2010 National Hurricane Center Forecast Verification Report

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

8 April 2011

Updated 5 May 2011 to correct Table 5a

ABSTRACT

The 2010 Atlantic hurricane season had above normal activity, with 404 official forecasts issued. The NHC official track forecast errors in the Atlantic basin were similar to the previous 5-yr means from 12-36 h, but up to 26% smaller beyond 36 h, and set a record for accuracy at 120 h. On average, the skill of the official forecasts was very close to that of the TCON/TVCN consensus models, as well as to the best performing of the dynamical models. The EMXI and GFSI exhibited the highest skill, and the EGRI performed well at longer forecast times. The NGPI and GFNI were the poorer performing major dynamical models. Among the consensus models, FSSE (a corrected consensus model) performed the best overall for the second year in a row. The corrected versions of TCON, TVCN, and GUNA, however, did not perform as well as their parent models. The Government Performance and Results Act of 1993 (GPRA) track goal was met.

Official intensity errors for the Atlantic basin in 2010 were above the 5-yr means at 12-48 h, but below the 5-yr means at the remaining lead times. Decay-SHIFOR errors in 2010 were above their 5-yr means at all forecast times, indicating the season's storms were more difficult to forecast than normal. Regarding the individual intensity guidance models, the consensus models ICON/IVCN were among the best performers at 12-48 h, with the LGEM showing similar or superior skill at 72-120 h. The FSSE skill was very near that of the ICON/IVCN at 12-72 h, but decreased sharply beyond that. The dynamical models were the worst performers, but did show competitive skill at longer forecast times. The GPRA intensity goal was not met.

There were 161 official forecasts issued in the eastern North Pacific basin in 2010, although only 29 of these verified at 120 h. This level of forecast activity was well below normal. NHC official track forecast errors set new records for accuracy at all forecast times, albeit for a small sample. Track forecast skill was also at an all-time high at all forecast times except 120 h. The official forecast skill was very close to the TVCN consensus and better than all of the dynamical models. Among the guidance models with sufficient availability, EMXI and GHMI were the best performers, with GFSI faring least well. There was a significant east-northeastward bias in the official forecast and some of the more reliable models at 96-120 h.

For intensity, the official forecast set records for accuracy at 72-96 h and skill was at record highs at all forecast times, again for a small sample. The official forecasts, in general, performed better than all of the eastern Pacific guidance through 48 h, with

FSSE being the most skillful model during that time. LGEM and DSHP were the best performers at 72-120 h. HWFI was an outlier and had negative skill at all forecast times, except 12 h.

Quantitative probabilistic forecasts of tropical cyclogenesis (i.e., the likelihood of tropical cyclone formation from a particular disturbance within 48 h) were made public for the first time in 2010. Forecasts were expressed in 10% increments and in terms of categories ("low", "medium", or "high"). Results for the four-year period 2007-10 indicate that the numerical probability forecasts are quite reliable in the Atlantic basin, but have a low (under-forecast) bias in the eastern North Pacific basin.

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.

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

[7] On very rare occasions, operational CLIPER or SHIFOR runs are missing from forecast databases. To ensure a complete 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.

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, so 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 h and 12 h interpolated models.[9]

A list of models is given in Table 1. In addition to their timeliness, models are characterized by their complexity or structure; this information is contained in the table for reference. Briefly, *dynamical* models forecast by solving the physical equations governing motions in the atmosphere. Dynamical models may treat the atmosphere either as a single layer (two-dimensional) or as having multiple layers (three-dimensional), and their domains may cover the entire globe or be limited to specific regions. The interpolated versions of dynamical model track and intensity forecasts are also sometimes referred to as dynamical models. *Statistical* models, in contrast, do not consider the characteristics of the current atmosphere explicitly but instead are based on historical relationships between storm behavior and various other parameters. *Statistical-dynamical* models are statistical in structure but use forecast parameters from dynamical models as predictors. *Consensus* models are not true forecast models *per se*, but are merely combinations of results from other models. One way to form a consensus is to simply average the results from a collection (or "ensemble") of models, but other, more

---

[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.

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 22 February 2011 for the Atlantic basin, and on 27 January 2011 for the eastern 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, which is a program that began experimentally in 2007 and became operational in 2010. Section 5 discusses the Hurricane Forecast Improvement Project (HFIP) Stream 1.5 activities in 2010. Section 6 summarizes the key findings of the 2010 verification and previews anticipated changes for 2011.

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

2. Atlantic Basin

a. 2010 season overview – Track

Figure 1 and Table 2 present the results of the NHC official track forecast verification for the 2010 season, along with results averaged for the previous 5-yr period, 2005-2009. In 2010, the NHC issued 404 Atlantic basin tropical cyclone forecasts[11], a number above the average over the previous five years (322). Mean track errors ranged from 34 n mi at 12 h to 187 n mi at 120 h. It is seen that mean official track forecast errors in 2010 were similar to the previous 5-yr mean from 12-36 h, but up to 26% smaller beyond 36 h. In addition, the 120 h official track forecast error set a record for accuracy. Over the past 15 years or so, 24-72 h track forecast errors have been reduced by about 50% (Fig. 2). Track forecast error reductions of about 40% have occurred over the past 10 years for the 96-120 h forecast periods. Vector biases were mostly eastward in 2010 (i.e., the official forecast tended to fall to the east of the verifying position). An examination of the track errors shows that the biases were primarily along-track and slow, but there was a cross-track bias as well. Track forecast skill in 2010 ranged from 36% at 12 h to 61% at 48 h and 72 h (Table 2). Note that the mean official error in Fig. 1 is not precisely zero at the 0 h (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

[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.

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 representative error.

Table 3a presents a homogeneous[12] verification for the official forecast along with a selection of early models for 2010. In order to maximize the sample size for comparison with the official forecast, a guidance model had to be available at least two-thirds of the time at both 48 h and 120 h. For the early track models, this requirement resulted in the exclusion of CMCI. Vector biases of the guidance models are given in Table 3b. This table shows that the official forecast had smaller biases than most of the model guidance. Results in terms of skill are presented in Fig. 3. The figure shows that official forecast skill was very close to that of the consensus models TVCN, TCON, and FSSE. In the Atlantic basin it is not uncommon for the best of the dynamical models to beat TVCN, and such was the case in 2010 beyond 36 h. The best-performing dynamical model in 2010 was EMXI, followed by GFSI. The GHMI, GF5I, HWFI, and EGRI made up the "second tier" of three-dimensional dynamical models, and NGPI and GFNI performed less well, with skill comparable to or even lower than the two-dimensional BAM collection.

A separate homogeneous verification of the primary consensus models is shown in Fig. 4. The figure shows that skill was about equal among the models through 72 h.

---

[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.

FSSE, a corrected-consensus model, performed the best at 96-120 h. The other corrected-consensus models (TCCN, TVCC, and CGUN) showed less skill than their respective parent models in 2010. In general, it has proven difficult to use the past performance of models to derive operational corrections: the sample of forecast cases is too small, the range of meteorological conditions is too varied, and model characteristics are insufficiently stable to produce a robust developmental data sample on which to base the corrections.

The GFS ensemble mean (AEMI) trailed the respective deterministic model (GFSI) at all time periods during 2010 (Fig. 3). While multi-model ensembles continue to provide consistently useful tropical cyclone guidance, the same cannot yet be said for single-model ensembles (although a four-year comparison of AEMI and GFSI shows roughly equivalent skill at 120 h).

Atlantic basin 48-h official track error, evaluated for all tropical cyclones[13], is a forecast measure tracked under the Government Performance and Results Act of 1993 (GPRA). In 2010, the GPRA goal was 90 n mi and the verification for this measure was 89.1 n mi.

b. 2010 season overview – Intensity

Figure 5 and Table 4 present the results of the NHC official intensity forecast verification for the 2010 season, along with results averaged for the preceding 5-yr period. Mean forecast errors in 2010 ranged from about 8 kt at 12 h to about 19 kt at 120 h. These errors were above the 5-yr means at 12-48 h but below the 5-yr means at the remaining lead times. Official forecasts had a moderate over-forecast bias in 2010.

[13] Prior to 2010, the GPRA measure was evaluated for tropical storms and hurricanes only.

Decay-SHIFOR5 errors were above their 5-yr means at all forecast times, indicating the season's storms were unusually difficult to forecast. Figure 6 shows that there has been virtually no net change in error over the past 15-20 years, although forecasts during the current decade have been more skillful than those from the previous one.

Table 5a presents a homogeneous verification for the official forecast and the primary early intensity models for 2010. Intensity biases are given in Table 5b, and forecast skill is presented in Fig. 7. The statistical-dynamical and consensus models had relatively similar skill. The consensus models ICON and IVCN were the most skillful in general, but were edged out by FSSE at 24 h and 48 h. FSSE, however, had much less skill at 96-120 h. LGEM showed similar or superior skill to ICON and IVCN at 72-120 h. The official forecasts on average beat all the guidance only at 12 h. The dynamical models were among the worst performers, but did show an increase in skill at longer forecast times. Table 5b shows that intensity model biases were small in 2010. Therefore, the moderate high bias in the official forecast was not due to poor model guidance.

An evaluation over the three years 2008-10 indicates that ICON is superior to all of the guidance, followed closely by LGEM (Fig. 8). Over this same period HWFI, a model that has underperformed during the past couple of years, contributed positively to the ICON consensus at most forecast times, and so will be retained as a component of ICON (and IVCN) in 2011.

The 48-h official intensity error, evaluated for all tropical cyclones, is another GPRA measure for the NHC. In 2010, the GPRA goal was 13 kt and the verification for this measure was 15.5 kt. Failure to reach the GPRA goal can be attributed in part to high forecast difficulty in 2010; as noted above, Decay-SHIFOR5 errors were well above their 5-yr means. The primary problem, however, is the GPRA goal itself, which was established based on the assumption that the HWRF model would immediately lead to forecast improvements. This has not occurred, of course, and only in 2003 were seasonal mean errors as low as the current GPRA goal of 13 kt. (And as it happens, the forecast *skill* in 2003 was not particularly high.) It is reasonable to assume that until there is some modeling or conceptual breakthrough, annual official intensity errors are mostly going to rise and fall with forecast difficulty, and therefore routinely fail to meet GPRA goals.

c. Verifications for individual storms

Forecast verifications for individual storms are given in Table 6. Of note are the large track errors at 72-120 h for Lisa, which were more than double the long-term means. These large errors were associated with the difficulty in predicting the slow eastward and then northward motion from 22 to 25 September; most of the guidance models predicted a northwestward track during this time. Intensity forecasts for Karl had large errors, primarily due to missing the rapid intensification and the incorrect timing of the rapid weakening over land. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2010atlan.shtml.

3. Eastern North Pacific Basin

a. 2010 season overview – Track

The NHC track forecast verification for the 2010 season in the eastern North Pacific, along with results averaged for the previous 5-yr period is presented in Figure 9 and Table 7. There were 161 forecasts issued for the eastern Pacific basin in 2010, although only 29 of these verified at 120 h. This level of forecast activity was well below average. In fact, 2010 was the least active year in terms of the number of tropical storms and hurricanes since routine satellite reconnaissance of the basin began in 1971. Track forecast accuracy for 2010 was exceptionally good. Mean track errors ranged from 26 n mi at 12 h to 145 n mi at 120 h, and were unanimously lower, between 15% and 39%, than the 5-yr means. New records were set at all forecast times; however, the sample in 2010 was unusually small so these results could be anomalous. In addition, CLIPER5 errors were below their long-term means, implying that forecast difficulty in 2010 was lower than normal. Forecast biases were near average through 48 h but above average at 72-120 h; biases at the latter times were about 60% of the mean error magnitude, and directed toward the east-northeast. Blas and Frank were major contributors to these biases.

Figure 10 shows recent trends in track forecast accuracy and skill for the eastern North Pacific. Errors have been reduced by roughly 35-60% for the 24-72 h forecasts since 1990 a somewhat smaller but still substantial, improvement relative to what has occurred in the Atlantic. Forecast skill in 2010 was about 10-15 % higher than in 2009 and set new records at all forecast times except 120 h, albeit again for a small sample.

Nevertheless, this season has accelerated the upward trend in forecast skill that began in the late 1990's.

Table 8a presents a homogeneous verification for the official forecast and the early track models for 2010, with vector biases of the guidance models given in Table 8b. Skill comparisons of selected models are shown in Fig. 11. Note that the sample becomes very small (only 9 cases) by 120 h. Several models (EGRI, CMCI, GUNA, and TCON) were eliminated from this evaluation because they did not meet the two-thirds availability threshold. Among the surviving three-dimensional dynamical models, the EMXI had the lowest errors from the 12-72 h period. The GFNI and HWFI showed increased skill at the longer ranges and fared best among the individual models at 96-120 h. In contrast to its strong performance in the Atlantic, the GFSI showed the least skill overall among the dynamical models in the eastern Pacific. The multi-model consensus TVCN provided value over the models it comprises; indeed, the power of a multi-model consensus traditionally is much stronger for the eastern North Pacific than for the Atlantic. The skill of the official forecasts was very close to that of TVCN.

A separate verification of the primary consensus aids is given in Figure 12. TVCN performed best overall – better than either of the corrected consensus models (FSSE and TVCC), and significantly better than the GFS ensemble mean (AEMI). AEMI, however, was superior to its deterministic run at all forecast times and was comparable to FSSE and TVCC at 120 h in 2010. An evaluation over the three-years 2008-10 (not shown) indicates that the superior performance of AEMI over the GFSI in 2010 was an anomaly and the models appear to have nearly equivalent skill.

b. 2010 season overview – Intensity

Figure 13 and Table 9 present the results of the NHC eastern North Pacific intensity forecast verification for the 2010 season, along with results averaged for the preceding 5-yr period. Mean forecast errors were 6 kt at 12 h and increased to 18 kt by 120 h. The errors were lower than the 5-yr means, up to 15%, at all forecast times. The SHIFOR5 forecast errors were substantially higher than their 5-yr means; this implies that forecast difficulty in 2010 was higher than normal. Despite this, all-time low errors were set at 72-96 h. A review of error and skill trends (Fig. 14) indicates that skill improved dramatically in 2010 and is at the highest point on record at all forecast times, albeit for a unusually small sample. Intensity forecast biases in 2010 were small and generally positive.

Figure 15 and Table 10a present a homogeneous verification for the primary early intensity models for 2010. Forecast biases are given in Table 10b. The official forecasts, in general, were more skillful than all of the eastern Pacific guidance through 48 h, with the FSSE being the best model during that time. The statistical-dynamical guidance (DSHP and/or LGEM) were the best models at 72-120 h. The GHMI performed reasonably well, having skill at all forecast times, but showed less skill than the statistical-dynamical and consensus models. The outlier was the HWFI, which only had skill at 12 h and had negative biases of 30 kt or higher at 96-120 h. The slight high bias in the official forecast at 36-48 h was generally lower than that of the Decay-SHIFOR5, but higher than the FSSE and LGEM. Interestingly, most of the dynamical and consensus models had strong negative biases at 72-120 h.

c. Verifications for individual storms

Forecast verifications for individual storms are given for reference in Table 11. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2010epac.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 during the following 48 hours. In 2007, the NHC began producing in-house (non-public) experimental probabilistic tropical cyclone genesis forecasts. 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 subjective determination of the chance of TC formation during the 48 h period following the nominal TWO issuance time. These probabilities became available to the public in 2010. Verification was 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 for the Atlantic and eastern North Pacific basins for 2010 are given in Table 12 and illustrated in Fig. 16. In the Atlantic basin, the forecasts were extremely well calibrated at the low and high probabilities, but forecasters were not able to discern gradations in threat from 40-70%. In the eastern Pacific, once the forecast likelihood exceeded 20%, there appeared to be minimal correlation between the forecast and verifying rates (although the sample is small at those frequencies).

Combined results for the four-year period 2007-10 are given in Table 13. For the four-year sample, there is a slight over-forecast bias at 50% and higher forecast likelihood for the Atlantic basin forecasts, but is otherwise very well calibrated. Results for the eastern North Pacific are not quite as good, with a general under-forecast bias in the middle probabilities. Even so, the forecasters were clearly able to distinguish gradations in genesis likelihood (evidenced by the nearly monotonic increase of the verifying percentage with forecast percentage). 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". Two models, GF5I and AHWI, were provided to NHC in 2010 under Stream 1.5. GF5I represented a set of upgrades to the operational GFDL model that occurred too late to be implemented operationally. AHWI derives from version 3.2 of the WRF ARW. In 2010 it was run with a 1.33 km resolution inner nest with 36 vertical levels, and was initialized using an ensemble Kalman filter approach.

The Stream 1.5 GF5I performance overall was disappointing. Figure 17 shows the track skill for GF5I in both basins, along with the skill of the operational GFDL (GHMI). It is seen that the Stream 1.5 model performance in the Atlantic was virtually identical to that of the model it was meant to improve upon, while in the eastern Pacific the Stream 1.5 model actually performed worse than its operational counterpart. For intensity, GF5I improved upon the operational model at 96 and 120 h but was less skillful at the shorter forecast leads in the Atlantic. In the eastern Pacific, the Stream 1.5 model was less skillful than the operational version at all time periods (Fig. 18).

The AHWI was run only for the Atlantic basin, and only a small sample of forecasts (44) was provided. The performance of AHWI was also disappointing. Figure 19 shows how the AHWI compared against the traditionally top-performing track models. AHWI lagged all of them, except for GHMI through 36 h. Figure 20 compares the AHWI intensity forecasts against the best operational intensity guidance. AHWI lacked skill, trailing the statistical models badly and beating GHMI only at 96 h.

6. Looking Ahead to 2011

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-year sample fall within the circle. The circle radii defining the cones in 2011 for the Atlantic and eastern North Pacific basins (based on error distributions for 2006-10) are in Table 14. In the Atlantic basin, the cone circles will be up to 16% smaller than they were last year, with the biggest decrease at 96-120 h. In the eastern Pacific basin, the cone circles will be similar to what they were last year and are now about the same size as in the Atlantic basin.

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., GUNA), while others are less restrictive, requiring only two or more members to be present (e.g., TVCN). 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 has changed for 2011. The NGPI model has degraded the skill of the multi-model consensus models TVCN and TCON during the past few seasons in the Atlantic basin, but has contributed positively to those models in the eastern North Pacific basin. To account for this discrepancy, two versions of the TVCN and TCON will be implemented for 2011. TVCA and TCOA, which do not include the NGPI model, are intended for use in the Atlantic basin. TVCE and TCOE, which do include the NGPI model, are intended for use in the eastern Pacific basin. A summary of the consensus model composition is given in Table 15.

Acknowledgments:

The authors gratefully acknowledge Chris Sisko of NHC, keeper 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
NGPS	Navy Operational Global Prediction System	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
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
NGPI	Previous cycle NGPS, 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, NGPI, and GFSI	Consensus	E	Trk
CGUN	Version of GUNA corrected for model biases	Corrected consensus	E	Trk
AEMI	Previous cycle AEMN, adjusted	Consensus	E	Trk, Int

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
FSSE	FSU Super-ensemble	Corrected consensus	E	Trk, Int
TCON	Average of GHMI, EGRI, NGPI, 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 NGPI GHMI HWFI GFNI 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

Table 2. Homogenous comparison of official and CLIPER5 track forecast errors in the Atlantic basin for the 2010 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
2010 mean OFCL error (n mi)	34.2	54.2	71.6	89.1	129.4	166.0	186.7
2010 mean CLIPER5 error (n mi)	53.0	107.5	165.9	229.8	332.2	400.1	469.7
2010 mean OFCL skill relative to CLIPER5 (%)	35.5	49.6	56.8	61.2	61.0	58.5	60.3
2010 mean OFCL bias vector (°/n mi)	052/4	027/7	019/8	034/10	120/8	116/23	098/32
2010 number of cases	365	327	292	259	198	149	115
2005-2009 mean OFCL error (n mi)	31.8	53.4	75.4	96.8	143.8	195.6	252.1
2005-2009 mean CLIPER5 error (n mi)	46.9	97.3	155.4	211.6	304.8	387.9	467.8
2005-2009 mean OFCL skill relative to CLIPER5 (%)	32.2	45.1	51.5	54.3	52.8	49.6	46.1
2005-2009 mean OFCL bias vector (°/n mi)	307/5	307/11	311/18	314/24	308/22	347/13	034/31
2005-2009 number of cases	1459	1300	1145	1013	802	618	479
2010 OFCL error relative to 2005-2009 mean (%)	7.5	1.5	-5.0	-8.0	-10.0	-15.1	-25.9
2010 CLIPER5 error relative to 2005-2009 mean (%)	13.0	10.5	6.8	8.6	9.0	3.1	0.4

Table 3a. Homogenous comparison of Atlantic basin early track guidance model errors (n mi) for 2010. 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	31.0	50.9	68.5	84.1	126.2	175.0	195.9
OCD5	48.0	100.3	156.6	209.1	305.2	384.2	428.9
GFSI	**30.8**	**49.8**	**66.9**	**84.0**	130.0	185.9	236.3
GHMI	35.8	58.8	82.4	100.9	143.0	207.2	288.4
GF5I	36.5	57.6	80.1	98.5	138.3	208.6	309.7
HWFI	36.7	62.9	89.1	112.8	160.8	204.0	263.6
GFNI	40.5	75.9	110.9	138.9	187.6	278.2	373.1
NGPI	38.4	68.2	98.5	125.3	181.7	251.2	329.8
EGRI	37.1	60.2	85.1	106.7	143.9	185.5	**190.9**
EMXI	31.5	**50.0**	**66.0**	**81.8**	**113.2**	**160.6**	215.0
AEMI	32.5	54.1	74.4	92.3	139.7	204.8	266.8
FSSE	**29.9**	**49.1**	**67.4**	85.2	**122.2**	**164.6**	208.5
TCON	31.2	51.0	69.0	86.1	**125.0**	176.5	225.9
TCCN	32.0	51.3	69.9	87.6	130.6	189.0	228.5
TVCN	**30.3**	**50.0**	**68.4**	84.1	**120.2**	**169.9**	225.9
TVCC	31.2	**50.3**	68.7	85.2	127.7	185.6	234.4
LBAR	42.3	81.2	126.9	175.9	233.3	272.0	297.2
BAMS	52.4	96.6	139.5	170.2	228.5	249.8	307.3
BAMM	42.9	76.8	110.6	134.0	177.6	205.8	249.0
BAMD	45.8	84.6	119.7	156.8	201.0	263.8	354.8
# Cases	230	202	181	152	115	85	54

Table 3b. Homogenous comparison of Atlantic basin early track guidance model bias vectors (º/n mi) for 2010.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	043/005	023/010	011/013	016/014	089/013	087/019	073/040
OCD5	348/002	308/004	247/013	242/016	093/014	072/056	064/092
GFSI	046/007	032/010	031/014	034/017	077/027	072/031	049/080
GHMI	097/012	069/019	052/025	043/034	037/032	043/052	046/132
GF5I	098/012	067/019	051/027	042/033	044/040	052/067	050/147
HWFI	042/015	029/029	027/041	031/049	057/063	068/081	080/087
GFNI	284/009	278/023	277/035	273/043	244/058	319/025	040/116
NGPI	338/008	317/015	298/018	272/017	173/027	142/047	087/094
EGRI	336/004	309/009	292/012	268/016	208/024	201/051	206/051
EMXI	059/003	022/004	025/005	026/006	064/024	062/050	067/099
AEMI	016/009	006/018	006/027	010/033	034/038	041/048	044/095
FSSE	057/007	035/011	031/015	033/014	116/015	128/019	086/036
TCON	042/007	018/012	015/016	018/017	080/019	093/030	070/067
TCCN	089/009	088/014	092/018	106/022	135/038	161/048	167/058
TVCN	028/005	358/009	351/012	350/013	087/009	079/026	063/077
TVCC	087/008	086/011	094/011	110/019	136/039	141/040	111/074
LBAR	046/017	018/039	009/056	007/076	018/078	061/079	087/065
BAMS	284/009	273/022	260/037	260/051	239/051	232/033	140/051
BAMM	046/005	025/010	011/007	355/010	108/010	072/023	053/061
BAMD	066/016	054/035	051/047	049/062	058/069	049/105	031/172
# Cases	230	202	181	152	115	85	54

Table 4. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the Atlantic basin for the 2010 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
2010 mean OFCL error (kt)	7.6	12.0	13.9	15.5	16.7	18.4	18.6
2010 mean Decay-SHIFOR5 error (kt)	8.9	13.4	16.9	19.9	23.2	25.4	26.3
2010 mean OFCL skill relative to Decay-SHIFOR5 (%)	14.6	10.4	17.8	22.1	28.0	27.6	29.3
2010 OFCL bias (kt)	0.2	1.4	2.4	3.2	3.8	2.9	3.1
2010 number of cases	365	327	292	259	198	149	115
2005-9 mean OFCL error (kt)	7.0	10.7	13.1	15.2	18.6	18.7	20.1
2005-9 mean Decay-SHIFOR5 error (kt)	8.6	12.5	15.8	18.2	21.0	22.7	21.7
2005-9 mean OFCL skill relative to Decay-SHIFOR5 (%)	18.6	14.4	17.1	16.5	11.4	17.6	7.4
2005-9 OFCL bias (kt)	0.1	0.5	0.0	-0.5	-0.8	-2.1	-2.0
2005-9 number of cases	1459	1300	1145	1013	802	618	479
2010 OFCL error relative to 2005-9 mean (%)	7.9	12.1	6.1	2.0	-10.2	-1.6	-7.5
2010 Decay-SHIFOR5 error relative to 2005-9 mean (%)	3.5	7.2	7.0	9.3	10.5	11.9	21.2

Table 5a. Homogenous comparison of selected Atlantic basin early intensity guidance model errors (kt) for 2010. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	8.1	12.2	14.0	15.9	16.7	18.6	18.1
OCD5	9.3	13.4	16.6	19.7	22.6	24.6	23.5
HWFI	9.8	13.3	15.5	17.0	19.6	23.7	20.6
GHMI	9.4	13.3	16.1	17.5	19.1	19.8	18.1
GF5I	9.7	13.7	16.6	18.5	19.4	**18.1**	**16.8**
GFNI	10.2	13.9	16.0	17.5	20.3	20.9	**17.9**
DSHP	8.7	12.5	14.5	15.9	**16.2**	**16.7**	**16.4**
LGEM	8.6	**12.1**	14.2	**15.2**	**15.7**	**16.0**	**15.6**
ICON	8.5	**11.8**	**13.7**	**14.8**	**15.8**	**17.2**	**15.7**
IVCN	8.6	**11.7**	**13.3**	**14.6**	**15.7**	**17.1**	**15.5**
FSSE	8.5	**11.7**	**13.4**	**14.4**	**16.0**	19.7	20.0
# Cases	285	254	225	190	140	107	78

Table 5b. Homogenous comparison of selected Atlantic basin early intensity guidance model biases (kt) for 2010. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	0.7	2.4	3.7	3.7	4.8	3.8	5.3
OCD5	**-0.2**	**1.1**	**1.0**	**-0.8**	**-4.4**	-8.0	-8.4
HWFI	-4.0	-5.2	-4.7	-4.6	**-2.7**	**-2.1**	**-1.3**
GHMI	**-0.5**	**0.6**	**2.6**	**2.5**	**1.5**	**0.9**	**1.6**
GF5I	-0.8	**0.9**	**3.0**	**2.8**	**2.8**	**2.8**	**1.9**
GFNI	-1.7	-3.0	**-3.2**	**-3.3**	-4.9	**-3.5**	**0.7**
DSHP	**0.6**	3.0	4.5	4.6	**4.4**	**1.4**	**1.1**
LGEM	**0.3**	**1.9**	**2.6**	**2.0**	**2.8**	**1.2**	**1.3**
ICON	**-0.6**	**0.3**	**1.5**	**1.4**	**1.7**	**0.6**	**0.9**
IVCN	-0.8	**-0.2**	**0.7**	**0.6**	**0.5**	**-0.1**	**0.9**
FSSE	**-0.3**	**1.4**	**3.0**	**2.2**	**0.9**	**-2.7**	-5.8
# Cases	285	254	225	190	140	107	78

Table 6. Official Atlantic track and intensity forecast verifications (OFCL) for 2010 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: AL012010 ALEX

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	26	6.3	6.3	26	2.5	3.1
012	24	34.3	53.3	24	5.4	7.3
024	22	51.4	96.8	22	6.8	9.7
036	20	67.2	129.6	20	5.3	14.4
048	18	84.3	161.4	18	6.4	14.4
072	14	105.8	183.7	14	10.7	26.2
096	10	104.7	179.5	10	17.5	36.7
120	6	103.8	246.8	6	25.0	32.5

Verification statistics for: AL022010 TWO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	4	16.2	22.7	4	0.0	0.0
012	3	72.9	85.7	3	5.0	6.3
024	1	65.0	124.3	1	15.0	18.0
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: AL032010 BONNIE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	10	1.9	1.9	10	0.5	0.5
012	8	35.2	55.0	8	5.0	6.4
024	6	51.7	142.6	6	12.5	10.3
036	4	69.5	269.5	4	16.3	17.3
048	2	98.4	382.7	2	17.5	20.0
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: AL042010 COLIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	16	15.9	16.2	16	0.9	1.3
012	14	55.9	82.4	14	4.3	5.3
024	12	104.6	174.0	12	9.6	8.6
036	10	136.8	274.6	10	13.5	12.6
048	8	162.2	369.8	8	16.3	15.6
072	5	154.6	529.7	5	8.0	16.0
096	5	133.6	491.2	5	9.0	26.2
120	4	180.6	368.2	4	15.0	34.5

Verification statistics for: AL052010 FIVE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	2	12.1	12.1	2	0.0	2.5
012	0	-999.0	-999.0	0	-999.0	-999.0
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: AL062010 DANIELLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	37	9.4	9.4	37	2.6	2.7
012	35	38.7	62.3	35	7.4	8.6
024	33	65.7	130.2	33	12.7	13.0
036	31	80.6	188.8	31	14.8	16.2
048	29	92.7	257.3	29	13.6	17.1
072	25	142.7	270.6	25	12.2	15.7
096	21	233.8	262.1	21	13.8	18.0
120	17	275.1	415.7	17	15.6	27.2

Verification statistics for: AL072010 EARL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	42	13.7	13.8	42	1.9	2.3
012	40	23.1	36.4	40	7.3	8.3
024	38	34.8	74.4	38	10.5	12.7
036	36	49.4	121.3	36	12.1	15.4
048	34	69.4	174.7	34	14.0	19.0
072	30	115.0	289.3	30	15.2	24.4
096	26	173.0	406.8	26	15.8	28.7
120	22	232.9	535.3	22	16.8	25.2

Verification statistics for: AL082010 FIONA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	17.7	17.7	17	2.6	2.9
012	15	39.1	61.9	15	4.7	8.4
024	13	48.2	115.5	13	7.3	12.3
036	11	54.7	184.2	11	7.3	13.0
048	9	74.5	279.9	9	5.6	14.6
072	5	106.7	422.5	5	5.0	32.4
096	1	164.3	660.6	1	5.0	11.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL092010 GASTON

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	24.4	26.8	5	3.0	3.0
012	3	28.6	73.5	3	8.3	9.3
024	1	61.1	128.6	1	5.0	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: AL102010 HERMINE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	9	14.7	16.2	9	3.3	3.3
012	9	43.2	77.0	9	3.9	5.8
024	9	64.1	139.1	9	4.4	8.0
036	9	99.2	194.1	9	3.3	12.0
048	7	126.8	289.5	7	2.9	20.4
072	3	206.9	461.4	3	0.0	29.7
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL112010 IGOR

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	53	10.2	10.5	53	1.6	1.8
012	51	29.8	42.8	51	9.8	9.5
024	49	45.9	88.5	49	15.0	13.8
036	47	63.4	146.4	47	17.4	16.6
048	45	82.8	202.5	45	19.0	20.4
072	41	124.0	315.3	41	22.3	25.3
096	37	139.4	384.6	37	23.5	30.1
120	33	158.2	442.4	33	22.1	31.8

Verification statistics for: AL122010 JULIA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	33	11.5	11.8	33	2.9	2.9
012	31	34.0	61.9	31	6.9	9.0
024	29	57.1	129.5	29	12.8	15.3
036	27	71.7	190.8	27	16.7	22.1
048	25	79.1	245.2	25	18.0	27.8
072	21	92.6	325.5	21	17.6	35.4
096	17	116.4	390.1	17	10.3	27.8
120	13	128.1	396.1	13	5.8	23.2

Verification statistics for: AL132010 KARL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	2.1	2.1	15	1.0	1.0
012	13	25.7	30.1	13	14.2	12.2
024	11	54.2	57.2	11	20.9	22.2
036	9	83.6	89.1	9	31.7	31.0
048	7	120.2	124.9	7	33.6	42.3
072	3	159.2	273.3	3	30.0	29.7
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL142010 LISA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	14.3	13.8	23	3.7	3.9
012	21	39.2	47.4	21	9.5	10.8
024	19	74.8	97.7	19	17.9	18.6
036	17	122.6	163.6	17	19.7	20.1
048	15	183.3	247.8	15	17.3	16.5
072	11	332.2	370.5	11	10.5	11.0
096	7	439.3	500.7	7	12.1	15.4
120	3	464.4	568.7	3	13.3	18.0

Verification statistics for: AL152010 MATTHEW

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	12	16.7	16.7	12	1.3	1.3
012	10	54.9	54.9	10	5.5	6.8
024	8	94.0	102.3	8	5.6	5.3
036	6	129.2	148.6	6	10.0	5.7
048	4	201.8	221.6	4	31.3	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: AL162010 NICOLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	20.2	20.2	5	7.0	7.0
012	3	62.8	44.0	3	1.7	6.0
024	1	97.4	92.7	1	0.0	0.0
036	0	-999.0	-999.0	0	-999.0	-999.0
048	0	-999.0	-999.0	0	-999.0	-999.0
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: AL172010 OTTO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	16	11.2	11.2	16	2.5	2.2
012	14	20.7	54.8	14	6.4	9.6
024	12	32.7	152.5	12	8.3	11.4
036	10	46.3	291.1	10	8.0	12.5
048	8	42.5	452.8	8	10.0	10.8
072	4	72.6	672.6	4	15.0	15.3
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: AL182010 PAULA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	15	5.7	6.1	15	4.0	4.0
012	13	21.5	40.6	13	7.7	11.5
024	11	33.5	97.7	11	9.1	19.7
036	9	49.3	160.9	9	8.3	25.4
048	7	71.8	224.0	7	10.7	28.3
072	3	218.8	347.7	3	23.3	4.7
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL192010 RICHARD

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	20	9.3	10.1	20	1.5	1.8
012	18	20.6	33.0	18	5.3	7.0
024	16	29.5	70.3	16	10.0	7.9
036	14	33.5	130.2	14	10.0	12.9
048	12	38.8	194.6	12	9.2	15.5
072	8	72.0	373.5	8	7.5	20.9
096	4	124.0	556.1	4	17.5	24.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL202010 SHARY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	4.5	5.2	7	4.3	4.3
012	5	64.9	161.9	5	13.0	12.8
024	3	104.0	364.9	3	20.0	21.7
036	1	139.2	529.9	1	20.0	12.0
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: AL212010 TOMAS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	37	12.1	12.6	37	3.0	3.1
012	35	37.5	53.3	35	9.4	11.3
024	33	57.1	100.7	33	13.9	15.5
036	31	68.1	155.4	31	15.8	17.9
048	29	70.2	230.7	29	21.0	21.6
072	25	100.8	409.1	25	25.8	21.6
096	21	130.5	572.8	21	29.3	17.3
120	17	120.4	633.4	17	26.2	16.0

Table 7. Homogenous comparison of official and CLIPER5 track forecast errors in the eastern North Pacific basin in 2010 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
2010 mean OFCL error (n mi)	26.0	40.1	48.6	54.7	85.3	119.3	145.4
2010 mean CLIPER5 error (n mi)	36.5	73.2	110.8	143.5	204.4	255.3	259.4
2010 mean OFCL skill relative to CLIPER5 (%)	28.8	45.2	56.1	61.9	58.3	53.3	43.9
2010 mean OFCL bias vector (°/n mi)	244/4	266/6	301/5	050/9	068/39	074/69	077/97
2010 number of cases	138	115	97	83	63	43	29
2005-2009 mean OFCL error (n mi)	30.8	51.5	71.6	89.6	120.9	155.0	192.0
2005-2009 mean CLIPER5 error (n mi)	38.9	75.3	115.7	155.8	226.9	275.1	321.5
2005-2009 mean OFCL skill relative to CLIPER5 (%)	20.8	31.4	38.0	42.5	47.2	44.6	41.5
2005-2009 mean OFCL bias vector (°/n mi)	308/3	297/5	283/7	282/12	260/12	241/5	109/5
2005-2009 number of cases	1323	1160	1001	863	628	448	310
2010 OFCL error relative to 2005-2009 mean (%)	-15.6	-22.1	-33.0	-39.0	-29.4	-23.0	-24.3
2010 CLIPER5 error relative to 2005-2009 mean (%)	-6.2	-2.8	-4.2	-7.9	-9.9	-7.2	-19.3

Table 8a. Homogenous comparison of eastern North Pacific basin early track guidance model errors (n mi) for 2010. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	20.0	32.5	42.7	52.0	74.2	95.7	122.9
OCD5	30.8	66.7	109.3	150.0	208.0	259.2	191.9
GFSI	30.0	54.1	77.3	99.4	156.2	201.9	222.3
GHMI	24.7	40.8	54.6	73.6	108.4	142.2	204.1
HWFI	27.5	49.7	68.1	95.2	125.2	152.2	**120.5**
GFNI	35.5	61.8	82.8	103.5	108.7	99.9	**111.7**
NGPI	29.3	51.4	68.1	90.9	112.7	127.4	137.3
EMXI	23.5	38.4	52.9	66.1	92.9	126.1	190.6
FSSE	21.4	34.5	43.3	55.5	79.0	112.5	165.6
AEMI	28.8	49.7	69.8	88.8	114.1	146.1	**120.7**
TVCN	21.6	34.1	42.7	53.7	**69.0**	**90.0**	**107.6**
TVCC	24.4	34.7	47.4	63.4	97.9	131.7	152.1
LBAR	31.0	75.6	129.7	195.2	292.7	332.7	274.9
BAMD	40.7	76.0	105.2	145.0	205.9	274.5	328.2
BAMM	32.7	59.4	83.2	118.5	186.1	246.8	268.2
BAMS	31.7	56.6	78.5	106.5	184.7	276.1	331.0
# Cases	78	72	64	58	39	24	13

Table 8b. Homogenous comparison of eastern North Pacific basin early track guidance model bias vectors (º/n mi) for 2010.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	334/002	257/002	255/003	305/004	043/023	046/033	048/060
OCD5	336/008	332/021	326/040	322/065	327/097	335/109	047/139
GFSI	199/008	193/018	190/026	194/036	188/048	233/075	279/090
GHMI	067/003	062/007	016/007	340/017	356/043	008/069	034/113
HWFI	303/014	293/022	295/030	296/048	320/063	328/080	352/053
GFNI	317/013	274/017	252/026	253/034	281/021	216/018	040/025
NGPI	296/007	270/015	271/022	279/036	298/047	287/051	289/057
EMXI	084/007	107/013	119/018	118/024	099/065	091/105	090/187
FSSE	310/003	261/004	271/005	299/010	009/031	017/054	045/134
AEMI	200/011	194/022	196/034	204/047	192/053	206/050	121/027
TVCN	306/004	257/006	243/009	256/014	355/007	340/011	043/039
TVCC	011/011	215/005	328/013	306/025	335/068	303/097	347/094
LBAR	302/018	301/061	296/114	295/177	295/163	281/311	265/248
BAMD	308/028	301/055	294/079	291/111	284/134	266/176	250/174
BAMM	318/018	303/034	291/047	284/071	274/077	261/103	266/106
BAMS	359/008	331/012	304/018	289/032	272/031	279/047	289/100
# Cases	78	72	64	58	39	24	13

Table 9. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the eastern North Pacific basin for the 2010 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
2010 mean OFCL error (kt)	6.1	9.3	12.4	13.5	15.6	15.9	17.8
2010 mean Decay-SHIFOR5 error (kt)	8.0	12.7	17.4	20.0	23.8	24.7	31.2
2010 mean OFCL skill relative to Decay-SHIFOR5 (%)	26.7	24.8	33.9	41.0	38.0	36.2	37.3
2010 OFCL bias (kt)	1.2	1.6	2.1	1.6	1.0	2.0	-1.2
2010 number of cases	138	115	97	83	63	43	29
2005-9 mean OFCL error (kt)	6.2	10.4	13.7	15.4	17.2	18.6	18.4
2005-9 mean Decay-SHIFOR5 error (kt)	7.1	11.6	15.0	17.5	18.8	19.5	19.2
2005-9 mean OFCL skill relative to Decay-SHIFOR5 (%)	12.7	10.3	8.7	12.0	8.5	4.6	4.2
2005-9 OFCL bias (kt)	0.6	1.1	1.1	0.0	-0.5	-2.6	-4.7
2005-9 number of cases	1296	1135	978	842	611	435	301
2010 OFCL error relative to 2005-9 mean (%)	-1.6	-10.6	-9.5	-12.3	-9.3	-14.5	-3.3
2010 Decay-SHIFOR5 error relative to 2005-9 mean (%)	12.7	9.5	16.0	14.3	26.6	26.7	62.5

Table 10a. Homogenous comparison of eastern North Pacific basin early intensity guidance model errors (kt) for 2010. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	6.3	10.6	12.3	11.8	14.2	16.9	20.0
OCD5	8.6	14.1	18.6	20.0	22.9	26.5	31.9
HWFI	7.3	13.4	19.8	22.7	29.3	34.3	37.9
GHMI	7.5	11.1	13.2	13.9	18.3	23.8	27.9
GFNI	8.5	13.4	15.0	15.4	19.4	21.9	27.3
DSHP	7.5	11.5	13.4	14.5	14.6	**16.5**	**17.7**
LGEM	7.7	12.3	13.7	14.2	**12.6**	**14.7**	22.1
FSSE	6.6	**10.4**	**11.9**	12.7	**13.1**	20.0	28.0
ICON	6.7	11.0	13.1	14.1	15.4	20.3	24.7
IVCN	6.7	11.1	12.9	13.7	15.4	20.3	24.3
# Cases	84	77	65	57	37	26	15

Table 10b. Homogenous comparison of eastern North Pacific basin early intensity guidance model biases (kt) for 2010. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	1.8	3.4	6.2	6.3	2.8	2.7	0.0
OCD5	2.4	6.8	13.7	16.9	9.1	**0.6**	-12.8
HWFI	-2.3	**-3.2**	**-3.7**	**-6.1**	-19.1	-30.1	-32.3
GHMI	-2.3	**-3.3**	**-5.0**	-6.8	-14.4	-20.3	-20.2
GFNI	-3.9	-6.7	-8.6	-10.3	-16.8	-19.2	-20.7
DSHP	**0.7**	**2.7**	**6.0**	7.4	3.5	**1.4**	-4.7
LGEM	**0.1**	**0.6**	**2.4**	**2.5**	**-0.6**	**-0.4**	-2.6
FSSE	**-0.2**	**0.4**	**1.2**	**0.1**	-7.6	-14.6	-20.7
ICON	**-0.7**	**-0.5**	**0.2**	**-0.4**	-7.3	-12.2	-14.9
IVCN	**-1.3**	**-1.6**	**-1.4**	**-2.3**	-9.2	-13.5	-16.0
# Cases	84	77	65	57	37	26	15

Table 11. Official eastern North Pacific track and intensity forecast verifications (OFCL) for 2010 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: EP012010 AGATHA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	16.0	16.0	5	2.0	2.0
012	3	70.8	74.1	3	10.0	8.7
024	1	119.2	159.9	1	25.0	17.0
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: EP022010 TWO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	3	9.4	9.4	3	1.7	1.7
012	1	26.7	55.2	1	5.0	3.0
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: EP032010 BLAS

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	14.8	14.7	17	1.2	1.5
012	15	26.4	31.7	15	6.7	7.0
024	13	45.3	60.9	13	9.6	9.9
036	11	58.3	92.1	11	11.8	12.9
048	9	72.5	123.8	9	11.1	11.3
072	5	149.2	215.2	5	5.0	10.2
096	1	316.7	437.6	1	10.0	11.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP042010 CELIA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	38	7.9	7.5	38	2.1	2.4
012	36	15.3	28.3	36	7.9	11.6
024	34	23.9	62.8	34	9.6	15.9
036	32	29.7	97.3	32	11.3	18.2
048	30	35.5	130.8	30	11.2	20.5
072	26	56.9	178.2	26	15.8	28.3
096	22	90.8	212.0	22	15.0	29.2
120	18	131.4	212.8	18	17.5	34.9

Verification statistics for: EP052010 DARBY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	5.7	5.4	23	2.6	2.8
012	21	24.9	39.6	21	7.6	7.6
024	19	43.2	93.3	19	14.5	13.7
036	17	56.8	151.9	17	20.3	21.3
048	15	67.2	216.0	15	25.0	27.3
072	11	83.2	326.6	11	28.2	30.1
096	7	103.8	424.5	7	30.7	22.9
120	3	72.3	394.0	3	16.7	23.7

Verification statistics for: EP062010 SIX

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	13.4	13.4	7	2.1	2.1
012	6	63.1	79.8	6	4.2	7.7
024	4	114.2	138.1	4	5.0	13.8
036	2	142.0	181.0	2	10.0	21.0
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: EP072010 ESTELLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	9.2	9.2	17	1.5	1.5
012	15	24.3	29.4	15	3.7	5.5
024	13	39.2	59.5	13	5.8	8.4
036	11	50.3	92.1	11	8.6	14.2
048	9	56.0	110.6	9	9.4	15.0
072	5	113.1	162.5	5	12.0	20.6
096	1	180.4	174.0	1	15.0	23.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP082010 EIGHT

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	9.2	9.2	6	0.0	0.0
012	4	31.4	45.2	4	2.5	6.0
024	2	54.3	71.6	2	5.0	14.5
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: EP092010 FRANK

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	28	7.5	8.1	28	1.1	0.9
012	26	23.9	35.8	26	5.6	7.3
024	24	38.9	70.8	24	8.8	12.5
036	22	51.8	105.4	22	10.9	18.2
048	20	65.5	131.9	20	11.3	20.0
072	16	104.1	172.5	16	10.9	17.4
096	12	158.9	227.6	12	9.6	18.7
120	8	204.4	314.0	8	18.8	25.8

Verification statistics for: EP102010 TEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	14.1	15.9	5	0.0	0.0
012	3	48.8	35.2	3	1.7	3.3
024	1	65.7	24.6	1	5.0	9.0
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: EP112010 ELEVEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	4	0.0	1.5	4	0.0	0.0
012	2	44.3	59.4	2	7.5	10.5
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: EP122010 GEORGETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	8.8	9.9	8	0.0	0.0
012	6	26.5	30.2	6	0.8	3.0
024	4	48.1	90.3	4	1.3	3.0
036	2	89.8	169.1	2	5.0	0.5
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

Table 12a Verification of experimental in-house probabilistic genesis forecasts for the Atlantic basin in 2010.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	0	77
10	7	220
20	9	116
30	37	62
40	59	51
50	42	31
60	51	43
70	25	24
80	74	19
90	86	7
100	-	0

Table 12b. Verification of experimental in-house probabilistic genesis forecasts for the eastern North Pacific basin in 2010.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	5	19
10	10	93
20	18	61
30	43	21
40	60	15
50	75	12
60	38	16
70	45	11
80	63	8
90	0	1
100	-	0

Table 13a. Verification of experimental in-house probabilistic genesis forecasts for the Atlantic basin for the period 2007- 2010.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	2	481
10	6	801
20	15	391
30	32	224
40	48	128
50	43	91
60	55	100
70	53	59
80	70	43
90	78	23
100	100	1

Table 13b. Verification of experimental in-house probabilistic genesis forecasts for the eastern North Pacific basin for the period 2007-2010.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	3	171
10	16	462
20	29	295
30	42	126
40	58	86
50	72	83
60	73	62
70	74	47
80	71	34
90	83	6
100	100	1

Table 14. NHC forecast cone circle radii (n mi) for 2011. Change from 2010 values (n mi) given in parentheses.

Track Forecast Cone Two-Thirds Probability Circles (n mi)		
Forecast Period (h)	**Atlantic Basin**	**Eastern North Pacific Basin**
12	36 (0)	33 (-3)
24	59 (-3)	59 (0)
36	79 (-6)	79 (-3)
48	98 (-10)	98 (-4)
72	144 (-17)	134 (-4)
96	190(-30)	187 (13)
120	239 (-46)	230 (10)

Table 15. Composition of NHC consensus models for 2011. 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 2011			
Model ID	**Parameter**	**Type**	**Members**
GUNA	Track	Fixed	GFSI EGRI NGPI GHMI
TCOA	Track	Fixed	GFSI EGRI GHMI HWFI
TCOE*	Track	Fixed	GFSI EGRI NGPI GHMI HWFI
ICON	Intensity	Fixed	DSHP LGEM GHMI HWFI
TVCA	Track	Variable	GFSI EGRI GHMI HWFI GFNI EMXI
TVCE**	Track	Variable	GFSI EGRI NGPI GHMI HWFI GFNI EMXI
IVCN	Intensity	Variable	DSHP LGEM GHMI HWFI GFNI
CGUN	Track	Fixed (corrected)	GFSI EGRI NGPI GHMI
TCCN	Track	Fixed (corrected)	GFSI EGRI NGPI GHMI HWFI
TVCC	Track	Variable (corrected)	GFSI EGRI NGPI GHMI HWFI GFNI EMXI

* 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

18. Comparison of intensity skill for the HFIP Stream 1.5 GF5I and operational GFDL (GHMI) in the Atlantic and eastern North Pacific basins.
19. Comparison of track skill for the HFIP Stream 1.5 AHWI and the traditionally top-performing track models in the Atlantic basin.
20. Comparison of the intensity skill for the HFIP Stream 1.5 AHWI and the best performing intensity models in the Atlantic basin.

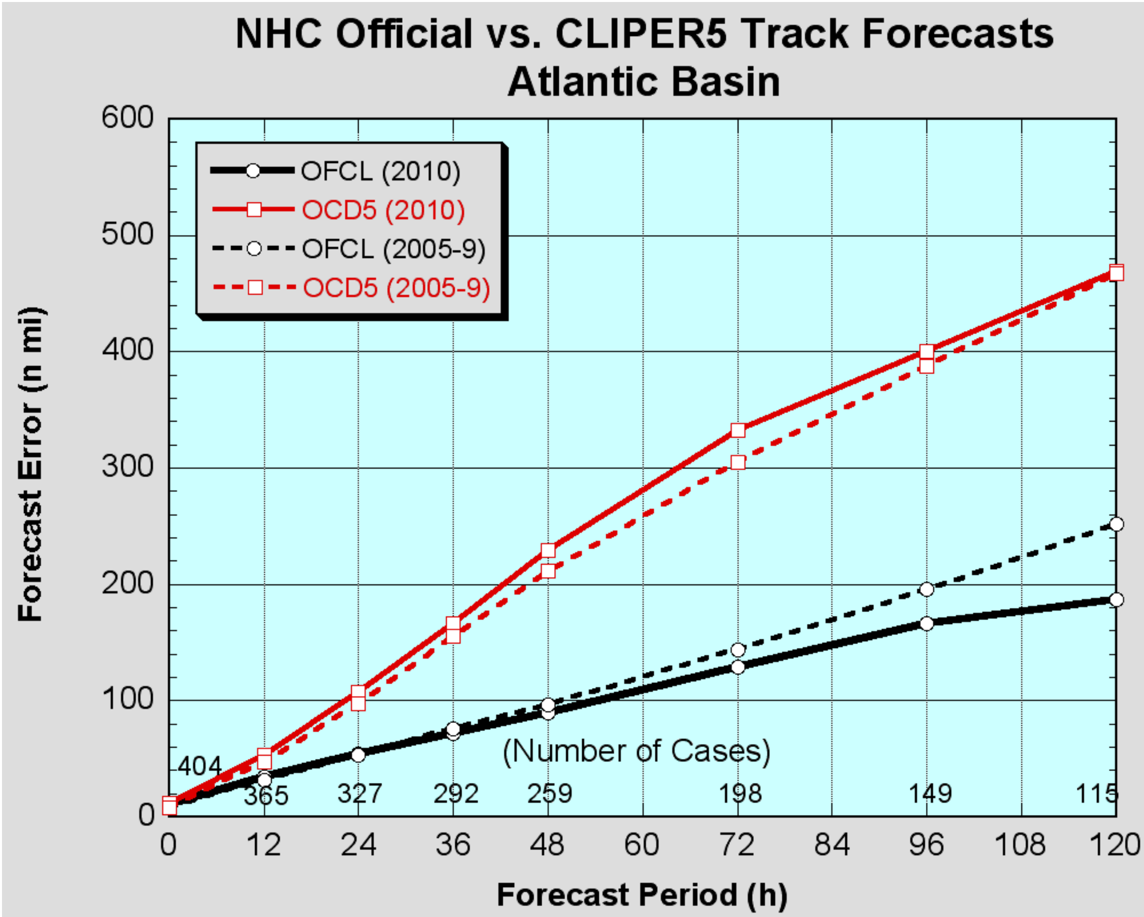
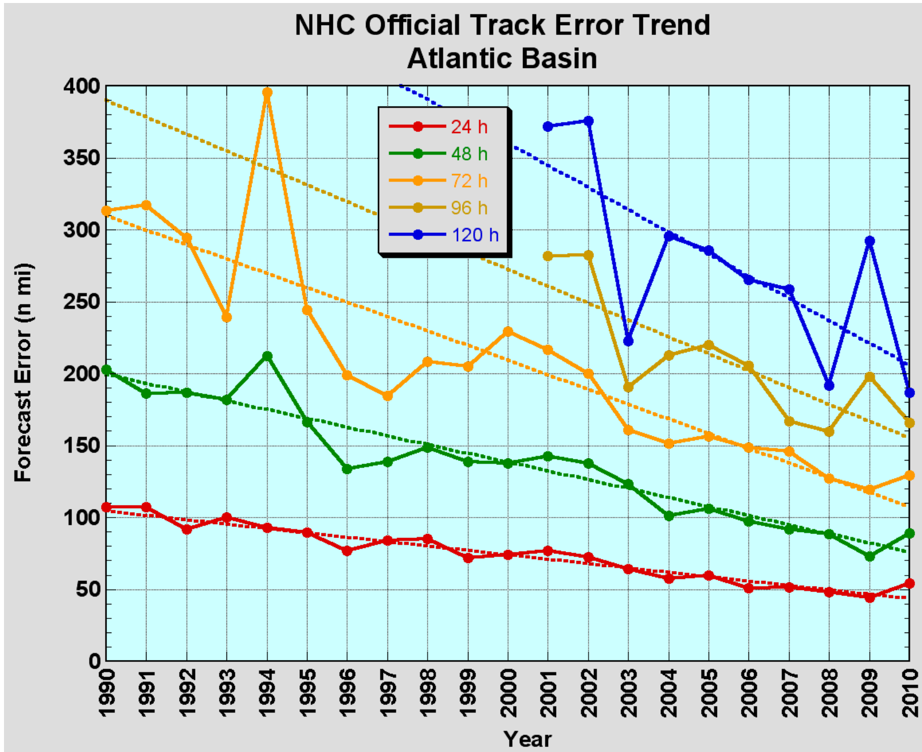


Figure 1. NHC official and CLIPER5 (OCD5) Atlantic basin average track errors for 2010 (solid lines) and 2005-2009 (dashed lines).



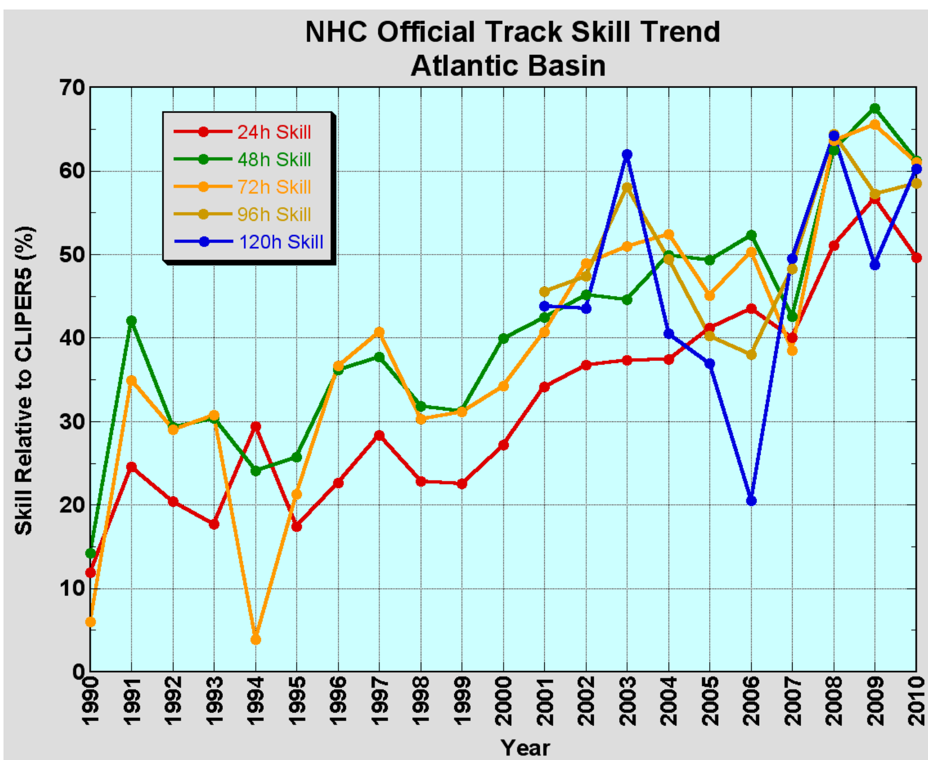


Figure 2. Recent trends in NHC official track forecast error (top) and skill (bottom) for the Atlantic basin.

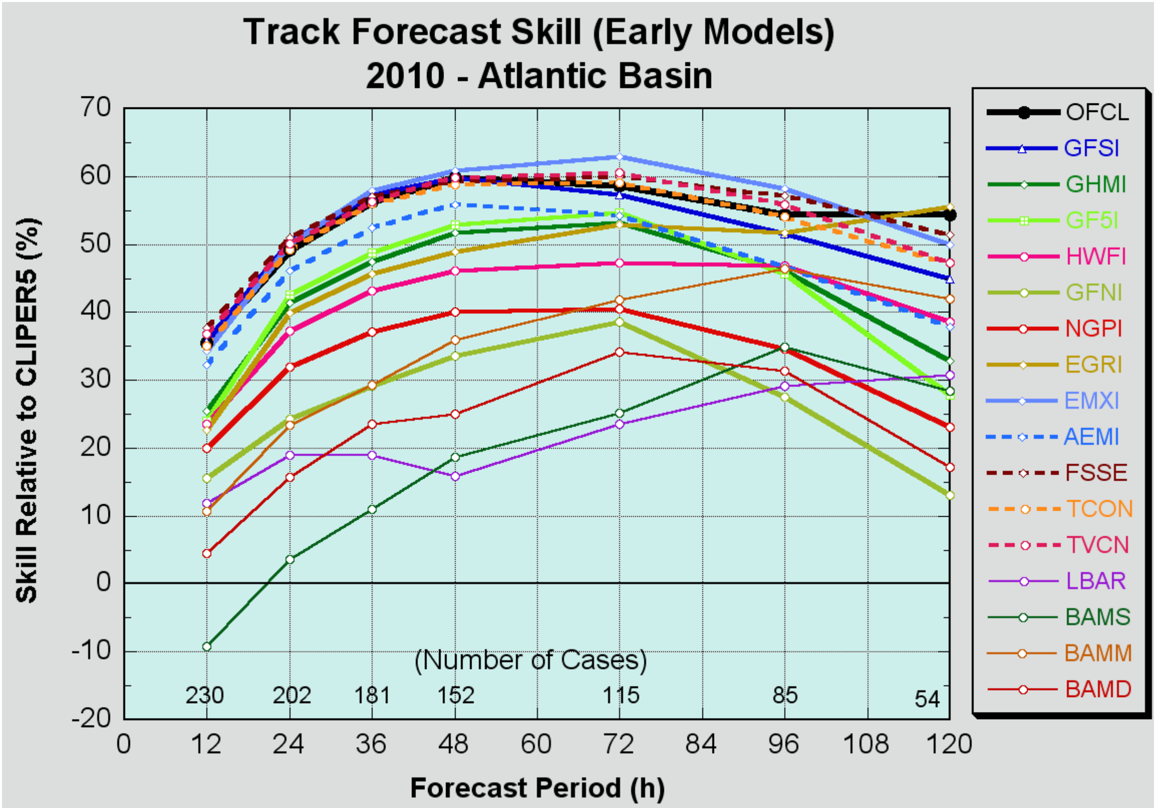


Figure 3. Homogenous comparison for selected Atlantic basin early track models for 2010. This verification includes only those models that were available at least 2/3 of the time (see text).

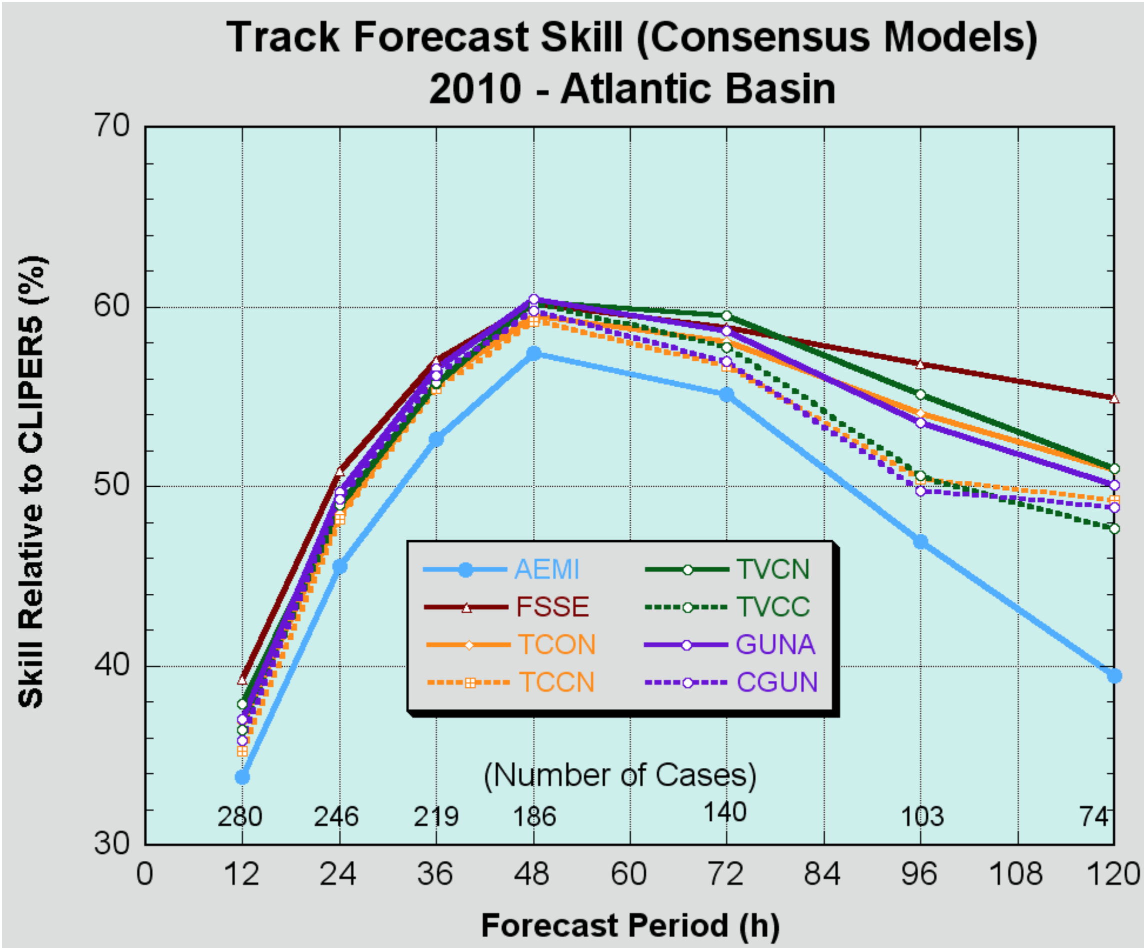


Figure 4. Homogenous comparison of the primary Atlantic basin track consensus models for 2010.

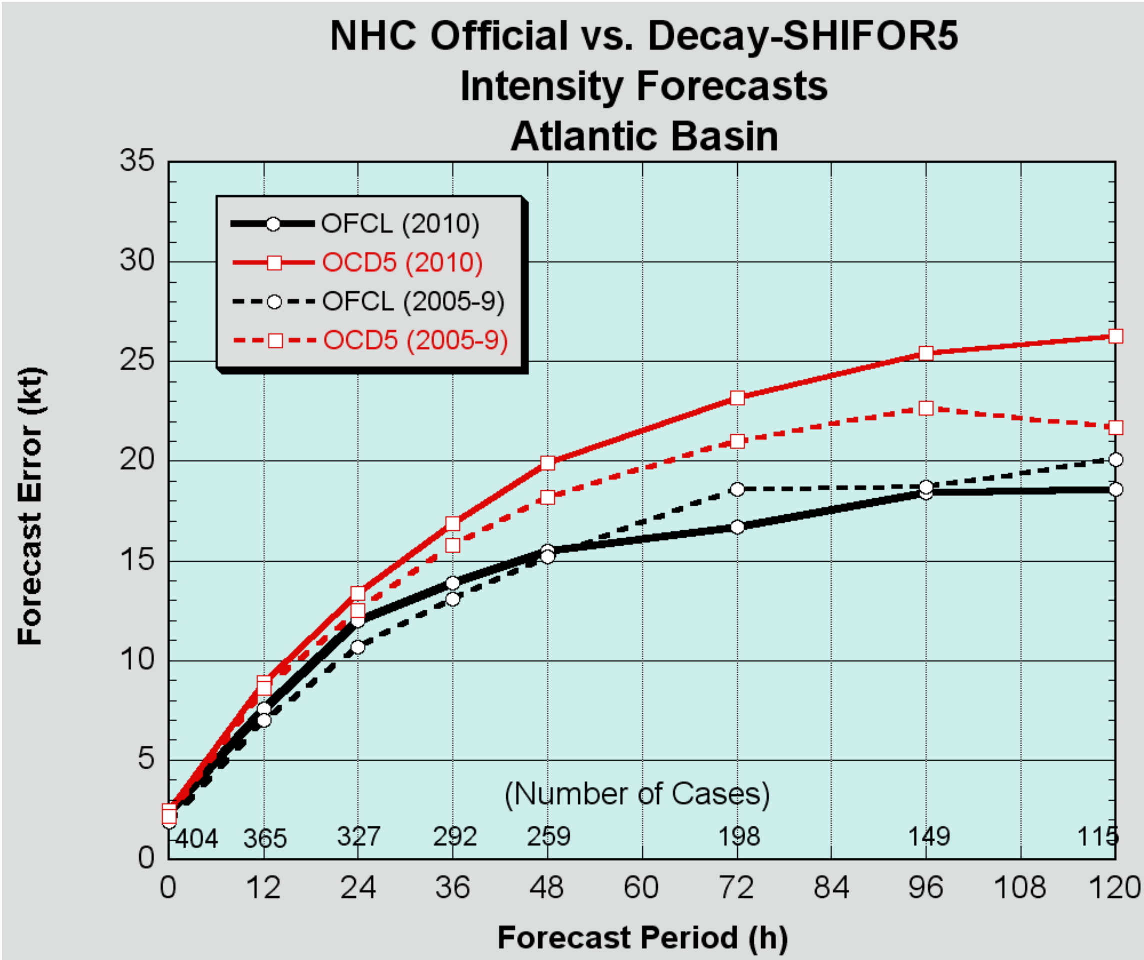


Figure 5. NHC official and Decay-SHIFOR5 (OCD5) Atlantic basin average intensity errors for 2010 (solid lines) and 2005-2009 (dashed lines).

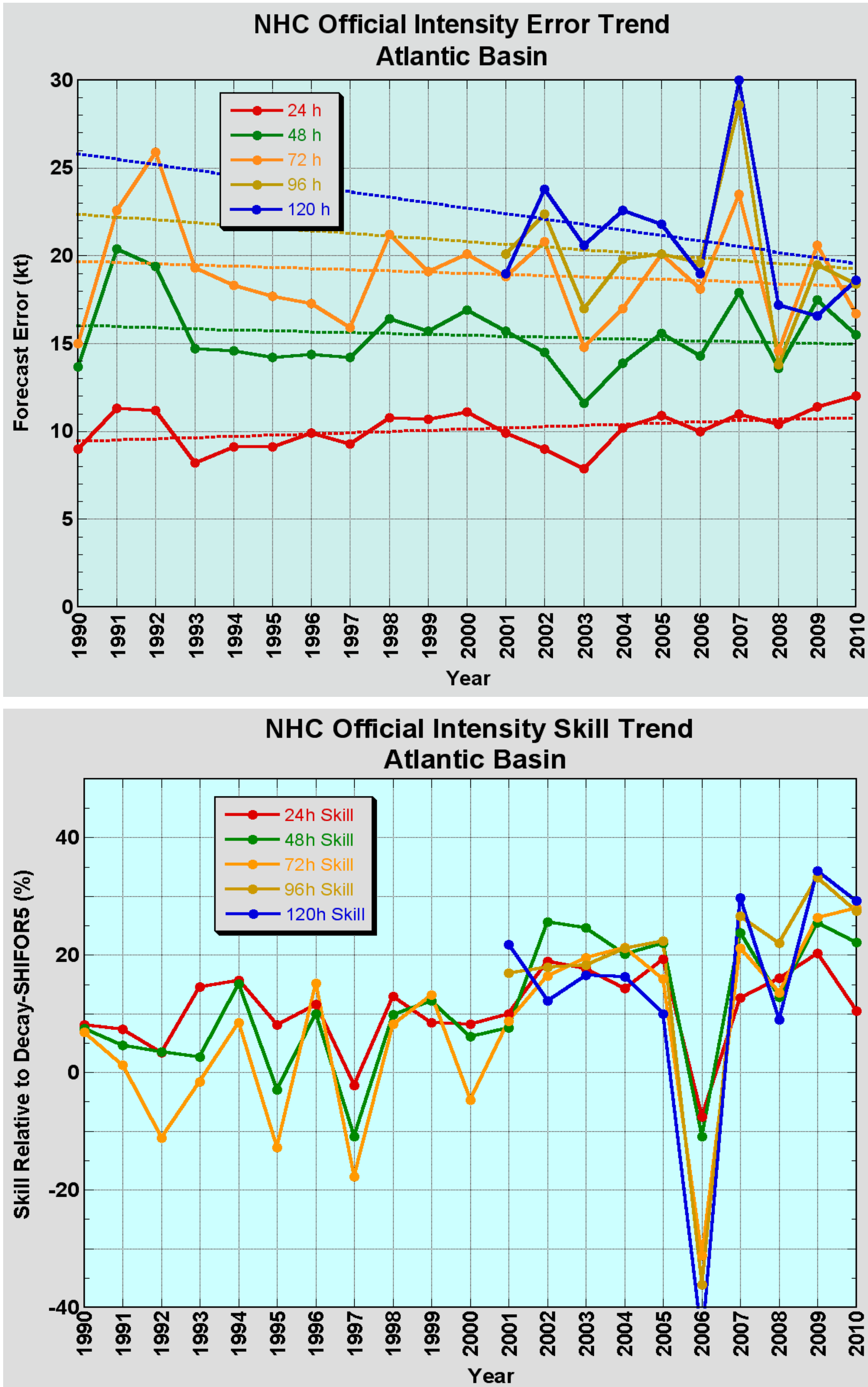


Figure 6. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the Atlantic basin.

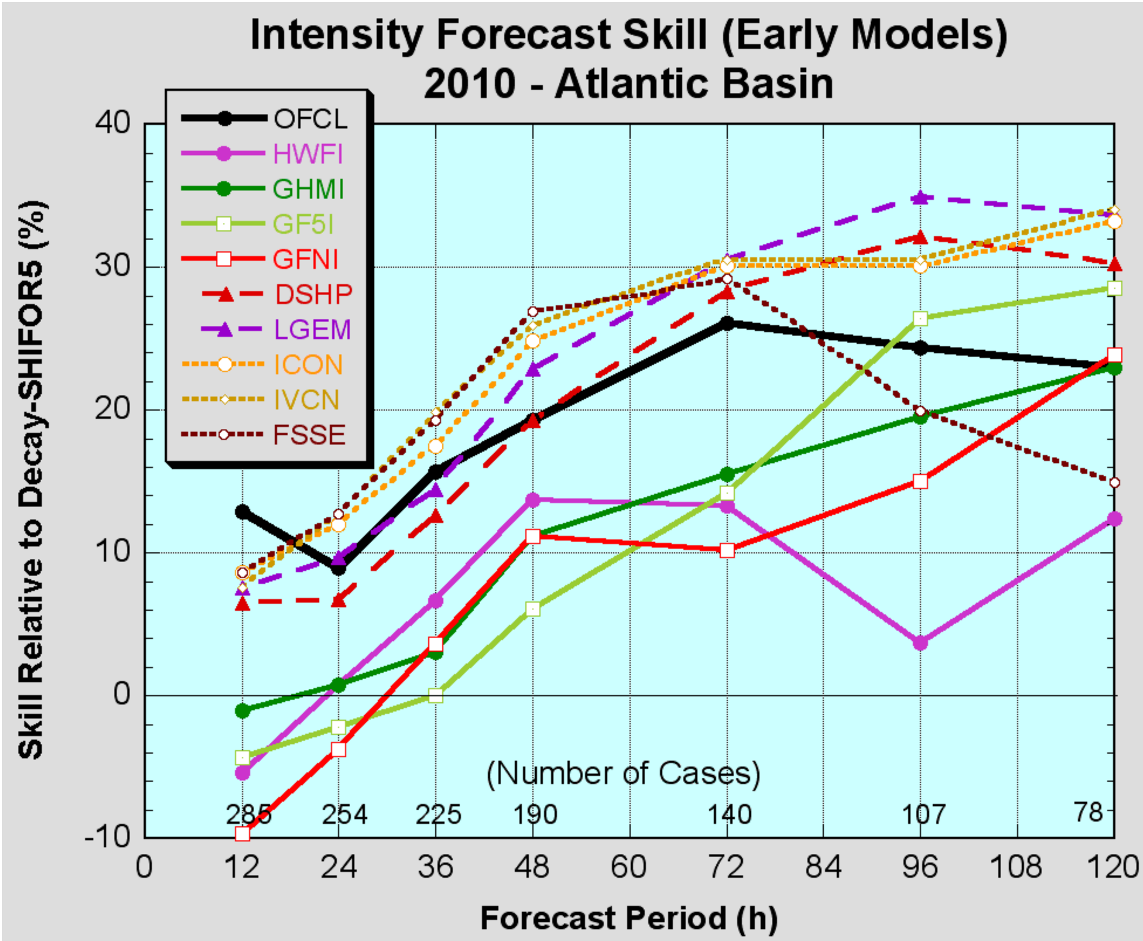


Figure 7. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2010.

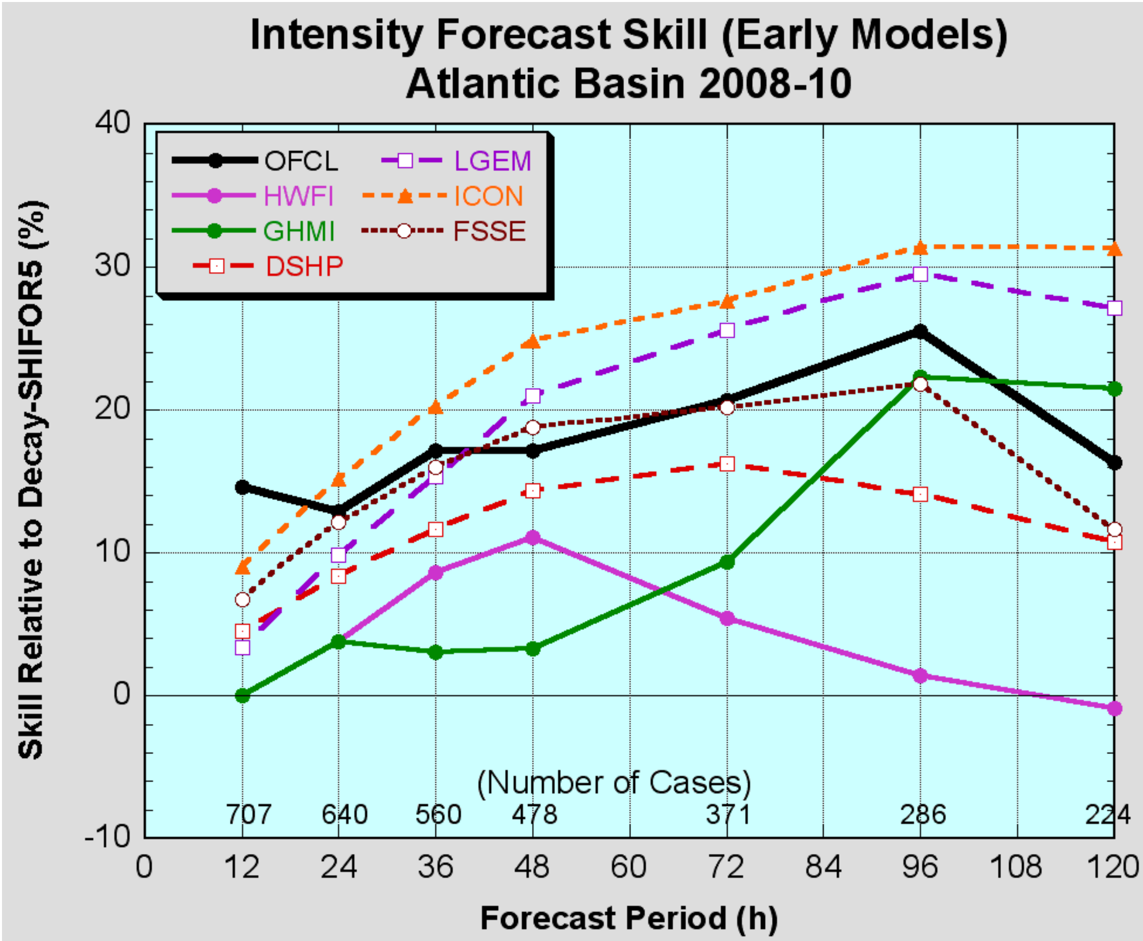


Figure 8. Homogenous comparison for selected for Atlantic basin early intensity guidance models for 2008-2010.

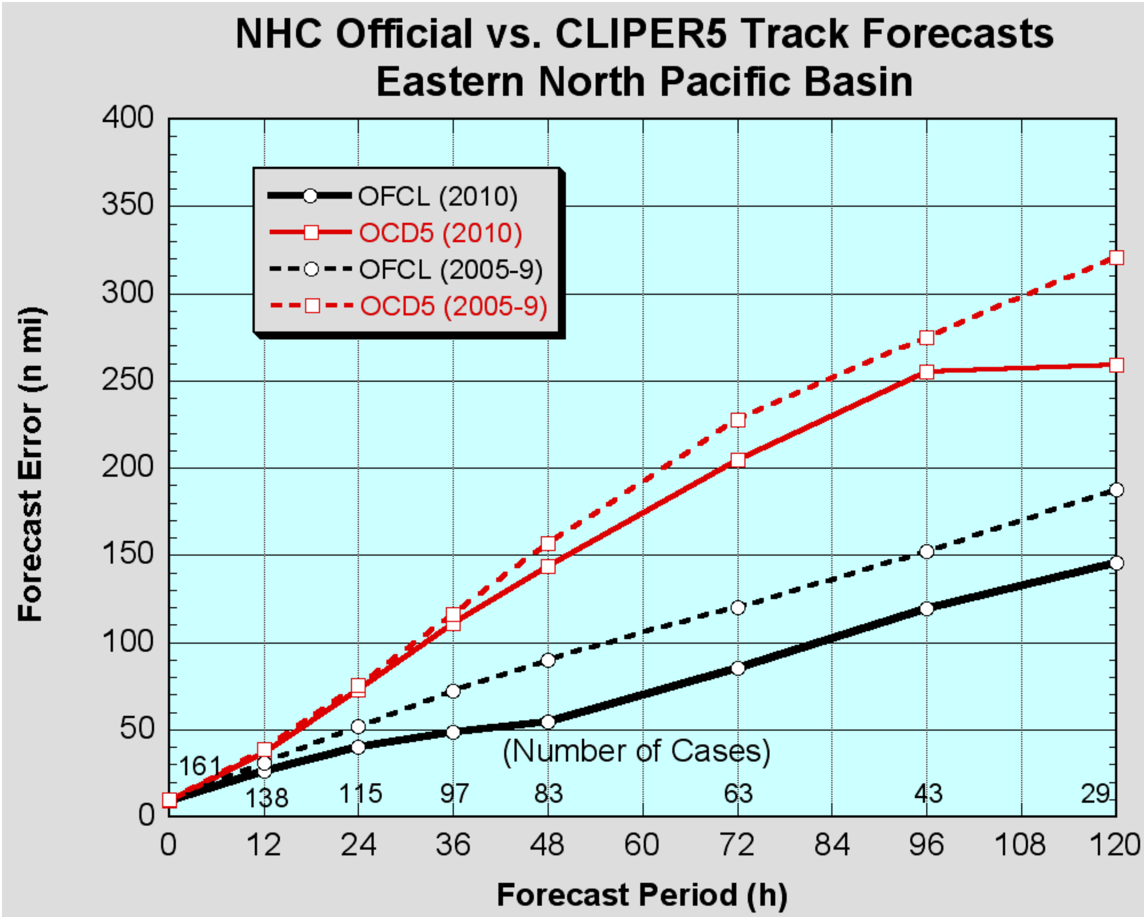
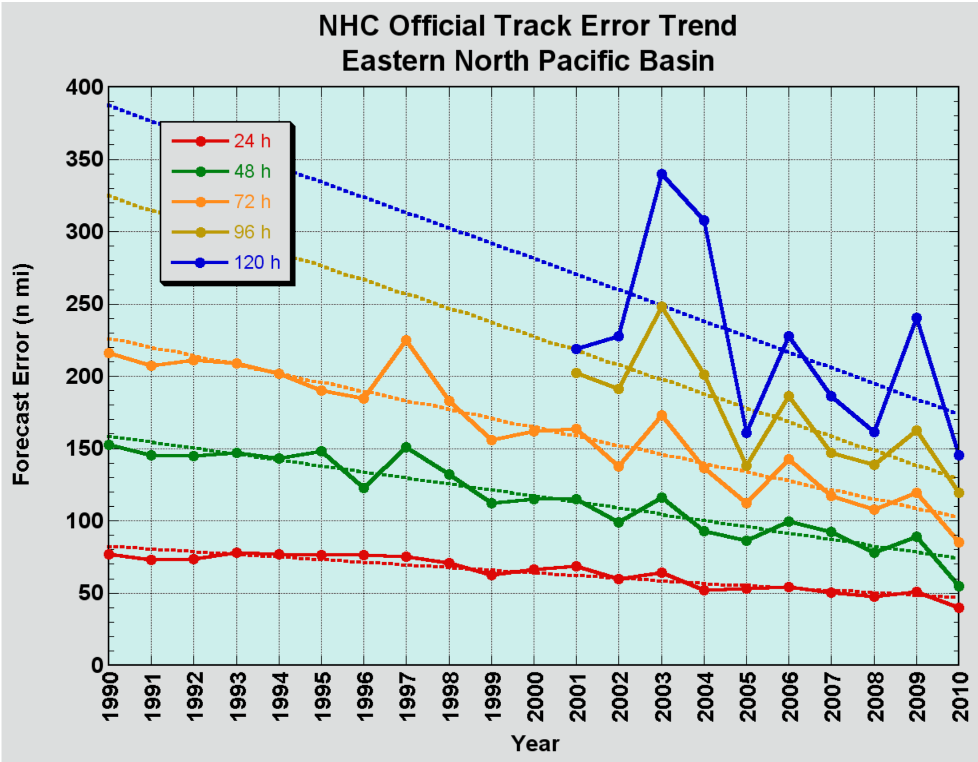


Figure 9. NHC official and CLIPER5 (OCD5) eastern North Pacific basin average track errors for 2010 (solid lines) and 2005-2009 (dashed lines).



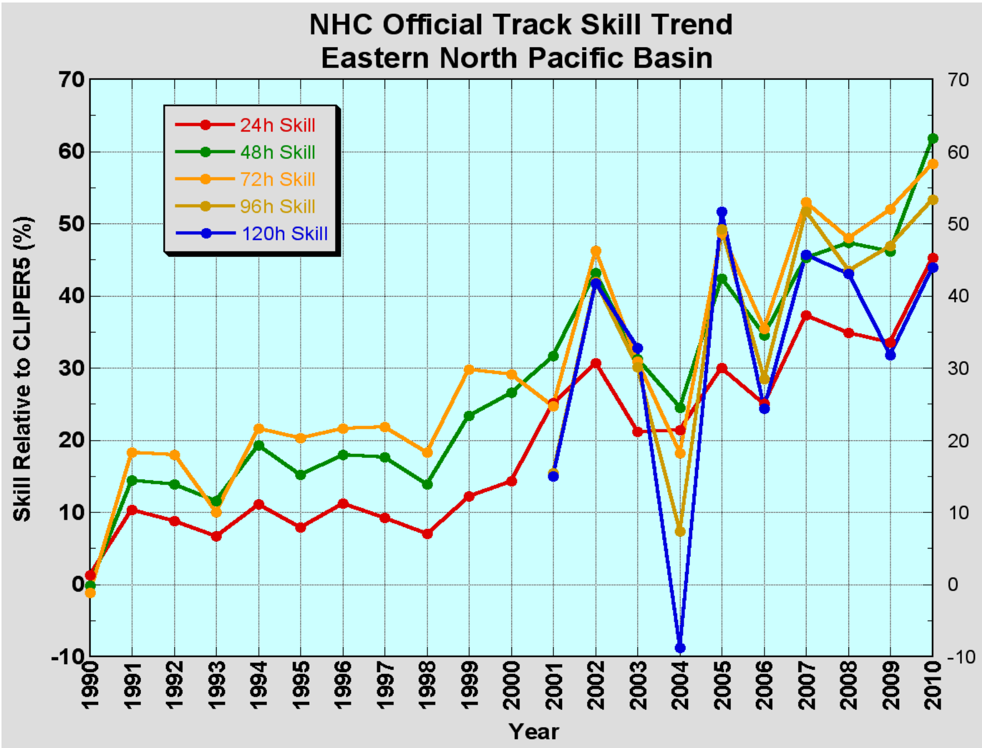


Figure 10. Recent trends in NHC official track forecast error (top) and skill (bottom) for the eastern North Pacific basin.

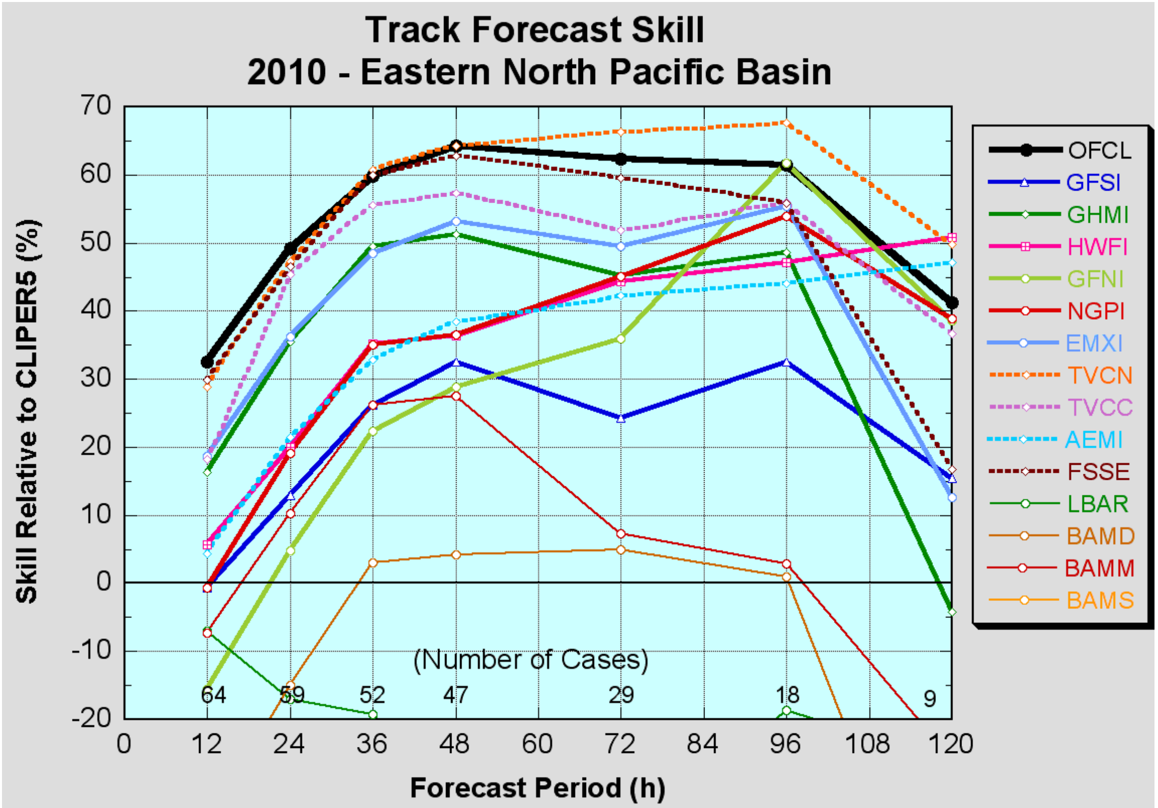


Figure 11. Homogenous comparison for selected eastern North Pacific early track models for 2010. This verification includes only those models that were available at least 2/3 of the time (see text).

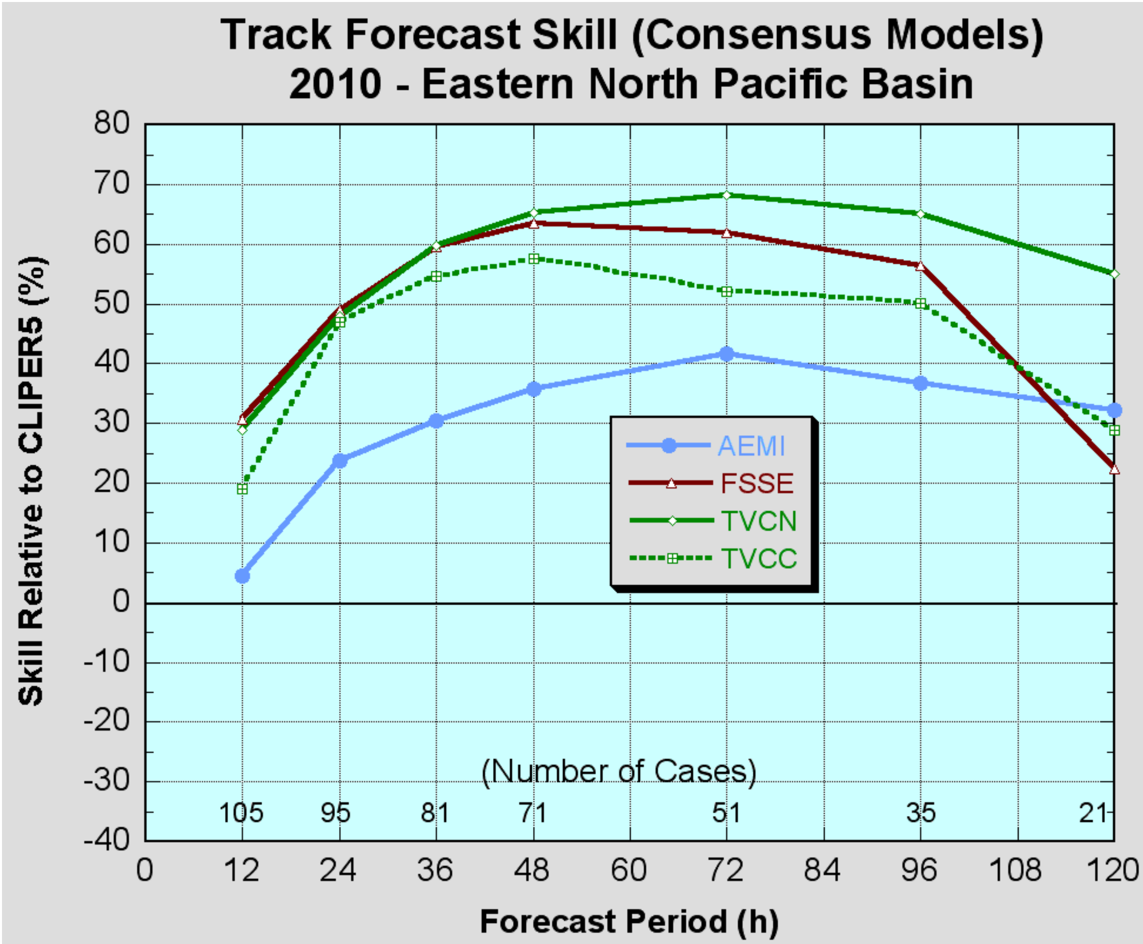


Figure 12. Homogenous comparison of the primary eastern North Pacific basin track consensus models for 2010.

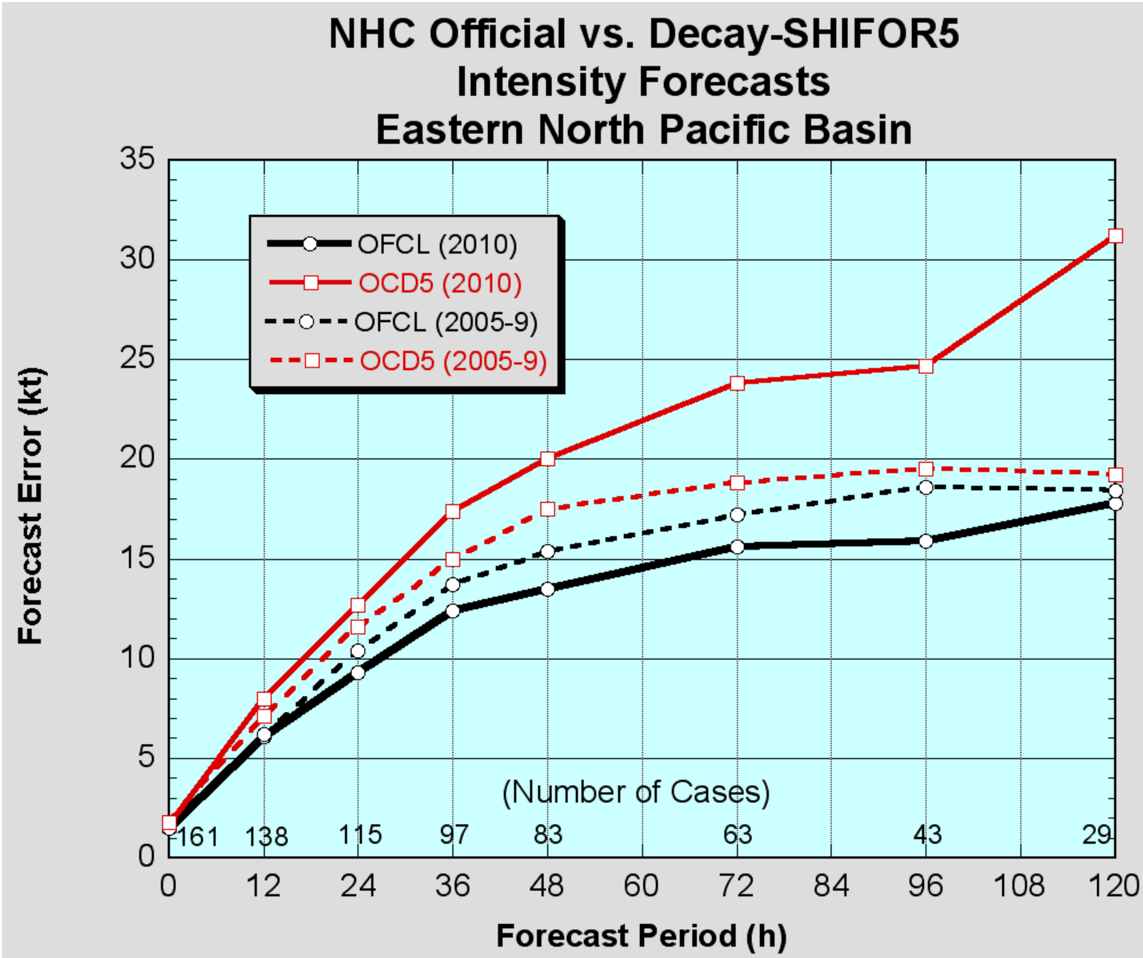
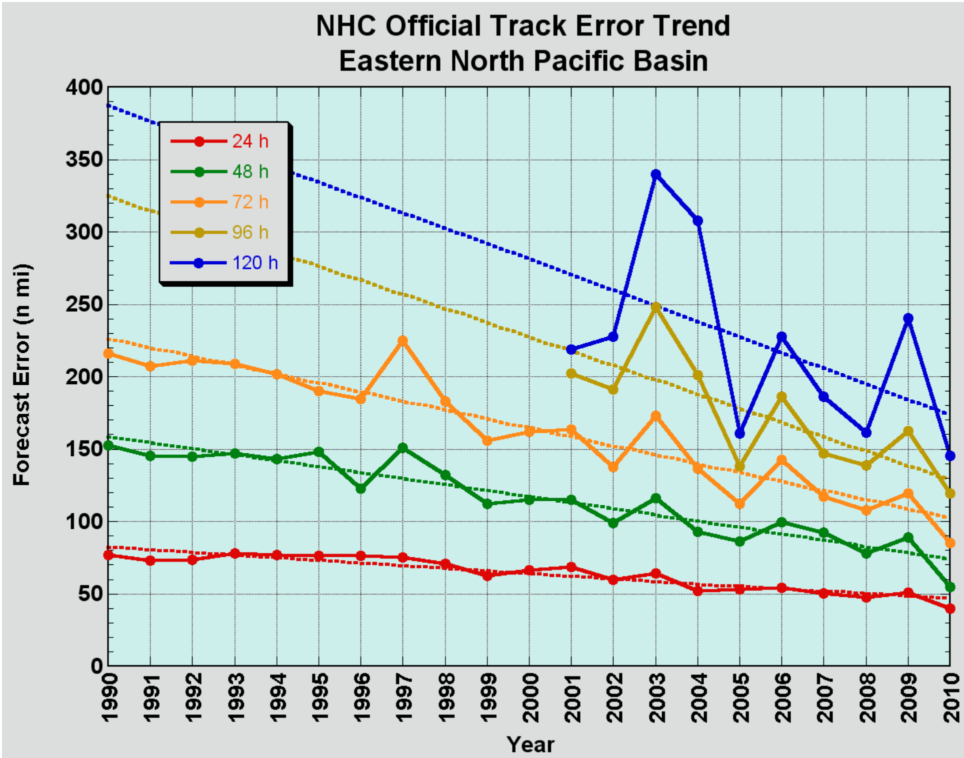


Figure 13. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2010 (solid lines) and 2005-2009 (dashed lines).



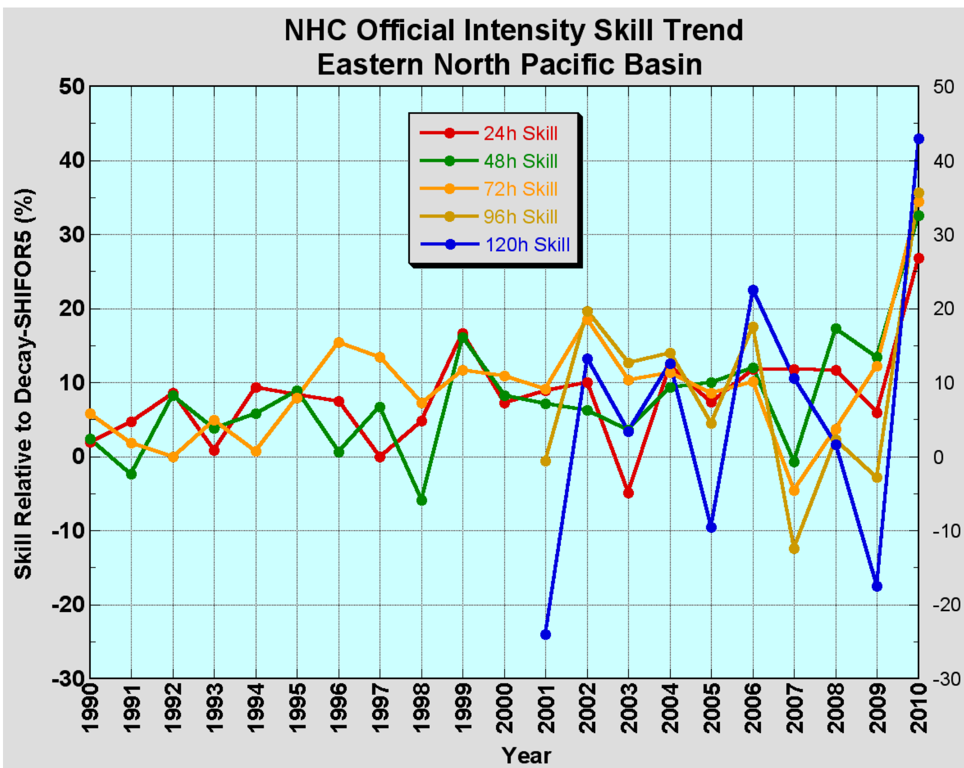


Figure 14. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

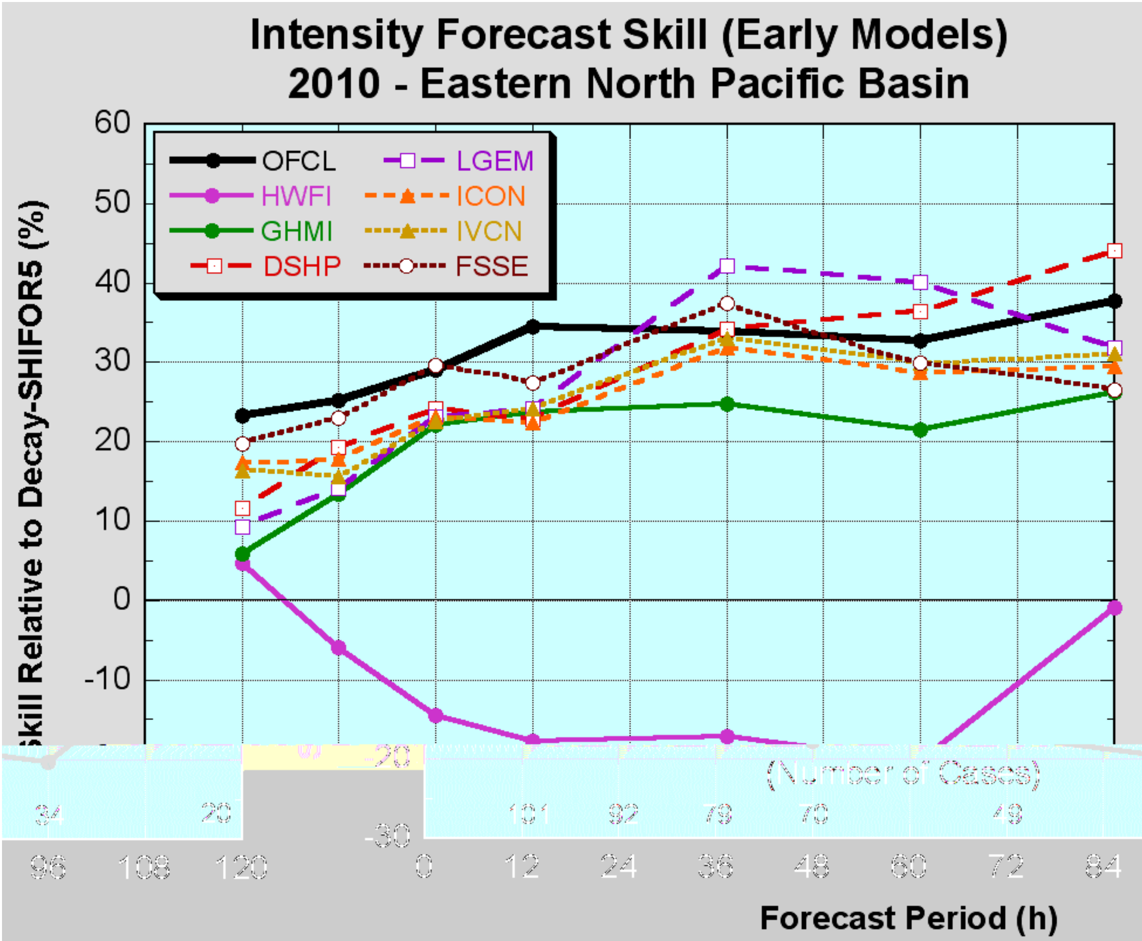


Figure 15. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2010.

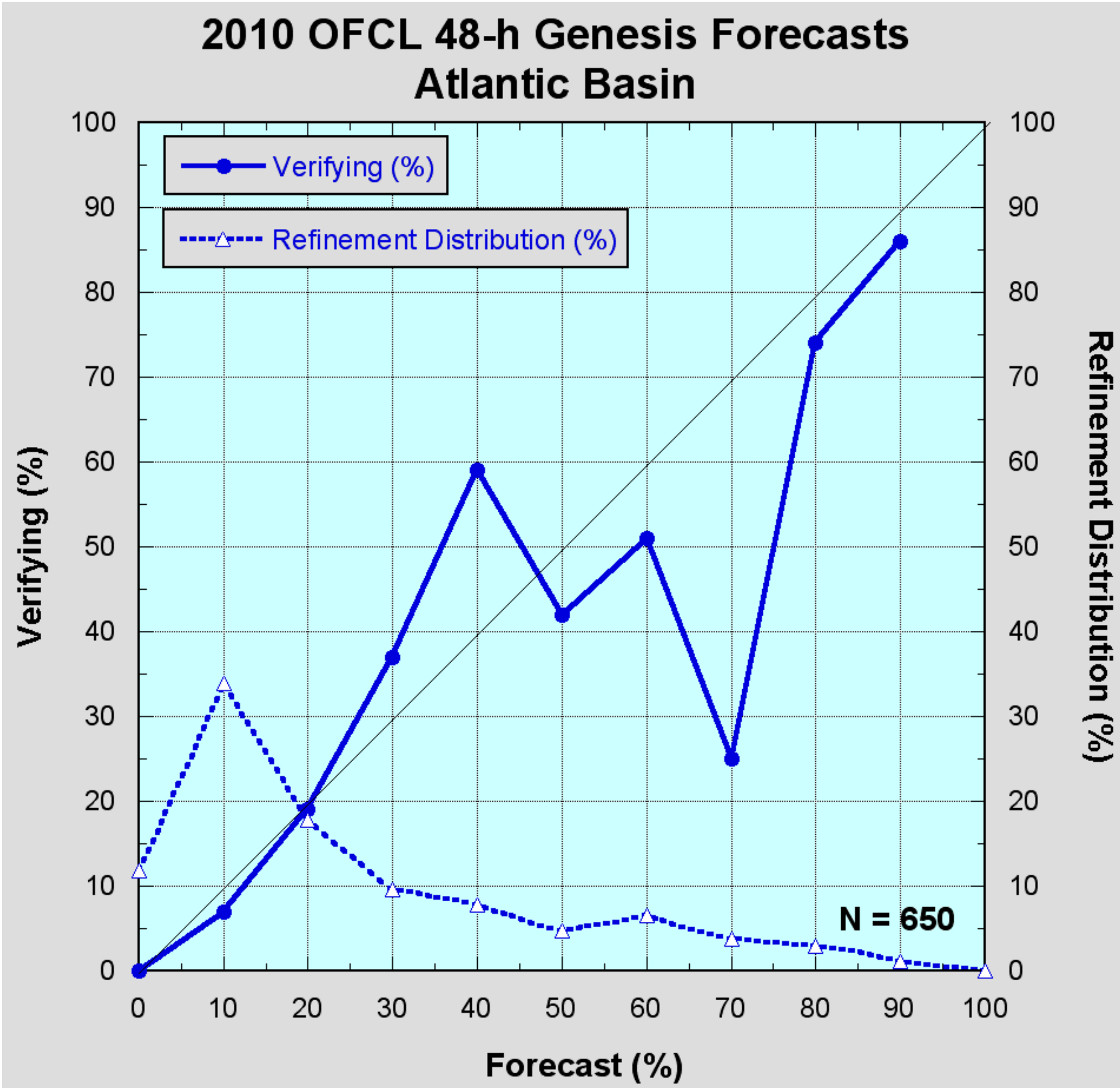


Figure 16a. Reliability diagram for Atlantic probabilistic tropical cyclogenesis forecasts for 2010. The solid blue line indicates the relationship between the forecast and verifying genesis percentages, with perfect reliability indicated by the thin diagonal black line. The dashed blue line indicates how the forecasts were distributed among the possible forecast values.

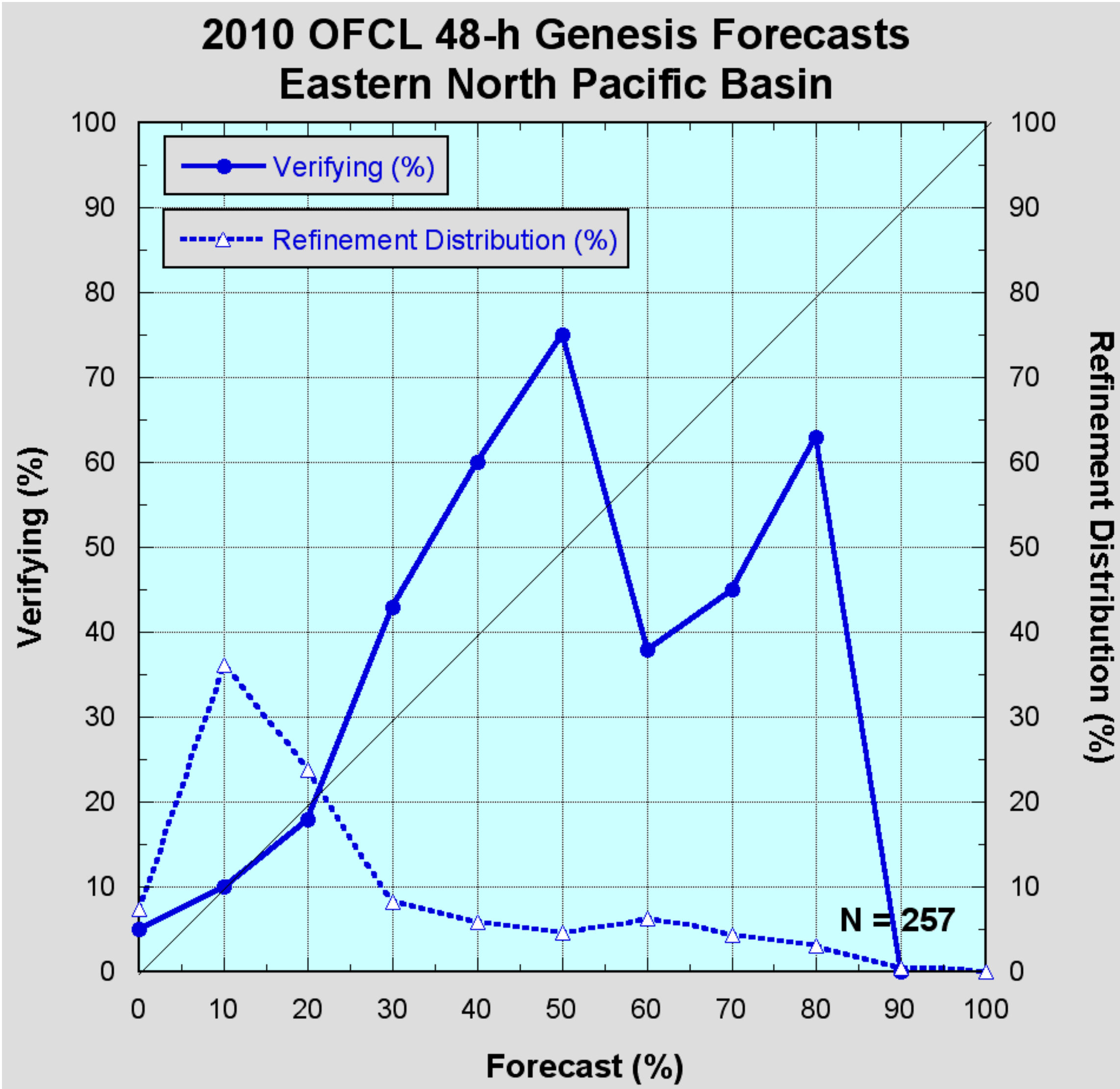


Figure 16b. As described for Fig. 16a, except for the eastern North Pacific basin.

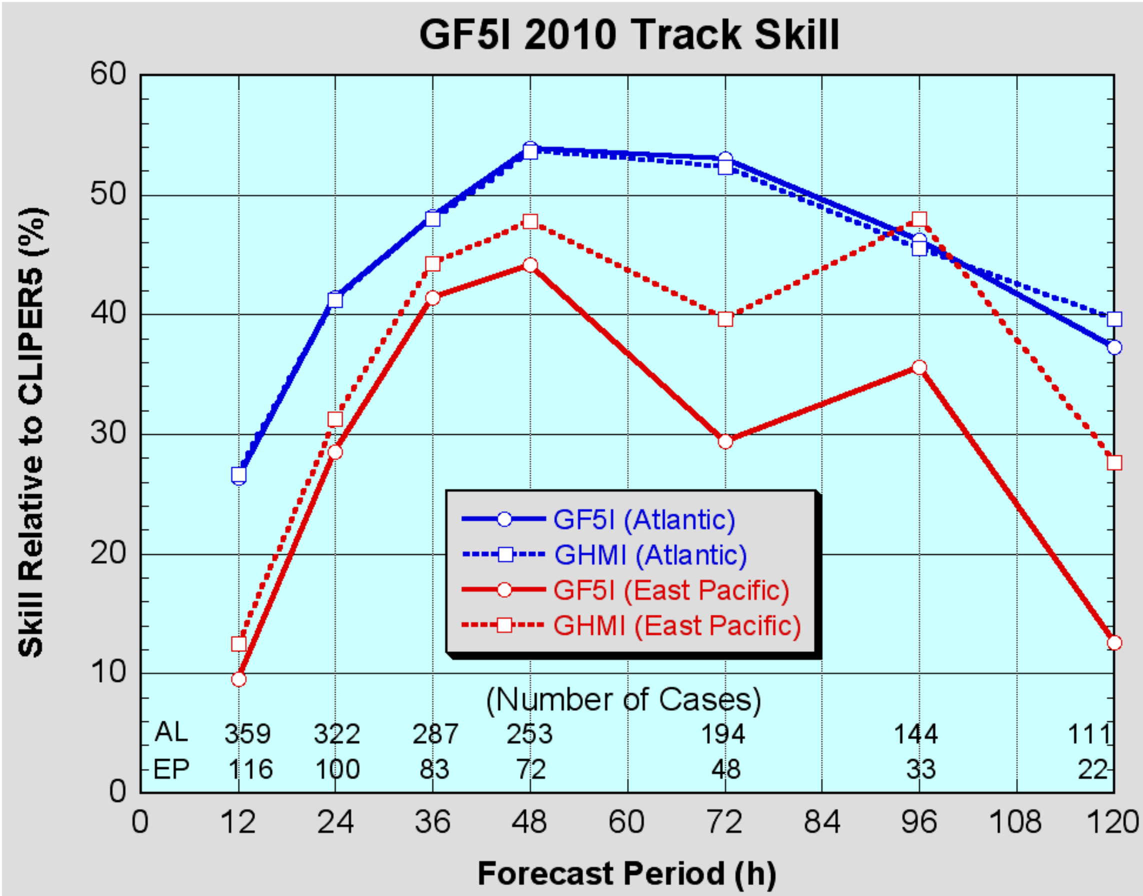


Figure 17. Comparison of track skill for the HFIP Stream 1.5 GF5I and operational GFDL (GHMI) in the Atlantic and eastern North Pacific basins.

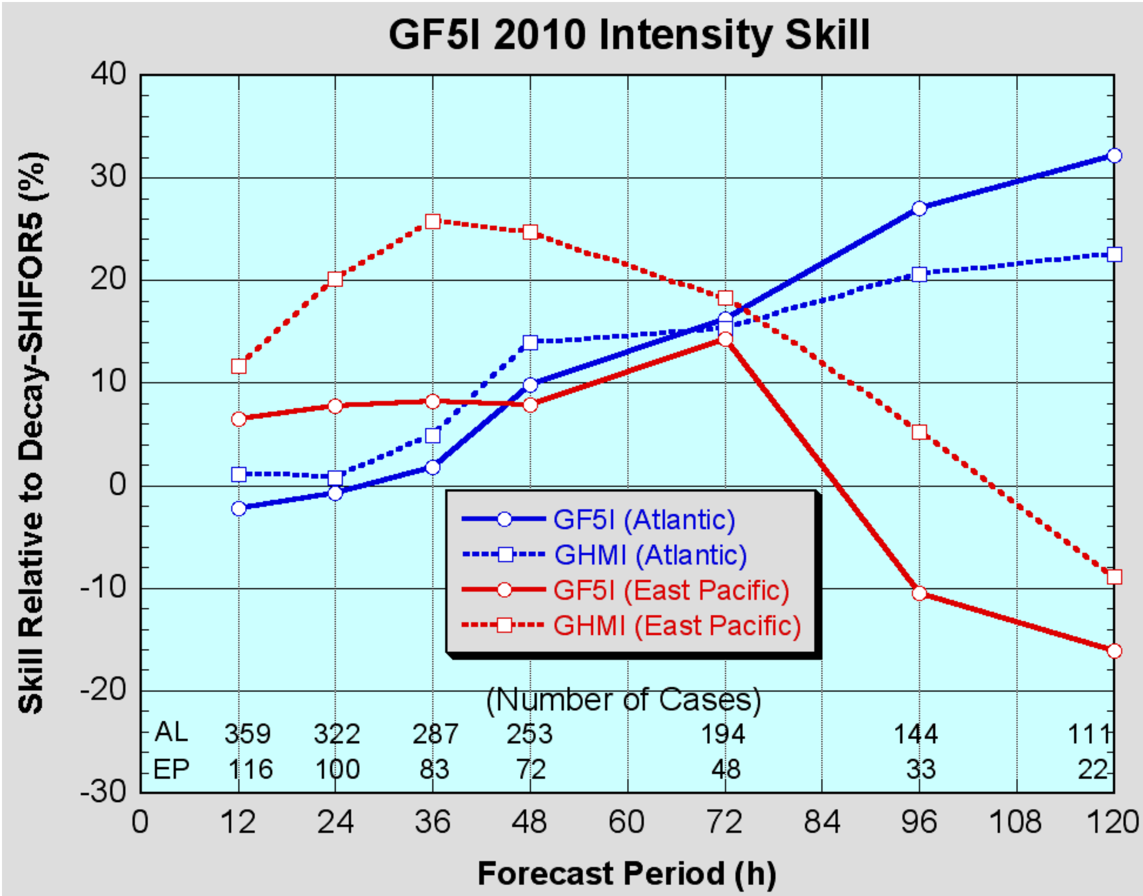


Figure 18. Comparison of intensity skill for the HFIP Stream 1.5 GF5I and operational GFDL (GHMI) in the Atlantic and eastern North Pacific basins.

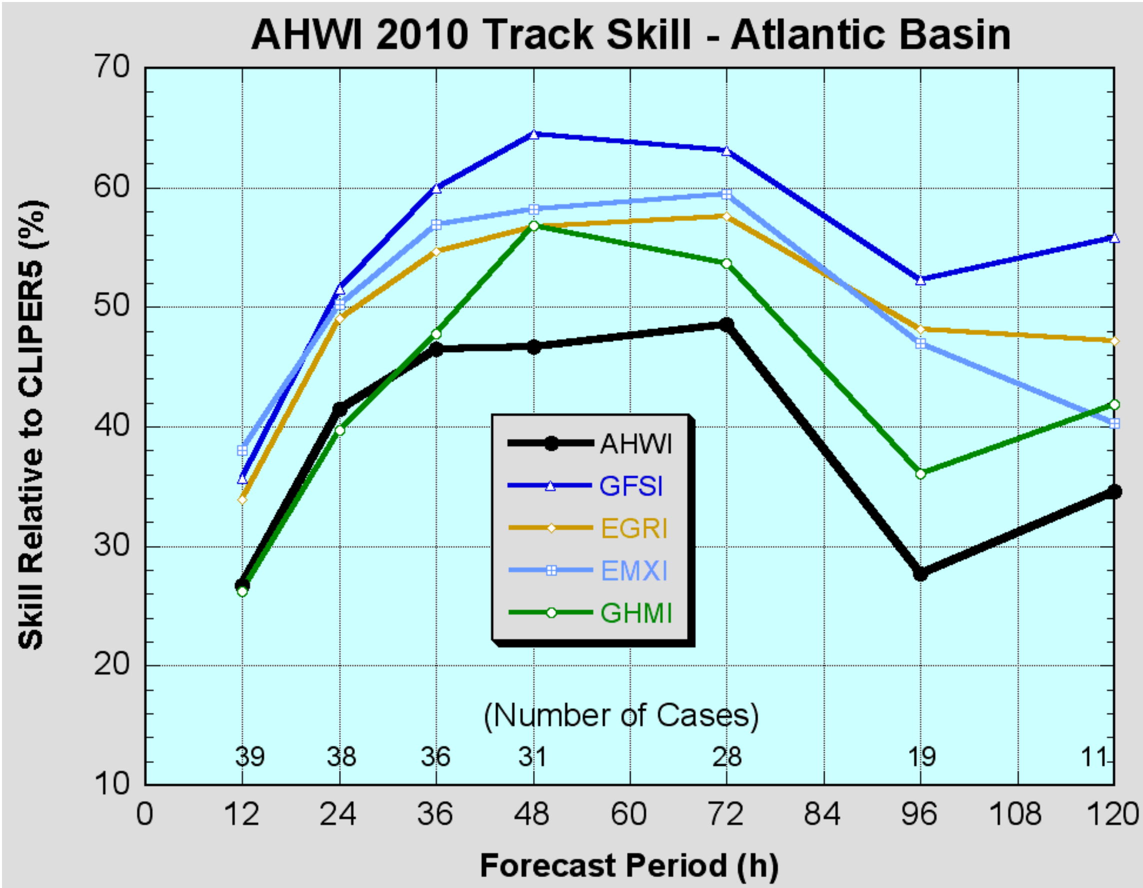


Figure 19. Comparison of track skill for the HFIP Stream 1.5 AHWI and operational GFDL (GHMI) in the Atlantic basin.

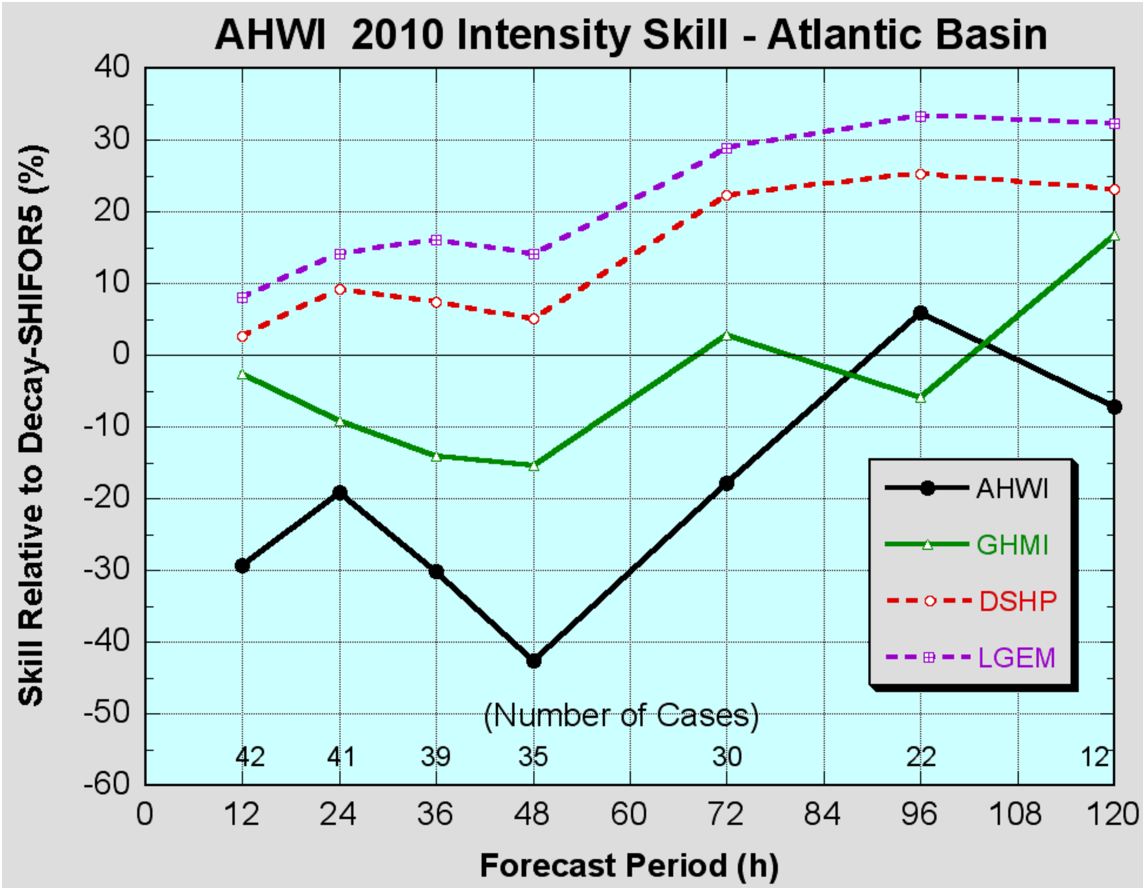


Figure 20. Comparison of intensity skill for the HFIP Stream 1.5 AHWI and operational GFDL (GHMI) in the Atlantic basin.