Objective Tropical Cyclone Intensity Estimation Using Analogs of Spatial Features in Satellite Data

GHOLAMREZA FETANAT AND ABDOLLAH HOMAIFAR
North Carolina Agricultural and Technical State University, Greensboro, North Carolina

KENNETH R. KNAPP
NOAA/National Climatic Data Center, Asheville, North Carolina

(Manuscript received 31 December 2012, in final form 19 June 2013)

ABSTRACT

An objective method for estimating tropical cyclone (TC) intensity using historical hurricane satellite data (HURSAT) is developed and tested. This new method, referred to as feature analogs in satellite imagery (FASI), requires a TC's center location to extract azimuthal brightness temperature (BT) profiles from current imagery as well as BT profiles from imagery 6, 12, and 24 h prior. Instead of using regression techniques, the estimated TC intensity is determined from the 10 closest analogs to this TC based on the BT profiles using a $k$-nearest-neighbor algorithm. The FASI technique was trained and validated using intensity data from aircraft reconnaissance in the North Atlantic Ocean, where the data were restricted to include storms that are over water and south of 45°N. This subset comprised 2016 observations from 165 storms during 1988–2006. Several tests were implemented to statistically justify the FASI algorithm using $n$-fold cross validation. The resulting average mean absolute intensity error was 10.9 kt (50% of estimates are within 10 kt, 1 kt = 0.51 m s$^{-1}$) or 8.4 mb (50% of estimates are within 8 mb); its accuracy is on par with other objective techniques. This approach has the potential to provide global TC intensity estimates that could augment intensity estimates made by other objective techniques.

1. Introduction

Tropical cyclones (TCs) are a significant threat to life and property. Developing and improving objective techniques to estimate a TC’s intensity remain a challenge. The Dvorak technique (DT) is the state-of-the-art method that has been used over three decades for estimating the intensity of TCs (see Velden et al. 2006 for a complete review). The DT subjectively estimates TC intensity based on visible and infrared satellite images (Dvorak 1972, 1975, 1984). Based on the success of the DT, the objective Dvorak technique (ODT) was derived, which used computer-based analyses to provide an objective estimate of TC intensity (Velden and Olander 1998). To overcome the limitations of the ODT, such as the manual selection of the storm center or the inability to operate on weak storms, the advanced objective Dvorak technique (AODT) was developed. The most recent version of ODT is the advanced Dvorak technique (ADT; Olander and Velden 2007). Unlike the ODT and AODT, whose focuses were on mimicking the subjective technique, the ADT concentrates on extending the method beyond the original application and constraints (Velden et al. 2006).

In light of the historical use of the various Dvorak techniques, investigation of other approaches to help increase the accuracy and precision of automated estimation of TCs’ intensities is needed. A recently developed method, called the deviation angle variance (DAV) technique (Piñeros et al. 2008, 2011), uses the gradient of the brightness temperature (BT) field to determine the level of symmetry of the TC’s cloud structure, which correlates with the intensity of the TC. A modification of the DAV technique in Ritchie et al. (2012) uses the National Hurricane Center’s best-track database to constrain the technique.

The algorithm described herein estimates TC intensity from the intensities of TCs that are analogous to it. Analogous storms are determined from BT profiles at the current time as well as during recent development.
This new technique—the feature analogs in satellite imagery (FASI) technique—was inspired by the availability of satellite imagery for TCs from the historical hurricane satellite (HURSAT) dataset. Our goal was to develop a new objective technique for estimating the intensity of TCs using historical satellite imagery centered on TCs.

The remainder of the paper is organized as follows: Section 2 describes the satellite data used for this study followed by a description of the FASI technique in section 3. Section 4 discusses the validation process, followed by a discussion of the results within the context of other objective algorithms in section 5. Section 6 provides conclusions and a discussion of potential future work.

2. Satellite data

Hurricane satellite data (HURSAT-B1, version 05), described in Knapp and Kossin (2007), provides infrared window imagery (\( \sim 11 \mu m \)) for global TCs from 1978 to 2009. This dataset covers TCs in the northern Atlantic and the eastern and western Pacific, the Southern Hemisphere, and the Indian Ocean at 8 km with 3-hourly resolution. HURSAT-B1 data files contain storm-centered images in network common data form (NetCDF) format, in which each file provides a snapshot of the storm from one of the international geostationary weather satellites.

Multiple files are possible when a storm is viewed by two satellites at the same time. Only the satellite images with the best view (i.e., the smallest view zenith angle) were considered for this study. Following Kossin et al. (2007), the training data were restricted to include imagery when the TC is over water and is south of 45°N. We considered the most accurate intensity estimates to be those within 12 h of an aircraft reconnaissance. This subset comprised 2016 measurements in 165 storms from 1988 to 2006. A stricter temporal threshold (e.g., within 6 h of reconnaissance) might mean more accurate intensity estimates, but this temporal assumption significantly reduces the available training samples to a point that the estimation error would be too large.

From HURSAT imagery, we derived the mean (\( M \)) and standard deviation (SD) of BT (K) for 14 azimuthal rings at 10-km intervals from the storm center (5, 15, 25, . . . , 695 km).

3. The FASI technique

Data mining has attracted a great deal of attention in the information industry due to the availability of large amounts of data and the urgent need to extract useful knowledge from such data. This study applied the techniques that are routinely used in data mining. The process for estimating intensity can be expressed as two functions applied in time and space: \( \text{INT} = f[g(x,y), t] \). The intensity (INT) is estimated from spatial analysis of satellite image \( g(x,y) \) that is constrained in time \( t \) by some function \( f \). For example, the Dvorak technique is the estimate of a final T number based on spatial analysis (e.g., the \( g \) function), which is then constrained in time and by other rules (e.g., the \( f \) function). This study focused solely on the analysis of satellite imagery (\( g \)). Further constraints on time are left to future endeavors.

The cloud patterns of TCs evolve through recognizable stages as the intensity of the cyclone changes (Dvorak 1984). As the BT changes, different cloud patterns form around the center of the storm. We developed our technique to discover these BT patterns in the rings around the center of the storms and to compare these patterns with the existing historical BT patterns. We hypothesized that similar patterns would correspond with similar intensities. Because patterns in each ring may vary considerably during the evolution of the storm, we hypothesize that the \( M \) and SD of each ring’s BT would have less variability between TCs with similar intensity. The FASI technique used the center location estimate (provided by HURSAT-B1 data) and imagery from the current time along with imagery from 6, 12, and 24 h before the current time to estimate the intensity.
a. FASI algorithm description

In data mining, the $k$-nearest-neighbor algorithm is an algorithm for classifying objects based on other objects from some training set that are the closest in the feature space. In this problem, the intensity of each snapshot of a TC is determined by averaging the intensity of its $k$ nearest neighbors (NNs). The feature vector when comparing a query image with a candidate image has 112 dimensions, which is derived from four images (the current time and three prior images), two values per azimuthal ring ($M$ and SD) for 14 azimuthal rings ($4 \times 2 \times 14 = 112$). Nearest neighbors to the query are determined using a Euclidean distance metric, which is a commonly used distance metric. The Euclidian distance between a query image, $X = \{x_1, x_2, \ldots, x_d\}$, and a sample image, $Y = \{y_1, y_2, \ldots, y_d\}$, is defined as

$$D(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_d - y_d)^2}. \quad (1)$$
where $d = 112$. The nearest neighbor is the sample image from the training set that minimizes $D$.

The FASI technique for estimating intensity is illustrated in Fig. 1. First, the data in the training set are organized according to selected features. Second, for each query (i.e., a TC image with unknown intensity) the query feature vector was constructed. Third, training dataset entries were sorted based on the Euclidian distance between the query entry and the entries in the training set. Fourth, we applied the $k$-nearest-neighbor algorithm to choose $k$ entries with the shortest Euclidian distance. These $k$ entries were the $k$ “nearest neighbors” that are closest to the unknown query entry. Last, we estimate the intensity of the query entry as the average intensity of the 10 NNs. The estimates of intensity can be expressed in terms of maximum sustained wind speed (MSW) or minimum sea level pressure (MSLP). Hence, the value of MSLP was estimated simultaneously without using an empirical formula to calculate MSLP from MSW.

Since the FASI intensity algorithm returns a continuous intensity rather than a categorical label (e.g., category 2), it is difficult to conclude exactly whether the estimated intensity is accurate. To overcome this...
difficulty, the accuracy of FASI is measured in terms of the difference between the estimated and the best-track intensities using mean absolute error (MAE) and root-mean-square error (RMSE) (following Han and Kamber 2006).

b. FASI technique derivation

A sequential forward selection algorithm (Koutroumbas 1999) was used to find the optimal number of azimuthal rings. Each ring was sequentially added to an empty candidate set until the addition of further rings did not significantly decrease the MAE or RMSE. Fourteen rings around the center of the storm were found to minimize the MAE and RMSE (see section 4a for more details). These 14 rings were consecutive and started from the center of the storm (i.e., centered on 5–135 km).

The algorithm’s performance is affected by how many NNs ($k$) are selected. If $k$ was small, then noise could affect the algorithm’s performance. If $k$ was too large, then the set of NN images might belong to widely different storm types. By varying the value of $k$ from 1 to 200 for several validation processes, we found that the optimal value is 10, which minimized both MAE and RMSE.

Deciding how to best combine intensity estimates from the 10 NNs might include a simple average, a weighted average, the median, NN, and mode. These were all tested; the simple average of the 10 NNs minimized the error (see section 4a for more details).

Two options for NNs are to use a certain number of the closest neighbors (as decided above, 10) or using all storm images that were within a certain distance. As shown in Fig. 2, by considering a distance threshold, only NNs within a certain radial distance of the query can be considered for estimating intensity. The following is an analysis comparing the use of a threshold on the Euclidean distance rather than just the 10 NNs (whatever their distance may be).
We normalized the feature vectors by their maximum possible values for each feature. Again, the query vector $\mathbf{X} = \{x_1, x_2, \ldots, x_d\}$ and training sample vector $\mathbf{Y} = \{y_1, y_2, \ldots, y_d\}$ with dimension of $d$ (112) can be normalized using

$$\mathbf{NX} = \left\{ \frac{x_1}{c_1}, \frac{x_2}{c_2}, \ldots, \frac{x_d}{c_d} \right\} \quad \text{and} \quad \mathbf{NY} = \left\{ \frac{y_1}{c_1}, \frac{y_2}{c_2}, \ldots, \frac{y_d}{c_d} \right\},$$

(3)

where $\mathbf{NX}$ and $\mathbf{NY}$ are the normalized feature vectors and $c_1, c_2, \ldots, c_d$ are the maximum values of each feature. The Euclidian distance between these two normalized points based on Eq. (1) is

$$D(\mathbf{NX}, \mathbf{NY}) = \sqrt{\left(\frac{x_1}{c_1} - \frac{y_1}{c_1}\right)^2 + \left(\frac{x_2}{c_2} - \frac{y_2}{c_2}\right)^2 + \cdots + \left(\frac{x_d}{c_d} - \frac{y_d}{c_d}\right)^2}.$$

(4)

c. An example from Hurricane Katrina

An example of the FASI technique is provided for Hurricane Katrina. The sample task is to estimate the intensity of the storm at 0000 UTC 28 August 2005 (Fig. 3). Hurricane Katrina’s feature vector is provided in Figs. 4a and 4b, which shows the $M$ and SD of the BT azimuthal profiles of the present time and the priors (6, 12, and 24 h prior). Using the method described, the NN
in the training set to Hurricane Katrina at that date and time was Hurricane Bonnie at 0000 UTC 23 August 1998 (Fig. 5). The corresponding feature vector for Bonnie is shown in Figs. 4c and 4d. This set of azimuthal profiles is the closest in Euclidean space to the feature vector of Katrina. Bonnie’s corresponding intensity of 90 kt (1 kt = 0.51 m s\(^{-1}\)) is just 10 kt from the reported intensity of Katrina (100 kt). Figure 6 shows the other nine NNs. Then, averaging the intensity of the 10 NNs results in a wind speed of 96.5 kt, only 3.5 kt from the reported intensity.

This example is not necessarily characteristic of all retrievals, but helps us to understand how the algorithm functions. The following analysis discusses the performance of the technique across the entirety of the validation dataset.

4. Results

Several tests were done using \(n\)-fold cross validation for statistical justification of the FASI technique (often referred as \(k\)-fold cross validation but referred to here as \(n\)-fold so as not to cause confusion with \(k\) nearest neighbors). The initial data were partitioned into \(n\) mutually exclusive subsets or folds, \(S_1, S_2, \ldots, S_n\), where \(n\) is the number of storms. Since 165 storms with 2016 samples from 1988 to 2006 were used as the training data, then \(n = 165\). The entire validation was performed \(n\) times, each time leaving an entire storm out of the training dataset. For example, for a particular iteration \(j\), none of the imagery of that storm was included in the training set. All images from storm \(j\) were then used as query images to obtain an estimated intensity from the rest of the training set. This was repeated for each storm. Thus, no imagery from the query storm was used in the training set. All images from storm \(j\) were then used as query images to obtain an estimated intensity from the rest of the training set. This was repeated for each storm. Thus, no imagery from the query storm was used in the training set. Note that each storm image was used the same number of times as the training data and once for validation. The error estimate was computed as the total loss from the \(n\) iterations divided by the total number of initial subsets.

The distribution of the error for MSW estimation is shown in Fig. 7a. Error was defined as the absolute
difference between the FASI-estimated and best-track intensities for best-track observations within 12 h of a reconnaissance flight. Brown and Franklin (2004) compared the Dvorak intensity estimates (for 1977–2003) with best-track intensities that were concurrent with aircraft reconnaissance; they found that 90%, 75%, and 50% of their mean absolute errors were less than 18, 12, and 5 kt, respectively. Corresponding errors in the FASI technique were less than 18, 14, and 10 kt, respectively. The averaged MAE is 10.9 kt, the RMSE is 12.7 kt, and the bias is −1.1 kt for the 2016 samples of the storms. Here, the term bias means the average differences between the estimated intensity and the best-track intensity. The validation results for the 165 storms are shown in Figs. 7b–d.

Biases and errors can be shown as a function of intensity to investigate the under- or overestimates and variations of errors in different intensities. One of the limitations of the training data was that the data were not distributed uniformly. Figure 8a shows the frequency of different intensities in the training set. As shown, the numbers of training snapshots in intense (more than 130 kt) and weak (less than 25 kt) systems were very small. Thus, more errors can be expected in those ranges since the numbers of similar instances were

![Fig. 10. Frequency of various MSLPs in the training dataset.](image)

![Fig. 11. Time series of best-track intensity and that estimated from the FASI technique for (a) Hurricane Allison (1995, MAE = 3.9 kt), (b) Hurricane Erika (1997, MAE = 7 kt), (c) Hurricane Floyd (1999, MAE = 7.5 kt), and (d) Hurricane Katrina (2005, MAE = 7.9 kt).](image)
too few. The results of the validation were compiled in overlapping bins following Knaff et al. (2010). For instance, the first and second bins had intensity ranges of 20–35 and 25–45 kt, respectively. The last bin was extended from 127 to 170 kt as one bin. Figures 8b–d show the bias, MAE, and RMSE associated with the FASI technique as a function of intensity. The results showed that the FASI technique generated overestimates in cases of low intensities (less than 40 kt) and underestimates in cases of higher intensities (more than 130 kt).

The distribution of the error for estimating MSLP is shown in Fig. 9. It shows that 50% of the estimates have an MAE of less than 8 mb, 75% are within 10.7 mb, 90% are within 14.6 mb, and the maximum absolute error is 23.9 mb. Moreover, the averaged MAE, RMSE, and bias were 8.4, 9.8, and −1.1 mb, respectively, for all 165 storms. Detailed validation results for the 165 storms are shown in Figs. 9b–d. Again, the nonuniform distribution of MSLP values was a limitation of the training dataset. Figure 10 shows the frequency of different MSLPs in the training dataset. As shown, the number of training snapshots for intense systems (less than 940 mb) was very small. Thus, more errors were expected in that range since the numbers of analogous instances were too few. Conversely, the algorithm can likely be improved by expanding the training dataset to more instances of intense storms.

We present, in Fig. 11, examples of the FASI estimation of intensity for Hurricanes Allison (1995), Erika (1997), Floyd (1999), and Katrina (2005). These time series show that the estimated FASI intensity follows the best-track intensity values closely. These plots are comparable with results of other objective techniques (which is discussed more in section 5), as in Fig. 12, which shows values of the estimated intensity for Hurricane Katrina (2005) from several techniques including ADT and the subjective Dvorak technique. Further analysis of the technique’s accuracy and precision are needed in different global basins and storm types.

**Sensitivity analysis of the predictors and the selected parameters**

This section discusses the impacts of the various components of the technique on the overall algorithm accuracy. The following are discussed: number of prior

<table>
<thead>
<tr>
<th>Predictors</th>
<th>RMSE (kt)</th>
<th>MAE (kt)</th>
<th>Bias (kt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current image</td>
<td>16</td>
<td>14</td>
<td>−3.1</td>
</tr>
<tr>
<td>Current image</td>
<td>15</td>
<td>13</td>
<td>−2.4</td>
</tr>
<tr>
<td>6-h-prior image</td>
<td>14</td>
<td>12</td>
<td>−1.7</td>
</tr>
<tr>
<td>12-h-prior image</td>
<td>13</td>
<td>11</td>
<td>−1.1</td>
</tr>
</tbody>
</table>

**Table 1. FASI technique error statistics for various predictors sets.**
images, number of similar NNs, how intensities of the NNs were combined, and how NNs were selected. The sequential forward selection algorithm was used to determine the sensitivity of intensity estimation by performing the same $n$-fold cross validation using different sets of the predictors. The procedure initially considered only the current image of the storm, for which the values of RMSE and MAE are indicated in Table 1. The procedure was repeated by adding the prior 6-h image information and so on. The results, shown in Table 1, show that maximum accuracy is achieved when all of the predictors are included.

Fig. 13. Variations in the average RMSE and MAE vs $k$ (number of NNs) and number of rings.
The optimal number of NN (or similar) images \( k \) and the number of azimuthal rings used to estimate a TC's intensity were determined by minimizing MAE and RMSE. This number of NNs was varied from 1 to 200 and the number of rings likewise varied from 1 to 70. The distributions of MAE and RMSE are shown in Fig. 13. Based on these variations, optimal values for parameters were selected as \( k = 10 \) and the 14 innermost rings (5–135 km) from the center of the storm.

As described above, the average intensity of the NNs is used to estimate intensity in this technique. Other approaches for estimating the intensity of TCs from the 10 NNs were also investigated: weighted average, median, mode, and NN were each evaluated. The corresponding average RMSE results using the \( n \)-fold cross validations are provided in Table 2. The results show that the lowest RMSEs were found by averaging the intensity of the NNs. However, the RMSEs of intensity estimates based on a weighted average and median were very close to the RMSEs of the average value.

By considering a distance threshold, only NNs within a certain distance of the query feature were considered for estimating intensity. Consequently the distance threshold provided a quantitative value of how far the NNs were from the query. Figure 14 shows how the MAE and RMSE in \( n \)-fold cross validation change when the distance threshold vary from 1% to 20% of the maximum possible distance. Figure 14 shows that the optimal value for the decision threshold appears to be 13% of the maximum possible distance. Figure 15 shows the distribution of the error for \( n \)-fold cross validation using the NNs within 13% of the maximum possible distance. The averaged MAE, RMSE, and bias of this decision technique were 11, 12.9, and 1.3 kt, respectively, for all 165 storms. Table 3 shows the percentage MAE results derived by averaging 10 NNs' intensities and averaging NNs' intensities within a radius of 13% of the maximum possible distance. The results showed that the accuracy of both techniques is similar.

5. Comparison with existing objective techniques

The following is a short comparison of statistics resulting from the validation of existing algorithms for estimating TC intensity objectively. Since the validation sets used for development and verification of the FASI technique are different than those used to validate other objective techniques, one can only draw broad conclusions on whether the FASI technique can contribute new information to the objective analysis of TC intensity.

The original and modified objective DAV techniques described in Piñeros et al. (2008, 2011) and Ritchie et al. (2012) are considered. The DAV technique uses the gradient of the BT field to determine the level of symmetry of the TC's cloud structure. The departure of the cloud structure from axisymmetry is correlated with the intensity of the storm. Table 4 shows the results obtained from using DAV techniques (from Table 2 of Ritchie et al. 2012) employing samples from the period 2004–10 compared to the results of the FASI technique. The results show that the FASI technique and DAV have similar RMSE values.

A technique introduced by Kossin et al. (2007) used principal components of azimuthal averages and storm latitude to estimate TC intensity, developed using 1940

<table>
<thead>
<tr>
<th>Approach</th>
<th>RMSE (kt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Averaging</td>
<td>12.69</td>
</tr>
<tr>
<td>Weighted averaging</td>
<td>12.73</td>
</tr>
<tr>
<td>Median</td>
<td>12.73</td>
</tr>
<tr>
<td>Mode</td>
<td>14.66</td>
</tr>
<tr>
<td>NN</td>
<td>16.70</td>
</tr>
</tbody>
</table>
observations from 137 storms between 1989 and 2004 as training and testing data. Data within 3 h of aircraft measured intensities were used for training data. The algorithm introduced by Kossin et al. (2007) was not geared toward real-time analysis but for historical re-analysis. Nonetheless, many are familiar with its application so it is included here. Table 5 compares the MAE and RMSE between the techniques. Again, while the validation period is different, the FASI technique does appear to be on par with error estimates of Kossin et al. (2007).

The FASI technique can also be compared with the ADT and DT results (Olander and Velden 2007). The operational DT intensity estimates were obtained from three operational forecast centers (OFCs): the Satellite Analysis Branch (SAB) of National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service (NOAA/NESDIS) in Washington, D.C.; the Tropical Analysis and Forecast Branch (TAFB) of NOAA/National Centers for Environmental Prediction (NCEP) at the Tropical Prediction Center (TPC) in Miami, Florida; and the Air Force Weather Agency (AFWA) at Offutt Air Force Base in Omaha, Nebraska. Most of the cases have multiple OFC estimates. In these cases, the resulting OFC estimates used in the comparisons were averaged and represent a consensus of the available DT estimates (Table 7 of Olander and Velden 2007). Table 6 compares the ADT and OFC results (using 2093 samples, from 1996 to 2005) with the FASI technique. It indicates that the FASI technique’s error statistics are in the range of the ADT and OFC error statistics, based on RMSE.

The FASI technique uses a different approach to estimate the intensity of the TCS compared to other subjective and objective techniques. This different approach makes intensity estimates from the FASI technique statistically independent (different) than other techniques. Using the FASI technique with other independent intensity estimation techniques will reduce the mean error of an equally weighted consensus. The importance of independence in an equally weighted consensus is presented in appendix B of Sampson et al. (2008).

One of the limitations of the training dataset was that the data were not distributed uniformly. There were very few training snapshots for extremely intense and weak systems. Thus, more errors were expected in those ranges since the numbers of nearby analogs were small. The results show that the FASI technique is accurate in estimating the intensities for tropical depressions, tropical storms, and hurricane categories 1–3, as shown in Table 7. In comparison the DAV technique accurately estimates the intensities for tropical storms and hurricane categories 1–3 (Ritchie et al. 2012). However, the subjective Dvorak technique is most skillful (has the least error) in an intensity range from 90 to 125 kt, which is approximately for hurricane categories 2–4 (Knaff et al. 2010).

6. Conclusions and future work

We have described a new technique for estimating the intensity of TCS using the feature analogs in satellite

<table>
<thead>
<tr>
<th>Approach</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed number (kt)</td>
<td>10.0</td>
<td>13.8</td>
<td>18.0</td>
</tr>
<tr>
<td>Distance threshold (kt)</td>
<td>10.2</td>
<td>13.4</td>
<td>18.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technique</th>
<th>MAE (kt)</th>
<th>RMSE (kt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kossin et al. (2007)</td>
<td>13.2</td>
<td>16.7</td>
</tr>
<tr>
<td>FASI</td>
<td>10.9</td>
<td>12.7</td>
</tr>
</tbody>
</table>

FIG. 15. Cumulative cross-validation error results using NNs intensities within 13% of the maximum possible distance.
imagery (FASI) technique. This technique uses a TC’s center location, and the $M$ and SD of satellite BTs in 14 azimuthal rings. This information from a current image, as well as that from the preceding 6, 12, and 24 h, are used as predictors to estimate the TC intensity. Instead of regression techniques, the intensities of the 10 closest analogs (determined using a $k$-nearest-neighbor algorithm) were averaged to estimate the intensity.

Several tests were done to statistically justify the design of the algorithm using $n$-fold cross validation. The resulting averaged MAE was 10.9 kt (50% of points are within 10 kt) or 8.4 mb (50% of points are within 8 mb). By considering a distance threshold of 13% of the maximum possible distance, only NNs within this radius of the query were considered for decision making. The results showed that the accuracies for averaging the intensities of the 10 NNs and averaging the intensities of the NNs within a radius of 13% of the maximum possible distance were very similar. It was found that the numbers of very intense and very weak storms in training data were too few to find accurate analogs for similar storms. Therefore, increasing the number of analog intense storms is a goal of future work. Also, an important step would be to perform a homogeneous comparison of FASI with other techniques using the same validation set.

In our approach, we considered only satellite imagery as predictors, but other analogs could be included. While adding more predictors to increase the accuracy of the FASI technique may seem reasonable at first glance, it may also limit the size of search space. The training set would contain fewer storms that are true analogs for each query entry. For example, few snapshots of category 5 hurricanes exist in the current training set; if sea surface temperatures (SSTs) were included as new predictors, there would be even fewer category 5 storms for a given SST. Likewise, one could envision including data from higher temporal resolution geostationary imagery. HURSAT is limited to 3-hourly data; some information might be gained by including the other times as predictors. Nonetheless, adding predictors needs to be done carefully so as not to dilute the number of nearby analogs.

Further improvements of FASI are also possible. Given the global coverage of HURSAT, it is possible that a training set could be constructed for the western North Pacific during the period of aircraft reconnaissance (1980–87). Such an approach would be unique to FASI since other objective algorithms mentioned here have been developed over the North Atlantic. Doing so would also allow a climatological application: a global estimate of TC activity based on FASI that could be validated in two separate basins.

Temporal constraints on changes in intensity could also be applied to the algorithm in the future. The FASI algorithm as described here has no dependence upon time except that which is implied by the use of features from previous satellite images (6, 12, and 24 h prior). Other objective algorithms use temporal constraints on intensity changes (e.g., the ADT and subjective DT). Including similar constraints in FASI would likely decrease some of the random error found herein.

Another application of the FASI technique would be to include it in a satellite consensus (SatCon) technique, which incorporates strengths and weaknesses of several objective estimates to compute an intensity estimate that is better than the any of its individual parts (Herndon and Velden 2008). The FASI technique may well provide independent information for such an approach.

Finally, the FASI technique can be combined with an algorithm to detect the TC center of circulation (e.g., Wimmers and Velden 2010) for full automation of the technique. But again, this would require some further analysis of the center location algorithm’s performance using HURSAT data. For example, the DAV technique was originally allowed to derive its own TC center, but

### Table 6. Comparison of algorithm RMSE for MSLP using ADT, DT (OFCs), and the FASI technique.

<table>
<thead>
<tr>
<th>Technique</th>
<th>RMSE (mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>12.5</td>
</tr>
<tr>
<td>DT (OFCs)</td>
<td>9.9</td>
</tr>
<tr>
<td>FASI</td>
<td>9.8</td>
</tr>
</tbody>
</table>

### Table 7. FASI technique errors (MAE and RMSE) using $n$-fold cross validation stratified by TC intensity. Values with and without an asterisk (*) presented the results using the distance threshold and the fixed number (10) of NNs, respectively.

<table>
<thead>
<tr>
<th>Bin</th>
<th>No. of samples</th>
<th>MAE (kt)</th>
<th>MAE* (kt)</th>
<th>RMSE (kt)</th>
<th>RMSE* (kt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropical depression ($\leq 34$ kt)</td>
<td>132</td>
<td>14.1</td>
<td>14.2</td>
<td>16.4</td>
<td>16.5</td>
</tr>
<tr>
<td>Tropical storms ($\geq 35$ kt and $\leq 63$ kt)</td>
<td>861</td>
<td>9.2</td>
<td>9.2</td>
<td>11.9</td>
<td>12.3</td>
</tr>
<tr>
<td>Hurricane category 1 ($\geq 64$ kt and $\leq 82$ kt)</td>
<td>448</td>
<td>11.2</td>
<td>10.6</td>
<td>13.8</td>
<td>12.9</td>
</tr>
<tr>
<td>Hurricane category 2 ($\geq 83$ kt and $\leq 95$ kt)</td>
<td>190</td>
<td>12.1</td>
<td>10.3</td>
<td>15.8</td>
<td>14.1</td>
</tr>
<tr>
<td>Hurricane category 3 ($\geq 96$ kt and $\leq 112$ kt)</td>
<td>164</td>
<td>13.8</td>
<td>13.7</td>
<td>17.5</td>
<td>16.8</td>
</tr>
<tr>
<td>Hurricane category 4 ($\geq 113$ kt and $\leq 136$ kt)</td>
<td>183</td>
<td>14.7</td>
<td>17.1</td>
<td>19.7</td>
<td>21.5</td>
</tr>
<tr>
<td>Hurricane category 5 ($\geq 137$ kt)</td>
<td>38</td>
<td>27.6</td>
<td>33.6</td>
<td>30.5</td>
<td>35.4</td>
</tr>
</tbody>
</table>
the modified algorithm uses best-track TC centers for more accurate intensity estimation.

In short, the FASI technique provides a unique approach to estimating the intensity of tropical cyclones. The error statistics of the technique are on par with those of other algorithms and more work remains to make the FASI technique operational.

Acknowledgments. This work was supported by the Expeditions in Computing Program of the National Science Foundation under Award CCF-1029731. Special thanks are given to two anonymous reviewers for their helpful comments, as well as Jim Kossin and Carl Schreck who provided valuable insight during the development of this work. The opinions expressed in this paper are those of the authors. They do not necessarily reflect the official views or policies of NSF, NOAA, Department of Commerce, or the U.S. government.

REFERENCES


