2013 National Hurricane Center Forecast Verification Report

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

23 April 2014

Updated 5 May 2014 to correct Table 1

ABSTRACT

The 2013 Atlantic hurricane season had below-normal activity, with 196 official forecasts issued. The mean NHC official track forecast errors in the Atlantic basin were similar to or slightly larger than the previous 5-yr means from 12 to 96 h, but lower than the 5-yr mean at 120 h. A record for accuracy was set at 120 h. The official track forecasts were highly skillful and performed close to the best-performing models. The TVCA had the highest skill and beat the official forecast at several time periods. FSSE, AEMI, GFSI, and EMXI were also strong performers, and HWFI was the best model at the longer range. GHMI was in the second tier of models, while CMCI performed less well. The Government Performance and Results Act of 1993 (GPRA) track goal was not met.

Mean official intensity errors for the Atlantic basin in 2013 were below the 5-yr means at all lead times. Decay-SHIFOR errors in 2013 were also lower than their 5-yr means from 12 to 36 h, but larger than average after 36 h. The consensus models IVCN and FSSE were the best performers, and beat the official forecast at several time periods. HWFI was a good performer early, but its skill leveled off after 48 h. Conversely, DSHP and LGEM were the least skillful of the models early but the skill of those models increased at the longer forecast times. The GPRA intensity goal was met.

There were 288 official forecasts issued in the eastern North Pacific basin in 2013, although only 56 of these verified at 120 h. This level of forecast activity was near normal. No records for accuracy were set in 2013. 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 lower than the 5-yr means at all times. Decay-SHIFOR errors in 2013 were slightly lower than their 5-yr means at all forecast times, indicating the season's storms were a little easier to forecast than normal. The official forecast was better than all of the guidance early, but it was outperformed by the consensus aids IVCN and FSSE at 36 h and beyond. HWFI and DSHP were skillful throughout the forecast period and were the next best models, while LGEM and GHMI had little skill.

Quantitative probabilistic forecasts of tropical cyclogenesis were publically 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 2013 indicate that the 48-h probabilistic forecasts had a low (under-forecast) bias at the low and medium probabilities. The 120-h probabilities were slightly more reliable, though only a small sample exists at the high probabilities. A slight under-forecast bias was present at most ranges in the 48-h probabilistic forecasts in the eastern North Pacific basin. An under-forecast bias was observed at the low and middle probabilities for the 120-h forecasts.

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 2013 was improved somewhat from 2012. The FM9I was competitive with the top-tier dynamical models for track, and UW4I was a good-performing intensity model. In addition, SPC3 had more intensity skill than its individual components DSHP and LGEM at 36 h and beyond.

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'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 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. 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 6 February 2014 for the Atlantic basin, and on 12 February 2014 for the eastern North

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

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 discusses the Hurricane Forecast Improvement Project (HFIP) Stream 1.5 activities in 2013. Section 6 summarizes the key findings of the 2013 verification and previews anticipated changes for 2014.

2. Atlantic Basin

a. 2013 season overview – Track

Figure 1 and Table 2 present the results of the NHC official track forecast verification for the 2013 season, along with results averaged for the previous 5-yr period, 2008-2012. In 2013, the NHC issued 196 Atlantic basin tropical cyclone forecasts[11], a number well below the average over the previous 5 yr (350). Mean track errors ranged from 29 n mi at 12 h to 166 n mi at 96 h. It is seen that the mean official track forecast errors in 2013 were close to or slightly larger than the previous 5-yr mean from 12 to 96 h, and lower than the 5-yr mean at 120 h. The season's storms were about of average difficulty from 12 to 48 h, but more difficult to forecast than average from 72 to 120 h. The official track forecast errors set a record for accuracy at 120 h, although it should be noted that the sample of 120-h forecasts (19) was very small. The official track forecast vector biases were small through 48 h, but increased beyond that time and were westward to northwestward from 72 to 120 h (i.e., the official forecast tended to fall to the west or northwest of the verifying position). Further examination of the track errors shows that there was a slow along-track bias, and a leftward cross-track bias in 2013. Track forecast

[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.

skill in 2013 ranged from 40% at 12 h to 69% at 120 h (Table 2). The track errors in 2013 generally increased from the 2012 values, but over the past 15-20 yr, 24−72-h track forecast errors have been reduced by about 60% (Fig. 2). Track forecast error reductions of about 50% have occurred over the past 10 yr for the 96- and 120-h forecast periods. One possible reason for the increase in track errors in 2013 is the small amount of hurricanes that occurred that year. On average, the NHC track errors steadily decrease as the initial intensity of a cyclone increases (Fig. 3).

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 2013. 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

[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.

included in this comparison. The performance of the official forecast and the early track models in terms of skill are presented in Fig. 4. Note that the sample becomes small at 72 h and beyond. The figure shows that the official forecast was highly skillful, and near the accuracy of the best models. The consensus models TVCA, FSSE, and AEMI were the most skillful and beat the official forecast at a few time periods. GFSI and EMXI were the best-performing individual dynamical models in the short term in 2013, but HWFI was the best model at 96 and 120 h. It should be noted that the HWFI results shown here are a mixture of the 2012 and 2013 model versions, since the upgrade was made on 2 July 2013.

The GHMI made the second tier of the three-dimensional dynamical models, while CMCI performed less well. The more simplistic BAMS was a good performer in the 72 to 120 h forecast period, and beat many of the three-dimensional models at those times. An evaluation over the three years 2011-13 (Fig. 5) indicates that FSSE and TVCA are the best-performing models and have about equal skill from 12 to 72 h, with TVCA being a little more skillful at 96 and 120 h. The GFSI, EMXI, and AEMI are the next best models. The official forecasts are as good as the best-performing models from 12 to 72 h, and better than all of the guidance at 96 and 120 h. Vector biases of the guidance models for 2013 are given in Table 3b. The table shows that the official forecast had similar biases to TVCA and FSSE, but the biases were generally smaller than most of the model guidance. Among the typically high-performing models, the EMXI had a larger northwestward bias than the official forecast from 72 to 120 h. GFSI had one of the smallest biases among the models through 72 h.

A separate homogeneous verification of the primary consensus models for 2013 is shown in Fig. 6. The figure shows that the consensus models were equally skillful through 48 h, with TVCA most skillful after that. An examination of AEMI over the past couple of 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 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 2013, the GPRA goal was 81 n mi and the verification for this measure was 102.5 n mi. The reasons why the goal was not met in 2013 are: (1) the small sample makes the mean error susceptible to random fluctuations, (2) there was little tropical cyclone activity in the deep tropics, where forecasts tend to be easier and (3) storms were weaker than average, for which errors tend to be larger.

b. 2013 season overview – Intensity

Figure 7 and Table 4 present the results of the NHC official intensity forecast verification for the 2013 season, along with results averaged for the preceding 5-yr period. Mean forecast errors in 2013 ranged from about 5 kt at 12 h to about 13 kt at 120 h. These errors were below the 5-yr means at all forecast times, especially at 72 and 96 h where the 2013 errors were nearly 50% smaller than the 5-yr average. The official forecasts had a high bias at most forecast times. Decay-SHIFOR5 errors were below their 5-yr means from 12 to 36 h, but larger than average after that time. Figure 8 shows that there has been a notable decrease in the intensity errors over the past few years;

however, these recent improvements are likely 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 virtually no net change in error at the shorter leads, although forecasts during the current decade, on average, have been more skillful than those from the previous one. Although skill was at an all-time high in 2013, the small sample and few cases of rapid intensification make the high skill most likely non-representative. It is also worth noting that the DSHIFOR errors in 2013 were extraordinarily high, probably due to the unclimatological descent that limited development. The high DSHIFOR errors resulted in a high apparent skill, but this is probably non-representative also.

Table 5a presents a homogeneous verification for the official forecast and the primary early intensity models for 2013. Intensity biases are given in Table 5b, and forecast skill is presented in Fig. 9. The intensity models were much more skillful in 2013 than they were in recent years, although for the reasons noted above this likely overstates the true improvement in the numerical model guidance. The best performers were the consensus aids IVCN and FSSE, which both beat the official forecasts at several times. The LGEM and DSHP were the least skillful of the models in the short term, but skill increased at the latter forecast times. Conversely, HWFI was a strong performer through 48 h, but leveled off to the middle of the model pack with GHMI beyond that time. The top-performing global models, GFSI and EMXI, were included in the intensity verification for completeness, although they are typically not considered by forecasters. GFSI and EMXI were competitive with the better performing models early, but skill sharply decreased beyond 48 h. An inspection of the intensity biases (Table 5b) indicates that nearly all of the guidance suffered from a high bias in 2013, with LGEM and DSHP having a positive bias of about 14 kt at 48 h. GHMI and HWFI had the lowest biases through 72 h. An evaluation over the three years 2011-13 (Fig. 10) 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 slightly more skillful than 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 2013, the GPRA goal was 14 kt and the verification for this measure was 10.4 kt. This was the third time in six 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 Tropical Storm Jerry, which were two to three times larger than the long-term mean from 36 to 96 h. The primary reason for the unusually large errors for this storm was that the official forecasts did not anticipate the significant decrease in the cyclone's forward speed. Large track errors also occurred for Tropical Storm Karen in the Gulf of Mexico; these errors resulted from forecasts calling for Karen to be a deep system and move toward the north or northeast into the mid-latitudes, which did not occur.

Ingrid and Karen were the sources of the largest intensity errors. In the case of Ingrid, there were two sources of error in the intensity forecasts. The early forecasts for Ingrid did not predict that the cyclone would become a hurricane due to westerly shear, and later forecasts did not capture the slow decay that occurred after Ingrid reached its peak strength. In the case of Karen, the majority of the error was associated with a high bias in the official forecasts, which called for Karen to become a hurricane. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2013atlan.shtml.

3. Eastern North Pacific Basin

a. 2013 season overview – Track

The NHC track forecast verification for the 2013 season in the eastern North Pacific, along with results averaged for the previous 5-yr period is presented in Figure 11 and Table 7. There were 288 forecasts issued for the eastern Pacific basin in 2013, although only 56 of these verified at 120 h. This level of forecast activity is about average. Mean track errors ranged from 25 n mi at 12 h to 167 n mi at 120 h, and were lower than the 5-yr means at all times except for 120 h. No records were set for forecast accuracy in 2013. CLIPER5 errors were slightly larger than their long-term means at most forecast times, implying that the season's storms were a little more difficult to forecast than average. Little bias was present in the official forecast through 48 h, with a moderate northeastward bias present from 72 to 120 h.

Figure 12 shows recent trends in track forecast accuracy and skill for the eastern North Pacific. Errors have been reduced by roughly 45-60% 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 2013 set a new record high at 24 h and was near all-time highs at the 48-, 72-, and 96-h forecast times.

Table 8a presents a homogeneous verification for the official forecast and the early track models for 2013, with vector biases of the guidance models given in Table 8b. Skill comparisons of selected models are shown in Fig. 13. Note that the sample becomes rather small by 120 h (only 22 cases). FSSE and TVCE were the only two models that consistently outperformed the official forecast. EMXI was the best individual model, but it had about a little less skill than TVCE/FSSE. That model also beat the official forecast at 96 and 120 h. GSFI, HWFI, and AEMI made up the second tier of models, with GHMI not very far behind. CMCI was a poor performer and had similar skill to the simple BAMM and BAMS models.

A separate verification of the primary consensus aids is given in Figure 14. TVCE and FSSE had comparable skill from 12 to 72 h, but TVCE trailed FSSE slightly at the longer forecast times. The skill of AEMI was about 10-15 % smaller than that of FSSE and TVCE.

b. 2013 season overview – Intensity

Figure 15 and Table 9 present the results of the NHC eastern North Pacific intensity forecast verification for the 2013 season, along with results averaged for the preceding 5-yr period. Mean forecast errors were 5 kt at 12 h and increased to 15 kt at 96 h. The errors were lower than the 5-yr means at all times, by up to 25%. The Decay-SHIFOR5 forecast errors were also lower than their 5-yr means (by up to 19%); this implies that forecast difficulty in 2013 was lower than normal. A review of error and skill trends (Fig. 16) indicates that the intensity errors have decreased slightly over the past 15-20 yr, especially at 48 h and beyond. Forecast skill generally decreased in 2013. Intensity forecast biases in 2013 were slightly high from 36 to 96 h.

Figure 17 and Table 10a present a homogeneous verification for the primary early intensity models for 2013. Forecast biases are given in Table 10b. The official forecasts performed better than all of the guidance at 12 and 24 h, but was outperformed by the consensus aids IVCN and FSSE at 36 h and beyond. HWFI and DSHP were skillful throughout the forecast period and made up the second tier of models; the former beat the official forecast at couple of forecast leads. LGEM and GHMI had little skill in 2013. GFSI and especially EMXI were not competitive with the standard intensity guidance in this basin. The model biases were not very significant; though LGEM suffered from a slight high bias at all forecast times and HWFI had a consistent low bias.

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/2013epac.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 2013 are given in Table 12 and illustrated in Fig. 18. In the Atlantic basin, a total of 641 48-h genesis forecasts were made. These forecasts exhibited a slight under-forecast (low) bias in the low and medium probabilities. The sample is too small at the high probabilities to draw meaningful conclusions. In the eastern Pacific, a total of 527 48-h genesis forecasts were made. The forecasts had a slight low bias at all probabilities, but overall were well calibrated and improved from recent years. 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.

Verification of the 120-h outlook for the Atlantic and eastern North Pacific basins for 2013 are given in Table 13 and illustrated in Fig. 18. In the Atlantic basin, 528 120-h genesis forecasts were made. Although a fair amount of noise exists at the middle and high probabilities, due in part to a relatively small sample, the forecasts are fairly reliable. In the eastern North Pacific, a slight low bias exists at the low and middle probabilities.

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".

Seven models/modeling systems were provided to NHC in 2013 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 FM9I track forecasts but not the FM9I intensity forecasts. Two HFIP Stream 1.5 consensus aids were constructed: the track consensus TV15 comprised the operational models GFSI, EGRI, GHMI, HWFI, EMXI and the Stream 1.5 model APSI, while the intensity consensus IV15 comprised the operational models DSHP, LGEM, GHMI, HWFI and the Stream 1.5 models APSI, UW4I, and CXTI.

Figure 19 presents a homogeneous verification of the primary operational models against the Stream 1.5 track models FM9I and GPMI. The figure shows that in 2013 the FM9I was a very good performer and was competitive with or slightly better than the best performing operational models. GPMI (GFDL ensemble mean) was also a good performer through 72 h, but it had notably less skill than the best models for a small sample at 96 and 120 h. Figure 20 shows that there was very little impact from adding the Stream 1.5 models to the track consensus through 72 h, and then a slight positive effect (for a tiny sample) at 96 h. The reason why the sample was so small is because of the limited availability of APSI.

Figure 21 presents the track and intensity forecast skill of GHMI and the Stream 1.5 GPMI. Regarding track, GPMI had a little more skill than the deterministic run. For intensity, GPMI was slightly more skillful than GHMI through 48 h, after that the skill values are similar. Figure 22 presents the track and intensity forecast skill of HWFI and the Stream 1.5 HWRF ensemble mean (HWMI). For track, HWMI was better than the deterministic run at nearly all forecast times, with the largest improvement noted in the short term. Regarding intensity, HWMI had a little more skill than HWFI through 48 h, and the skill was about the same beyond that forecast time.

Intensity results are shown in Fig. 23, for a sample that excludes the APSI runs due to limited availability. The Stream 1.5 model UW4I was a very good performer throughout and its skill was near the high end of the models. CXTI had little skill early, but skill increased beyond 36 h bringing it into the middle of the model pack. The impact of the Stream 1.5 models to the intensity consensus was slightly positive (Fig. 24).

Figure 25 presents the intensity forecast skill of DSHP, LGEM, and the Stream 1.5 model SPC3. SPC3 has more skill than DSHP and LGEM at 36 h and beyond. This result is not surprising, given that it represents an intelligent consensus of DSHP and LGEM.

Overall the performance of the Stream 1.5 models in 2013 was improved from 2012. The FM9I did show about equivalent skill to the high performing operational models for track. For intensity, UW4I was a good performer, and the dynamical-statistical consensus SPC3 did have more skill than its individual components.

6. Looking Ahead to 2014

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 2014 for the Atlantic and eastern North Pacific basins (based on error distributions for 2009-13) are given in Table 15. In the Atlantic basin, the cone circles will be unchanged through 48 h, and then slightly smaller (by up to 4 %) from 72 to 120 h. In the eastern Pacific basin, the cone circles will be up to 6 % smaller than they were in 2013.

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 2014 is unchanged from 2013, and is shown in Table 16. NVGI and GFNI were examined, but they did not improve upon the consensus and therefore will not be added to the track or intensity consensus models in 2014.

Acknowledgments:

The authors gratefully acknowledge Michael Brennan and Monica Bozeman 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.

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
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 2013 season.

Table 2. Homogenous comparison of official and CLIPER5 track forecast errors in the Atlantic basin for the 2013 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
2013 mean OFCL error (n mi)	28.5	49.4	71.6	102.5	141.3	165.5	165.2
2013 mean CLIPER5 error (n mi)	47.6	101.3	161.0	224.5	383.0	498.5	531.0
2013 mean OFCL skill relative to CLIPER5 (%)	40.1	51.2	55.5	54.3	63.1	66.8	68.9
2013 mean OFCL bias vector (°/n mi)	082/001	215/002	227/007	251/015	274/055	279/081	328/082
2013 number of cases	162	131	107	83	40	22	19
2008-2012 mean OFCL error (n mi)	28.6	45.7	62.0	78.4	116.5	158.6	205.1
2008-2012 mean CLIPER5 error (n mi)	47.5	99.3	161.0	224.0	330.5	419.4	495.9
2008-2012 mean OFCL skill relative to CLIPER5 (%)	39.8	54.0	61.5	65.0	64.8	62.2	58.6
2008-2012 mean OFCL bias vector (°/n mi)	320/003	309/007	296/010	294/013	273/014	340/007	026/022
2008-2012 number of cases	1513	1345	1186	1049	844	668	531
2013 OFCL error relative to 2008-2012 mean (%)	-0.3	8.0	15.5	30.6	21.3	4.4	-19.5
2013 CLIPER5 error relative to 2008-2012 mean (%)	0.2	0.2	0.0	-0.2	15.9	18.9	7.1

Table 3a. Homogenous comparison of Atlantic basin early track guidance model errors (n mi) for 2013. 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	29.2	51.6	73.8	104.2	109.7	115.3	166.0
OCD5	49.4	102.1	170.4	255.8	442.1	540.0	518.1
GFSI	31.5	53.4	78.7	105.7	121.8	150.1	191.5
GHMI	36.5	64.3	89.0	117.6	152.5	182.7	250.7
HWFI	35.4	61.8	88.5	115.7	131.7	125.8	**136.9**
EMXI	31.6	53.0	**72.2**	105.0	135.8	177.1	212.5
CMCI	40.0	73.3	109.8	151.2	200.3	305.6	370.5
AEMI	31.0	52.0	**71.0**	**93.2**	119.9	141.0	197.1
FSSE	29.5	**48.4**	**68.5**	**96.3**	117.7	145.4	175.1
TVCA	29.6	**49.8**	**68.7**	**95.1**	112.6	117.0	**135.7**
LBAR	42.4	78.1	115.9	154.8	243.4	367.9	392.6
BAMD	51.1	90.3	135.8	189.7	301.3	376.6	591.5
BAMM	40.6	68.1	103.7	150.1	217.4	215.0	282.4
BAMS	44.2	74.8	100.9	127.4	174.9	135.3	**162.6**
TCLP	48.7	103.3	172.6	255.4	397.1	519.7	497.3
# Cases	112	89	74	55	19	11	10

Table 3b. Homogenous comparison of Atlantic basin early track guidance model bias vectors (º/n mi) for 2013.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	093/002	100/002	346/003	316/009	304/043	328/053	314/081
OCD5	161/005	173/009	176/013	154/018	040/059	042/302	042/404
GFSI	102/004	121/010	108/013	087/010	323/016	021/038	017/025
GHMI	136/003	343/005	326/018	314/035	313/068	347/151	008/175
HWFI	124/005	167/011	193/015	199/028	230/032	226/013	251/056
EMXI	045/003	038/005	034/007	358/003	316/057	326/111	303/152
CMCI	331/004	293/014	285/026	281/041	282/113	303/153	315/203
AEMI	158/001	193/004	248/004	281/009	312/024	008/077	354/083
FSSE	081/004	091/005	051/006	019/005	325/037	350/077	338/065
TVCA	098/002	090/001	322/005	298/012	303/044	337/085	332/097
LBAR	088/014	065/017	049/016	035/017	049/097	072/308	081/390
BAMD	076/018	068/027	058/045	059/060	062/112	087/353	104/545
BAMM	315/005	303/013	316/021	315/027	306/049	036/140	044/132
BAMS	287/022	281/046	275/064	270/077	240/100	053/007	313/009
TCLP	213/008	227/021	227/039	222/052	150/018	073/248	071/324
# Cases	112	89	74	55	19	11	10

Table 4. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the Atlantic basin for the 2013 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
2013 mean OFCL error (kt)	5.3	7.5	9.3	10.4	8.0	9.1	12.6
2013 mean Decay-SHIFOR5 error (kt)	6.3	9.6	13.3	16.3	19.0	23.1	29.6
2013 mean OFCL skill relative to Decay-SHIFOR5 (%)	15.9	17.7	30.1	36.2	57.9	60.6	57.4
2013 OFCL bias (kt)	0.8	2.4	4.6	6.0	3.3	7.7	10.5
2013 number of cases	162	131	107	83	40	22	19
2008-12 mean OFCL error (kt)	6.6	10.1	12.2	14.0	15.5	15.2	16.3
2008-12 mean Decay-SHIFOR5 error (kt)	7.8	11.6	14.1	15.6	17.9	18.2	18.1
2008-12 mean OFCL skill relative to Decay-SHIFOR5 (%)	15.4	12.9	13.5	10.3	13.4	16.5	10.0
2008-12 OFCL bias (kt)	-0.3	0.4	0.8	1.3	1.7	1.3	1.5
2008-12 number of cases	1513	1345	1186	1049	844	668	531
2013 OFCL error relative to 2008-12 mean (%)	-19.7	-25.7	-23.8	-25.7	-48.4	-40.1	-22.7
2013 Decay-SHIFOR5 error relative to 2008-12 mean (%)	-19.2	-17.2	-5.7	4.5	6.1	26.9	63.5

Table 5a. Homogenous comparison of selected Atlantic basin early intensity guidance model errors (kt) for 2013. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	5.7	8.1	9.9	10.8	7.9	9.6	12.5
OCD5	6.9	11.4	15.5	18.3	21.5	24.2	31.5
HWFI	6.0	**7.3**	**8.8**	**9.4**	12.4	13.9	17.7
GHMI	6.5	8.4	10.3	10.9	12.6	13.5	23.9
DSHP	6.0	8.7	11.8	14.9	14.4	13.4	**12.2**
LGEM	6.3	9.2	12.2	14.9	12.0	11.2	16.3
IVCN	5.9	**7.4**	**9.6**	**10.7**	9.1	**9.4**	15.8
FSSE	5.7	**7.0**	**8.9**	**8.9**	8.8	**6.3**	**10.6**
GFSI	6.0	8.1	10.2	11.8	17.2	22.4	27.9
EMXI	6. 0	**7.5**	**8.5**	**9.1**	17.5	19.6	23.1
TCLP	6.6	10.8	14.2	17.6	20.6	20.5	24.9
# Cases	119	93	76	56	21	13	14

Table 5b. Homogenous comparison of selected Atlantic basin early intensity guidance model biases (kt) for 2013. 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	4.5	7.1	9.4	5.5	8.1	10.4
OCD5	2.7	6.0	9.4	13.5	13.8	24.1	31.5
HWFI	**1.7**	**2.2**	**3.3**	**4.3**	**4.1**	9.3	15.9
GHMI	**1.5**	**-0.2**	**-0.3**	**1.8**	**3.0**	11.2	23.9
DSHP	2.8	6.4	10.1	13.7	9.2	**4.8**	**5.5**
LGEM	2.6	5.9	9.5	13.6	10.3	11.2	16.3
IVCN	2.5	**3.9**	**6.1**	**8.8**	7.0	9.4	15.5
FSSE	**1.7**	**2.7**	**4.1**	**6.0**	**3.9**	**5.5**	10.4
GFSI	**0.6**	**2.1**	**4.9**	**8.3**	12.9	22.4	27.9
EMXI	**-0.4**	**-0.5**	**0.8**	**2.3**	8.5	17.5	23.1
TCLP	2.7	6.1	9.4	13.6	13.7	20.5	24.9
# Cases	119	93	76	56	21	13	14

Table 6. Official Atlantic track and intensity forecast verifications (OFCL) for 2013 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: AL012013 ANDREA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	4.0	4.0	8	1.3	1.9
012	6	29.7	68.0	6	7.5	7.0
024	4	38.3	178.8	4	3.8	7.5
036	2	38.9	370.1	2	0.0	4.5
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: AL022013 BARRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	14	12.8	13.8	14	2.1	2.1
012	12	27.5	38.3	12	4.2	3.3
024	10	41.7	78.5	10	7.0	3.5
036	8	57.5	99.6	8	8.1	6.4
048	6	59.8	95.3	6	9.2	7.0
072	2	25.6	111.1	2	5.0	5.5
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: AL032013 CHANTAL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	11	12.3	11.8	11	2.3	2.7
012	9	46.0	70.6	9	6.7	7.8
024	7	79.6	144.6	7	9.3	9.1
036	5	94.7	232.5	5	11.0	12.0
048	3	101.7	293.4	3	15.0	10.0
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: AL042013 DORIAN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	16	14.0	13.9	16	0.6	1.3
012	12	18.9	42.2	12	3.8	5.2
024	10	32.2	92.6	10	6.0	10.2
036	8	39.5	143.2	8	9.4	13.5
048	6	42.8	194.1	6	10.0	16.8
072	2	11.1	270.0	2	5.0	21.5
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: AL052013 ERIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	12.1	13.2	13	1.9	1.5
012	11	25.3	37.5	11	2.7	3.4
024	9	41.3	65.3	9	4.4	4.9
036	7	66.4	80.8	7	7.1	7.3
048	5	103.1	90.4	5	12.0	10.8
072	1	186.3	66.2	1	15.0	16.0
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: AL062013 FERNAND

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	4	11.0	11.6	4	6.3	6.3
012	2	18.6	45.7	2	5.0	9.5
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	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL072013 GABRIELLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	3.3	3.3	18	2.2	2.5
012	14	19.9	48.2	14	6.8	7.8
024	10	44.0	131.0	10	7.5	9.4
036	8	77.2	239.0	8	10.6	17.8
048	6	102.1	371.0	6	13.3	24.2
072	2	132.0	575.8	2	15.0	30.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	2	164.3	410.0	2	15.0	29.5

Verification statistics for: AL082013 EIGHT

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	3	2.7	2.7	3	0.0	0.0
012	1	20.6	47.5	1	0.0	4.0
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
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: AL092013 HUMBERTO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	33	9.8	10.8	33	2.1	2.4
012	29	29.3	51.9	29	5.2	7.0
024	25	47.5	107.2	25	7.2	10.4
036	23	61.3	168.3	23	7.6	12.9
048	21	74.9	255.8	21	6.9	14.7
072	17	91.3	524.5	17	7.6	20.1
096	17	133.3	565.2	17	9.1	24.1
120	17	165.4	545.2	17	12.4	29.6

Verification statistics for: AL102013 INGRID

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	6.7	9.5	18	1.1	1.1
012	16	28.7	47.8	16	8.4	7.7
024	14	43.9	87.0	14	14.3	15.5
036	12	47.6	124.6	12	15.8	23.3
048	10	80.6	174.4	10	17.5	29.6
072	6	80.3	134.5	6	13.3	24.8
096	2	100.0	152.7	2	20.0	21.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL112013 JERRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	19	10.6	11.4	19	1.1	1.6
012	17	36.0	46.1	17	3.2	4.2
024	15	73.4	91.8	15	4.3	6.8
036	13	126.2	125.1	13	4.2	8.4
048	11	204.8	163.2	11	5.0	9.9
072	7	333.6	262.1	7	3.6	13.9
096	3	391.7	351.0	3	1.7	19.3
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL122013 KAREN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	6.7	6.7	13	3.5	3.8
012	11	31.2	34.5	11	6.8	9.5
024	9	54.7	36.1	9	11.1	17.8
036	7	89.6	62.5	7	20.7	26.9
048	5	148.4	107.0	5	25.0	36.6
072	1	288.1	298.4	1	20.0	37.0
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: AL132013 LORENZO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	12	8.9	8.9	12	1.3	1.7
012	10	23.2	39.8	10	3.0	6.2
024	8	39.0	91.5	8	3.8	10.1
036	6	55.5	152.6	6	2.5	10.5
048	4	79.1	229.3	4	2.5	14.8
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: AL142013 MELISSA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	14	10.0	10.1	14	2.1	2.5
012	12	29.3	53.9	12	6.3	6.1
024	10	50.9	159.6	10	8.0	7.4
036	8	83.2	319.5	8	10.6	8.9
048	6	128.5	495.2	6	9.2	4.8
072	2	236.1	742.3	2	0.0	2.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Table 7. Homogenous comparison of official and CLIPER5 track forecast errors in the eastern North Pacific basin in 2013 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
2013 mean OFCL error (n mi)	25.3	40.2	51.3	63.1	89.9	125.5	166.6
2013 mean CLIPER5 error (n mi)	39.4	80.3	124.2	166.0	253.9	324.7	387.0
2013 mean OFCL skill relative to CLIPER5 (%)	25.4	49.9	58.7	62.0	64.6	61.3	57.1
2013 mean OFCL bias vector (°/n mi)	282/002	245/004	235/003	336/001	048/019	043/051	015/085
2013 number of cases	255	222	189	159	111	78	56
2008-2012 mean OFCL error (n mi)	27.0	43.1	57.8	71.9	101.7	137.2	165.9
2008-2012 mean CLIPER5 error (n mi)	37.4	73.0	114.9	158.3	238.4	313.5	389.1
2008-2012 mean OFCL skill relative to CLIPER5 (%)	27.8	41.0	49.7	54.6	57.3	56.2	57.4
2008-2012 mean OFCL bias vector (°/n mi)	117/000	121/002	127/004	111/007	107/013	100/024	094/038
2008-2012 number of cases	1161	1017	882	755	540	366	224
2013 OFCL error relative to 2008-2012 mean (%)	-6.3	-6.7	-11.3	-12.2	-11.6	-8.5	0.4
2013 CLIPER5 error relative to 2008-2012 mean (%)	5.3	14.4	8.1	4.9	6.5	3.6	0.5

Table 8a. Homogenous comparison of eastern North Pacific basin early track guidance model errors (n mi) for 2013. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	26.4	38.5	49.0	62.1	90.6	126.5	168.7
OCD5	42.1	80.9	125.3	175.6	271.1	347.1	347.4
GFSI	29.3	47.5	63.8	79.1	113.5	138.0	217.1
GHMI	31.0	50.1	70.2	85.8	126.9	175.5	215.9
HWFI	29.4	46.6	62.1	77.1	107.1	147.7	246.9
EMXI	29.1	43.1	52.8	65.1	93.1	**124.3**	**129.1**
CMCI	36.4	57.8	83.1	104.1	145.5	177.8	226.1
AEMI	28.8	46.2	59.9	76.0	112.6	146.5	197.9
FSSE	**24.9**	**34.8**	**44.8**	**56.5**	**84.5**	**109.6**	**133.9**
TVCE	**24.4**	**35.1**	**44.8**	**56.8**	**80.8**	**109.4**	**159.6**
LBAR	38.2	80.2	129.5	185.1	315.9	558.8	682.6
BAMD	40.4	73.1	104.0	145.2	220.9	296.3	423.0
BAMM	33.2	58.3	79.4	100.4	145.6	178.6	275.3
BAMS	36.7	58.1	82.6	108.3	151.3	206.7	350.8
TCLP	41.2	80.3	125.6	176.5	266.1	356.0	384.2
# Cases	160	132	116	96	73	51	22

Table 8b. Homogenous comparison of eastern North Pacific basin early track guidance model bias vectors (º/n mi) for 2013.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	261/003	055/001	038/002	026/004	049/024	035/049	035/049
OCD5	306/002	015/010	006/023	003/051	015/114	022/193	027/244
GFSI	332/003	020/005	055/006	098/012	078/031	038/041	321/071
GHMI	189/002	039/007	016/018	015/020	027/045	027/083	022/072
HWFI	007/003	077/005	110/005	167/010	191/031	205/059	235/140
EMXI	266/006	218/006	202/010	198/008	096/021	085/041	356/036
CMCI	295/007	321/014	327/021	338/031	335/062	338/102	325/132
AEMI	271/006	300/007	296/008	271/006	020/016	357/049	307/103
FSSE	260/006	261/006	261/007	240/010	157/008	123/009	263/062
TVCE	274/003	343/001	336/002	289/002	059/006	025/014	289/052
LBAR	017/014	360/050	357/093	004/138	020/246	039/482	045/620
BAMD	022/014	026/035	025/059	029/090	033/141	028/209	026/344
BAMM	339/008	356/018	352/027	355/034	009/045	359/059	347/139
BAMS	306/015	312/027	305/043	300/065	294/090	280/116	268/170
TCLP	272/005	336/006	339/017	345/037	006/084	019/148	024/194
# Cases	160	132	116	96	73	51	22

Table 9. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the eastern North Pacific basin for the 2013 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
2013 mean OFCL error (kt)	5.2	9.5	12.0	13.0	14.3	14.9	13.1
2013 mean Decay-SHIFOR5 error (kt)	7.0	11.5	13.6	15.4	16.5	17.7	19.3
2013 mean OFCL skill relative to Decay-SHIFOR5 (%)	25.7	17.4	12.3	15.6	13.3	15.8	32.1
2013 OFCL bias (kt)	-0.3	0.2	1.2	1.9	3.4	1.9	0.3
2013 number of cases	255	222	189	159	111	78	56
2008-12 mean OFCL error (kt)	6.3	10.5	13.4	14.5	15.3	17.0	17.3
2008-12 mean Decay-SHIFOR5 error (kt)	7.6	12.5	16.5	18.8	20.4	20.3	20.6
2008-12 mean OFCL skill relative to Decay-SHIFOR5 (%)	17.1	16.0	18.8	22.9	25.0	16.3	16.0
2008-12 OFCL bias (kt)	0.0	-0.1	-0.9	-1.9	-1.4	-1.0	0.2
2008-12 number of cases	1161	1017	882	755	540	366	224
2013 OFCL error relative to 2008-12 mean (%)	-17.5	-9.5	-10.4	-10.4	-6.5	-12.4	-24.6
2013 Decay-SHIFOR5 error relative to 2008-12 mean (%)	-7.9	-8.0	-17.6	-18.1	-19.1	-12.8	-6.3

Table 10a. Homogenous comparison of eastern North Pacific basin early intensity guidance model errors (kt) for 2013. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	5.8	9.8	12.7	14.2	14.3	14.8	12.1
OCD5	8.2	12.4	14.7	16.1	16.3	16.9	21.1
HWFI	7.8	11.8	12.9	**13.5**	15.4	**14.7**	20.3
GHMI	8.8	13.5	16.5	17.2	16.9	18.5	17.8
DSHP	7.4	11.2	13.8	15.0	16.3	16.0	12.2
LGEM	7.5	10.7	13.6	15.8	17.0	17.9	14.4
IVCN	7.4	10.8	**12.6**	**13.5**	**13.3**	**12.3**	**10.4**
FSSE	6.9	10.1	**12.0**	**12.6**	**13.1**	**13.4**	14.3
GFSI	10.0	14.8	17.8	19.8	20.2	19.0	13.5
EMXI	10.4	16.3	20.6	24.4	26.6	26.1	20.1
TCLP	8.0	12.7	16.3	18.4	17.1	16.9	16.2
# Cases	171	144	122	101	77	54	24

Table 10b. Homogenous comparison of eastern North Pacific basin early intensity guidance model biases (kt) for 2013. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	0.4	2.3	3.3	3.5	4.4	4.3	2.5
OCD5	-0.5	2.3	**2.8**	**2.2**	**2.5**	**2.6**	**-2.2**
HWFI	-1.4	**-1.3**	**-1.3**	**-2.3**	-4.9	-8.2	-11.3
GHMI	-1.6	-3.7	-4.7	**-2.8**	**2.0**	**3.8**	7.8
DSHP	0.8	3.0	5.4	6.4	7.9	6.5	5.7
LGEM	0.6	**1.7**	**2.7**	**2.7**	4.9	5.4	2.6
IVCN	**-0.1**	**0.1**	**0.7**	**1.2**	**2.5**	**1.9**	**1.3**
FSSE	1.2	3.5	4.6	3.6	**1.4**	**-2.4**	-8.1
GFSI	-3.5	-3.8	**-2.7**	**0.8**	**4.0**	6.8	7.6
EMXI	-2.0	**-1.3**	**0.7**	**3.0**	6.9	9.9	9.9
TCLP	0.4	**1.9**	**2.3**	**1.6**	**2.2**	**1.6**	**-2.0**
# Cases	171	144	122	101	77	54	24

Table 11. Official eastern North Pacific track and intensity forecast verifications (OFCL) for 2013 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: EP012013 ALVIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	10.2	10.2	7	0.7	0.7
012	5	35.5	39.5	5	9.0	8.4
024	3	45.1	43.8	3	11.7	17.7
036	1	71.3	97.6	1	20.0	27.0
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: EP022013 BARBARA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	9.3	10.7	7	1.4	1.4
012	5	26.7	56.3	5	5.0	10.6
024	3	66.8	161.8	3	10.0	19.3
036	1	91.5	306.2	1	5.0	15.0
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: EP032013 COSME

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	16	7.6	7.9	16	2.5	2.5
012	14	25.7	44.4	14	2.5	6.3
024	12	38.5	98.6	12	5.0	10.9
036	10	48.8	156.1	10	6.0	10.3
048	8	62.0	234.6	8	6.3	8.9
072	4	70.3	449.4	4	3.8	7.8
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0
144	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP042013 DALILA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	29	5.7	6.3	29	1.2	1.2
012	27	17.1	31.2	27	4.4	5.7
024	25	28.3	64.1	25	6.6	8.6
036	23	40.3	103.9	23	7.2	11.7
048	21	50.0	139.5	21	10.7	14.8
072	17	73.2	208.4	17	14.1	19.4
096	13	87.7	300.9	13	12.3	13.0
120	9	134.9	403.4	9	11.1	8.7

Verification statistics for: EP052013 ERICK

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	19	57.7	6.4	19	0.8	0.3
012	16	15.6	23.7	16	2.2	3.8
024	14	25.1	43.3	14	6.8	7.8
036	12	25.4	59.3	12	11.3	11.8
048	10	37.2	83.5	10	11.0	14.7
072	6	94.5	132.2	6	9.2	14.8
096	2	140.2	165.0	2	2.5	9.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP062013 FLOSSIE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	11	13.7	14.2	11	0.5	0.5
012	11	31.8	50.6	11	3.6	3.6
024	11	48.9	95.7	11	5.5	5.5
036	11	72.0	151.8	11	7.3	7.9
048	11	97.0	212.0	11	9.5	9.5
072	10	153.2	347.8	10	9.0	6.9
096	6	270.8	514.3	6	6.7	13.0
120	2	410.0	814.6	2	0.0	7.0

Verification statistics for: EP072013 GIL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	27	5.6	5.9	27	1.1	1.7
012	27	21.0	22.7	27	5.2	6.4
024	26	35.9	42.8	26	10.2	9.7
036	24	45.8	69.5	24	13.5	13.2
048	22	56.2	107.4	22	14.1	16.6
072	18	73.6	189.4	18	14.4	23.2
096	14	96.6	250.3	14	13.6	26.0
120	10	135.6	328.9	10	8.0	24.7

Verification statistics for: EP082013 HENRIETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	10.2	10.2	23	3.0	3.0
012	23	23.1	33.5	23	7.2	8.3
024	23	36.7	71.9	23	9.6	11.7
036	23	46.7	113.4	23	11.5	13.5
048	23	53.8	155.9	23	12.8	15.3
072	21	81.4	237.0	21	14.8	14.4
096	17	108.3	253.6	17	17.1	15.5
120	13	143.5	237.4	13	18.5	17.0

Verification statistics for: EP092013 IVO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	13.7	13.7	13	1.2	1.2
012	11	47.4	59.3	11	3.2	4.1
024	9	69.1	106.7	9	3.9	6.8
036	7	61.9	146.4	7	5.0	5.3
048	5	49.0	170.5	5	2.0	6.2
072	1	62.5	224.4	1	5.0	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: EP102013 JULIETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	2.6	2.6	5	5.0	6.0
012	3	24.6	35.0	3	10.0	8.3
024	1	17.3	68.5	1	0.0	10.0
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	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP112013 KIKO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	10	5.8	6.5	10	2.5	2.5
012	8	28.0	42.6	8	6.9	11.6
024	6	55.8	90.4	6	18.3	23.3
036	4	71.6	140.6	4	17.5	17.8
048	2	84.0	139.0	2	7.5	4.0
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: EP122013 LORENA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	10	10.9	10.9	10	3.0	3.0
012	8	31.3	65.2	8	1.9	5.0
024	6	66.1	136.2	6	0.8	1.7
036	4	93.4	206.2	4	5.0	5.8
048	2	102.7	255.5	2	2.5	14.5
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: EP132013 MANUEL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	20	10.0	10.9	20	2.3	2.5
012	16	23.2	36.6	16	8.4	10.4
024	12	37.6	77.8	12	13.3	16.3
036	8	53.5	121.6	8	11.3	15.4
048	4	78.4	150.9	4	13.8	19.3
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: EP142013 NARDA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	9.3	9.3	15	0.7	1.0
012	13	29.9	45.4	13	6.9	7.3
024	11	53.9	91.0	11	15.5	14.0
036	9	68.8	131.8	9	21.1	17.3
048	7	82.4	162.3	7	29.3	23.7
072	3	77.1	120.1	3	33.3	27.3
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: EP152013 OCTAVE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	11	14.7	16.3	11	1.4	1.4
012	9	38.6	52.5	9	3.9	5.2
024	7	71.8	123.2	7	6.4	9.1
036	5	105.2	219.5	5	6.0	7.4
048	3	141.3	302.5	3	5.0	4.0
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: EP162013 PRISCILLA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	10	6.4	6.4	10	0.0	0.0
012	8	21.5	36.2	8	2.5	4.6
024	6	26.3	105.5	6	7.5	8.3
036	4	51.6	175.6	4	12.5	11.5
048	2	126.8	248.7	2	15.0	15.0
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: EP172013 RAYMOND

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	42	9.1	9.9	42	2.4	2.4
012	40	22.0	38.3	40	7.1	9.5
024	38	32.3	79.3	38	15.0	17.7
036	36	43.5	135.3	36	19.0	21.5
048	34	58.4	200.7	34	18.4	20.5
072	30	94.9	313.4	30	16.7	16.3
096	26	136.4	391.6	26	18.3	18.8
120	22	185.1	456.1	22	14.3	23.7

Verification statistics for: EP182013 SONIA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	10.0	10.0	13	1.5	1.5
012	11	35.8	63.5	11	1.8	4.0
024	9	49.4	128.9	9	5.0	6.3
036	7	57.2	174.8	7	5.7	5.7
048	5	78.9	171.7	5	3.0	8.4
072	1	176.5	201.0	1	10.0	6.0
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 2013.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	255
10	12	169
20	28	88
30	52	48
40	44	32
50	32	25
60	58	12
70	33	12
80	-	0
90	-	0
100	-	0

Table 12b. Verification of 48-h probabilistic genesis forecasts for the eastern North Pacific basin in 2013.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	1	162
10	17	125
20	20	87
30	43	54
40	44	27
50	73	33
60	67	15
70	80	10
80	80	10
90	100	3
100	100	1

Table 13a Verification of 120-h probabilistic genesis forecasts for the Atlantic basin in 2013.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	26
10	6	210
20	38	94
30	41	68
40	47	32
50	65	43
60	32	22
70	88	17
80	43	14
90	100	2
100	-	0

Table 13b Verification of 120-h probabilistic genesis forecasts for the eastern North Pacific basin in 2013.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	11
10	18	76
20	35	108
30	45	64
40	63	24
50	58	19
60	56	25
70	56	25
80	80	25
90	80	10
100	-	0

Table 14. HFIP Stream 1.5 models for 2013.

ID	Description	Parameter	NHC Application
APSI	PSU ARW with radar data assimilated. Early version of APSU.	Trk, Int	Direct use. Include in TV15 and IV15 consensus.
FM9I	ESRL FIM 15-km global model. Early version of FIM9.	Trk	Direct use.
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	Include in IV15 consensus.
GPMI	GFDL ensemble mean. Early version of GPMN.	Trk, Int	Direct use.
HWMI	HWRF ensemble mean. Early version of HWMN.	Trk, Int	Direct use.

Table 15. NHC forecast cone circle radii (n mi) for 2014. Change from 2013 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	33 (0: 0%)	30 (0: 0%)
24	52 (0: 0%)	46 (-3: -6%)
36	72 (0: 0%)	62 (-4: -6%)
48	92 (0: 0%)	79 (-3: -4%)
72	125 (-3: -2%)	105 (-6: -5%)
96	170 (-7: -4%)	154 (-3: -2%)
120	226 (-3: -1%)	190 (-7: -4%)

Table 16. Composition of NHC consensus models for 2014. 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 2014			
Model ID	Parameter	Type	Members
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

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

17. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2013. This verification includes only those models that were available at least 2/3 of the time (see text).

18. Reliability diagram for Atlantic (a) and eastern North Pacific (b) probabilistic tropical cyclogenesis forecasts for 2013. 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.

19. Homogeneous comparison of HFIP Stream 1.5 track models and selected operational models for 2013.

20. Impact of adding Stream 1.5 models to the variable track consensus TVCA.

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

22. Homogeneous comparison of HFIP Stream 1.5 HWRF ensemble mean (HWMI) and HWFI for track (top) and intensity (bottom).

23. Homogeneous comparison of HFIP Stream 1.5 intensity models and selected operational models for 2013.

24. Impact of adding Stream 1.5 models to the variable intensity consensus IVCN.

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

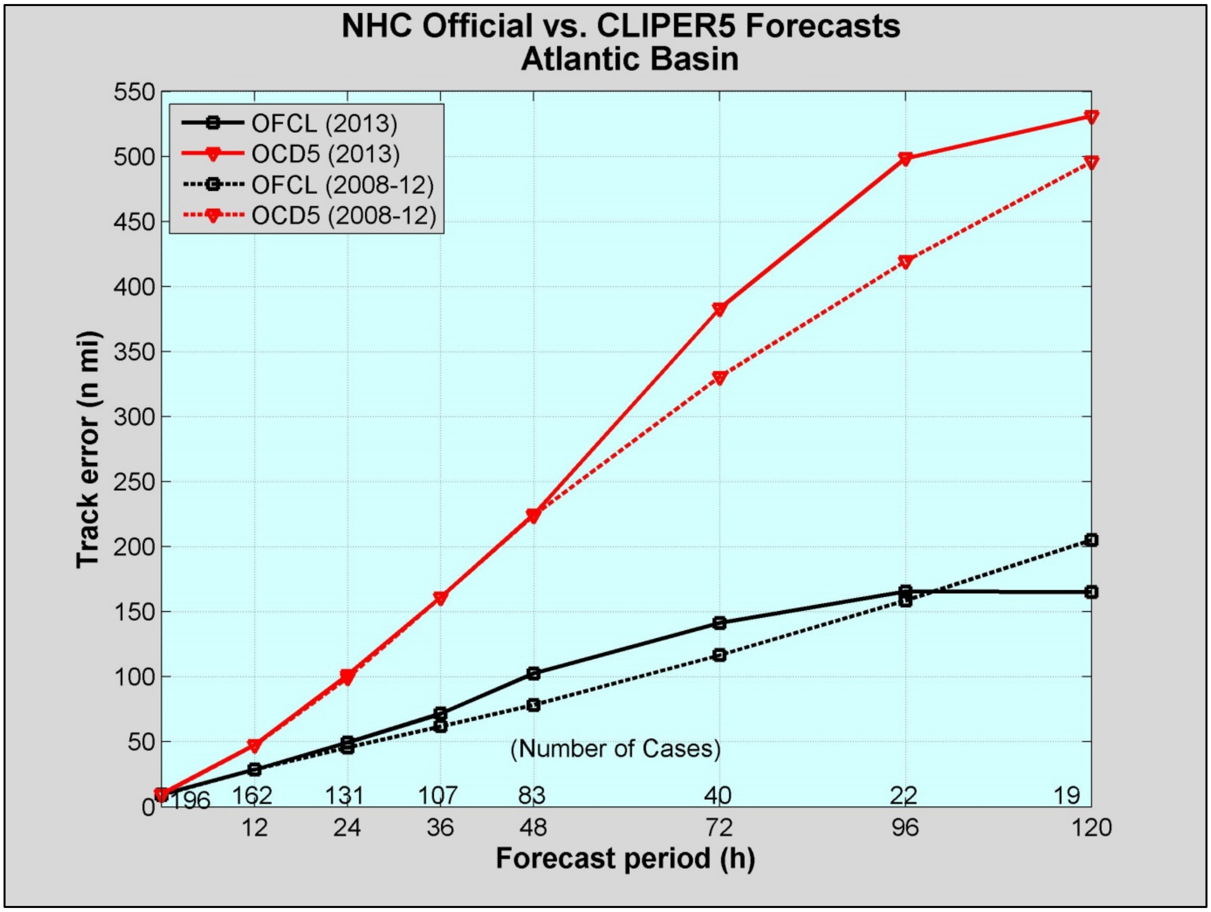
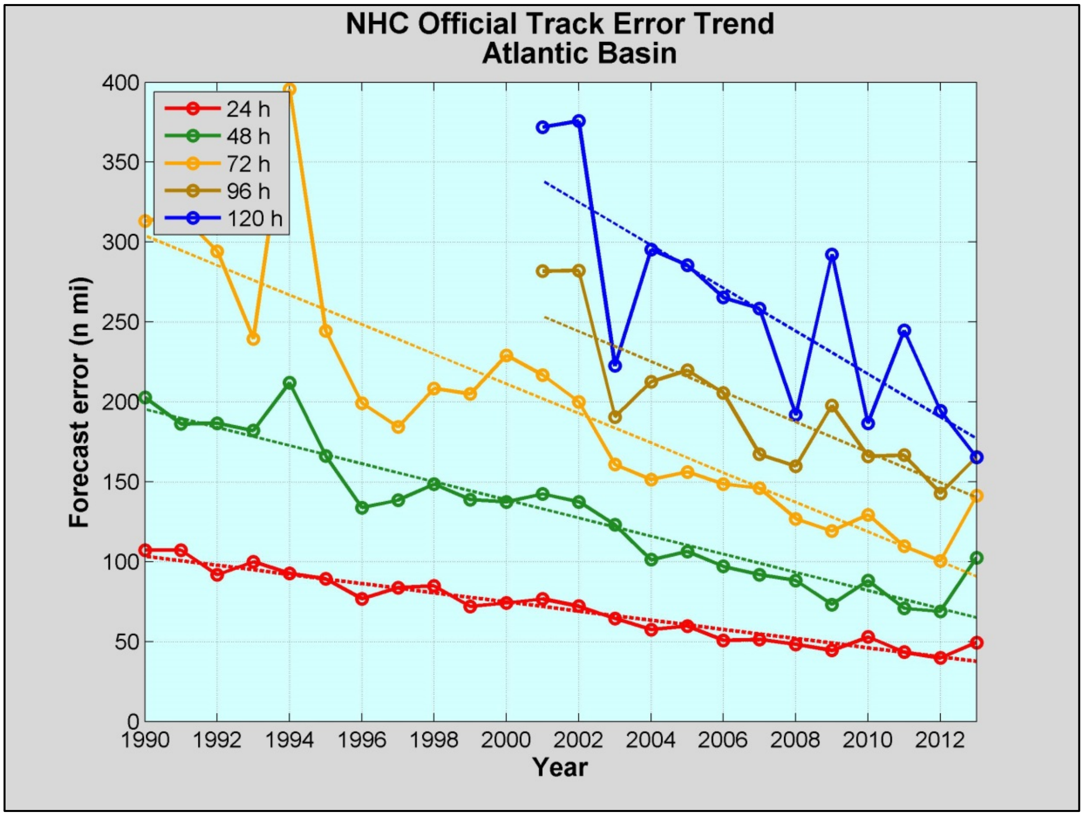


Figure 1. NHC official and CLIPER5 (OCD5) Atlantic basin average track errors for 2013 (solid lines) and 2008-2012 (dashed lines).



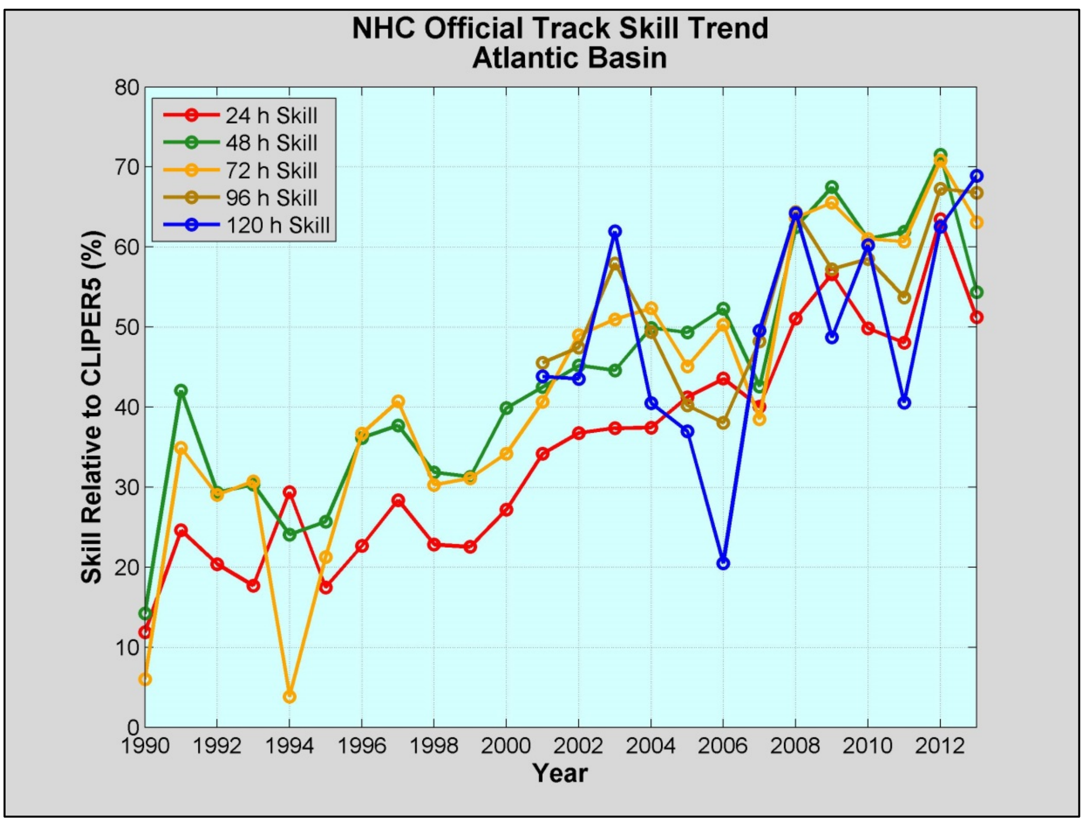


Figure 2. Recent trends in NHC official track forecast error (top) and skill (bottom) for the Atlantic basin.

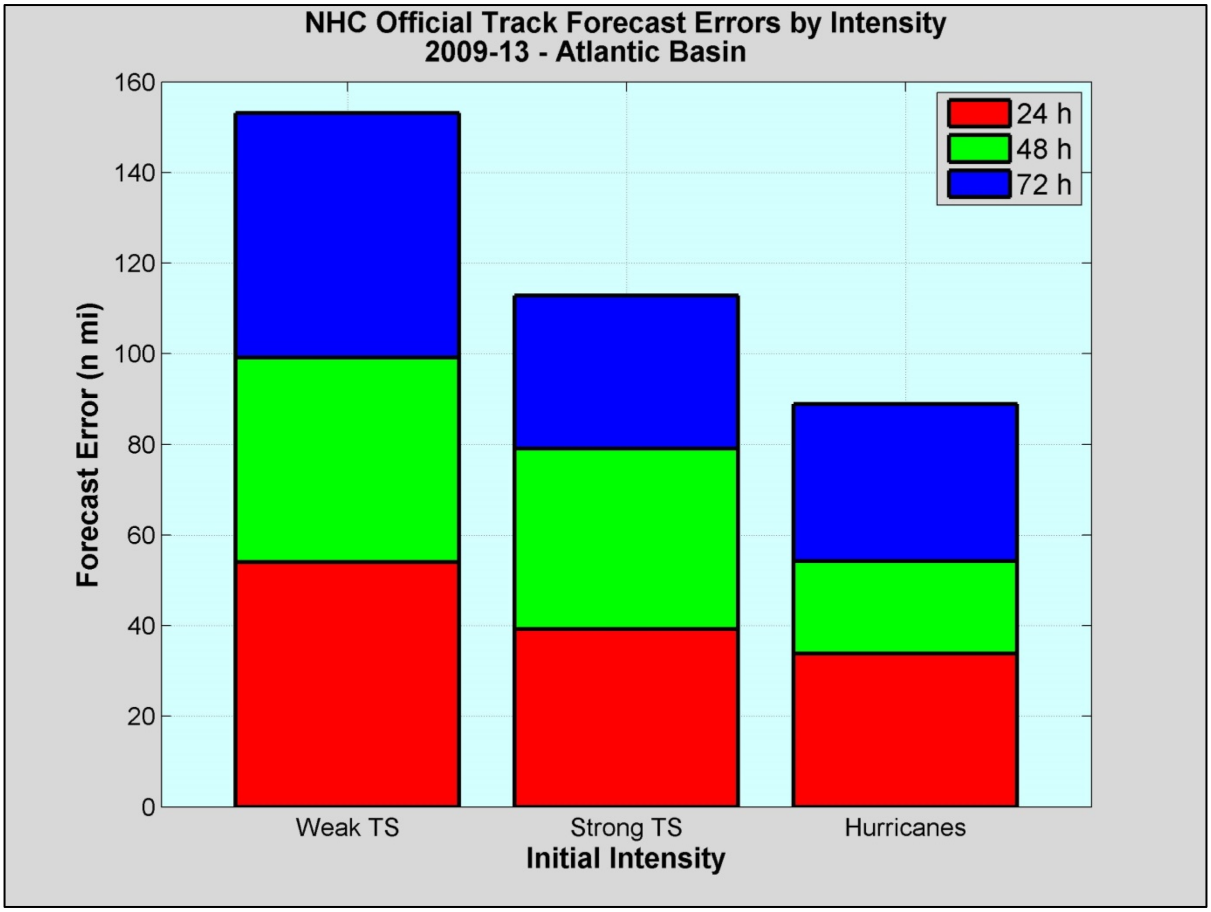


Figure 3. 2009-13 NHC official track forecast error binned by initial intensity for the Atlantic basin. Weak tropical storms are in the 35-45 kt range and strong tropical storms are in the 50-60 kt range.

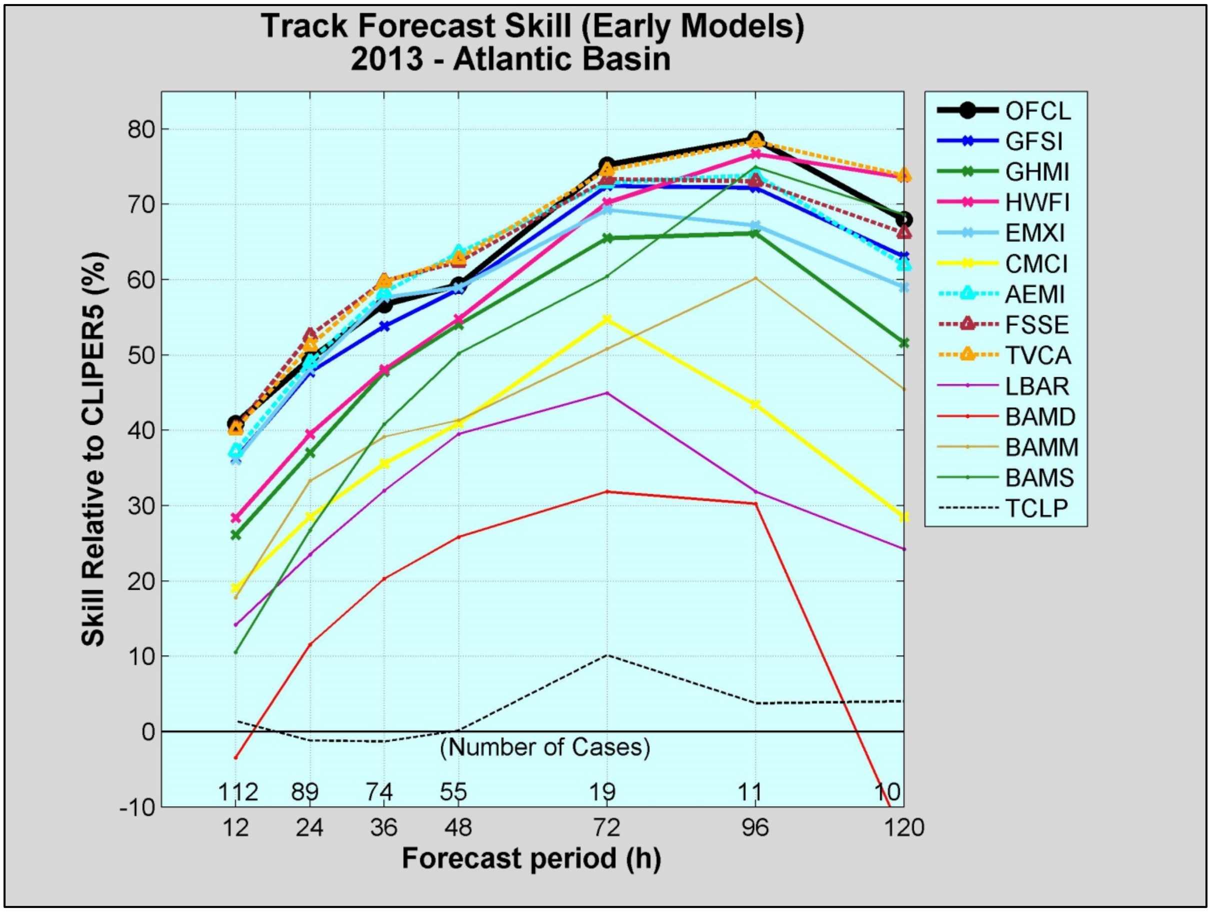


Figure 4. Homogenous comparison for selected Atlantic basin early track models for 2013. This verification includes only those models that were available at least 2/3 of the time (see text).

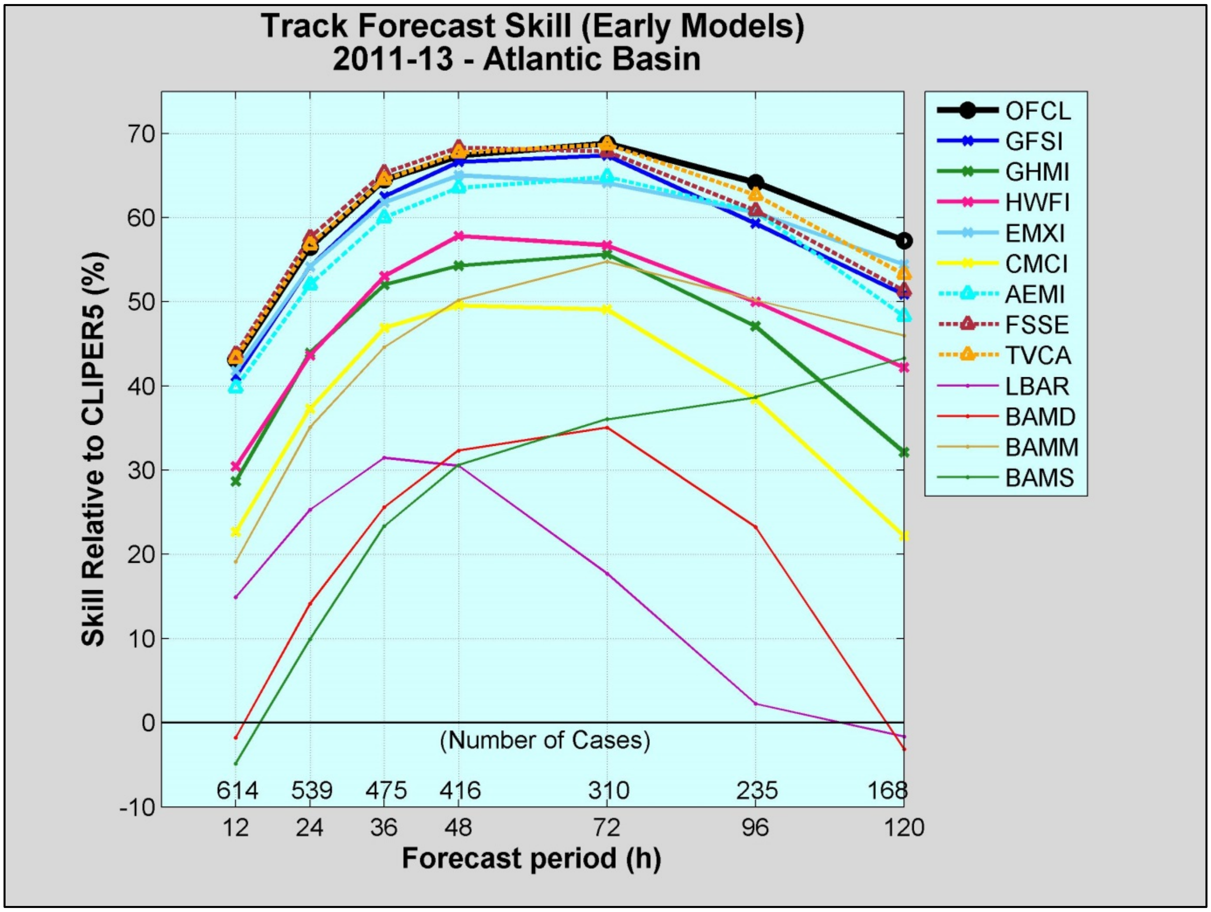


Figure 5. Homogenous comparison for selected Atlantic basin early track models for 2011-2013.

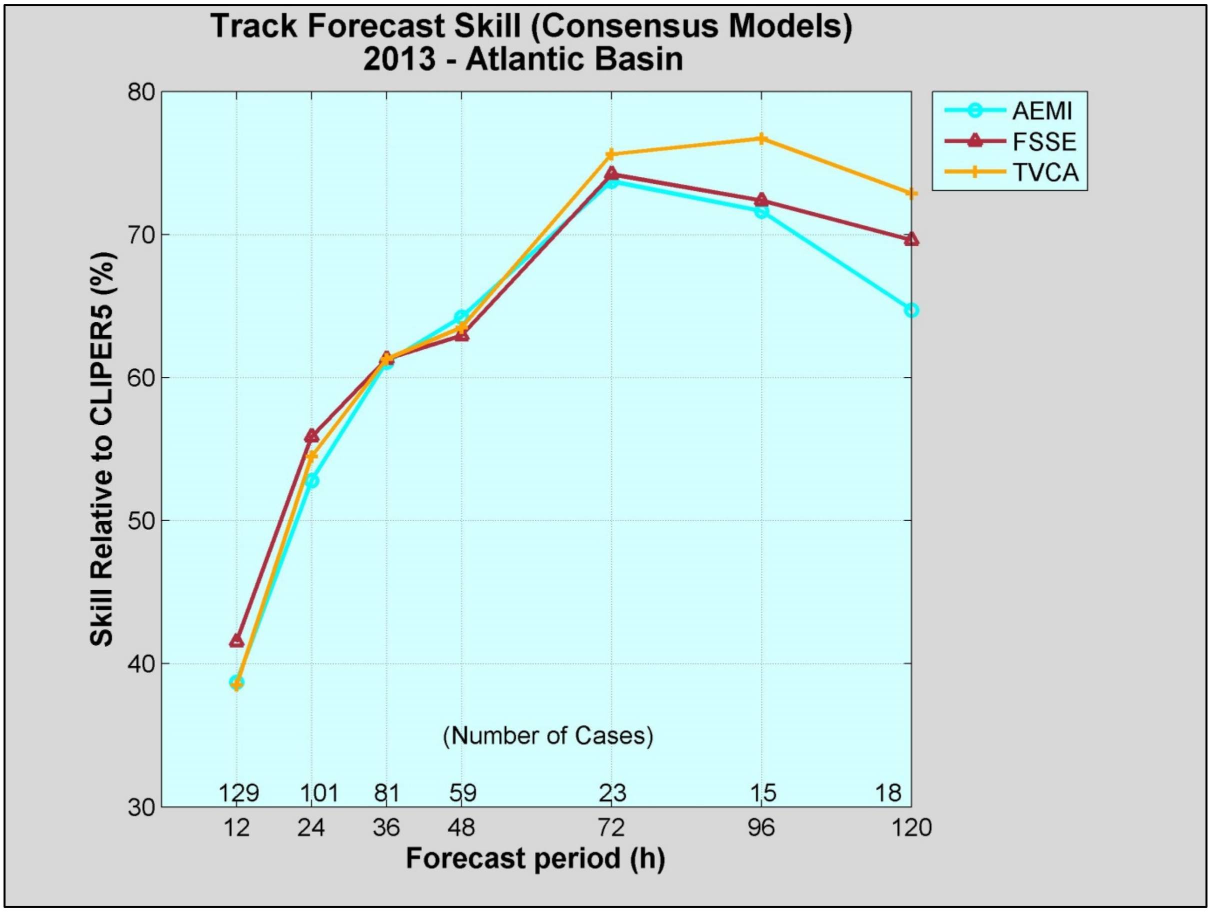


Figure 6. Homogenous comparison of the primary Atlantic basin track consensus models for 2013.

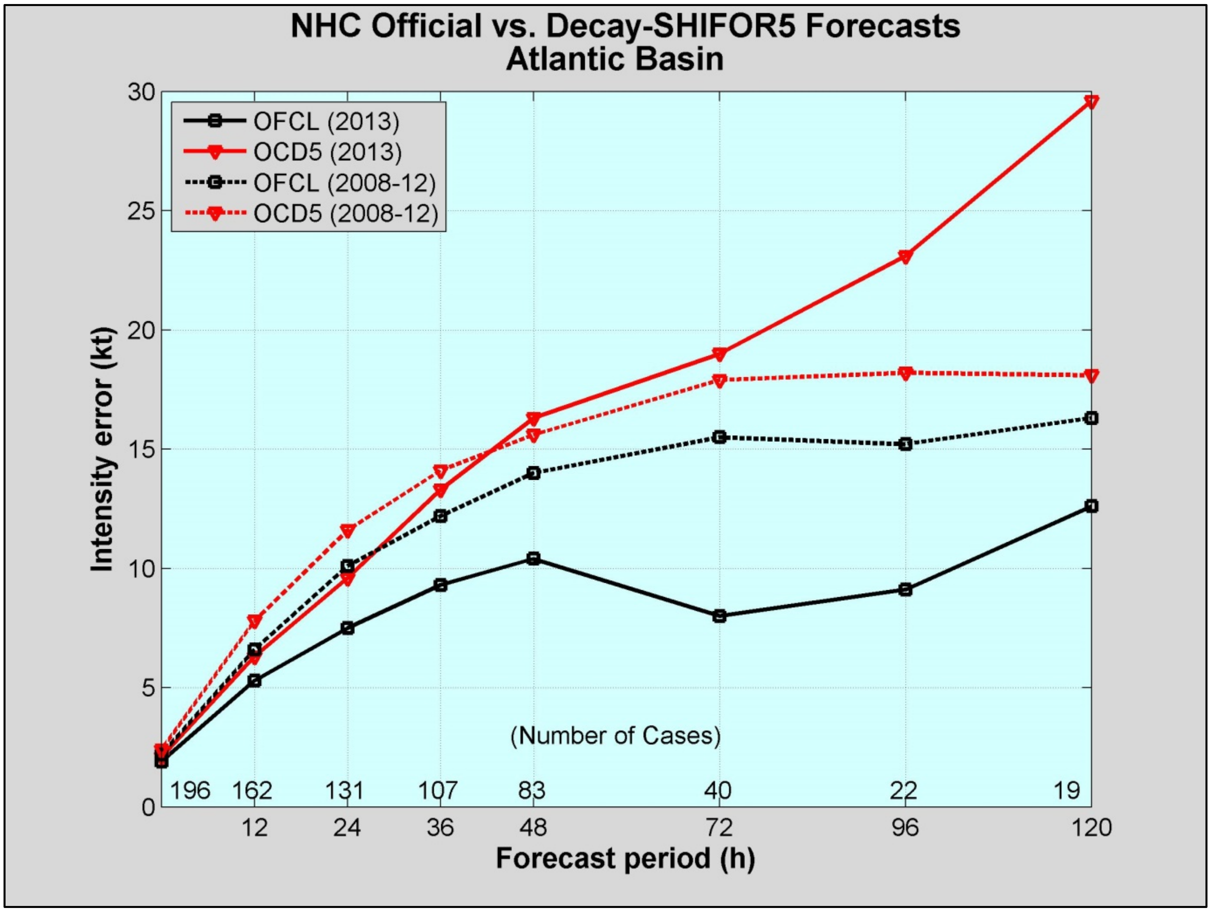


Figure 7. NHC official and Decay-SHIFOR5 (OCD5) Atlantic basin average intensity errors for 2013 (solid lines) and 2008-2012 (dashed lines).

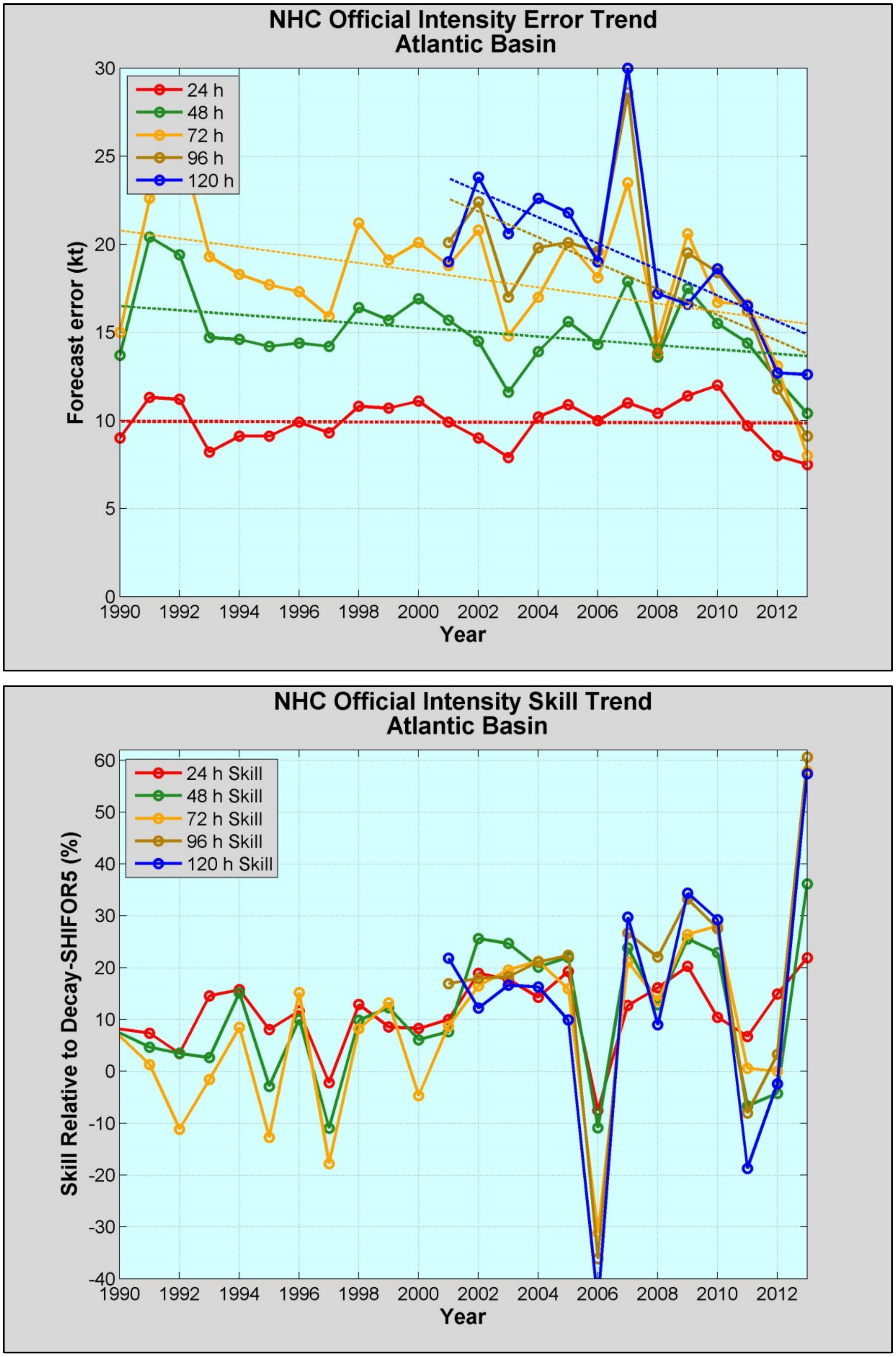


Figure 8. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the Atlantic basin.

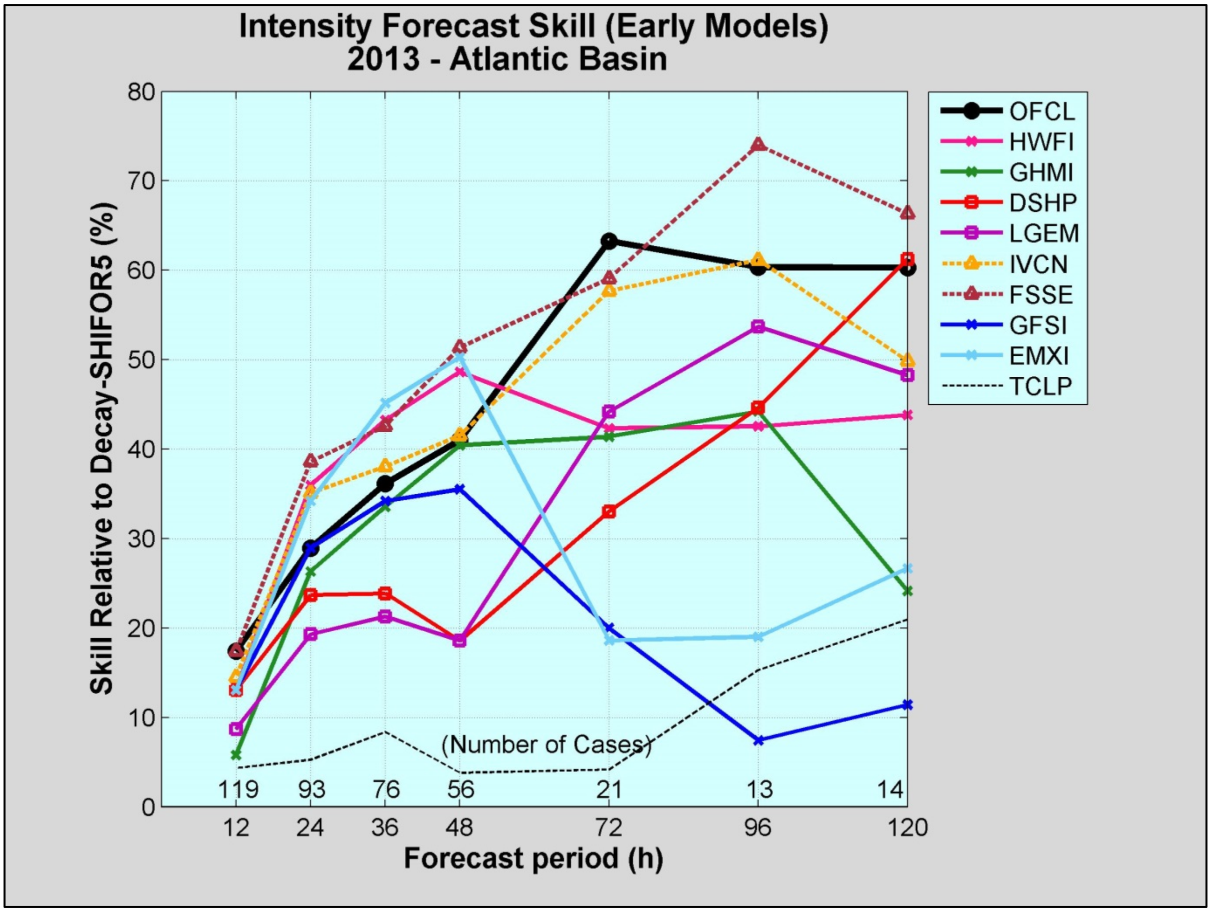


Figure 9. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2013.

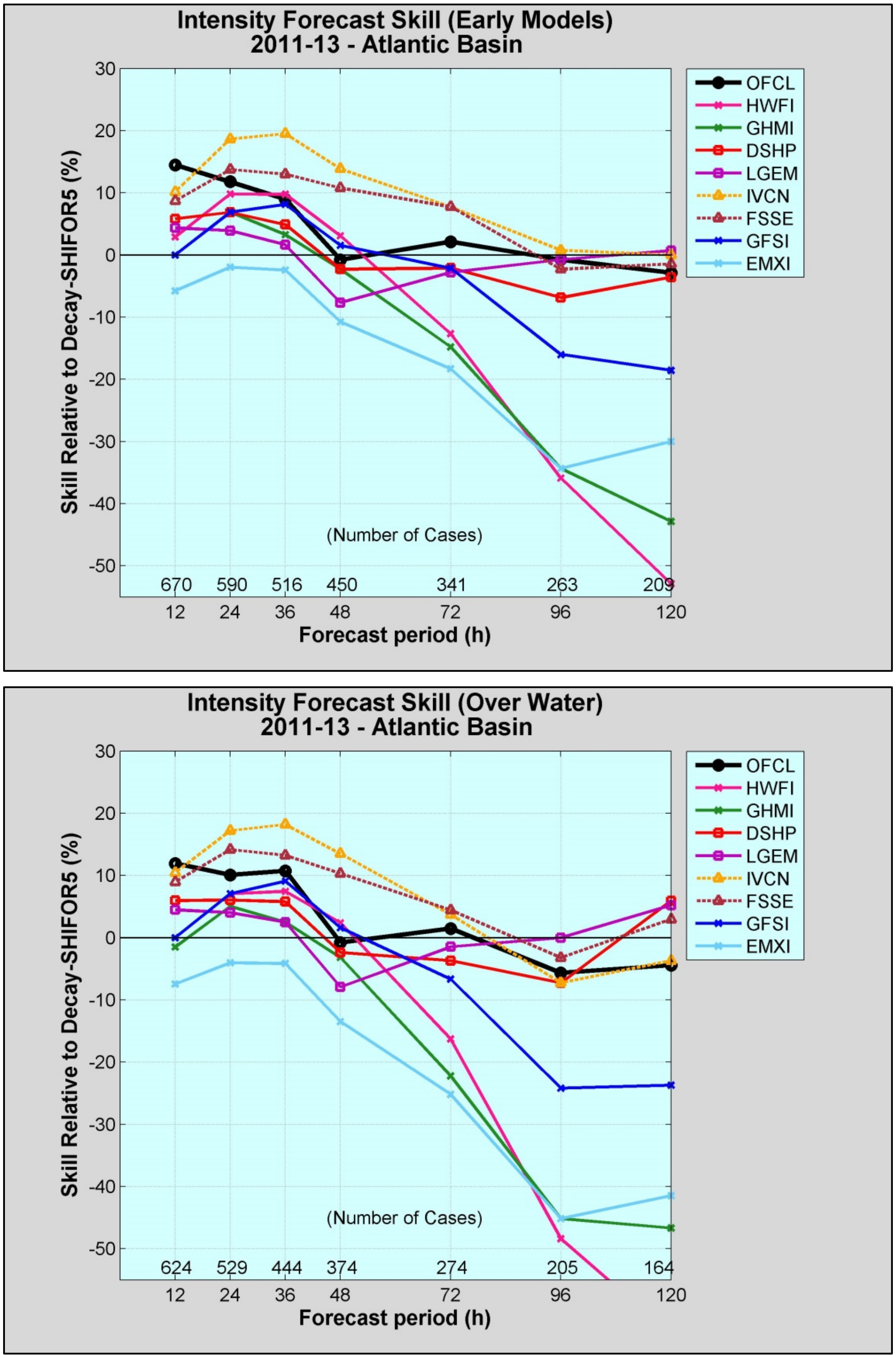


Figure 10. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2011-2013 (top) and for pre-landfall verifications only from 2011-13 (bottom). The pre-landfall verification sample is defined by excluding any portion of a model forecast that occurs after either the model forecast track or the verifying best track encounters land.

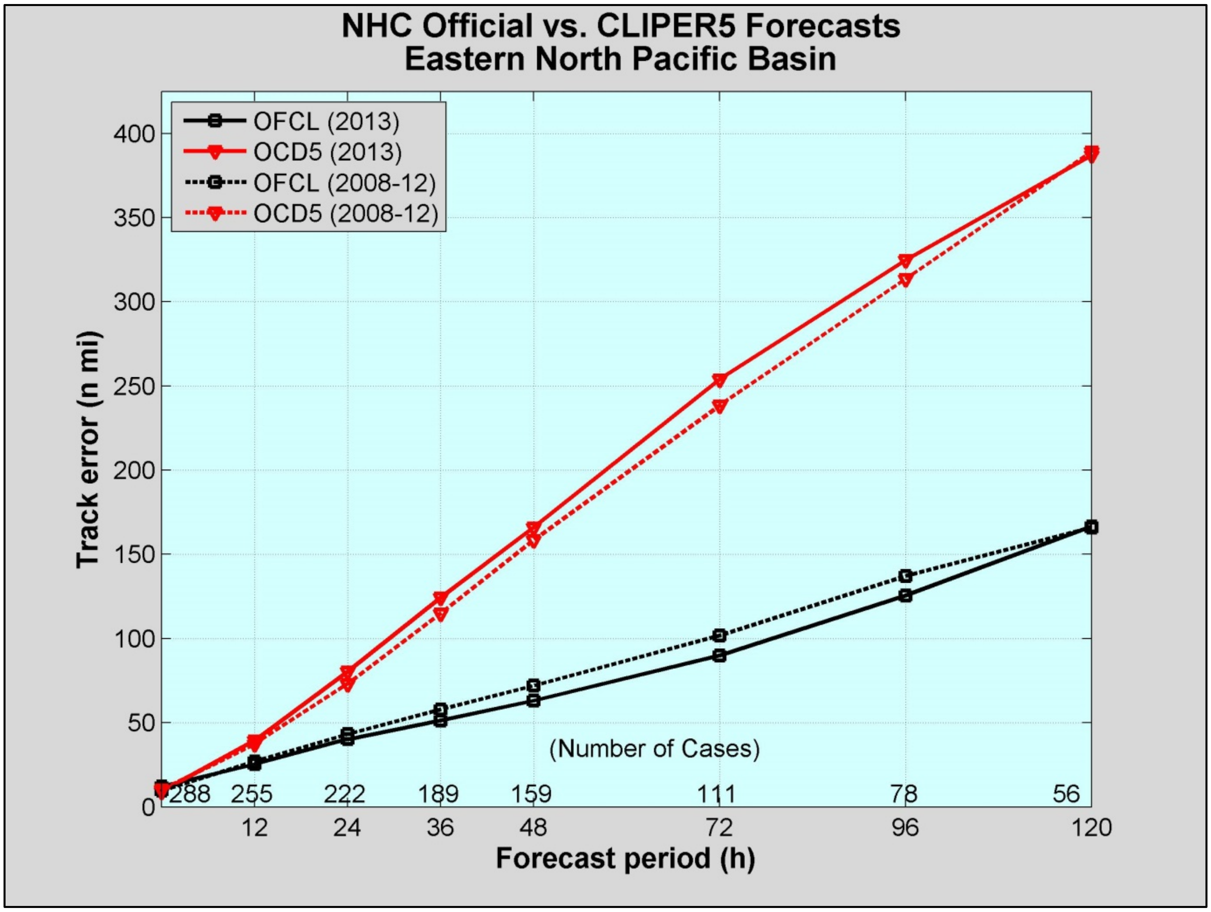


Figure 11. NHC official and CLIPER5 (OCD5) eastern North Pacific basin average track errors for 2013 (solid lines) and 2008-2012 (dashed lines).

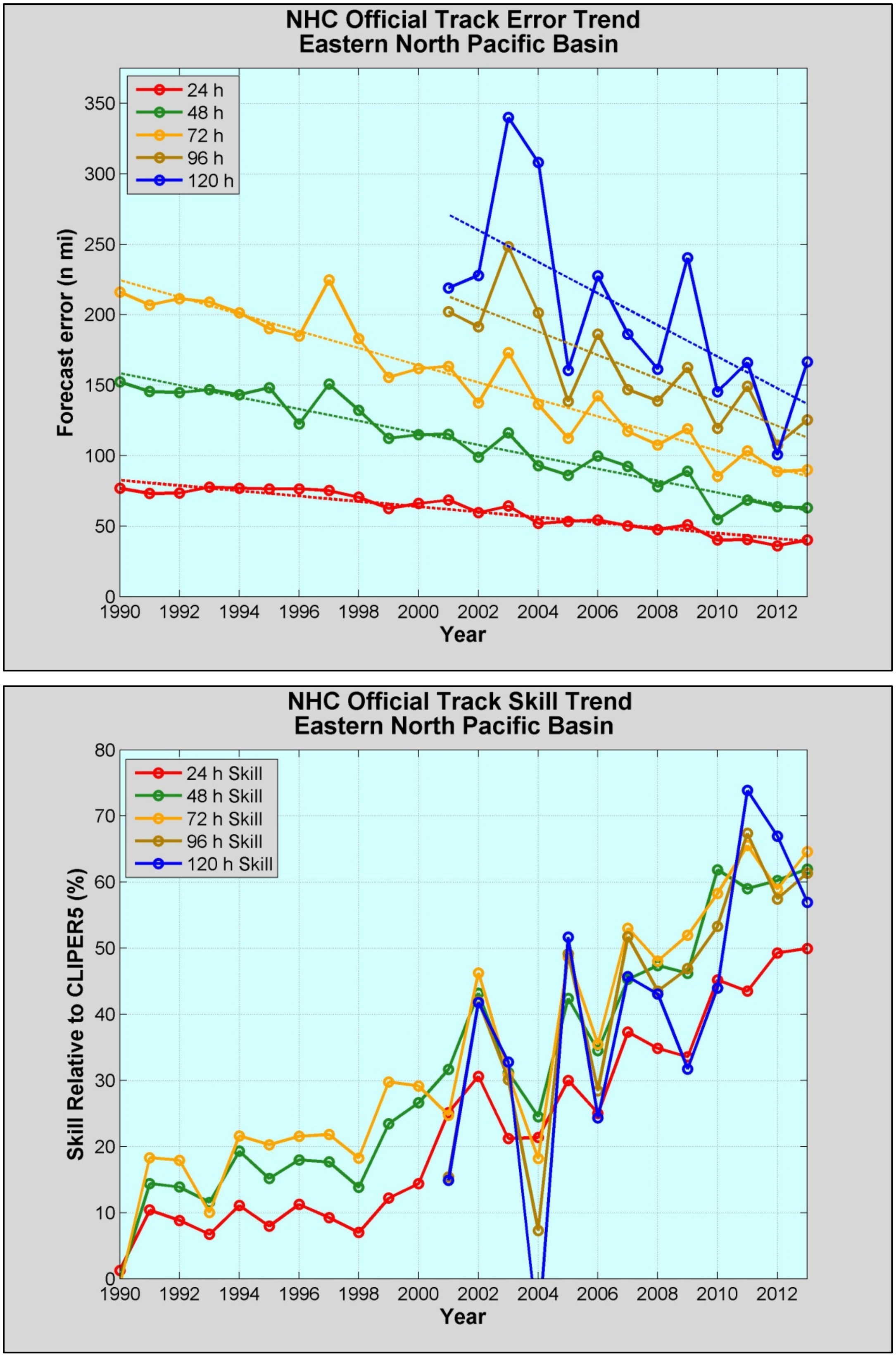


Figure 12. Recent trends in NHC official track forecast error (top) and skill (bottom) for the eastern North Pacific basin.

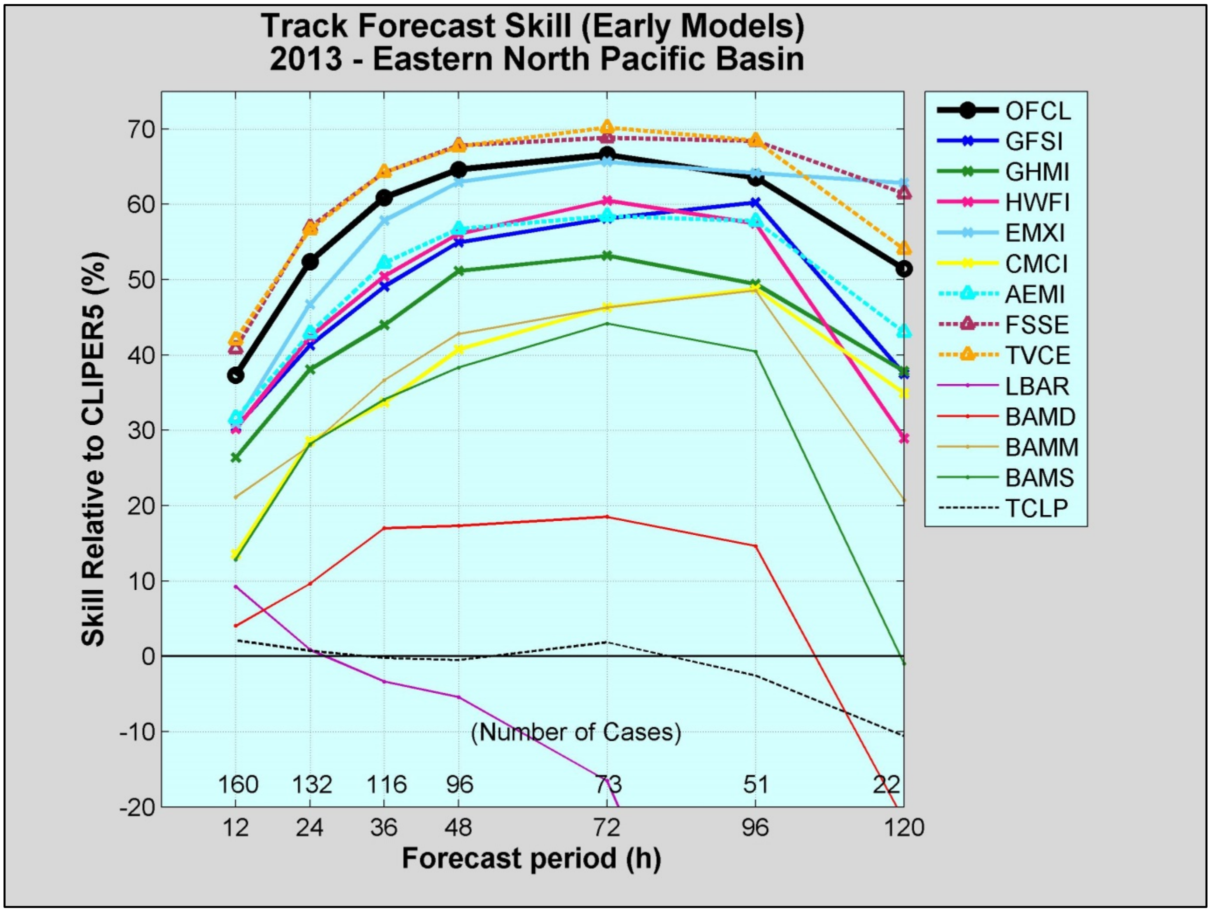


Figure 13. Homogenous comparison for selected eastern North Pacific early track models for 2013. This verification includes only those models that were available at least 2/3 of the time (see text).

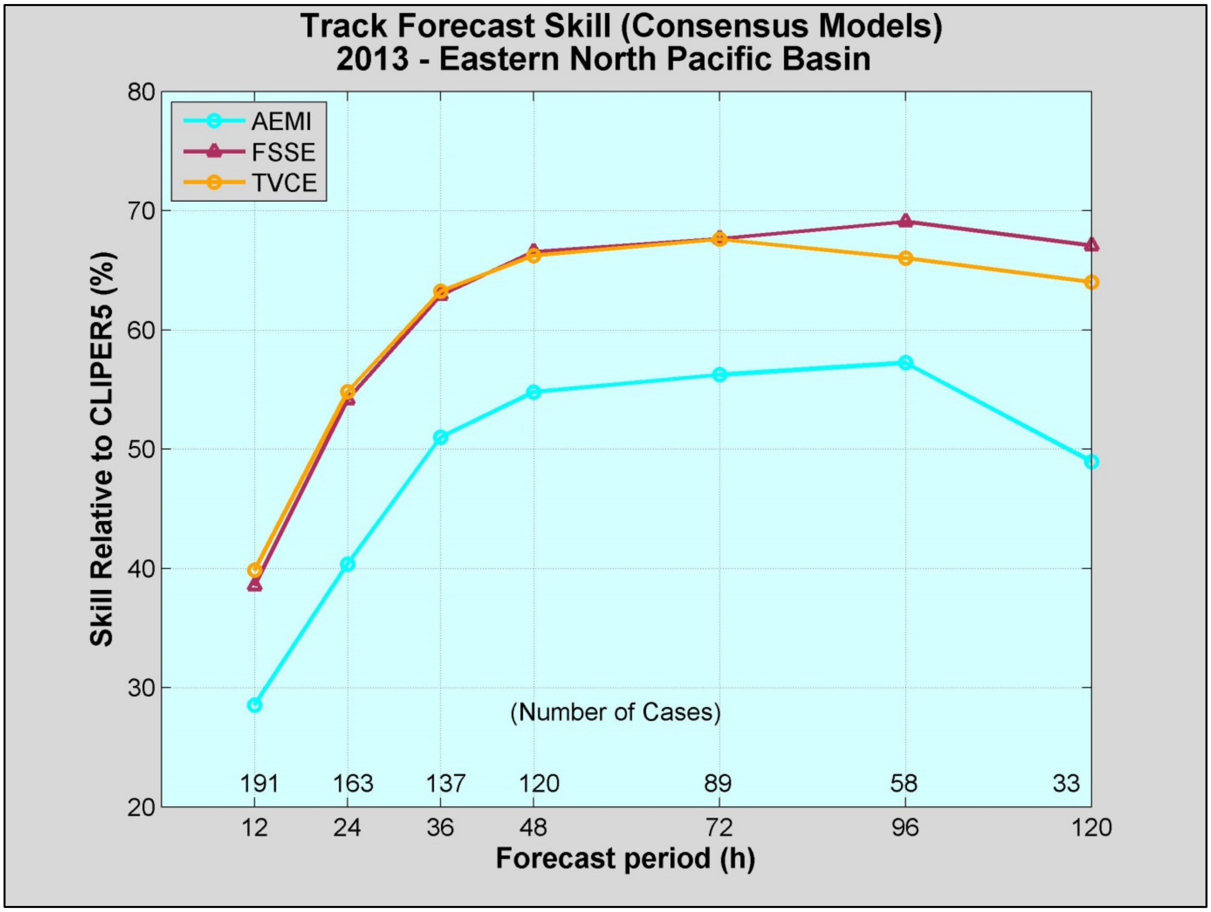


Figure 14. Homogenous comparison of the primary eastern North Pacific basin track consensus models for 2013.

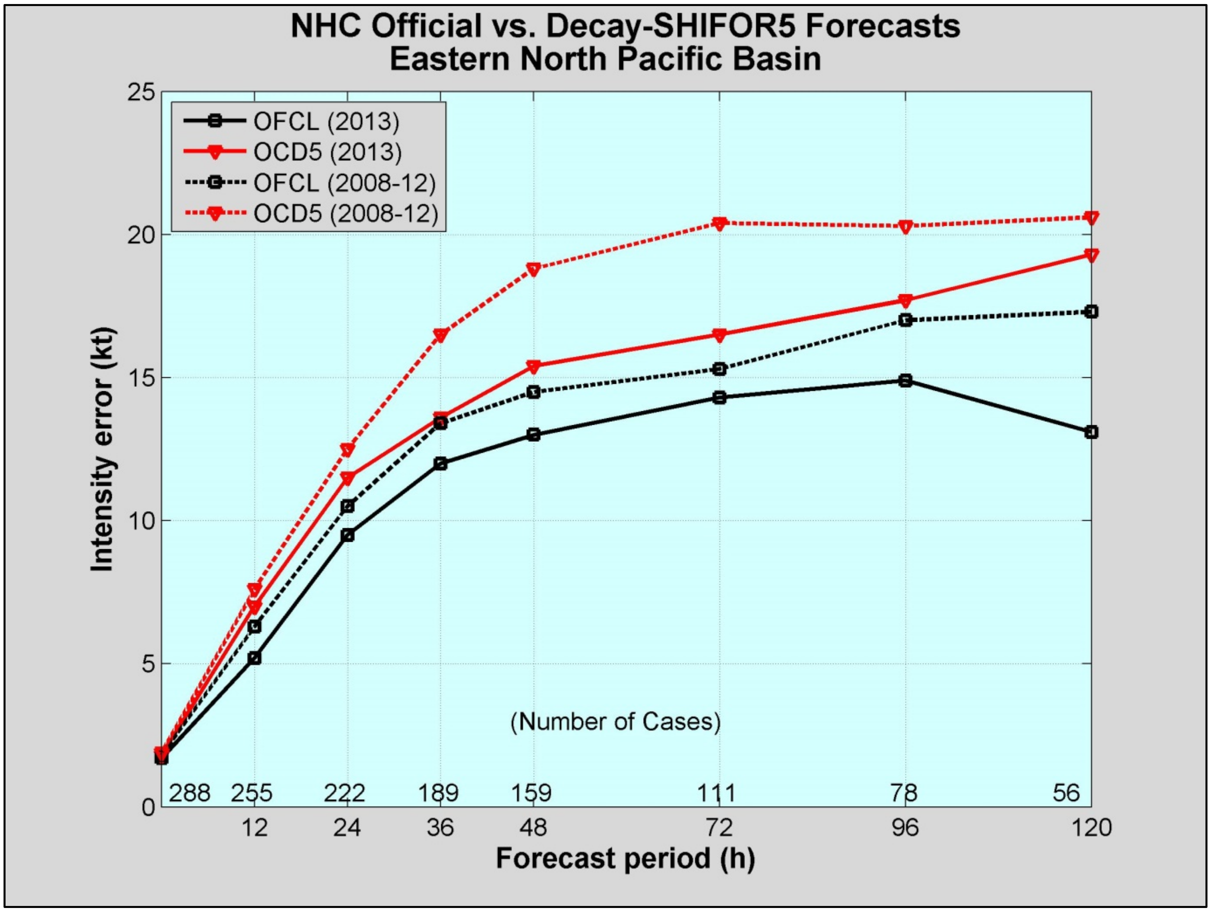
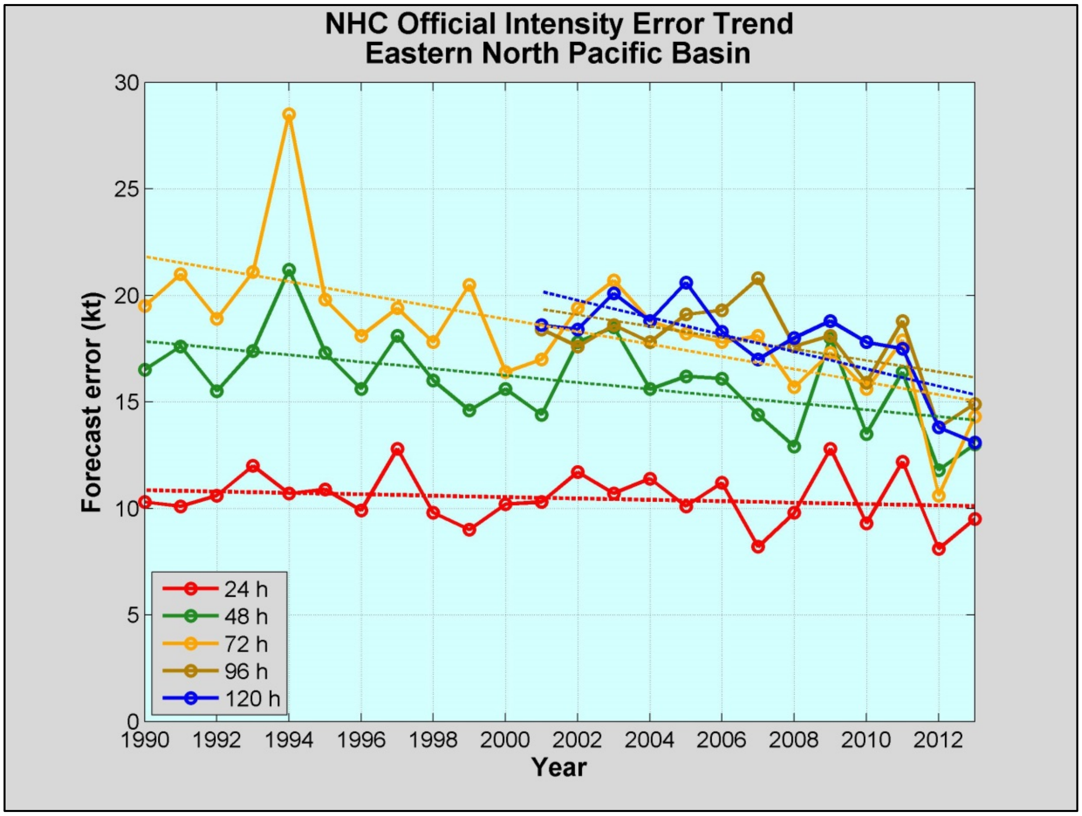


Figure 15. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2013 (solid lines) and 2008-2012 (dashed lines).



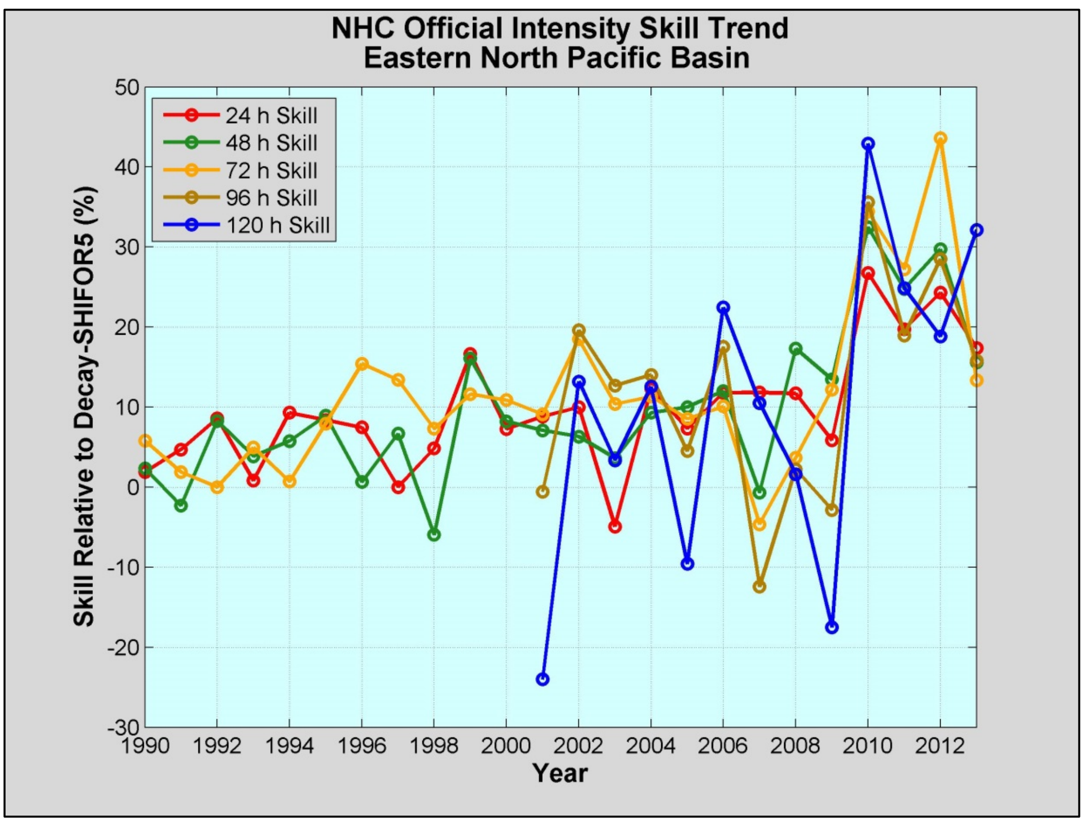


Figure 16. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

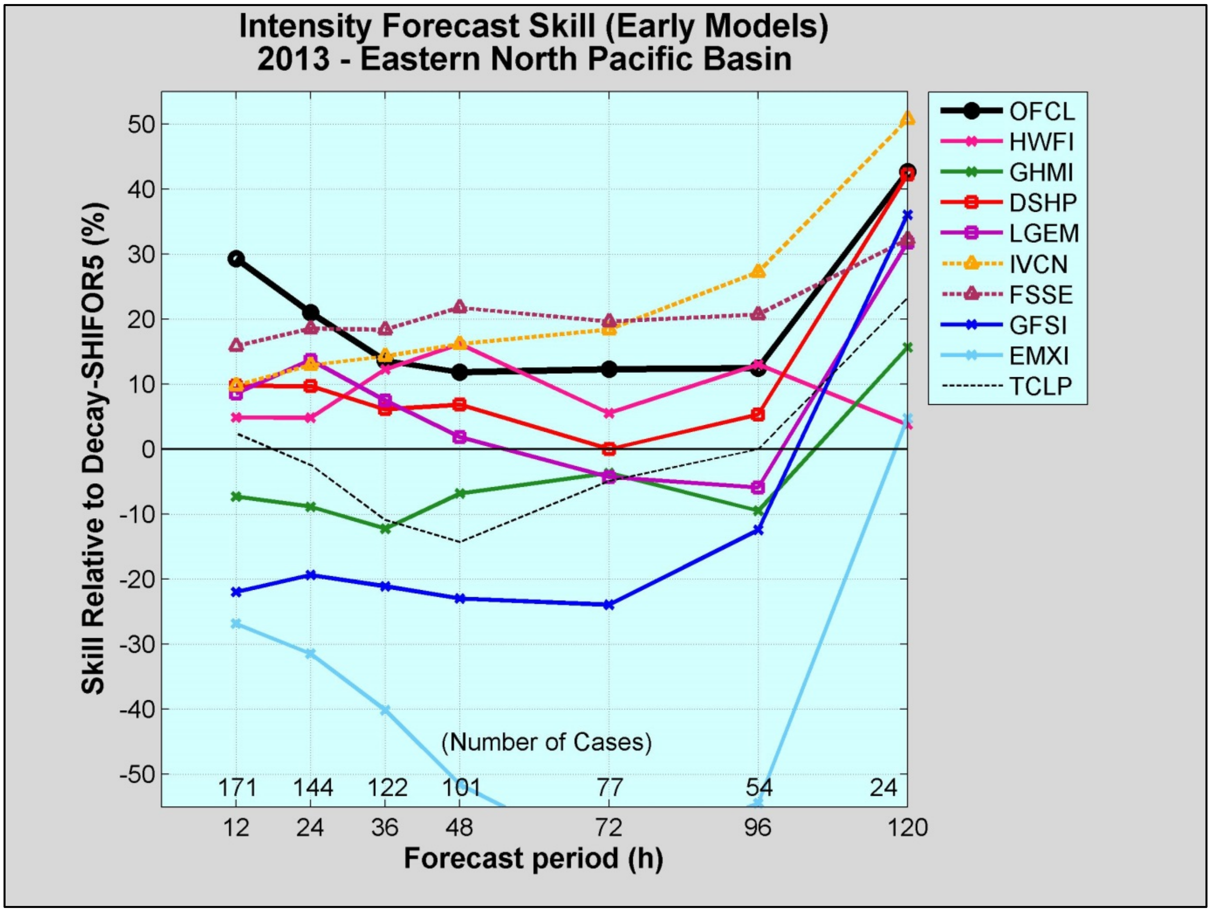


Figure 17. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2013.

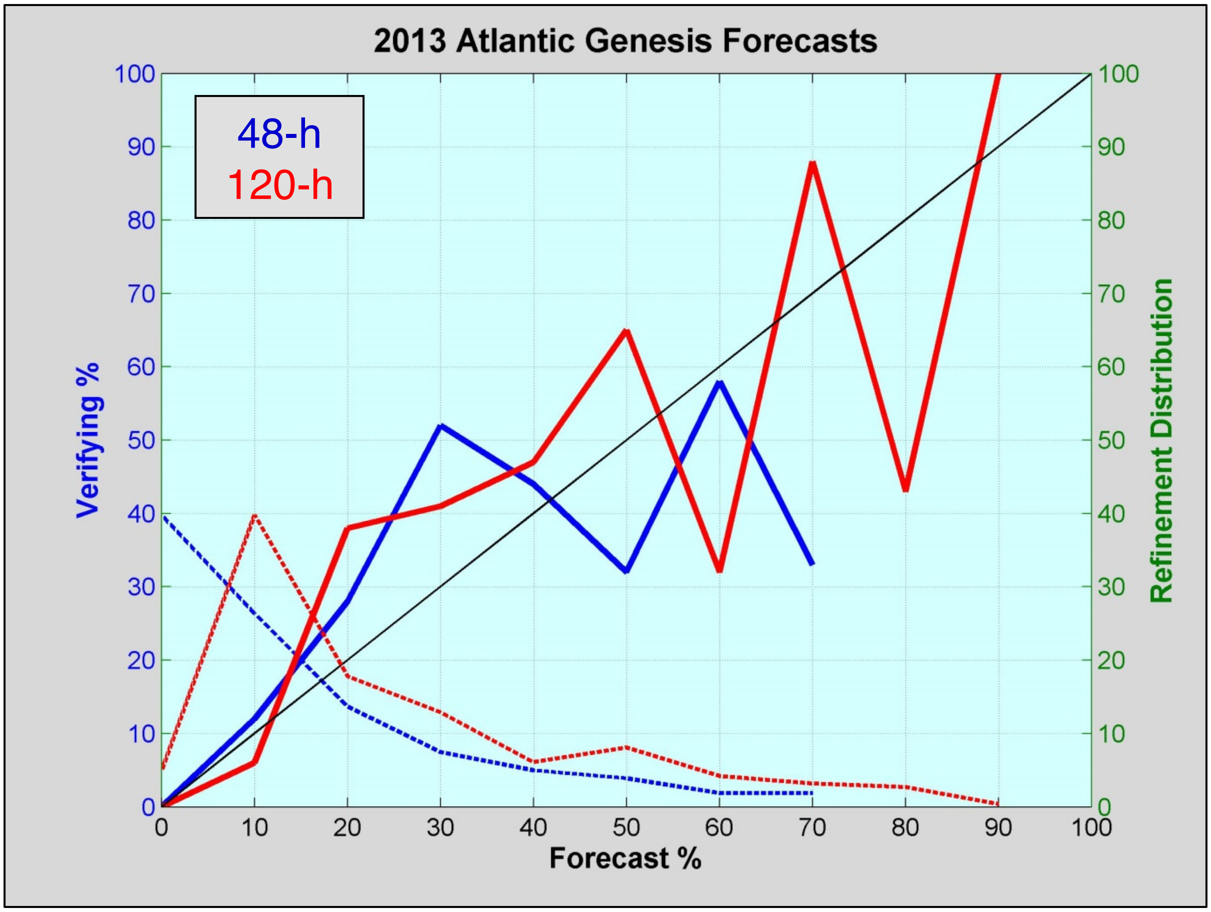


Figure 18a. Reliability diagram for Atlantic probabilistic tropical cyclogenesis forecasts for 2013. The solid lines indicate the relationship between the forecast and verifying genesis percentages, with perfect reliability indicated by the thin diagonal black line. The dashed lines indicate how the forecasts were distributed among the possible forecast values.

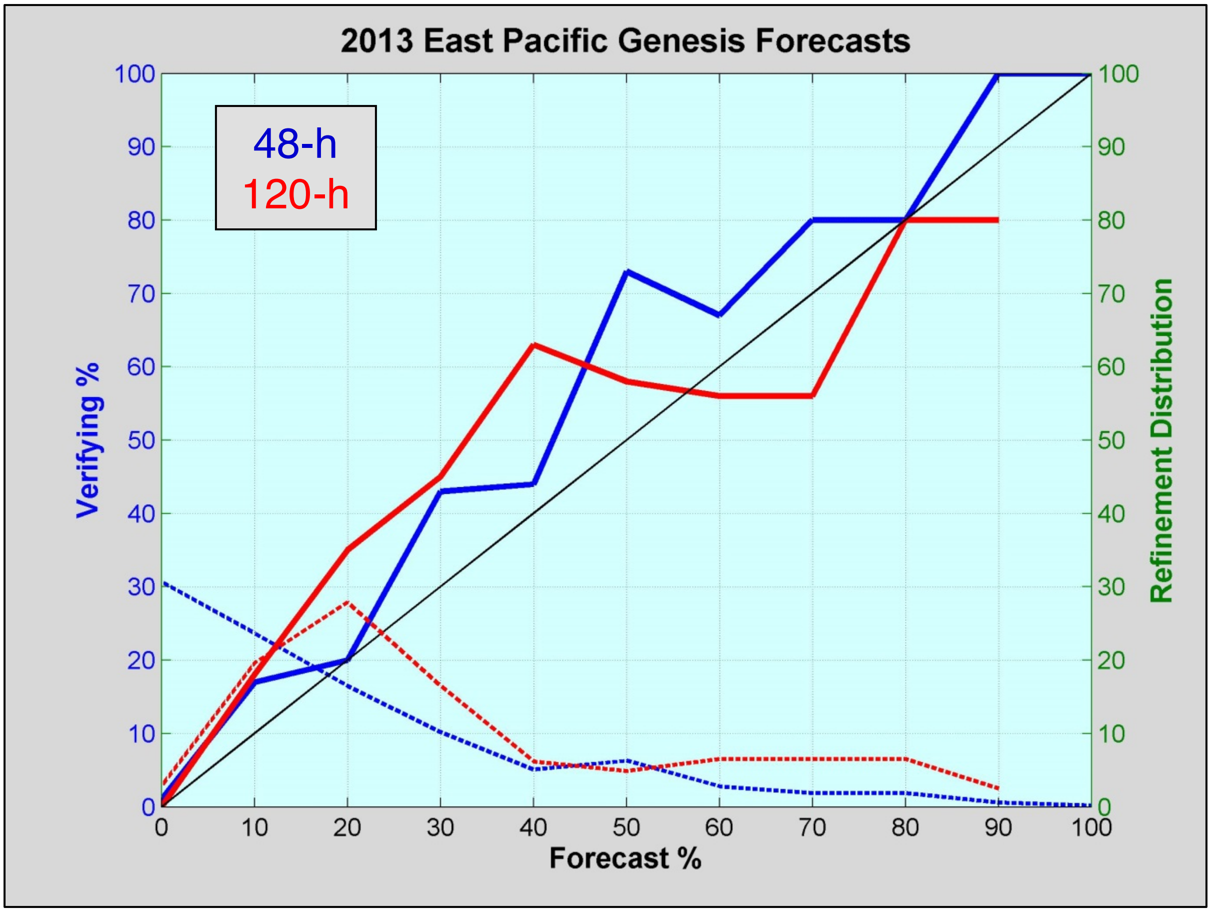


Figure 18b. As described for Fig. 18a, except for the eastern North Pacific basin.

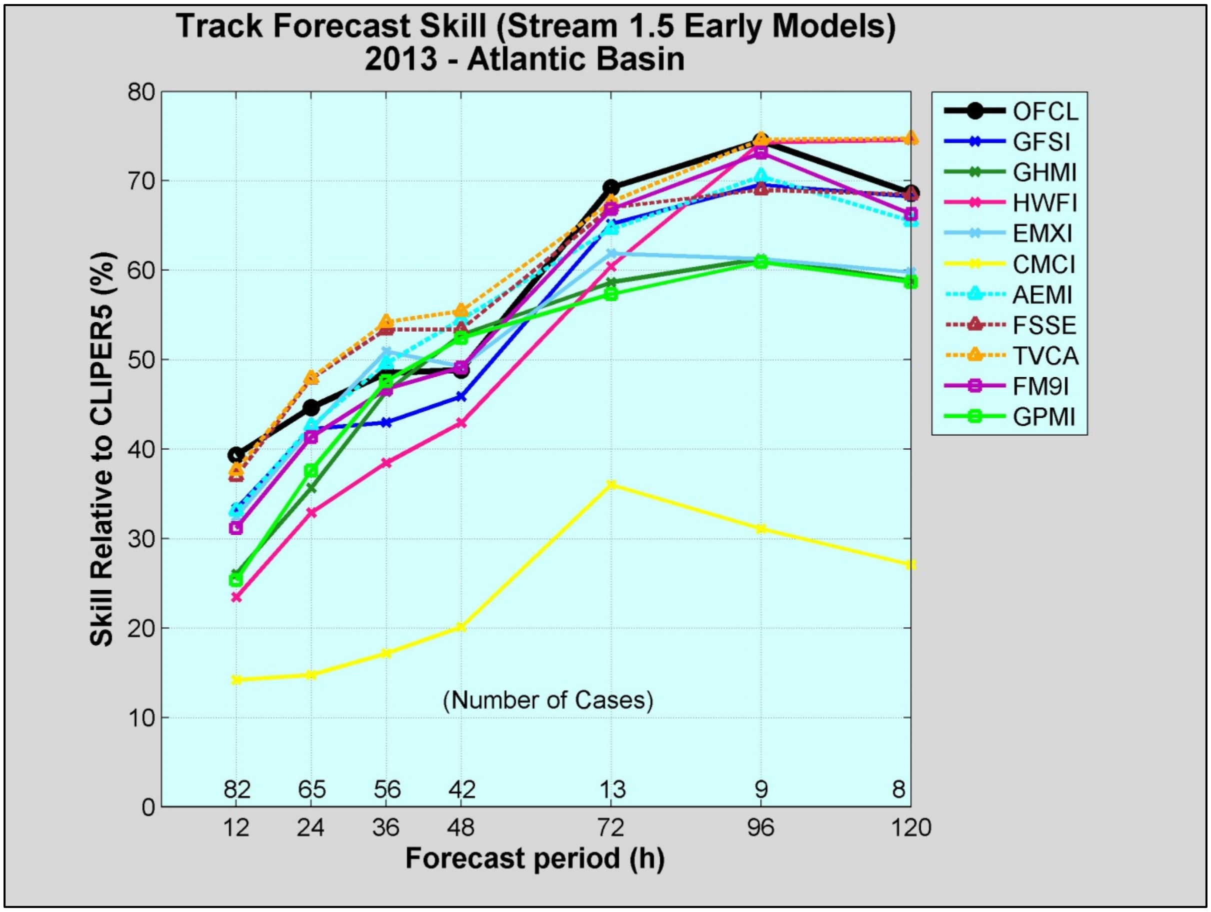


Figure 19. Homogeneous comparison of HFIP Stream 1.5 track models and selected operational models for 2013.

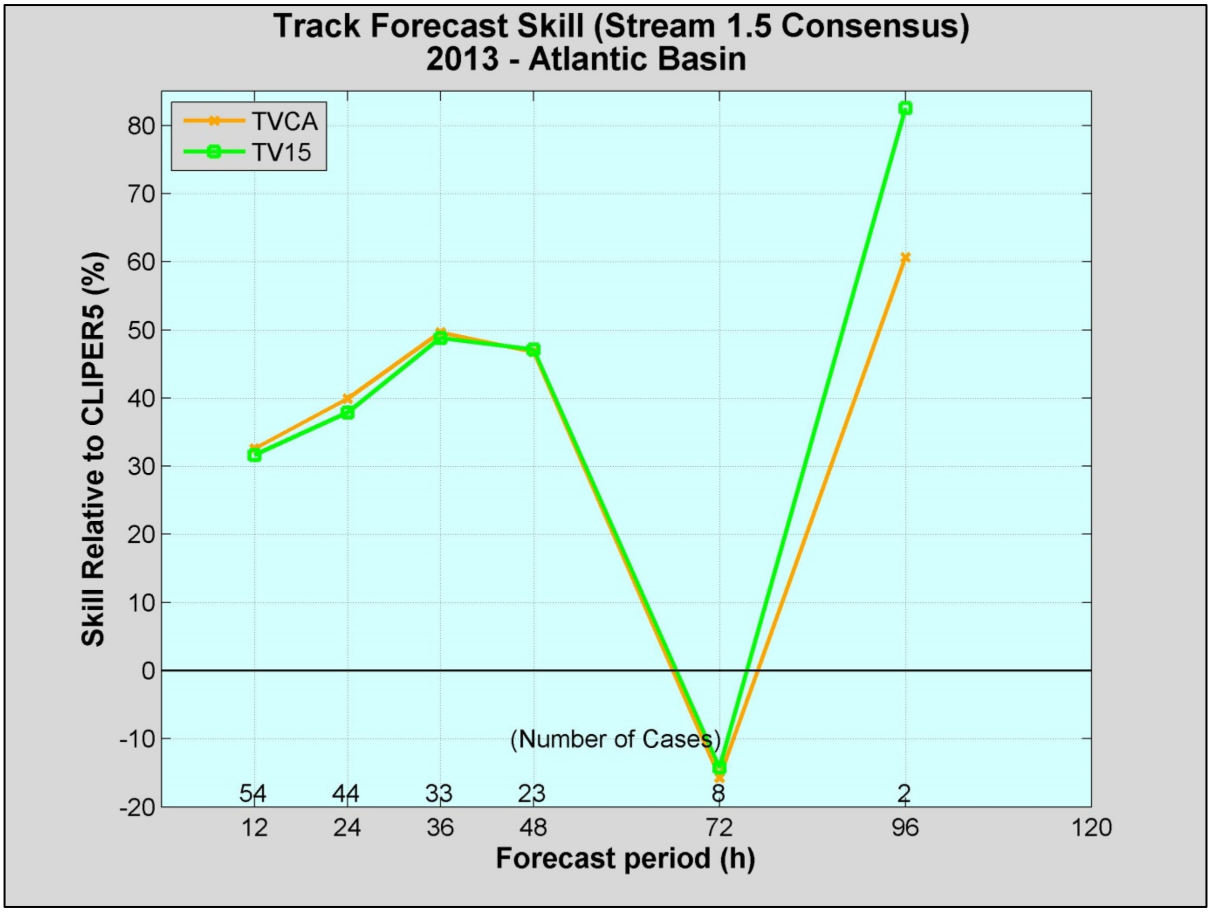


Figure 20. Impact of adding Stream 1.5 models to the variable track consensus TVCA.

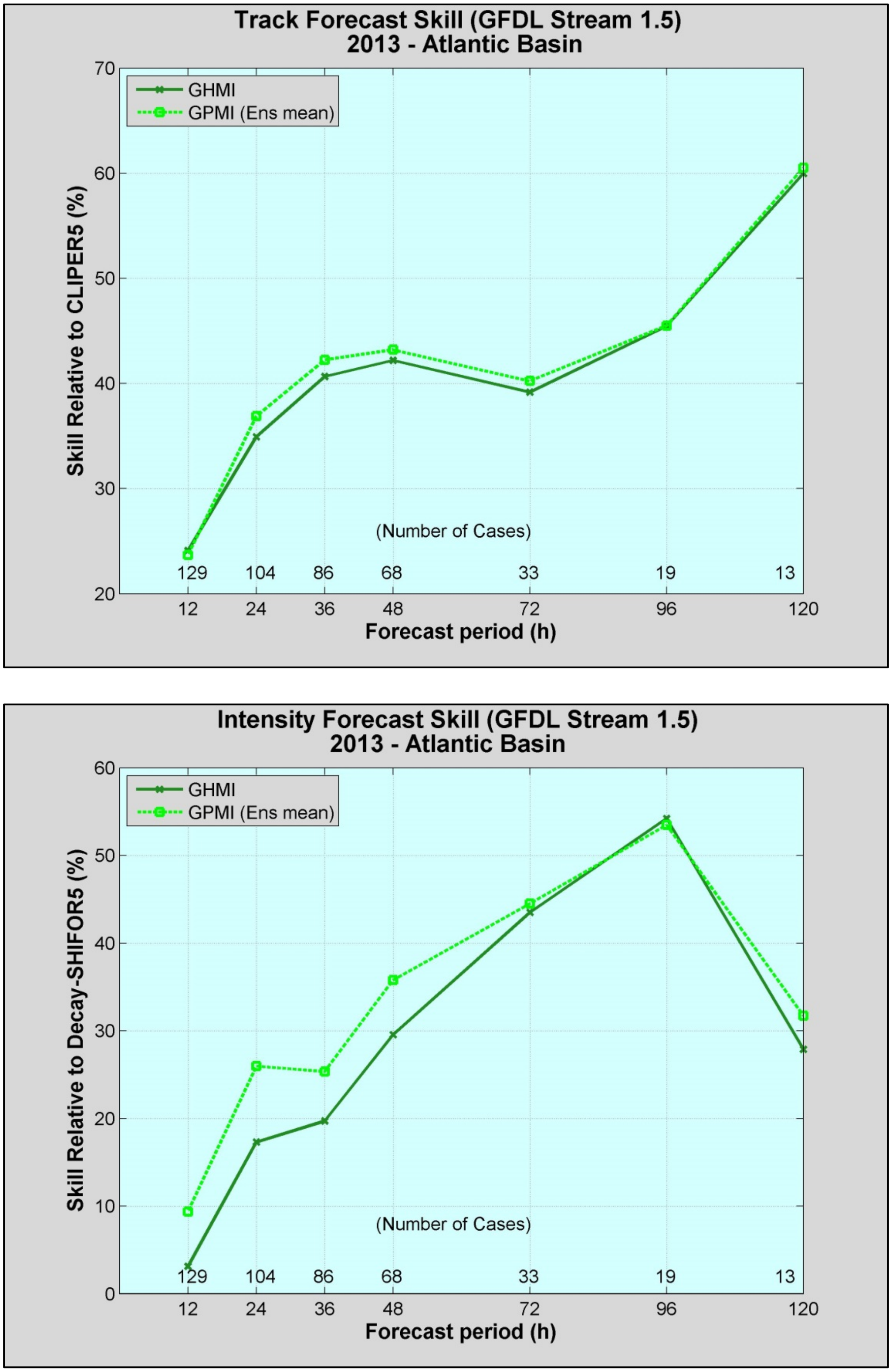
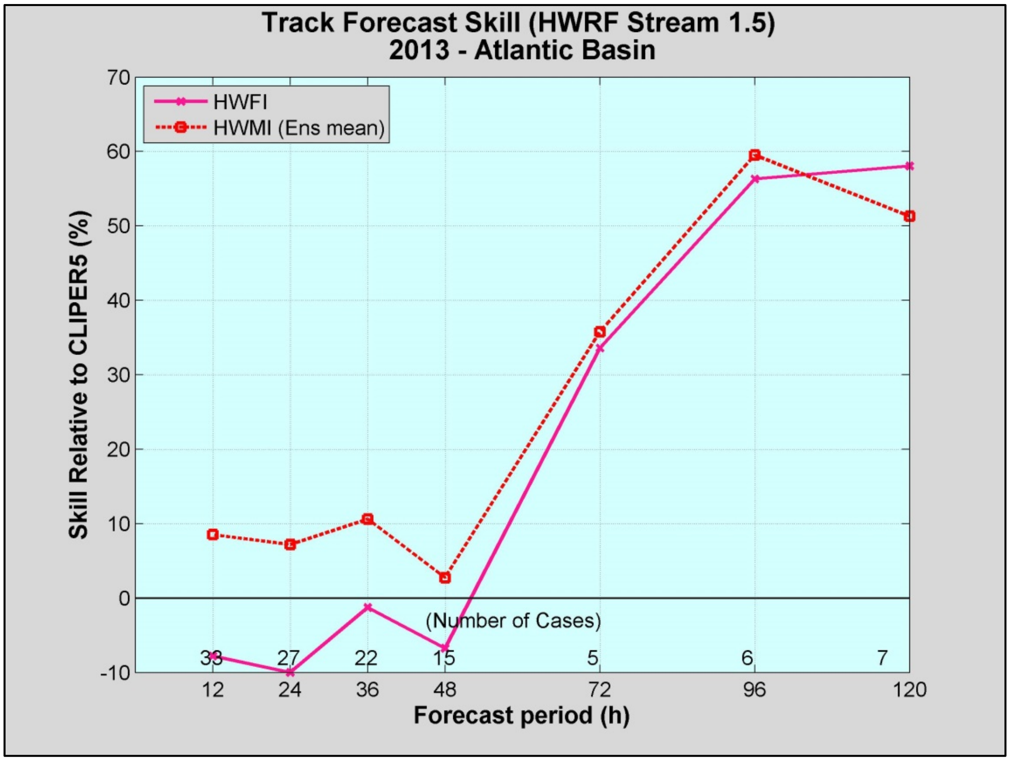


Figure 21. Homogeneous comparison of HFIP Stream 1.5 GFDL ensemble mean (GPMI) and GHMI for track (top) and intensity (bottom).



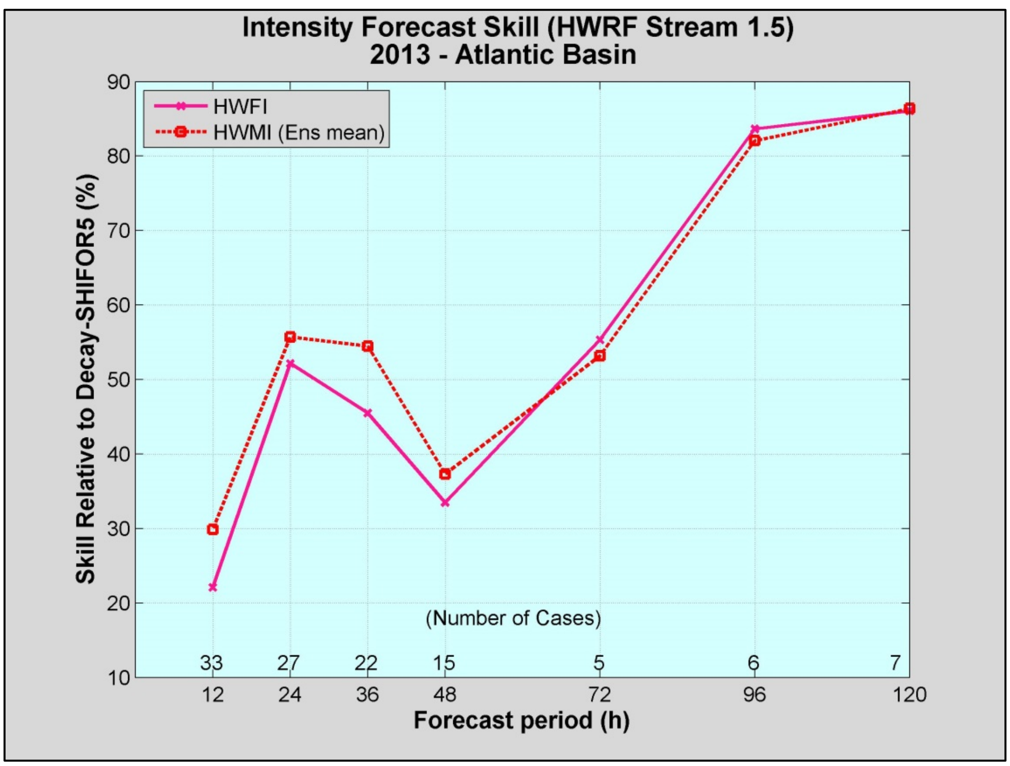


Figure 22. Homogeneous comparison of HFIP Stream 1.5 HWRF ensemble mean (HWMI) and HWFI for track (top) and intensity (bottom).

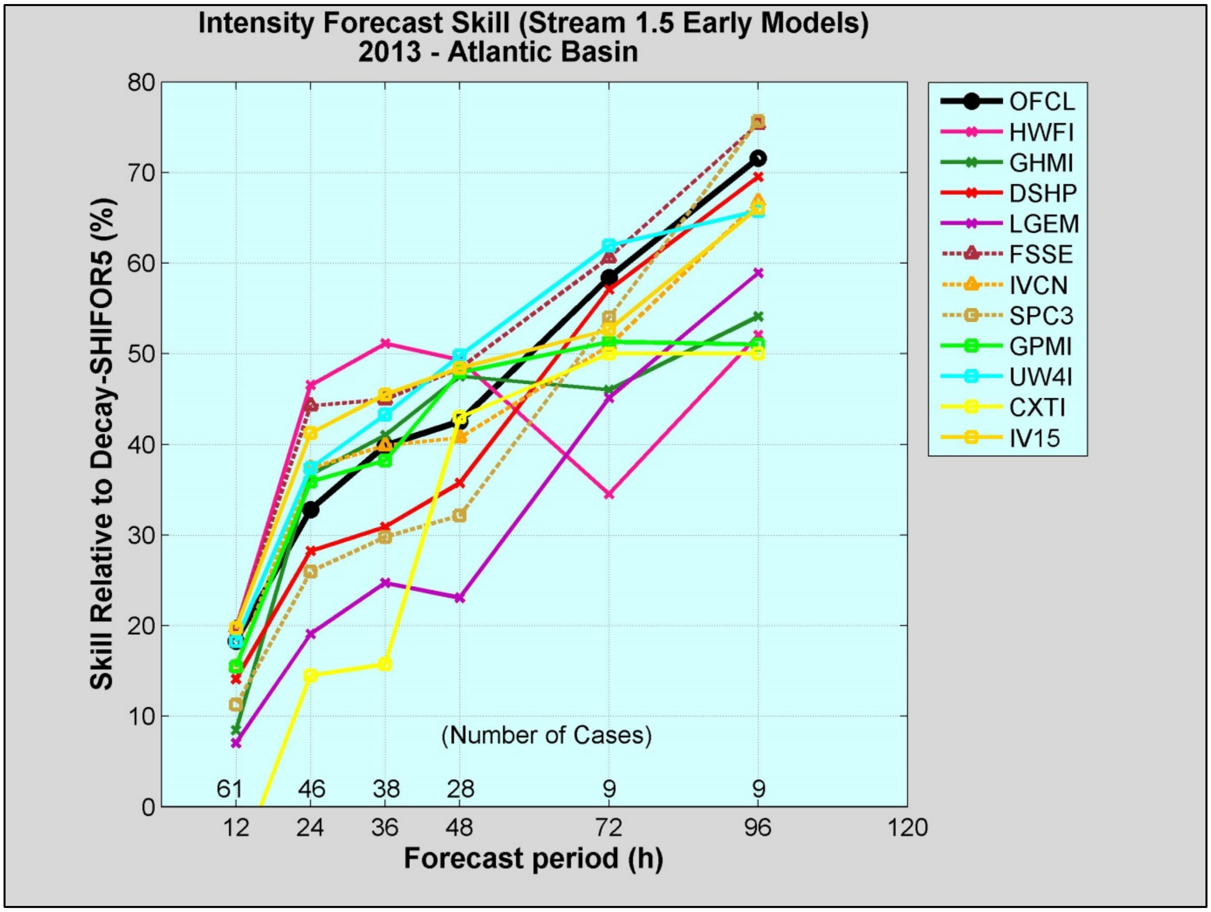


Figure 23. Homogeneous comparison of HFIP Stream 1.5 intensity models and selected operational models for 2013.

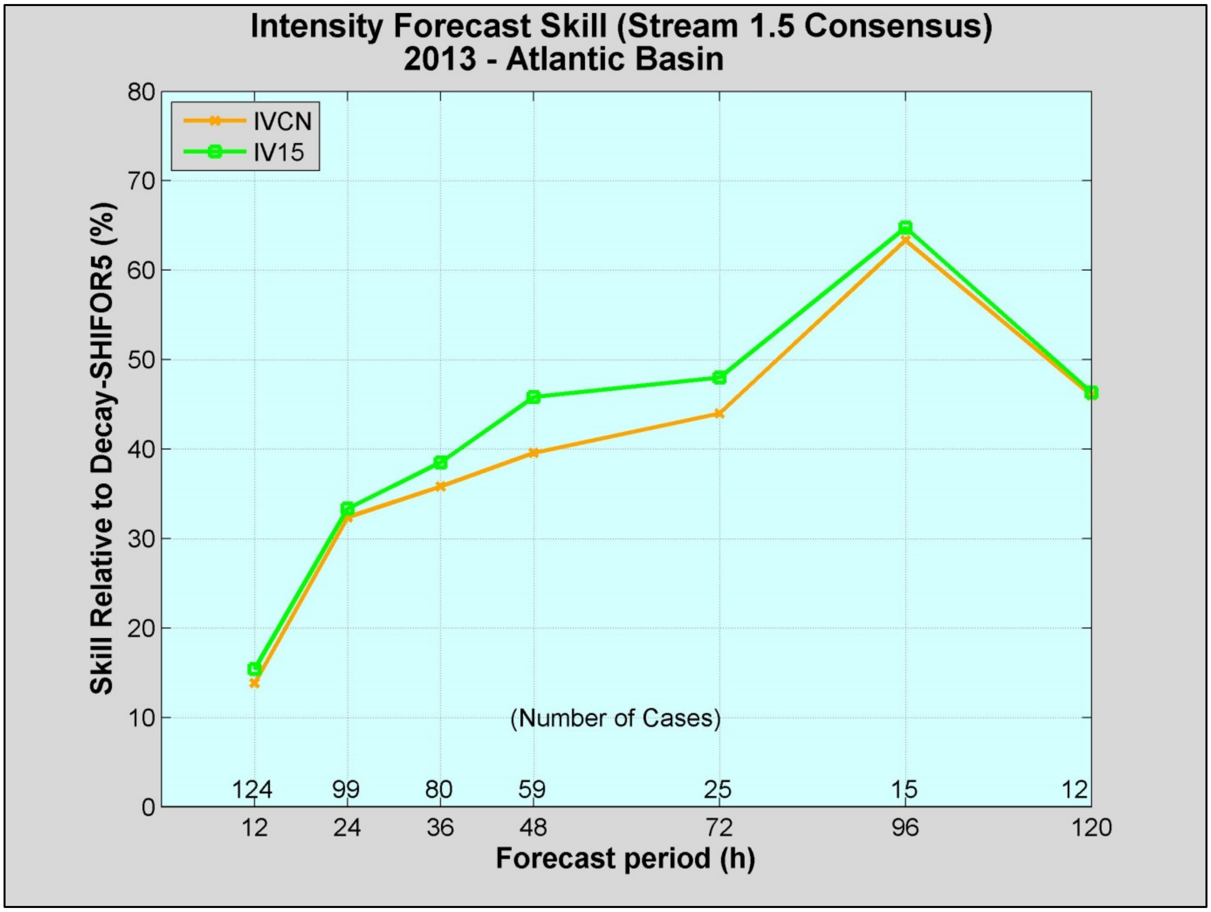


Figure 24. Impact of adding Stream 1.5 models to the variable intensity consensus IVCN.

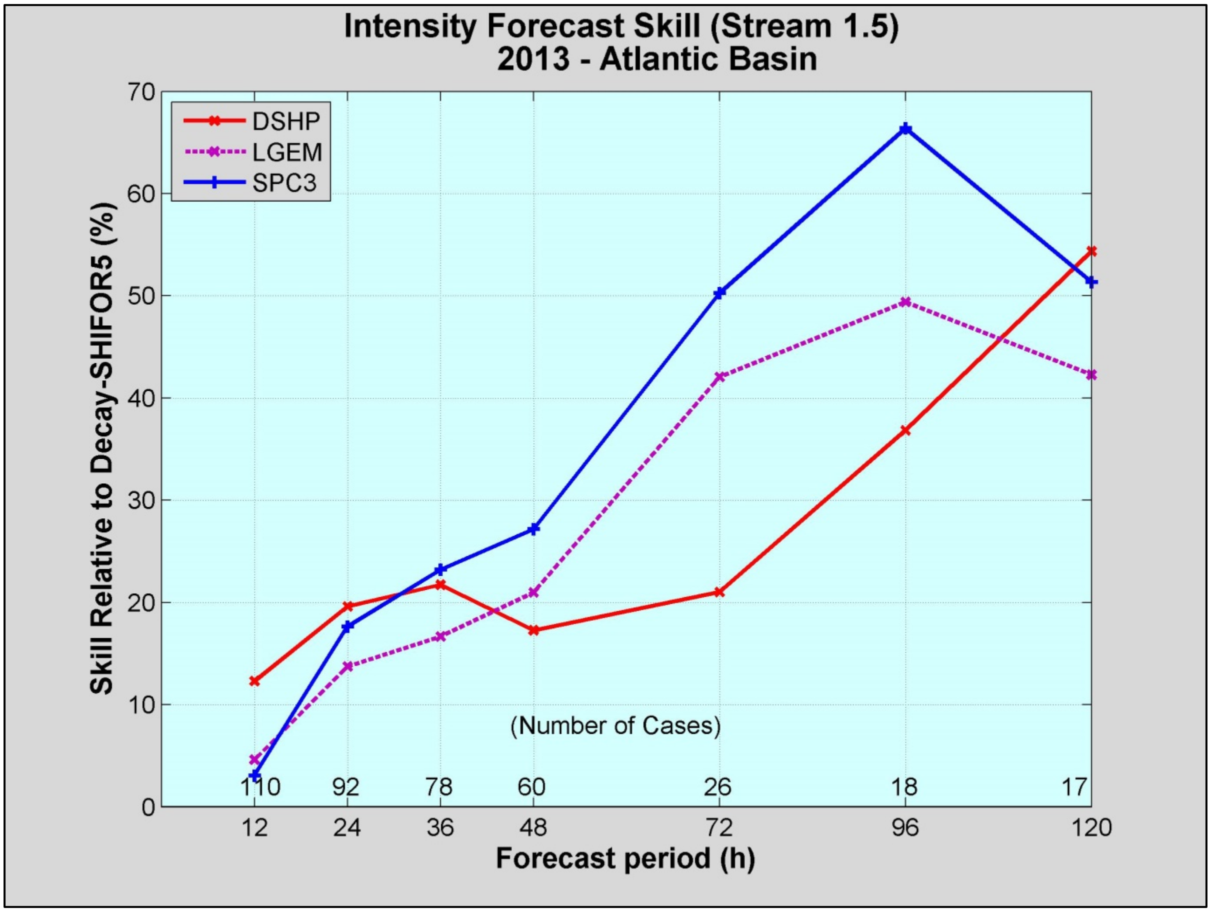


Figure 25. Homogeneous comparison of HFIP Stream 1.5 model SPC3 and operational models DSHP and LGEM.