; (b) RI45 and RI55; and (c) RI56, RI70, and RI64.

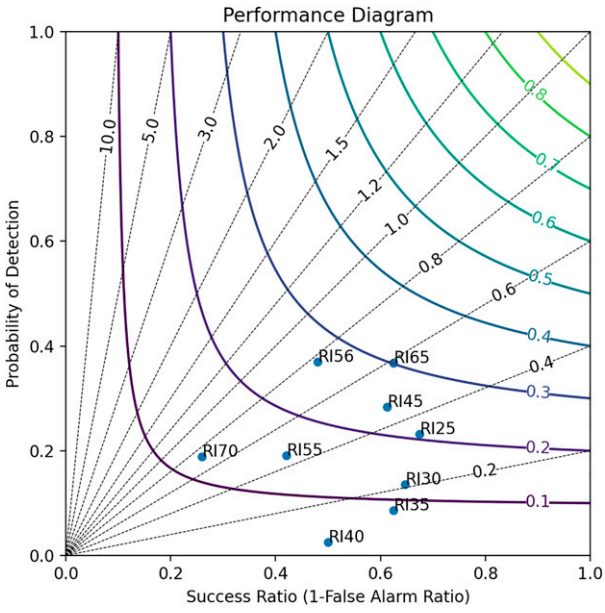

FIG. 6. Performance diagram for RIDE deterministic forecasts made in JTWC's AOR during the 2021 season. Shown are RI25, RI30, RI35, RI40, RI45, RI55, RI56, RI70, and RI65. RI20 is excluded as no probabilistic forecasts exceeded 40% probability.

Developmental data presented shows that these models are well calibrated and fit to the input data. Generally, the most important inputs, based on normalized logistic regression coefficients are the current intensity capped at 75 kt, and intensity changes in HWFI, CTCI, DSHA, and LGEA. It is important to emphasize that the importance of the model-based predictors is with respect to the logit function, and does not imply similar or any model weighting is appropriate for linear combinations, i.e., an intensity consensus.

In using the JTWC (2021) season for verification, there are worthy findings. First, using 40% to trigger a deterministic forecast suggests that RIDE has good reliability and discrimination. And, for the independent forecasts examined, the 40% trigger probabilities capture large observed intensity changes (Fig. 8). However, those forecasts also indicate that RIDE tends to underpredict RI events, has relatively high critical success indices, and exhibits a low false alarm ratio. This suggests that when RIDE forecasts RI events, it is often correct to do so. Second, RIDE has little skill in 12-h RI forecasting but is skillful between 24 and 48 h. Third, RIDE's deterministic forecasts had near zero biases. These results suggest that RIDE deterministic forecasts are likely useful in operations (e.g., when you see a RIDE deterministic forecast, pay attention), and a suitable candidate for addition to intensity consensus aids that are typically low biased during RI events.

This paper describes the first effort to employ methods used in DTOPS to the JTWC AOR and likely requires further development. For example, the AVNI aid is now exhibiting the capability to predict RI on occasion. The development set in RIDE likely does not capture the increased capability of GFS or any of our other NWP models, so these likely

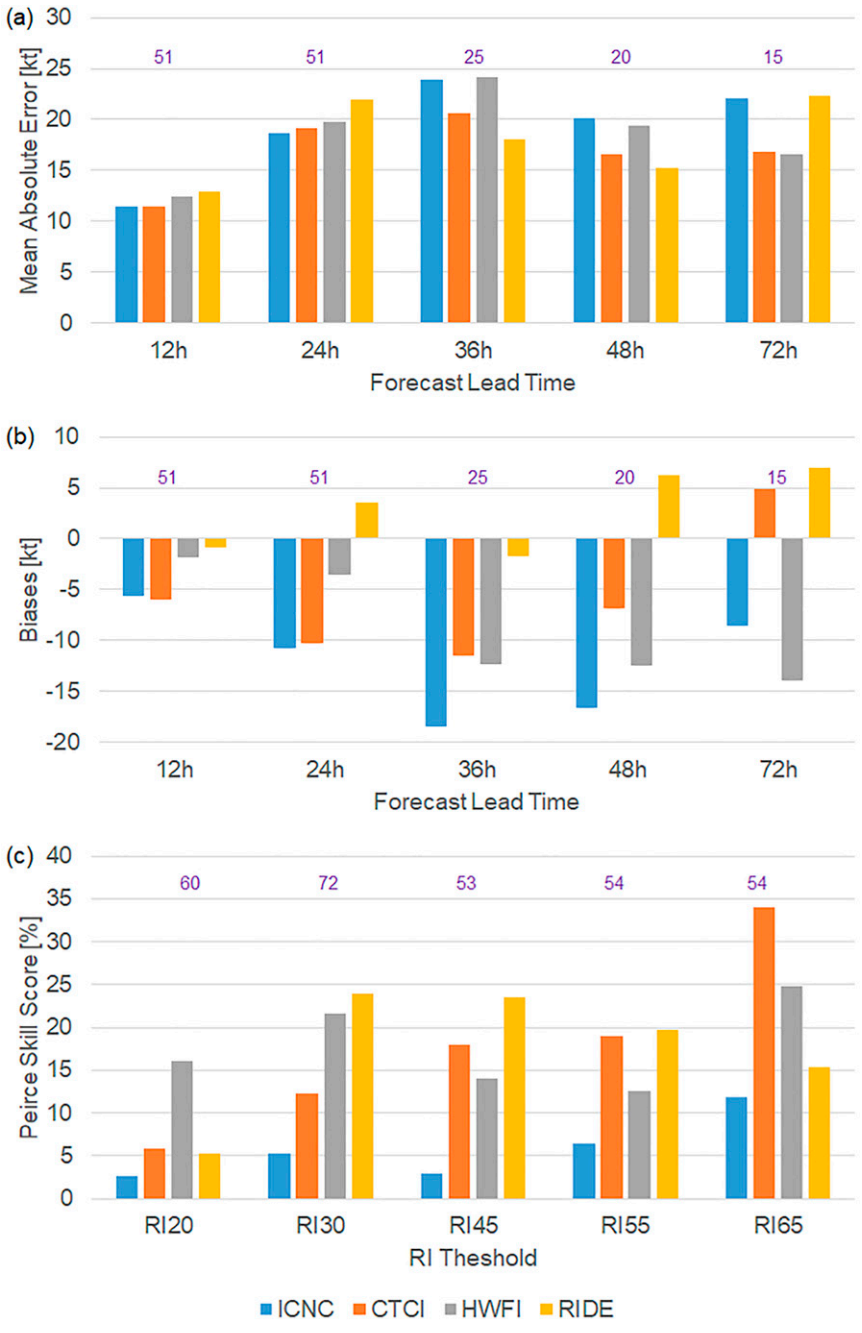

FIG. 7. JTWC AOR 2021 independent data (a) errors and (b) biases of the intensity consensus (ICNC), early COAMPS-TC (CTCI), early HWRF (HWFI), and RIDE forecasts; and (c) PSSs associated with RI20, RI30, RI45, RI55, and RI65 from these same guidance models. The number of cases used to create these statistics is shown at the top of each panel.

should be revisited annually. The objective of this work is to improve these tools so that quality RI guidance is available to JTWC forecasters in operations. And, while RIDE targets a need at JTWC, documentation and development of tools of this type are also relevant to other TC forecasters working in basins served by JTWC. Aids like RIPA and the new Forest Rapid Intensification Aid (FRIA; Slocum 2021) complement the effort in this paper, and between the difficulty and limited time to make RI forecasts, these efforts are becoming increasingly important.

Acknowledgments. The authors would like to acknowledge that much of the funding for this work was provided by the Office of Naval Research under a grant to CIRA

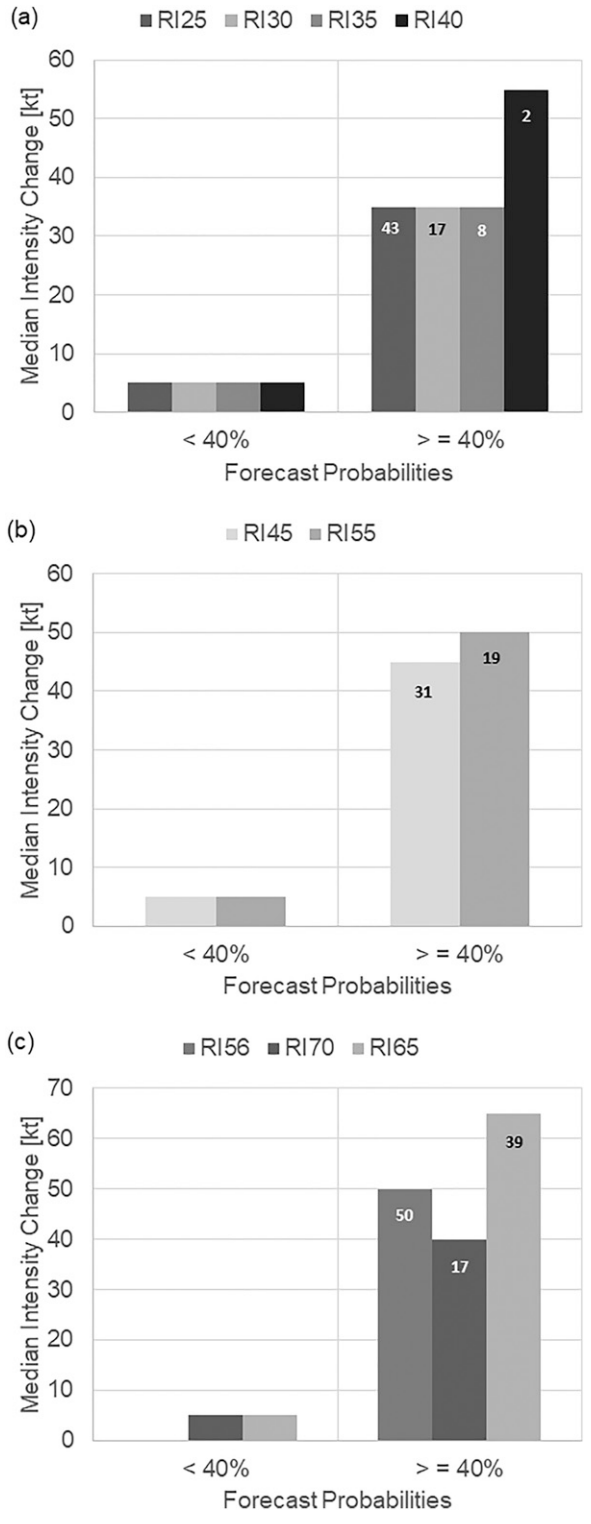

FIG. 8. Median intensity changes associated with RI forecasts less than 40% and greater than 40% for each of the RI thresholds based on the JTWC AOR 2021 season independent sample. Shown are (a) RI25, RI30, RI35, and RI40; (b) RI45 and RI55; and (c) RI56, RI70, and RI65. The number of cases exceeding 40% is shown on the individual bars.

(Grant N00173-21-1-G008) and ONR Award N0001420WX00517 to NRL Monterey. We would also like to thank Jonathan Martinez and Marie McGraw for reviewing this manuscript before submission, and acknowledge feedback we have received from Steve Barlow and Levi Cowan at JTWC. Finally, we thank Joe Courtney and two anonymous reviewers for their constructive reviews. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

Data availability statement. The best track data used in this study are freely available from the DoD's Joint Typhoon Warning Center (https://www.metoc.navy.mil/jtwc/jtwc.html?best-tracks). Forecasts used are in the Joint Typhoon Warning Center's forecast database (a-deck) and are available upon request at https://pzal.metoc.navy.mil/php/rds/login.php.

REFERENCES

Biswas, M. K., and Coauthors, 2018: Hurricane Weather Research and Forecasting (HWRF) model: 2018 scientific documentation. Developmental Testbed Center Doc., 112 pp., https://dtcenter.org/sites/default/files/community-code/hwrf/docs/scientific_documents/HWRFv4.0a_ScientificDoc.pdf.

CSIRO, 2022: Software from Alan J. Miller. Commonwealth Scientific and Industrial Research Organisation, accessed 1 June 2022, https://wp.csiro.au/alanmiller/.

Cummings, J. A., 2005: Operational multivariate ocean data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3583–3604, https://doi.org/10.1256/qj.05.105.

DeMaria, M., 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.*, **137**, 68–82, https://doi.org/10.1175/2008MWR2513.1.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvement to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, https://doi.org/10.1175/WAF862.1.

——, J. L. Franklin, M. J. Onderlinde, and J. Kaplan, 2021: Operational forecasting of tropical cyclone rapid intensification at the National Hurricane Center. *Atmosphere*, **12**, 683, https://doi.org/10.3390/atmos12060683.

Doyle, J. D., and Coauthors, 2014: Tropical cyclone prediction using COAMPS-TC. *Oceanography*, **27**, 104–115, https://doi.org/10.5670/oceanog.2014.72.

Emanuel, K., 2023: The Coupled Hurricane Intensity Prediction System (CHIPS). MIT, 5 pp., https://wind.mit.edu/~emanuel/CHIPS.pdf.

Francis, A. S., and B. R. Strahl, 2022: Joint Typhoon Warning Center annual tropical cyclone report 2020. JTWC Tech. Rep., 146 pp., https://www.metoc.navy.mil/jtwc/products/atcr/2020atcr.pdf.

GFS, 2021: GFS Global Forecast System. NCEP and NOAA, accessed 1 April 2022, https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/gfs.php.

Goerss, J. S., and C. R. Sampson, 2014: Prediction of consensus tropical cyclone intensity forecast error. *Wea. Forecasting*, **29**, 750–762, https://doi.org/10.1175/WAF-D-13-00058.1.

Hogan, T. F., and Coauthors, 2014: The Navy global environmental model. *Oceanography*, **27**, 116–125, https://doi.org/10.5670/oceanog.2014.73.

JTWC, 2021: JTWC 2020 operational highlights, challenges, and future changes. *75th Interdepartmental Hurricane Conf.*, Miami, FL, JTWC, 16 pp., https://www.icams-portal.gov/meetings/TCORF/tcorf21/01-Session/s1-03_jtwc.pdf.

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

——, C. R. Sampson, and K. D. Musgrave, 2018: An operational rapid intensification prediction aid for the western North Pacific. *Wea. Forecasting*, **33**, 799–811, https://doi.org/10.1175/WAF-D-18-0012.1.

——, ——, and B. R. Strahl, 2020: A tropical cyclone rapid intensification prediction aid for the Joint Typhoon Warning Center's areas of responsibility. *Wea. Forecasting*, **35**, 1173–1185, https://doi.org/10.1175/WAF-D-19-0228.1.

Komaromi, W. A., P. A. Reinecke, J. D. Doyle, and J. R. Moskaitis, 2021: The Naval Research Laboratory's Coupled Ocean–Atmosphere Mesoscale Prediction System-Tropical Cyclone Ensemble (COAMPS-TC ensemble). *Wea. Forecasting*, **36**, 499–517, https://doi.org/10.1175/WAF-D-20-0038.1.

Lee, C.-Y., M. K. Tippett, A. H. Sobel, and S. J. Camargo, 2016: Rapid intensification and the bimodal distribution of tropical cyclone intensity. *Nat. Commun.*, **7**, 10625, https://doi.org/10.1038/ncomms10625.

Malkus, J. S., 1958: Tropical weather disturbances—Why do so few become hurricanes. *Weather*, **13**, 75–89, https://doi.org/10.1002/j.1477-8696.1958.tb02330.x.

Manzato, A., 2007: A note on the maximum Peirce skill score. *Wea. Forecasting*, **22**, 1148–1154, https://doi.org/10.1175/WAF1041.1.

Marchok, T. P., 2002: How the NCEP tropical cyclone tracker works. *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P1.13, https://ams.confex.com/ams/25HURR/webprogram/Paper37628.html.

——, 2021: Important factors in the tracking of tropical cyclones in operational models. *J. Appl. Meteor. Climatol.*, **60**, 1265–1284, https://doi.org/10.1175/JAMC-D-20-0175.1.

Mautner, D. A., and C. P. Guard, 1993: 1992 annual tropical cyclone report. Joint Typhoon Warning Center Rep., 280 pp., https://www.metoc.navy.mil/jtwc/products/atcr/1992atcr.pdf.

Miller, A., 2002: *Subset Selection in Regression*. 2nd ed. CRC Press, 256 pp.

Mundell, D. B., 1990: Prediction of tropical cyclone rapid intensification. M.S. thesis, Dept. of Atmospheric Science, Colorado State University, 186 pp., https://apps.dtic.mil/sti/pdfs/ADA227668.pdf.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/2008WAF2222159.1.

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.

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

——, J. Cummings, J. A. Knaff, M. DeMaria, and E. A. Serra, 2022: An upper ocean thermal field metrics dataset. *Meteorology*, **1**, 327–340, https://doi.org/10.3390/meteorology1030021.

Slocum, C. J., 2021: What can we learn from random forest in the context of the tropical cyclone rapid intensification problem? *Second NOAA Workshop on leveraging AI in Environmental Sciences*, online, NOAA, https://www.star.nesdis.noaa.gov/star/documents/meetings/2020AI/presentations/202101/20210128_Slocum.pptx.

Vigh, J. L., J. A. Knaff, and W. H. Schubert, 2012: A climatology of hurricane eye formation. *Mon. Wea. Rev.*, **140**, 1405–1426, https://doi.org/10.1175/MWR-D-11-00108.1.

Willoughby, H. E., 1990: Temporal changes of the primary circulation in tropical cyclones. *J. Atmos. Sci.*, **47**, 242–264, https://doi.org/10.1175/1520-0469(1990)047<0242:TCOTPC>2.0.CO;2.

Yanai, M., 1961: A detailed analysis of typhoon formation. *J. Meteor. Soc. Japan*, **39**, 187–214, https://doi.org/10.2151/jmsj1923.39.4_187.

Zhang, Z., and Coauthors, 2022: Intensity changes: Operational perspectives. *10th WMO Int. Workshop on Tropical Cyclones (IWTC-10)*, Bali, Indonesia, World Meteorological Organization, 2.3, https://community.wmo.int/en/iwtc-10-reports.