An Examination of Model and Official National Hurricane Center Tropical Cyclone Size Forecasts

JOHN P. CANGIALOSI AND CHRISTOPHER W. LANDSEA

NOAA/NWS/NCEP/National Hurricane Center, Miami, Florida

(Manuscript received 15 November 2015, in final form 3 June 2016)

ABSTRACT

While the National Hurricane Center (NHC) has been issuing analyses and forecasts of tropical cyclone wind radii for several years, little documentation has been provided about the errors in these forecasts. A key hurdle in providing routine verification of these forecasts is that the uncertainty in the wind radii best tracks is quite large for tropical cyclones that are well away from land and unmonitored by aircraft reconnaissance. This study evaluates the errors of a subset of NHC and model 34-, 50-, and 64-kt (1 kt = 0.514 m s^{-1}) wind radii forecasts from 2008 through 2012 that had aircraft reconnaissance available at both the initial and verification times. The results show that the NHC wind radii average errors increased with forecast time but were skillful when compared against climatology and persistence. The dynamical models, however, were not skillful and had errors that were much larger than the NHC forecasts, with substantial negative (too small) biases even after accounting for their initial size differences versus the tropical cyclone's current wind radii. Improvements in wind radii forecasting will come about through a combination of better methods for observing tropical cyclone size as well as enhanced prediction techniques (dynamical models, statistical methods, and consensus approaches).

1. Introduction

The National Hurricane Center (NHC) analyzes and forecasts the size of tropical cyclones (TCs) in four quadrants of the cyclone (northeast, southeast, southwest, and northwest) out to 72 h at the 34- and 50-kt (1 kt = 0.514 m s^{-1}) wind thresholds, and additionally out to 36 h at the 64-kt wind threshold. These values are currently analyzed and predicted to a precision of 10 n mi (1 n mi = 1.852 km) for 34- and 50-kt radii and 5 n mi for 64-kt radii. These radii indicate the maximum (not average) extent of these winds in each quadrant (Fig. 1). It is well known that the size of the TC wind field varies significantly from storm to storm and during an individual TC life cycle. Predictions of TC wind radii are directly incorporated into watches and warnings at both land and sea, so that these forecasts are a critical part of the warning-decision-making process. The NHC has been forecasting the wind radii of TCs for several decades. The 50-kt wind radii forecasts were first introduced in 1958, followed by the 34-kt wind radii forecasts in 1979, and the 64-kt wind radii forecasts in 1995. Even though the NHC wind radii forecasts have been produced for quite some time, there has not been an annual or standard verification of this metric similar to what is done routinely for TC track and intensity forecasts (Cangialosi and Franklin 2015). Knaff and Sampson (2015), however, examined NHC's 24-, 48-, and 72-h wind radii forecasts of 34-kt wind radii for all Atlantic basin TCs for the period 2004–13 and found average errors of about 26, 31, and 37 n mi, respectively, with about a 5 n mi bias (too small) present in the forecasts.

NHC does not produce a formal verification of its wind radii forecasts because of the lack of data to support a reliable poststorm analysis to serve as ground truth. While NHC has been providing a poststorm reassessment (‘‘best track'') of TC size since 2004, for most TCs over the open ocean, surface observations are extremely limited with only an occasional ship, moored buoy, or coastal station measurement (Landsea and Franklin 2013). The majority of the wind radii information for TCs over the open ocean comes from satellite-based scatterometer passes, which are infrequent and often do not sample the entire TC circulation (Brennan et al. 2009). In addition, rain contamination makes those data difficult to interpret at times,

---

Corresponding author address: John. P. Cangialosi, National Hurricane Center, 11691 SW 17th St., Miami, FL 33165.
E-mail: john.p.cangialosi@noaa.gov

DOI: 10.1175/WAF-D-15-0158.1

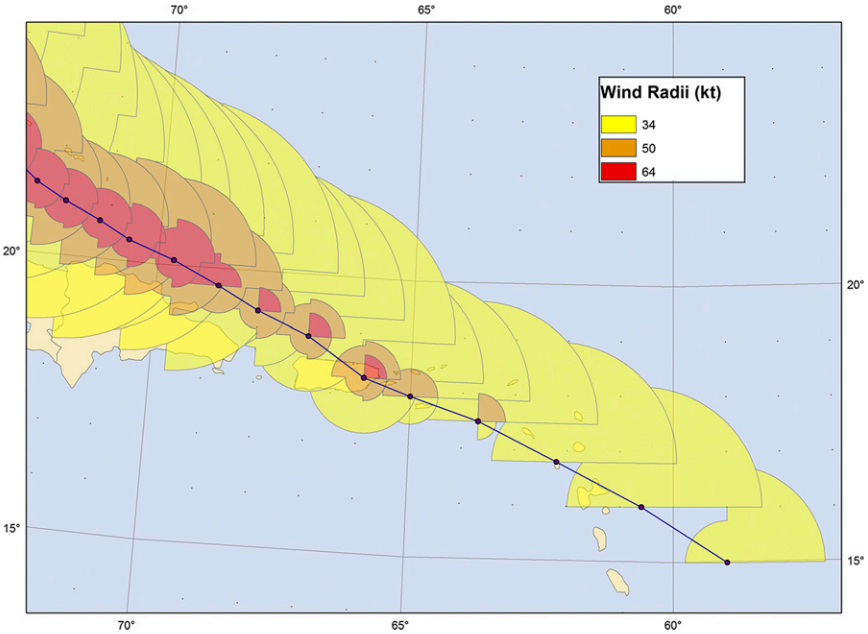


FIG. 1. An example of NHC's 34-, 50-, and 64-kt wind radii forecasts from Hurricane Irene (2011).

and they are typically only useful for the 34-kt and occasionally 50-kt wind radii of TCs (Zeng and Brown 1998; Yueh et al. 2003) because of inadequate resolution as well as the reduced sensitivity of the measurement at high wind speeds. There do exist other satellite-based techniques with some ability to assess TC size, such as the Advanced Microwave Sounding Unit (Demuth et al. 2006) and the multiplatform satellite analysis (Kossin et al. 2007; Knaff et al. 2011), but these methods have not yet been fully incorporated into NHC operations. Over the open ocean, the wind radii best tracks are likely to have an uncertainty of around 40, 30, and 25 n mi for 34-, 50-, and 64-kt winds (Landsea and Franklin 2013). Given that these large uncertainties are on the order of about one-third to one-half of the values they are depicting, routinely verifying NHC size forecasts with such limited verification data is not justifiable at this time.

Within the Atlantic basin (i.e., the North Atlantic Ocean, Gulf of Mexico, and Caribbean Sea), however, the availability of aircraft reconnaissance with accurate flight-level winds and Stepped Frequency Microwave Radiometer surface winds (Uhlhorn et al. 2007; Klotz and Uhlhorn 2014) allows for improved assessments of TC size. Landsea and Franklin (2013) estimate that the wind radii best tracks have a reduced uncertainty—around 30, 25, and 15 n mi for 34-, 50-, and 64-kt winds, respectively—when aircraft reconnaissance data are available. Thus, to mitigate the lack of reliable data for the whole population of forecasts, this study focused on the cases when reconnaissance aircraft were investigating a TC, which provides improved measures of TC size.

The results of this verification can help address questions commonly asked in regard to the wind radii forecasts. Some of these questions are as follow:

- How accurate are the NHC wind radii forecasts? Are they skillful when compared to the climatological and persistence model?
- Do the dynamical models have skill in predicting wind radii? Do they have the potential to guide forecasters?

The first two questions were addressed by Knaff and Sampson (2015) for the entire basin, but the accuracy of the wind radii at the initial and verification time was limited. This paper attempts to answer these questions within the context of a subset of the model and NHC's wind radii forecasts for which better ground truth is available.

2. Methodology

To address the questions noted above, we used a reconnaissance dataset that contains only cases when either the U.S. Air Force C-130s or the NOAA Orion P-3s were conducting center fix missions both at the initial and verification times. This sample is viewed as being relatively well surveyed, and the NHC best tracks for these TCs will contain the most accurate wind radii information routinely available today. Despite the fact that the Atlantic basin is the only one globally with routine aircraft

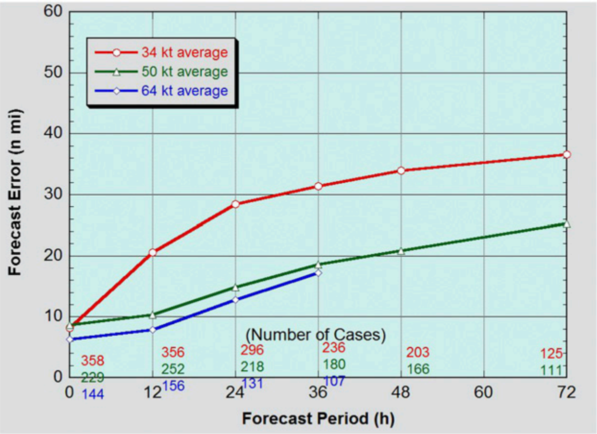


FIG. 2. NHC errors for the 34-, 50-, and 64-kt wind radii for the reconnaissance-only dataset.

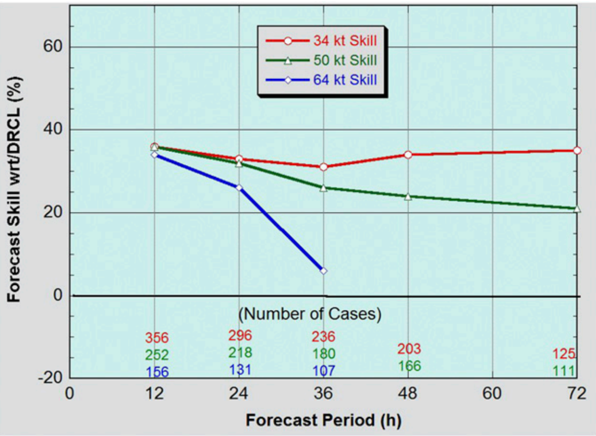


FIG. 3. NHC wind radii skill for the reconnaissance-only dataset.

reconnaissance currently available, only about 30% of the time is there an aircraft providing center fixes (Rappaport et al. 2009). This is a result of many TCs remaining over the open ocean away from land and having limited aircraft reconnaissance resources. The reconnaissance dataset we used extends from 2008 through 2012, which was a generally active period for TC development in the Atlantic basin, in order to have a meaningfully large sample available for verification.

Models operationally available for wind radii forecasting include a simple climatology and persistence statistical model, the DeMaria Climatology and Persistence Model Intensity Forecast (DRCL; Knaff et al. 2007), as well as a few dynamical models: the European Centre for Medium-Range Weather Forecasts (ECMWF) model (EMXI/EMX; ECMWF 2012), the interpolated Global Forecast System (GFS) model (GFSI; Kanamitsu 1989), and the interpolated Hurricane Weather Research and Forecasting (HWRF) model (HWFI; Tallapragada et al. 2013). Guidance models are characterized as either ''early'' or ''late,'' depending on whether or not they are available to the forecaster during the forecast cycle. For example, consider the 1200 UTC forecast cycle, which begins with the 1200 UTC synoptic time and ends with the release of an official forecast at 1500 UTC. The 1200 UTC run of the GFS model is not complete and available to the forecaster until about 1600 UTC, or about an hour after the NHC forecast is released. Consequently, the 1200 UTC GFS would be considered a late model since it could not be used to prepare the 1200 UTC official forecast.

This report focuses on the verification of the models mentioned above (early and late) and the official NHC forecast. Multilayer dynamical models are always late models. Fortunately, a technique exists to take the most recent available run of a late model and adjust its forecast to apply to the current synoptic time and initial conditions. In the example above, forecast data for hours 6–126 from the previous (0600 UTC) run of the GFS would be smoothed and then adjusted, or shifted, such that the 6-h forecast (valid at 1200 UTC) would match the observed 1200 UTC wind radii of the TC. The adjustment process creates an early version of the GFS model for the 1200 UTC forecast cycle that is based on the most current available guidance (initialized at 0600 UTC). The adjusted versions of the late models are known, mostly for historical reasons, as interpolated models. The adjustment algorithm is invoked as long as the most recent available late model is not more than 12 h old; for example, a 0000 UTC late model could be used to form an interpolated model for the subsequent 0600 or 1200 UTC forecast cycles, but not for the subsequent 1800 UTC cycle (Cangialosi and Franklin 2015).

TABLE 1. The 95% confidence intervals of the mean skill of the NHC official wind radii forecasts for the reconnaissance dataset.

Forecast time (h)	34-kt confidence interval (%)	50-kt confidence interval (%)	64-kt confidence interval (%)
12	30.1–39.9	29.1–40.9	25.6–40.4
24	28.1–38.9	26.2–38.7	18.5–33.5
36	25.6–37.4	21.0–34.0	1.5–10.5
48	28.4–41.6	17.5–30.5	—
72	27.1–43.9	13.4–28.6	—

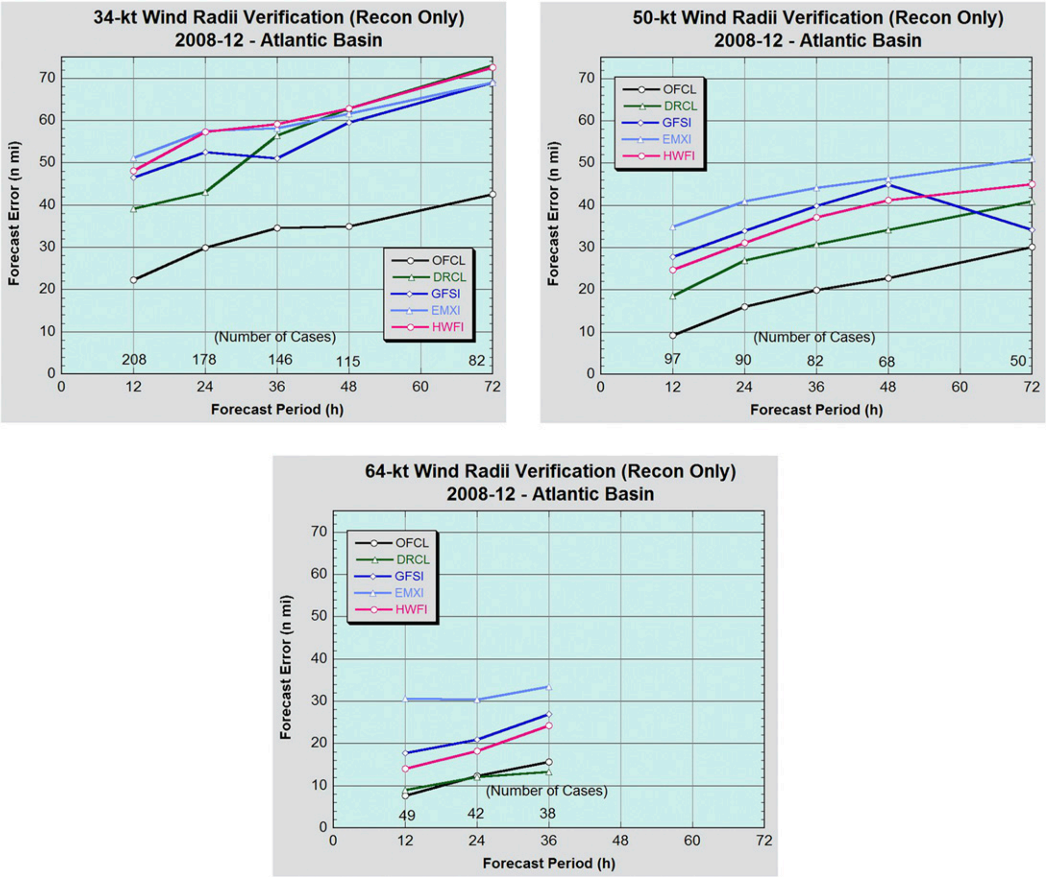


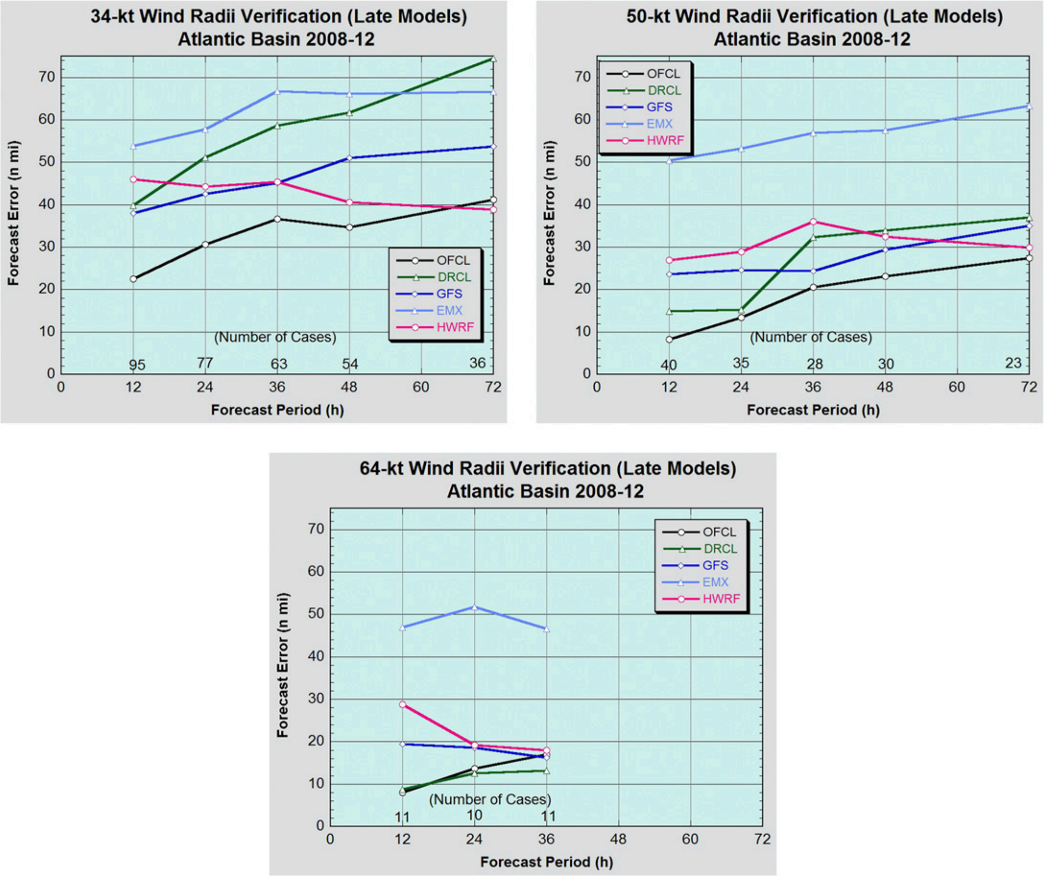
FIG. 4. NHC and early model errors for the 34-, 50-, and 64-kt wind radii for the reconnaissance-only dataset.

Verification procedures here make no distinction between 6- and 12-h interpolated models. Note that for simplicity, the results illustrated throughout this article are averages of the four quadrants for each of the wind radii thresholds. Additionally, nonexistent wind radii in the best track are treated as missing values (i.e., not a forecast). That is, if the best-track intensity is less than 64 kt, then no 64-kt radii verification is performed for that particular time. Note that one could alternatively include into the wind radii forecast verification cases where there is no best-track wind radius as a result of the best-track intensity being below that particular wind radii threshold. However, it is not believed that this would appreciably change the results obtained. Similarly, if the forecast maximum wind speed was for 60 kt, there were no 64-kt radii forecasts made. Assuming that the forecast and verifying maximum wind speeds are both above the threshold, the radii are simply differenced. A forecast of 60 n mi with an actual result of 0 n mi is a 60 n mi error.

As is common practice within the tropical cyclone community, skill in forecasts is determined by improvement (or degradation) against a climatology–persistence "no skill" prediction (Rappaport et al. 2009; Cangialosi and Franklin 2015). With respect to wind radii forecasts, the appropriate benchmark is DRCL (Knaff et al. 2007). In some of the analyses presented below (e.g., Fig. 3), the results are in the form of a skill diagram with errors shown relative (percent larger or smaller) to DRCL.

3. Results

Figure 2 shows the NHC wind radii verification of the 34-, 50-, and 64-kt wind radii forecasts for the reconnaissance dataset. For each wind speed threshold (34, 50, and 64 kt), the average errors increase with forecast time. The 34-kt wind radii average forecast errors range from about 8 n mi at the initial time to 36 n mi at 72 h. Errors at the initial time are a result of changes from the operational estimates of the wind radii analyses to the best-track values. The 50-kt wind radii errors are smaller than the 34-kt wind radii errors, and range from about 8 n mi at the initial time to about 25 n mi at 72 h.


34-kt Wind Radii Verification (Late Models) Atlantic Basin 2008-12

50-kt Wind Radii Verification (Late Models) Atlantic Basin 2008-12

64-kt Wind Radii Verification (Late Models) Atlantic Basin 2008-12

FIG. 5. As in Fig. 4, but for late models.

The smallest errors exist for the 64-kt wind radii, and these errors range from about 6 n mi at the initial time to about 18 n mi at 36 h. It is expected that the average errors decrease for the higher-value wind radii thresholds, given that the 50- and 64-kt wind radii are usually notably smaller than the 34-kt wind radii. For context, the long-term average values for 34-, 50-, and 64-kt wind radii are 120, 76, and 50 n mi, respectively (Kimball and Mulekar 2004).

Figure 3 illustrates that the NHC forecasts are skillful when compared against DRCL. The 34-kt wind radii errors are the most skillful with skill values ranging from about 30% to 35% from 12 to 72 h. Skill decreases with increasing wind radii thresholds. The NHC 50-kt wind radii skill values are around 35% at 12 h but fall to around 20% at 72 h. The hurricane-force wind radii skill values are a little over 30% at 12 h but decrease to only about 5% at 36 h. These skill values are statistically significant at the 95% confidence level (Table 1).

The verification of the early models that are primarily used as guidance for the 34-, 50-, and 64-kt wind radii forecasts is shown in Fig. 4. For the 34- and 50-kt wind thresholds, the NHC forecasts are far superior to the guidance with errors almost half the size of the models. The dynamical models shown are similar to, or have larger errors than, DRCL, implying that they are not skillful on average. Among the dynamical models, GFSI performed best at 34 kt, and HWFI was the best model for the 50-kt radii, except at 72 h, where it was bested by GFSI. Regarding the 64-kt wind radii, the NHC forecast errors are still considerably lower than the dynamical models but are similar to those of DRCL. Since forecasters and other users infer size directly from the model fields, a verification of the late models is also shown in Fig. 5. Notice that the sample size of the late model verification is considerably smaller than the early models since the ECMWF is only run twice per day. Similar to the early model verification, the NHC forecasts are better than all of the available guidance for the 34- and 50-kt wind radii size predictions. However, unlike the early models, the late versions of the GFS and HWRF are skillful when compared to DRCL at most forecast times for the 34- and

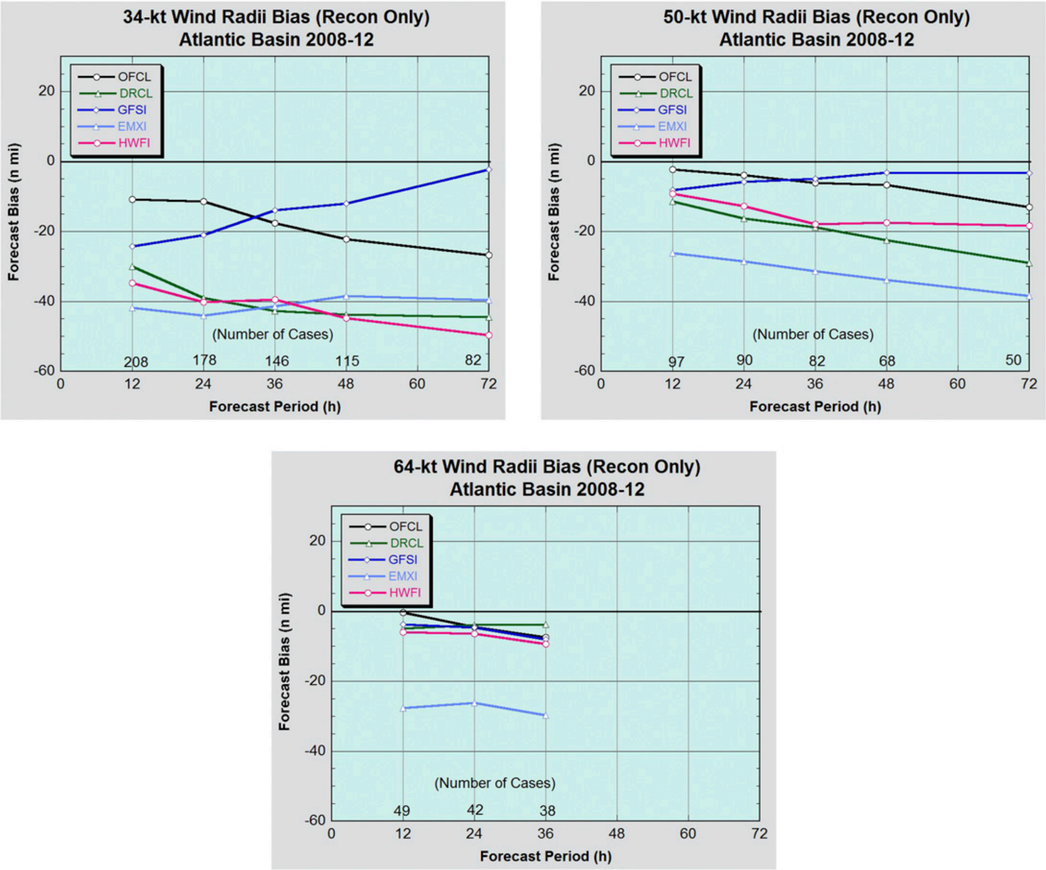


FIG. 6. NHC and early model bias for the 34-, 50-kt, and 64-kt wind radii for the reconnaissance-only dataset.

50-kt wind radii forecasts, but these models are not skillful for the size of the hurricane-force wind predictions. EMX has the largest errors at all radii forecasts and is not skillful for any wind threshold.

Figure 6 shows the wind radii bias for the three wind speed thresholds. For the 34-kt wind radii, the NHC forecasts have a negative bias (forecast radii are too small) that becomes slightly worse with longer forecast times. The DRCL, HWFI, and EMXI models have a more substantial negative bias, while GFSI has a slight negative bias that decreases with forecast time. In general, the same pattern exists for the 50-kt wind radii bias, with DRCL, HWFI, and EMXI having a substantial negative bias with the NHC forecasts and GFSI showing less bias. The NHC forecasts and the guidance all have negligible bias for the hurricane-force wind radii, with the exception of EMXI, which has a substantial negative bias. A verification of the wind radii bias for the late models is shown in Fig. 7. There are some notable differences between the early and late model results. The negative biases in the early versions of the GFS and HWRF models are replaced with neutral to slightly positive biases in the late versions. Conversely, the negative bias in the early version of the ECMWF was even more strongly negative in the late version of that model.

4. Summary

This study produced a formal verification of a subset of the NHC wind radii forecasts and selected guidance. The verification of the reconnaissance-only dataset, which was used to get the most accurate ‘‘ground truth’’ information, showed that the NHC wind radii average errors increased with forecast time and were skillful when compared against climatology and persistence. The 34-kt wind radii forecast errors presented here for the reconnaissance-based sample are about 10% larger than those reported by Knaff and Sampson (2015) for the entire basin. Given that Knaff and Sampson (2015) had similar verification rules as used here, this small difference may be due to the reconnaissance-based wind radii best tracks containing more variability (i.e., less smoothed than wind radii best tracks based on satellite data only). The dynamical models, however, were generally not skillful and had errors that were much larger than the NHC

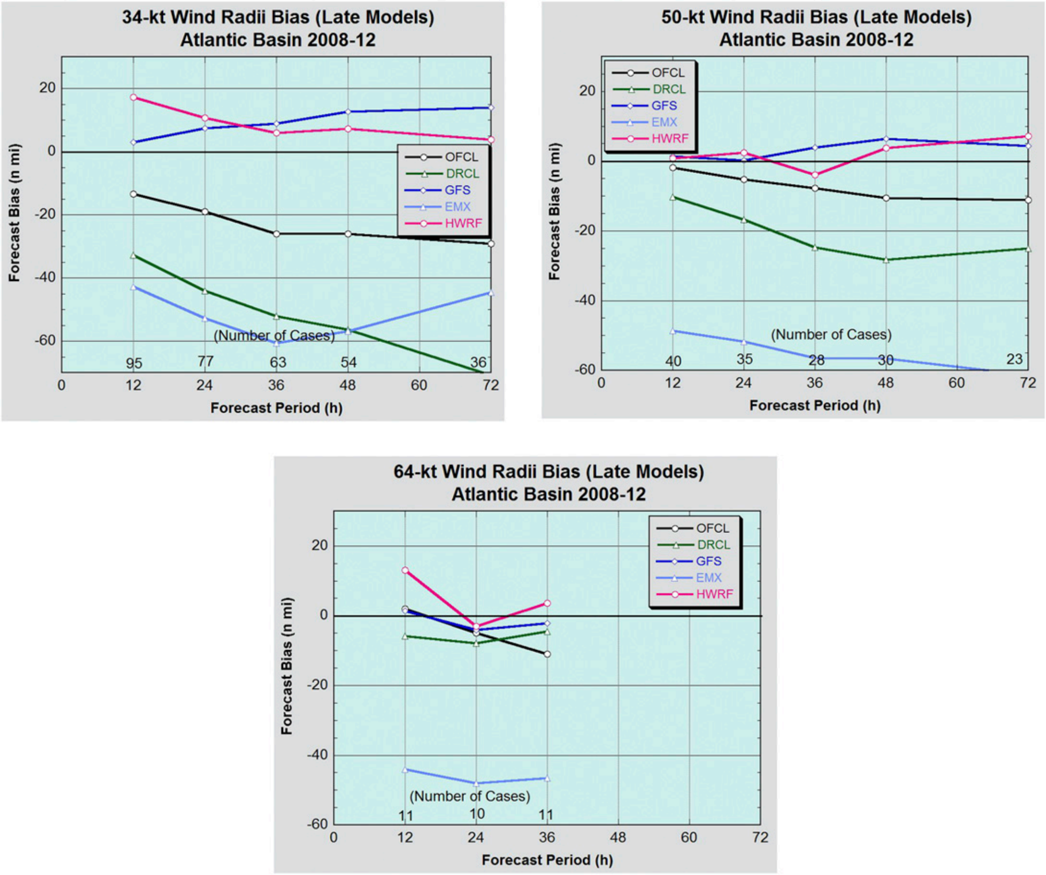


FIG. 7. As in Fig. 6, but for late models.

forecasts and mainly had negative biases. It is not known why the dynamical models have difficulty in providing skillful wind radii predictions. A better understanding of this limitation would be of substantial assistance in operations and potentially could help in improving the models' predictions of wind radii. We believe that the NHC wind radii forecasts outperform the guidance since the forecasters use a consensus or a blend of the guidance and apply their own understanding of TC characteristics.

It is worth noting, however, that the magnitude of these NHC wind radii errors—especially for the short term (12- and 24-h forecasts) for this reconnaissance-only verification—were about the same size or smaller than the uncertainty in the best-track values themselves (Landsea and Franklin 2013). Thus, until the accuracy of the wind radii best tracks improves, it may be difficult for the errors in NHC's wind radii forecasts at these short-term lead times to decrease. As such, we encourage continued development of observational techniques, which may better assist efforts for both operational and best-track assessments of the tropical storm-force and hurricane-force wind radii. In addition, NHC forecasts of wind radii can be improved by better guidance being made available to the forecasters. This would include improved explicit representations of the TC's wind field in both global and mesoscale hurricane models, statistical–dynamical approaches, as well as consensus techniques (Sampson and Knaff 2015).

Acknowledgments. This work took place at the National Hurricane Center. The authors would like to James Franklin, Mark DeMaria, and Edward Rappaport for reviewing this work.

REFERENCES

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, doi:10.1175/2008WAF2222188.1.

Cangialosi, J. P., and J. L. Franklin, 2015: 2014 National Hurricane Center Forecast Verification Report. NOAA/National Hurricane Center, 82 pp. [Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2014.pdf.]

Demuth, J. L., M. DeMaria, and J. A. Knaff, 2006: Improvement of Advanced Microwave Sounding Unit tropical cyclone intensity and size estimation algorithms. *J. Appl. Meteor. Climatol.*, **45**, 1573–1581, doi:10.1175/JAM2429.1.

ECMWF, 2012: ECMWF annual reports. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2013/16286-annual-report-2012.pdf.]

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. *Wea. Forecasting*, **4**, 335–342, doi:10.1175/1520-0434(1989)004<0335:DOTNGD>2.0.CO;2.

Kimball, S. K., and M. S. Mulekar, 2004: A 15-year climatology of North Atlantic tropical cyclones. Part I: Size parameters. *J. Climate*, **17**, 3555–3575, doi:10.1175/1520-0442(2004)017<3555:AYCONA>2.0.CO;2.

Klotz, B. W., and E. W. Uhlhorn, 2014: Improved stepped frequency microwave radiometer tropical cyclone surface winds in heavy precipitation. *J. Atmos. Oceanic Technol.*, **31**, 2392–2408, doi:10.1175/JTECH-D-14-00028.1.

Knaff, J. A., and C. R. Sampson, 2015: After a decade are Atlantic tropical cyclone gale force wind radii forecasts now skillful? *Wea. Forecasting*, **30**, 702–709, doi: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, doi:10.1175/WAF1026.1.

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

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, doi:10.1175/WAF985.1.

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, doi:10.1175/MWR-D-12-00254.1.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, doi:10.1175/2008WAF2222128.1.

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

Tallapragada, V., and Coauthors, 2013: Hurricane Weather Research and Forecasting (HWRF) model: 2013 scientific documentation. Developmental Testbed Center, 99 pp. [Available online at http://www.dtcenter.org/HurrWRF/users/docs/.]

Uhlhorn, E. W., P. G. Black, J. L. Franklin, M. Goodberlet, J. Carswell, and A. S. Goldstein, 2007: Hurricane surface wind measurements from an operational Stepped Frequency Microwave Radiometer. *Mon. Wea. Rev.*, **135**, 3070–3085, doi:10.1175/MWR3454.1.

Yueh, S. H., B. W. Stiles, and W. T. Liu, 2003: QuikSCAT wind retrievals for tropical cyclones. *IEEE Trans. Geosci. Remote Sens.*, **41**, 2616–2628, doi:10.1109/TGRS.2003.814913.

Zeng, L., and R. A. Brown, 1998: Scatterometer observations at high wind speeds. *J. Appl. Meteor.*, **37**, 1412–1420, doi:10.1175/1520-0450(1998)037<1412:SOAHWS>2.0.CO;2.