# Verification of Tropical Cyclone Genesis Forecasts from Global Numerical Models: Comparisons between the North Atlantic and Eastern North Pacific Basins

DANIEL J. HALPERIN

*Department of Atmospheric and Environmental Science, University at Albany, State University of New York, Albany, New York*

HENRY E. FUELBERG AND ROBERT E. HART

*Department of Earth, Ocean, and Atmospheric Science, Florida State University, Tallahassee, Florida*

JOSHUA H. COSSUTH

*Naval Research Laboratory, Monterey, California*

(Manuscript received 12 November 2015, in final form 16 February 2016)

ABSTRACT

Accurately forecasting tropical cyclone (TC) genesis is an important operational need, especially since the National Hurricane Center's Tropical Weather Outlook product has been extended from 2 to 5 days. A previous study by the coauthors verified North Atlantic TC genesis forecasts from five global models out to 4 days during 2004–11. This study expands on the previous research by 1) verifying TC genesis forecasts over both the Atlantic and eastern North Pacific basins, 2) extending the forecast window to 5 days, and 3) updating the analysis period through 2014. Verification statistics are presented and compared between the two basins. Probability of detection and critical success indices generally are greater over the eastern North Pacific basin compared to the North Atlantic. There is a trade-off between models that exhibit a greater probability of detection and a greater false alarm ratio, and models that exhibit a smaller false alarm ratio and a smaller probability of detection. Results also reveal that the models preferentially miss TCs over the North Atlantic (eastern North Pacific) that have a relatively small radius of the outer closed isobar (radius of maximum wind) at the forecast genesis time. Overall, global models have become a more reliable source of TC genesis guidance during the past few years compared to the early years in the dataset.

## 1. Introduction

Global numerical weather prediction models are a primary source of guidance for operational tropical cyclone (TC) genesis forecasts. Numerous studies have investigated the degree to which model-indicated TC genesis forecasts are reliable (e.g., Briegel and Frank 1997; Chan and Kwok 1999; Cheung and Elsberry 2002; Pratt and Evans 2009; Tsai et al. 2011; Halperin et al. 2013). Halperin et al. (2013, hereafter H13) verified historical model-indicated TC genesis forecasts out to 96 h over the North Atlantic (NATL) basin during 2004–11. They compared the performance of five global models and found notable spatial and temporal biases in each. The goal of the present study is to expand on H13 so the spatial and temporal domains are consistent with those of the National Hurricane Center's (NHC) Tropical Weather Outlook (TWO) product. That is, genesis forecasts are verified out to 120 h for the NATL and eastern North Pacific (EPAC) basins.

We compare results from Environment Canada's (now known as Environment and Climate Change Canada) Global Environmental Multiscale Model (CMC; Côté et al. 1998a,b), the European Centre for Medium-Range Weather Forecasts global model (ECMWF 2016), the National Centers for Environmental Prediction Global Forecast System (GFS; Kanamitsu 1989), and the Met Office global model (UKMET; Cullen 1993). The Navy Operational Global

*Corresponding author address*: Daniel J. Halperin, Dept. of Atmospheric and Environmental Sciences, University at Albany, State University of New York, 1400 Washington Ave., Albany, NY 12222.
E-mail: danieljhalperin@gmail.com

DOI: 10.1175/WAF-D-15-0157.1

© 2016 American Meteorological Society
Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

TABLE 1. TC genesis threshold values for each model for the NATL (EPAC) basin.

| Model | 850-hPa max $\zeta$ ($\times 10^{-5}$ s$^{-1}$) | 250–850-hPa max thickness (m) | 925-hPa max wind speed (m s$^{-1}$) |
|---|---|---|---|
| CMC | 15.3 (14.9) | 9478.5 (9483.3) | 17.0 (15.8) |
| ECMWF | 20.3 (20.6) | 9473.0 (9491.3) | 15.4 (15.9) |
| GFS | 13.9 (12.7) | 9471.1 (9480.2) | 16.1 (13.7) |
| UKMET | 14.4 (14.2) | 9469.3 (9490.0) | 15.5 (15.1) |

Atmospheric Prediction System (NOGAPS; Rosmond 1992), which was included in H13, is not considered here because the model was decommissioned in 2013. The Navy Global Environmental Model (NAVGEM; Hogan et al. 2014), which replaced NOGAPS, is not included because there is an insufficient sample of genesis forecasts to verify. Select model upgrades since 2012 are included in the appendix. Model upgrades during 2004–11 are given in appendix A of H13.

## 2. Methodology

Model data were available during 2004–14 (2007–14 for the ECMWF). The definition of a TC in the model forecast fields, including the threshold values for maximum 850-hPa relative vorticity $\zeta$, 250–850-hPa thickness, and 925-hPa wind speed were the same as defined in H13 for the NATL. Threshold values for the EPAC basin were calibrated separately using the H13 methodology (Table 1). Except for the ECMWF, the 850-hPa $\zeta$ and 925-hPa wind speed maxima are weaker over the EPAC compared to the NATL. TCs generally are smaller over the EPAC (Chavas and Emanuel 2010; Knaff et al. 2014); thus, the models may have been unable to resolve the observed maximum values. Conversely, the higher-resolution ECMWF may have been able to capture the observed maximum values, thus yielding larger threshold values over the EPAC. The thickness threshold is larger over the EPAC compared to the NATL for all models. This likely is the result of the EPAC genesis events originating at a more equatorward latitude compared to the NATL.

The verification criteria employed here differ slightly from those defined in H13. Here, a successful genesis forecast (i.e., ''hit'') is defined when best-track (Jarvinen et al. 1984; McAdie et al. 2009; Landsea and Franklin 2013) TC genesis occurred within 120 h of the model initialization time and when the model forecast genesis location was within 5° latitude and longitude of the best-track location at the corresponding time. Torn and Snyder (2012) and Landsea and Franklin (2013) discuss the uncertainties in the best-track data. For model genesis forecasts with valid times prior to the best-track genesis time, combined automated response to query (CARQ) entries in the Automated Tropical Cyclone Forecasting (ATCF) system a-deck files (Sampson and Schrader 2000) were used to verify the forecast TC location. CARQ entries are the hurricane specialists' operational initial best estimates of various parameters of a tropical disturbance or cyclone (e.g., location, intensity, etc.) (H13). A genesis forecast that did not result in best-track genesis was classified as a false alarm (FA). This methodology is better aligned with NHC's 120-h TWO forecast verification method, and it results in a greater temporal tolerance compared to the definition of a hit in H13. Thus, it is expected that success ratios will be greater in the present study than in H13.

The 0000 and 1200 UTC initializations of each model were analyzed with a forecast window of 120 h. Thus, there were 10 opportunities for a model to predict the genesis of each best-track TC (i.e., two model runs per day for 5 days). A miss was defined when a model had the opportunity, but did not predict genesis for a best-track TC.

As in H13, the probability of detection (POD), success ratio (SR), bias, and critical success index (CSI) were calculated using the following definitions:

$$\mathrm{POD} = \frac{\mathrm{hit}}{\mathrm{hit} + \mathrm{miss}}, \tag{1}$$

$$\mathrm{SR} = 1 - \frac{\mathrm{false\ alarm}}{\mathrm{hit} + \mathrm{false\ alarm}}, \tag{2}$$

$$\mathrm{bias} = \frac{\mathrm{hit} + \mathrm{false\ alarm}}{\mathrm{hit} + \mathrm{miss}}, \quad \text{and} \tag{3}$$

$$\mathrm{CSI} = \frac{\mathrm{hit}}{\mathrm{hit} + \mathrm{false\ alarm} + \mathrm{miss}}. \tag{4}$$

Greater values of POD, SR, and CSI indicate better model-indicated TC genesis forecasts. Bias values greater (less) than unity indicate the model is over- (under-) forecasting TC genesis. If the model-indicated TC genesis forecasts were perfect, all metrics in Eqs. (1)–(4) would equal unity.

## 3. Results

### *a. North Atlantic*

Several metrics were employed to assess the performance of the model-indicated genesis forecasts. A comparison of the yearly mean SR over the NATL is presented in Fig. 1a.

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

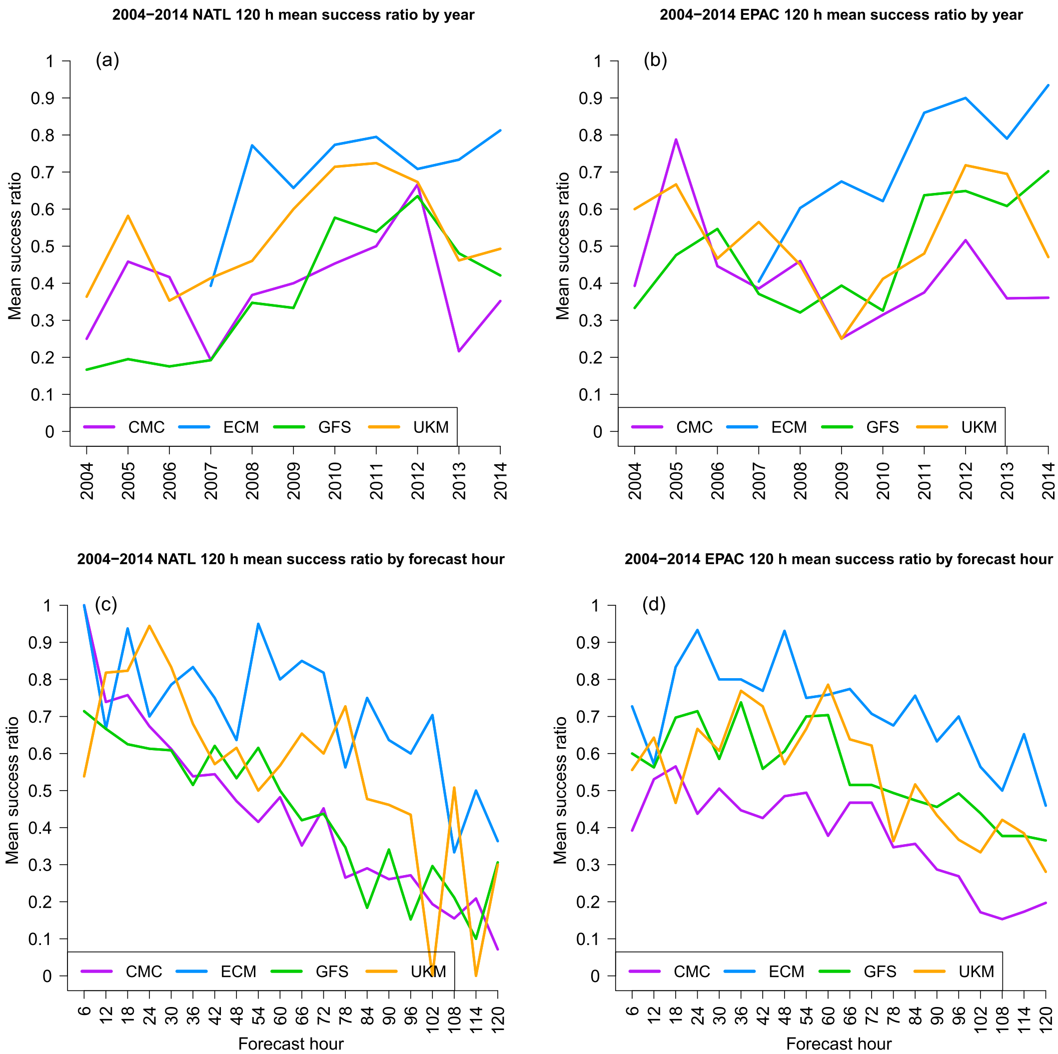


FIG. 1. Mean success ratio of model-indicated TC genesis forecasts within 120 h of the initialization time (a) by year for the NATL basin, (b) by year for the EPAC basin, (c) by forecast hour for the NATL basin, and (d) by forecast hour for the EPAC basin.

The ECMWF ranks best since 2008 over the NATL. The GFS and UKMET exhibit their greatest SR values during 2010–12, with some degradation in SR during 2013–14. H13 showed that the GFS and UKMET preferentially forecast genesis over the main development region. The relatively less active seasons of 2013–14, especially over the main development region, may explain some of the observed degradation.

H13 revealed that mean SR decreases as forecast hour increases. This result is useful and highly intuitive from an operational perspective, as it indicates that forecasters should have more confidence in a TC genesis forecast that occurs early in the forecast cycle compared to one that occurs after several days of model integration. Figure 1c shows the mean SR by forecast hour over the NATL. After forecast hour 72, SR is less than 0.5 for GFS and CMC genesis forecasts over the NATL (Fig. 1c). The decline is not as dramatic for the ECMWF and UKMET models. By forecast hour 120, SR is less than 0.4 for all models over the NATL. Note that UKMET data at forecast hours 102 and 114 only were available during 2014, and that the sample size at those forecast hours is quite small. The SR at forecast hours 6–12 is smaller than one might expect for the GFS and UKMET. The sample size at these forecast hours is relatively less than at longer lead times. Thus, only a few false alarm events can result in small SRs. Some of the false alarms at the 6–12-h lead times are of disturbances that never develop. Others, however, would correspond to a best-track TC, but genesis

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

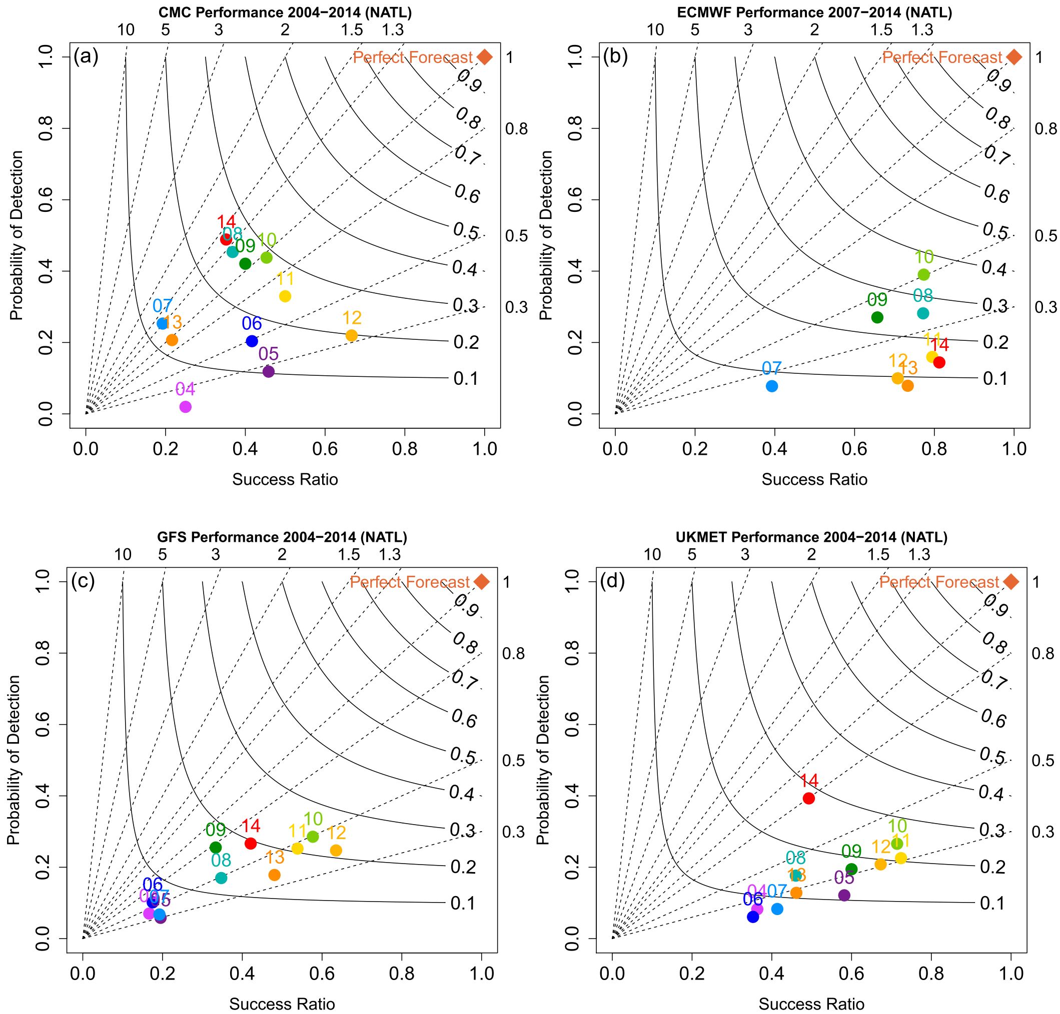


FIG. 2. Performance diagrams for the (a) CMC, (b) ECMWF, (c) GFS, and (d) UKMET for the NATL basin. Success ratio is given along the $x$ axis; probability of detection is along the $y$ axis. Bias values are indicated by the dashed lines, and CSI values are indicated by the curved, solid lines. A ''perfectly'' performing model would be in the top-right corner of the plot. Genesis events from all forecast hours (6–120) are included. Mean statistics for each year are plotted with the two-digit year above each point. Cool (warm) colors indicate the earlier (later) years in the dataset.

is predicted in the model before the system is tracked in the ATCF a or b decks. Thus, there are insufficient verification data to match the forecast TC with an observed TC (or its precursor disturbance). This contributes to the small SR observed at 6–12-h lead times for the GFS and UKMET.

While SR is useful for conducting a simple model comparison, additional metrics such as POD and CSI should be considered. CSI is particularly useful for situations where correct null events far outnumber hit, false alarm, and miss events (Wilks 2006). It is difficult to quantify correct null events in this situation, but it can be inferred that the correct null outcome is dominant since the base state of the atmosphere is for TCs to not exist.

Performance diagrams (Roebber 2009) display several verification metrics over the NATL (Fig. 2). The CMC has a relatively small SR, but a relatively large POD compared to the other models (Fig. 2a). The CMC also exhibits a large range of yearly mean CSI values. With the exception of 2007, the ECMWF consistently has a large SR (~0.65–0.82). The POD has decreased to values below 0.2 during 2011–14 (Fig. 2b). The GFS performance has improved since 2008 compared to 2004–07 (Fig. 2c). While its SR varies from year to year, the POD recently has been nearly constant at ~0.2–0.25. The performance of the UKMET model (Fig. 2d) is interesting. Bias is approximately constant throughout 2004–13, with values near 0.3. The SR decreases during

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

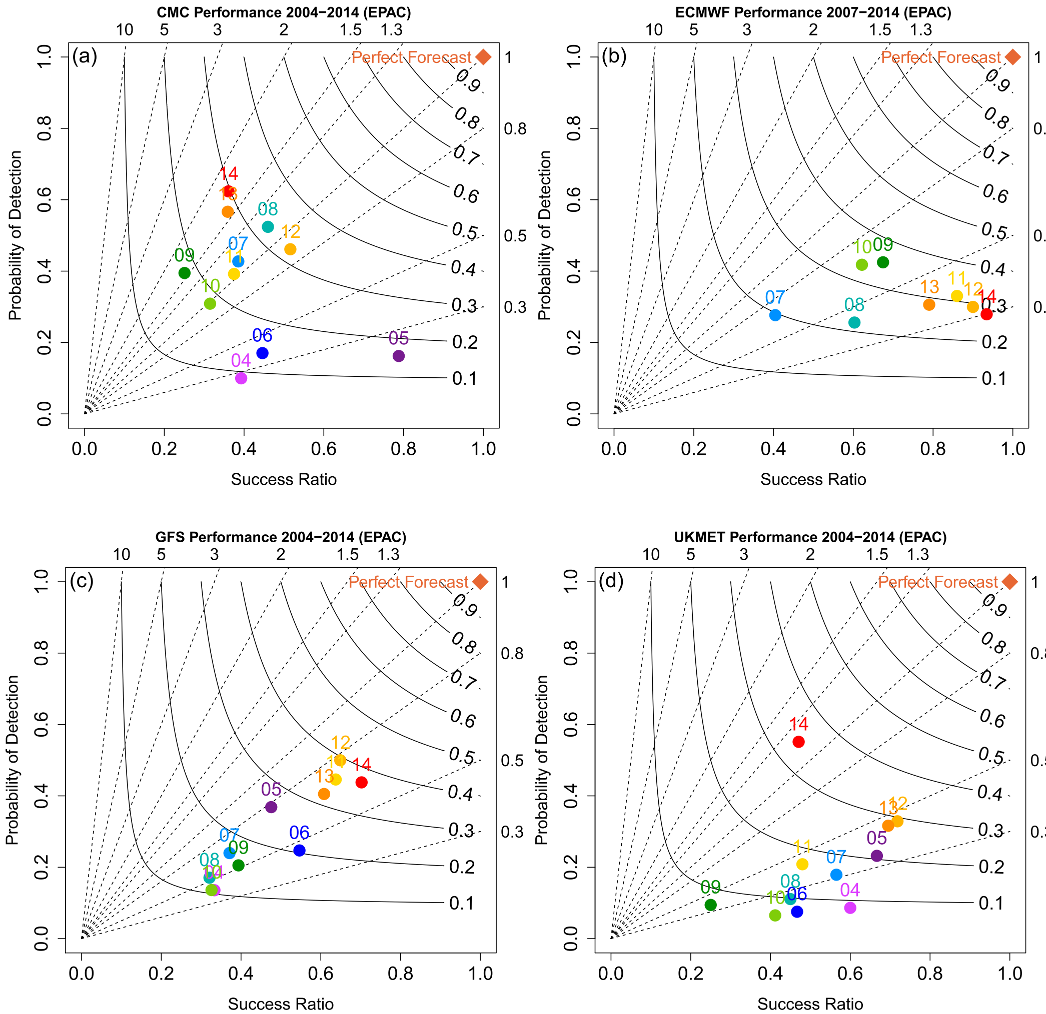


FIG. 3. As in Fig. 2, but for the EPAC basin.

2014 relative to 2009–12. Nevertheless, the increase in POD during 2014 results in the greatest CSI during 2014. A model upgrade may have contributed to the changed level of performance of the UKMET during 2014. Heming (2014) found that the upgraded UKMET model was able to intensify TCs in the forecast fields more than previous configurations. While genesis was not explicitly examined, it is possible that the model also was more prone to intensifying weak disturbances into TCs, thus increasing the overall number of genesis forecasts. CSI scores for all models rarely exceed 0.3. The single best-performing model and year is the ECMWF during 2010, with a CSI near 0.35.

The relatively small POD values over the NATL are due mostly to TCs that the models miss (generally smaller TCs). However, the small POD also is partly due to the models rarely being able to capture genesis in every forecast cycle during the 5 days before genesis occurs. For example, if a model correctly forecasts genesis 6 times out of 10 during the 5 days leading up to best-track genesis, that generally would be considered a successful forecast, but it only yields a POD of 0.6.

### *b. Comparison of eastern North Pacific with North Atlantic*

The EPAC verification statistics share some commonalities with those of the NATL. For example, ranking the models by yearly mean SR reveals that the ECMWF again performs best since 2008 (Fig. 1b). In fact, the ECMWF has been quite reliable in predicting genesis during 2011–14 over the EPAC (SR $\geq$ 0.75). Its POD and CSI also are larger than over the NATL during this time period. Meanwhile, the CMC (Fig. 3a) exhibits some of the greatest POD values (~0.6), but has a smaller mean SR value than the other models. The

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

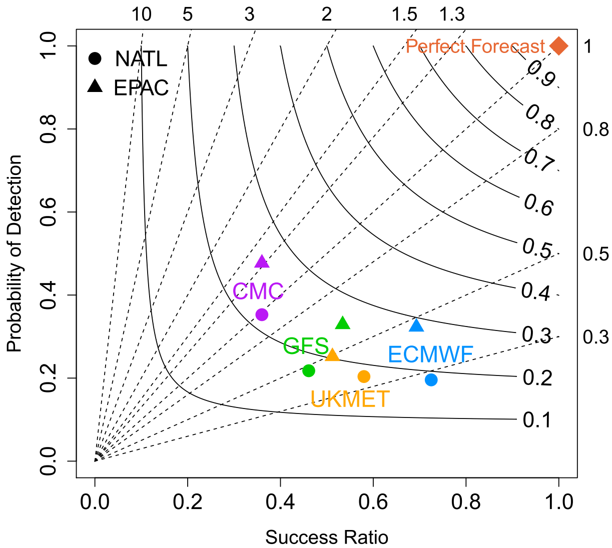


FIG. 4. Performance diagram of 2007–14 mean forecast performance for each model over the NATL (circle) and EPAC (triangle) basins.

GFS is the only model to exhibit CSI values greater than 0.3 during each of the past four years. The UKMET exhibits its greatest SR during 2012 and 2013. As over the NATL, 2014 performance shows a decrease in SR; however, the notable increase in POD results in the model's greatest yearly CSI. These results are consistent with findings from Heming (2014) (i.e., the UKMET simulates stronger TCs as a result of the 2014 upgrade). The best score for any model and year is the GFS during 2012 with a CSI of 0.39.

The EPAC also exhibits a decline in model performance with increasing forecast hour (Fig. 1d). However, the forecasts have a greater SR for slightly longer in the forecast cycle than over the NATL. For example, all models except the CMC have SRs near or exceeding 0.5 out to forecast hour 90. Over the NATL, this occurs only until forecast hour 60.

In addition to the aforementioned year-to-year variability, there also is basin-to-basin variability. A given model may perform well in one basin, but poorly in the other during the same year. For example, the degradation during 2013–14 observed over the NATL for the GFS is not apparent over the EPAC. This suggests that the lack of genesis events over the main development region may have contributed to the relatively poor performance over the NATL. The interbasin variability also implies that model upgrades alone do not determine performance.

The performance diagram of mean statistics during 2007–14 (Fig. 4) reveals that the mean CSI over the NATL is near 0.2 for each model. This implies that there is a trade-off between a relatively large POD but small SR (e.g., CMC), and a relatively small POD but large SR (e.g., ECMWF). Values of SR over the EPAC are comparable to those of the NATL, but the greater POD values over the EPAC lead to better CSI scores. McTaggart-Cowan et al. (2013) found that the nonbaroclinic TC genesis pathway is dominant over the EPAC, but a nontrivial percentage of TCs develop from baroclinic pathways over the NATL. The larger POD over the EPAC suggests that the models may better capture genesis from the nonbaroclinic pathway compared to the various baroclinic pathways of McTaggart-Cowan et al. (2013). It also is possible that our TC genesis criteria are calibrated in favor of TCs developing from the nonbaroclinic pathway.

TABLE 2. Median values of RMW and ROCI for NATL best-track TCs that were detected by the models at least once vs best-track TCs that never were detected by the models. Boldface font indicates the difference in medians is statistically significant at the 95% confidence interval according to a two-sided Wilcoxon rank sum test.

| Model | Median RMW (n mi) | | Median ROCI (n mi) | |
|---|---|---|---|---|
| | TCs detected | TCs not detected | TCs detected | TCs not detected |
| CMC | 60 | 50 | **175** | **150** |
| ECMWF | **60** | **50** | **180** | **150** |
| GFS | 60 | 50 | **175** | **150** |
| UKMET | **60** | **45** | **175** | **150** |

### c. *Examination of miss events*

While the SR generally has improved over time, the same cannot be said for the POD. To determine possible explanations for the relatively small POD, we split our dataset of best-track TCs into two groups: 1) best-track TCs whose genesis was predicted at least once by a given model and 2) best-track TCs whose genesis was never predicted by the same model. Given the relatively coarse resolution of the models, TC size at genesis time is an obvious metric to test. The Wilcoxon rank sum test [equivalent to the Mann–Whitney $U$ test; Wilks (2006)] was used to compare the medians of the best-track radius of maximum wind (RMW) and the radius of the outer closed isobar (ROCI) at the time of model-indicated genesis. The results are presented in Tables 2 and 3. RMW and ROCI data at the genesis time are unavailable for many TCs during 2004.

All four models exhibit a statistically significant difference (at the 95% confidence level) between the medians of ROCI for NATL best-track TCs that are

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

TABLE 3. As in Table 2, but for the EPAC.

| Model | Median RMW (n mi) | | Median ROCI (n mi) | |
|---|---|---|---|---|
| | TCs detected | TCs not detected | TCs detected | TCs not detected |
| CMC | **45** | **30** | 150 | 150 |
| ECMWF | **50** | **45** | **172.5** | **150** |
| GFS | 45 | 40 | 150 | 150 |
| UKMET | **50** | **40** | 150 | 150 |

detected at least once versus NATL best-track TCs that never are detected by the models (Table 2). The detected TCs are larger, with median ROCIs of 175–180 n mi (1 n mi = 1.852 km). The detected TCs also are larger according to median RMW. However, the difference in medians only is significant for the ECMWF and UKMET models.

There are notable differences over the EPAC (Table 3) compared to the NATL. The median ROCI values are equal for TCs detected versus TCs not detected for the CMC, GFS, and UKMET. The ECMWF does exhibit a significant difference in median ROCI. Meanwhile, differences in median RMW are significant for all models except the GFS. It is interesting that median values for the detected TCs over the EPAC are approximately equivalent to median values for the nondetected TCs over the NATL. This likely is a result of the smaller average size of TCs over the EPAC (Chavas and Emanuel 2010; Knaff et al. 2014). Since at least one size metric generally is significant when comparing detected and missed TCs, it is likely that model resolution has some impact on the probability of detection.

## 4. Summary and conclusions

This study expanded on the work of H13. TC genesis forecasts from four global models (CMC, ECMWF, GFS, and UKMET) were verified out to 5 days over the NATL and EPAC basins during 2004–14. Each TC genesis forecast was classified as a hit or false alarm.

Results revealed both notable similarities and differences among the models and basins. The CMC generally exhibited the greatest year-to-year variability over both basins. It typically had the smallest success ratio, but the greatest probability of detection compared to the other models. Conversely, the ECMWF had the greatest mean success ratio, but the smallest mean probability of detection of the models examined. The mean critical success index over the NATL was near 0.2 for all four models, confirming the observed trade-off between success ratio and probability of detection. Furthermore, the similar mean critical success indices among the models may suggest a limit to the predictability of TC genesis over the NATL. The mean values of the critical success index over the EPAC generally are greater than over the NATL as a result of comparable success ratios and larger probabilities of detection. The success ratio expectedly decreases with increasing forecast hour over both basins, although the EPAC appears to have a smaller rate of decay.

Miss events briefly were investigated. It appears that a TC's size does impact whether it is detected by the models (larger TCs generally are more likely to be detected). There was a statistically significant difference between the median radius of the outer closed isobar for TCs that were detected at least once compared to TCs that were never detected by all four models over the NATL. All models except the GFS exhibited a significant difference in the medians of radius of maximum wind between detected and missed TCs over the EPAC.

Overall, the results indicate that the verification statistics improved during the second half of the study period compared to the first half, suggesting that global models have become a more reliable source of TC genesis forecast guidance. A future paper will describe how these verified TC genesis forecasts can be used to predict the genesis probability of real-time model forecast TCs.

*Acknowledgments.* The authors thank Julian Heming of the Met Office for providing the archived UKMET model output and upgrade information. ECMWF output was obtained from the TIGGE archive (ECMWF portal). We thank the AMS anonymous reviewers for their constructive feedback. The authors benefited from discussions on this topic with Jeff Chagnon at FSU and Joint Hurricane Testbed NHC points of contact, Eric Blake, Todd Kimberlain, Chris Landsea, Craig Mattocks, and Richard Pasch. This research was supported by NOAA Grant NA13OAR4590185.

## APPENDIX

### Select Model Upgrades since 2012

*a. CMC (Buehner et al. 2015)*

- February 2013—horizontal grid spacing changed from 0.45° × 0.3° to 0.35° × 0.23°.
- February 2013—introduced vertical staggering.
- 2014—changed to a four-dimensional ensemble–variational data assimilation system.
- 2014—analysis increment horizontal grid spacing changed from 0.9° × 0.9° to 0.45° × 0.45°.

*b. ECMWF (ECMWF 2016)*

- June 2012—modified convective downdraft entrainment.

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

- June 2012—modified cloud ice fall speed.
- June 2013—changed from 91 to 137 vertical levels.

*c. GFS (NCEP/EMC 2016)*

- May 2012—hybrid ensemble Kalman filter–three-dimensional variational (EnKF–3DVAR) data assimilation introduced.
- January 2015—horizontal resolution increased from T574 (~27 km) to T1534 (~13 km) out to 240 h.
- January 2015—changed from Eulerian to semi-Lagrangian dynamics.
- January 2015—switched from 1° Reynolds 7-day sea surface temperature (SST) analysis to 5′ daily real-time global SST.
- January 2015—reduced drag coefficient at high wind speeds.

*d. UKMET (Heming 2014; J. Heming 2016, personal communication)*

- July 2012—switched off the wind-based tropical cyclone initialization scheme.
- January 2013—hybrid data assimilation upgraded to use 44-member N400 global version of the Met Office Global and Regional Ensemble Prediction System (MOGREPS-G) dataset.
- July 2014—new dynamical core (ENDGame) implemented.
- July 2014—horizontal grid spacing changed from ~25 to ~17 km at midlatitudes.
- July 2014—horizontal grid spacing for data assimilation changed from ~60 to ~40 km.
- July 2014—increased entrainment rate in deep convection.
- July 2014—reduced turbulent mixing and revised stability functions in the boundary layer.
- July 2014—improved cloud erosion, improved treatment of cirrus, and smoother phase change for condensate detrained from convection.

REFERENCES


Briegel, L. M., and W. M. Frank, 1997: Large-scale influences on tropical cyclogenesis in the western North Pacific. *Mon. Wea. Rev.*, **125**, 1397–1413, doi:10.1175/1520-0493(1997)125<1397:LSIOTC>2.0.CO;2.

Buehner, M., and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part I: The global system. *Mon. Wea. Rev.*, **143**, 2532–2559, doi:10.1175/MWR-D-14-00354.1.

Chan, J. C., and R. H. Kwok, 1999: Tropical cyclone genesis in a global numerical weather prediction model. *Mon. Wea. Rev.*, **127**, 611–624, doi:10.1175/1520-0493(1999)127<0611:TCGIAG>2.0.CO;2.

Chavas, D. R., and K. A. Emanuel, 2010: A QuikSCAT climatology of tropical cyclone size. *Geophys. Res. Lett.*, **37**, L18816, doi:10.1029/2010GL044558.

Cheung, K. K., and R. L. Elsberry, 2002: Tropical cyclone formations over the western North Pacific in the Navy Operational Global Atmospheric Prediction System forecasts. *Wea. Forecasting*, **17**, 800–820, doi:10.1175/1520-0434(2002)017<0800:TCFOTW>2.0.CO;2.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998a: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395, doi:10.1175/1520-0493(1998)126,1373:TOCMGE.2.0.CO;2.

——, J.-G. Desmarais, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998b: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part II: Results. *Mon. Wea. Rev.*, **126**, 1397–1418, doi:10.1175/1520-0493(1998)126,1397:TOCMGE.2.0.CO;2.

Cullen, M., 1993: The unified forecast/climate model. *Meteor. Mag.*, **122**, 81–94.

ECMWF, 2016: ECMWF IFS documentation. [Available online at http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation.]

Halperin, D. J., H. E. Fuelberg, R. E. Hart, J. H. Cossuth, P. Sura, and R. J. Pasch, 2013: An evaluation of tropical cyclone genesis forecasts from global numerical models. *Wea. Forecasting*, **28**, 1423–1445, doi:10.1175/WAF-D-13-00008.1.

Heming, J. T., 2014: The impact on tropical cyclone predictions of a major upgrade to the Met Office global model. Preprints, *31st Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 11A.3. [Available online at https://ams.confex.com/ams/31Hurr/webprogram/Paper243428.html.]

Hogan, T. F., and Coauthors, 2014: The Navy Global Environmental Model. *Oceanography*, **27** (3), 116–125, doi:10.5670/oceanog.2014.73.

Jarvinen, B., C. Neumann, and M. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC 22, 21 pp. [Available online at http://www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.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.

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

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.

McAdie, C., C. Landsea, C. Neumann, J. David, E. Blake, and G. Hammer, 2009: Tropical cyclones of the North Atlantic Ocean, 1851–2006. Historical Climatology Series 6-2, NCDC, 238 pp. [Available online at http://www.nhc.noaa.gov/pdf/TC_Book_Atl_1851-2006_lowres.pdf.]

McTaggart-Cowan, R., T. J. Galarneau Jr., L. F. Bosart, R. W. Moore, and O. Martius, 2013: A global climatology of baroclinically influenced tropical cyclogenesis. *Mon. Wea. Rev.*, **141**, 1963–1989, doi:10.1175/MWR-D-12-00186.1.

NCEP/EMC, 2016: Global Forecast System: Recent implementations. [Available online at http://www.emc.ncep.noaa.gov/GFS/impl.php.]


Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC

Pratt, A. S., and J. L. Evans, 2009: Potential impacts of the Saharan air layer on numerical model forecasts of North Atlantic tropical cyclogenesis. *Wea. Forecasting*, **24**, 420–435, doi:10.1175/2008WAF2007090.1.

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

Rosmond, T. E., 1992: The design and testing of the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting*, **7**, 262–272, doi:10.1175/1520-0434(1992)007<0262:TDATOT>2.0.CO;2.

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

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

Tsai, H.-C., K.-C. Lu, R. L. Elsberry, M.-M. Lu, and C.-H. Sui, 2011: Tropical cyclone-like vortices detection in the NCEP 16-day ensemble system over the western North Pacific in 2008: Application and forecast evaluation. *Wea. Forecasting*, **26**, 77–93, doi:10.1175/2010WAF2222415.1.

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


Brought to you by NOAA Central Library | Unauthenticated | Downloaded 08/13/24 08:03 PM UTC