NCEP NOTES

Tropical Cyclone Gale Wind Radii Estimates, Forecasts, and Error Forecasts for the Western North Pacific

CHARLES R. SAMPSON,[a] JAMES S. GOERSS,[b] JOHN A. KNAFF,[c] BRIAN R. STRAHL,[d] EDWARD M. FUKADA,[e] AND EFREN A. SERRA[f]

[a] *Naval Research Laboratory, Monterey, California*
[b] *SAIC, Monterey, California*
[c] *NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado*
[d] *JTWC, Pearl Harbor, Hawaii*
[e] *General Dynamics Institute of Technology, Monterey California*
[f] *DeVine Consulting, Monterey, California*

(Manuscript received 17 October 2017, in final form 8 March 2018)

ABSTRACT

In 2016, the Joint Typhoon Warning Center extended forecasts of gale-force and other wind radii to 5 days. That effort and a thrust to perform postseason analysis of gale-force wind radii for the ''best tracks'' (the quality controlled and documented tropical cyclone track and intensity estimates released after the season) have prompted requirements for new guidance to address the challenges of both. At the same time, operational tools to estimate and predict wind radii continue to evolve, now forming a quality suite of gale-force wind radii analysis and forecasting tools. This work provides an update to real-time estimates of gale-force wind radii (a mean/consensus of gale-force individual wind radii estimates) that includes objective scatterometer-derived estimates. The work also addresses operational gale-force wind radii forecasting in that it provides an update to a gale-force wind radii forecast consensus, which now includes gale-force wind radii forecast error estimates to accompany the gale-force wind radii forecasts. The gale-force wind radii forecast error estimates are computed using predictors readily available in real time (e.g., consensus spread, initial size, and forecast intensity) so that operational reliability and timeliness can be ensured. These updates were all implemented in operations at the Joint Typhoon Warning Center by January 2018, and more updates should be expected in the coming years as new and improved guidance becomes available.

1. Introduction

Forecasting tropical cyclone (TC; see Table 1 for this and other acronyms used in this paper) surface wind structure has been one of the challenges of the forecast process at the Joint Typhoon Warning Center (JTWC). Wind structure analyses and forecasts are provided in terms of the ''wind radii''.[1] Wind radii are defined as the maximum extent of 34- (R34), 50-, and 64-kt winds in the four compass quadrants (northeast, southeast, southwest, and northwest) surrounding the TC. In operations, an intense TC can require up to 12 estimates (four quadrants for each of the three radii thresholds) for an analysis (0 h) and then 12 estimates for each of the seven forecast periods (12, 24, 36, 48, 72, 96, and

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

Corresponding author: Buck Sampson, buck.sampson@nrlmry.navy.mil

[1] The operational units for wind radii distance and wind speeds at JTWC are specified in nautical miles (1 n mi = 1.85 km) and knots (1kt = 0.514 m s^{-1}), respectively, so these units will be used for the remainder of this paper.

DOI: 10.1175/WAF-D-17-0153.1

TABLE 1. List of acronyms used in this paper.

Acronym	Definition
AMSU	Advanced Microwave Sounding Unit
ASCII	American Standard Code for Information Exchange
ASCT	Objective R34 scatterometer fixes
ATCF	Automated Tropical Cyclone Forecast System
AVNO/AHNI	Global Forecast System (GFS) model radii analyses/interpolated forecasts
CHTI	COAMPS-TC model radii/interpolated forecast
CIRW	Cooperative Institute for Research in the Atmosphere (CIRA) wind radii estimates
COAMPS	Coupled Ocean–Atmosphere Mesoscale Prediction System
COAMPS-TC	COAMPS Tropical Cyclone model
DSHA	Statistical–dynamical wind radii forecasts based on GFS model data
DVRK	Dvorak wind radii
EMXI	European Centre for Medium-Range Weather Forecasts model R34 interpolated forecast
GFDL	Geophysical Fluid Dynamics Laboratory
GFDT/GFTI	GFDL model radii/interpolated forecast
GPCE	Goerss-predicted consensus error
HWRF/HHNI	Hurricane Weather Research and Forecasting Model radii/interpolated forecast
INTF	Official intensity forecast (kt), 6 h old interpolated or consensus
JTWC	Joint Typhoon Warning Center
JTWI	JTWC interpolated forecast
LATC	Forecast latitude change (°)
MetOp	Meteorological Operational satellite
NHC	The National Hurricane Center
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical weather prediction
NRL	Naval Research Laboratory
OBTK	Objective R34, an equally weighted average of R34 estimates
RVCN	R34 forecast consensus
RVCN 2014	Original version of RVCN = AHNI + HHNI + GFTI + EMXI
RVCN 2017	New version of RVCN = AHNI + HHNI + EMXI + CHTI + DSHA
R34	Radii of 34-kt winds, also known as gale-force wind radii
SHIPS	Statistical Hurricane Intensity Prediction System
SK15	Sampson and Knaff (2015)
SPRD	Average spread of nonzero R34 estimates (n mi)
TC	Tropical cyclone
WRE	Wind radius predicted error (n mi)
WRCHG	Wind radius change (n mi)

120 h) for a total of 96 wind radii. Production of such a large number of estimates in real time can become a time-consuming task. Nonetheless, wind radii estimates are important to postprocessed guidance such as wind speed probabilities (DeMaria et al. 2013), storm surge forecasts for the National Hurricane Center (NHC 2016), wave forecasts (Sampson et al. 2010), modeling of potential infrastructure damages (e.g., Quiring et al. 2014), wind and wave/surge damage potential (Powell and Reinhold 2007), and Department of Defense danger swaths and Tropical Cyclone Conditions of Readiness (see Sampson et al. 2012). They are also used to initialize numerical weather prediction with some benefit in forecast error reduction (Tallapragada et al. 2015; Kunii 2015; Bender et al. 2017).

Postseason reanalysis of JTWC western North Pacific TC wind radii (see Sampson et al. 2017) has recently become available (JTWC 2017), and can serve as "ground truth" for developing and updating guidance, including satellite-based estimates of wind radii/surface winds (Demuth et al. 2004, 2006; Mueller et al. 2006; Kossin et al. 2007; Knaff et al. 2011; Knaff et al. 2016; Dolling et al. 2016; Meissner et al. 2017; Reul et al. 2017), the wind radii climatology and persistence model (Knaff et al. 2007), and consensus forecasts (Sampson and Knaff 2015, hereafter SK15). Errors in the best-track R34 have been estimated in the past to be as high as 10%–40% (Knaff and Harper 2010; Landsea and Franklin 2013; Knaff and Sampson 2015), depending on the quality and quantity of the available observational data.

Forecasts of R34 have been improving as well, and skill has been found in individual forecast models (e.g., Tallapragada et al. 2014; Cangialosi and Landsea 2016; Knaff et al. 2017), consensus forecasts (SK15), and forecasts from the operational centers (Knaff and Sampson 2015). As real-time forecaster estimates of R34 become

more accurate, they contribute to improvements in real-time intensity and structure guidance forecasts [Bender et al. (2017) and Knaff et al. (2017), respectively], which then contribute to official forecasts and postprocessed guidance that are so critical for emergency managers and ships at sea.

The goals of this work are to 1) document updates to an objective method that provides R34 estimates for use in real-time and postseason analyses with comments on the observation error, 2) describe a forecast consensus in use at the JTWC, and 3) describe a new real-time estimate of consensus forecast error based on parameters readily available in the Automated Tropical Cyclone Forecasting system (ATCF; Sampson and Schrader 2000). The datasets and processes used to create R34 estimates are described in section 2. Results for the 2014–16 western North Pacific R34 analyses, R34 forecast consensus, and R34 predicted consensus error are discussed in section 3. Finally, we summarize our findings and provide needs and suggestions for future development in section 4.

2. Data and methods

The wind radii best tracks and forecasts used in this work are derived from the databases of the ATCF (see https://www.nrlmry.navy.mil/atcf_web/docs/database/new/database.html for formats). The database used for analysis includes the 2014, 2015, and 2016 western North Pacific TC seasons. As with other studies (e.g., SK15; Bender et al. 2017; Sampson et al. 2017), this study will focus on R34 verification since R34 values are the most-often analyzed and likely the best-observed wind radii. Scatterometry is used extensively for verification as it is one of the best methods to construct wind radii analyses around TCs, especially when aircraft-based data are not available. Bentamy et al. (2008), Brennan et al. (2009), and Chou et al. (2013) all suggest that scatterometer winds can be used specifically for R34 analysis. The scatterometer passes cover large areas of the ocean and generally provide high quality estimates of wind speeds less than approximately 50 kt when they are available.

a. *Ground truth*

Two different ground truth estimates of R34 are used in this study. The first are subjective estimates constructed by forecasters as described in appendix B of Sampson et al. (2017). The procedure for producing postseason subjective estimates is relatively new to JTWC, so the authors also use a set of objectively determined R34 values from an average of up to four satellite-based and three model-based estimates (OBTK; Sampson et al. 2017) as a second source of ground truth. The availability of two estimates of R34 (subjective and objective) also provides an opportunity to comment further on the uncertainty in R34 estimates discussed in Sampson et al. (2017).

The first satellite-based method used in the OBTK algorithm employs data from the Advances Microwave Sounding Unit (AMSU) instrument on NOAA and European satellites (hereafter referred to as the AMSU method) and is described in Demuth et al. (2006). The second satellite method (hereafter referred to as DVRK) uses Dvorak (1984) satellite intensity, position, and motion estimates along with matching digital storm infrared imagery and a climatological estimate of the radius of maximum winds to create estimates of wind radii, as described in detail in Knaff et al. (2016). A third satellite-based estimate is the multiplatform TC surface wind analysis described in Knaff et al. (2011) and hereafter referred to as CIRW. The satellite-based technique estimates are supplemented by 6-hourly R34 model forecasts provided by the Geophysical Fluid Dynamics Laboratory (GFDL) Vortex Tracker (Marchok 2002, 2016). The R34 forecasts are currently limited to the Global Forecasting System (named AVNO in the ATCF files) and the Hurricane Weather Research and Forecasting Model (named HWRF in the ATCF files), and the GFDL Hurricane Model (named GFDT in the ATCF files).

A new addition for this study are objective R34 estimates from scatterometer wind swaths over the TCs (referred to hereafter as ASCT; see Fig. 1) since they are derived from the Advanced Scatterometer sensor on the European Space Agency MetOp satellites. These ASCT R34 estimates are computed for our entire 2014–16 dataset (782 cases), as described in the appendix. Since the scatterometer estimates are considered to be ''ground truth'' by forecasters, they are given approximately 30 times (the maximum available with the current algorithm) the weight of the others (all of which are equal weight). Individual estimates available ±3 h from the synoptic times (0000, 0600, 1200, and 1800 UTC) are used in the average. Although the ASCT fixes are a great addition to the OBTK, their frequency does not warrant use on their own as a ground truth dataset even with the 3-yr record. The 782 cases are mostly off synoptic time, are partial passes, and are frequently grouped into periods of frequent (dependent) observations followed by long gaps without any observations. The 782 cases represent only 291 passes. In addition, many of the TCs have no ASCT fixes, which means they would be pruned from an evaluation using only ASCT. Finally, the center of the fix is assigned using bilinear interpolation of the best track, which is sometimes misaligned with the ASCT wind centers. Given the issues with ASCT described above, the authors chose to use the ASCT fixes in OBTK rather than as its own dataset.

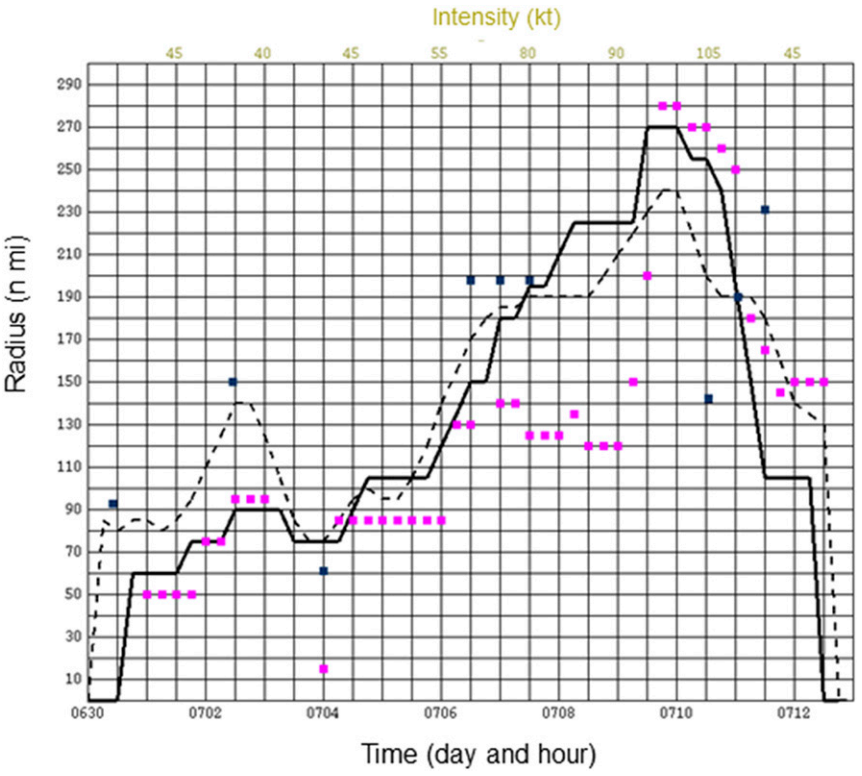

FIG. 1. R34 estimates for the northwest quadrant as a function of time for Chan-Hom, the 15th tropical cyclone of the 2015 western North Pacific season. Dark blue squares are NRL objective ASCT estimates, purple squares are real-time JTWC estimates, the dashed line is the average of many estimates (OBTK), and the solid line is a subjective reanalysis based on scatterometry, imagery, and OBTK. The best-track intensity is shown at the top of the chart.

A three-point center-weighted binomial filter is applied 10 times on the resultant R34 OBTK to smooth the R34 values through time. This smoothing is especially important for the times when a TC's intensity is near 34 kt or the TC is becoming extratropical, as it provides stability in the R34 estimates through time. The filter also reduces real variability in the R34 estimates, so the filter should be adjusted to suit the operational needs at the centers.

b. *Forecast consensus and forecast consensus verification*

Consensus forecasts are constructed by averaging forecasts from available NWP models as described in SK15. Results of NWP model forecasts for a given start time are delayed so as not to be available for the official forecast at that start time; thus, they are considered "late models" by the operational centers. The NWP model forecast output is generally delayed by 6–12 h, and so they are "interpolated" [a process that relabels the output to appear to be at the current forecast time and bias corrects part of the forecast; see Goerss et al. (2004) for more information] to produce guidance that is ready for use in operations. SK15 found that most of the forecast wind radii in the NHC basins can be bias corrected so that the initial wind radii match the current analysis but that in most cases this bias correction should be phased out within 12 h. An exception is made for wind radii estimates from the European Centre for Medium-Range Weather Forecasts model (EMXI), which tended to be smaller than the verification R34 through the entire forecast. For EMXI, the bias correction is applied throughout the entire forecast. The equally weighted R34 consensus forecast defined in SK15, RVCN = AHNI + HHNI + GFTI + EMXI (this RVCN is hereafter named RVCN 2014; see Table 1 for definitions), is applied unaltered in this work so that we avoid tuning our data and maintain independence in the evaluation. As in SK15, the consensus average only includes nonzero wind radii estimates for each forecast time with a minimum of one forecast value per quadrant per forecast time.

In our verification statistics, values of R34 in each quadrant are compared to the best-track values for each estimated value. The occurrence of zero-valued wind

radii in one or more of the storm quadrants introduces an added complication when verifying wind radii. Zero-valued R34 values typically occur when the maximum wind speeds in storms are near the 34-kt intensity threshold or when storm translation speeds are large (i.e., greater than $8\,m\,s^{-1}$). The following verification strategy is adopted: if a quadrant in the best track has a nonzero wind radius, any model or consensus with a nonzero value is verified. Using this strategy, we average the error values from quadrants with nonzero wind radii to form a single measurement of the mean absolute error and bias. The authors define mean bias as the average of an estimated value minus the ground truth. To keep the verification brief, we present the statistics for combined quadrants (i.e., the errors in all quadrants are averaged). Errors are also calculated in homogeneous sets (i.e., they include the same cases). The R34 probability of detection and false alarm results are not presented here. In SK15 we found the consensus probability of detection to be nearly 100% (as we designed it to be by requiring only one nonzero radius value to produce a forecast). Also following SK15, we consider false alarms to be beneficial as they provide additional information on TC size to the forecaster. For example, the forecaster would be informed of the approximate size of the wind radii ahead of its intensification to 35 kt.

c. *Goerss-predicted consensus error*

As is the case with the track or intensity Goerss-predicted consensus error (GPCE; Goerss 2007; Goerss and Sampson 2014), predictors of wind radii GPCE are limited to parameters available in the real-time ATCF files at operational centers for use in the operational forecast. This is a practical consideration to avoid complexity (a sure way to generate support issues) and ensure that the algorithm can be executed in real time with success. Consensus model spread is defined to be the average of the absolute differences between the nonzero member R34 forecasts and RVCN. The predictors examined in this study are latitude and longitude, current TC intensity, forecast TC intensity and intensity change, TC speed of motion, R34 forecast and R34 forecast change from the initial R34 estimate, average R34 spread from the consensus mean, and the number of model members available. Forecast TC intensity and intensity change are determined using the interpolated official forecasts (JTWI). All cases where an official intensity forecast was made and verified against best-track values were used in the GPCE analysis.

After regression analysis of the abovementioned predictors, we used stepwise linear regression (Draper and Smith 1966) and the pool of predictors from the 2015 western North Pacific basins regression models to predict the RVCN absolute forecast error at each forecast period. We required that a predictor explain at least 3% of the variance before allowing it to be used by the final regression equation in order to avoid overfitting. All of the final regression coefficients were found to be significantly different from zero at well above the 99% level using an F test.

d. *Development sets and independent data*

Questions frequently arise regarding how best to develop and evaluate tropical cyclone forecast products (e.g., the RVCN and RVCN absolute forecast error described in this manuscript) that are intended for use in real-time operations. Cross validation is one option, but there are many obstacles that must be avoided in order to choose samples for cross validation. If information from the omitted dependent dataset is used at any point in the procedure that selects the rule used in hindcasting, then the algorithm is not truly cross validated and forecast skill will be overestimated (Elsner and Schmertmann 1994). One way to avoid these obstacles is to apply the development and evaluation as will be done in real time, developing on past data and withholding the future data for independent evaluation (Von Storch and Zwiers 1999). Doing this not only ensures that the evaluation does not overestimate the skill, but also tests the viability of the implementation in operations. As such, the authors elect to develop in this manner.

3. Results

Figure 2 shows an evaluation of a 3-yr dataset of western North Pacific TCs for individual R34 estimates and the OBTK against the subjective best tracks. The evaluation is limited to data coincident with scatterometer passes, as those are considered to be of the highest quality available in the western North Pacific. By design, the OBTK availability is nearly 100% and is higher than for any of the individual estimates. This availability is preferable for JTWC operations, since they most always have guidance even if they decide not to assign a particular radius. All the estimates have average errors of 20–40 n mi, which equates to approximately 15%–30% of the climatological mean R34 (130 n mi) for this dataset. The ASCT estimates are nearly zero biased, and the mean error relative to the subjective best tracks is 26 n mi. The OBTK has a slightly high bias and has estimates closest to the subjective best-track estimates. All three NWP model estimates are high biased for this dataset, which is different than in Sampson et al. (2017) in that the GFDT biases in that study were near zero and slightly negative. The high bias is expected

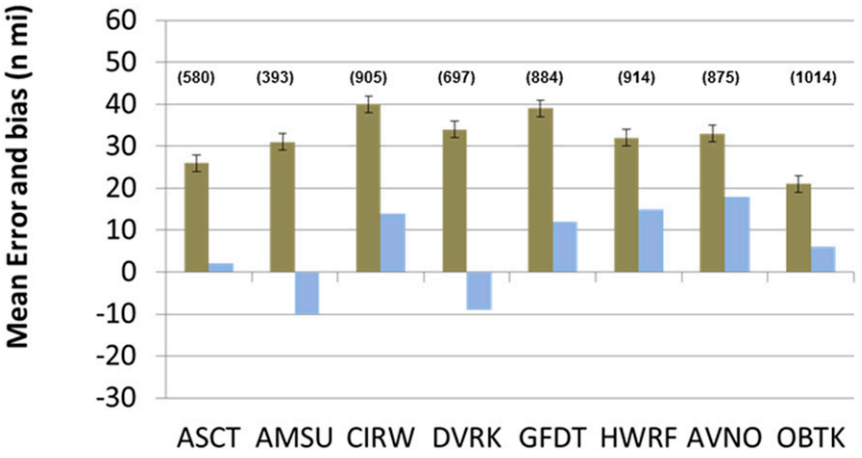

FIG. 2. The 34-kt wind radii fix absolute mean errors (brown) and biases (blue) relative to JTWC 2014–16 best tracks coincident with ASCT. OBTK has the lowest mean error of all estimates. The standard error of the mean is shown as the black bars and whiskers; the number of cases is shown in parentheses. Acronyms are as defined in Table 1.

since the 2016 Geophysical Fluid Dynamics Laboratory model had been modified to compensate for small-biased R34 estimates from JTWC (M. Bender 2017, personal communication), a modification that was subsequently removed for 2017 when the JTWC estimates were found to have become nearly zero biased.

To answer questions about the independence of the R34 in subjective best tracks, the OBTK, and the objective ASCT fixes, we computed correlation coefficients between the three datasets. We find the correlation coefficient between the subjective R34 and OBTK to be 0.84 using the 578 individual radii in the 2014–16 western North Pacific seasons available from three methods, which is high but certainly not perfect. We find the correlation between the subjective R34 and the objective ASCT fixes to be 0.76. These are well correlated but certainly less correlated than the subjective best track and the OBTK wind radii. The correlation between the OBTK and the ASCT fixes is found to be 0.92, which is high and expected since the OBTK is largely influenced by these fixes. There are several reasons the two best-track datasets are not perfectly correlated with the ASCT fixes including time windows (the ASCT fixes are produced with a moving ±3-h window of data around the synoptic time, which can include multiple passes), subjective analysis, and different scatterometer data processing and display formats. The above discussion and the display of results from the three different methods shown in Fig. 1 illustrate uncertainty in where the R34 should be located. Both the subjective analysis and the OBTK smoothly transition from small to large and back again, which is a desirable quality in operational forecasting where these changes can impact emergency planning and ship-routing decisions as well as the credibility of the warning agency. Smoothing is also employed for determining position and intensity at the operational centers. For example, neither the subjective nor OBTK analysis strictly adheres to the ASCT fixes in the northwest quadrant in Fig. 1. These traces should instead reflect gradual changes in time, which are more acceptable to end users.

If the OBTK and subjective best track are assumed to be of equal quality, we can estimate the uncertainty in R34. In the mean, the estimates are within 21 n mi of each other. But it is the standard deviation of the errors, considered in Torn and Snyder (2012) to be a measure of uncertainty in track and intensity, that can be used on our dataset to estimate R34 uncertainty. The standard deviation of the errors for OBTK using the subjective best tracks as ground truth is 18 n mi and is larger than the 14 n mi estimate Sampson et al. (2017) found for the 2014–15 NHC data; however, the estimates as a percentage of the climatological sizes are both approximately 15%.

Now that we have R34 estimates that can serve as ground truth (the objective best-track OBTK or the subjective postseason analyses), we can evaluate the performance of the R34 forecasts and the equally weighted average of the consensus (RVCN). Figure 3 shows the mean forecast errors and biases of the individual members and the consensus using the subjective postseason analyses as ground truth. There are two consensus forecasts shown: the first (RVCN 2014) is the consensus described in SK15 with the four input forecasts (AHNI + HHNI + GFTI + EMXI) and the second (RVCN 2017) includes guidance that was made available since, namely SHIPS-based wind radii based on the National Centers for Environment Prediction Global Forecast System dynamic input and tracks (DSHA; see Knaff et al. 2017) and the COAMPS-TC

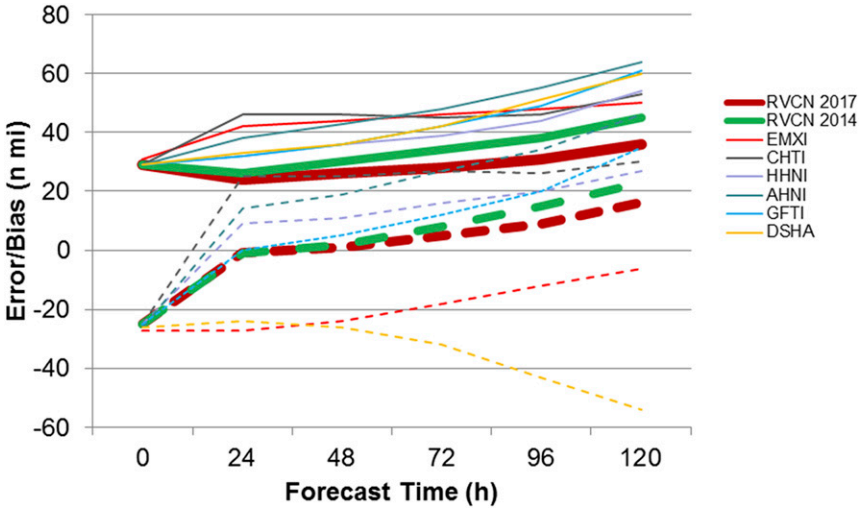

FIG. 3. RVCN 34-kt wind radii forecast performance using reanalyzed JTWC best tracks as ground truth for 2014–16 western North Pacific seasons. Solid lines show mean absolute error, and dashed lines show bias. Acronyms are as defined in Table 1.

forecasts (CHTI; Doyle et al. 2012, 2014) run with U.S. Navy global model initial and boundary conditions. As was the case in SK15, the consensus forecasts have among the smallest mean forecast errors and biases when compared to their member aids. The initial forecast bias for individual and consensus aids is approximately 26 n mi too small. All these aids begin as the JTWC real-time estimate so the small bias is an artifact of the real-time JTWC estimates being small compared to the subjective best tracks. Again, with most individual aids the real-time JTWC estimate is phased out by 12 h, so the negative bias is diminished over time. The exceptions are EMXI and DSHA. EMXI shows bias throughout its forecast, as prescribed by the interpolator described in SK15. The DSHA aid retains part of the initial bias throughout the forecast as the model predicts changes in TC size from the initial conditions. The low bias in EMXI is also expected as it was present in the NHC basins (SK15). The large negative biases in DSHA at 96 and 120 h are likely compounded by negative forecast intensity biases in the SHIPS (DeMaria et al. 2005) dataset used for the computation of the DSHA wind radii. Although unintentional, these biases offset the gradual increase in the bias of the remaining members with forecast time, which are possibly related to an overdeepening of TCs in the subtropics, as discussed by Heming (2016). Finally, errors and biases computed using the OBTK as ground truth (not shown) appear similar to those shown in Fig. 3, but with 5 n mi more negative initial biases and 5 n mi higher initial errors. Since the OBTK R34 is somewhat independent of the subjective best tracks (correlation coefficient of 0.84), it provides us with more confidence in the results.

Since we have the R34 forecasts with verification, we can also attempt to estimate errors associated with the forecasts, as is done with track consensus forecasts in Goerss (2007) and with intensity consensus forecasts in Goerss and Sampson (2014). First, we illustrate the relationships between some of the possible predictors and the RVCN forecast error. For RVCN in the 2015 western North Pacific basin, the consensus absolute forecast wind radius change (WRCHG) is found to be positively correlated with the consensus model TC wind radius absolute forecast error for all forecast lengths. This relationship is illustrated in the scatterplot in Fig. 4 (top left), where we see that, in general, there is a positive correlation between consensus forecast absolute wind radius error and consensus absolute forecast wind radius change. The correlation between WRCHG and absolute wind radius forecast error ranges from 0.71 for the 12-h RVCN forecasts (Fig. 4, top left) to 0.32 for the 96-h RVCN forecasts (not shown). WRCHG is chosen as the leading predictor for the 12–72-h forecasts with correlations ranging from 0.71 to 0.38 and is chosen as the second leading predictor for the 96- and 120-h forecasts with correlations of 0.32 and 0.36, respectively. One caveat, which will be seen in the evaluation below, is that the WRCHG is computed from postseason R34 analyses rather than those estimated in real time. This is done because changes are being implemented in operational procedures that adhere more closely with the postseason tracks than the real-time tracks from our sample.

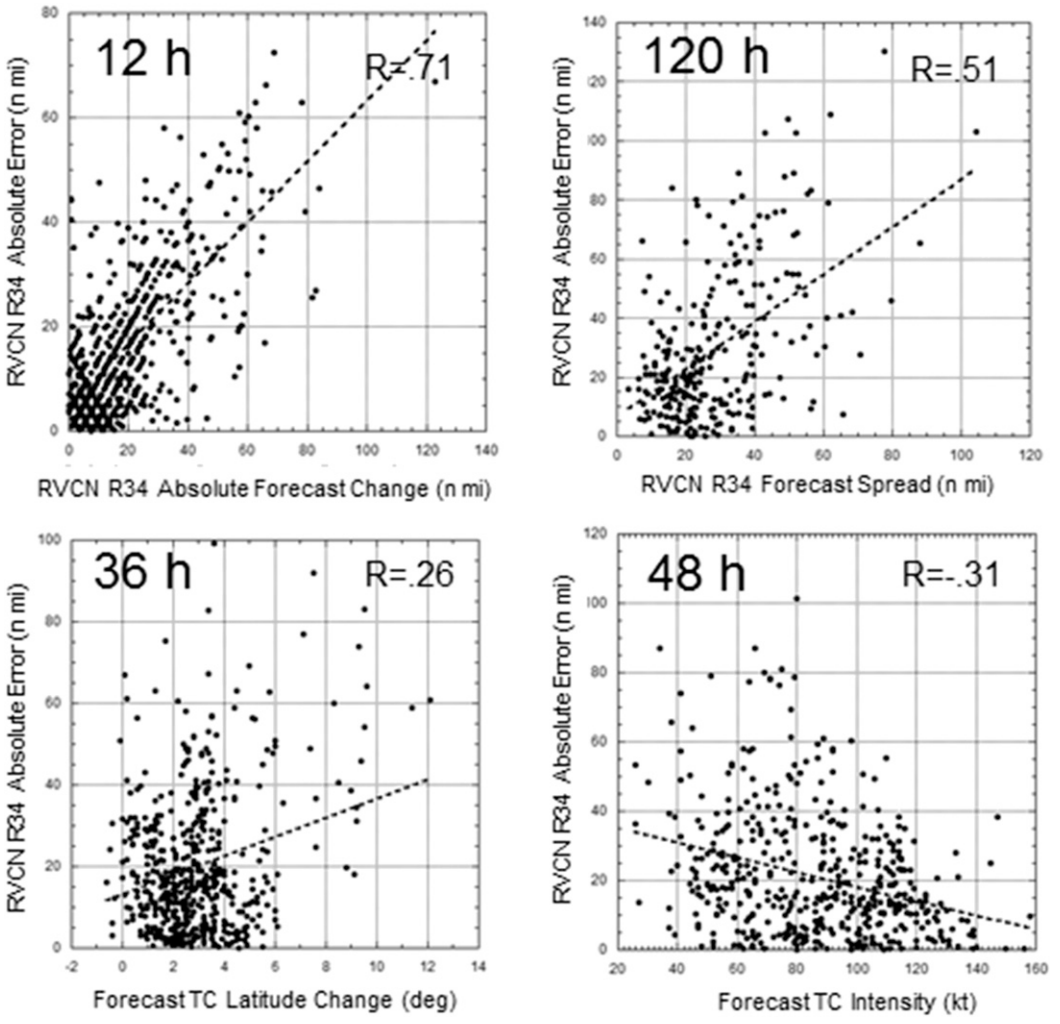

FIG. 4. R34 error vs top GPCE predictors at the forecast lead time shown at the top left of each panel for the 2015 western North Pacific season. Correlation coefficients are shown in the top-right corner of each panel.

Examples of the other leading predictors are shown in the other scatterplots in Fig. 4. The consensus model spread (SPRD; Fig. 4, top right) and the forecast change in TC latitude (LATC; Fig. 4, bottom left) are found to be positively correlated with forecast absolute wind radius error while the forecast TC intensity (INTF; Fig. 4, bottom right) is found to be negatively correlated. Consensus model spread is chosen as the leading predictor for the 96- and 120-h RVCN forecasts. The correlations between SPRD and forecast absolute wind radius error range from 0.34 for the 72-h RVCN forecasts to 0.51 for the 120-h RVCN forecasts. Forecast latitude change is chosen as the third leading predictor for the 24- and 36-h RVCN forecasts with correlations of 0.25 at 24 h and 0.26 at 36 h. Forecast TC intensity is chosen as the second-leading predictor for the 24–72-h RVCN forecasts and the third-leading predictor for the 96- and 120-h RVCN forecasts with correlations ranging from −0.31 at 48 h to −0.17 at 96 h.

Next, we want to translate these results into a form that has meaning to the NHC and JTWC forecasters. To do so, we construct ranges centered on the consensus wind radius forecasts that contain the verifying TC wind radius roughly 67% of the time. For the 2015 western North Pacific season, which we consider the dependent data, ranges are determined by trial and error, adding a constant at each forecast length to the predicted RVCN TC absolute wind radius forecast error derived from the linear regression. The constants computed are 9, 22, 31, 37, 30, 29, and 31 n mi at 12, 24, 36, 48, 72, 96, and 120 h, respectively. The final regression equations describing the relationship between RVCN absolute wind radius forecast error and predicted error for all forecast lengths are shown in Table 2. The correlation coefficients range

TABLE 2. Regression equations derived for the 2015 western North Pacific season with variance explained for dependent (2015) and independent (2016) data. Acronyms are defined in Table 1.

Wind radii forecast error equation	Cases (2015)	Variance explained (2015)	Cases (2016)	Variance explained (2016)
$WRE_{12} = (0.581 \times WRCHG) + 9$	630	0.089	372	0.171
$WRE_{24} = (0.307 \times WRCHG) - (0.107 \times INTF) + (2.22 \times LATC) + 22$	603	0.140	347	0.197
$WRE_{36} = (0.205 \times WRCHG) - (0.161 \times INTF) + (2.09 \times LATC) + 31$	569	0.137	315	0.161
$WRE_{48} = (0.237 \times WRCHG) - (0.213 \times INTF) + 37$	531	0.125	278	0.118
$WRE_{72} = (0.441 \times SPRD) + (0.217 \times WRCHG) - (0.238 \times INTF) + 30$	443	0.236	206	0.076
$WRE_{96} = (0.489 \times SPRD) + (0.223 \times WRCHG) - (0.274 \times INTF) + 29$	366	0.341	143	0.068
$WRE_{120} = (0.542 \times SPRD) + (0.249 \times WRCHG) - (0.310 \times INTF) + 31$	288	0.412	93	0.262

from a low of 0.298 at 12 h (expected as the equations were derived from best tracks) to a high of 0.642 at 120 h. Despite using real-time estimates of R34 instead of postseason best tracks, the percent variance explained is still in the 9%–41% range. This is more variance explained than was found in the original track GPCE (Goerss 2007) and intensity GPCE (Goerss and Sampson 2014).

Tests on independent 2016 data also shown in Table 2 indicate that the correlation coefficients at the shortest forecast periods have increased. We suspect that this is partly due to ongoing improvements in operational real-time R34 estimates at JTWC and is why we decided to train the algorithm on best-track R34 data rather than the real-time forecaster estimates when we developed the algorithm. The number of cases in the independent sample drops under 300 for the 48–120-h forecast periods, which is not optimal, and could be responsible for the wide variation in correlation coefficients for those forecast periods. Performance of the GPCE for RVCN 2017 on the same dataset was similar to that of RVCN 2014 shown in Table 2, with lower correlation coefficients for the 12-, 24-, 36-, and 48-h forecasts and higher correlation coefficients for the 72-, 96-, and 120-h forecasts. This is without any tuning to the addition of new models and provides an indication that the algorithm is useable even if models are added, removed, or changed. Finally, we expect the coefficients to be recomputed with more data before the start of each new season similar to what is done for the track and intensity GPCEs.

For further verification of the wind radii GPCE, we choose to compute the predicted average half-range (two half-ranges equals one confidence range), the minimum and maximum half-ranges, the percent of the time that the verifying R34 is included within the confidence range centered on the RVCN forecast wind radius, and the number of forecasts (Table 3). We can see that, for the 24-h wind radius forecasts for example, the size of the confidence ranges vary from ±11 to ±76 n mi rather than a fixed or climatological confidence interval based on the average error for the 2015 season (29 n mi). Variable confidence ranges allow a forecaster to place more or less confidence on the RVCN wind radius forecasts. Independent results for the 2016 season are also shown at the bottom of Table 3. Most important, the performance on the independent data is similar to that of the dependent data. Finally, the performance of the RVCN 2017 was similar to that of RVCN 2014. The percent of forecast wind radii included in the half-range jumps 2%–12%, the average half-range is 0–2 n mi larger, the minimum half-ranges are 0–5 n mi larger, and

TABLE 3. GPCE verification for the 2015 and 2016 western North Pacific seasons. Average, min, and max are for half-ranges so the percentages included are computed using the RVCN value plus or minus the half-range.

Forecast (h)	12	24	36	48	72	96	120
	Dependent data (2015)						
Average (n mi)	30	29	30	33	34	36	45
Min (n mi)	9	11	13	12	6	9	9
Max (n mi)	128	76	61	71	81	93	109
Percentage included	67	69	69	69	68	66	66
No. of forecasts	649	619	575	535	442	367	288
	Independent data (2016)						
Average (n mi)	32	31	33	35	37	40	53
Min (n mi)	9	9	11	6	5	16	24
Max (n mi)	119	70	63	71	70	73	113
Percentage included	60	62	67	65	66	68	73
No. of forecasts	372	347	315	278	206	143	93

the maximum half-ranges are as much as 14 n mi smaller. Our hope is that the RVCN 2017 performance can be improved further as we gather enough data to derive new regression coefficients, specifically for that suite of models and its consensus.

4. Conclusions and recommendations

In this paper, we updated our equally weighted R34 estimates (OBTK; Sampson et al. 2017) to include objectively analyzed scatterometer fixes. For a 3-yr dataset of global scatterometer wind observations, we were able to locate and estimate R34 for 782 cases for TCs in the NHC and JTWC archives. Since R34 estimates from the scatterometer are considered to be of high quality and are often used as ground truth, they were given extra weight in the OBTK algorithm. Somewhat independently (i.e., not beholden to OBTK estimates), Sampson et al. (2017) developed a subjective analysis of R34 for the western North Pacific 2014–16 seasons. We used both datasets for RVCN evaluation, and the results were similar and provided us with a degree of confidence that RVCN performs well. When compared to its members, the RVCN consensus was the top performer in both mean forecast error and bias. One concern for the future is that the experimental GFDL model run as part of the Hurricane Forecast Improvement Project was discontinued on 16 July 2017. Consensus forecasts become erratic when there are few member model forecasts, so we are constantly looking for more consensus members to add to the suite. RVCN 2014 was implemented at JTWC in 2015, and OBTK was implemented at JTWC in 2016. Both were updated (the OBTK to include scatterometer fixes and the RVCN to RVCN 2017) at the end of the 2017 season and will continue to be upgraded as improvements are made. The RVCN error estimates were implemented at the end of the 2017 season, and graphical display for forecasting will follow in 2018.

We can also use the two sets of estimates to provide a rough estimate of uncertainty by calculating the standard deviation of the differences between the two. This uncertainty estimation follows the calculations used in Torn and Snyder (2012). The standard deviation for OBTK relative to the subjective R34 is 18 n mi for the estimates coincident with objective scatterometer fixes and represents estimated uncertainty for the best-observed cases. This is higher than the 14 n mi estimate we found for the 2014–15 NHC data; however, the estimates as a percentage of climatological R34 in the datasets are both approximately 15%.

This work used existing operational real-time datasets for the ground truth, which could be further improved with more independent estimates using data from recently launched satellites (Morris and Ruf 2017; Meissner et al. 2017; Reul et al. 2017) and other observations. The improved ground truth could in turn be employed to improve guidance, either actively [e.g., redeveloping the Wind Radii Climatology and Persistence Model of Knaff et al. (2007)] or passively [e.g., using R34 from the OBTK, as in Bender et al. (2017)]. OBTK can and is being deployed for other basins and is expected to be especially useful in areas with few in situ observations (e.g., most of the Southern Hemisphere and Indian Ocean). RVCN has already been employed with success in the NHC basins and is also now being run for the Southern Hemisphere and Indian Ocean. The addition of other skillful models to the suite will be pursued as this not only potentially adds skill but also reduces the erratic behavior between quadrants and with respect to time that is present in a consensus of few models. The authors intend to investigate if the wind radii GPCE can be applied to other basins with existing coefficients or if new coefficients are required in each basin.

The TC community continues to address issues with R34 estimates, but attention has shifted to the more difficult problem of ascertaining the inner-core structure of TCs. Although this is an inherently difficult problem for several reasons (e.g., sparse observations and insufficient NWP model resolution), efforts to observe and explore the inner core are already yielding results, and hopes are even higher for the future as costs for observing platforms such as microsatellites and remotely piloted vehicles drop and capabilities improve.

Acknowledgments. The authors would like to acknowledge the staff at NHC and JTWC, the NOAA GOES and GOES-R program offices for support of TC structure studies, three anonymous reviewers, and the important behind-the-scenes efforts of Kim Richardson and Mike Frost. This publication was graciously funded by the Office of Naval Research, Program Element 0602435N. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. government position, policy, or decision.

APPENDIX

NRL Monterey Objective Scatterometer Fix Algorithm

The NRL Monterey objective scatterometer fix algorithm was originally developed as a part of a structure

analysis for use in COAMPS-TC initial conditions; however, parts of the algorithm were also applicable to an objective scheme to find R34 in scatterometer passes. The scheme was developed using four years of scatterometer passes over tropical cyclones in the Atlantic and the entire North Pacific basins. Each R34 fix was individually inspected and the algorithm adjusted to avoid assigning extremely large (greater than 240 n mi) or extremely small (smaller than 40 n mi) estimates, as well as estimates that appeared to be separated from the TC flow. The algorithm was designed to make estimates only in cases where we had confidence that those estimates were comparable to available subjective estimates from NHC and/or passed visual inspection. The following is the resulting algorithm. It misses many of the fixes, but we have confidence in its results:

- Read in best track.
- Read in scatterometer wind observations for the entire day; select a ±3-h window.
- Move wind observations to best-track time using current TC movement.
- Find data that is within 330 n mi of the estimated center of the TC.
- Divide area up into 8 n mi wide annuli, then quarter using compass directions.
- Find maximum wind speed in each quarter annulus.
- Start R34 search algorithm.
  - Search from 40 to 240 n mi using the 8 n mi quarter annuli from above.
  - Find where observations cross the 34-kt threshold.
  - Only retain locations where the slope is negative (i.e., from high to low).
  - The 34-kt winds cannot have gaps > 75 n mi going out from center.[A1]
  - First wind speed > 33 kt needs to be at $r < 120$ n mi.
- Write scatterometer fix in ATCF ASCII format and ingest into ATCF.

We diagnosed a total of 1037 cases using the 2013–16 best tracks from the Atlantic and North Pacific and scatterometer fixes, many of those fixes having multiple R34 estimates (up to four per fix). We also scrutinized fixes in near–real time in 2017 and found it produced R34 estimates that were generally similar to the subjective estimates from NHC.

[A1] This restriction eliminates fixes to the wind radii that are from partial passes and/or relatively flat wind fields where it is difficult to determine whether a 34-kt wind is associated with a TC or some other unassociated feature. Subjective scrutiny plays an important role, which is why we chose to eliminate those fixes.

REFERENCES

Bender, M. A., T. Marchok, C. Sampson, J. Knaff, and M. Morin, 2017: Impact of storm size on prediction of storm track and intensity using the 2016 operational GFDL hurricane model. *Wea. Forecasting*, **32**, 1491–1508, https://doi.org/10.1175/WAF-D-16-0220.1.

Bentamy, A., D. Croize-Fillon, and C. Perigaud, 2008: Characterization of ASCAT measurements based on buoy and QuikSCAT wind vector observations. *Ocean Sci.*, **4**, 265–274, https://doi.org/10.5194/os-4-265-2008.

Brennan, M. J., C. C. Hennon, and R. D. Knabb, 2009: The operational use of QuikSCAT ocean surface vector winds at the National Hurricane Center. *Wea. Forecasting*, **24**, 621–645, https://doi.org/10.1175/2008WAF2222188.1.

Cangialosi, J. P., and C. W. Landsea, 2016: An examination of model and official National Hurricane Center tropical cyclone size forecasts. *Wea. Forecasting*, **31**, 1293–1300, https://doi.org/10.1175/WAF-D-15-0158.1.

Chou, K.-H., C.-C. Wu, and S.-Z. Lin, 2013: Assessment of the ASCAT wind error characteristics by global dropwindsonde observations. *J. Geophys. Res. Atmos.*, **118**, 9011–9021, https://doi.org/10.1002/jgrd.50724.

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

——, and Coauthors, 2013: Improvements to the operational tropical cyclone wind speed probability model. *Wea. Forecasting*, **28**, 586–602, https://doi.org/10.1175/WAF-D-12-00116.1.

Demuth, J. L., M. DeMaria, J. A. Knaff, and T. H. Vonder Haar, 2004: Evaluation of Advanced Microwave Sounder Unit (AMSU) tropical cyclone intensity and size estimation algorithm. *J. Appl. Meteor.*, **43**, 282–296, https://doi.org/10.1175/1520-0450(2004)043<0282:EOAMSU>2.0.CO;2.

——, ——, and ——, 2006: Improvement of Advanced Microwave Sounding Unit tropical cyclone intensity and size estimation algorithms. *J. App. Meteor. Climatol.*, **45**, 1573–1581, https://doi.org/10.1175/JAM2429.1.

Dolling, K., E. A. Ritchie, and J. S. Tyo, 2016: The use of the deviation angle variance technique on geostationary satellite imagery to estimate tropical cyclone size parameters. *Wea. Forecasting*, **31**, 1625–1642, https://doi.org/10.1175/WAF-D-16-0056.1.

Doyle, J. D., and Coauthors, 2012: Real time tropical cyclone prediction using COAMPS-TC. *Adv. Geosci.*, **28**, 15–28, https://doi.org/10.1142/9789814405683_0002.

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

Draper, N. R., and H. Smith, 1966: *Applied Regression Analysis*. John Wiley and Sons, 407 pp.

Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, 47 pp., http://satepsanone.nesdis.noaa.gov/pub/Publications/Tropical/Dvorak_1984.pdf.

Elsner, J. B., and C. P. Schmertmann, 1994: Assessing forecast skill through cross validation. *Wea. Forecasting*, **9**, 619–624, https://doi.org/10.1175/1520-0434(1994)009<0619:AFSTCV>2.0.CO;2.

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985–1993, https://doi.org/10.1175/MWR3390.1.

——, 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.

——, ——, and J. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638, https://doi.org/10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2.

Heming, J. T., 2016: Met Office Unified Model tropical cyclone performance following major changes to the initialization scheme and a model upgrade. *Wea. Forecasting*, **31**, 1433–1449, https://doi.org/10.1175/WAF-D-16-0040.1.

JTWC, 2017: JTWC tropical cyclone best track data site. Joint Typhoon Warning Center, http://www.metoc.navy.mil/jtwc/jtwc.html.

Knaff, J. A., and B. A. Harper, 2010: Tropical cyclone surface wind structure and wind-pressure relationships. *Seventh Int. Workshop on Tropical Cyclones*, La Reunion, France, WMO, KN1, https://www.wmo.int/pages/prog/arep/wwrp/tmr/otherfileformats/documents/KN1.pdf.

——, and C. R. Sampson, 2015: After a decade are Atlantic tropical cyclone gale force wind radii forecasts now skillful? *Wea. Forecasting*, **30**, 702–709, https://doi.org/10.1175/WAF-D-14-00149.1.

——, ——, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence. *Wea. Forecasting*, **22**, 781–791, https://doi.org/10.1175/WAF1026.1.

——, M. DeMaria, D. A. Molenar, C. R. Sampson, and M. G. Seybold, 2011: An automated, objective, multisatellite platform tropical cyclone surface wind analysis. *J. Appl. Meteor. Climatol.*, **50**, 2149–2166, https://doi.org/10.1175/2011JAMC2673.1.

——, C. J. Slocum, K. D. Musgrave, C. R. Sampson, and B. 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.

——, C. R. Sampson, and G. Chirokova, 2017: A global statistical–dynamical tropical cyclone wind radii forecast scheme. *Wea. Forecasting*, **32**, 629–644, https://doi.org/10.1175/WAF-D-16-0168.1.

Kossin, J. P., J. A. Knaff, H. I. Berger, D. C. Herndon, T. A. Cram, C. S. Velden, R. J. Murnane, and J. D. Hawkins, 2007: Estimating hurricane wind structure in the absence of aircraft reconnaissance. *Wea. Forecasting*, **22**, 89–101, https://doi.org/10.1175/WAF985.1.

Kunii, M., 2015: Assimilation of tropical cyclone track and wind radius data with an ensemble Kalman filter. *Wea. Forecasting*, **30**, 1050–1063, https://doi.org/10.1175/WAF-D-14-00088.1.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, https://doi.org/10.1175/MWR-D-12-00254.1.

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/pdfpapers/37628.pdf.

——, 2016: Scientific documentation for the community release of the GFDL Vortex Tracker: Version 3.7b. Developmental Testbed Center, 22 pp., http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/stand_alone_tracker_SD_v3.7b.pdf.

Meissner, T., L. Ricciardulli, and F. Wentz, 2017: Capability of the SMAP mission to measure ocean surface winds in storms. *Bull. Amer. Meteor. Soc.*, **98**, 1660–1677, https://doi.org/10.1175/BAMS-D-16-0052.1.

Morris, M., and C. S. Ruf, 2017: Determining tropical cyclone surface wind speed structure and intensity with the CYGNSS satellite constellation. *J. Appl. Meteor. Climatol.*, **56**, 1847–1865, https://doi.org/10.1175/JAMC-D-16-0375.1.

Mueller, K. J., M. DeMaria, J. A. Knaff, J. P. Kossin, and T. H. Vonder Haar, 2006: Objective estimation of tropical cyclone wind structure from infrared satellite data. *Wea. Forecasting*, **21**, 990–1005, https://doi.org/10.1175/WAF955.1.

NHC, 2016: Introduction to storm surge. Storm Surge Unit, National Hurricane Center, 5 pp., http://www.nhc.noaa.gov/surge/surge_intro.pdf.

Powell, M. D., and T. A. Reinhold, 2007: Tropical cyclone destructive potential by integrated kinetic energy. *Bull. Amer. Meteor. Soc.*, **88**, 513–526, https://doi.org/10.1175/BAMS-88-4-513.

Quiring, S., A. Schumacher, and S. Guikema, 2014: Incorporating hurricane forecast uncertainty into decision support applications. *Bull. Amer. Meteor. Soc.*, **95**, 47–58, https://doi.org/10.1175/BAMS-D-12-00012.1.

Reul, N., and Coauthors, 2017: A new generation of tropical cyclone size measurements from space. *Bull. Amer. Meteor. Soc.*, **98**, 2367–2385, https://doi.org/10.1175/BAMS-D-15-00291.1.

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.

——, 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.

——, P. A. Wittmann, and H. L. Tolman, 2010: Consistent tropical cyclone wind and wave forecasts for the U.S. Navy. *Wea. Forecasting*, **25**, 1293–1306, https://doi.org/10.1175/2010WAF2222376.1.

——, and Coauthors, 2012: Objective guidance for use in setting tropical cyclone conditions of readiness. *Wea. Forecasting*, **27**, 1052–1060, https://doi.org/10.1175/WAF-D-12-00008.1.

——, E. M. Fukada, J. A. Knaff, B. R. Strahl, M. J. Brennan, and T. Marchok, 2017: Tropical cyclone gale wind radii estimates for the western North Pacific. *Wea. Forecasting*, **32**, 1007–1028, https://doi.org/10.1175/WAF-D-16-0196.1.

Tallapragada, V., C. Q. Kieu, Y. C. Kwon, S. G. Trahan, Q. Liu, Z. Zhang, and I.-H. Kwon, 2014: Evaluation of storm structure from the operational HWRF Model during 2012 implementation. *Mon. Wea. Rev.*, **142**, 4308–4325, https://doi.org/10.1175/MWR-D-13-00010.1.

——, and Coauthors, 2015: Hurricane Weather Research and Forecasting (HWRF) Model: 2015 scientific documentation, August 2015 – HWRF v3.7a. NCAR Developmental Testbed Center, 123 pp., http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRF_v3.7a_SD.pdf.

Torn, R., 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.

Von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.