P1.20 TROPICAL CYCLONE INTENSITY ESTIMATION USING TEMPORAL ANALYSIS AND SPATIAL FEATURES IN SATELLITE DATA

Gholamreza Fetanat, Abdollah Homaifar*
North Carolina A&T State University, Greensboro, North Carolina,
Kenneth Knapp
NOAA's National Climatic Data Center, Ashville, North Carolina

1. Introduction

Tropical cyclones (TCs) are a significant threat to populations. Accurate estimation and prediction of TC's intensity will save lives and property. The Dvorak TC intensity estimation technique has been the primary method applied for more than 30 years (Velden et al. 2006a) in the world. The Dvorak technique (DT) subjectively estimate TC intensity based on visible and infrared satellite images (Dvorak 1975, 1984). Improvement of the original DT has evolved into the objective Dvorak technique (ODT). The main goal of the ODT is to use computer based objective method to limit the subjectivity for human interpretation of DT (Velden and Olander 1998). To overcome the limitations of the ODT such as manual selection of the storm center or the inability to operate on weak storms, advanced objective Dvorak technique (AODT) was developed. The most recent version of ODT is the Advanced Dvorak Technique (ADT). Unlike the ODT and AODT whose focuses were to mimic the subjective technique, the ADT concentrates on extending the method beyond the original application and constraints (Olander and Velden 2007).

In spite of wide usage of the DT for TC analysis it has some limitations. First, this method infers intensity from cloud features and patterns instead of using direct measurement of wind. This leads to an error from natural variability between the observed wind speed and remotely sensed cloud patterns. Second, the DT is a group of practical rules and has not theoretical foundation. Third, this method does not use the valuable historical data mainly because of great challenges on computing and human resources.

This research is inspired by the availability of tropical cyclone satellite imagery. Developing a new automated TC estimation technique using Hurricane Satellite (HURSAT) data is still a challenge. We hypothesize that discovering unknown regularities and abnormalities that may exist in the large group of past observations could help human experts interpret TC intensity changes from various points of view.

Our goal is to provide a data mining tool that increases the ability of human experts to analyze huge amount of historical data for TC intensity estimation. This line of research discovers a set of facts and guidelines with the statistical justification.

The proposed intensity estimation algorithm has two parts: temporal constraints and image feature analysis. This paper focuses on the temporal constraints. Temporal information provides a priori estimates of storm intensity (in terms of wind speed) prior to using any satellite analysis.

The remainder of the paper is organized as follows. Section 2 describes the methodology of temporal analysis. Section 3 provides the details of preliminary work on spatial data processing and results. Section 4 discusses the validation process for temporal proposed technique, and section 5 provides discussion and future work.

2. Methodology

This section is divided into four main phases. The first phase describes the database which is used for training and validation process, the second phase describes the initial data mining steps. The third phase describes extracted features, and the fourth phase outlines the procedure for intensity estimation.

2.1 Database

Hurricane Satellite data (HURSAT–B1) described in Knapp and Kossin (2007) includes best-track intensity used as a ground truth data (see detailed specification in Table 1). This data spans from 1978 through 2009

* *Corresponding author address*: Abdollah Homaifar, NC A&T University, Dept. of Electrical and Computer Engineering, Greensboro, NC 27411;
e-mail: Homaifar@ncat.edu.

and provides TC coverage of Northern Atlantic, Eastern and Western Pacific, Southern Hemisphere and Indian Oceans at 8-km and 3-hour resolution. HURSAT–B1 data files are in a Network Common Data Form (NetCDF) format. Each file is a snapshot of the storm from one of the satellites. Multiple NetCDF files are possible when a storm is visible by two satellites at the same time. In this study, we focus on Northern Atlantic storms from 1978 – 2006. For intensity estimation different groups of storms are considered as training and testing sets to prove robustness of proposed technique.

Product	**Hursat-B1**	**HURSAT-AVHRR**	**HURSAT-MW**	**GridSat**
Temporal span	1978 - 2009	1978 - 2009	1988 - 2009	1979 - 2009
Spatial span	Storm-centric: 10.5° from center for all global TCs	Storm - centric: 10.5° from storm for all global TCs	Storm-centric: 10.5° from center for all global TCs	Global
Temporal resolution	3 hourly	Varying (6-12 hourly)	Varying (6-12 hourly)	3 hourly
Gridding resolution	8km	4km	8km	8km
Data source	ISCCP B1	AVHRR GAC	DMSP SSM/I	ISCCP B1
Channels available	IRWIN (11μm) IRWVP (6.7μm) (0.65μm)	All AVHRR channels	All SSM/I channels	IRWIN(11μm) IRWVP(6.7μm) (0.65μm)
Calibration	Clim. – IRWIN, ISCCP - IRWVP,	Climate calibrated	Operational calibration	Clim. – IRWIN, ISCCP–IRWVP,
Yearly size (GB)	< 6.5	40-60	4	200
Format	NetCDF	NetCDF	NetCDF	NetCDF
Current version	4.0	Beta	Beta	Beta
Imagery	Movies	BD Imagery	Imagery	Planned

Table 1: Detailed specification of Hursat data (Knapp and Kossin 2007)

2.2 **Data Mining**

Data mining has attracted a great deal of attention in the information industry due to availability of huge amounts of data and the urgent need for changing such data into useful knowledge. Every data mining system contains of an iterative sequence of the following steps (Han and Kamber 2006):

1. Data cleaning: Remove incomplete data (e.g., that has not complete history information to be analyzed).
2. Data integration: All Northern Atlantic storms from 1978–2006 are combined for further analysis.
3. Data selection: For each entry the features that described in section 2.3 are retrieved from integrated data. It includes 6, 12, 24 prior intensities and the duration of each entry. To access the origin of the data, the NetCDF file name of each entry is also saved.
4. Data transformation: Data are transformed into matrix forms which are appropriate for data mining.
5. Data mining: Machine learning methods are applied in order to extract an estimation of the intensity of the query entry. These methods are described in section 2.4.
6. Estimation evaluation: Measure the estimator accuracy. Estimation accuracy is discussed in section 4.
7. Knowledge presentation: Results are presented to facilitate understanding of the mined knowledge. This step is covered in section 4.

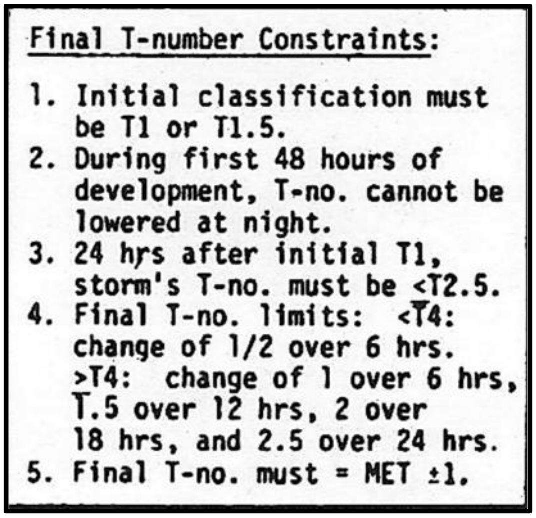
It is important to mention that, classification and estimation are two forms of data analysis that can be used to extract models describing important data classes or to estimate future data trends. Our problem to estimate the intensity of TC considered as prediction of a continuous valued functions and it is different from classification which predicts categorical discrete labels.

2.3 **Features**

In order to select appropriate features (that is, predictors) for intensity estimation, a careful review of the DT revealed a number of interesting correlations between the T numbers (Table 2), the constraint on TC wind speed, duration and prior intensity of the storm. Figure 1 shows the step 8 of DT. This step provides the constraints on final T number in terms of duration, time of the day and prior intensities. One of the important features that strongly related to the intensity of the storm is duration. For each snapshot of the storm, duration

T number	**Wind speed (kt)**
1 .0& 1.5	**25**
2.0	**30**
2.5	**35**
3.0	**45**
3.5	**55**
4.0	**65**
4.5	**77**
5.0	**90**
5.5	**102**
6.0	**115**
6.5	**127**
7	**140**
7.5	**155**
8	**170**

Table 2: Dvorak T numbers



Final T-number Constraints:

1. Initial classification must be T1 or T1.5.
2. During first 48 hours of development, T-no. cannot be lowered at night.
3. 24 hrs after initial T1, storm's T-no. must be <T2.5.
4. Final T-no. limits: <T4: change of 1/2 over 6 hrs. >T4: change of 1 over 6 hrs, 1.5 over 12 hrs, 2 over 18 hrs, and 2.5 over 24 hrs.
5. Final T-no. must = MET ±1.

Figure 1: Step 8 of Dvorak technique (Dvorak 1984)

means the time elapsed between current (time of intensity estimation) and the starting time of the storm. For example constraint No. 1 states that for storm just started, duration= 0, T# is 1 or 1.5.

The other important features that related directly to intensity estimation are prior intensities of the storm up to 24 hour before the estimation time. For example, constraints No. 3 and 4 (Figure 1) are directly related to prior intensities of TC till estimation time.

Finally, 4 features including 6, 12, 24 hour prior intensities and the duration were selected (after testing all combinations of prior intensities) as the most dominant features in intensity estimation. These features are extracted from the NetCDF files of the ground truth data. Figure 2 shows the extracted features for sample query with 12 days duration from storm Kate (2003).

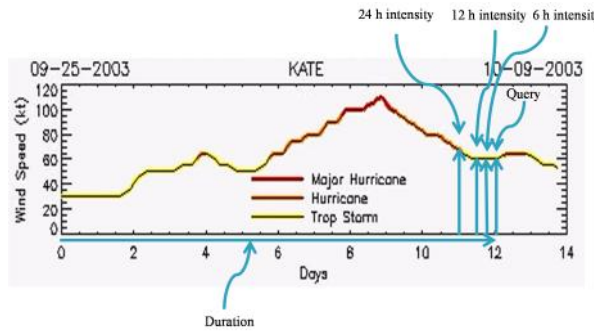


Figure 2: Visualizing extracted features for sample query with 12 days duration from starting point of the storm Kate (2003)

2.4 **Procedure for Intensity Estimation**

Figure 3 illustrates the procedure for the proposed technique. First, all the data in the ground truth data base are organized according to selected features. Second, the features (6, 12, 24 prior intensities and the duration) of each query entry (unknown intensity entry) are extracted. The third and the forth steps are used for sorting similar entries based on duration and prior intensities for a given query. Since the units of the features; duration (hour) and the intensities (kt) are

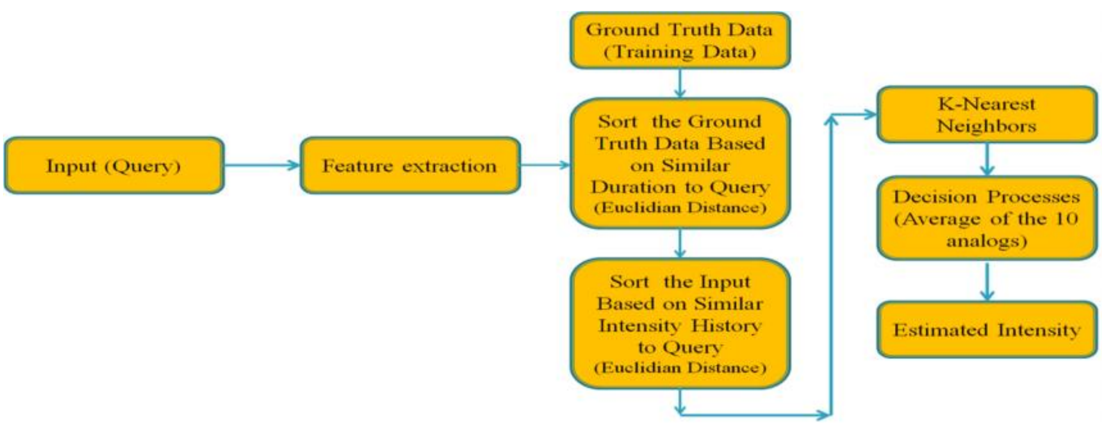


Figure 3: Block diagram of the proposed technique

different. Similarity is defined in terms of Euclidian distance between the query entry and all of the training entries. The Euclidian distance between two entries $X=\{x_1,x_2,...,x_n\}$ and $Y=\{y_1,y_2,...,y_n\}$ is defined as:

$$Dis(X,Y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+...+(x_n-y_n)^2}$$

In the third step, the data base is searched based on similar duration. All entries in the ground truth data are sorted in ascending order based on the computed Euclidian distance between duration (in hours) of query entry and entries from the ground truth data. The entry from the ground truth data with shortest distance (most similar duration; shortest time) is set as the first entry. In the fourth step, the current entries are sorted based on the Euclidian distance between the prior intensities (kt) of the query entry and the ground truth entries. Fifth, we apply k-nearest neighbor algorithm to classify k entries with shortest Euclidian distance. For example, if each entry is described by an *n* attributes (or features). Each entry represents a point in an *n*-dimensional space. In this way, all of the entries are stored in an *n*-dimensional space. When presented with a query with an unknown intensity, a k-nearest-neighbor classifier searches the space for the k entries that are closest to the unknown entry based on duration and prior intensities. These k entries are the k “nearest neighbors” to the unknown query entry. “Closeness” is defined in terms of a metric Euclidean distance. Sixth, the average intensity value of the 10 nearest neighbors is the estimated intensity of the query entry. The algorithm’s performance is affected by the choice of k. If k is small, then the algorithm can be affected by noisy points. If k is too large then the nearest neighbors could belong to different classes. Therefore, k=10 is selected as an optimum value for k based on estimation accuracy after several testing.

An example is provided to clarify our proposed technique. Consider the query entry is the snapshot of the storm Rita for a given date (2005.09.19) at 21:00 UTC with real intensity of 60 kt. Figure 4 shows how Rita was evolved and the corresponding snapshot query entry. The goal is to estimate the intensity corresponding to the given query entry at the given date and time. Assuming that duration and the 6, 12, and 24 hours prior intensities for the query entry are 42 hour, and 58, 53, 40 kt respectively. For k=1, the nearest neighbor storm snapshot in the ground truth data entries is the snapshot of storm Lili (1996.10.17) at 9:00 UTC. Lili’s current and 6, 12, 24 hour prior intensities were 63 kt and 58, 53, 40 kt respectively with 66 hour duration as shown in Figure 5. Finally the estimated intensity value is 63 kt which has negligible error of 3 kt compare to real intensity of 60 kt.

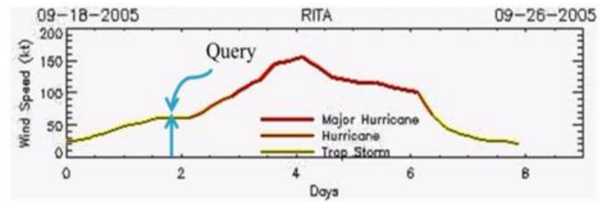


Figure 4: Query entry selected from Rita storm (2005)

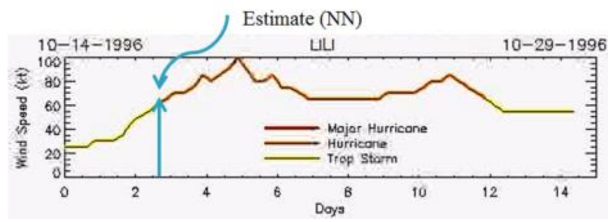


Figure 5: Nearest neighbor from Lili storm (1996)

3. Spatial data processing

The Dvorak technique formulation is based on “Tropical cyclones have characteristic evolutions of cloud pattern that correspond to stages of development and certain intensities” (Dvorak 1984). In this paper we present preliminary results using image processing and pattern recognition techniques to investigate the relation between TC intensity and TC satellite images. And remove the human subjectivity of the processes.

A general technique that has been applied to many tasks in image processing is the 2 dimensional discrete wavelet transform (DWT2), Stark (2005). An important aspect of the DWT2 is its ability to compute and use data in compressed parameters which are often called features. The TC satellite images can be compressed into fewer parameters describing the condition of the TC. It is particularly important to use a smaller set of features for recognition. The DWT2 is capable of representing images in different resolutions by dilating and compressing its basis functions. The compressed functions hold fast activities while the dilated functions adjust to slow activities. DWT2 works like a band pass filter and can be applied to an image at several levels.

Each level decomposes the input image into a low frequency *c*omponent (*a*pproximations-cA) and a high frequency *c*omponent (*v*ertical-cV, *h*orizontal-cH and *d*iagonal-cD) of the initial image. Sample decomposition is shown in Figure 6. The choice of a wavelet function usually application dependent and the criteria used is based on maximum accuracy.

In this study, Daubechies wavelet function of level 1 is used. The smoothing feature of the Daubechies wavelet of order 5 (db5) made it more suitable to detect changes in the satellite image. The wavelet coefficients are computed using the MATLAB software package. Ignoring the high frequency (detail) components (cV, cH and cD), *a*pproximation wavelet coefficients of the input image is used as the feature vector representing the input images. Each satellite image has dimension of 301 by 301 (90601 discrete data points) and only using the approximation wavelet component has reduced the dimension to 24025 (26.5% of input image data points).

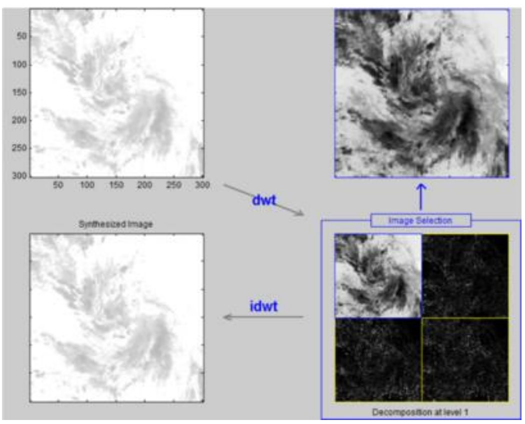


Figure 6: Sample DWT2 of a satellite image, the approximation component is selected.

The procedure for classification of the satellite image is shown in Figure 7. The features of query image are extracted and compared with ground truth data features based on Euclidean distance. The k-nearest neighbor algorithm is used to assign the nearest neighbor (k=1) class to query image.

For validation process, ground truth data and testing data are selected randomly from images in HURSAT-B1 data during 1995-2005 from different classes defined base on Saffir-Simpson hurricane scale (Table 3). The 6 classes are Tropical depression (TD) or Tropical storm (TS) with intensities less than Cat1 plus Cat 1 through Cat5. Table 4 shows the number of entries chosen for validation process with 80% and 20% ratios as Ground truth data and testing data respectively.

Table 5 shows the tabulated result in a confusion matrix. The average accuracy of classification result is 71%. First, in standard TC intensity classification, the intensity accuracy is given in terms of knots. Therefore,

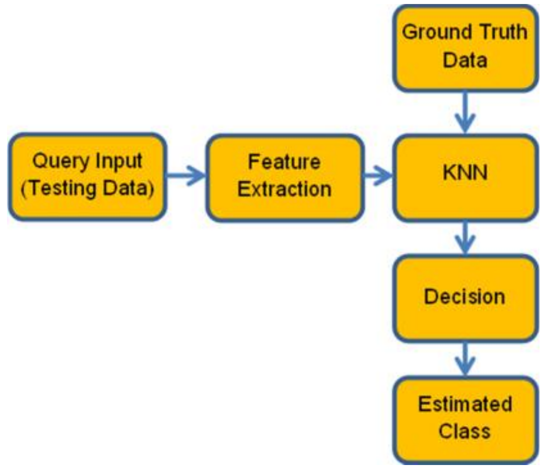


Figure 7: Block diagram of satellite image classification

it is hard to compare the results. Second some spill over between categories (e.g., a storm that was just above the threshold of Cat 3 can be classified as Cat 2) may exist, the error in terms of category is 1-level difference, but the wind speed difference in knots may be small. Third, some large outliers exist. For instance, the Cat 0 column has interpreted 1% of those storms as Cat 5 (2 images) and 5.5% as Cat 4 (11 images). That is a pretty significant difference from Cat 0. Temporal constraints can be used as well to improve the accuracy. For instance, knowing what the category was just an image or two prior can help limit the miss-classifications.

However, when classes are combined in such a way that class TS/TD remain the same as before, classes 1 and 2 are combined as the second class and classes 3, 4,5 are combined as the third class. Then, the average accuracy of the classification has improved and is 81.3% (Table 6). These results show that the classification technique works better at high wind speeds (second and third classes) than the TS/TD class. As anticipated the results indicate that a relationship exists between the images and intensity. Increasing the accuracy of the image classification is the subject of our future work.

4. Validation and Results of Temporal Analysis

Since estimators return a continuous value rather than a categorical label, it is difficult to conclude exactly whether the predicted values are correct or not. Instead of focusing on whether the estimated values are exact match with the real-values, the accuracy is measured in terms of how far off is the predicted values from the reported values. Loss functions measure the error among actual values and the estimated values. The most common loss functions are mean absolute error (MAE) and root mean square error (RMSE) (Han and Kamber 2006). Several tests are done to validate the proposed technique.

Category	**Wind speed (mph)**
Five	**>= 156**
Four	**131-155**
Three	**111-130**
Two	**96-110**
One	**74-95**

Table 3: Saffir-Simpson Hurricane scale

Class	TS/TD	1	2	3	4	5
Ground truth	800	800	336	291	244	60
Testing	200	200	84	73	61	15
Total	1000	1000	420	364	305	75

Table 4: Number of training and testing images from Northern Atlantic storms 1995 – 2005

Class	TS/TD %	Cat1 %	Cat2 %	Cat3 %	Cat4 %	Cat5 %
TS/TD	71	7.5	2.4	0	0	0
Cat1	17.5	77.5	20.2	6.8	3.3	0
Cat2	2.5	8.5	59.5	13.7	0	0
Cat3	2.5	1.5	11.9	64.4	9.8	0
Cat4	5.5	5	6.0	15.1	83.6	33.3
Cat5	1	0	0	0	3.3	66.7

Table 5: Confusion matrix (average accuracy 71%)

Class	TS/TD	C1/C2	C3
TS/TD	71%	6%	0%
C1/C2	20%	84%	11%
C3(Cat3, 4,5)	9%	10%	89%

Table 6: Combined results

In order to provide a quantitative comparison with DT for intensity error estimation the data used are, from 1978 – 1996 and 2004 – 2006 storms as the ground truth data, and from 1997-2003 as a testing data. Table 7 provides the number of data corresponding to different T numbers used for the ground truth and testing data. The proportion of the training and testing data in Table 7 are about 70% and 30% respectively. The MAE of intensity estimation based on DT is shown in Figure 8 (Velden et al. 2006a) with the same period as testing data. This Figure shows that, 50% of the Dvorak intensity estimates are within 5 kt of the best track intensity, 75% are within 12 kt and 90% are within 18 kt. The distribution of intensity estimation errors of the proposed technique, in Figure 9, shows that 50% of the estimates have an MAE less than 2.4 kt, 75% are within 4.4 kt and 90% are within 7.5 kt. It clearly, shows that the proposed technique has an average improvement in MAE intensity estimation of 55% compared to DT.

Class	T1	T1.5	T2	T2.5	T3	T3.5	T4	T4.5	T5	T5.5	T6	T6.5	T7	T7.5	T8
Total	**1759**	**16**	**1854**	**1323**	**2474**	**1759**	**1530**	**1043**	**964**	**343**	**363**	**203**	**105**	**49**	**6**

(a)

Class	T1	T1.5	T2	T2.5	T3	T3.5	T4	T4.5	T5	T5.5	T6	T6.5	T7	T7.5	T8
Total	**649**	**5**	**652**	**606**	**951**	**788**	**632**	**398**	**420**	**172**	**196**	**82**	**57**	**11**	**0**

(b)

Table 7: (a) Classified Ground truth data from North Atlantic storms 1978-1996 & 2004 – 2006 storms, total is 13791 entries (b) Classified Query (testing) data from North Atlantic storms 1997 – 2003 storms total is 5619 entries.

For statistical justification of the proposed technique several tests are done using k-Fold Cross-Validation and compared it with DT. The initial data are partitioned into k mutually exclusive subsets or folds, $S_1, S_2, \ldots, S_k$. Since the time interval consists of 29 years (1978 – 2006) then k=29 and each subset considered as storms occurred during each year from 1978 – 2006. The entire testing is performed k times. In iteration (test) j, partition S_j is reserved as the test set, and the remaining partitions are collectively used as ground truth data. For example, in the first iteration, subsets S_2, ..., S_k collectively serve as the ground truth set, and S_1 is used as test data; the process is repeated similarly. Note that, each sample is used the same number of times as ground truth data and once for testing data. The error estimate is computed as the total loss from the k iterations divided by the total number of initial subsets. Figure 10 shows the results of the proposed method and DT for BIAS, MAE, and RMSE. The term BIAS means the average differences between estimated values and the best track values. The Dvorak intensity estimates in Figure 10 d, e, f are from the National Environmental Satellite, Data, and Information Service's (NESDIS) Satellite Analysis Branch (SAB) and the Tropical Analysis and Forecast Branch (TAFB), which is part of the Tropical Prediction Center/National Hurricane Center (Knaff et al. 2010). Although the data are used for verification of proposed technique and the one used for DT shown in Figure 10 are not the same, but it gives an overall view of the accuracy of each technique. Mean Bias values of the proposed technique are almost

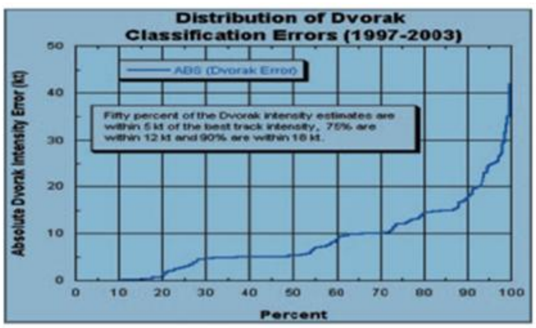


Figure 8: Distribution of Dvorak classification errors (1997 - 2003) in the Atlantic basin (Velden et al. 2006a)

zero but for DT the Bias values changes from -8 to 4 kt. MAEs and RMSEs of DT during 1989 - 2008, are with mean values of approximately 8 and 11 kt respectively. But MAEs and RMSEs of proposed technique during 1978 – 2006 are with mean values of approximately 3 and 5 kt respectively. The results show that proposed technique has greater accuracy than DT.

Biases and errors can be showed as a function of intensity to study the underestimate/overestimates and variation of errors in different intensities. The results of validation are compiled in overlapping bins with endpoints that correspond to the T number versus intensity as in Table 1. For instance, the first and second bins have intensity ranges of 20 – 35 and 25 – 45 kt, respectively. The last bin extends from 127 to 170 kt as one bin (Knaff et al. 2010). Figure 11 shows the biases, MAEs, and RMSEs associated with the proposed technique and DT intensity estimates. The biases show that the DT underestimates intensities when TCs have intensities between 35 and 55 kt and greater than 125 kt. On the other hand, overestimation of intensities occurs between 75 and 105 kt. However, for the proposed technique the underestimates occur especially for intensities greater than 115 kt. The MAEs and RMSEs shown in Figure 11 are lower for weak storms and larger for the higher intensities which are similar for both techniques.

Studying the effect of noise is important because measurements are often corrupted and tend to propagate noise on estimations. We considered noise with a Gaussian distribution having a zero mean with 5 kt and 6 hour standard deviations for prior intensities and duration, respectively. We add Gaussian noise to all features and then, the k-Fold Cross-Validation

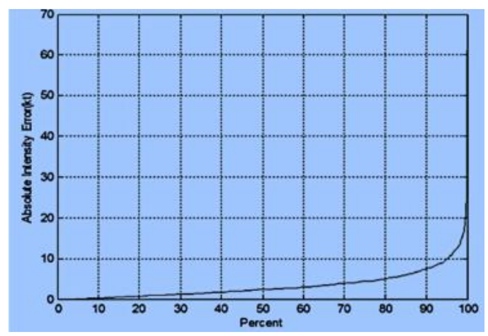


Figure 9: Distribution of proposed technique classification errors (1997 - 2003) in the Atlantic basin

are performed with new values. Figure 12 indicates the new results. The biases are not changed since the mean of the noise is zero. However, the errors (MAE and RMSE) are both increased. In spite of noise, the average RMSE of the proposed technique is around 8.2 kt which is still less than Dvorak error of 11.7 kt.

5. Discussion and future work

In summary, we hypothesize that discovering unknown regularities and abnormalities that may exist in the large group of past observations could help human experts interpret TC intensity changes from various points of view. Our goal is to provide a data mining tool that increases the ability of human experts to analyze

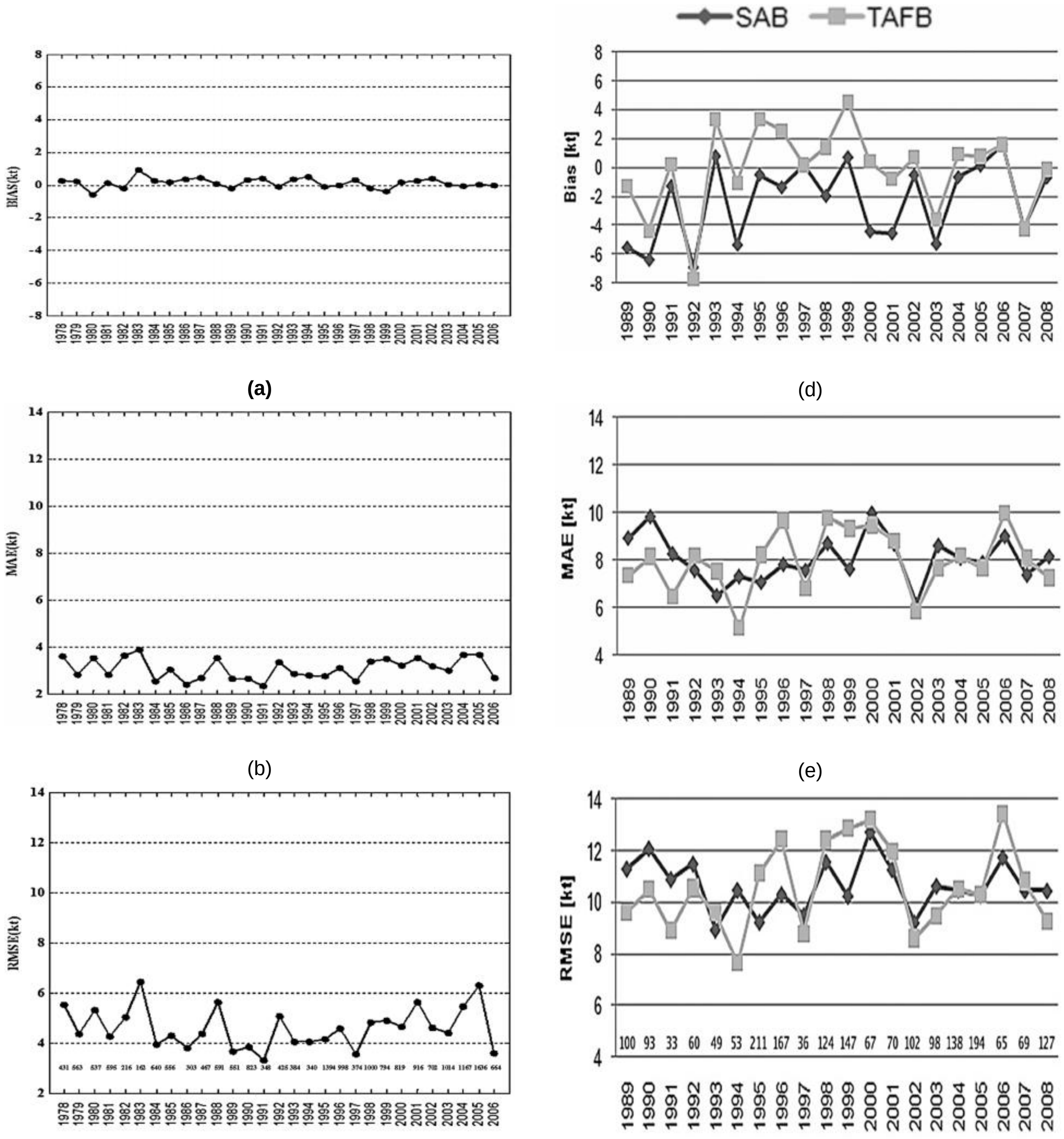


Figure 10: K-Fold Cross–Validation for North Atlantic base in from 1978 – 2006, (a),(b), (c)Bias, MAE, RMSE of proposed technique (d), (e), (f) DT results (Knaff et al. 2010), the number of cases is provided in the bottom panel. (Satellite Analysis Branch (SAB), Tropical Analysis and Forecast Branch (TAFB))

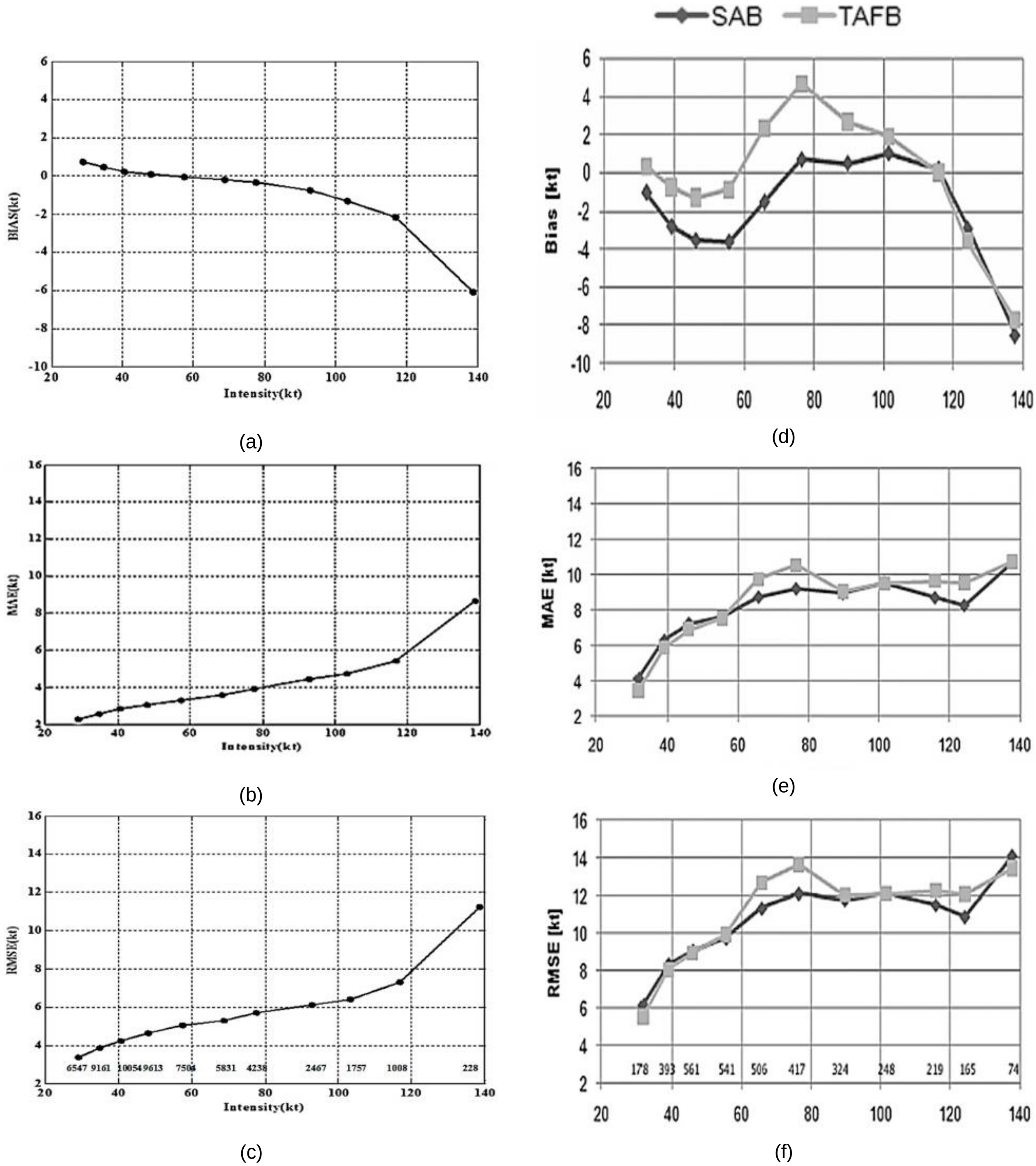
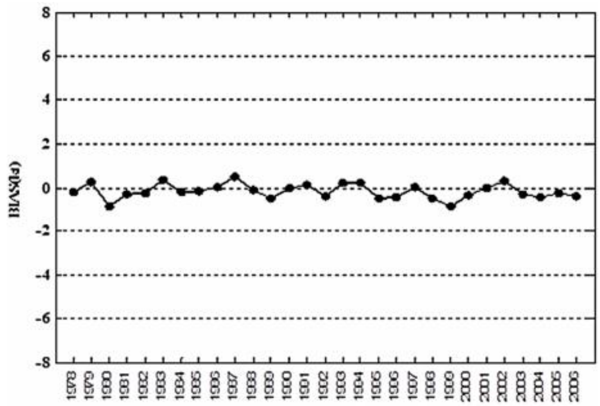
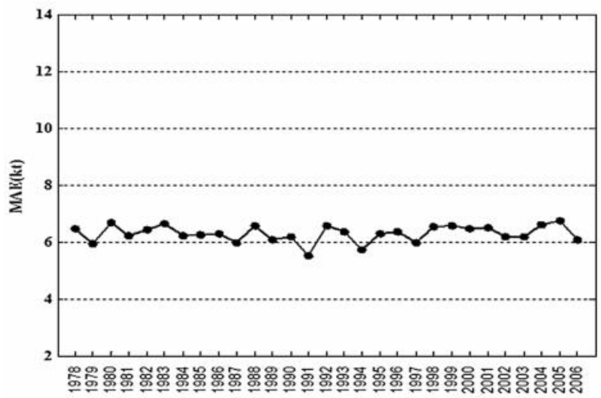


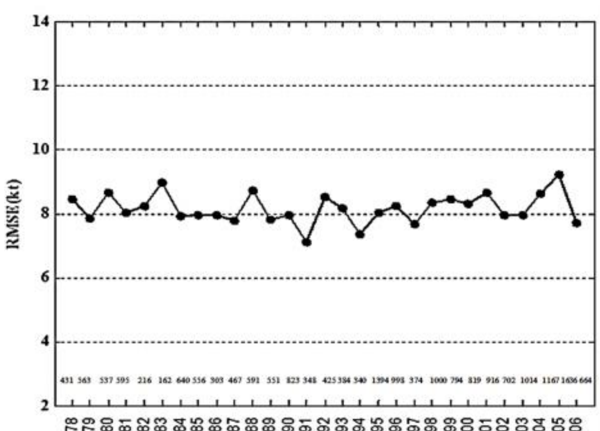
Figure 11: K-Fold Cross –Validation for North Atlantic basin from 1978–2006 showed as a function of intensity. (a),(b),(c) Bias, MAE, RMSE of proposed technique (d), (e), (f) Bias, MAE, RMSE of DT (Knaff et al. 2010). The number of cases is provided in the bottom panel. (Satellite Analysis Branch (SAB) Tropical Analysis and Forecast Branch (TAFB)).



(a)



(b)



(c)

Figure 12: Studying the effect of the noise (a) Bias (b) MAE (c) RMSE. The number of cases is provided in the bottom panel.

huge amount of historical data for TC intensity estimation.

Temporal information provides a priori estimates of TC intensity before using any satellite analysis. The temporal analysis uses the duration, 6, 12 and 24 hours prior intensities of TC as predictors of the expected intensity. The algorithm used 70% and 30% of the data as the ground truth data and verification data respectively. Instead of regression techniques, the 10 closest analogs (determined using a k-nearest-neighbor (K-NN) algorithm) are averaged to estimate the intensity. Such an estimate has a 4.8 kt RMSE (50% of points are within 2.4 kt). Several tests were implemented to statistically justify the proposed algorithm using k-Fold Cross-Validation. The resulting average RMSE is 4.6 kt. We considered noise as having a zero mean Gaussian distribution with 5 kt and 6 hour standard deviations for prior intensities and duration, respectively to study its' impact. The results indicate that the average RMSE is around 8.2 kt.

The proposed technique has the potential to provide new temporal constraints on satellite analyses (e.g., the Dvorak technique). The current analysis has the potential to decrease the Dvorak noise. Once spatial analysis of image features is included, the noise will likely be less. For example, our current spatial technique analyzes 71% of storms at the correct Saffir-Simpson category. The next step in algorithm development is to combine the temporal analysis with the satellite image analysis.

Acknowledgments

This work is supported by the Expeditions in Computing by the National Science Foundation under Award CCF-1029731.

Disclaimer

The opinions expressed in this paper are those of the authors. They do not necessarily reflect the official views or policies of NOAA, Department of Commerce, or the US Government.

References

Dvorak, V. F., 1984: Tropical Cyclone Intensity Analysis Using Satellite Data. NOAA Technical Report *NESDIS*, 11, 1–47.

Dvorak, V. F., 1975: Tropical cyclone intensity analysis and forecasting from satellite imagery. Mon. Wea. Rev., 103, 420–430.

Han, J., and M. Kamber, 2006: Data Mining: Concepts and Techniques. Morgan Kaufmann, 743 pp.

Knaff, J. A., D. P. Brown, J. Courtney, G. M. Gallina, and J. L. Beven II, 2010: An Evaluation of Dvorak Technique–Based Tropical Cyclone Intensity Estimates. Wea. Forecasting, 25, 1362-1379.

Knapp, K. R. and J. P. Kossin, 2007: New global tropical cyclone data from ISCCP B1 geostationary satellite observation. Journal of Applied Remote Sensing, 1, 013505.

Olander, T. L., and C. S. Velden, 2007: The advanced Dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. Wea. Forecasting, 22, 287–298.

Velden, C. S., and Coauthors, 2006a: The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. Bull. Amer. Meteor. Soc., 87, 1195 – 1210.

Velden, C. S., and T. L. Olander, 1998: Development of an Objective Scheme to Estimate Tropical Cyclone Intensity from Digital Geostationary Satellite Infrared Imagery. Wea. Forecasting,13, 172-186.

Stark, H.-G., 2005: Wavelets and signal processing. Springer-Verlag, 150 pp.