2007 National Hurricane Center Forecast Verification Report

James L. Franklin
National Hurricane Center
NOAA/NWS/NCEP/Tropical Prediction Center

28 February 2008

ABSTRACT

NHC official track forecasts in the Atlantic basin set records for accuracy from 36-96 h in 2007. They beat or matched the consensus models at most time periods, but generally trailed the best of the dynamical models. Examination of trends suggests that there has been little net change in forecast skill over the past several years. Among the consensus models, CGUN (the corrected version of GUNA) performed the best overall. The GFSI and UKMI/EGRI provided the best dynamical track guidance, while the GFDI and NGPI performed relatively poorly. The performance of the EMXI in 2007 was mediocre.

The 2007 Atlantic season, which featured two category 5 hurricanes and several episodes of rapid deepening, presented some unusual challenges. Intensity forecast difficulty, as measured by Decay-SHIFOR, was highly elevated compared to the previous 5-year mean, and official intensity errors in 2007 were also larger than normal. Skill levels, however, were higher in 2007 than in 2006. The statistical DSHP and LGEM models provided the best objective guidance.

In the eastern North Pacific, official track errors set records at 12-36 h. Forecast errors were below the previous 5-year mean even though the CLIPER error in 2007 was higher than its 5-year mean. The official forecast beat the individual dynamical models on average but trailed the consensus guidance. Among the dynamical models, EMXI provided the best guidance by a wide margin.

Eastern North Pacific official intensity errors were well below the 5-year averages at many time periods, setting accuracy records at 12-48 and 120 h. Despite the low errors in 2007, there has been little or no overall trend in intensity error since 1990; skill, however, appears to have increased slightly during this time. Either DSHP or LGEM, both statistical models, provided the best intensity guidance at each time period.

The 2007 season marked the first year of operational availability of the HWRF regional hurricane model. The model generally lagged its GFDL benchmark for intensity, although it significantly outperformed the GFDL for track forecasts in the Atlantic. A combination of the two models, however, generally was superior to either one alone.

Also initiated in 2007 were in-house probabilistic forecasts of tropical cyclogenesis. The verification was sufficiently favorable to begin experimental public genesis forecasts in 2008.

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 hours, and contain projections valid 12, 24, 36, 48, 72, 96, and 120[1] h after the forecast's nominal initial time (0000, 0600, 1200, or 1800 UTC)[2]. 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)[3] 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[4]. 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[5]. Except where noted to the contrary, all verifications in this report include the depression stage.

[1] NHC began making 96 and 120 h forecasts in 2001, although they were not released publicly until 2003.
[2] 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.
[3] For the remainder of this report, the term "tropical cyclone" shall be understood to also include subtropical cyclones.
[4] Possible classifications in the best track are: Tropical Depression, Tropical Storm, Hurricane, Subtropical Depression, Subtropical Storm, Extratropical, Disturbance, Wave, and Low.
[5] 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 has not been consistent over the years. The current practice of retaining and verifying the original advisory forecast began in 2005.

It is important to distinguish between *forecast error* and *forecast skill*. Track forecast error is defined as the great-circle distance between a cyclone's forecast position and the best track position at the forecast verification time. Skill, on the other hand, represents a normalization of forecast error against some standard or baseline, and 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)[6]. Errors from the CLIPER5 model are taken to represent a "no-skill" level of accuracy that can be used as a baseline for evaluating other forecasts[7]. 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 operational counterpart, is generally used to analyze lengthy historical records for which operational inputs are unavailable. Forecasters, of course, see only the operational

---

[6] CLIPER5 and SHIFOR5 are 5-day versions of the original 3-day CLIPER and SHIFOR models.

[7] To be sure, some "skill", or expertise, is required to properly initialize the CLIPER model.

version of CLIPER5, and therefore this version is the more appropriate one for the verifications discussed below.[8]

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). 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 current 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 virtually meaningless. No verifications of NHC wind radii are therefore included in this

[8] 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. If a forecaster made multiple estimates of the storm's initial motion, location, etc., over the course of the forecast cycle, then these retrospective runs may differ slightly from the operational runs in the forecast database.

report. In 2008, it is expected that the entire fleet of reconnaissance aircraft will be equipped with Stepped Frequency Microwave Radiometer (SFMR) instruments, which measure surface winds below the aircraft flight track. In time, as increasing numbers of SFMR data sets are obtained, it may be possible to do 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 forecast is released - thus 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, although some late model information is included.

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[9]. 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 at 12Z, but not at 18Z. Verification procedures here make no distinction between 6 h and 12 h interpolated models.[10]

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, but a complete description of the various model types is beyond the scope of this report. Briefly, *dynamical* models forecast by solving the physical equations governing motions in the atmosphere. These 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 model is to simply average the results from a collection of models, but other, more complex techniques can give better

---

[9] 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".

[10] The UKM and EMX models are only run out to 120 h twice a day (at 0000 and 1200 UTC). Consequently, roughly half the interpolated forecasts from these models are 12 h old.

results. The FSU super-ensemble, for example, combines its individual components on the basis of past performance in an attempt to correct for biases in those components. A consensus model that considers past error characteristics can be described as a "weighted" or "corrected" consensus[11]. Additional information about the guidance models used at the NHC can be found at http://www.nhc.noaa.gov/modelsummary.shtml.

A new dynamical model, the Hurricane Weather Research and Forecasting model (HWRF) became operational in 2007. The HWRF covers a limited area with its domain, and horizontal and vertical resolutions comparable to those of the GFDL prediction system (Bender et al. 2007). The HWRF initialization is more realistic than that of the GFDL system and allows for the inclusion of real-time observations of the inner core of a tropical cyclone.

The verifications described in this report are based on forecast and best track data sets taken from the Automated Tropical Cyclone Forecast (ATCF) System on 29 January 2008[12]. Verifications for the Atlantic and eastern North Pacific basins are given in Sections 2 and 3 below, respectively. Section 4 discusses NHC's first attempt at probabilistic genesis forecasts, conducted in house during 2007. Section 5 summarizes the key findings of the 2007 verification and previews some verification-related topics for 2008.

2. Atlantic Basin

[11] It has been argued that "consensus" is not an appropriate term for a combination of models, since consensus is defined as a general agreement among all the members of a group. One could imagine however, that if a group of disparate models were to sit down and politely settle their differences, some combination of their collective viewpoints might well be the result. In any event, the term consensus has a long history of use in meteorology for this purpose and will be retained here.

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

a. *2007 season overview – Track*

Figure 1 and Table 2 present the results of the NHC official track forecast verification for the 2007 season, along with results averaged for the previous 5-yr period 2002-2006. In 2007, the NHC issued 208 tropical cyclone forecasts, a number well below normal (about 50% of normal at 12 h and about 15% of normal at 120 h). Two storms (Dean and Noel) accounted for all of the 120-h forecasts. Mean track errors ranged from 33 n mi at 12 h to 258 n mi at 120 h. It is seen that mean official track forecast errors were smaller in 2007 than during the previous 5-yr period (by 7%-24%), and in fact, the 36-96 h forecast projections established new all-time lows. Since 1990, 24-72 h track forecast errors have been reduced by a little more than 50% (Fig. 2). Substantial vector biases at the longer ranges were noted in 2007; at 120 h the official forecast bias was 162 n mi to the east-northeast of the verifying position. These vector biases largely were caused by forecasts for Hurricane Dean that had a persistent slow (and slightly northward) bias. Examination of Table 3b reveals that official forecast biases closely tracked those of the GUNA consensus.

Track forecast skill in 2007 was comparable to skill levels over the previous 5-yr period (Table 2). An examination of skill trends (Fig. 2) suggests that after a sharp increase in skill around the beginning of the decade, there has been little change in skill since.

Table 3a presents a homogeneous[13] verification for the official forecast along with a selection of early models for 2007. In order to maximize the sample size for comparison with the official forecast, a guidance model had to be available at least two-

[13] 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.

thirds of the time at both 48 h and 120 h. For the early track models, this requirement resulted in the exclusion of GFNI, AEMI, and FSSE. Vector biases of the guidance models are given in Table 3b. Results in terms of skill are presented in Fig. 3. The figure shows that official forecast skill was generally close to that of the consensus models, beating the consensus models at 24-48 h and trailing them slightly at 72-120 h. The best dynamical models in 2007 were GFSI and EGRI, the UKMET model with subjective quality control applied to the vortex tracker. This was the first year of ATCF availability for EGRI, and its substantial improvement over UKMI through 72 h suggests some significant issues exist with the objective UKMET tracker. It's worth noting that the UKMET's strong performance in 2007 follows a year in which it was last among the major dynamical models. Trailing GFSI and EGRI in performance were the HWFI and EMXI, with the poorest performers in 2007 being NGPI and GHMI[14]. The simple trajectory model BAMM had a very strong year, with forecast skill comparable to the dynamical model consensus. This should not be a surprising result, given that the BAM models are based on the GFS, which had a very good year, and that Dean, whose long track was remarkably straight, dominated the season's sample.

Perhaps as a consequence of the year's small sample that did not allow the statistics to stabilize, there was an unusually large range in the skill of the various models in 2007. Regardless of the cause, the large variation in skill produced the unusual result that the consensus models lagged the best performing dynamical models; or put less charitably, the GFDL and NOGAPS errors were sufficiently large in 2007 that they brought down the consensus. A separate homogeneous verification of the primary consensus models is shown in Fig. 4. Keeping in mind that the sample size was quite

[14] For track, GHMI is identical to GFDI (see Table 1).

small, it can be seen that the FSSE had a mixed year, with success at the earlier forecast times but poorer performance at the longer ranges. CGUN, the corrected version of GUNA, did well. GENA, a modification of GUNA in which EGRI is substituted for UKMI, outperformed GUNA; based on this result and other anecdotal issues with the objective UKMET tracker over the past few seasons, a redefinition of GUNA is planned for 2008 (see section 5d). Although not shown here, the GFS ensemble mean (AEMI) trailed its control run by a wide margin, and the ECMWF ensemble mean also trailed its control run. While multi-model ensembles continue to provide useful tropical cyclone guidance, the same cannot yet be said for single-model ensembles.

Although late models are not available to meet forecast deadlines, for completeness a verification for a selection of these models is given in Table 4. Performance of the late models was largely similar to that of the interpolated-dynamical models discussed above. Because the season's storms were short lived, and because some of the late models are run only twice a day, this sample is exceedingly small and the results are unworthy of further comment.

Atlantic basin 48-h official track error, evaluated for tropical storms and hurricanes only, is a forecast metric tracked under the Government Performance and Results Act of 1993 (GPRA). In 2007, the GPRA goal was 110 n mi and the verification for this metric was 86 n mi.

b. *2007 season overview – Intensity*

Figure 5 and Table 5 present the results of the NHC official intensity forecast verification for the 2007 season, along with results averaged for the preceding 5-yr period. Mean forecast errors in 2007 ranged from about 8 kt at 12 h to nearly 30 kt at 96 and 120 h. These errors were considerably above the 5-year means - by 25% or more at all time periods except 24 and 36 h. Large negative forecast biases occurred at 96 and 120 h, and the biases were negative at all time periods. In contrast, long-term intensity forecast biases are near zero. It is interesting that these large errors and negative biases occurred in a year for which there were many instances of rapid strengthening[15] (11.9% of all 24 h intensity changes qualified, which is more than twice the climatological rate, and nearly four times the rate observed in 2006). This led to decay-SHIFOR errors that were well above normal; in short, this year's storms posed unusual forecast challenges. Because the decay-SHIFOR errors were so large, intensity forecast skill in 2007 was at or above the levels of recent seasons (Fig. 6).

Table 6a presents a homogeneous verification for the official forecast and the primary early intensity models for 2007. Intensity biases are given in Table 6b, and the results in terms of skill are presented in Fig. 7. In spite of the large official absolute errors discussed above, the official forecasts on average were superior to virtually all of the guidance, trailing only DSHP at 96 h and LGEM at 120 h. As has normally been the case, the best-performing intensity guidance model at each time period was a statistical model. Of the two regional hurricane models, GHMI was mostly superior to the new HWFI. Overall, the guidance was much more skillful in 2007 than in 2006, when none of the models showed skill beyond 48 h. The large low bias in the official forecasts at the longer projections essentially mirrored a low bias in the guidance.

[15] Following Kaplan and DeMaria (2003), rapid intensification is defined as a 30 kt increase in maximum winds in a 24 h period, and corresponds to the 5th percentile of all intensity changes in the Atlantic basin.

The above sample excludes FSSE because it did not meet the two-thirds availability requirement. However, a homogeneous comparison of FSSE against a simple average of the four intensity models HWFI/GHMI/DSHP/LGEM (not shown) indicated that the FSSE errors exceeded those of the simple consensus by 15%-20% in 2007.

c. Verifications for individual storms

Forecast verifications for individual storms are given in Table 7. Track errors were relatively constant over the course of the season, with no storms standing out as unusually well or poorly forecast. For intensity, forecast errors for Felix were particularly large, due in part to early track forecasts that kept the cyclone over water longer than actually occurred, and in part to missing Felix's rapid intensification. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2007atlan.shtml.

3. Eastern North Pacific Basin

a. 2007 season overview – Track

Figure 8 and Table 8 present the NHC official track forecast verification for the 2007 season in the eastern North Pacific, along with results averaged for the previous 5-yr period 2002-6. Mean track errors ranged from 30 n mi at 12 h to 186 n mi at 120 h, and were roughly 10%-20% below the 5-year means. New records for accuracy were set at 12-48 h and at 120 h. What is remarkable about these low errors is that they occurred in a year when CLIPER errors were 5%-10% above their long-term means. Figure 9 shows recent trends in track forecast accuracy and skill for the eastern North Pacific.

Errors have been reduced by roughly 20-40% for the 24-72 h forecasts since 1990, a somewhat smaller improvement than what has occurred in the Atlantic over this period, but still substantial. Forecast skill in 2007 established new records at most time periods, continuing a generally upward trend that began near the end of the last decade. Interestingly, although the track errors were relatively small in 2007, forecast biases were considerably larger than average.

Table 9a presents a homogeneous verification for the official forecast and the early track models for 2007, with vector biases of the guidance models given in Table 9b. Skill comparisons of selected models are shown in Fig. 10. Several models (UKMI, EGRI, AEMI, FSSE, and GUNA) were eliminated from the sample because they did not meet the two-thirds availability threshold. Among the surviving dynamical models, the EMXI performed best overall by a wide margin, largely on the strength of its forecasts of Kiko. GHMI and GFSI came in second and third, respectively, while the HWFI and NGPI performed relatively poorly. The BAMM, which had performed as well or better than the more sophisticated dynamical models during the past two seasons, was not as successful in 2007. Once again, the multi-model consensus CONU provided significant value over the models it comprises. (The same could not be said about the GFS ensemble mean [AEMI], which had nearly identical mean errors to GFSI in 2007 [not shown]).

A separate verification of the primary multi-model consensus aids is given in Figure 11. No single model stood out among this group. As was the case in the Atlantic, GENA was superior to GUNA, seemingly indicating issues with the UKMET's objective tracker.

A verification of late track models is given in Table 10. The results mirror the verification of the early models.

b. 2007 season overview – Intensity

Figure 12 and Table 11 present the results of the NHC eastern North Pacific intensity forecast verification for the 2007 season, along with results averaged for the preceding 5-yr period. Mean forecast errors started near 5 kt at 12 h and reached a high of 21 kt at 96 h. These errors were generally below the 5-year means. Decay-SHIFOR5 forecast errors in 2007 were also lower than their 5-year means, indicating that the season's storms were somewhat less difficult to forecast than average. A review of annual errors and skill scores (Fig. 13) indicates little net change in intensity error since 1990, although there has been a slight increase in forecast skill. Eastern North Pacific intensity forecasts have traditionally had a high bias, and this was true again in 2007.

Figure 14 and Table 12a present a homogeneous verification for the primary early intensity models for 2007. The official forecast beat all the individual guidance models through 48 h, but was beaten by one or more of the guidance models at the longer ranges. LGEM provided the best guidance overall, and at every time period the most accurate guidance model was statistical in nature. Examination of model biases (Table 12b) shows that DSHP had the largest positive biases, while the LGEM had the largest negative biases. The HWFI and GHMI biases were similar, except at 120 h, suggesting a possible different response of the two models to colder waters commonly experienced by eastern North Pacific cyclones near the end of their life cycles.

The above sample excludes FSSE because it did not meet the two-thirds availability requirement. However, a homogeneous comparison of FSSE against a simple average of the four intensity models HWFI/GHMI/DSHP/LGEM (not shown) indicated that in 2007 the FSSE outperformed the simple consensus from 12-72 h by 5%-10%. The FSSE also showed some modest skill at 24-48 h. The average errors of these two consensus techniques at longer projections were very similar.

c. Verifications for individual storms

Forecast verifications for individual storms are given for reference in Table 13. Additional discussion on forecast performance for individual storms can be found in NHC Tropical Cyclone Reports available at http://www.nhc.noaa.gov/2007epac.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 next 48 hours. In 2007, the NHC began producing in-house 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 NHC forecaster's subjective determination of the chance of TC formation during the 48 h period following the nominal TWO issuance time.

Verification was based on NHC best-track data, with the time of genesis defined to be the first tropical (or subtropical) point appearing in the best track. Verifications for the Atlantic and eastern North Pacific basins are given in Table 14. In the Atlantic, there was a very good correlation between the forecast and verifying genesis percentages (with the exception of an anomaly at 50%), and only a modest over-forecast bias. In the eastern Pacific, however, actual genesis rates were well above the forecasted rates. In addition, once the forecasted likelihood exceeded 30%, there appeared to be minimal correlation between the forecast and verifying rates.

These results suggest that division of the probability space into 10%-wide bins is too fine for the existing level of skill for a public product (at least for the eastern Pacific). A division into three bins, however, does appear to offer sufficient separation to be useful (Table 15). Based on this result, a three-tiered categorical genesis forecast will be issued publicly on an experimental basis in 2008.

5. Summary and Concluding Remarks

a. Atlantic Summary

- OFCL track forecasts established new records for accuracy from 36-96 h. They beat or matched the consensus models at most time periods, but generally trailed the best of the dynamical models.
- Among the consensus models, CGUN (the corrected version of GUNA) performed the best overall. The GFSI and UKMI/EGRI provided the best dynamical track guidance, while the GFDI and NGPI performed relatively poorly. The performance of EMXI in 2007 was mediocre.

- Atlantic official intensity errors were higher than the 5-year means, largely due to above average storm intensity and frequency of rapid deepening. Skill levels were higher in 2007 than they had been the year before, and slightly above the 5-year means. The official forecast mostly beat the objective guidance, the best of which were the statistical DSHP and LGEM.

b. Eastern North Pacific Summary

- Official track errors in the eastern North Pacific set records for accuracy at 12-36 h. Forecast errors were below the previous 5-year mean even though the CLIPER5 error in 2007 was higher than its 5-year mean. The official forecast beat the individual dynamical models on average but trailed the consensus guidance.
- The consensus model CONU in the eastern North Pacific was better than any of its components. Among the dynamical models, EMXI provided the best guidance by a wide margin.
- Eastern North Pacific official intensity errors were well below the 5-year averages at many time periods, setting accuracy records at 12-48 and 120 h. A statistical model provided the best intensity guidance for every time period.

c. Track Forecast Cone Sizes for 2008

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 a 5-year sample fall within the circle. The circle radii defining the cones in 2008 for the Atlantic and eastern North Pacific basins (based on error distributions for 2003-7) are given below. In the Atlantic, the 96 and 120 h circles will be about 20 n mi smaller than they were last year, while the differences at other times will be relatively small. The eastern North Pacific circles will be essentially unchanged for 2008.

Track Forecast Cone Two-Thirds Probability Circles for 2008 (n mi)		
Forecast Period (h)	**Atlantic Basin**	**Eastern North Pacific Basin**
12	39	36
24	67	66
36	92	92
48	118	115
72	170	161
96	233	210
120	305	256

d. Looking Ahead

Some changes are planned to the content and nomenclature of the consensus models used by the NHC in 2008 and beyond. The new system defines a set of consensus model identifiers that will 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., CONU). 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 new consensus nomenclature scheme defines the following consensus models for 2008:

NHC Consensus Model Definitions For 2008			
Model ID	**Parameter**	**Type**	**Members**
TCON	Track	Fixed	GFSI EGRI NGPI GHMI HWFI
ICON	Intensity	Fixed	DSHP LGEM GHMI HWFI
TVCN	Track	Variable	GFSI EGRI NGPI GHMI HWFI GFNI EMXI
IVCN	Intensity	Variable	DSHP LGEM GHMI HWFI GFNI
TCCN	Track	Fixed (corrected)	GFSI EGRI NGPI GHMI HWFI
TVCC	Track	Variable (corrected)	GFSI EGRI NGPI GHMI HWFI GFNI EMXI

In addition to the models listed above, GUNA (and its corrected version, CGUN) will continue to be computed, except that EGRI will replace UKMI, when available. CONU and CCON will no longer be computed, being replaced by TVCN and TVCC, respectively.

Experimental quantitative forecasts of tropical cyclone genesis will continue in 2008. Although the quantitative forecasts will not be publicly disseminated, they will form the basis of the categorical (i.e., low/medium/high likelihood) genesis forecasts that will be issued as part of an experimental Graphical Tropical Weather Outlook.

Acknowledgments:

The author gratefully acknowledges Chris Sisko of TPC, keeper of the NHC forecast databases, and Hurricane Specialist Dan Brown for maintaining the genesis forecast database.

6. References

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

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, T. Marchok, 2007: The operational GFDL coupled hurricane-ocean prediction system and a summary of its performance. *Mon. Wea. Rev*., **135**, 3965-3989.

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.

Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting,* **18**, 1093-1108.

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.

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
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
A98E	NHC98 (Atlantic)	Statistical-dynamical	E	Trk
P91E	NHC91 (Pacific)	Statistical-dynamical	E	Trk

ID	Name/Description	Type	Timeliness (E/L)	Parameters forecast
CLP5	CLIPER5 (Climatology and Persistence model)	Statistical (baseline)	E	Trk
SHF5	SHIFOR5 (Climatology and Persistence model)	Statistical (baseline)	E	Int
DSF5	DSHIFOR5 (Climatology and Persistence model)	Statistical (baseline)	E	Int
OCD5	CLP5 (track) and DSF5 (intensity) models merged	Statistical (baseline)	E	Trk, Int
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
GUNA	Average of GFDI, UKMI, NGPI, and GFSI	Consensus	E	Trk
GENA	Average of GFDI, EGRI, NGPI, 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
CONU	Average of at least 2 of GFDI, UKMI, NGPI, GFSI, and GFNI	Consensus	E	Trk
CCON	Version of CONU corrected for model biases	Corrected consensus	E	Trk
FSSE	FSU Super-ensemble	Corrected consensus	E	Trk, Int

Table 2. Homogenous comparison of official and CLIPER5 track forecast errors in the Atlantic basin for the 2007 season for all tropical cyclones. Averages for the previous 5-year period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2007 mean OFCL error (n mi)	32.8	51.2	70.7	91.9	146.0	167.2	258.4
2007 mean CLIPER5 error (n mi)	45.3	85.4	121.5	160.1	237.4	323.0	512.3
2007 mean OFCL error relative to CLIPER5 (%)	-28	-40	-42	-43	-38	-48	-50
2007 mean OFCL bias vector (°/n mi)	341/3	001/7	026/17	035/34	046/75	059/107	069/162
2007 number of cases	177	145	116	93	62	39	23
2002-2006 mean OFCL error (n mi)	35.3	61.0	86.3	111.8	161.6	220.9	290.0
2002-2006 mean CLIPER5 error (n mi)	48.0	100.3	159.6	215.6	318.4	418.5	509.7
2002-2006 mean OFCL error relative to CLIPER5 (%)	-26	-39	-46	-48	-49	-47	-43
2002-2006 mean OFCL bias vector (°/n mi)	309/6	316/14	322/21	324/27	321/24	354/19	035/39
2002-2006 number of cases	1852	1686	1519	1362	1100	885	723
2007 OFCL error relative to 2002-2006 mean (%)	-7	-16	-18	-17	-9	-24	-10
2007 CLIPER5 error relative to 2002-2006 mean (%)	-6	-15	-24	-26	-25	-23	1

Table 3a. Homogenous comparison of Atlantic basin early track guidance model errors (n mi) for 2007. 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	29.8	48.9	73.3	94.8	132.2	149.6	229.2
OCD5	41.6	79.9	119.5	161.2	242.0	361.7	586.4
GFSI	35.9	53.9	74.6	**91.0**	**114.1**	**123.7**	**147.9**
GHMI	32.7	56.3	86.8	122.0	193.2	305.9	417.5
HWFI	36.6	63.0	90.7	111.9	151.0	193.0	296.0
UKMI	35.0	71.5	101.3	**81.2**	**110.6**	**132.6**	**166.6**
EGRI	34.0	53.2	**69.9**	**76.9**	**103.0**	**140.6**	**170.5**
EMXI	40.9	67.1	91.2	121.1	178.5	219.9	**219.2**
NGPI	39.6	70.2	106.3	149.6	222.9	252.6	329.3
GUNA	**29.4**	52.3	78.7	96.1	**124.9**	**147.9**	**217.2**
CONU	30.4	53.1	81.7	99.7	**126.5**	**143.9**	**213.9**
BAMS	47.6	85.2	120.0	151.5	180.8	221.9	**206.2**
BAMM	34.8	55.5	75.6	97.3	**130.9**	162.9	**184.1**
BAMD	39.6	66.7	91.8	114.2	156.0	197.9	**214.4**
# Cases	108	90	76	61	39	23	14

Table 3b. Homogenous comparison of Atlantic basin early track guidance model bias vectors (º/n mi) for 2007.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	035/4	023/17	032/34	039/50	043/80	041/100	048/181
OCD5	114/4	039/10	054/32	057/75	059/208	061/325	063/582
GFSI	329/9	356/17	023/29	036/37	073/052	068/56	069/127
GHMI	066/5	037/16	031/31	036/49	031/93	021/182	020/274
HWFI	010/10	015/26	026/48	043/74	060/116	074/174	075/286
UKMI	031/6	009/32	010/60	017/35	310/47	315/56	348/120
EGRI	039/5	021/19	008/31	007/31	312/40	307/59	344/100
EMXI	087/11	077/27	076/49	082/79	082/132	094/137	095/193
NGPI	052/15	052/40	059/70	066/111	067/187	075/221	082/288
GUNA	032/7	028/25	033/45	048/55	049/76	045/101	045/163
CONU	048/7	036/23	039/44	051/57	048/77	044/94	043/156
BAMS	300/19	314/039	317/56	315/63	303/62	288/49	081/118
BAMM	300/7	338/17	345/25	358/27	020/38	044/50	076/117
BAMD	069/12	058/28	056/47	058/74	056/127	053/174	039/183
# Cases	108	90	76	61	39	23	14

Table 4. Homogenous comparison of Atlantic basin late track guidance model errors (n mi) for 2007. Errors from OCD5, an early model, are shown for comparison. The smallest error at each time period is displayed in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
GFDL	34.8	**55.4**	77.0	112.8	182.7	273.6	393.8
HWRF	**30.5**	62.1	89.8	115.6	158.2	208.1	260.5
UKM	41.3	60.3	91.4	143.2	**100.0**	185.1	**165.2**
NGPS	40.0	66.7	97.4	132.2	217.5	272.5	321.9
GFSO	36.2	56.8	**70.0**	**90.2**	120.9	**130.1**	195.3
EMX	48.6	75.7	95.5	121.4	190.6	252.7	231.0
OCD5	42.8	83.6	120.7	160.8	240.8	344.9	553.6
# Cases	62	50	42	34	22	15	9

Table 5. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the Atlantic basin for the 2007 season for all tropical cyclones. Averages for the previous 5-year period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2007 mean OFCL error (kt)	8.1	11.0	14.0	17.9	23.5	28.6	30.0
2007 mean Decay-SHIFOR5 error (kt)	9.8	12.6	17.4	23.5	29.8	39.0	42.7
2007 mean OFCL error relative to Decay-SHIFOR5 (%)	-17	-13	-20	-24	-21	-27	-30
2007 OFCL bias (kt)	-0.5	-1.1	-1.3	-0.4	-1.4	-4.5	-12.6
2007 number of cases	177	145	116	93	62	39	23
2002-6 mean OFCL error (kt)	6.4	9.8	12.0	14.1	18.3	19.8	21.8
2002-6 mean Decay-SHIFOR5 error (kt)	7.6	11.5	14.8	17.6	21.3	23.7	24.3
2002-6 mean OFCL error relative to Decay-SHIFOR5 (%)	-16	-15	-19	-20	-14	-17	-10
2002-6 OFCL bias (kt)	0.3	0.7	0.5	0.0	-0.2	-1.0	-0.8
2002-6 number of cases	1852	1686	1519	1362	1100	885	723
2007 OFCL error relative to 2002-6 mean (%)	26	12	17	27	28	44	38
2007 Decay-SHIFOR5 error relative to 2002-6 mean (%)	29	10	18	34	40	65	76

Table 6a. Homogenous comparison of selected Atlantic basin early intensity guidance model errors (kt) for 2007. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	8.4	11.2	14.2	18.0	23.3	28.8	30.0
OCD5	9.9	12.7	17.6	23.5	29.7	39.4	42.7
HWFI	10.0	13.2	16.4	22.6	26.7	30.9	39.0
GHMI	10.1	12.8	17.5	20.8	25.6	30.2	34.0
DSHP	9.7	11.8	14.3	19.6	24.1	**27.9**	31.6
LGEM	10.0	12.4	15.2	19.9	23.7	30.8	**27.3**
# Cases	167	139	113	91	61	38	23

Table 6b. Homogenous comparison of selected Atlantic basin early intensity guidance model biases (kt) for 2007. Biases smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	-0.5	-1.4	-1.5	-0.7	-2.0	-4.1	-12.6
OCD5	-0.9	-1.6	**-0.7**	-1.3	-4.2	-13.4	-29.3
HWFI	-4.3	-8.4	-10.6	-11.3	-11.8	-10.4	-12.8
GHMI	-2.1	-3.6	-3.6	-4.3	**1.4**	4.9	**-8.6**
DSHP	-0.9	**-0.4**	**0.2**	**0.4**	**0.3**	-4.3	-23.6
LGEM	-1.9	-3.0	-2.9	-2.4	**0.9**	**0.8**	**-9.8**
# Cases	167	139	113	91	61	38	23

Table 7. Official Atlantic track and intensity forecast verifications (OFCL) for 2007 by storm. CLIPER5 and Decay-SHIFOR5 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: AL012007 ANDREA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	5.9	5.9	7	0.7	0.7
012	5	22.0	40.5	5	1.0	4.4
024	3	46.1	113.2	3	5.0	2.7
036	1	67.0	203.3	1	5.0	8.0
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: AL022007 BARRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	3.6	3.6	5	3.0	2.0
012	3	76.6	126.6	3	5.0	11.3
024	1	50.2	170.9	1	15.0	24.0
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: AL032007 CHANTAL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	8.7	8.7	5	1.0	2.0
012	3	36.3	74.8	3	6.7	8.7
024	1	0.0	167.7	1	10.0	8.0
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: AL042007 DEAN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	39	3.7	3.8	39	3.6	3.3
012	37	21.4	31.5	37	9.6	9.5
024	35	38.8	65.2	35	12.0	12.4
036	33	57.9	102.9	33	13.5	16.5
048	31	81.6	145.6	31	13.7	19.7
072	27	140.8	252.8	27	21.5	34.1
096	23	197.8	391.9	23	32.6	51.0
120	19	274.0	566.5	19	32.1	46.9

Verification statistics for: AL052007 ERIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	13.2	16.6	7	0.7	0.7
012	7	41.9	55.4	7	5.7	6.1
024	6	69.7	112.1	6	5.0	6.8
036	4	96.6	181.7	4	7.5	13.0
048	2	110.0	323.2	2	2.5	23.5
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: AL062007 FELIX

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	19	8.5	8.2	19	4.7	3.7
012	17	21.1	26.0	17	17.9	22.4
024	15	36.6	63.6	15	25.7	26.9
036	13	52.7	111.6	13	35.4	41.9
048	11	70.5	175.3	11	53.2	60.4
072	7	106.0	276.2	7	55.7	52.3
096	3	120.1	418.0	3	35.0	24.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL072007 GABRIELLE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	14	8.4	8.4	14	1.4	2.1
012	12	19.6	34.5	12	5.0	7.4
024	10	31.5	86.9	10	6.0	7.4
036	8	46.5	123.7	8	11.9	15.9
048	6	61.9	141.6	6	15.8	26.0
072	2	154.5	170.0	2	15.0	42.5
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: AL082007 INGRID

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	19	12.6	12.6	19	1.3	1.3
012	17	28.0	32.1	17	1.8	2.8
024	15	52.3	62.1	15	4.0	6.3
036	13	88.5	100.5	13	4.2	9.9
048	11	115.0	132.4	11	7.7	14.5
072	7	144.9	150.1	7	17.9	20.9
096	3	154.4	215.0	3	28.3	28.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL092007 HUMBERTO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	0.9	1.8	6	4.2	5.8
012	5	25.6	41.0	5	18.0	20.6
024	3	50.5	99.0	3	11.7	14.7
036	1	88.8	141.6	1	5.0	3.0
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: AL102007 TEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	3	8.2	8.2	3	0.0	0.0
012	1	42.3	59.9	1	10.0	13.0
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: AL112007 JERRY

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	7	3.3	3.3	7	1.4	1.4
012	5	27.6	56.4	5	3.0	4.0
024	3	51.5	138.5	3	10.0	9.7
036	1	100.5	238.5	1	10.0	20.0
048	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL122007 KAREN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	10.8	10.8	18	1.7	2.5
012	16	51.6	55.8	16	5.6	8.6
024	14	72.3	76.5	14	11.4	16.2
036	12	96.3	86.8	12	12.9	21.8
048	10	106.1	93.4	10	14.5	21.3
072	6	142.2	139.3	6	21.7	21.5
096	2	185.9	201.9	2	25.0	28.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: AL132007 LORENZO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	13	3.7	5.0	13	2.3	2.3
012	11	30.7	41.1	11	13.2	13.6
024	9	44.6	80.6	9	11.1	15.8
036	7	51.4	124.2	7	18.6	16.6
048	5	67.8	171.6	5	24.0	20.0
072	1	180.1	217.7	1	35.0	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: AL142007 MELISSA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	9	8.6	8.6	9	1.1	2.2
012	7	33.7	42.4	7	4.3	6.0
024	5	40.8	63.3	5	5.0	7.8
036	3	22.3	32.8	3	5.0	12.0
048	1	72.0	23.1	1	10.0	27.0
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: AL152007 FIFTEEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	5	5.7	5.7	5	0.0	0.0
012	3	29.4	81.5	3	3.3	7.3
024	1	78.5	214.9	1	5.0	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: AL162007 NOEL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	24	11.8	12.3	24	0.4	1.5
012	22	51.1	65.0	22	8.0	9.0
024	20	78.0	131.1	20	10.3	10.4
036	18	97.4	185.5	18	11.7	10.0
048	16	119.7	230.1	16	12.2	13.3
072	12	179.5	292.7	12	13.8	13.8
096	8	97.1	159.9	8	15.6	16.9
120	4	184.1	255.0	4	20.0	22.3

Verification statistics for: AL172007 OLGA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	8	12.4	12.4	8	2.5	2.5
012	6	47.1	64.1	6	6.7	8.7
024	4	61.4	83.5	4	8.8	8.0
036	2	51.8	146.8	2	5.0	2.0
048	0	-999.0	-999.0	0	-999.0	-999.0
072	0	-999.0	-999.0	0	-999.0	-999.0
096	0	-999.0	-999.0	0	-999.0	-999.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Table 8. Homogenous comparison of official and CLIPER5 track forecast errors in the eastern North Pacific basin for the 2007 season for all tropical cyclones. Averages for the previous 5-year period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2007 mean OFCL error (n mi)	30.0	50.2	71.4	92.5	117.2	146.9	186.3
2007 mean CLIPER5 error (n mi)	39.9	80.1	124.6	169.1	249.5	304.3	343.0
2007 mean OFCL error relative to CLIPER5 (%)	-24.8	-37.3	-42.7	-45.3	-53.0	-51.7	-45.7
2007 mean OFCL bias vector (°/n mi)	281/7	279/17	275/30	269/41	258/44	231/22	112/37
2007 number of cases	208	182	156	140	108	77	52
2002-6 mean OFCL error (n mi)	33.1	56.8	79.1	98.9	139.6	188.1	233.1
2002-6 mean CLIPER5 error (n mi)	39.4	76.8	117.8	155.1	225.2	286.7	351.4
2002-6 mean OFCL error relative to CLIPER5 (%)	-16.0	-26.0	-32.9	-36.2	-38.0	-34.4	-33.7
2002-6 mean OFCL bias vector (°/n mi)	319/12	312/3	310/6	309/12	301/10	283/6	270/17
2002-6 number of cases	1349	1192	1039	897	655	465	311
2007 OFCL error relative to 2002-6 mean (%)	-9.4	-11.6	-9.7	-6.5	-16.0	-21.9	-20.1
2007 CLIPER5 error relative to 2002-6 mean (%)	1.3	4.3	5.8	9.0	10.8	6.1	-2.4

Table 9a. Homogenous comparison of eastern North Pacific basin early track guidance model errors (n mi) for 2007. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	26.7	44.5	64.5	82.1	107.5	143.6	172.5
OCD5	35.8	70.1	109.3	147.9	223.2	273.1	297.0
GFSI	35.0	61.4	88.6	115.2	158.7	175.1	200.5
GHMI	31.8	55.7	81.0	103.8	144.1	163.6	188.3
HWFI	36.8	66.1	91.8	120.7	178.3	221.4	271.9
NGPI	33.2	55.5	80.6	109.5	169.5	263.4	371.7
EMXI	29.5	48.5	65.2	87.8	116.5	145.5	223.5
CONU	27.3	**43.7**	**59.6**	**76.9**	**105.0**	145.0	189.9
LBAR	40.3	86.4	143.4	197.1	293.4	377.6	429.6
BAMD	45.4	84.0	118.9	148.3	205.1	254.6	365.9
BAMM	39.2	71.2	104.0	132.3	189.8	234.6	277.6
BAMS	36.8	68.7	102.4	136.7	190.0	213.5	239.4
# Cases	129	112	101	87	65	41	26

Table 9b. Homogenous comparison of eastern North Pacific basin early track guidance model bias vectors (º/n mi) for 2007.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	312/5	292/11	281/20	271/31	268/23	207/13	095/15
OCD5	323/3	266/6	253/14	252/37	246/37	171/14	066/49
GFSI	269/16	259/31	254/48	248/71	236/76	195/70	174/84
GHMI	005/7	343/14	323/24	321/29	000/60	018/90	029/102
HWFI	313/19	298/38	287/57	282/078	290/103	288/124	282/157
NGPI	349/2	217/3	220/6	227/10	212/6	034/16	011/50
EMXI	124/6	163/13	174/21	175/34	172/47	149/38	242/57
CONU	302/4	267/8	263/15	257/22	279/15	025/17	061/50
LBAR	331/14	319/49	312/90	307/127	319/172	339/155	027/178
BAMD	313/24	305/47	298/67	288/88	290/97	284/72	307/50
BAMM	332/21	316/41	305/63	294/83	285/109	254/111	242/111
BAMS	341/16	320/30	307/48	293/69	281/89	249/72	254/72
# Cases	129	112	101	87	65	41	26

Table 10. Homogenous comparison of eastern North Pacific basin late track guidance model errors (n mi) for 2007. Errors from CLP5, an early model, are shown for comparison. The smallest errors at each time period are displayed in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
GFDL	36.2	56.4	80.7	106.9	139.7	157.6	182.2
GFDN	38.5	64.6	92.1	115.2	172.6	222.4	280.7
NGPS	39.5	58.8	80.9	104.0	154.1	233.4	279.0
GFSO	43.6	66.6	84.0	106.7	150.4	172.6	196.4
EMX	**33.3**	**48.0**	**65.1**	**82.1**	**111.7**	**136.9**	**153.6**
CLP5	38.7	73.8	114.2	156.4	237.1	301.3	352.7
# Cases	70	61	56	50	36	25	13

Table 11. Homogenous comparison of official and Decay-SHIFOR5 intensity forecast errors in the eastern North Pacific basin for the 2007 season for all tropical cyclones. Averages for the previous 5-year period are shown for comparison.

	Forecast Period (h)						
	12	24	36	48	72	96	120
2007 mean OFCL error (kt)	5.1	8.2	11.6	14.4	18.1	20.8	17.0
2007 mean Decay-SHIFOR5 error (kt)	5.9	9.3	12.0	14.3	17.3	18.5	19.0
2007 mean OFCL error relative to Decay-SHIFOR5 (%)	-13.6	-11.8	-3.3	0.7	4.6	12.4	-10.5
2007 OFCL bias (kt)	1.2	2.3	3.9	4.4	3.8	1.3	-2.6
2007 number of cases	208	182	156	140	108	77	52
2002-6 mean OFCL error (kt)	6.3	11.0	14.6	16.9	18.9	18.5	19.3
2002-6 mean Decay-SHIFOR5 error (kt)	7.2	12.0	15.7	18.4	21.5	21.5	21.1
2002-6 mean OFCL error relative to Decay-SHIFOR5 (%)	-12.5	-8.3	-7.0	-8.2	-12.1	-14.0	-8.5
2002-6 OFCL bias (kt)	0.7	1.9	2.8	2.6	4.1	3.9	1.4
2002-6 number of cases	1349	1192	1039	896	655	465	311
2007 OFCL error relative to 2002-6 mean (%)	-19	-25	-20	-15	-4	12	-12
2007 Decay-SHIFOR5 error relative to 2002-6 mean (%)	-18	-22	-24	-22	-19	-14	-10

Table 12a. Homogenous comparison of eastern North Pacific basin early intensity guidance model errors (kt) for 2007. Errors smaller than the NHC official forecast are shown in boldface.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	5.3	8.5	11.7	14.5	19.0	22.2	19.5
OCD5	6.0	9.5	12.4	14.6	**18.8**	**19.2**	**18.5**
HWFI	7.4	11.7	15.2	18.1	20.5	27.0	26.7
GHMI	7.3	11.9	16.3	18.1	19.1	**20.3**	**19.2**
DSHP	5.9	9.8	13.2	16.8	20.5	**22.0**	**17.8**
LGEM	6.1	10.0	13.1	16.4	**18.5**	**19.6**	20.0
# Cases	165	144	126	110	85	61	42

Table 12b. Homogenous comparison of eastern North Pacific basin early intensity guidance model biases (kt) for 2007.

	Forecast Period (h)						
Model ID	12	24	36	48	72	96	120
OFCL	0.9	1.6	2.8	2.5	0.9	-1.6	-4.3
OCD5	0.8	1.3	1.9	1.4	2.2	0.0	2.0
HWFI	-0.6	-0.9	0.2	1.5	-0.6	-2.7	-7.0
GHMI	-0.7	0.0	1.1	1.4	-0.1	-0.5	-0.1
DSHP	1.4	3.4	4.8	5.9	3.4	2.2	0.5
LGEM	0.1	-0.5	-1.5	-2.3	-5.8	-8.2	-8.2
# Cases	165	144	126	110	85	61	42

Table 13. Official eastern North Pacific track and intensity forecast verifications (OFCL) for 2007 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: EP012007 ALVIN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	20	11.1	11.0	20	0.8	0.8
012	18	32.4	37.7	18	2.8	5.0
024	16	53.4	74.8	16	6.3	6.8
036	14	70.6	114.4	14	11.4	9.7
048	12	90.1	154.4	12	13.3	11.8
072	8	90.8	276.4	8	16.9	14.3
096	4	139.1	331.2	4	25.0	18.5
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP022007 BARBARA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	7.5	8.8	17	0.6	0.9
012	15	32.4	44.1	15	6.7	6.8
024	13	61.8	105.6	13	10.0	11.0
036	11	97.9	176.3	11	12.3	11.8
048	9	141.4	252.5	9	19.4	16.4
072	5	226.0	453.7	5	29.0	10.0
096	1	333.2	598.9	1	60.0	19.0
120	0	-999.0	-999.0	0	-999.0	-999.0

Verification statistics for: EP032007 THREE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	7.2	7.2	6	0.0	0.8
012	4	27.8	38.6	4	3.8	4.0
024	2	53.0	86.0	2	7.5	9.5
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: EP042007 FOUR

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	18.7	18.7	6	0.0	0.8
012	4	35.2	39.4	4	2.5	4.0
024	2	58.5	66.2	2	0.0	8.0
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: EP052007 FIVE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	9.4	9.4	6	0.0	0.0
012	4	25.8	30.8	4	5.0	4.8
024	2	23.0	55.1	2	7.5	12.0
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: EP062007 COSME

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	18	12.3	12.6	18	1.1	1.4
012	18	32.6	42.8	18	6.4	7.4
024	18	50.0	83.0	18	8.9	12.2
036	18	64.1	122.3	18	11.4	14.8
048	18	75.4	163.1	18	10.3	15.1
072	18	101.9	264.6	18	8.3	16.6
096	17	154.6	380.8	17	8.2	17.1
120	13	232.2	503.3	13	4.2	19.5

Verification statistics for: EP072007 DALILA

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	23	11.9	12.9	23	1.7	1.3
012	21	31.9	43.9	21	3.8	5.6
024	19	47.6	79.3	19	6.3	8.5
036	17	69.3	120.6	17	5.3	8.4
048	15	98.8	170.8	15	8.0	9.4
072	11	142.4	264.4	11	11.4	11.2
096	7	164.6	401.2	7	15.0	18.0
120	3	186.7	534.6	3	16.7	26.0

Verification statistics for: EP082007 ERICK

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	12.1	14.2	6	0.8	1.7
012	4	25.9	35.6	4	2.5	4.5
024	2	49.4	74.7	2	5.0	8.5
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: EP092007 FLOSSIE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	12	7.5	7.5	12	2.9	4.2
012	12	25.4	29.8	12	12.5	13.2
024	12	40.8	50.4	12	21.7	22.6
036	12	49.9	66.2	12	32.9	31.5
048	12	52.0	78.1	12	45.8	38.9
072	12	52.2	97.2	12	56.3	43.8
096	12	63.2	105.6	12	48.3	38.2
120	12	104.5	102.6	12	34.6	30.8

Verification statistics for: EP102007 GIL

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	17	7.9	7.9	17	0.9	1.2
012	15	24.7	28.1	15	2.7	3.2
024	13	43.3	50.2	13	5.0	5.7
036	11	66.5	66.1	11	6.4	11.0
048	9	86.5	88.0	9	6.1	13.3
072	5	104.5	178.6	5	0.0	22.6
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: EP112007 HENRIETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	27	11.4	11.9	27	1.5	1.9
012	25	33.4	42.7	25	5.6	5.1
024	23	57.9	83.4	23	7.2	5.9
036	21	80.4	132.1	21	10.5	8.0
048	19	100.7	178.1	19	8.7	7.5
072	15	126.5	220.2	15	6.7	8.5
096	11	172.4	240.6	11	8.6	14.3
120	7	195.5	253.6	7	7.9	21.3

Verification statistics for: EP122007 IVO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	21	6.7	7.7	21	1.9	2.1
012	19	18.4	28.1	19	7.4	7.3
024	17	25.7	58.7	17	12.6	10.2
036	15	38.0	104.7	15	16.3	12.7
048	13	48.3	164.6	13	20.0	16.2
072	9	52.4	308.9	9	22.2	21.9
096	5	97.3	430.5	5	20.0	24.6
120	1	103.0	456.7	1	5.0	30.0

Verification statistics for: EP132007 THIRTEEN

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	6	5.8	7.8	6	0.8	0.8
012	4	20.8	40.6	4	3.8	4.5
024	2	44.1	92.8	2	7.5	4.5
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: EP142007 JULIETTE

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	12	17.0	17.0	12	2.5	2.5
012	11	28.1	48.2	11	3.2	6.5
024	9	39.9	94.7	9	3.9	8.3
036	7	48.4	163.0	7	3.6	6.1
048	5	82.7	257.3	5	4.0	7.2
072	1	212.6	479.7	1	10.0	27.0
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: EP152007 KIKO

VT (h)	NT	OFCL	OCD5	NI	OFCL	OCD5
000	36	9.5	11.4	36	0.7	0.8
012	34	35.1	47.5	34	4.4	4.3
024	32	63.6	100.9	32	6.1	7.5
036	30	94.0	154.7	30	8.8	9.9
048	28	121.6	196.8	28	11.8	11.3
072	24	152.8	257.3	24	17.5	12.0
096	20	174.9	308.2	20	21.0	8.7
120	16	211.5	389.3	16	19.1	6.7

Table 14a. Verification of experimental in-house probabilistic genesis forecasts for the Atlantic basin in 2007.

Atlantic Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	1	192
10	6	197
20	12	129
30	24	76
40	29	38
50	15	20
60	57	23
70	62	13
80	75	8
90	88	8
100	100	1

Table 14b. Verification of experimental in-house probabilistic genesis forecasts for the eastern North Pacific basin in 2007.

Eastern North Pacific Basin Genesis Forecast Reliability Table		
Forecast Likelihood (%)	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0	3	68
10	7	111
20	30	105
30	63	30
40	83	12
50	100	15
60	87	15
70	100	5
80	80	5
90	100	3
100	100	1

Table 15a. Verification of experimental in-house probabilistic genesis forecasts for the Atlantic basin in 2007.

Atlantic Basin Genesis Forecast Reliability Table			
Forecast Likelihood (%)	**Expected Genesis Occurrence Rate (%)**	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0-10	5	3	389
20-50	28	18	263
60-100	71	66	53

Table 15b. Verification of experimental in-house probabilistic genesis forecasts for the eastern North Pacific basin in 2007.

Eastern North Pacific Basin Genesis Forecast Reliability Table			
Forecast Likelihood (%)	**Expected Genesis Occurrence Rate (%)**	**Verifying Genesis Occurrence Rate (%)**	**Number of Forecasts**
0-10	6	6	179
20-50	26	47	162
60-100	70	90	29

List of Figures

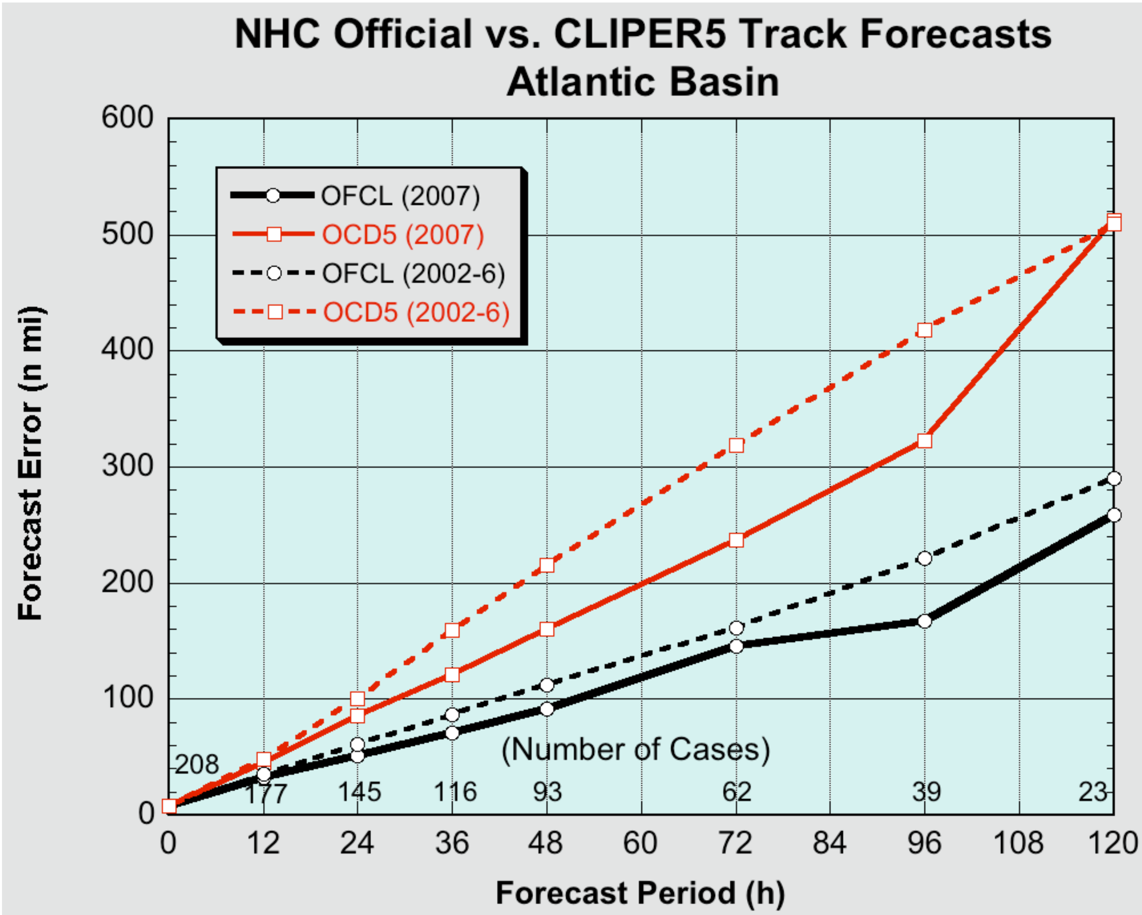


Figure 1. NHC official and CLIPER5 (OCD5) Atlantic basin average track errors for 2007 (solid lines) and 2002-2006 (dashed lines).

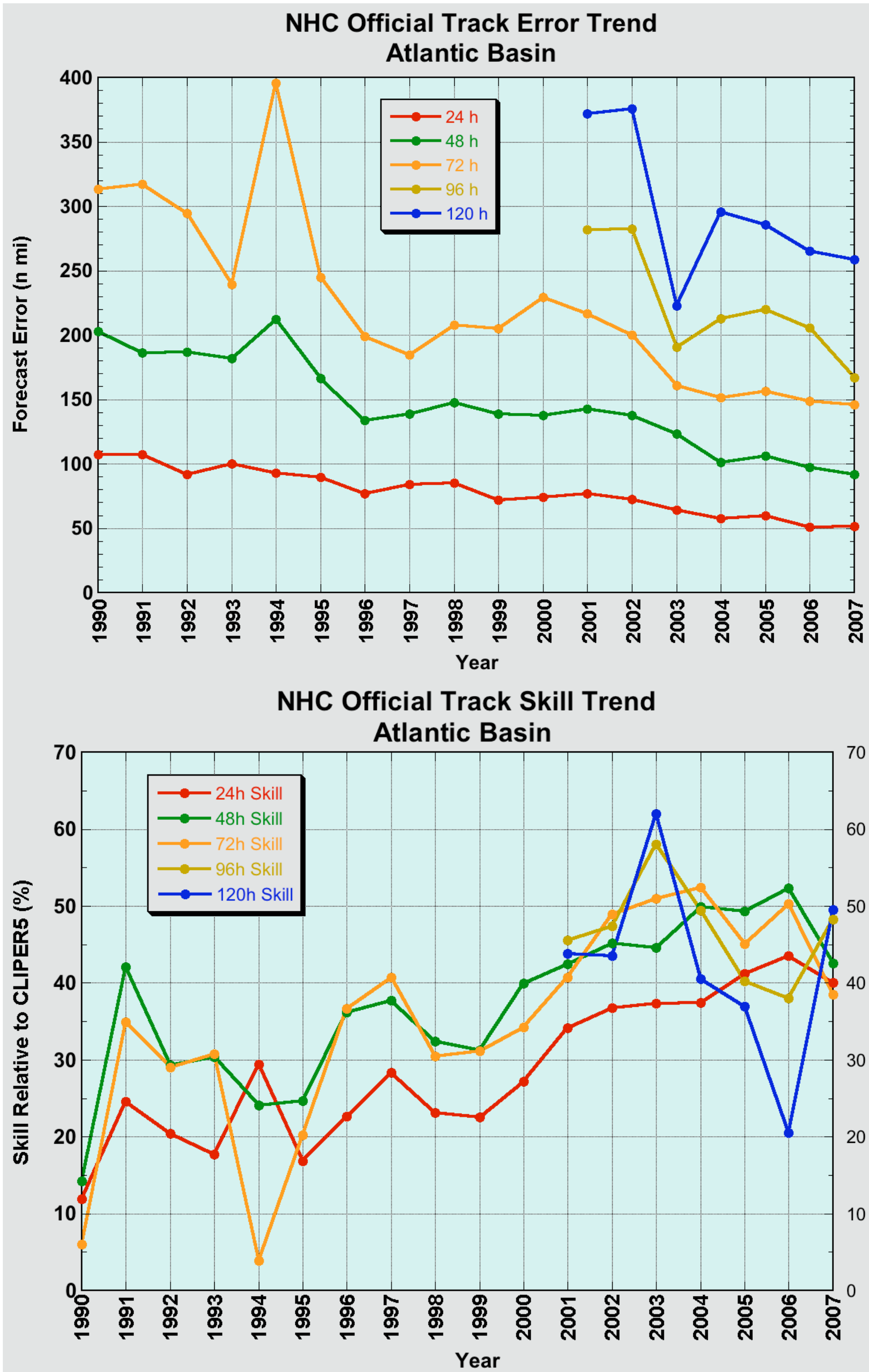


Figure 2. Recent trends in NHC official track forecast error (top) and skill (bottom) for the Atlantic basin.

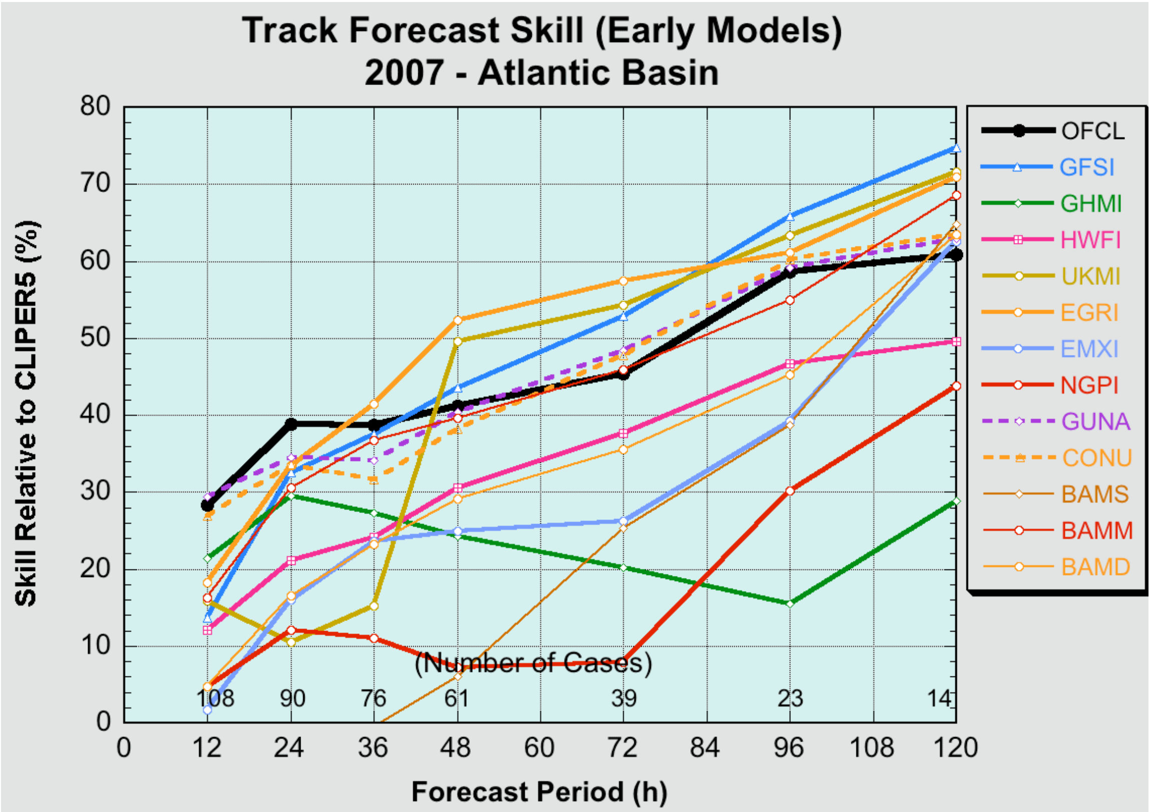


Figure. 3. Homogenous comparison for selected Atlantic basin early track guidance models for 2007. This verification includes only those models that were available at least 2/3 of the time (see text).

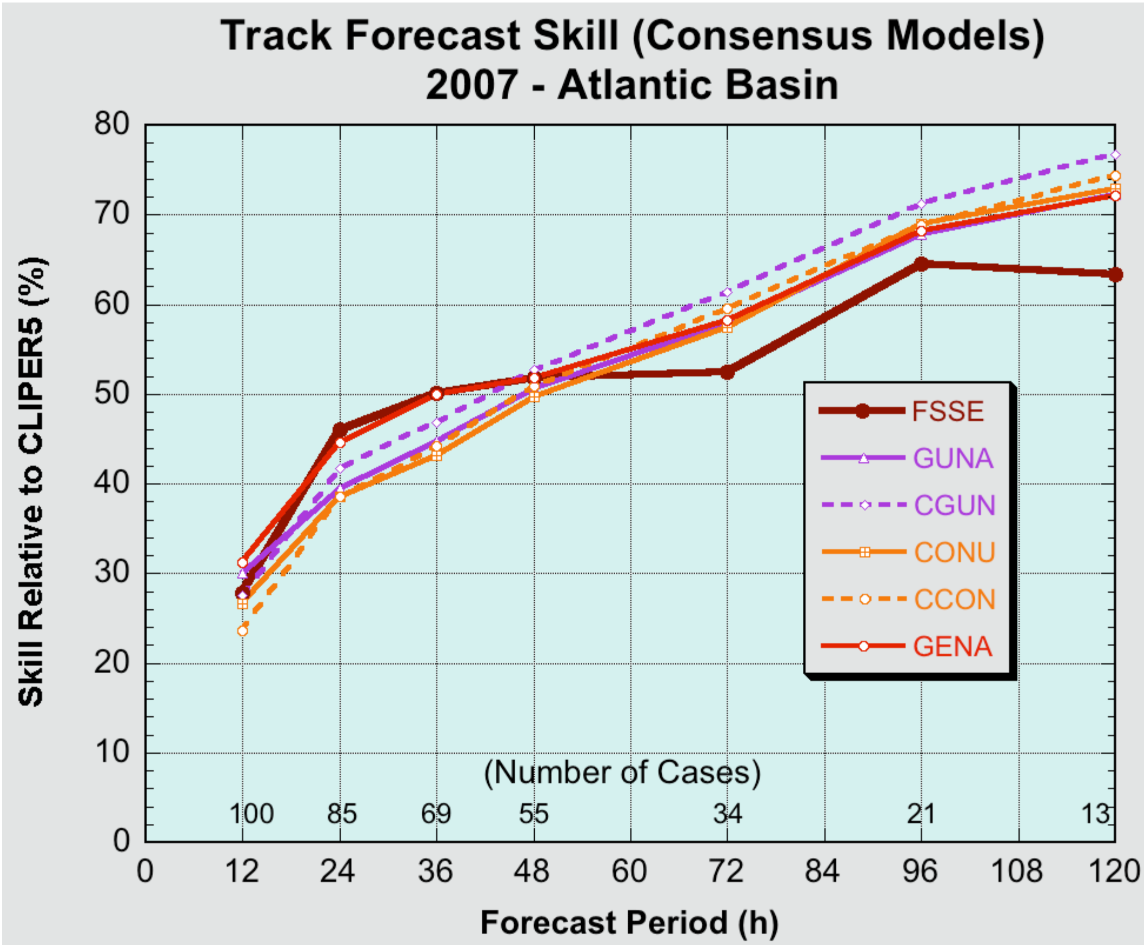


Figure 4. Homogenous comparison of the primary Atlantic basin track consensus models for 2007.

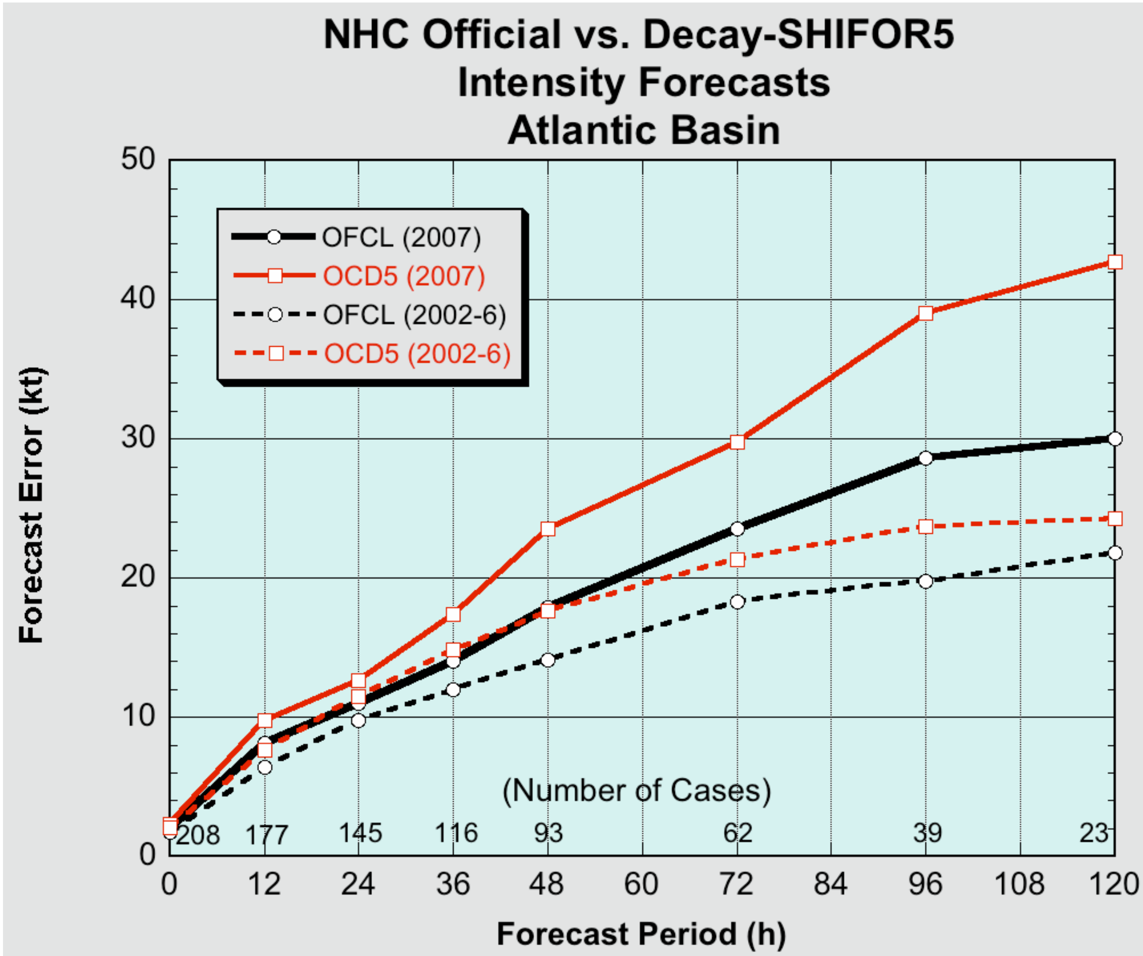


Figure 5. NHC official and Decay-SHIFOR5 (OCD5) Atlantic basin average intensity errors for 2007 (solid lines) and 2002-2006 (dashed lines).

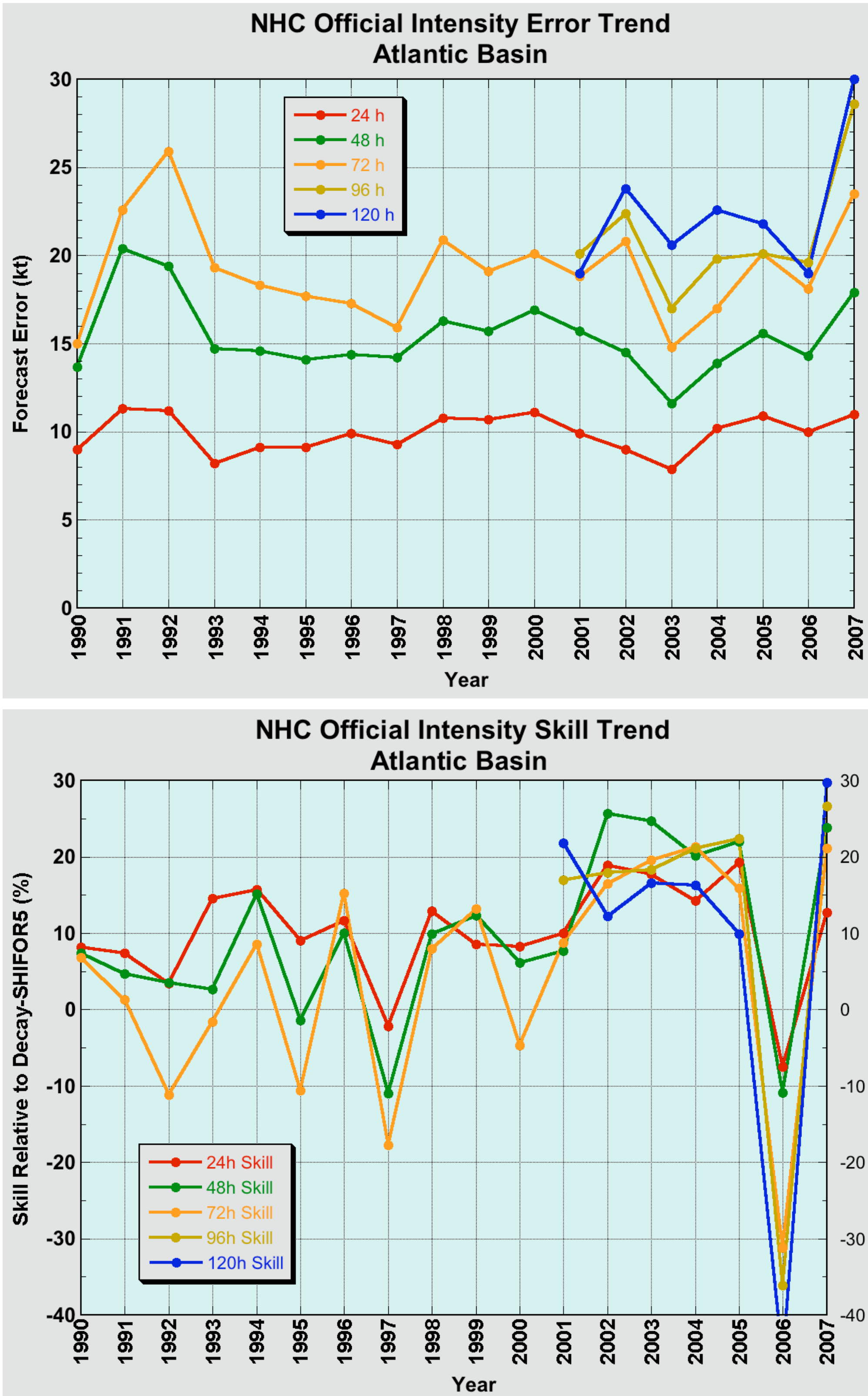


Figure 6. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the Atlantic basin.

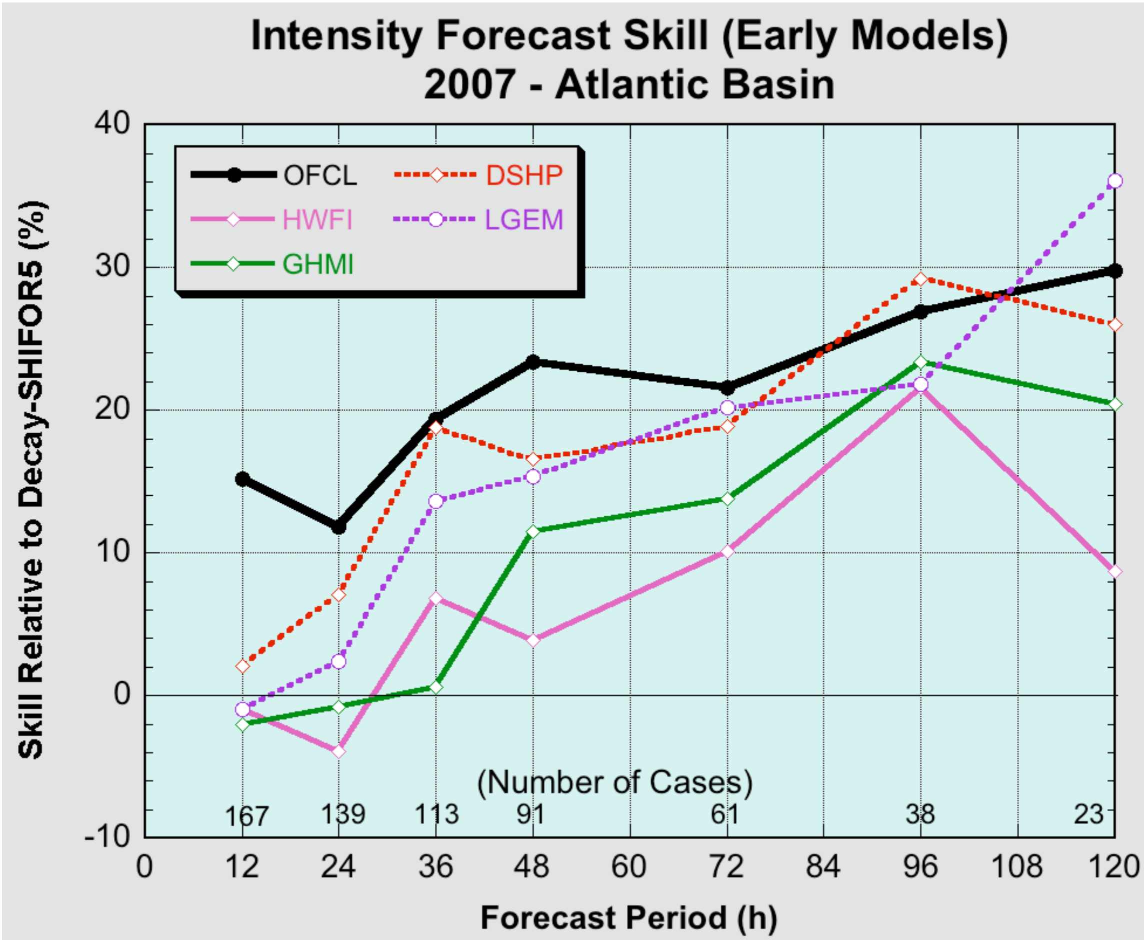


Figure. 7. Homogenous comparison for selected Atlantic basin early intensity guidance models for 2007.

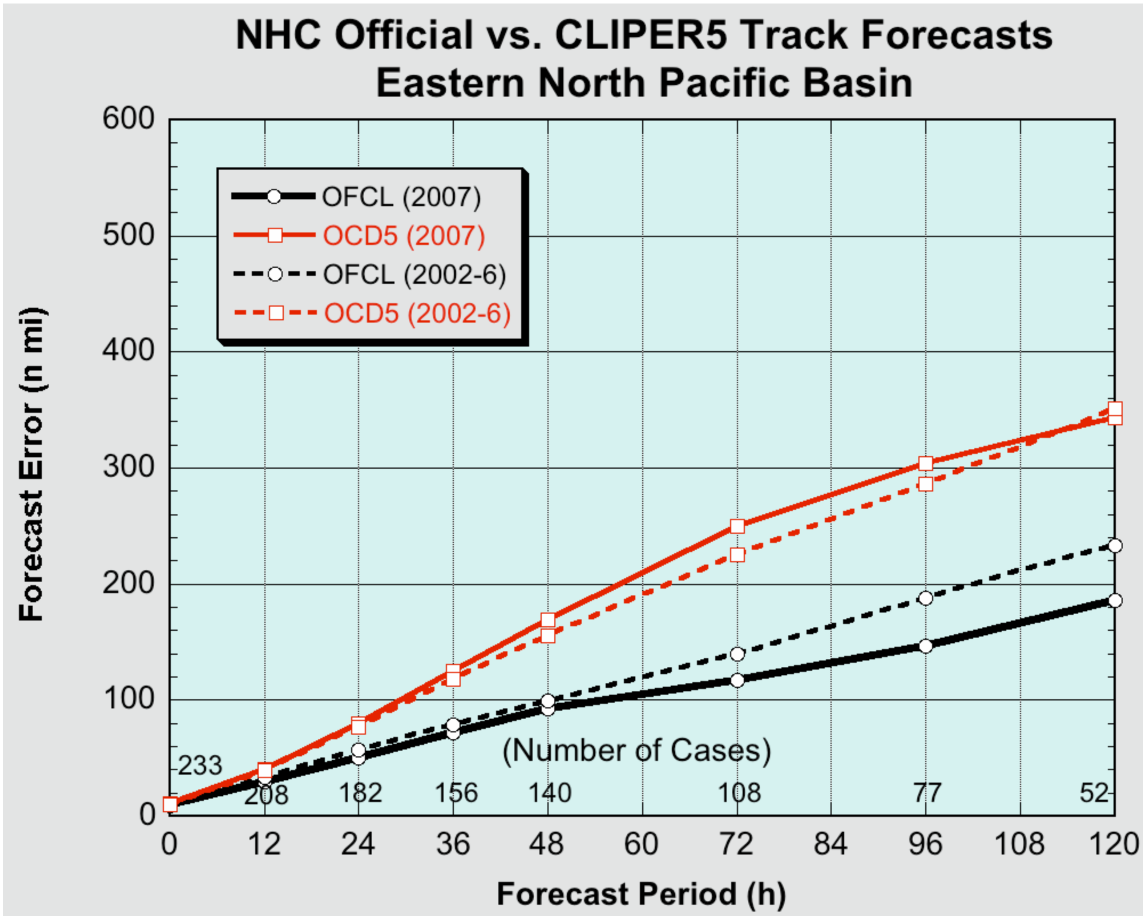
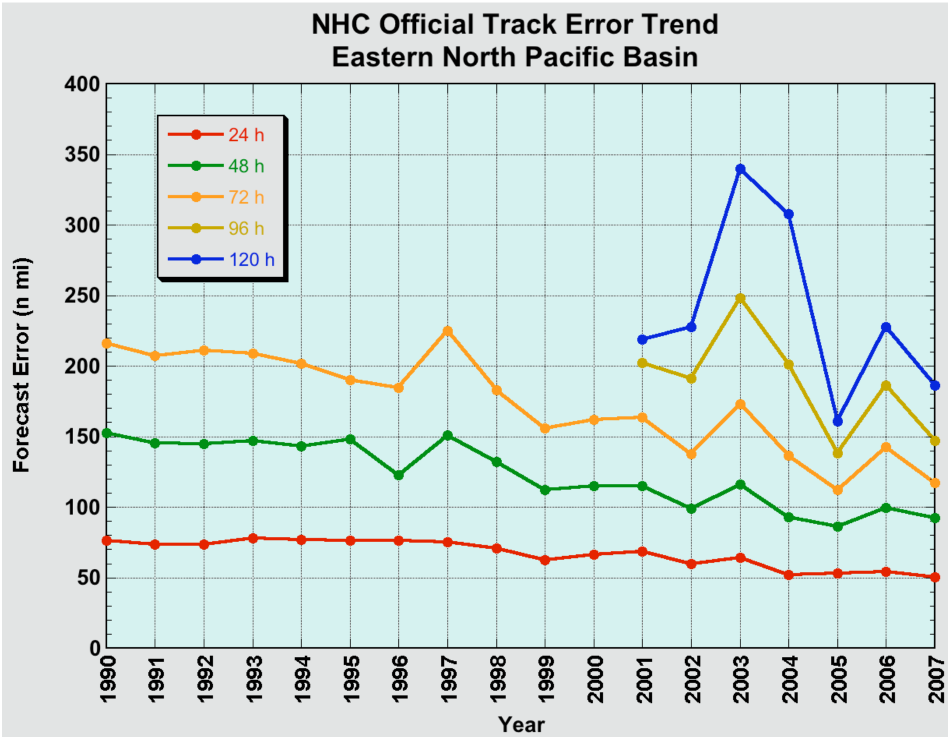


Figure 8. NHC official and CLIPER5 (OCD5) eastern North Pacific basin average track errors for 2007 (solid lines) and 2002-2006 (dashed lines).



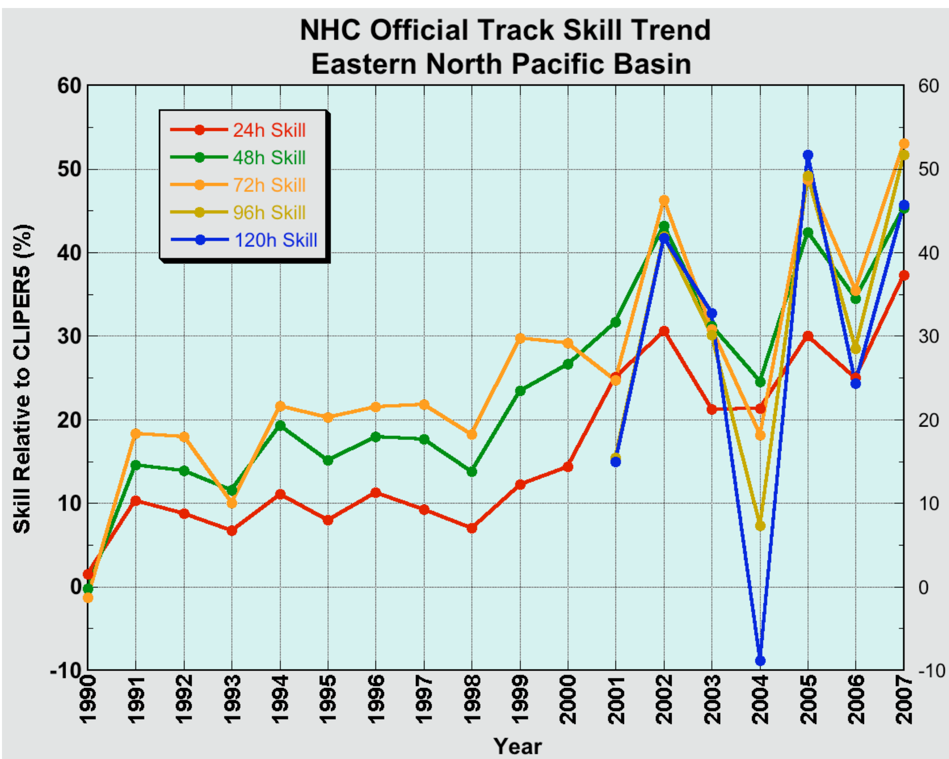


Figure 9. Recent trends in NHC official track forecast error (top) and skill (bottom) for the eastern North Pacific basin.

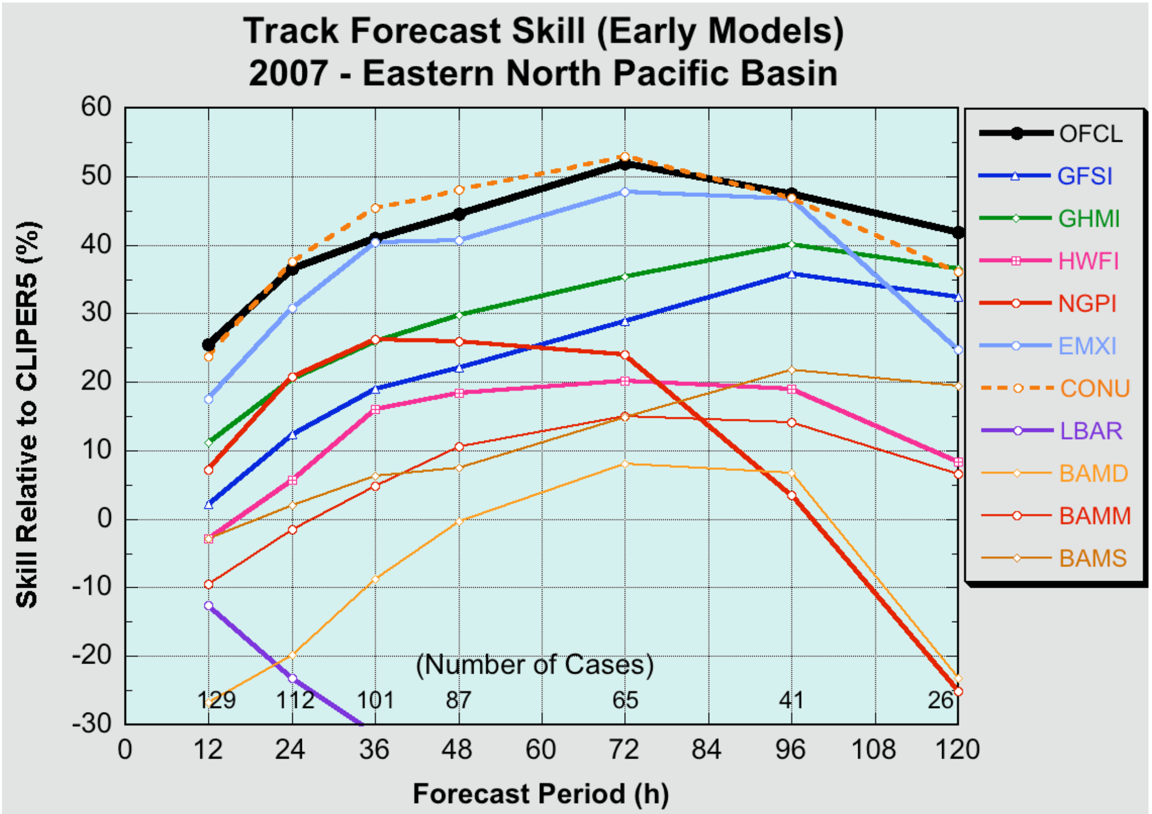


Figure. 10. Homogenous comparison for selected eastern North Pacific early track models for 2007. This verification includes only those models that were available at least 2/3 of the time (see text).

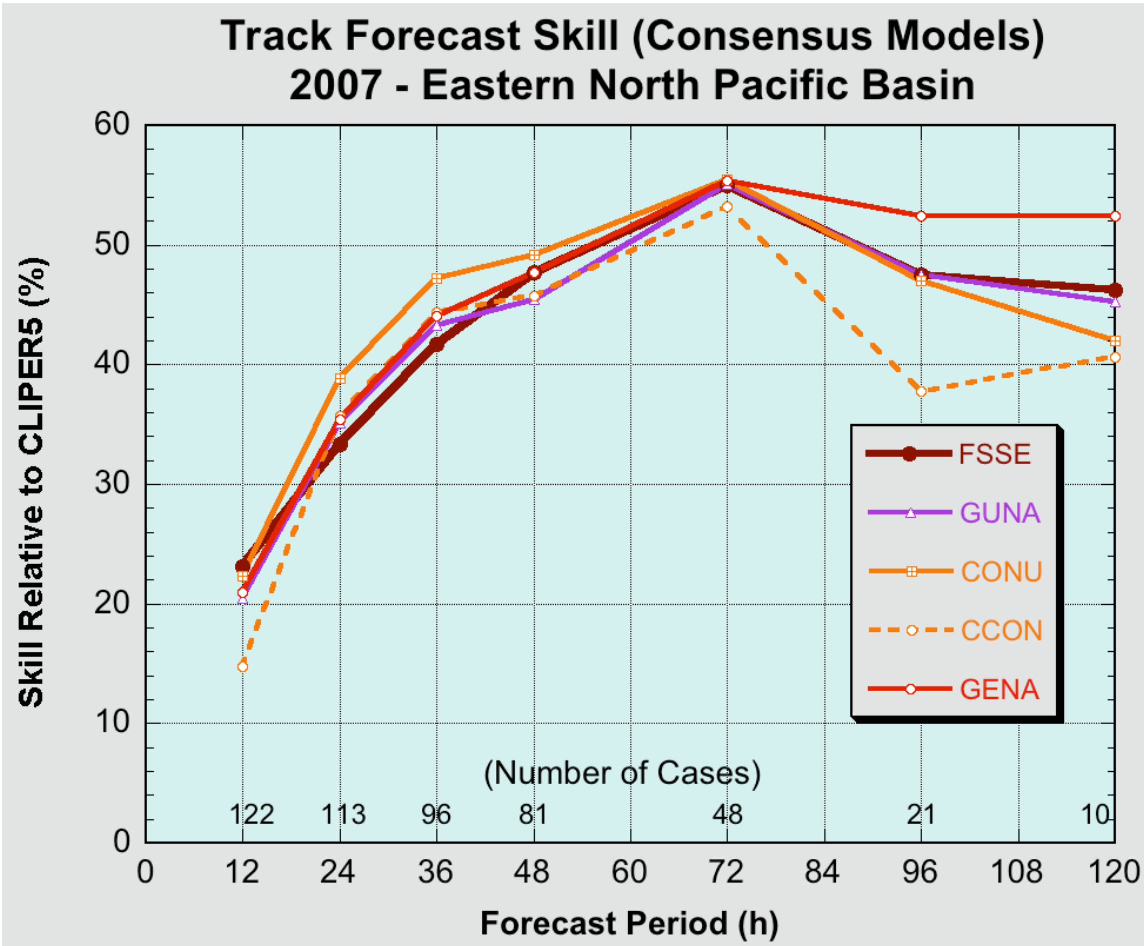


Figure 11. Homogenous comparison of the primary eastern North Pacific basin track consensus models for 2007.

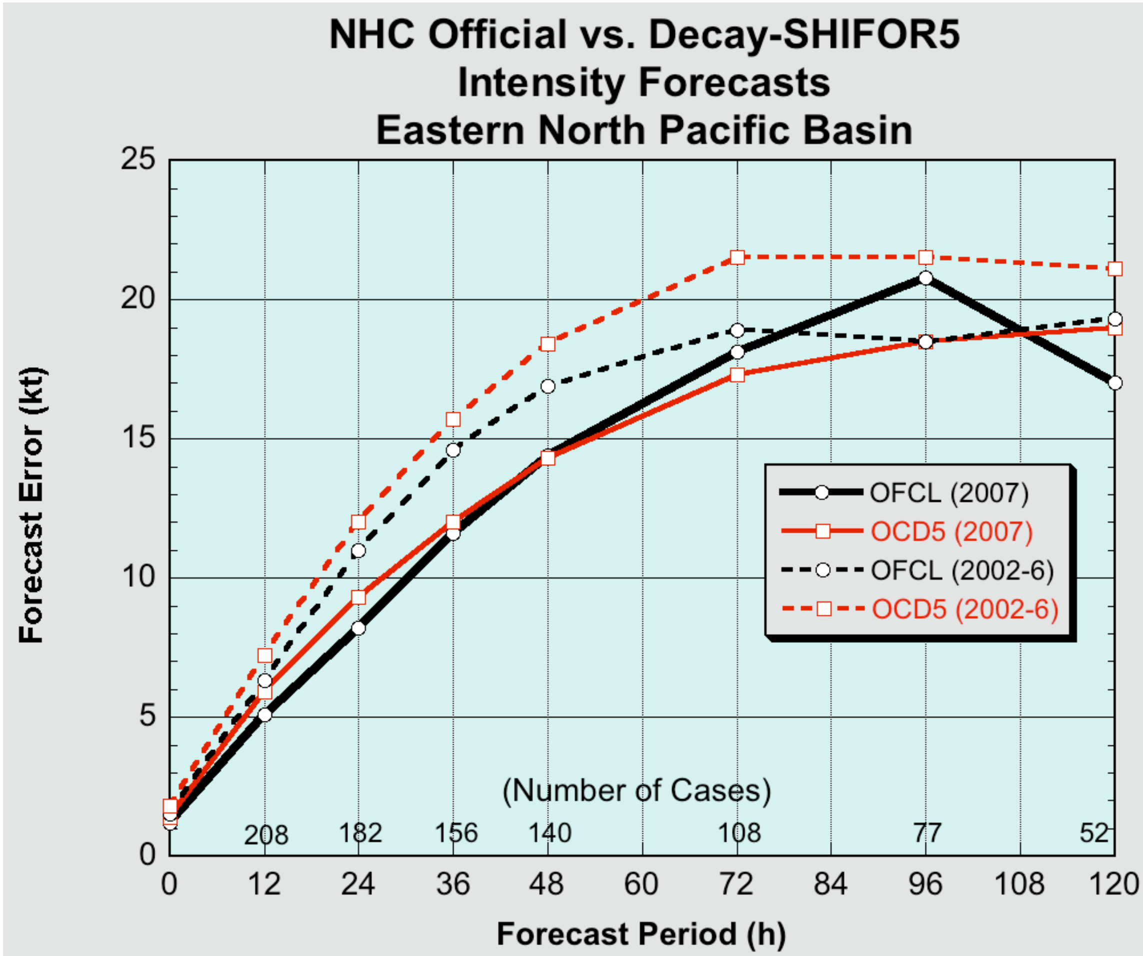
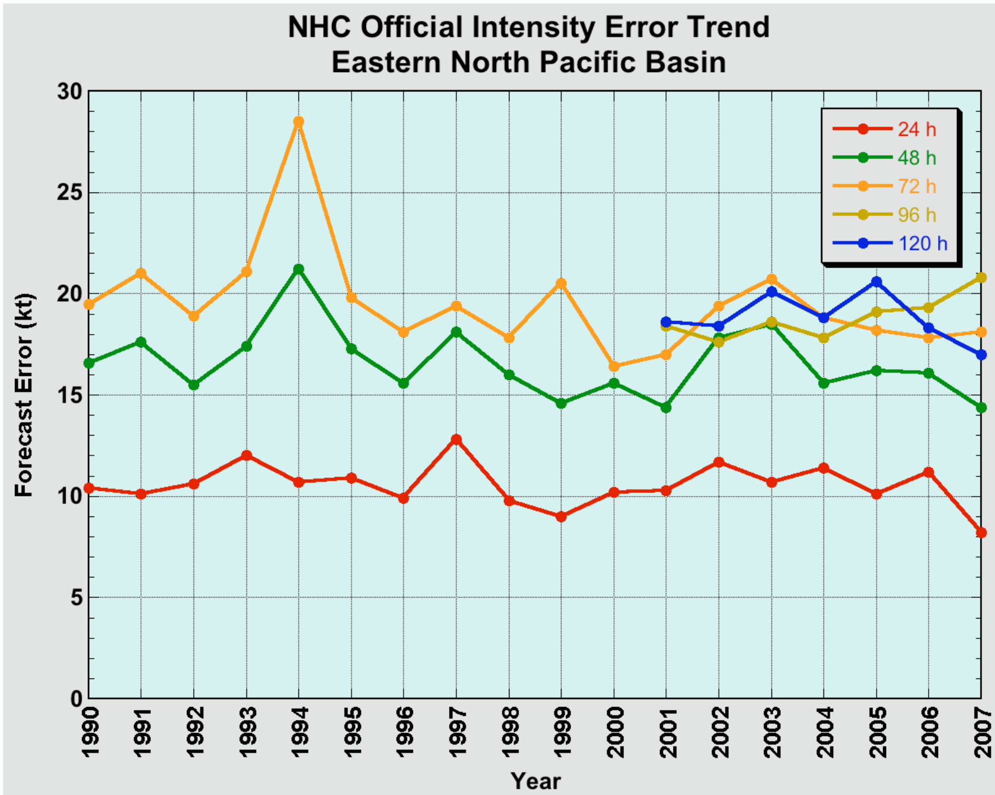


Figure 12. NHC official and Decay-SHIFOR5 (OCD5) eastern North Pacific basin average intensity errors for 2007 (solid lines) and 2002-2006 (dashed lines).



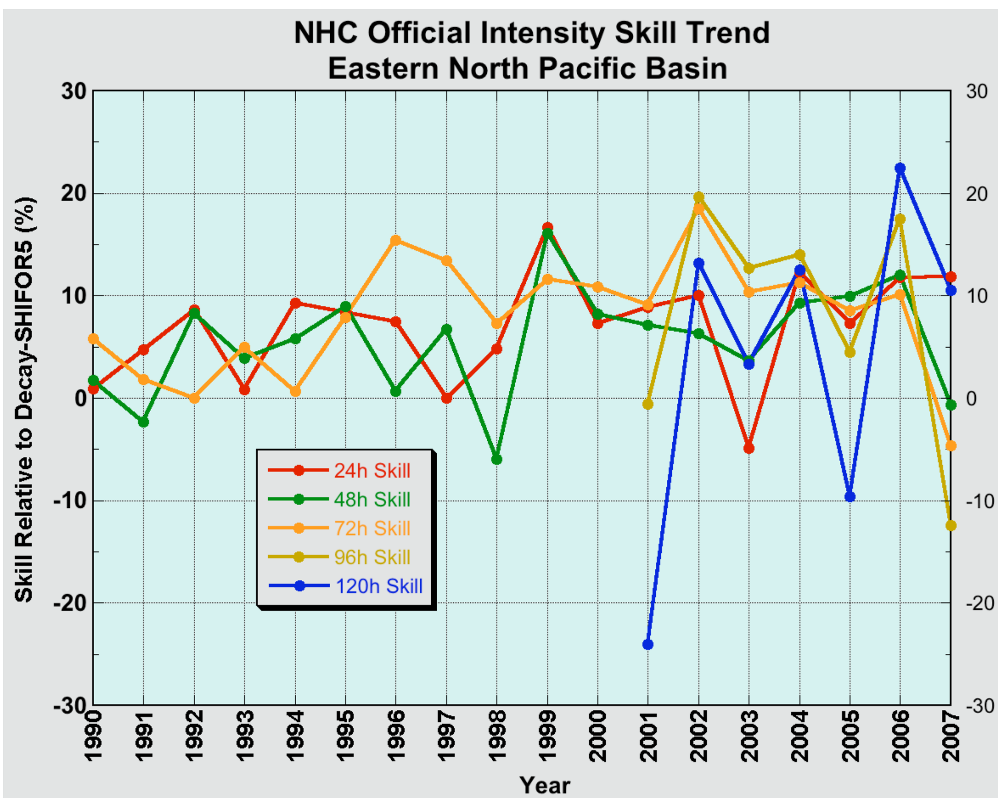


Figure 13. Recent trends in NHC official intensity forecast error (top) and skill (bottom) for the eastern North Pacific basin.

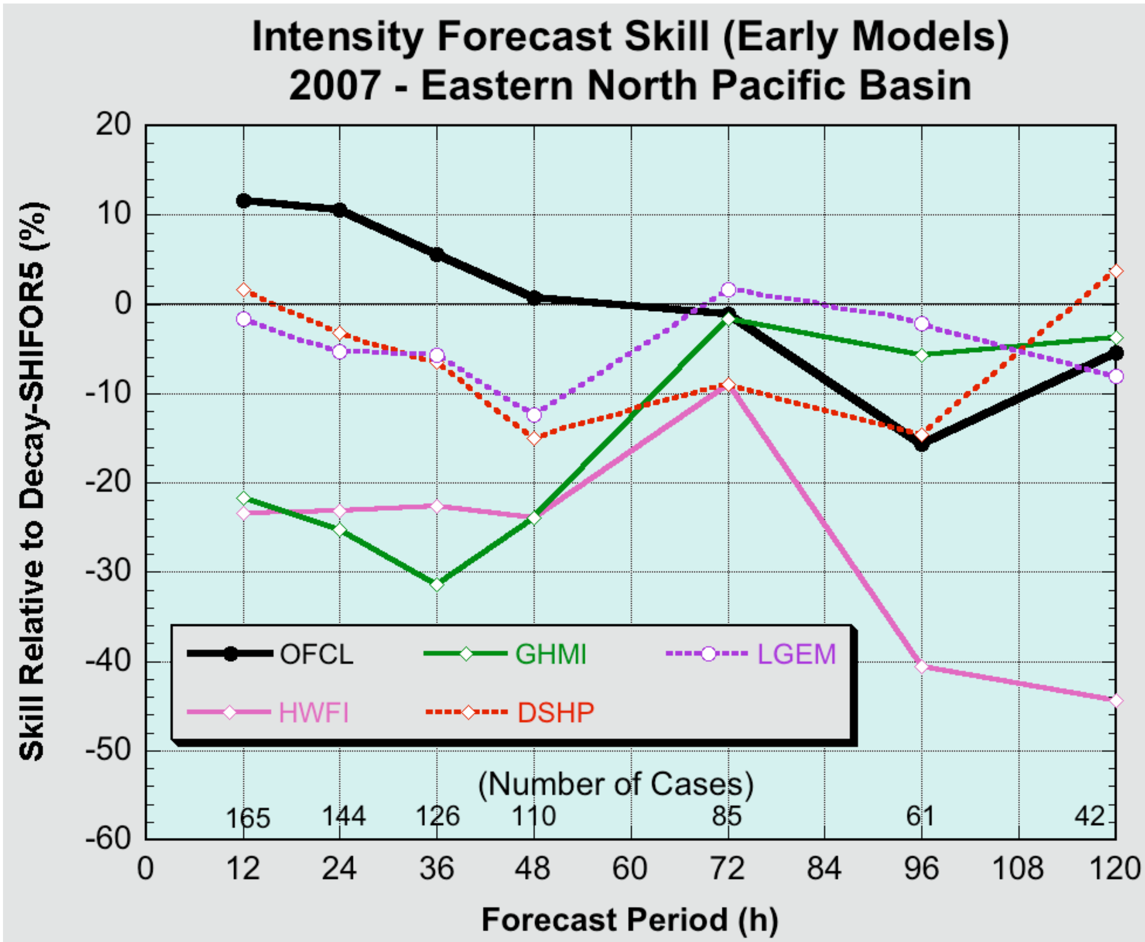


Figure. 14. Homogenous comparison for selected eastern North Pacific basin early intensity guidance models for 2007.