An Automated Method to Estimate Tropical Cyclone Intensity Using SSM/I Imagery

RICHARD L. BANKERT AND PAUL M. TAG

Naval Research Laboratory, Monterey, California

(Manuscript received 7 March 2001, in final form 2 December 2001)

ABSTRACT

An automated method to estimate tropical cyclone intensity using Special Sensor Microwave Imager (SSM/I) data is developed and tested. SSM/I images (512 km × 512 km) centered on a given tropical cyclone (TC), with a known best-track intensity, are collected for 142 different TCs (1988–98) from the North Pacific, Atlantic, and Indian Oceans. Over 100 characteristic features are computed from the 85-GHz (H-pol) imagery data and the derived rain-rate imagery data associated with each TC. Of the 1040 sample images, 942 are selected as training samples. These training samples are examined in a feature-selection algorithm to select an optimal subset of the characteristic features that could accurately estimate TC intensity on unknown samples in a *K*-nearest-neighbor (*K*-NN) algorithm. Using the 15 selected features as the representative vector and the best-track intensity as the ground truth, the 98 testing samples (taken from four TCs) are presented to the *K*-NN algorithm. A root-mean-square error (rmse) of 19.8 kt is produced. This ''snapshot'' approach is enhanced (rmse is 18.1 kt) when a TC intensity history feature is added to 71 of the 98 samples. Reconnaissance data are available for two recent (1999) Atlantic hurricanes, and a comparison is made in the rmse using those data as ground truth versus best track. For these two TCs (17 SSM/I images), an rmse of 15.6 kt is produced when best track is used and an rmse of 19.7 kt is produced when reconnaissance data are used as the ground truth.

1. Introduction

Analysis of remotely sensed data has become an increasingly valuable tool for determining tropical cyclone (TC) location and structure, formulating TC intensity estimates, and predicting strength and movement. Aircraft reconnaissance missions stopped in 1987 in the western Pacific, leaving the Atlantic basin as the only region for which TCs are monitored routinely by aircraft. Critical decision making that includes public and marine advisories and evacuation plans is impacted by TC analysis (intensity, location, structure, etc.). Inaccurate intensity estimates and predictions are potentially costly in both monetary and human terms.

Dvorak (1975, 1984) has provided a reliable method to estimate TC intensity manually from visible and infrared (vis/IR) satellite imagery by using subjective pattern recognition and a set of applicable rules. Velden et al. (1998) introduced an objective Dvorak technique (ODT) to eliminate much of the subjectivity in the standard Dvorak method. ODT intensity estimates have been shown to be competitive with or superior to three independent Tropical Analysis Center estimates (Olander and Velden 1999).

One of the main limitations of the Dvorak technique (and any vis/IR technique) is that low- and midlevel clouds can be obscured by upper-level clouds. Unknown low-level structure and circulation make estimating intensity and locating the circulation center difficult. For Special Sensor Microwave Imager (SSM/I) passive microwave channel images, most upper-level (nonprecipitating) clouds are essentially transparent (Hawkins et al. 1998). Using SSM/I images to examine TC structure has an advantage when compared with the limitations of other types of imagery (Hawkins et al. 2001). Rainbands and a TC center (when it exists) can be seen in the 85-GHz channel images when the banding structure is often obscured by upper-level clouds as seen in vis/IR imagery. Figure 1 provides an example comparison, illustrating how significant structural information is available in 85-GHz data and is not available in coincident vis/IR imagery.

Developing a manual technique that uses SSM/I imagery to estimate TC intensity is possible; however, an automated approach that provides a quick and accurate intensity estimate at the time of an image is even more desirable in an operational environment. In developing this kind of analysis/decision aid, image characteristics are extracted and compared with those characteristics taken from historical SSM/I images for which the TC intensity is known in order to arrive at an intensity estimate for the current TC.

May et al. (1997) computed empirical orthogonal functions (EOFs) from SSM/I images and used their values as inputs to a trained neural network to automate

Corresponding author address: Richard L. Bankert, Naval Research Laboratory, Monterey, CA 93943-5502.
E-mail: bankert@nrlmry.navy.mil

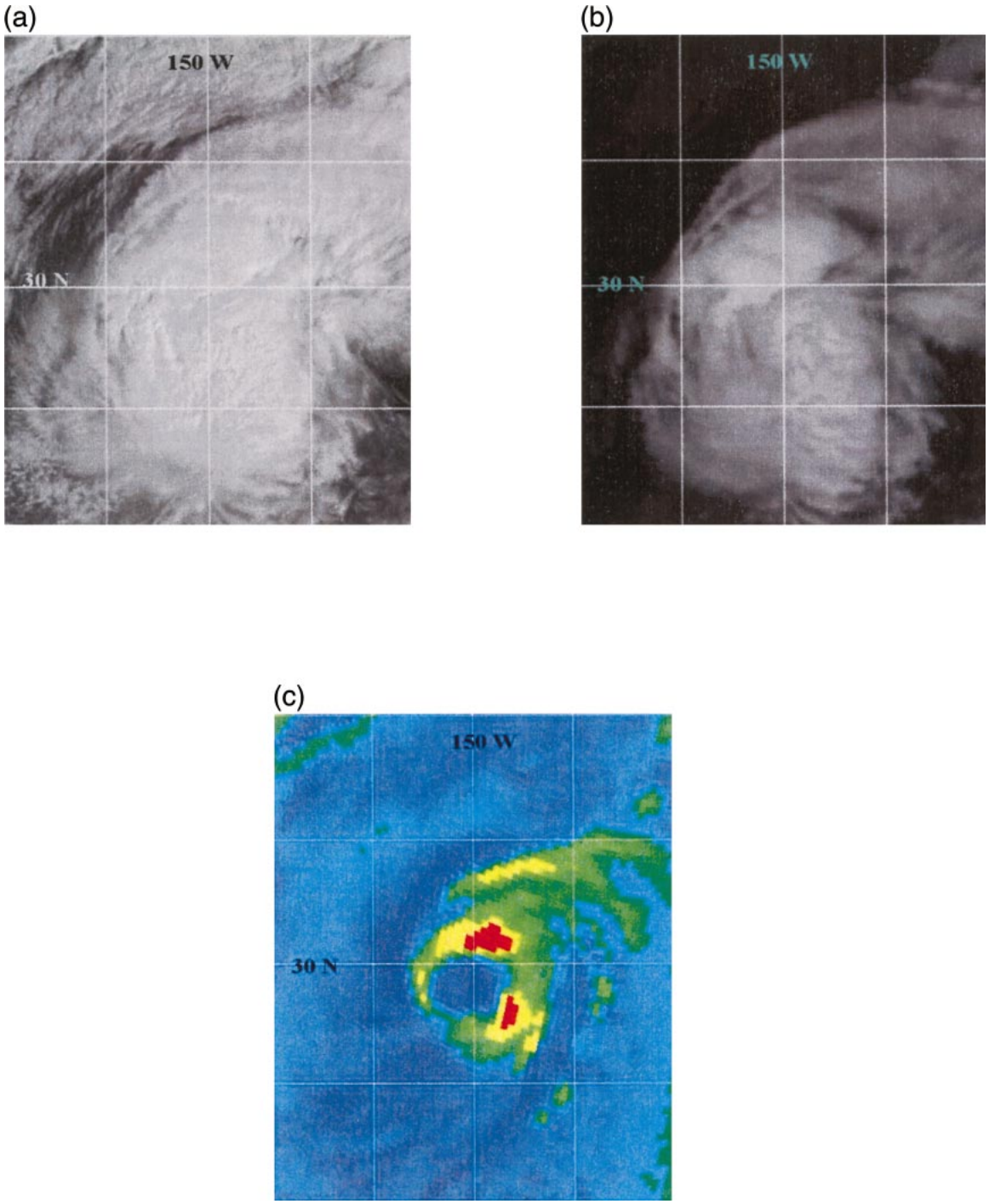

FIG. 1. The structure of Typhoon Joan (2232 UTC 22 Oct 1997) is obscured by high clouds in Geostationary Operational Environmental Satellite *GOES-9* (a) visible and (b) IR imagery. This structure (including the eye) is apparent in the (c) SSM/I 85-GHz image. The red areas in (c) indicate low brightness temperatures (high scattering due to large ice particles—convective precipitation), and blue areas indicate higher brightness temperatures (no or limited convection).

TC intensity estimation. The research described here examines the viability of SSM/I image characteristic features (minimum and mean pixel brightness temperature, texture measurements, etc.) and local information (time, location, etc.) to estimate TC intensity as measured by the maximum wind speed. A subset of these features is used to represent a TC, at the time of the image, in a *K*-nearest-neighbor (*K*-NN) algorithm (Duda and Hart 1973) that provides an intensity estimate. The *K* represents the number of training samples that are

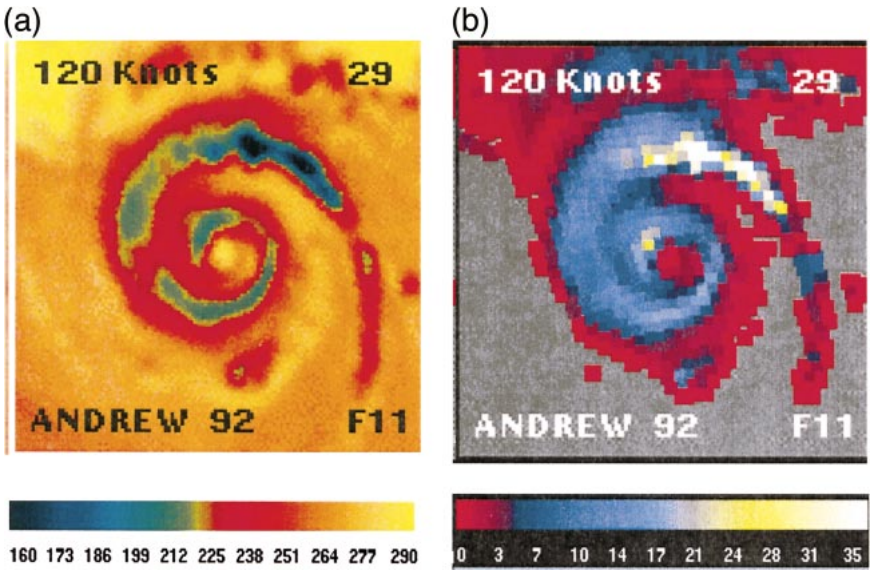

FIG. 2. (a) 512 × 512 SSM/I 85-GHz image (K) and (b) 512 × 512 SSM/I-derived rain-rate image (mm h^{-1}) for Hurricane Andrew (2252 UTC 25 Aug 1992).

used to estimate the intensity of any given testing sample. Using available data, including Dvorak estimates, surface observations, and reconnaissance reports (Atlantic and eastern Pacific), a best-track maximum wind speed has been determined for each TC at a given time by either the Joint Typhoon Warning Center (JTWC) or the National Hurricane Center (NHC). This best-track intensity will serve as the ground truth and has an estimated error range of ±10–15 kt.

A description of the SSM/I data, image characteristics related to TC intensity, and other potential features is presented in section 2. Reduction of the feature vector and the application of *K*-NN are discussed in section 3. Section 4 includes test results and algorithm evaluations that include completely independent testing data consisting of multiple images of individual TCs. A summary and concluding discussion follow in section 5.

2. SSM/I data and TC characteristics

The SSM/I instrument resides on a polar-orbiting platform within the Defense Meteorological Satellite Program and uses a suite of passive microwave channels (19, 22, 37, and 85 GHz). Image characteristics of the 85-GHz (H-pol) channel provide one dataset to be used within this study. In addition to having the highest resolution, this channel has an advantage over other microwave channels with its ability to detect convective precipitation due to high scattering from ice particles (e.g., TC rainbands) that produce low brightness temperature signatures. The 85-GHz image data have a spatial resolution of 13 km × 15 km. The interpolation algorithm of Poe (1990) provides a mechanism to map the oversampled 85-GHz image data to 1–2-km resolution to create a product that enhances TC structure and can be compared more easily with vis/IR images. Another valuable piece of data for TC analysis is a rain-rate ''image'' that takes advantage of the information extracted from multiple SSM/I channels (19, 22, and 85 GHz; Ferraro 1997). Examples of 85-GHz and rain-rate images are shown in Fig. 2.

SSM/I data have been collected for 142 different TCs (1988–98) from the Atlantic (19 TCs), eastern Pacific (47 TCs), western Pacific (73 TCs), and Indian (3 TCs) Ocean basins. One thousand forty 512 × 512 (kilometers or pixels) images that have been centered on the TC center and include the associated best track TC intensity in knots are processed to search for and to evaluate representative features and to test an intensity estimation algorithm.

General and specific TC image characteristics, along with local features (101 in total) are computed or measured. A subset of these features or attributes will be used to represent the TC at the time of the image. The goal is to find the ''optimum'' combination of features to estimate TC intensity accurately on unknown cases. Some of the more general attributes include spectral characteristics that are simple statistical measurements (maximum, minimum, mean, median, and mode pixel values, range of pixel values, and standard deviation) of both 85-GHz and rain-rate imagery. Textural feature values of the 85-GHz data are also computed. These textural characteristics include contrast, entropy, angular second moment, local homogeneity, and so on and are computed over the entire 512 × 512 pixel image (Bankert 1994). They provide a representation of the spatial distribution of the brightness temperatures within the image. Figure 3 is a display of an 85-GHz image with high contrast compared with one with low contrast (different TC).

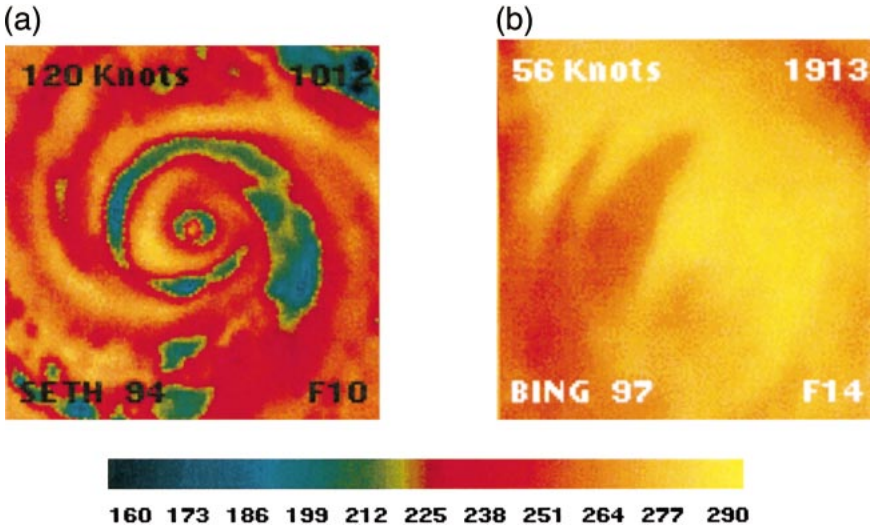

FIG. 3. (a) An 85-GHz image (Typhoon Seth) with relatively high contrast in comparison with (b) Tropical Storm Bing, which has relatively low contrast after the TC has lost all deep convection and contains only low cloud structure (extratropical—44°N).

The areal extent (512 km × 512 km) of both image types (85 GHz and rain rate) is divided into four quadrants (NE, SE, NW, SW relative to center pixel) over the entire (512 × 512) image and over the inner 150 × 150 pixel region to understand in a general sense the distribution or organization of the convective areas within the TC environment. For all quadrants (16 total—both image types with two quadrant sets), the summation of all pixel values within the quadrant is saved as a feature. In addition, the summations of all pixel values for the 512 × 512 and the inner 150 × 150 pixel region are computed.

To determine more specific TC characteristics within an image, the data are segmented. For 85-GHz data, pixels with brightness temperatures of 255 K or less and rain-rate pixels with values greater than 3 mm h^{-1} are segmented as ''TC pixels.'' This segmentation puts the focus on the convective areas of the image. See Fig. 4

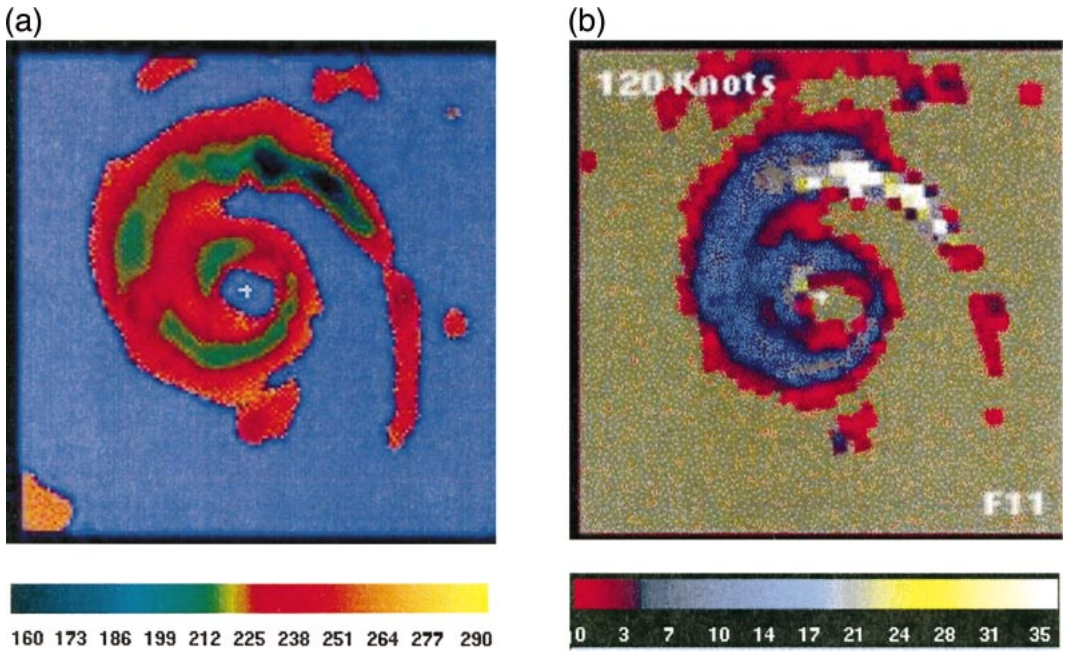

FIG. 4. (a) 512 × 512 segmented 85-GHz (K) and (b) rain-rate (mm h^{-1}) images for Hurricane Andrew as shown in Fig. 2. The 85-GHz image in (a) is segmented at 255 K or less (bright blue >255 K). The rain-rate image in (b) is segmented above 3 mm h^{-1} (tan ≤ 3 mm h^{-1}).

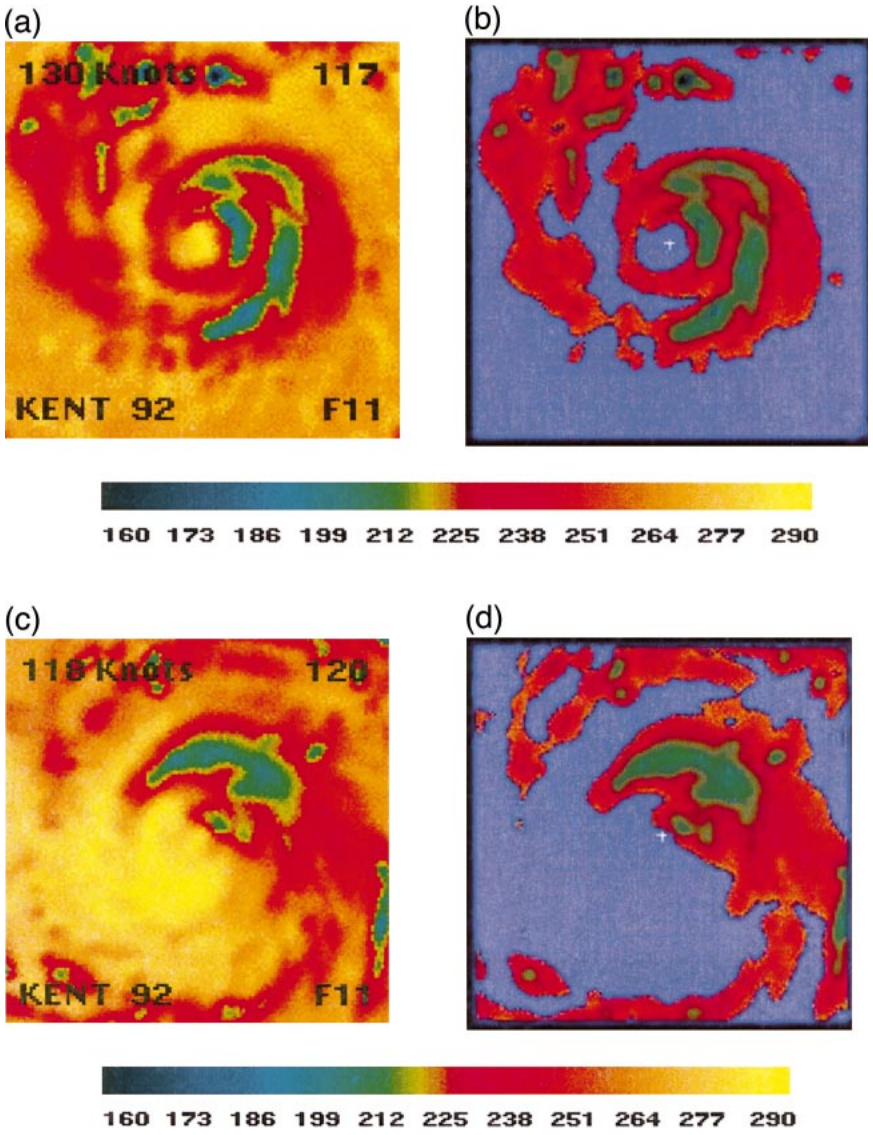

FIG. 5. Enclosed eye feature examples. Typhoon Kent at 1942 UTC 11 Aug 1992 [(a) unsegmented image] with enclosed eye at 255-K threshold [(b) segmented image] and later at 1929 UTC 12 Aug 1992 [(c) unsegmented image] with no enclosed eye at 255-K threshold [(d) segmented image].

for examples of images segmented at these thresholds. A pixel count is performed over the entire segmented image. The E–W and N–S extent of segmented pixels are measured to estimate the size of the TC.

The most relevant extracted image features to define the TC include the existence of an enclosed eye at the 255-K threshold (yes/no feature). The presence of an enclosed eye at this threshold may indicate a more intense TC. Examples of TCs with and without enclosed eyes are shown in Fig. 5. The size of this eye (if it exists) is computed also. To obtain an even more specific representation of the brightness temperature and extent of the eyewall, the maximum and minimum brightness temperature thresholds for which an enclosed eye exists are computed, along with the associated range (maximum–minimum). Enclosed eye features are analogously computed from the rain-rate data (using 3 mm h^{-1} as the initial threshold).

Following the technique of Velden et al. (1998), an algorithm has been developed to represent the degree of organization of the TC. The warmest pixel value (WPC) within a 40-km radius of the center is saved as one feature. The pixel temperatures on 1-km-wide concentric rings (24–136-km radius from center) are analyzed such that the coldest pixel value (ST) is determined from the set of extracted warmest pixel values on each ring, with ST saved as a feature. WPC − ST is a third feature derived from this method as a repre-

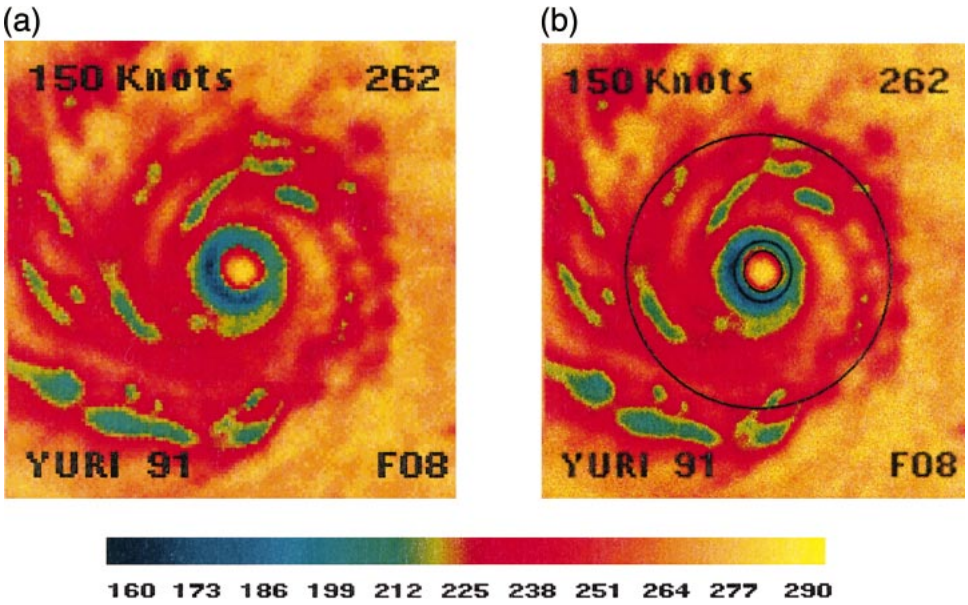

FIG. 6. (a) Typhoon Yuri (0845 UTC 27 Nov 1991) has a WPC − ST value of 73 K. The typhoon center has a representative (WPC) brightness temperature of 285 K and a representative (ST) surrounding temperature of 212 K. (b) The radius limits for computing WPC and ST are marked. WPC is found within the area inside the middle ring (40-km radius), and ST is found between the innermost ring (24-km radius) and the outer ring (136-km radius).

sentation of the difference between the eye temperature and the ''surrounding'' temperature (Fig. 6). Six other attributes associated with 1-km-wide concentric rings surrounding the center of the TC are computed to represent the banding structure and convective nature of the TC at the time of the image. The banding structure and organization and the strength of the convection around the TC center are known to be related to the intensity. The summations of the pixel values that are below a certain threshold in the 85-GHz data and above a certain threshold in the rain-rate data are computed for each 1-km ring. The maximum summation value (normalized by the radius) is a characteristic feature of the imagery. The summation is computed for three different thresholds in both image types: 225, 240, and 255 K and 3, 10, and 17 mm h^{-1}. Along those same lines, other representative features are extracted, including the shortest radius (from the center) to the 1-km-wide concentric ring that has 60% or more of its pixels with a brightness temperature less than 250 K. To determine the extent of this cold-banding area, each successive 1-km concentric ring is examined in this manner until the rule (60% at <250 K) is no longer satisfied. The ratio of the inner radius and the computed thickness is saved as a feature.

As a representative value of the TC data (relative to peak activity), |yearday − 237| is included as a local feature. Other local features include latitude, longitude, and sea surface temperature (climatological data from the National Centers for Environmental Prediction). Table 1 is a summary of the features discussed in this section. They provide the total feature set from which a subset is selected through a feature selection process described in the next section. That subset is subsequently used in a TC intensity estimation algorithm.

TABLE 1. Summary list of extracted image (85 GHz and rain rate) features.

General
Spectral
Textures
Pixel value summation (total, quadrant, inner region)
Local
Latitude
Longitude
|Yearday − 237|
Sea surface temperature (climatological values)
Segmented image
Pixel count
E–W, N–S extent
TC specific
Enclosed eye (yes/no)
Size of enclosed eye
Max/min/range of thresholds for enclosed eyes
Warmest pixel value in TC center (WPC)
Representative surrounding temperature (ST)
WPC − ST
Banding and concentric ring features

3. Feature selection and *K* nearest neighbor

Presenting redundant and irrelevant features to a pattern recognition algorithm has been found to degrade the algorithm's performance. Finding the optimum subset of the features (section 2) that allows the algorithm

to perform with highest possible accuracy on unseen samples is a difficult problem.

Aha and Bankert (1995) examined various combinations of search and evaluation algorithms for reducing the size of feature vectors. Using the same function in the feature subset evaluation as was used in the automated pattern recognition algorithm increases classification accuracy. Bankert and Aha (1996) demonstrated the importance of the choice of feature selection algorithms when working on cloud classification from satellite imagery. These previous uses of feature selection algorithms required the data to be organized in discrete classes, but for this research the algorithm has been modified to make use of the best-track (postanalysis consensus) intensity associated with each training image (i.e., continuous data). Instead of searching for a feature subset that maximizes classification accuracy, the search is for a subset that minimizes the root-mean-square error (rmse measured in knots) of the TC intensity estimates.

The 1040 images are separated into training (942 samples) and testing (98 samples) datasets. The separation is partially determined by storm name (e.g., all of the Hurricane Andrew images are placed in the training set). The training set supplies the feature data and best-track intensities associated with each image to be used in the feature selection algorithm. The testing set consists of four storms: Hurricane Guillermo (August 1997), Supertyphoon Oliwa (September 1997), Supertyphoon Paka (December 1997), and Supertyphoon Babs (October 1998).

The feature selection algorithm consists of a variation of the backward sequential selection (BSS) algorithm and a *K*-NN that serves as the evaluation function. The *K*-NN algorithm also serves as the pattern recognition algorithm and uses the feature subset, found by the search algorithm, that maximizes the evaluation function. In general, BSS begins with all features and repeatedly removes the feature that, after being removed, allows the maximum performance to increase. Using BSS as the search algorithm is not preferable when there is a large number of features. BSS is biased toward yielding a large percentage of the total number of features because, frequently, removing only one feature at a time may not improve performance; removing multiple features at one time might prove useful, however. To this end, a variation of BSS, called beam, is used. The beam variant modifies BSS so that it can perform more reasonably when there are numerous features. For beam, the feature space is randomly sampled for a fixed number of iterations to save the best-performing (based on the evaluation function output) subsets of a given size (e.g., 25 features). The number of chosen subsets is fixed. For a beam search, the fixed-size queue of subsets is maintained. Subsets in the queue are subsequently reduced in size and evaluated, with the queue updated each time the search is exhausted and the best state (i.e., a feature subset of one of the initial subsets) is found.

TABLE 2. Selected feature list.

Latitude
Longitude
|Yearday − 237|
85 GHz: pixel summation—NE quadrant (512 × 512)
85 GHz: pixel summation—NE quadrant (inner 150 × 150)
85 GHz: segmented—normalized mean radius
85 GHz: range of pixel values (512 × 512)
85 GHz: range of temperature thresholds for which enclosed eye exists
Rain rate: average pixel value for those > 0 mm h^{-1} (512 × 512)
Rain rate: No. pixels > 0 mm h^{-1} (inner 150 × 150)
Rain rate: pixel summation—SE quadrant (512 × 512)
Rain rate: pixel summation—NW quadrant (inner 150 × 150)
Rain rate: mean pixel value (512 × 512)
Rain rate: mode pixel value (512 × 512)
Rain rate: banding feature—max summation of pixels on concentric rings (>3 mm h^{-1})

The *K*-NN is the pattern recognition algorithm of choice because of its nonlinear, nonparametric nature; therefore, it requires no assumptions to be made about the distribution of the data. The *K*-NN algorithm computes the similarity distances (1) in feature space between the testing sample and each training sample:

$$\sum (\text{testing feature}_i - \text{training feature}_i)^2. \quad (1)$$

Each feature value used in (1) has been normalized, and the summation is over all features. Using the single nearest-neighbor distance as the standard for inclusion, those training samples within a distance factor (1.5 × nearest-neighbor distance) are used to estimate the testing sample intensity. A simple averaging technique is performed on those *K*-nearest-neighbor best track intensities. A leave-one-out cross-validation test is applied to the training set, and the rmse is computed relative to each feature subset. For this test, each training sample is presented, via feature subset representation, to the *K*-NN algorithm, with the similarity distance to each of the remaining 941 training samples being computed. The error for any given training sample is the difference between the *K*-NN estimated intensity and the best-track intensity. The reduced subset that produces the minimum rmse after the search and *K*-nearest-neighbor evaluation is the final selected set. These selected features are listed in Table 2.

4. Results

Using the 15 selected features as a representative vector for each sample and the best-track intensity as the ground truth, the leave-one-out test of the 942 training samples produced an rmse of 14.9 kt, an average absolute error (AAE) of 10.8 kt, and an average percentage error (APE) of 16.2%. Although this leave-one-out cross-validation test gives some indication of the algorithm's potential performance on unknown cases, testing on completely independent samples is more useful in evaluating the performance on unseen instances. The

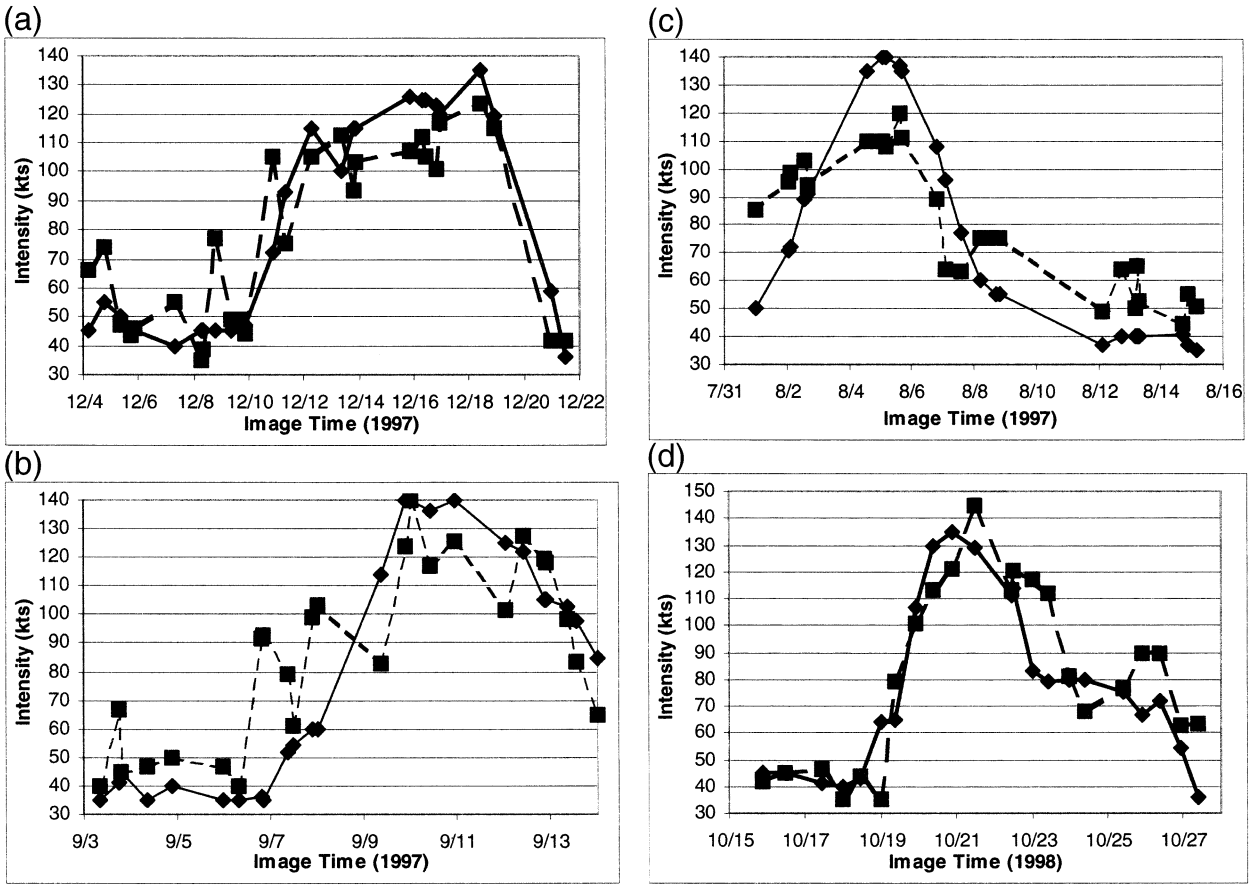

FIG. 7. Time series plots of best-track intensity (solid line, diamond points) and *K*-NN estimated (dashed line, square points) intensity for the four testing TCs: (a) Paka, (b) Oliwa, (c) Guillermo, and (d) Babs.

selected features (Table 2) are computed for the 98 testing samples. An estimated intensity for those samples, represented by the selected feature vector, is determined using the *K*-nearest-neighbor algorithm in association with the best-track intensity values of the training samples. The overall rmse for the testing samples is 19.8 kt. A majority (54%) of the samples had an intensity estimate within 15 kt of the best-track intensity, but 12% had an error greater than 30 kt. Percentage errors for the testing samples are as high as 166%; however, a majority (58%) of the samples have much more acceptable errors of less than 20%. For individual TCs, rms errors ranged from 15.4 kt for Supertyphoon Paka to 24.6 kt for Supertyphoon Oliwa. See Table 3 for a summary of the results for the four testing TCs.

TABLE 3. *K*-NN testing results (98 total samples).

Tropical cyclone	Rmse (kt)	AAE (kt)	APE (%)
Oliwa (25 samples)	24.6	19.0	34.5
Guillermo (24 samples)	21.4	19.7	30.6
Paka (27 samples)	15.4	12.7	18.3
Babs (22 samples)	16.5	12.7	18.4
Overall	19.8	16.0	25.5

Figures 7a–d are time series plots of the best-track intensity versus *K*-nearest-neighbor estimation for the testing samples. For Oliwa (Fig. 7b), the high variability along the *K*-NN-predicted timeline is clearly evident, especially when compared with the much smoother best-track timeline. As shown in Figs. 8a and 8b, two Oliwa samples (1852 and 2007 UTC 6 September 1997) during the early stages of cyclone development (36 and 35 kt, respectively) have extreme errors (55.7 and 58.0 kt, respectively) that are probably a result of poor interpretation of the convective area around the center of the cyclone by the automated algorithm. Even Paka (Fig. 7a), which has the lowest rmse, has ''spikes'' along the timeline that are inconsistent with expectations. These expectations are based upon the best-track data however, and some of these spikes could be representing actual strengthening/weakening cycles not captured in the smooth best-track data.

The TC in Fig. 9a has an acceptable *K*-NN intensity estimate error of 7.3 kt. However, two other images (Figs. 9b,c), with best-track intensities that are nearly to Fig. 9a have associated significant intensity estimate errors. These results give an indication of the difficulty in determining a set of characteristics that can describe

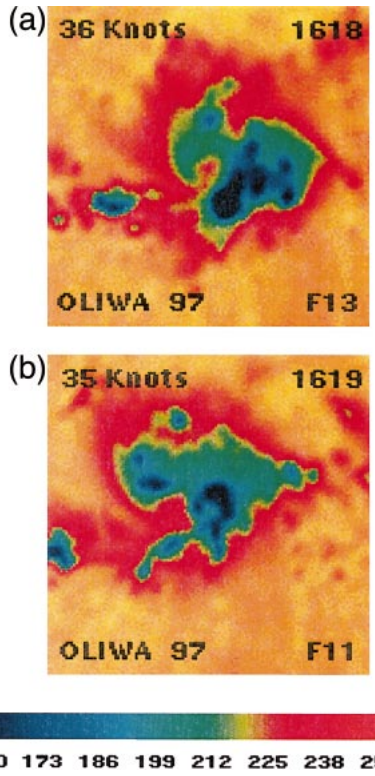

FIG. 8. Tropical Cyclone Oliwa 85-GHz images (512 × 512) at a time of relatively low intensity with extensive areas of cold brightness temperatures (blue and green areas) that are contributing to the misclassification of the cyclone as being much more intense than the best-track analysis: (a) 1852 and (b) 2007 UTC 6 Sep 1997.

images with similar intensities. The *K*-NN-predicted intensity for the weakening TC in Fig. 9b is 17.3 kt less than the best-track intensity, and the estimated intensity for the strengthening TC in Fig. 9c was 43.3 kt greater than the best-track intensity.

To evaluate the use of a feature selection algorithm, a comparison is made between the *K*-NN algorithm's performance using the selected features and its performance using specific feature types. Overall, rms errors (testing samples) ranged from 24.8 kt for general 85-GHz features to 38.7 kt for local features (latitude, longitude, date, SST). A complete summary of test results is shown in Table 4. As expected, a mixture of feature types, as chosen through a selection algorithm that includes the *K*-NN as the evaluation function, produces the lowest rmse.

None of the feature subsets described above takes into account the history of the TC prior to the time of the sample image. Having this additional prior information should produce better results than using ''snapshot'' characteristics alone. In a real-time setting, the training data, which provide the nearest-neighbor matches for the testing set, can contain a priori information; therefore, keeping track of the *K*-NN-estimated intensity at specific times for a given testing sample and using that data as an additional feature to augment the image characteristics is possible. To test this procedure, the 12-h (previous to image time) best-track estimate is added as a feature for each of the training samples. For the testing set, when appropriate, the *K*-NN-predicted intensity from approximately 12 hours previous is added as a feature for a given testing sample (a total of 16 features); otherwise, the snapshot algorithm (15 features) is applied. The result of this application produced 27 testing samples used in the 15-feature algorithm and 71 testing samples used in the 16-feature algorithm. A set of results analogous to those of Table 3 is presented in Table 5. Time series plots for each of the four storms are provided in Figs. 10a–d. Points along the *K*-NN timeline that are associated with 15 features only (no history) are marked with a triangle. The overall rmse of 18.1 kt is an improvement when compared with the 19.8-kt rmse that is produced for the snapshot-only (15 features) version (Table 3). This 8.6% decrease in rmse goes with a 10.6% decrease in AAE and a 12.2% decrease in APE. In terms of improvement of the 71 samples that used the history feature, rmse, AAE, and APE had an 11.7%, 14.7%, and 16.0% reduction, respectively, when compared with using no additional history feature. Thirty-seven of the 71 samples that added the history feature had a more accurate intensity estimate, with no change in the accuracy for 23 samples. Of the remaining 11 samples that had a higher error, 6 are Oliwa samples for which a large error early in the storm history skewed subsequent testing samples with the addition of these largely incorrect history features.

An evaluation of the TC intensity estimation algorithm using ground-truth data other than best-track intensities is possible with the availability of reconnaissance data for Atlantic Hurricanes Dennis and Floyd from 1999. Seventeen SSM/I images associated with these TCs were matched with reconnaissance data (minimum central pressure) within 6 h of the SSM/I data times. The pressure observation was converted to wind speed using the pressure–wind relationship of Holliday (1969):

$$\text{max wind speed (mi h}^{-1}) = 1431 - (1.37 \times \text{central pressure}), \quad (2)$$

where central pressure is in hectopascals. Best-track intensities from the NHC are also available for these TC samples. The first two columns in Table 6 consist of the reconnaissance and best-track TC intensities. There are a few disagreements in both Dennis and Floyd, with reconnaissance intensity noticeably higher for those samples. Using the SSM/I intensity estimation algorithm on these 17 images produced an rmse of 19.7 kt when considering the reconnaissance data as ground truth and an rmse of 15.6 kt when using best-track intensities as ground truth. The last column in Table 6 consists of the SSM/I algorithm estimate for each sample. A complete summary of testing results is presented in Table 7. Given the fact that the training data in the SSM/I algorithm

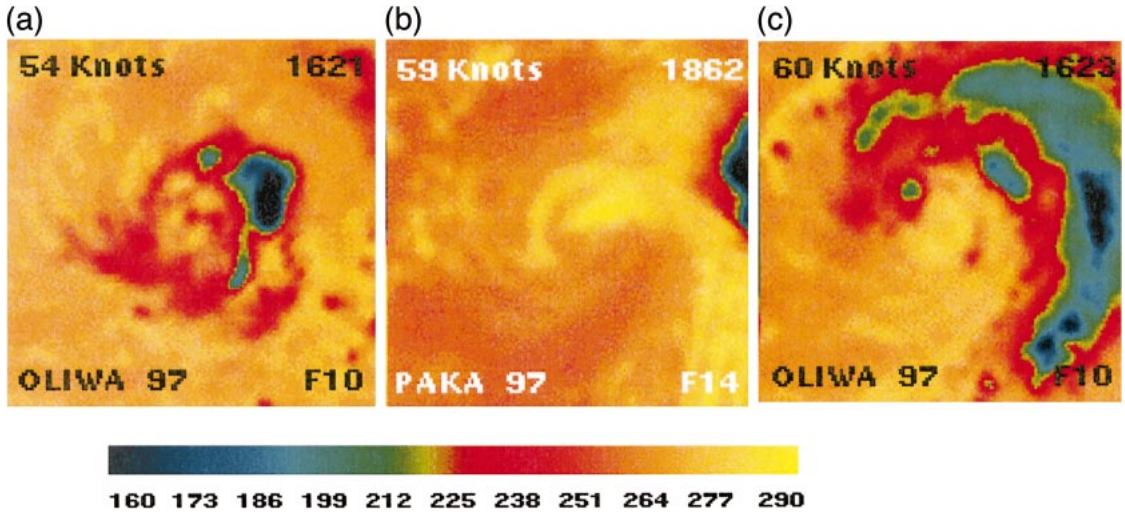

FIG. 9. 85-GHz images of tropical cyclones with similar best-track intensities with much different feature characteristics, contributing to *K*-NN intensity estimation errors: (a) Oliwa, 1126 UTC 7 Sep 1997; (b) Paka, 0004 UTC 21 Dec 1997; and (c) Oliwa, 2351 UTC 7 Sep 1997 [12 h after (a)].

consist of samples with only best-track intensities, it is not too surprising that the rmse when using the best-track intensity is less than the rmse associated with the reconnaissance intensity. The sample size here is relatively small, however, and the overall result may be heavily influenced by the large error during the high-intensity period of Floyd. The reconnaissance rmse for Dennis is much lower (8.5 kt) than the best-track rmse (15.4 kt). An additional nine images from Dennis and Floyd have associated best-track intensities only (no reconnaissance data), and an rmse of 16.6 kt (using best track as the ground truth) was produced for the 26 total samples.

5. Summary and discussion

SSM/I data can provide a more detailed examination of TC structure than can visible or infrared imagery. An attempt is made here to demonstrate the potential of estimating TC intensity automatically by using extracted characteristics from SSM/I imagery. From a total of 101 features, 15 are used as a result of their selection within a feature selection algorithm that uses *K*-nearest-neighbor similarity as the evaluation function. When presented to the intensity estimation algorithm, the selected set of features was shown to be superior to artificially selected subsets composed of general feature types. A set of 942 training samples and the *K*-NN algorithm are applied to 98 independent testing samples. When compared with the best-track intensities of each testing sample, an rmse of 19.8 kt is produced. This result can be viewed as the potential performance of the algorithm on unseen cases.

The snapshot approach can be improved upon when adding past intensity to the set of features. An algorithm is developed to make use of the previous *K*-NN-estimated intensities during the growth and decay of a TC. This strategy could be employed in operations. With 71 samples using this additional feature in the testing set, an rmse of 18.1 kt is produced.

At a given best-track intensity, the variety of possible TC structures together with the surrounding environment (as seen in the 85-GHz and rain-rate imagery) poses a significant problem. This problem becomes evident both in terms of developing and selecting appropriate features and in eliminating large *K*-NN intensity estimation errors. Using historical data for a given TC is an initial step in overcoming this problem.

There are no performance values available for comparing the algorithm's performance on this testing set with other TC intensity estimation algorithms. For an

TABLE 4. *K*-NN testing results (98 total samples) using various feature subsets as the TC representative vector.

Feature set	Rmse (kt)	AAE (kt)	APE (%)
Local	38.7	31.6	40.9
85-GHz textures	31.7	25.6	39.1
Image segmentation	27.2	19.2	33.2
General rain rate	26.9	19.5	31.7
TC specific	25.0	18.1	29.0
General 85 GHz	24.8	19.0	30.6
15 selected features	19.8	16.0	25.5

TABLE 5. *K*-NN testing results using 15 selected features and a 12-h history feature (for 71 of the 98 samples). Twenty-seven samples included in these results used only the 15 selected features.

Tropical cyclone	Rmse (kt)	AAE (kt)	APE (%)
Oliwa (25 samples)	23.1	18.3	31.1
Guillermo (24 samples)	19.6	18.0	27.4
Paka (27 samples)	12.4	9.3	14.1
Babs (22 samples)	15.8	11.8	17.4
Overall	18.1	14.3	22.4

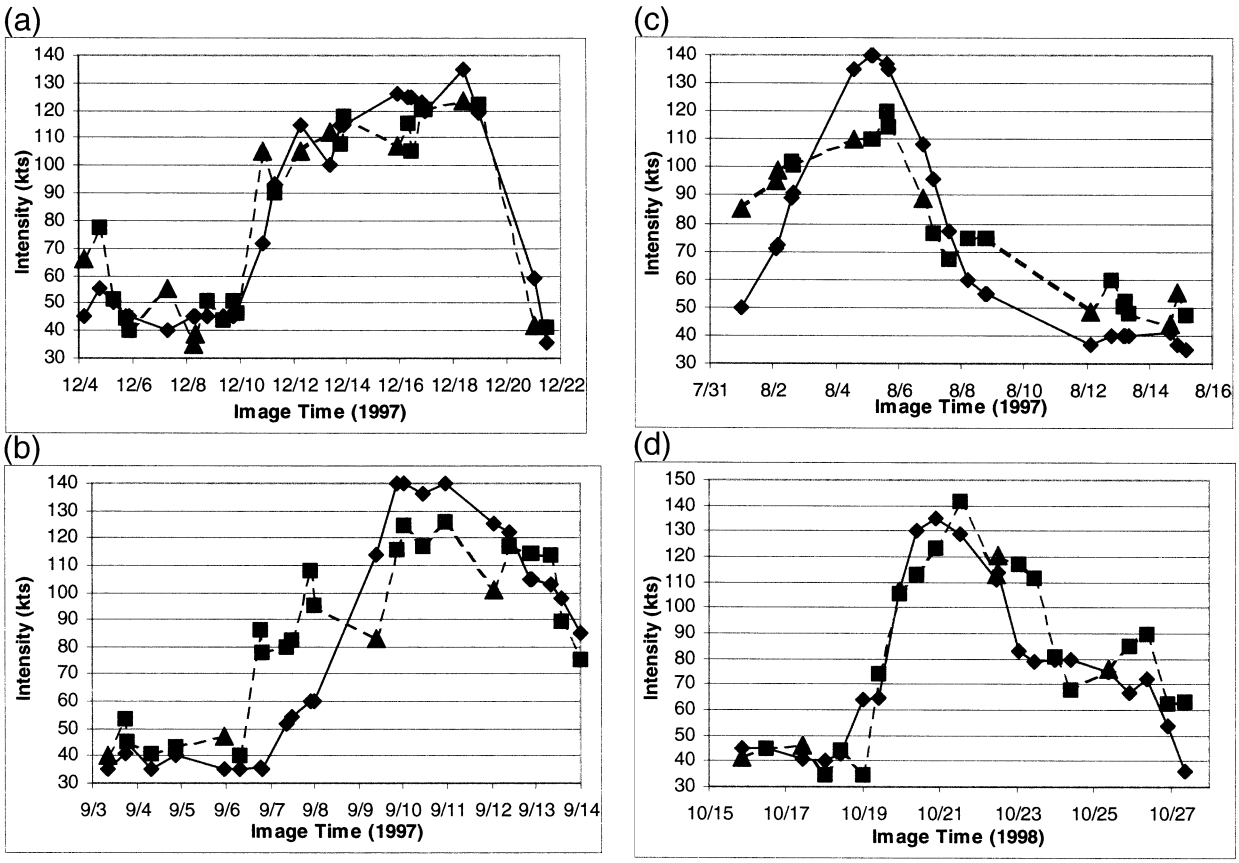

FIG. 10. Time series plots of best-track intensity (solid line, diamond points) and *K*-NN estimated (dashed line, square and triangle points) intensity for the four testing TCs that includes a history feature: (a) Paka, (b) Oliwa, (c) Guillermo, and (d) Babs.

Atlantic-only TC dataset, however, Velden et al. (1998) noted an rmse of 8.3 hPa [9.9 kt—from (2)] for the objective Dvorak technique and an rmse of 10.6 hPa (12.6 kt) for the standard Dvorak method. There obviously is still room for improvement in the TC intensity estimation methodology described in this paper, but it is important to note that the algorithm is dependent upon the best-track intensity accuracy of the training samples. These intensities are estimated to have an error range of ±10–15 kt. Also, confidence in K-NN intensity estimations during operations could be enhanced by adding the 98 testing samples to the training set (>10%

TABLE 6. Best-track, reconnaissance (recon), and SSM/I (*K*-NN) intensity estimates at 17 SSM/I image times for Dennis and Floyd.

TC name and image time–date group	Best-track intensity (kt)	Recon intensity (kt)	SSM/I intensity (kt)
Dennis 0007 UTC 27 Aug 1999	65	62	71
Dennis 1356 UTC 27 Aug 1999	65	67	89
Dennis 2246 UTC 27 Aug 1999	68	74	79
Dennis 1115 UTC 28 Aug 1999	83	85	90
Dennis 2234 UTC 28 Aug 1999	90	90	87
Dennis 2211 UTC 30 Aug 1999	81	87	85
Dennis 1240 UTC 3 Sep 1999	45	68	67
Dennis 1407 UTC 3 Sep 1999	46	68	68
Dennis 0128 UTC 4 Sep 1999	51	71	70
Floyd 0008 UTC 10 Sep 1999	60	59	56
Floyd 1241 UTC 10 Sep 1999	70	66	87
Floyd 1003 UTC 11 Sep 1999	95	98	103
Floyd 2337 UTC 12 Sep 1999	124	13	104
Floyd 0113 UTC 13 Sep 1999	127	135	108
Floyd 1344 UTC 13 Sep 1999	132	146	104
Floyd 2236 UTC 13 Sep 1999	117	143	110
Floyd 1106 UTC 14 Sep 1999	105	137	106

TABLE 7. *K*-NN testing results for Dennis and Floyd.

Tropical cyclone	Rmse (kt)	AAE (kt)	APE (%)
Dennis (9 samples) reconnaissance	8.5	5.4	7.7
Dennis (9 samples) best track	15.4	13.1	23.7
Floyd (8 samples) reconnaissance	27.3	24.0	19.9
Floyd (8 samples) best track	15.7	13.0	12.4
Overall reconnaissance	19.7	14.2	13.5
Overall best track	15.6	13.1	18.4

increase in current training set). Given the wide variety of possible imagery characteristics, increasing the number of unique training samples is one effort that will improve the estimation algorithm. Other efforts that could provide improvement include examining additional feature characteristics from other SSM/I channels or channel combinations, manipulating the *K*-NN computations by using different formulations to arrive at the estimated intensity, and researching other ways to make use of TC history or other ancillary information. In addition, given their unique characteristics, data obtained from other satellite sensors should also help. These sensors include Tropical Rainfall Measuring Mission, SSM/I Sounder, Advanced Microwave Sounding Unit-A/B, Advanced Microwave Scanning Radiometer, and Global Precipitation Mission. Last, this TC intensity estimation method could be merged in some fashion with ODT to take advantage of each technique to produce a more accurate intensity estimate. The potential of accurately estimating TC intensity from SSM/I data has been demonstrated, and, with further enhancements, a useful tool for estimating TC intensity will be available.

Acknowledgments. The support of the sponsor, the Office of Naval Research under Program Element 0602435N, is gratefully acknowledged. Much appreciation is extended to Marla Heleveston (Analysis and Technology, Inc.) for processing and supplying the digital imagery and to Jeff Hawkins (Naval Research Laboratory, Monterey) for recentering all imagery, his insightful comments on tropical cyclones, and his thoughtful suggestions during the course of this research. The authors also thank Roger Edson and Chip Guard of JTWC for relating their vast experiences and providing suggestions during this research effort.

REFERENCES

Aha, D. W., and R. L. Bankert, 1995: A comparative evaluation of sequential feature selection algorithms. *Learning from Data: Artificial Intelligence and Statistics V*, D. Fisher and H.-J. Lenz, Eds., Springer-Verlag, 199–206.

Bankert, R. L., 1994: Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *J. Appl. Meteor.*, **33,** 909–918.

——, and D. W. Aha, 1996: Improvement to a neural network cloud classifier. *J. Appl. Meteor.*, **35,** 2036–2039.

Duda, R. O., and P. E. Hart, 1973: *Pattern Classification and Scene Analysis.* John Wiley and Sons, 482 pp.

Dvorak, V. F., 1975: Tropical cyclone intensity analysis and forecasting from satellite imagery. *Mon. Wea. Rev.*, **103,** 909–918.

——, 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, 47 pp.

Ferraro, R. R., 1997: Special Sensor Microwave Imager derived global rainfall estimates for climatological applications. *J. Geophys. Res.*, **102,** 16 715–16 735.

Hawkins, J. D., D. A. May, J. Sandidge, R. Holyer, and M. J. Helveston, 1998: SSM/I-based tropical cyclone structural observations. Preprints, *9th Conf. on Satellite Meteorology and Oceanography*, Vol. 1, Paris, France, Amer. Meteor. Soc., 230–233.

——, T. F. Lee, J. Turk, C. Sampson, J. Kent, and K. Richardson, 2001: Real-time Internet distribution of satellite products for tropical cyclone reconnaissance. *Bull. Amer. Meteor. Soc.*, **82,** 567–578.

Holliday, C. R., 1969: On the maximum sustained winds occurring in Atlantic hurricanes. Weather Bureau Southern Region Tech. Memo. WBTH-SR-45, 6 pp. [NTIS PB 184609.]

May, D. A., J. Sandidge, R. Holyer, and J. D. Hawkins, 1997: SSM/I derived tropical cyclone intensities. Preprints, *22d Conf. on Hurricanes and Tropical Meteorology*, Fort Collins, CO, Amer. Meteor. Soc., 27–28.

Olander, T. L., and C. S. Velden, 1999: UW-CIMSS objective Dvorak technique (ODT). Preprints, *23d Conf. on Hurricanes and Tropical Meteorology*, Dallas, TX, Amer. Meteor. Soc., 435–436.

Poe, G., 1990: Optimum interpolation of imaging microwave radiometer data. *IEEE Trans. Geosci. Remote Sens.*, **GE-28,** 800–810.

Velden, C. S., T. L. Olander, and R. M. Zehr, 1998: Development of an objective scheme to estimate tropical cyclone intensity from digital geostationary satellite infrared imagery. *Wea. Forecasting*, **13,** 172–186.