

# HAVE WE REACHED THE LIMITS OF PREDICTABILITY FOR TROPICAL CYCLONE TRACK FORECASTING?

CHRISTOPHER W. LANDSEA AND JOHN P. CANGIALOSI

Recent hurricane seasons in the Western Hemisphere suggest that improvements in track forecasting have slowed or perhaps even come to a halt.

**T**rack forecasting improvements for tropical cyclones (inclusive of subtropical cyclones in this study) have been one of the most remarkable accomplishments in meteorology, with errors decreasing by about two-thirds in just a generation. For example, in the Atlantic basin (including the North Atlantic Ocean, Gulf of Mexico, and Caribbean Sea) the National Hurricane Center's (NHC) 3-day forecast track error of the tropical cyclone's center<sup>1</sup> averaged about 300 n mi (1 n mi = 1.852 km) in 1990 compared with just 100 n mi in 2016 (Fig. 1; Cangialosi and Franklin 2017). Likewise, the 3-day

forecast track error in the eastern North Pacific (east of 140°W) dropped from 225 n mi in 1990 down to 75 n mi in 2016 (Fig. 1). Such changes are most assuredly due to increasingly realistic and detailed global modeling, advanced data assimilation techniques maximizing dramatically increased frequency, type, and density of geostationary and polar-orbiting satellite data, as well as regular use of multimodel track consensus techniques. These improvements have allowed NHC to begin publicly issuing 5-day predictions beginning in 2003 (Rappaport et al. 2009) as well as to extend the lead time for tropical storm and hurricane watches/warnings out an additional 12 h to 48 and 36 h, respectively, beginning in 2010.

At public presentations, it is common for NHC forecasters to highlight these incredible advancements. We also point out that a simple linear extrapolation of trends in the track errors would

**AFFILIATIONS:** LANDSEA AND CANGIALOSI—NOAA/NWS/NCEP/ National Hurricane Center, Miami, Florida

**CORRESPONDING AUTHOR:** Christopher W. Landsea, [chris.landsea@noaa.gov](mailto:chris.landsea@noaa.gov)

*The abstract for this article can be found in this issue, following the table of contents.*

DOI:10.1175/BAMS-D-17-0136.1

In final form 9 June 2018

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

<sup>1</sup> From the NHC Glossary: "Generally speaking, the vertical axis of a tropical cyclone, usually defined by the location of minimum wind or minimum pressure. The cyclone center position can vary with altitude. In advisory products, refers to the center position at the surface."

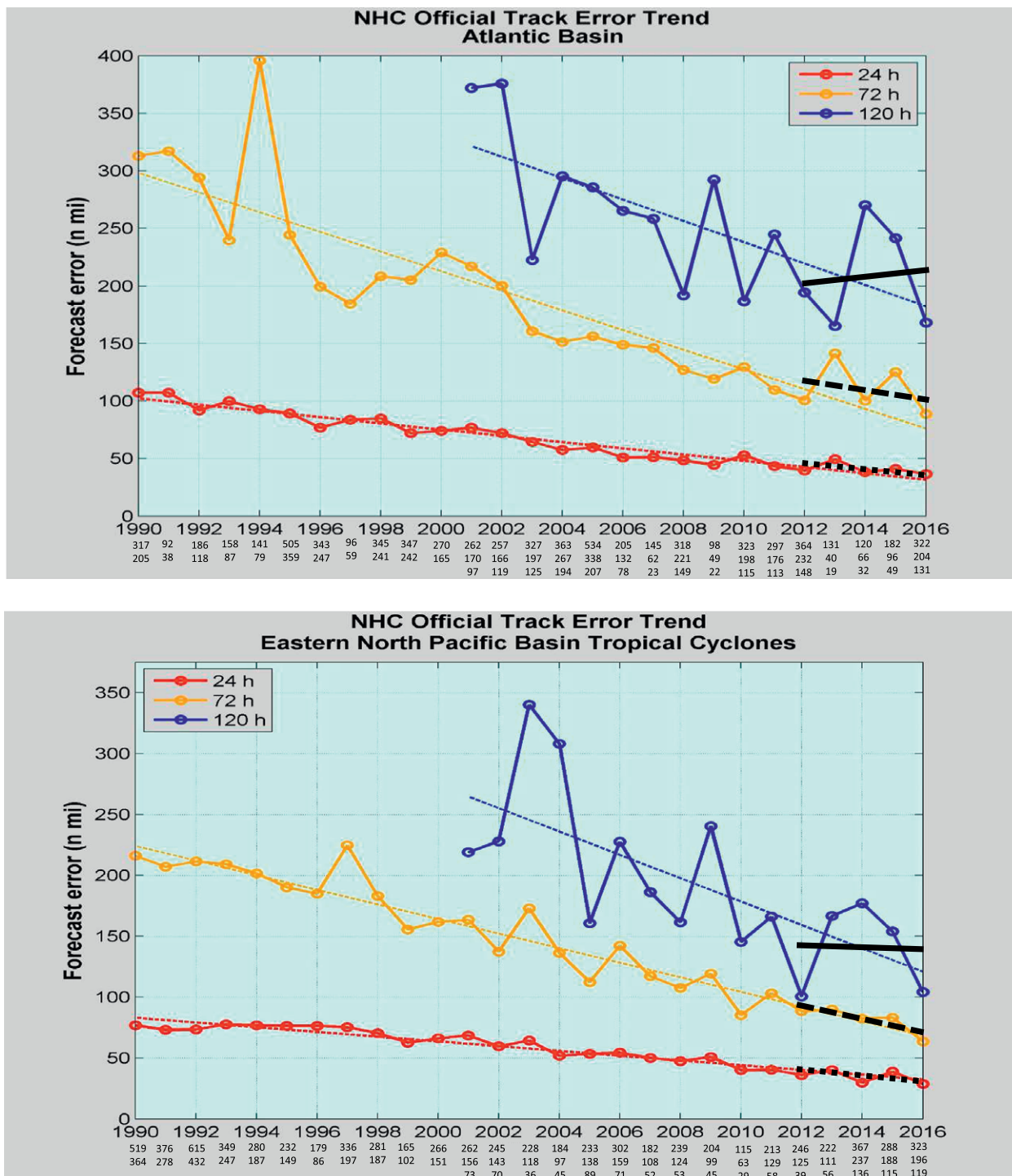
lead to perfect predictions in just another decade or two. However, such improvements, of course, cannot continue until perfection is reached. As first demonstrated by Lorenz (1969), there exists a limit of predictability beyond which further increases of skill are unattainable due to the cascading of errors up to the synoptic scale.<sup>2</sup> There were a number of studies examining the limits of predictability for tropical cyclone tracks conducted in the late 1980s to the early 2000s, but these either relied on climatology, statistical techniques, or simplified dynamical models. Moreover, the limits suggested in most of these earlier studies—for example, track errors of about 120 n mi for 72-h forecasts (Leslie et al. 1998)—have already been surpassed (Fig. 1). One paper (Aberson 1998) calculated an *e*-folding growth of errors for track forecasting in the Atlantic basin of about 2.5 days for a doubling of errors roughly every 45 h. The only new study in the last decade and a half on the topic—Plu (2011)—re-examined the issue using state-of-the-art, highly skillful global models from the European Centre for Medium-Range Weather Forecasts (ECMWF), Météo-France, and the Met Office (though these models have further evolved since the publication of this paper). He found that the doubling time of small errors in track prediction is on the order of 40 h. This result is generally longer than what most

studies had first suggested but is quite close to the finding by Aberson (1998).

Of interest, then, is the appearance of a leveling off of the track forecast improvements at NHC during the last few years. Figure 1 includes a linear best fit to the official NHC track errors for the 5-yr period of 2012–16 in addition to the long-term (1990–2016) trend lines. (A 5-yr period is chosen to minimize interannual variability of tropical cyclone track errors as well as to best represent recent performance in this metric.) While the errors at 24 and 72 h for both basins still are showing some improvements in the last 5 years, the 120-h track forecast error trend appears to be flat. The 2012–16 error trend was flatter than 19 of the previous 23 five-year periods for the Atlantic basin at 24 h, 16 of 23 for 72 h, and 9 of 12 for 120 h. In the eastern North Pacific, the 2012–16 error trend was flatter than 13 of the previous 23 five-year periods at 24 h, 9 of 23 for 72 h, and 9 of 12 for 120 h. It is of note that many of these flatter trends in error reductions were in the 1990s before the huge improvements from global models became available. Removal of the data points before 2000 results in the 2012–16 error trend being flatter than 11 of the previous 15 five-year periods for the Atlantic basin at 24 h and 12 of 15 for 72 h. Likewise, for the eastern North Pacific, data points from 2000 onward indicate the error trend was flatter in 2012–16 than 10 of the 15 previous 5-yr trends at 24 h and 5 of 15 at 72 h.

However, examination of raw track errors—especially for short-term fluctuations—suffers from variability in the tracks themselves. For example, years with more straight-running tropical cyclones tend to have smaller errors than years with primarily recurving or erratic tracks. One accepted approach to account for this variability is to normalize the track errors by comparing them with a simple climatology and persistence model—CLIPER5 (Aberson 1998). Figure 2 provides trends in the official track skill, relative to the errors from CLIPER5. These still show some continued improvement in the eastern North Pacific for 24 and 72 h during the last 5 years. However, in the Atlantic basin for 24, 72, and 120 h, as well as in the eastern North Pacific basin for 120 h, there appears to be no further improvements in track forecasting skill from 2012 to 2016. (The sample size varies from year to year for both the Atlantic and eastern North Pacific basin and is smallest at the 120-h point. This variability can result in added noise for the overall trends, as can be seen especially for the 120-h errors and skill time series shown in Figs. 1 and 2.) The 2012–16 skill trend was flatter than 22 of the

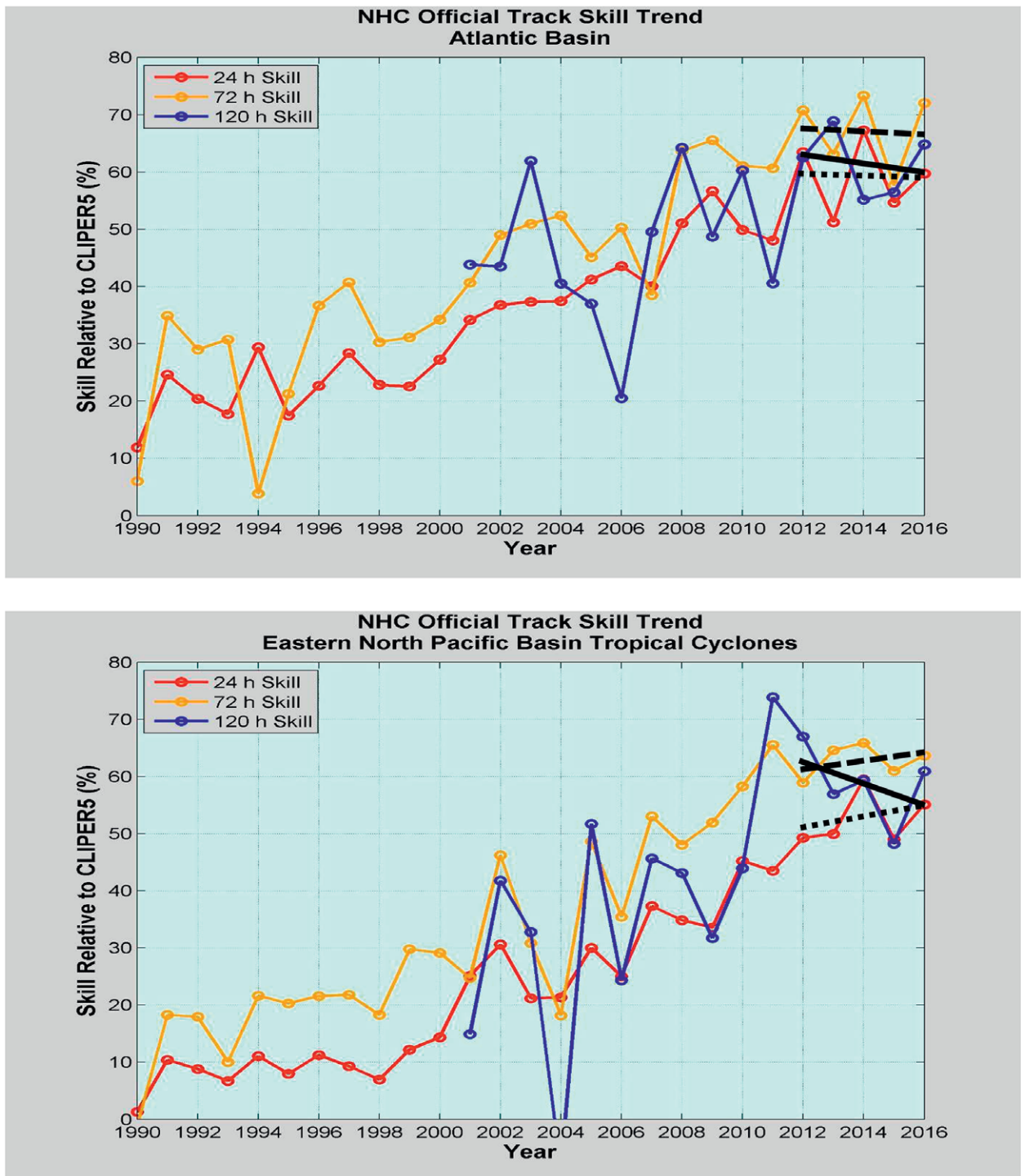
<sup>2</sup> One of the formal reviewers of this paper astutely pointed out that “the location of the center of a tropical cyclone... is neither directly measured by any instrumentation nor is it directly forecast by any sophisticated numerical model (the exceptions being the climatology and persistence model and the beta and advection models and their current incarnations). It is therefore a derived quantity, estimated from observations of other fields and forecast by processing of forecasts of other fields. It is clear from Lorenz that predicted quantities themselves have a certain predictability limit, but it is unclear how that specifically extends to predictability of derived parameters.” One of the other formal reviewers of this paper also insightfully noted that “more recent literature (e.g., Tribbia and Baumhefner 2004; Rotunno and Snyder 2008) suggest that the Lorenz (1969) concept of rapid upscale transfer of errors limiting the time scale of predictability is oversimplified. Fast upscale transfer is generally appropriate when the kinetic-energy spectrum follows a  $k^{-5/3}$  power law (mesoscales), but not when the spectrum is  $k^{-3}$  (synoptic scales). How does the fast upscale transfer at the mesoscales seeds the slower-growing larger scales operate specifically in the tropics? That I am not sure, but what I feel confident in saying is that our field has progressed to a much more mature understanding of predictability.”



**FIG. 1.** Track forecast errors (in n mi) from the NHC for (top) the Atlantic basin and (bottom) the eastern North Pacific basin for the 24- (red), 72- (yellow), and 120-h (blue) predictions. The best fit linear trend is indicated for each time frame in the colored dotted lines. The 2012–16 best fit linear trend is appended for each time series in black. Samples sizes for the 24-, 72-, and 120-h forecasts are indicated (from top to bottom) along the x axis.

previous 23 five-year periods for the Atlantic basin at 24 h, 17 of 23 for 72 h, and 7 of 12 for 120 h. In the eastern North Pacific, the 2012–16 skill trend was flatter than 14 of the previous 23 five-year periods at

24 h, 19 of 23 for 72 h, and 10 of 12 for 120 h. Removal of the data points before 2000 results in the 2012–16 skill trend being flatter than all of the previous 15 five-year periods for the Atlantic basin at 24 h and 12

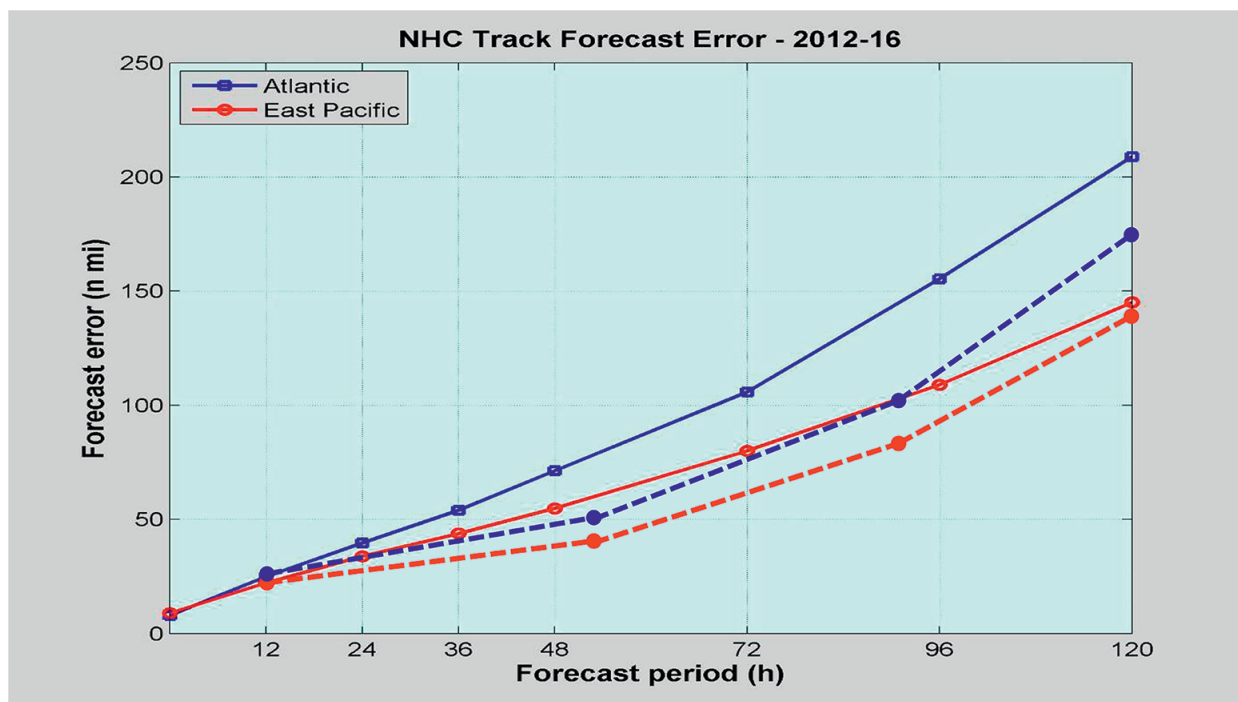


**FIG. 2.** Track forecast skill (improvements relative to CLIPER5) from the NHC for (top) the Atlantic basin and (bottom) the eastern North Pacific basin for the 24- (red), 72- (yellow), and 120-h (blue) predictions. The 2012–16 best fit linear trend is appended for each time series in black: 24 h, dotted; 72 h, dashed; and 120 h, solid.

of 15 for 72 h. Likewise, for the eastern North Pacific, data points from 2000 onward indicate the skill trend was flatter in 2012–16 than 12 of the 15 previous 5-yr trends at 24 h and 12 of 15 at 72 h.

While this sample of 5 years is too short for meaningful statistical significance testing, these results do suggest that there has been either a slower rate of improvement or perhaps no additional advances in





**FIG. 3.** NHC official track forecast errors (in n mi) for the Atlantic basin (blue) and eastern North Pacific basin (red) from 12 through 120 h for the 2012–16 hurricane seasons. The nonzero value at the initial time represents the initial position error. The dashed black line represents the 40-h doubling time errors suggested by Plu (2011) applied initially to the 12-h forecast error value for the Atlantic and eastern North Pacific basins separately.

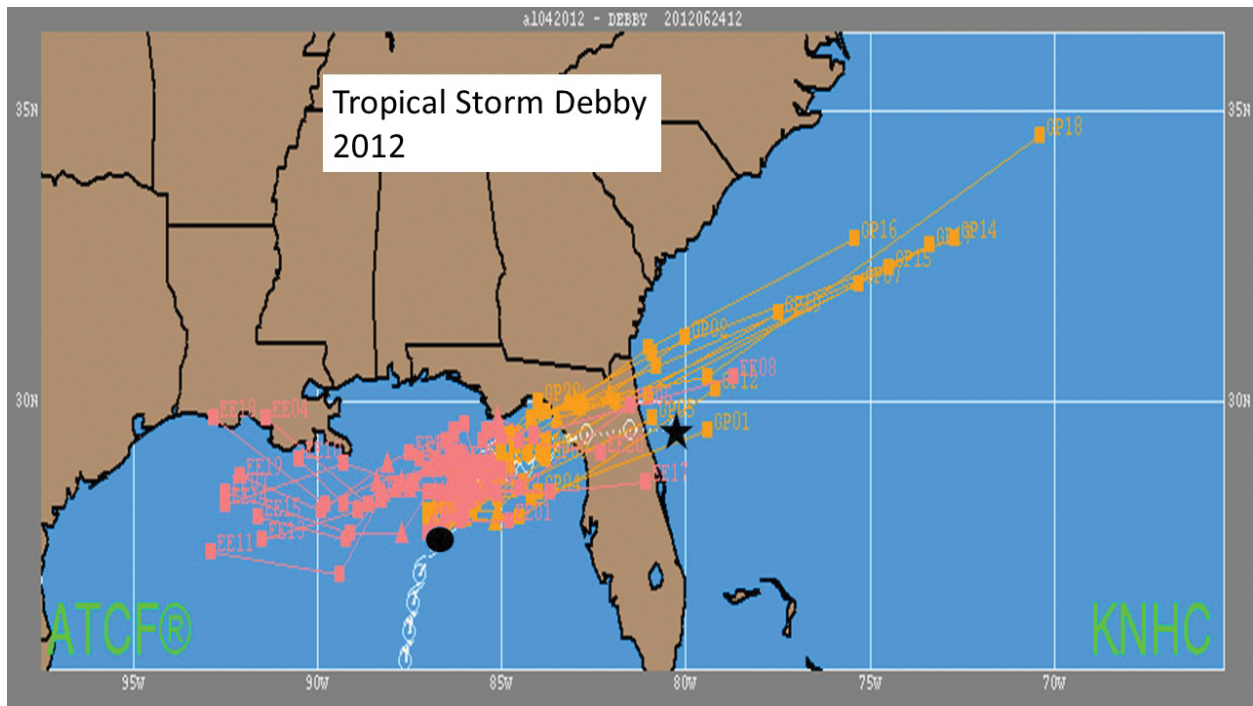
2012–16 for track forecasting.<sup>3</sup> (Similar testing of error and skill trends using 4- and 6-yr periods yielded nearly the same results.)

How, then, does this possible slowdown in track predictability improvement compare against the expectations of the limits of predictability shown in Plu (2011)? Figure 3 applies the 40-h doubling time errors as suggested by Plu (2011), beginning at the 12-h forecast time for the Atlantic and eastern North Pacific separately.<sup>4</sup> Applying Plu's 40-h error doubling provides projected limits of predictability that are slightly lower than the current errors at 24, 72, and 120 h, suggesting

that some improvement is still possible. However, simply extrapolating the linear long-term trends shown in Fig. 1 suggests that these limits may be reached within the next *five* hurricane seasons, if those long-term trends were to continue. Of course, if the limits of predictability are actually being approached, one would expect that the trend of improvement would gradually—not abruptly—slow. Empirically, this might be what has been observed during 2012–16. Moreover, if the estimates of doubling time for tropical cyclone track errors are overstated, then it is possible that the limits of predictability have already been reached for the eastern

<sup>3</sup> Testing of error and skill trends were also done for the track variable consensus model (TVCN) available back to 2004. The ECMWF—the single best deterministic track model guidance in recent years—has been providing explicit tropical cyclone track forecasts to NHC only back to 2007. The results from both TVCN and ECMWF, while not extending as far back in time, show similar variations to the official NHC predictions.

<sup>4</sup> The errors at the 12-h forecast point would be due to a combination of the initial position error along with the first usable prediction errors. The initial position errors were relatively large in the 1990s (14 n mi in the Atlantic and 15 n mi in the eastern North Pacific), decreased around 2000 because of improved access to microwave satellite imagery (Landsea and Franklin 2013), and have remained steady during the 2000s (8 n mi in the Atlantic and 10 n mi in the eastern North Pacific) and 2010s (9 n mi in both the Atlantic and eastern North Pacific). There is no indication of continuing decreases in the initial position errors in the last 15 years. Thus, it is unlikely that the 12-h forecast point errors will be able to be reduced much in the future unless improvements become available in our ability to observe the initial location of tropical cyclones. Some small improvements might be possible based upon new observation platforms such as the new Geostationary Operational Environmental Satellite R (GOES-R) satellite series and the high-resolution microwave imagery from the new Joint Polar Satellite System, but there is no expectation of dramatic advancements.



**FIG. 4.** 72-h track forecasts for Tropical Storm Debby issued at 1200 UTC 24 Jun 2012 from the 20 GEFS ensemble members (orange) and from 20 of the ECMWF ensemble members (pink). The white symbols represent the best track positions for Debbie. The initial point at 1200 UTC 24 Jun is shown as a black dot, while the verification point at 1200 UTC 27 Jun is shown as a black star.

North Pacific (if the doubling time is about 35 h) and for the Atlantic (if the doubling time is about 30 h).

One could argue that substantial improvements are still possible if there are outliers with large track errors on occasion that can be reduced. Figure 4 shows an example from an infamous case—Tropical Storm Debby in 2012. Shown here are the 20 ensemble members from the Global Ensemble Forecast System (GEFS) along with a representative 20 of the 51 ensemble members from the ECMWF model for a 72-h forecast from a 1200 UTC 24 June initial time. In this case, there was a bifurcation in possible forecast tracks that was apparent over several forecast cycles, with nearly all Global Forecast System (GFS) ensemble members showing an east-northeast trajectory and most ECMWF members showing little movement or a trajectory toward the west. These types of forecast scenarios can result in very large forecast errors for both the individual models as well as the NHC official forecast. Kimberlain (2013) wrote that “official forecast track errors were larger [for Debbie] than the mean official errors for the previous 5-yr period at all forecast times through 72 h, and considerably so after 24 h. Official track errors were about double the 5-yr mean at 36–48 h and triple that at 72 h.” Other examples of model bifurcation causing large track forecast errors in

recent years include Hurricane Sandy in 2012 and Hurricane Joaquin in 2015.

Tracks with such substantial forecast errors caused by model bifurcation, however, may be inherently unpredictable as relatively small differences in the tropical cyclone vortex or in the environment may cause large changes in the projected path. Thus, it is likely that one will continue to see, on occasion, very large forecast track errors that cannot be improved by better observations, better data assimilation, better numerics, and better model postprocessing techniques.<sup>5</sup>

Despite incredible improvements in tropical cyclone track forecast errors and skill, it is well accepted that making perfect forecasts will never happen. Evidence has been presented hinting that the approaching limit of predictability for tropical cyclone track prediction is near or has already been reached. If indeed track prediction will cease to improve, this conclusion is of critical importance for planning by all users of tropical cyclone track forecast

<sup>5</sup> Examination of the outliers represented by the 90th percentile largest errors reveals that errors for the outliers have decreased substantially during the last couple of decades for both the Atlantic and eastern North Pacific basins for all time periods. However, these—like the average errors—show signs of plateauing during the last few years.

information, including the emergency management community/governments, the media, the private sector, and the general public. Given the lack of achieving perfection in track predictions (and other attributes of tropical cyclones), emphasis should be changed therefore to provide well-calibrated probabilistic forecasting information for tropical cyclone impacts. One can use the spread in both the multi-model (deterministic) and single-model ensembles to provide an appropriate confidence in conjunction with an official track forecast to inform users about the uncertainty. This approach has been pursued at NHC for wind speed (DeMaria et al. 2013) and storm surge probabilities (Morrow et al. 2015) and will continue to be refined for these and other tropical cyclone impacts in the future.

## REFERENCES

- Aberson, S. D., 1998: Five-day tropical cyclone track forecasts in the North Atlantic basin. *Wea. Forecasting*, **13**, 1005–1015, [https://doi.org/10.1175/1520-0434\(1998\)013<1005:FDTCTF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<1005:FDTCTF>2.0.CO;2).
- Cangialosi, J. P., and J. L. Franklin, 2017: National Hurricane Center forecast verification report: 2016 hurricane season. National Hurricane Center Rep., 72 pp., [www.nhc.noaa.gov/verification/pdfs/Verification\\_2016.pdf](http://www.nhc.noaa.gov/verification/pdfs/Verification_2016.pdf).
- DeMaria, M., 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>.
- Kimberlain, T. B., 2013: Tropical cyclone report: Tropical Storm Debby (AL042012). National Hurricane Center Rep., 51 pp., [www.nhc.noaa.gov/data/tcr/AL042012\\_Debby.pdf](http://www.nhc.noaa.gov/data/tcr/AL042012_Debby.pdf).
- 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>.
- Leslie, L. M., R. F. Abbey, and G. J. Holland, 1998: Tropical cyclone track predictability. *Meteor. Atmos. Phys.*, **65**, 223–231, <https://doi.org/10.1007/BF01030790>.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, [https://doi.org/10.1175/1520-0469\(1969\)26<636:APARBN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2).
- Morrow, B. H., J. K. Lazo, J. Rhome, and J. Feyen, 2015: Improving storm surge risk communication: Stakeholder perspectives. *Bull. Amer. Meteor. Soc.*, **96**, 35–48, <https://doi.org/10.1175/BAMS-D-13-00197.1>.
- Plu, M., 2011: A new assessment of the predictability of tropical cyclone tracks. *Mon. Wea. Rev.*, **139**, 3600–3608, <https://doi.org/10.1175/2011MWR3627.1>.
- Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, <https://doi.org/10.1175/2008WAF2222128.1>.
- Rotunno, R., and C. Snyder, 2008: A generalization of Lorenz's model for the predictability of flows with many scales of motion. *J. Atmos. Sci.*, **65**, 1063–1076, <https://doi.org/10.1175/2007JAS2449.1>.
- Tribbia, J. J., and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Mon. Wea. Rev.*, **132**, 703–713, [https://doi.org/10.1175/1520-0493\(2004\)132<0703:SIAAP A>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0703:SIAAP A>2.0.CO;2).

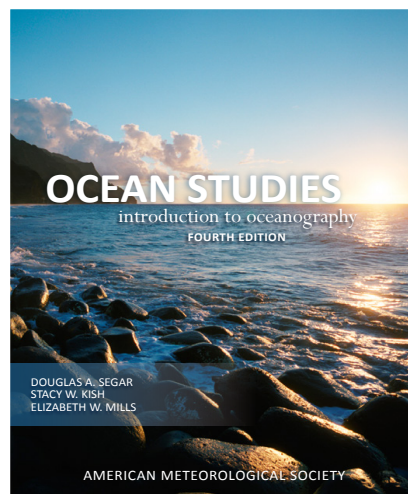


# NEW SIGNIFICANTLY REVISED TEXTBOOK



*AMS Ocean Studies* is an introductory, undergraduate-level curriculum package developed, licensed, and nationally implemented by the American Meteorological Society. *AMS Ocean Studies* consists of the new *Ocean Studies: Introduction to Oceanography* textbook, investigations manual, RealTime Ocean Portal, and course management system-compatible files. The 4<sup>th</sup> edition *Ocean Studies* textbook underwent a major revision with significant updates to content and imagery. It is authored by Douglas A. Segar, Stacy W. Kish, and Elizabeth W. Mills, and includes 15 chapters exploring the world ocean from an Earth system perspective.

Instructors can use these materials in combinations that cater to face-to-face or online lecture or laboratory courses. Explore the Introductory Chapter of *Ocean Studies* at [ametsoc.org/OceanTextbook](http://ametsoc.org/OceanTextbook)



A robust,  
introductory-level  
oceanography  
course

Flexible course  
for undergraduate  
classrooms

Your adoption of these materials contributes to our ability to provide professional development opportunities in weather, ocean, and climate science to K-12 teachers.

Request a free exam copy at [ametsoc.org/OceanStudies](http://ametsoc.org/OceanStudies)