Satellite-Derived Tropical Cyclone Intensity in the North Pacific Ocean Using the Deviation-Angle Variance Technique

ELIZABETH A. RITCHIE AND KIMBERLY M. WOOD
Department of Atmospheric Sciences, The University of Arizona, Tucson, Arizona

OSCAR G. RODRÍGUEZ-HERRERA, MIGUEL F. PIÑEROS, AND J. SCOTT TYO
College of Optical Sciences, The University of Arizona, Tucson, Arizona

(Manuscript received 7 February 2013, in final form 14 November 2013)

ABSTRACT
The deviation-angle variance technique (DAV-T), which was introduced in the North Atlantic basin for tropical cyclone (TC) intensity estimation, is adapted for use in the North Pacific Ocean using the “best-track center” application of the DAV. The adaptations include changes in preprocessing for different data sources [Geostationary Operational Environmental Satellite-East (GOES-E) in the Atlantic, stitched GOES-E–Geostationary Operational Environmental Satellite-West (GOES-W) in the eastern North Pacific, and the Multifunctional Transport Satellite (MTSAT) in the western North Pacific], and retraining the algorithm parameters for different basins. Over the 2007–11 period, DAV-T intensity estimation in the western North Pacific results in a root-mean-square intensity error (RMSE, as measured by the maximum sustained surface winds) of 14.3 kt (1 kt ≈ 0.51 m s⁻¹) when compared to the Joint Typhoon Warning Center best track, utilizing all TCs to train and test the algorithm. The RMSE obtained when testing on an individual year and training with the remaining set lies between 12.9 and 15.1 kt. In the eastern North Pacific the DAV-T produces an RMSE of 13.4 kt utilizing all TCs in 2005–11 when compared with the National Hurricane Center best track. The RMSE for individual years lies between 9.4 and 16.9 kt. The complex environment in the western North Pacific led to an extension to the DAV-T that includes two different radii of computation, producing a parametric surface that relates TC axisymmetry to intensity. The overall RMSE is reduced by an average of 1.3 kt in the western North Pacific and 0.8 kt in the eastern North Pacific. These results for the North Pacific are comparable with previously reported results using the DAV for the North Atlantic basin.

1. Introduction
a. Intensity estimation from remote sensing data

Unlike land-based surface observations, in situ surface observations over the tropical oceans are sparse. This lack of data results in a heavy reliance on satellite observations in order to monitor severe weather systems such as tropical cyclones (TCs). Using these satellite measurements, it is possible to estimate TC intensity when direct measurements are not available. The Dvorak technique (Dvorak 1975) and its extended versions, the objective Dvorak technique (Velden et al. 1998) and the advanced Dvorak technique (Olander and Velden 2007), are well-known examples of methodologies to estimate the maximum sustained surface winds of TCs from satellite imagery. In the original Dvorak technique, the analyst considers patterns and cloud-top temperatures in a series of satellite images, classifying the scenes according to predefined patterns and assigning a T number based on the features of the classification. The performance of the technique in the North Atlantic and eastern North Pacific basins is described in Knaff et al. (2010). Other procedures are based on the Advanced Microwave Sounding Unit as described in Demuth et al. (2004) and Spencer and Braswell (2001). Microwave frequencies can reveal important structural features that are obscured by high-level cirrus and thus cannot be observed...
by infrared imagers. However, limitations exist for each of these techniques, and, in some cases, their performance depends on the skill of the user. It should also be noted that the absence of aircraft reconnaissance in the western and eastern North Pacific basins makes it difficult to compare the intensity estimates with in situ observations.

b. The deviation-angle variance technique

The deviation-angle variance technique (DAV-T), a method that can be used to estimate the intensity of TCs, was described and applied to the North Atlantic basin by Pineres et al. (2008, 2011) and Ritchie et al. (2012). Here, intensity is defined as the maximum 1-min sustained 10-m surface wind speed near the core of the TC. This technique quantifies the axisymmetry of TCs in IR satellite imagery by performing a statistical analysis of the direction of the gradient of the IR brightness temperatures. The brightness temperature gradient vectors are more aligned along radials of an intense TC than the vectors computed from a weak, disorganized tropical system. This degree of storm organization can be quantified by calculating the variance of the angle between the gradient vector and the radial at each pixel in the scene. Previous results for the North Atlantic basin show that the variance of this deviation angle is negatively correlated with the TC’s intensity (Pineres et al. 2008). Although the original application of the technique was completely objective with an automated center-finding routine in place, there were significant errors in the intensity estimation for a very small number of highly sheared cases of TCs (Pineres et al. 2011). An updated version, which used the National Hurricane Center’s best-track archive to center the algorithm, yielded significant improvements in the specific sheared cases. Root-mean-square (RMS) intensity errors using the best-track center version of the DAV-T for the North Atlantic basin over the period 2004–10 were 12.9 kt (1 kt $= 0.51$ m s$^{-1}$) (Ritchie et al. 2012), an improvement of 1.8 kt over the automated center routine.

In this paper, this same DAV-T for intensity estimation is adapted to the eastern and western North Pacific basins. The best parameters for DAV-T are determined, and the method is expanded to accommodate the more diverse storm characteristics that are present in these basins. In section 2, this methodology is briefly discussed, and the method for stitching imagery from two separate geostationary satellites to be used in the eastern North Pacific is described. The results found when applying the technique to the western North Pacific are described in section 3. Section 4 discusses the performance of the DAV-T in the eastern North Pacific. Finally, section 5 presents some conclusions and details future development work.

2. Methodology

a. IR imagery used for the DAV-T

This paper presents results from two different intensity estimation studies: one in the western North Pacific and one in the eastern North Pacific. The data for the western North Pacific portion of this study are derived from longwave (10.7 $\mu$m) IR satellite imagery from the Japan-based Multifunctional Transport Satellite (MTSAT) over the western North Pacific basin. The images used in the western North Pacific study are resampled to a spatial resolution of 10 kilometers per pixel and encompass an area from 0° to 40°N and from 100°E to 180°E.

The data for the eastern North Pacific study are derived from longwave (10.7 $\mu$m) IR satellite imagery obtained from the Geostationary Operational Environmental Satellite-West (GOES-W) and -East (GOES-E). Data are obtained from both satellites as GOES-W does not cover the extreme eastern part of the eastern North Pacific basin. These images are also resampled to a spatial resolution of 10 kilometers per pixel. To cover the longitudinal extent of the basin, GOES-W images are stitched with GOES-E images after resampling in order to generate the final data used in the study. The resampled image from each satellite is blended using a column of pixels centered near 104.7°W and averaging the data from each image out to four pixels on either side of this center pixel in order to account for slight shifts resulting from the 15-min time difference between each satellite image. The final stitched images encompass an area from 0° to 40°N and from 170° to 80°W.

Similar to previous versions of the technique, those samples when the TC center is over land (continental land and large islands) in each basin are removed from the analysis. This removal does not affect the ability to use the technique near and over small islands. Furthermore, RMS intensity errors for near-land (0–250 km) samples are very close to the overall RMS intensity error for all samples (data not shown).

The western North Pacific study considers images that include existing TCs from the 2007–11 typhoon seasons and comprise a total of 22,552 unique hourly images. The resulting dataset includes 12 supertyphoons, 32 typhoons, 27 tropical storms, and 18 tropical depressions. Furthermore, of the 22,552 images, 47% are tropical depression intensity, 27% are tropical storm intensity, 24% are typhoon intensity, and 2% are super typhoon intensity. The eastern North Pacific study examines images with existing TCs from the 2005–11 hurricane seasons and comprises a total of 20,213 unique half-hourly images. The resulting dataset includes 21 major hurricanes, 25 hurricanes, and 44 tropical storms. The North Atlantic hurricane database (HURDAT) does not include
systems that peaked below tropical storm (34 kt) intensity; thus, none of these cases was included in the analysis. Of the 20,213 images, 38% are tropical depression intensity, 40% are tropical storm intensity, 15% are hurricane intensity, and 7% are major hurricane intensity.

Examples of the IR image data analyzed in this study are shown in Fig. 1. The western North Pacific image (Fig. 1a) contains Typhoons Parma and Melor (2009), and the eastern North Pacific image (Fig. 1b) includes Hurricanes Celia and Darby (2010). The results of applying the DAV-T to these data are compared against the best-track databases from the Joint Typhoon Warning Center (JTWC; http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/) and the National Hurricane Center (NHC; http://www.nhc.noaa.gov/data/#hurdat).

b. Tropical cyclone intensity estimation

The DAV-T was introduced in Piñeros et al. (2008) as a procedure to objectively estimate the intensity of TCs. The procedure quantifies the level of axisymmetry of TC brightness temperatures by computing the gradient of the brightness temperatures from the infrared images and then calculating the deviation angle of the gradient vectors with respect to a radial projected from a defined center. The amount of deviation from the perfect radial indicates the level of alignment of the gradient vectors. An example of this calculation for a single gradient vector, when the center is located in the eye of the vortex, is shown in Fig. 2a. The variance of the distribution of angles within a defined radius from the center then quantifies the axisymmetry of the cyclone relative to that center point (Fig. 2b). The DAV decreases at a given center point as the system becomes more axisymmetric about that point, since the gradient vectors within the calculation radius are pointing more uniformly toward or away from the center. In contrast, the DAV increases when the orientation of the gradient vectors becomes increasingly asymmetric with respect to that center point. This asymmetry is characteristic.
of nondeveloping cloud clusters, sheared TCs, and other disorganized cloud structures. Whereas the fully automated system used the minimum variance point to perform the intensity estimation (Piñeros et al. 2011), Ritchie et al. (2012) updated the intensity estimation algorithm by using a known center pixel as a reference for the calculation. The system can now be operated using automatic centers, best-track centers, or any other center point defined by the user. Recent results in the North Atlantic basin showed that the use of best-track TC fixes provided measurable improvements in the intensity estimate over the original automated center-finding methodology (Ritchie et al. 2012). Therefore, in this study the best-track positions are used in both basins to constrain the intensity estimation.

For intensity estimation, the DAV values are obtained at the single pixel location identified in the best-track archives of the operational centers. For the western North Pacific study, MTSAT IR images are available hourly, and thus the 6-h best-track positions are interpolated using a cubic spline interpolation to match the times of the IR images. For the eastern North Pacific study, GOES images are available half-hourly, and the 6-h best-track positions are again interpolated to match the temporal resolution of the IR imagery. The western North Pacific (eastern North Pacific) best-track intensity values are linearly interpolated to hourly (half hourly) resolution to properly compare these data with the DAV estimates.

The DAV-T has been well tested in the North Atlantic, leading to annual RMS intensity errors of 10–14 kt (Ritchie et al. 2012). In the majority of the years in the North Atlantic, the performance of the DAV-T is optimized at a radius of calculation of 200–250 km. However, because there are differences in both the large-scale environment and the modes of genesis between the North Atlantic and western North Pacific basins (e.g., Lander 1994; Ritchie and Holland 1999; Briegel and Frank 1997; Ritchie et al. 2003), with considerably more variety in the sizes and structures of TCs in the western North Pacific (e.g., Lander 1994), it is expected that the DAV-T parameters would be different for the North Pacific from those chosen for the North Atlantic. Therefore, the DAV-T is tested over eight different radii of DAV calculation, varying from 150 to 500 km in steps of 50 km, in order to capture the large size variability of the TCs in the North Pacific.

Similar to the North Atlantic technique, the DAV time series is smoothed in both basins using a low-pass filter (impulse response of $e^{-t/T}$) with a filter time constant of 100 h. The filtering reduces high-frequency oscillations in the DAV time series allowing for better comparison with the highly smoothed interpolated best-track intensity estimates. However, the filter also introduces a time delay that can lead to errors in rapidly intensifying cases (Ritchie et al. 2012).

After computing the DAV values, the western and eastern North Pacific datasets are divided into two subsets: a training set comprising all samples from all but one of the years, and an independent testing set comprising all samples from the excluded year. The training set is used to calculate a parametric curve that relates the TC intensity with the DAV signal for each basin, and the testing set is used to independently test the accuracy of the wind speed estimate produced by this parametric curve. A sigmoid curve unique to each basin is fitted to the training data. The curve for the western North Pacific is given by

$$f(\sigma^2) = \frac{140}{1 + \exp[\alpha(\sigma^2 + \beta) - 25 \text{ kt}],}$$

and the curve for the eastern North Pacific is given by

$$\text{Fig. 2.}$$

(a) Deviation-angle calculation for a gradient vector, where the center is located in the eye of the tropical cyclone. (b) Distribution of angles for the calculation using the center (eye) pixel. The variance of the histogram is $(1216^2)^2$. (c) Map of variances $(\theta^2$ calculated by performing the DAV calculation using each pixel in turn as the center pixel and mapping the variance value back to that pixel.)
\[ f(\sigma^2) = \frac{130}{1 + \exp[\alpha(\sigma^2 + \beta) + 25 \text{ kt}]} \quad (2) \]

where \( \sigma^2 \) is the DAV value and \( \alpha \) and \( \beta \) are two free parameters that describe the relation between the axisymmetry parameter and the best-track intensity estimates. The parameters \( \alpha \) and \( \beta \) from Eqs. (1) and (2) are fit to minimize the mean square error from the interpolated best-track data. Note that Eq. (1) constrains the minimum intensity that can be estimated in the western North Pacific to 25 kt and the maximum intensity to 165 kt, and Eq. (2) restricts the eastern North Pacific to a range of 25–155 kt.

Examples of the two-dimensional histogram of the filtered DAV samples and the best-track intensity for both basins are provided in Fig. 3. The western North Pacific training set is composed of all 89 TCs during the period 2007–11 shown for a radius of 300 km (Fig. 3a), and the eastern North Pacific training set is composed of all 90 TCs during the period 2005–11 shown for a radius of 200 km (Fig. 3b). Finally, the RMS intensity errors for both basins are calculated by comparing the DAV-T-derived intensity estimates to the best-track archived intensity estimates from the JTWC (western North Pacific) and the NHC (eastern North Pacific).

3. Results for the western North Pacific basin

a. Tropical cyclone intensity estimation

Testing for the intensity estimation was accomplished in two ways. First, an overall test result was calculated by training on all five years of data and then testing on those same five years for all eight radii (150–500 km) of the DAV calculation. A second result was obtained by withholding each year, training on the other four years to create the DAV parametric curve, and using the resulting curve to estimate wind speed intensities for the withheld year. The DAV intensity estimates were then compared with the interpolated JTWC best-track intensity records by calculating the RMS intensity error between the two intensity signals. The result of training on 2007–10 and testing on 2011 is shown in Fig. 4a as an example. Average RMS intensity errors for the test on all five of the years and for each test year using several different radii of DAV calculation are shown in Table 1.

As given in Table 1, training on all 89 storms and then testing on those same 89 storms produces a best overall RMS intensity error of 14.3 kt using a radius of calculation of 300 km. Furthermore, RMS intensity errors vary between 12.9 and 15.1 kt for the individual years (Table 1, Fig. 4a). These results are not particularly sensitive to small changes in the sigmoid parameters. In addition, the technique performs best for a calculation radius of 300 km, and this is consistent through the years of the study. This suggests that the axisymmetric structure of the clouds within about 300 km of the center as measured by the DAV technique generally has the highest correlation with the intensity of the TC in the western North Pacific basin.

Further analysis shows that western North Pacific TCs tend to partition into two groups of storms: those that are best characterized by a radius of calculation of 200–300 km, and those large storms that are better characterized by a radius of calculation of 450–500 km (e.g., Fig. 5b). We suspect that this is a result of the more complex mechanisms that modulate TC size and structure in the monsoon environment of the western North Pacific.
Pacific compared to those in the North Atlantic (e.g., Chang and Song 2006; Harr et al. 1996a, b). This variability results in TCs that range in size from “midgets,” with radii less than 200 km [e.g., Typhoon Ellie 1991; see Lander (1994) and Joint Typhoon Warning Center (1991)], to “giants,” with radii greater than 1000 km [e.g., Supertyphoon Tip 1979; see Joint Typhoon Warning Center (1979)]. In many cases, the core of the TC may be surrounded by disorganized cloud clusters associated with the monsoon trough. These cloud clusters are not directly associated with the TC, but they adversely impact the technique when they are included in the DAV calculation. In these cases, a small radius of DAV calculation is desirable. Alternatively, for those larger TCs that form in the monsoon trough with associated banding structures, a larger radius of DAV

![Graph showing intensity estimates and best-track intensities for 2007–10 and 2005–07 to calculate the DAV-T-intensity parametric curve.](image)

**Table 1.** RMS intensity error (kt) for every radius of DAV calculation (150–450 km), number of storms, and number of samples for each individual year for the western North Pacific. The lowest annual RMS intensity errors are highlighted in italics.

<table>
<thead>
<tr>
<th>Year</th>
<th>150 km</th>
<th>200 km</th>
<th>250 km</th>
<th>300 km</th>
<th>350 km</th>
<th>400 km</th>
<th>450 km</th>
<th>TCs</th>
<th>No. of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>19.4</td>
<td>17.0</td>
<td>14.5</td>
<td>13.7</td>
<td>14.3</td>
<td>14.7</td>
<td>15.2</td>
<td>23</td>
<td>6368</td>
</tr>
<tr>
<td>2008</td>
<td>18.1</td>
<td>17.1</td>
<td>15.8</td>
<td>15.1</td>
<td>15.4</td>
<td>15.7</td>
<td>16.2</td>
<td>14</td>
<td>3821</td>
</tr>
<tr>
<td>2009</td>
<td>17.5</td>
<td>15.9</td>
<td>15.1</td>
<td>16.0</td>
<td>16.8</td>
<td>16.9</td>
<td>17.1</td>
<td>15</td>
<td>5042</td>
</tr>
<tr>
<td>2010</td>
<td>18.7</td>
<td>16.5</td>
<td>14.9</td>
<td>14.0</td>
<td>14.1</td>
<td>14.6</td>
<td>14.9</td>
<td>19</td>
<td>4915</td>
</tr>
<tr>
<td>2011</td>
<td>18.6</td>
<td>15.8</td>
<td>14.2</td>
<td>12.9</td>
<td>14.1</td>
<td>14.9</td>
<td>16.0</td>
<td>18</td>
<td>2406</td>
</tr>
<tr>
<td>All</td>
<td>18.5</td>
<td>16.5</td>
<td>14.9</td>
<td>14.3</td>
<td>14.8</td>
<td>15.3</td>
<td>15.7</td>
<td>89</td>
<td>22,552</td>
</tr>
</tbody>
</table>
calculation is more appropriate. Figure 5 shows an example of those features that introduce noise to the DAV calculation for small TCs embedded within monsoon cloudiness but are desirable for large TCs. In Fig. 5a the circular core of Supertyphoon Sepat is detached from the cloudiness to the south by a narrow dark moat region, which is associated with monsoonal southwesterlies. The size of Sepat, as measured by the radius of 34-kt winds at this time, is approximately 200 km according to the JTWC best-track archive, and the entire system is enclosed by the dashed black circle (250-km radius). In contrast, in Fig. 5b it is more difficult to distinguish the surrounding clouds from the inner core of Typhoon Man-Yi. At this time, the JTWC best-track archive lists the size of the storm as approximately 270 km. Here, the rainband features are inseparable from the core and appear to be an integral part of the TC structure. This analysis will be expounded upon further in section 3b.

Finally, categorizing by intensity bins (Table 2) results in a similar trend as is found in the North Atlantic (Ritchie et al. 2012). For the highest intensities (above 64 kt), the combined RMS intensity error is 19.5 kt, and more storms are underestimated than overestimated. For tropical storms, the RMS intensity error is 13.8 kt, and the majority of samples are again underestimated. The RMS intensity error for tropical depression intensities is 11.0 kt. This sample constitutes 47% of the population, and 82% of these samples are overestimated by the technique. As discussed in Ritchie et al. (2012) for the North Atlantic basin, these biases are at least partly due to the choice of the parametric curve used to fit the training data (e.g., Fig. 3a). The curve tends to underestimate the higher-intensity values and overestimate the lower-intensity values (Table 2). A lack of training data at the highest end of the intensity spectrum compounds the problem. We are currently working on other parametric methods to address this problem.

b. Tropical cyclone size variability and the two-dimensional training surface

As mentioned in the previous section, the large variety of TC sizes and structures that exists in the predominantly monsoon trough environment of the western North Pacific presents a challenge for a scheme that measures the structure of the clouds. One solution is to find a way to automatically take the size of the storm into account, either as an average over its life or at any instant in time. However, this is extremely challenging as there are no

<table>
<thead>
<tr>
<th>Bin</th>
<th>No. of samples</th>
<th>RMSE (kt)</th>
<th>No. overestimated</th>
<th>No. underestimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropical depression (&lt;34 kt)</td>
<td>10621</td>
<td>11.0</td>
<td>8693 (82%)</td>
<td>1928 (18%)</td>
</tr>
<tr>
<td>Tropical storm (34–63 kt)</td>
<td>6197</td>
<td>13.8</td>
<td>2065 (33%)</td>
<td>4132 (67%)</td>
</tr>
<tr>
<td>Typhoon (64–129 kt)</td>
<td>5324</td>
<td>19.5</td>
<td>1755 (33%)</td>
<td>3569 (67%)</td>
</tr>
<tr>
<td>Supertyphoons (&gt;129 kt)</td>
<td>410</td>
<td>19.5</td>
<td>46 (11%)</td>
<td>364 (89%)</td>
</tr>
</tbody>
</table>
reliable, continuous measurements of the storm size currently available. We are studying methods that allow the optimum calculation radius to be objectively estimated from the brightness temperature data directly, but to date have not found a robust solution.

Figure 5 shows two examples of storms from 2007 (best overall performance radius of 300 km) that were best characterized by the DAV calculated over very different radii. Supertyphoon Sepat (Fig. 5a) was a small-to-medium-sized TC characterized by a circular core of cold brightness temperatures, separated from but embedded in monsoonal southwesterlies, and it is best represented by a radius of 250 km. Using this small radius to train and test the DAV results in an RMS intensity error over the storm’s lifetime of 13.3 kt, in comparison to the RMS intensity error of 18 kt found from the overall best 300-km radius of calculation. Typhoon Man-Yi (Fig. 5b) was a medium- to large-sized TC characterized by spiral bands extending more than 500 km from the center of the storm. The best performance from the DAV is achieved by using a radius of calculation of 400 km, which produced a lifetime RMS intensity error of 14.9 kt in comparison with 16 kt found from a 300-km radius of calculation. Note that, for Supertyphoon Sepat, a large radius of 400 km yields an RMS intensity error of 32.3 kt, while a small radius of 250 km yields an RMS intensity error of 17.2 kt for the larger Typhoon Man-Yi. It appears that the size of the individual storm may have a bearing on the best radius over which the DAV analysis should be performed. It is even plausible that, as storms change size during their lifetime, a suitably varying radius of calculation may produce the most accurate DAV intensity estimates. This is an ongoing subject of research.

Sifting through the entire dataset reveals that the western North Pacific TCs tend to fall into one of three categories: 1) those that are small and surrounded by peripheral monsoon cloudiness and only perform well with a radius of DAV calculation of 200–250 km, 2) those that are large and only perform well with a radius of DAV calculation of 400–500 km, and 3) those systems that show no preference for a small or large radius of DAV calculation. To accommodate the first two classes of tropical cyclones, a set of DAV training data is specified using a combination of two radii calculations: a small radius and a large radius. A rectangular grid surrounding the entire set of training data is defined as an array of nodes to describe a two-dimensional fitted surface. The fitting method works by expressing the value of the surface at points within the rectangular grid as a regularized linear regression problem, where the regularization is used to assure the existence of the solution when the number of grid points (i.e., the number of nodes) is larger than the number of sampling points. To build the original problem matrix, the surface value at the sampling points within the grid is expressed as a linear combination of the values that the fitted surface would have at the nodes surrounding the points. The linear combination is then written as a matrix, A, given by a relation of the form

$$\mathbf{Ax} = \mathbf{y},$$

where x is a vector with the value of the surface at the nodes and y is a vector with the value of the surface at the sampling points. The regularization matrix is obtained by forcing the first partial derivatives of the surface in neighboring regions to be equal at the nodes. This condition may be written as

$$\mathbf{Bx} = 0,$$

where B is the regularization matrix. The two matrices are then coupled and the resulting system of linear equations is solved to find the surface value at the nodes by finding the vector x that minimizes the coupled system of linear equations (D’Errico 2013). Mathematically, this may be expressed as

$$(\mathbf{Ax} - \mathbf{y})^2 + \lambda \mathbf{Bx}^2,$$

where $\lambda$ is a parameter that can be used to control the smoothness of the fitted surface. The resulting two-dimensional parametric surface is shown in Fig. 6 for a training set comprising all years and calculated using the DAV radii of calculation of 250 and 500 km.

For testing purposes, all possible combinations of radii are tested to find the best performance. Table 3 shows the best-performing radii when training and testing on all 89 storms as well as when performing independent testing on each year. Overall, the RMS intensity error improves from 14.3 kt with a training radius of 300 km to 13.0 kt with a combination of 250- and 500-km training radii, and a combination could be found that improves every individual year’s performance except for 2011. However, no single combination proves to be best for all years and only a slight worsening of the RMS intensity error (by 0.1–0.2 kt) typically occurs when one of the two radii of calculation contributing to the 2D parametric surface is changed by a 50-km increment. Significance testing using a Student’s t test suggests that the results using only a slightly different parametric surface is typically not significantly different at the 90% confidence level. However, when the radii of calculation contributing to the parametric surface are more than 150 km different in one direction or 100 km in both directions, then the differences are statistically significant. This suggests that the results are robust, and small changes in the chosen
parametric surface will not alter the RMS intensity errors obtained using the parametric surface. Calculation of RMS intensity errors for the two previous example TCs using radii of 250 and 500 km yields results of 12.7 kt for Typhoon Man-Yi (compared to a DAV best performance of 14.9 kt using a single 400 km radius of calculation) and 21.2 kt for Supertyphoon Sepat (compared to a DAV best performance of 13.3 kt using a 250 km radius). Clearly for Sepat, the use of a 2D surface is an impressive failure. However, for Man-Yi, the improvement over the overall best training radius RMS intensity error of 16 kt represents a respectable improvement, and there is an overall improvement for 2007 from 13.7 to 12.6 kt, despite the degradation in Sepat’s performance.

4. Results for the eastern North Pacific basin

The two methods for intensity estimation testing discussed for the western North Pacific in section 3a are also followed in the eastern North Pacific using the NHC best-track intensities for comparison. The results for testing on the individual year 2008 are shown in Fig. 4b. Average RMS errors for intensity derived from the overall test and from each testing year using several radii for the DAV calculation are shown in Table 4.

By training and testing on all 90 storms, the overall best RMS intensity error obtained is 13.4 kt using a radius of calculation of 200 km. When individual years are withheld for testing, RMS intensity errors range from 9.4 to 16.9 kt. The overall best radius of 200 km is also the best radius for two of the seven years, and four of the seven years have a best radius of 250 km. Note that all four of these years had only a slight increase in RMS error for the 200-km-radius calculation. The outlier year, 2007, had a best radius of 300 km. This outlier year will be discussed in more detail later in this section.

Compared to the North Atlantic study, which had an overall RMS intensity error of 12.9 kt, the eastern North Pacific basin result is approximately 0.5 kt higher for similar DAV radii [section 3a; Ritchie et al. (2012)]. While not the only factor, the relative lack of aircraft reconnaissance in the eastern North Pacific compared with the North Atlantic may have contributed to the increased RMS intensity error. Furthermore, no estimation bias due to environmental vertical wind shear was observed, a result also found in the North Atlantic study.

The DAV estimates are again categorized by intensity bins in order to examine the technique’s performance against storm intensity (Table 5). The majority of the tropical depression samples are overestimated, with an

![Fig. 6. The 2D parametric surface using training data from 2007 to 2011 calculated using both a 250-km radius of calculation and a 500-km radius of calculation. To estimate intensity, the DAV for an individual sample is calculated at both radii and then the intensity is extracted from the surface.](image-url)
RMS intensity error of 10.0 kt. Recall that the parametric curve for DAV intensity estimation is constrained to 25 kt, which is partially responsible for this overestimation. The tropical storm bin has an RMS intensity error of 11.5 kt, and fewer than half (39%) of the samples are overestimated. The proportion of underestimated samples continues to increase for category 1–2 (16.6 kt) and major hurricane (26.1 kt) samples. Again, the parametric curve continues to underestimate higher intensity and overestimate lower-intensity values. The lack of major hurricane samples also contributes to this underestimation issue. As mentioned earlier, 2007 has the largest best radius for calculation, 300 km, and also has the lowest RMS intensity error of 9.4 kt. This year has no samples above category 1 intensity, which contributes to its low error value.

As a result of the error improvements in intensity estimates provided by the two-dimensional surface technique when used in the western North Pacific (section 3b), as well as the greater radius and low RMS intensity error found when testing 2007, this method is also applied to the eastern North Pacific. Again, all combinations of radii are tested on each individual year as well as on all years. Table 6 gives the results for the best combination for each year and for all years. The overall RMS intensity error improves from 13.4 kt for a radius of 200 km to 12.7 kt for a combination of 200- and 500-km radii. The individual year RMS intensity errors improve for all years except 2005 and 2006. The RMS intensity error for each of the remaining two years increases by 0.2 kt or less. As with the western North Pacific, no single combination is best for all seven years. It is perhaps not surprising that the two-radii surface had less impact in the eastern North Pacific since the size variation of TCs in this basin is lower and the close proximity of cloudiness not directly associated with the TC is less of a factor compared with the western North Pacific.

5. Summary and conclusions

In this paper, an objective technique called the deviation-angle variance technique (DAV-T), which was developed for the North Atlantic basin to estimate intensity of TCs from infrared imagery, is applied to the eastern and western North Pacific basins. We note that this application of the DAV-T represents a best case scenario for the technique. The main difference in the technique as applied to the western North Pacific is the use of MTSAT data instead of the GOES-E imagery that is used in the Atlantic basin. Also, instead of solely using GOES-E imagery in the eastern North Pacific, stitched GOES-W–GOES-E imagery is employed. Furthermore, the DAV technique is applied to a slightly larger region, encompassing 0°–40°N and 100°E–180° in the western North Pacific and 0°–40°N and 170°–80°W in the eastern North Pacific. Finally, the parametric training curve used in the previous studies to estimate intensity is extended to a two-dimensional parametric surface in order to better accommodate the larger size and structure range of TCs in the western North Pacific. This surface is also applied to the eastern North Pacific, though the resultant error reduction is much more modest in this basin. TC intensity estimation statistics are calculated for the years 2007–11 in the western North Pacific and 2005–11 in the eastern North Pacific. A comparison with statistics in the North Atlantic basin is included.

In the western North Pacific, the DAV-T shows similar results in estimating the intensity of TCs as those described in Ritchie et al. (2012) for the North Atlantic basin. A parametric sigmoid-based curve that describes the relationship between the DAV calculated from the MTSAT image and the current intensity of the TC is obtained by fitting to training data calculated from MTSAT IR images for 89 TCs from 2007 to 2011. The

### Table 5: The number of samples, RMSE (kt), and the number of samples either overestimated or underestimated for all eastern North Pacific samples categorized into bins based on the NHC’s scale of tropical cyclone intensities.

<table>
<thead>
<tr>
<th>Bin</th>
<th>No. of samples</th>
<th>RMSE (kt)</th>
<th>No. overestimated</th>
<th>No. underestimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropical depression (&lt;34 kt)</td>
<td>10254</td>
<td>10.0</td>
<td>8796 (86%)</td>
<td>1458 (14%)</td>
</tr>
<tr>
<td>Tropical storm (34–63 kt)</td>
<td>10759</td>
<td>11.5</td>
<td>4198 (39%)</td>
<td>6561 (61%)</td>
</tr>
<tr>
<td>Categories 1 and 2 (64–95 kt)</td>
<td>4122</td>
<td>16.6</td>
<td>1405 (34%)</td>
<td>2717 (66%)</td>
</tr>
<tr>
<td>Categories 3–5 (&gt;95 kt)</td>
<td>1912</td>
<td>26.1</td>
<td>213 (11%)</td>
<td>1699 (69%)</td>
</tr>
</tbody>
</table>

### Table 6: RMSE (kt) for the best combination of radii to create the 2D parametric intensity curve in the eastern North Pacific.

<table>
<thead>
<tr>
<th>Year</th>
<th>RMSE (kt)</th>
<th>Radii</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>11.4</td>
<td>200 and 500 km</td>
</tr>
<tr>
<td>2006</td>
<td>14.5</td>
<td>200 and 300 km</td>
</tr>
<tr>
<td>2007</td>
<td>8.7</td>
<td>200 and 500 km</td>
</tr>
<tr>
<td>2008</td>
<td>9.9</td>
<td>250 and 500 km</td>
</tr>
<tr>
<td>2009</td>
<td>15.4</td>
<td>250 and 500 km</td>
</tr>
<tr>
<td>2010</td>
<td>13.6</td>
<td>200 and 250 km</td>
</tr>
<tr>
<td>2011</td>
<td>15.9</td>
<td>150 and 300 km</td>
</tr>
<tr>
<td>All</td>
<td>12.7</td>
<td>200 and 500 km</td>
</tr>
</tbody>
</table>
We would also like to thank Dr. Haiyan Jiang for her preprocessed MTSAT infrared imagery for this study. We also thank Mr. Rich Bankert, and Mr. Kim Richardson from the Naval Research Laboratory, Monterey, California, who provided preprocessed MTSAT infrared imagery for this study. We would also like to thank Dr. Haiyan Jiang for her preprocessed MTSAT infrared imagery for this study.

Acknowledgments. TC best-track data were obtained from the Joint Typhoon Warning Center (http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/) and the National Hurricane Center (http://www.nhc.noaa.gov/data/#hurdat). This study has been supported by the NOPP program under Office of Naval Research Grant N00014-10-1-0146. Many thanks to Mr. Jeff Hawkins, Mr. Rich Bankert, and Mr. Kim Richardson from the Naval Research Laboratory, Monterey, California, who provided preprocessed MTSAT infrared imagery for this study.

Future work includes exploring the use of the two-dimensional parametric curve for the North Atlantic basin. In addition, the added benefit of including a variable-size DAV calculation for intensity, particularly for the western North Pacific, is being explored. While it is currently difficult to obtain reasonable size estimates of TCs from observations, we have been exploring this capability through both the use of simulated TC data (K. E. Ryan et al. 2013, unpublished manuscript), and also by developing size parameters based on the IR imagery. In addition, a key component of an intensity estimation technique is to be able to test it in an operational setting over several years in order to collect robust statistics of its performance under these conditions. To achieve this goal, there are several small steps that need to be accomplished: the 6-hourly best-track center needs to be replaced by an hourly or half-hourly real-time fix, the first 24 h of the filter for the DAV signal needs to be calculated differently, and finally we have to determine the most effective way to report the estimated intensity from the DAV technique since the half-hourly DAV-based value oscillates compared to the smooth 3- or 6-h operational estimates. Finally, because there is no aircraft reconnaissance program in the western North Pacific, the JTWC best-track archive is heavily reliant on Dvorak estimates of intensity. Similarly, infrequent aircraft reconnaissance in the eastern North Pacific also increases the reliance on Dvorak intensity estimates in this basin. Our intention is to explore methods of validating our results using periods when the best-track archives were enhanced by other observations of intensity, including those times when scatterometer data and data from special field campaigns such as the Tropical Cyclone Structure-2008 (TCS-08) and Impacts of Typhoons on the Ocean in the Pacific (ITOP) programs were available.
suggestions for some statistical analysis and also two other anonymous reviewers for their comments, which have helped to strengthen this manuscript.

REFERENCES


