NCEP NOTES

Prediction of Consensus Tropical Cyclone Intensity Forecast Error

JAMES S. GOERSS

IES, San Diego, and SAIC, Monterey, California

CHARLES R. SAMPSON

Naval Research Laboratory, Monterey, California

(Manuscript received 22 May 2013, in final form 16 December 2013)

ABSTRACT

The extent to which the tropical cyclone (TC) intensity forecast error of IVCN and S5YY, consensus models routinely used by forecasters at the National Hurricane Center and the Joint Typhoon Warning Center, respectively, can be predicted is determined. A number of predictors of consensus intensity forecast error, which must be quantities that are available prior to the official forecast deadline, were examined for the Atlantic and eastern North Pacific basins for 2008–11 and the western North Pacific basin for 2012. Leading predictors were found to be forecast TC intensity and intensity change, initial intensity and latitude of the TC, and consensus model spread, defined to be the average of the absolute intensity differences between the member forecasts and the consensus forecast. Using stepwise linear regression and the full pool of predictors, regression models were found for each forecast length to predict the IVCN and S5YY TC intensity forecast errors. Using the regression models, intervals were determined centered on the IVCN and S5YY forecasts that contained the verifying TC intensity about 67% of the time. Based on the size of these intervals, a forecaster can determine the confidence that can be placed upon the IVCN or S5YY forecasts. Independent data testing yielded results only slightly degraded from those of dependent data testing, highlighting the capability of these methods in practical forecasting applications.

1. Introduction

Consensus tropical cyclone (TC) intensity forecast aids formed using TC intensity forecasts from statistical models and regional numerical weather prediction models have become increasingly important in recent years as guidance for forecasters at both the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC). Two of these consensus forecast aids are IVCN, which is available for forecasting in the Atlantic and eastern North Pacific, and S5YY, which is available for forecasting in the western North Pacific. For this study, IVCN is a consensus that is computed when intensity forecasts from at least two of the following five models are available: the interpolated version of the Geophysical Fluid Dynamics Laboratory (GFDL) model (GFDI), the Statistical Hurricane Intensity Prediction Scheme (SHIPS) with an inland decay components (Decay-SHIPS or DSHP), the Logistic Growth Equation Model (LGEM), the intermediate Hurricane version of the Weather Research and Forecasting (WRF) model (HWFI), and the U.S. Navy–GFDL interpolated model (GFNI).[1] S5YY is a consensus that is computed when at least two of nine intensity forecast aids run for the western North Pacific are available. IVCN is well known in the tropical cyclone community, but S5YY is new as of 2012 and so we discuss the algorithm in detail in the appendix.

Although historically much of the TC consensus work has been focused on attempts to reduce ensemble mean forecast error through postprocessing (Goerss 2000; Vijaya Kumar et al. 2003; Reynolds et al. 2011; Sampson et al. 2008), others have attempted to predict probabilities

Denotes Open Access content.

Corresponding author address: Charles R. Sampson, NRL, 7 Grace Hopper Ave., Stop 2, Monterey, CA 93943-5502.
E-mail: sampson@nrlmry.navy.mil

DOI: 10.1175/WAF-D-13-00058.1

[1] The composition of IVCN can change from year to year, as can the composition of S5YY.

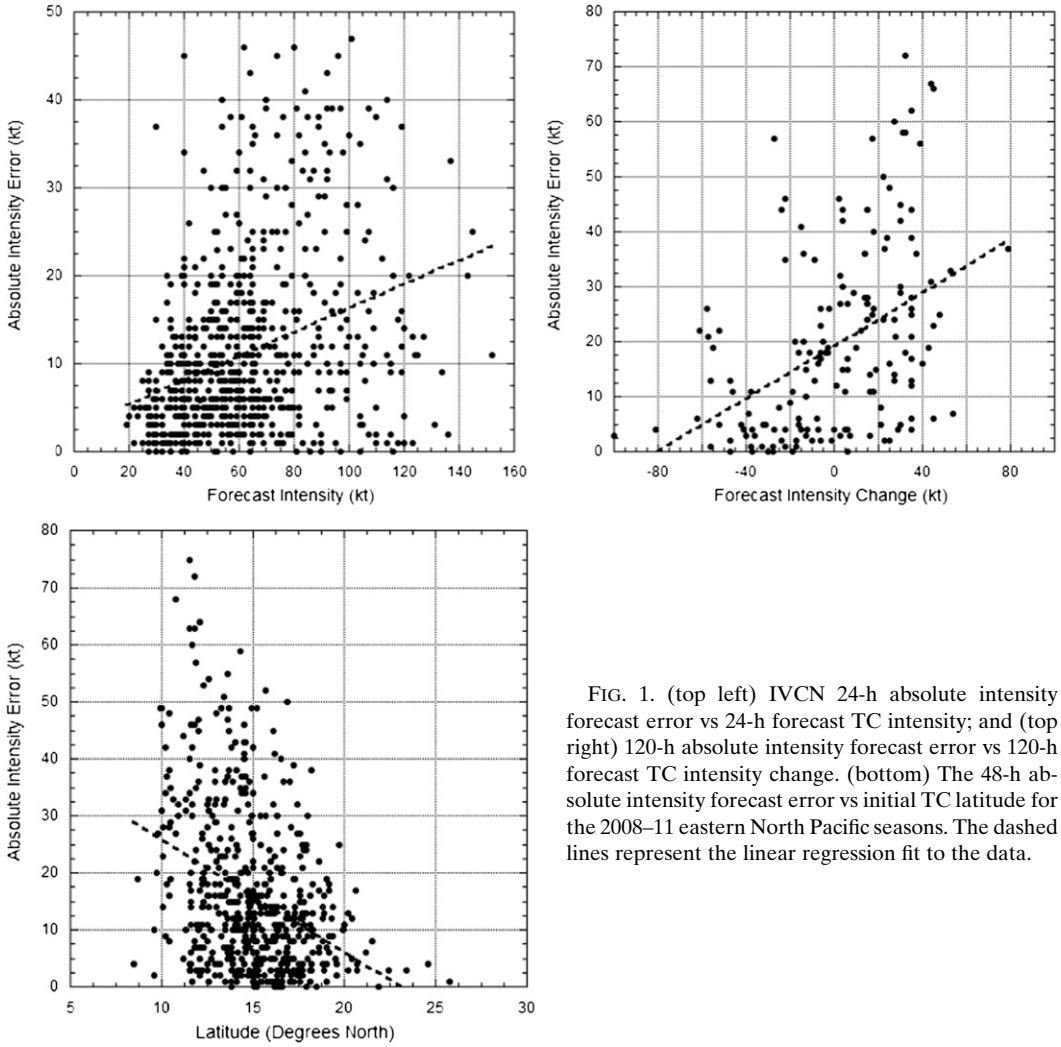

FIG. 1. (top left) IVCN 24-h absolute intensity forecast error vs 24-h forecast TC intensity; and (top right) 120-h absolute intensity forecast error vs 120-h forecast TC intensity change. (bottom) The 48-h absolute intensity forecast error vs initial TC latitude for the 2008–11 eastern North Pacific seasons. The dashed lines represent the linear regression fit to the data.

based on ensemble spread and other parameters available to forecasters (Weber 2005; Goerss 2007; Majumdar and Finocchio 2010). Almost all of this work has been focused on track. The purpose of this study is to determine to what extent the TC absolute intensity forecast error of the consensus models, IVCN and S5YY, can be predicted prior to the time when official forecasts must be issued. The techniques used for the Goerss predicted consensus error (GPCE; Goerss 2007) will be employed to predict consensus absolute intensity error. Predictors of consensus forecast error must be quantities that are available in the real-time files of the Automated Tropical Cyclone Forecasting System (ATCF; Sampson and Schrader 2000) prior to official forecast construction at the operational centers. Consensus model spread is defined to be the average of the absolute differences between the member intensity forecasts and the consensus intensity forecast. The possible predictors examined in this study are consensus model spread, forecast TC intensity and intensity change, initial TC intensity and position, TC speed of motion, and the number of members available to the consensus model. Forecast TC intensity and intensity change are determined using the interpolated official forecasts (OFCI and JTWI). All cases where an official intensity forecast was made and verified were used in this study.

In the next section, we describe how these predictors are used to estimate IVCN and S5YY TC forecast intensity errors for the three basins. In section 3, we outline the results of independent data testing of the technique. In the final section, we summarize the results of this research and discuss our future research plans.

2. Estimation of consensus model absolute intensity forecast error

First, we illustrate the relationships between some of the possible predictors and the IVCN and S5YY TC

absolute intensity forecast errors. For IVCN in the Atlantic basin, forecast intensity (INTF) was found to be the leading predictor for the 12–36-h forecasts with correlations ranging from 0.19 to 0.22. The leading predictor for the 48–120-h forecasts was forecast intensity change (INTC) with correlations ranging from 0.21 to 0.22. Initial TC latitude (LATI) was found to be the second-leading predictor for the 24- and 36-h forecasts.

For IVCN in the eastern North Pacific basin, forecast intensity was found to be the leading predictor for the 12- and 24-h forecasts with correlations of 0.32 and 0.37, respectively, and the second-leading predictor for the 36–72-h forecasts. Initial TC latitude was found to be the leading predictor for the 36–72-h forecasts with correlations ranging from −0.28 to −0.36 and the second-leading predictor for the 12- and 24-h forecasts. Forecast intensity change was found to be the leading predictor for the 96- and 120-h forecasts with correlations of 0.38 and 0.44, respectively, while consensus model spread (SPR) was found to be the second-leading predictor. Examples of the leading predictors for the eastern North Pacific are shown in the scatterplots displayed in Fig. 1. The 24-h INTF (Fig. 1, top left) and the 120-h INTC (Fig. 1, top right) were found to be positively correlated with forecast absolute intensity error with correlations of 0.33 and 0.44, respectively, while the LATI (Fig. 1, bottom) was found to be negatively correlated with correlation of −0.36. The 120-h consensus model spread (Fig. 2, top) was found to be positively correlated with forecast absolute intensity error with a correlation of 0.31.

For S5YY in the western North Pacific basin, initial TC latitude was found to be the leading predictor for the 24–72-h forecasts with correlations ranging from −0.27 to −0.34 and the second-leading predictor for the 12- and 96-h forecasts. Initial TC longitude (LONI) was found to be the leading predictor for the 12-h forecast with a correlation of −0.20 and the second-leading predictor for the 24–48-h forecasts. Consensus model spread was found to be the leading predictor for the 96- and 120-h forecasts with correlations of 0.27 and 0.24, respectively, and the second-leading predictor for the 72-h forecasts. The relationship between the 96-h consensus model spread and forecast absolute intensity error is shown in Fig. 2 (bottom).

Using stepwise linear regression (Draper and Smith 1966) and the pool of predictors from the Atlantic and eastern North Pacific basins for 2008–11, regression models were found for each forecast length to predict the IVCN TC absolute intensity forecast error. To avoid overfitting the dependent dataset, we required that

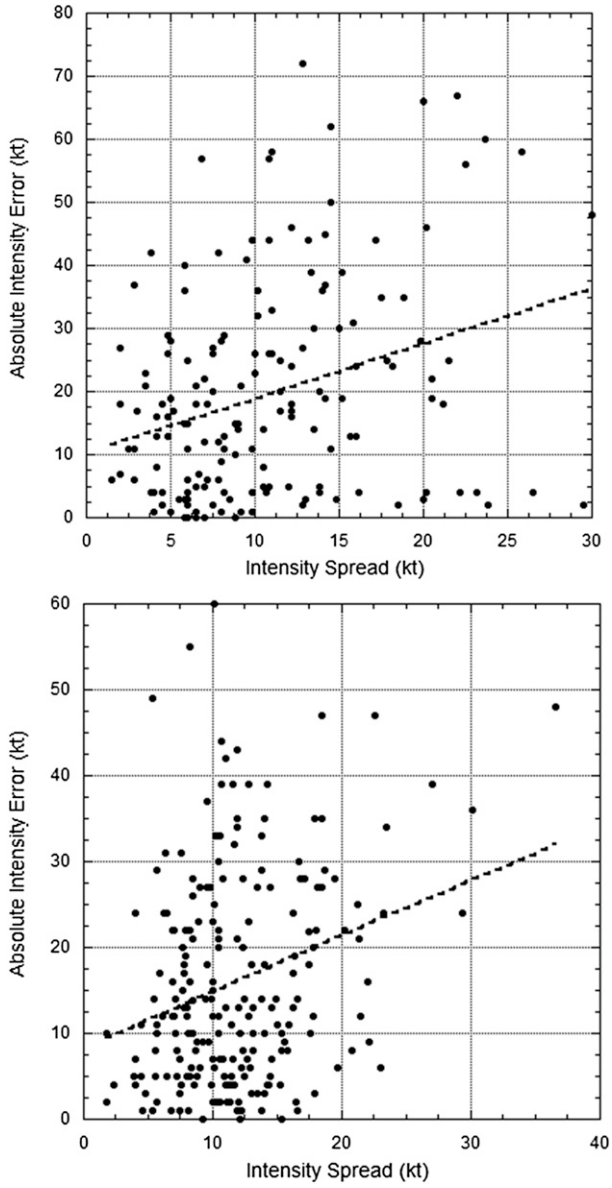

FIG. 2. (top) IVCN 120-h absolute intensity forecast error vs spread for the 2008–11 eastern North Pacific seasons and (bottom) S5YY 96-h absolute intensity forecast error vs spread for the 2012 western North Pacific season. The dashed lines represent the linear regression fit to the data.

a predictor explain at least 3% of the variance before allowing it to be used by the final regression equation. All of the final regression coefficients were found to be significantly different from zero at well above the 99% level. Similarly, using the pool of predictors from the western North Pacific basin for 2012, regression coefficients were found to predict the S5YY absolute intensity error. The regression equations and scatterplots displaying the relationship between IVCN absolute intensity forecast error and predicted error for the

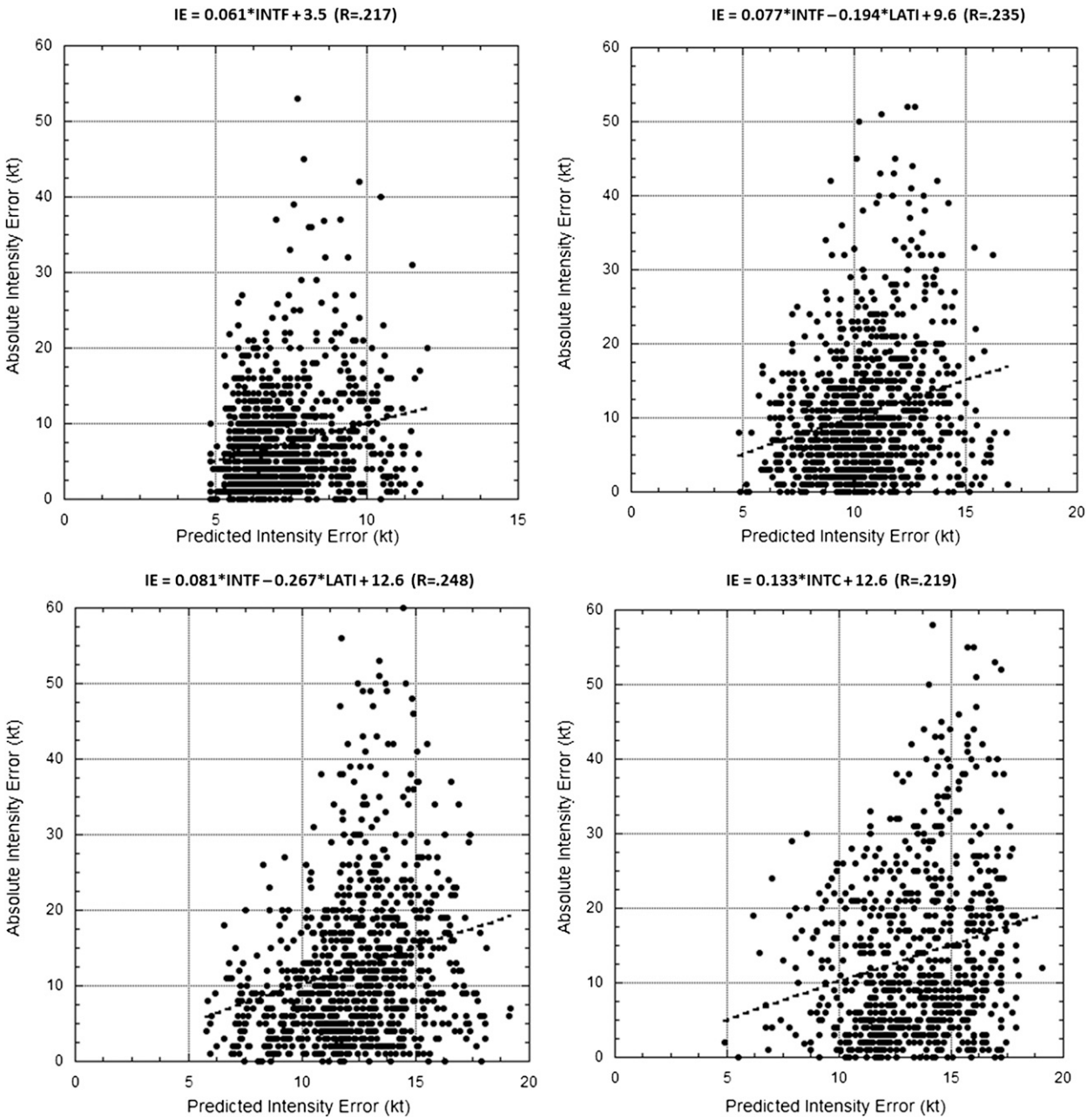

FIG. 3. IVCN (top left) 12-, (top right) 24-, (bottom left) 36-, and (bottom right) 48-h absolute intensity forecast error vs predicted error for the Atlantic basin (2008–11). The equation for the predicted error (IE) is shown at the top of each panel. The dashed lines represent the linear regression fit to the data.

shorter-range forecasts for the Atlantic and eastern North Pacific basins are shown in Figs. 3 and 4, respectively, while those for S5YY for the western North Pacific basin are shown in Fig. 5. The regression equations and the correlations of the regression fit are also displayed in these figures. The correlations for the Atlantic basin ranged from 0.22 to 0.25 while those for the eastern and western North Pacific basins ranged from 0.36 to 0.41 and 0.25 to 0.39, respectively. Using these linear regression models, the percent variance of IVCN TC absolute intensity forecast error that could be explained for the 2008–11 Atlantic seasons ranged from 5% to 6% for the shorter-range forecasts while that for the 2008–11 eastern North Pacific seasons ranged from 13% to 16%. The percent variance of the S5YY absolute intensity forecast error that could be

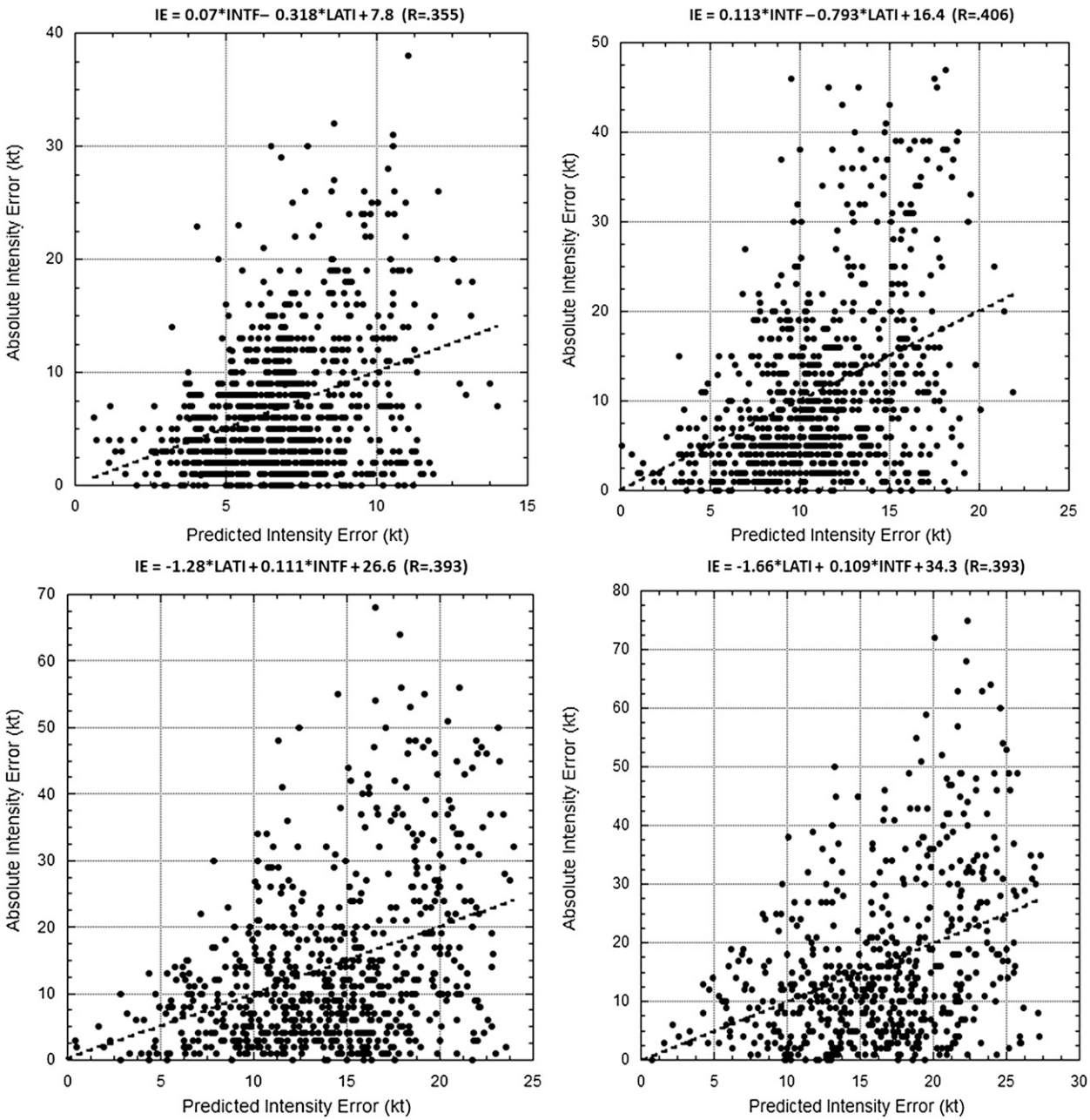

FIG. 4. As in Fig. 3, but for the eastern North Pacific basin (2008–11).

explained for the 2012 western North Pacific season ranged from 6% to 15%.

The regression equations and scatterplots displaying the relationship between IVCN absolute intensity forecast error and predicted error for the longer-range forecasts for the Atlantic and eastern North Pacific basins are shown in Figs. 6 and 7, respectively, while those for S5YY for the western North Pacific basin are shown in Fig. 8. The correlations for the Atlantic basin ranged from 0.21 to 0.22 while those for the eastern and western North Pacific basins ranged from 0.33 to 0.49 and 0.24 to 0.44, respectively. The percent variance of IVCN absolute intensity forecast error that could be explained for the 2008–11 Atlantic seasons ranged from 4% to 5% for the longer-range forecasts while those for the IVCN forecast error for the 2008–11 eastern North Pacific and for the S5YY forecast error for the 2012 western North Pacific seasons ranged from 11% to 24% and 6% to 19%, respectively.

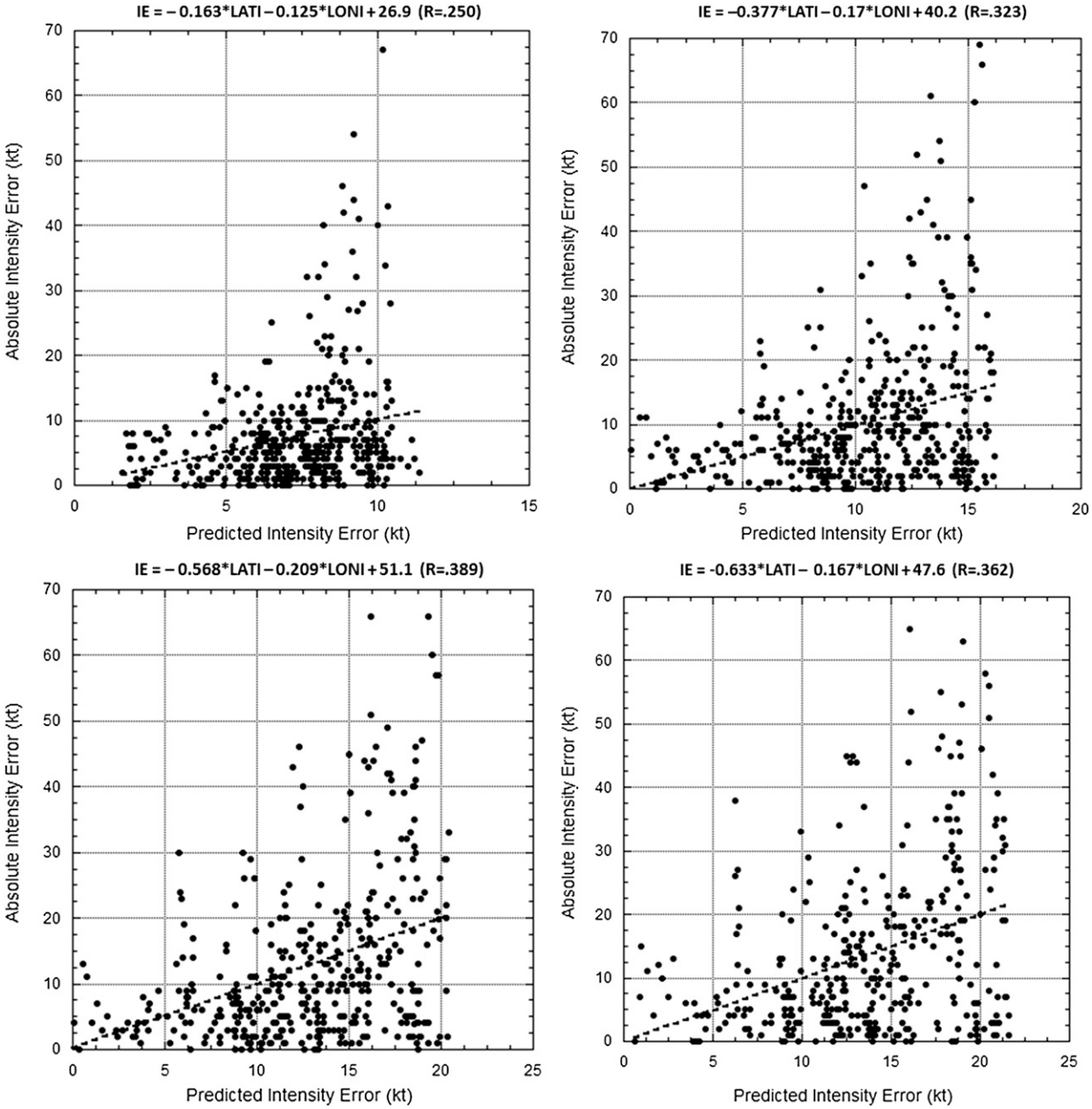

FIG. 5. As in Fig. 3, but for S5YY and for the western North Pacific (2012).

To put the percent variance of the consensus absolute intensity forecast error into context, we compare it with that for consensus track forecast error. For the 2012 season, the percent variance of the consensus track forecast error explained by the GPCE predicted track forecast error ranged from 12% to 23% and 6% to 19% for TVCN in the Atlantic and eastern North Pacific basins, respectively, and from 8% to 22% for the consensus version W (CONW) in the western North Pacific basin. For the 2008–11 seasons, the percent variance of the IVCN absolute intensity forecast error explained ranged from 4% to 6% and 11% to 24% for the Atlantic and eastern North Pacific basins, respectively. For the 2012 season, the percent variance of the S5YY absolute intensity forecast error explained ranged from 6% to 19%. Except for the Atlantic basin, the variances explained for intensity and track were quite comparable.

Next, we want to translate these results into a form that has meaning to NHC and JTWC forecasters. We would like to construct intervals centered on the

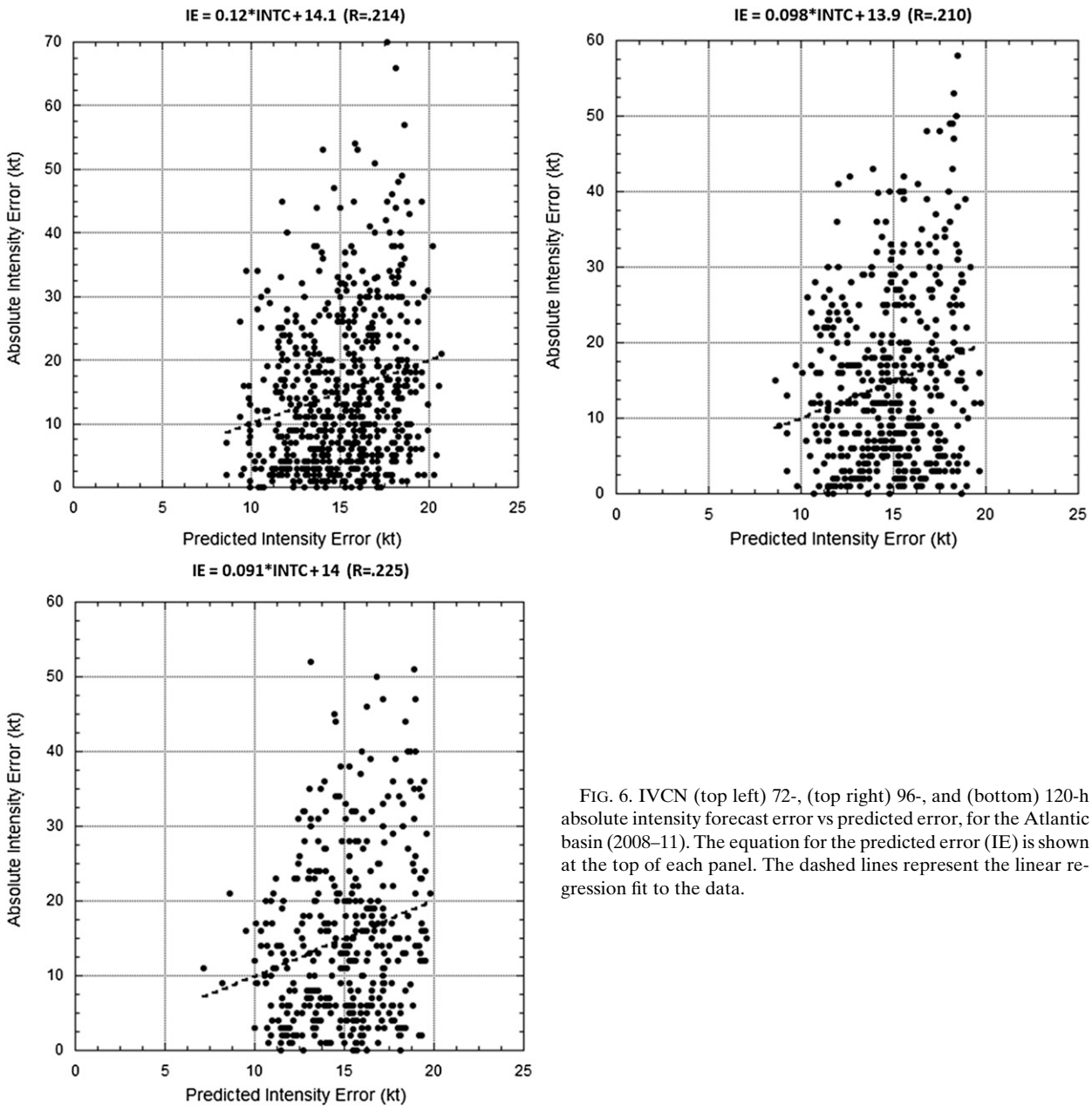

FIG. 6. IVCN (top left) 72-, (top right) 96-, and (bottom) 120-h absolute intensity forecast error vs predicted error, for the Atlantic basin (2008–11). The equation for the predicted error (IE) is shown at the top of each panel. The dashed lines represent the linear regression fit to the data.

consensus intensity forecasts that will contain the verifying TC intensity roughly 67% of the time. For the 2008–11 Atlantic hurricane seasons, intervals were determined whose half-widths were computed by adding a constant varying with forecast length to the predicted IVCN TC absolute intensity forecast error derived using the linear regression equations. The constants [1.2 kt (1 kt = 0.51 m s^{-1}) at 12 h, 1.6 kt at 24 h, 1.7 kt at 36h, 3 kt at 48 h, 3.4 kt at 72 h, 3.6 kt at 96 h, and 3.1 kt at 120 h] were chosen so that the verifying TC intensity would be contained within the interval centered on the IVCN forecast intensity roughly 67% of the time. The results of applying these intervals to the IVCN forecasts for the 2008–11 Atlantic seasons are summarized in Table 1. For this dependent sample for the Atlantic basin, the average size of the predicted half-width for each of these intervals is displayed in Table 1 along with the minimum and maximum size, the percentage of the time that the verifying intensity was included within the interval

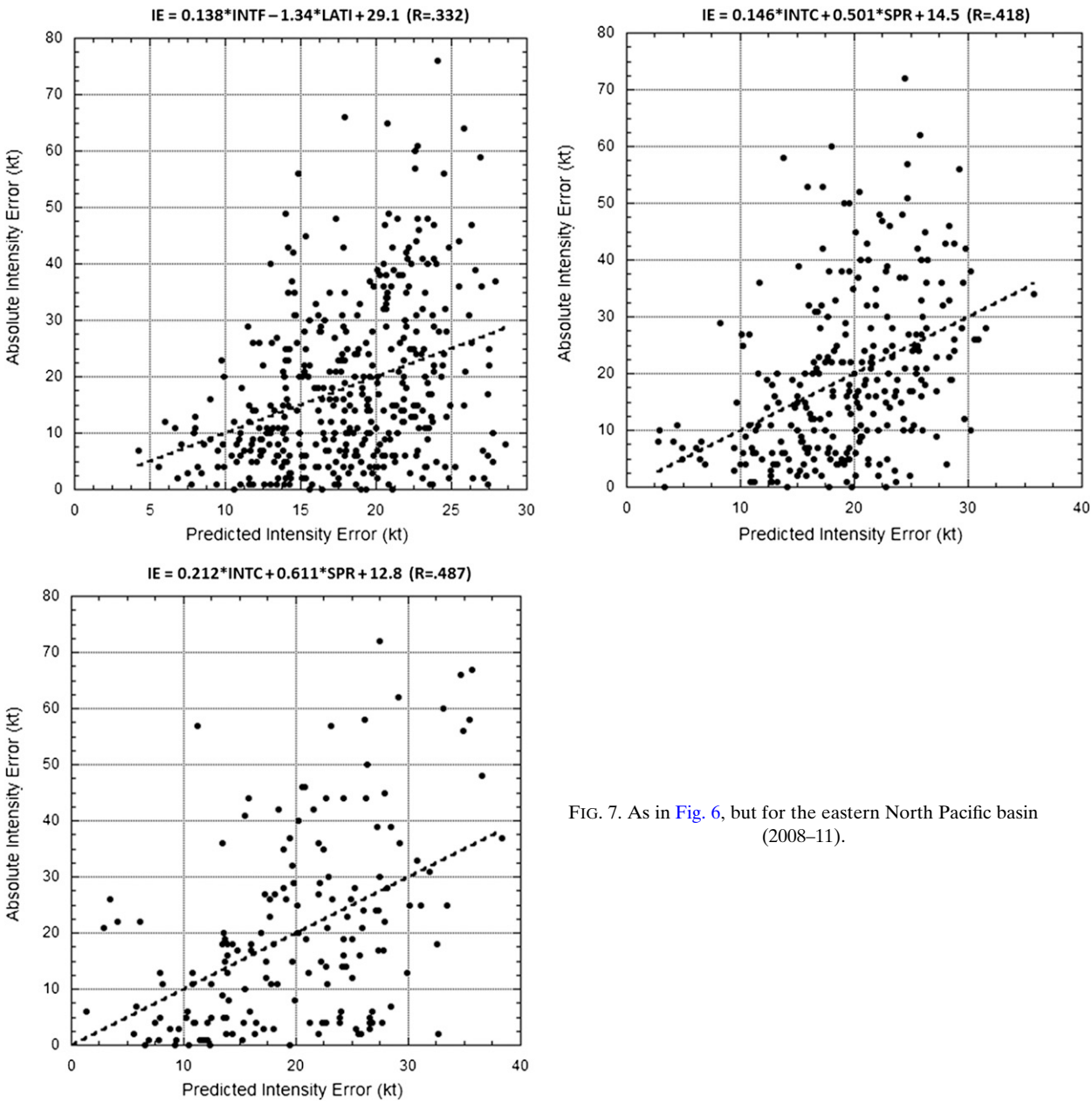

FIG. 7. As in Fig. 6, but for the eastern North Pacific basin (2008–11).

centered on the IVCN forecast intensity, and the number of forecasts. From Table 1 we can see, for example, that for the 24-h intensity forecasts the size of the confidence intervals determined range from ±6 to ±18 kt. Thus, we have a range in the size of the confidence intervals rather than a fixed confidence interval of ±12 kt that would be determined strictly based on the average error for the 2008–11 seasons. Based on the size of these confidence intervals, forecasters can place more or less confidence in the IVCN intensity forecasts.

Similarly, the constants chosen for the 2008–11 eastern North Pacific seasons for the respective forecast lengths were 1.6, 1.2, 2.8, 2.8, 4, 3.1, and 3.7 kt. The results of applying these intervals to the IVCN forecasts for the 2008–11 eastern North Pacific seasons are summarized in Table 2. And, finally, the constants chosen for the 2012 western North Pacific season for the respective forecast lengths were 0.5, 1.5, 1.4, 2.4, 2.9, 3.6, and 1.4 kt. The results of applying these intervals to the S5YY forecasts for the 2012 western North Pacific season are summarized in Table 3.

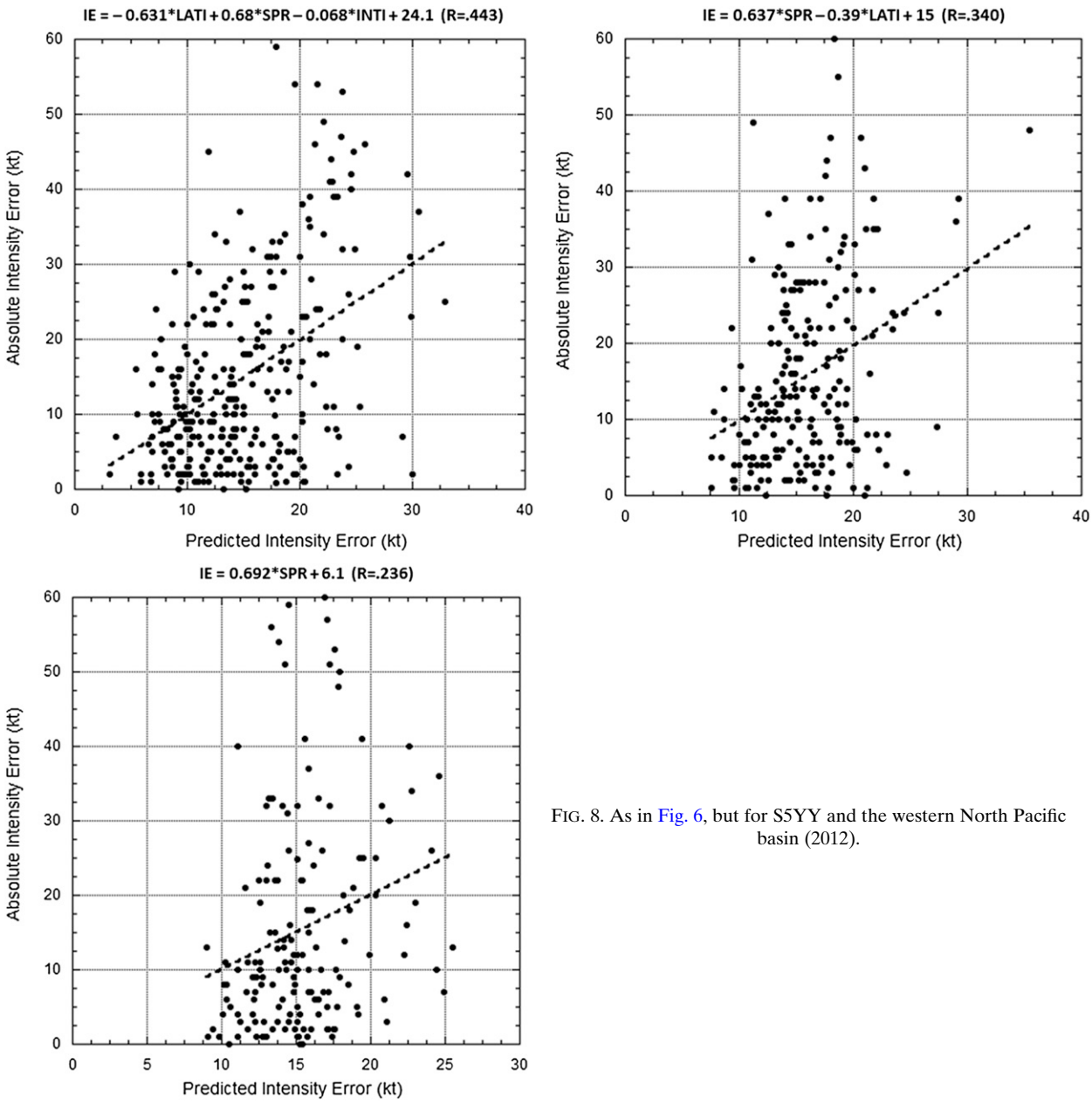

FIG. 8. As in Fig. 6, but for S5YY and the western North Pacific basin (2012).

3. Independent data testing

Evaluating performance on dependent data is good, but forecast guidance should be evaluated on independent data whenever possible. We cannot evaluate our western North Pacific guidance on independent data since S5YY has only been running for one season; however, we can do this for our IVCN guidance in the Atlantic and eastern North Pacific basins. For each forecast length, regression equations were computed using stepwise linear regression and the pool of predictors from the two basins for 2008–11. The predicted TC absolute intensity forecast errors were then computed for the 2012 Atlantic and eastern North Pacific seasons. For the Atlantic basin, the percent variance of the IVCN TC absolute intensity forecast error that could be explained for this independent sample ranged from 2% to 5% compared with 4% to 6% for the dependent sample. Just as was done previously, predicted intervals were constructed to be centered on the IVCN forecast intensities. The results of the independent testing for the 2012 Atlantic season are summarized in Table 4. We

TABLE 1. Verification summary for IVCN for the 2008–11 Atlantic seasons.

Forecast length (h)	12	24	36	48	72	96	120
Avg ½ range (kt)	8	12	14	16	19	18	18
Min ½ range (kt)	6	6	8	8	12	12	11
Max ½ range (kt)	13	18	21	22	25	24	23
Forecasts within range (%)	67	67	67	67	67	67	67
No. of forecasts	1183	1093	927	819	651	520	406

TABLE 3. Verification summary for S5YY for the 2012 western North Pacific season.

Forecast length (h)	12	24	36	48	72	96	120
Avg ½ range (kt)	7	12	14	16	17	19	17
Min ½ range (kt)	3	3	3	3	5	11	10
Max ½ range (kt)	11	18	22	24	35	39	29
Forecasts within range (%)	67	67	67	67	67	67	67
No. of forecasts	489	448	413	375	296	220	159

see that the predicted intervals were too large, as they included the verifying intensities from 71% to 81% of the time, compared with our target of 67%. Comparing Tables 1 and 4, we see that the average sizes of the half-widths for the confidence intervals for the various forecast lengths are quite similar, as is the range of sizes of the intervals. For the independent tests and operational implementation on the ATCF we set the minimum ½ range to 3 kt, which seems an appropriate minimum since the official forecasts are in increments of 5 kt. The primary reason for the overprediction is the unusually small size of the IVCN forecast errors for 2012. For the 12–48-h forecasts they ranged from 5.5 to 11 kt compared with a range of 7–13 kt for the 2008–11 dependent sample from which the regression equations were derived. For the 72–120-h forecasts, they ranged from 13 to 14 kt compared with a range of 15–16 kt for the dependent sample. This is not unexpected since we have observed the same thing when verifying the GPCE areas centered on consensus track forecasts. One final note on the Atlantic evaluation: we also went through the same exercise of excluding land cases with very little change in either the derived coefficients or the independent evaluation (not shown).

For the eastern North Pacific, the percent variance of IVCN TC absolute intensity forecast error that could be explained for the 2012 season ranged from 10% to 24% compared with 11%–24% for the 2008–11 dependent sample. The results of this independent testing are summarized in Table 5. Like we saw for the Atlantic, the predicted intervals were too large, as they include the verifying intensity from 73% to 85% of the time. For the 12–48-h forecasts, the IVCN forecast errors for 2012 ranged from 6 to 12 kt compared with a range of 7–16 kt for the dependent sample. For the 72–120-h forecasts, they ranged from 12 to 13 kt compared with a range of 18–19 kt for the dependent sample.

Comparing Tables 2 and 5, we again see that the results found from independent testing compare quite favorably with those found from dependent testing. We conclude from our independent testing that the regression equations derived from previous seasons appear to be stable and that coefficients derived from prior seasons can be effectively applied to the next season, which is how we intend to implement this aid in operations.

Finally, DeMaria et al. (2013) found that the track GPCE values, when used to stratify the track errors into terciles, have added value to their operational wind speed probability product. For both the Atlantic and eastern North Pacific basins for 2012, we have used the predicted ½ ranges to stratify the independent samples into terciles. The verifications based on this stratification are summarized for the respective basins in Tables 6 and 7. For example, from Table 4, we see that the IVCN 24-h intensity forecasts for the Atlantic basin were within the predicted range 76% of the time overall. From the tercile verification in Table 6 we see that for predicted ½ ranges less than 10 kt, from 10 to 11 kt, and greater than 11 kt the IVCN 24-h intensity forecasts were within the predicted range 71%, 74%, and 81% of the time, respectively. For the Atlantic and for the 12- and 24-h forecasts in the eastern North Pacific, none of the tercile percentages were more than 7% larger or smaller than the overall percentage for that forecast length. With the exception of the 96-h forecasts, none of the tercile percentages for the longer-range forecasts in the eastern North Pacific were more than 14% larger or

TABLE 2. As in Table 1, but for the eastern North Pacific.

Forecast length (h)	12	24	36	48	72	96	120
Avg ½ range (kt)	9	12	17	19	23	23	23
Min ½ range (kt)	3	3	3	3	8	11	3
Max ½ range (kt)	16	23	27	31	34	41	43
Forecasts within range (%)	67	67	67	67	67	67	67
No. of forecasts	919	809	699	602	429	300	195

TABLE 4. Verification summary for IVCN for the 2012 Atlantic season.

Forecast length (h)	12	24	36	48	72	96	120
Avg ½ range (kt)	8	11	12	16	18	18	18
Min ½ range (kt)	6	7	7	11	14	14	13
Max ½ range (kt)	11	14	17	20	24	24	23
Forecasts within range (%)	79	76	75	81	79	73	71
No. of forecasts	393	354	313	278	224	189	150

TABLE 5. Verification summary for IVCN for the 2012 eastern North Pacific season.

Forecast length (h)	12	24	36	48	72	96	120
Avg 1/2 range (kt)	8	11	16	18	21	22	21
Min 1/2 range (kt)	3	3	3	4	11	16	5
Max 1/2 range (kt)	14	20	26	29	32	32	38
Forecasts within range (%)	78	73	78	77	85	80	82
No. of forecasts	300	266	232	200	140	89	44

smaller than the overall percentage. The small sample sizes for the 96- and 120-h forecasts (89 and 44 cases, respectively) in this basin would certainly contribute to this variability. The performance in this analysis indicates that the terciles could be used in other applications such as the NHC wind speed probability product.

4. Summary and conclusions

The purpose of this study was to determine to what extent the TC absolute intensity forecast error of the consensus models, IVCN and S5YY, could be predicted prior to the forecast deadline. The possible predictors examined in this study were consensus model spread, forecast TC intensity and intensity change, initial TC intensity and position, TC speed of motion, and the number of members available to the consensus model.

For IVCN in the Atlantic basin, forecast intensity and forecast intensity change were found to be the leading predictors for the 12–36-h and 48–120-h forecasts, respectively. For IVCN in the eastern North Pacific basin, forecast intensity and forecast intensity change were found to be the leading predictors for the 12- and 24-h and the 96- and 120-h forecasts, respectively. Initial TC latitude was found to be the leading predictor for the 36–72-h forecasts. For S5YY in the western North Pacific basin, initial TC latitude and initial TC longitude were found to be the leading predictors for the 24–72-h and 12-h forecasts, respectively. Consensus model spread was found to be the leading predictor for the 96- and 120-h forecasts. Forecast intensity, forecast intensity change, and consensus model spread were found to be positively correlated with forecast absolute intensity error while initial TC latitude and longitude were found to be negatively correlated.

Using stepwise linear regression and the pool of predictors for the 2008–11 Atlantic and eastern North Pacific TC seasons, regression models were found to predict IVCN absolute intensity forecast error for each forecast length. Similarly for the 2012 western North Pacific TC season, regression models were found to predict S5YY absolute intensity error. The regression models explained 5%–6%, 13%–16%, and 6%–15% of the absolute intensity forecast error variance for the shorter forecast lengths (12–48 h) for the Atlantic, eastern North Pacific, and western North Pacific basins, respectively. For the longer forecast lengths (72–120 h), the regression models were found to explain only 4%–5% of the intensity absolute forecast error variance for the Atlantic basin but the variance explained for the eastern and western North Pacific basins ranged from 11% to 24% and 6% to 19%, respectively. Predicted confidence intervals were derived by adding a constant, which varied with respect to forecast length, to the absolute intensity forecast error predicted by the regression models. The additive constants were chosen so that the verifying TC intensity was contained within the interval centered on the consensus forecast intensity about 67% of the time for the dependent samples. For each forecast length and for each basin, the size of these predicted confidence intervals varied considerably. For example, for IVCN in the eastern North Pacific basin the half-widths of these confidence intervals varied from 3 to 16 kt for the 12-h forecasts and from 3 to 43 kt for the 120-h forecasts. This indicates that there is potential for this algorithm to provide dynamic confidence intervals for use in operational forecasting.

We also performed independent data testing when possible. Using the 2008–11 Atlantic and eastern North Pacific seasons as dependent datasets, regression models were derived and applied to the 2012 Atlantic and eastern North Pacific seasons. We found that the results from the independent testing compared quite favorably with those found from dependent testing. Thus, we concluded that we should be able to effectively use the regression models determined from previous seasons to produce guidance to be used during the next season.

TABLE 6. Tercile verification summary for IVCN for the 2012 Atlantic season. The predicted 1/2 ranges (kt) are in parentheses followed by percent of forecasts within range.

Forecast length (h)	12	24	36	48	72	96	120
Lower tercile	(<8) 74	(<10) 71	(<11) 72	(<16) 77	(<18) 85	(<18) 78	(<18) 77
Middle tercile	(8) 82	(10–11) 74	(11–13) 71	(16–17) 82	(18–19) 80	(18–19) 71	(18) 67
Upper tercile	(>8) 81	(>11) 81	(>13) 82	(>17) 82	(>19) 74	(>19) 71	(>18) 68

TABLE 7. As in Table 6, but for the eastern North Pacific.

Forecast length (h)	12	24	36	48	72	96	120
Lower tercile	(<8) 79	(<10) 79	(<15) 90	(<17) 91	(<19) 96	(<21) 57	(<19) 94
Middle tercile	(8) 85	(10–13) 74	(15–17) 83	(17–19) 81	(19–22) 80	(21–23) 88	(19–25) 69
Upper tercile	(>8) 75	(>13) 66	(>17) 66	(>19) 63	(>22) 80	(>23) 94	(>25) 83

We suspect that forecasters can use confidence intervals in real time to determine how much (or little) confidence can be ascribed to individual consensus intensity forecasts. We also hope this new guidance can be employed in the Monte Carlo Wind Speed Probability Product (DeMaria et al. 2009) to reduce or expand the wind probability plume for given situations. Efforts applying the GPCE track output to wind probabilities have proven successful (DeMaria et al. 2013) both in performance and acceptance by forecasters, which gives us some hope that the same is true for the intensity GPCE.

Acknowledgments. This research was funded by the Office of Naval Research and the Hurricane Forecast Improvement Program. The ATCF system is a trademarked product of the Naval Research Laboratory, Monterey, California. The authors wish to thank Ann Schrader and Mike Frost for their efforts with the ATCF and the forecasters at both the National Hurricane Center and Joint Typhoon Warning Center for their diligent management of archived tropical cyclone data.

APPENDIX

The S5YY Consensus

The first step in creating the S5YY intensity consensus is to construct interpolated track forecasts using numerical weather prediction (NWP) model forecasts. We do this because the NWP model forecasts initiated from a given initial time are available too late to be used by forecasters during that particular forecast cycle so they are postprocessed in order that they are available for the next forecast cycle. The postprocessor we use is called the interpolator (see Goerss et al. 2004; Sampson et al. 2008) and it generates both track and intensity forecasts that are then available at approximately synoptic time + 1.5 h, in time to be used for the official forecast. The interpolated intensity forecasts, if skillful, can then be used in a consensus such as S5YY. We found three of these that provided reasonable skill (the first three aids in Table A1).

We also chose to run an ensemble of SHIPS (DeMaria et al. 2005) and LGEM (DeMaria 2009) members instead of a single run of each, similar to what was done for the Statistical Typhoon Intensity Prediction Scheme (STIPS; Knaff et al. 2005) and its ensemble (Sampson et al. 2008). To do this, we chose commonly available, skillful members of the operational track consensus used at JTWC, all interpolated to the current synoptic time. Ideally, the ensemble members should be run through SHIPS and LGEM with thermodynamic and dynamic input from the model corresponding to the interpolated track. This would provide the most independence in the members, which should lead to a somewhat larger reduction in the consensus mean. It would also provide model fields with a vortex structure collocated with the interpolated track and, thus, should provide for more realistic SHIPS–LGEM computations (e.g., shear

TABLE A1. A list of tropical cyclone intensity forecast aids used to form a simple intensity consensus (S5YY) for the western Pacific basin. The first column provides the member name, the second a description of the aid.

Forecast aid ID	Description
CHII	Interpolated Coupled Hurricane Intensity Prediction System (CHIPS; Emanuel et al. 2004)
COTI	Interpolated Coupled Ocean–Atmosphere Mesoscale Prediction System-Tropical Cyclone (COAMPS-TC; Doyle et al. 2012)
GFNI	Interpolated U.S. Navy GFDL model (Rennick 1999)
LGEN*	LGEM, NOGAPS fields, and Next Generation Product Identification (NGPI) track
DSHN*	SHIPS, NOGAPS fields, and NGPI track
LGEA*	LGEM, GFS/NOGAPS fields, and AVNI track
DSHA*	SHIPS, GFS/NOGAPS fields, and AVNI track
LGEW*	LGEM, NOGAPS dynamic fields, and WBAI track
DSHW*	SHIPS, NOGAPS dynamic fields, and WBAI track

* SHIPS–LGEM ensemble member.

computation) for that member. Since we could not obtain complete model field input for all of the member models, we constructed a compromise solution. For the GFS interpolated model tracks (GFS/Aviation Interpolated, AVNI), SHIPS–LGEM is run with dynamic fields (u and v components of the wind) from Global Forecast System (GFS) and with fields from the Navy Operational Global Atmospheric Prediction System (NOGAPS; Hogan and Pauley 2007) for the other input data (temperature, relative humidity, and geopotential height). For the Weber barotropic model (WBAR; Weber 2001) and NOGAPS interpolated tracks (WBAI and NGPI, respectively), SHIPS–LGEM is run with NOGAPS field data input. A summary of the SHIPS–LGEM ensemble members and the S5YY consensus members is shown in Table A1.

REFERENCES

DeMaria, M., 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.,* **137,** 68–82, doi:10.1175/2008MWR2513.1.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvement to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting,* **20,** 531–543, doi:10.1175/WAF862.1.

——, J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting,* **24,** 1573–1591, doi:10.1175/2009WAF2222286.1.

——, and Coauthors, 2013: Improvements to the operational tropical cyclone wind speed probability model. *Wea. Forecasting,* **28,** 586–602, doi:10.1175/WAF-D-12-00116.1.

Doyle, J. D., and Coauthors, 2012: Real time tropical cyclone prediction using COAMPS-TC. *Atmospheric Science (AS) and Ocean Science (OS),* C.-C. Wu and J. Gan, Eds., Advances in Geosciences, Vol. 28, World Scientific, 15–28.

Draper, N. R., and H. Smith, 1966: *Applied Regression Analysis.* John Wiley and Sons, 407 pp.

Emanuel, K., C. DesAutels, C. Holloway, and R. Korty, 2004: Environmental control of tropical cyclone intensity. *J. Atmos. Sci.,* **61,** 843–858, doi:10.1175/1520-0469(2004)061<0843:ECOTCI>2.0.CO;2.

Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.,* **128,** 1187–1193, doi:10.1175/1520-0493(2000)128<1187:TCTFUA>2.0.CO;2.

——, 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.,* **135,** 1985–1993, doi:10.1175/MWR3390.1.

——, C. R. Sampson, and J. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting,* **19,** 633–638, doi:10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2.

Hogan, T. F., and R. L. Pauley, 2007: The impact of convective momentum transport on tropical cyclone track forecasts using the Emanuel cumulus parameterization. *Mon. Wea. Rev.,* **135,** 1195–1207, doi:10.1175/MWR3365.1.

Knaff, J., C. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting,* **20,** 688–688, doi:10.1175/WAF863.1.

Majumdar, S. J., and P. M. Finocchio, 2010: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities. *Wea. Forecasting,* **25,** 659–680, doi:10.1175/2009WAF2222327.1.

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, doi:10.1175/1520-0434(1999)014<0297:POTNST>2.0.CO;2.

Reynolds, C. A., J. G. McLay, J. S. Goerss, E. A. Serra, D. Hodyss, and C. R. Sampson, 2011: Impact of resolution and design on ensemble performance in the tropics. *Mon. Wea. Rev.,* **139,** 2145–2155, doi:10.1175/2011MWR3546.1.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.,* **81,** 1231–1240, doi:10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2.

——, J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting,* **23,** 304–312, doi:10.1175/2007WAF2007028.1.

Vijaya Kumar, T. S. V., T. N. Krishnamurti, M. Fiorino, and M. Nagata, 2003: Multimodel superensemble forecasting of tropical cyclones in the Pacific. *Mon. Wea. Rev.,* **131,** 574–583, doi:10.1175/1520-0493(2003)131<0574:MSFOTC>2.0.CO;2.

Weber, H. C., 2001: Hurricane track prediction with a new barotropic model. *Mon. Wea. Rev.,* **129,** 1834–1858, doi:10.1175/1520-0493(2001)129<1834:HTPWAN>2.0.CO;2.

——, 2005: Probabilistic prediction of tropical cyclones. Part II: Intensity. *Mon. Wea. Rev.,* **133,** 1853–1864, doi:10.1175/MWR2950.1.