2014 National Hurricane Center Forecast Verification Report

John P. Cangialosi and James L. Franklin
NOAA/NWS/NCEP/National Hurricane Center

25 March 2015

ABSTRACT

The 2014 Atlantic hurricane season had below-normal activity, with 159 official forecasts issued. The mean NHC official track forecast errors in the Atlantic basin were slightly smaller than the previous 5-yr means from 12 to 72 h, but higher than the 5-yr mean at 120 h. A record for accuracy was set from 24 to 72 h. The official track forecasts were highly skillful and performed close to the best-performing models. TVCA and HWFI had the highest skill and beat the official forecast at a few time periods. FSSE and EMXI were strong performers in the short term and EGRI was the best model at the longer range. GFSI and AEMI were fair performers while GHMI, CMCI, NVGI, and GFNI 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 2014 were below the 5-yr means at all lead times. Decay-SHIFOR errors in 2014 were also lower than their 5-yr means from 12 to 36 h and at 120 h, but larger than average at the other forecast times. The official forecasts beat all of the intensity guidance in 2014. Among the guidance, IVCN and DSHP were the best performers, followed by FSSE and LGEM. GHMI and HWFI were not skillful through 36 h, but were competitive with the statistical and consensus aids at the longer ranges. The GPRA intensity goal was met.

There were 439 official forecasts issued in the eastern North Pacific basin in 2014, although only 136 of these verified at 120 h. This level of forecast activity was considerably above average. Records for track accuracy were set from 12 to 72 h in 2014. The official forecast was very skillful, but it was outperformed by FSSE and TVCE at all times. EMXI was the best individual model, which also beat the official forecast at 96 and 120 h, but it had a little less skill than TVCE and FSSE. GFSI, AEMI, and HWFI performed fairly well and made up the second tier of models.

For intensity, the official forecast errors in the eastern North Pacific basin were slightly lower than the 5-yr means at most times. Decay-SHIFOR errors in 2014 were similar to their 5-yr means at most forecast times, indicating the season's storms were about of average difficulty to predict. The official forecast was as good as or better than all of the guidance early, but it was outperformed by the consensus aid IVCN at 72 h and beyond. HWFI was good performer for all leads, while GHMI showed increased skill at the longer forecast times. DSHP and LGEM were skillful as well, except at 120 h. All of the guidance and the official forecast had a low bias in 2014.

Quantitative probabilistic forecasts of tropical cyclogenesis were publicly expanded from 48 to 120 h in August of 2013. Forecasts were expressed in 10%

increments and in terms of categories ("low", "medium", or "high"). In the Atlantic basin, results from 2014 indicate that the 48-h probabilistic forecasts were quite reliable at the low probabilities, but had a low (under-forecast) bias at medium probabilities. The 120-h probabilities also had a low bias at probabilities higher than 40%. A slight under-forecast bias was present at most ranges in the 48- and 120-h probabilistic forecasts in the eastern North Pacific basin.

The Hurricane Forecast Improvement Project (HFIP) and the National Hurricane Center agreed in 2009 to establish a pathway to operations known as "Stream 1.5". The performance of the Stream 1.5 models in 2014 was mixed. The GFDL ensemble mean, GPMI, performed better than its deterministic parent model for both track and intensity, but the intensity model SPC3 was not quite as skillful as DSHP. CXTI was a poor intensity model in 2014, and the performance of the HFIP consensus aids was about equal to the standard consensus models.

Table of Contents

1. Introduction

For all operationally designated tropical or subtropical cyclones 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, 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

---

[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.

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.

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

[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.

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

[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.

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 baselines for skill, in the context of examining forecast performance over the course of a season (or longer). However, they are 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 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 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 (Landsea and Franklin 2013). 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.

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 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]

---

[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".

[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.

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 FSU "super-ensemble", for example, combines its individual components on the basis of past performance and attempts to correct for biases in those components (Williford et al. 2003). 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 based on forecast and best track data sets taken from the Automated Tropical Cyclone Forecast (ATCF) System[10] on 19 February 2015 for the Atlantic basin, and on 6 March 2015 for the eastern North Pacific

[10] In ATCF lingo, these are known as the "a decks" and "b decks", respectively.

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 discusses the Hurricane Forecast Improvement Project (HFIP) Stream 1.5 activities in 2014. Section 6 summarizes the key findings of the 2014 verification and previews anticipated changes for 2015.

2. Atlantic Basin

a. 2014 season overview – Track

Figure 1 and Table 2 present the results of the NHC official track forecast verification for the 2014 season, along with results averaged for the previous 5-yr period, 2009-2013. In 2014, the NHC issued 159 Atlantic basin tropical cyclone forecasts[11], a number well below the average over the previous 5 yr (314). In fact, 2014 had the lowest amount of forecasts since 2009 and the fourth lowest since 1990 (Fig. 2). Mean track errors ranged from 26 n mi at 12 h to 270 n mi at 120 h. It is seen that the mean official track forecast errors in 2014 were slightly smaller than the previous 5-yr mean from 12 to 72 h, and higher than the 5-yr mean at 120 h. The season's storms were a little more difficult to forecast than average from 24 to 120 h. The official track forecast errors set a record for accuracy from 24 to 72 h. The official track forecast vector biases were small through 72 h, but increased beyond that time and were south-southwestward at 96 and 120 h (i.e., the official forecast tended to fall to the south-southwest of the verifying position). Further examination of the track errors shows that there was a slow along-track bias, but little cross-track bias in 2014. Track forecast skill ranged from 50% at 12

[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.

h to 73% at 48 and 72 h (Table 2). The track errors in 2014 were improved over the 2013 values (except at 120 h), and over the past 15-20 yr, 24−72-h track forecast errors have been reduced by about 70% (Fig. 2). Track forecast error reductions of about 50% have occurred over the past 10 yr for the 96- and 120-h forecast periods. On average, the NHC track errors steadily decrease as the initial intensity of a cyclone increases (Fig. 4).

Note that the mean official error in Fig. 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 Fig. 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 2014. 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 Fig. 5. Note that the sample becomes

[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.

exceedingly small (less than 50 cases) at 72 h and beyond, and in fact at no time period is the sample large enough to be particularly meaningful. That said, the figure shows that the official forecast was highly skillful, and near the accuracy of the best models. The consensus model TVCA and HWFI were the most skillful and beat the official forecast at a few time periods. FSSE and EMXI were strong performers in the short term, but trailed off beyond 48 h. Conversely, the skill of EGRI increased with forecast time and it was one of the best models at 96 and 120 h. GFSI and AEMI were not quite as good and made up the second tier of models. GHMI, CMCI, NVGI, and GFNI were the poorest-preforming dynamical models. An evaluation over the three years 2012-14 (Fig. 6) indicates that FSSE and TVCA are the best-performing models, with FSSE having a little more skill than the official forecasts (although it is worth noting that one member of FSSE is OFCI). GFSI and AEMI were strong performers for this larger sample and were the next best models behind the consensus aids. EMXI had skill values close to GFSI at 12-24 h and at 96-120 h, but lagged a little at the other forecast times. HWFI, GHMI, CMCI, and EGRI performed less well.

Vector biases of the guidance models for 2014 are given in Table 3b. The table shows that the official forecast had similar biases to TVCA, but the biases were generally smaller than most of the model guidance. Among the typically high-performing models, EMXI had little bias through 72 h, but it did exhibit a large south-southwestward bias at 96 and 120 h, albeit for a very small sample. GFSI and AEMI had a north to northeast bias that increased with forecast time. HWFI had a relatively small north bias from 12 to 72 h, and a more notable southwestward bias at 96 and 120 h.

A separate homogeneous verification of the primary consensus models for 2014 is shown in Fig. 7. The figure shows that FSSE, TVCA, and GFEX (consensus of GFSI and EMXI) were equally skillful through 48 h, with TVCA being the most skillful after that time. AEMI had less skill than the other consensus aids, except at 120 h where it was the best model for a very small sample. An examination of AEMI over the past few years (not shown) indicates that the ensemble mean has become increasingly skillful in the Atlantic basin, and it is even more skillful than the deterministic run from 72 to 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 2014, the GPRA goal was 77 n mi and the verification for this measure was 64.8 n mi.

b. 2014 season overview – Intensity

Figure 8 and Table 4 present the results of the NHC official intensity forecast verification for the 2014 season, along with results averaged for the preceding 5-yr period. Mean forecast errors in 2014 ranged from about 5 kt at 12 h to about 13 kt at 72 h. These errors were below the 5-yr means at all forecast times, especially at 120 h where the 2014 errors were nearly 60% smaller than the 5-yr average. The official forecasts had a low bias at all forecast times, and explained about 40-50% of the total error from 48 to 96 h. Decay-SHIFOR5 errors were below their 5-yr means from 12 to 36 h and at 120 h, but larger than average at the other forecast times. Figure 9 shows that there has been a notable decrease in the intensity errors over the past few years; however, these recent improvements are likely in part due to a lack of rapidly intensifying hurricanes, which are typically the source of the largest forecast errors. Over the long term there has been only a slight decrease in error, although forecasts during the current decade, on average, have been more skillful than those from the previous one. In 2014, skill was at an all-time high at the shorter leads and at 120 h, however, the small sample (especially at 120 h) and small number of rapid deepening storms make the high skill most likely non-representative.

Table 5a presents a homogeneous verification for the official forecast and the primary early intensity models for 2014. Intensity biases are given in Table 5b, and forecast skill is presented in Fig. 10. The official forecasts showed increasing skill throughout the forecast period and beat all of the guidance at all forecast times in 2014, which is quite an unusual occurrence. Many of the models also showed skill in 2014. The best model performers were DSHP and the consensus aid IVCN, followed by LGEM and FSSE. GHMI and HWFI had no skill through 36 h, but were competitive with the statistical and consensus aids beyond that time. GFNI was not as good as GHMI and HWFI beyond 48 h. The top-performing global models, GFSI and EMXI, were included in the intensity verification for completeness, although they are typically not considered by forecasters, and unsurprisingly, GFSI and especially EMXI were not competitive with the better performing models. An inspection of the intensity biases (Table 5b) indicates that nearly all of the guidance suffered from a low bias in 2014, with FSSE having a substantial low bias of about 14 kt at 72 h. DSHP had the smallest biases through 96 h, and was the only model that consistently had less bias than the official forecasts. An evaluation over the three years 2012-14 (Fig. 11) indicates that the consensus models have been superior to all of the individual models throughout the forecast period.

However, a separate verification including only the pre-landfall cases reveals that DSHP and LGEM are as skillful as the consensus models at the longer forecast times when land interactions are not involved.

The 48-h official intensity error, evaluated for all tropical cyclones, is another GPRA measure for the NHC. In 2014, the GPRA goal was 14 kt and the verification for this measure was 10.4 kt. This was the fourth time in seven years that the intensity goal was met.

c. Verifications for individual storms

Forecast verifications for individual storms are given in Table 6. Of note are the large track errors for Hurricane Cristobal, which were higher than the long-term mean at all forecast times. The primary reason for the higher than average errors for this storm was that the official forecasts did not anticipate the abrupt northeastward to east-northeastward acceleration of the cyclone following a period of very slow motion. Conversely, the official track forecast errors were quite low for Hurricane Edouard. The track errors for that storm where about 20-35% lower than the long-term mean at all forecast times.

For intensity prediction, Hurricane Fay was one of the more challenging cyclones in 2014. The official forecast errors were larger than the long-term mean at most forecast times. The primary reason for the higher than average errors for Fay was because the official forecasts underestimated the amount of strengthening that occurred, since strong vertical wind shear was expected to hamper significant intensification. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2014atlan.shtml.

3. Eastern North Pacific Basin

a. 2014 season overview – Track

The NHC track forecast verification for the 2014 season in the eastern North Pacific, along with results averaged for the previous 5-yr period is presented in Figure 12 and Table 7. There were 439 forecasts issued for the eastern Pacific basin in 2014, although only 136 of these verified at 120 h. This level of activity is considerably above average. Mean track errors ranged from 20 n mi at 12 h to 177 n mi at 120 h, and were lower than the 5-yr means at all times except for 120 h. Records for forecast accuracy were set from 12 to 72 h in 2014. CLIPER5 errors were close to their long-term means, implying that the season's storms were about of average difficulty to predict. Little bias was present in the official track forecast at most lead times.

Figure 13 shows recent trends in track forecast accuracy and skill for the eastern North Pacific. Errors have been reduced by roughly 60% for the 24 to 72 h forecasts since 1990, a somewhat smaller but still substantial improvement relative to what has occurred in the Atlantic. Forecast skill in 2014 set a new record high at the 24-, 48- and 72-h forecast times.

Table 8a presents a homogeneous verification for the official forecast and the early track models for 2014, with vector biases of the guidance models given in Table 8b. Skill comparisons of selected models are shown in Fig. 14. FSSE and TVCE were the only two models that consistently outperformed the official forecast. EMXI was the best individual model, but it had less skill than TVCE/FSSE. EMXI also beat the official forecast at 96 and 120 h. GSFI, HWFI, and AEMI made up the second tier of models, with GHMI and EGRI not very far behind. CMCI was a poor performer and had similar skill to the simple BAMM model.

A separate verification of the primary consensus aids is given in Figure 15. TVCE and FSSE had comparable skill from 12 to 72 h and at 120 h, but TVCE bested FSSE slightly at the other forecast times. GFEX did not perform quite as well as TVCE/FSSE and AEMI was the least skillful of the consensus aids shown.

b. 2014 season overview – Intensity

Figure 16 and Table 9 present the results of the NHC eastern North Pacific intensity forecast verification for the 2014 season, along with results averaged for the preceding 5-yr period. Mean forecast errors were 6 kt at 12 h and increased to 17 kt at 72 h. The errors were lower than the 5-yr means at most forecast time periods and the intensity forecast biases were also low for all lead times. The Decay-SHIFOR5 forecast errors were similar to their 5-yr means, except at 120 h where the errors were smaller than the long-term mean. A review of error and skill trends (Fig. 17) indicates that the intensity errors have decreased slightly over the past 15-20 yr, especially at 48 h and beyond. Forecast skill appeared to increase in 2010 but has changed little during the last few years.

Figure 18 and Table 10a present a homogeneous verification for the primary early intensity models for 2014. Forecast biases are given in Table 10b. The official forecasts performed as well as or better than all of the guidance from 12 to 48 h, but were outperformed by IVCN at 72 h and beyond. HWFI was skillful at all forecast times, though not as much as the consensus aids, and was one of the best individual models. GHMI showed increasing skill over time and was the best individual model at 96 and 120 h. DSHP and LGEM were fair performers as well, but they were not skillful at 120 h. EMXI was not competitive with the standard intensity guidance, but GFSI had some skill at the longer leads. All of the models suffered from a low bias in 2014, with GHMI having the most significant bias (about 14 kt at 48 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/2014epac.shtml.

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. Since 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 TC 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 out to 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 2014 are given in Table 12 and illustrated in Fig. 19. In the Atlantic basin, a total of 559 genesis forecasts were made. These 48-h forecasts were quite reliable for the low probabilities but exhibited an under-forecast (low) bias in the medium probabilities. The sample is too small at the high probabilities to draw meaningful conclusions. In the eastern Pacific, a total of 716 genesis forecasts were made. The forecasts had a modest low bias at most probabilities, but overall were well calibrated.

Verification of the 120-h outlook for the Atlantic and eastern North Pacific basins for 2014 are given in Table 13 and illustrated in Fig. 20. In the Atlantic basin, the 120-h forecasts were well-calibrated from 0 to 40%, but a low bias existed at probabilities greater than 40%. In the eastern North Pacific, a slight low bias exists at all probabilities. The diagrams also show the refinement distribution, which indicates how often the forecasts deviated from (a perceived) climatology. Sharp peaks at climatology indicate low forecaster confidence, while maxima at the extremes indicate high confidence; the refinement distributions shown here suggest an intermediate level of forecaster confidence.

5. HFIP Stream 1.5 Activities

The Hurricane Forecast Improvement Project (HFIP) and the National Hurricane Center agreed in 2009 to establish a pathway to operations known as “Stream 1.5”. Stream 1.5 covers improved models and/or techniques that the NHC, based on prior assessments, wants to access in real-time during a particular hurricane season, but which cannot be made available to NHC by the operational modeling centers in conventional production mode. HFIP's Stream 1.5 supports activities that intend to bypass operational limitations by using non-operational resources to move forward the delivery of guidance to NHC by one or more hurricane seasons. Stream 1.5 projects are run as part of HFIP's annual summertime "Demo Project".

Four models/modeling systems were provided to NHC in 2014 under Stream 1.5; these are listed in Table 14. Note that most models were admitted into Stream 1.5 based on the models' performance forecasting either track or intensity, but generally not both. For example, forecasters were instructed to consult the CXTI intensity forecasts but not the CXTI track forecasts. Two HFIP Stream 1.5 consensus aids were constructed: the track consensus TV15 comprised the operational models GFSI, EGRI, HWFI, EMXI and the Stream 1.5 model GPMI, while the intensity consensus IV15 comprised the operational models DSHP, LGEM, HWFI and the Stream 1.5 models UW4I, CXTI, and GPMI.

Figure 21 presents a homogeneous verification of the primary operational models against the Stream 1.5 track model GPMI. The figure shows that in 2014 GPMI (GFDL ensemble mean) performed better than its deterministic model, but it had notably less skill than the best models. Figure 22 shows that there was very difference between the Stream 1.5 consensus TV15 and TVCA.

Figure 23 presents a direct comparison of the track and intensity forecast skill of GHMI and the Stream 1.5 model GPMI. Regarding track, GPMI had a little more skill than the deterministic model at all forecast times. For intensity, GPMI was slightly more skillful than GHMI at all forecast times except 36 and 96 h, where the skill values were similar.

Intensity results of the Stream 1.5 models are shown in Fig. 24. The Stream 1.5 model CXTI was a poor performer and was the least skillful model among all of the guidance shown. SPC3 and GPMI were fair performers and were near the middle of the pack. Like for the track consensus, there was little difference in skill between the Stream 1.5 model consensus IV15 and IVCN (Fig. 25).

Figure 26 presents the intensity forecast skill of DSHP, LGEM, and the Stream 1.5 model SPC3. In 2014, SPC3 had more skill than LGEM but less than DSHP in the short range. In the longer range, however, the skill of SPC3 decreased and was similar to or lower than LGEM.

6. Looking Ahead to 2015

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 2015 for the Atlantic and eastern North Pacific basins (based on error distributions for 2010-14) are given in Table 15. In the Atlantic basin, the cone circles will be slightly smaller (by up to 3 %) at a few forecast time periods. In the eastern Pacific basin, the cone circles will be considerably smaller (up to 13 %) than they were in 2014.

b. Consensus Models

In 2008, NHC changed the nomenclature for many of its consensus models. The new system defines a set of consensus model identifiers that remain fixed from year to year. The specific members of these consensus models, however, 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., TCOA), while others are less restrictive, requiring only two or more members to be present (e.g., TVCA). 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 2015 is unchanged from 2014, and is shown in Table 16.

Acknowledgments:

The authors gratefully acknowledge Michael Brennan, Monica Bozeman, and Craig Mattocks of NHC, managers of the NHC forecast databases.

7. References

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

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.

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.

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

Williford, C.E., T. N. Krishnamurti, R. C. Torres, S. Cocke, Z. Christidis, and T. S. V. Kumar, 2003: Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999. *Mon. Wea. Rev.*, **131**, 1878-1894.

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
GFDL	NWS/Geophysical Fluid Dynamics Laboratory model	Multi-layer regional dynamical	L	Trk, Int
HWRF	Hurricane Weather and Research Forecasting 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
NVGM	Navy Global Environmental Model	Multi-layer global dynamical	L	Trk, Int
GFDN	Navy version of GFDL	Multi-layer regional 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
AFW1	Air Force MM5	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
BAMS	Beta and advection model (shallow layer)	Single-layer trajectory	E	Trk
BAMM	Beta and advection model (medium layer)	Single-layer trajectory	E	Trk
BAMD	Beta and advection model (deep layer)	Single-layer trajectory	E	Trk
LBAR	Limited area barotropic model	Single-layer regional dynamical	E	Trk
CLP5	CLIPER5 (Climatology and Persistence model)	Statistical (baseline)	E	Trk

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
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
GFDI	Previous cycle GFDL, adjusted	Interpolated-dynamical	E	Trk, Int
GHMI	Previous cycle GFDL, adjusted using a variable intensity offset correction that is a function of forecast time. Note that for track, GHMI and GFDI are identical.	Interpolated-dynamical	E	Trk, Int
HWFI	Previous cycle HWRF, 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
EGRI	Previous cycle EGRR, adjusted	Interpolated-dynamical	E	Trk, Int
NVGI	Previous cycle NVGM, adjusted	Interpolated-dynamical	E	Trk, Int
GFNI	Previous cycle GFDN, 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
GUNA	Average of GFDI, EGRI, and GFSI	Consensus	E	Trk

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
CGUN	Version of GUNA corrected for model biases	Corrected consensus	E	Trk
AEMI	Previous cycle AEMN, adjusted	Consensus	E	Trk, Int
FSSE	FSU Super-ensemble	Corrected consensus	E	Trk, Int
GFEX*	Average of GFSI and EMXI	Consensus	E	Trk
TCON*	Average of GHMI, EGRI, GFSI, and HWFI	Consensus	E	Trk
TCCN*	Version of TCON corrected for model biases	Corrected consensus	E	Trk
TVCN*	Average of at least two of GFSI EGRI GHMI HWFI EMXI	Consensus	E	Trk
TVCA*	Average of at least two of GFSI EGRI GHMI HWFI EMXI	Consensus	E	Trk
TVCE*	Average of at least two of GFSI EGRI GHMI HWFI EMXI	Consensus	E	Trk
TVCC*	Version of TVCN corrected for model biases	Corrected consensus	E	Trk
ICON*	Average of DSHP, LGEM, GHMI, and HWFI	Consensus	E	Int
IVCN*	Average of at least two of DSHP LGEM GHMI HWFI GFNI	Consensus	E	Int

* The composition of the consensus aids can change from year to year; the table lists the composition used during the 2014 season.

Table 2. Homogenous comparison of official and CLIPER5 track forecast errors in the Atlantic basin for the 2014 season for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2014 mean OFCL error (n mi)	25.7	38.3	51.4	64.8	100.4	161.8	270.2
2014 mean CLIPER5 error (n mi)	51.6	116.2	178.7	242.8	375.5	475.7	602.0
2014 mean OFCL skill relative to CLIPER5 (%)	50.2	67.0	71.2	73.3	73.3	66.0	55.1
2014 mean OFCL bias vector (°/n mi)	072/003	085/005	109/008	142/009	181/026	206/075	227/160
2014 number of cases	139	120	104	89	66	46	32
2009-2013 mean OFCL error (n mi)	28.8	45.6	61.4	78.0	114.8	160.3	209.7
2009-2013 mean CLIPER5 error (n mi)	48.2	100.1	160.2	220.8	326.6	410.7	479.4
2009-2013 mean OFCL skill relative to CLIPER5 (%)	40.2	54.4	61.7	64.7	64.8	61.0	56.3
2009-2013 mean OFCL bias vector (°/n mi)	346/002	326/005	308/006	304/007	279/008	016/007	039/025
2009-2013 number of cases	1390	1213	1054	914	695	537	417
2014 OFCL error relative to 2009-2013 mean (%)	-10.8	-16.0	-16.3	-16.9	-12.5	0.9	28.9
2014 CLIPER5 error relative to 2009-2013 mean (%)	7.1	16.1	11.5	10.0	15.0	15.8	25.6

Table 3a. Homogenous comparison of Atlantic basin early track guidance model errors (n mi) for 2014. Errors smaller than the NHC official forecast are shown in bold-face.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	24.6	34.2	47.8	61.0	87.9	110.1	171.5
OCD5	57.4	123.1	188.1	228.9	362.0	404.8	536.1
GFSI	26.5	35.8	54.8	68.1	116.3	171.9	**166.9**
GHMI	30.3	48.5	76.1	94.7	149.5	183.1	242.3
HWFI	27.6	36.1	**47.2**	**54.6**	**79.2**	**107.5**	175.8
EMXI	24.6	**34.0**	**45.7**	64.9	108.7	150.5	303.9
CMCI	33.9	56.9	76.5	88.7	126.4	198.3	316.4
EGRI	26.6	39.9	51.9	68.8	93.0	**97.7**	**164.9**
NVGI	32.4	57.5	81.6	95.2	163.7	208.2	305.0
GFNI	32.9	59.2	84.0	107.0	197.5	320.4	438.0
AEMI	25.9	37.2	57.5	71.9	103.0	138.3	**117.3**
FSSE	**23.4**	**30.3**	**43.4**	**57.1**	103.7	125.6	183.3
TVCA	**23.0**	**30.1**	**43.4**	**52.4**	**83.0**	**104.2**	**163.7**
LBAR	39.3	70.4	101.0	129.8	196.7	298.0	546.4
BAMD	39.3	68.7	100.8	118.5	170.2	327.5	243.9
BAMM	35.8	58.7	87.1	100.4	140.5	212.9	187.8
BAMS	49.5	84.1	109.9	114.0	144.5	191.4	213.6
TCLP	64.0	144.4	235.4	307.1	510.3	616.5	819.0
# Cases	72	64	58	50	38	24	15

Table 3b. Homogenous comparison of Atlantic basin early track guidance model bias vectors (º/n mi) for 2014.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	017/009	020/015	024/015	030/016	037/026	247/044	234/102
OCD5	276/013	252/032	239/070	231/110	228/245	228/301	223/497
GFSI	007/013	002/018	005/030	016/031	027/055	341/048	336/049
GHMI	334/005	299/009	311/013	294/014	296/032	279/074	285/100
HWFI	358/013	358/015	007/021	016/018	018/008	247/033	235/079
EMXI	024/004	073/002	181/002	187/006	132/007	219/108	226/250
CMCI	045/008	078/013	077/023	078/031	083/042	190/081	215/203
EGRI	046/012	056/017	046/026	044/037	045/061	358/013	259/100
NVGI	331/013	328/017	320/020	327/015	279/006	218/084	209/204
GFNI	274/013	261/027	266/037	259/054	267/084	256/173	255/239
AEMI	011/012	007/009	016/033	025/041	034/063	339/034	236/049
FSSE	022/010	028/014	029/024	030/033	017/074	326/058	271/118
TVCA	012/009	010/009	011/015	018/016	024/024	262/037	250/097
LBAR	301/001	259/012	261/027	256/025	211/077	202/252	210/524
BAMD	318/007	263/010	243/018	229/026	239/052	255/138	347/034
BAMM	225/010	212/024	206/034	194/028	175/025	246/074	314/011
BAMS	249/021	237/043	229/059	222/050	205/059	225/118	209/153
TCLP	258/028	247/082	242/162	239/235	236/449	239/572	233/781
# Cases	72	64	58	50	38	24	15

Table 4. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the Atlantic basin for the 2014 season for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2014 mean OFCL error (kt)	5.2	8.2	10.0	10.4	12.5	11.3	6.6
2014 mean Decay-SHIFOR5 error (kt)	6.8	10.6	13.4	16.4	19.4	21.5	16.2
2014 mean OFCL skill relative to Decay-SHIFOR5 (%)	23.5	22.6	25.4	36.6	35.6	47.4	61.7
2014 OFCL bias (kt)	-1.9	-2.9	-3.6	-4.6	-5.8	-5.7	-2.8
2014 number of cases	139	120	104	89	66	46	32
2009-13 mean OFCL error (kt)	6.3	9.7	11.9	13.9	15.2	15.2	15.6
2009-13 mean Decay-SHIFOR5 error (kt)	7.4	11.1	13.8	15.7	18.3	18.2	18.1
2009-13 mean OFCL skill relative to Decay-SHIFOR5 (%)	14.9	12.6	13.8	11.5	16.9	16.5	13.8
2009-13 OFCL bias (kt)	-0.3	0.4	1.0	1.6	1.6	1.7	2.2
2009-13 number of cases	1390	1213	1054	914	695	537	417
2014 OFCL error relative to 2009-13 mean (%)	-17.5	-15.5	-16.0	-25.2	-17.8	-25.7	-57.7
2014 Decay-SHIFOR5 error relative to 2009-13 mean (%)	-8.1	-4.5	-2.9	4.5	6.0	18.1	-10.5

Table 5a. Homogenous comparison of selected Atlantic basin early intensity guidance model errors (kt) for 2014. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	5.1	7.9	9.8	10.0	11.3	10.5	5.5
OCD5	6.4	9.4	12.3	15.0	18.7	22.6	15.4
HWFI	7.4	10.5	13.0	14.8	14.2	14.2	11.3
GHMI	8.0	10.9	12.8	14.4	13.4	14.4	10.6
GFNI	7.3	10.1	13.1	14.4	15.5	19.6	14.7
DSHP	5.9	8.1	9.8	11.7	12.3	13.0	8.6
LGEM	6.2	8.4	10.9	12.9	14.5	15.3	6.4
IVCN	6.2	8.6	10.7	12.4	12.3	12.8	6.3
FSSE	6.0	9.0	11.7	13.3	15.4	15.3	9.4
GFSI	7.0	11.1	14.2	17.9	21.5	20.5	13.4
EMXI	6.6	11.3	16.8	24.4	31.4	31.8	29.2
TCLP	7.4	11.3	14.5	17.5	20.7	19.5	12.0
# Cases	84	75	66	58	44	32	19

Table 5b. Homogenous comparison of selected Atlantic basin early intensity guidance model biases (kt) for 2014. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	-2.0	-3.2	-3.9	-5.3	-6.5	-5.2	0.3
OCD5	-2.4	-3.9	-6.8	-10.4	-13.4	-15.7	-10.3
HWFI	-4.0	-7.6	-9.0	-9.2	**-6.2**	**-4.7**	-3.2
GHMI	-5.5	-8.9	-11.1	-11.3	-8.5	-6.5	2.8
GFNI	-4.7	-6.8	-8.5	-9.6	-7.1	-5.5	7.7
DSHP	**-1.9**	**-2.1**	**-2.0**	**-2.0**	**0.4**	**1.7**	6.1
LGEM	-2.9	-5.2	-7.7	-9.9	-11.0	-8.7	-4.5
IVCN	-3.3	-5.7	-7.3	-7.9	**-6.2**	**-4.3**	0.4
FSSE	-3.9	-6.9	-8.8	-10.7	-13.8	-12.5	-6.6
GFSI	-3.3	-5.1	-6.7	-7.5	-7.2	-8.5	-7.5
EMXI	-3.3	-4.6	-6.6	-7.8	-8.7	-12.4	-14.6
TCLP	-4.2	-7.8	-11.6	-15.1	-16.4	-14.1	-7.5
# Cases	84	75	66	58	44	32	19

Table 6. Official Atlantic track and intensity forecast verifications (OFCL) for 2014 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: AL012014 ARTHUR

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	4.3	4.1	18	2.5	2.5
012	16	13.7	30.2	16	5.6	9.3
024	14	27.0	85.7	14	11.4	15.6
036	12	41.2	161.5	12	11.7	20.8
048	10	55.9	239.8	10	11.5	25.7
072	6	117.0	422.8	6	15.8	26.7
096	2	179.4	724.8	2	5.0	27.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL022014 TWO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	0.8	1.5	8	0.0	0.0
012	6	28.7	36.6	6	0.0	2.3
024	4	62.4	68.4	4	0.0	6.0
036	2	115.8	99.3	2	0.0	11.0
048	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: AL032014 BERTHA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	6.3	6.3	23	2.0	1.7
012	21	27.2	54.6	21	5.5	7.8
024	19	38.3	122.0	19	8.9	12.5
036	17	39.4	206.0	17	10.6	14.9
048	15	43.0	315.6	15	11.0	15.1
072	11	66.1	564.0	11	15.5	12.7
096	7	154.3	829.9	7	15.0	6.1
120	3	253.4	1031.9	3	1.7	6.0

Verification statistics for: AL042014 CRISTOBAL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	6.8	6.8	23	1.1	1.1
012	21	33.7	69.4	21	3.6	6.3
024	19	53.7	149.1	19	5.5	7.6
036	17	73.3	205.7	17	6.8	5.9
048	15	103.2	225.0	15	7.0	8.1
072	11	150.7	340.7	11	3.6	9.5
096	7	277.2	387.8	7	2.1	13.7
120	3	825.9	739.3	3	3.3	12.0

Verification statistics for: AL052014 DOLLY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	14.7	14.7	7	2.1	2.1
012	5	71.4	92.2	5	8.0	10.4
024	3	61.0	107.5	3	6.7	9.7
036	1	30.5	18.8	1	5.0	2.0
048	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: AL062014 EDOUARD

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	33	5.6	6.0	33	1.1	1.1
012	31	20.5	53.6	31	2.7	4.1
024	29	31.3	128.2	29	5.7	6.6
036	27	42.2	190.5	27	7.2	10.4
048	25	50.1	246.0	25	8.2	12.8
072	21	73.5	355.9	21	12.6	20.0
096	17	118.8	401.0	17	12.6	22.1
120	13	169.0	398.7	13	10.0	16.3

Verification statistics for: AL072014 FAY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	11	12.6	13.1	11	3.6	4.1
012	9	35.7	58.7	9	7.8	6.4
024	7	48.9	115.2	7	9.3	10.3
036	5	101.8	174.8	5	15.0	11.0
048	3	134.2	266.7	3	16.7	18.0
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: AL082014 GONZALO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	29	5.3	5.5	29	3.6	3.6
012	27	18.3	41.3	27	8.5	9.1
024	25	31.6	98.9	25	12.0	14.1
036	23	44.5	148.5	23	14.3	18.8
048	21	65.0	197.9	21	13.8	22.9
072	17	117.5	283.5	17	15.0	26.6
096	13	157.2	391.7	13	13.5	32.2
120	9	263.6	662.3	9	6.1	23.7

Verification statistics for: AL092014 HANNA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	12.6	12.6	7	0.7	0.7
012	3	33.1	34.0	3	6.7	3.3
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
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	4	210.1	702.1	4	2.5	9.5

Table 7. Homogenous comparison of official and CLIPER5 track forecast errors in the eastern North Pacific basin in 2014 for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2014 mean OFCL error (n mi)	20.3	29.9	38.5	50.1	82.4	119.4	177.0
2014 mean CLIPER5 error (n mi)	36.0	73.8	113.8	155.5	241.5	326.8	435.1
2014 mean OFCL skill relative to CLIPER5 (%)	43.6	59.5	66.2	67.8	65.9	63.5	59.3
2014 mean OFCL bias vector (°/n mi)	317/004	300/006	279/006	249/007	204/013	178/011	254/018
2014 number of cases	405	367	327	293	237	188	136
2009-2013 mean OFCL error (n mi)	25.7	41.4	55.0	68.6	97.8	134.2	167.1
2009-2013 mean CLIPER5 error (n mi)	37.2	74.7	118.0	162.5	249.4	332.6	413.3
2009-2013 mean OFCL skill relative to CLIPER5 (%)	30.9	44.6	53.3	57.8	60.8	59.7	59.6
2009-2013 mean OFCL bias vector (°/n mi)	264/000	140/001	125/004	099/007	078/015	063/028	046/039
2009-2013 number of cases	1140	999	865	739	527	359	227
2014 OFCL error relative to 2009-2013 mean (%)	-21.0	-27.8	-30.0	-27.0	-15.7	-11.0	5.9
2014 CLIPER5 error relative to 2009-2013 mean (%)	-3.2	-1.2	-3.6	-4.3	-3.2	-1.7	5.3

Table 8a. Homogenous comparison of eastern North Pacific basin early track guidance model errors (n mi) for 2014. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	18.3	27.6	36.8	47.7	81.2	126.1	196.0
OCD5	33.2	68.7	109.0	154.0	250.2	373.3	531.0
GFSI	21.1	33.8	45.3	58.3	99.6	166.2	260.8
GHMI	22.3	37.6	51.9	68.2	106.1	176.8	285.1
HWFI	21.7	34.8	47.8	62.0	100.6	156.2	216.3
EMXI	20.3	31.5	41.4	51.7	88.0	**122.7**	**185.1**
EGRI	22.6	38.9	56.2	75.2	115.2	165.1	246.9
CMCI	28.4	50.0	72.4	94.4	139.3	216.3	324.7
AEMI	21.2	32.8	43.1	55.8	94.7	147.8	210.9
FSSE	**17.1**	**26.1**	**35.1**	**45.3**	**80.2**	**121.3**	**179.6**
TVCE	**16.8**	**25.7**	**34.6**	**44.2**	**75.3**	**118.8**	**180.0**
LBAR	29.0	61.1	99.4	141.2	225.7	329.1	437.3
BAMD	33.3	61.6	89.9	116.3	168.6	261.1	377.9
BAMM	28.8	51.2	73.7	97.3	142.7	201.3	275.3
BAMS	35.1	64.2	91.5	117.8	167.6	230.2	287.5
TCLP	32.7	68.2	108.1	155.5	258.5	376.3	531.4
# Cases	296	275	251	229	173	125	69

Table 8b. Homogenous comparison of eastern North Pacific basin early track guidance model bias vectors (º/n mi) for 2014.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	323/004	300/005	270/006	245/009	203/015	151/005	344/038
OCD5	311/005	286/012	269/026	259/043	243/086	253/142	274/272
GFSI	319/006	315/007	315/006	311/004	023/002	032/030	010/093
GHMI	088/002	093/004	079/007	079/009	071/024	051/065	038/133
HWFI	345/009	338/009	310/008	285/011	227/018	210/017	253/037
EMXI	320/003	272/004	233/009	215/016	202/036	192/039	257/022
EGRI	270/006	261/015	260/025	260/035	257/051	262/053	302/085
CMCI	042/002	096/004	097/008	090/012	095/031	070/061	064/087
AEMI	297/006	300/009	304/009	305/009	329/010	014/032	002/074
FSSE	322/003	303/004	281/004	264/007	221/014	262/006	326/039
TVCE	320/004	293/005	273/007	256/010	228/014	338/001	352/046
LBAR	353/013	331/039	320/068	314/096	310/150	316/210	313/293
BAMD	338/015	337/029	333/043	327/055	312/071	321/097	305/159
BAMM	336/016	330/028	324/040	315/050	296/068	306/073	299/086
BAMS	348/014	336/024	325/035	314/047	291/074	287/090	290/102
TCLP	284/005	258/016	250/036	245/060	234/113	244/183	264/333
# Cases	296	275	251	229	173	125	69

Table 9. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the eastern North Pacific basin for the 2014 season for all tropical cyclones. Averages for the previous 5-yr period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2014 mean OFCL error (kt)	6.0	9.7	12.2	14.3	16.8	16.3	13.8
2014 mean Decay-SHIFOR5 error (kt)	7.8	13.0	16.5	19.3	21.0	20.1	16.0
2014 mean OFCL skill relative to Decay-SHIFOR5 (%)	23.1	25.4	26.1	25.9	20.0	18.9	13.8
2014 OFCL bias (kt)	-1.1	-1.7	-3.4	-5.3	-7.1	-5.5	-2.5
2014 number of cases	405	367	327	293	237	188	136
2009-13 mean OFCL error (kt)	6.1	10.4	13.4	14.5	15.0	16.4	16.1
2009-13 mean Decay-SHIFOR5 error (kt)	7.7	12.7	16.4	18.8	20.5	20.3	20.8
2009-13 mean OFCL skill relative to Decay-SHIFOR5 (%)	20.8	18.1	18.3	22.9	26.8	19.2	22.6
2009-13 OFCL bias (kt)	-0.2	-0.2	-0.5	-0.8	0.8	2.2	3.0
2009-13 number of cases	1140	999	865	739	527	359	227
2014 OFCL error relative to 2009-13 mean (%)	-1.6	-6.7	-9.0	-1.4	12.0	-0.6	-14.3
2014 Decay-SHIFOR5 error relative to 2009-13 mean (%)	1.3	2.4	0.7	2.7	2.4	-1.0	-23.1

Table 10a. Homogenous comparison of eastern North Pacific basin early intensity guidance model errors (kt) for 2014. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	6.6	10.4	12.9	14.7	16.4	15.7	14.9
OCD5	8.4	13.9	17.3	19.5	20.0	18.7	15.1
HWFI	7.3	11.4	14.4	16.1	16.8	15.8	**13.3**
GHMI	8.4	13.1	17.1	17.6	16.9	**15.1**	**12.6**
DSHP	7.6	12.0	14.4	16.0	17.4	17.0	17.5
LGEM	7.8	12.5	15.4	17.7	19.1	16.6	15.9
IVCN	7.1	11.1	13.7	15.2	**15.9**	**13.9**	**12.2**
FSSE	7.3	10.6	**12.7**	15.0	16.5	15.8	16.5
GFSI	8.6	13.9	17.6	20.0	19.1	16.6	**12.9**
EMXI	10.2	17.4	22.9	26.3	25.9	21.0	16.9
TCLP	8.1	13.7	17.9	20.7	21.9	19.7	18.6
# Cases	332	304	271	246	196	149	96

Table 10b. Homogenous comparison of eastern North Pacific basin early intensity guidance model biases (kt) for 2014. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	-1.2	-1.7	-3.4	-4.8	-5.9	-3.9	-2.0
OCD5	-1.6	-3.0	-5.4	-7.0	-7.9	-4.1	-2.3
HWFI	-3.2	-5.8	-8.1	-9.3	-8.0	-5.1	-3.3
GHMI	-3.8	-8.8	-13.3	-13.6	-10.6	-5.7	-2.3
DSHP	-1.3	**-1.6**	**-2.5**	**-3.3**	**-5.3**	**-3.5**	-3.4
LGEM	-1.7	-3.7	-6.3	-7.9	-9.8	-8.1	-7.1
IVCN	-2.3	-4.8	-7.3	-8.3	-8.2	-5.3	-3.9
FSSE	-1.7	-3.2	-5.2	-6.9	-8.9	-7.8	-8.7
GFSI	-2.6	-2.8	**-2.8**	**-2.4**	**-1.6**	**1.3**	3.0
EMXI	-1.8	-2.2	**-2.6**	**-2.6**	**-2.3**	**1.5**	3.8
TCLP	-2.2	-5.2	-8.4	-10.2	-10.9	-6.9	-5.4
# Cases	332	304	271	246	196	149	96

Table 11. Official eastern North Pacific track and intensity forecast verifications (OFCL) for 2014 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: EP012014 AMANDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	28	5.9	6.2	28	3.2	3.2
012	26	20.0	33.7	26	8.7	12.3
024	24	24.9	70.3	24	14.8	21.0
036	22	33.2	121.3	22	20.0	27.0
048	20	39.4	164.4	20	25.0	32.7
072	16	61.1	296.1	16	27.2	39.1
096	12	58.9	425.1	12	20.0	29.8
120	8	79.0	550.4	8	11.3	15.0

Verification statistics for: EP022014 BORIS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	13.0	13.0	8	1.3	1.3
012	6	24.3	34.0	6	2.5	4.0
024	4	44.7	72.1	4	6.3	7.3
036	2	69.4	116.7	2	15.0	18.0
048	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: EP032014 CRISTINA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	22	6.2	6.5	22	1.6	2.0
012	20	18.0	21.2	20	10.8	14.7
024	18	25.1	40.4	18	17.2	22.9
036	16	34.5	59.3	16	20.0	27.4
048	14	45.4	75.7	14	23.6	34.9
072	10	57.6	89.9	10	19.0	32.6
096	6	89.8	81.5	6	15.0	22.3
120	2	127.6	77.2	2	7.5	5.5

Verification statistics for: EP042014 DOUGLAS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	26	9.3	9.6	26	1.3	1.3
012	24	19.0	37.1	24	4.8	4.5
024	22	22.1	70.6	22	6.1	5.7
036	20	31.1	110.7	20	8.0	6.6
048	18	49.7	163.8	18	8.1	8.6
072	14	94.5	303.2	14	6.8	4.6
096	10	156.2	475.4	10	8.5	5.8
120	6	234.4	632.9	6	8.3	6.7

Verification statistics for: EP052014 ELIDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	2.5	2.5	7	0.0	0.0
012	5	24.7	57.1	5	5.0	5.0
024	3	61.1	170.8	3	10.0	12.3
036	1	95.5	301.5	1	15.0	20.0
048	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: EP062014 FAUSTO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	11.3	11.3	7	4.3	4.3
012	5	30.4	36.8	5	10.0	13.0
024	3	34.9	67.8	3	18.3	23.3
036	1	36.0	109.5	1	25.0	38.0
048	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: EP072014 GENEVIEVE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	9	4.5	4.5	9	0.0	0.0
012	9	24.1	43.0	9	2.8	4.9
024	7	36.0	95.0	7	7.1	10.7
036	5	41.2	116.8	5	6.0	18.2
048	3	72.2	124.8	3	3.3	22.7
072	3	70.0	147.5	3	3.3	16.7
096	7	80.8	211.6	7	2.1	22.7
120	7	125.8	292.7	7	5.0	24.9

Verification statistics for: EP082014 HERNAN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	10.0	10.5	13	0.4	0.4
012	11	17.9	39.7	11	5.0	7.0
024	9	34.5	85.7	9	10.0	16.9
036	7	42.5	138.0	7	10.7	19.6
048	5	54.4	204.3	5	4.0	12.6
072	1	82.8	377.5	1	5.0	17.0
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: EP092014 ISELLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	21	5.0	5.0	21	1.7	1.7
012	21	12.9	22.3	21	6.7	8.6
024	21	21.3	44.4	21	12.1	15.7
036	21	25.1	69.5	21	17.4	21.5
048	21	31.4	92.3	21	21.2	24.2
072	21	47.1	144.5	21	24.3	27.3
096	18	77.4	171.2	18	26.1	28.7
120	14	103.4	216.8	14	22.5	22.2

Verification statistics for: EP102014 JULIO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	5.3	5.3	18	2.8	2.8
012	18	19.4	34.1	18	7.8	9.8
024	18	28.2	58.9	18	10.0	13.6
036	18	38.2	90.8	18	9.7	15.6
048	18	53.8	135.9	18	12.8	19.4
072	18	99.3	239.8	18	16.7	26.2
096	18	167.9	344.6	18	16.1	26.6
120	18	251.6	434.7	18	11.1	21.1

Verification statistics for: EP112014 KARINA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	55	6.9	7.1	55	1.8	1.7
012	53	21.9	38.4	53	5.4	5.9
024	51	31.4	81.6	51	7.8	9.4
036	49	42.4	137.0	49	10.4	11.8
048	47	51.8	204.4	47	12.4	12.7
072	43	96.4	352.7	43	15.9	13.3
096	39	152.2	518.5	39	15.5	12.4
120	35	230.7	687.7	35	15.1	10.7

Verification statistics for: EP122014 LOWELL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	25	8.6	8.9	25	0.6	0.6
012	24	19.7	27.2	24	3.5	5.0
024	22	30.5	48.8	22	4.8	7.5
036	20	34.4	73.3	20	5.0	9.5
048	18	43.6	101.8	18	6.7	11.9
072	14	65.6	149.1	14	6.4	9.7
096	10	94.2	219.1	10	6.0	7.2
120	6	134.1	296.9	6	3.3	14.5

Verification statistics for: EP132014 MARIE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	27	6.5	6.5	27	0.4	0.7
012	25	17.7	38.7	25	7.2	10.1
024	23	23.5	73.9	23	12.6	15.9
036	21	24.4	91.9	21	12.4	20.3
048	19	35.4	109.2	19	13.9	22.9
072	15	69.7	163.8	15	19.3	23.5
096	11	96.9	212.2	11	16.4	16.2
120	7	134.0	296.4	7	22.9	13.9

Verification statistics for: EP142014 NORBERT

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	22	5.9	6.1	22	0.5	0.5
012	20	17.9	37.0	20	6.3	8.9
024	18	27.7	85.2	18	11.7	17.4
036	16	33.2	135.8	16	15.9	23.6
048	14	31.1	169.1	14	18.6	28.9
072	10	34.2	196.0	10	24.0	35.1
096	6	68.2	255.5	6	23.3	37.7
120	2	145.6	502.9	2	12.5	23.0

Verification statistics for: EP152014 ODILE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	31	6.8	6.8	31	2.6	2.6
012	30	20.9	34.2	30	6.7	7.2
024	28	37.3	70.2	28	10.4	11.9
036	26	56.9	117.9	26	14.6	13.1
048	24	82.3	170.1	24	15.0	15.8
072	20	152.2	275.1	20	14.3	20.1
096	16	203.2	377.2	16	19.4	29.7
120	12	252.2	441.1	12	20.4	22.0

Verification statistics for: EP162014 SIXTEEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	12.2	12.2	15	0.0	0.0
012	13	26.1	66.8	13	1.2	2.9
024	11	34.9	151.3	11	2.7	3.9
036	7	45.4	267.3	7	5.0	4.3
048	5	73.1	420.8	5	4.0	4.0
072	1	159.5	702.1	1	5.0	3.0
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: EP172014 POLO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	25	10.0	9.7	25	1.4	1.8
012	23	20.3	32.5	23	3.9	4.6
024	21	27.2	56.4	21	6.7	7.7
036	19	31.5	67.0	19	9.2	10.4
048	17	40.0	77.2	17	10.9	13.8
072	13	53.2	109.6	13	13.5	19.4
096	9	67.4	118.2	9	19.4	23.1
120	5	56.7	130.0	5	16.0	21.6

Verification statistics for: EP182014 RACHEL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	24	7.1	7.4	24	1.5	1.7
012	22	19.2	32.3	22	4.8	6.0
024	20	30.0	65.7	20	6.3	10.3
036	18	41.2	86.6	18	6.9	12.1
048	16	50.7	108.6	16	10.6	14.4
072	12	60.4	187.9	12	15.4	13.4
096	8	90.9	215.3	8	15.0	9.5
120	4	102.6	253.5	4	11.3	7.5

Verification statistics for: EP192014 SIMON

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	25	7.9	7.9	25	1.8	2.2
012	23	17.2	32.2	23	7.0	11.2
024	21	23.7	67.1	21	11.4	19.0
036	19	38.0	111.8	19	14.5	23.3
048	17	54.7	164.9	17	16.8	27.3
072	13	112.8	246.5	13	20.8	25.2
096	9	141.5	251.3	9	13.3	13.0
120	5	155.5	207.9	5	4.0	3.0

Verification statistics for: EP202014 TRUDY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	6.1	6.1	6	4.2	4.2
012	4	31.0	33.3	4	11.3	10.8
024	2	60.9	76.7	2	12.5	2.5
036	0	-999.0	-999.0	0	-999.0	-999.0
048	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: EP212014 VANCE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	25	11.3	11.2	25	1.0	0.8
012	23	26.9	51.3	23	6.5	8.3
024	21	43.4	120.0	21	10.0	14.7
036	19	53.8	205.7	19	13.2	19.9
048	17	67.6	269.1	17	15.3	23.2
072	13	79.7	341.2	13	16.9	21.8
096	9	48.8	330.4	9	18.3	26.6
120	5	63.8	195.4	5	10.0	23.8

Table 12a Verification of 48-h probabilistic genesis forecasts for the Atlantic basin in 2014.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	3	257
10	11	161
20	19	57
30	19	32
40	80	10
50	50	16
60	100	11
70	55	11
80	100	4
90	-	0
100	-	0

Table 12b. Verification of 48-h probabilistic genesis forecasts for the eastern North Pacific basin in 2014.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	3	301
10	20	128
20	26	100
30	29	58
40	52	27
50	65	37
60	81	26
70	88	17
80	100	9
90	100	6
100	100	7

Table 13a Verification of 120-h probabilistic genesis forecasts for the Atlantic basin in 2014.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	32
10	9	221
20	22	137
30	21	62
40	41	27
50	82	17
60	100	23
70	92	24
80	100	13
90	100	3
100	-	0

Table 13b Verification of 120-h probabilistic genesis forecasts for the eastern North Pacific basin in 2014.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	9
10	22	143
20	43	164
30	38	99
40	63	40
50	51	71
60	85	52
70	93	41
80	94	50
90	100	35
100	100	12

Table 14. HFIP Stream 1.5 models for 2014.

ID	Description	Parameter	NHC Application
UWNI	University of Wisconsin non-hydrostatic. Early version of UWN8.	Int	Include in IV15 consensus.
SPC3	CIRA statistical intensity consensus.	Int	Direct use.
CXTI	NRL COAMPS-TC regional model. Early version of COTC.	Int	Use explicitly and include in IV15 consensus.
GPMI	GFDL ensemble mean. Early version of GPMN.	Trk, Int	Direct use and replaces GHMI in TV15 and IV15

Table 15. NHC forecast cone circle radii (n mi) for 2015. Change from 2014 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**
12	32 (-1: -3%)	26 (-4: -13%)
24	52 (0: 0%)	42 (-4: -9%)
36	71 (-1: -1%)	54 (-8: -13%)
48	90 (-2: -2%)	69 (-10: -13%)
72	122 (-3: -2%)	100 (-5: -5%)
96	170 (0: 0%)	143 (-11: -7%)
120	225 (-1: -0.5%)	182 (-8: -4%)

Table 16. Composition of NHC consensus models for 2015. It is intended that TCOA/TVCA would be the primary consensus aids for the Atlantic basin and TCOE/TVCE would be primary for the eastern Pacific.

NHC Consensus Model Definitions For 2015			
Model ID	**Parameter**	**Type**	**Members**
GFEX	Track	Fixed	GFSI EMXI
TCOA	Track	Fixed	GFSI EGRI GHMI HWFI
TCOE*	Track	Fixed	GFSI EGRI GHMI HWFI
ICON	Intensity	Fixed	DSHP LGEM GHMI HWFI
TVCA	Track	Variable	GFSI EGRI GHMI HWFI EMXI
TVCE**	Track	Variable	GFSI EGRI GHMI HWFI EMXI
IVCN	Intensity	Variable	DSHP LGEM GHMI HWFI

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

List of Figures

17. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.
18. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2014. This verification includes only those models that were available at least 2/3 of the time (see text).
19. Reliability diagram for Atlantic (top) and eastern North Pacific (bottom) probabilistic tropical cyclogenesis 48-h forecasts for 2014. 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.
20. As described for Fig. 19, but for 120-h forecasts.
21. Homogeneous comparison of HFIP Stream 1.5 track models and selected operational models for 2014.
22. Comparison of Stream 1.5 consensus TV15 and TVCA.
23. Homogeneous comparison of HFIP Stream 1.5 GFDL ensemble mean (GPMI) and GHMI for track (top) and intensity (bottom).
24. Homogeneous comparison of HFIP Stream 1.5 intensity models and selected operational models for 2014.
25. Comparison of Stream 1.5 consensus IV15 and IVCN.
26. Homogeneous comparison of HFIP Stream 1.5 model SPC3 and operational models DSHP and LGEM.

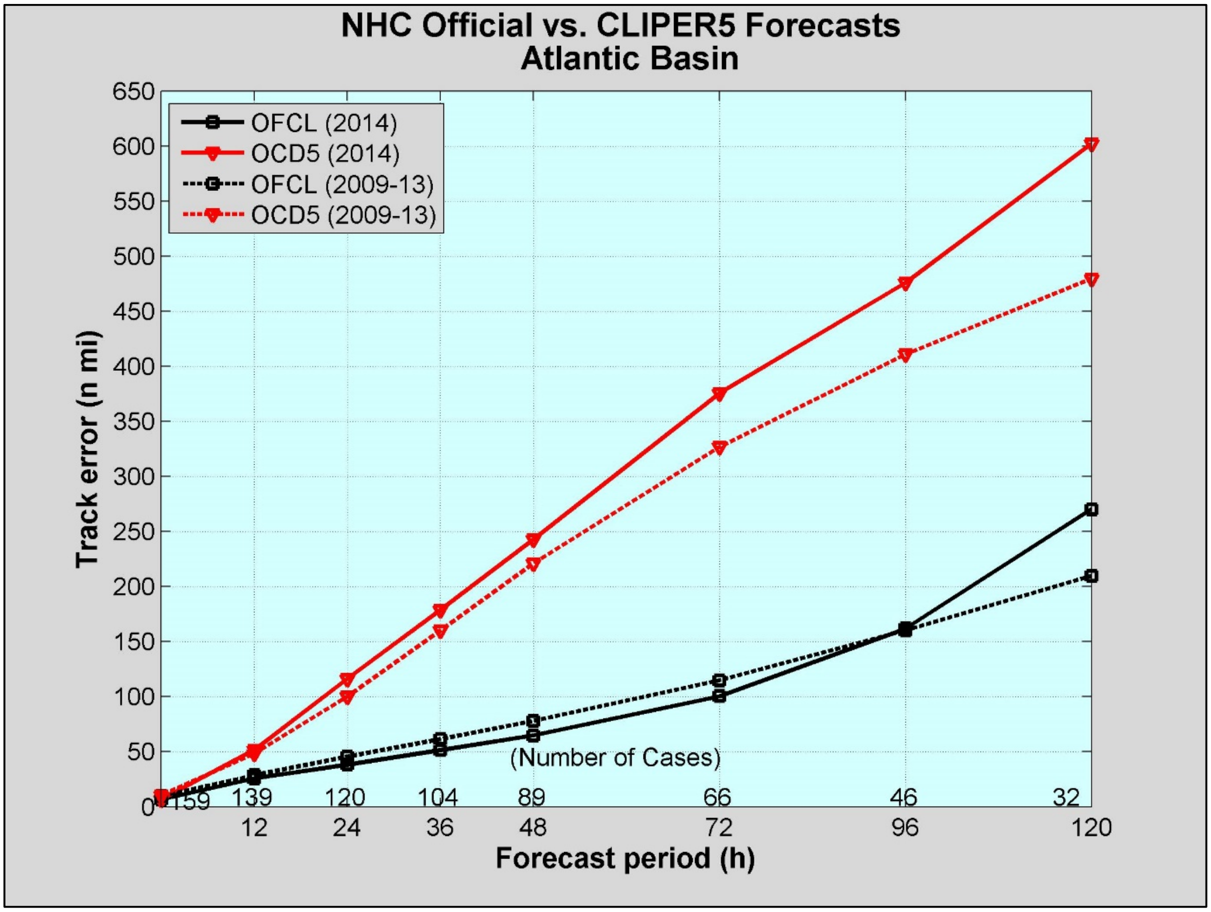


Figure 1. NHC official and CLIPER5 (OCD5) Atlantic basin average track errors for 2014 (solid lines) and 2009-2013 (dashed lines).

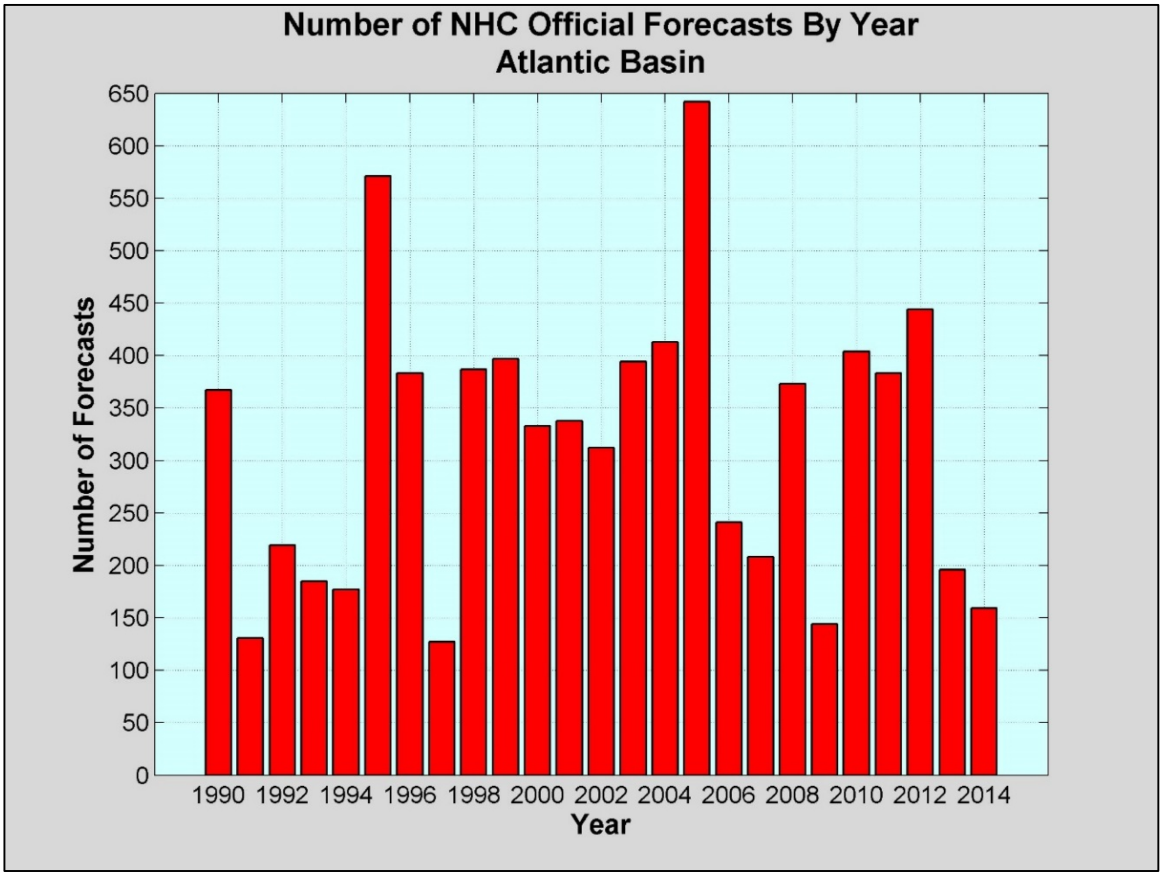
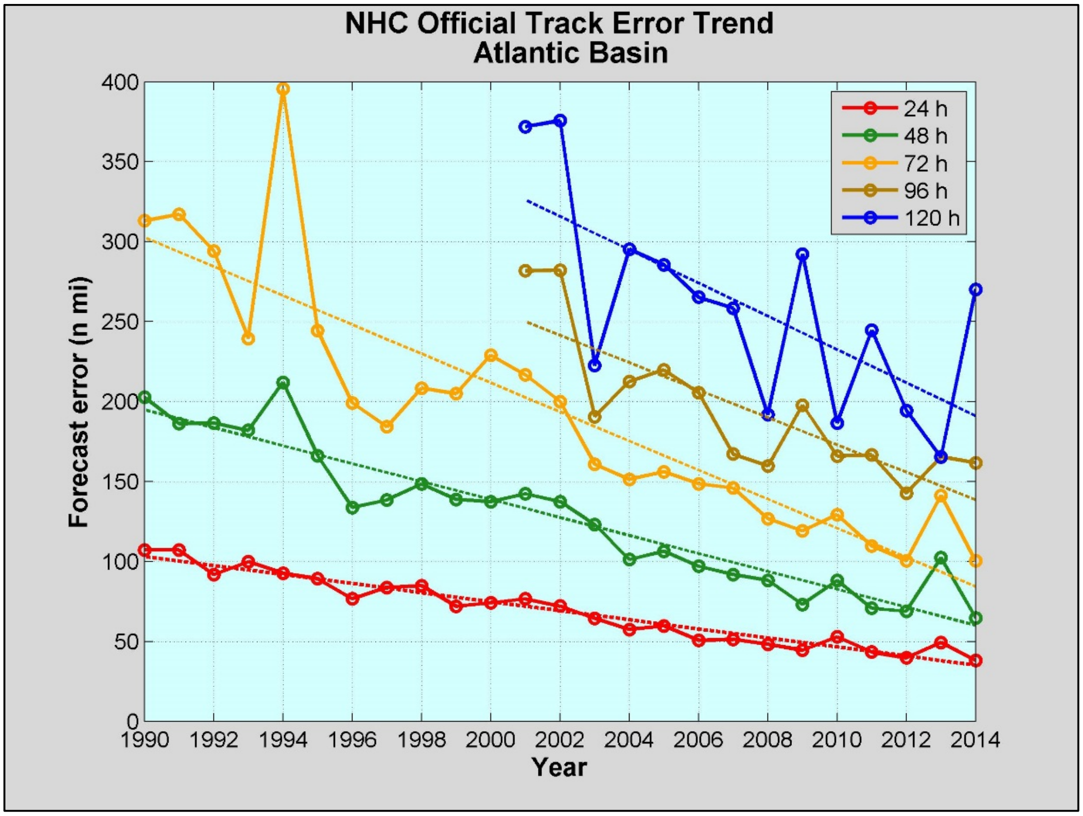


Figure 2. Number of NHC official forecasts for the Atlantic basin stratified by year.



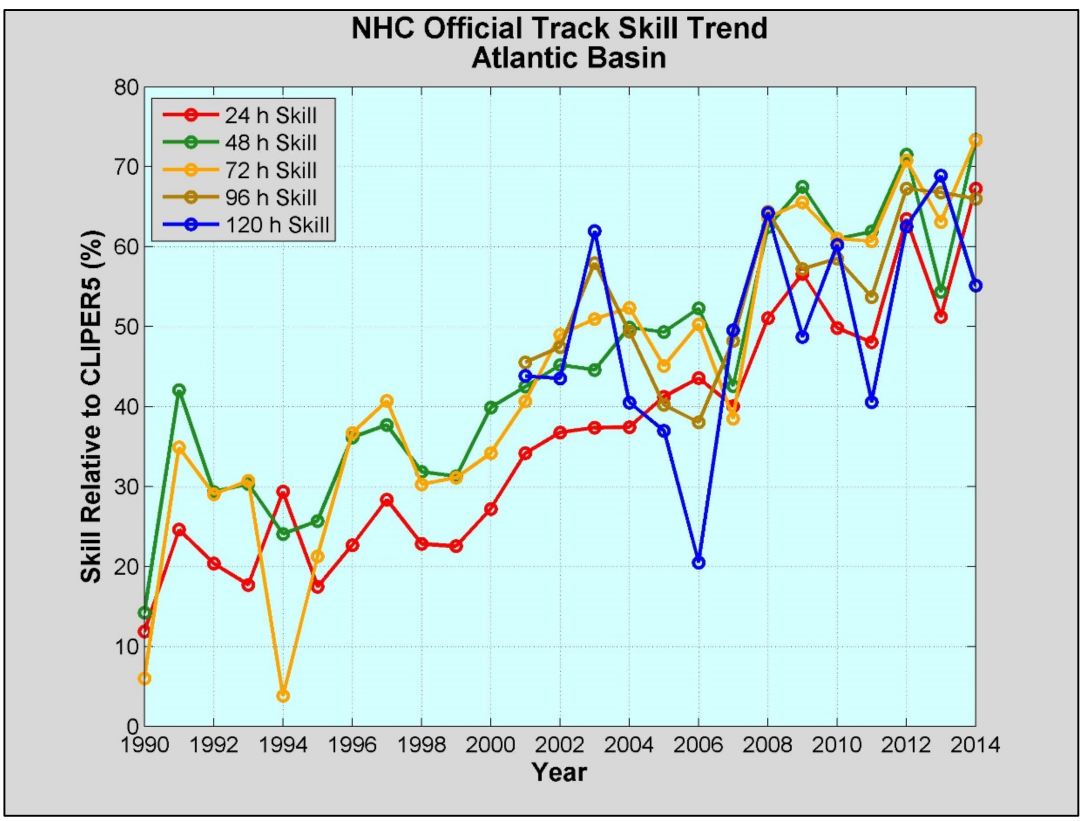


Figure 3. Recent trends in NHC official track forecast error (top) and skill (bottom) for the Atlantic basin

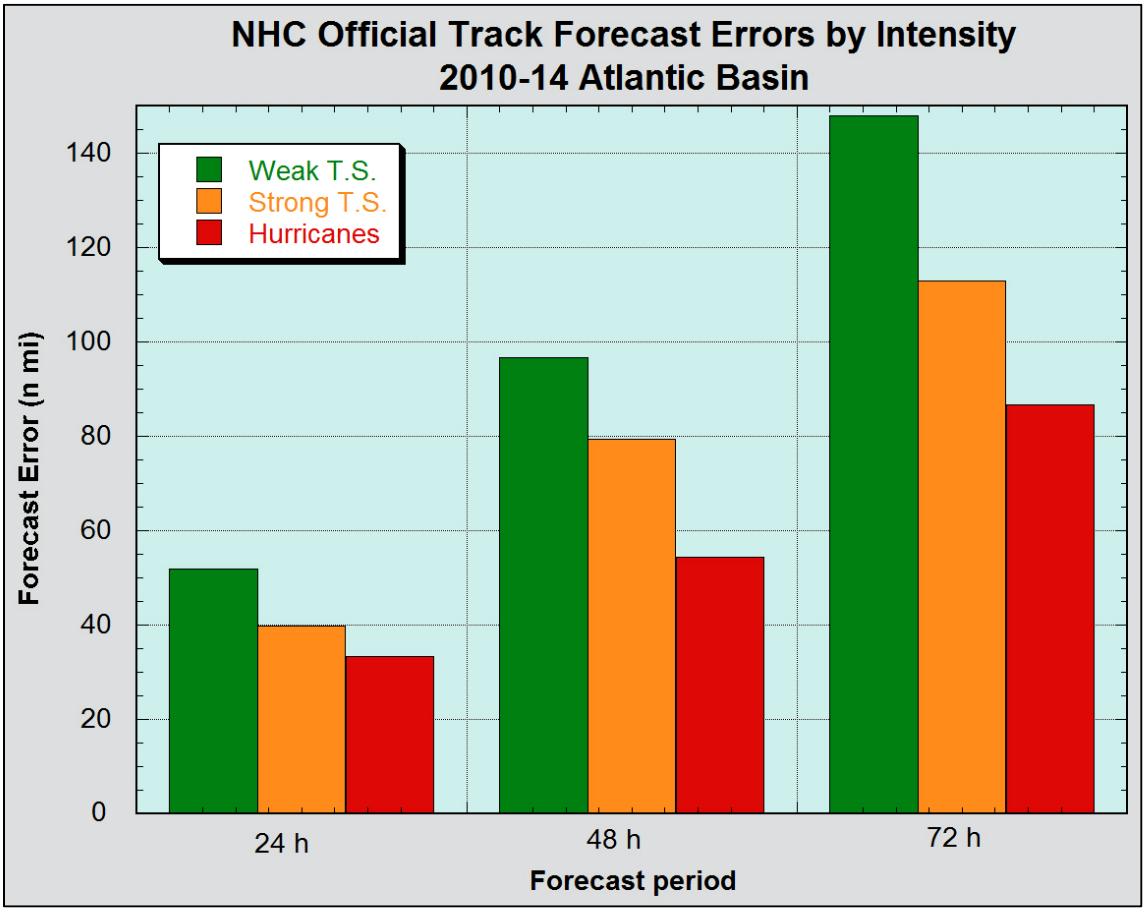


Figure 4. 2010-14 NHC official track forecast error binned by initial intensity for the Atlantic basin. Weak tropical storms are in the 35-45 kt range and strong tropical storms are in the 50-60 kt range.

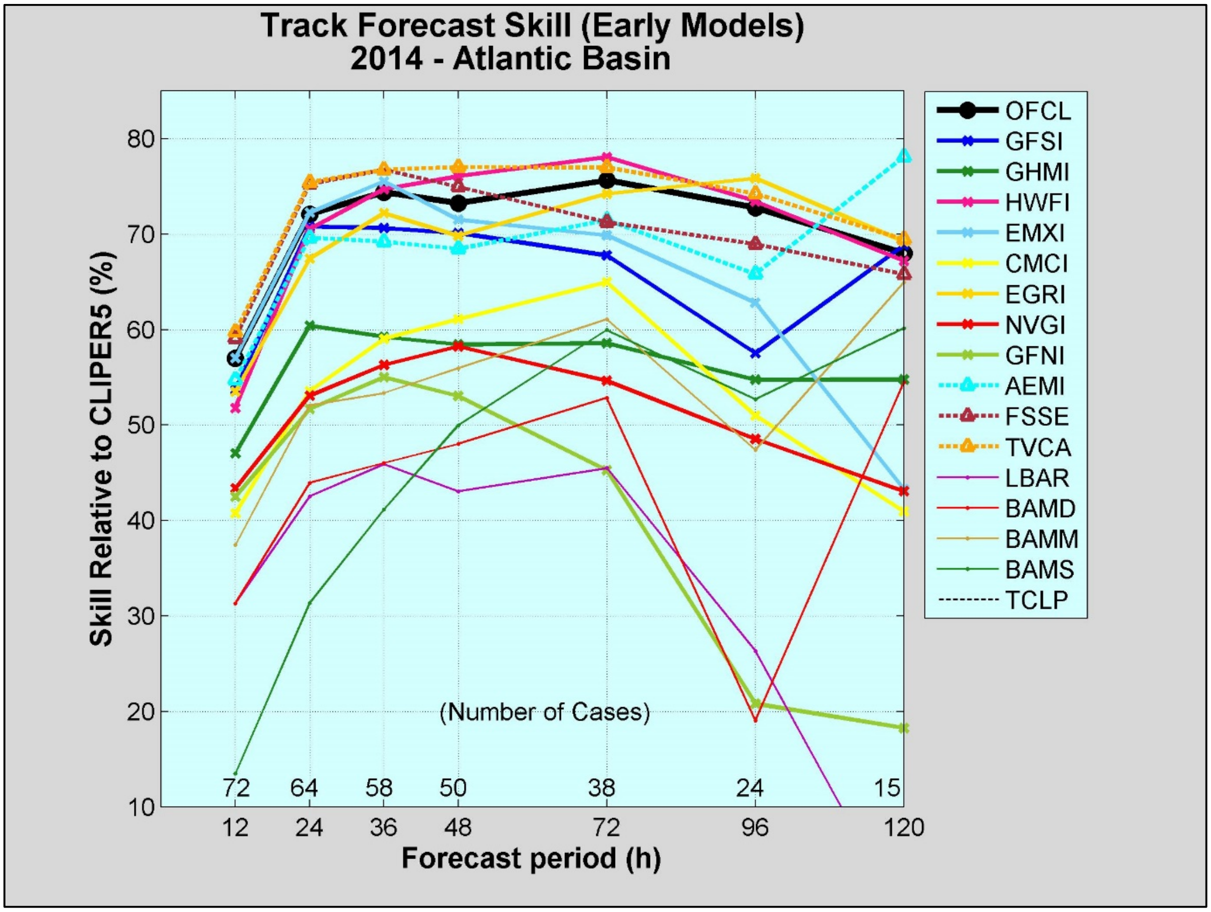


Figure 5. Homogenous comparison for selected Atlantic basin early track models for 2014. This verification includes only those models that were available at least 2/3 of the time (see text).

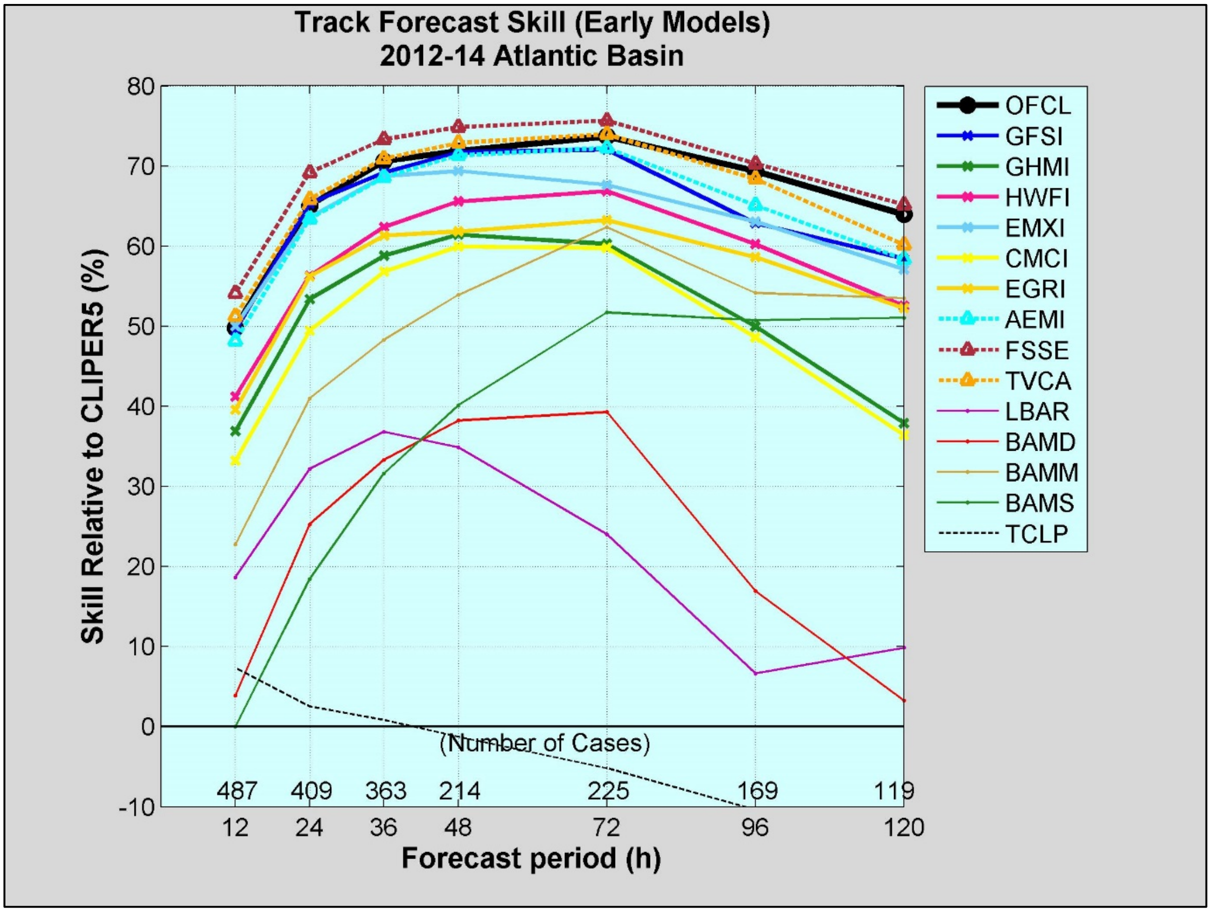


Figure 6. Homogenous comparison for selected Atlantic basin early track models for 2012-2014.

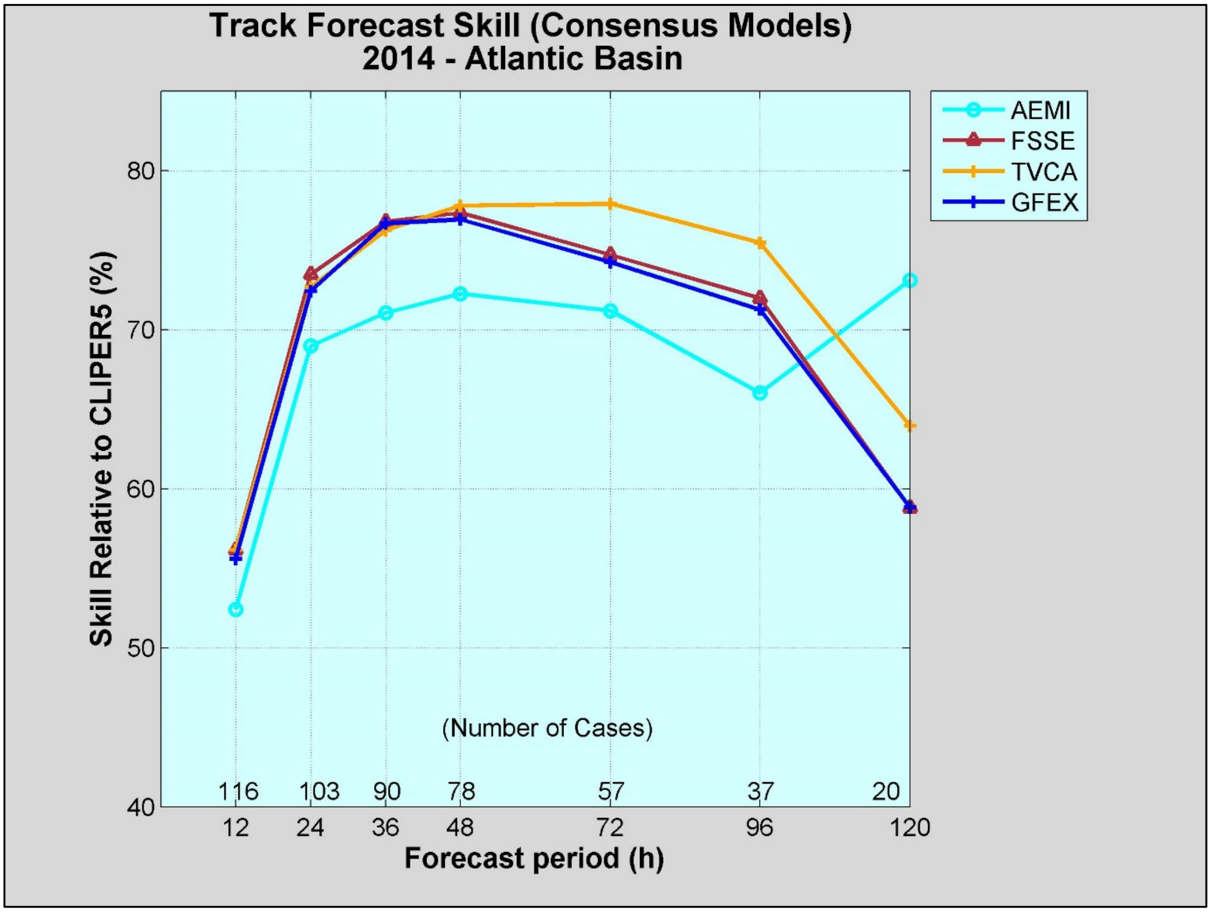


Figure 7. Homogenous comparison of the primary Atlantic basin track consensus models for 2014.

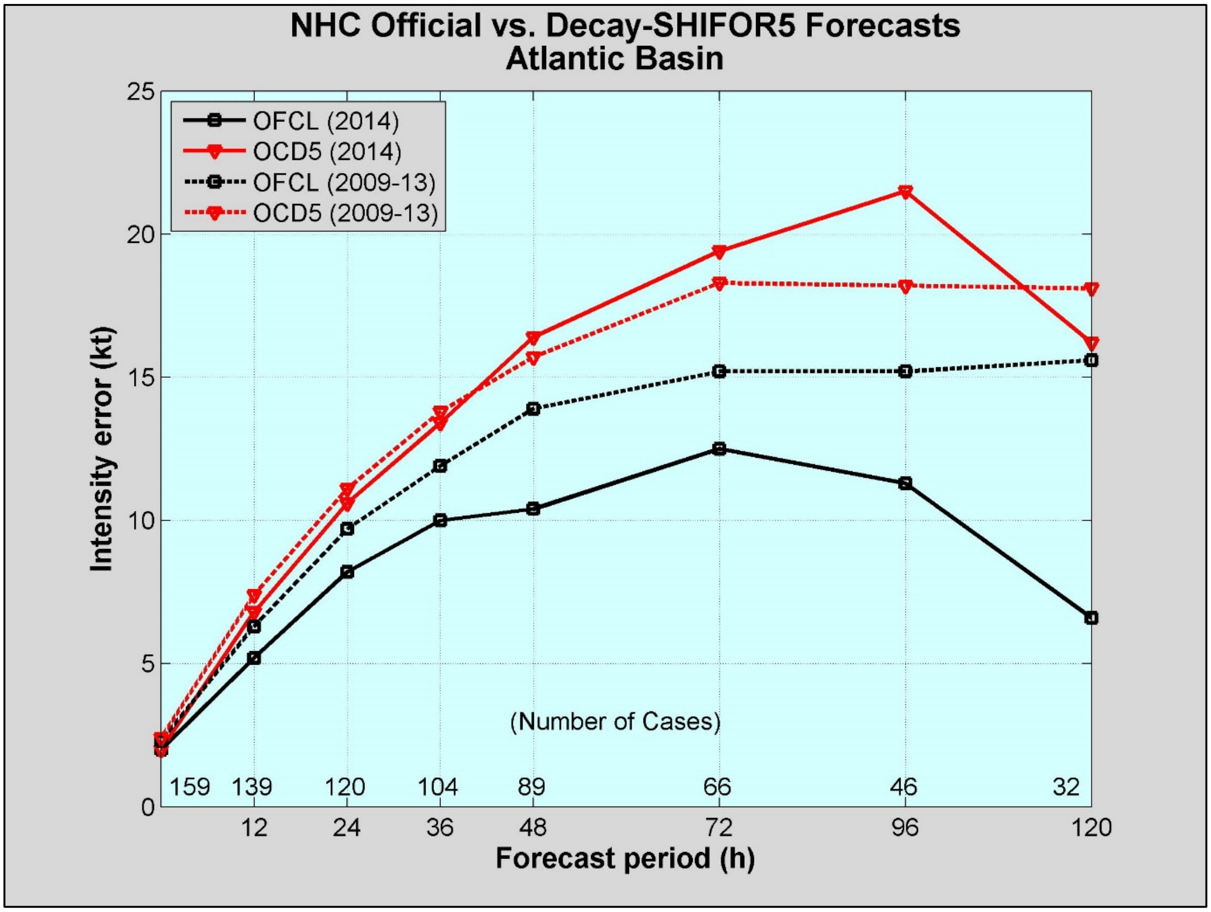
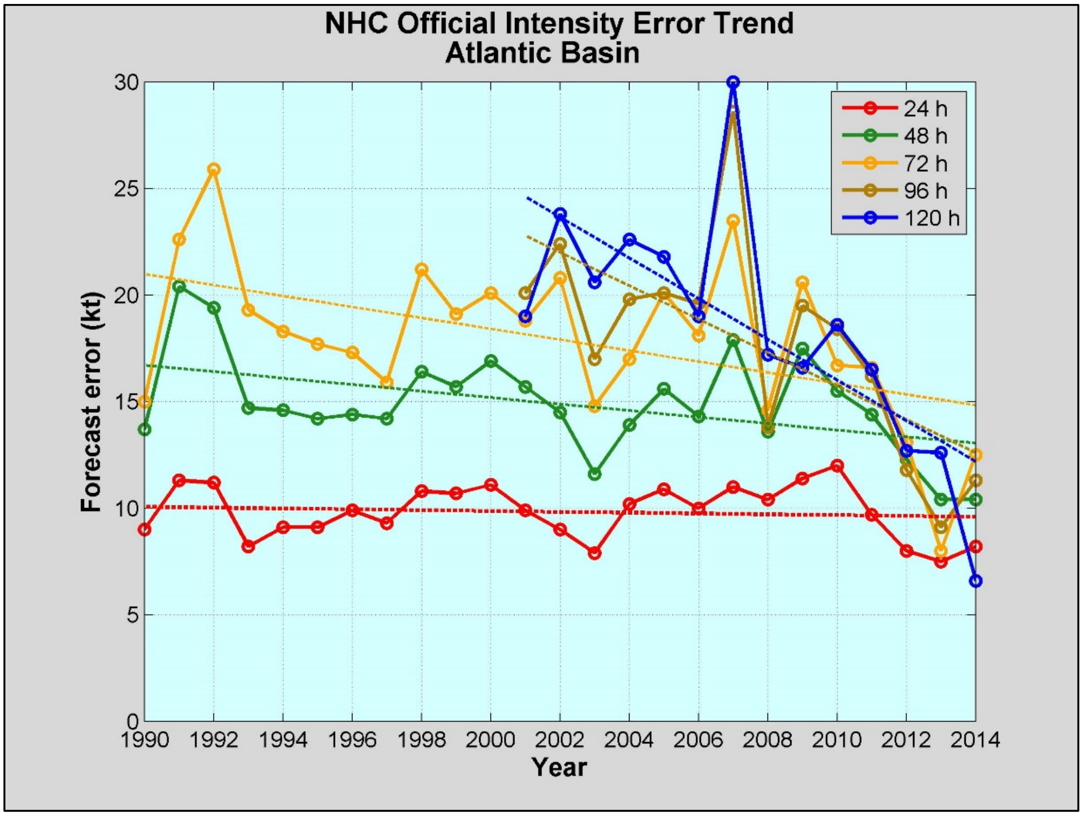


Figure 8. NHC official and Decay-SHIFOR5 (OCD5) Atlantic basin average intensity errors for 2014 (solid lines) and 2009-2013 (dashed lines).



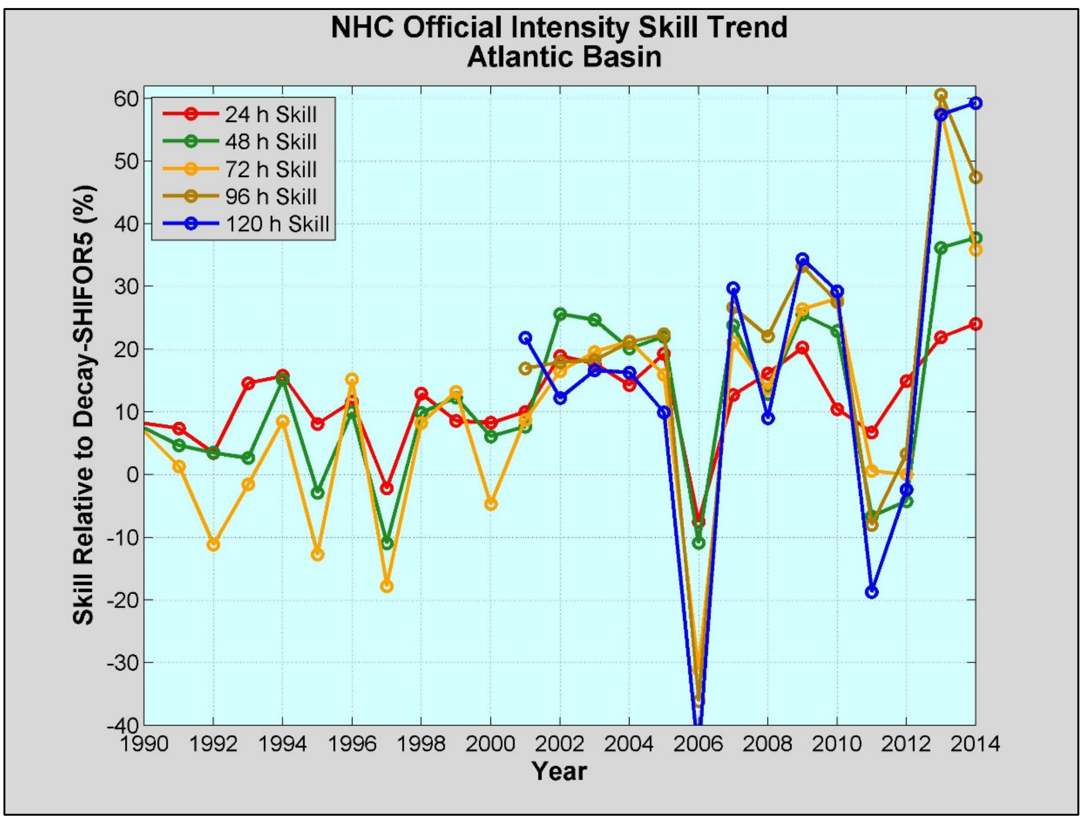


Figure 9. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the Atlantic basin.

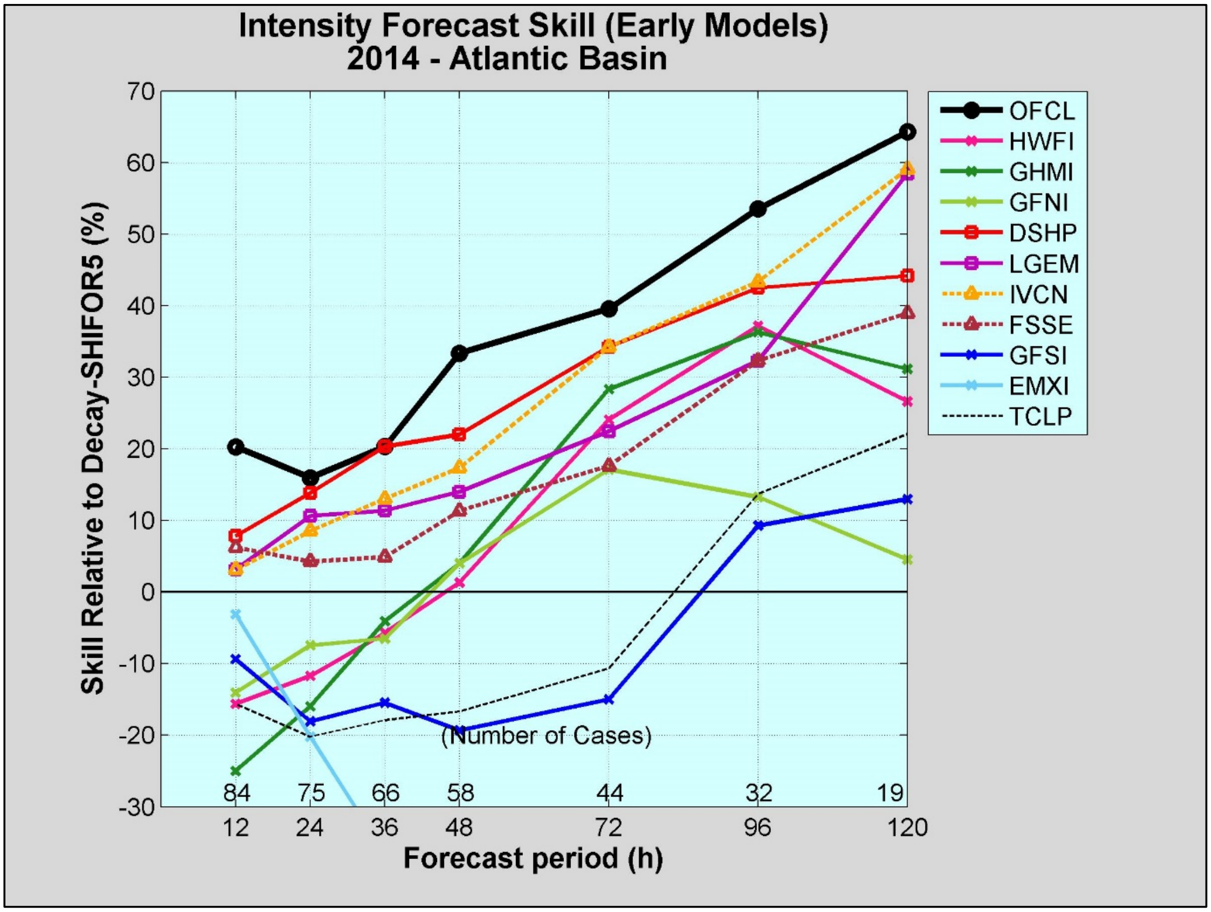
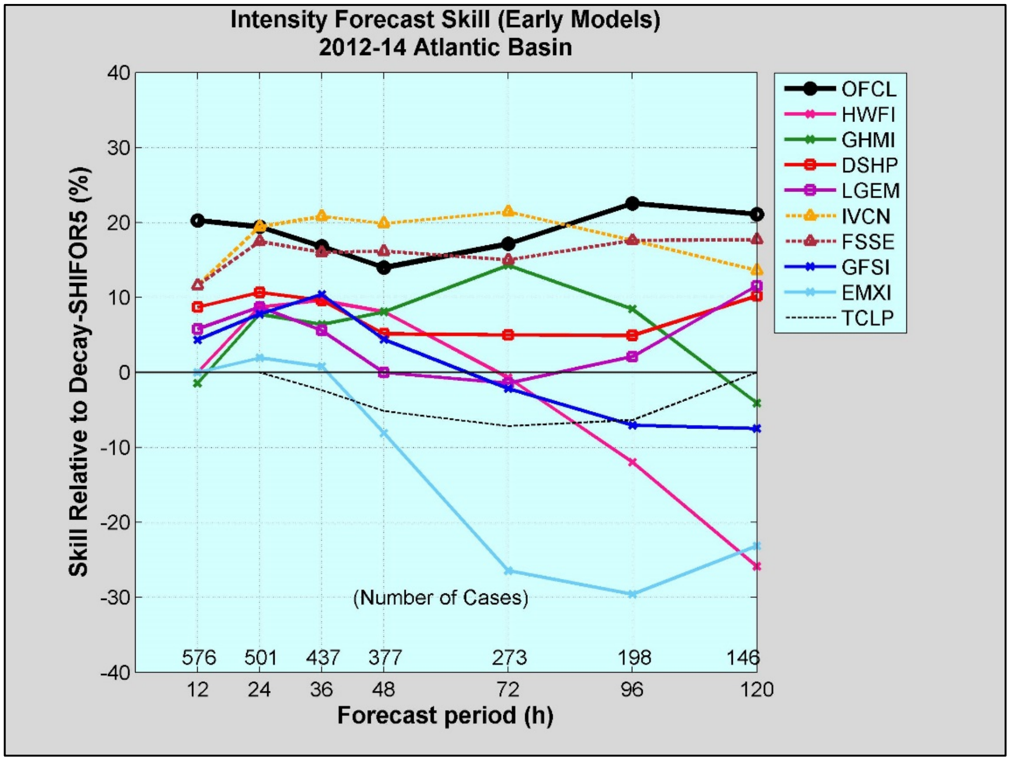


Figure 10. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2014.



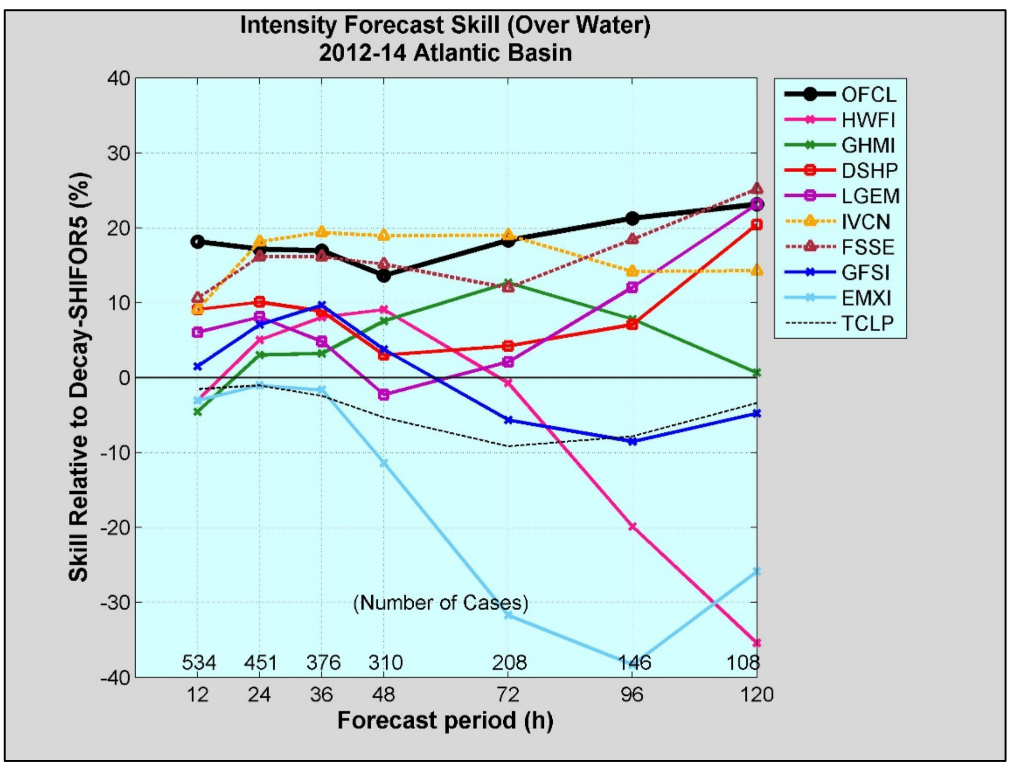


Figure 11. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2012-2014 (top) and for pre-landfall verifications only from 2012-14 (bottom). The pre-landfall verification sample is defined by excluding any portion of a model forecast that occurs after either the model forecast track or the verifying best track encounters land.

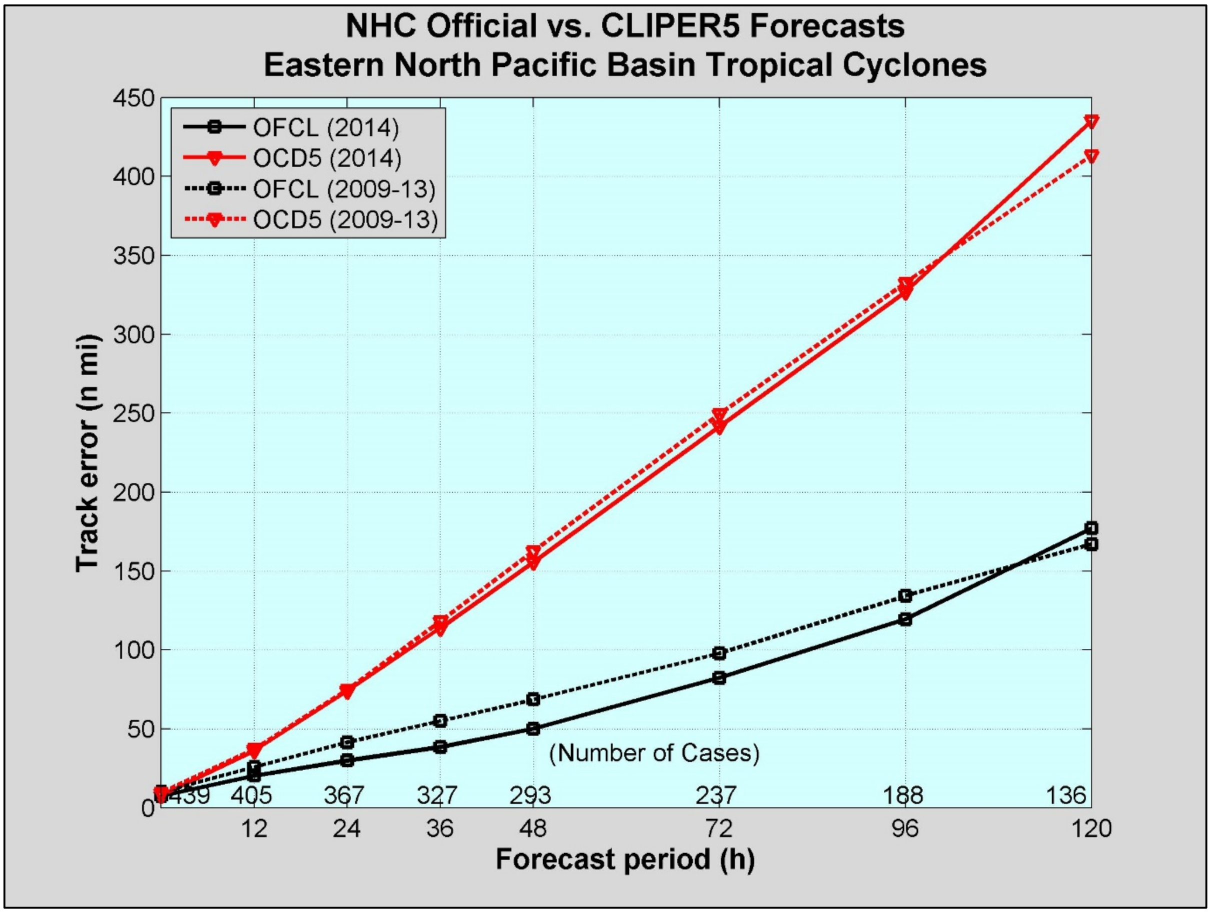


Figure 12. NHC official and CLIPER5 (OCD5) eastern North Pacific basin average track errors for 2014 (solid lines) and 2009-2013 (dashed lines).

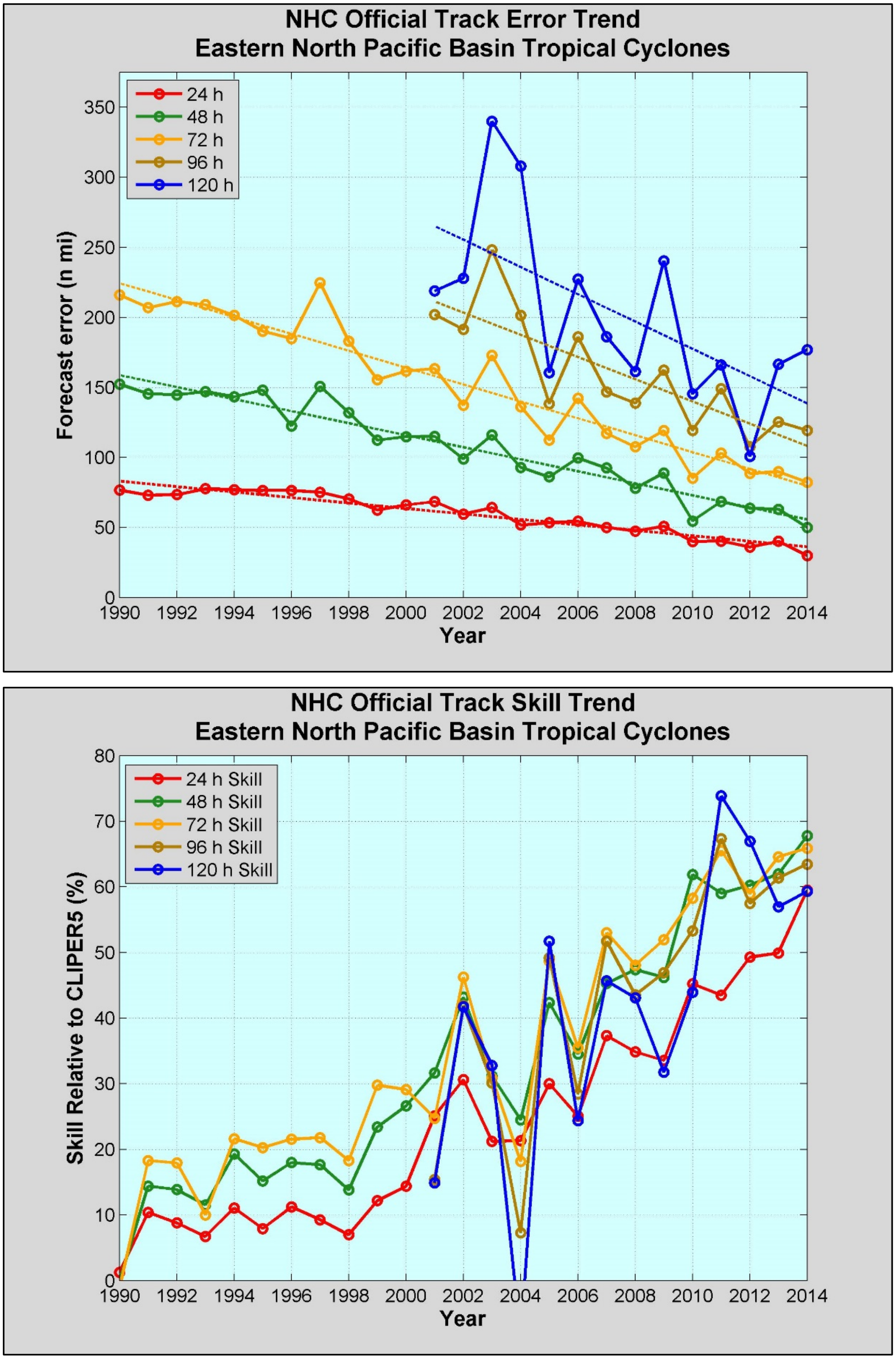


Figure 13. Recent trends in NHC official track forecast error (top) and skill (bottom) for the eastern North Pacific basin.

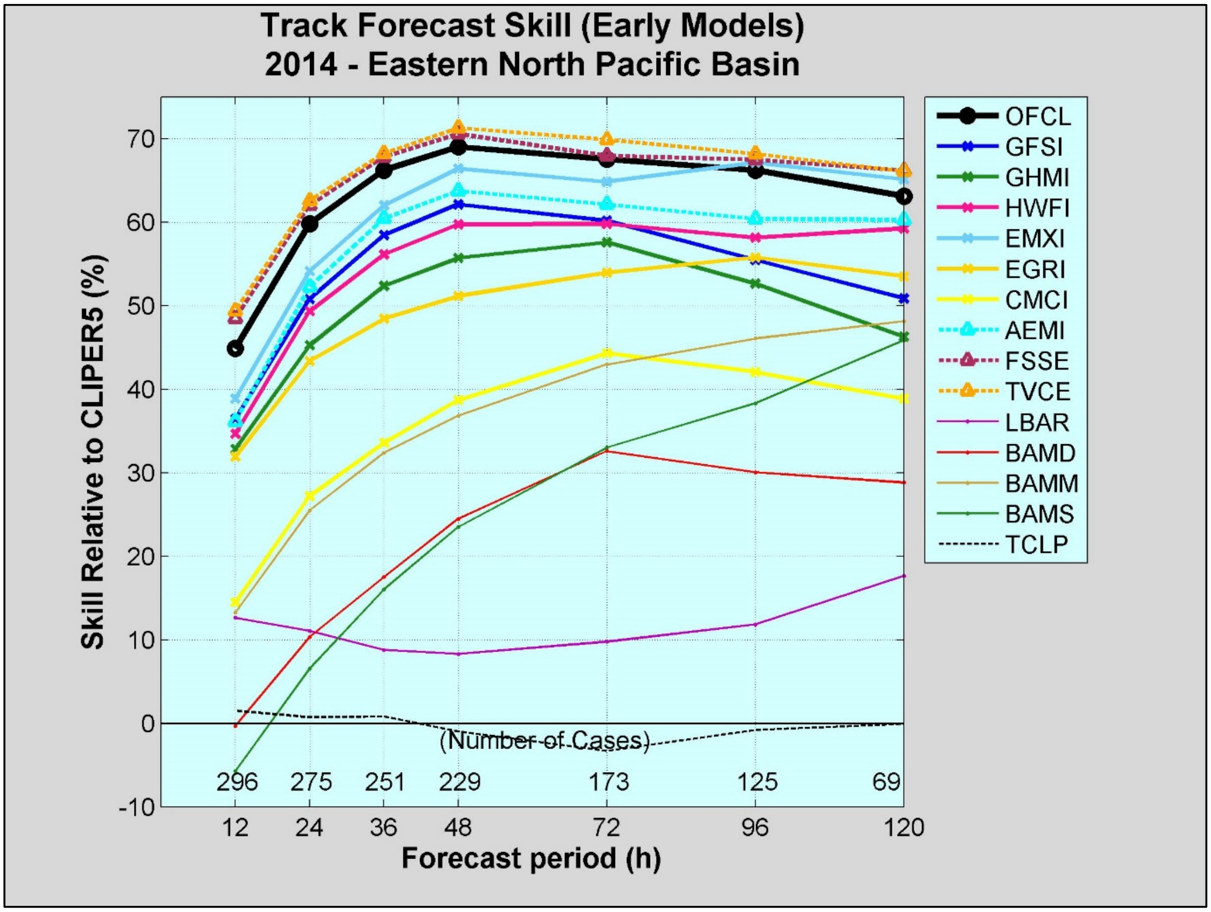


Figure 14. Homogenous comparison for selected eastern North Pacific early track models for 2014. This verification includes only those models that were available at least 2/3 of the time (see text).

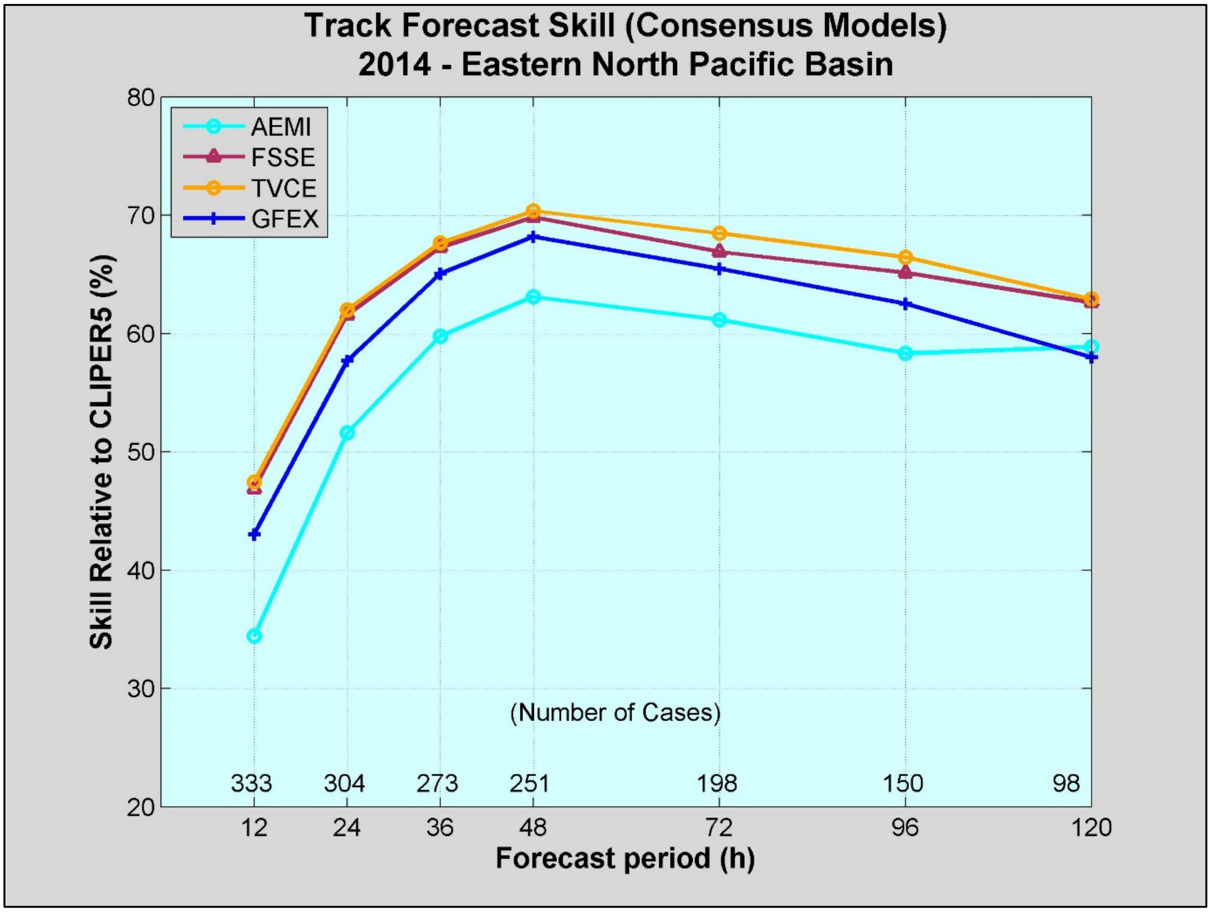


Figure 15. Homogenous comparison of the primary eastern North Pacific basin track consensus models for 2014.

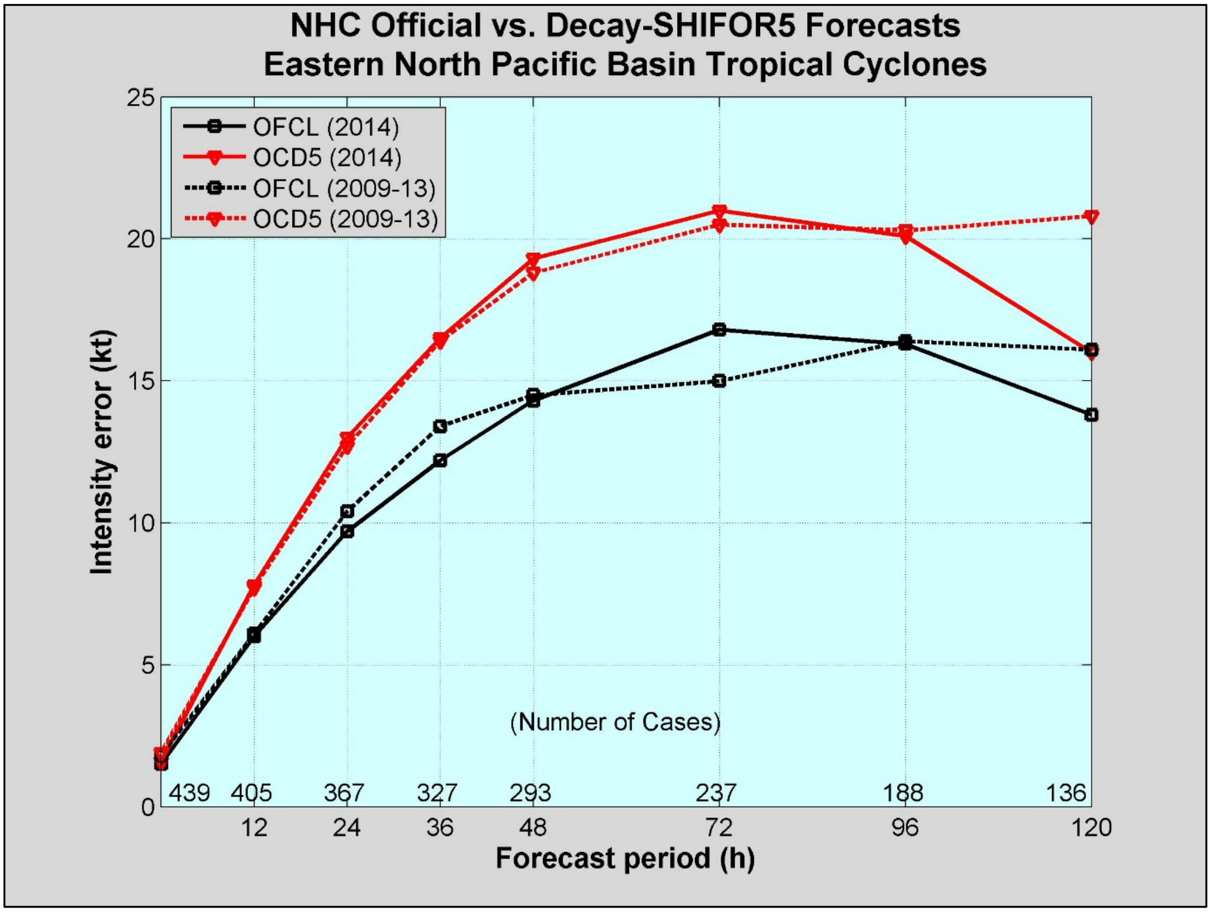
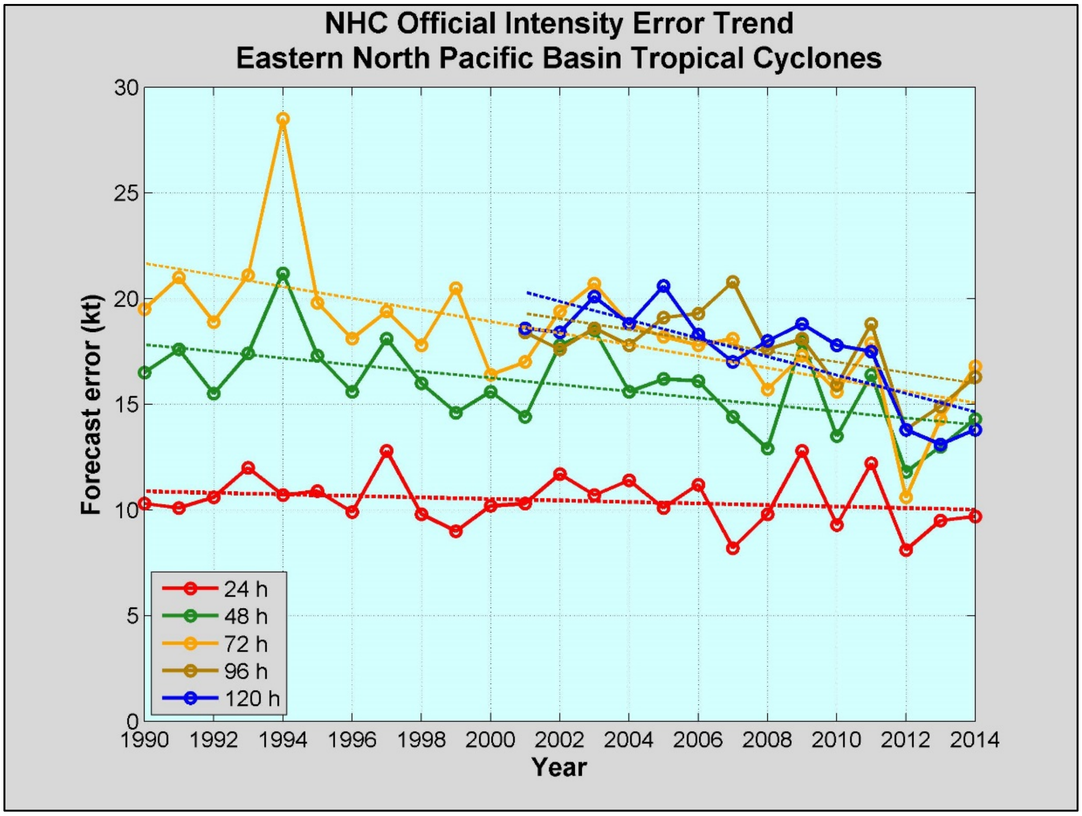


Figure 16. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2014 (solid lines) and 2009-2013 (dashed lines).



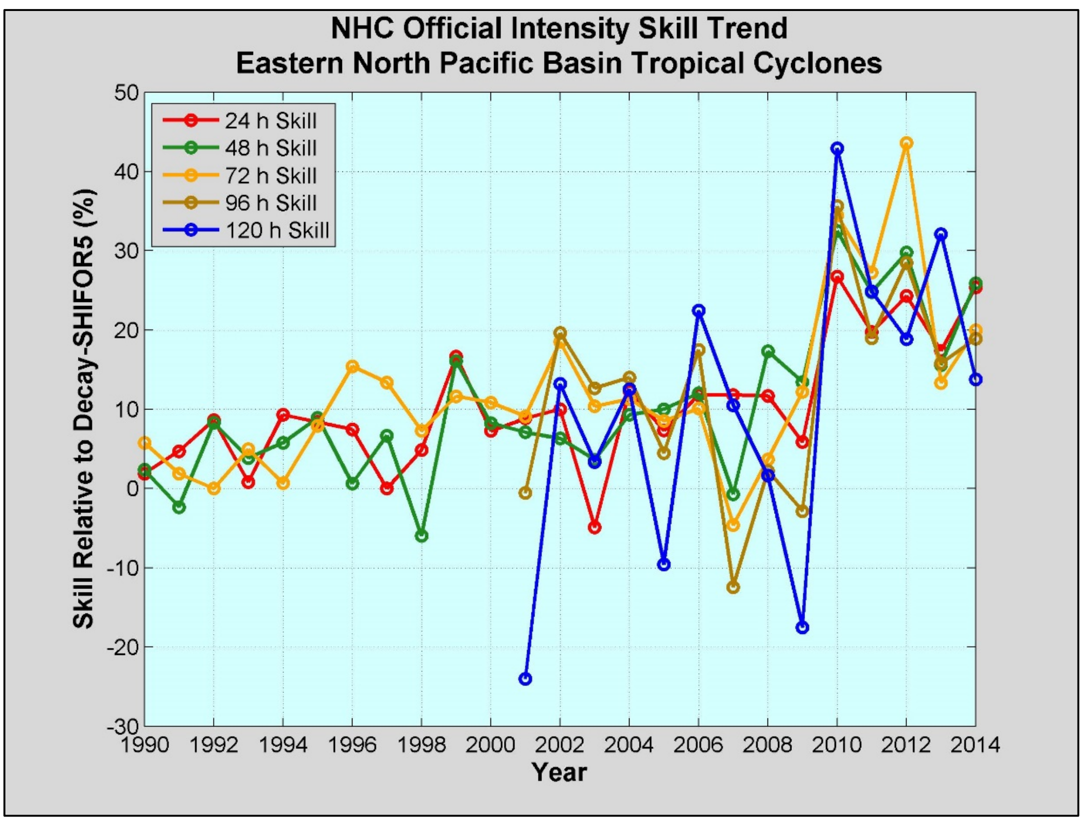


Figure 17. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

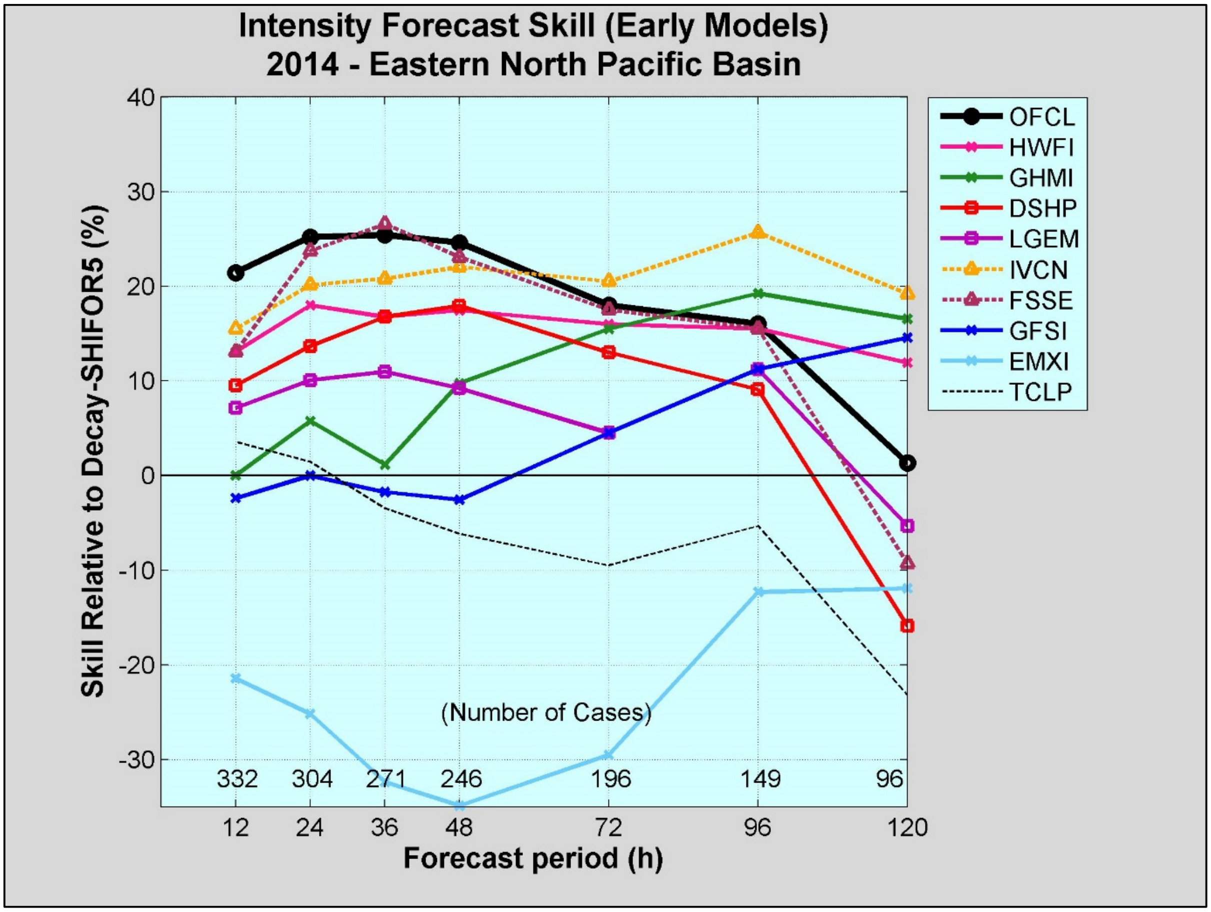
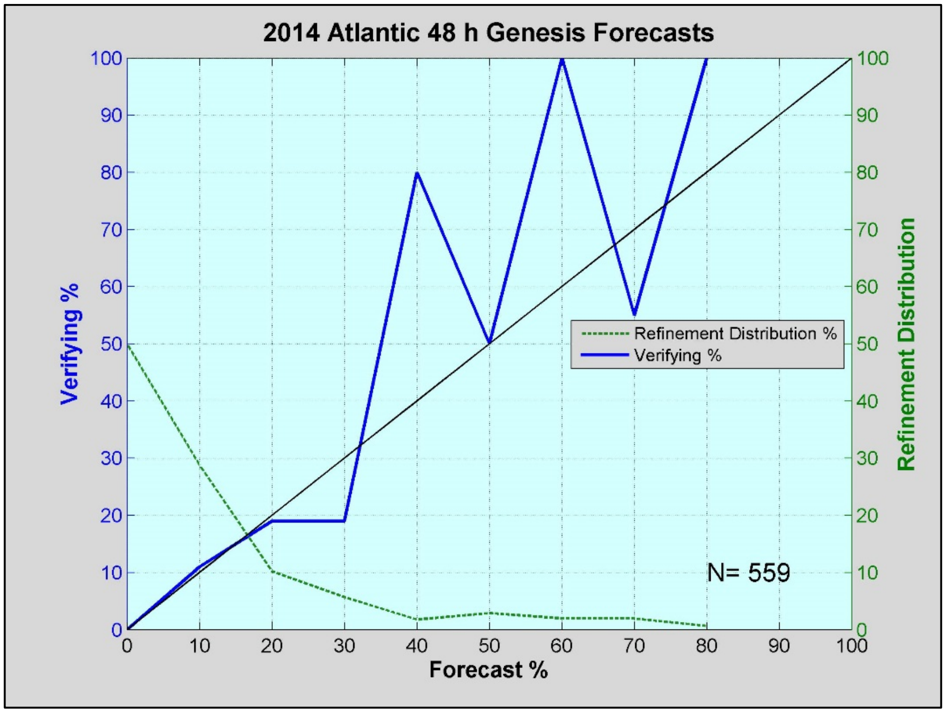


Figure 18. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2014.



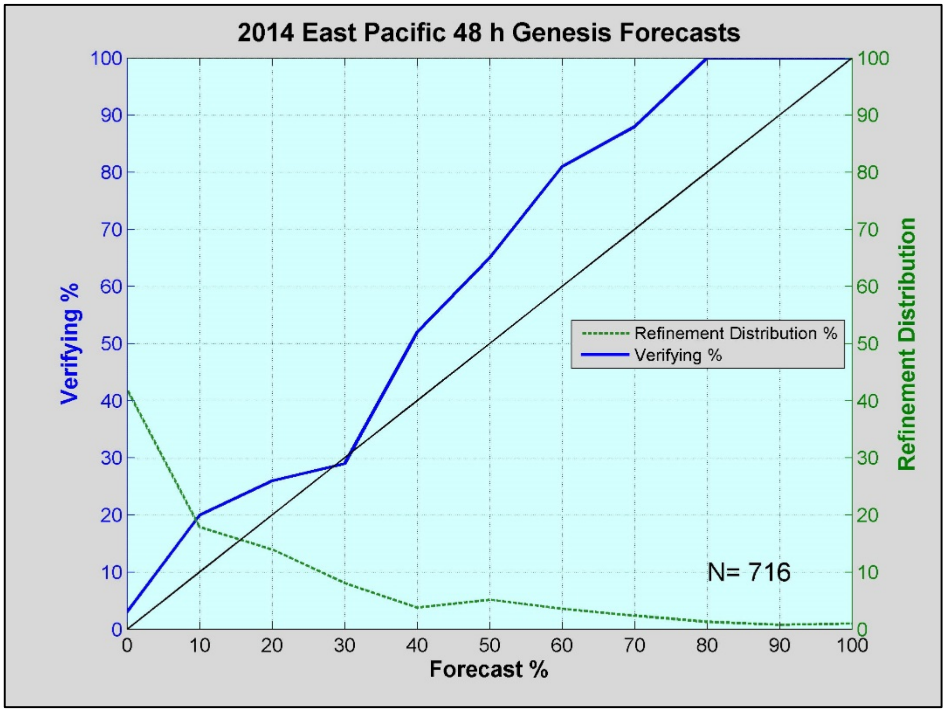
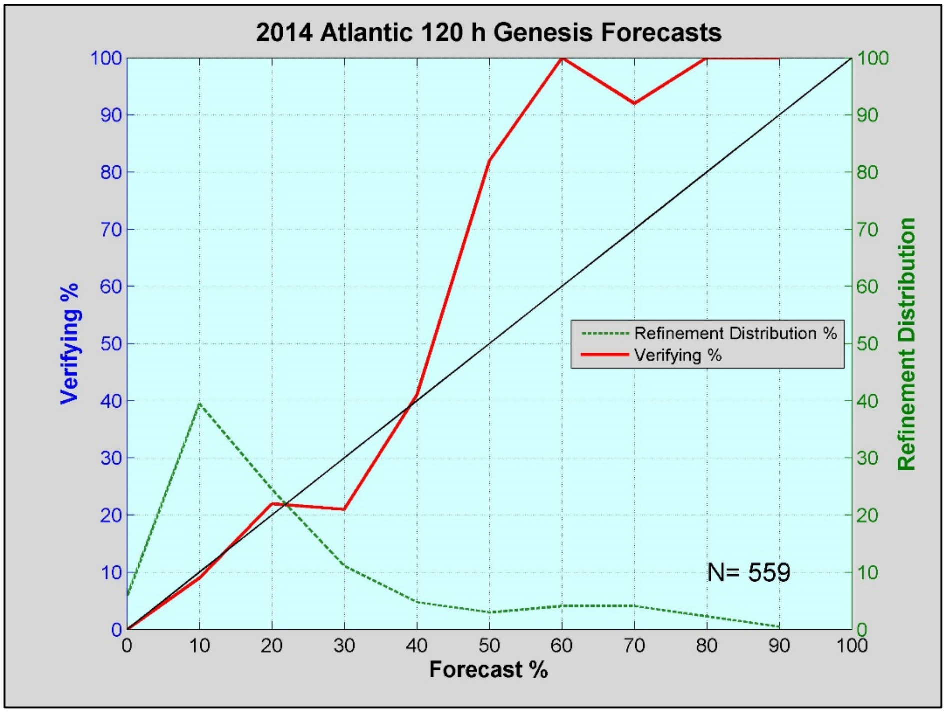


Figure 19. Reliability diagram for Atlantic (top) and eastern North Pacific (bottom) probabilistic tropical cyclogenesis 48-h forecasts for 2014. The solid lines indicate the relationship between the forecasts and verifying genesis percentages, with perfect reliability indicated by the thin diagonal black line. The dashed lines indicate how the forecasts were distributed among the possible forecast values.



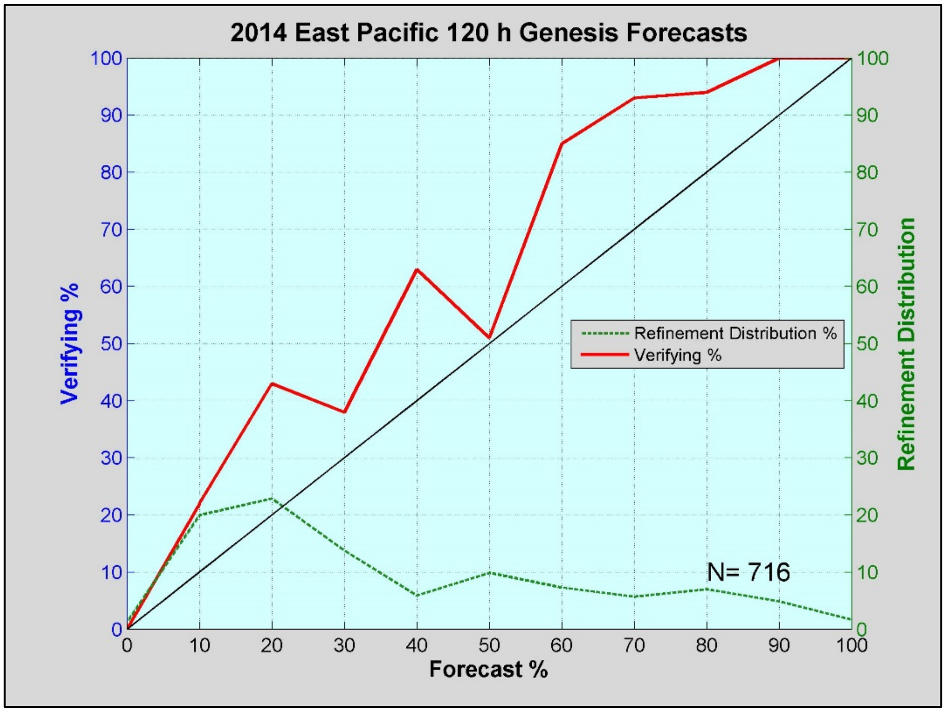


Figure 20. As described for Fig. 19, except for 120-h forecasts.

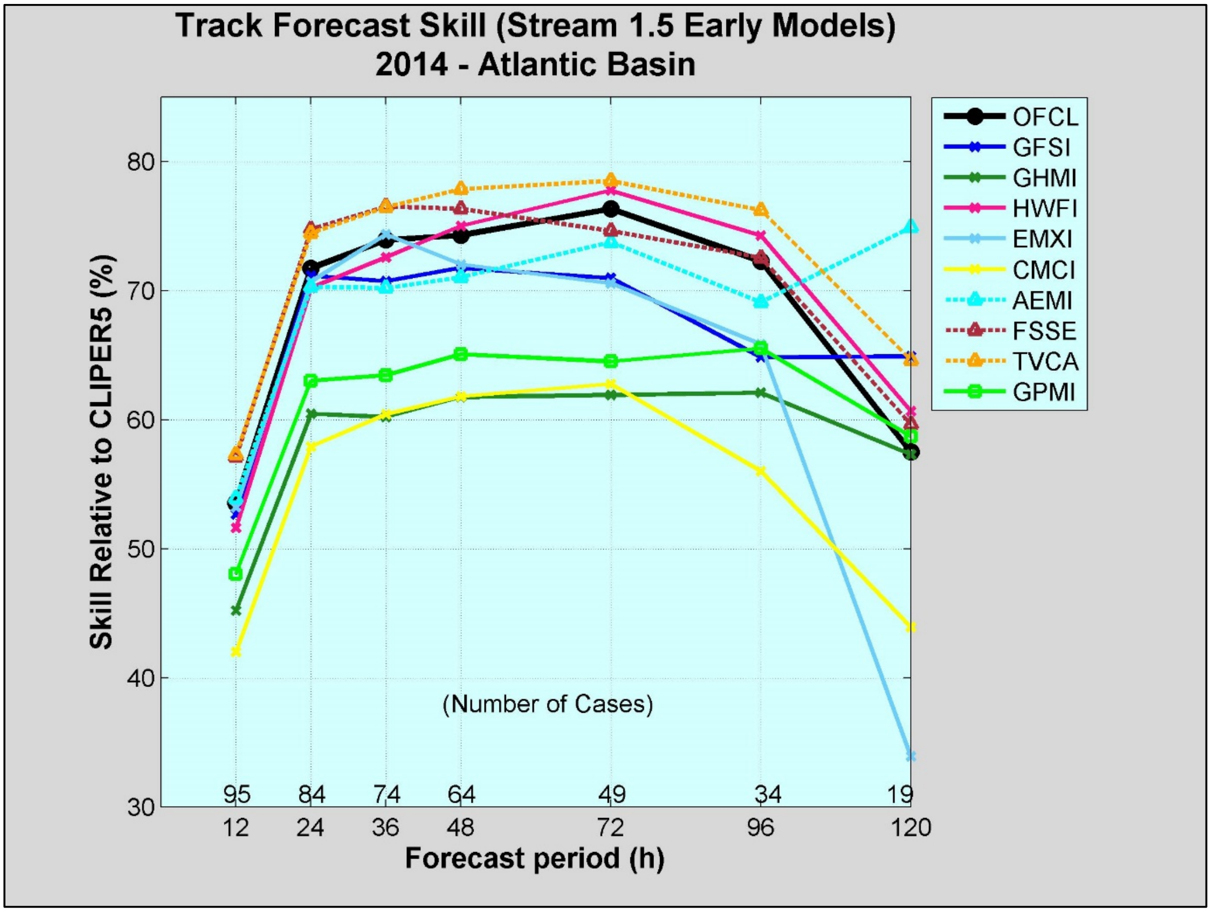


Figure 21. Homogeneous comparison of HFIP Stream 1.5 track models and selected operational models for 2014.

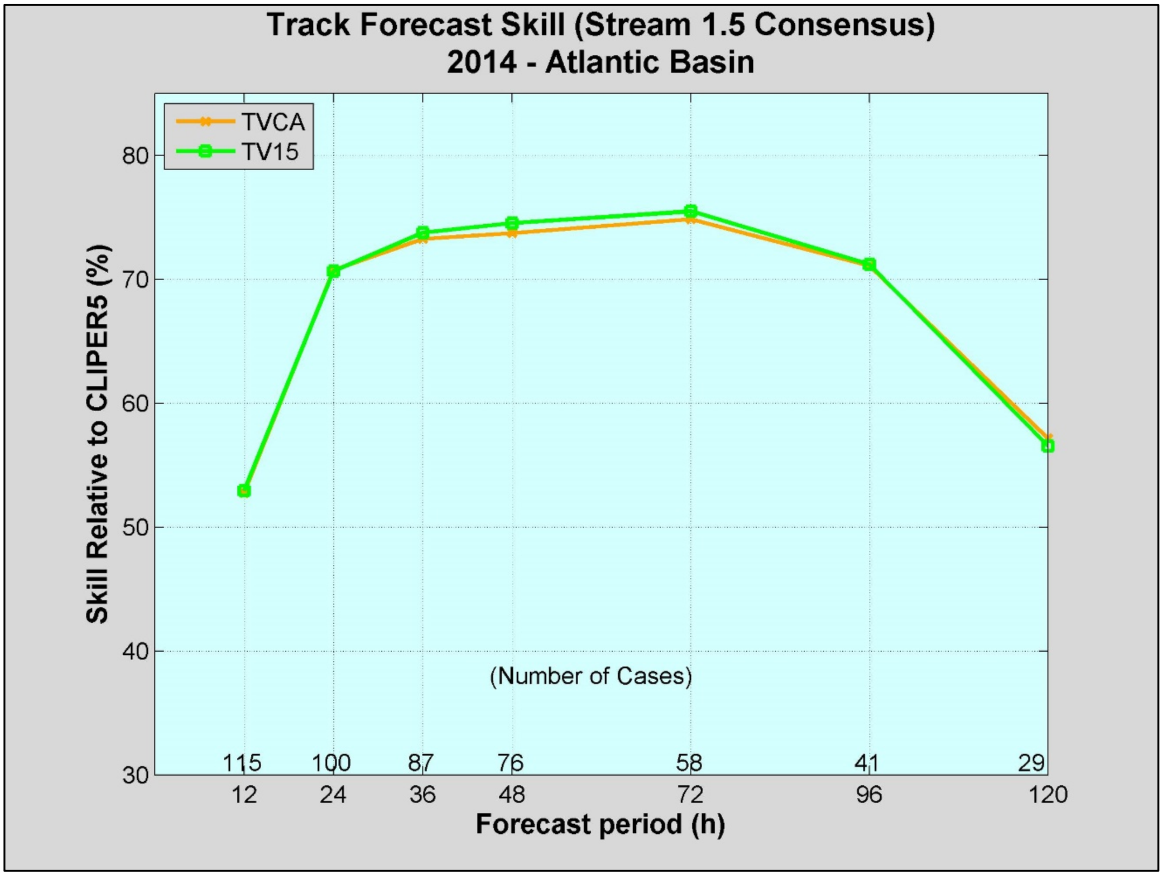
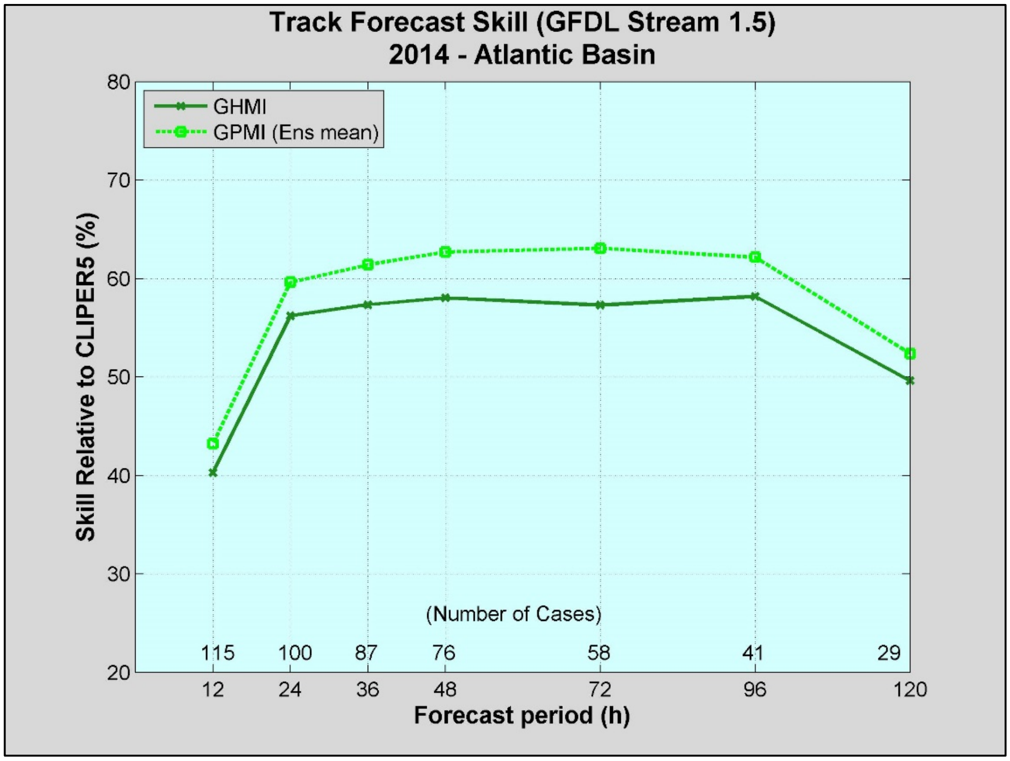


Figure 22. Comparison of Stream 1.5 consensus TV15 and TVCA.



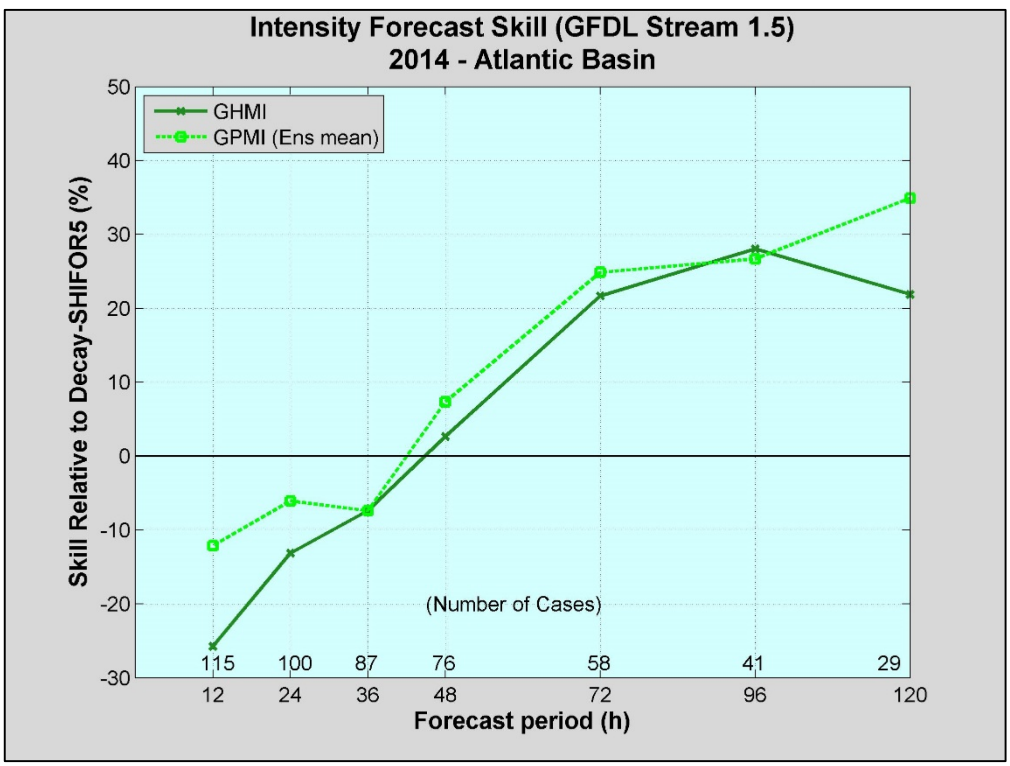


Figure 23. Homogeneous comparison of HFIP Stream 1.5 GFDL ensemble mean (GPMI) and GHMI for track (top) and intensity (bottom).

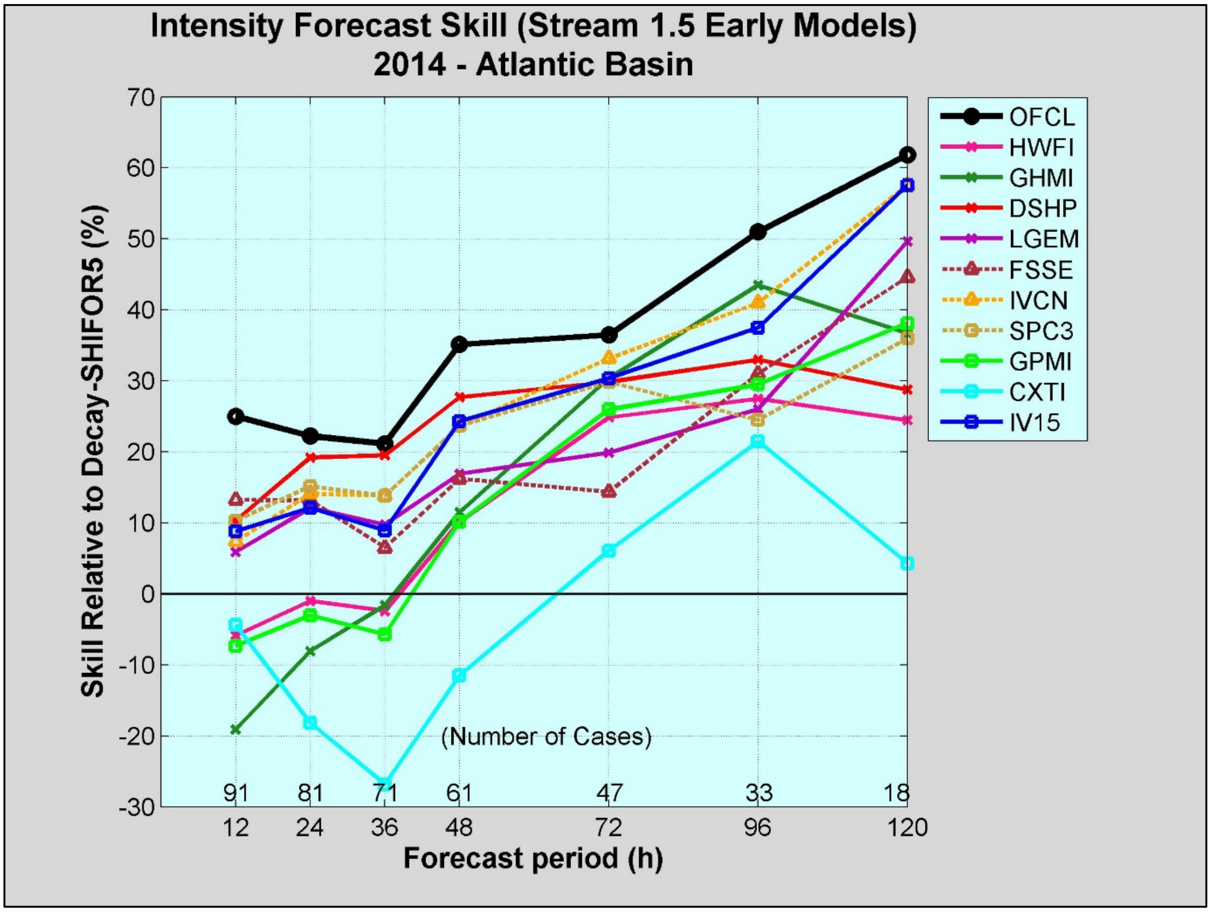


Figure 24. Homogeneous comparison of HFIP Stream 1.5 intensity models and selected operational models for 2014.

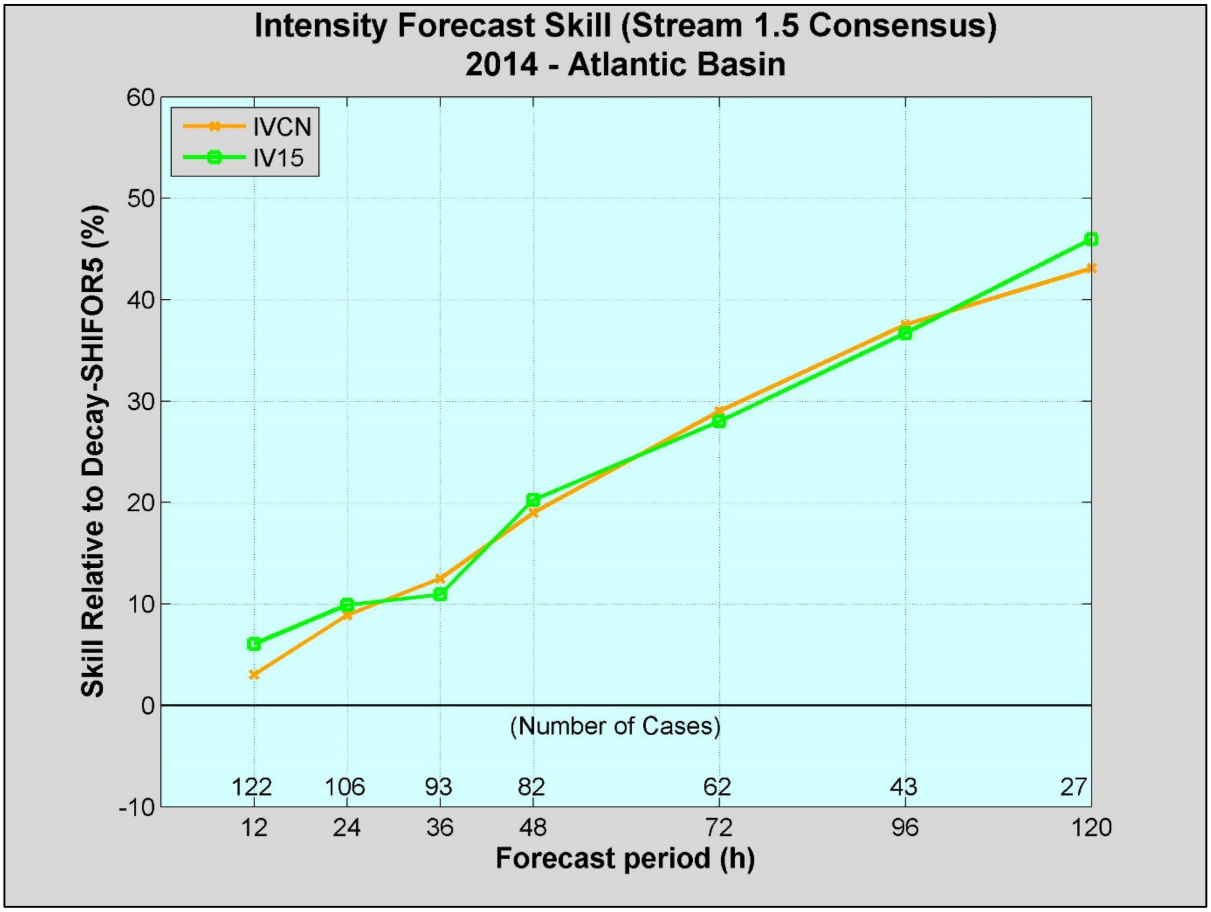


Figure 25. Comparison of Stream 1.5 consensus IV15 and IVCN.

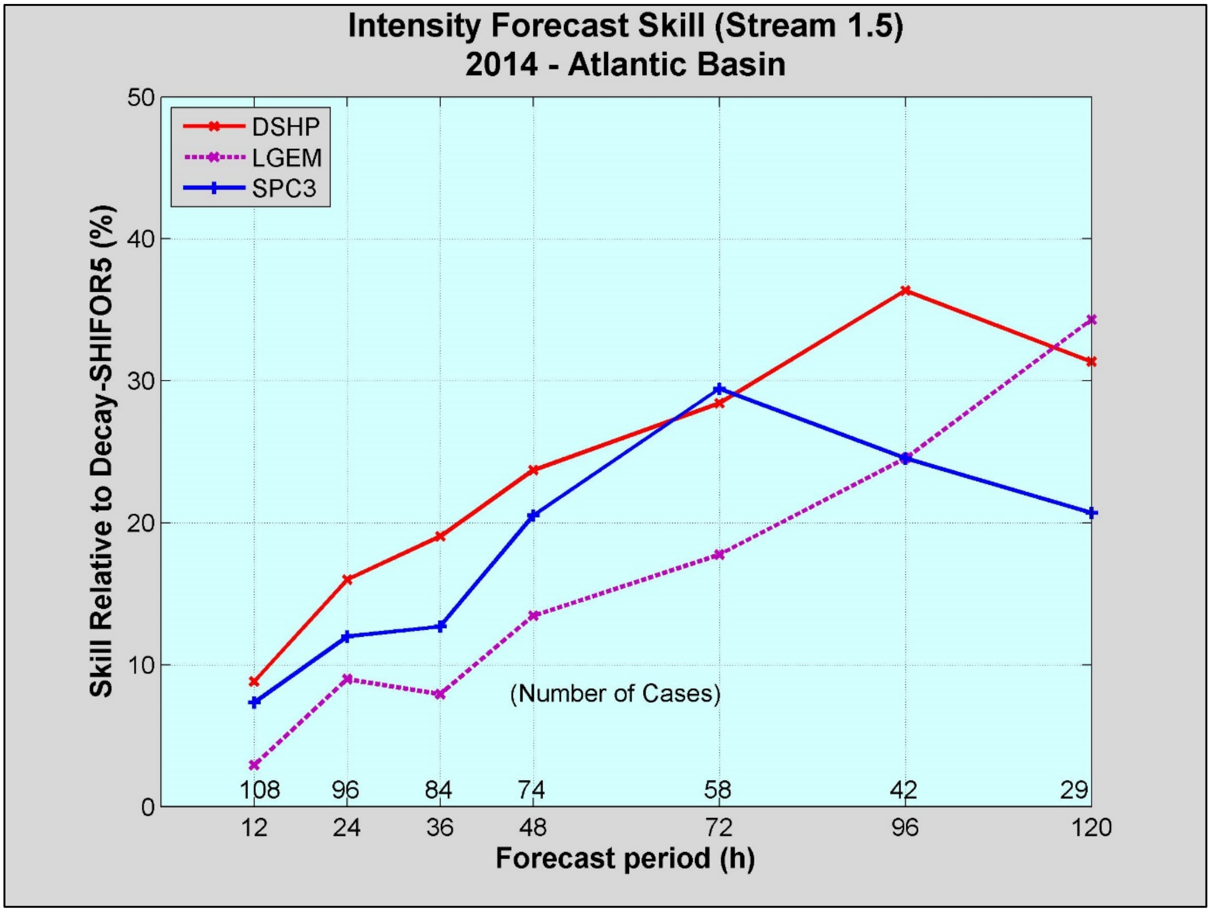


Figure 26. Homogeneous comparison of HFIP Stream 1.5 model SPC3 and operational models DSHP and LGEM.