Passive-Microwave-Enhanced Statistical Hurricane Intensity Prediction Scheme

THOMAS A. JONES AND DANIEL CECIL

Department of Atmospheric Science, University of Alabama in Huntsville, Huntsville, Alabama

MARK DEMARIA

NOAA/NESDIS, Fort Collins, Colorado

(Manuscript received 26 January 2005, in final form 23 September 2005)

ABSTRACT

The formulation and testing of an enhanced Statistical Hurricane Intensity Prediction Scheme (SHIPS) using new predictors derived from passive microwave imagery is presented. Passive microwave imagery is acquired for tropical cyclones in the Atlantic and eastern North Pacific basins between 1995 and 2003. Predictors relating to the inner-core (within 100 km of center) precipitation and convective characteristics of tropical cyclones are derived. These predictors are combined with the climatological and environmental predictors used by SHIPS in a simple linear regression model with change in tropical cyclone intensity as the predictand. Separate linear regression models are produced for forecast intervals of 12, 24, 36, 48, 60, and 72 h from the time of a microwave sensor overpass. Analysis of the resulting models indicates that microwave predictors, which provide an intensification signal to the model when above-average precipitation and convective signatures are present, have comparable importance to vertical wind shear and SST-related predictors. The addition of the microwave predictors produces a 2%–8% improvement in performance for the Atlantic and eastern North Pacific tropical cyclone intensity forecasts out to 72 h when compared with an environmental-only model trained from the same sample. Improvement is also observed when compared against the current version of SHIPS. The improvement in both basins is greatest for substantially intensifying or weakening tropical cyclones. Improvement statistics are based on calculating the forecast error for each tropical cyclone while it is held out of the training sample to approximate the use of independent data.

1. Introduction

Current statistical tropical cyclone intensity forecast models are primarily based on climatology, persistence, and synoptic-environmental parameters (DeMaria and Kaplan 1994 hereafter DK94; DeMaria and Kaplan 1999, hereafter DK99; DeMaria et al. 2005; Knaff et al. 2005). Missing from these models is explicit information about the characteristics of a tropical cyclone's internal structure. Passive microwave imagery has been suggested as a very useful tool for deriving these characteristics (e.g., Rao and MacArthur 1994; Cecil and Zipser 1999). These characteristics include the presence and intensity of eyewall convection and rainfall, the areal coverage of tropical cyclone–related clouds and precipitation, the location and size of rainbands, and their relative homogeneity. Passive microwave imagery is able to observe these phenomena due to the emission and scattering properties of atmospheric water vapor, raindrops, and precipitation ice. The primary goal of this work is the creation of an improved tropical cyclone intensity forecast model based on incorporating microwave-derived tropical cyclone characteristics into the statistical forecasting process. This work focuses on creating microwave-enhanced models for the Atlantic and eastern North Pacific tropical cyclone basins.

Statistical intensity forecast models such as Climatology and Persistence (CLIPER) and the Statistical Hurricane Intensity Forecast (SHIFOR) were developed using only climatology and persistence parameters to forecast tropical cyclone intensity change using a simple linear regression model (Neumann 1972; Jarvinen and Neumann 1979; Knaff et al. 2003). The primary opera-

Corresponding author address: Thomas Jones, Dept. of Atmospheric Science, University of Alabama in Huntsville, Rm. 4015, 320 Sparkman Dr., Huntsville, AL 35805.
E-mail: tjones@nsstc.uah.edu

tional hurricane intensity model after SHIFOR was created by DK94 and labeled the Statistical Hurricane Intensity Prediction Scheme (SHIPS). SHIPS incorporated oceanic and environmental (in additional to climatological) parameters into intensity change forecasts for tropical cyclones out to 72 h. Since 1994, SHIPS has been improved significantly with additional predictors, better environmental analyses, larger training datasets, increased forecast times out to 120 h, and the addition of an eastern Pacific model (DK99; DeMaria et al. 2005).

A further refinement to intensity prediction is the addition of tropical cyclone observations to the statistical models. Fitzpatrick (1997) was the first to attempt the feat on a large scale by adding infrared satellite observations of western North Pacific tropical cyclones into a linear regression scheme similar to that used by DK99. The resulting model showed significant improvement in forecast skill over a model derived solely from climatology, persistence, and environmental parameters. DeMaria et al. (2005) proceeded to include infrared (10.7 μm) imagery and ocean heat content (OHC) into SHIPS with operational implementation beginning during the 2004 Atlantic and eastern Pacific tropical cyclone seasons. However, these additions are currently only able to provide a 3%–4% improvement in operational intensity forecast skill for the Atlantic basin and a 4%–8% improvement for the eastern Pacific (DeMaria et al. 2005). In the Atlantic, only half of this improvement can be attributed to the infrared data alone. The eastern Pacific model does not include OHC data; hence, all of its improvement is derived from the infrared data.

One significant limitation of using infrared data is that low- and midlevel tropical cyclone characteristics are often obscured by cirrus outflow (Lee et al. 2002). Passive microwave imagery on the other hand has the ability to observe tropical cyclone characteristics at several levels and scales. Several previous works focused on associating 85-GHz information with current and future intensity (Glass and Felde 1992; Rao and McCoy 1997; Cecil and Zipser 1999; Bankert and Tag 2002; Lee et al. 2002). The primary reasons for using 85-GHz brightness temperatures T_b are their relatively high spatial resolution (compared with other microwave channels) and their sensitivity to precipitation ice such as graupel and hail. Precipitation ice is produced in regions of deep convection where updrafts carry water drops above the freezing level. Larger concentrations of precipitation ice indicate regions of stronger updrafts and more vigorous convection whose magnitude can be related to current and future tropical cyclone intensity. Smaller ice particles such as those associated with cirrus outflow do not significantly scatter microwave radiation and are transparent in this part of the microwave spectrum. The results of these works suggest a modest correlation with 24-h tropical cyclone intensity change using mean 85-GHz brightness temperature values calculated within the first 1.0° (or ~100 km) from the center.

Other works used information from multiple microwave frequencies (e.g., 19 and 37 GHz) to derive rainfall rate information associated with tropical cyclones (Spencer et al. 1989; Rao and MacArthur 1994; Rodgers et al. 1994; Rodgers and Pierce 1995; Bankert and Tag 2002). Rainfall rate has been noted as an estimator of the amount of diabatic heating occurring within the inner-core region. Diabatic heating resulting from latent heat release is a key factor associated with the intensification process. The addition of latent heat to a tropical cyclone acts to warm the column, increasing thicknesses in the midtroposphere while hydrostatically decreasing pressure near the surface. The pressure decrease near the center increases the pressure gradient between the center and the surrounding environment. The "response" to the increased pressure gradient is an increase in the velocities associated with the low-level circulation so that near-gradient balance is maintained. Modest correlations with tropical cyclone intensity change were observed using rainfall rate information within 1.0° from the center. In some cases, the correlation between rainfall rate characteristics and intensity change was higher than that between 85-GHz characteristics and intensity change.

2. Data

a. Best-track data

Intensity and location data for each tropical cyclone are taken from postanalysis "best track" files produced by the National Hurricane Center (NHC). Several central Pacific tropical storms are included in the eastern Pacific sample to correspond with operational practice. Postanalysis combines all available tropical cyclone data (e.g., surface observations, ship and buoy reports, aircraft measurements, dropsonde measurements, and satellite observations) to produce files that contain maximum sustained wind speed (MSW), location, and minimum sea level pressure for the lifespan of a tropical cyclone from depression stage to either its dissipation or transition into an extratropical system. Data are smoothed and reported in 6-h intervals at 0000, 0600, 1200, and 1800 UTC with wind speed values rounded to the nearest 5-kt (1 kt-0.5144 m s^{-1}) interval. For the purposes of this work, "intensity" is classified as the wind speed of a tropical cyclone at a particular point in

time. The 1-min-average wind speed 10 m above the surface defines intensity for all observations.

In the Atlantic, a significant portion of tropical cyclone observations is based on measurements from aircraft and dropsondes. Otherwise, most intensity observations are based on the Dvorak (1984) technique. Franklin et al. (2003) note the uncertainties associated with estimating intensity from aircraft flight level or dropsonde measurements. Brown and Franklin (2004) report an 11-kt rms error for Dvorak estimates when compared with more reliable aircraft-based estimates. Location errors of ~20 km are also common due to smoothing and/or lack of a well-defined center. An attempt is made to correct for location errors when well-defined eyes are present, though this is only possible for ~10% of the sample. Intensity estimates are not modified and are accepted as is.

b. Environmental data

Environmental data are taken from National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis or archived model output and written to diagnostic files designed for SHIPS training (DK94; DeMaria et al. 2005). Diagnostic data used by the 2004 version of SHIPS exist back to 1982 for the Atlantic and eastern Pacific. Climatological and environmental data associated with each tropical cyclone are sampled at 12-h intervals at 0000 and 1200 UTC, respectively. Each environmental parameter is computed at the observation time and for "future" times at 12-h intervals out to 120 h. In real-time operation, these parameters are derived from the NCEP Global Forecasting System (DeMaria et al. 2005). All environmental fields are modified such that the circulation signature of the tropical cyclone itself is removed to produce a better estimate of a tropical cyclone's environment and remove errors associated with the often poor representation of tropical cyclone circulations within numerical models (DK99).

Environmental predictors are based on spatial averages of the resulting environmental data fields out to a radius of 1000 km from a tropical cyclone's center. Many environmental predictors are time averaged along the projected track of the tropical cyclone so that the conditions it may encounter along its path may be factored into the forecast. The "perfect prog" assumption is made implying that best-track and environmental data within the training sample are a "perfect" representation of the actual event (Wilks 1995). This is not the case in real-time operation, resulting in a forecast whose accuracy is very dependent on the quality of the input data.

Data available in the diagnostic files can be broken down into several major categories, which include climatology, persistence, and environmental parameters (Table 1). A more detailed explanation of these parameters as well as the rationale for their use can be found in DK94, DK99, Knaff et al. (2005), and DeMaria et al. (2005).

c. Microwave imagery

Passive microwave imagery collected from the Defense Meteorological Satellite Program (DMSP) and Tropical Rainfall Measuring Mission (TRMM) satellites forms the basis of this work. The relevant sensors include the Special Sensor Microwave Imager (SSM/I; Hollinger et al. 1987), aboard the DMSP satellites and the TRMM Microwave Imager (TMI; Kummerow et al. 1998). Both sensors measure upwelling microwave radiation emitted from the earth's surface and atmosphere. The measured radiances are then converted into brightness temperature values using a form of Planck's law. Raindrops emit (absorb) radiation over a large portion of the microwave spectrum. Water vapor emits microwave radiation near 22 GHz. Areas of significant rain and high water vapor content would appear warmer than the surrounding environment due to the greater atmospheric emission of microwave radiation as compared with a low-emissivity ocean background at these frequencies. Precipitation ice and large raindrops scatter (in the Mie regime) upwelling radiation at higher frequencies such as 37 and 85 GHz. Increased scattering occurs in areas of large raindrop and/or precipitation ice concentrations resulting in these areas appearing colder than the surrounding environment.

SSM/I data from four DMSP satellites (*F10*, *F13*, *F14*, and *F15*) from 1995 to 2003 are collected for this research. Passive microwave channels observed by the SSM/I sensor include horizontally (H) and vertically (V) polarized 19.35-, 22.235- (V only), 37.0-, and 85.5-GHz frequencies. Hereafter, these channels are referred to as H- or V19, 22, 37, and 85 GHz for simplicity. Pixel resolution is a function of frequency with 85 GHz being the highest at 16 km × 14 km and 19 GHz being the lowest at 70 km × 45 km. Data are oversampled by the satellite producing approximately 12 km × 12 km high-resolution pixels for 85-GHz data with the remaining three channels oversampled to low-resolution 24 km × 24 km pixels.

TMI data are collected from December 1997 (launch) to 2003. The TRMM satellite was launched in a unique orbit to maximize observations of tropical regions, but has the limitation of not being able to observe latitudes beyond ±38°. TRMM's lower altitude improves the spatial resolution to 7 km × 5 km at 85

GHz and to 30 km × 18 km at 19 GHz (oversampled to 5 km × 5 km and 10 km × 10 km, respectively).

For both instruments, vertically and horizontally polarized 85- and 37-GHz T_b are combined to form polarization corrected brightness temperatures (PCT85, PCT37), which are defined by Spencer et al. (1989) and Cecil et al. (2002). Polarization correction reduces the effect of the background (e.g., land or water) on the meteorological signal being studied.

d. Microwave parameters

Many parameters derived from the raw TMI and SSM/I microwave brightness temperature data are considered. All parameters are calculated from data within a radius of 0–100 km centered around the best-track-defined tropical cyclone center. Parameters are also calculated from larger radii, but they are disregarded after confirming they have lower correlation values with intensity and intensity change than do the 0–100-km parameters. This result is in agreement with Rao and McCoy (1997) and Cecil and Zipser (1999). Data for the 0–100-km area are considered valid if at least 90% of the possible pixels within this area exist and are not over land. The highly variable emission properties of land in the microwave spectrum lead to the contamination of the meteorological signals being observed.

Parameters considered for use include mean and minimum (or maximum) brightness temperature values from each channel. These parameters roughly correspond to the average and maximum concentration of rainfall (19, 37 GHz), water vapor (22 GHz), and precipitation ice (85 GHz) within the 0–100-km area. The same parameters are also calculated for the four quadrants (northeast, northwest, southeast, southwest) of the 0–100-km area to give an estimate of tropical cyclone symmetry. The standard deviation of the four quadrant means relative to the overall mean is calculated and used as the measure of symmetry because the standard deviation of the raw brightness temperatures would be dependant on the horizontal resolution of a particular sensor. TMI minimum and maximum parameters are spatially averaged to approximate the SSM/I resolution.

An automated eye-detection algorithm is developed and used to correct tropical cyclone best-track locations so the microwave data are centered on an eye at the satellite overpass time. The algorithm determines the presence of an eye by searching for a ring of PCT85 below 260 K surrounding at least 75% of a tropical cyclone's center. The radius at which the eye detection criteria are satisfied is defined as the eyewall radius. This parameter usually represents the inner edge of eyewall precipitation in the mid- to upper troposphere. Data within the eyewall radius are considered to be in the "eye region" of a tropical cyclone. The 0–100-km parameters discussed above are recalculated using the recentered location with pixels representing the eye region removed from the averaging process. Removing them gives a better measure of the rainfall and convective characteristics associated with the eyewall and surrounding convection.

All microwave tropical cyclone observations are "snapshot" observations. No time continuity is currently taken into account. In many cases, microwave predictors are inadequately and irregularly sampled for the reliable calculation of microwave-derived time rate of change predictors.

3. Methodology

a. Valid forecasts and sample size

Intensity change forecasts are only computed for times when tropical cyclones are located over water, as the effects of land on intensity change are beyond the scope of this work. A decay algorithm currently exists to account for land interaction, which is applied to SHIPS output in real-time operation (DeMaria et al. 2005). Tropical cyclone observations located within 100 km of a significant landmass at the time of a satellite overpass are removed due to the requirement of non-contaminated microwave data, which is discussed in section 2. Training samples include observations of all tropical cyclone stages from depression to hurricane strength including subtropical storms. Data are valid until either the final dissipation of the tropical cyclone or to the point where it has been declared extratropical.

Best-track and environmental data are interpolated to the overpass time corresponding to each microwave tropical cyclone observation. Environmental parameters vary somewhat smoothly over a period of a few hours, but microwave observations may change significantly during the same period of time. For this reason, other interpolation schemes such as interpolating the microwave overpasses to the nearest synoptic time (0000 or 1200 UTC) are not pursued. Using microwave overpasses "as is" in a time window (~3 h) before a synoptic time also proved ineffective.

To ensure that each microwave overpass in the sample represents some degree of independent information, an overpass that lies within a certain time window to another overpass is removed from the sample. The time window is determined by comparing the autocorrelation coefficients (r) for several microwave and environmental predictors. The goal is to calculate the interval between overpasses where the microwave autocorrelation roughly corresponds to the 12-h environ-

mental autocorrelation. The rapidly changing nature of inner-core characteristics results in the microwave autocorrelation decreasing much more rapidly than for the environmental predictors. The environmental autocorrelation value at 12 h is used as a threshold since this is the maximum degree of autocorrelation currently present within SHIPS. The autocorrelation for important microwave predictors reaches the 12-h environmental threshold ($r \sim 0.65$) for an overpass interval of between 4 and 5 h. Being conservative, all overpasses that occur within 6 h of each other for a particular tropical cyclone are removed from the sample. For a 6-h interval, the approximate microwave autocorrelation is ~0.6, resulting in a reduction in sample size of approximately 30%. Still, autocorrelation must be taken into account prior to the formulation and interpretation of statistical models by using the effective sample size in place of the actual sample size (Aberson and DeMaria 1994).

The effective sample size N_e of the training sample is determined by calculating the variance inflation factor and multiplying by the actual sample size N [Eq. (1); Wilks (1995)]. The average time between observations after the removal of serially correlated scans is approximately 14 h indicating that the autocorrelation between 12-hourly observations would provide a conservative estimate of the variance inflation factor. The 12-h interval also corresponds to the environmental data spacing and the model forecast interval. The autocorrelation of 12-h intensity change results in a 50% reduction in sample size applied to all forecast times. The effective sample size replaces the actual sample size in all calculations of statistical significance:

$$N_e = N\frac{(1-r)}{(1+r)}. \quad (1)$$

b. Regression

Tropical cyclone intensity change forecast models are produced in a fashion similar to that of DK94 and DK99. Separate forecast models are created for forecast intervals of 12, 24, 36, 48, 60, and 72 h from the microwave overpass time using a multiple linear regression technique [Eq. (2)]. Coefficients b_K are multiplied by predictor values x_K to produce a forecast change in intensity for a tropical cyclone y, where K is the number of predictors used in the regression:

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_Kx_K. \quad (2)$$

Models extending beyond 72 h are not derived since a snapshot of tropical cyclone characteristics is unlikely to have a significant impact on intensity change beyond 3 days (Fitzpatrick 1997). Prior to regression, predictors and the predictand are normalized by subtracting their means and dividing by their standard deviations. This allows the direct comparison of the coefficients resulting from the regression process.

About 40 predictors, selected based on their physical relevance to tropical cyclone intensity change and their independence from other predictors, are fed into a backward stepwise linear regression process (Table 1). Different forecast intervals produce regression models with different sets of significant predictors. For simplicity of physical interpretation, it is desired that models for a particular basin contain the same predictors over all forecast times. For the final regression models, predictors are retained for *all* times if they are significant to the 99% confidence level for at least two forecast intervals. Remaining predictors that only marginally pass this threshold are subtracted from the model on a predictor-by-predictor basis to determine whether or not the predictor adds significant skill. If the predictor does not, it is removed from the final regression, keeping with the goal of having a model with as few predictors as possible. The high threshold for predictor retention and a degree of subjectivity are necessary to keep the total number of predictors to a minimum, reducing artificial skill (Neumann et al. 1977). The use of more objective means such as principle component analysis failed to produce superior results. Once the final set of predictors is chosen, regression models are recomputed for all forecast intervals. Hereafter, the microwave-enhanced intensity forecast model is referred to as SHIPS-MI. Separate SHIPS-MI models are derived for the Atlantic and eastern Pacific basins.

This methodology differs from the residual adjustment produced by DeMaria et al. (2005). DeMaria et al. (2005) added infrared and oceanic heat content predictors to the forecast *after* the SHIPS forecast had been produced. The adjustment is calculated by regressing the new predictors against the residual errors present in the statistical–dynamical-only model. This technique has the advantage of retaining the information from tropical cyclones prior to 1995, where the satellite imagery is not available for model training. The same adjustment technique is also applied here with microwave data to compare the advantages and disadvantages of both the pure regression and the adjustment modeling techniques. Hereafter, the microwave form of the adjustment model is referred to as SHIPS-MA.

c. Verification

To approximate an independent dataset, a jackknifing procedure is used as in DK94 and DK99. A tropical cyclone is withheld from the training sample and the

TABLE 1. Climatological, environmental, and microwave predictors available from the SHIPS diagnostic files or computed from microwave data. "Time avg" defines whether or not (Y or N) the predictors are averaged from the initial observation to the forecast time. Predictors retained in the final Atlantic and eastern Pacific models are denoted with an A or E, respectively. Operational predictors are denoted with an O. Those used in stepwise regression that are not retained are denoted with an R. Predictors with no denotation were subjectively removed prior to stepwise regression. Microwave parameters retained by SHIPS-MI include MAXH19 (both basins), MEANH19 (Atlantic), and MEANPCT85 and STDQM85 (eastern Pacific).

Predictor	Description	Units	Time avg	Model
Climatology and persistence				
MSW0	Initial maximum sustained wind speed (1-min avg)	kt	N	A, E, O
WCG12	12-h change in intensity prior to MSW0	kt	N	E, O
USPD	Zonal component of storm motion (12-h avg)	$m\ s^{-1}$	N	O
VSPD	Meridional component of storm motion (12-h avg)	$m\ s^{-1}$	N	E
LAT	Storm-centered latitude	°	N	R, A, E
LON	Storm-centered longitude	°	N	R
ADAY	Absolute value of Julian day − peak activity Julian day	Day	N	O
Environmental				
D200	200-hPa divergence	$10^{-5}\ s^{-1}$	N	O
T200	200-hPa temperature	°C	Y	A, O
U200	200-hPa zonal wind speed	kt	Y	E
REFC	200-hPa eddy momentum flux	$m\ day\ s^{-1}$	N	R
E000	1000-hPa equivalent potential temperature	°C	Y	
EPOS	θ_e difference between lifted surface parcel and environment	°C	Y	A, E, O
RHLO	850–700-hPa relative humidity	%	Y	R
RHHI	500–300-hPa relative humidity	%	Y	O
SHRD	850–200-hPa vertical shear magnitude	kt	Y	A, E, O
SHRS	850–500-hPa vertical shear magnitude	kt	Y	R
SHRG	Generalized 850–200-hPa shear (average over all levels)	kt	Y	
USHRD	Zonal component of SHRD	kt	Y	R
VSHRD	Meridional component of SHRD	kt	Y	R
USHRS	Zonal component of SHRS	kt	Y	R
VSHRS	Meridional component of SHRS	kt	Y	R
Z850	850-hPa absolute vorticity	$10^{-5}\ s^{-1}$	Y	A, O
PSLV	Pressure steering level	hPa	N	O
DTL	Distance from land	km	N	R
MPI	Maximum potential intensity	kt	Y	
POT	MPI-MSW0	kt	Y	A, E, O
Nonlinear				
MPI2	Square of MPI	kt^2	Y	
VMPI	MSW0 times MPI	kt^2	Y	
POT2	Square of POT	kt^2	Y	A, E, O
VPER	MSW0 times WCG12	$kt^2\ s^{-1}$	Y	A, O
SHRDLAT	SHRD times SIN(LAT)	kt	Y	A, E, O
SHRSLAT	SHRS times SIN(LAT)	kt	Y	R
SHRGLAT	SHRG times SIN(LAT)	kt	Y	
MSWSHRD	SHRD times MSW	kt^2	Y	A, E, O
MSWSHRS	SHRS times MSW	kt^2	Y	R
MSWSHRG	SHRG times MSW	kt^2	Y	
Microwave (0–100 km)				
MEAN (H19, V19, V22, 37, 85)	Mean T_b	K	N	A, E
STD (H19, V19, V22, 37, 85)	Standard deviation T_b	K	N	
MIN (37, 85)	Minimum T_b (modified)	K	N	R
MAX (H19, V19, V22)	Maximum T_b (modified)	K	N	A, E
D19	Mean difference between V19 and H19 T_b	K	N	R
STDQM (H19, V19, V22, 37, 85)	Standard deviation of four-quadrant mean T_b	K	N	E

regression coefficients are recomputed without it. Forecasts and associated errors are produced for that tropical cyclone. Then the process is repeated for every other tropical cyclone. When forecast intensity falls below 20 kt, the forecast time at which that threshold is crossed and all those thereafter are removed from consideration in error calculations. This is done to correspond with operational practice since a tropical cyclone is considered "dissipated" when the forecast intensity falls below 20 kt. Errors derived this way are compared

TABLE 2. A list of the various model acronyms used in this work followed by a brief description of each.

Model	Description
SHIPS	Operational SHIPS coefficients for 2005 applied to training sample
SHIPS-ENV	Operational SHIPS with infrared and heat content adjustment removed
SHIFOR	Operational SHIFOR coefficients for 2005 applied to training sample
SHIPS-MI	Microwave-enhanced regression models
SHIPS-MA	Microwave enhancement using adjustment scheme
SHIPS-85	Microwave-enhanced regressions using only PCT85 data
Base	Microwave predictors in SHIPS-MI removed
Base-IR	Microwave predictors in SHIPS-MI replaced by infrared predictors

for several different forecast models listed in Table 2 and described below.

Comparisons with the operational SHIPS model are limited because of the small homogeneous sample size. We do not have the necessary infrared predictors for all microwave overpass times. Some comparisons are made, but any involving infrared data use a homogeneous sample that is about ⅔ the size of the full set of microwave-based forecasts. To make use of the full sample size, we also consider SHIPS forecasts that exclude the infrared and oceanic heat content adjustment. These are referred to as SHIPS-ENV. In both cases, data from the interpolated environmental sample are applied to the operational coefficients; hence, forecast errors for SHIPS and SHIPS-ENV are not computed using a jackknifing procedure.

Most comparisons utilize a "base" regression that excludes the microwave predictors from SHIPS-MI. Since both the base regression and SHIPS-MI use the same training sample and the same climatological and environmental predictors, this comparison provides a direct measure of the usefulness of microwave predictors. This is not the case when comparing SHIPS-MI with SHIPS-ENV. Similarly, the microwave adjustment model, SHIPS-MA, is trained on base model errors to solely examine the effect of the addition of microwave data.

To directly compare the effectiveness of microwave versus infrared predictors, a model is created that replaces the microwave predictors in SHIPS-MI with the operational infrared predictors (percentage of 50–200-km radius pixels below −20°C and standard deviation of T_b over a 100–300-km radius). This model is labeled Base-IR. (Note that this application does not use the adjustment scheme.) These two models are compared to determine whether the infrared or microwave predictors provide the greatest increase in forecast skill. However, the limited sample size prevents these comparisons from being statistically significant.

4. Atlantic results

a. Model interpretation

Valid tropical cyclone overpasses in the Atlantic number 1034 for the 12-h forecast sample and decrease to 552 for the 72-h forecast sample. Out of a total of 144 possible tropical cyclones, at least one valid 12-h forecast exists for 120 tropical cyclones while only 68 have valid forecasts out to 72 h. Stepwise regression and subsequent application of threshold tests for significant predictors across all forecast times results in a linear regression model with 13 separate predictors (Table 3).

Climatological predictors include persistence multiplied by initial intensity (VPER), initial intensity alone (MSW0), and the tropical cyclone's initial latitude (LAT). Synoptic-environmental predictors include the following: 200-hPa temperature (T200), deep-layer shear (SHRD), potential (POT), θ_e difference between a lifted surface parcel and its environment (EPOS), and 850-hPa absolute vorticity (Z850). Several combinations of environmental and climatological predictors also exist to partially correct for nonlinear relationships between the various predictors and the intensity change (POT2, SHRDLAT, MSWSHRD).

Several differences exist between environmental predictors present in SHIPS-MI and those present in SHIPS. SHIPS-MI includes initial latitude as a predictor whereas SHIPS only includes it within a nonlinear shear predictor. Other differences are cases where marginal SHIPS predictors fall out of the SHIPS-MI regression (e.g., ADAY, USPD). Persistence (WCG12) and relative humidity (RHHI) in SHIPS are replaced by VPER and EPOS in SHIPS-MI. Finally, 200-hPa divergence (D200) and the pressure at the steering level (PSLV) predictors are subjectively removed since they do not provide any meaningful increase in skill by their inclusion. These differences are primarily a result of using a different training dataset than SHIPS and the added influence of the microwave predictors. With the exception of persistence, these predictors rarely contribute more than ±5 kt to intensity change forecasts; thus, the overall impact of their removal is limited. The persistence contribution is retained by the nonlinear VPER predictor in the base regression and SHIPS-MI.

Microwave predictors retained after stepwise regression include mean and maximum H19 T_b for the area 0–100 km from the tropical cyclone center (MEANH19, MAXH19), which represent the mean and maximum rainfall and latent heat release present

TABLE 3. Final Atlantic SHIPS-MI predictor standardized coefficient values at each forecast time (h) with the explained variance (R^2) statistic at the bottom.

Predictor	$t = 12$	$t = 24$	$t = 36$	$t = 48$	$t = 60$	$t = 72$
MEANH19	0.39	0.50	0.47	0.35	0.22	0.11
MAXH19	0.07	0.03	0.04	0.09	0.15	0.15
POT	1.14	1.13	1.22	1.18	1.00	0.86
POT2	−0.77	−0.75	−0.78	−0.67	−0.55	−0.44
SHRD	−0.19	−0.38	−0.50	−0.57	−0.70	−0.77
SHRDLAT	0.41	0.67	0.88	0.91	1.02	1.01
MSWSHRD	−0.38	−0.49	−0.63	−0.66	−0.65	−0.54
MSW0	−0.03	−0.10	0.00	0.09	0.05	0.00
VPER	0.31	0.22	0.16	0.15	0.10	0.09
LAT	−0.07	−0.18	−0.30	−0.33	−0.39	−0.41
EPOS	0.07	0.13	0.10	0.08	0.12	0.13
T200	−0.13	−0.17	−0.19	−0.19	−0.19	−0.18
Z850	0.14	0.15	0.16	0.19	0.18	0.21
R^2	0.42	0.51	0.57	0.63	0.64	0.64

within the inner core of tropical cyclones. Predictors relating to tropical cyclone symmetry did not stand out as significant. Predictors from the H19 channel are favored by the stepwise regression, and the use of similar predictors from other channels results in a somewhat less skilled model.

As with SHIPS, predictors related to sea surface temperature and vertical wind shear (POT, SHRD, and related nonlinear terms) are the most important individual predictors at most forecast times (DK94; DK99), although there is often some cancellation between related predictors. The relative importance of every predictor can be defined by comparing the absolute value of the normalized coefficients (Table 3). For forecast intervals out to 48 h, the most important predictor beyond POT and SHRD (and related nonlinear predictors) is MEANH19. The magnitude of the MEANH19 coefficient is such that substantial intensity change (±10 kt) forecasts are possible from this one predictor. Coefficient values for MEANH19 are positive, indicating that warmer 19-GHz T_b are associated with future intensification. This term increases the forecast intensity for tropical cyclones with abundant inner-core rainfall, and decreases the forecast intensity for those with relatively little rain signature present. The MAXH19 predictor acts much the same way as the MEANH19 predictor, but is much less important than MEANH19 out to 60 h. By the 72-h forecast interval, the importance of both 19-GHz predictors is small.

Since the POT, SHRD, and microwave signals are present in multiple and sometimes offsetting predictors, the true importance of these physical signals cannot be determined by the coefficient value alone. To make this determination, the total intensity change contributions are calculated for POT + POT2, SHRD + SHRDLAT + MSWSHRD, and MEANH19 + MAXH19. The resulting intensity change contributions indicate that out of these three physical signals, microwave exceeds both potential and shear for 12- and 24-h forecasts. At 24 h, the microwave predictors contribute an average of 29% of the total forecast. The potential and shear-related predictors contribute 25% and 20%, respectively. Beyond 36 h, shear becomes the dominant signal followed by potential with the influence of microwave decreasing rapidly. While each of these contributions seems large, the removal of any one only decreases model performance by a small margin. For example, removing all the shear terms would only increase the mean absolute errors by a few percent. In this case, the remaining physical contributors take on a greater importance and offset some of the lost information. Following these physical signals in overall importance at the 12- and 24-h forecast intervals is persistence (VPER). Refer to DK94, DK99, and DeMaria et al. (2005) for a more in-depth discussion of the physical interpretation of the remaining nonmicrowave predictors.

For the final calculation of forecast intensity change, the sample mean intensity change must be factored in. The mean intensity change in the Atlantic is positive due to the larger number of intensifying versus weakening tropical cyclones in the sample. The sample mean intensity change increases from 2 kt at 12 h to over 11 kt at 72 h. The implication of the substantial positive mean is that SHIPS-MI will forecast 11 kt of intensification at 72 h even if the total predictor contribution is zero. When interpreting the final forecast, this contribution must be taken into account.

Variance explained by SHIPS-MI, defined by the R^2 coefficient, increases as a function of time from 42% at 12 h to 61% at 72 h (Table 3). These values are similar to those reported for SHIPS by DeMaria et al. (2005). The increase in variance explained does not represent

TABLE 4. Jackknifed mean absolute errors (kt) for SHIPS-ENV, SHIPS-MI, SHIPS-85, and SHIPS-MA forecasts for valid Atlantic tropical cyclone overpasses (N).

Forecast (h)	TCs	N	Base	SHIPS-ENV	SHIPS-MI	SHIPS-85	SHIPS-MA
12	120	1034	5.4	5.3	5.2	5.3	5.3
24	111	916	9.0	8.6	8.4	8.6	8.7
36	99	793	11.8	11.4	10.8	11.1	11.3
48	87	702	13.7	13.2	12.9	13.3	13.3
60	75	620	15.8	15.0	15.0	15.4	15.4
72	68	552	17.4	16.2	16.6	17.0	17.0

an increase in model skill at longer forecast times. Instead, it is simply a reflection of the increased variability of intensity change as a function of time.

b. Model performance

Table 4 and Fig. 1a give mean absolute errors for various SHIPS models for the Atlantic sample. All SHIPS-like models substantially outperform SHIFOR, indicating that adding environmental data to climatological and best-track predictors does increase forecast skill (e.g., DK94). SHIPS-ENV, which is simply the use of operational SHIPS coefficients without any adjustment from infrared or oceanic heat content predictors, outperforms the base model for all forecast times despite the latter's inclusion of the latitude predictor. This is because of the larger training sample used in SHIPS-ENV and the greater overall number of predictors (16).

SHIPS-MI outperforms both SHIPS-ENV and the base model out to 60 h. Figure 1b shows percent improvement relative to the base regression model for SHIPS-MI, SHIPS-85, SHIPS-MA, and SHIPS-ENV. Comparison against the base model best highlights the skill made available by the addition of the microwave predictors. SHIPS-MI has mean absolute errors ranging from 5.2 to 16.6 kt between 12 and 72 h compared with 5.4 and 17.4 kt for the base model, representing a 4%–8% improvement. Maximum improvement occurs at the 24- and 36-h forecast intervals, which is consistent with both physical and mathematical expectations. Using a paired Student's t-test statistic (Wilks 1995) with the effective sample size, the improvement in mean absolute error between the base model and SHIPS-MI is statistically significant to the 95% significance level for 24–48-h forecasts. SHIPS-MI also outperforms SHIPS-ENV by 2%–5% out to 60 h.

For the Atlantic, MEANH19 and MAXH19 are replaced with MEANPCT85 and MINPCT85 in SHIPS-MI to form SHIPS-85. Sample size is not expanded to include the near-land observations so that error comparisons are all based upon the same sample size as before. SHIPS-85 outperforms the base model by 3%–4% at all forecast times. However, SHIPS-85 is not as skillful as SHIPS-MI, indicating that some predictive ability is lost by the substitution of 85-GHz predictors in place of 19-GHz predictors. Given the lower correlation between 85-GHz predictors and intensity relative to 19-GHz predictors, the decreased performance resulting from using 85-GHz predictors is of no surprise. SHIPS-85 shows virtually no skill when compared against SHIPS-ENV.

SHIPS-MA is the poorest performing microwave-enhanced model, only showing a 2%–3% improvement over the base model, which is not statistically significant. Given the nature of the adjustment technique, the limited improvement is not surprising. In fact, SHIPS-MA differs from the base model by 5 kt or greater for less than 13% of all forecasts (which compares with ~35% for SHIPS-MI forecasts).

To assess whether SHIPS-MI has skill relative to the fully operational SHIPS model, Fig. 2a shows relative errors for the subset of forecasts where infrared and oceanic heat content predictors are available. Here, forecasts for SHIPS-MI, SHIPS, and SHIPS-ENV are computed using regression coefficients from the full SHIPS-MI (Table 3) and operational SHIPS and SHIPS-ENV (DeMaria et al. 2005). Hence, forecast errors in Fig. 2a are not computed using a jackknifing procedure. SHIPS-MI outperforms SHIPS at all forecast times with a maximum improvement of 6% at 36 h. However, the differences in model training samples, the number of predictors between SHIPS and SHIPS-MI, and the use of an adjustment technique in SHIPS complicate the interpretation of this comparison. Also note that the difference between SHIPS-ENV and SHIPS (with infrared) is only on the order of 1%–2% for this limited sample.

Both SHIPS-MI and Base-IR outperform the base regression model, each computed from the smaller training sample (Fig. 2b). Here, all errors are computed using a jackknifing procedure. The microwave predictors in SHIPS-MI provide almost twice as much improvement as the infrared terms in Base-IR. This result indicates that the advantages of microwave data (resolving inner-core precipitation features) compared

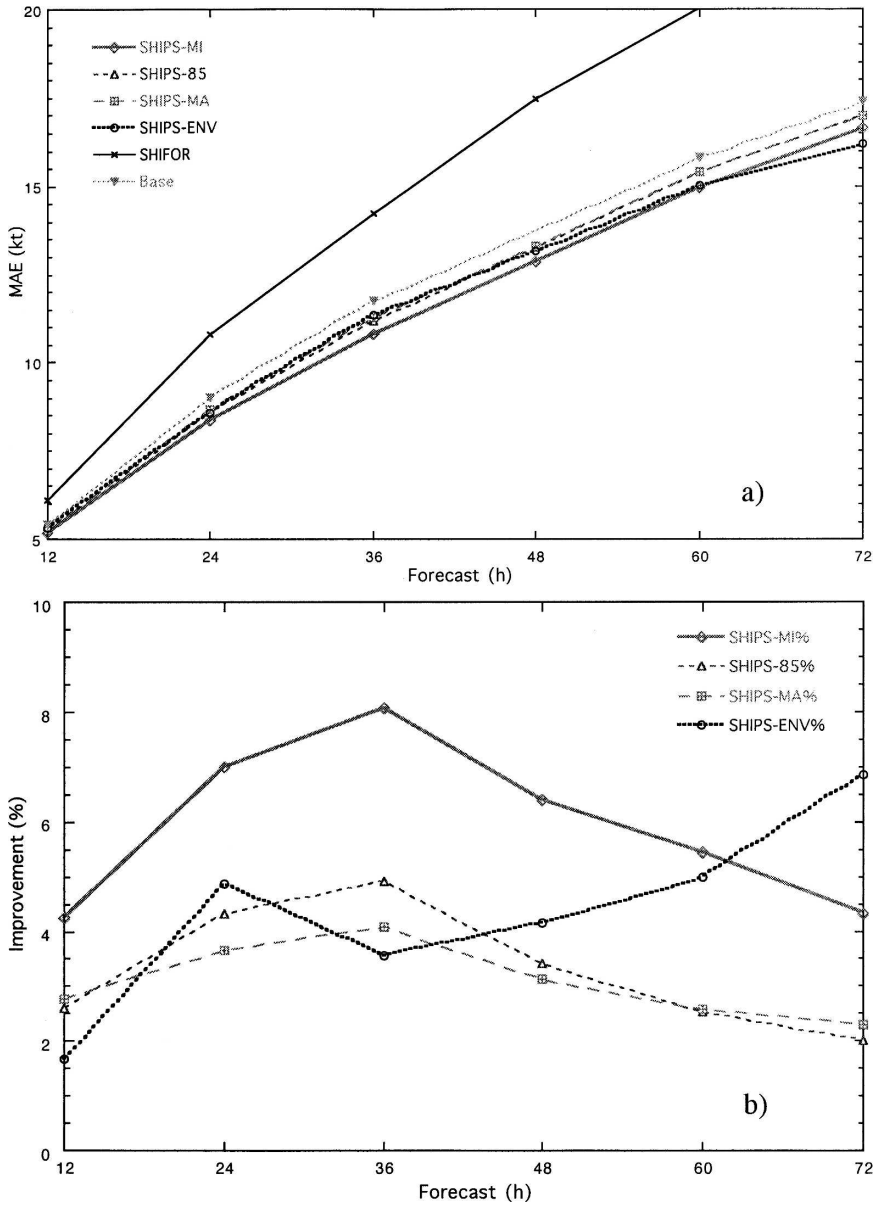

FIG. 1. (a) Atlantic mean absolute errors for SHIFOR, SHIPS-ENV, SHIPS-MI, SHIPS-85, and SHIPS-MA and (b) percent improvement of SHIPS-ENV and various microwave-enhanced models normalized against the base regression.

with infrared data (resolving cloud-top features) *do* translate into improved model performance on a homogeneous sample.

Mean absolute error and forecast improvement are stratified by initial storm intensity and subsequent intensity change to determine model performance as a function of storm type. Figure 3a shows 24-h mean absolute error binned in 5-kt intervals as a function of initial intensity. SHIPS-MI outperforms the base model for all initial intensities with the exception of very strong (MSW $>$ 130 kt) hurricanes. However, the sample size for this type is very small. Improvement is maximized for tropical depressions and for hurricanes below 110 kt. Strong convective bursts near the center of forming and intensifying tropical depressions provide a positive signal to the forecast, thereby increasing skill. The strong precipitation signal for intensifying hurricanes produces similar results. The improvement for tropical storms is somewhat less and is a result of larger variability in intensity change for this portion of the sample. Figure 3b shows the same 24-h mean absolute error, this time binned as a function of subsequent

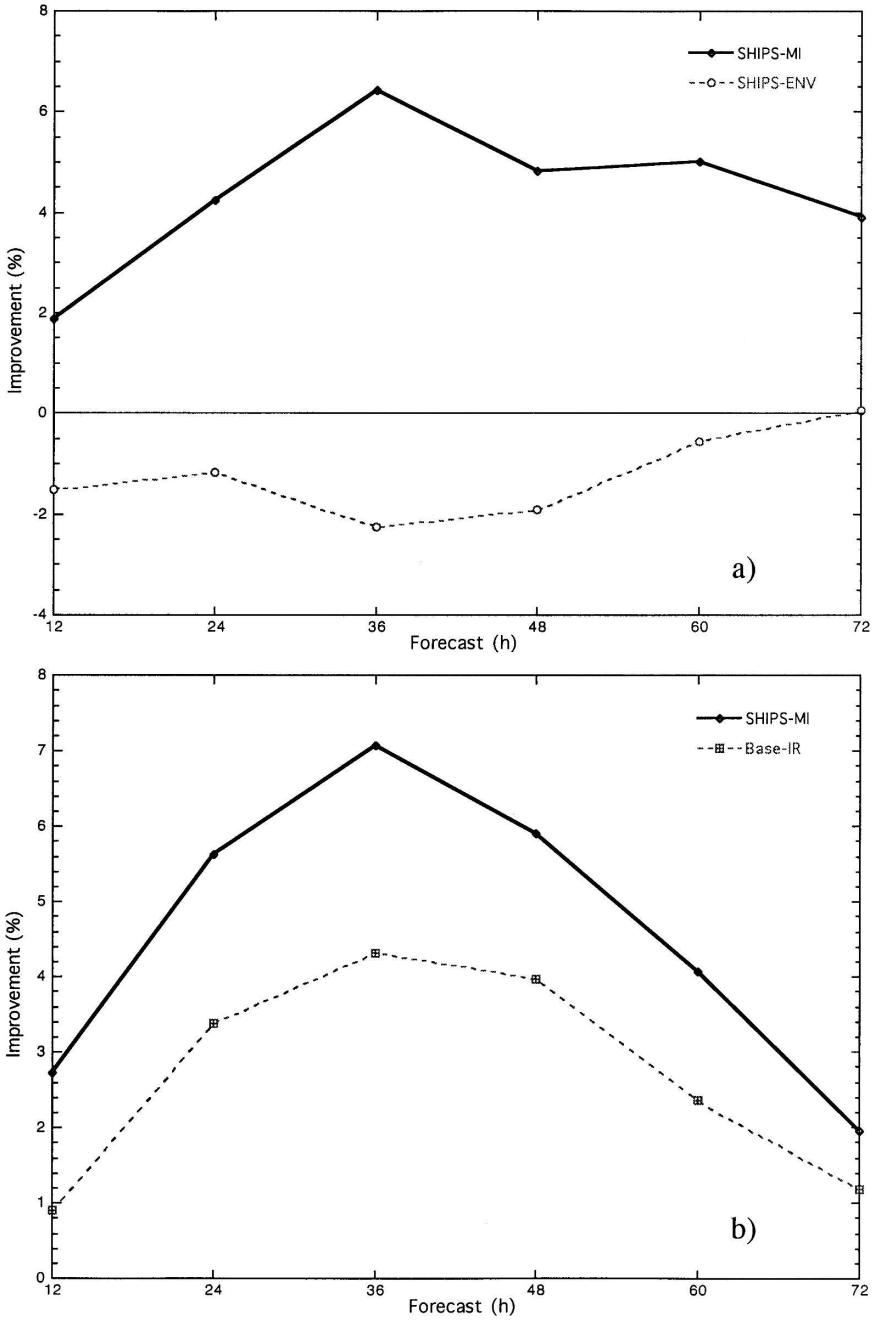

FIG. 2. Comparisons of mean absolute errors for the smaller, homogeneous, sample where infrared and oceanic heat content are available. (a) Percent improvement of SHIPS-MI normalized against SHIPS, which includes the infrared and oceanic heat content adjustment. SHIPS-ENV (no infrared of heat content adjustment applied) is shown for comparison. Regression coefficients are derived from the full training samples with no jackknifing applied here. (b) Improvements of SHIPS-MI and Base-IR normalized against the base model. Coefficients for these three models are computed from the homogenous sample and jackknifing is applied.

intensity change from the overpass time to a particular forecast time. Over 80% of the observations at 24 h fall within an intensity change range of ±20 kt. Mean absolute error for SHIPS-MI is smaller than for the base model for all intensity change bins except 0–10 kt.

Base model and SHIPS-MI 24-h forecast errors are also stratified in 5-kt bins for intensifying, weakening,

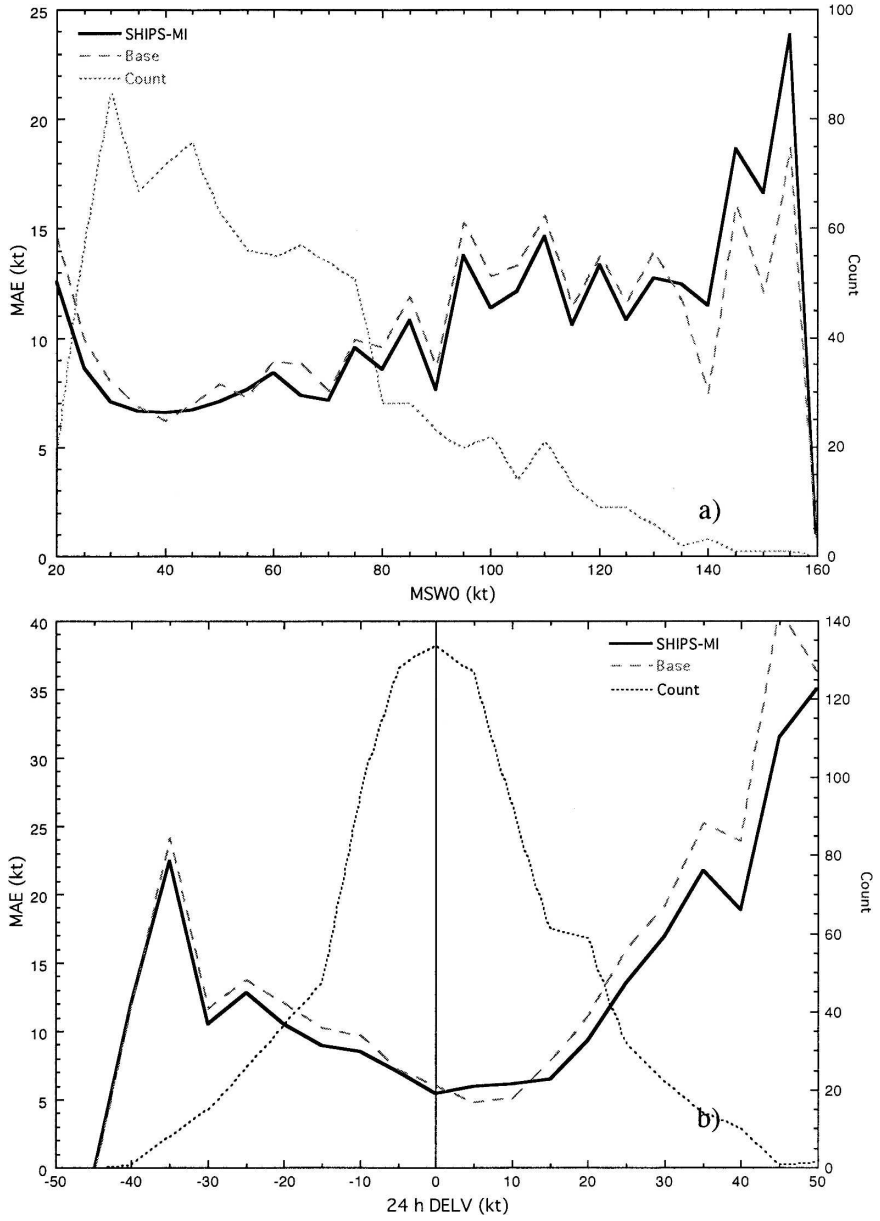

FIG. 3. Base model and SHIPS-MI mean absolute error (a) stratified by best-track initial intensity (MSW0) in 5-kt bins for 24-h forecasts and (b) stratified by subsequent intensity change (DELV) in 5-kt bins. Lower values of mean absolute error represent better forecasts. Dotted lines represent the number of valid observations within a particular bin. Note the small sample size beyond DELV $> \pm 20$ kt.

and relatively steady state tropical cyclone types (Fig. 4). Relatively steady state tropical cyclones are those that undergo less than ± 10 kt intensity change over a particular forecast duration. Tropical cyclones that undergo greater than ± 10 kt intensity change are classified as those that "intensify" and "weaken." For all types, fewer SHIPS-MI errors fall into the largest error bins (Fig. 4a). Thus, SHIPS-MI errors are generally slightly smaller than corresponding base model errors. Both models have a substantial underforecasting bias for intensifying tropical cyclones, with SHIPS-MI's bias being smaller (Fig. 4b). However, SHIPS-MI does somewhat overforecast 24-h intensity compared with the base model in some instances. Similar results were observed for weakening tropical cyclones (Fig. 4c). Models tended to not forecast enough weakening, with SHIPS-MI again having a larger number of smaller errors. Errors are generally small and show a slight posi-

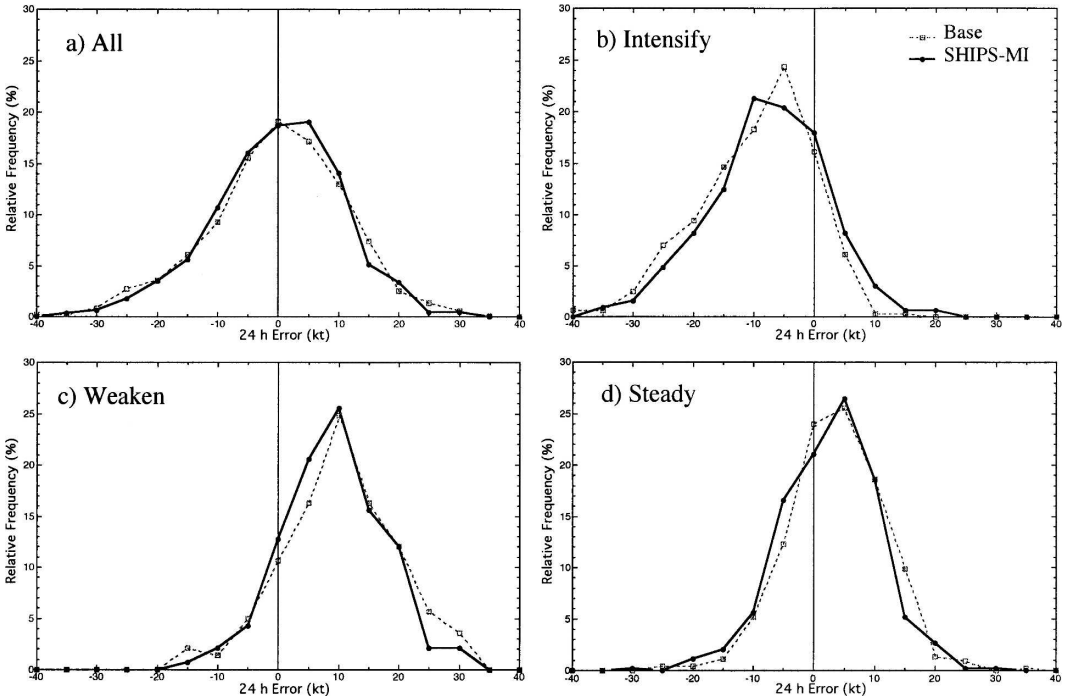

FIG. 4. Base model (dashed lines) and SHIPS-MI (solid lines) 24-h forecast errors binned in 5-kt intervals and plotted relative to the total sample size for a particular type. Types plotted include (a) all cases, (b) intensifying cases (DELV $>$ 10 kt), (c) weakening cases (DELV $<$ -10 kt), and (d) relatively steady state cases (-10 kt $\leq$ DELV $\leq$ 10 kt). Positive values indicate overforecasting while negative values indicate underforecasting.

tive bias for relatively steady state tropical cyclones (Fig. 4d).

For about 55% of all 24-h forecasts, SHIPS-MI has smaller errors than the base model (Fig. 5a). In most cases, however, SHIPS-MI and the base model generally do agree with each other and have similar errors. When the forecasts are substantially different, SHIPS-MI is usually the more accurate. In 14% of the forecasts, SHIPS-MI is at least 5 kt more accurate than the base model. The base model is at least 5 kt more accurate than SHIPS-MI in only 7% of all forecasts. This discrepancy is more pronounced in substantially intensifying (Fig. 5b) or substantially weakening storms (Fig. 5c). SHIPS-MI is at least 5 kt more accurate in 18% of the intensifying cases and 18% of the weakening cases. The base model is at least 5 kt more accurate in only 8% of the intensifying cases and 5% of the weakening cases.

The percent improvement of SHIPS-MI over the base model is also stratified as a function of subsequent intensity change and time (Fig. 6). Improvement is maximized for 12-h forecasts of intensifying cyclones and 24–48-h forecasts of weakening cyclones. Some improvement for steady state tropical cyclones occurs for 24–48-h forecasts. Generally, the improvement is greater for intensifying or weakening tropical cyclones than for relatively steady state cyclones, and SHIPS-MI is never significantly worse than the base regression model. Also, where base model and SHIPS-MI are most different, SHIPS-MI generally produces the better forecasts indicating that the addition of microwave data more often helps rather than hurts the forecast. Overall, SHIPS-MI performs best for tropical cyclones that show substantial changes in intensity over a 24–48-h period.

c. *Forecast example*

An example of the forecast improvement that SHIPS-MI provides is given by a specific forecast for Hurricane Lenny on 16 November 1999. The microwave data for this forecast were taken from the *DMSP F14* SSM/I at 1350 UTC (Fig. 7). A strong liquid rain signature surrounds Lenny's center with H19 T_b in excess of 260 K. Collocated with the high H19 T_b are low PCT85 indicating the presence of ice aloft associated with deep convection, though not as low as the PCT85 associated with the rainband convection to the east of the inner core.

Lenny was undergoing rapid intensification from 85 kt at the time of the overpass to 128 kt 36 h later (Fig. 8). Forecasts are calculated by applying the perfect prog data for Lenny to the SHIPS and SHIPS-MI

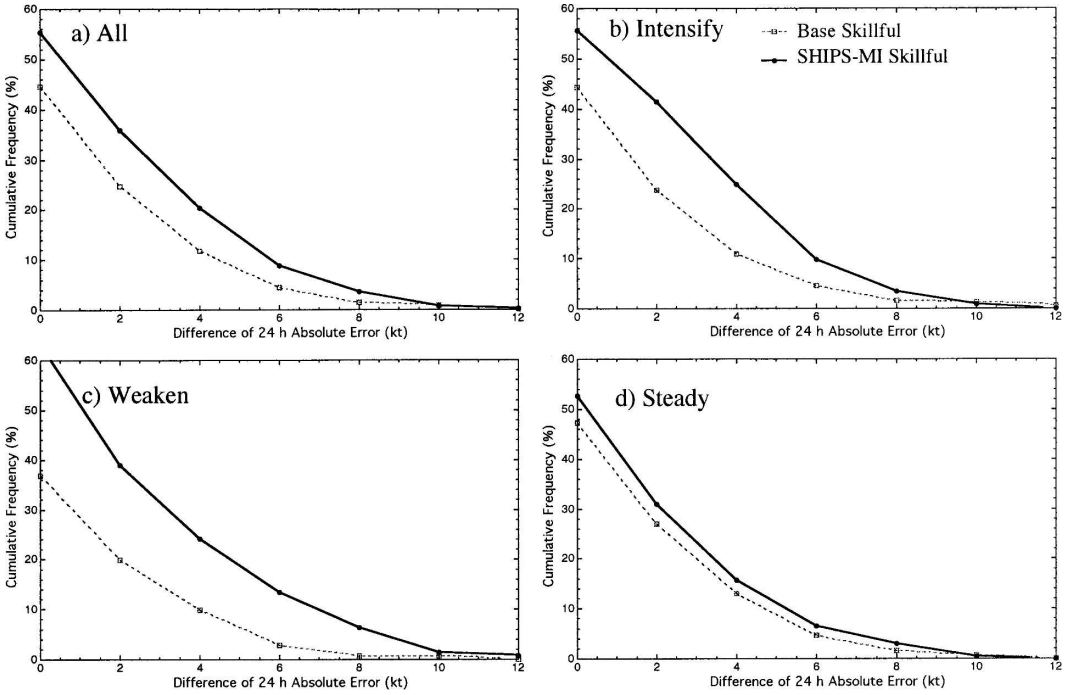

FIG. 5. Percentage of forecasts where SHIPS-MI is more accurate (solid lines) vs those where the base model is more accurate (dashed lines). Types plotted include (a) all cases, (b) intensifying cases (DELV $>$ 10 kt), (c) weakening cases (DELV $<$ -10 kt), and (d) relatively steady state cases (-10 kt $\leq$ DELV $\leq$ 10 kt).

model coefficients. SHIPS, with the infrared adjustment included, only forecasts intensification in the first 24 h with a forecast intensity of 104 kt at 24 h followed by modest weakening with a 48-h forecast intensity of only 97 kt. The base model forecasts intensification out to 36 h before weakening Lenny and is generally stronger than SHIPS. This difference is partially a result of the training sample used to train the base model (and

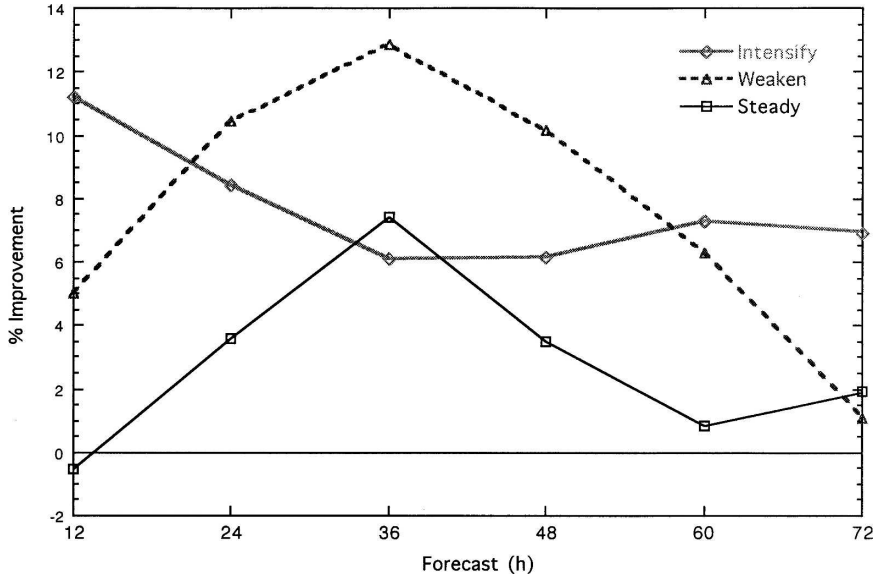

FIG. 6. Percent improvement relative to the base model for SHIPS-MI forecasts of intensifying (DELV $>$ 10 kt), weakening (DELV $<$ -10 kt), and relatively steady state (-10 $\leq$ DELV $\leq$ 10 kt) cases.

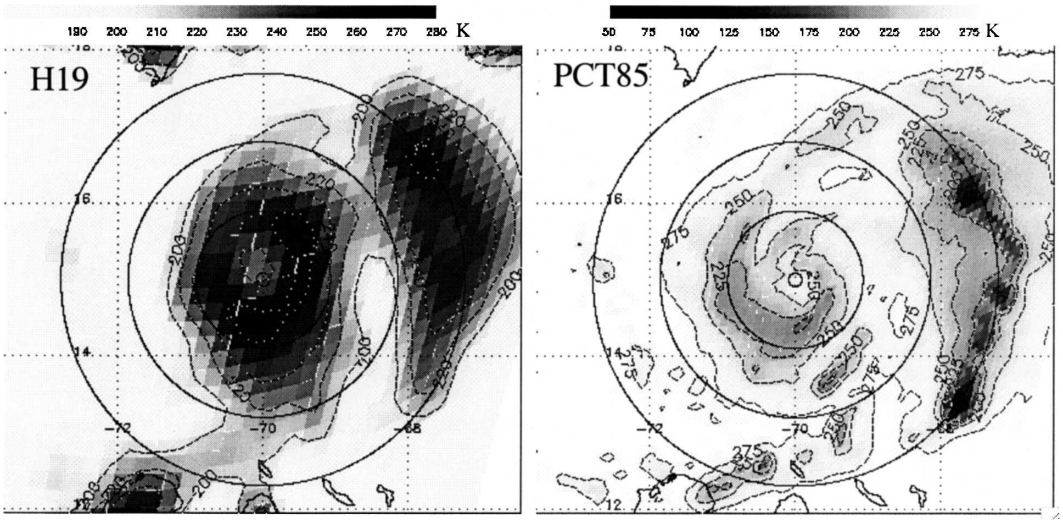

FIG. 7. *DMSP F14* SSM/I (left) H19 and (right) PCT85 for Hurricane Lenny at 1350 UTC 16 Nov 1999. The small ring at the center represents the interpolated best-track position at this time. Large rings indicated radials of 100 km from the center. H19 and PCT85 are contoured at 20 and 25 K, respectively. PCT85 contours below 200 K are not shown to increase clarity. All microwave parameters are calculated by using data within the first 100-km ring. At this time, the interpolated intensity for Lenny was 85 kt.

SHIPS-MI) having a slightly greater proportion of intensifying versus weakening tropical cyclones than does the sample used to train SHIPS. SHIPS-MI outperformed both the base model and SHIPS with the forecast intensity reaching 113 kt at 36 h. Analyzing the contributions from each predictor in SHIPS-MI, reveals a 18-kt intensification signal from the MEANH19 predictor at 36 h (Table 5). The difference between the base model and SHIPS-MI is only 5 kt because the microwave predictors take over some of the contribu-

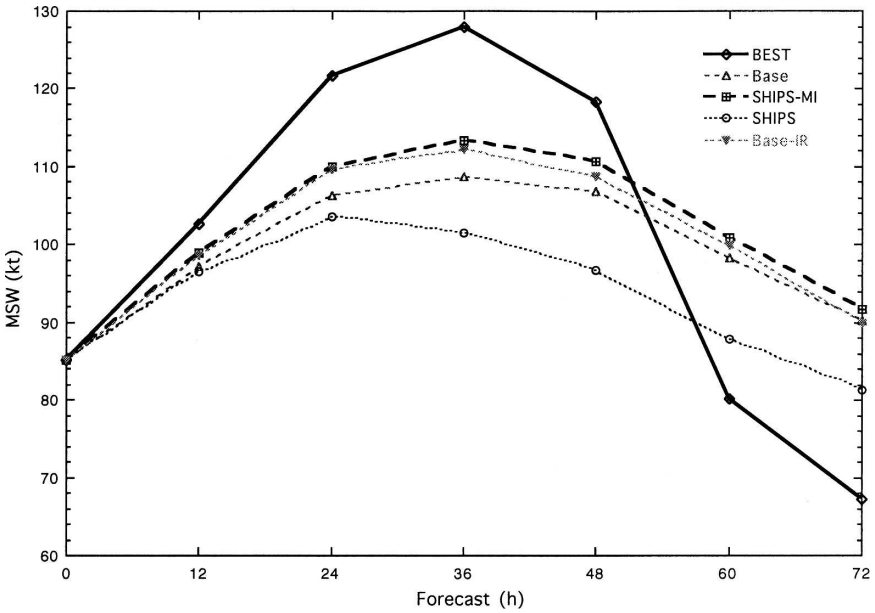

FIG. 8. Interpolated best-track intensity for Lenny beginning at the overpass time of 1350 UTC 16 Nov 1999. The 0–72-h base model, SHIPS-MI, SHIPS, and Base-IR forecasts are also plotted. The SHIPS forecast represents the forecast produced by applying reanalysis data to the 2005 operational coefficients with infrared adjustment and oceanic heat content and not the operational forecast produced in real time in 1999.

TABLE 5. Predictor contributions (kt) to the SHIPS-MI Hurricane Lenny forecast originating at 1350 UTC 16 Nov 1999. Note that the MEANH19 predictor provides a strong intensification signal for the 24–60-h forecasts.

Predictor	$t = 12$	$t = 24$	$t = 36$	$t = 48$	$t = 60$	$t = 72$
MEANH19	6.6	14.2	18.0	16.1	11.8	7.2
MAXH19	0.5	0.3	0.7	1.8	3.2	3.6
POT	−5.0	−9.4	−13.8	−16.5	−16.7	−16.6
POT2	4.7	8.3	11.8	12.4	12.2	11.4
SHRD	2.4	7.9	9.6	9.5	7.4	3.2
SHRDLAT	−4.2	−11.9	−17.1	−17.9	−18.4	−16.7
MSWSHRD	2.6	5.3	3.6	−0.3	−6.1	−10.0
MSW0	−0.3	−1.4	−0.2	1.7	0.8	−0.8
VPER	3.7	4.7	4.5	4.9	3.5	3.1
LAT	0.6	2.7	5.9	7.1	9.0	10.1
EPOS	0.8	1.8	1.8	1.5	2.4	2.8
T200	−0.8	−1.8	−2.3	−3.0	−3.9	−4.2
Z850	0.6	1.3	1.8	2.7	3.1	4.4
MEAN DELV	2.4	4.5	6.6	8.3	10.1	11.1
TOTAL DELV	14.4	26.5	30.8	28.3	18.3	8.6

tion from the initial intensity (MSW) related predictors in the base model.

After 48 h, Lenny began to weaken as it encountered greater shear. All models correctly capture this weakening signal in SHRD-related terms (Table 5). In addition, some of the weakening is likely due to the interaction with the Lesser Antilles islands after the 48-h forecast interval. Land interaction is not taken into account by either SHIPS or SHIPS-MI in this forecast. By the 72-h forecast, MEANH19 and MAXH19 are still contributing a 10-kt positive signal, causing SHIPS-MI not to weaken Lenny as quickly as SHIPS or as actually happened. The end result is a 25-kt overforecast by SHIPS-MI at 72 h with SHIPS overforecasting by only 10 kt. The intensification signal produced by SHIPS-MI (>15 kt at 36 and 48 h) is clearly significant for this forecast based on the physical association between the microwave signature and the subsequent intensification. Both SHIPS-85 and SHIPS-MA performed similarly to SHIPS-MI for this example. The infrared version of the base model, Base-IR, does forecast greater intensification than the base model alone, but not as much as the microwave-enhanced versions. The infrared adjustment applied to SHIPS-ENV adds ~3 kt to 36- and 48-h forecasts.

5. Pacific results

a. Model interpretation

In a fashion similar to that used for the Atlantic SHIPS-MI model, an eastern Pacific version of SHIPS-MI is produced. Valid forecast tropical cyclone overpasses in the eastern Pacific number 913 for the 12-h forecast sample and decrease to 425 for the 72-h forecast sample. Out of a total of 140 possible tropical cyclones, at least one valid forecast at 12 h exists for 128 tropical cyclones while only 83 have valid forecasts out to 72 h. Several of the missed tropical cyclones are a result of central Pacific depressions not being included in the SHIPS training data. While the number of tropical cyclones occurring in the eastern Pacific is roughly equal to the number in the Atlantic during the same period (1995–2003), eastern Pacific tropical cyclones tended to be shorter lived than their Atlantic counterparts resulting in the smaller eastern Pacific sample size. This also results in a larger number of dissipation forecasts, which further reduces sample size as forecast time increases. The resulting eastern Pacific model includes 14 predictors and is similar to, but not identical to the Atlantic model (Table 6).

Among the environmental and climatological predictors, differences between the Atlantic and east Pacific models are a result of T200 and Z850 predictors falling out and being replaced by meridional storm motion (VSPD) and the zonal 200-hPa wind speed (U200). Persistence (WCG12) also replaces VPER. The remaining climatological and environmental predictors correspond well with those in the Atlantic model in terms of importance and physical interpretation (Table 3). Three microwave predictors remain in the eastern Pacific model: MEANPCT85, STDQM85, and MAXH19. MEANPCT85 replaces the MEANH19 term found in the Atlantic model. The 85-GHz scattering signal from precipitation ice, instead of the 19-GHz emission signal from liquid rain, produces a more important microwave term for the eastern Pacific. Since lower PCT85 represent more intense convection, the coefficient values are negative. STDQM85 has a negative coefficient showing

TABLE 6. Final eastern Pacific SHIPS-MI predictor standardized coefficient values at each forecast time (h) with the explained variance (R^2) statistic at the bottom.

Predictor	$t = 12$	$t = 24$	$t = 36$	$t = 48$	$t = 60$	$t = 72$
MEANPCT85	−0.39	−0.30	−0.20	−0.11	−0.04	−0.01
STDQM85	−0.16	−0.18	−0.15	−0.09	−0.06	−0.03
MAXH19	0.03	0.07	0.06	0.08	0.14	0.13
POT	0.85	1.10	1.23	1.15	1.02	0.97
POT2	−0.68	−0.71	−0.71	−0.65	−0.48	−0.48
SHRD	−0.26	−0.52	−0.87	−0.77	−0.69	−0.68
SHRDLAT	0.36	0.56	0.83	0.62	0.52	0.47
MSWSHRD	−0.16	−0.13	−0.07	−0.01	0.01	0.07
MSW0	−0.17	−0.08	−0.06	−0.17	−0.21	−0.33
WCG12	0.37	0.25	0.15	0.09	0.05	0.05
LAT	−0.18	−0.28	−0.37	−0.35	−0.31	−0.27
EPOS	0.09	0.10	0.11	0.15	0.18	0.20
VSPD	0.04	0.07	0.01	0.13	0.13	0.14
U200	−0.03	−0.04	−0.06	−0.08	−0.11	−0.13
R^2	0.56	0.58	0.63	0.66	0.71	0.75

that low values (indicating greater symmetry) produce an intensification signal. Still, the symmetry predictor is less important that the overall mean value. Some tropical cyclones can and do intensify despite highly sheared and asymmetric convective signatures, which reduces the effectiveness of this type of symmetry predictor. Liquid rain is still represented in the eastern Pacific model with the MAXH19 term and it performs much the same way as it does in the Atlantic model. The 85-GHz signals are most important at the 12-h forecast time and decrease thereafter while MAXH19 slowly increases in importance as the forecast time increases.

The mean intensity change for the eastern Pacific sample is less than ±2 kt for all forecast times; it does not have the built-in bias toward intensification that the Atlantic version has. Explained variance ranges from 55% to 75% between 12- and 72-h forecast models (Table 6). While these values are higher than for the Atlantic, it is not an indicator of the relative performance of SHIPS-MI between basins.

b. Model performance

The performance of all forms of eastern Pacific SHIPS is somewhat worse than the corresponding Atlantic versions (Table 7; Fig. 9a). Unlike the Atlantic, the base model performs better than SHIPS-ENV for 24–72-h forecasts (Fig. 9b). This improvement exceeds 8% for forecasts beyond 48 h. Eastern Pacific predictors for SHIPS were defined as the same ones used in the Atlantic version of SHIPS for simplicity (DK99). As a result, SHIPS is less tailored to eastern Pacific tropical cyclones. Conversely, the predictors in the base model are specific to the eastern Pacific. The largest difference between the base model and SHIPS-ENV is the inclusion of latitude in the base model. Initial tropical storm latitude shows a strong negative correlation with intensity change, resulting in a statistically significant predictor. The latitude predictor in the eastern Pacific is a reflection of the sharp meridional temperature gradient near 15°N. Tropical cyclones that move north of this latitude often weaken rapidly. Ideally, the potential terms would capture this signal, but this turns out not to be the case. As forecast time increases, so does the importance of latitude, which is reflected by the increasing skill of the base model over SHIPS-ENV.

SHIPS-MI outperforms the base model for all forecast times out to 72 h with a slight overforecasting bias in the overall forecasts. Mean absolute error for SHIPS-MI increases from 5.5 kt at 12 h to 16.2 kt at 72 h compared with 5.8 and 16.3 kt for the base model (Table 7). Improvement is maximized at 12 h near 5% and decreases to near 2% for 36-h forecasts and beyond (Fig. 9b). The improvement is statistically significant to a 95% confidence level for only 12-h forecasts. SHIPS-85 is similar to SHIPS-MI out to 48 h after which it performs slightly worse. The only difference between SHIPS-MI and SHIPS-85 is the replacement of MAXH19 by MIN85. Since the contribution from MEANH19 increases as a function of time, so does the difference between SHIPS-MI and SHIPS-85. As with the Atlantic, SHIPS-MA is only able to provide a 1%–2% improvement to the base model at all forecast times.

As in the Atlantic, SHIPS-MI outperforms SHIPS (with infrared adjustment applied) at all forecast times (Fig. 10a). Improvement increases as a function of time, primarily due to the inclusion of latitude in SHIPS-MI,

TABLE 7. Jackknifed mean absolute errors (kt) for the base model, SHIPS-ENV, SHIPS-MI, SHIPS-85, and SHIPS-MA forecasts for valid eastern Pacific tropical cyclone overpasses (N).

Forecast (h)	TCs	N	Base	SHIPS-ENV	SHIPS-MI	SHIPS-85	SHIPS-MA
12	128	913	5.8	5.8	5.5	5.5	5.8
24	118	794	10.2	10.3	9.8	9.8	10.0
36	108	690	13.6	14.3	13.3	13.3	13.4
48	102	607	15.8	17.0	15.6	15.6	15.5
60	97	514	16.9	18.2	16.6	16.8	16.5
72	83	425	16.3	18.1	16.2	16.5	16.0

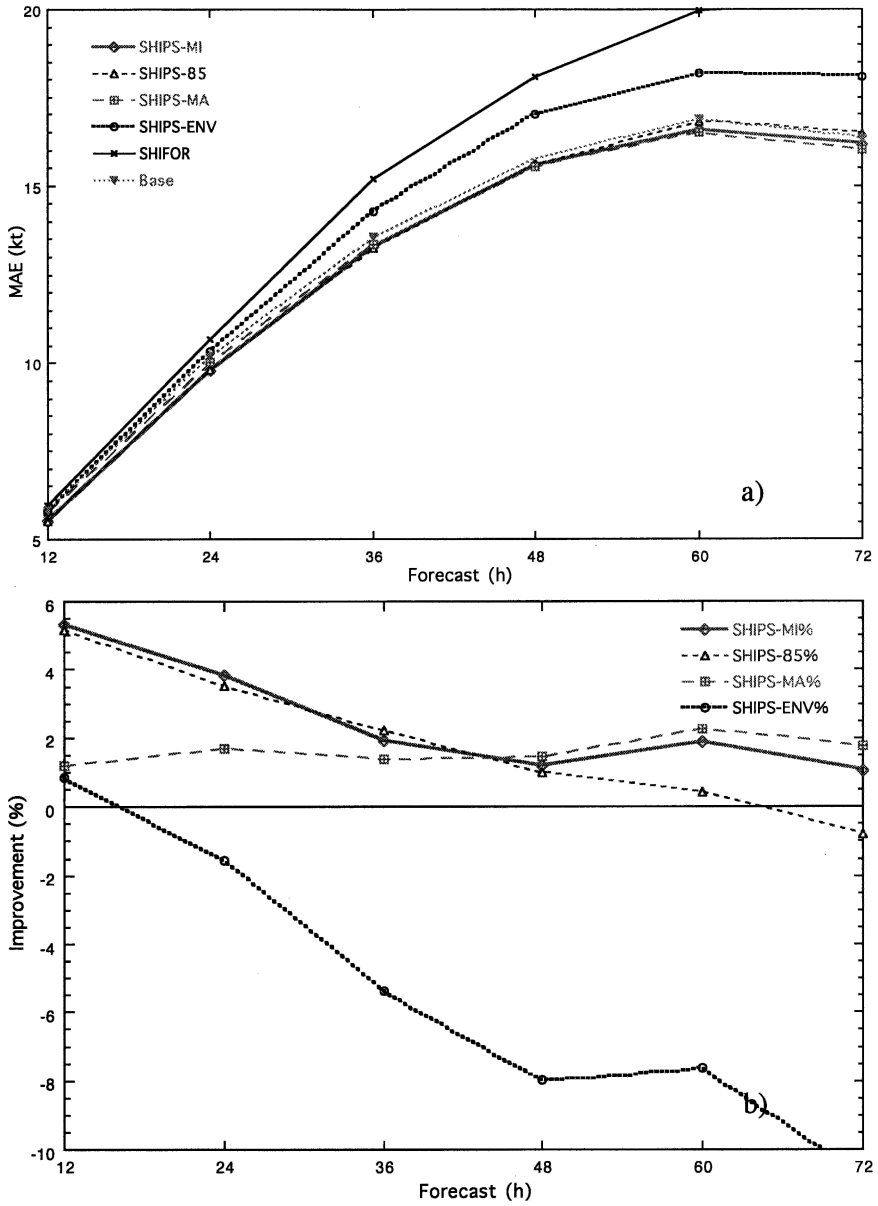

FIG. 9. (a) Eastern Pacific mean absolute errors for SHIFOR, SHIPS-ENV, SHIPS-MI, SHIPS-85, and SHIPS-MA and (b) percent improvement of SHIPS-ENV and various microwave-enhanced models normalized against SHIPS.

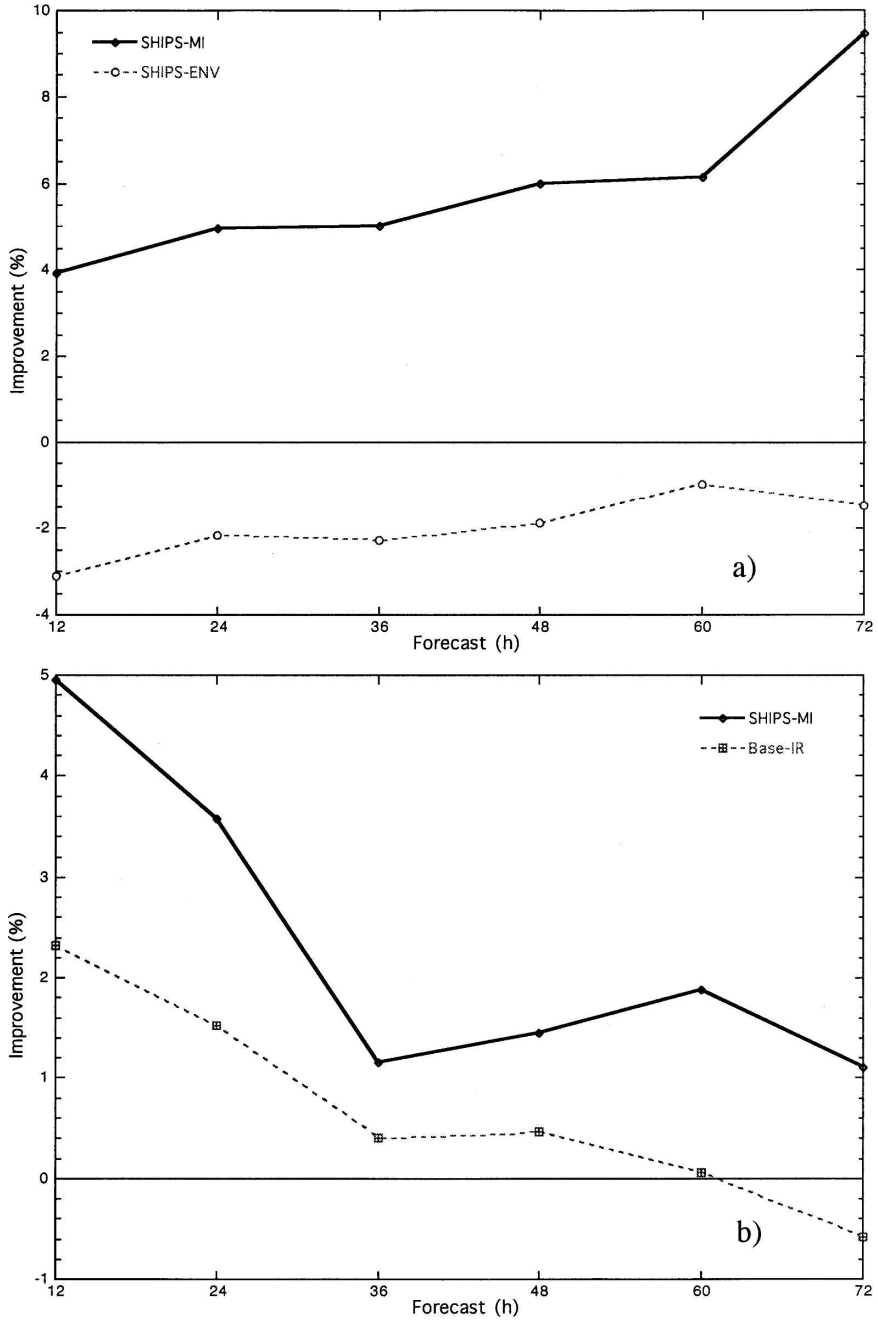

FIG. 10. Same as in Fig. 2 but for the eastern Pacific. The scaling is modified to better highlight model characteristics for this basin. Note that no oceanic heat content adjustment is applied to forecasts in Fig. 10a.

but not SHIPS. The relative skill of the microwave versus infrared enhancement is difficult to discern. However, comparing SHIPS-MI against base model and Base-IR forecasts does show that the microwave-enhanced model outperforms the infrared enhancement on a homogeneous sample (Fig. 10b). Infrared predictors are slightly less significant in the eastern Pacific providing less that 50% of the increase in skill produced by microwave predictors when compared against SHIPS-MI. In fact, the addition of infrared predictors actually decreases skill below that of the base model for 72-h forecasts.

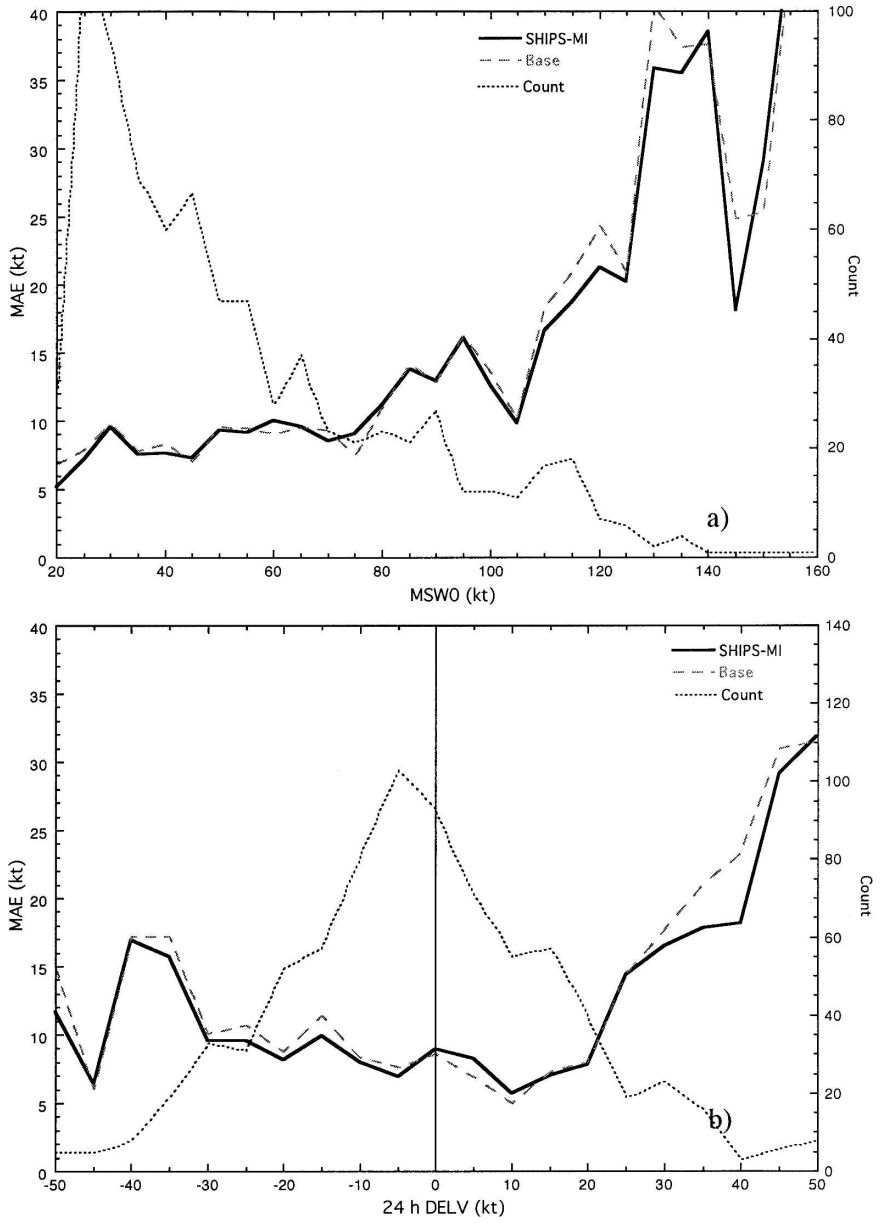

FIG. 11. Same as in Fig. 3 but for the eastern Pacific.

Analysis of 24-h mean absolute error for SHIPS-MI stratified into 5-kt intensity change and initial intensity bins reveals similar results to that seen for the Atlantic version of SHIPS-MI (Fig. 11). The greatest intensity forecast improvement is again observed in substantially intensifying or weakening tropical cyclones. The eastern Pacific SHIPS-MI generally reduces error for tropical cyclone of all initial intensities, but large reductions (or increases) in error for a particular intensity are not apparent. Overall, SHIPS-MI performs best relative to the base model for weakening tropical cyclones (Fig. 12). Substantial improvement is also observed for intensifying tropical cyclones for 12- and 24-h forecasts. In both cases, maximum improvement occurs for the 12-h forecast and decreases thereafter. Little difference exists between base model and SHIPS-MI forecasts for steady-state tropical cyclones.

The eastern Pacific version of SHIPS-MI is most skillful for 12–24-h forecasts of substantially weakening or intensifying tropical cyclones of any initial intensity. Compared with the Atlantic, the improvement is less and is concentrated at earlier forecast times. The latter

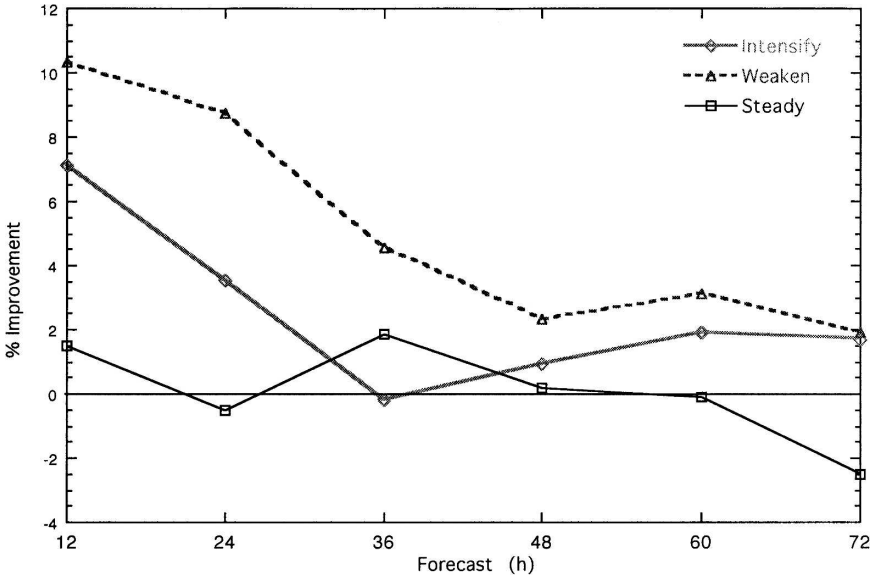

FIG. 12. Same as in Fig. 6 but for the eastern Pacific.

may be a reflection of the shorter life cycle of many eastern Pacific tropical cyclones.

6. Conclusions

The addition of microwave imagery provides a 4%–8% percent improvement in intensity forecasts when compared with a base regression incorporating only environmental and climatological predictors trained using the same training sample. The improvement is greatest for tropical cyclones with an intensity change greater than ±10 kt over a 24–48-h time period. Beyond 48 h, the importance of the microwave data to the forecasts decreases as environmental factors become dominant. The improvement produced by SHIPS-MI is a result of the inclusion of a crude representation of rainfall, convective vigor, and most importantly latent heat release into the intensity forecasts. The result of this added information is a forecast of greater intensification in tropical cyclones that have above-average rainfall and convective signatures. When combined with favorable environmental conditions for intensification, the microwave predictors allow SHIPS-MI to better capture the true intensification rate of tropical cyclones. The ability to predict the onset of rapid intensification has been one of the biggest weaknesses in current intensity forecast models and SHIPS-MI does represents a partial solution to this problem.

The pure regression form of the microwave-enhanced SHIPS model, SHIPS-MI, proved to be the best performer for both basins out to 60 h. SHIPS-MI outperforms SHIPS in both basins; however, differences inherent in the training sample and predictors included in each are responsible for a portion of this improvement. The substitution of 85-GHz predictors in place of 19-GHz predictors reduced skill by some degree, but still left a model superior to the base regression at most forecast times. This is an important operational consideration as 19-GHz predictors cannot be used in close proximity to land where intensity forecasts are critical. The microwave adjustment model, SHIPS-MA, shows only marginal improvement over the base regression. The inclusion of microwave data into the pure regression allows for much larger improvements to be made.

The most important factor that must be taken into account when incorporating microwave data into operational intensity forecast models is the temporal availability of the microwave data. Timely microwave observations of a tropical cyclone may not be available at or near the synoptic forecast times of 0000, 0600, 1200, and 1800 UTC. The question of how "old" a microwave tropical cyclone overpass is before the information present becomes dated and unrepresentative remains open. Given that SSM/I and TMI data latency is on the order of 2–3 h, the longer an overpass can still be considered representative increases the usefulness of microwave data in an operational setting.

Despite this concern, preliminary results from SHIPS-MI tested in near–real time during the 2004 tropical cyclone season suggest microwave availability for nearly 50% of all possible forecasts for tropical cy-

clones not over land assuming a best-case scenario for data reception. The results of this testing also indicate that when microwave data are available, improvement in the intensity forecast *does* occur for several tropical cyclone types and conditions.

The difference in SHIPS-MI between the Atlantic and eastern Pacific basins provides an avenue for future research. It is currently unclear whether the differences in the number and types of microwave predictors are a result of statistical randomness, or have their basis in differing physical processes between the basins. One possibility for the replacement of MEANH19 by MEAN85 in the eastern Pacific model is that the shorter life cycle of eastern Pacific storms is better captured by the shorter-term convective signatures present in the PCT85 data. The techniques used in this work will be expanded to tropical cyclones in the western North Pacific, Indian Ocean, and Southern Hemisphere to answer this question as well as provide microwave-enhanced models to those basins.

Additional work to extend the SSM/I dataset back to 1988 is being undertaken to produce a training sample comparable in size to that used by SHIPS. The smaller sample size used to train SHIPS-MI complicates comparisons with SHIPS and limits direct comparison of their forecasts. Increasing the SHIPS-MI sample size should help alleviate this concern.

Acknowledgments. Funding for this research was made possible by NASA Grant NAG-512563. SSM/I data were provided by Global Hydrology Resource Center (GHRC), and TMI data were provided by Goddard Space Flight Center. We would like to acknowledge John Kaplan and John Knaff, and the staff of GHRC and the University of Alabama in Huntsville, for their assistance in making this work possible. The views, opinions, and findings in this report are those of the authors and should not be construed as an official NOAA and or U.S. government position, policy, or decision.

REFERENCES

Aberson, S. M., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). *Mon. Wea. Rev.,* **122,** 2804–2815.

Bankert, R. L., and P. M. Tag, 2002: An automated method to estimate tropical cyclone intensity using SSM/I imagery. *J. Appl. Meteor.,* **41,** 461–472.

Brown, D. P., and J. L. Franklin, 2004: Dvorak tropical cyclone wind speed biases determined from reconnaissance-based "best track" data (1997–2003). Preprints, *26th Conf. on Hurricanes and Tropical Meteorology,* Miami, FL, Amer. Meteor. Soc., CD-ROM, 3D.5.

Cecil, D. J., and E. J. Zipser, 1999: Relationships between tropical cyclone intensity and satellite based indicators of inner core convection: 85-Ghz ice-scattering and lightning. *Mon. Wea. Rev.,* **127,** 103–123.

——, ——, and S. W. Nesbitt, 2002: Reflectivity, ice scattering, and lightning characteristics of hurricane eyewalls and rainbands. Part I: Quantitative description. *Mon. Wea. Rev.,* **130,** 769–784.

DeMaria, M., and J. Kaplan, 1994: A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting,* **9,** 209–220.

——, and ——, 1999: An updated Statistical Hurricane Intensity Prediction Scheme for the Atlantic and eastern North Pacific basins. *Wea. Forecasting,* **14,** 326–337.

——, M. Mainelli, L. K. Shay, J. Knaff, and J. Kaplan, 2005: Further improvements to the updated Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting,* **20,** 531–543.

Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Memo. NESDIS 11, 47 pp.

Fitzpatrick, P. J., 1997: Understanding and forecasting tropical cyclone intensity change with the Typhoon Intensity Prediction Scheme (TIPS). *Wea. Forecasting,* **12,** 826–846.

Franklin, J. L., M. L. Black, and K. Valde, 2003: GPS dropwindsonde wind profiles in hurricanes and their operational implications. *Wea. Forecasting,* **18,** 32–44.

Glass, M., and G. W. Felde, 1992: Intensity estimation of tropical cyclones using SSM/I brightness temperatures. Preprints, *Sixth Conf. on Satellite Meteorology and Oceanography,* Atlanta, GA, Amer. Meteor. Soc., J8–J10.

Hollinger, J., R. Lo, G. Poe, R. Savage, and J. Peirce, 1987: Special Sensor Microwave/Imager user's guide. Naval Research Laboratory, Washington, DC, 177 pp.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity. NOAA Tech. Memo. NWS NHC-10, 22 pp.

Knaff, J. A., M. DeMaria, C. R. Sampson, and J. M. Gross, 2003: Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting,* **18,** 80–92.

——, C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting,* **20,** 688–699.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817.

Lee, T. F., F. J. Turk, J. Hawkins, and K. Richardson, 2002: Interpretation of TRMM TMI images of tropical cyclones. *Earth Interactions,* **6**. [Available online at http://EarthInteractions.org.]

Neumann, C. J., 1972: An alternate to the HURRAN tropical forecast system. NOAA Tech. Memo. NWS NHC-34, 34 pp.

——, M. B. Lawrence, and E. L. Caso, 1977: Monte Carlo significance testing as applied to statistical tropical cyclone prediction models. *J. Appl. Meteor.,* **16,** 1165–1174.

Rao, G. V., and P. D. MacArthur, 1994: The SSM/I estimated rainfall amounts of tropical cyclones and their potential in predicting the cyclone intensity changes. *Mon. Wea. Rev.,* **122,** 1568–1574.

——, and J. H. McCoy, 1997: SSM/I measured microwave brightness temperatures (TB's), anomalies of TB's and their relationship to typhoon intensification. *Nat. Hazards,* **15,** 1–19.

Rodgers, E. B., and H. F. Pierce, 1995: A satellite observational study of precipitation characteristics in western North Pacific tropical cyclones. *J. Appl. Meteor.,* **34,** 2587–2599.

——, S. W. Chang, and H. F. Pierce, 1994: A satellite observational and numerical study of precipitation characteristics in western North Atlantic tropical cyclones. *J. Appl. Meteor.,* **33,** 129–139.

Spencer, R. W., H. M. Goodman, and R. E. Hood, 1989: Precipitation retrieval over land and ocean with SSM/I: Identification and characteristics of the scattering signal. *J. Atmos. Oceanic Technol.,* **6,** 254–273.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.