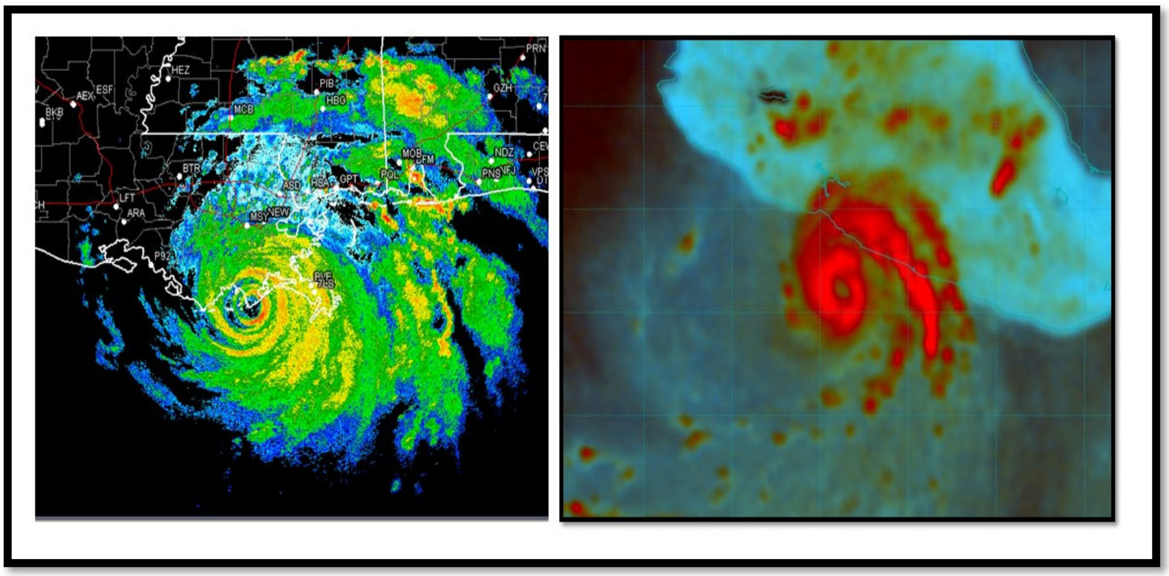
NATIONAL HURRICANE CENTER FORECAST VERIFICATION REPORT

2021 HURRICANE SEASON

John P. Cangialosi
National Hurricane Center

25 April 2022



NATIONAL WEATHER SERVICE DOPPLER RADAR IMAGE OF HURRICANE IDA MAKING LANDFALL IN LOUISIANA (LEFT). 89-GHZ SSMIS IMAGE OF HURRICANE RICK SEVERAL HOURS BEFORE LANDFALL IN MEXICO (RIGHT). IMAGE COURTESY OF THE NAVAL RESEARCH LABORATORY.

ABSTRACT

There were 394 official forecasts issued during the 2021 Atlantic hurricane season, which is above the long-term average number of forecasts and a similar level of activity as the 2016-2018 seasons. The mean NHC official track forecast errors in the Atlantic basin were close to or below their previous 5-yr means. Records for track accuracy were set from 48-72 h in 2021. Track forecast errors have decreased significantly over the long term, but there has been less improvement during the past several years. The official track forecasts were slightly outperformed by the consensus models at some time periods. GFSI was the best performing track model, with EMXI and AEMI being the next best. HWFI, HMNI, CMCI, CTCI, and NVGI performed less well. The Government Performance and Results Act of 1993 (GPRA) track goal was met.

Mean official intensity errors for the Atlantic basin in 2021 were lower than the previous 5-yr means at all forecast times. Decay-SHIFOR errors in 2021 followed a similar pattern to the official forecasts. The official forecasts were quite skillful and beat all of the models at 12-36 h and 72 h. Records for intensity accuracy were set from 12-60 h in 2021. Although there is a considerable amount of year-to-year variability, the intensity forecast errors have been gradually decreasing over the past decade or so. Among the guidance, IVCN and HCCA were the best performers. HMNI and HWFI were also quite skillful and were the best individual models. LGEM and DSHP were fair performers and GFSI was competitive with those models for some of the forecast times. EMXI was generally less skillful. The GPRA intensity goal was also met.

There were 314 official forecasts issued in the eastern North Pacific basin in 2021, although only 43 of these verified at 120 h. This level of forecast activity was a little below average. The mean NHC official track forecast errors in the east Pacific basin were a little lower than the previous 5-yr means at most forecast times. No records for track accuracy were set in 2021. The official track forecasts were very skillful, but they were outperformed by HCCA, FSSE and TVCE at the short lead times. EMXI, GFSI, and AEMI were the best individual track models in this basin. HWFI, HMNI, CMCI, and EGRI were fair performers, while NVGI lagged behind.

For intensity, the official forecast errors in the eastern North Pacific basin were slightly lower than the 5-yr means for the short lead times, but higher than the means for the longer lead times. Decay-SHIFOR errors followed a similar trend, indicating that predicting the intensity of the season's storms at the longer lead times was more challenging than average. No records for intensity accuracy were set in 2021. The official forecasts were close to the consensus models and were skillful at all times. HMNI was the best individual model for the short lead times while DSHP was best at the longer forecast times.

An evaluation of track performance during the 2019-21 period in the Atlantic basin indicates that HCCA and TVCA were the best models. EMXI was the best individual model in the short term, and GFSI and AEMI had the most skill at the long lead times. The official track forecasts for the 3-yr sample had skill that was quite close to the best aids throughout the forecast period. For intensity in the Atlantic basin, the official forecasts performed quite well and had skill that was comparable to the best guidance, the consensus models. HWFI and DSHP were the best individual models.

A three-year evaluation from 2019-21 in the eastern North Pacific indicates that the official track forecasts were very skillful, and had skill levels close to the consensus models. Regarding intensity, the official forecasts during the 3-yr sample performed as good as or better than the consensus aids. HMNI was the best individual model for the short lead times, and DSHP was best at 96 and 120 h.

Quantitative probabilistic forecasts of tropical cyclogenesis are expressed in 48 and 120 h time frames in 10% increments and in terms of categories ("low", "medium", or "high"). In the Atlantic basin, results from 2021 indicate that the probabilistic forecasts were generally well calibrated at most probabilities for both the 48- and 120-h forecasts. In the eastern North Pacific basin, a slight low bias existed at most ranges for the 120-h probabilistic forecasts.

TABLE OF CONTENTS

1. Introduction

For all operationally designated tropical or subtropical cyclones, or systems that could become tropical or subtropical cyclones and affect land within the next 48 h in the Atlantic and eastern North Pacific basins, the National Hurricane Center (NHC) issues an official forecast of the cyclone's center location and maximum 1-min surface wind speed. Forecasts are issued every 6 h, and contain projections valid 12, 24, 36, 48, 60, 72, 96, and 120 h after the forecast's nominal initial time (0000, 0600, 1200, or 1800 UTC)[1]. At the conclusion of the season, forecasts are evaluated by comparing the projected positions and intensities to the corresponding post-storm derived "best track" positions and intensities for each cyclone. A forecast is included in the verification only if the system is classified in the final best track as a tropical (or subtropical[2]) cyclone at both the forecast's initial time and at the projection's valid time. All other stages of development (e.g., tropical wave, [remnant] low, extratropical) are excluded[3]. For verification purposes, forecasts associated with special advisories do not supersede the original forecast issued for that synoptic time; rather, the original forecast is retained[4]. All verifications in this report include the depression stage.

It is important to distinguish between *forecast error* and *forecast skill*. Track forecast error, for example, is defined as the great-circle distance between a cyclone's forecast position and the best track position at the forecast verification time. Skill, on the other hand, represents a normalization of this forecast error against some standard or baseline. Expressed as a percentage improvement over the baseline, the skill of a forecast s_f is given by

$$s_f\,(\%) = 100 * (e_b - e_f) / e_b$$

where e_b is the error of the baseline model and e_f is the error of the forecast being evaluated. It is seen that skill is positive when the forecast error is smaller than the error from the baseline.

To assess the degree of skill in a set of track forecasts, the track forecast error can be compared with the error from CLIPER5, a climatology and persistence model that contains no information about the current state of the atmosphere (Neumann 1972, Aberson 1998)[5]. Errors from the CLIPER5 model are taken to represent a "no-skill" level of accuracy that is used as the baseline (e_b) for evaluating other forecasts[6]. If CLIPER5 errors are unusually low during a given season, for example, it indicates that the year's storms were inherently "easier" to forecast than normal or otherwise unusually well behaved. The current version of CLIPER5 is based on developmental data from 1931-2004 for the Atlantic and from 1949-2004 for the eastern Pacific.

---

1 The nominal initial time represents the beginning of the forecast process. The actual advisory package is not released until 3 h after the nominal initial time, i.e., at 0300, 0900, 1500, and 2100 UTC.
2 For the remainder of this report, the term "tropical cyclone" shall be understood to also include subtropical cyclones.
3 Possible classifications in the best track are: Tropical Depression, Tropical Storm, Hurricane, Subtropical Depression, Subtropical Storm, Extratropical, Disturbance, Wave, and Low.
4 Special advisories are issued whenever an unexpected significant change has occurred or when watches or warnings are to be issued between regularly scheduled advisories. The treatment of special advisories in forecast databases changed in 2005 to the current practice of retaining and verifying the original advisory forecast.
5 CLIPER5 and SHIFOR5 are 5-day versions of the original 3-day CLIPER and SHIFOR models.
6 To be sure, some "skill", or expertise, is required to properly initialize the CLIPER model.

Particularly useful skill standards are those that do not require operational products or inputs, and can therefore be easily applied retrospectively to historical data. CLIPER5 satisfies this condition, since it can be run using persistence predictors (e.g., the storm's current motion) that are based on either operational or best track inputs. The best-track version of CLIPER5, which yields substantially lower errors than its operational counterpart, is generally used to analyze lengthy historical records for which operational inputs are unavailable. It is more instructive (and fairer) to evaluate operational forecasts against operational skill benchmarks, and therefore the operational versions are used for the verifications discussed below.[7]

Forecast intensity error is defined as the absolute value of the difference between the forecast and best track intensity at the forecast verifying time. Skill in a set of intensity forecasts is assessed using Decay-SHIFOR5 (DSHIFOR5) as the baseline. The DSHIFOR5 forecast is obtained by initially running SHIFOR5, the climatology and persistence model for intensity that is analogous to the CLIPER5 model for track (Jarvinen and Neumann 1979, Knaff et al. 2003). The output from SHIFOR5 is then adjusted for land interaction by applying the decay rate of DeMaria et al. (2006). The application of the decay component requires a forecast track, which here is given by CLIPER5. The use of DSHIFOR5 as the intensity skill benchmark was introduced in 2006. On average, DSHIFOR5 errors are about 5-15% lower than SHIFOR5 in the Atlantic basin from 12-72 h, and about the same as SHIFOR5 at 96 and 120 h.

It has been argued that CLIPER5 and DSHIFOR5 should not be used for skill benchmarks, primarily on the grounds that they were not good measures of forecast difficulty. Particularly in the context of evaluating forecaster performance, it was recommended that a model consensus (see discussion below) be used as the baseline. However, an unpublished study by NHC has shown that on the seasonal time scales at least, CLIPER5 and DSHIFOR5 are indeed good predictors of official forecast error. For the period 1990-2009 CLIPER5 errors explained 67% of the variance in annual-average NHC official track forecast errors at 24 h. At 72 h the explained variance was 40% and at 120 h the explained variance was 23%. For intensity the relationship was even stronger: DSHIFOR5 explained between 50 and 69% of the variance in annual-average NHC official errors at all time periods. Given this, CLIPER5 and DSHIFOR5 appear to remain suitable, if imperfect, baselines for skill, in the context of examining forecast performance over the course of a season (or longer). However, they're probably less useful for interpreting forecast performance with smaller samples (e.g., for a single storm).

The trajectory-CLIPER (TCLP) model is an alternative to the CLIPER and SHIFOR models for providing baseline track and intensity forecasts (DeMaria, personal communication). The input to TCLP [Julian Day, initial latitude, longitude, maximum wind, and the time tendencies of position and intensity] is the same as for CLIPER/SHIFOR, but rather than using linear regression to predict the future latitude, longitude and maximum wind, a trajectory approach is used. For track, a monthly climatology of observed storm motion vectors was developed from a 1982-2011 sample. The TCLP storm track is determined from a trajectory of the climatological motion vectors starting at the initial date and position of the storm. The climatological motion vector is modified

---

[7] On very rare occasions, operational CLIPER or SHIFOR runs are missing from forecast databases. To ensure a completely homogeneous verification, post-season retrospective runs of the skill benchmarks are made using operational inputs. Furthermore, if a forecaster makes multiple estimates of the storm's initial motion, location, etc., over the course of a forecast cycle, then these retrospective skill benchmarks may differ slightly from the operational CLIPER/SHIFOR runs that appear in the forecast database.

by the current storm motion vector, where the influence of the current motion vector decreases with time during the forecast. A similar approach is taken for intensity, except that the intensity tendency is estimated from the logistic growth equation model (LGEM) with climatological input. Similar to track, the climatological intensity tendency is modified by the observed tendency, where the influence decreases with forecast time. The track used for the TCLP intensity forecast is the TCLP track forecast. When the storm track crosses land, the intensity is decreased at a climatological decay rate. A comparison of a 10-yr sample of TCLP errors with those from CLIPER5 and DSHIFOR5 shows that the average track and intensity errors of the two baselines are within 10% of each other at all forecast times out to five days for the Atlantic and eastern North Pacific. One advantage of TCLP over CLIPER5/DSHIFOR5 is that TCLP can be run to any desired forecast time.

NHC also issues forecasts of the size of tropical cyclones; these “wind radii” forecasts are estimates of the maximum extent of winds of various thresholds (34, 50, and 64 kt) expected in each of four quadrants surrounding the cyclone. Unfortunately, there is insufficient surface wind information to allow the forecaster to accurately analyze the size of a tropical cyclone’s wind field. As a result, post-storm best track wind radii are likely to have errors so large as to render a verification of official radii forecasts unreliable and potentially misleading; consequently, no verifications of NHC wind radii are included in this report. In time, as our ability to measure the surface wind field in tropical cyclones improves, it may be possible to perform a meaningful verification of NHC wind radii forecasts (Cangialosi and Landsea 2016).

Numerous objective forecast aids (guidance models) are available to help the NHC in the preparation of official track and intensity forecasts. 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 (12Z) forecast cycle, which begins with the 12Z synoptic time and ends with the release of an official forecast at 15Z. The 12Z run of the National Weather Service/Global Forecast System (GFS) model is not complete and available to the forecaster until about 16Z, or about an hour after the NHC forecast is released. Consequently, the 12Z GFS would be considered a late model since it could not be used to prepare the 12Z official forecast. This report focuses on the verification of early models.

Multi-layer dynamical models are generally, if not 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 (06Z) run of the GFS would be smoothed and then adjusted, or shifted, such that the 6-h forecast (valid at 12Z) would match the observed 12Z position and intensity of the tropical cyclone. The adjustment process creates an “early” version of the GFS model for the 12Z forecast cycle that is based on the most current available guidance. The adjusted versions of the late models are known, mostly for historical reasons, as *interpolated* models[8]. The adjustment algorithm is invoked as long as the most recent available late model is not more than 12 h old, e.g., a 00Z late model could be used to form an interpolated model for

[8] When the technique to create an early model from a late model was first developed, forecast output from the late models was available only at 12 h (or longer) intervals. In order to shift the late model’s forecasts forward by 6 hours, it was necessary to first interpolate between the 12 h forecast values of the late model – hence the designation “interpolated”.

the subsequent 06Z or 12Z forecast cycles, but not for the subsequent 18Z cycle. Verification procedures here make no distinction between 6 and 12 h interpolated models.[9]

A list of models is given in Table 1. In addition to their timeliness, models are characterized by their complexity or structure; this information is contained in the table for reference. Briefly, *dynamical* models forecast by solving the physical equations governing motions in the atmosphere. Dynamical models may treat the atmosphere either as a single layer (two-dimensional) or as having multiple layers (three-dimensional), and their domains may cover the entire globe or be limited to specific regions. The interpolated versions of dynamical model track and intensity forecasts are also sometimes referred to as dynamical models. *Statistical* models, in contrast, do not consider the characteristics of the current atmosphere explicitly but instead are based on historical relationships between storm behavior and various other parameters. *Statistical-dynamical* models are statistical in structure but use forecast parameters from dynamical models as predictors. *Consensus* models are not true forecast models *per se*, but are merely combinations of results from other models. One way to form a consensus is to simply average the results from a collection (or "ensemble") of models, but other, more complex techniques can also be used. The Hurricane Forecast Improvement Program Corrected Consensus Approach (HCCA), for example, combines its individual components on the basis of past performance and attempts to correct for biases in those components (Simon et al. 2018). A consensus model that considers past error characteristics can be described as a "weighted" or "corrected" consensus. Additional information about the guidance models used at the NHC can be found at http://www.nhc.noaa.gov/modelsummary.shtml.

The verifications described in this report are for all tropical cyclones in the Atlantic and eastern North Pacific basins. These statistics are based on forecast and best track data sets taken from the Automated Tropical Cyclone Forecast (ATCF) System[10] on 18 March 2022 for the Atlantic basin, and on 10 March 2022 for the eastern North Pacific basin. Verifications for the Atlantic and eastern North Pacific basins are given in Sections 2 and 3 below, respectively. Section 4 discusses NHC's probabilistic genesis forecasts. Section 5 summarizes the key findings of the 2021 verification and previews anticipated changes for 2022.

2. Atlantic Basin

a. 2021 season overview – Track

Figure 1 and Table 2 present the results of the NHC official track forecast verification for the 2021 season, along with results averaged for the previous 5-yr period, 2016-2020. In 2021, the NHC issued 394 Atlantic basin tropical cyclone forecasts[11], a number notably above the long term mean (325) and comparable to the 2016-2018 seasons (Fig. 2). Mean track errors ranged from 25 n mi at 12 h to 176 n mi at 120 h. The mean official track forecast errors in 2021 were

[9] The UKM and EMX models are only available through 120 h twice a day (at 0000 and 1200 UTC). Consequently, roughly half the interpolated forecasts from these models are 12 h old.
[10] In ATCF lingo, these are known as the "a decks" and "b decks", respectively.
[11] This count does not include forecasts issued for systems later classified to have been something other than a tropical cyclone at the forecast time.

similar to, or slightly higher than, the previous 5-yr means at 12, 24, and 120 h, but 8-22% lower than the means at the other forecast times. The CLIPER errors for 2021 were a little higher than their previous 5-yr means at 12 and 24 h, but significantly lower than their long-term means at the other forecast times, indicating that the track of the season's storms were easier than normal to predict at long range. Records for track accuracy were set at 48, 60, and 72 h in 2021. The official track forecast vector biases were northeastward and increased with forecast time (i.e., the official forecast tended to fall to the northeast of the verifying position). Track forecast skill ranged from 48% at 12 h to 73% at 60 h (Table 2). The track errors in 2021 decreased from the 2020 values at most forecast times, and over the past few decades the 24−72-h track forecast errors have been reduced by 70 to 75% (Fig. 3a). Track forecast error reductions of about 60% have occurred over the past 20 years for the 96- and 120-h forecast periods. An evaluation of track skill indicates that there has been a gradual increase in skill over the long term (Fig. 3b). Although the long-term trends are quite well established, the improvements in track error and skill have leveled off some during the past several years. Figure 4 indicates that on average the NHC track errors decrease as the initial intensity of a cyclone increases, and that relationship holds true through the 120-h forecast period.

Note that the mean official error in Figure 1 is not precisely zero at 0 h (the analysis time). This non-zero difference between the operational analysis of storm location and best track location, however, is not properly interpreted as "analysis error". The best track is a subjectively smoothed representation of the storm history over its lifetime, in which the short-term variations in position or intensity that cannot be resolved in a 6-hourly time series are deliberately removed. Thus the location of a strong hurricane with a well-defined eye might be known with great accuracy at 1200 UTC, but the best track may indicate a location elsewhere by 5-10 miles or more if the precise location of the cyclone at 1200 UTC was unrepresentative. Operational analyses tend to follow the observed position of the storm more closely than the best track analyses, since it is more difficult to determine unrepresentative behavior in real time. Consequently, the t=0 "errors" shown in Figure 1 contain both true analysis error and representativeness error.

Table 3a presents a homogeneous[12] verification for the official forecast along with a selection of early models for 2021. In order to maximize the sample size, a guidance model had to be available at least two-thirds of the time at both 48 and 120 h to be included in this comparison. The performance of the official forecast and the early track models in terms of skill are presented in Figure 5. The figure shows that the official forecasts were highly skillful, and near the best models throughout the forecast period. The best models were the consensus aids FSSE, HCCA, and TVCA, but HCCA was less good at 120 h. Among the individual models, GFSI was a very good performer and the best model at all forecast times. AEMI and EMXI were competitive with GFSI, but they had less skill at all times. CMCI and CTCI were quite skillful and competitive with the best models for the short lead times, but were notably less skillful at long range. Conversely, HWFI and HMNI were less good at the short lead times, but fair performers at 96 and 120 h. NVGI lagged behind all of the models. An evaluation over the three years 2019-21 (Fig. 6) indicates that HCCA and TVCA were the best models, and the official forecasts had about the same skill levels as those models throughout the forecast period. EMXI was the best individual model through 72 h, but GFSI and AEMI had more skill at 120 h. EGRI, HWFI, and

[12] Verifications comparing different forecast models are referred to as *homogeneous* if each model is verified over an identical set of forecast cycles. Only homogeneous model comparisons are presented in this report.

HMNI were fair performers and had skill levels about 10-15% lower than the best performing models and official forecasts.

Vector biases of the guidance models for 2021 are given in Table 3b. The table shows that the official forecast had similar biases to the consensus aids, which all had a general northeast to east bias at most forecast times. Conversely, EMXI had a large east-southeast to south bias from 72 to 120 h in 2021. Figure 7 shows a homogenous comparison of the 120-h biases of the official forecasts, GFSI, EXMI, and EGRI from 2019-21 in the Atlantic basin. It can be seen that mean biases (denoted by the red X) were generally westward in the models and NHC forecasts for this sample.

A separate homogeneous verification of the primary consensus models for 2021 is shown in Figure 8. The figure shows that TVDG was the most skillful model overall, but TVCA, TVCX, and GFEX had only slightly less skill. HCCA and FSSE had similar skill levels to the best aids early, but trailed some after 60 h. AEMI was notably less skillful through 72 h, but was comparable to HCCA and FSSE at 96 and 120 h.

Atlantic basin 48-h official track error, evaluated for all tropical cyclones, is a forecast measure tracked under the Government Performance and Results Act of 1993 (GPRA). In 2021, the GPRA goal was 57 n mi, and the verification for this measure was met at 53.3 n mi.

b. 2021 season overview – Intensity

Figure 9 and Table 4 present the results of the NHC official intensity forecast verification for the 2021 season, along with results averaged for the preceding 5-yr period. Mean forecast errors in 2021 ranged from 5 kt at 12 h to 12 kt at 120 h. These errors were 13-24% lower than the previous 5-yr means at all forecast times Records for accuracy were set for the 12-60 h forecast periods in 2021. The official forecasts had little bias from 12 to 96 h, but a low bias existed at 120 h. The Decay-SHIFOR5 errors exhibited a similar pattern to the official forecasts, with the errors being 11-25% lower than their 5-yr means. Figure 10 indicates that the NHC official intensity errors decreased at all forecast times from their 2020 values, and over the long-term, despite year-to-year variability, there has been a notable decrease in error that began around 2010. It appears that the intensity predictions are gradually improving as the forecasts are generally more skillful in the past 10 years or so than they were in the 1990s and the first decade of the 2000s (Cangialosi et. al 2020).

Table 5a presents a homogeneous verification for the official forecasts and the primary early intensity models for 2021. Intensity biases are given in Table 5b, and forecast skill is presented in Figure 11. The official forecasts were quite skillful, and they beat all of the models from 12-36 h and 72 h. The consensus models IVCN and HCCA were the best aids and outperformed the official forecasts at the other time periods. FSSE had similar skill to IVCN and HCCA for the short lead times, but it trailed off after 72 h. Among the individual models, HMNI was the best model at most time periods, and it had increasing skill with forecast time. HWFI was next best, though it had notably less skill than HMNI. DSHP, LGEM, and CTCI had skill at all forecast times, but were generally less good. GFSI was competitive with the standard intensity models through 96 h, but it had no skill at 120 h. Conversely, EMXI had no skill through 60 h, but was competitive with the more typical intensity models at long range. An inspection of the intensity biases (Table 5b) indicates that most of the models had small biases, except for GFSI and EMXI, which had large low biases. Despite its high skill, HMNI also had a low bias at all forecast time periods.

An evaluation over the three years 2019-21 (Fig. 12) indicates that the official forecasts have been consistently performing quite well, and had skill values close to the best aids IVCN and HCCA. For this sample, HWFI was the best individual model from 12 to 60 h, and at 96 and 120 h. DSHP was best at 72 h. HMNI and LGEM were fair performers, but they were generally not as skillful as HWFI and DSHP. GFSI had skill throughout the forecast period, but EMXI was only skillful beyond 24 h and was less competitive. It should be noted that even though the global models are not as good as the statistical-dynamical and regional hurricane models, they have been improving over the past several years and are gradually becoming more competitive.

The 48-h official intensity error, evaluated for all tropical cyclones, is another GPRA measure for the NHC. In 2021, the GPRA goal was 11 kt and the verification for this measure was met at 9.4 kt.

c. Verifications for individual storms

Forecast verifications for individual storms are given in Table 6. Of note are the large track errors for Hurricane Grace and Tropical Storm Wanda, especially at the longer forecast lead times. For Grace, NHC had a large right-of-track bias as the storm remained shallower than expected in the early portion of its life cycle, and therefore, was steered more westward by the low- to mid-level flow. In the case of Wanda, very large errors occurred at long range. A complex steering pattern over the northern Atlantic Ocean caused Wanda to meander for several days and take an unusual track, which made for very challenging forecasts. Conversely, the official track forecast errors were quite low for Hurricanes Elsa and Larry, and were well below NHC's 5-yr means for both of those tropical cyclones. Figure 13 shows an illustration of the official track errors stratified by storm.

With regards to intensity, Hurricane Henri, Hurricane Grace, and Tropical Storm Victor were among the more challenging cyclones to predict in 2021. For all of these tropical cyclones, the official intensity forecast errors were notably higher than the 5-yr averages at most forecast times. A high bias was present for Henri, as the storm did not intensify as much as expected before landfall in southern New England. Regarding Grace, the aforementioned poor track forecasts resulted in higher-than-normal intensity errors due to poorly predicted land interaction. For Victor, a high bias was also present as some of the forecasts anticipated that the cyclone would reach hurricane strength. Figure 14 shows an illustration of the official intensity errors stratified by storm. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/data/tcr/index.php?season=2021&basin=atl

3. Eastern North Pacific Basin

a. 2021 season overview – Track

The NHC track forecast verification for the 2021 season in the eastern North Pacific, along with results averaged for the previous 5-yr period is presented in Figure 15 and Table 7. There were 314 forecasts issued for the eastern Pacific basin in 2021, which is slightly below the long-term mean of about 330 forecasts (Fig. 16). Since most of the tropical cyclones in the basin were short-lived, only 43 of the forecasts verified at 120 h. Mean track errors ranged from 22 n mi at 12 h to 114 n mi at 120 h. These errors were a little lower than the 5-yr means at 24, 36, 48, 60, 96, and 120 h, but slightly higher than the means at 12 and 72 h. The CLIPER errors were 12-20% higher than their 5-yr means, indicating that the track of the season's storms were more challenging to predict than normal. No records for accuracy were set for track in this basin in 2021. The official track forecast vector biases were small through 48 h, but a more notable south to southwest bias existed from 60 to 120 h. A closer examination of the NHC track forecasts indicate that significantly higher errors occurred for tropical cyclones close to the coast of Mexico. Conversely, a separate verification for systems west of 115°W for 2021 indicate that the NHC track errors were much lower than the means at days 4 and 5, where the errors were around 75 and 90 n mi, respectively (not shown).

Figure 17 shows recent trends in track forecast accuracy and skill for the eastern North Pacific. Track errors decreased at all forecast times compared to 2020, and over the long term, track errors have been dramatically reduced by about 70% for the 24 to 72 h forecasts since 1990. It should be noted, however, that like in the Atlantic basin there has been a slower rate of change during the past five years or so. At the 96 and 120 h forecast times, errors have dropped by about 60% since 2001, but like the short lead times, the error trends have been flatter during the past few years. Forecast skill in 2021 was higher than the 2020 values at all forecast times, and are near all-time highs.

Table 8a presents a homogeneous verification for the official forecast and the early track models for 2021, with vector biases of the guidance models given in Table 8b. Skill comparisons of selected models are shown in Fig. 18. The official forecasts were very skillful and near the best models, the consensus aids. HCCA and FSSE were the best aids overall, and they beat the official forecasts from 12 to 72 h. However, the official forecasts beat all of the models at 120 h. TVCE was the next best model, also beating the official forecasts from 12 to 60 h. Among the individual models, GFSI, AEMI, and EMXI were competitive with one another, but EMXI was a stronger performer at 120 h. In fact, EMXI was as good as or better than the consensus aids at the long lead times. HWFI was the next best model, followed by CMCI and HMNI. NVGI and EGRI were poor performers and not competitive with the remainder of the models at most forecast times. An evaluation of the three years 2019-21 (Fig. 19) indicates that the official forecasts were very skillful, and they were near the performance of the consensus models. HCCA and FSSE slightly bested the official forecasts in the short term, but they had slightly less skill than the official forecast for the longer lead times. Regarding the individual models, EMXI was the best performer at all forecast times, with AEMI not too far behind. GFSI was next best, followed by HWFI, HMNI, CMCI, and EGRI. The official forecasts had similar biases to HCCA and FSSE at most forecast times. Among the individual models, EGRI had a large westward bias from 60 to 120 h while GFSI and AEMI had the lowest biases among the models.

A separate verification of the primary consensus aids is given in Figure 20. The skill of the consensus models was tightly clustered, but HCCA and FSSE were generally best for the short lead times and GFEX was best at 96 and 120 h. AEMI was less skillful (about 5-10% lower skill) than the highest performers.

b. 2021 season overview – Intensity

Figure 21 and Table 9 present the results of the NHC eastern North Pacific intensity forecast verification for the 2021 season, along with results averaged for the preceding 5-yr period. Mean forecast errors were 5 kt at 12 h and increased to 21 kt at 96 h. The errors were slightly lower than the previous 5-yr means at 12 and 24 h, but higher than the means from 36 to 120 h. In fact, the 2021 errors were 30% higher than the mean at 96 h. The Decay-SHIFOR forecast errors followed a similar trend. Their errors were lower than the 5-yr means at 12 and 24 h, but higher than their means after that time and were 34% larger than average at 96 h. No records for accuracy were set in 2021. A review of error and skill trends (Fig. 22) indicates that although there is considerable year-to-year variability in intensity errors, there has been a decrease in error over the past couple of decades at most forecast times. Forecast skill has changed little during the last several years, however. Intensity forecast biases were slightly high from 12 to 72 h, and slightly low at 120 h.

Figure 23 and Table 10a present a homogeneous verification for the primary early intensity models for 2021. Forecast biases are given in Table 10b. The official forecasts were skillful throughout the period with skill levels increasing from 12 to 60 h, and then gradually declining after that time. The consensus models HCCA, IVCN, FSSE, and NNIC outperformed the official forecasts for the middle forecast times, generally from 36 to 72 h, but their skill levels also declined at the longer lead times. Among the individual models, HMNI was the best performer for the short lead times while DSHP was best at the longer forecast periods. HWFI and LGEM were slightly less skillful. Although GFSI and EMXI were not skillful at 12 h, they had increasing levels of skill through the forecast period, and were competitive with the more standard intensity guidance at the longer lead times. Most of the models had a low bias, especially HMNI, GFSI, and EMXI at the longer forecast times. DSHP had the least bias overall. An evaluation over the three years 2019-21 (Fig. 24) indicates a different result than the 2021 sample. The official forecasts were skillful through 72 h, but had no skill at 96 and 120 h. A similar pattern was seen among most of the models. Similar to the 2021 results, HMNI was the best individual model for the short lead times and DSHP was best at 96 and 120 h.

c. Verifications for individual storms

Forecast verifications for individual storms are given for reference in Table 11. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/data/tcr/index.php?season=2021&basin=epac.

4. Genesis Forecasts

The NHC routinely issues Tropical Weather Outlooks (TWOs) for both the Atlantic and eastern North Pacific basins. The TWOs are text products that discuss areas of disturbed weather and their potential for tropical cyclone development. Beginning in 2007, forecasters subjectively assigned a probability of genesis (0 to 100%, in 10% increments) to each area of disturbed weather described in the TWO, where the assigned probabilities represented the forecaster's determination of the chance of tropical cyclone formation during the 48-h period following the nominal TWO issuance time. In 2009, the NHC began producing in-house (non-public) experimental probabilistic tropical cyclone forecasts through 120 h, which became public in August of 2013. Verification is based on NHC best-track data, with the time of genesis defined to be the first tropical cyclone point appearing in the best track.

Verifications of the 48-h outlook for the Atlantic and eastern North Pacific basins for 2021 are given in Table 12 and illustrated in Figure 25. In the Atlantic basin, a total of 911 genesis forecasts were made. These 48-h forecasts were well calibrated, except for a slight high bias for the 20-50% probabilities. In the eastern Pacific, a total of 727 genesis forecasts were made. The forecasts in this basin were also well calibrated, with only a slight low bias apparent for the 60-80% probabilities.

Verification of the 120-h outlook for the Atlantic and eastern North Pacific basins for 2021 are given in Table 13 and illustrated in Figure 26. In the Atlantic basin, the 120-h forecasts had a slight low bias at the high probabilities, but were otherwise quite reliable. In the eastern North Pacific, the genesis forecasts had a slight low bias at most probabilities. It should be noted that these results have been steadily improving over the past several years. The diagrams also show the refinement distribution, which indicates the sample size at each probabilistic bin.

Over the past few years, NHC has been experimenting with extending TWO forecasts out to day 7 (168 h). Figure 27 shows the verification of the 2-, 5-, and experimental 7-day genesis forecasts in the Atlantic basin for the 2019-21 period. It can be seen that a slight low bias exists at most ranges, but in general the forecasts are fairly calibrated for all lead times. The results in the eastern North Pacific basin for the same period are similar (not shown).

5. Looking Ahead to 2022

a. Track Forecast Cone Sizes

The National Hurricane Center track forecast cone depicts the probable track of the center of a tropical cyclone, and is formed by enclosing the area swept out by a set of circles along the forecast track (at 12, 24, 36 h, etc.). The size of each circle is set so that two-thirds of historical official forecast errors over the most-recent 5-yr sample fall within the circle. The circle radii defining the cones in 2022 for the Atlantic and eastern North Pacific basins (based on error distributions for 2017-21) are given in Table 14. In the Atlantic basin, the cone circles will be up to 6% smaller in 2022 from 12 to 96 h, but slightly larger than last year at 120 h. In the eastern Pacific basin, the cone circles will be slightly larger, especially at the longer lead times where the cone will be about 5% bigger in 2022. It should be noted that 60-h cone circles are based on interpolation of the 48- and 72-h cone sizes since the sample size at that forecast time is smaller.

b. Consensus Models

A set of NHC consensus model identifiers remain fixed from year to year. However, the specific members of these consensus models will be determined at the beginning of each season and may vary from year to year.

Some consensus models require all of their member models to be available in order to compute the consensus (e.g.,GFEX, ICON), while others are less restrictive, requiring only two or more members to be present (e.g., TVCA, IVCN). The terms “fixed” and “variable” can be used to describe these two approaches, respectively. In a variable consensus model, it is often the case that the 120-h forecast is based on a different set of members than the 12-h forecast. While this approach greatly increases availability, it does pose consistency issues for the forecaster.

The consensus model composition for 2022 is given in Table 15. The consensus models are unchanged from their compositions in 2021.

Acknowledgments

The author gratefully acknowledges Michael Brennan and members of the Technology and Science Branch of NHC, managers of the NHC forecast databases.

6. References

Aberson, S. D., 1998: Five-day tropical cyclone track forecasts in the North Atlantic basin. Wea. Forecasting, 13, 1005-1015.

Cangialosi. J.P., and C.W. Landsea, 2016: An examination of model and official National Hurricane Center tropical cyclone size forecasts. Wea. And Forecasting, 31, 1293-1300.

Cangialosi, J.P., E. Blake, M. DeMaria, A. Penny, A. Latto, and E. Rappaport, 2020: Recent progress in tropical cyclone intensity forecasting at the National Hurricane Center. *Wea. Forecasting*, **35**, 1913-1922.

DeMaria, M., J. A. Knaff, and J. Kaplan, 2006: On the decay of tropical cyclone winds crossing narrow landmasses, J. Appl. Meteor., 45, 491-499.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity for the North Atlantic basin. NOAA Tech. Memo. NWS NHC-10, 22 pp.

Knaff, J.A., M. DeMaria, B. Sampson, and J.M. Gross, 2003: Statistical, five-day tropical cyclone intensity forecasts derived from climatology and persistence. Wea. Forecasting, 18, 80-92.

Neumann, C. B., 1972: An alternate to the HURRAN (hurricane analog) tropical cyclone forecast system. NOAA Tech. Memo. NWS SR-62, 24 pp.

Simon, A., A. B. Penny, M. DeMaria, J. L. Franklin, R. J. Pasch, E. N. Rappaport, and D. A. Zelinsky, 2018: A description of the real-time HFIP Corrected Consensus Approach (HCCA) for tropical cyclone track and intensity guidance. Wea. Forecasting, 33, 37–57, https://doi.org/10.1175/WAF-D-17-0068.1.

List of Tables

Table 1. National Hurricane Center forecasts and models.

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
OFCL	Official NHC forecast			Trk, Int
HWRF	Hurricane Weather and Research Forecasting Model	Multi-layer regional dynamical	L	Trk, Int
HMON	Hurricanes in a Multi-scale Ocean-coupled Non-hydrostatic model	Multi-layer regional dynamical	L	Trk, Int
GFSO	NWS/Global Forecast System (formerly Aviation)	Multi-layer global dynamical	L	Trk, Int
AEMN	GFS ensemble mean	Consensus	L	Trk, Int
UKM	United Kingdom Met Office model, automated tracker	Multi-layer global dynamical	L	Trk, Int
EGRR	United Kingdom Met Office model with subjective quality control applied to the tracker	Multi-layer global dynamical	L	Trk, Int
UEMN	UKMET ensemble mean	Consensus	L	Trk, Int
NVGM	Navy Global Environmental Model	Multi-layer global dynamical	L	Trk, Int
CMC	Environment Canada global model	Multi-level global dynamical	L	Trk, Int
NAM	NWS/NAM	Multi-level regional dynamical	L	Trk, Int
CTX	COAMPS-TC using GFS initial and boundary conditions	Multi-layer regional dynamical	L	Trk, Int
EMX	ECMWF global model	Multi-layer global dynamical	L	Trk, Int
EEMN	ECMWF ensemble mean	Consensus	L	Trk

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
TABS	Beta and advection model (shallow layer)	Single-layer trajectory	E	Trk
TABM	Beta and advection model (medium layer)	Single-layer trajectory	E	Trk
TABD	Beta and advection model (deep layer)	Single-layer trajectory	E	Trk
CLP5	CLIPER5 (Climatology and Persistence model)	Statistical (baseline)	E	Trk
SHF5	SHIFOR5 (Climatology and Persistence model)	Statistical (baseline)	E	Int
DSF5	DSHIFOR5 (Climatology and Persistence model)	Statistical (baseline)	E	Int
OCD5	CLP5 (track) and DSF5 (intensity) models merged	Statistical (baseline)	E	Trk, Int
TCLP	Trajectory-CLIPER model	Statistical (baseline)	E	Trk, Int
SHIP	Statistical Hurricane Intensity Prediction Scheme (SHIPS)	Statistical-dynamical	E	Int
DSHP	SHIPS with inland decay	Statistical-dynamical	E	Int
OFCI	Previous cycle OFCL, adjusted	Interpolated	E	Trk, Int
HWFI	Previous cycle HWRF, adjusted	Interpolated-dynamical	E	Trk, Int
HMNI	Previous cycle HMON, adjusted	Interpolated-dynamical	E	Trk, Int
CTCI	Previous cycle CTCX, adjusted	Interpolated-dynamical	E	Trk, Int
GFSI	Previous cycle GFS, adjusted	Interpolated-dynamical	E	Trk, Int
UKMI	Previous cycle UKM, adjusted	Interpolated-dynamical	E	Trk, Int

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
EGRI	Previous cycle EGRR, adjusted	Interpolated-dynamical	E	Trk, Int
NVGI	Previous cycle NVGM, adjusted	Interpolated-dynamical	E	Trk, Int
EMXI	Previous cycle EMX, adjusted	Interpolated-dynamical	E	Trk, Int
CMCI	Previous cycle CMC, adjusted	Interpolated-dynamical	E	Trk, Int
AEMI	Previous cycle AEMN, adjusted	Consensus	E	Trk, Int
UEMI	Previous cycle UEMN, adjusted	Consensus	E	Trk, Int
FSSE	FSU Super-ensemble	Corrected consensus	E	Trk, Int
GFEX	Average of GFSI and EMXI	Consensus	E	Trk
TVCN	Average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus	E	Trk
TVCA	Average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus	E	Trk
TVCE	Average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus	E	Trk
TVCX	EMXI and average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus	E	Trk
TVCC	Version of TVCN corrected for model biases	Corrected consensus	E	Trk
TVDG	GFSI (double weight) EMXI (double weight) EGRI (double weight) CTCI HWFI	Corrected consensus	E	Trk
HCCA	Weighted average of AEMI, GFSI, CTCI, DSHP, EGRI, EMNI, EMXI,HWFI, HMNI LGEM	Corrected consensus	E	Trk, Int

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
ICON	Average of DSHP, LGEM, CTCI, HMNI and HWFI	Consensus	E	Int
IVDR	CTCI (double weight) HWFI (double weight) HMNI (double weight) GFSI DSHP LGEM	Consensus	E	Int
IVCN	Average of at least two of DSHP LGEM HWFI HMNI CTCI	Consensus	E	Int
NNIC	Average of at least two of HWFI GFSI DSHP LGEM	Corrected consensus	E	Int

Table 2. Homogenous comparison of official and CLIPER5 track forecast errors in the Atlantic basin in 2021 for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)							
	12	24	36	48	60	72	96	120
2021 mean OFCL error (n mi)	24.9	35.3	43.5	53.3	65.7	77.9	117.8	175.9
2021 mean CLIPER5 error (n mi)	48.2	99.1	145.4	190.7	240.0	280.1	352.9	394.9
2021 mean OFCL skill relative to CLIPER5 (%)	48.3	64.4	70.1	72.1	72.6	72.2	66.6	55.5
2021 mean OFCL bias vector (°/n mi)	020/004	033/006	046/010	046/015	049/024	045/034	052/047	056/082
2021 number of cases	351	309	272	244	216	189	151	115
2016-2020 mean OFCL error (n mi)	23.6	35.9	48.5	63.1	83.8	93.3	128.6	170.6
2016-2020 mean CLIPER5 error (n mi)	44.9	96.9	157.0	216.7	258.0	324.3	415.4	499.7
2016-2020 mean OFCL skill relative to CLIPER5 (%)	47.4	63.0	69.0	70.9	67.5	71.2	69.0	65.9
2016-2020 mean OFCL bias vector (°/n mi)	358/002	307/002	286/003	287/003	290/003	232/002	177/006	214/017
2016-2020 number of cases	1873	1681	1504	1338	481	1041	802	623
2021 OFCL error relative to 2016-2020 mean (%)	5.6	-1.7	-10.5	-15.5	-21.6	-16.5	-8.4	3.1
2021 CLIPER5 error relative to 2016-2020 mean (%)	7.3	2.3	-7.4	-12.0	-7.0	-13.6	-15.0	-21.0

Table 3a. Homogenous comparison of Atlantic basin early track guidance model errors (n mi) for 2021. Errors smaller than the NHC official forecast are shown in boldface.

Model ID	Forecast Period (h)							
	12	24	36	48	60	72	96	120
OFCL	21.3	29.2	35.9	42.4	50.4	61.9	97.0	119.4
OCD5	44.1	93.7	150.1	202.5	244.2	279.0	317.8	292.9
GFSI	**21.0**	32.0	41.8	51.4	60.0	71.5	99.3	123.2
HMNI	23.2	37.5	53.1	64.9	69.5	77.1	119.9	163.9
HWFI	24.7	37.7	50.2	59.5	67.1	80.9	119.3	167.6
EMXI	21.4	32.4	44.3	55.1	69.1	78.5	111.9	137.1
CMCI	23.3	35.5	47.2	59.4	77.6	97.8	148.1	181.2
NVGI	29.9	50.3	69.8	85.6	102.2	130.7	200.3	220.7
CTCI	21.8	34.4	46.9	59.6	77.6	102.3	168.1	193.0
AEMI	22.1	35.1	46.2	52.8	64.3	80.7	120.8	131.1
HCCA	**19.6**	**28.4**	**35.6**	42.6	51.0	63.7	104.4	134.0
TVCA	**19.9**	**28.0**	**35.6**	**42.0**	**49.5**	**60.8**	97.2	**106.2**
FSSE	**19.9**	**27.7**	**35.5**	44.3	54.0	68.1	99.0	**108.7**
Forecasts	195	176	162	143	125	112	83	55

Table 3b. Homogenous comparison of Atlantic basin early track guidance model bias vectors (º/n mi) for 2021.

	Forecast Period (h)							
Model ID	12	24	36	48	60	72	96	120
OFCL	007/004	031/006	033/008	031/011	045/015	053/018	088/024	339/007
OCD5	129/002	082/013	081/026	078/041	084/059	096/068	121/063	228/114
GFSI	018/006	035/010	039/015	041/018	051/020	059/026	076/041	059/038
HMNI	353/005	005/010	005/015	357/017	354/016	002/017	047/044	043/049
HWFI	028/006	036/010	031/013	028/013	059/008	116/009	146/034	180/041
EMXI	331/004	322/004	327/003	032/007	068/010	108/105	159/041	201/079
CMCI	331/005	355/010	010/017	020/027	026/038	033/049	048/079	034/106
NVGI	039/006	069/011	082/016	088/022	080/030	074/036	052/066	336/020
CTCI	028/007	038/012	041/016	040/019	044/024	053/029	083/033	201/054
AEMI	358/005	016/010	020/014	023/017	036/024	047/031	063/049	018/041
HCCA	006/002	038/004	028/006	012/008	011/009	029/011	084/016	296/022
TVCA	007/004	022/006	026/008	026/010	044/009	068/012	108/023	204/023
FSSE	349/003	003/005	025/006	026/011	034/016	045/020	068/032	007/020
Forecasts	195	176	162	143	125	112	83	55

Table 4. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the Atlantic basin for the 2021 season for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)							
	12	24	36	48	60	72	96	120
2021 mean OFCL error (kt)	4.8	6.7	7.5	9.4	9.8	10.3	11.3	12.1
2021 mean Decay-SHIFOR5 error (kt)	6.4	9.8	12.4	14.7	16.9	17.8	18.3	17.8
2021 mean OFCL skill relative to Decay-SHIFOR5 (%)	25.0	31.6	39.5	36.1	42.0	42.1	38.3	32.0
2021 OFCL bias (kt)	0.6	0.8	0.6	0.2	0.0	-0.4	-1.1	-3.6
2021 number of cases	351	309	272	244	216	189	151	115
2016-20 mean OFCL error (kt)	5.5	8.2	9.9	11.1	11.5	12.4	13.8	15.4
2016-20 mean Decay-SHIFOR5 error (kt)	7.2	11.3	14.7	17.2	18.5	20.2	22.3	23.6
2016-20 mean OFCL skill relative to Decay-SHIFOR5 (%)	23.6	27.4	32.7	35.5	37.8	38.6	38.1	34.7
2016-20 OFCL bias (kt)	-0.1	-0.4	-0.8	-1.0	0.3	-0.9	-2.1	-4.8
2016-20 number of cases	351	309	272	244	216	189	151	115
2021 OFCL error relative to 2016-20 mean (%)	-12.7	-18.3	-24.2	-15.3	-14.8	-16.9	-18.1	-21.4
2021 Decay-SHIFOR5 error relative to 2016-20 mean (%)	-11.1	-13.3	-15.6	-14.5	-8.6	-11.9	-17.9	-24.6

Table 5a. Homogenous comparison of selected Atlantic basin early intensity guidance model errors (kt) for 2021. Errors smaller than the NHC official forecast are shown in boldface.

Model ID	Forecast Period (h)							
	12	24	36	48	60	72	96	120
OFCL	5.6	7.4	7.5	9.5	9.8	9.4	11.0	10.1
OCD5	7.2	10.7	12.8	14.8	16.8	16.9	17.4	17.6
HWFI	6.5	8.5	9.6	11.7	13.1	12.8	12.9	11.2
HMNI	6.6	8.5	8.8	10.7	11.3	11.8	**10.8**	10.8
CTCI	6.7	8.3	9.7	11.1	12.6	14.1	17.0	15.3
DSHP	6.7	9.4	9.9	11.2	12.4	12.5	13.2	13.6
LGEM	6.9	9.7	10.4	11.6	12.6	12.3	14.6	17.0
IVCN	5.9	7.6	7.8	**8.9**	9.9	9.8	**9.8**	**9.2**
FSSE	5.9	7.4	7.7	**8.5**	9.8	9.9	12.2	14.1
HCCA	5.9	7.5	7.9	**8.3**	**9.3**	9.4	**9.7**	**9.7**
GFSI	7.0	9.6	10.4	11.4	12.9	13.8	15.4	19.2
EMXI	8.4	12.6	13.8	15.3	17.3	16.8	14.0	14.3
NNIC	6.0	7.9	8.2	9.9	12.0	12.3	12.5	13.5
Forecasts	207	186	169	150	133	116	86	61

Table 5b. Homogenous comparison of selected Atlantic basin early intensity guidance model biases (kt) for 2021. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)							
Model ID	12	24	36	48	60	72	96	120
OFCL	1.3	1.7	2.1	1.9	1.6	1.9	3.0	-0.2
OCD5	**-0.4**	**-1.2**	**-0.1**	**-0.1**	**0.2**	**0.2**	**0.3**	-9.0
HWFI	-2.5	-3.0	-2.4	**-1.6**	**-0.9**	**-1.6**	**-2.6**	-3.8
HMNI	-1.8	-3.3	-4.6	-5.9	-6.2	-6.3	-5.4	-6.5
CTCI	-1.6	-3.0	-2.6	**-1.3**	**0.5**	2.3	**1.4**	0.5
DSHP	**0.3**	**0.3**	**0.6**	**0.2**	**-0.3**	**-0.5**	**0.6**	-1.7
LGEM	**-0.4**	**-1.3**	**-1.1**	**-1.0**	**-0.9**	**-1.1**	**-0.3**	-6.0
IVCN	**-0.9**	-1.8	**-1.7**	**-1.7**	**-1.3**	**-1.1**	**-1.0**	-3.5
FSSE	**-0.6**	**-1.1**	**-1.2**	**-1.4**	-1.7	-1.9	**-0.6**	-4.9
HCCA	**-0.3**	**-0.4**	**-0.2**	**0.4**	**0.8**	**1.1**	**2.9**	1.6
GFSI	-1.5	-3.0	-4.2	-5.1	-6.4	-8.1	-10.8	-18.3
EMXI	-1.5	-2.8	-4.5	-6.1	-8.5	-9.0	-7.1	-8.6
NNIC	**0.7**	**1.5**	**2.0**	4.4	5.0	4.6	5.1	-1.3
Forecasts	207	186	169	150	133	116	86	61

Table 6. Official Atlantic track and intensity forecast verifications (OFCL) for 2021 by storm. CLIPER5 (CLP5) and SHIFOR5 (SHF5) forecast errors are given for comparison and indicated collectively as OCD5. The number of track and intensity forecasts are given by NT and NI, respectively. Units for track and intensity errors are n mi and kt, respectively.

Verification statistics for: AL012021 ANA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	3.0	3.0	6	1.7	2.5
012	4	30.8	65.8	4	3.8	4.5
024	2	45.2	187.1	2	2.5	6.5
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL022021 BILL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	17.1	17.1	6	3.3	5.8
012	4	53.9	115.1	4	5.0	11.0
024	2	70.9	251.4	2	5.0	9.5
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL032021 CLAUDETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	14.4	14.4	13	1.5	1.9
012	11	21.1	51.2	11	2.7	4.7
024	9	35.7	135.4	9	1.7	7.1
036	7	49.1	210.8	7	2.1	7.9
048	5	70.7	279.3	5	5.0	10.4
060	3	129.7	444.4	3	8.3	12.7
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL042021 DANNY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	3	4.6	4.6	3	5.0	5.0
012	1	24.5	31.4	1	5.0	1.0
024	0	-999.0	-999.0	0	-999.0	-999.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL052021 ELSA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	36	6.9	7.5	36	1.5	1.7
012	34	26.3	48.7	34	5.4	8.8
024	32	37.8	103.9	32	8.0	12.6
036	30	40.4	162.6	30	8.7	13.7
048	28	41.9	215.8	28	10.4	15.1
060	26	48.1	272.8	26	7.3	18.3
072	24	57.0	329.8	24	6.9	19.0
096	20	72.2	411.0	20	4.8	24.0
120	16	100.6	491.0	16	8.1	14.2

Verification statistics for: AL062021 FRED

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	22	10.4	10.3	22	1.1	0.9
012	18	25.7	38.9	18	3.3	3.4
024	14	34.9	58.7	14	3.2	6.9
036	10	37.7	71.2	10	2.5	8.1
048	8	52.5	76.3	8	2.5	7.6
060	6	88.0	153.8	6	5.8	9.3
072	6	111.8	214.4	6	7.5	8.8
096	10	104.3	226.5	10	6.0	14.3
120	8	85.0	231.0	8	4.4	11.6

Verification statistics for: AL072021 GRACE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	33	6.1	6.6	33	3.2	3.2
012	31	27.9	37.9	31	6.9	8.5
024	29	42.4	71.2	29	9.8	13.1
036	27	62.7	117.7	27	12.2	16.2
048	25	80.4	169.1	25	15.4	17.4
060	23	97.9	235.0	23	15.0	19.5
072	21	120.3	304.8	21	14.3	18.4
096	17	212.5	451.8	17	20.3	18.2
120	13	365.3	643.5	13	25.0	18.5

Verification statistics for: AL082021 HENRI

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	31	3.5	3.8	31	0.3	0.3
012	29	23.9	48.6	29	4.0	4.6
024	27	37.6	129.0	27	5.7	6.7
036	25	45.6	204.3	25	8.0	9.4
048	23	45.9	249.3	23	9.8	10.6
060	21	61.9	309.8	21	11.0	11.5
072	19	97.5	370.2	19	12.4	11.5
096	15	183.7	450.7	15	16.0	12.7
120	11	295.6	525.9	11	21.8	14.5

Verification statistics for: AL092021 IDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	24	4.4	4.4	24	1.9	1.9
012	22	14.5	32.2	22	6.4	10.5
024	20	20.4	70.5	20	8.8	14.1
036	18	26.4	122.6	18	4.7	15.4
048	16	35.8	180.1	16	9.4	19.4
060	14	44.7	223.5	14	11.4	25.6
072	12	56.3	238.1	12	10.8	22.9
096	8	102.9	270.3	8	11.9	11.2
120	4	181.7	305.4	4	5.0	18.5

Verification statistics for: AL102021 KATE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	8.5	8.5	18	0.6	0.6
012	16	27.4	47.4	16	3.1	4.3
024	14	45.5	83.2	14	3.6	7.7
036	12	64.7	114.0	12	3.3	10.2
048	10	88.2	144.3	10	4.0	13.1
060	8	108.1	191.8	8	6.2	13.4
072	6	108.7	236.3	6	6.7	16.7
096	2	94.6	542.4	2	0.0	21.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL112021 JULIAN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	7.4	7.4	5	1.0	1.0
012	4	40.6	113.2	4	6.2	10.0
024	2	53.9	309.0	2	2.5	10.5
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL122021 LARRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	43	8.0	8.2	43	1.9	1.9
012	41	17.4	39.1	41	4.8	5.9
024	39	22.8	84.0	39	7.7	8.5
036	37	28.8	122.6	37	7.8	9.8
048	35	35.3	155.8	35	8.9	11.6
060	33	43.3	180.3	33	9.7	11.8
072	31	50.1	202.2	31	11.1	12.7
096	27	63.7	228.8	27	13.1	13.3
120	23	87.0	204.5	23	12.8	15.9

Verification statistics for: AL132021 MINDY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	0.0	0.0	5	2.0	1.0
012	3	20.2	48.9	3	1.7	0.3
024	1	39.0	72.7	1	0.0	0.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL142021 NICHOLAS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	16.3	15.9	13	1.2	1.5
012	11	30.6	52.3	11	3.2	6.2
024	9	39.9	110.3	9	5.6	7.9
036	7	57.7	159.5	7	5.0	9.3
048	5	77.7	185.0	5	2.0	7.8
060	3	100.3	224.3	3	5.0	12.3
072	1	108.8	363.2	1	5.0	35.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL152021 ODETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	3	33.5	33.5	3	1.7	1.7
012	1	56.5	130.2	1	0.0	2.0
024	0	-999.0	-999.0	0	-999.0	-999.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL162021 PETER

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	16	20.5	16.7	16	0.6	0.9
012	14	40.1	48.8	14	3.9	5.3
024	12	52.4	76.5	12	6.2	11.2
036	10	72.5	112.2	10	5.5	19.2
048	8	104.1	154.6	8	5.0	24.1
060	6	123.2	200.5	6	2.5	29.8
072	4	152.6	250.7	4	3.8	32.8
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL172021 ROSE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	9.1	9.1	13	0.8	0.8
012	11	25.9	48.3	11	3.6	5.6
024	9	40.8	87.8	9	3.3	6.8
036	7	52.6	116.2	7	3.6	9.3
048	5	62.1	153.6	5	4.0	12.2
060	3	67.5	194.3	3	3.3	13.0
072	1	50.0	296.6	1	0.0	19.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL182021 SAM

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	50	4.2	4.2	50	1.1	1.1
012	48	17.0	34.4	48	6.1	7.7
024	46	25.3	78.6	46	7.5	12.4
036	44	33.1	118.4	44	8.8	16.2
048	42	41.4	153.8	42	11.0	18.9
060	40	50.8	192.4	40	11.2	20.5
072	38	57.9	226.7	38	12.0	22.2
096	34	77.6	335.1	34	11.5	22.2
120	30	105.5	399.5	30	10.2	27.2

Verification statistics for: AL192021 TERESA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	4	10.5	10.5	4	1.2	1.2
012	2	61.0	125.7	2	10.0	15.0
024	0	-999.0	-999.0	0	-999.0	-999.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL202021 VICTOR

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	20	11.3	11.3	20	0.2	0.2
012	18	31.5	42.4	18	2.5	5.1
024	16	52.7	67.6	16	5.9	10.8
036	14	68.8	91.9	14	10.4	15.0
048	12	92.2	88.3	12	14.2	22.6
060	10	113.4	112.6	10	19.0	27.6
072	8	152.5	163.7	8	20.0	30.1
096	4	232.5	301.0	4	18.8	31.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL212021 WANDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	30	5.2	5.2	30	2.3	2.5
012	28	27.8	85.9	28	5.2	3.9
024	26	36.9	188.4	26	6.2	4.7
036	24	34.4	272.1	24	5.8	6.3
048	22	42.1	377.5	22	6.4	8.1
060	20	57.4	434.4	20	4.5	9.5
072	18	68.3	454.0	18	3.3	11.6
096	14	188.2	452.5	14	3.2	18.9
120	10	405.3	364.8	10	4.0	7.4

Table 7. Homogenous comparison of official and CLIPER5 track forecast errors in the eastern North Pacific basin in 2021 for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)							
	12	24	36	48	60	72	96	120
2021 mean OFCL error (n mi)	21.8	31.4	42.4	53.7	66.9	77.6	90.1	113.7
2021 mean CLIPER5 error (n mi)	39.9	78.3	120.4	166.7	218.2	260.4	337.7	399.0
2021 mean OFCL skill relative to CLIPER5 (%)	45.4	59.9	64.8	67.8	69.3	70.2	73.3	71.5
2021 mean OFCL bias vector (°/n mi)	217/001	180/005	179/009	177/014	176/021	186/027	200/040	236/034
2021 number of cases	275	235	197	161	132	108	68	43
2016-2020 mean OFCL error (n mi)	21.1	33.0	44.1	54.7	79.5	76.2	96.2	117.2
2016-2020 mean CLIPER5 error (n mi)	33.9	69.2	107.9	147.1	187.1	220.1	281.1	342.5
2016-2020 mean OFCL skill relative to CLIPER5 (%)	37.8	52.3	59.1	62.8	57.5	65.4	65.8	65.8
2016-2020 mean OFCL bias vector (°/n mi)	338/002	355/002	332/001	309/002	222/007	324/003	006/003	343/009
2016-2020 number of cases	1505	1323	1162	1019	223	784	597	453
2021 OFCL error relative to 2016-2020 mean (%)	3.3	-4.8	-3.9	-1.8	-15.8	1.8	-6.3	-3.0
2021 CLIPER5 error relative to 2016-2021 mean (%)	17.7	13.2	11.6	13.3	16.6	18.3	20.1	16.5

Table 8a. Homogenous comparison of eastern North Pacific basin early track guidance model errors (n mi) for 2021. Errors smaller than the NHC official forecast are shown in boldface.

Model ID	Forecast Period (h)							
	12	24	36	48	60	72	96	120
OFCL	18.9	29.8	40.1	49.6	63.5	70.0	85.2	81.1
OCD5	36.4	76.4	120.4	162.1	209.7	237.6	302.0	362.6
GFSI	21.3	34.3	47.8	58.3	70.5	75.6	94.5	107.0
HWFI	23.2	38.4	52.2	60.3	74.9	95.9	141.1	156.1
HMNI	23.5	35.8	50.8	69.8	90.8	109.5	157.5	197.0
EMXI	20.9	33.8	47.4	59.1	74.4	84.8	89.0	85.0
EGRI	25.3	43.3	67.2	91.7	116.7	132.8	172.3	158.6
CMCI	24.7	39.6	55.2	70.5	83.3	101.1	162.7	176.0
NVGI	28.9	51.2	66.7	83.3	102.5	116.3	154.7	197.9
AEMI	21.2	34.9	47.3	55.7	69.3	75.7	92.6	98.9
FSSE	**17.7**	**27.8**	**36.7**	**44.0**	**57.0**	**67.9**	**83.1**	93.4
TVCE	**18.2**	**28.6**	**39.0**	**49.0**	**62.4**	72.4	95.9	102.0
HCCA	**17.6**	**27.8**	**36.9**	**44.4**	**55.7**	**66.1**	92.1	95.6
Forecasts	178	147	124	103	80	59	35	20

Table 8b. Homogenous comparison of eastern North Pacific basin early track guidance model bias vectors (º/n mi) for 2021.

	Forecast Period (h)							
Model ID	12	24	36	48	60	72	96	120
OFCL	161/002	169/006	182/010	189/016	200/020	222/029	224/044	282/026
OCD5	189/003	212/011	233/025	235/039	245/063	256/074	320/091	028/210
GFSI	099/005	088/011	096/015	101/018	114/019	143/013	210/018	044/020
HWFI	121/007	122/013	135/016	152/017	191/012	247/026	240/057	260/075
HMNI	139/004	128/009	142/009	185/008	252/013	269/037	264/110	268/180
EMXI	261/006	239/014	233/025	220/032	213/036	210/047	206/068	237/054
EGRI	252/006	239/015	246/026	251/038	258/061	270/092	282/127	277/130
CMCI	130/006	125/016	125/028	125/040	121/051	124/068	134/127	106/126
NVGI	329/002	250/006	256/017	262/031	263/054	271/069	288/083	329/135
AEMI	093/005	102/009	126/013	141/018	162/022	180/025	198/039	165/009
FSSE	190/002	155/006	154/010	149/016	155/019	184/023	209/043	247/040
TVCE	153/002	163/005	187008	199/012	221/017	240/029	247/057	265/070
HCCA	156/003	165/007	178/012	184/017	201/022	218/028	226/047	257/031
Forecasts	178	147	124	103	80	59	35	20

Table 9. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the eastern North Pacific basin for the 2021 season for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)							
	12	24	36	48	60	72	96	120
2021 mean OFCL error (kt)	5.2	8.7	11.3	13.3	15.6	17.8	20.6	20.0
2021 mean Decay-SHIFOR5 error (kt)	6.4	11.7	16.0	21.0	23.4	25.5	28.6	24.5
2021 mean OFCL skill relative to Decay-SHIFOR5 (%)	18.8	25.6	29.4	36.7	33.3	30.2	28.0	18.4
2021 OFCL bias (kt)	1.4	2.2	3.4	3.7	3.0	2.1	0.6	-2.1
2021 number of cases	275	235	197	161	132	108	68	43
2016-20 mean OFCL error (kt)	5.6	9.0	10.9	12.6	15.0	15.3	15.9	16.6
2016-20 mean Decay-SHIFOR5 error (kt)	7.3	12.1	15.4	17.7	18.6	20.6	21.4	20.8
2016-20 mean OFCL skill relative to Decay-SHIFOR5 (%)	23.3	25.6	29.2	28.8	19.4	25.7	25.7	20.2
2016-20 OFCL bias (kt)	-0.1	-0.2	-0.4	-0.7	2.7	-0.9	-1.6	-4.3
2016-20 number of cases	1505	1323	1162	1019	223	784	597	453
2021 OFCL error relative to 2016-20 mean (%)	-7.1	-3.3	3.7	5.6	4.0	16.3	29.6	20.5
2021 Decay-SHIFOR5 error relative to 2016-20 mean (%)	-12.3	-3.3	3.9	18.6	25.8	23.8	33.6	17.8

Table 10a. Homogenous comparison of eastern North Pacific basin early intensity guidance model errors (kt) for 2021. Errors smaller than the NHC official forecast are shown in boldface.

Model ID	Forecast Period (h)							
	12	24	36	48	60	72	96	120
OFCL	5.7	9.8	13.1	15.1	17.5	19.2	18.6	19.3
OCD5	6.8	12.0	16.0	19.4	23.2	24.3	23.5	21.7
HWFI	8.2	11.1	**12.9**	15.6	17.9	19.9	21.5	23.6
HMNI	6.5	**8.9**	**11.4**	**14.5**	17.8	19.6	20.2	**17.2**
DSHP	6.6	11.3	15.5	17.0	17.7	**18.5**	**17.3**	**15.7**
LGEM	6.8	11.2	14.8	16.9	17.6	19.2	21.7	22.9
IVCN	6.2	**9.5**	**12.3**	**14.4**	**16.2**	**17.5**	**18.5**	**18.1**
HCCA	6.2	9.9	13.6	15.2	**16.2**	**16.8**	**18.1**	20.5
FSSE	6.2	**9.6**	**12.6**	**15.0**	**17.0**	**18.0**	**18.2**	**17.3**
GFSI	7.3	11.3	14.5	17.8	20.2	19.8	**18.1**	**17.2**
EMXI	8.2	12.1	14.9	17.8	19.6	19.5	**16.6**	**16.8**
NNIC	6.3	10.1	13.1	**14.1**	**15.6**	**16.9**	22.3	26.4
Forecasts	204	177	148	120	97	73	45	27

Table 10b. Homogenous comparison of eastern North Pacific basin early intensity guidance model biases (kt) for 2021. Biases smaller than the NHC official forecast are shown in boldface.

Model ID	Forecast Period (h)							
	12	24	36	48	60	72	96	120
OFCL	1.8	2.8	4.4	4.0	2.2	-0.7	-4.6	-8.5
OCD5	**0.1**	**1.2**	**1.7**	**1.2**	**1.4**	**-1.2**	-5.9	-11.0
HWFI	-4.5	-4.8	-4.2	-4.4	-4.3	-3.9	**-1.1**	**-3.3**
HMNI	**-0.2**	**-1.0**	**-1.1**	**-2.0**	-4.6	-7.9	-9.9	-14.6
DSHP	**0.2**	**0.7**	**1.8**	**1.5**	**0.4**	**0.2**	**-0.4**	**-0.6**
LGEM	**-0.7**	**-2.2**	**-3.1**	-4.3	-5.9	-6.9	-9.0	-11.8
IVCN	**-1.1**	**-1.5**	**-1.1**	**-1.5**	-2.7	-4.0	**-4.3**	**-7.1**
HCCA	**0.5**	**1.0**	**0.8**	**-0.5**	-2.6	-4.6	-6.5	-9.2
FSSE	**0.2**	**0.8**	**0.7**	**-0.1**	**-1.9**	-4.3	-6.3	-9.0
GFSI	-1.8	-3.4	-4.6	-7.1	-9.2	-11.0	-12.7	-12.6
EMXI	-2.8	-4.1	-4.7	-4.6	-5.4	-7.0	-8.9	-11.8
NNIC	**0.1**	**0.4**	**1.2**	**-0.3**	**-1.5**	-1.8	**-0.4**	**-4.8**
Forecasts	204	177	148	120	97	73	45	27

Table 11. Official eastern North Pacific track and intensity forecast verifications (OFCL) for 2021 by storm. CLIPER5 (CLP5) and SHIFOR5 (SHF5) forecast errors are given for comparison and indicated collectively as OCD5. The number of track and intensity forecasts are given by NT and NI, respectively. Units for track and intensity errors are n mi and kt, respectively.

Verification statistics for: EP012021 ANDRES

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	11.5	13.3	8	1.2	1.2
012	6	25.6	34.1	6	1.7	4.2
024	4	29.6	63.2	4	1.2	8.2
036	2	32.3	90.9	2	2.5	14.5
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP022021 BLANCA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	9.3	9.4	17	0.9	0.9
012	15	19.8	35.4	15	4.0	6.0
024	13	25.4	72.3	13	5.4	10.4
036	11	26.9	112.3	11	5.5	11.6
048	9	30.2	148.7	9	7.8	16.0
060	7	30.2	207.2	7	10.0	19.4
072	5	28.7	267.1	5	12.0	18.6
096	1	40.9	164.5	1	20.0	23.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP032021 CARLOS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	6.7	7.4	15	1.0	1.0
012	13	25.0	34.9	13	3.8	5.2
024	11	55.7	84.5	11	5.0	7.9
036	9	109.8	144.9	9	8.9	11.8
048	7	180.9	204.0	7	10.7	16.4
060	5	272.8	257.8	5	11.0	27.0
072	3	370.1	321.4	3	11.7	24.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP042021 DOLORES

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	12.2	12.2	8	0.0	0.0
012	6	39.2	80.9	6	5.0	8.0
024	4	72.3	183.1	4	5.0	12.5
036	2	130.2	320.0	2	17.5	11.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP052021 ENRIQUE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	22	5.0	5.0	22	0.5	0.7
012	20	18.4	36.4	20	5.0	5.6
024	18	28.1	72.1	18	9.2	10.1
036	16	38.0	102.7	16	11.2	13.6
048	14	52.4	152.6	14	12.5	15.8
060	12	69.3	225.1	12	10.8	17.2
072	10	84.5	290.9	10	10.0	16.6
096	6	140.3	428.7	6	10.8	10.3
120	2	222.3	488.5	2	10.0	11.0

Verification statistics for: EP062021 FELICIA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	26	3.1	3.1	26	1.2	1.2
012	24	9.6	26.2	24	7.1	9.0
024	22	17.3	60.9	22	16.1	18.6
036	20	23.9	97.3	20	23.8	28.0
048	18	31.2	126.4	18	28.9	33.7
060	16	44.6	146.3	16	30.6	35.6
072	14	58.3	157.5	14	32.1	37.6
096	10	89.5	230.1	10	27.5	32.9
120	6	83.3	303.9	6	20.0	13.8

Verification statistics for: EP072021 GUILLERMO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	10	13.3	13.3	10	2.0	2.0
012	8	23.8	43.7	8	6.9	5.0
024	6	34.7	73.3	6	10.8	12.2
036	4	39.6	62.4	4	15.0	10.0
048	2	47.7	59.5	2	20.0	3.5
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP082021 HILDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	26	7.1	7.5	26	2.1	2.1
012	24	20.7	29.6	24	4.8	6.0
024	22	31.1	57.1	22	7.0	8.9
036	20	36.4	85.4	20	7.0	10.8
048	18	41.5	119.0	18	6.1	12.6
060	16	48.9	155.4	16	7.5	17.6
072	14	57.5	202.0	14	8.2	19.5
096	10	98.0	283.1	10	9.5	21.9
120	6	160.1	357.9	6	9.2	17.5

Verification statistics for: EP092021 JIMENA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	9.1	9.0	15	0.0	0.0
012	11	27.9	30.6	11	1.8	3.1
024	7	29.6	39.3	7	2.9	4.0
036	3	27.9	31.5	3	1.7	2.7
048	1	51.0	51.0	1	0.0	4.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	2	133.0	196.4	2	10.0	31.0
120	6	182.5	314.7	6	13.3	27.3

Verification statistics for: EP102021 IGNACIO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	9	11.7	10.7	9	0.0	0.0
012	7	31.9	46.9	7	0.0	4.0
024	5	55.3	73.0	5	0.0	6.4
036	3	86.9	117.8	3	0.0	9.0
048	1	197.1	237.1	1	0.0	13.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP112021 KEVIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	20	14.0	13.2	20	1.5	1.5
012	18	21.5	35.9	18	3.6	5.2
024	16	21.1	59.3	16	8.1	8.4
036	14	29.6	84.3	14	13.2	12.5
048	12	42.9	114.3	12	19.2	16.7
060	10	63.5	141.5	10	23.5	22.1
072	8	76.1	168.3	8	25.0	23.6
096	4	82.3	195.8	4	26.2	23.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP122021 LINDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	38	6.1	6.1	38	2.4	2.4
012	36	16.0	35.8	36	5.8	6.6
024	34	23.0	74.6	34	9.3	12.0
036	32	29.0	113.8	32	9.8	14.1
048	30	32.3	147.1	30	11.3	16.9
060	28	35.2	183.0	28	14.1	20.5
072	26	39.1	217.1	26	17.7	24.3
096	22	49.8	286.4	22	22.3	28.2
120	18	65.4	336.8	18	24.2	30.3

Verification statistics for: EP132021 MARTY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	4	6.5	6.5	4	2.5	2.5
012	2	27.6	43.2	2	2.5	2.0
024	0	-999.0	-999.0	0	-999.0	-999.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP142021 NORA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	19	27.5	27.7	19	3.7	3.7
012	18	34.7	65.0	18	8.1	4.6
024	16	43.2	112.1	16	12.5	7.6
036	14	54.9	164.1	14	15.7	12.8
048	12	61.9	221.7	12	18.8	20.8
060	10	88.0	315.1	10	24.0	21.1
072	8	124.1	388.6	8	26.2	25.4
096	4	176.3	458.7	4	40.0	42.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP152021 OLAF

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	14	7.9	7.9	14	1.1	1.1
012	12	27.3	46.4	12	8.3	10.9
024	10	49.8	106.6	10	12.5	19.2
036	8	74.2	171.6	8	13.8	23.1
048	6	89.6	225.3	6	12.5	23.0
060	4	91.8	265.3	4	8.8	13.5
072	2	88.3	203.2	2	15.0	8.5
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP162021 PAMELA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	7.9	7.9	15	1.3	1.3
012	13	22.6	52.7	13	7.7	6.2
024	11	31.2	113.9	11	10.5	5.1
036	9	33.9	202.2	9	18.3	8.0
048	7	55.7	303.7	7	20.7	12.0
060	5	97.1	430.1	5	29.0	12.0
072	3	161.9	540.9	3	38.3	15.7
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP172021 RICK

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	14	8.9	8.9	14	2.5	2.5
012	12	22.9	44.4	12	9.6	9.6
024	10	30.0	85.9	10	14.5	15.5
036	8	41.2	134.3	8	11.9	14.2
048	6	51.2	180.3	6	6.7	12.0
060	4	68.9	279.7	4	12.5	23.2
072	2	77.6	433.4	2	12.5	20.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP182021 TERRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	25	7.8	7.7	25	2.2	2.2
012	23	17.4	37.1	23	3.5	8.1
024	21	31.4	84.9	21	4.8	19.2
036	19	49.6	143.7	19	4.5	30.8
048	17	69.7	232.9	17	5.6	45.8
060	15	87.0	301.4	15	6.3	36.8
072	13	94.2	376.1	13	9.2	37.9
096	9	108.0	642.5	9	18.9	41.1
120	5	142.2	851.9	5	30.0	26.8

Verification statistics for: EP192021 SANDRA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	9	11.3	11.3	9	0.6	0.6
012	7	32.0	56.9	7	0.7	4.1
024	5	31.6	64.1	5	3.0	9.4
036	3	50.2	80.2	3	3.3	13.7
048	1	79.2	148.6	1	0.0	13.0
060	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Table 12a. Verification of 48-h probabilistic genesis forecasts for the Atlantic basin in 2021.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	2	432
10	11	152
20	14	96
30	20	59
40	36	36
50	41	39
60	57	23
70	77	26
80	67	21
90	100	27
100	-	0

Table 12b. Verification of 48-h probabilistic genesis forecasts for the eastern North Pacific basin in 2021.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	2	401
10	13	62
20	23	60
30	32	34
40	38	39
50	63	19
60	73	33
70	79	24
80	89	37
90	83	18
100	-	0

Table 13a. Verification of 120-h probabilistic genesis forecasts for the Atlantic basin in 2021.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	12	103
10	14	257
20	20	136
30	39	117
40	38	61
50	49	55
60	64	47
70	90	41
80	95	41
90	98	53
100	-	0

Table 13b. Verification of 120-h probabilistic genesis forecasts for the eastern North Pacific basin in 2021.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	9	80
10	19	127
20	37	81
30	54	79
40	41	88
50	62	47
60	]67	48
70	81	59
80	92	51
90	96	67
100		0

Table 14. NHC forecast cone circle radii (n mi) for 2022. Change from 2021 values expressed in n mi and percent are given in parentheses.

Track Forecast Cone Two-Thirds Probability Circles (n mi)		
Forecast Period (h)	**Atlantic Basin**	**Eastern North Pacific Basin**
3	16 (0: 0%)	16 (0: 0%)
12	26 (-1: -4%)	25 (0: 0%)
24	39 (-1: -3%)	38 (1: 3%)
36	52 (-3: -6%)	51 (0: 0%)
48	67 (-2: -3%)	65 (1: 2%)
60	84 (-2: -2%)	79 (2: 3%)
72	100 (-2: -2%)	93 (4: 4%)
96	142 (-6: -4%)	120 (6: 5%)
120	200 (2: 1%)	146 (8: 5%)

Table 15. Composition of NHC consensus models for 2022. It is intended that TVCA would be the primary consensus aids for the Atlantic basin and TVCE would be primary for the eastern Pacific.

NHC Consensus Model Definitions For 2022			
Model ID	**Parameter**	**Type**	**Members**
GFEX	Track	Fixed	GFSI EMXI
ICON	Intensity	Fixed	DSHP LGEM HWFI CTCI HMNI
TVCA**	Track	Variable	GFSI EGRI HWFI EMXI CTCI
TVCE	Track	Variable	GFSI EGRI HWFI EMXI CTCI HMNI EMNI
TVDG	Track	Variable	GFSI (double weight) EMXI (double weight) EGRI (double weight) CTCI HWFI
TVCX	Track	Variable	EMXI (double weight) GFSI EGRI HWFI
IVCN	Intensity	Variable	DSHP LGEM HWFI CTCI HMNI
IVDR	Intensity	Variable	CTCI (double weight) HWFI (double weight) HMNI (double weight) GFSI DSHP LGEM

** TVCN will continue to be computed and will have the same composition as TVCA. GPCE circles will continue to be based on TVCN.

LIST OF FIGURES

21. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2021 (solid lines) and 2016-2020 (dashed lines).

22. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

23. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2021. This verification includes only those models that were available at least 2/3 of the time (see text).

24. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2019-21.

25. Reliability diagram for Atlantic (top) and eastern North Pacific (bottom) probabilistic tropical cyclogenesis 48-h forecasts for 2021. The solid lines indicate the relationship between the forecasts and verifying genesis percentages, with perfect reliability indicated by the thin diagonal black line. The dashed lines indicate how the forecasts were distributed among the possible forecast values.

26. As described for Fig. 25, but for 120-h forecasts.

27. Reliability diagram for Atlantic probabilistic tropical cyclonegensis 2-, 5-, and experimental 7-day forecasts from 2019-21.

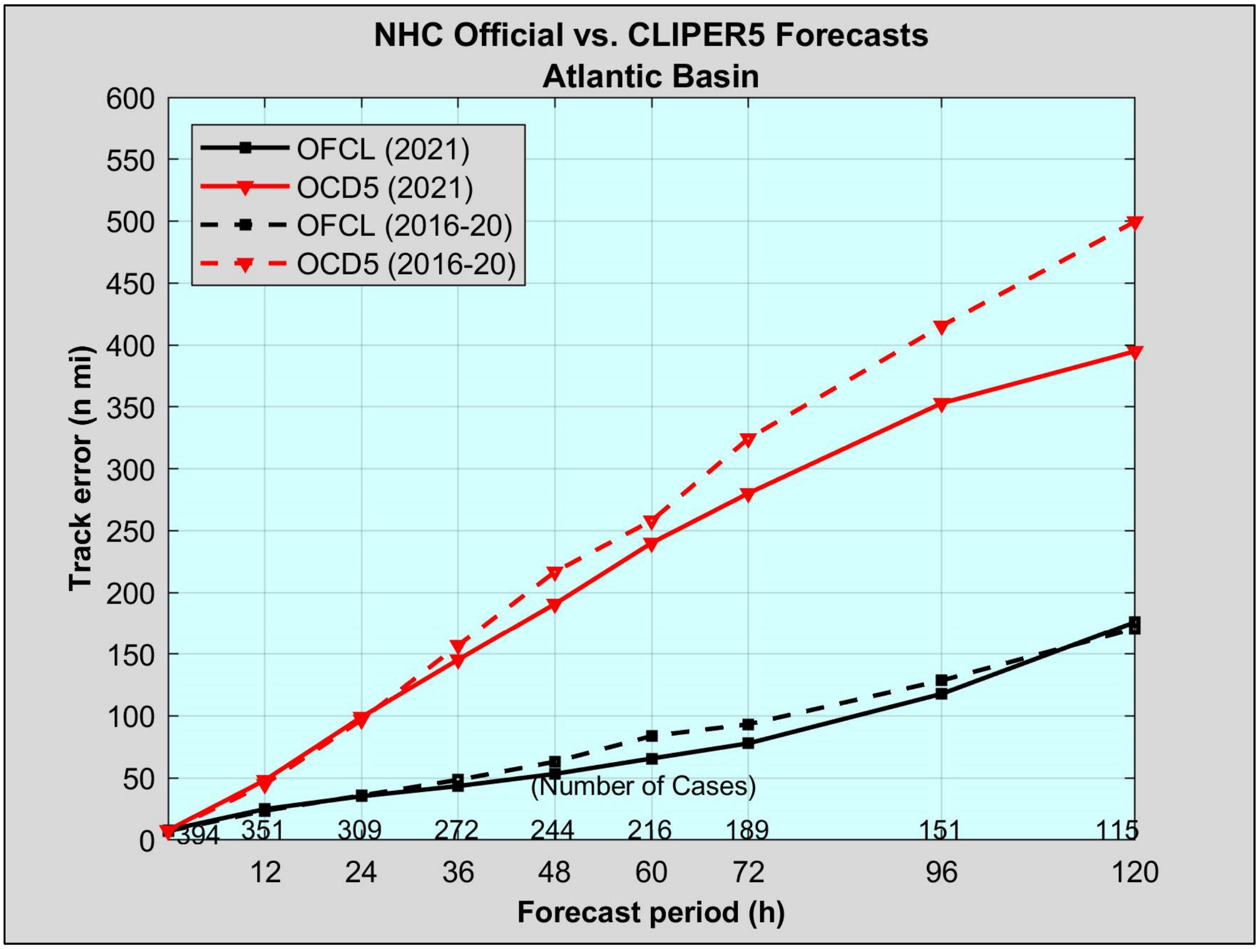


Figure 1. NHC official and CLIPER5 (OCD5) Atlantic basin average track errors for 2021 (solid lines) and 2016-2020 (dashed lines).

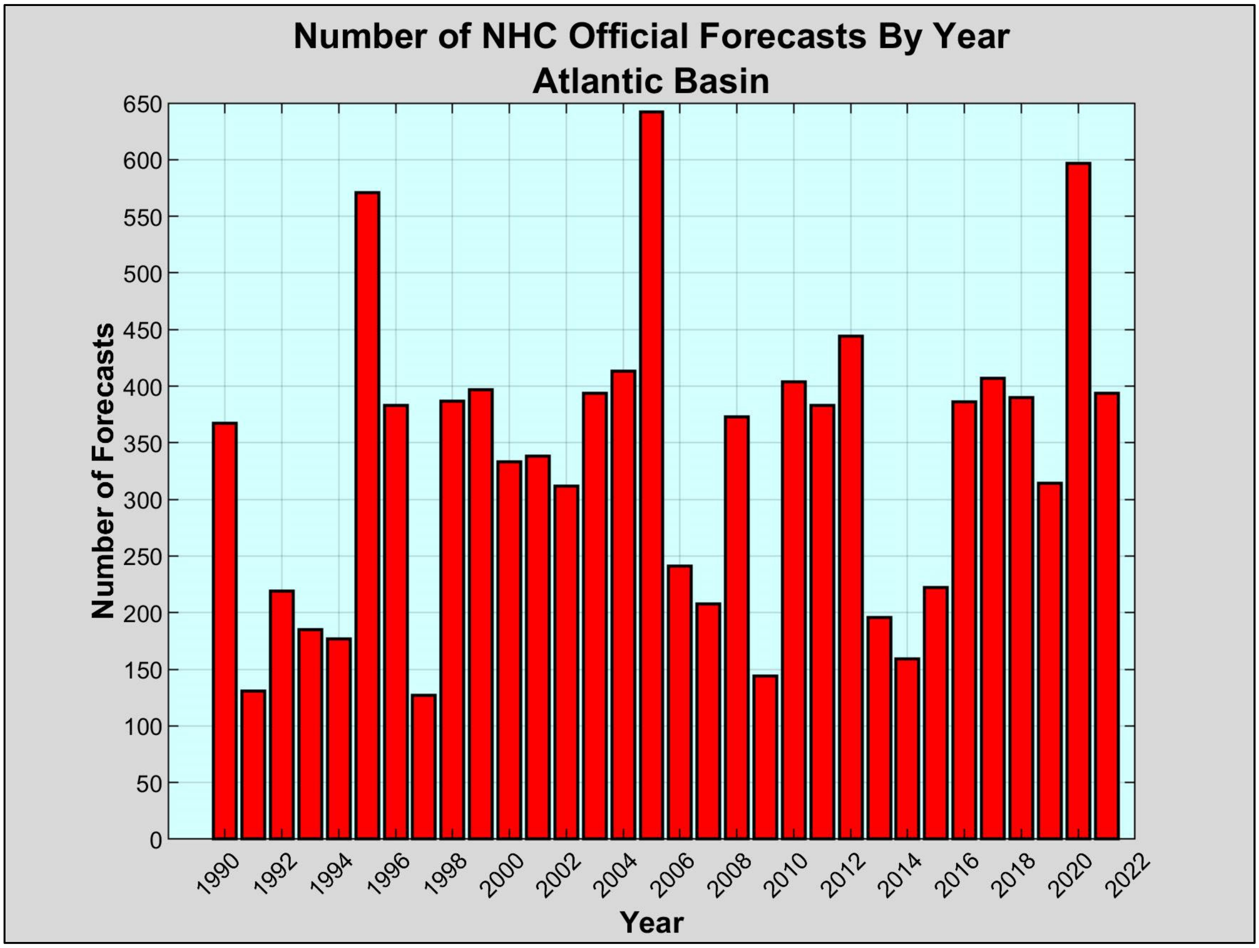


Figure 2. Number of NHC official forecasts for the Atlantic basin stratified by year.

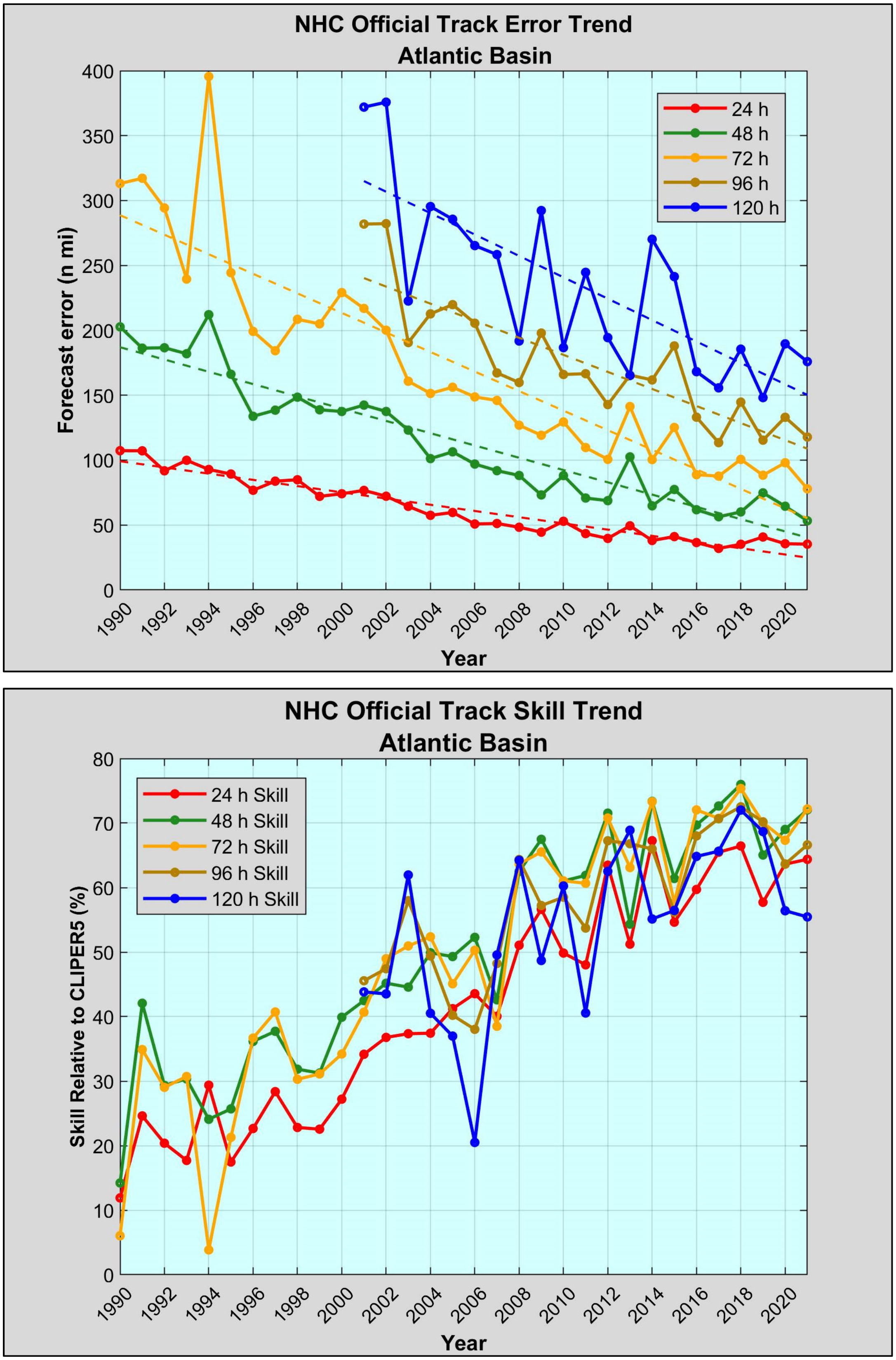


Figure 3. Recent trends in NHC official track forecast error (top) and skill (bottom) for the Atlantic basin.

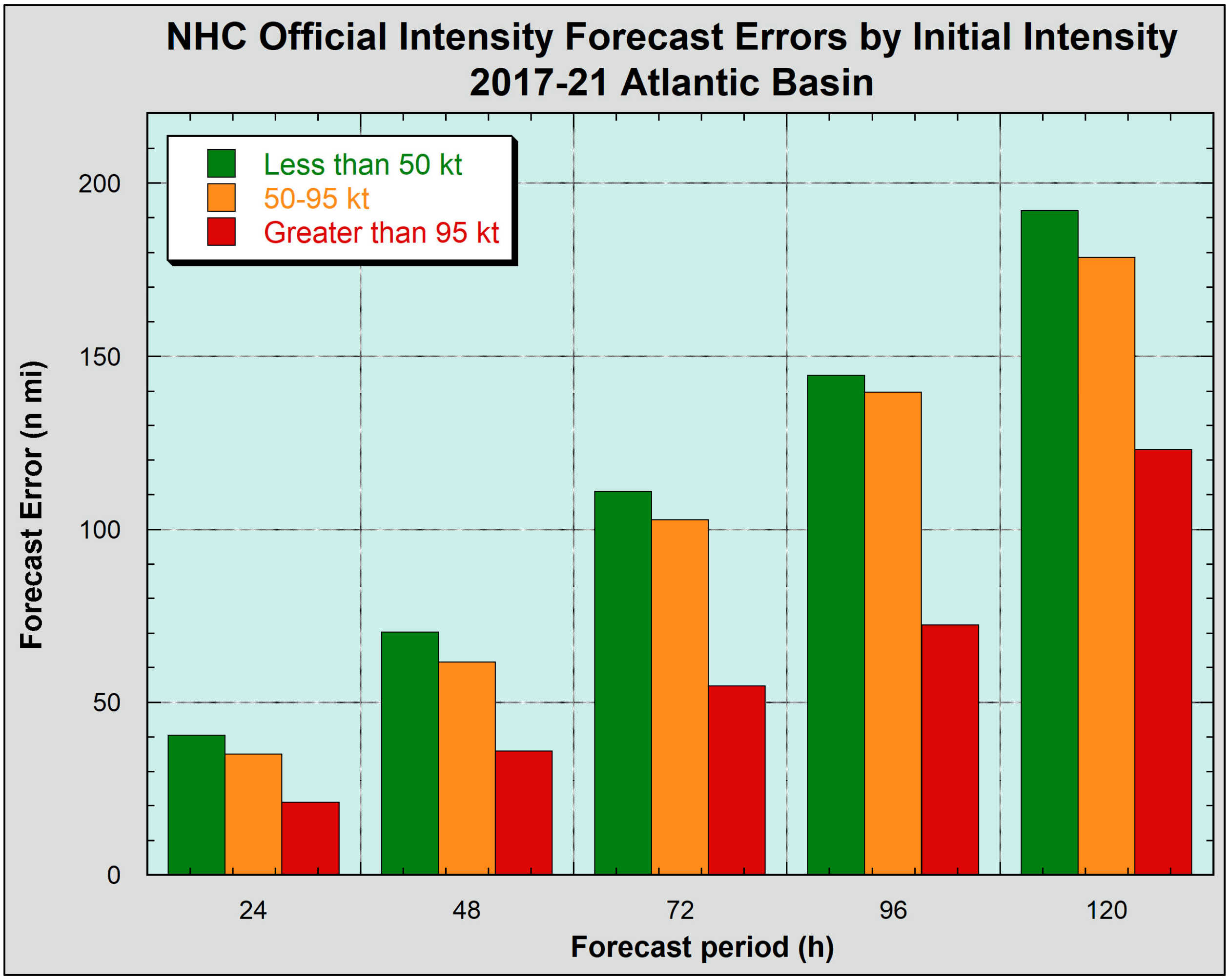


Figure 4. 2017-21 NHC official track forecast error binned by initial intensity for the Atlantic basin.

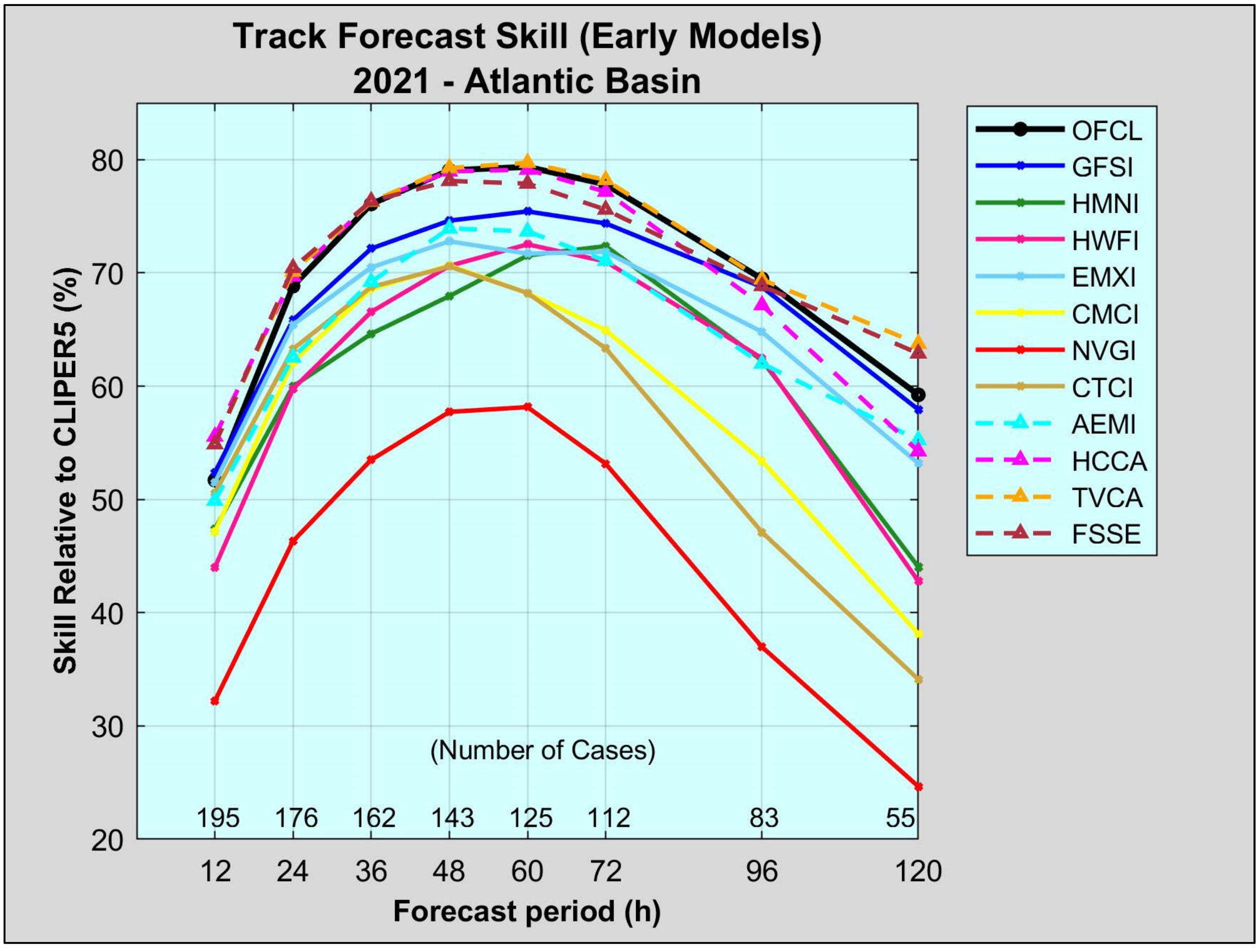


Figure 5. Homogenous comparison for selected Atlantic basin early track models for 2021. This verification includes only those models that were available at least 2/3 of the time (see text).

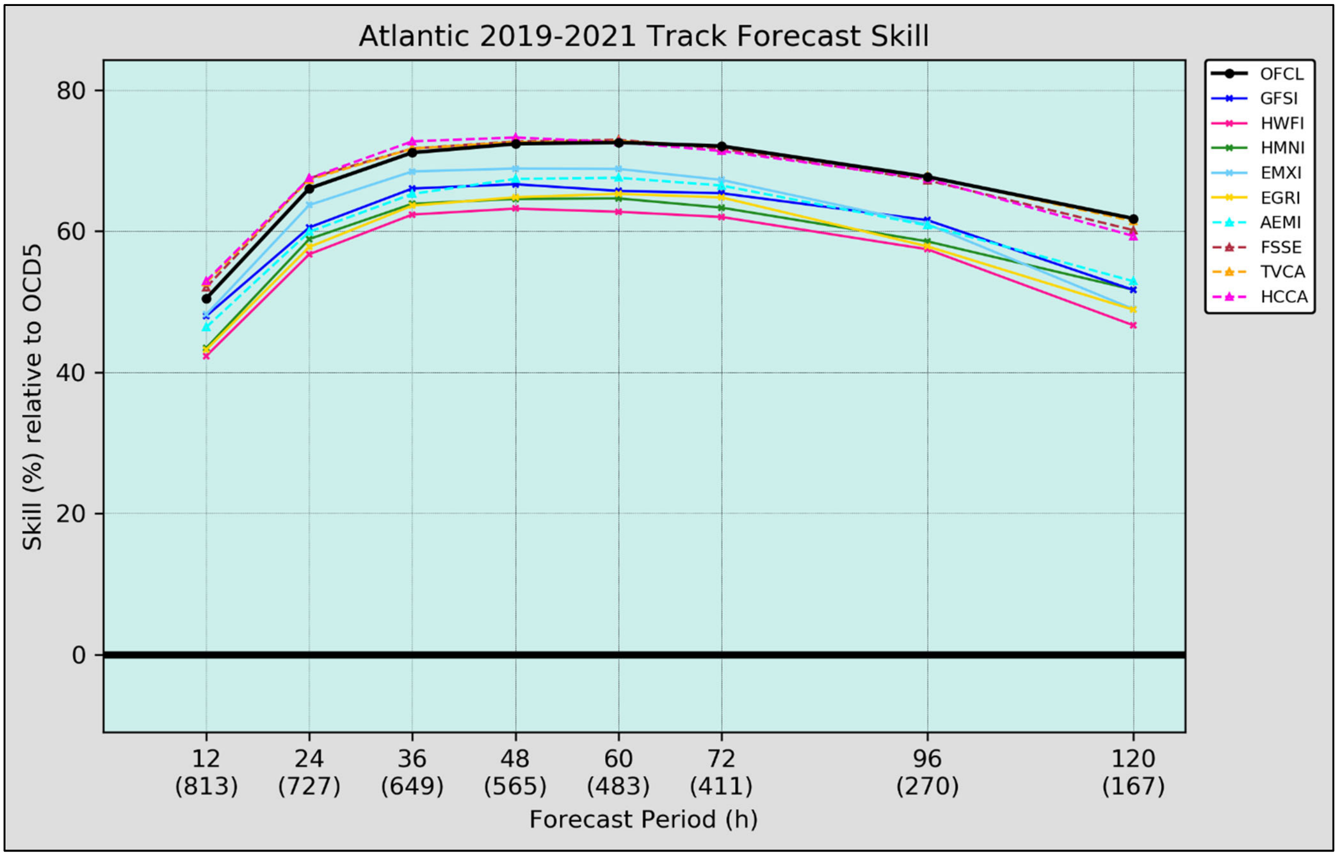


Figure 6. Homogenous comparison for selected Atlantic basin early track models for 2019-2021.

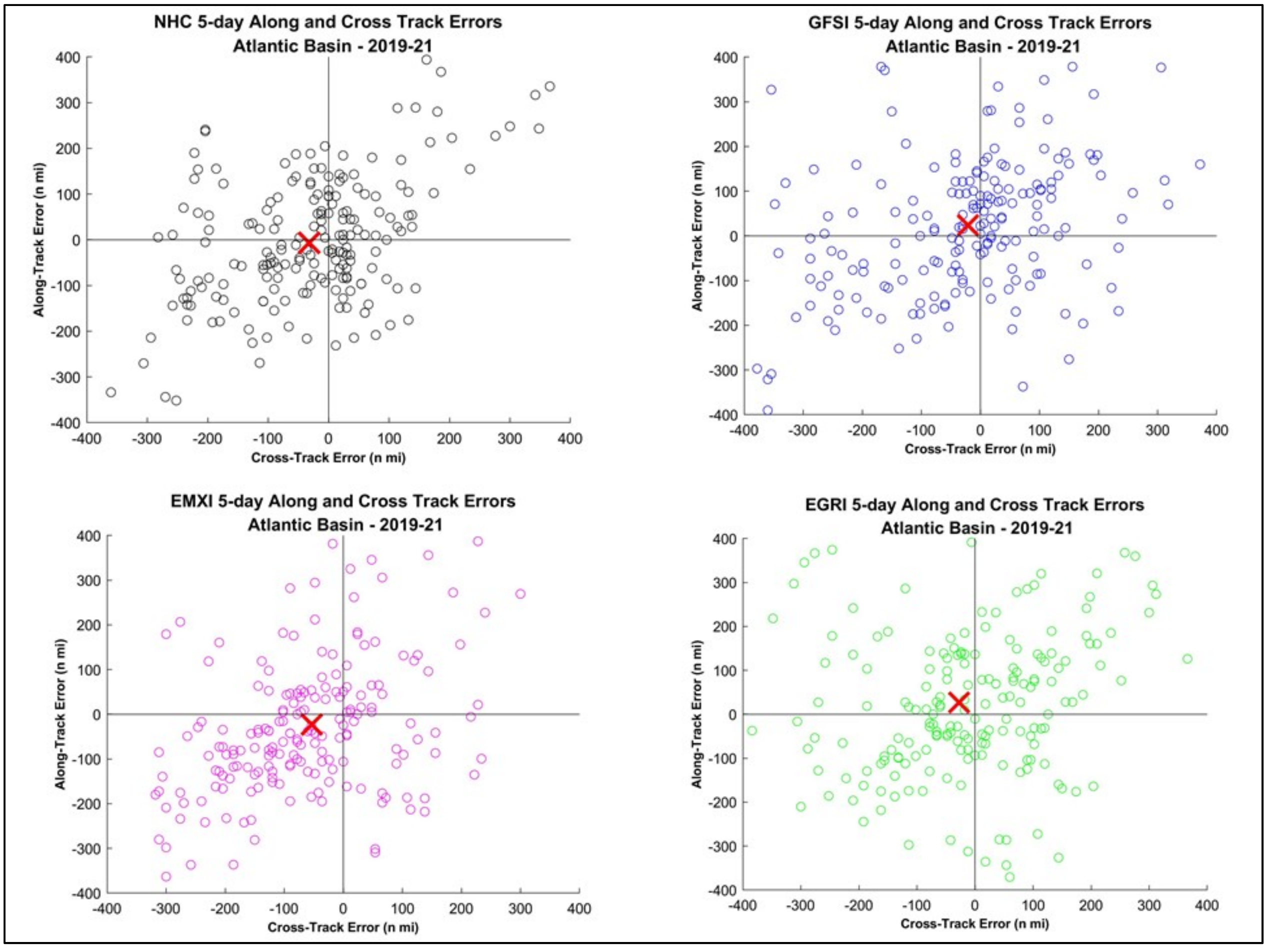


Figure 7. Homogenous comparison of OFCL, GFSI, EMXI, EGRI model track biases (n mi) at verifying 120-h forecasts for the Atlantic basin during the 2019-21 period. The red 'X' depicts the mean bias for each model.

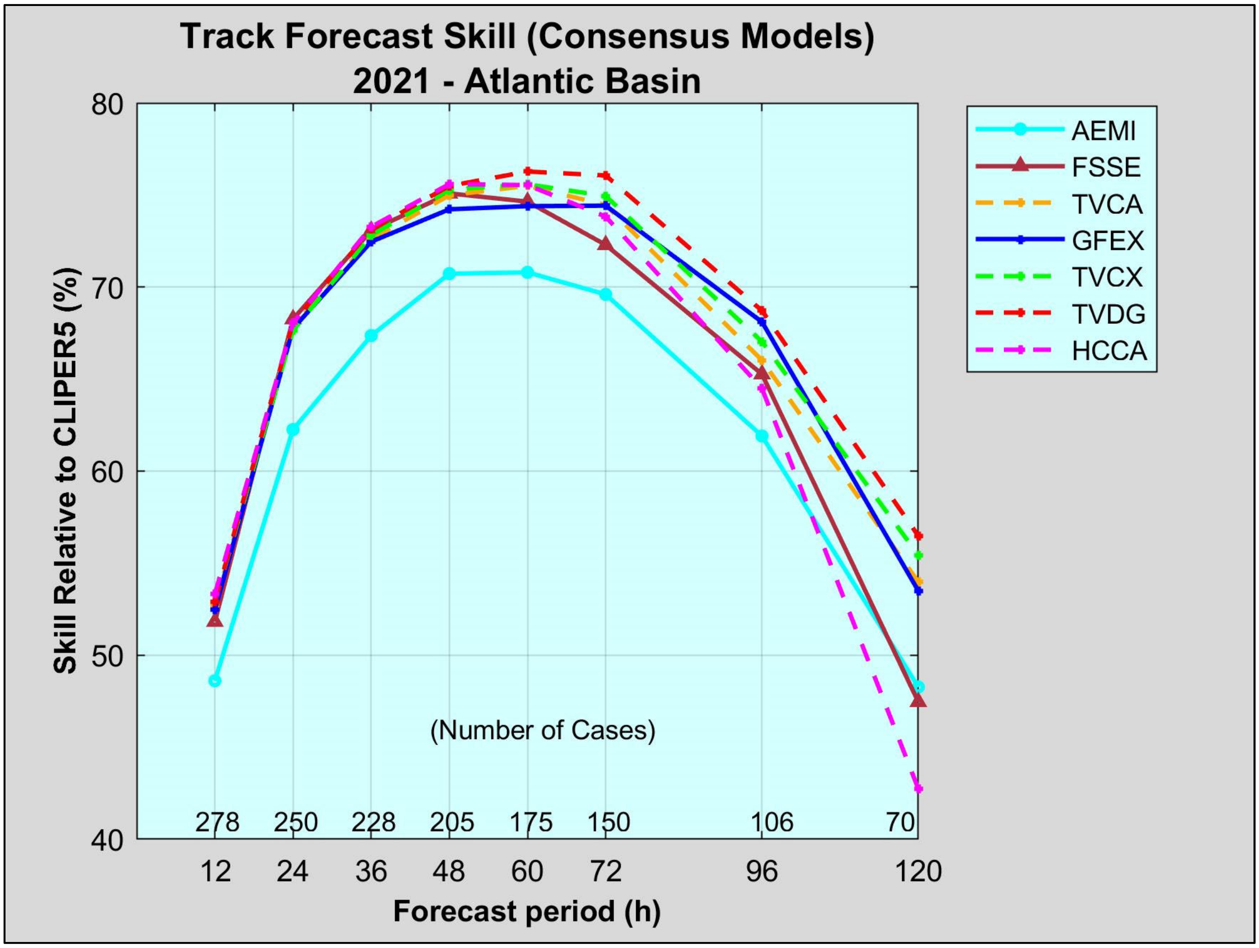


Figure 8. Homogenous comparison of the primary Atlantic basin track consensus models for 2021.

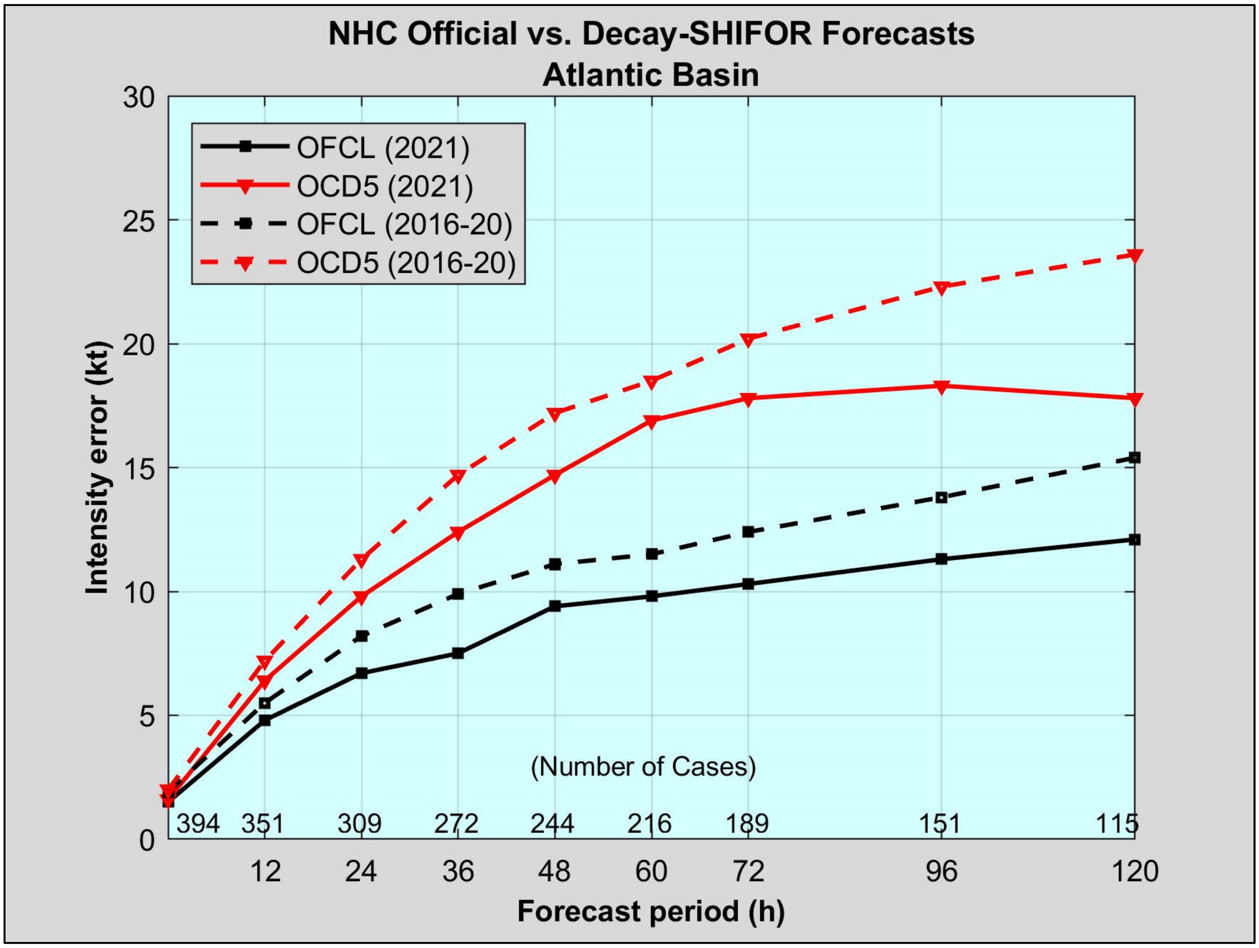
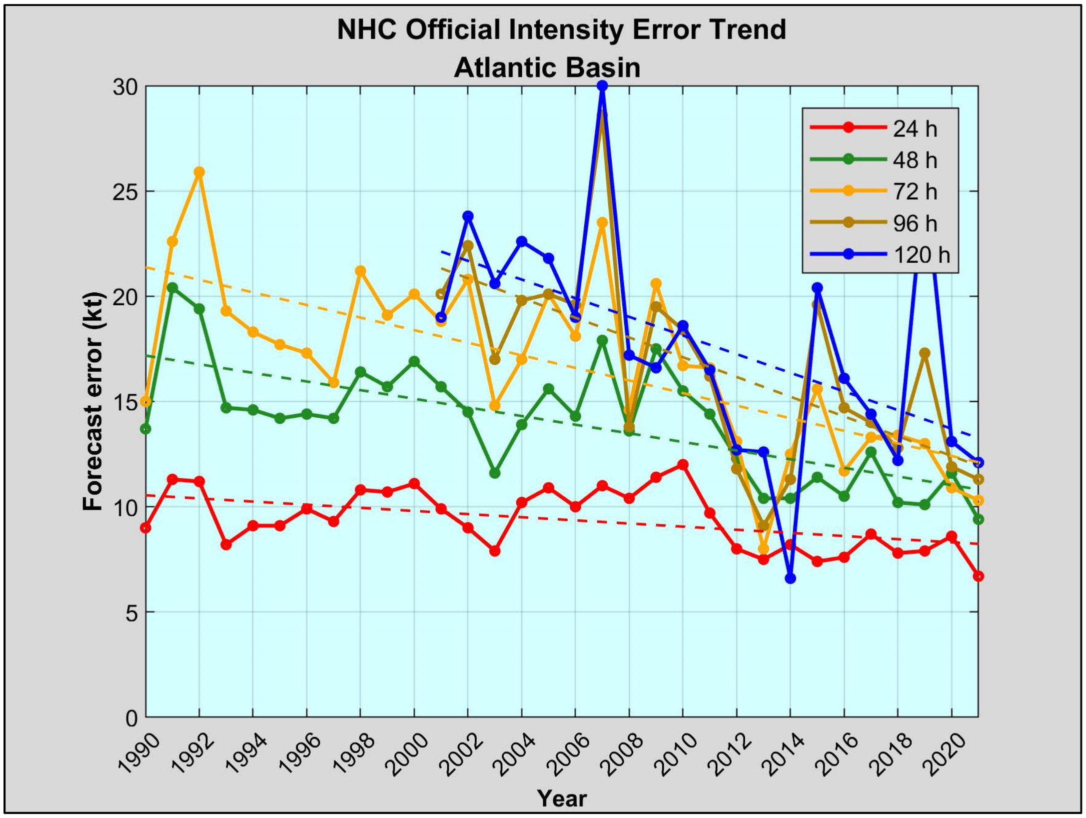


Figure 9. NHC official and Decay-SHIFOR5 (OCD5) Atlantic basin average intensity errors for 2021 (solid lines) and 2016-2020 (dashed lines).



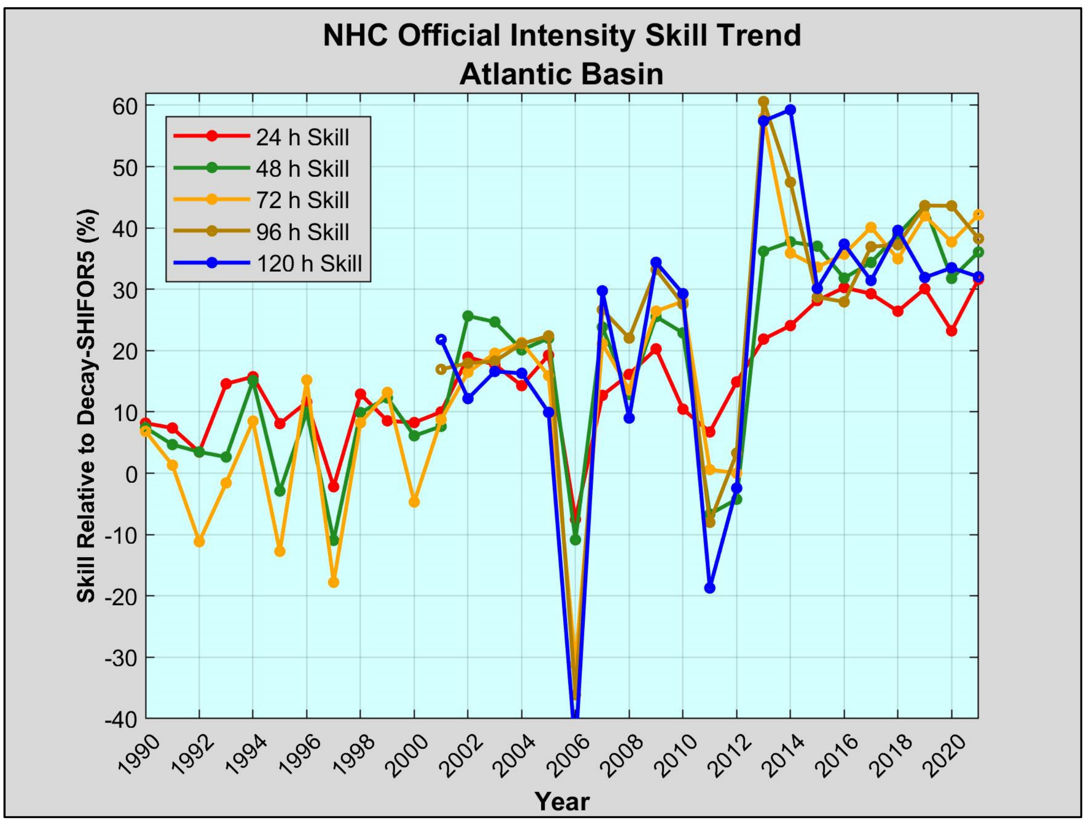


Figure 10. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the Atlantic basin.

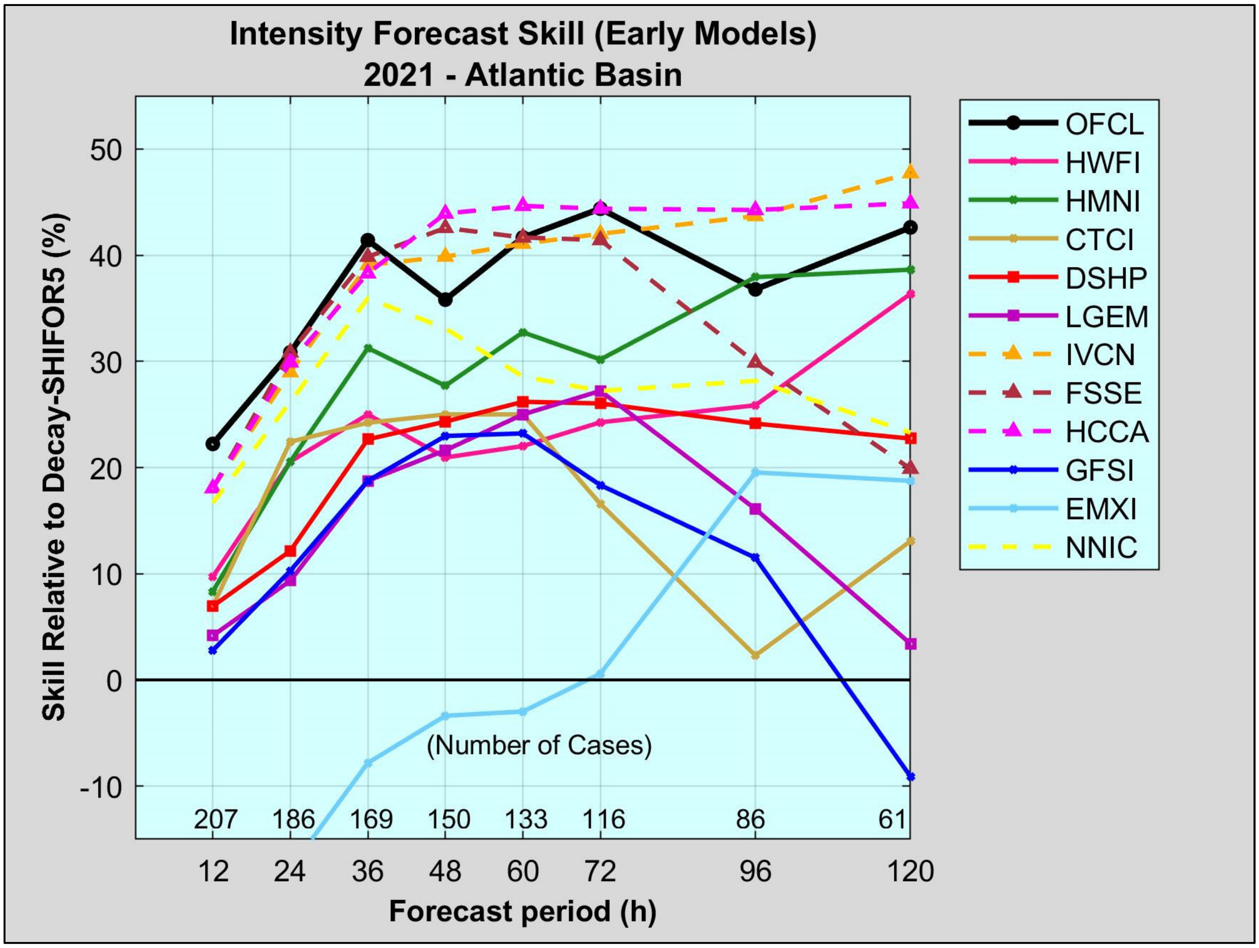


Figure 11. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2021.

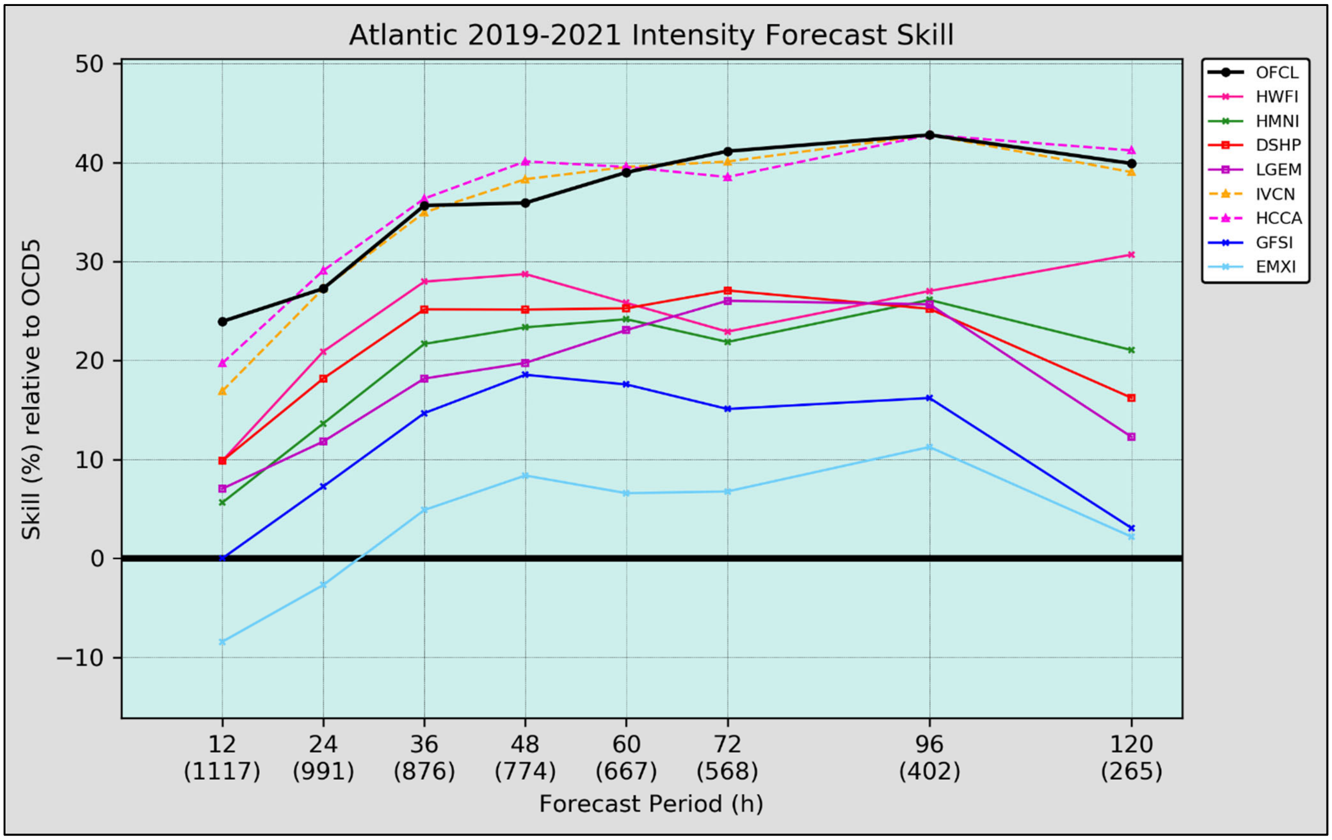


Figure 12. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2019-2021.

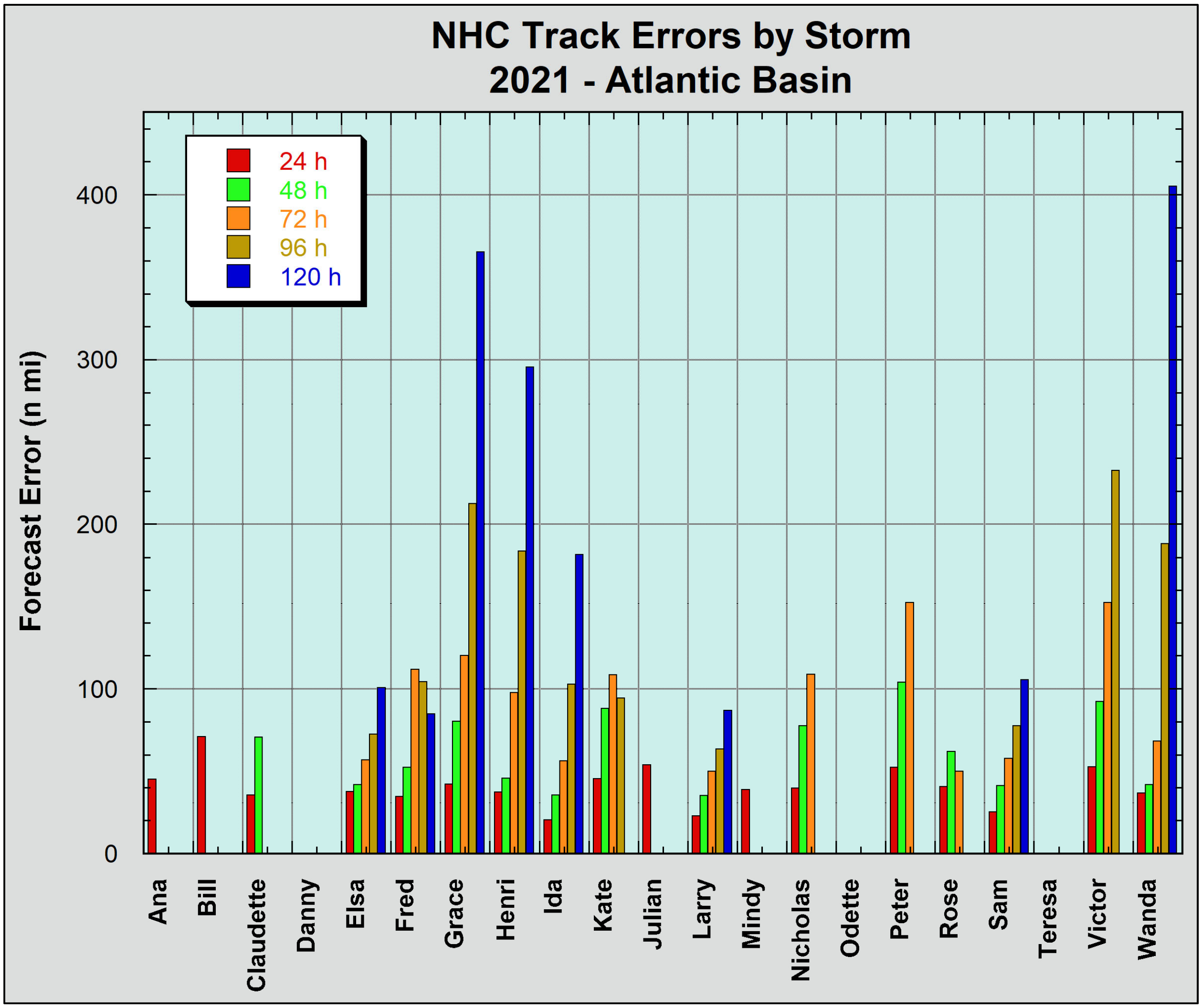


Figure 13. 2021 NHC official track errors by tropical cyclone at 24, 48, 72, 96 and 120 h.

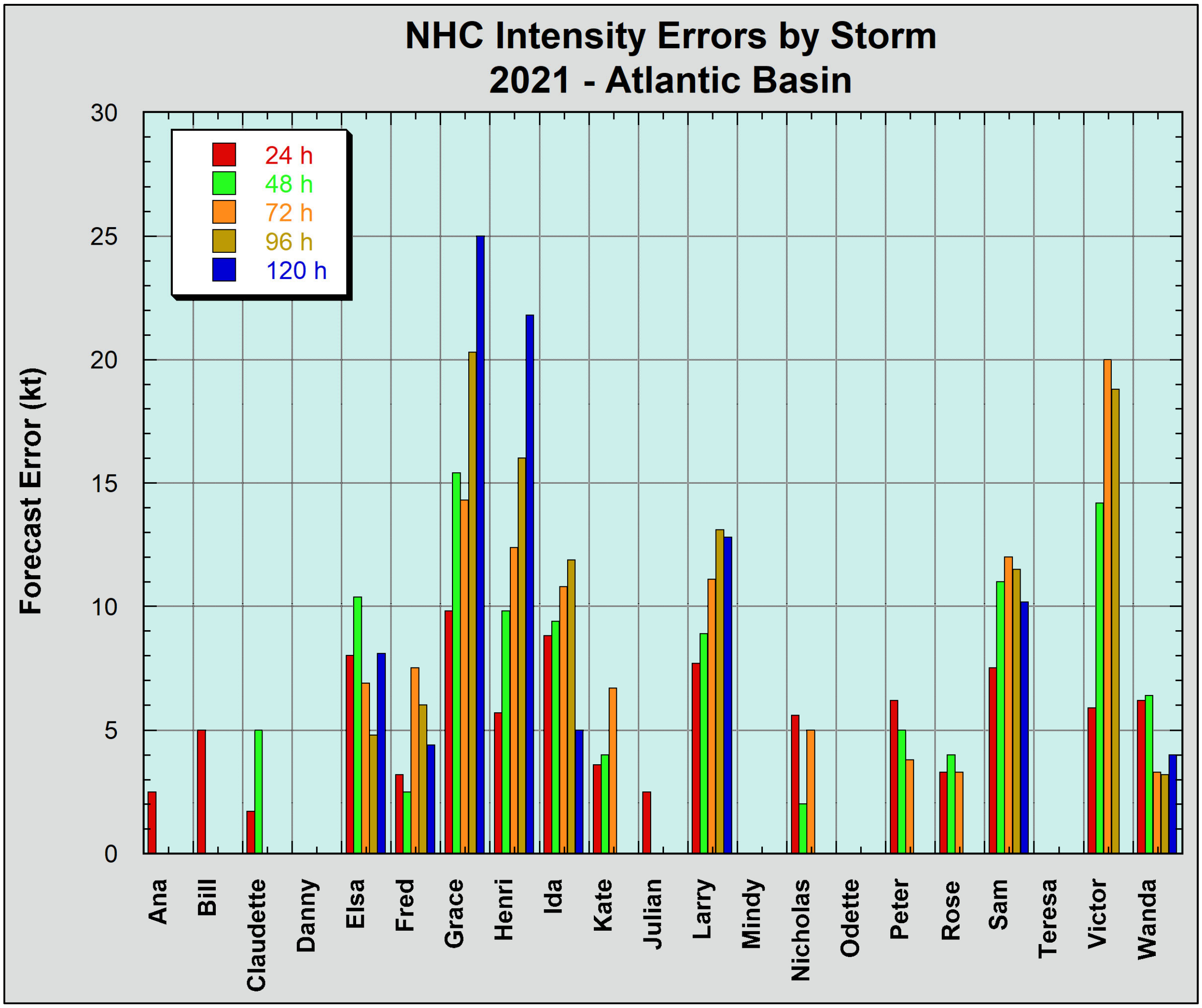


Figure 14. 2021 NHC official intensity errors by tropical cyclone at 24, 48, 72, 96 and 120 h.

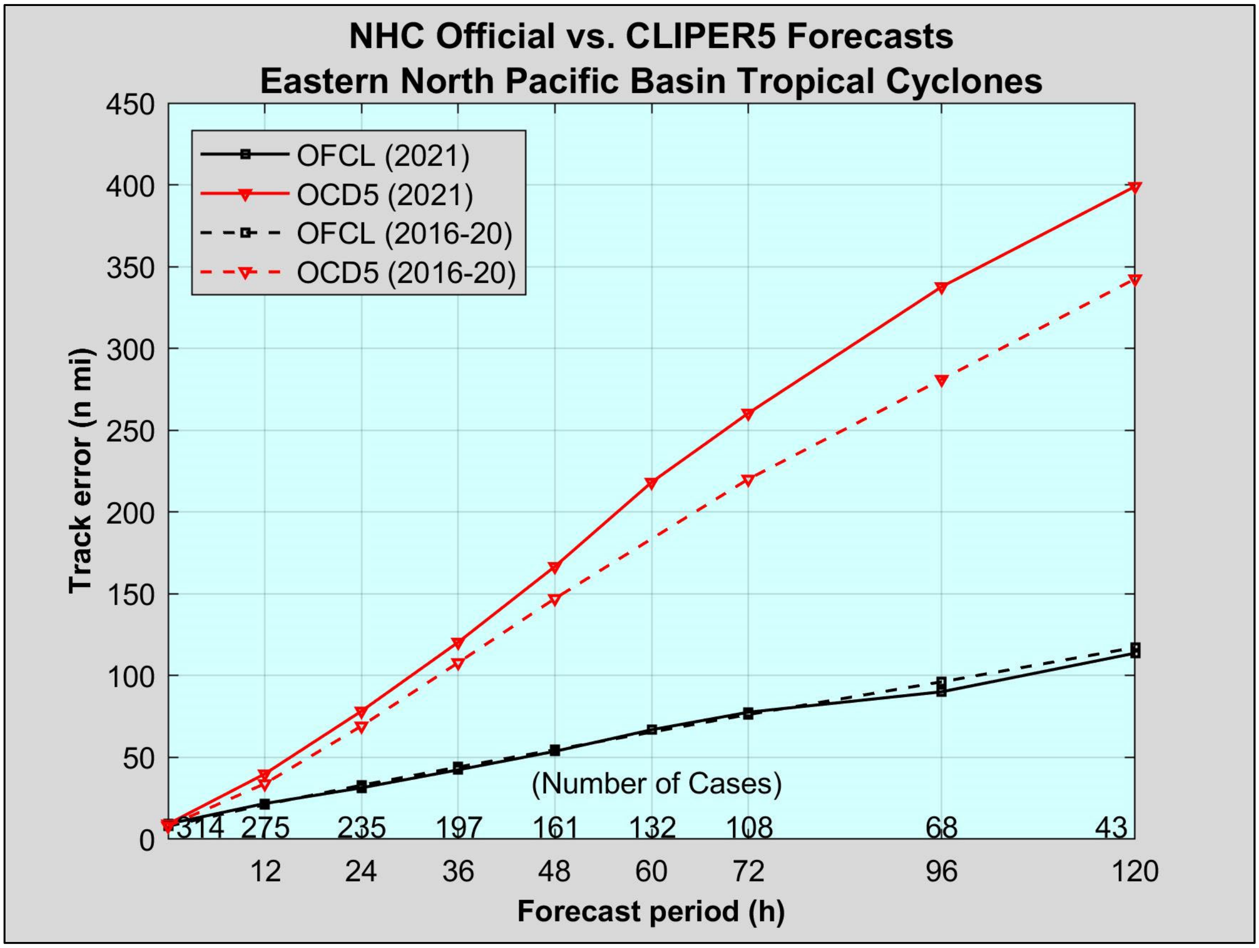


Figure 15. NHC official and CLIPER5 (OCD5) eastern North Pacific basin average track errors for 2021 (solid lines) and 2016-2020 (dashed lines).

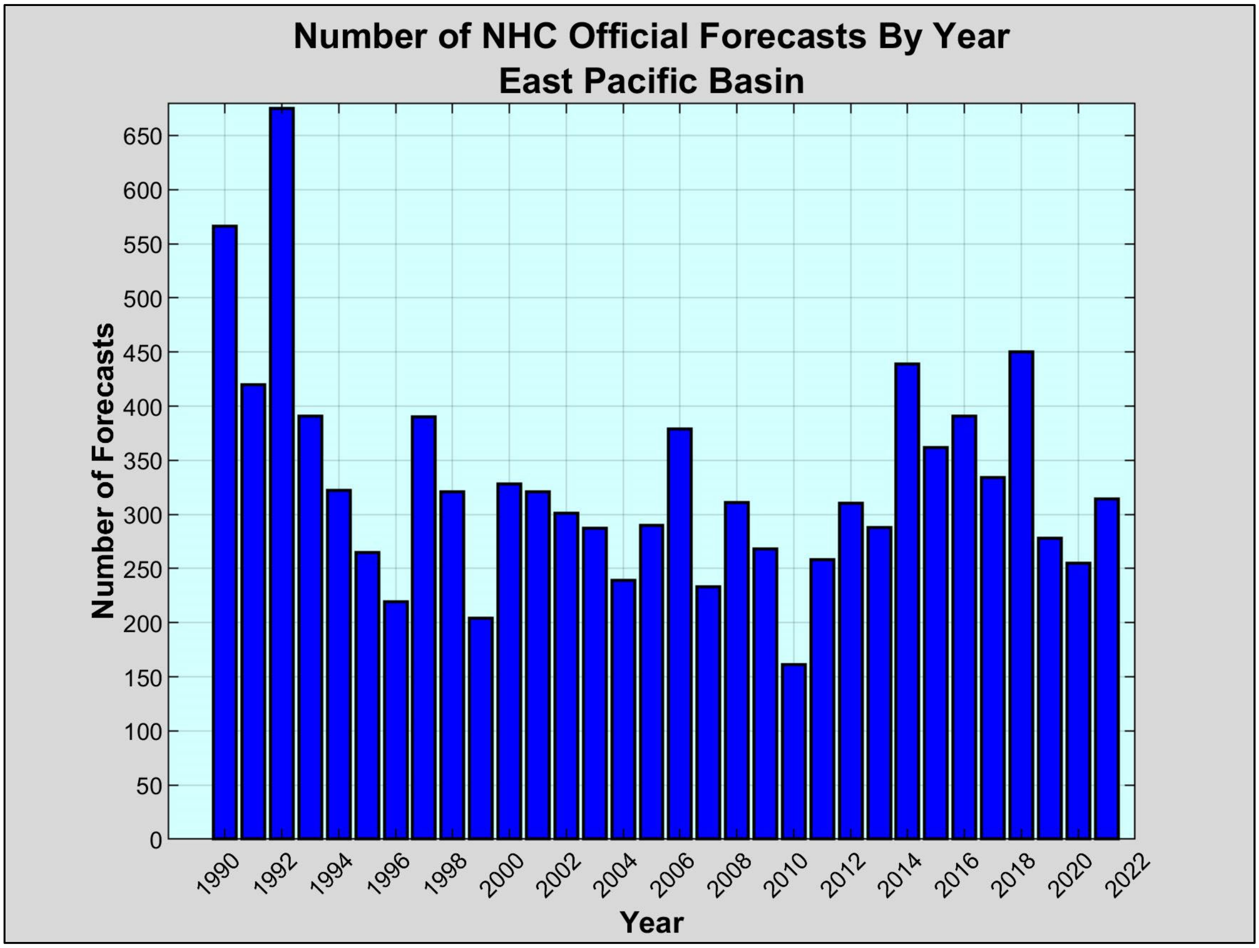


Figure 16. Number of NHC official forecasts for the eastern North Pacific basin stratified by year.

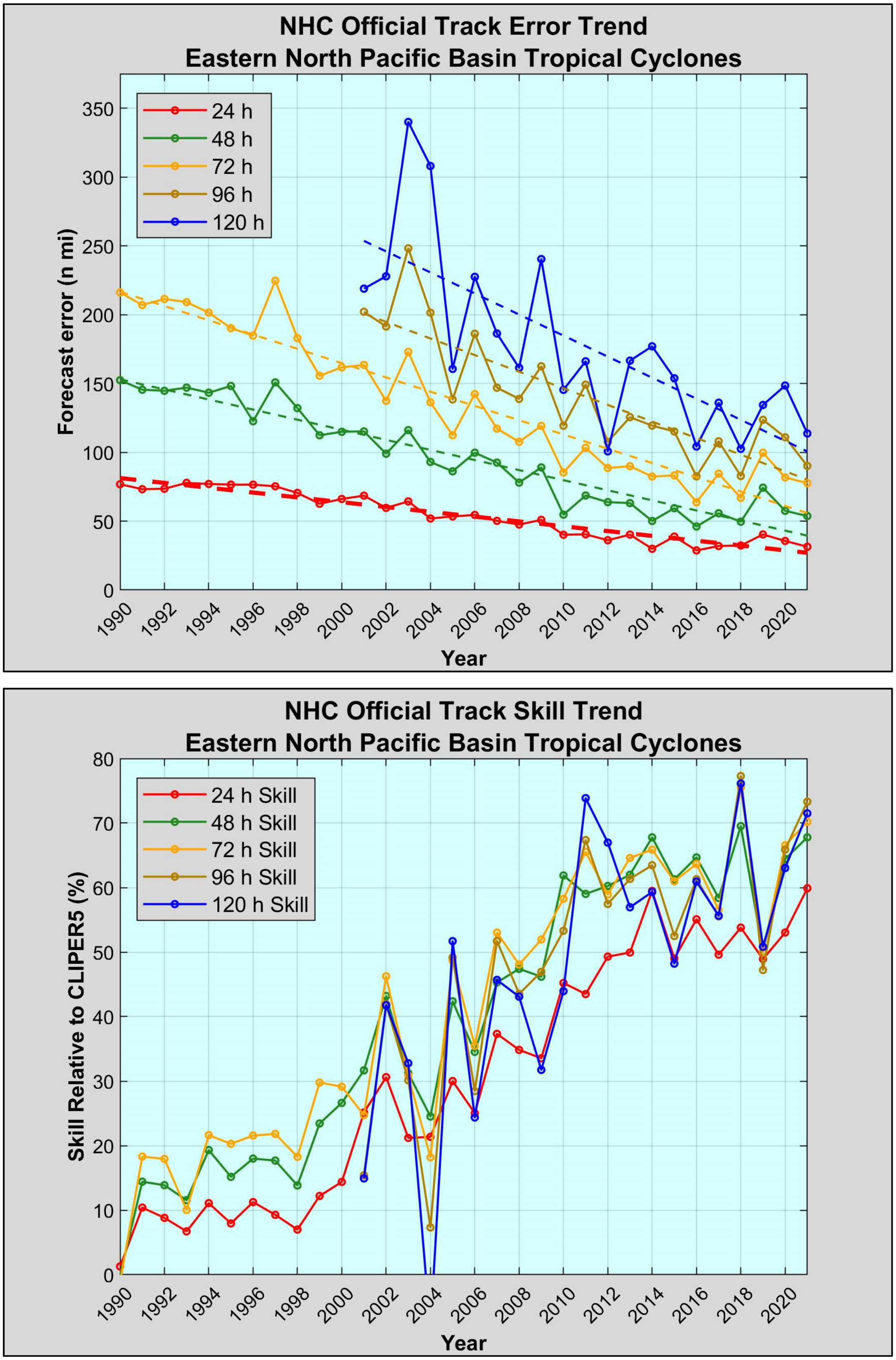


Figure 17. Recent trends in NHC official track forecast error (top) and skill (bottom) for the eastern North Pacific basin.

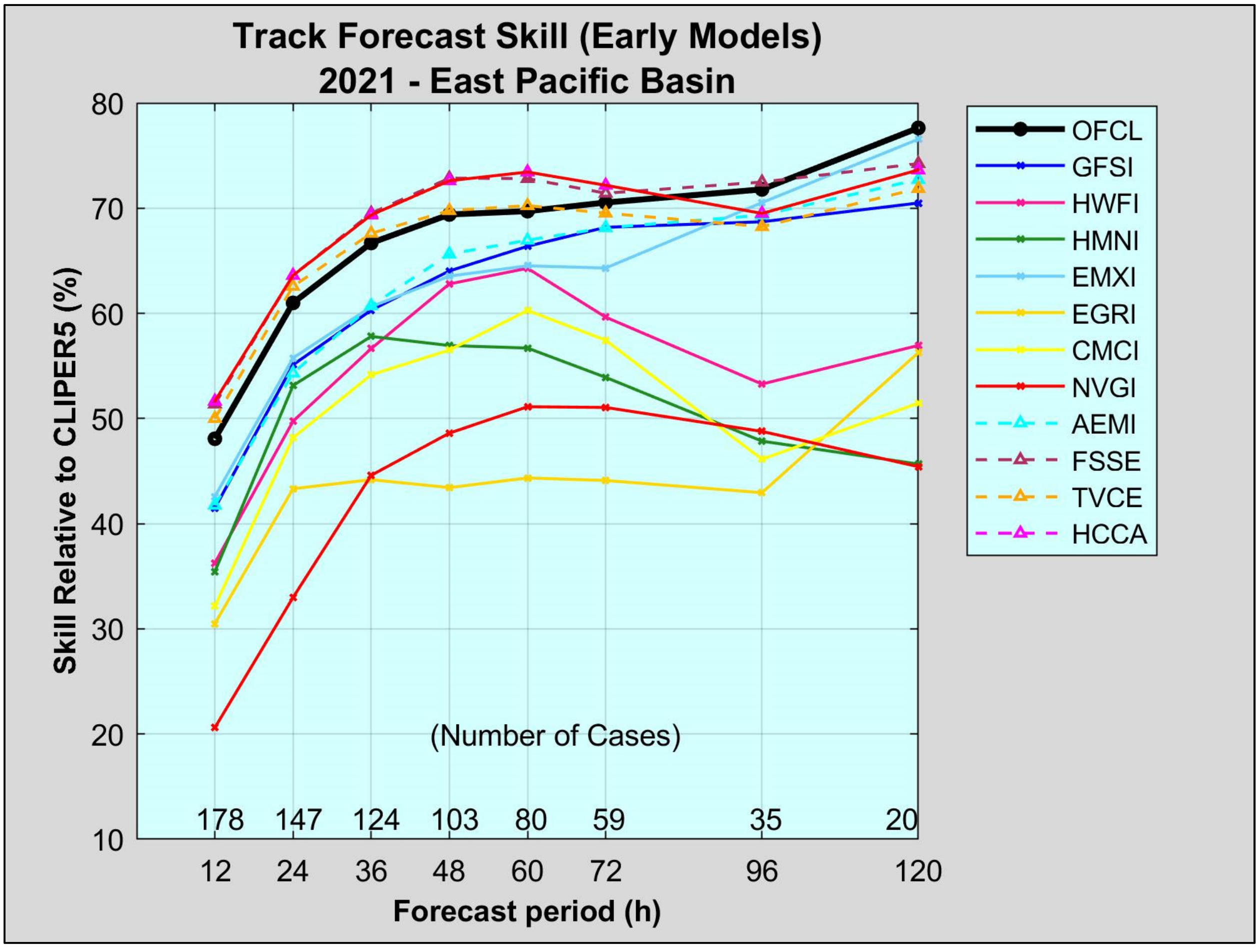


Figure 18. Homogenous comparison for selected eastern North Pacific early track models for 2021. This verification includes only those models that were available at least 2/3 of the time (see text).

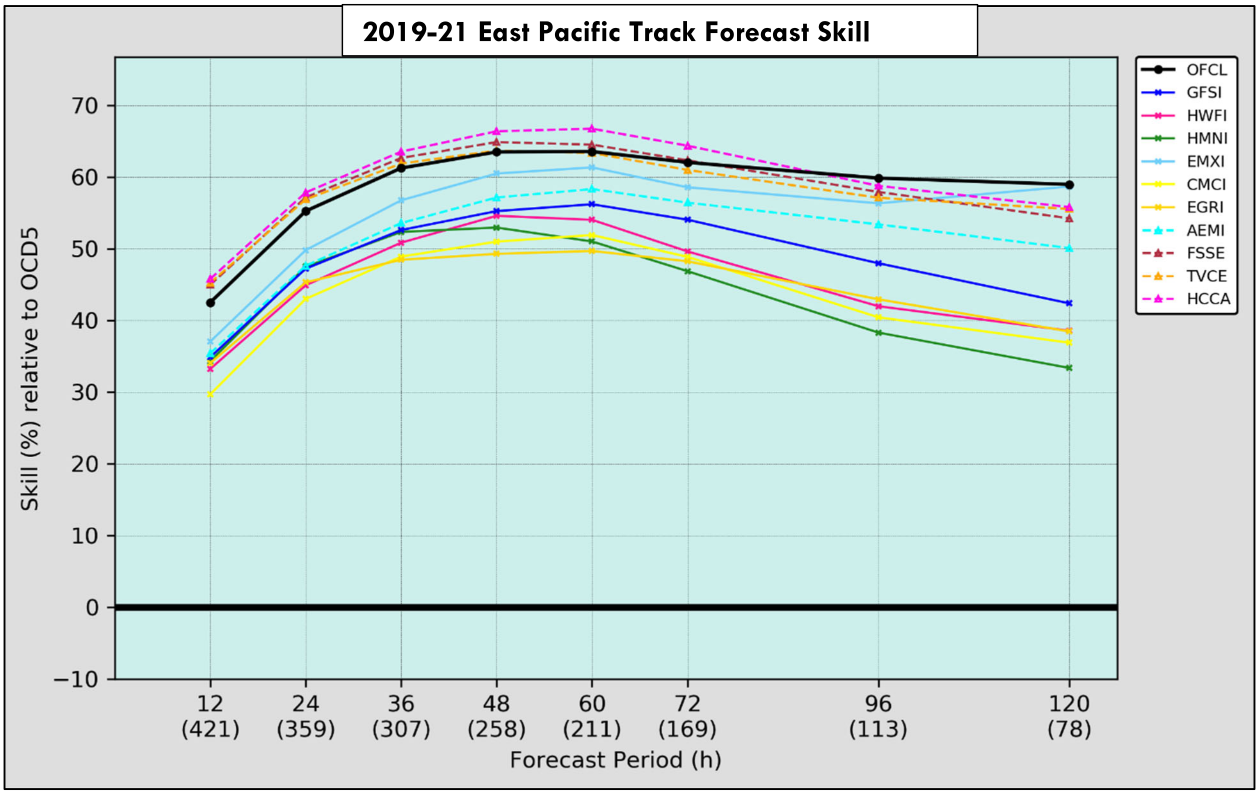


Figure 19. Homogenous comparison for selected eastern North Pacific basin early track models for 2019-2021.

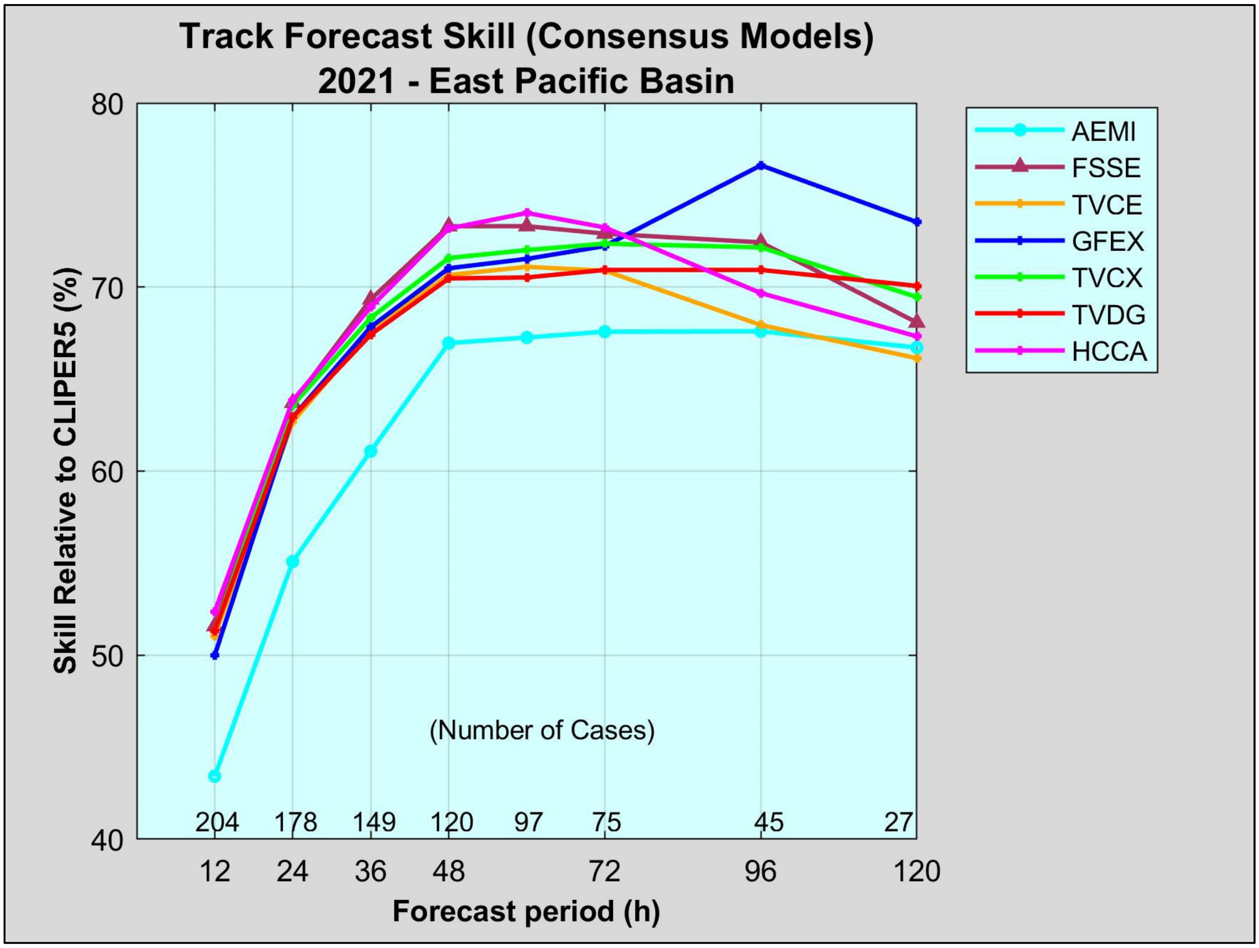


Figure 20. Homogenous comparison of the primary eastern North Pacific basin track consensus models for 2021.

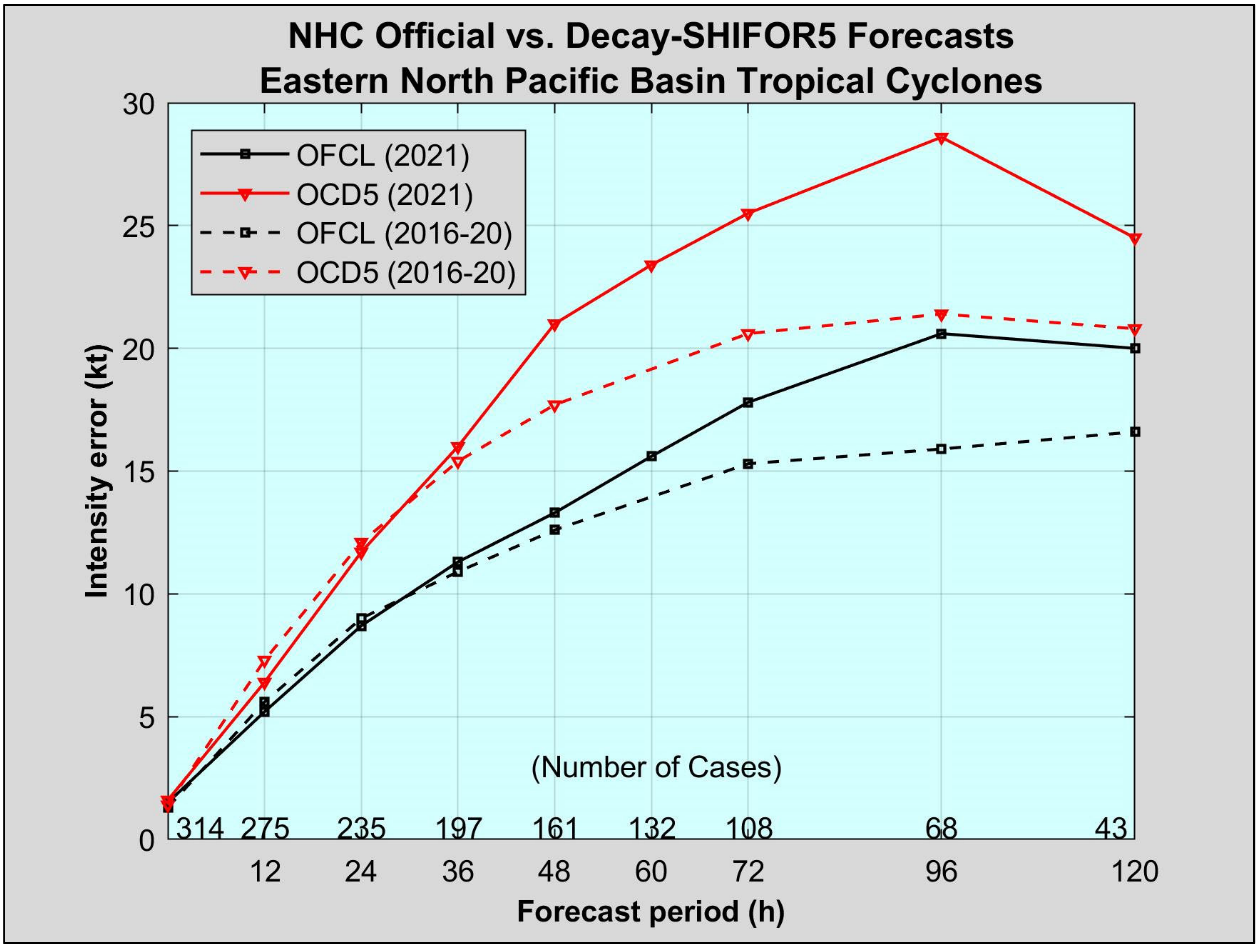


Figure 21. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2021 (solid lines) and 2016-2020 (dashed lines).

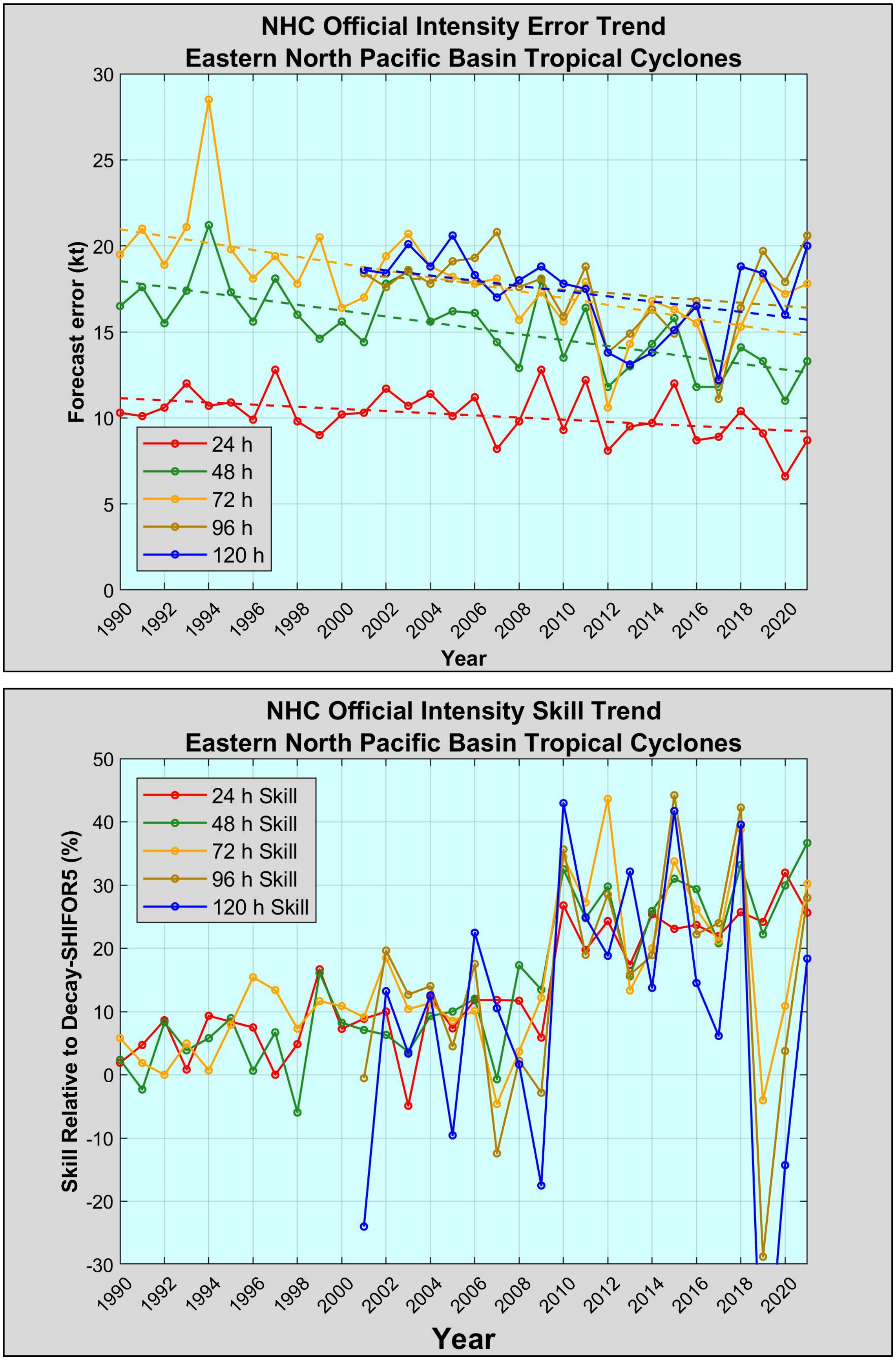


Figure 22. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

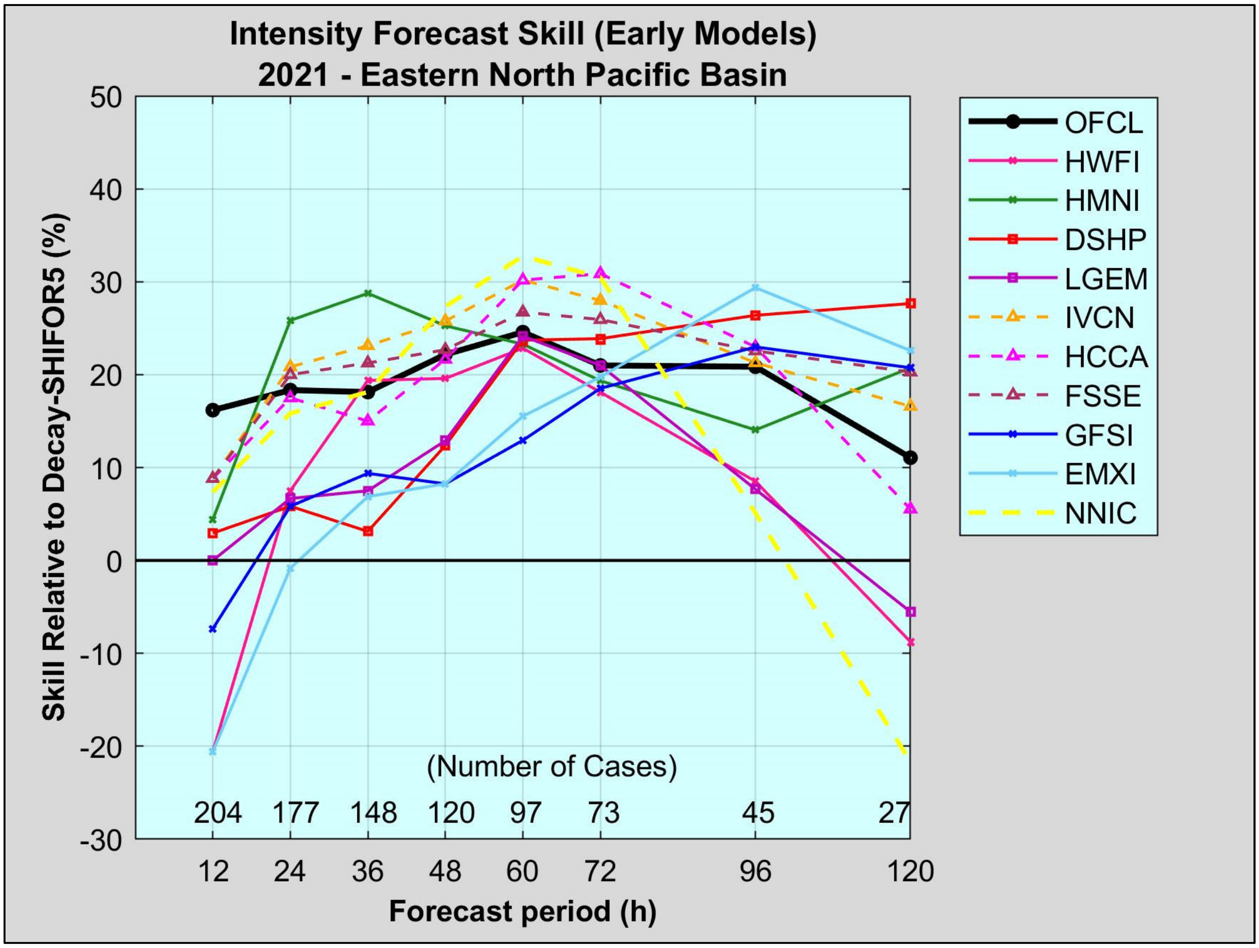


Figure 23. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2021.

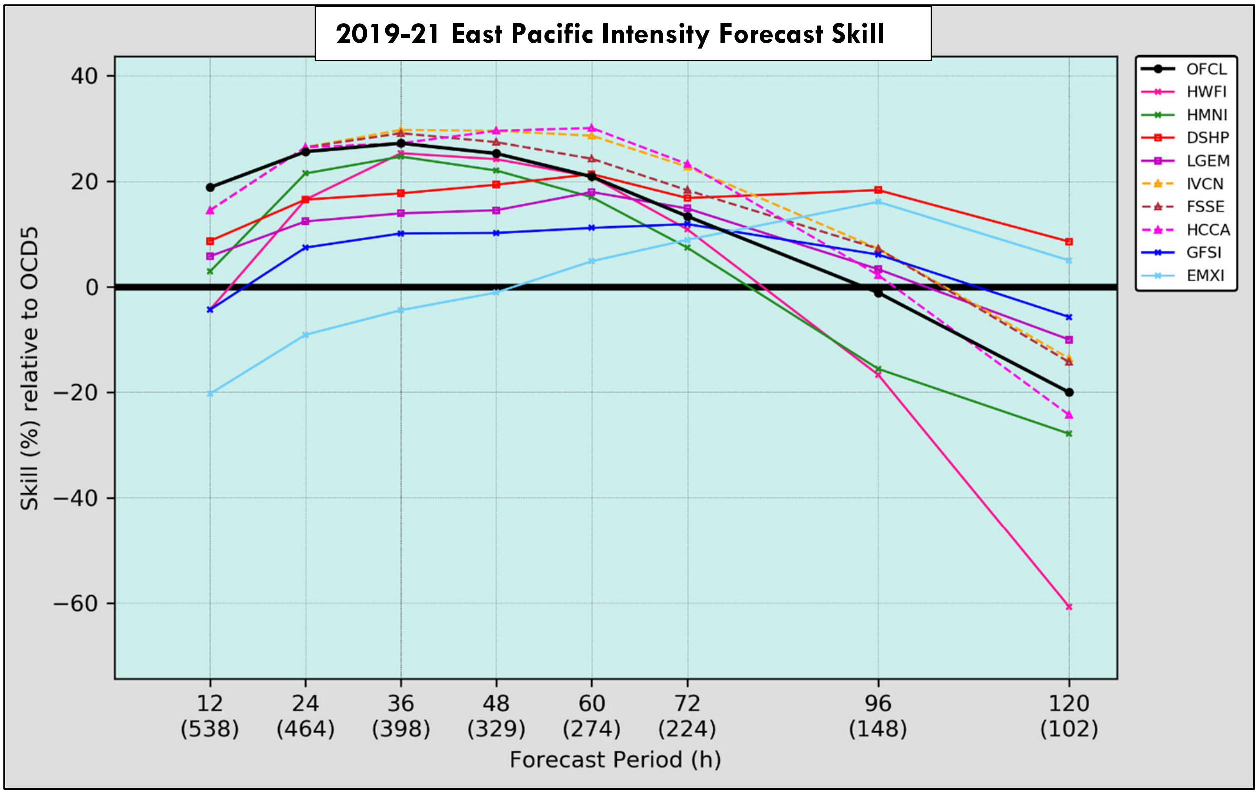


Figure 24. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2019-2021.

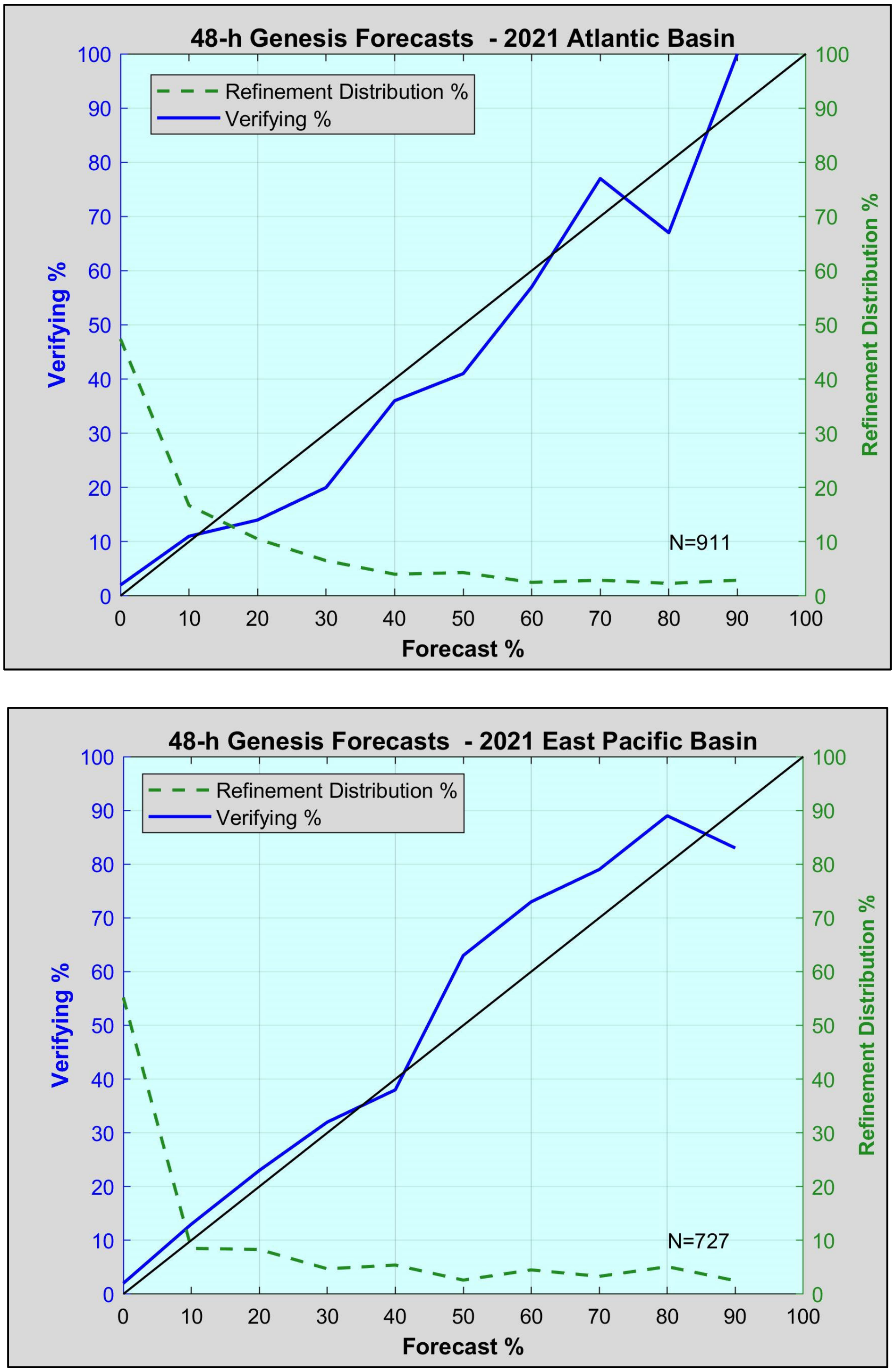
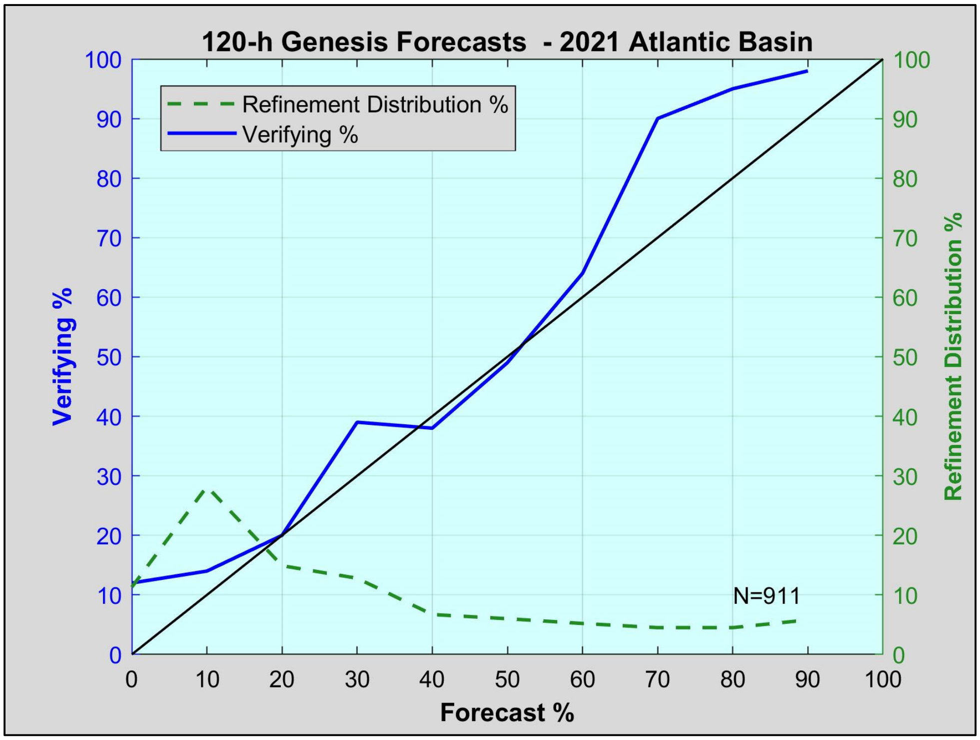


Figure 25. Reliability diagram for Atlantic (top) and eastern North Pacific (bottom) probabilistic tropical cyclogenesis 48-h forecasts for 2021. The solid lines indicate the relationship between the forecasts and verifying genesis percentages, with perfect reliability indicated by the thin diagonal black line. The dashed lines indicate how the forecasts were distributed among the possible forecast values.



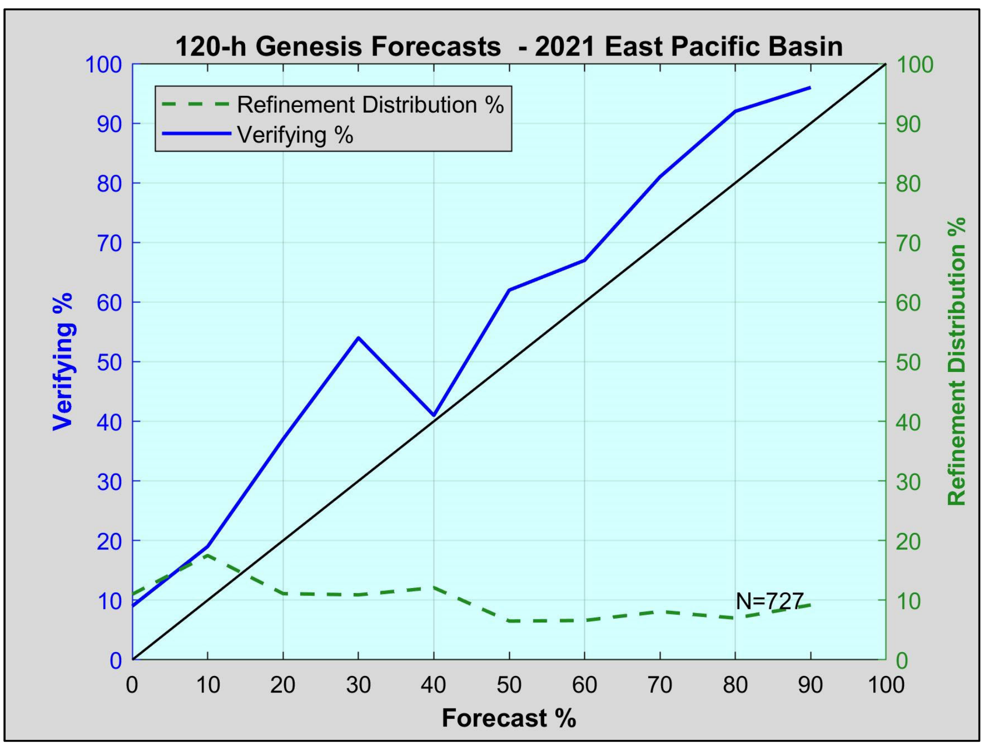


Figure 26. As described for Fig. 25, except for 120-h forecasts.

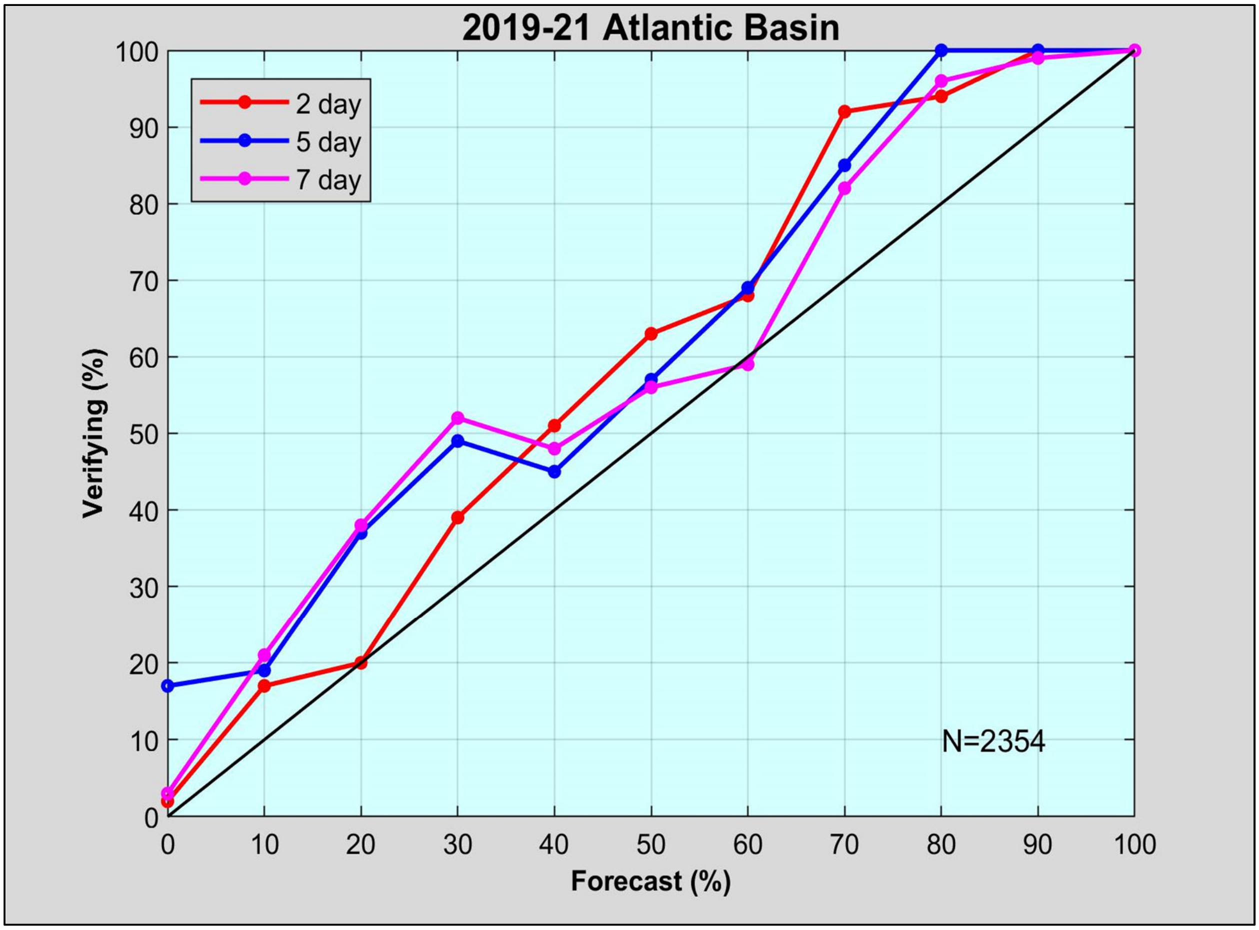


Figure 27. Reliability diagram for Atlantic probabilistic tropical cyclonegensis 2-, 5-, and experimental 7-day forecasts from 2019-21.