2005-09

Accuracy of Atlantic and Eastern North Pacific tropical cyclone intensity guidance

Lambert, Tara Denise Barton

Monterey, California. Naval Postgraduate School

---

http://hdl.handle.net/10945/1958

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ACCURACY OF ATLANTIC AND EASTERN NORTH PACIFIC TROPICAL CYCLONE INTENSITY GUIDANCE

by

Tara Denise Barton Lambert

September 2005

Thesis Advisor: Russell L. Elsberry
Second Reader: Mark A. Boothe

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2005	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE: Accuracy of Atlantic and Eastern North Pacific Tropical Cyclone Intensity Guidance	**5. FUNDING NUMBERS**
6. AUTHOR(S) Lambert, Tara D. B.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Five statistical and dynamical tropical cyclone intensity guidance techniques available at the National Hurricane Center during the 2003 and 2004 Atlantic and eastern North Pacific seasons were evaluated within three intensity phases: (i) formation; (ii) early intensification; and (iii) decay.

During the formation phase, the Decay Statistical Hurricane Intensity Prediction (DSHIPS) technique was the best technique in both basins. When the forecast errors during formation exceed +/- 10 kt, the statistical techniques tend to over-forecast and the dynamical models tend to under-forecast. Whereas DSHIPS was also the best technique in the Atlantic during the early intensification stage, the Geophysical Fluid Dynamics Laboratory model was the best in the eastern North Pacific. All techniques under-forecast periods of rapid intensification and the peak intensity, and have an overall poor performance during decay-reintensification cycles in both basins. Whereas the DSHIPS was the best technique in the Atlantic during decay, none of the techniques excelled during the decay phase in the eastern North Pacific. All techniques tend to decay the tropical cyclones in both basins too slowly, except that the DSHIPS performed well (13 of 15) during rapid decay events in the Atlantic. Similar error characteristics had been found in the western North Pacific.

14. SUBJECT TERMS Tropical Meteorology; Tropical Cyclone Intensity; Tropical Cyclone Intensity Change Forecast Techniques; Numerical and Statistical Intensity Change Guidance	**15. NUMBER OF PAGES** 141
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

ACCURACY OF ATLANTIC AND EASTERN NORTH PACIFIC TROPICAL CYCLONE INTENSITY GUIDANCE

Tara D. Lambert
Lieutenant, United States Navy
B.S., Jacksonville University, 1999

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY AND PHYSICAL OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
September 2005

Author: Tara Denise Barton Lambert

Approved by: Professor Russell L. Elsberry
Thesis Advisor

Mark Boothe
Second Reader

Professor Philip A. Durkee
Chairman, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Five statistical and dynamical tropical cyclone intensity guidance techniques available at the National Hurricane Center during the 2003 and 2004 Atlantic and eastern North Pacific seasons were evaluated within three intensity phases: (i) formation; (ii) early intensification; and (iii) decay.

During the formation phase, the Decay Statistical Hurricane Intensity Prediction (DSHIPS) technique was the best technique in both basins. When the forecast errors during formation exceed +/- 10 kt, the statistical techniques tend to over-forecast and the dynamical models tend to under-forecast. Whereas DSHIPS was also the best technique in the Atlantic during the early intensification stage, the Geophysical Fluid Dynamics Laboratory model was the best in the eastern North Pacific. All techniques under-forecast periods of rapid intensification and the peak intensity, and have an overall poor performance during decay-reintensification cycles in both basins. Whereas the DSHIPS was the best technique in the Atlantic during decay, none of the techniques excelled during the decay phase in the eastern North Pacific. All techniques tend to decay the tropical cyclones in both basins too slowly, except that the DSHIPS performed well (13 of 15) during rapid decay events in the Atlantic. Similar error characteristics had been found in the western North Pacific.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank Professor Russell Elsberry for his guidance, dedication, professionalism, and patience with me and this thesis. It has been an honor and a pleasure to work with him and learn from him. I would like to thank Mary Jordan for her guidance during this thesis and during my time at the Naval Postgraduate School. She always took the time to help me no matter the day or the hour. I would also like to thank LCDR David Brown for his help with my thesis, thesis presentation, and his mentorship. I would also like to acknowledge two of my friends - Mark Butler and Elizabeth Scheidecker. Their friendship and laughter always kept a smile on my face even through the difficult days.

Most importantly, I would like to express my sincere gratitude and appreciation to my husband, Bobby Lambert. He has stood by me for 14 years of marriage and throughout my naval career. Thank you for being my rock. My two children, Bobbi and Taran – I work, pray, and live for the two of you. Finally, to my mother who has always been the wind beneath my wings.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. MOTIVATION

1. Necessity for Accurate Intensity Forecast

The formation and intensification of a tropical cyclone (TC) in the Atlantic Ocean can quickly become an immediate threat to the vast population in the Islands of the Caribbean, Central America, the Gulf of Mexico, and the east coast of the United States. An intensity forecast determines whether to evacuate an entire Gulf Coast region or sortie (relocate) U.S. military vessels and aircraft. An over-forecast storm intensity may cost cities and states millions of dollars for an unnecessary evacuation and disaster planning/preparation. More importantly, an under-forecast intensity may cause loss of life and contribute to property damages in the billions of dollars and force states to declare a state of emergency.

The Department of Defense (DOD), and especially the U.S. Navy, is affected by the accuracy of intensity forecasts. The Navy has over 20 naval air stations, submarine bases, and weapon stations along the eastern and Gulf coasts of the U.S. including Norfolk, VA, which is the largest naval base in the world. Navy vessels can better avoid a storm of hurricane strength when they are out to sea, and yet maintain the capability to respond to any national crisis. A greater potential for damage from high winds and storm surges exists when a ship is moored and the ship is battered against the pier. It costs the Navy millions of dollars to sortie (relocate) its ships, evacuate its airplanes/helicopters, and to secure vast amounts of equipment before the onset of destructive winds (magnitude may be different for each military base). For example, it cost $105 million when the Navy sortied approximately 40 vessels and 150 planes from the Norfolk area because of Hurricane Isabel in September 2003. None of the Navy ships or planes was damaged by the storm. An accurate intensity forecast can save lives and millions of dollars.

2. Intensity Guidance

The tendency in hurricane (typhoon) forecasting is that operational hurricane forecast models are more skillful in predicting the TC track than the intensity. The track errors have been cut in half in the Atlantic and by one-third in the eastern North Pacific in the last 15 years (1990 to 2004), but the official intensity errors have been reduced little in either basin (Franklin 2005). The Atlantic and eastern North Pacific forecast models are no exception as DeMaria et al. (2002) indicated that the official National Hurricane Center (NHC) intensity forecasts are relatively much less skillful compared to climatology and persistence than the official NHC track forecasts. More recently, Franklin (2005) summarized the 2004 Atlantic hurricane season as having a downward trend in official track errors through 72 h, while the official intensity errors continued to show little improvement. Franklin's (2005) results for the eastern North Pacific were similar to the Atlantic with a modest improvement trend in the official track errors while the official intensity errors show little improvement.

The improvement in track forecast skill is primarily due to the gain in accuracy of the dynamical TC track forecast guidance and the application of consensus forecasting techniques (Elsberry 2004). The improved skills of the dynamical models led to the development of the Systematic Approach Forecast Aid (SAFA) by Carr and Elsberry (1994) that enables the forecaster to improve on the accuracy and consistency of the dynamical models and provide an official track forecast that is meteorologically sound. The Joint Typhoon Warning Center (JTWC) began using SAFA in 2000, which has helped JTWC to three straight record track forecasting seasons from 2000 to 2002 (Jeffries and Fukada webpage reference, cited 2005). However, the large errors of the intensity forecast models do not allow forecasters to apply the same systematic approach used with the dynamical track forecast guidance.

3. Example of Intensity Model Errors

Reducing the large errors in the intensity models is the first step in the process of trying to implement a systematic approach to intensity forecasting. Figure 1.1 is an example of intensity forecast errors produced by the NHC with Hurricane Isis (TC12-E) in the eastern North Pacific during September 2004. One of the primary reasons why the NHC official (OFCL) forecast errors were large in this case is likely the faulty guidance (Figure 1.2) provided to the NHC forecasters by the Statistical Hurricane Prediction Scheme (SHIPS) intensity forecast technique. The early version of SHIPS demonstrated some skill relative to climatology and persistence in both the eastern North Pacific and Atlantic basins since 1997 (DeMaria et al. 2002), and is now the most skillful intensity guidance. However, the intensity forecasts in Figure 1.2 indicate that the SHIPS guidance is not always good, and it would be useful to the forecasters to know when it is good or not good.

The disturbance that developed into Hurricane Isis was initially a tropical wave that entered the eastern North Pacific basin on 3 September 2004. The disturbance was designated by the NHC as a tropical disturbance at 0600 UTC 8 September when it was approximately 460 n mi south of Cabo San Lucas, Mexico. By 1800 UTC 8 September, the system strengthened into a tropical storm, but was downgraded back to a tropical depression two days later as the system moved under easterly vertical wind shear. Isis re-strengthened to a tropical storm on 12 September and maintained winds of around 45 kt until 1200 UTC 14 September when it underwent a period of decay down to 35 kt followed by rapid intensification to hurricane strength at 1200 UTC 15 September. Isis then underwent rapid decay and became a tropical depression within 24 h of becoming a hurricane as it turned northwestward into a region with cooler ocean water and a stable low-level environment and then became stationary. By 1800 UTC 16 September, Isis was a remnant low located 1300 n mi west of Cabo San Lucas, and it dissipated on 21 September. Note that Hurricane Isis maintained a mostly westward track during its eight-day lifespan and its closest approach to land was 460 n mi while it briefly achieved hurricane status.

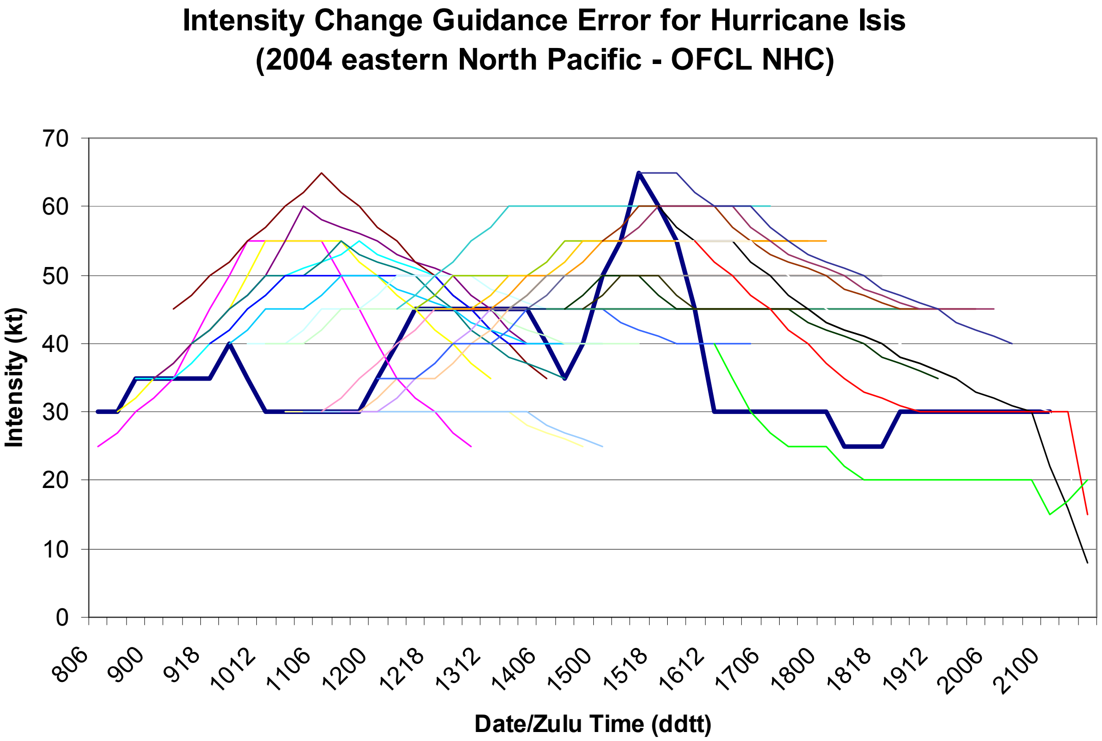


Figure 1.1 National Hurricane Center official intensity guidance (OFCL) for Hurricane Isis (Storm #12) during 2004. The heavy blue line represents the observed (best track) storm intensity. Colored lines illustrate intensity forecasts each 6 h. These errors may be described as: early over-intensification; missed decay; missed secondary decay; missed 'rapid' intensification; and missed 'rapid' decay.

Because of the fairly simple track entirely over water and the moderate strength of Isis, one might have expected that the intensity forecast guidance should be accurate. However, the SHIPS technique (Figure 1.2) consistently over-intensifies the storm for the first 36 hours with a forecast of near-hurricane strength (62 kt) when Isis actually verified as a tropical depression (30 kt). Thus, the SHIPS technique misses an early period of decay from tropical storm to tropical depression. The SHIPS technique did provide useful guidance as to a short period of intensification from 30 kt to 45 kt, but these forecasts continued intensification to about 60 kt, whereas the storm actually had constant intensity and even a short period of decay to 35 kt. Then the SHIPS technique missed the rapid intensification to a hurricane in 24 h, and then the rapid decay to a tropical depression.

The different intensity errors throughout the life stages of this TC are an example of the lack of skillful guidance given to the forecasters. It demonstrates a necessity for more skillful intensity guidance. That is, a consensus approach for intensity guidance can only be accomplished when multiple skillful intensity forecast models are available. It is at this point when short-term 'corrective' efforts such as use of this guidance should be considered.

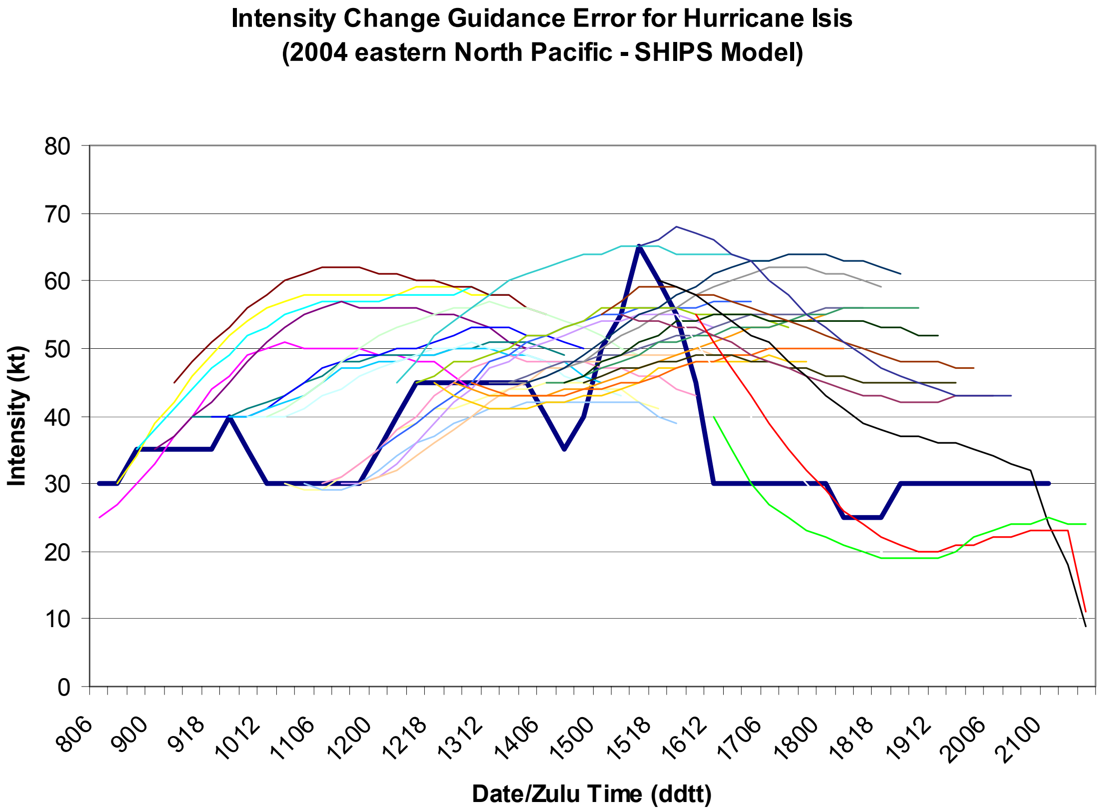

Intensity Change Guidance Error for Hurricane Isis
(2004 eastern North Pacific - SHIPS Model)

Figure 1.2 As in Figure 1.1, except for Statistical Hurricane Intensity Prediction Scheme (SHIPS) guidance for Hurricane Isis (Storm #12) during 2004.

B. CHAPTER OVERVIEW

The objective of this thesis is to evaluate TC intensity guidance for the Atlantic and eastern North Pacific tropical cyclones during 2003 and 2004 using a conceptual intensity model for different stages of a TC life cycle. Chapter II explores causes for the intensification, re-intensification, and decay of a TC and contains a summary of the 2003 and 2004 Atlantic and eastern North Pacific tropical seasons. The idealized intensity and decay cycle of a TC, and the various TC intensity guidance techniques are also introduced in Chapter II. The intensity guidance evaluations are described in Chapter III. Chapter IV summarizes the key findings of this study and describes the potential strengths and weaknesses of the intensity change techniques that will provide forecasters with additional knowledge on how to interpret intensity guidance.

II. FRAMEWORK

A. INTENSITY PHASES OF A TROPICAL CYCLONE

The conceptual model used to analyze the intensity guidance models is based on the intensity changes during the life cycle of the TC (Figure 2.1) used by Blackerby (2005). The life cycle model is divided into three intensity phases: (i) formation stage to tropical storm (34 kt); (ii) early intensification through tropical storm (intensification phase); and (iii) the decay phase. Each intensity forecast technique (described in section D) is evaluated to determine its accuracy during each intensity phase to provide guidance to the forecasters on when the technique is likely to be (not be) accurate.

The formation stage of the TC is defined as a tropical weather system with organized convection (pre-TC or tropical disturbance) to the formation of a closed circulation with maximum surface sustained winds (using the U.S. 1-minute average standard) of 33 kt or less. Several of the intensity forecast techniques do not provide guidance during the formation stage as the storm is usually too weak for the models to discern. Therefore, this study primarily examines the intensity changes following TC formation.

In Phase II, the intensification change can either be rapid, typical, or slow due to the TC interaction with the ocean, its inner-core processes, or environmental interactions, but these interactions may not always cause the TC to intensify. Warmer than normal tropical ocean waters in the mixed layer can contribute to a positive TC intensity change. However, many authors (e. g., Bender and Ginis 2000) have shown that upwelling and vertical mixing of the cool seasonal thermocline water by the TC vortex will produce a negative feedback between the ocean and the atmosphere that tends to cause the TC to decay. As noted by Willoughby et al. (1982), concentric eyewalls develop a secondary ring of convection around an existing inner eyewall, and if the secondary ring contracts and replaces the inner eyewall, a marked change in TC intensity may occur. The collapse of the inner eyewall can result in an initial decay in the TC while further contraction of the outer eye can produce rapid intensification. Such TC intensity changes are not exclusive to Phase II, but may also occur in Phase III as a TC can reintensify and then decay, which makes the conceptual model in Figure 2.1 too simple. Therefore, a more realistic conceptual model (Figure 2.2) is introduced that incorporates the possibility of such TC intensity changes.

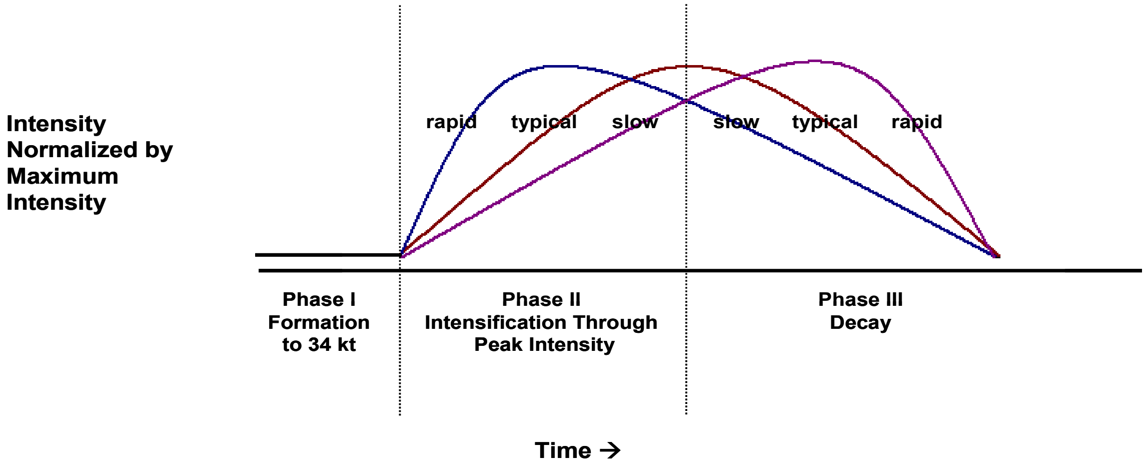


Figure 2.1 Idealized intensity traces showing intensification and decay profiles following the formation phase (Phase I). Phase II intensification may be described as slow (modest increase in intensity for a given time period), typical (an average rate of intensification for a given ocean basin and time period), or rapid (an above-average rate of intensification for a given time period). Phase III may be described in a similar manner: slow, typical, and rapid rates of decay (from Blackerby 2005).

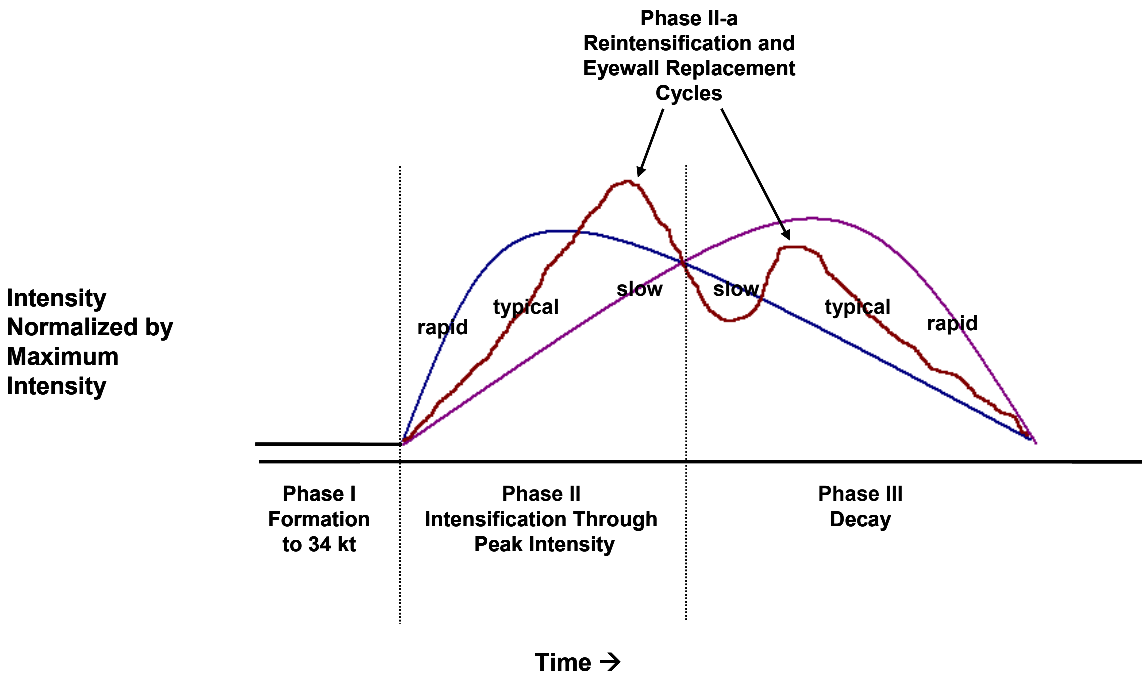


Figure 2.2 Similar to Figure 2.1, except with the addition of Phase IIa to represent reintensification following an initial period of decay, especially during eyewall replacement cycles observed in intense TCs (from Blackerby 2005).

B. PREVIOUS WORK

Blackerby (2005) analyzed the western North Pacific intensity guidance techniques using the same conceptual models in Figures 2.1 and 2.2 for the 2003 and 2004 tropical cyclone seasons. He evaluated the performance of six intensity guidance techniques available to forecasters at the Joint Typhoon Warning Center (JTWC) and also the official intensity forecasts. The six intensity techniques evaluated by Blackerby (2005) are as follows: Statistical 5-Day Typhoon Intensity Forecast—5-day Model (ST5D); Statistical Typhoon Intensity Prediction Scheme (STIPS); Geophysical Fluid Dynamics Model–Navy (interpolated) (GFNI); Air Force Weather Agency Mesoscale Meteorological Model-5 (interpolated) (AFWI); Coupled Hurricane–Ocean Intensity Prediction System (CHIPS); and Japan Typhoon Model (interpolated) (JTYI). A total of 59 storms were used for the intensity analysis.

The results of his evaluation suggest that intensity change guidance in the western North Pacific do not skillfully predict future intensity changes. Blackerby (2005) concluded that all of the various intensity guidance techniques and the JTWC (on average) under-forecast decay rates during Phase III, many did not predict intensity oscillations during Phase IIa, nor did they capture periods of rapid intensification and decay. A summary of the basic results are listed below in Table 2.1 from the Blackerby (2005) thesis. In the evaluation (see Chapter 3) for the Atlantic and eastern North Pacific intensity guidance, a more detailed comparison of analogous intensity prediction techniques will be made with these western North Pacific evaluations for various phases of the life cycle.

Table 2.1 Summary of Phase I through Phase III model forecasts. If forecasters anticipate a certain intensification/decay rate, then the Xs in each column indicate which model is most likely to produce at that forecast (from Blackerby 2005).

24h Reasoning Indicates…	**Rate (kt)**	**ST5D**	**STIP**	**GFNI**	**AFWI**	**CHIP**	**JTYI**
Nearly Steady Intensity (Esp. TD in Phase I)	**0 to 5**				**X**	**X**	
Slow Intensification Rate (Esp. Moving WNW)	**5 to 10**				**X**		**X**
Average Intensification Rate (Esp. Moving WNW)	**10 to 20**	**X**	**X**				
Rapid Intensification Period (Esp. Just After Formation)	**30+**			**X**		**X**	
Intensity Greater Than 110 kt	**Any**		**X**	**X**		**X**	
Intensity Oscillations (Phase IIa)	**Any**			**X**			
Slow Decay Rate (Esp. early in Phase III)	**5 to 10**			**X**	**X**		**X**
Average Decay Rate	**10 to 20**	**X**	**X**				
Decay Past STR Axis	**Any**	**X**	**X**			**X**	
Rapid Decay at Landfall	**30+**					**X**	**X**
Rapid Decay Moving Poleward	**30+**	**X**				**X**	

C. HYPOTHESIS

The hypothesis for this study is that similar techniques and numerical models used for intensity forecasting in the Atlantic and eastern North Pacific will have a similar performance of the various intensity guidance models that was found in the western North Pacific by Blackerby (2005). Some of the techniques used in these basins are identical, such as GFNI, or are quite similar but are named differently (interchanging typhoon and hurricane) such as STIPS becoming SHIPS (see acronym definitions below) and ST5D becoming SHF5, and have the same model physics and design. Therefore, the techniques should have similar strengths and deficiencies and perform the same in each of the three basins.

D. DATA SOURCE

The data for this study are from the NHC Automated Tropical Cyclone Forecasting System (ATCF) (Sampson and Schrader 2000) files that include the “best track” data, intensity forecasts by the various techniques, and the NHC official intensity forecasts for the 2003 and 2004 Atlantic and eastern North Pacific tropical cyclone seasons. A total of 69 storms in 2003 and 2004 for both basins were used to apply the life cycle framework for intensity analysis to evaluate the performance of the intensity guidance models. The 2003 tropical cyclone season included 21 TCs in the Atlantic (Table 2.2) and 16 TCs in the eastern North Pacific (Table 2.3). For 2004, both the Atlantic and eastern North Pacific had 16 TCs (Table 2.4 and Table 2.5).

This thesis evaluates the performance of five intensity guidance techniques available directly to the forecasters at the NHC. The five intensity guidance models are as follows: Statistical Hurricane Intensity Forecast 5-day Model (SHF5); Statistical Hurricane Intensity Prediction Scheme (SHIPS); Decay Statistical Hurricane Intensity Prediction Scheme (DSHIPS); Geophysical Fluid Dynamics Model (interpolated) (GFDI); and Geophysical Fluid Dynamics Model–Navy (interpolated) (GFNI). This list is not all inclusive, e.g., the CHIPS model of Emanuel (1995a) is not included because it is only available via a website. However, these five intensity techniques are expected to have some skill during the various intensity phases of the life cycle model in Figure 2.2.

Each of the above intensity techniques may be classified as statistical, statistical-dynamical, or dynamic. The SHF5 is a statistical model based on climatological intensity change average and a persistence forecast that the current intensity trend will continue throughout the forecast period.

Table 2.2 Tropical cyclones in the Atlantic Ocean basin during 2003. A total of 21 tropical cyclones occurred: 5 tropical depressions, 9 tropical storms, and 7 hurricanes. The information was gathered from the NHC website http://www.nhc.noaa.gov

NAME	MONTH	WARNING DURATION (days)	EST. MAX. SFC WIND (kt)	LOWEST PRESSURE (mb)
TS Ana	APR	4	50	994
TD Two	JUN	¾	30	1008
TS Bill	JUN/JUL	3 ¼	50	997
Hurricane Claudette	JUL	8 ¼	80	979
Hurricane Danny	JUL	4 ½	65	1000
TD Six	JUL	1 ¾	30	1010
TD Seven	JUL	1 ½	30	1016
Hurricane Erika	AUG	2 ¼	65	986
TD Nine	AUG	¾	30	1007
Hurricane Fabian	AUG/SEP	11 ¾	125	939
TS Grace	AUG/SEP	2 ¾	35	1007
TS Henri	SEP	4 ¾	50	997
Hurricane Isabel	SEP	13 ¼	145	915
TD Fourteen	SEP	2 ¼	30	1007
Hurricane Juan	SEP	5	90	969
Hurricane Kate	SEP/OCT	12	110	952
TS Larry	OCT	4 ½	55	993
TS Mindy	OCT	3 ¼	40	1002
TS Nicholas	OCT	10 ¾	60	996
TS Odette	DEC	3	55	993
TS Peter	DEC	3 ¼	60	990

Table 2.3 Tropical cyclones in the eastern North Pacific Ocean basin during 2003. A total of 16 tropical cyclones occurred which included 9 tropical storms and 7 hurricanes. The information was gathered from the NHC website http://www.nhc.noaa.gov.

NAME	MONTH	WARNING DURATION (days)	EST. MAX. SFC WIND (kt)	LOWEST PRESSURE (mb)
TS Andres	MAY	5 ½	50	997
TS Blanca	JUN	5 ¼	50	997
TS Carlos	JUN	1 ½	55	996
TS Dolores	JUL	1 ¾	35	1005
TS Enrique	JUL	3 ¼	55	993
TS Felicia	JUL	5 ½	45	1000
TS Guillermo	AUG	5 ¼	50	997
TS Hilda	AUG	4 ¼	35	1004
Hurricane Ignacio	AUG	5	90	970
Hurricane Jimena	AUG/SEP	8 ½	90	970
TS Kevin	SEP	2 ½	35	1000
Hurricane Linda	SEP	4	65	987
Hurricane Marty	SEP	6	85	970
Hurricane Nora	OCT	7 ½	90	969
Hurricane Olaf	OCT	4 ¾	65	987
Hurricane Patricia	OCT	5 ½	70	984

Table 2.4 Tropical cyclones in the Atlantic Ocean basin during 2004. A total of 16 tropical cyclones occurred: 1 tropical depression, 1 subtropical storm, 5 tropical storms, and 9 hurricanes. The information was gathered from the NHC website http://www.nhc.noaa.gov.

NAME	MONTH	WARNING DURATION (days)	EST. MAX. SFC WIND (kt)	LOWEST PRESSURE (mb)
Hurricane Alex	JUL/AUG	5 ¾	105	957
TS Bonnie	AUG	10 ¼	55	1001
Hurricane Charley	AUG	5 ¼	130	941
Hurricane Danielle	AUG	8	95	964
TS Earl	AUG	2	45	1009
Hurricane Frances	AUG/SEP	14 ¾	120	935
Hurricane Gaston	AUG/SEP	4 ¾	65	985
TS Hermine	AUG	3 ½	50	1002
Hurricane Ivan	SEP	21 ½	145	910
TD Ten	SEP	1 ¾	30	1009
Hurricane Jeanne	SEP	15	105	950
Hurricane Karl	SEP	8 ½	125	938
Hurricane Lisa	SEP/OCT	13 ¼	65	987
TS Matthew	OCT	2 ¼	40	997
Subtropical Storm Nicole	OCT	1 ½	45	986
TS Otto	NOV/DEC	3 ½	45	995

Table 2.5 Tropical cyclones in the eastern North Pacific basin during 2004. A total of 16 tropical cyclones occurred: 4 tropical depressions, 6 tropical storms, and 6 hurricanes. The information was gathered from the NHC website http://www.nhc.noaa.gov

NAME	MONTH	WARNING DURATION (days)	EST. MAX. SFC WIND (kt)	LOWEST PRESSURE (mb)
TS Agatha	MAY	2 ¼	50	997
TD Two - E	JUL	1 ¼	30	1007
TS Blas	JUL	2 ½	55	991
Hurricane Celia	JUL	6 ¾	75	981
Hurricane Darby	JUL/AUG	5 ¾	105	957
TD Six - E	AUG	¾	25	1008
TS Estelle	AUG	5 ¼	60	989
Hurricane Frank	AUG	2 ¾	75	979
TD Nine – E	AUG	2 ¾	30	1005
TS Georgette	AUG	4	55	995
Hurricane Howard	AUG/SEP	6	120	943
Hurricane Isis	SEP	8 ¼	65	987
Hurricane Javier	SEP	8 ¾	130	930
TS Kay	SEP	1 ½	40	1004
TS Lester	OCT	1 ¾	45	1000
TD Sixteen - E	OCT	1 ¼	30	1004

It is used by the NHC as a “control” forecast to assess the skill of the other intensity forecast models. The SHIPS and DSHIPS techniques are statistical-dynamical techniques that combine climatological, persistence, and atmospheric predictors to forecast intensity changes. The DSHIPS adds an empirical decay model to the SHIPS model to account for the intensity change in storms that move over land. The GFDI and the GFNI dynamical models are full physics models that predict the structure and track of the TC and can be coupled with a sophisticated ocean model. The six-hour interpolated versions of the dynamical models are evaluated here due to their delayed availability beyond the time the NHC forecast must be issued, because full physics models require more than 2.5 hours to be completed. A comprehensive description of each of the intensity techniques design, including their general characteristics, is found in Appendices A through E.

III. RESULTS

A. SEASONAL INTENSITY ERROR SUMMARIES

1. Error Definitions

A typical analysis of intensity forecast techniques is to calculate the average magnitude of the intensity error and the intensity bias. That is, the average intensity error is an average of the absolute value of the intensity difference (intensity forecast minus observed value), which gives the magnitude of error. The intensity forecast bias is just the average of the intensity errors, i.e., if the intensity is under-(over-) forecast, the bias is negative (positive). If the technique has as many over-forecasts as under-forecasts, the bias should be near zero. Normally, homogeneous samples are used in the comparison of the different intensity techniques to assure that the same forecasts are being compared, as some techniques are not available at every forecast interval and some forecasts are easier than others.

2. Baseline Verification

The average intensity error and bias were computed from the NHC ATCF-format files (i.e., the data source) for the 2003 and 2004 Atlantic and eastern North Pacific hurricane seasons. The primary purpose of these intensity error and bias compilations is to ensure that the overall statistics for this study were very similar to the 2003 and 2004 NHC verifications reported by Gross (2004) and Franklin (2005) at the Interdepartmental Hurricane Conferences. The NHC errors and bias during 2004 were from the NHC official intensity forecasts that consist of the average intensity errors out to 120 h and the bias error at 24 h, 72 h, and 120 h for the Atlantic (Table 3.1) and eastern North Pacific (Table 3.2). The average intensity errors and bias from the NHC and this study are essentially identical at these forecast intervals. The 2003 NHC verification reports consist of the average intensity error (for the Atlantic only and with no bias) for the NHC official intensity forecasts. The average intensity errors calculated in this study are within +/- 1 kt of the 2003 NHC verification results.

Thus, the sample used in this study, and the forecast error algorithm, are consistent with the NHC verification study. Similar verifications will now be done for the intensity forecast techniques described in Chapter II and the Appendices with subsamples for the phases in the intensity life cycle (Figure 2.2).

Table 3.1 NHC Intensity Error - The average intensity error and bias from the 2004 NHC Verification Report for the Atlantic basin compared with the average intensity error and bias from the sample used in this study

Valid Time (hr)	The NHC Intensity Error (kt)	This Study Intensity Error (kt)	Bias (NHC) (kt)	Bias (This Study) (kt)
00	2.0	2.0		-0.1
12	7.4	7.4		+0.8
24	10.2	10.2	+ 0.5	+ 0.7
36	12.4	12.5		+0.5
48	13.9	14.0		+0.3
72	17.0	17.3	- 0.4	+0.1
96	19.8	20.2		-1.1
120	22.6	22.6	- 3.1	- 3.1

.

Table 3.2 Same as Table 3.1, except for the eastern North Pacific.

Valid Time (hour)	The NHC Intensity Error (kt)	This Study Intensity Error (kt)	Bias (NHC) (kt)	Bias (This Study) (kt)
00	1.7	1.6		-0.2
12	6.6	6.6		+1.1
24	11.4	11.4	+ 2.3	+ 2.8
36	14.4	14.4		+3.5
48	15.6	15.9		+3.0
72	18.8	18.9	+ 3.6	+ 4.0
96	17.8	17.8		+5.6
120	18.8	18.8	+ 7.0	+ 7.0

3. Average Error and Biases during 2003 Season

The biases for the five intensity techniques in the Atlantic basin during the 2003 season were small and negative (under-forecasting intensity), with the GFDI model having a larger negative bias at all forecast times except 120 h (Figure 3.1). The climatology and persistence technique SHF5 and the statistical-dynamical model SHIPS have small biases (of less than one knot) except at 120 h which suggests that this sample of intensity forecasts is similar to the developmental sample for these statistical techniques.

The intensity errors for all the techniques increase with increasing forecast interval (Figure 3.2). However, the average errors are generally small (less than 15 kt at most forecast intervals for all the techniques). As indicated above, the SHF5 is a measure of skill for the intensity models since it requires no special meteorological knowledge. Defining skill as smaller intensity errors than SHF5, only the SHIPS and DSHIPS techniques had skill relative to SHF5 during the 2003 season.

The biases for the five intensity techniques in the eastern North Pacific were all positive during the 2003 season (Figure 3.3). That is, all the intensity techniques had a tendency to over-forecast storms in this basin. Notice that the three statistically-based techniques (SHF5, SHIPS, and DSHIPS) all had large biases, which suggest that this sample of storm had different characteristics from the developmental samples, since such statistically-based systems should not have a bias. Using SHF5 as a measure of skill, the larger biases for the SHIPS technique would suggest it did not have skill. Although the DSHIPS technique had smaller biases than the SHIPS techniques, it also did not have skill relative to SHF5. Conversely, the GFDI had the smallest biases at all forecast intervals and did have skill relative to SHF5. Although the GFNI had larger biases than the GFDI, it still has small skill relative to SHF5. The average intensity errors in the eastern North Pacific increased with each forecast interval except for the DSHIPS at 72 h (Figure 3.4). The techniques had large average intensity errors, which should be expected since the biases were large (Figure 3.3).

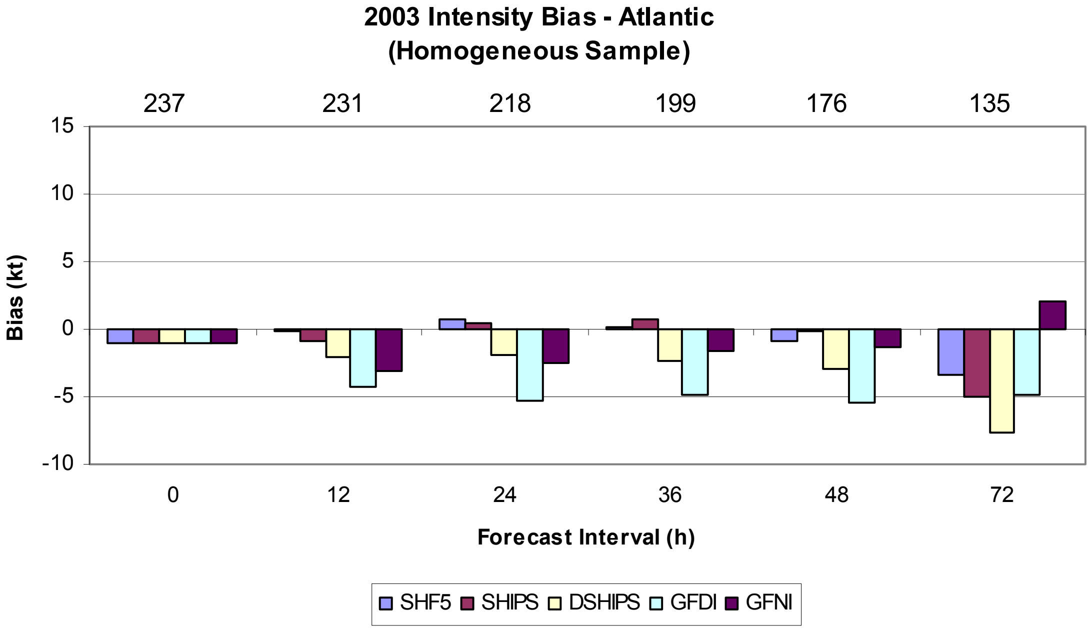


Figure 3.1 Biases of the intensity forecasts by five techniques at analysis time (time = 0) and at 12, 24, 36, 48, and 72 h during the 2003 Atlantic TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

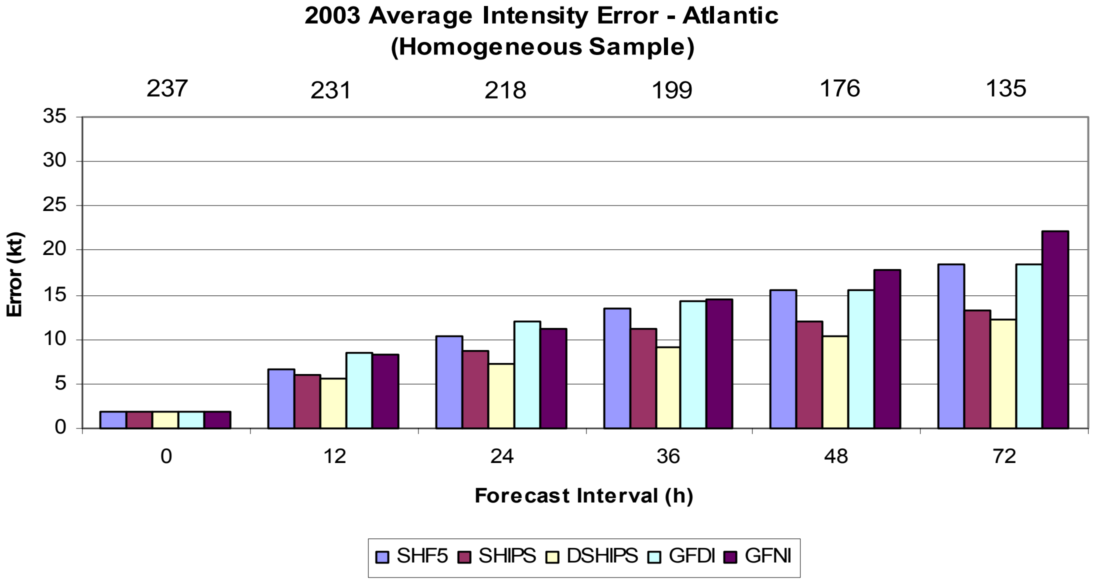


Figure 3.2 Average magnitude of the intensity forecast errors (kt) for five techniques at analysis time (time = 0) and at 12, 24, 36, 48, and 72 h during the 2003 Atlantic TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

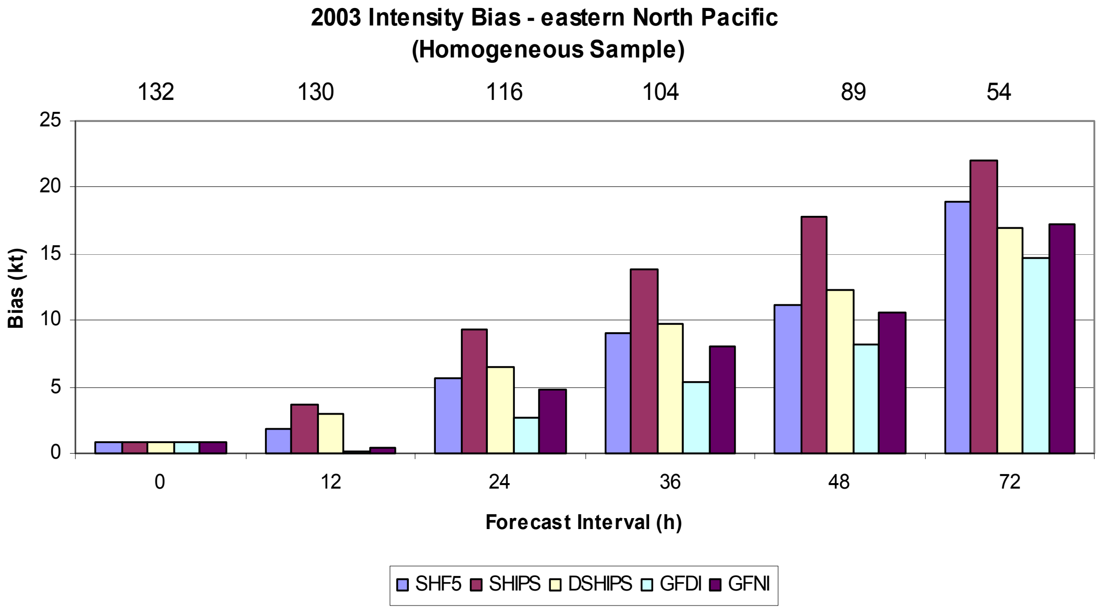


Figure 3.3 Bias of the intensity forecast as in Figure 3.1, except for the 2003 eastern North Pacific TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

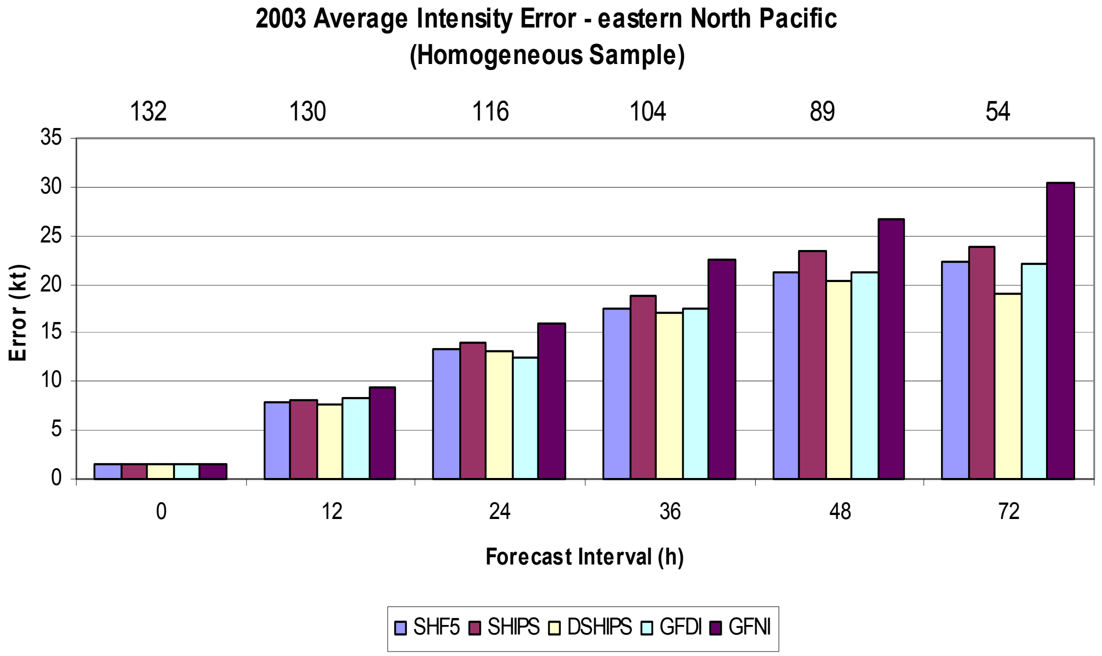


Figure 3.4 Average magnitude of the intensity forecast error as in Figure 3.2, except for the 2003 eastern North Pacific TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

Although the SHIPS technique did not have skill relative to SHF5, the DSHIPS technique did have skill, especially at 72 h. The GFDI had smaller average errors than the GFNI at all forecast intervals. At 24 h and beyond, the GFDI did have skill relative to SHF5. By contrast, the GFNI did not have skill at any forecast interval.

4. Average Error and Biases during 2004 Season

For the 2004 Atlantic season, the intensity biases of the two statistical-dynamical techniques, the SHIPS and DSHIPS, had opposite tendencies (Figure 3.5). While the SHIPS had all positive biases, the DSHIPS technique had all negative biases from the 12-h to the 120-h forecasts (note the longer forecast interval in 2004 relative to only 72 h in 2003). The 2004 Atlantic tropical cyclone season had many storms that moved over land or near small islands. The SHIPS technique, which does not take into account the decay of storms over land, will over-forecast the intensity of storms over land. Whereas the DSHIPS technique does decay the storms when the official forecast is across land, there was a small tendency to decay TCs too much over land especially at 96 h and 120 h. Considering the larger track errors at these times, it could be that the official track forecasts brought the TC over land too much or too soon.

The average intensity error for the 2004 Atlantic season was three to five knots higher than in 2003 for all techniques (compare Figure 3.6 and Figure 3.2). In contrast to the Atlantic and eastern North Pacific during 2003 (Figures 3.2 and 3.4), all of the techniques had skill relative to SHF5 through the 120-h forecast, except the DSHIPS model had no skill relative to SHF5 at the 96-h and 120-h forecast intervals. This lack of skill at the 96 h and 120 h by the DSHIPS technique again suggests that the larger track errors at day 4 and day 5 may contribute to an increase in the DSHIPS intensity errors.

In general, the intensity biases (Figure 3.7) for the eastern North Pacific during 2004 are small and positive with the exception of the GFDI model that had an increasingly negative bias with forecast time. Bender et al. (2003) stated that a tendency of the GFDI is to under-predict the intensity of weak systems, and

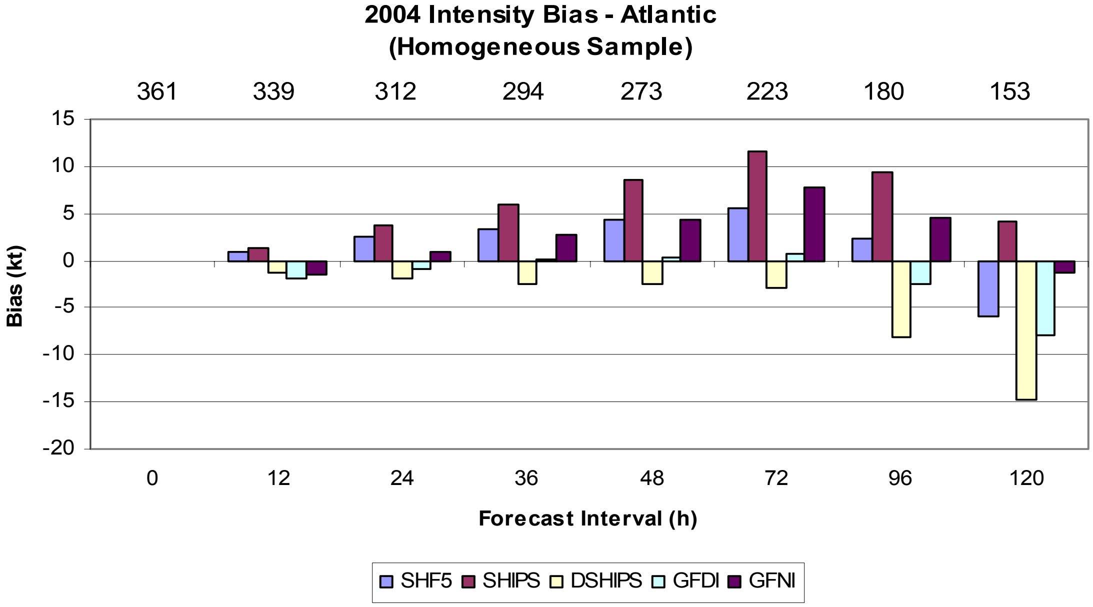


Figure 3.5 Biases of the intensity forecasts as in Figure 3.1, except for the 2004 Atlantic TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

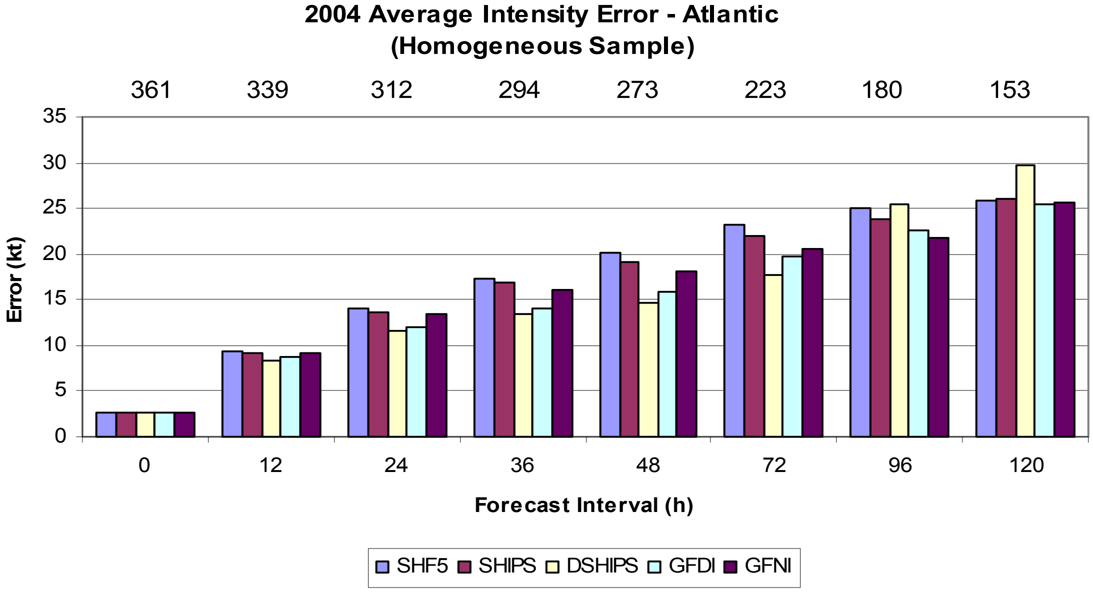


Figure 3.6 Average magnitude of the intensity forecast errors as in Figure 3.2, except for the 2004 Atlantic TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

during the 2004 eastern North Pacific TC season only three of the 16 TCs had maximum surface winds above 75 kt. It is interesting that the analogous GFNI model during 2004 season for the eastern North Pacific had only a small intensity bias.

For the eastern North Pacific average intensity errors (Figure 3.8), the SHF5 technique and the GFNI model have either the largest or second to largest values for 12-h through 72-h forecast intervals. Therefore, except for the GFNI, the techniques demonstrated skill out to the 72-h forecast. The SHIPS and DSHIPS, which have almost identical errors, have the lowest errors and have skill relative to SHF5 at all forecast interval. All the techniques have a general pattern of increasing average intensity errors with increasing forecast interval except at the 120-h forecast interval.

5. Bias and Average Error Comparisons among Basins

The hypothesis stated earlier is that similar techniques and numerical models used for intensity forecasting in the Atlantic and eastern North Pacific will have a similar performance as the intensity guidance models for the western North Pacific evaluated by Blackerby (2005). To examine this hypothesis, the biases and average intensity errors of the intensity techniques are compared among the three basins. The intensity techniques examined here have the same technique design in each basin: the SHF5, SHIPS, and GFNI used in the Atlantic and eastern North Pacific by the NHC and the ST5D, STIPS, and GFNI used in the western North Pacific by the JTWC. Whereas the western North Pacific and the 2003 Atlantic and eastern North Pacific biases and average intensity errors are only to 72 h, the 2004 Atlantic and eastern North Pacific forecasts are extended to 120 h.

a. Biases

For this comparison, the seasonal bias errors are averaged over the forecast interval for each intensity technique, and are characterized as small bias (< +/- 5 kt), moderate bias (+/- 5 kt to +/- 10 kt), and large bias (> +/- 10 kt). Table 3.3 is a summary of the biases for all three basins for 2003 and 2004.

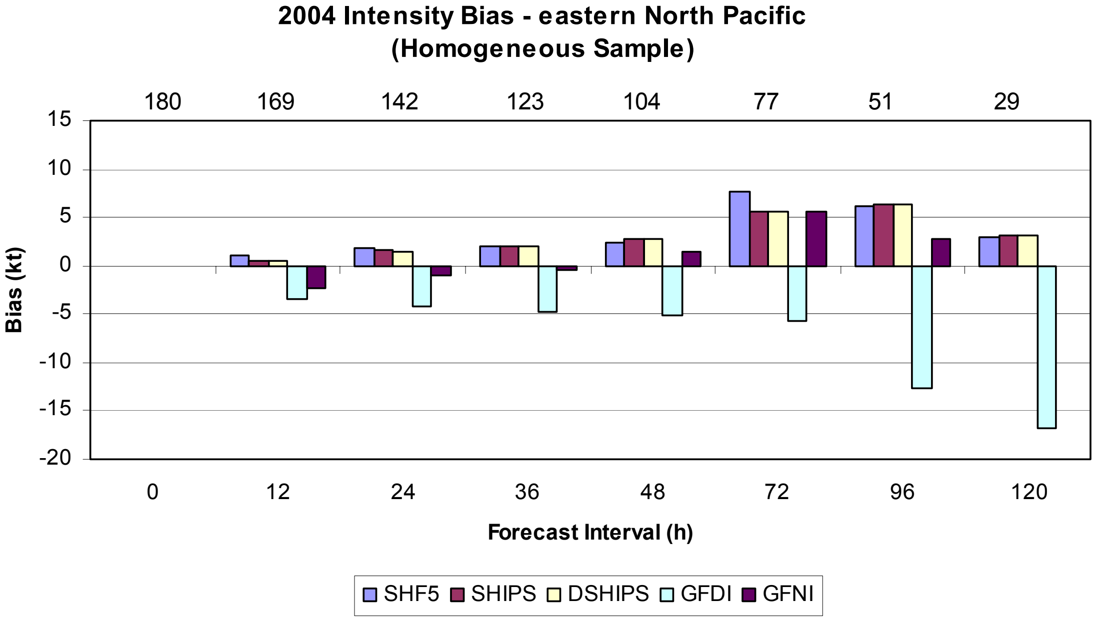


Figure 3.7 Biases of the intensity forecasts as in Figure 3.1, except for the 2004 eastern North Pacific TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

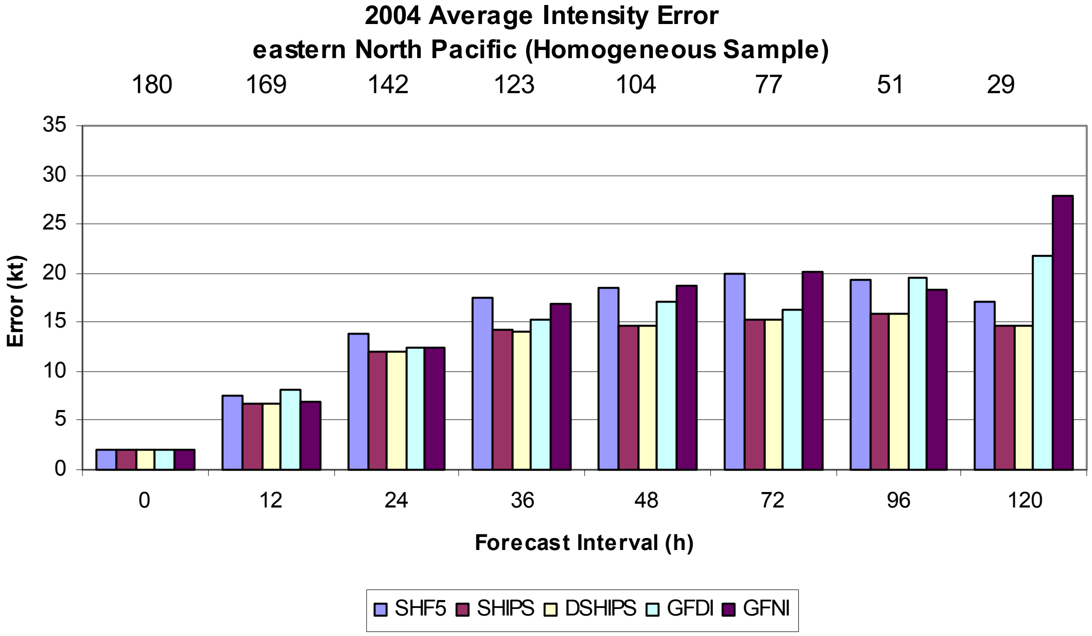


Figure 3.8 Average magnitude of the intensity forecast errors as in Figure 3.2, except for the 2004 eastern North Pacific TC season. Numbers along the top of the graph represent the sample sizes of homogeneous cases.

Both of the climatology and persistence techniques (SHF5 and ST5D) have small negative biases in the Atlantic and the western North Pacific in 2003. In 2004, the SHF5 and ST5D technique biases are small and positive for all three basins. The GFNI model also has similar biases in different basins, i.e., small positive biases for the Atlantic, eastern North Pacific, and western North Pacific in 2004. Conversely, the statistical-dynamical techniques SHIPS and STIPS biases are not consistent among the basins. For example, the 2003 SHIPS has a small negative bias in the Atlantic, a large positive bias in the eastern North Pacific, and a large negative bias in the western North Pacific. However, the biases in 2004 are more consistent. Based on this comparison of biases, a generally consistent pattern of biases is found for the climatology and persistence (SHF5 and ST5D) and the dynamical GFNI (except for the eastern North Pacific where the sample size is small during 2003).

Table 3.3 The biases are the average bias per intensity technique (from 0-h to 72/120-h forecast interval) per year. The techniques without bold-type are for the 2003 season and those in bold-type and highlighted are the biases for 2004.

BIASES	**ATLANTIC**	**EASTERN NORTH PACIFIC**	**WESTERN NORTH PACIFIC**
Small negative (< -5 kt)	SHF5 SHIPS GFNI		ST5D
Small positive (< +5 kt)	***SHF5*** ***GFNI***	***SHF5*** ***SHIPS*** ***GFNI***	GFNI ***GFNI*** ***ST5D***
Moderate negative (- 5 kt to -10 kt)			***STIPS***
Moderate positive (+ 5 kt to +10 kt)	***SHIPS***	SHF5 GFNI	
Large negative (> -10 kt)			STIPS
Large positive (> +10 kt)		SHIPS	

b. Average Intensity Errors

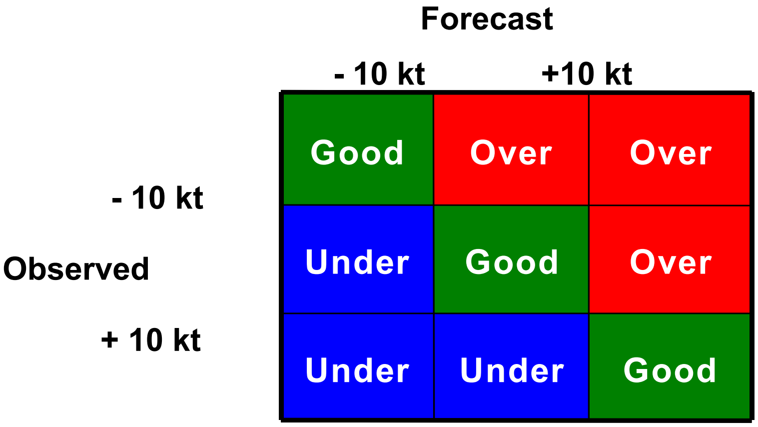
The seasonal average intensity errors are averaged over the forecast interval for each intensity technique, and are categorized in increments of 5 kt. Table 3.4 is a summary of the average intensity errors for all three basins for 2003 and 2004. The average intensity errors for all the techniques increased with each forecast interval in all three basins for both years. That is, the guidance of the intensity techniques becomes less reliable to the NHC and JTWC as the forecast time increases. The eastern North Pacific and the western North Pacific techniques have consistent average intensity errors when averaged over the forecast interval for the 2003 and 2004 seasons. Thus, the eastern North Pacific and the western North Pacific techniques performance do support the hypothesis that intensity guidance will have a similar performance in the different basins. The climatology and persistence techniques (SHF5 and ST5D) average intensity errors are also consistent in all three basins during 2003. In the Atlantic, the SHIPS and GFNI techniques have smaller average intensity errors in 2003 and larger average intensity errors during 2004. A consistent pattern exists with the average intensity errors for the climatology and persistence techniques (SHF5 and ST5D) as was found with the biases.

Table 3.4 The average intensity errors are the average errors per intensity technique (from 0-h to 72/120-h forecast interval) per year. The techniques without bold-type are for the 2003 season and those in bold-type and highlighted are the biases for 2004.

AVERAGE INTENSITY ERRORS (kt)	**ATLANTIC**	**EASTERN NORTH PACIFIC**	**WESTERN NORTH PACIFIC**
0 – 5			
6 – 10	SHIPS		
11 -15	SHF5 GFNI	SHF5 SHIPS ***SHF5 SHIPS GFNI***	ST5D STIPS **ST5D STIPS GFNI**
16 - 20	***SHF5 SHIPS GFNI***	GFNI	GFNI

B. GUIDANCE TREND PERFORMANCE DURING DIFFERENT INTENSITY PHASES

Following Blackerby (2005), a contingency table is used to evaluate the performances of the different intensity guidance techniques during each intensity phase (Phase I to Phase III in Figure 2.2). The contingency table is based on whether a guidance technique can forecast the general intensity trends: increasing, remaining nearly constant, or decaying. For example, when the intensity technique forecasts an intensity increase of 20 kt or more when the observed intensity increase is any value less than 10 kt, an Over (O) intensity trend performance is assigned to the intensity technique for that particular forecast (Figure 3.9). Similarly, an Under (U) forecast is assigned if the technique forecasts an intensity change that is at least 10 kt smaller than the observed intensity change over the forecast interval. A Good (G) intensity trend is then assigned if the forecast intensity change is within +/- 10 kt of the observed intensity change over the forecast interval.



Good	Over	Over
Under	Good	Over
Under	Under	Good

Figure 3.9 Sample contingency table with Good (G) intensity trend forecasts along the diagonal. A Good forecast was assigned if the forecast intensity change fell within +/- 10 kt of the observed intensity change for each 12-h forecast interval, an Under (U) trend represents a forecast intensity change that is too small by more than 10 kt, and an Over (O) trend represents a forecast intensity change that is too large by more than 10 kt (from Blackerby 2005).

Using this procedure, the intensity trend analyses are calculated for all forecast intervals in the three intensity phases. The purpose is to determine how well each guidance technique and the NHC forecasts the intensity trend within +/- 10 kt throughout the storm life cycle. That is, even if the technique can not forecast the actual intensity value, can it forecast the intensity trend within +/- 10 kt?

C. PHASE I: FORMATION

1. Phase I: Good (G) Intensity Trends

a. Atlantic

All of the intensity techniques and the NHC have a high percentage (mostly 90%) of G intensity trends for the 12-h forecast period during the formation stage, but the percentage drastically declines (~ 45%) by 24 h (Figure 3.10). The percentage of good forecasts then increases at longer forecast intervals. One interpretation is that for this sample in which a tropical cyclone is more than likely to form (because the NHC would otherwise not be following it), the techniques will tend to predict such a formation because they were designed to do so from Phase I conditions. The statistical-dynamical techniques would be expected to have a good trend performance at 12 h because they contain a persistence of past 12-h trend predictor. However, continuing that trend (predominating zero or small increase) out to 24 h will miss the formation about 50% of the time. The dynamical models also do well with only a 12-h forecast, but also miss more than half of the formation cases by 24 h.

Whereas the SHIPS technique has skill relative to the SHF5 only from 24 h to 60 h, the DSHIPS technique has skill relative to SHF5 from the 24-h through the 96-h forecast interval. Also, the DSHIPS out-performs the SHIPS from the 48-h to 96-h forecast interval. The NHC official forecast was skillful relative to SHF5 (note that NHC forecasts are not archived at 60 h, 84 h, and 108 h) from the 36-h to 96-h forecast intervals with greater than 15% better than the SHF5 at 36 h and 72 h. None of the techniques has skill relative to the SHF5 beyond 96 h. The GFDI has skill relative to the SHF5 at the 24-h and 36-h

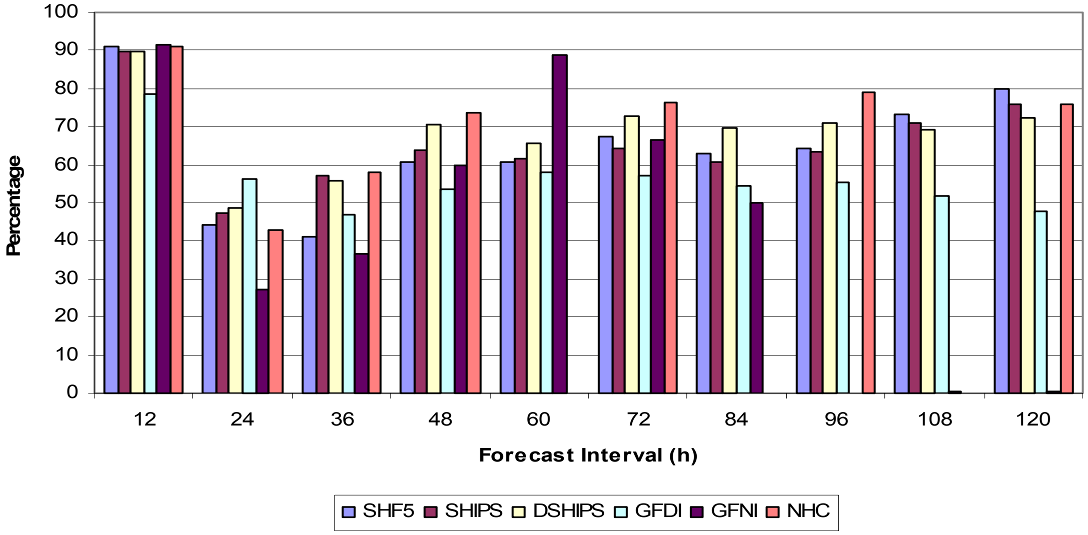


Figure 3.10 Percentage of Good intensity trends as defined in Figure 3.9 during storm formation (Phase I) for the combined 2003 and 2004 Atlantic TC seasons. A Good forecast during the formation period indicates that the magnitude of the forecast intensity change is within +/- 10 kt of the actual intensity change. Although sample sizes of the verified forecasts range from over 350 at 12 h to 225 at the 96-h interval, this is not a homogeneous sample and the dynamical models in particular have many fewer forecasts.

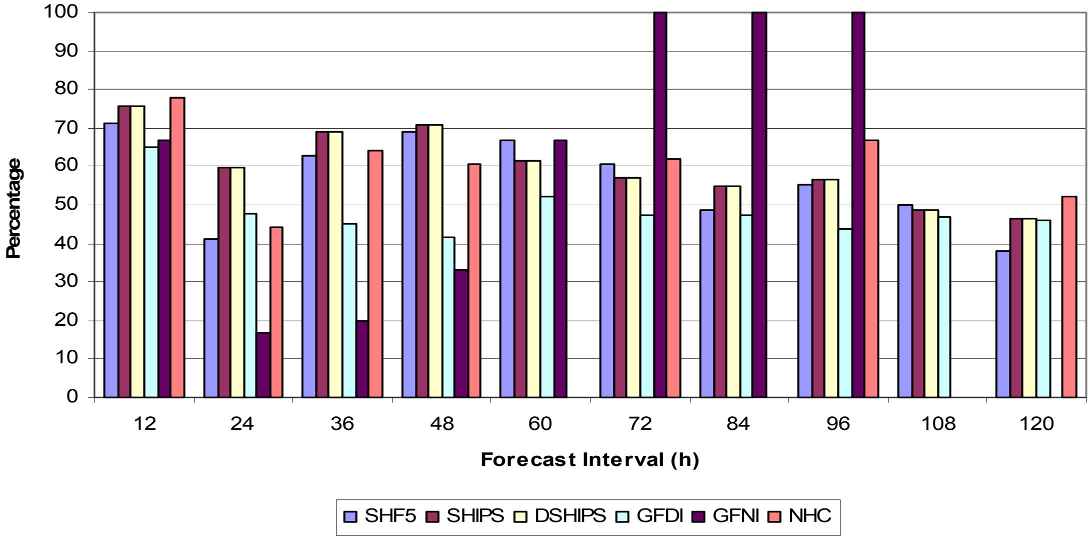


Figure 3.11 Same as Figure 3.10, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from 290 at 12 h to less than 130 by the 96 h, but again very much smaller sample sizes are available for the GFNI.

forecast intervals and out-performs the other techniques and the NHC for the 24-h forecasts (although as noted above, these percentages are small). For the remainder of the forecast, the GFDI has the lowest G percentages. The GFNI technique has a very small sample size of 12 at 12 h and decreases to two by the 96-h forecast interval. Therefore, the graphs do not adequately represent the GFNI trend performance during Phase I for the Atlantic.

b. Eastern North Pacific

Consistent with the Atlantic, high percentages of G intensity trends occur at the 12-h forecast interval, although the values are lower than the percentages in the Atlantic (Figure 3.11). The SHIPS and DSHIPS techniques have identical G percentages throughout the forecast intervals and are skillful relative to SHF5 from 12 h to 48 h and at 84, 96, and 120 h. The cause of these identical forecasts is mainly due to the TC tracks in the eastern North Pacific Ocean. Very few of these TCs moved over land during 2003 and 2004. Therefore, the DSHIPS forecast will be identical to the SHIPS. The NHC official Good (G) forecast trend was marginally skillful relative to SHF5 at all forecast intervals except at 48 h (again the NHC does not provide forecasts at 60-, 84-, or 108-h forecast periods). The GFDI model has only two G percentages above 50 % (12 h and 60 h), but has skill relative to SHF5 at the 24-h and 120-h forecast intervals. Since the GFNI technique has a very small sample size (6 or less) the graphs of Phase I for the eastern North Pacific do not accurately represent the GFNI performance.

2. Phase I: Under (U) Intensity Trends

a. Atlantic

High U percentages during Phase I indicate that intensity techniques may predict zero intensity changes or decreases even though the storm is intensifying during Phase I (Figure 3.12). Because so many of the 12-h forecasts had a Good trend forecast, the possible percentages of U can not be large at 12 h. The measure of skill for the intensity techniques is to have smaller percentages of U forecast than the SHF5, which has large U percentages at 24 h and 36 h, but then has small percentages at longer forecast intervals. In

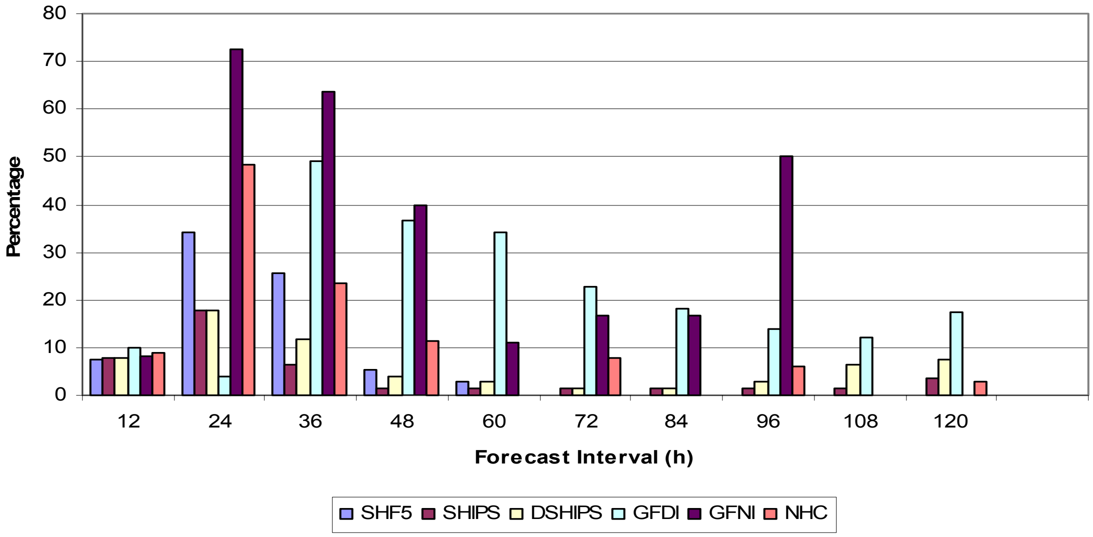


Figure 3.12 Percentage of Under intensity trends during Phase I as defined in Figure 3.9 during storm formation for the combined 2003 and 2004 Atlantic TC seasons. An Under forecast during formation indicates that the magnitude of the forecast intensification rate is less than the magnitude of the actual intensification rate by at least 10 kt. Sample sizes of the verified forecasts range from over 350 at 12 h to 225 at 96 h, except that the GFNI has much smaller sample sizes.

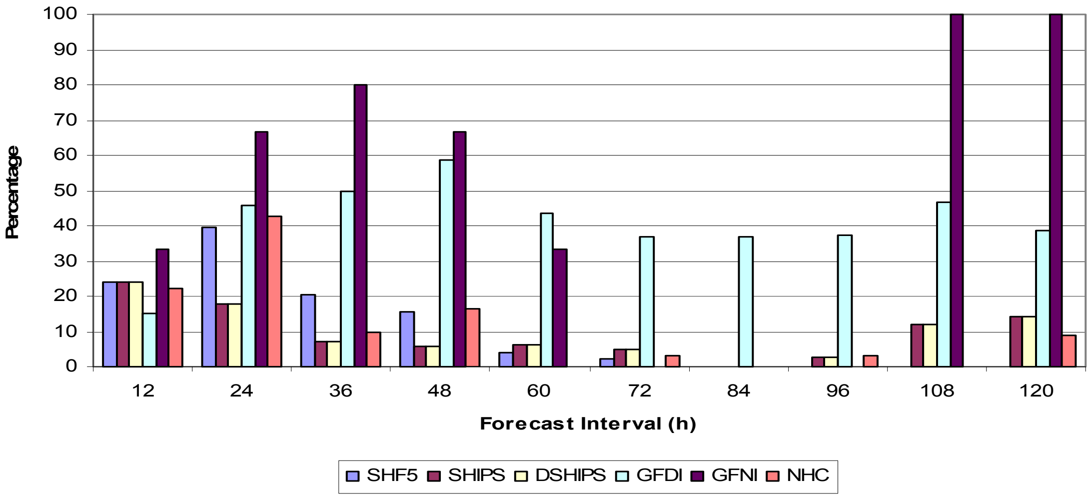


Figure 3.13 Same as Figure 3.12, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from 290 at 12 h to less than 130 by 96 h and specifically comparatively small GFDI and GFNI forecasts are available.

particular, the SHF5 has no under-forecasts (0%) of the intensity trend during Phase I from 72 h to 120 h. The NHC has a high U percentage at the 24-h forecast interval (48%), but this percentage declines gradually to 3% for 120-h forecasts. Thus, the NHC has skill relative to the SHF5 in the U forecast only at 36 h. The SHIPS and DSHIPS techniques (which are almost identical) have low percentages of under-forecasts except at the 24-h and 36-h forecast intervals. Thus, the SHIPS and DSHIPS show skill relative to SHF5 in the 24-h to 60-h forecast intervals. The GFDI model has a high U percentage (greater than 22%) from 36-h to 60-h forecast interval and is skillful relative to SHF5 only at the 24-h forecast interval. Again, the sample sizes for the GFNI model are too small to be reliable.

b. Eastern North Pacific

The SHF5 technique has low U percentages during Phase I except at the 24-h forecast period, with no under-forecasts from 84 h to 120 h (Figure 3.13). The GFDI technique's inability to forecast intensification in the formation phase is again evident in the eastern North Pacific basin. Although based on a small sample, this model continues to have high U percentage (above 30%) at all forecast intervals except at 12 h and has skill relative to the SHF5 only at 12 h. The relatively low U percentages (0% at 84 h) by the SHIPS and DSHIPS techniques indicate that these techniques rarely produce an under-forecast of intensity change during the formation Phase I, which was also demonstrated in the Atlantic basin by both the SHIPS and DSHIPS techniques. The low percentages of under-forecasts by the SHF5, SHIPS, and DSHIPS techniques reveal that they normally generate forecast of intensification when the storm is weak. The NHC U trends in Phase I are over 20% at 12 h, over 40% at 24 h, but they have a U percentage below 20% with relatively low percentages on days 3, 4, and 5. However, only at the 12-h and 36-h forecast intervals is the NHC skillful relative to the SHF5.

3. Phase I: Over (O) Intensity Trends

Given the Good- and Under-forecasts for the intensity trends during Phase I discussed above, the Over-forecasts are the remaining forecasts since the total must equal 100%.

a. Atlantic

The SHF5 technique has a relatively high percentage of O intensity trends during Phase I beyond the 12-h forecast interval. The SHIPS technique also has a general pattern of over-forecasting TC intensities during the formation stage (Figure 3.14) with the majority of the percentages over 30%, except at 12 h (2.6%). The SHIPS technique has the highest percentage of O intensity trends at each forecast period from 24 h to 84 h and has skill relative to the SHF5 only at the 96-h forecast interval. The DSHIPS pattern is almost identical to the SHIPS with no over-forecasting of the TC in the first 12 h, a dramatic increase at 24 h, and then a consistent pattern of relatively high O percentages for the longer forecasts. Unlike the SHIPS, the DSHIPS has skill relative to the SHF5 from the 36-h to the 108-h forecast intervals. The GFDI has few cases of over-forecasting intensity during the formation phase until day 3 when the percentage is consistently over 20%. The NHC does not (0%) over-forecast intensity during the first 12 h and maintains a modest O percentage of approximately 15% throughout the forecast interval.

b. Eastern North Pacific

After the first 12 hours, the percentage of Over-forecasts during Phase I for the SHF5, SHIPS, and DSHIPS techniques increase (Figure 3.15) and have similar patterns (the SHIPS and DSHIPS are identical). The SHIPS and DSHIPS techniques only have skill relative to SHF5 at 12 h and after 72 h. Whereas the GFDI has a high percentage of O forecasts at 12 h, the percentages decrease through 60 h, and then an increase in over-forecast percentage occurs, but not above 20%. The GFDI has skill relative to the SHF5 from the 24-h to the 120-h forecast intervals. The NHC trend resembles the SHF5, SHIPS, and DSHIPS techniques, which is to be expected since these techniques are an important part of the forecaster’s guidance.

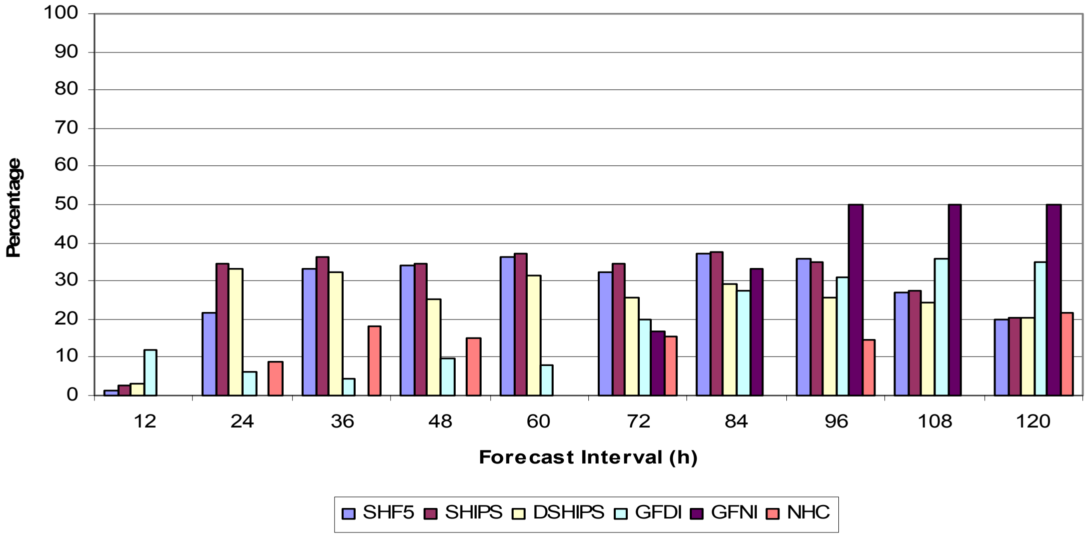


Figure 3.14 Percentage of Over intensity trends as defined in Figure 3.9 during storm formation for the combined 2003 and 2004 Atlantic TC seasons. An over-forecast during Phase I indicates that the magnitude of the forecast intensification rate exceeds the magnitude of the actual intensification rate by at least 10 kt. Sample sizes of the verified forecasts range from over 350 at 12 h to 225 at the 96-h interval.

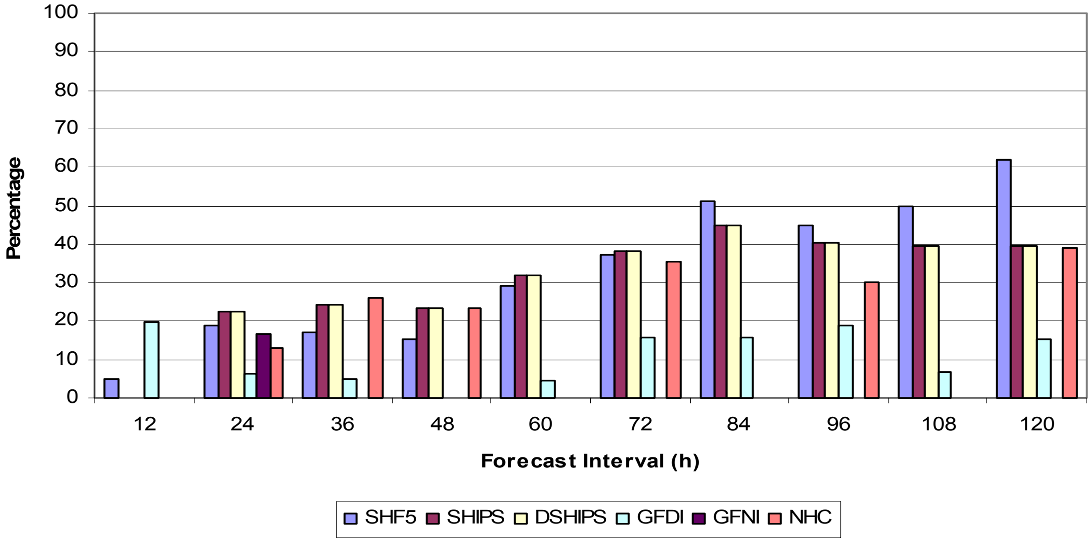


Figure 3.15 Same as Figure 3.14, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from 290 at 12 h to less than 130 by the 96 h point.

4. Phase I: Summary

The climatological and persistence technique SHF5 and the statistical-dynamical techniques DSHIPS and SHIPS provided the most reliable intensity trend forecasts during Phase I in both the Atlantic and eastern North Pacific basins. The SHF5 is believed to be more reliable in the eastern North Pacific basin than in the Atlantic during the formation stage due to a drastically smaller area in which TCs can form compared to the Atlantic. Lower sea-surface temperatures to the north and south and large vertical wind shear in the western region of the eastern North Pacific Ocean allows TCs a somewhat constrained area for formation, and the paths are similar. Thus, a technique based on climatology is more apt to have a more accurate intensity forecast in a region in which most storms intensify in a similar manner. However, the SHF5, SHIPS, and DSHIPS techniques have more of a tendency to over-forecast than under-forecast the intensity trend of TCs in both basins after the first 12 hours during the formation phase. Conversely, the GFDI technique has more of a tendency to under-forecast TCs after the first 12 hours during the formation phase in both basins. However, the GFDI technique does have skill at the 24-h forecast period in the Atlantic since it has the highest percentage of G intensity trends (56%) and the lowest percentage of over- and under-forecasts. The NHC intensity trend errors resemble those of the DSHIPS and SHIPS techniques in both basins, but under-forecasts storms at the 24-h forecast interval, which is also a GFDI trend.

5. Phase I: Comparison with the Western North Pacific

Blackerby (2005) found that overall the comparable climatology and persistence ST5D and the statistical-dynamical STIP techniques provided the most reliable intensity change trends in the western North Pacific during Phase I, which is identical to the finding in this study. Blackerby (2005) also suggested that since the best guidance (ST5D and STIP) and the JTWC both have similar G intensity trend percentages (~75%) when averaged over all forecast intervals during Phase I, the value-added or skill is small. For the Atlantic, the G intensity trends averaged over all forecast intervals by the best techniques are 65% for SHF5, 66% for SHIPS, and 69% for DSHIPS. The NHC G intensity trend percentage when averaged over all forecast intervals is 71%, which suggests that the NHC follows their best guidance. Similar to the western North Pacific, the value-added is small. The average G intensity trends in the eastern North Pacific also suggest that the value-added or skill is small. The G intensity trends averaged for the best techniques are SHF5 (56%), SHIPS (60%), and DSHIPS (60%) in the eastern North Pacific, while the NHC has a percentage of 61%. In all three basins, the value-added is small during Phase I. The dynamical models also had similar performances in all three basins with a trend toward under-forecasting during the formation stage.

D. PHASE II: INTENSIFICATION

As noted in Chapter II section A, Phase II is intensification of the tropical storm through peak intensity in which the intensification can either be rapid, typical, or slow. Thus, Phase II contains any forecast from the time the TC is above 34 kt until the time of the first peak intensity, and may include all forecast intervals to 120 h (Figures 2.1 and 2.2). Therefore, the verification time for the longer forecasts could be in Phase IIa or Phase III. The same contingency table (Figure 3.9) is used to evaluate the performances of the intensity guidance techniques and the NHC for these intensity changes during intensification. Whereas the evaluation of the GFNI technique in Phase I was not representative due to a small sample size, the technique is evaluated during Phase II with sample sizes of over 100 in both basins.

1. Phase II: Good (G) Intensity Trends

a. Atlantic

The SHIPS and DSHIPS techniques and the NHC have skill relative to SHF5 in terms of the G intensity trends for all forecast intervals (Figure 3.16). The NHC has the highest percentage of Good (G) intensity trends at the 12-, 36-, and 72-h forecast intervals followed by the two statistical models. The DSHIPS technique has higher percentages of G forecast compared to SHIPS at all forecast intervals except 12 h, 108 h (identical), and 120 h and has the highest overall G percentages at 24 h, 48 h, and 60 h. The GFNI model does not have skill relative to SHF5 in term of G intensity changes during the first 72 hours of the intensification phase, and has the lowest G percentages from the 36-h to 72-h forecast intervals. However, the GFNI has skill relative to SHF5 from the 84-h to 120-h forecast intervals with the highest G percentage at the 96-, 108-, and 120-h forecast intervals. The GFDI model G percentage averages 51% during the entire forecast interval with only one forecast interval (12 h) above 60%. That is, the GFDI model provides intensity change guidance within +/- 10 kt during Phase II in only about 5 of 10 cases. The GFDI model still has skill relative to the SHF5 from the 36-h to 96-h forecast intervals.

b. Eastern North Pacific

In general, the percentages of Good intensity trend forecasts during Phase II in the eastern North Pacific are lower than in the Atlantic, with all of the techniques and the NHC having percentages around 50% for the 24-h through 48-h forecast intervals (Figure 3.17). None of the techniques or the NHC have G percentages above 62% except at the 12-h and 120-h forecast intervals. At 72 h and beyond, the percentages of G intensity trends for the SHF5 decrease, with only about 35% G (and thus 65% either U or O forecasts) at 96 h and 108 h. Thus, this skill measure is quite low. In contrast to the Atlantic, the GFDI model is better in the eastern North Pacific. Although the GFDI model averages 56% G intensity trends over all of the forecast intervals, the GFDI has the highest G percentages at the 60-h, 72-h, 108-h, and 120-h forecast intervals. The GFNI model also has relatively high G percentages from the 84-h to the 108-h forecast intervals, which is consistent with the Atlantic, but the G percentage declines drastically at 120 h. The SHIPS and DSHIPS percentages are identical or near identical, which is similar to Phase I in the eastern North Pacific. The DSHIPS G percentage is slightly higher than SHIPS (same for Phase II-Atlantic) and has skill relative to SHF5 at all forecast intervals while SHIPS does not have skill relative to SHF5 at the 12-h and 60-h forecast intervals. Although the NHC has the highest G percentages from the 12-h to 36-h forecast intervals, these percentages are only slightly higher than the intensity techniques. In summary, the NHC has little to no intensity trend guidance that is consistently good during the intensification phase in the eastern North Pacific.

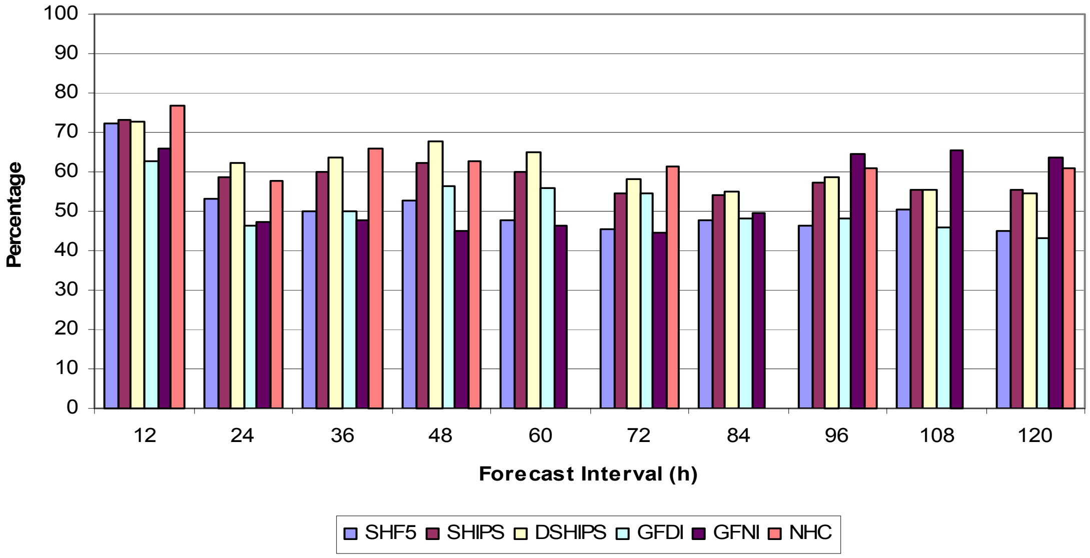


Figure 3.16 Percentage of Good intensity trends as defined in Figure 3.9 during storm intensification (Phase II) for the combined 2003 and 2004 Atlantic TC seasons. A Good forecast during intensification indicates that the magnitude of the forecast intensification rate is within +/- 10 kt of the actual intensification rate. Although, sample sizes of the verified forecasts range from over 1400 at 12 h to over 700 at the 96-h interval, this is not a homogeneous sample and the dynamical models in particular have many fewer forecasts.

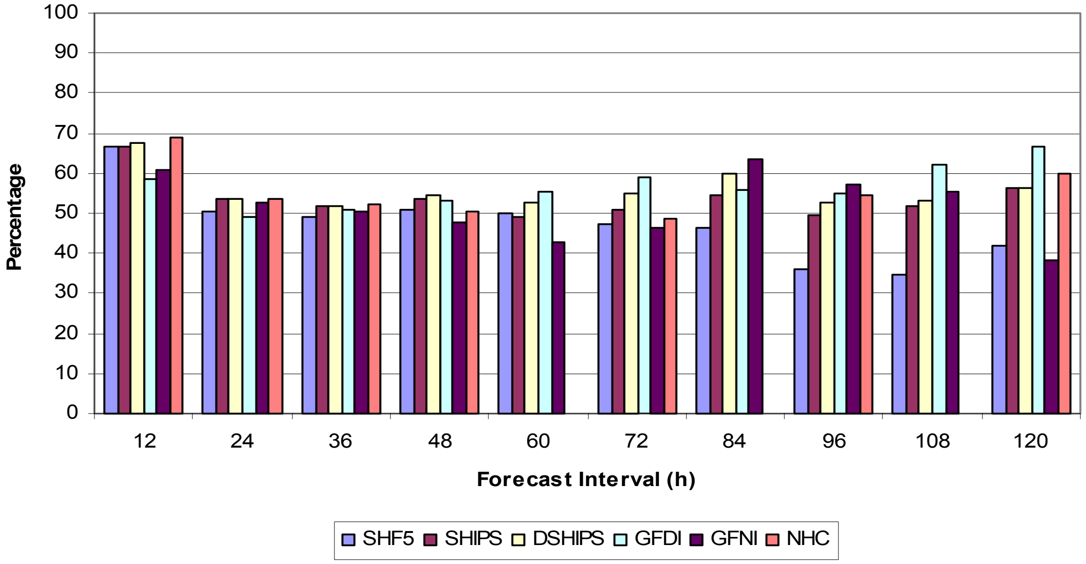


Figure 3.17 Same as Figure 3.16 except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 950 at 12 h to over 280 by the 96-h interval.

2. Phase II: Under (U) Intensity Trends

a. Atlantic

Most of the intensity guidance techniques and the NHC consistently have Under percentages of about 20% at all forecast intervals during the intensification phase in the Atlantic (Figure 3.18). The GFDI and GFNI have the highest U percentage (some over 30%) from the 12-h to 72-h forecast intervals. That is, both models have a tendency to under-forecast intensity changes during the intensification phase. The GFDI maintains this trend through the 120-h forecast interval while the GFNI percentage decreases to below 7% at 96 h and 108 h before increasing to 16% at 120 h. Consequently, the GFDI does not have skill relative to SHF5 in terms of having smaller percentages of U forecasts during the entire forecast interval. In contrast, the SHIPS technique has skill relative to SHF5 from the 12-h to 72-h forecast intervals and has the lowest U percentages through the 60-h forecast period. The DSHIPS resembles the SHIPS technique, but has slightly higher U percentages at all forecast periods. When averaged over all of the forecast intervals, the NHC average U percentage is approximately 19%, which is slightly better than the average U percentage (22.5%) of all of the techniques. This difference suggests that the NHC adds value to the intensity forecasts by the various techniques in avoiding some of the U forecasts, but the value added is small.

b. Eastern North Pacific

Every intensity technique and the NHC have U percentages over 20% at the 12-h forecast interval during the intensification phase with only the SHIPS and DSHIPS having skill relative to SHF5 (Figure 3.19). After the 12-h forecast interval, the two statistical-dynamical techniques and the two dynamical models have contrasting U trends. For the SHIPS and DSHIPS techniques, the U percentages decrease below 7% by the 48-h forecast interval, are below 2% by 84 h, and both have zero percentages at the 108-h forecast interval. Both the SHIPS and DSHIPS have skill relative to SHF5 through the entire forecast interval with the SHIPS having a slightly lower (or identical) U percentage as the

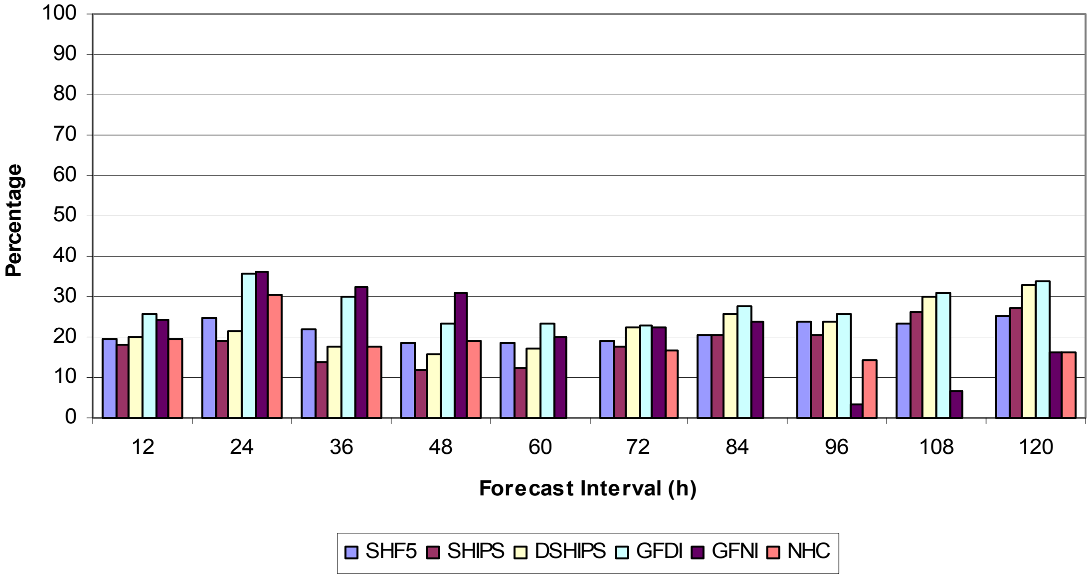


Figure 3.18 Percentage of Under intensity trends during Phase II as defined in Figure 3.9 for the combined 2003 and 2004 Atlantic TC seasons. An Under forecast during intensification indicates that the magnitude of the forecast intensification rate is less than the magnitude of the actual intensification rate by at least 10 kt. Sample sizes of the verified forecasts range from over 1400 at 12 h to over 700 at the 96-h interval, except that the dynamical models have smaller sample sizes.

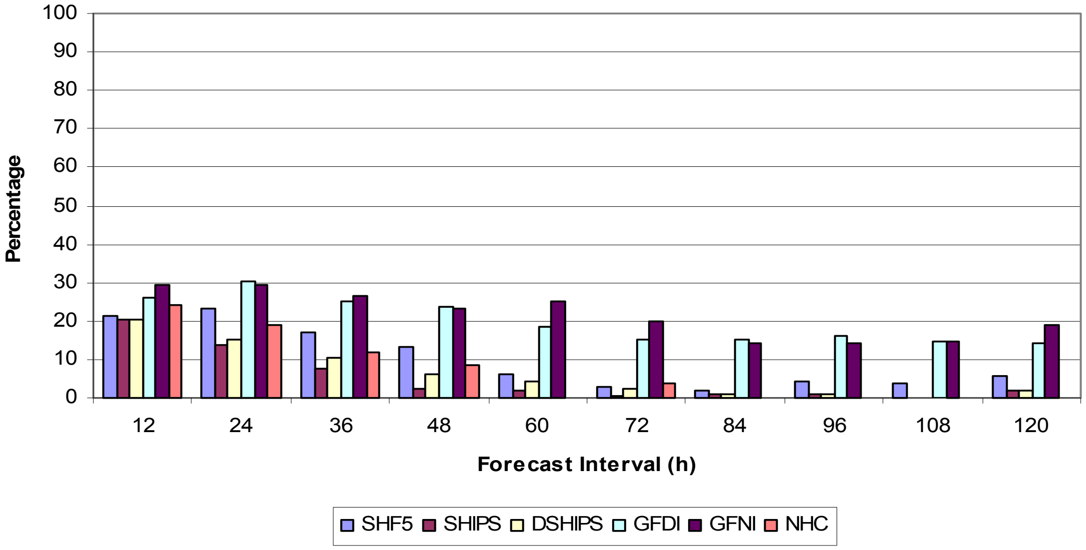


Figure 3.19 Same as Figure 3.18, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 950 at 12 h to over 280 by the 96-h forecast interval.

DSHIPS technique. The two dynamical models have U percentages above 23% from the 12-h to 48-h forecast intervals and continue to have U percentages over 14% through the entire forecast interval. The SHF5 has relatively high U percentages from the 12-h forecast to the 36-h forecast interval and then decreases to less than 3% by the 72-h forecast. The NHC percentages of U intensity trends in Phase II are very similar to the SHIPS and DSHIPS, but the NHC percentages are slightly higher than both models at every forecast interval. This comparison suggests that the NHC relies more on the statistical-dynamical techniques during the intensification stage, but no value is added.

3. Phase II: Over (O) Intensity Trends

As in Phase I, the Good, Under, and Over percentages must sum to 100. Therefore, the Over (O) intensity change percentages are the remaining portion of the forecast.

a. Atlantic

Except at 12 h, most of the techniques have O percentages of 20% or more at all forecast intervals during Phase II for the Atlantic tropical cyclones (Figure 3.20). Whereas the NHC over-forecast percentage increases with increasing forecast interval except at 120 h, they have the lowest Over (O) percentages at 12, 24, and 36 h. Although the NHC has skill relative to SHF5 at all forecast intervals, this is in part because the SHF5 technique has the highest O percentage in six of the 10 forecast intervals. Similarly, the DSHIPS technique has skill relative to SHF5 at all forecast intervals and has the lowest O percentages from the 48-h to 120-h forecast intervals excluding 72 h. The SHIPS technique has more of a tendency than the DSHIPS to over-forecast intensity during Phase II especially during the first 60 h of a tropical storm intensification. The difference in performance between the DSHIPS and SHIPS techniques is the ability of the DSHIPS technique to utilize the decay formula that provides more of a decrease in the intensity of a TC that is forecast to move over land. Therefore, the DSHIPS will have less of a tendency to over-forecast than the SHIPS technique when a TC is forecast to move over land. Consequently,

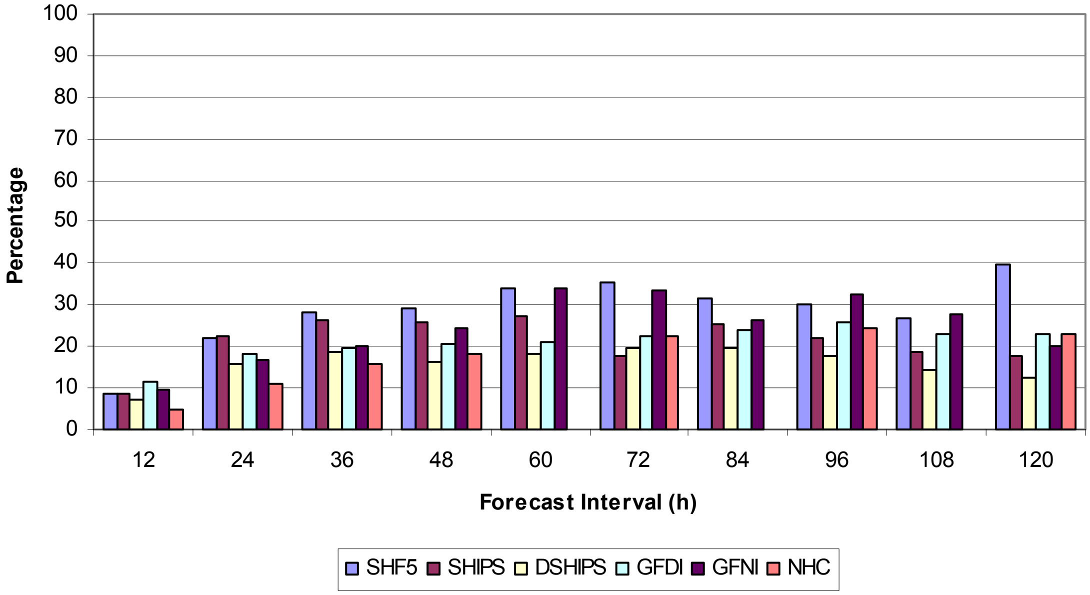


Figure 3.20 Percentage of Over intensity trends as defined in Figure 3.9 during storm intensification for the combined 2003 and 2004 Atlantic TC seasons. An over-forecast during Phase II indicates that the magnitude of the forecast intensification rate is greater by at least 10 kt than the magnitude of the actual intensification rate. Sample sizes of the verified forecasts range from over 1400 at 12 h to over 700 at the 96-h interval.

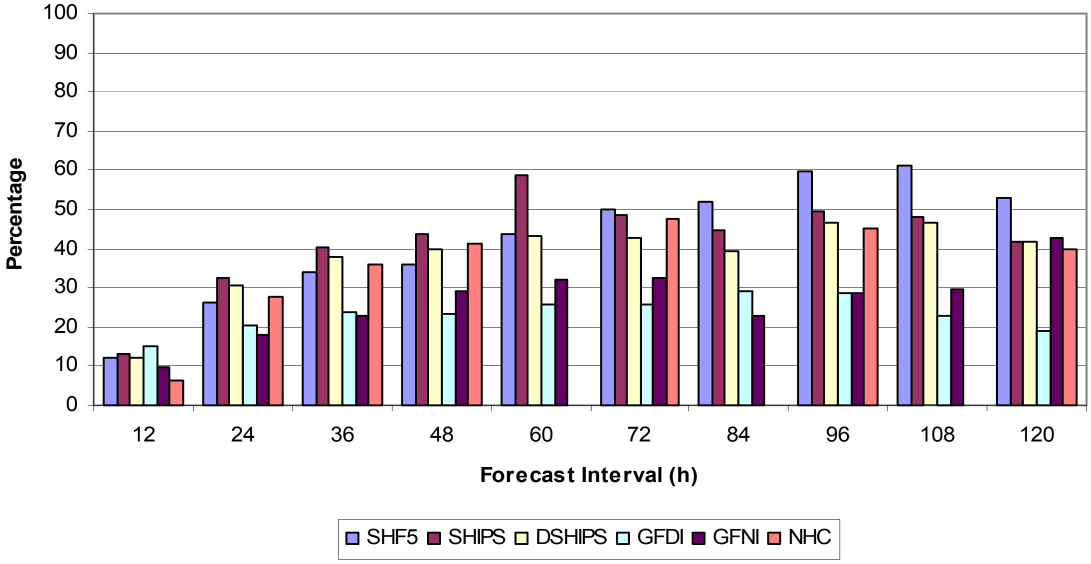


Figure 3.21 Same as Figure 3.20, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 950 at 12 h to over 280 by the 96-h forecast interval.

the SHIPS technique has no skill relative to the SHF5 at 12 h and 24 h and is over 25% from the 36-h to 60-h forecast intervals. The 12-h forecast is the only interval in which the GFDI model has no skill relative to the SHF5 in terms of the percentages of O intensity trend forecast in Phase II. Still, the model has approximately a 20% tendency to over-forecast intensity during the intensification phase. Similarly, the GFNI model over-forecast percentages increase with increasing in time until they reach a peak of 33% at 60 h and are then over 25% until the 120-h forecast interval.

b. Eastern North Pacific

Whereas most of the techniques and the NHC have relatively small percentages of U forecasts after 24 h in the eastern North Pacific (Figure 3.19), a clear dominance of O forecasts is evident during Phase II (Figure 3.21). The SHF5 over-forecast percentages reach 60% at 96 h and 108 h and are around 50% at 72, 84, and 120 h. In contrast to the GFNI performance in the Atlantic, the model consistently has the lowest or second-lowest O percentage (except at 120 h), and has skill relative to SHF5 at all forecast intervals. The GFDI performance is similar to the GFNI in having the lowest or second-lowest O percentage at all forecast intervals excluding 12 h. The statistical-dynamical techniques have a tendency to over-forecast intensity during Phase II in the eastern North Pacific. The SHIPS has the highest O percentage from the 24-h to 60-h forecast intervals in which the technique has over a 58% tendency to over-forecast intensity. After the 60-h forecast interval, the SHIPS technique is only second to the SHF5 with the highest O percentages. The DSHIPS technique O percentages are only slightly lower than for the SHIPS at all forecast intervals. The NHC O percentage is close to the statistical-dynamical techniques (except at 12 h), which suggests the NHC tends to follow the intensity guidance that is more likely to over-intensify a TC in the eastern North Pacific during the initial intensification.

4. Phase II: Summary

For the Atlantic and the eastern North Pacific, the statistical-dynamical techniques SHIPS and DSHIPS have a tendency to over-forecast, while the dynamical models GFDI and GFNI have a tendency to under-forecast the intensity of TCs during the intensification phase. This tendency was the same as in Phase I in both basins. All of the techniques and the NHC have skill relative to the SHF5 for eastern North Pacific storms when comparing the Good (G) intensity trend percentages. When averaged over all of the forecast intervals, the GFDI model has the highest G intensity trend percentages (56.6%), the DSHIPS technique has the next highest average (55.75%), and the NHC is third highest (55.43%) when averaged over all of the forecast intervals. This comparison implies that the NHC is following their intensity guidance and does not add value.

The NHC does add value in the Atlantic as it has the highest G intensity trend percentage (63.68%) when averaged over all of the forecast intervals. The DSHIPS technique G intensity trend percentage is second at 61.34%, and the SHIPS technique is next with 59.11%. All of the techniques had skill relative to SHF5. In contrast to the eastern North Pacific, the GFDI model has the second to lowest G intensity trend percentage at 51.16%, which was only slightly above the SHF5 with 51.03%.

5. Phase II: Comparison with the Western North Pacific

The comparable statistical-dynamical techniques continue to be the better performers in all three basins. Whereas the STIPS technique in the western North Pacific was the best performer of all the techniques (Blackerby 2005) during Phase II, the closely related DSHIPS was either the best or second best technique in the Atlantic and eastern North Pacific. Blackerby (2005) stated that the GFNI had the third best G intensity trend percentage (~66%) during Phase II in the western North Pacific when averaged over all of the forecast intervals. This model has the second lowest G intensity trend percentage (51.4%) in the eastern North Pacific and third lowest (53.9%) in the Atlantic. The JTWC and NHC forecast out-performed guidance from every technique during Phase II in the western North Pacific and the Atlantic, respectively, but the NHC was second to the DSHIPS technique in the eastern North Pacific. The climatology and persistence techniques in both the Atlantic and the eastern North Pacific have the worst G intensity trend percentages of all the techniques when averaged over all of the forecast intervals. By contrast, the ST5D technique had the second-best G intensity trend percentage for Phase II forecasts in the western North Pacific.

6. Peak Intensity

An important aspect to consider in Phase II is the magnitude and timing of the tropical storm peak intensity. Tropical storms are further from their peak intensity than hurricanes and are better organized than tropical depressions and thus have more potential for more rapid intensification. Therefore, it is of interest whether the intensity guidance techniques can accurately predict the peak intensity at the end of Phase II.

To examine the techniques performance in forecasting peak intensity, the average intensity errors for the various techniques for the forecasts preceding the time of maximum intensity were computed to quantify intensity guidance reliability. Using the 2003-2004 database for the 37 storms in the Atlantic and 32 in the eastern North Pacific, the date-time-group (DTG) of the first peak intensity was determined for the series of intensity predictions verifying at this time of peak intensity (-120 h, -96 h,….-24 h). The averages of all predictions of peak intensity minus the actual intensity for the various forecast intervals that could be verified are determined for the Atlantic and the eastern North Pacific. A non-homogeneous sample is used to maximize the sample sizes. However, the sample size is 25 cases or less for each intensity guidance technique and the NHC per forecast interval in both basins.

a. Atlantic

All of the intensity guidance techniques under-forecast peak intensity (Figure 3.22). This pattern of under-forecasting is evident even 24 h before peak intensification. Only the SHIPS technique with an average error of -9 kt has skill relative to the SH5F at the 24-h forecast interval. The GFNI model has the highest average error of 15 kt at the 24-h forecast interval. The average error increases 48 h before peak intensity to around 18 kt for all of the techniques and the NHC. Again, only the SHIPS technique has skill relative to the SH5F, and the GFNI model has the highest average error at the 48-h forecast interval. Although the average error continues to increase (~ 30 kt) at the 72-h forecast interval, all of the techniques and the NHC have skill relative to the SH5F. As stated earlier, the sample sizes become small as few 72-h to 120-h forecasts prior to peak intensity are validated, and therefore minimum confidence should be placed in these calculations. However, it is evident that all of the techniques under-forecast the tropical storm/hurricane maximum intensity and thus provide the NHC with little to no guidance as to the peak intensity.

b. Eastern North Pacific

The pattern of under-forecasting intensity prior to peak intensity in the Atlantic is also evident in the eastern North Pacific among all the techniques and the NHC (Figure 3.23). The errors 24 h before peak intensity are under-forecast by at least 10 kt by all of the techniques and the NHC. By 48 h prior to peak intensity, the under-forecasting errors are even larger (at least -25 kt) by all the techniques and the NHC. That is, an intensity guidance technique in the eastern North Pacific may forecast that the TC will become a strong tropical storm within the next 48 h when it will actually be a category two hurricane (Saffir-Simpson Hurricane Scale). Still, the NHC forecasters are able to add value to the intensity guidance as the NHC has the lowest error at 24 h and 72 h and has skill relative to the SHF5 technique at all forecast hours. The SHIPS and DSHIPS errors are identical at all forecast hours and have the lowest error at 24 h, with skill relative to the SHF5 at 24 h and 48 h. However, the SHIPS and DSHIPS errors increase dramatically at 72 h (-60 kt). The GFDI and GFDI models have the highest errors at 24 h and 48 h and have no skill relative to the SHF5 technique. The eastern North Pacific errors verifying at peak intensity are only through 72 h as no forecasts of peak intensity 96 h or 120 h were available for validation.

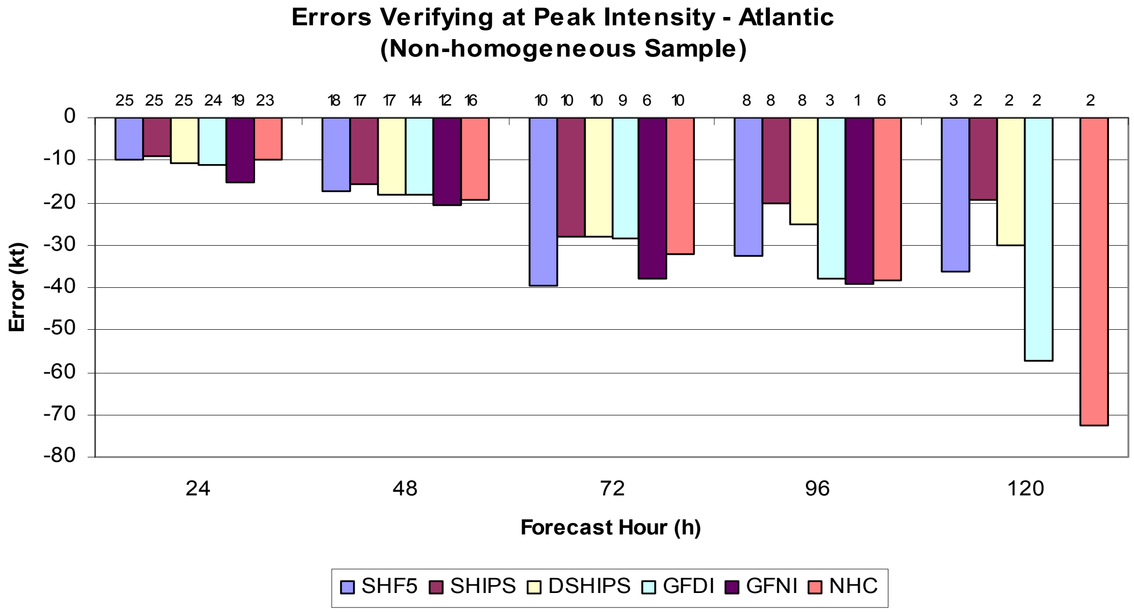


Figure 3.22 Error (kt) in the intensity forecasts verifying at the date/time of the first peak intensity of each of the 37 TCs during the 2003 and 2004 Atlantic hurricane seasons by the various techniques and the NHC. The small numbers above each bar indicate the number of cases.

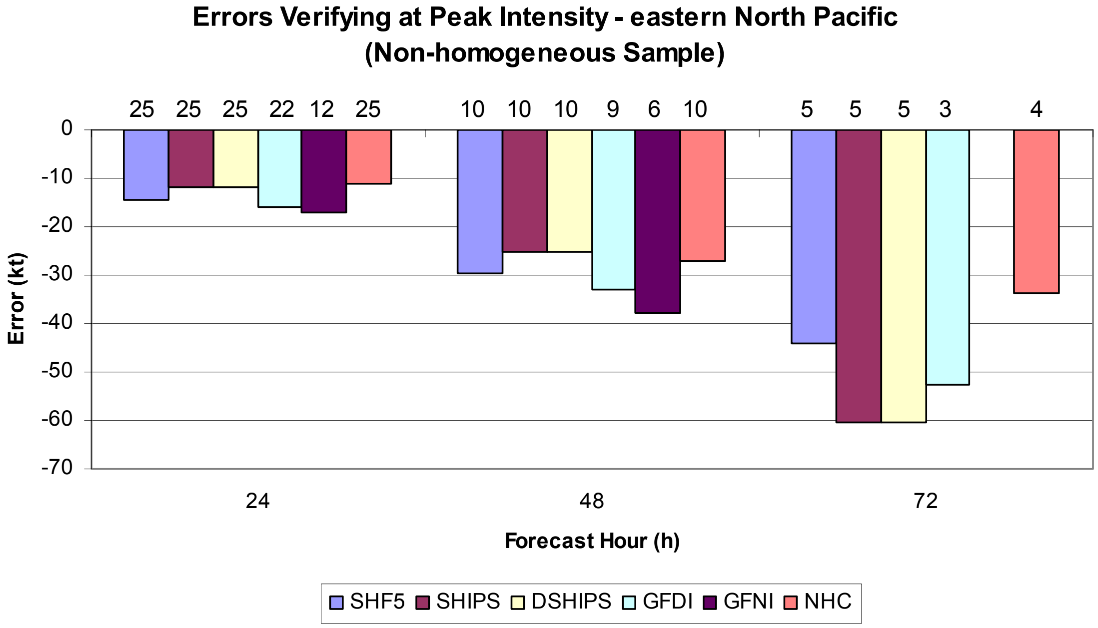


Figure 3.23 Same as Figure 3.22, except for each of the 32 eastern North Pacific TCs during the 2003 and 2004 hurricane seasons. The small numbers above each bar indicate the number of cases.

c. Comparison with the Western North Pacific

As in the Atlantic and eastern North Pacific, Blackerby (2005) found a distinct trend of under-forecasting prior to peak intensity in the western North Pacific with the under-forecasting errors becoming larger the longer the forecast interval prior to peak intensity. Whereas the GFNI model in the Atlantic and the eastern North Pacific under-forecast error is consistently among the highest of all the techniques, the GFNI model in the western North Pacific had smaller under-forecast errors relative to the other techniques. The comparable statistical-dynamical techniques STIPS and SHIPS have the lowest error at 24 h and 48 h among all of the techniques for all three basins. This pattern continues for the STIPS and SHIPS techniques in the Atlantic and western North Pacific from 72 h to 120 h before peak intensity.

7. Rapid Intensification

Another important aspect of TC intensification is rapid intensification. Kaplan and DeMaria (2003) define rapid intensification of TCs in the Atlantic basin as a maximum sustained surface wind speed increase of 30 kt over a 24-h period. Kaplan and DeMaria also noted that approximately 31% of all TCs in the Atlantic basin between 1989 and 2000 underwent rapid intensification at least once during their lifetimes. Because of the frequent occurrence of rapid intensification, the intensity guidance techniques and the NHC are evaluated in both the Atlantic and the eastern North Pacific on their capability to predict rapid intensification.

The ‘best track’ intensity values from the NHC ATCF-format files were used to determine which TCs underwent rapid intensification during the 2003 and 2004 tropical seasons for both basins. The intensity forecasts from the NHC aids files for the various techniques and for the NHC of the TCs that underwent rapid intensification were examined for all times +/- 12 h relative to the actual time of rapid intensification. That is, if a TC intensity at DTG ‘A’ is 35 kt and increases 24 h later to 65 kt, the intensity techniques forecast starting at DTG ‘A’ should increase 30 kt or more within the next 48 h. If the technique’s predictions matched or exceeded the rapid intensification threshold (30 kt/day) this was regarded as a ‘hit.’ However, if the technique prediction did not meet the threshold at the correct time or within +/- 12 h, it was recorded as a ‘miss.’

a. Atlantic

In the Atlantic, 10 of the 37 cases underwent rapid intensification during 2003 and 2004 (Figure 3.24). That is, approximately 27% of the TCs during 2003 and 2004 underwent rapid intensification, which agrees closely with the 31% found by Kaplan and DeMaria (2003) between 1989 and 2000. All the techniques forecast two or three of the 10 rapid intensifications in the Atlantic, and thus have a tendency to under-forecast the rapid intensity changes. That is, the intensity guidance on average will miss rapid intensification in seven of 10 TCs in the Atlantic. Based on this guidance, the NHC was only able to forecast three cases of rapid intensification.

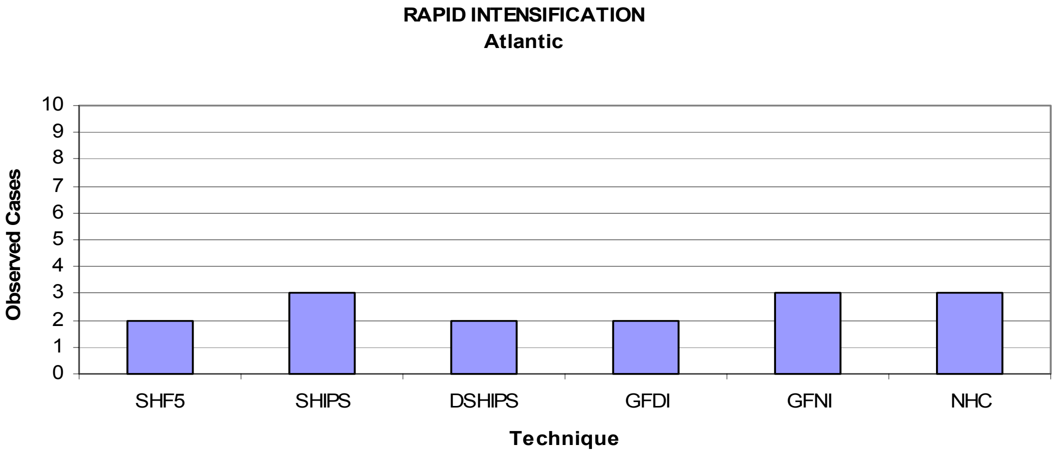


Figure 3.24 The number out of 10 cases that a forecast technique or NHC predicted an observed period of rapid intensification at the correct time or within +/- 12 h for the Atlantic basin during the 2003 and 2004 seasons.

b. Eastern North Pacific

Misses of rapid intensification events were also common in the eastern North Pacific during 2003 and 2004 (Figure 3.25). The GFDI model forecast four of the 11 cases of rapid intensification, which was the best performance among the intensity guidance techniques. The two statistical-dynamical techniques did not provide useful guidance in forecasting rapid intensification with either one or zero correct predictions of the 11 observed cases of rapid intensification. The NHC forecast four of the 11 rapid intensification events, which is the same as their best intensity guidance. This evaluation suggests that the NHC has little skillful intensity guidance in both basins, and their performance is only slightly better than SHF5.

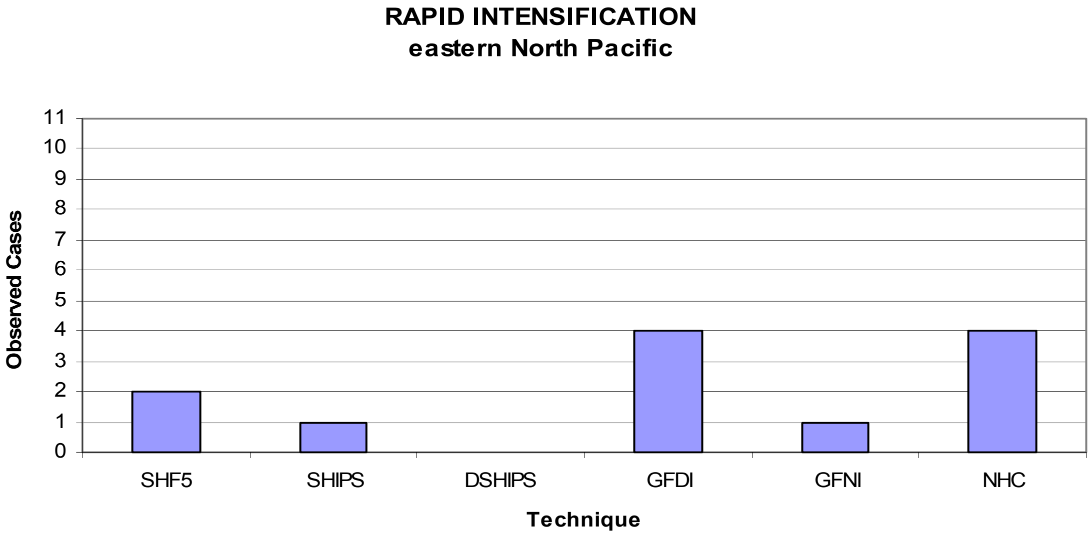


Figure 3.25 Same as Figure 3.26, except for the eastern North Pacific during 2003 and 2004.

c. Comparison with the Western North Pacific

The GFNI model predicted 10 of the 28 rapid intensification events in the western North Pacific (highest among all techniques) with a tendency to over-predict intensification (Blackerby 2005). Although the GFNI model in the Atlantic was one of three intensity guidance techniques with the most predictions of rapid intensification (3 of 10), the model slightly under-predicts the intensification. In contrast, the GFNI model performance in the eastern North Pacific was poor as only all one case of rapid intensification was predicted within +/- 12 h.

The statistical-dynamical techniques STIPS and SHIPS had poor performances in both the eastern and western North Pacific (predicted only one of 28 cases of rapid intensification), but were somewhat better in the Atlantic in that they predicted three of 10 rapid intensifications. Whereas the SHF5 techniques in the eastern North Pacific and the Atlantic predicted two rapid intensifications events, the comparable ST5D technique in the western North Pacific did not predict any of the 28 events of rapid intensification.

8. Intensity Change Distribution during Phase II

A skillful intensity guidance technique will have the distribution of intensity changes similar to the distribution of observed intensity changes. Thus, the analysis of how the techniques forecast intensity changes within a certain time frame will indicate their range of forecast intensifications and give insight as to why the techniques miss the majority of rapid intensification cases as noted in the previous section. Although other time intervals could be described, the intensity distribution here is for the first 48 h after the initial intensity of 35 kt (beginning of Phase II), which is the time period when a TC is more likely to have a variety of intensity changes (i.e., intensification, rapid intensification, or decay). Therefore, the observed distribution of 48-h intensity changes from the beginning of Phase II (initial intensity of 35 kt) will be compared with the predicted distributions by the various techniques in the Atlantic (Figure 3.26) and the eastern North Pacific (Figure 3.27).

a. Atlantic

The observed intensity changes during the first 48 h of Phase II have a wide range of values from -15 kt to 65 kt. The majority of the intensity changes (66%) are positive. That is, the first 48 h after a TC reaches tropical storm strength it will usually intensify in the Atlantic. A cluster of intensity changes (44%) are between 5 kt and 40 kt, and another 22% between 50 kt and 65 kt. However, approximately 25% of the TCs have decayed after the first 48 h of Phase II and 7% have no change

Nearly 97% of the SHF5-predicted 48-h intensity changes are between 5 kt and 35 kt, with 27% at 15 kt and 24% at 25 kt (Figure 3.26a). This narrow distribution indicates that the SHF5 technique does not provide guidance

48-h Intensity Change Forecast for TCs with Initial Intensity of 35 kt in Phase II - **ATLANTIC**

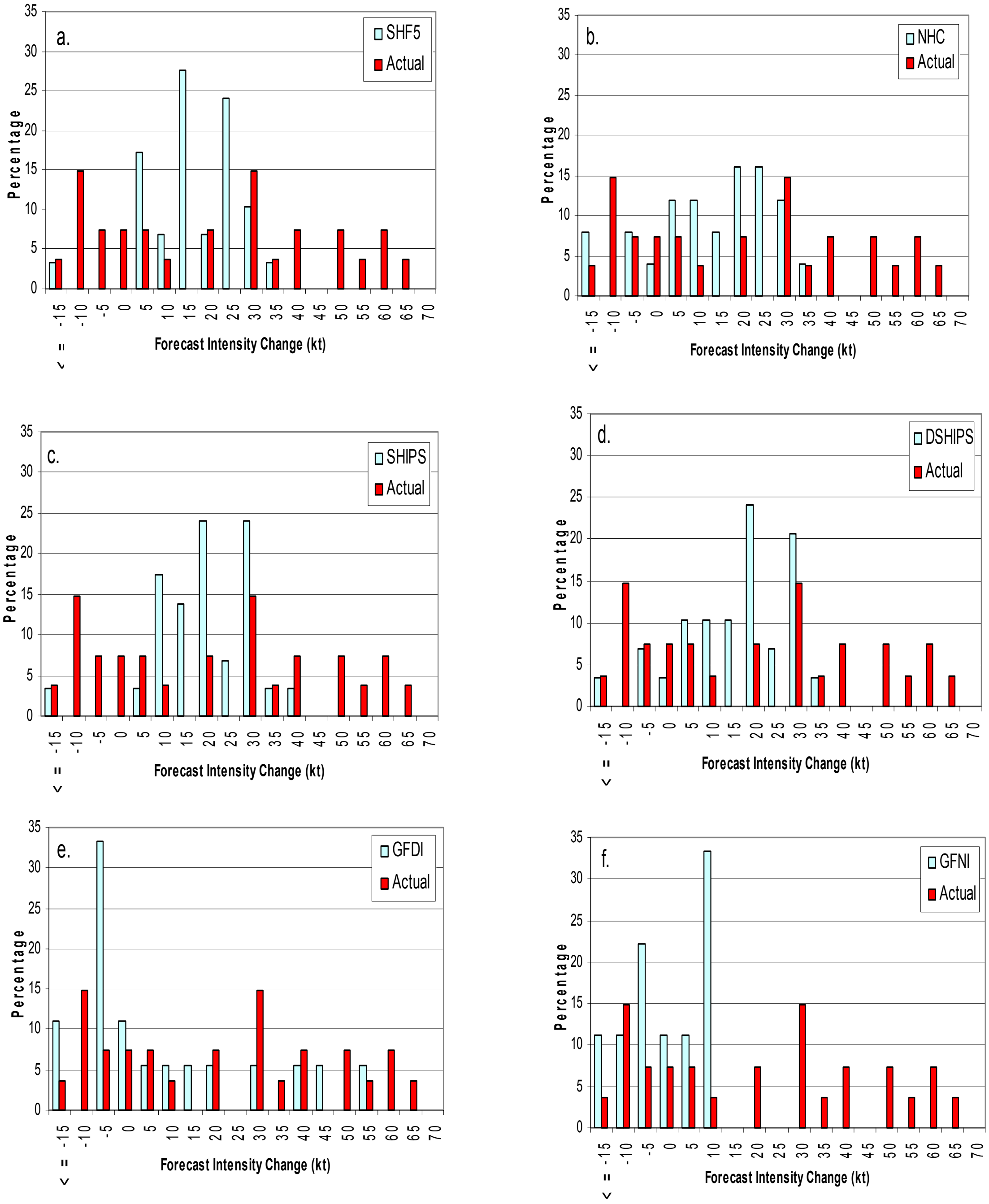


Figure 3.26 (a-f) Observed 48-h intensity changes (solid bars) during Phase II and the predicted 48-h intensity changes (blue bars) by the various techniques and by the NHC for the 2003 and 2004 Atlantic seasons.

on larger intensity changes during the first 48 h in Phase II, and the technique has a bias toward smaller intensity changes. Also, the SHF5 technique has difficulty in predicting the non-intensification or decay of a TC when it is observed in 25% of TCs after the beginning of Phase II. This distribution suggests that the SHF5 may predict that a TC is at tropical storm strength when it is actually is a tropical depression.

The largest 48-h intensity change forecast by the SHIPS is 40 kt (Figure 3.26c). Similar to the SHF5 technique, the majority of the intensity change distribution (93%) is between 5 kt and 35 kt and the technique rarely decays a tropical storm during the 48 h after the intensity is 35 kt. Thus, the technique under-intensifies TCs that are rapidly intensifying and continues to intensify a TC when it is not changing intensity or even decaying.

The DSHIPS technique (Figure 3.26d) has a very similar 48-h intensity change distribution as the SHF5 and SHIPS techniques, i.e., it has a bias toward smaller 48-h intensity changes. The technique has 86% of the intensity changes between 5 kt and 40 kt. However, the DSHIPS does capture more of the decay events than do the SHF5 and SHIPS techniques during the first 48 h after reaching tropical storm strength.

Whereas the SHF5, SHIPS, and DSHIPS techniques have a limited range of intensity changes, the GFDI model predicts 48-h intensity changes from -30 kt to +55 kt (Figure 3.26e). These ranges are similar to the actual 48-h intensity change distribution of -15 kt to +65 kt. Although the model distribution is centered at -5 kt with 44% of its intensification rates between -5 kt and 0 kt, the GFDI model has the ability to predict some of the larger intensity changes as well as the cases in which the TC had decayed.

The GFNI 48-h model intensity changes only range between -15 kt and 10 kt from a small sample size (Figure 3.26f), which indicates that the model has little skill in the early intensification of Phase II. Indeed, the GFNI has a tendency to decay a TC in Phase II. If the TC does intensify during this period, the intensity change forecast by the GFNI model is likely to be much too small, or may even forecast the TC to decay. Thus, the GFNI model under-intensifies the majority of TCs during the first 48 h after a TC reaches 35 kt.

The NHC distribution of 48-h intensity changes resembles that of the statistical and statistical-dynamical techniques (Figure 3.26b). Approximately 80% of the intensity changes are between 5 kt and 35 kt, which are similar to the SHIPS and DSHIPS intensity changes. The NHC intensity change distribution has a more limited range from -15 kt to 35 kt relative to the observed distribution, and therefore has a bias toward smaller 48-h intensity changes during the beginning of Phase II. Thus, the NHC distribution of intensity change forecasts is too narrow, as is the intensity change distribution from many of its intensity guidance techniques. In particular, the NHC does not predict the ≥ 40 kt intensity changes in 48 h.

b. Eastern North Pacific

Although a surprising number (23%) of 48-h intensity changes are decreases of -5 kt, the observed intensity changes during the first 48 h of Phase II is positively skewed in the eastern North Pacific (Figure 3.27). Whereas the distribution of intensity changes is somewhat Gaussian between 0 kt and 30 kt, outliers of rapid intensity changes from 40 kt to 70 kt also are observed. That is, nearly a quarter of the TCs in the eastern North Pacific either decays slightly during the first 48 h as a tropical storm, or has rapid intensity changes.

The SHF5 technique 48-h intensity change distribution is generally similar to the observed distribution in that it is positively skewed with values as large as 45 kt (Figure 3.27a). However, the SHF5 does not forecast a TC to decay during the first 48 h at the beginning of Phase II, which occurred in 23% of the TCs in the eastern North Pacific. The SHF5 technique intensity forecasts do tend to cluster at smaller positive increases with 75% of the 48-h intensity change forecasts between 0 kt and 25 kt. These smaller intensity changes for the SHF5 arise because a climatology and persistence technique is unable to provide guidance on the large intensity changes.

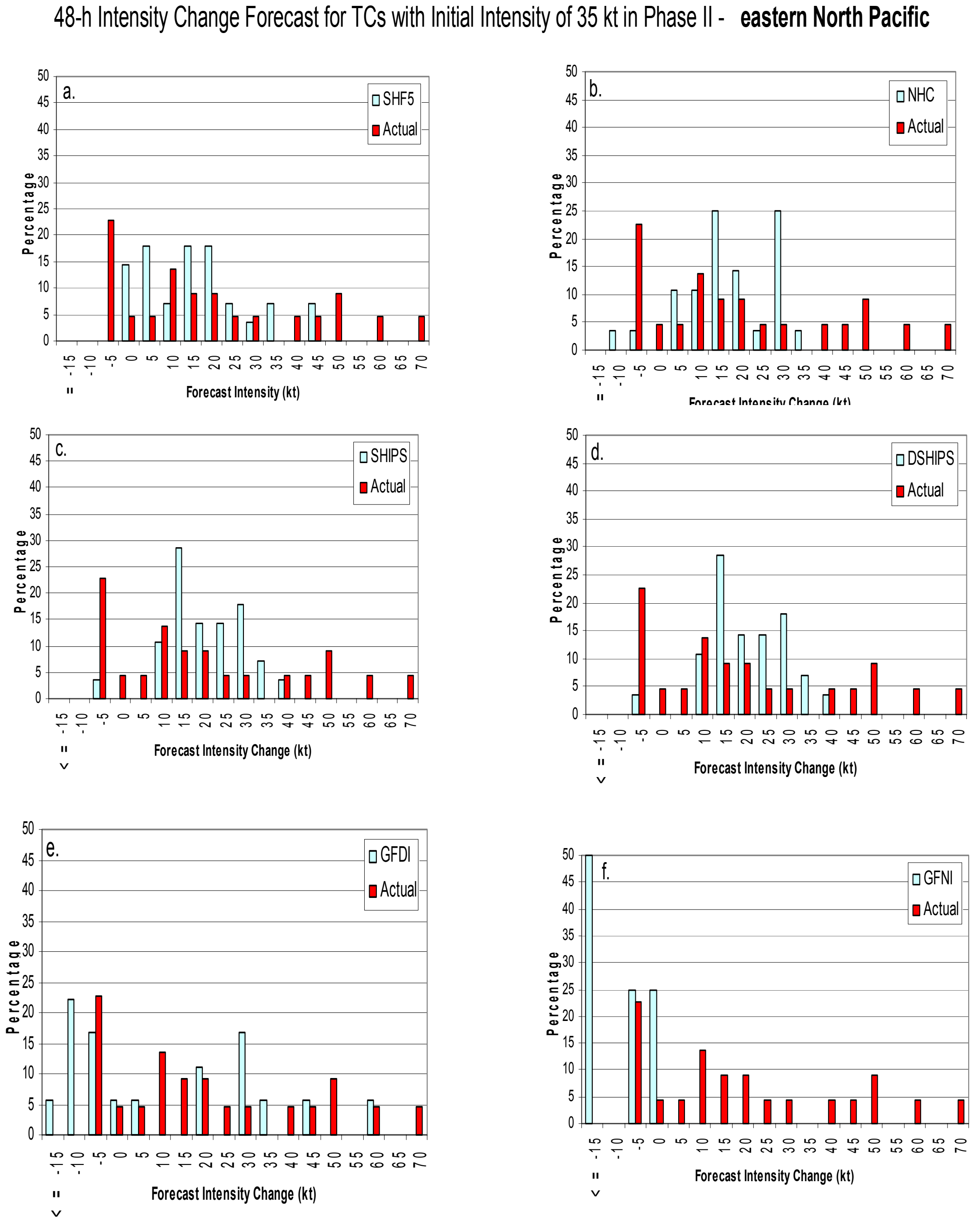


Figure 3.27 (a-f) Same as Figure 3.26 (a-f), except for the 2003 and 2004 eastern North Pacific seasons.

The SHIPS and DSHIPS techniques have identical 48-h intensity change distributions in the eastern North Pacific (Figures 3.27c-d). These techniques do not forecast an intensification of more than 40 kt in 48 h at the beginning of Phase II and thus missed the higher intensification changes (above 45 kt) that occurred in 27% of the observed events. Whereas these techniques have 75% of their intensity changes between 15 kt and 30 kt, the observed 48-h intensity change distribution has only 27% between 15 kt and 30 kt. Thus, the SHIPS and DSHIPS techniques do not provide good guidance for either the non-intensifiers/weak decayers or the largest intensification rates during Phase II.

The GFDI model was able to forecast rapid intensifications during the beginning of Phase II (Figure 3.27e). It was also able to forecast the slight decay of TCs in the eastern North Pacific after first reaching tropical storm strength. However, the GFDI model decayed TCs too strongly and too often with 27% of the intensity change distribution between at -10 kt and -15 kt.

As noted in the Phase II trend summary (Chapter III. D), the GFNI model under-forecast intensity in the eastern North Pacific. Keeping in mind that the GFNI model has a smaller sample size than the other techniques, the pattern of under-forecasting is apparent in the initial 48 h of Phase II with no intensity changes by the GFNI model above 0 kt (Figure 3.27f). Thus, the GFNI model has no skill in the early intensification of Phase II.

Similar to the SHIPS and DSHIPS techniques, most (71%) of the NHC 48-h intensity forecasts in the eastern North Pacific are between 15 kt and 30 kt (Figure 3.27b). The maximum intensity change forecast by NHC is only 35 kt, which is 5 kt lower than the SHIPS and DSHIPS techniques. This distribution suggests that the NHC follows its statistical-dynamical intensity guidance that is unable to provide guidance on the larger intensity changes and there is no value added relative to the SHIPS technique.

c. Intensity Change Distribution Summary

The SHF5, SHIPS, and DSHIPS techniques have 48-h intensity change distributions that have too-narrow distributions with a bias toward small

48-h intensity changes in both basins. While the maximum 48-h intensity change that is forecast by the three techniques is only 45 kt, intensity changes of 50 kt or greater are observed in at least 18% of the cases in both basins. The GFNI model had a much smaller sample size than the other three techniques, but the trend of under-forecasting during Phase II is evident in the distribution of intensity changes. Thus, these techniques provided to the NHC do not predict the typical distribution of intensification rates or the largest values. The GFDI model does have a wider range of 48-h intensity changes that was closer to the observed 48-h intensity change distribution in both basins. However, the GFDI model has a tendency to decay TCs too much during the first 48 h at the beginning of Phase II in both basins.

d. Comparison with the Western North Pacific

These climatology and persistence techniques and the statistical-dynamical techniques intensity change distributions in the Atlantic and eastern North Pacific were very similar to those found by Blackerby (2005) in the western North Pacific. The comparable climatology and persistence techniques ST5D and SHF5 have positively skewed peaks and the intensity change distribution is too narrow relative to the observed distribution. In the eastern and western North Pacific respectively, the SHF5 and ST5D techniques intensity change distributions were constrained between 0 and 45 kt, and the SHF5 for the Atlantic had only one intensity forecast that decayed a TC and one maximum intensity change of 35 kt. The comparable STIPS and SHIPS also had positively skewed peaks with a too-narrow intensity change distribution compared to the observed distribution. The intensity change forecasts by these statistical-dynamical techniques in all three basins had a maximum of 40 kt and only a few cases of decaying a TC during the first 48 h in Phase II. The GFNI model in the western North Pacific had a wider (-10 kt to 50 kt) intensity change distribution than the GFNI forecasts used in the Atlantic (-15 kt to 10 kt) and in the eastern North Pacific (-30 kt to 0 kt). It is noteworthy that the GFNI model in the Atlantic and the eastern North Pacific had similar tendencies toward a very limited intensity change range and very negatively skewed peaks.

9. Average and Standard Deviation of the Intensity Change Distribution

Another summary-type representation of the 48-h intensity change distribution during Phase II is shown in Figure 3.28 (Atlantic) and Figure 3.29 (eastern North Pacific). Both the mean and +/-1 standard deviation of the intensity change distributions for each technique and the NHC are compared with these values for the observed distribution. These summaries illustrate in a simple way the biases and ranges of the techniques and the NHC discussed in the previous sections.

a. Atlantic

For the Atlantic (Figure 3.28), the more limited range of the intensity change distribution for all of the techniques (except the GFDI model) and the NHC relative to the actual distribution is quite evident. Specifically, the GFNI model range is too narrow and the bias is large because the mean intensity is too low. The GFDI model does have a range of intensity changes that is comparable to the actual intensity changes, but has a clear bias of under-estimating the mean intensification rate from the beginning of Phase II. In contrast, the SHF5, SHIPS, and DSHIPS have a small (within seven kt) bias of under-estimating the mean intensity changes. However, the range is much too narrow. Therefore, these techniques will miss cases of rapid intensification and any cases of decay. The NHC mean intensity change is between the statistical-dynamical techniques and dynamical models and the range is too narrow, which means it under-estimates rapid intensification during Phase II.

b. Eastern North Pacific

The mean 48-h intensity change during Phase II in the eastern North Pacific (Figure 3.29) is similar to the Atlantic, i.e., all of the intensity guidance techniques (except again GFDI model) and the NHC have a limited intensity change distribution and therefore are unable to forecast cases of rapid intensification as well as some decays over the 48-h period. The GFNI model average intensity change has almost no overlap with the actual intensity change distribution, which clearly demonstrates the inability of the GFNI model in

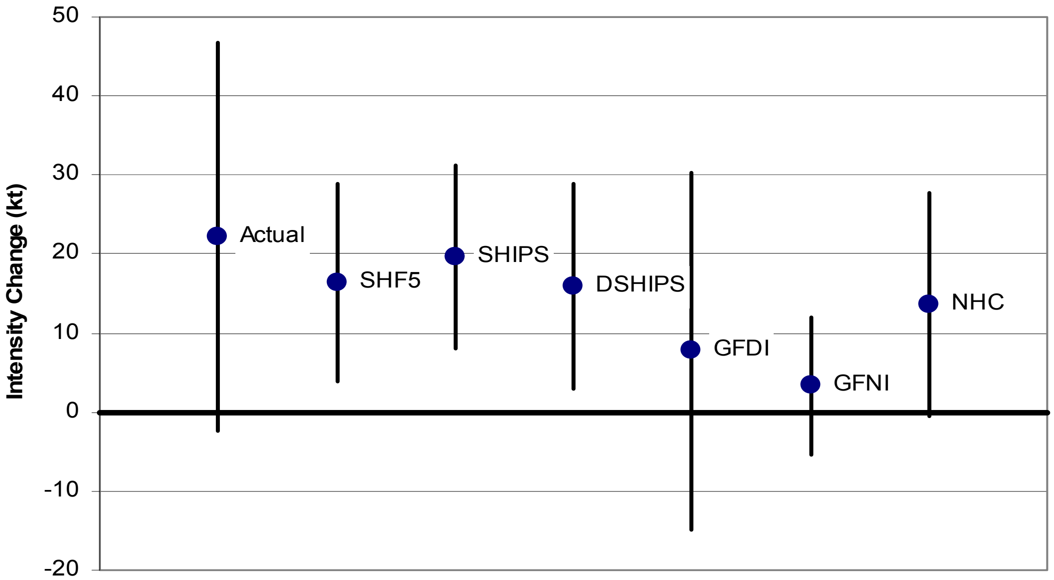


Figure 3.28 Actual and predicted 48-h intensity change distributions during Phase II for 37 TCs during the 2003 and 2004 Atlantic seasons beginning with an intensity of 35 kt. The heavy dot represents the mean intensity change value (kt) while the length of the sticks represents plus or minus one standard deviation about the mean.

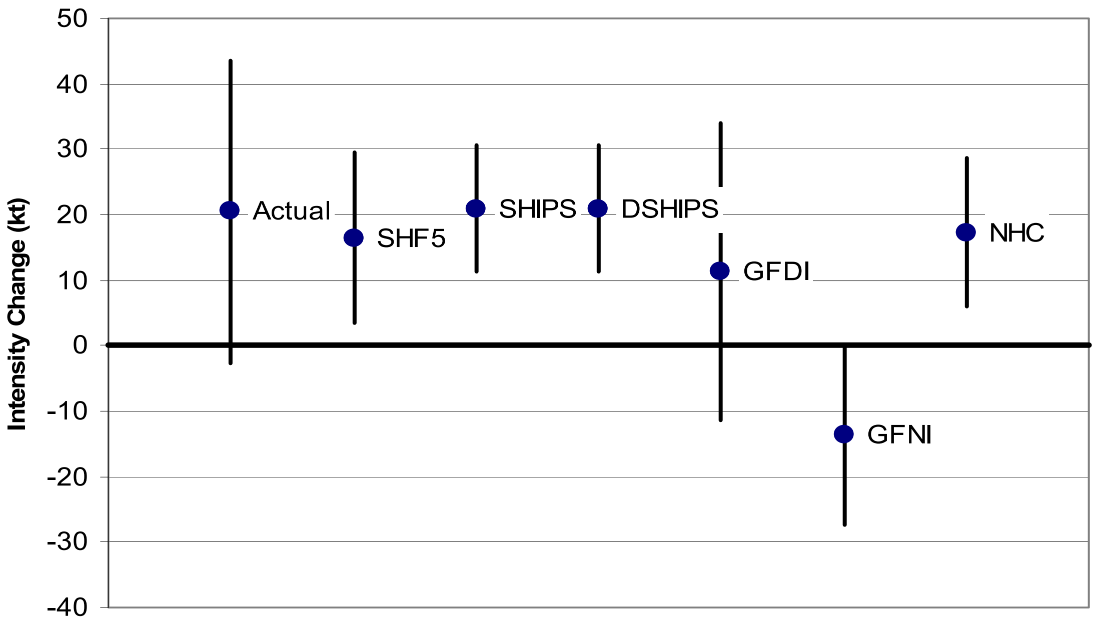


Figure 3.29 Same as Figure 3.28, except for the 32 TCs in the eastern North Pacific during the 2003 and 2004 seasons.

predicting intensity changes during the beginning of Phase II. It is noted that the statistical-dynamical techniques SHIPS and DSHIPS mean intensity change value (20.9 kt) is slightly above the actual intensity change (20.4 kt). Surprisingly, the intensity ranges for these techniques are smaller than for the climatology and persistence technique SHF5.

c. Comparison with the Western North Pacific

Blackerby (2005) had found similar mean 48-h intensity changes during Phase II in the western North Pacific. The GFNI model has a low intensity change bias and too-narrow ranges of intensity change forecasts in all three basins. The SHIPS and STIPS techniques both have very limited intensity ranges, but the mean intensity change is the closest to the actual mean intensity change among the various techniques in all three basins. In the Atlantic and western North Pacific, the SHIPS and STIPS techniques have a bias of under-estimating the rates of intensification during the beginning of Phase II.

10. Decay and Reintensification Cycles

A common occurrence among the more intense TCs is an intensity decay and subsequent reintensification, especially during eyewall replacement cycles. Such a reintensification cycle can happen more than once during the lifecycle of a TC. Thus, intensity guidance techniques should have skill in forecasting the decay and reintensification cycles that may occur in a TC. In this study, the decay-reintensification cycle corresponds to Phase IIa of the intensity framework (Figure 2.2), and is defined as a reintensification of 10 kt or greater following an initial decay cycle of 10 kt or greater.

The predictions by each technique were examined for all times +/- 12 h relative to the actual time of the decay and reintensification cycle. If the prediction by any of the techniques reasonably matched the phase (timing) and magnitude of intensity oscillations, it was regarded as a 'hit.' If none of the predictions within +/- 12 h reasonably matched the phase and magnitude of intensity oscillations, the event was recorded as being a 'miss.' Since many of the TCs had more than one decay-reintensification cycle, a 'hit' is recorded if the technique or the NHC was able to predict at least one of the cycles. The eastern North Pacific only had two occurrences of secondary peaks during the 2003 and 2004 hurricane seasons and is not evaluated due to this inadequate sample size.

a. Atlantic

Ten cases of decay-reintensification cycles occurred in the Atlantic basin during the 2003 and 2004 seasons (Figure 3.30), with five of the 10 cases having multiple decay/reintensification cycles. The best performance among the techniques and the NHC was the GFNI model, although it only forecast two of the 10 cases. Whereas the DSHIPS, GFDI, and the NHC predicted one of 10 decay-reintensification cycles, the SHIPS and SHF5 did not forecast any of the cycles. The failure by the climatology and persistence technique SHF5 could be expected as the climatological average would not include such short-term, rapid intensity changes. However, the three-dimensional dynamical models also did not predict the decay-reintensification cycles within +/- 12 h of occurrence. An example of the lack of skill during multiple decay-reintensification cycles is displayed in Figure 3.31. Thus, the NHC intensity guidance techniques have no skill in forecasting decay-reintensification cycles and the NHC adds no value.

b. Comparison with the Western North Pacific

The GFNI model forecast three of 12 decay-reintensification cycles in the western North Pacific, which was the highest among all of the techniques (Blackerby 2005). Thus, the GFNI performed better relative to the other techniques in both the Atlantic and the western North Pacific. Whereas the STIPS technique forecast one of 12 cases in the western North Pacific, the corresponding SHIPS technique did not forecast any of the decay-reintensification cycles in the Atlantic. The climatology and persistence techniques in both basins did not predict any of the cycles.

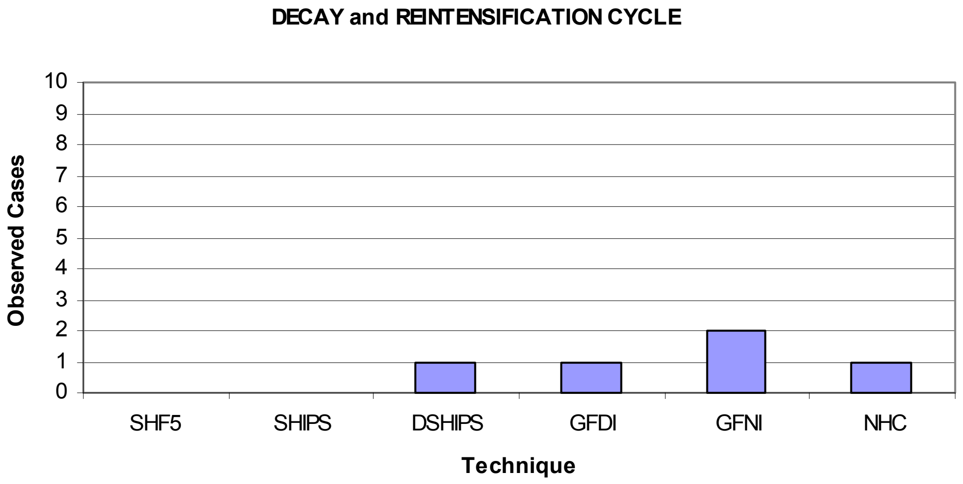


Figure 3.30 Number of cases (out of 10) a forecast technique or the NHC adequately predicted an observed decay and reintensification cycle in the Atlantic during the 2003 and 2004 seasons.

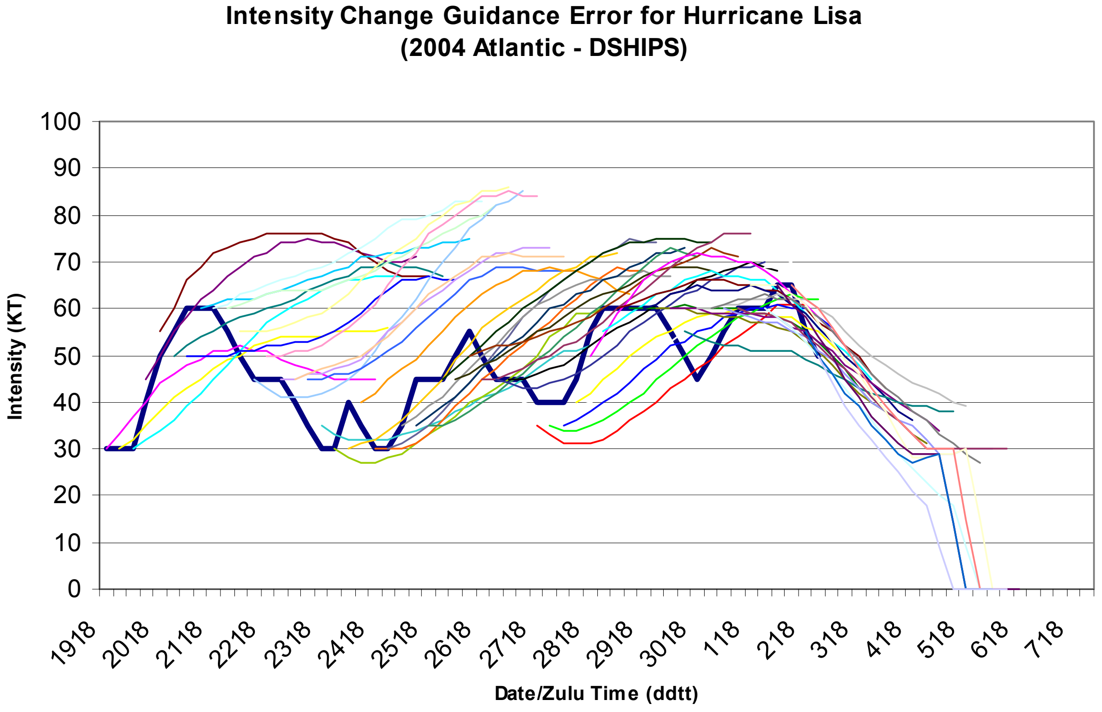


Figure 3.31 Decay Statistical Hurricane Intensity Prediction Scheme (DSHIPS) guidance for Hurricane Lisa (Storm #13) during 2004. The heavy blue line represents the observed (best-track) storm intensity. Colored lines illustrate intensity forecasts each 6 h.

E. PHASE III: DECAY

The last stage in the life cycle model (Figure 2.2) of a TC is the decay phase. In this study, the decay phase is defined from peak intensity through final decay. However, if the storm had a decay and reintensification cycle, i.e., Phase IIa, the decay phase is from the DTG of the secondary (last) peak intensity until final decay. The same contingency table (Figure 3.9) is used to evaluate the performances of the intensity guidance techniques and the NHC in the intensity changes during decay. The sample size for all the techniques and NHC from the 72-h to 120-h forecast interval during Phase III was 50 or less in the Atlantic and eastern North Pacific. Therefore, the trend percentages during the 72-h to 120-h forecast interval may not represent the true performance of the techniques and the NHC.

1. Phase III: Good (G) Intensity Trends

a. Atlantic

The SHF5 technique had G percentages of less than 50% at 24 h and 36 h (Figure 3.32). That is, 24 h to 36 h after peak intensity, the SHF5 is able to forecast the intensity decay rate within +/- 10 kt less than 50% of the time. After the 60-h forecast interval, the SHF5 had G percentages of 89% or higher. By contrast, the DSHIPS technique had G percentages of 70% or higher during the entire forecast interval. The DSHIPS G percentages were consistently higher than for the SHIPS technique and were 25% or higher than SHIPS at the 60-, 72-, and 84-h forecast intervals. This improvement over SHIPS demonstrates the usefulness of adding the decay-over-land aspect in the DSHIPS. The DSHIPS has skill relative to the SHF5 at all forecast intervals. The NHC had the highest G percentages at 12 h and 24 h and 100% G percentages at 96 h and 120 h (small sample size). The NHC had skill relative to SHF5 at all forecast intervals except 72 h. The GFNI model has G percentages of 60% or higher at all forecast intervals and is skillful relative to the SHF5 except at 72 h and 84 h. Although the GFDI model has no skill relative to the SHF5 at the 12-, 72-, and 84-h forecast intervals, it has G percentages of 66% or higher during the entire forecast interval.

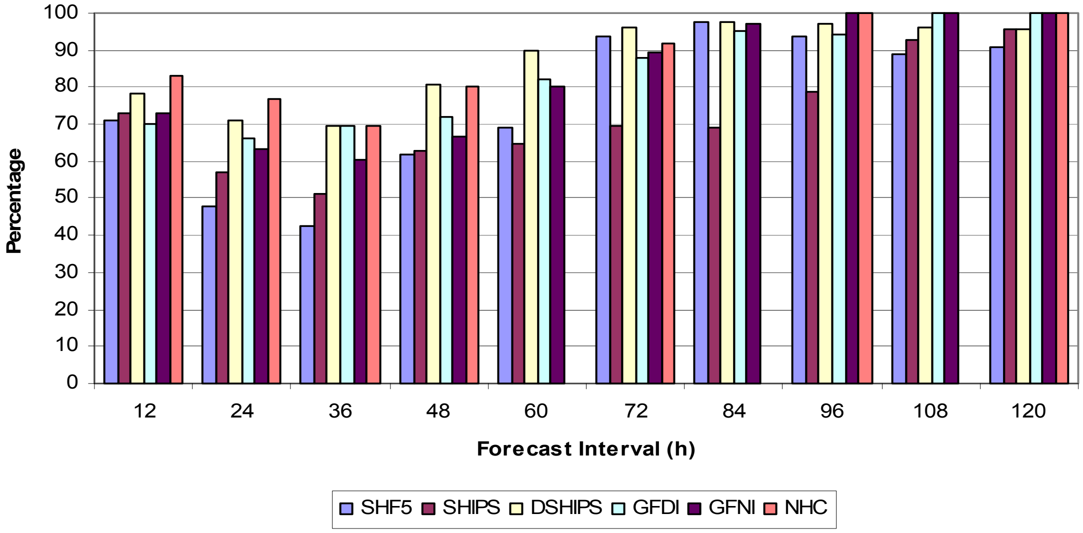


Figure 3.32 Percentages of Good intensity trends during the decay phase (Phase III) for the combined 2003 and 2004 Atlantic seasons. A Good forecast during decay indicates that the magnitude of the forecast intensity is within +/- 10 kt of the actual intensity. Sample sizes of the verified forecasts range from over 1200 at 12 h to 190 by the 96-h interval, except that the dynamical models have smaller samples.

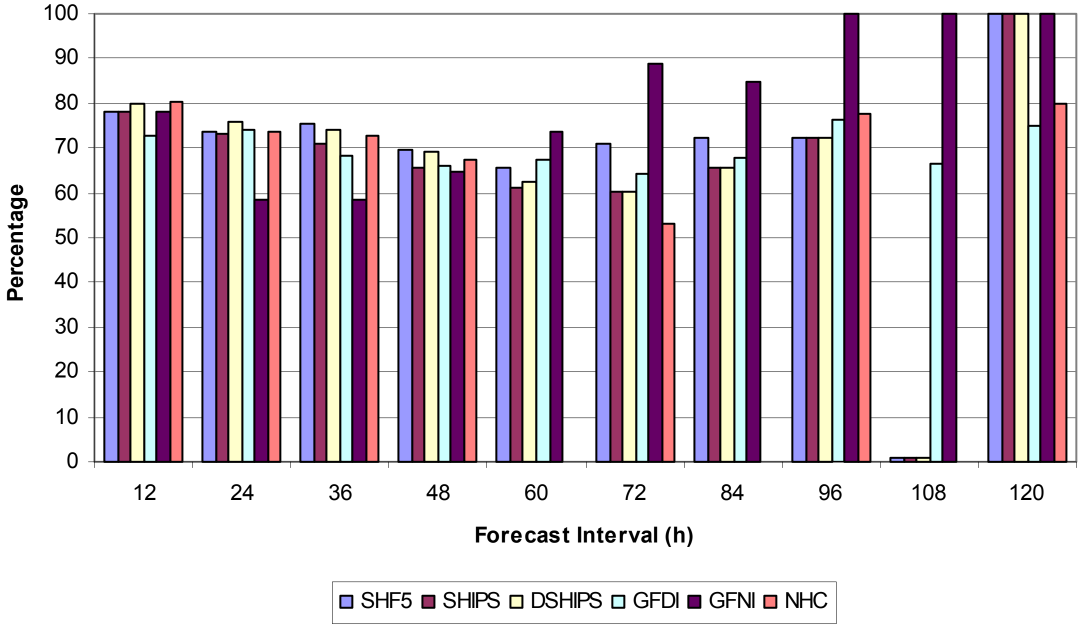


Figure 3.33 Same as Figure 3.32, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 1400 at 12 h to 95 by the 96-h interval.

b. Eastern North Pacific

The SHF5 technique has high G percentages of at least 70% during the entire forecast interval of Phase III except at the 60-h forecast interval (66%) (Figure 3.33). All of the techniques and the NHC have G percentages of 72% or higher at 12 h and maintain G percentages above 65% (excluding the GFNI at 24 h and 36 h) through the 48-h forecast interval. However, none of the techniques or the NHC has skill relative to SHF5 at the 36-h and 48-h forecast intervals and only DSHIPS has skill relative to SHF5 at 24 h. Only the two dynamical models have skill relative to SHF5 at the 60-h forecast interval. Although the GFNI model G percentages increase to 100% at the 96-h to 120-h forecast intervals, these forecast intervals have a very small sample size.

2. Phase III: Under (U) Intensity Trends

Notice that an Under (U) intensity in Phase III is defined here to be that the forecast intensity is at least 10 kt too low, which is to say that decay rate from the peak intensity was forecast to be too strong.

a. Atlantic

The Under (U) intensity trend percentages for all of the techniques and the NHC are 12% or less except for the GFDI model (Figure 3.34). From the 60-h to 120-h forecast intervals, the U percentages are 5% or less for all of the techniques and the NHC. However, the sample sizes are small in these longer forecast intervals. The GFDI model has the largest U percentages from the 12-h (18%) to 60-h (16%) forecast intervals, and the GFNI model has the next highest percentages. Only the NHC has skill relative to SHF5 at 12 h and 48 h and only the NHC and the SHIPS technique have skill relative to SHF5 at 24 h and 36 h. Although the DSHIPS technique U percentages are 10% or less during the entire forecast interval, the technique had no skill relative to the SHF5. Even though the dynamical models have more of a tendency to decay a TC more quickly than the other techniques and the NHC in the Atlantic, these percentages are relatively small.

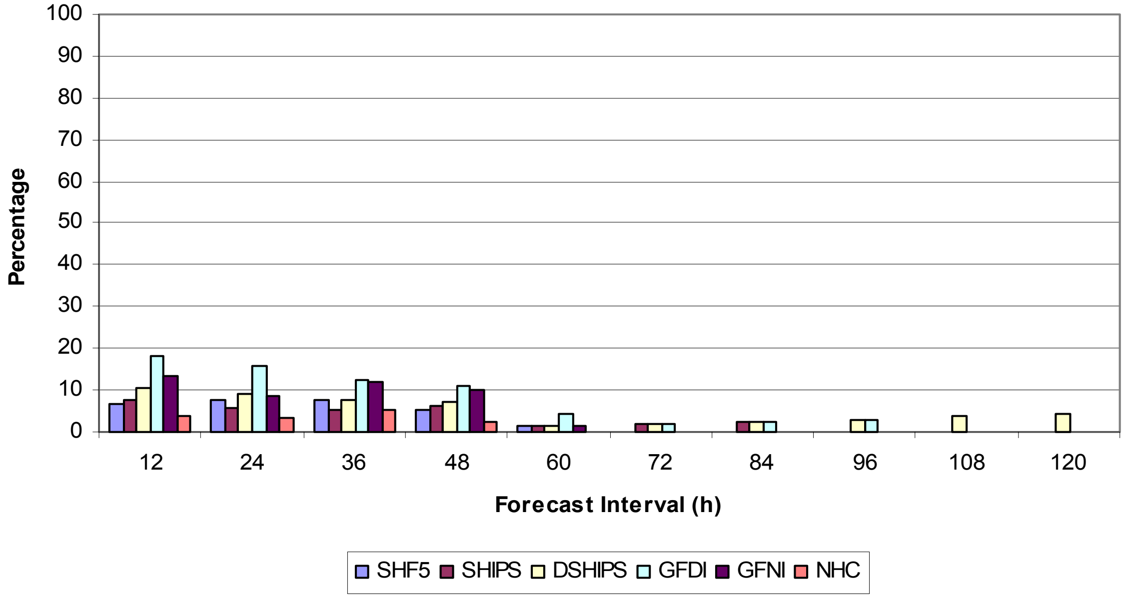


Figure 3.34 Percentages of Under intensity trends during the decay phase (Phase III) for the combined 2003 and 2004 Atlantic seasons. An under-forecast during decay indicates that the magnitude of the forecast intensity is less than the actual intensity by at least 10 kt (decay too fast). Sample sizes of the verified forecasts range from over 1200 at 12 h to 190 by the 96-h interval.

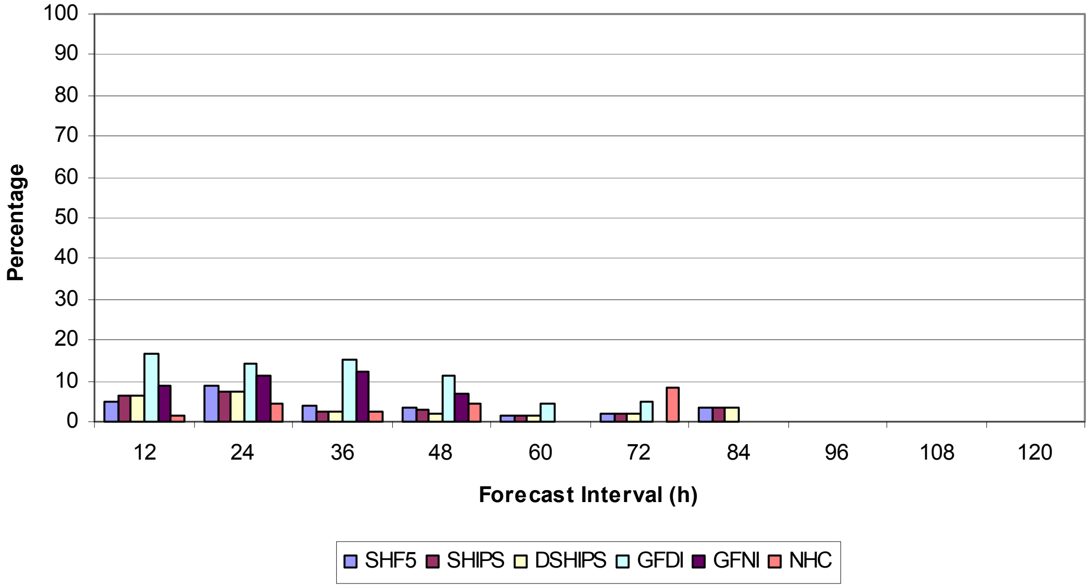


Figure 3.35 Same as Figure 3.34, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 1400 at 12 h to 95 by the 96-h interval.

b. Eastern North Pacific

Similar to the Atlantic, the U intensity trend percentages during Phase III in the eastern North Pacific are small through the entire forecast interval (Figure 3.35). The GFDI model has the majority of the highest U percentages, and the GFNI model has the next highest percentages. From the 12-h to the 60-h forecast intervals, the GFDI has the highest U percentages ranging from 11% to 17%. The NHC has skill relative to SHF5 at 12-, 24-, 36-h forecast intervals. The SHIPS and DSHIPS techniques have skill relative to SHF5 from the 24-h to 48-h forecast intervals. Although the sample sizes are small beyond the 60-h forecast, all of the techniques and the NHC have 0% U percentages from the 96-h to 120-h forecast interval, which means that none of these forecasts was for a too-large decay of the storm in Phase III.

3. Phase III: Over (O) Intensity Trends

The definition of an Over (O) intensity in Phase III is that the forecast intensity is too high, which is to say that the decay rate from the peak intensity was forecast to be too small. Given that the percentages have to sum to 100%, and that the U intensity forecasts seldom occurred (see previous subsection), most of the intensity errors greater than +/- 10 kt during the decay phase are expected to be O intensity forecast errors.

a. Atlantic

The SHF5 and the SHIPS most frequently over-forecast the intensity during the decay phase (Figure 3.36). That is, these techniques do not decay a TC fast enough. The SHF5 has the highest O (Over) percentage from the 12-h to 48-h forecast intervals, and the SHIPS technique has the next highest percentages. Consequently, all the techniques and the NHC are skillful relative to the SHF5 through the first 48 h during Phase III. From the 60-h to 108-h forecast intervals, the SHIPS technique has the highest O percentages. Conversely, the DSHIPS has either the lowest or second lowest O percentages during the entire forecast interval with 0% from the 84-h to 120-h forecast intervals. The GFDI O intensity trend is similar to the DSHIPS technique with the lowest O percentages at 24 h and 36 h. Although the GFNI O intensity

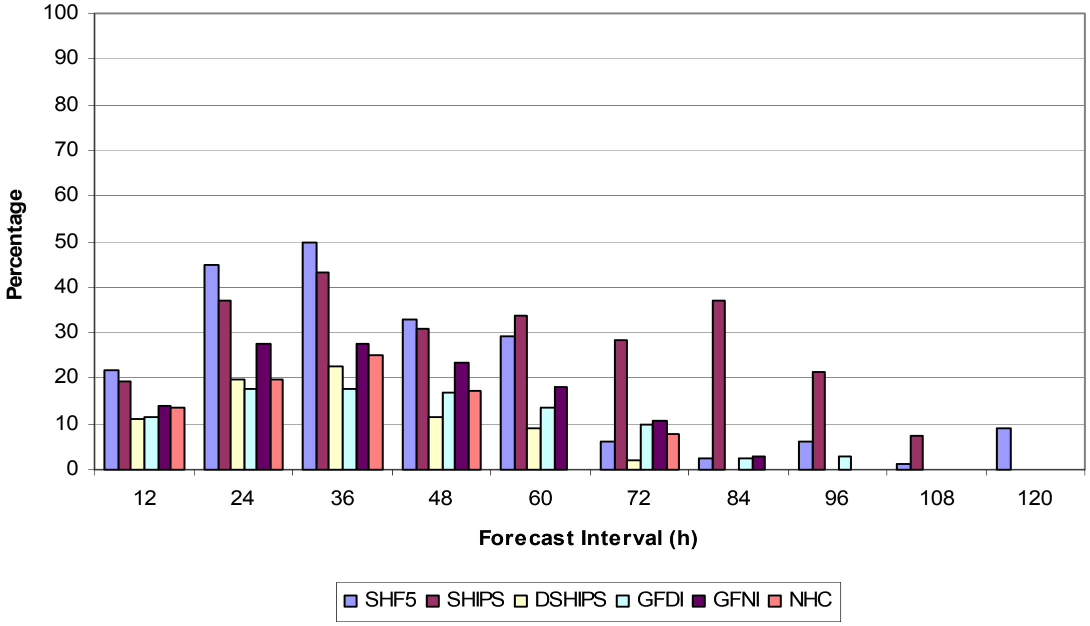


Figure 3.36 Percentage of Over intensity trends during the decay phase (Phase III) for the combined 2003 and 2004 Atlantic seasons. An over- forecast during Phase III indicates that the magnitude of the forecast intensity is greater than the actual intensity (decaying too slowly). Sample sizes of the verified forecasts range from over 1200 at 12 h to 190 by the 96-h interval.

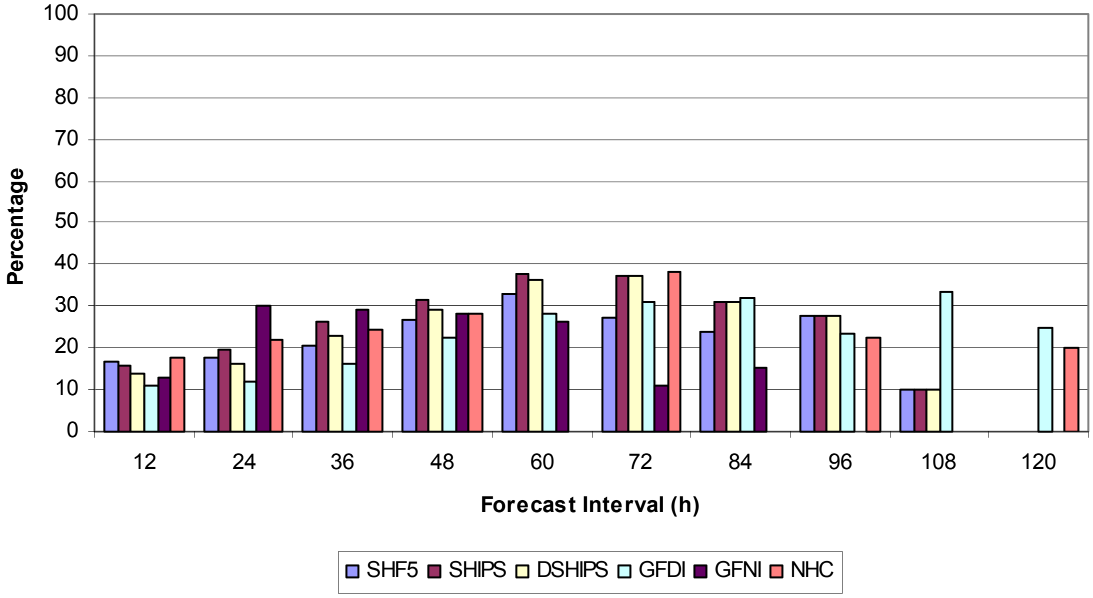


Figure 3.37 Same as Figure 3.36, except for the combined 2003 and 2004 eastern North Pacific TC seasons. Sample sizes of the verified forecasts range from over 1400 at 12 h to 95 by the 96-h interval.

percentages are not quite as low as for the GFDI model, the GFNI model also has skill relative to SHF5 through 60 h in the decay phase. Whereas the NHC has similar O percentages as DSHIPS at the 12-h and 24-h forecast intervals, the NHC does not add value relative to DSHIPS at longer forecast intervals.

b. Eastern North Pacific

The O percentages for the various techniques and the NHC increase with increasing forecast intervals until 60 h (Figure 3.37). After 60 h, the O percentages for the techniques and the NHC then decrease gradually. It is striking that the SHF5 and SHIPS techniques have significantly fewer O forecasts of the intensity during Phase III than in the Atlantic (Figure 3.36). The GFDI model has skill relative to SHF5 from the 12-h to the 60-h forecast interval, while the GFNI model has skill from the 60-h to the 108-h forecast interval (both the SHF5 and GFNI have 0% O percentages at 120 h). Although the SHIPS technique only has skill relative to SHF5 at the 12-h forecast interval, the DSHIPS has skill relative to SHF5 at the 12-h and 24-h forecast intervals. Similarly, the NHC has skill relative to SHF5 in terms of smaller O percentages only at the 96-h forecast interval. As in the eastern North Pacific during Phase II, the NHC follows the intensity guidance of its statistical-dynamical techniques, which have a tendency to over-forecast TCs during the decay phase, and thus the NHC does not add value relative to climatology and persistence.

4. Phase III: Summary

In the decay Phase III, all the techniques have a tendency to over-forecast intensity and only the GFNI model has a slight tendency to under-forecast intensity at 12 h and 24 h. However, the DSHIPS technique had the highest Good (G) intensity trend percentage at 87% for Phase III in the Atlantic followed by the NHC at 86% when averaged over all of the forecast intervals. The DSHIPS technique is less likely to over-forecast a TC during the decay phase than the NHC or the other techniques, and has a 10% or less chance of under-forecasting the intensity. Therefore, the DSHIPS technique has the best performance in the Atlantic during the decay phase. The NHC on average does not add value relative to the DSHIPS technique.

In the eastern North Pacific, none of the techniques or the NHC clearly out-performed the others during Phase III. Since the small sample sizes during the 72-h to the 120-h forecast intervals do not properly represent the techniques and the NHC performance, more emphasis is placed on the 12-h to the 60-h forecast intervals. The NHC has essentially the same (within 1%) G intensity trend percentages from the 12-h to the 60-h forecast intervals as the SHF5 and DSHIPS techniques. However, these techniques have more of a tendency to over-forecast the intensity changes during the decay phase in the eastern North Pacific. As in Phase II in this region, the NHC evidently uses guidance from the statistical-dynamical techniques, which means the value added is small.

5. Phase III: Comparison with the Western North Pacific

The comparable statistical-dynamical SHIPS and STIPS techniques and the GFNI model had a tendency to over-forecast the intensity of a TC during the decay stage in all three basins. The climatology and persistence ST5D technique was the most reliable intensity change guidance during the decay phase in the western North Pacific (Blackerby 2005), but had average performance in the eastern North Pacific and in the Atlantic. Blackerby (2005) also noted that the JTWC forecast added value relative to many of the techniques, but infrequently added value to the ST5D. Although the NHC does add value in the eastern North Pacific relative to all of the techniques, the value-added is small.

6. Intensity Forecasts Verifying at the 45-kt Decay Point

Forecasting an accurate and timely decay of a TC is just as important as forecasting its intensification. Predicting a TC will maintain its strength when it is actually decaying may cause an unnecessary military sortie or a massive evacuation. In contrast, decaying a TC too quickly can lead to a lack of proper preparation that results in extensive property damage and even human lives. Therefore, it is of interest whether the intensity techniques can provide accurate guidance during the decay of a TC.

One approach in evaluating the intensity techniques' ability to forecast during the decay phase is to compare the forecasts that verify at the 45-kt decay point. The 45-kt value is used because many storms are not tracked during extratropical transition phase or after landfall where intensity abruptly ends above 45 kt. Thus, the DTG of each storm that had decayed down to 45 kt was extracted, which included 15 storms in the Atlantic and 22 storms in the eastern North Atlantic during the 2003 and 2004 hurricane seasons. All forecasts (i.e., -120 h, -96 h, …-24 h) that could be verified at the DTG of the 45-kt decay point were collected. The difference between the predicted intensity and the actual 45 kt intensity is defined as the error. If the error is positive, then the intensity technique decayed the TC too slowly (forecast intensity was too high). If the error is negative, then the intensity technique decayed the TC too quickly (forecast intensity was too low). The TCs that had intensities of 50 kt or less (weak storms) were excluded in this analysis as they do not exhibit a well-defined decay phase or were already at the decay threshold. Therefore the sample sizes for the Atlantic basin from 72 h to 96 h are less than 10, and the sample size for the eastern North Pacific at 120 h is also less than 10 and considerably lower for the GFNI and GFDI. Thus the small sample may not give an accurate representation of the intensity techniques performances.

a. Atlantic

The tendency in the Atlantic is to over-forecast the intensity during the decay to 45 kt (Figure 3.38). The SHF5 and SHIPS techniques have the highest errors for all forecasts (except SHF5 at 120 h) prior to the 45-kt decay point, which was also evident during the Phase III (decay) analysis in the previous section. The large over-forecasting errors by the SHIPS technique suggest that many of these decays are associated with landfalls because the DSHIPS technique has the lowest errors during all forecast hours (except NHC at 120 h). On average, the DSHIPS technique is less than 2 kt from exactly matching the observed intensity decrease during the 24 h before the TC intensity decays to 45 kt. Thus, the DSHIPS technique out-performed all of the techniques and the NHC. The dynamical models GFNI and GFDI maintained a

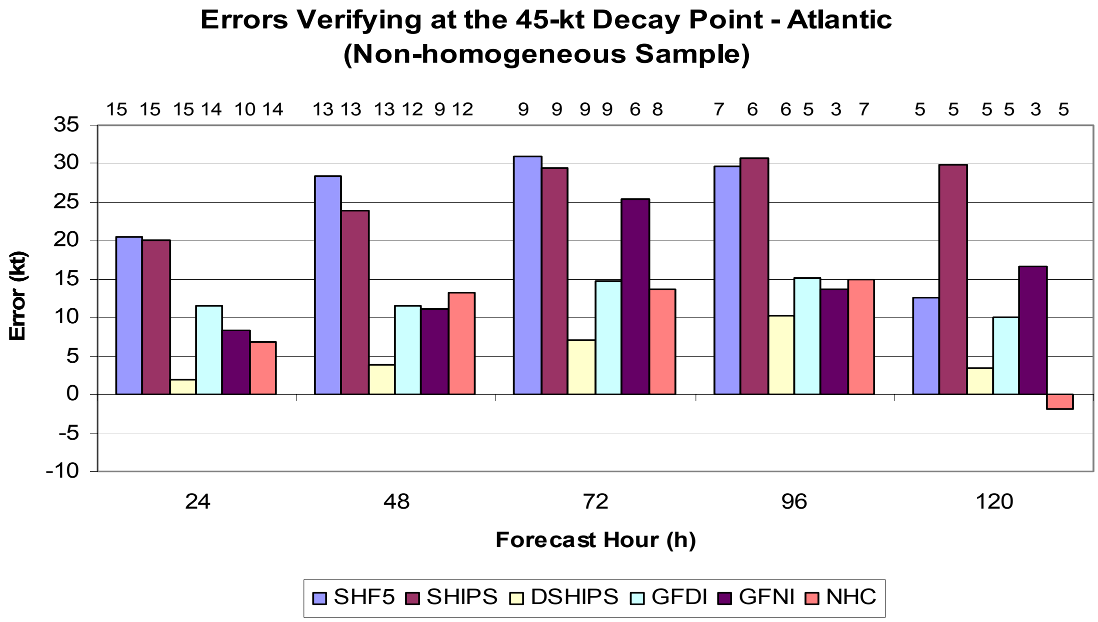


Figure 3.38 Observed mean intensity errors (kt) verifying at the date/time of decay down to 45 kt for the 2003 and 2004 Atlantic hurricane seasons. The small numbers above each bar represent the number of cases. A non-homogeneous sample is allowed to maximize the cases for analysis.

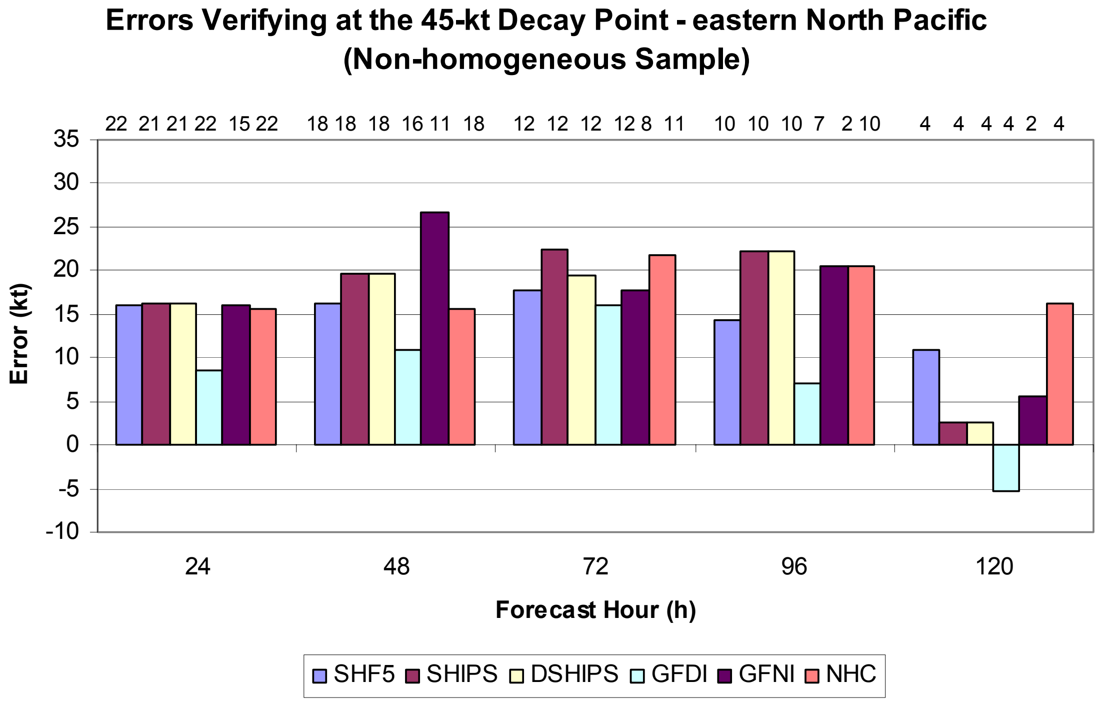


Figure 3.39 Same as Figure 3.38, except for the 2003 and 2004 eastern North Pacific hurricane seasons.

modest over-forecast intensity of 12.5 kt and 15 kt, respectively, during the decay to 45 kt, which is the same average maintained by the NHC from 96 h to 24 h prior to the decay to 45 kt. It is noteworthy that the NHC has a tendency to forecast a TC to decay too quickly five days before the TC reaches the 45 kt point. As noted earlier, the sample sizes are small at 72 h, 96 h, and 120 h.

b. Eastern North Pacific

The tendency of over-forecasting the intensity during the decay to 45 kt is also evident in the eastern North Pacific (Figure 3.39). The GFDI model has the lowest errors during all the forecast hours prior to the 45-kt decay point (except SHIPS and DSHPS at 120 h), especially at 24 h, 48 h, and 96 h. The SHIPS and DSHIPS technique have an error of at least 16 kt at all the forecast hours except 120 h. Similarly, the NHC has errors of at least 15 kt at all the forecast hours prior to the 45-kt decay point, and only has skill relative to SHF5 at 24 h and 48 h. Although the GFNI model has a 26 kt error just 48 h prior to the 45-kt decay point, the model has skill relative to SHF5 at the 24-, 72-, and 120-h forecast intervals. Since the SHIPS and DSHIPS techniques errors at 120 h are much smaller compared to the other forecast periods, this is probably indicative of too small sample sizes of forecasts that can be validated.

c. Comparison with the Western North Pacific

The comparable climatology and persistence techniques SHF5 and STD5 have similar error values (within 5 kt) at each forecast hour during the decay to 45 kt in the eastern and western North Pacific, respectively. The SHF5 technique has much larger errors in the Atlantic errors. The SHIPS and STIPS in the Atlantic and western North Pacific decay a TC too slowly with average errors of 25 kt and 30 kt. The GFNI model average errors in the western North Pacific are well above the values in the Atlantic and eastern North Pacific.

7. Rapid Decay

Similar to rapid intensification, the ‘best track’ intensity values from the NHC ATCF-format files were used to determine which TCs underwent rapid decay (30 kt/day) during the 2003 and 2004 tropical seasons for both basins. The intensity forecasts for the various techniques and for the NHC of the TCs that underwent rapid decay were examined for all times +/- 12 h relative to the actual time of rapid decay. If the technique's predictions matched or exceeded the rapid decay threshold, this was regarded as a 'hit.' If the technique prediction did not meet the threshold at the correct time or within +/- 12 h, it is recorded as a 'miss.'

a. Atlantic

The DSHIPS technique forecast 13 of the 15 rapid decay events in the Atlantic during 2003 and 2004 seasons (Figure 3.40). Conversely, the SHIPS technique only forecast 3 of 15 rapid decay events. This large discrepancy between the DSHIPS and SHIPS techniques in forecasting rapid decay demonstrates the usefulness of adding the empirical decay formula over land to the DSHIPS technique. It is noteworthy that the DSHIPS technique had a tendency to decay many of the TCs too much (5 to 10 kt faster), especially for the stronger hurricanes. The NHC also forecast 13 of the 15 rapid decay events. Recall that the decay factor is triggered when the official NHC forecast crosses land. Of course, the NHC forecaster who made the landfall forecast will then make the intensity forecast begin a decay at the same time as the DSHIPS technique. The other techniques forecast less than 50% of the rapid decay events in the Atlantic.

b. Eastern North Pacific

None of the techniques forecast many of the rapid decay events in the eastern North Pacific (Figure 3.41). Of the 11 rapid decay events that occurred in the eastern North Pacific during the 2003 and 2004 seasons, the most forecast by any technique or the NHC was three events. The two dynamical models and the DSHIPS all forecast three of 11 rapid decay events while SHIPS forecast only two of 11 events. The SHF5 technique did not forecast any of the rapid decay events. The NHC only forecast one of the 11 events and therefore did not add value. Thus, all of the techniques and the NHC performed poorly in forecasting rapid decay in the eastern North Pacific during 2003 and 2004.

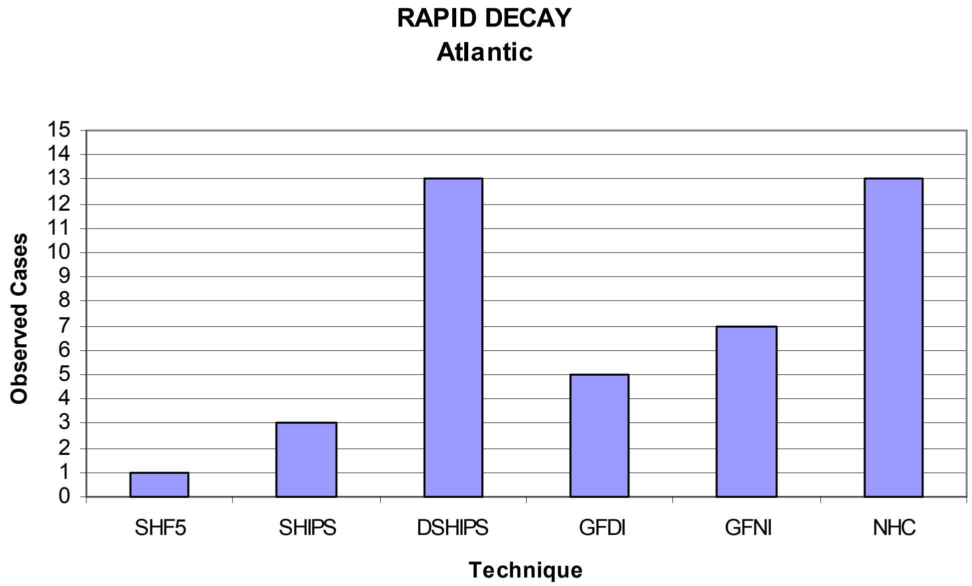


Figure 3.40 Number out of 15 cases in the Atlantic basin during the 2003 and 2004 seasons that a forecast technique or NHC predicted an observed period of rapid decay at the correct time or within +/- 12 h.

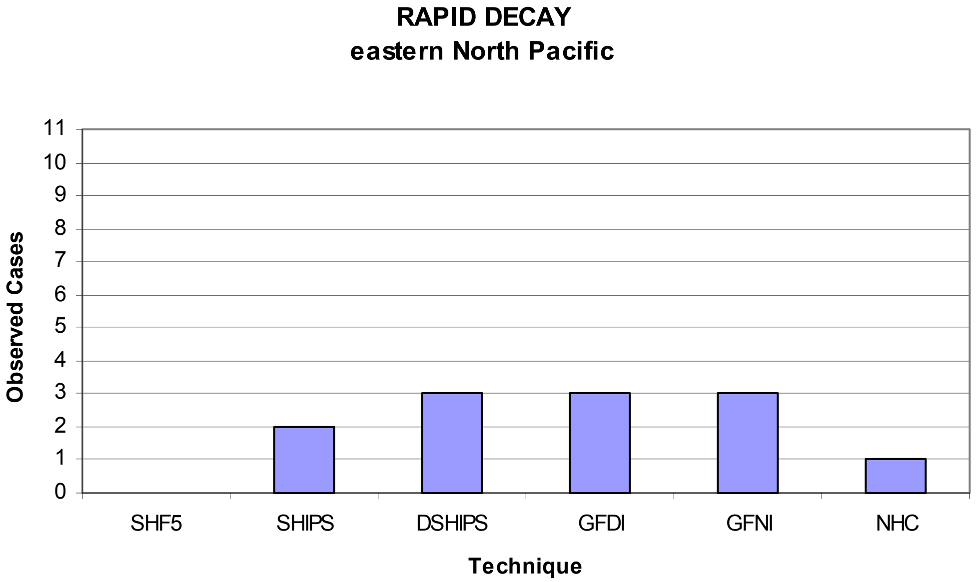


Figure 3.41 Same as Figure 3.40, except for 11 cases in the eastern North Pacific during 2003 and 2004.

c. Comparison with the Western North Pacific

Whereas the climatology and persistence technique ST5D in the western North Pacific forecast 10 of 28 events of rapid decay, the SHF5 only forecast one rapid decay event in the Atlantic and none in the eastern North Pacific. This difference suggests that the climatology (average intensity change) in the western North Pacific has more occurrences of rapid decay, perhaps due to the stronger TCs (on average) in the region than in the Atlantic and eastern North Pacific. The GFNI predicted approximately 47% of the rapid decay events in the Atlantic and 43% western North Pacific, but predicted less than 30% in the eastern North Pacific. The statistical-dynamical SHIPS and STIPS performed poorly in predicting rapid decay in all three basins as the STIPS technique only forecast six of 28 events in the western North Pacific.

8. Intensity Change Distribution during Phase III

As noted earlier (Chapter III Section D.6), an important limitation of the intensity guidance techniques was revealed from the distribution of intensification rates at the beginning of Phase II. Equally as important is the distribution of intensity changes of each technique from the beginning of Phase III. Thus, the 48-h distribution of predicted intensity changes that started at the DTG of the peak intensity (or the last secondary peak) was analyzed to quantitatively determine the distributions of predicted intensity changes in the Atlantic (Figure 3.42) and the eastern North Pacific (Figure 3.43) for the 2003 and 2004 tropical seasons.

a. Atlantic

The observed intensity changes, which are repeated in each panel of Figure 3.42, range mainly from -75 kt to - 10 kt, which should be expected since the TCs are in the decay phase. Approximately 82% of the observed intensity changes were between -35 kt and -10 kt with the highest percentages at -30, -25, and -10 kt, which indicates that the range of the actual 48-h intensity changes was relatively narrow. Although one TC during 2003 had a slight reintensification (< 10 kt) after an initial 5 kt decay from the peak intensity this did not qualify as a decay-reintensification cycle with a secondary peak.

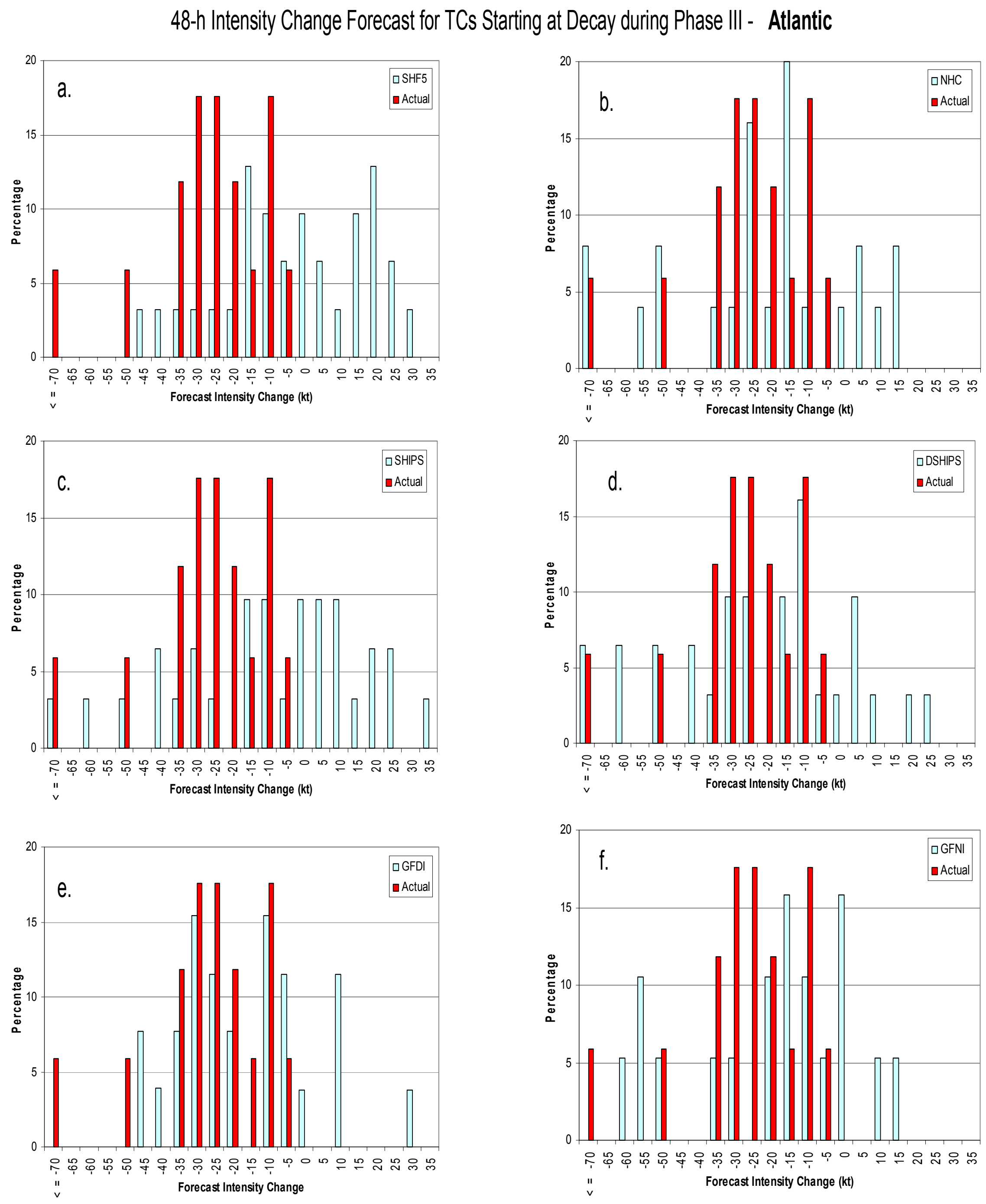


Figure 3.42 (a-f) Observed 48-h intensity changes (solid bars) for TCs starting at peak intensity during Phase III and the predicted 48-h intensity changes (blue bars) by the various techniques and by the NHC for the 2003 and 2004 Atlantic seasons.

Over half of the SHF5 48-h intensity changes were between 0 kt and 30 kt (Figure 3.42a). That is, the SHF5 will forecast either no change or an increase in intensity in 50% of the cases when the TC is actually decaying. Thus, the climatology and persistence technique has difficulties in predicting the intensity changes at the beginning of the decay phase, which suggests that the tracks in the developmental sample used to determine the techniques climatological average did not have landfalls that would cause decay during the 48 h after peak intensity. The climatological and persistence techniques depend on location, time of year, current and past intensity. Consequently, a TC that encounters a landfall when many of the climatological tracks from the initial location did not have a landfall and subsequent decay can make the SHF5 intensity forecast erroneous in such situations.

The SHIPS technique 48-h intensity change distribution was similar to SHF5 with nearly half of the intensity changes between 0 kt and 35 kt instead of a decay (Figure 3.42c). Although SHIPS was able to forecast some of the stronger decay rates, the percentage is small compared to the actual percentage. Thus, the SHIPS technique with inclusion of atmospheric predictors was able to overcome some of the erroneous intensifications in the climatology and persistence SHF5 technique.

Even though the DSHIPS technique did forecast some intensifications instead of decays, the majority (77%) of the 48-h intensity change distribution are decays (Figure 3.42d). The DSHIPS also forecast larger decay rates, which was demonstrated earlier as the technique performed well during rapid decay. This improved performance relative to the SHIPS also suggests that the empirical decay formula used by DSHIPS once a TC is forecast to move over land is valuable. However, the DSHIPS forecast some larger decays than observed and had one decay of -95 kt that was 20 kt greater than the largest observed intensity change. If the NHC track forecast moving over or remaining on land is not accurate, the DSHIPS technique will erroneously decay the TC or decay it too much.

The GFDI model has 80% of its 48-h intensity change distribution in Phase III between -45 kt and -5 kt, which is similar to the actual intensity change distribution (Figure 3.42e). In only a few cases did the GFDI model forecast a decaying TC to intensify or have no change in intensity. This distribution suggests that the GFDI model physics are capable of predicting the factors that lead a TC to decay, such as lower sea-surface temperatures or strong vertical wind shear. Still, the model is unable to predict the strongest cases of decay after the beginning of Phase III.

Unlike the GFDI, the GFNI model was able to forecast some of the stronger decay rates (Figure 3.42f). However, the model had very few forecast intensity changes at -30 kt and -25 kt, which is the range of the majority of actual decay rates. By contrast, the GFNI had its highest percentage at -15 kt and 0 kt intensity changes. Thus, the GFNI model forecasts of 48-h intensity change after peak intensity are usually conservative, with an occasional forecast stronger than -25 kt.

Nearly 68% of the NHC 48-h intensity change forecasts after the peak intensity are between -25 kt and 25 kt, with 24% being a positive intensity change instead of decay (Figure 3.42b). Approximately 20% of the NHC forecast intensity changes are at -15 kt and 16% are at -25 kt. Whereas the observed 48-h distribution had 29% of the intensity changes at -35 kt and -30 kt, only 8% of the NHC forecasts are at -35 kt and -30 kt. As demonstrated in Figure 3.40, the forecasters at the NHC did accurately forecast 13 of the 15 rapid decays. This NHC distribution suggests that the NHC forecaster will decay a TC too slowly a majority of the time, but still predict a strong decay when the TC makes landfall.

b. Eastern North Pacific

Similar to the Atlantic, all of the observed 48-h intensity distribution changes in Phase III are negative except in one case when a TC experienced reintensification back to its original peak intensity, which resulted in no change in intensity from 48 h (Figure 3.43). Over 80% of the 48-h intensity distribution changes are between -55 kt and -10 kt with no intensity changes of -5 kt. These intensity changes are distributed over a wide range with no single maximum percentage in intensity change. During the eastern North Pacific hurricane seasons during 2003 and 2004, nearly 28% of the TCs experienced a -50 kt or stronger decay in intensity during the 48 h after peak intensity.

The SHF5 48-h intensity change distribution had a range from -35 kt to 25 kt (Figure 3.43a). About 21% of the 48-h intensity changes in Phase III predicted by the SHF5 were no change in intensity. Because the SHF5 technique did not forecast any decays greater than -35 kt, the technique did not forecast any rapid decay events (Figure 3.41).

As in the Atlantic, the SHIPS technique intensity change distribution was skewed toward positive intensity (Figure 3.43c). Nearly 50% of the intensity changes were between 0 kt and 30 kt rather than decays. Thus, the technique has a tendency to intensify a TC beyond its maximum intensity. The SHIPS had only one forecast with a decay greater than -35 kt, so the technique fails in forecasting stronger decay rates.

The DSHIPS technique has fewer positive intensity changes from the beginning of Phase III and a few more large decay rates than does SHIPS (Figure 3.43d). However, the DSHIPS technique still has 85% of the intensity changes between -25 kt and 20 kt. Thus, the DSHIPS technique is essentially similar to SHIPS in predicting either an intensification or a modest decay in the 48 h after a TC reaches peak intensity.

In contrast to the statistical-dynamical techniques, the GFDI model predicts only 11% positive intensity changes (Figure 3.43e). Therefore, the model forecast the TCs in the eastern North Pacific to decay with a distribution close to the observed 48-h intensity change distribution. However, the largest model forecast decay rate is -50 kt, and therefore does not predict the extreme decay rates.

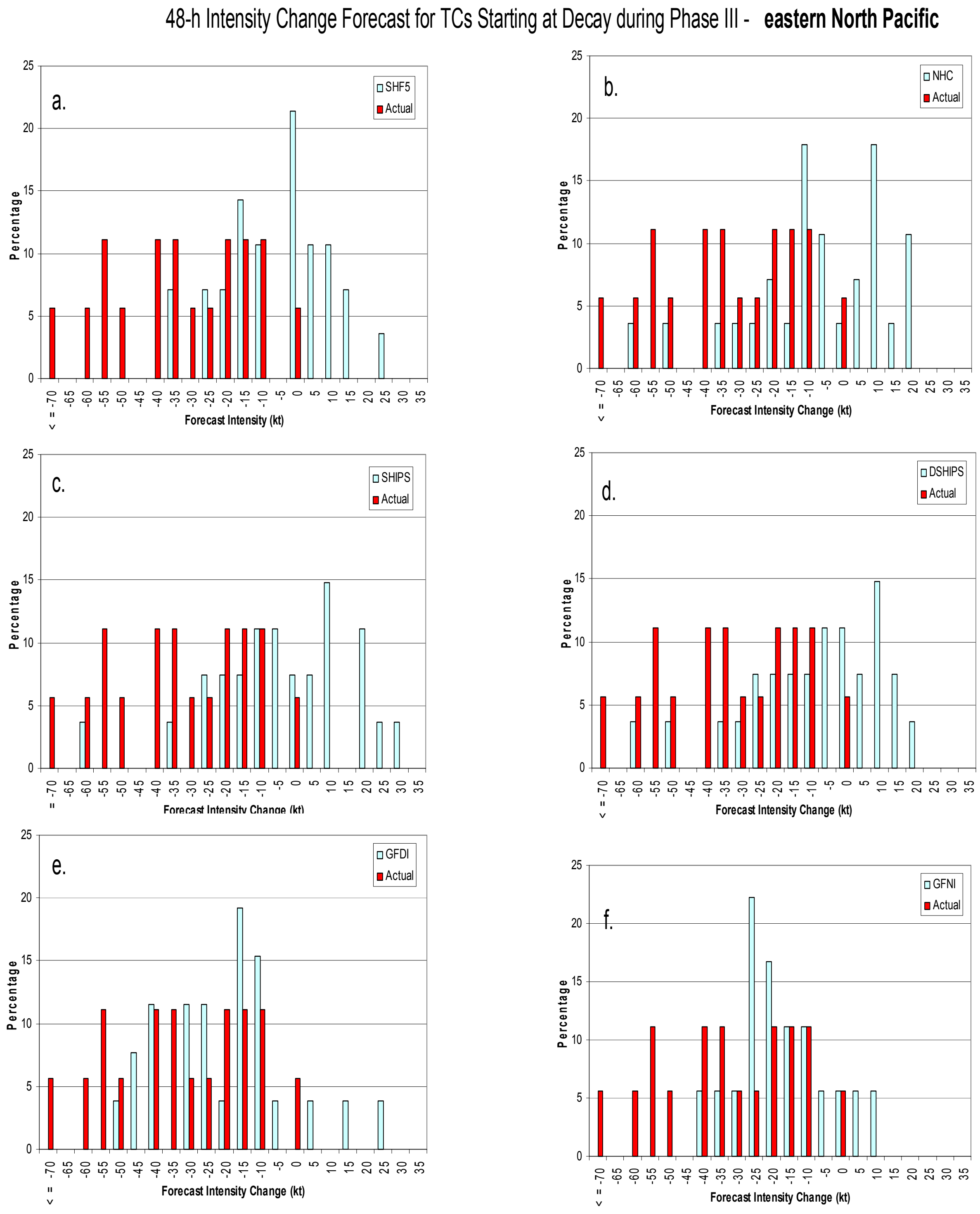


Figure 3.43 (a-f) Same as Figure 3.42 (a-f), except for the 2003 and 2004 eastern North Pacific seasons.

The GFNI model 48-h intensity changes in Phase III tend to cluster between -40 kt and 10 kt (Figure 3.43f). Nearly 67% of the intensity changes are between -25 kt and -5 kt, with 22% at -25 kt. Thus, the model is too conservative with its intensity changes and therefore is unable to forecast larger decay rates.

The NHC 48-h intensity distributions are similar to SHIPS and DSHIPS (Figure 3.43b). Approximately 93% of the intensity changes are between -35 kt and 20 kt. The NHC only made two forecasts of intensity changes larger than -35 kt. Many of the NHC forecasts were clustered between an intensification of 10 kt or a decay of 10 kt. The NHC forecasts in the eastern North Pacific are similar to their intensity guidance techniques - conservative and with rare forecasts of large decay rates.

c. Intensity Change Distribution Summary

In the Atlantic and the eastern North Pacific, the majority of the 48-h intensity change distributions following the beginning of Phase III were negative intensity changes. The climatology and persistence techniques were unable to predict most of the intensity changes at the beginning of Phase III in both basins. The SHIPS technique intensity change distributions in both basins were positively skewed, but the SHIPS was able to predict some of the largest decay rates. The DSHIPS technique (in the both basins) performed somewhat better than SHIPS, with the DSHIPS in the Atlantic forecasting much larger decay rates than in the eastern North Pacific where few landfalls occur. In both basins, the dynamical models had a 48-h intensity change distribution after the beginning of Phase III that was less positively skewed relative to the statistical and statistical-dynamical techniques. The reasons all of the techniques in the eastern North Pacific forecast smaller 48-h intensity changes after peak intensity than in the Atlantic was likely due to the greater chance a TC will encounter land in the Atlantic.

d. Comparison with the Western North Pacific

The climatology and persistence technique ST5D in the western North Pacific (Blackerby 2005) had a much wider spread of 48-h intensity changes (-70 kt to 35 kt) in Phase III than the comparable SHF5 technique in the Atlantic and the eastern North Pacific. However, the majority of the values were in the zero or positive range, which is similar to the values of SHF5 in the Atlantic and the eastern North Pacific. The statistical-dynamical techniques STIPS and SHIPS had a too broad range of 48-h intensity changes in all three basins. The western North Pacific and the Atlantic statistical-dynamical techniques had more negative intensity changes relative to the comparable technique in the eastern North Pacific. The STIPS technique in the western North Pacific had a high percentage (~23%) of 0 kt intensity changes, which was not the case in the Atlantic and eastern North Pacific. The GFNI models in both the Atlantic and eastern North Pacific had 48-h intensity change distributions more negatively skewed than the GFNI model in the western North Pacific. The Atlantic and western North Pacific GFNI models have wider ranges of intensity changes than the GFNI model in the eastern North Pacific, which is likely due to fewer landfalls in the eastern North Pacific.

9. Average and Standard Deviation of the Intensity Change Distribution during Phase III

A simple way to represent the 48-h intensity change distribution during Phase III is demonstrated in Figure 3.44 (Atlantic) and Figure 3.45 (eastern North Pacific). Both the mean and +/-1 standard deviation of the 48-h intensity change distributions for each technique and the NHC are compared with the values for the observed distribution. These summaries illustrate the biases and ranges for the techniques and the NHC as discussed in the previous sections.

a. Atlantic

All of the techniques and the NHC have similar or broader intensity change distributions than the observed standard deviation of (17 kt), which is quite different from the too-small intensity change distributions during Phase II. The DSHIPS technique and the NHC mean intensity changes are within 4 kt of the mean observed decay rate and, with the broader distribution ranges, are able to predict more cases of rapid decay than the other techniques. Similar to DSHIPS, the GFNI model also has a broader intensity change distribution than

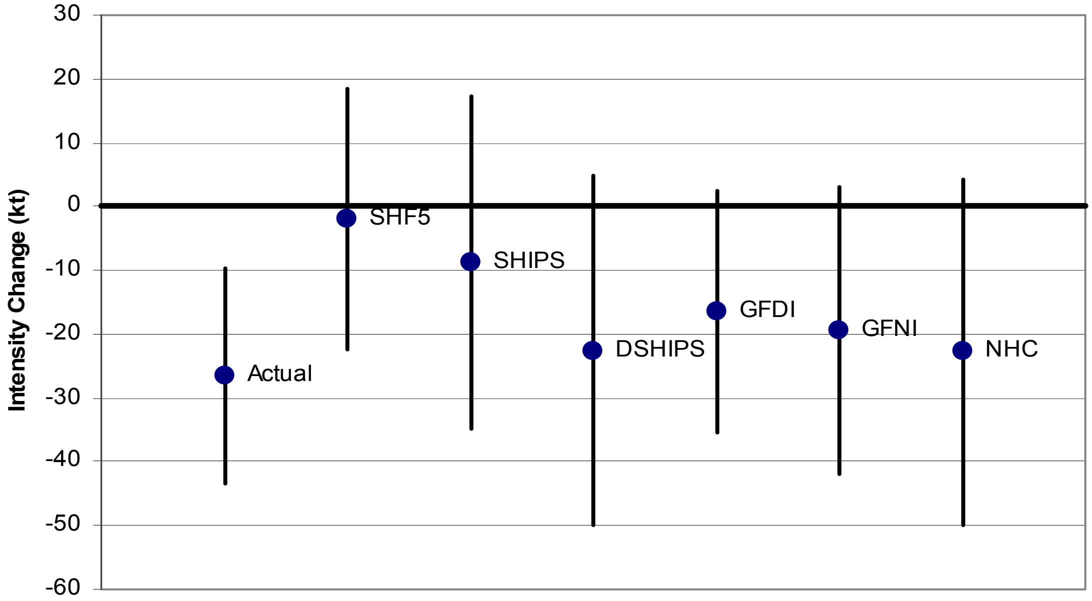


Figure 3.44 Actual and predicted 48-h intensity change distributions during Phase III starting at peak intensity through decay for 37 TCs during the 2003 and 2004 Atlantic seasons. The heavy dot represents the mean intensity change value (kt) while the lengths of the sticks represent plus or minus one standard deviation about the mean.

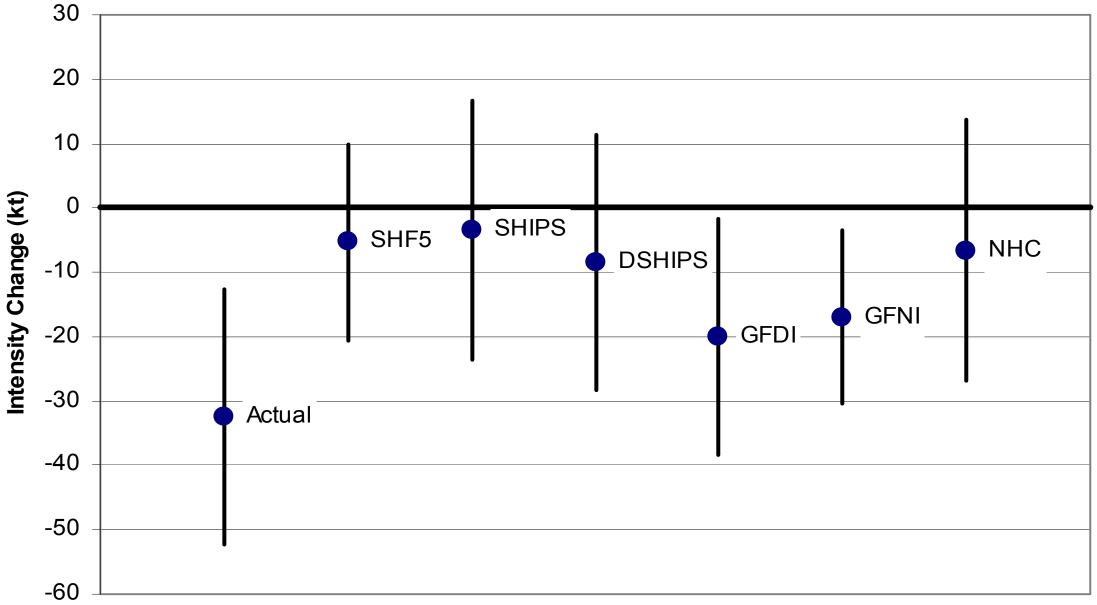


Figure 3.45 Same as Figure 3.44, except for the 32 TCs in the eastern North Pacific during the 2003 and 2004 seasons.

the observed standard deviation. Because, the GFNI model under-estimates the mean decay rate by 7 kt, the model does not predict some of the stronger decay rates, which is revealed by fewer cases of predicted rapid decay (7 of 15). The SHF5 and SHIPS distributions are shifted toward small decay rates with mean intensity changes of -2 kt and -9 kt, respectively, which is why the techniques are unable to forecast higher decay rates and have more of a tendency to decay TCs too slowly.

b. Eastern North Pacific

The actual mean 48-h intensity change (-32 kt) in Phase III is much larger than for the various techniques or the NHC. Whereas the SHIPS and DSHIPS techniques and the NHC have realistic intensity ranges, the mean decay rates are small because of the significant number of intensification predictions. Thus, neither of these techniques or the NHC forecast the typical decay rates or many of the cases of rapid decay in the eastern North Pacific during 2003 and 2004. This bias suggests that the atmospheric predictors in the statistical-dynamical techniques may not have enough influence in the prediction by the SHIPS and DSHIPS in the eastern North Pacific. Not only does the SHF5 technique have a near-zero (-5 kt) mean intensity change in Phase III, the standard deviation is also too small. The GFDI model 48-h intensity changes in Phase III come closest to representing the observed distribution except for a bias toward too small decay rates. The GFNI model has a similar bias and too-small of a range. In summary, none of the techniques or the NHC performed well when predicting intensity changes at the beginning of Phase III in the eastern North Pacific.

c. Comparison with the Western North Pacific

In all three basins, the comparable SHF5 and ST5D techniques have a bias of under-estimating the mean decay rate from the beginning of Phase III. In the Atlantic and western North Pacific, the standard deviations of the 48-h intensity changes from these climatology and persistence techniques were larger than the actual intensity change. The statistical-dynamical SHIPS and STIPS techniques in all three basins had larger standard deviations of the intensity changes than the actual intensity change, but they have a large bias of too small mean decay rates. The GFNI model also had a bias of too small mean intensity changes in all three basins. Thus, the comparable techniques in all three basins all had too small decay rates from the beginning of Phase III, which results in intensities that are too large 48 h after the final peak intensity.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. SUMMARY AND CONCLUSIONS

Selected intensity guidance techniques available at the National Hurricane Center during the 2003 and 2004 hurricane seasons for the Atlantic and eastern North Pacific have been evaluated to determine their accuracy during three intensity phases to provide guidance to the forecaster on when the technique is likely to be (not be) accurate. The three intensity phases are based on the intensity changes during the life cycle of the TC (Figure 2.2): (i) formation stage to tropical depression (33 kt); (ii) early intensification through tropical storm (intensification phase); and (iii) the decay phase. The hypothesis is that comparable intensity guidance techniques in these basins will have similar error characteristics as in the western North Pacific, which was studied by Blackerby (2005).

A contingency table (Figure 3.9) was used to evaluate the performances of the different intensity guidance techniques during each intensity phase. The contingency table was based on whether a guidance technique can forecast within +/- 10 kt the intensity trends: Over, Good, or Under. These Good, Under, or Over forecast trend performance were assigned to each intensity guidance technique for each forecast interval that could be verified. In Phase II and Phase III, each intensity technique was also analyzed for their errors during rapid intensification and decay, forecast verifying at peak intensity, intensity forecasts verifying at the 45-kt decay point, and the forecast intensity change distributions during Phase II and Phase III. Additionally, the NHC official forecasts were also evaluated during each phase to determine if their forecasts added value relative to the various intensity guidance techniques.

A. OVERALL SUMMARY

Summaries and basin comparisons of the intensity techniques have already been given at the conclusion of each Phase and the various intensity/decay analyses. To provide a summary of the error characteristics of the various techniques and the NHC, the errors are averaged over all forecast intervals. For example, the Good intensity trends in the Atlantic during the 2003 and 2004 seasons (Figure 3.10) averaged over all forecast intervals results in the summary diagram in Figure 4.1. Similar averages over all forecast intervals were calculated for the Under and Over intensity trends, and also in the eastern North Pacific to produce the overall summaries in Tables 4.1 to 4.4. The table consists of the 'best technique', over-forecast or under-forecast tendencies, and if the NHC forecast added-value to the 'best technique' for each phase and the various intensity/decay analyses.

The best technique is defined based on the overall highest percentage of Good trends when averaged over all forecast intervals based on the percentage of Good intensity trends as defined in Figure 3.9. For example, the DSHIPS technique had the overall highest percentage of Good intensity trends during Phase I in the Atlantic (Figure 4.1). Similarly, the over- (under-) forecast errors are determined from the highest percent of over (under) trends when averaged over all forecast intervals based on the percentage of Over (Under) intensity trends for each phase. For example, the SHF5 forecasts in both the Atlantic and in the eastern North Pacific, and the SHIPS forecasts in the Atlantic, have a strong tendency to over-forecast intensities during Phase I (Table 4.1). By contrast, the GFDI model has a tendency to under-forecast intensities in Phase I in both basins. The NHC value-added category is given a 'yes' in both basins (Table 4.1) during Phase I because the official forecast had a higher percentage of Good intensity trends than the 'best technique' when averaged over all forecast intervals.

During Phase II (Table 4.2), the best technique was the DSHIPS in the Atlantic and the GFDI model in the eastern North Pacific. Both the SHF5 and SHIPS have more of a tendency to have over-forecasts rather than under-forecasts when the errors exceed +/- 10 kt (i.e., not in the Good forecast category). Since few landfalls occurred in the eastern North Pacific, the DSHIPS has the same characteristics as the SHIPS in that basin. Both dynamical models

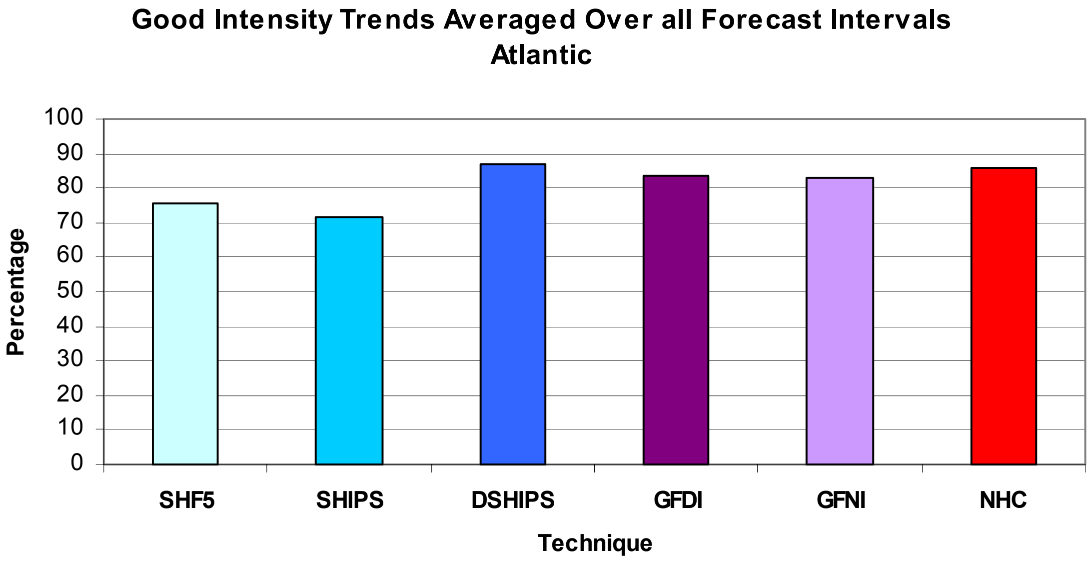


Figure 4.1 The Good intensity trends averaged over all forecast intervals from Figure 3.10 during Phase I in the Atlantic for the 2003 and 2004 hurricane seasons. This evaluation is used to determine the 'best technique' for each phase.

Table 4.1 Summary of Phase I intensity predictions for the Atlantic and eastern North Pacific during the 2003 and 2004 tropical seasons, except the GFNI model was not evaluated due to the small sample size.

	Atlantic	**Eastern North Pacific**
Best technique	DSHIPS	SHIPS and DSHIPS
Over-forecast	SHF5 and SHIPS	SHF5
Under-forecast	GFDI	GFDI
Does NHC add value compared to best technique?	Yes	Yes

have more under-forecasts than over-forecasts during Phase II in the Atlantic, but only GFNI has this characteristic in the eastern Pacific. Based on this guidance, NHC adds value over the DSHIPS in the Atlantic but does not add value over the GFDI model in the eastern North Pacific.

Table 4.2 Summary of Phase II intensity predictions for the Atlantic and eastern North Pacific during the 2003 and 2004 seasons.

	Atlantic	**Eastern North Pacific**
Best technique	DSHIPS	GFDI
Over-forecast	SHF5 and SHIPS	SHF5, SHIPS, DSHIPS
Under-forecast	GFDI and GFNI	GFNI
Does NHC add value compared to best technique?	Yes	No

The performance of the techniques in forecasting the various aspects during Phase II of the intensity cycle are summarized in Table 4.3. All of the techniques and the NHC have a bias of under-forecasting the peak intensity at the end of Phase II in both basins. While the bias is small (10 kt) at 24 h, the bias becomes progressively larger for longer-range forecasts that verify at the time of peak intensity. Part of this under-forecast bias arises from the overall poor performance in forecasting rapid intensification, which is true in both basins (Table 4.3). Again in the Phase IIa, which involves a decay and reintensification cycle, the performance of all of the techniques and the NHC must be rated as poor in both the Atlantic and eastern Pacific. Clearly the NHC requires better guidance for the existence, timing, and magnitude of such decay and reintensification cycles.

During the decay phase, the DSHIPS was clearly the best technique in the Atlantic, which can be related to the number of landfalls in that basin (Table 4.4).

Table 4.3 Summary of forecast performance for various aspects during Phase II for the Atlantic and eastern North Pacific during the 2003 and 2004 seasons.

Trend	**Atlantic**	**Eastern North Pacific**
Peak Intensity	Under-forecast by all techniques	Under-forecast by all techniques
Rapid Intensification	Under-forecast by all techniques (overall poor performance)	Under-forecast by all techniques (overall poor performance)
Decay-Reintensification Cycle	Overall poor performance	Overall poor performance

The SHF5 and SHIPS techniques have no information as to landfall timing and thus over-forecast the intensity during the decay phase. Although the dynamical models will decay the storm over land if the track forecast brings the storm over land at the correct time, the tendency to over-forecast by these models (Table 4.4) indicates the timing of the landfall forecast is more likely to be late than early. The landfall decay aspect did not apply in the eastern North Pacific for this sample, so none of the techniques (or the NHC) excels in that basin (Table 4.4). All of the techniques tend to over-forecast the intensity during the decay phase in that basin. Although the NHC performance is nearly as good as the DSHIPS in the Atlantic, it can not be said that NHC added value to the best technique in either basin.

Two other key aspects of the performance of the various techniques during Phase III are summarized in Table 4.5. The timing of the decay to 45 kt is important for resuming normal activities or beginning recovery operations. In both the Atlantic and eastern North Pacific, all of the techniques tended to decay the tropical cyclones too slowly. Rapid decay events were defined as a decrease of at least 30 kt in 24 h. The DSHIPS technique performed very well in predicting these events in the Atlantic because of its landfall decay factor (Table 4.5). Because landfall decay was not a factor for this sample of eastern North Pacific cyclones, DSHIPS did not perform any better than the other techniques in predicting rapid decay events in that basin, and the overall performance of all techniques was poor.

Table 4.4 Summary of Phase III intensity predictions for the Atlantic and eastern North Pacific during the 2003 and 2004 seasons.

Trend	**Atlantic**	**Eastern North Pacific**
Best technique	DSHIPS	None of the techniques or the NHC excels
Over-forecast	All other techniques	All techniques
Under-forecast	None	None
Does NHC add value compared to best technique?	No	No

Table 4.5 Summary of forecast performance for various aspects during Phase III for the Atlantic and eastern North Pacific during the 2003 and 2004 seasons.

Trend	**Atlantic**	**Eastern North Pacific**
Decay Prior to 45 kt	All techniques decayed TCs too slowly	All techniques decayed TCs too slowly
Rapid Decay	DSHIPS performed exceptionally well (13 of 15 events forecast)	Overall poor performance by all techniques

B. HYPOTHESIS VALIDATION

Blackerby (2005) analyzed the western North Pacific intensity guidance techniques during 2003 and 2004 using the same conceptual models in Figures 2.1 and 2.2. The hypothesis in this research was that similar techniques and numerical models used for intensity forecasting in the Atlantic and eastern North Pacific will have a similar performance in the western North Pacific. Thus, the comparable climatology and persistence SHF5 and ST5D techniques, the comparable statistical-dynamical SHIPS and STIPS techniques, and the GFNI models should perform similarly in all three basins. After each step in the evaluations of the intensity techniques in the Atlantic and eastern North Pacific, a comparison was made with the Blackerby (2005) evaluations for the western North Pacific. These comparisons are summarized in Table 4.6.

Table 4.6 Summary of the performance of comparable techniques/model in the Atlantic, eastern North Pacific, and western North Pacific during the 2003 and 2004 tropical seasons.

Climatology and Persistence (ST5D and SHF5) have a tendency to: • over-forecast during Phase I in all three basins • over-forecast during Phase II in the Atlantic and eastern North Pacific • over-forecast during Phase III in all three basins
Statistical-dynamical (SHIPS and STIPS) have a tendency to: • over-forecast during Phase I in all three basins • over-forecast during Phase II in all three basins • over-forecast during Phase III in all three basins
Dynamic model (GFNI) have a tendency to: • Not evaluated during Phase I in the Atlantic and eastern North Pacific due to a small sample size. Under-forecast in western North Pacific. • under-forecast during Phase II in the Atlantic and eastern North Pacific • over-forecast during Phase III in all three basins

The climatology and persistence technique ST5D in the western North Pacific and the SHF5 in the Atlantic and eastern North Pacific tend to fit an “average” or normal intensity cycle to each situation. Consequently, it is expected that these techniques will tend to predict the transition to 35 kt too early in Phase I, have a normal intensity curve in Phase II (see Figure 2.2), and an average decay rate in Phase III that will lead to an intensity that is too high. Thus, the only unusual entry in Table 4.6 for these techniques is the over-forecast during Phase II in the Atlantic and eastern North Pacific. This entry means that in this sample the SHF5 forecasts had more of a tendency to over-forecast the intensity when the intensity forecast errors exceeded +/- 10 kt during Phase II.

The statistical-dynamical techniques SHIPS in the Atlantic and eastern North Pacific and STIPS in the western North Pacific have the same tendency as the climatology and persistence techniques to over-forecast the formation Phase I and to decay the storms too slowly in Phase III. The landfall decay aspect of the DSHIPS corrects this aspect in the Atlantic, and likely would do so in the eastern North Pacific had more landfalls occurred in this sample. The decay version of STIPS in the western North Pacific was not evaluated by Blackerby (2005). The “over-forecast during Phase II” entry in Table 4.6 should be interpreted as for the SHF5 as described above, i.e., more over-forecasts than under-forecasts when the errors exceed +/- 10 kt.

The same dynamical model GFNI is integrated in all three basins. However, due to the small sample size, the GFNI model was not evaluated during Phase I in the Atlantic and eastern North Pacific (Table 4.6). The “under-forecast in Phase II”

entry for the GFNI in Table 4.6 indicates that this model is more likely to intensify the Atlantic and eastern North Pacific storms too slowly in Phase II. The GFNI has a tendency to maintain the intensity too high during the decay stage, perhaps because the convection is too strong as the storm recurves.

Thus, the hypothesis is generally validated in that similar techniques and numerical models used for intensity forecasting in the Atlantic and eastern North Pacific during the 2003 and 2004 season did have a similar performance in the western North Pacific.

C. FUTURE WORK

Tropical cyclone intensity guidance techniques are continuously updated. For example, the SHIPS and DSHIPS forecasts are verified at the end of each season to determine which predictors need modification, removal, or if an additional predictor should be added for the next season (see Appendix B). For the dynamical models, the model physics and horizontal/vertical resolution may be changed after each season (Appendix D and E). Thus, it is important that the accuracy of the intensity techniques continue to be evaluated either annually or bi-annually to provide operational forecasters guidance on intensity technique performance.

Because of the DSHIPS technique excellent performance during the decay phase in the Atlantic, evaluation is needed of how each technique predicts intensity for TCs that move over land during the decay phase. The empirical decay formula in DSHIPS needs to be compared with the full physics dynamical models and observations of decay. The effectiveness of the landfall decay factor for TCs that move over a peninsula or an island and return to the ocean should be evaluated.

Since intensity forecasting errors are inherently affected by track forecast errors, the intensity forecast errors arising from an inaccurate track forecast should be explored to determine what correlation exists between the two errors. Such a study of the intensity errors caused by track forecast errors could give guidance as to where investment in reducing track errors may also have a payoff in reduced intensity errors.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A: STATISTICAL HURRICANE INTENSITY FORECAST FIVE-DAY MODEL (SHF5)

A. TECHNIQUE DESIGN

The SHF5 is a 5-day tropical cyclone intensity forecast model derived from climatology and persistence (Knaff et al. 2003). Climatology refers to a forecast based on climatological averages and persistence is a forecast that the current intensity trend will continue throughout the forecast period. The SHF5 is one of the “control” forecasts used by the National Hurricane Center (NHC) in the Atlantic and eastern North Pacific Ocean basin to assess the skill of other intensity forecast models.

The model is developed using “best-track” data bases from each basin, i.e., the past history of all tropical cyclones. The data base contains the date, time, and location of tropical cyclones and subtropical cyclones that reached an intensity greater than 34 kt at sometime in its life cycle. The Atlantic SHF5 model was developed using storms from 1967 to 1999, but does not include non-developing depressions, i.e., a tropical depression that never reached tropical storm or subtropical storm status. The eastern North Pacific version was developed with storms from 1975 through 1999, and like the Atlantic version, does not include non-developing depressions in the best-track data.

The best-track data are used to create a set of multiple linear regression equations for the prediction. Each SHF5 has seven primary independent variables (predictors) and 28 secondary independent variables that are derived from the squares and cross products of the seven primary variables. The dependent variable that is predicted is the change in intensity each 12 h from the initial value. Except for how the date predictor is normalized, all predictors are the same in both basins. These seven primary predictors include date, latitude, longitude, zonal and meridional translation speeds of the storm, current intensity, and 12 h change in intensity. Of the possible 35 potential predictors only 15 are utilized in any of the forecast equations for the Atlantic and 13 for the eastern

North Pacific. The predictors are chosen using a combination of forward, backward, and stepwise regression schemes. Once chosen, the same predictors are used in the equations through the forecast period, except for the 108 h and 120 h time periods in the Atlantic and 84 h and 120 h time periods in the eastern North Pacific for which different predictors are entered in the regression equations.

B. TECHNIQUE CHARACTERISTICS

Knaff et al. (2002) discuss operational comparisons to 72 h with SHF5 and its Statistical Hurricane Intensity FORecast (SHIFOR) predecessor, using independent cases during 2000 and 2001. The mean absolute error is similar for both basins and both models in 2000 and 2001. However, the SHF5 biases (mean error) for 2001 in the Atlantic and for both years in the eastern North Pacific are larger and positive, suggesting that SHF5 intensity forecasts are on average larger than for SHIFOR. Also the SHF5 technique tends to predict greater intensity changes than does SHIFOR.

Knaff et al. (2002) also examined the error characteristics of SHF5 versus a simple persistence forecast through the 120-h period for 1997 to 2000. The Atlantic SHF5 has a tendency to under-forecast intensities through the 120 h time period. The eastern North Pacific bias is small through 60 h and then over-forecasts at longer leads, although not as much as in the Atlantic. No updates to SHF5 since its development in 2002 were found in the literature.

C. SUMMARY

The SHF5 is used by NHC as a no-skill threshold for tropical cyclone intensity forecasts in the Atlantic and eastern North Pacific basins. The model uses historical best-track track and intensity information to create a set of multiple linear regression equations for the intensity change prediction. Although the SHF5 has been noted to have certain forecast error tendencies, the model produces forecast error statistics that are statistically similar to its predecessor

SHIFOR (Knaff et al. 2002). It is important to note that even though the SHF5 is mainly used as a control model, it can be used to provide operational intensity forecasts.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B: STATISTICAL HURRICANE INTENSITY PREDICTION SCHEME (SHIPS)

A. TECHNIQUE DESIGN

The SHIPS is a statistical-dynamical intensity prediction technique used operationally by the National Hurricane Center (NHC) for the Atlantic basin since 1991 and the eastern North Pacific basin since 1996 (DeMaria et al. 2004). It combines climatological, persistence, and atmospheric predictors (vertical wind shear, sea-surface temperature, etc.) to forecast intensity changes using a multiple regression technique. The SHIPS provides intensity forecasts out to 120 h using the NHC track forecast from the previous six-hour forecast cycle that has been translated to match the current position of the storm.

A multiple linear regression approach is used to develop the equations in which the dependent variables are the intensity changes from zero to the forecast time (0-12 h, 0-24 h, etc.) through 72 h and the independent variables are parameters shown to be important for intensity changes (DeMaria et al. 2004). The synoptic predictors are evaluated from the National Centers for Environmental Prediction (NCEP) global model along a forecast track in real time. Predictors must be significant at the 1% level for at least one forecast period and the same predictors are used at each forecast interval for consistency. The same predictors, but with different coefficients, are used for the Atlantic and the eastern North Pacific basin. Each year the SHIPS forecast is verified at the end of the Atlantic tropical season to determine from the expanded data set which predictors need modifications, removal, or if an additional predictor should be added for the next tropical season. All predictors must still pass the 1% statistical significance test for at least one forecast interval if a new predictor is added or removed. Table B.1 lists the predictors used in the operational SHIPS model from 1997 to 2003. The first seven predictors listed have been used continuously in the SHIPS model from 1997 to 2003 and are presumed to be the most important predictors. DeMaria and Kaplan (1999) states that the difference between the maximum possible intensity and the current intensity is the most important predictor in the SHIPS model.

In 2004, three satellite predictors were added to the Atlantic version of SHIPS involving GOES imagery and TOPEX/Poseidon altimetry (DeMaria et al. 2004). The ocean heat content (OHC), which helps determine the ocean influence on intensity changes, was estimated from the TOPEX/Poseidon altimetry. The normalized predictor OHC above 50 kJ/cm^2 was utilized because the OHC usually exceeds 50 kJ/cm^2 in the Caribbean, the Loop Current, in warm-core eddies in the Gulf of Mexico, and along the Gulfstream (DeMaria et al. 2004). Two predictors from GOES channel 4 infrared (IR 10.7um) imagery provide information about the structure of the deep convection near the storm center. The predictors are the cold (blackbody temperature T_B < 253K) IR pixel count from 50 to 200 km from the storm center and the IR T_B standard deviation averaged from 100 to 300 km, which helps account for the variability of the T_B standard deviation as a function of storm intensity.

B. TECHNIQUE CHARACTERISTICS

The SHIPS forecasts have statistically significant skill for the period between 1997 and 2003 relative to a climatological-persistence technique called Statistical Hurricane Intensity FORecasts (SHIFOR) from 12 h to 72 h in the Atlantic and at 48 and 72 h in the eastern North Pacific. By contrast, the Geophysical Fluid Dynamics Laboratory (GFDL) numerical model did not have any significant forecast skill relative to SHIFOR during the same time period. The four- and five-day SHIPS forecasts were not skillful in the Atlantic but had modest skill in the eastern North Pacific. The 2004 SHIPS model has also shown forecast skill improvements with the addition of the new satellite inputs by as much as 9% at 48 h in the eastern North Pacific and 2% in the Atlantic (Figure B.1).

C. SUMMARY

The SHIPS technique has become one of the primary guidance tools that the NHC uses for their operational intensity forecasts. The SHIPS has skill compared to SHIFOR during 1997 through 2003. The new 2004 GOES and TOPEX/Poseidon altimetry predictors have resulted in operational intensity forecast improvements by as much as 9%.

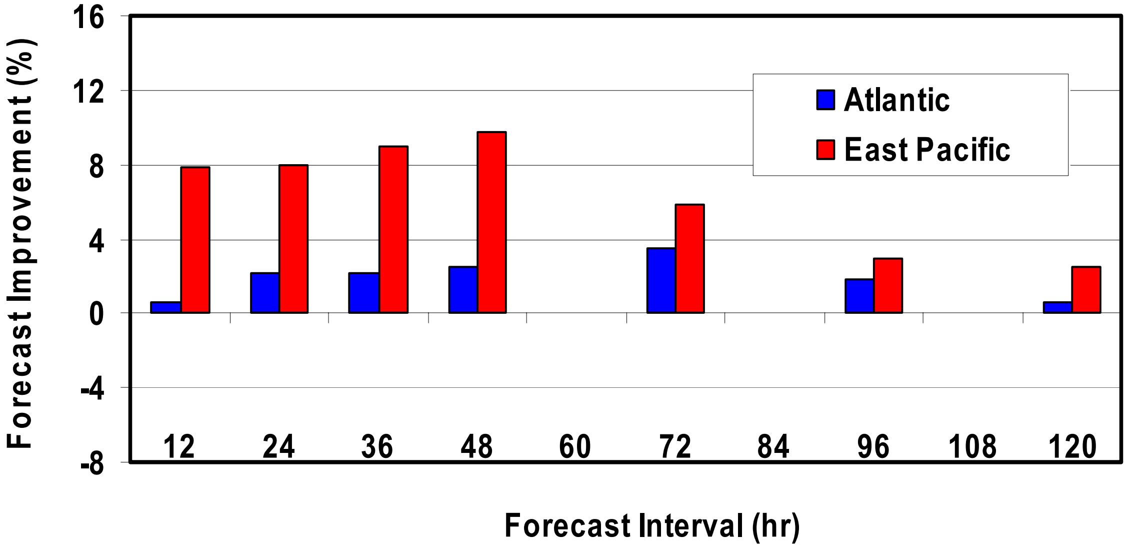


Figure B.1 SHIPS model improvements with satellite input for the 2004 operational forecast (from Mainelli et al. 2005).

Table B.1 Predictors used in the SHIPS operational model from 1997 to 2003.

	Predictor
1	Maximum potential intensity – current intensity
2	850 – 200 hPa vertical shear
3	Vertical shear times sine of storm latitude
4	Maximum wind change during the past 12 h
5	850 hPa relative vorticity
6	200 hPa temperature
7	Square of potential – current intensity
8	200 hPa divergence
9	200 hPa eddy momentum flux convergence (static)
10	200 hPa eddy momentum flux convergence (time dependent)
11	Absolute value of (Julian Day – peak season value)
12	Gaussian function of (Julian Day – peak value)
13	200 hPa zonal wind
14	Zonal component of storm motion
15	850 – 200 hPa relative humidity
16	500 – 300 hPa relative humidity
17	850 hPa relative vorticity
18	Surface – 200 hPa θ_e deviation of lifted parcel
19	Initial maximum winds
20	Pressure level of storm steering
21	Initial intensity time shear

APPENDIX C: DECAY STATISTICAL HURRICANE INTENSITY PREDICTION SCHEME (DSHIPS)

A. TECHNIQUE DESIGN

The DSHIPS technique (DeMaria et al. 2004) was developed in 2000 and is identical to the SHIPS model with the exception of applying an empirical decay model to account for the intensity change in storms that move over land. The procedure for using DSHIPS starts with the SHIPS model. The forecast track is interpolated to 1 h intervals and the SHIPS technique is applied to the entire track. An exponential decay formula is then applied to the portion of the track that is forecast to move over land. If the storm is forecast to move back over water, the remaining portion of the prediction is adjusted so that the intensity changes are the same as for the unadjusted SHIPS forecast.

The decay model reduces the maximum winds by multiplying the forecast value by 0.9 at each hour after landfall to account for land-sea differences in surface roughness. If the storm moves back over water, the forecast intensity is divided by 0.9 each hour, which further increases the storm intensity. The DSHIPS model also takes into account the forecast location of the storm when applying decay coefficients. When the forecast track is over land south of 30°N, the decay coefficients for the southeast U.S. are applied, the New England coefficients are used above 35°N, and a linear combination of the coefficients is used between the two latitudes (DeMaria et al. 2004).

B. TECHNIQUE CHARACTERISTICS

The reliability of DSHIPS depends critically on the accuracy of the forecast storm track, which determines if and when the storm is over land and DSHIPS will be used to modify the SHIPS forecast. DeMaria et al. (2004) determined that there is a positive correlation between the track and intensity errors. However, the DSHIPS forecasts have shown improvements relative to the SHIPS forecast for the Atlantic and the eastern North Pacific from 2000-2003. The Atlantic basin forecasts have improved by approximately 15% at 24 - 48 h and had statistically-significant improvements to 72 h. The Atlantic intensity forecasts are not improved at day 4 or 5, which DeMaria et al. (2004) thinks is likely due to the large landfall track forecast errors at these longer ranges, which then introduces a source of error that offsets any improvement due to the inclusion of land effects. The eastern North Pacific forecast improvements were 3% or less and were not statistically significant. The differences in improvement between the Atlantic and the eastern North Pacific are due to the smaller chance a storm will move over land in the eastern North Pacific, which means that DSHIPS adjustment will be less significant.

During the 2004 tropical cyclone season in the Atlantic basin, many storms moved over or near small islands and peninsulas, which made tendencies from DSHIPS more valuable. Mainelli et al. (2005) found that the DSHIPS model had a tendency to decay storms too much over narrow land masses and the decay rate was not proportional to the amount of circulation area over land.

C. TECHNIQUE SUMMARY

The DSHIPS technique adds an empirical decay model to the SHIPS technique to account for the decay of storms over land. This decay model has improved the operational intensity forecast and DSHIPS has more skill relative to SHIFOR than SHIPS since it was introduced operationally in 2000. However, there are still improvements needed for DSHIPS as the empirical decay model is too simple compared to the actual physical processes that occur when a storm moves over land. Additionally, the operational track forecasts are not perfect and there will inherently be errors in DSHIPS, especially for the 4- and 5- day forecasts when the track errors are largest.

APPENDIX D: GEOPHYSICAL FLUID DYNAMIC LABORATORY MODEL – (GFDI) INTERPOLATED

A. MODEL DESIGN

Since 1995, the GFDL model has been used operationally by NHC in both the Atlantic and the eastern North Pacific Ocean basins. During the 2002-2004 seasons, the dynamical model employed a double-nested, movable-mesh system to predict the structure and track of the tropical storm. The primitive equations governing the hurricane model are expressed in spherical coordinates (longitude, latitude), and sigma in the vertical to describe the time tendency of wind, temperature, pressure, and the water vapor mixing ratio (Kurihara et al. 1998).

The nested model utilizes two meshes to resolve the interior structure of the storm and the environment affecting the tropical cyclone. The outermost mesh has a 1/2-degree resolution and covers an area of 75-degrees longitude and latitude while the inner mesh has a resolution of 1/6-degree and extends 11-degrees in both latitude and longitude. The double-nested model was a change from the GFDL models prior to 2002 that incorporated a triply-nested mesh. The change in grid configuration was designed to better represent the interaction between the vortex and the environment by including: (i) an increase in resolution of the outer mesh from 1-degree resolution to a 1/2-degree resolution; (ii) an extension of the innermost (1/6-degree resolution) grid area from 5-degrees to 11-degrees latitude and longitude; and (iii) eliminating the third (medium) mesh of 1/3-degree resolution (Morris Bender, email August 8, 2005). The GFDL now also has 42 sigma levels with higher vertical resolution in the planetary boundary layer of the tropical storm.

The atmospheric model physics include diffusion, surface fluxes, radiation effects, a convective adjustment, and the prediction of land surface temperature. The effects of vertical diffusion are estimated with a level 2.5 Mellor-Yamada turbulent closure scheme and a nonlinear viscosity scheme is used to estimate the horizontal diffusion. The vertical fluxes of momentum and heat across the ocean and land surfaces are calculated using the Monin-Obukhov scheme. Radiation effects are evaluated by the Schwarzkopf and Fels (1991) infrared and Lacis and Hansen (1974) solar radiation parameterizations (Kurihara et al. 1998). The cumulus parameterization for the convective adjustment and a bulk subsurface layer scheme (Tuleya 1994) for predicting the land surface temperature are also utilized.

In 2001, the GFDL hurricane model was coupled with a high-resolution version of the Princeton Ocean Model (POM), which is a sigma vertical coordinate system, primitive-equation ocean model with three geographic domains (Gulf of Mexico, western Atlantic, and central Atlantic) of 1/6-degree resolution each. The three-dimensional ocean model has complete thermohaline dynamics, a turbulence sub-model, a free surface, and is formulated with an ocean bottom-following coordinate. The POM is initialized with the monthly averaged profiles of temperature and salinity called Generalized Digital Environmental Model (GDEM) that is produced by the Naval Oceanographic Office.

Recent upgrades in the GFDL model physics include an adoption of the Simplified Arakawa-Schubert (SAS) convective parameterization and a non-local boundary layer scheme, assimilation of a more realistic Gulf Stream structure in the ocean model, and improvements to the filter algorithms that remove the vortex from the NCEP global analysis (Bender et al. 2003).

B. MODEL CHARACTERISTICS

The GFDL hurricane prediction system is continuously being upgraded in both the model physics and vortex initialization to improve the skill in intensity prediction. However, many intensity problems still need to be addressed, such as the GFDL tendency to over-intensify systems in sheared environments, to under-predict the intensity of weak systems, and a spin-down and spin-up during the first 12 hours of the forecast (Bender et al. 2003). To address some of these problems, the 2005 model will return to a triply-nested mesh by including 1/12-

degree resolution grid that covers an area of 5-degrees latitude and longitude. The 2005 GFDL model will also incorporate an improved axi-symmetric vortex with identical moist and boundary layer physics between the axi-symmetric vortex spin-up used in the initialization and the three-dimensional model used in producing the forecast (Morris Bender, email August 8, 2005).

C. SUMMARY

The GFDL is a dynamical, three-dimensional, coupled hurricane system prediction model that is being upgraded in its model physics. Despite improvements, the 2003 GFDL model performance shows little skill relative to SHIPS and DSHIPS (Bender et al. 2003). The 2005 proposed changes to the model including implementation of a new vortex initialization, increased horizontal resolution, modification of the storm-size parameter, and improved surface roughness for computation of 10-m winds may increase the model intensity prediction skill.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E: GEOPHYSICAL FLUID DYNAMIC LABORATORY MODEL – NAVY MODEL (GFNI) INTERPOLATED

A. MODEL DESIGN

The GFDN model is a version of the GFDL dynamical hurricane prediction model that is produced operationally by the U.S. Navy and provided to NHC for intensity guidance in the Atlantic and eastern North Pacific ocean basins. One of the main differences between GFDN and GFDL models is that the former uses the Navy Operational Global Atmospheric Prediction System (NOGAPS) analyses and forecasts for providing the initial and lateral boundary conditions, while the GFDL utilizes the NCEP Global Forecast System (GFS).

The GFDN is a primitive-equation model with a triply-nested computational grid. The outermost grid covers an area of 75-degrees longitude and latitude with ½-degree resolution. The outer mesh is centered on the initial storm position with its poleward rigid boundary at 60N in the Northern Hemisphere and at 60S when applied in the Southern Hemisphere. The middle and inner meshes move with the storm to evaluate the storm environment and inner core with 1/3-degree and 1/6-degree resolution. The model physics include: a simplified Arakawa-Schubert cumulus parameterization, surface flux calculation based on Monin-Obukhov formulation, and Schwarzkoph and Fels (1991) infrared and Lacis and Hansen (1974) solar radiation parameterization with interactive radiative effects of clouds and a diurnal radiation cycle.

The GFDN model is separated into three segments: preparation of initial conditions, forecast, and post-processing. First, the preparation segment generates the model initial conditions that are obtained from the NOGAPS analyses, which are then interpolated onto each of the three nested meshes. The original vortex is then removed from the NOGAPS analysis within a certain region around the vortex center, and a synthetic vortex that has both axisymmetric and asymmetric components is then inserted back into the environment. Next, the forecast segment runs the forecast model and produces the intensity forecast. Finally, the post-processing segment creates graphics and prepares the output for transmission.

B. MODEL CHARACTERISTICS

Rennick (1999) noted that the GFDN model has a tendency to intensify storms in areas of large vertical wind shear, has difficulty in predicting the timing and location of recurvature, and has larger errors for forecasts of weak tropical cyclones. The GFDN track performance diminished in 2004 compared to 2003 in the Atlantic basin (Lerner et al. 2005) and may have been the cause of the GFDN intensity model lack of skill relative to SHF5 during 2004 in the Atlantic basin. The problem has been found to be due to an error in the initialization of the model, and has been corrected in June 2005.

C. SUMMARY

Improvements in the GFDN model are continuing, especially after having detected the problem with the model performance in 2003 and 2004. The recent correction in the model initialization step may improve the intensity skill since it had little skill relative to SHF5 in 2004 and less skill when compared to DSHIPS through 72 h in 2004.

LIST OF REFERENCES

Bender, M. A., and I. Ginis, 2000: Real-case simulations of hurricane-ocean interaction using a high-resolution coupled model: Effects on hurricane intensity. *Mon. Wea. Rev*., **128**, 917-943.

Bender, M. A., I. Ginis, T. P. Marchok, Pan, H.L., Thomas, B., and R. E. Tuleya, 2003: A summary of upgrades to the operational GFDL hurricane model for 2003. [HYPERLINK "http://ams.confex.com/ams/pdfpapers/75131.pdf." (Current as of 20 Sep 05)].

Bender, M. A., 2005. Email from August 8, 2005. Email address Morris.Bender@noaa.gov

Blackerby, J. S., 2005. The accuracy of western North Pacific tropical cyclone intensity guidance. M.S. thesis, Dept. of Meteorology, Naval Postgraduate School, Monterey, CA 93943-5114, 127 pp.

Carr, L. E., III, and R. L. Elsberry, 1994: Systematic and integrated approach to tropical cyclone track forecasting. Part I. Approach overview and description of meteorological basis. Tech. Rep. NPS-MR-94-002, Naval Postgraduate School, Monterey, CA 93943-5114, 273 pp.

DeMaria, M., and J. Kaplan, 1999: An updated hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific. *Wea. Forecasting,* **14,** 326-337.

DeMaria, M., R. M. Zehr, J. P. Kossin, and J. A. Knaff, 2002: The use of GOES imagery in statistical hurricane intensity prediction. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 120-121.

DeMaria, M., M. Mainelli, L. K. Shay, J. A., Knaff, and J. Kaplan, 2004: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). Manuscript submitted to *Wea. Forecasting*.

Elsberry, R. L. 2004: A global technique for tropical cyclone track prediction – the systematic approach. Chap. 4, Tropical Meteorology lecture notes, Naval Postgraduate School, Monterey, CA 93943-5114, 139 pp.

Emanuel, K. A., 1995a: The behavior of a simple hurricane model using a convective scheme based on subcloud-layer entropy equilibrium. *J. Atmos. Sci.,* **52,** 3959-3968.

Franklin, J. L. 2005: 2004 National Hurricane Center verification report. Preprints, 57th *Interdepartmental Hurricane Conference*, Miami, FL. [Available online at http://www.ofcm.gov/ihc05/Presentations/01%20session1/s1-03franklin.ppt (Current as of 20 Sep 2005)].

Gross, J. M. 2004: North Atlantic and East Pacific track and intensity verification for 2003. Preprints, *58th Interdepartmental Hurricane Conference*, Charleston, S.C. NHC/TPC. [Available online at http://www.ofcm.gov/ihc04/linking_file_ihc04.htm (Current as of 20 Sep 05)].

Jeffries, R. A., and E. M. Fukada/JTWC, cited 2005: Consensus approach to TC forecasting. [Available online at https://www.npmoc.navy.mil/jtwc/atcr/2002atcr/ch6/chap6_page1.html (Current as of 20 Sep 05)].

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, C. R. Sampson, and J. M. Gross, 2003: Statistical, five-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting,* **18,** 80-92.

Kurihara, Y., R. E. Tuleya, and M. A. Bender, 1998: The GFDL hurricane prediction system and its performance in the 1995 hurricane season. *Mon. Wea. Rev*., **126**, 1306-1322.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the earth's atmosphere. *J. Atmos. Sci*., **31**, 118-133.

Lerner, J., Y. Wang, C. Dickerman, J. Tesmer, and A. Hergert, 2005: Tropical cyclone Forecasting Support at Fleet Numerical Meteorology and Oceanography Center, Monterey, CA. Preprints, *59th Interdepartmental Hurricane Conference*, Jacksonville, FL. NHC/TPC. [Available online at http://www.ofcm.gov/ihc05/linking_file_ihc05.html (Current as of 20 Sep 05)].

Mainelli, M., M. DeMaria, and L. Shay, 2005: Results from the incorporation of Ocean Heat Content and GOES data in the operational SHIPS model during the 2004 hurricane season. Preprints, *59th Interdepartmental Hurricane Conference*, Jacksonville, FL. NHC/TPC. [Available online at http://www.ofcm.gov/ihc05/linking_file_ihc05.html (Current as of 20 Sep 05)].

Rennick, M. A., 1999: Performance of the Navy's tropical cyclone prediction model in the western North Pacific basin during 1996. *Wea. Forecasting,* **14,** 297-305.

Sampson, C. R., A. J. Schrader, 2000: The automated tropical cyclone forecasting system (Version 3.2). *Bull. Amer. Meteor. Soc*., **81**, 1231-1240.

Schwarzkoph, M. D., and S. B. Fels, 1991: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *J. Geophys. Res*., **96**, 9075-9096.

Tuleya, R. E. 1994: Tropical storm development and decay: Sensitivity to surface boundary conditions. *Mon. Wea. Rev*., **122**, 291-304.

Willoughby, H. E., J. A. Clos, and M. G. Shoreibah, 1982: Concentric eyewalls, secondary wind maxima, and the evolution of the hurricane vortex. *J. Atmos. Sci*., **39**, 395-411.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Commander, Naval Meteorology and Oceanography Command
   Stennis Space Center, Mississippi

4. National Hurricane Center
   Miami, Florida

5. Professor Russell L. Elsberry
   Naval Postgraduate School
   Monterey, California

6. Mark A. Boothe
   Naval Postgraduate School
   Monterey, California

7. Fleet Numerical Meteorology and Oceanography Center
   Monterey, California

8. Naval Atlantic Meteorology and Oceanography Center
   Norfolk, Virginia

9. LT Tara D. Lambert
   NAVPACMETOC DET
   Whidbey Island, Washington

10. Morris Bender
    Geophysical Fluid Dynamics Laboratory, NOAA
    Preston, New Jersey

11. Captain Jason Blackerby
    Joint Typhoon Warning Center
    Pearl Harbor, Hawaii