Objectively Determining the Rotational Center of Tropical Cyclones in Passive Microwave Satellite Imagery

ANTHONY J. WIMMERS AND CHRISTOPHER S. VELDEN

Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin—Madison, Madison, Wisconsin

(Manuscript received 4 February 2010, in final form 1 April 2010)

ABSTRACT

Precise center-fixing of tropical cyclones (TCs) is critical for operational forecasting, intensity estimation, and visualization. Current procedures are usually performed with manual input from a human analyst, using multispectral satellite imagery as the primary tools. While adequate in many cases, subjective interpretation can often lead to variance in the estimated center positions. In this paper an objective, robust algorithm is presented for resolving the rotational center of TCs: the Automated Rotational Center Hurricane Eye Retrieval (ARCHER). The algorithm finds the center of rotation using spirally oriented brightness temperature gradients in the TC banding patterns in combination with gradients along the ring-shaped edge of a possible eye. It is calibrated and validated using 85–92-GHz passive microwave imagery because of this frequency’s relative ubiquity in TC applications; however, similar versions of ARCHER are also shown to work effectively with other satellite imagery of TCs. In TC cases with estimated low to moderate vertical wind shear, the accuracy (RMSE) of the ARCHER estimated center positions is 17 km (9 km for category 1–5 hurricanes). In cases with estimated high vertical shear, the accuracy of the ARCHER estimated center positions is 31 km (21 km for category 2–5 hurricanes).

1. Introduction

Finding the rotational center of an existing tropical cyclone (TC) is an essential step in the analysis and forecasting of the event. Studies have shown the importance of accurate initial positioning to intensity estimation methods (Olander et al. 2004; Velden et al. 2006) and operational forecasts (J. L. Franklin 2009, personal communication). In addition, numerical model track predictions can benefit from synthetic (bogus) TC observations (Goerss 2009) and/or initial TC motion vectors (Aberson 2002) that rely on accurate positioning. Given that the majority of TCs are away from sources of direct observation, the primary substantial observations are from satellites [with the exception of the aircraft reconnaissance program in the western North Atlantic Ocean (NATL) TC basin].

Storm positions are routinely estimated and provided by TC analysis and warning agencies around the globe. The primary observing tool derives from visible (Vis) and infrared (IR) imagery provided by operational geostationary satellites. Static images and animated sequences are used to “fix” a TC center position based on the structure and organization of cloud features. Often a clear and well-formed eye is apparent, making the fixes generally reliable. However, the Vis–IR methods can struggle in situations where the eye appears ragged or is partially obscured, or the TC center is totally covered by a central dense overcast. In these circumstances, passive microwave (PMW) observations from polar-orbiting satellites can play a crucial role in revealing convective organization and eyewall structure that would otherwise be obscured by cloud tops (Velden et al. 1989; Hawkins et al. 2001; Patadia et al. 2004). The 85–92-GHz observations are also more frequent and of higher resolution than other PMW observations from various platforms, and they will continue to be more frequent in the future as the current generation of polar-orbiting satellites is replaced (Velden and Hawkins 2010). However, PMW observations can also present a unique set of challenges for applications to TC center positioning. These challenges include the irregular temporal image spacing (ranging from ~30 min to ~24 h but averaging 3–6 h for the current set of operational conical scanners), comparatively low resolution (currently 6–16 km for the
with Kalman filtering in Wong et al. (2004), but Kalman filtering requires a high temporal resolution lacking in PMW observations. The PMW method described here attempts to account for these shortcomings, and combines the following features: first, it can work independently and objectively on individual images; second, it accounts for both eye patterns and any banded circulation patterns; third, it includes a “default” option to fall back to a predefined first-guess position when the detection signal is determined to be too weak; and fourth, it has been tested and refined in operational applications for several years to insure robustness.

This paper introduces a PMW center-fixing method, which we call the Automated Rotational Center Hurricane Eye Retrieval (ARCHER). Section 2 describes the methodology and components of the algorithm; section 3 details the calibration and optimization of the technique using two Atlantic TC seasons’ worth of data. Section 4 validates the method with an independent dataset of the 2005 Atlantic TC season; section 5 explores the application of ARCHER to other kinds of satellite imagery; and section 6 concludes with highlights and a discussion of the inherent issues with TC center fixing. In addition, appendix A characterizes the operational errors in TC forecasting used to estimate the expected first-guess error corresponding to a given basin; and appendix B presents a method for correcting the algorithm error values to account for uncertainties in the best-track reference dataset.

2. Center-fixing methodology

The ARCHER algorithm finds the position in the image that is centered by the spiral-patterned image edges as well as the edges of a whole or partial eye. To accomplish this, two separate components are applied. A spiral-centering algorithm finds the optimal center of a spiral pattern in the image. Second, a ring-fitting algorithm finds the edges most likely associated with the inner eyewall. The two components are complementary: the spiral component guides the algorithm toward the correct candidate for the eye (and away from false eyes), and the eye-fitting component adds precision to the more holistic spiral estimate. The formulae presented here contain many weights and other parameters. The most sensitive parameters are described and optimized in the next section. However, there are too many parameters to fully optimize independently in this study, and therefore the other less sensitive parameters were determined instead by careful trial and error and are presented in Table 1.

The preferred unit of distance in this study is great circle degrees (1 great circle degree ≈ 111.18 km). However, for
brevity, this term is shortened to “degrees.” Thus the reader should note that distance described here in degrees is independent of direction, unlike in ordinary navigation.

a. Image preprocessing

Parallax is corrected by applying a uniform adjustment to the image. The navigation of the image is shifted ~12 km (with the exact number depending on the instrument’s zenith angle) toward the satellite to compensate for the parallax of features at a height of 10 km. Of course, the height of the 85–92-GHz signal only averages 10 km, and can vary about that height, but the error caused by this difference is on the order of only a few kilometers, and is therefore minor compared to other sources of error. After correcting for parallax, the brightness temperature image is interpolated to a regular grid (evenly spaced latitude and longitude coordinates) for input into the centering algorithm.

b. First-guess position

An initial estimate of the rotational center is a necessary input to the algorithm because the algorithm must operate on a limited-size domain and because a first-guess position can serve as a useful reference point to check the validity of the algorithm result. Operationally, the ARCHER algorithm obtains its first-guess position from the most recent forecast track provided by the National Hurricane Center (NHC) for TCs in the North Atlantic and east Pacific basins, and by the Joint Typhoon Warning Center (JTWC) for TCs in other basins. However, in the calibration and validation of the ARCHER algorithm the first-guess position is set at a controlled offset distance from the NHC best-track position (section 3).

c. Large-scale method: Spiral centering

The purpose of this component is to present an objective, computational algorithm to substitute for, and hopefully improve on, the longstanding practice of subjectively fitting a spiral pattern to a satellite image of a TC. Although spiral centering is traditionally performed on visible and infrared images of tropical cyclones (Dvorak 1973; Olander and Velden 2007), the curved rainband structure observed in 85–92-GHz PMW imagery is also well suited to this method, even at the typically lower spatial resolution afforded by these frequencies. Numerically, matching a digital image of brightness temperatures with a spiral pattern is accomplished by calculating the cross product between the image gradient and a spiral-shaped unit vector field. This is substantially different from previous approaches of isolating and identifying the spiral paths of banding structures. Instead, the algorithm uses gradients throughout the image that have been shaped by the overall spiral wind field of the TC.

<table>
<thead>
<tr>
<th>Function</th>
<th>Parameter name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided coarse spiral (left), guided fine spiral (right)</td>
<td>Regridding resolution(a)</td>
<td>0.05°</td>
</tr>
<tr>
<td></td>
<td>Spiral input radius(b)</td>
<td>3.0°</td>
</tr>
<tr>
<td></td>
<td>Spiral output radius(c)</td>
<td>2.0°</td>
</tr>
<tr>
<td></td>
<td>Output spacing(d)</td>
<td>0.25°</td>
</tr>
<tr>
<td></td>
<td>Penalty function weight</td>
<td>2°</td>
</tr>
<tr>
<td>Ring score</td>
<td>Spiral score input domain threshold(g)</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Input domain buffer(h)</td>
<td>0.25°</td>
</tr>
<tr>
<td></td>
<td>Min ring radius(i)</td>
<td>0.05°</td>
</tr>
<tr>
<td></td>
<td>Max ring radius(j)</td>
<td>0.40°</td>
</tr>
<tr>
<td></td>
<td>Ring increment(m)</td>
<td>0.05°</td>
</tr>
<tr>
<td></td>
<td>Fraction of ring data(n)</td>
<td>0.425</td>
</tr>
</tbody>
</table>

\(a\) Resolution of the regridded image grid (even latitude and longitude spacing).  
\(b\) Radius of data domain input into the function.  
\(c\) Radius of the range of output points.  
\(d\) Resolution of the output grid (heavily influences computational time).  
\(e\) Weight \(w_{D1}\) from Eq. (5).  
\(f\) Weight \(w_{D2}\) from Eq. (6).  
\(g\) Uses all grid cells for which the guided fine spiral score are within this range of the maximum score.  
\(h\) Further extends the input domain by this distance.  
\(i\) Lower limit of possible ring radius.  
\(j\) Upper limit of possible ring radius.  
\(k\) This was set low to improve computational speed. In an operational setting, the value should be 1.0°.  
\(m\) Increment of ring radius.  
\(n\) Minimum number of points on the ring that must not have missing data; otherwise ARCHER is not applied.
A spiral pattern of unit vectors represented in Cartesian coordinates \((x, y)\) and centered at the origin is represented as

\[
\mathbf{S}(x, y) = \frac{\alpha x \pm y}{[(1 + \alpha^2)(x^2 + y^2)]^{0.5}} \hat{x} + \frac{\alpha y \mp x}{[(1 + \alpha^2)(x^2 + y^2)]^{0.5}} \hat{y},
\]

(1)

where \(\alpha\) is the angle of inclination in radians; and \(\mp\), \(\mp\) indicate the rotation of the spiral, either inward-clockwise in the Northern Hemisphere or inward-clockwise in the Southern Hemisphere, respectively. In a good match between the spiral pattern and a collocated image, the image gradients are strongly orthogonal to the spiral unit vectors. In practice, this formula applies well not only to the alignment of spiral bands and the eyewall inner edge, but also to the finer-scale shearing patterns caused by contrasting features elongated in the direction of the local wind. In fact, our experience from using this approach has shown that this finescale directional pattern in the image is probably a more important guide to the rotational wind field than the direction of the large-scale banding structures.

The spiral centering is performed by application of a two-step process: 1) a large-domain, coarse spiral score (CSS) function, and 2) a relatively smaller-domain fine spiral score (FSS) function. The reason for having two functions operate rather than one is that a large-domain function is effective at determining a rough estimate of the rotational center even with a very poor first-guess position, whereas a smaller domain function would be more accurate if it is provided with a good initial position estimate (supplied by the coarse step).

The formula for the coarse spiral score is the following:

\[
\text{CSS}(\phi, \theta) = c_{SS} N^{-1} \sum_{i \in \text{disk}} \| \nabla \log(I_i) \times \mathbf{S}_i(\phi, \theta) \| - c_0,
\]

(2a)

where

\[
\mathbf{S}_i(\phi, \theta) = \frac{\alpha x_i \pm y_i}{[(1 + \alpha^2)(x_i^2 + y_i^2)]^{0.5}} \hat{x} + \frac{\alpha y_i \mp x_i}{[(1 + \alpha^2)(x_i^2 + y_i^2)]^{0.5}} \hat{y},
\]

(2b)

\[x_i = (\phi_i - \phi) \cos \theta,\]

(2c)

\[y_i = \theta_i - \theta,\]

(2d)

and

\[(c_{SS}, c_0)\) are coefficients \((15, 20)\) to scale the result to a range of \(0–50)\),

\((\phi, \theta)\) are \((\text{longitude, latitude})\) in the spiral score field, \(N\) is the number of points in the sample disk, \(i\) indicates a single a point inside the sample disk, \(\text{disk}\) is the domain of \(N\) grid points in a disk centered on the initial guess position (the “sample disk”), \(I\) is the microwave horizontal polarization brightness temperature image (and the log function on \(I\) deemphasizes the extreme gradients), \(S\) is the log-5 spiral vector field,

\(\alpha\) is the inclination of the spiral vector field in radians (zero indicates a circular vector field and a log-5 field is inclined \(5^\circ\), or 0.087 radians, outward from a circular vector field),

\(\pm\) means “+” in the Northern Hemisphere and “-” in the Southern Hemisphere, and \(\mp\) means the opposite, \((\hat{x}, \hat{y})\) are the unit vectors for north and east, respectively, and

\((x_i, y_i)\) are offset distances in the \((\hat{x}, \hat{y})\) directions, respectively, where \(x_i\) is normalized to be approximately equal in spacing to \(y_i\).

This function essentially computes a score that reflects the average alignment of the image boundaries with a spiral centered at a given point \((\phi, \theta)\). This alignment is expressed in terms of the cross product of the image gradient and its collocated element of the spiral field. The vector norm function \(\| \cdot \| \) is necessary to give equal prominence to gradients directed either radially inward or outward. We use a log-5 spiral, which is less than the inclination generally associated with TC spiral patterns, because the gradients associated with horizontal shear throughout the image are generally more circular than the patterns of spiral bands. (However, the algorithm is not very sensitive to the particular value of the spiral inclination.) In practice, scores from this function are calculated at a spacing of several grid cells over a pre-determined domain of points, and a successful result produces a contoured bull’s-eye pattern focused on or near the rotational center (Fig. 1).

Next, the formula for the fine spiral score is

\[
\text{FSS}(\phi, \theta) = c_{SS} N^{-1} \sum_{i \in \text{disk}} \left\{ 0.62 \| \nabla \log(I_i) \times \mathbf{S}_i(\phi, \theta) \| - c_0 \right\} \quad \text{where} \quad \| \nabla \log(I_i) \times \mathbf{S}_i(\phi, \theta) \| > 0
\]

(3)

\[
-0.62 \| \nabla \log(I_i) \times \mathbf{S}_i(\phi, \theta) \| - c_0 \right\} \quad \text{where} \quad \| \nabla \log(I_i) \times \mathbf{S}_i(\phi, \theta) \| < 0.
\]

The components are the same as those defined in Eq. (2), except for the vector magnitude of the cross product \(|| \cdot \||\), which can be positive or negative. The difference in weights between the two parts of Eq. (3) ensures that the
inner boundary of the convective eyewall receives more weight than the outer boundary. The constant 0.62 was determined empirically to be the optimum weight for this split approach. As with the CSS, we determine the field of scores by calculating the results over a spacing of several grid points, and a successful result produces a contoured bull’s-eye pattern focused at or near the rotational center.

d. Small-scale method: Ring centering

The purpose of this method is to computationally determine the presence of an eyewall as it appears in the TC image, and thus locate the rotational center within the eye. The result is a field of values called the ring score (RS). The underlying method is simple, but more computationally expensive than the spiral method because it iterates over both the spatial domain and over a range of possible radii. The approach is to find the highest average dot product between the image gradients and a set of radially oriented unit vectors encircling a given ring pattern (a presumed eyewall inner edge). When iterated over a domain of points and a range of radii \( r \), the maximum of this formula should occur at the eye center:

\[
RSR(\phi, \theta, r) = \frac{r^{0.1}}{N} \sum_{i=\text{ring}} \left[ -\nabla (I^{1/3}) \cdot \hat{r}_i \right] \quad (4a)
\]

\[
RS(\phi, \theta) = \max[RSP(\phi, \theta, r, r)] \quad (4b)
\]

where

- \( RSR \) is ring score radius, the ring score for each possible radius,
- \( (\phi, \theta) \) are (longitude, latitude) in the ring score field,
- \( c_{RSR} \) is a constant (value of 250) to scale the result to a range of 0–100,
- \( r \) is the length of the radius being evaluated (units: degrees),
- \( N \) is the number of points in the sample ring,
- \( i \) is the index for a point on the sample ring,
- \( I^{1/3} \) is the microwave horizontal polarization brightness temperature image to the power of \( \frac{1}{3} \) (to de-emphasize the higher gradients),
- \( \hat{r}_i \) is the unit vector pointed radially outward from the center of the ring to a point of index \( i \) on the ring, and
- \( \max[RSP(\phi, \theta, r, r)] \) is the maximum of the three-dimensional field \( RSP \) over the dimension \( r \).

The purpose of the weight \( r^{0.1} \) is to slightly favor the scores corresponding to larger radii, because larger eyes are more irregular and may not match the points on a perfect circle quite as well.

Of course not every TC image contains a well-defined eye. In such events, the ring score is normally quite low. This means that unless the corresponding spiral score is very high, the ARCHER algorithm will default to the first-guess position rather than use the position with the maximum score. This process is discussed further in the next subsection.

e. Optimized combination of spiral and ring centering

By design, the spiral-centering score field and the ring-centering score field combine into a weighted sum that takes advantage of the two methods’ complementary strengths. Specifically, the ring centering yields highly
 precise results for a well-defined eye, and the spiral-centering method is important both when an eye is not evident and when the ring-centering method finds one or more false eyes (i.e., inner-core, rain-free moat regions caused by convective downdrafts or environmental dry air intrusions). In the latter case, the spiral method provides a critically important additional weight to the feature nearest the rotational center.

One extra component in the combined scores is a distance penalty field that turns spiral score fields into “guided” spiral scores [Eqs. (5) and (6)]. The distance penalty is a field that varies with the square of the distance from the initial guess. Therefore it is minor near the first-guess position and dominant at the edges of the analysis domain. This acts as a check to insure that no matter how unusual the image, the final result will not stray unreasonably far from the first-guess position.

The scores are combined in a multistep algorithm as follows:

1) **STEP 1: GUIDED COARSE SPIRAL (GCS)**

\[
\text{GCS}(\phi, \theta, \phi_0, \theta_0) = \text{CSS}(\phi, \theta) - w_{\text{D1}} \text{dist}^2(\phi, \theta, \phi_0, \theta_0),
\]

where \((\phi_0, \theta_0)\) is the initial guess position and \text{dist}^2(\phi, \theta, \phi_0, \theta_0) is the square of the distance between the initial guess position and the point \((\phi, \theta)\) on the domain. The score field for GCS then sets the stage for the next step in the algorithm.

2) **STEP 2: GUIDED FINE SPIRAL (GFS)**

\[
\text{GFS}(\phi, \theta, \phi_0, \theta_0) = \text{FSS}(\phi, \theta) - w_{\text{D2}} \text{dist}^2(\phi, \theta, \phi_0, \theta_0).
\]

This is the gridded result used to calculate the overall spiral component of the combined score. The domain for step 2 is the area of points that are within a certain range of the maximum value of the GCS. The area is also expanded by a small amount. (The parameter values for this process are provided in Table 1).

3) **STEP 3: RING SCORE**

Calculate the RS from Eq. (4).

4) **STEP 4: COMBINED SCORE**

The combined score (CS) normally generates a gridded field that contours to a bull’s-eye pattern around the maximum value. For simplicity, the position of this maximum value is hereafter called the target position. The formula for this score is a weighted sum of the guided fine spiral (step 2) and the ring score (step 3). The maximum value of this gridded score field is then used to determine the target position and the final ARCHER position (demonstrated in Fig. 2):

\[
\text{CS}(\phi, \theta, \phi_0, \theta_0) = w_{\text{GFS}} \text{GFS}(\phi, \theta, \phi_0, \theta_0) + \text{RS}(\phi, \theta)
\]

\[
\text{MCS} = \max[\text{CS}(\phi, \theta, \phi_0, \theta_0)]
\]

\[
(\phi_f, \theta_f) = \begin{cases} (\phi_t, \theta_t) & \text{where } \text{CS}(\phi_t, \theta_t, \phi_0, \theta_0) = \text{MCS} \quad \text{if } \text{MCS} \geq \tau, \\ (\phi_0, \theta_0) & \text{if } \text{MCS} < \tau. \end{cases}
\]
where the scalar value MCS is the maximum combined score, \((\phi_f, \theta_f)\) is the target position, \((\phi_t, \theta_t)\) is the final ARCHER position, \((\phi_{0t}, \theta_{0t})\) is the default first-guess position, and \(\tau\) is the MCS threshold value for either applying the target position or defaulting to the first-guess position. This arrangement has two unresolved parameters: \(w_{GFS}\) and \(\tau\). These two parameters are the most naturally empirical of all the algorithm parameters, with \(w_{GFS}\) determining the relative importance of the spiral score and the ring score, and \(\tau\) determining the optimal distinction between effective and ineffective performance. Thus the next section is devoted to optimizing these parameters empirically, with special attention to how these values vary according to TC intensity.

### 3. Empirical calibration

The objective of this section is to find the optimum values of two parameters: the relative weight of the spiral score to the ring score and the optimum default threshold for the maximum combined score.


The calibration dataset is selected from microwave imagery of TCs during the 2007 and 2008 North Atlantic storm seasons (Table 2). Taken together, the data from these two years make up a representative sample of a wide variety of TC structures intensities. The “truth” position of the rotational center for each case is a cubic interpolation of the NHC best-track dataset. To insure that the truth dataset is as accurate as possible, only the best-track values \(\pm 3\) h from an aircraft reconnaissance position fix are used. This is the most limiting constraint on the sample size. Additional criteria for the microwave observations are the following: the best-track position must lie inside the satellite swath (allowing for partial coverage cases if the coverage of the TC exceeds 50%); all storms are required to be south of 40°N latitude to serve as a coarse exclusion of extratropical transition cases; and all observed TCs must have a surface wind speed greater than or equal to 34 kt (1 kt \(\approx 0.5\) m s\(^{-1}\)) because organized rotation is rarely evident in PMW imagery below tropical storm classifications.

It is desirable to partition the data sample by cases showing the effects of “weak vertical shear” and “strong vertical shear” because the displacement between surface center of rotation and image–height center of rotation can lead to difficulties in the application of the algorithm. Just as importantly, vertical shear can adversely effect the organization of the TC and disrupt the conventional spiral-eye configuration that makes the algorithm successful. Such a disruption naturally complicates the process of center-fixing. Unfortunately, systematic effects of vertical shear on vortex tilt and rotational asymmetries are not well understood or documented to the point that any particular static environmental analysis can serve as an effective measure to adequately stratify the sample. For example, the period of time a shear regime has been acting on a vortex might be just as important as the strength of the shear or the TC intensity at the analysis time. Consequently, several attempts to use objective environmental vertical shear statistics to partition the data did not yield very effective results.

However, it can be shown that the zonal translation velocity of the TC, which is easily extracted from the best track, can serve as a useful proxy for distinguishing TCs that are well-suited to the ARCHER algorithm from those in which vortex tilt and asymmetric structure are more likely to pose problems. Figure 3 shows that algorithm error increases sharply where TC eastward translation speeds \((V_x)\) exceed 5 kt. This increase in error is predominantly caused by recurvature during which stronger westerly wind shear regimes can often quickly transform an organized vortex into more distorted asymmetric and/or tilted structures. Consequently, the calibration and validation steps partition the observations into TCs in which \(V_x < 5\) kt (group A, including all westward-moving storms) and \(V_x \geq 5\) kt (group B). Only group A TCs are used in the calibration, because our investigation finds that the ARCHER algorithm is effective in only a minority of group B cases.

---

### Table 2. Satellite microwave instruments used in calibration and validation of the ARCHER algorithm.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Instrument</th>
<th>Frequency (GHz)</th>
<th>Orbit</th>
<th>Footprint (km)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSP-13</td>
<td>SSM/I</td>
<td>85.5, H(^*)</td>
<td>Polar, sun-synchronous</td>
<td>16 \times 14</td>
<td>Raytheon (2000)</td>
</tr>
<tr>
<td>TRMM</td>
<td>TMI</td>
<td>85.5, H</td>
<td>Equatorial, between 38°S and 38°N</td>
<td>7 \times 5</td>
<td>Kummerow et al. (1998)</td>
</tr>
<tr>
<td>Aqua</td>
<td>AMSR-E</td>
<td>89.0, H</td>
<td>Polar, sun-synchronous</td>
<td>6 \times 4</td>
<td>NASA MSFC (2001)</td>
</tr>
</tbody>
</table>

\(^*\) H: horizontal polarization.
validation of the ARCHER algorithm in group B cases is handled in a separate analysis.

b. Method

We approximate the real-time distribution of forecast error on the dataset by simulating 12 scenarios per image of a first-guess position displaced to varying offsets from the best-track position: 0.1°, 0.4°, and 0.7° in the four cardinal directions. Thus, the total number of cases (3972) in the calibration is the number of microwave images in the calibration dataset (331) times the number of scenarios per image (12). The statistics from these cases (such as RMS) are then combined into weighted averages to approximate the algorithm performance for the North Atlantic operational environment, which has a certain probability of small (0.1°), medium (0.4°), and large (0.7°) forecast (first guess) position errors. The weights have been determined according to the observed distribution of forecast position error in the North Atlantic operational environment (appendix A).

The guided fine spiral and ring score fields are calculated for each case, and the final positions determined for a range of spiral weights ($w_{GFS}$) and score threshold values ($\tau$). The errors of the final ARCHER positions form the basis of the performance statistics. With a set of statistics for the algorithm over this range of weights and thresholds, we can then judge the performance in order to identify the optimal weight–threshold combination, based on the following criteria:

1) RMS: Minimum possible RMS position error between the algorithm centers and the truth (best track) values.

2) Number of excessive errors: For all the scenarios, no case should have an algorithm position error exceeding the largest first-guess error in the simulation (0.7°).

3) Target usage: When presented with a range of parameter values that produce similar accuracy statistics, the value with the fewest defaults to first-guess positions will be selected. This is desirable because some amount of uncertainty in the application of the best-track dataset is inevitable; therefore the algorithm error is somewhat overstated by the statistics.

Several iterations of experimentation showed that three groupings of cases based on TC maximum surface wind speed ($V_{max}$) required separate weight–threshold combinations: tropical storm ($34 \leq V_{max} < 65$ kt), category 1 hurricane ($65 \leq V_{max} < 84$ kt), and category 2–5 hurricane ($V_{max} \geq 84$ kt). Counter to our expectations, the data organized itself in this “trimodal” pattern rather than any sort of continuous evolution with increasing $V_{max}$. The final results are organized in this fashion and the weight–threshold combinations optimized accordingly.

c. Results and discussion

Figure 4 presents the results of all the cases for group A TCs in the 2007–08 dataset. The results are organized to illustrate the optimal weight–threshold combination for each of the three $V_{max}$ groupings. However, one important note is that the threshold value is displayed as the scaled threshold:

$$\tau \text{ (scaled)} = \tau (1 + w_{GFS})^{-1}. \quad (8)$$

This conversion keeps the plotted trends reasonably “flat” as the spiral weight varies. Otherwise, the thresholds of interest would increase linearly (and out of scale) with the spiral weights.

Each row addresses one of the three criteria for parameter evaluation, weighted by the relative frequency of 0.1°, 0.4°, and 0.7° displacements in the North Atlantic (Table A1, described below in appendix A). The first row is the RMS error of the ARCHER position with respect to the best track. The second row of Fig. 4 identifies the number of excessive errors, indicated by the (weighted) number of cases in which the errors exceed the largest offset distance used in the simulations (this can be a fractional value lower than 1 because of the weighting scheme). The third row in this figure is the fraction of cases in which the algorithm position was applied, as opposed to defaulting to the first guess. Finally, the fourth row plots the distribution of distances between the best track and the ARCHER result. This distribution is not weighted because a non-weighted version is easier to read and interpret.
For each column, a magenta polygon outlines the area in weight–threshold space that achieves the best compromise between the three criteria: low RMS, few excessive errors, and a high fraction of applied targets. The values in these polygons show that one should expect the algorithm performance to be essentially the same throughout the range of weights and thresholds therein. However, a specific weight and threshold value must be used in the validation, and so a specific point is selected for each group (marked with a white circle), based on the robustness of this point in weight–threshold space. The values corresponding to these points are presented in Table 3.

4. Algorithm validation

The purpose of this section is to characterize the accuracy of the ARCHER algorithm with a dataset that is independent of the calibration. The results are meant to provide a record of performance to help users take the
TABLE 3. Calibrated weight and threshold values for the ARCHER algorithm.

<table>
<thead>
<tr>
<th>TC intensity (kt)</th>
<th>( w_{\text{GPS}}^a )</th>
<th>( \tau ) (scaled)(^b)</th>
<th>( \tau ), actual(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 34 \leq V_{\text{max}} &lt; 65 )</td>
<td>14.4</td>
<td>17.4</td>
<td>268</td>
</tr>
<tr>
<td>( 65 \leq V_{\text{max}} &lt; 84 )</td>
<td>14.4</td>
<td>13.6</td>
<td>209</td>
</tr>
<tr>
<td>( V_{\text{max}} \geq 84 )</td>
<td>38.0</td>
<td>8.0</td>
<td>312</td>
</tr>
</tbody>
</table>

\(^a\) Weight on the guided fine spiral score [Eq. (7)].
\(^b\) Scaled threshold value [Eq. (8)].
\(^c\) Threshold value [Eq. (7)].

The very active 2005 hurricane season of the North Atlantic was chosen for the independent validation dataset, with 241 separate 85–92-GHz PMW observations. The constraints for data quality were the same as with the calibration dataset: all observations must be \( \pm 3 \) h from an aircraft reconnaissance position fix, the best-track position must be inside the satellite swath, TCs must be south of 40°N latitude, and TCs must have maximum surface wind speeds greater than 34 kt. Of these 241 observations, approximately 40% are tropical storms, 20% are category 1 hurricanes, and 40% are category 2–5 hurricanes.

c. Algorithm performance for group A TCs

To interpret the performance of the algorithm, the results are first partitioned into a matrix of histograms (Fig. 5) that breaks the cases down by TC intensity (rows) and first-guess displacement (columns). A few trends and features in Fig. 5 are immediately apparent. The consequence of the algorithm defaulting to the first guess is a peak in the histogram at 0.1°, 0.4°, 0.7°, or 1.0°. The number of defaults decreases with increasing TC intensity, with defaults predominant in tropical-storm-strength TCs and almost nonexistent with category 2–5 hurricanes. Each group of TC intensities (the top three rows) shows a different feature that dominates the distribution of error. For tropical storms, the likelihood of a default is high for all values of first-guess position displacement. For category 1 hurricanes, the number of defaults increases with first-guess displacement (increasing from left to right), showing that the ARCHER score for the target center is more likely to fall below the threshold as the accuracy of the first-guess decreases. For category 2–5 hurricanes, the peak of the error distribution is unchanged with the accuracy of the first guess, and the shape remains similar except that the tail of the histogram grows slightly as the accuracy of the first-guess decreases.

Another significant feature of the category 2–5 hurricane row of histograms is that the error peaks at 0.05° rather than 0, as one might expect in a grouping where a majority of cases have well-defined eyes. This suggests that a portion of the error in the ARCHER product is due to an “intrinsic error” in the correspondence between the PMW images and the best-track dataset, which causes a departure of the best-track positions from the true center indicated by the image. Several factors contribute to the difference between best-track center and the center of rotation at the level of the microwave retrieval in a given image: vortex tilt caused by vertical shear or vortex precession, residuals in parallax correction, satellite navigational error, and best-track error.

This knowledge is important to consider in applications that use the ARCHER result because most applications (such as intensity estimation) require the image-relative center rather than the center of rotation at the surface. Therefore, it is highly useful to quantify this type of intrinsic error between the best track and the image-relative center if at all possible. Appendix B provides an estimation of the intrinsic error and shows...
that it can be approximated as an additive component of the ARCHER product’s variance with respect to the best track. In the following analysis, this approximation is used to complement the standard RMS error (relative to the best track) with a corresponding image-relative RMS error value.

Table 4 presents these complementary statistics. Results are sorted in columns by first-guess displacement, and then by estimates for the values in a North Atlantic basin environment in the fifth column and all other basins in the sixth column. The statistic “percentage applied” (percentage of ARCHER target positions applied rather than defaulting) varies greatly by TC intensity. For tropical storms, only about 17% of the cases use the ARCHER target position; whereas, about 84% of category 1 cases and 99.9% of the category 2–5 cases use the ARCHER target position. The statistic “percentage worsened” (percentage of cases in which the final position was in error by greater than 0.71° for first-guess displacements of 0.1°–0.7° and greater than 1.1° for first-guess displacements of 1.0°) was very low: less than 1% in all relevant groupings. (The value was 1% for one grouping when the first-guess displacement is 1.0°, which is already a very rare circumstance operationally.)

Trends in image-relative RMS are best explained visually (Fig. 6). The RMS of the ARCHER target positions (left plot) is encouragingly small (<0.05°) for category 1 and category 2–5 hurricanes when the first-guess
displacement is small (0.1°). The RMS for these two groups increases at roughly the same rate with increasing first-guess displacement. This increase with displacement is of the same order of magnitude as the RMS at the lowest first-guess displacement, showing that the sensitivity of the ARCHER algorithm to the first-guess displacement is significant but does not reduce the overall performance appreciably, because first-guess displacements of 0.4° and greater are considerably less common. However, the ARCHER target RMS for tropical-storm-strength TCs is significantly higher. The RMS at 0.1° does not even exceed accuracy of the default displacement. However, Table 4 shows that the tropical storm ARCHER target RMS for the North Atlantic basin (0.171°) and other

<table>
<thead>
<tr>
<th>TC intensity (kt)</th>
<th>Statistic</th>
<th>0.1°</th>
<th>0.4°</th>
<th>0.7°</th>
<th>1.0°</th>
<th>NATL basin</th>
<th>Other basins</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 ≤ V_max &lt; 65 (tropical storm)</td>
<td>No. cases</td>
<td>331.0</td>
<td>325.0</td>
<td>324.0</td>
<td>320.0</td>
<td>329.7</td>
<td>329.3</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>16.6</td>
<td>17.2</td>
<td>15.7</td>
<td>12.5</td>
<td>16.7</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.9</td>
<td>0.6</td>
<td>0.3</td>
<td>0.3</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.227</td>
<td>0.239</td>
<td>0.266</td>
<td>0.252</td>
<td>0.230</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.130</td>
<td>0.377</td>
<td>0.651</td>
<td>0.940</td>
<td>0.227</td>
<td>0.254</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.166</td>
<td>0.183</td>
<td>0.216</td>
<td>0.200</td>
<td>0.171</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.114</td>
<td>0.372</td>
<td>0.648</td>
<td>0.938</td>
<td>0.218</td>
<td>0.246</td>
</tr>
<tr>
<td>65 ≤ V_max &lt; 84 (category 1)</td>
<td>No. cases</td>
<td>178.0</td>
<td>168.0</td>
<td>168.0</td>
<td>162.0</td>
<td>175.9</td>
<td>175.2</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>84.8</td>
<td>84.5</td>
<td>80.4</td>
<td>74.1</td>
<td>84.6</td>
<td>84.6</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.0</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.160</td>
<td>0.175</td>
<td>0.191</td>
<td>0.194</td>
<td>0.164</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.152</td>
<td>0.225</td>
<td>0.354</td>
<td>0.536</td>
<td>0.176</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.043</td>
<td>0.082</td>
<td>0.112</td>
<td>0.118</td>
<td>0.055</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.055</td>
<td>0.175</td>
<td>0.326</td>
<td>0.519</td>
<td>0.102</td>
<td>0.115</td>
</tr>
<tr>
<td>V_max ≥ 84 (category 2–5)</td>
<td>No. cases</td>
<td>312.0</td>
<td>308.0</td>
<td>304.0</td>
<td>303.0</td>
<td>311.0</td>
<td>310.7</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>100.0</td>
<td>99.7</td>
<td>99.7</td>
<td>99.0</td>
<td>99.9</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.123</td>
<td>0.137</td>
<td>0.149</td>
<td>0.199</td>
<td>0.126</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.123</td>
<td>0.138</td>
<td>0.154</td>
<td>0.221</td>
<td>0.127</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.063</td>
<td>0.087</td>
<td>0.105</td>
<td>0.168</td>
<td>0.069</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.063</td>
<td>0.090</td>
<td>0.113</td>
<td>0.195</td>
<td>0.069</td>
<td>0.072</td>
</tr>
<tr>
<td>V_max ≥ 65 (category 1–5)</td>
<td>No. cases</td>
<td>490.0</td>
<td>476.0</td>
<td>472.0</td>
<td>465.0</td>
<td>486.9</td>
<td>485.9</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>94.5</td>
<td>94.3</td>
<td>92.8</td>
<td>90.3</td>
<td>94.4</td>
<td>94.4</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.9</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.136</td>
<td>0.150</td>
<td>0.163</td>
<td>0.197</td>
<td>0.139</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.134</td>
<td>0.174</td>
<td>0.245</td>
<td>0.363</td>
<td>0.146</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.053</td>
<td>0.082</td>
<td>0.104</td>
<td>0.152</td>
<td>0.061</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.056</td>
<td>0.124</td>
<td>0.213</td>
<td>0.343</td>
<td>0.080</td>
<td>0.087</td>
</tr>
<tr>
<td>V_max ≥ 34</td>
<td>No. cases</td>
<td>821.0</td>
<td>801.0</td>
<td>796.0</td>
<td>785.0</td>
<td>816.6</td>
<td>815.2</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>63.1</td>
<td>63.0</td>
<td>61.4</td>
<td>58.6</td>
<td>63.0</td>
<td>63.0</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.6</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.148</td>
<td>0.162</td>
<td>0.177</td>
<td>0.203</td>
<td>0.152</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.133</td>
<td>0.275</td>
<td>0.456</td>
<td>0.662</td>
<td>0.183</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.055</td>
<td>0.086</td>
<td>0.111</td>
<td>0.149</td>
<td>0.064</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.075</td>
<td>0.253</td>
<td>0.443</td>
<td>0.653</td>
<td>0.146</td>
<td>0.165</td>
</tr>
</tbody>
</table>

a Percentage of cases in which the ARCHER target position is applied.
b Percentage of cases in which the error exceeds 0.71° (or 1.0° if the first-guess displacement is 1.0°).
c RMS (distance in degrees) of “applied” cases (cases in which ARCHER target position is applied).
d RMS (distance in degrees) of all cases (both “applied” and default).
e Image-relative RMS for “applied” cases.
f Image-relative RMS for all cases.
g Calculated by applying the weights of Table A1 for NATL basin.
h Calculated by applying the weights of Table A1 for other basins.
basins (0.173°) is lower on the whole than if the default value were used every time.

The right side of Fig. 6 shows that the percentage of defaults in the ARCHER algorithm is still significant for tropical storm and category 1 hurricanes. However, it is important to note that the frequency of occurrence of these cases decreases sharply as the displacement increases, and so these errors at displacements of 0.4 and 0.7 have only a small impact on the general algorithm performance.

Table 4 \((V_{\text{max}} \geq 34, \text{All})\) provides the aggregate results for the ARCHER algorithm in the right two columns. Overall, the image-relative RMS of the applied ARCHER targets is 0.064° (7.1 km) for the North Atlantic basin and 0.067° (7.4 km) for all other basins. When including the cases in which ARCHER must default to the first-guess, the ARCHER image-relative RMS is 0.146° (16.2 km) for the North Atlantic basin and 0.165° (18.3 km) for all other basins. However, when considering only category 1–5 hurricanes (fourth grouping in Table 4), the overall performance is considerably better, with an image-relative RMS of 0.080° (8.8 km) in the North Atlantic and 0.087° (9.7 km) in the other basins.

d. Algorithm performance for group B TCs

The 2005 North Atlantic observations of group B TCs (eastward translation speed \(V_x \geq 5 \text{ kt}\)) verified by aircraft reconnaissance do not provide a large enough sample to perform a statistically significant validation (45, 8, and 45 cases for tropical storm, category 1, and category 2–5 hurricanes, respectively). Instead, we use a combined dataset from the 2005, 2007, and 2008 seasons to validate the ARCHER algorithm for these cases. This does not violate the independence of the calibration and validation datasets, because the group B cases for 2007–08 were not included in the calibration.

Figure 7 shows a histogram matrix corresponding to the group B validation; note that the x axes have changed in scale. In this figure, both the category 1 and category 2–5 hurricane rows show peaks in their distributions at 0.15°–0.20°, which is greater than the position of the peaks for group A storms. This suggests two possible
factors: 1) a larger intrinsic error, or 2) a systematic error in resolving the center of group B storms caused by possible effects from high-shear–tilt in extratropical transition events to be expected in this grouping. Our examination of all group B images for category 2–5 storms shows that both factors are likely in effect.

Throughout the group B sample we also observe a common, alternative mode of convective organization primarily responsible for the algorithm’s systematic error here, which we will refer to as the “hook” pattern. This systematic pattern appears to contribute significantly the histogram peaks at 0.15°–0.20°. Figure 8 shows a typical example of a hook pattern. In this image, convection is dominant in the north, curving cyclonically inward toward the best-track center of rotation. However, the guided fine spiral algorithm finds a center of the curvature to the south, indicated by low-level gradients and the other banding structure. Furthermore, the absence of any clear eye leads to a weak ring score field. Thus the guided fine spiral score field dominates in the combined result and the final position falls farther away from the best track than does the first guess. This may appear to be easily remedied by weighting the spiral score less, or setting the threshold higher, but each of these options tend to increase the error in other cases—for example, by choosing the wrong eye pattern or unnecessarily defaulting to the first guess.

Table 5 provides the numerical results of the validation for group B TCs. Statistics from this table show that for group B, the ARCHER target position is used less often for tropical storms and category 1 hurricanes than it is used for group A cases of corresponding intensities.
Although the target position is used in more than 50% of the cases of the category 1 grouping, the percentage of cases in which the algorithm error exceeds the 0.71° threshold (% worsened) is substantial: ~10%. Judging by inspection of the whole dataset, the largest displacements between best-track and the ARCHER target positions occur during convective hook patterns, whereas cases with coherent eyewalls are still resolved accurately.

5. Applications with other satellite imagery

Although ARCHER was designed for and calibrated to operate on 85–92-GHz microwave brightness temperatures, the algorithm can be easily adapted to imagery from other satellite instruments. The following examples demonstrate a similarly robust application of the ARCHER algorithm to longwave infrared and 37-GHz passive microwave imagery.

a. Geostationary longwave infrared

It is desirable to investigate the possibility of ARCHER working effectively on geostationary infrared (IR) imagery, because the temporal availability is much greater than that of PMW. TCs often share similar structural characteristics viewed in the IR as with the 85–92-GHz PMW imagery: curved gradients throughout the areas of convective banding, and an eye with a local brightness temperature maximum. The image values also vary in the same direction: convection causes lower brightness temperatures and subsidence causes higher brightness temperatures. The two main differences are that infrared gradients tend to be less sharp but more widespread (because IR gradients generally follow contours in cloud height rather than edges of convection–precipitation) and the eye is more often obscured. Fortunately, the ARCHER algorithm can operate well in these environments because weak gradients can have a significant effect on the spiral scores, and also, the combination of spiral scores and ring scores give the algorithm a measure of resilience to obstructions in the eye.

Figure 9 shows the performance of the ARCHER algorithm on Geostationary Operational Environmental Satellite (GOES-12) IR image during an intensifying stage of Hurricane Katrina (2005) in the Gulf of Mexico. This scenario, the first-guess position is located 0.7° (77 km) to the west. The guided fine spiral score is affected only by the gradients of the eyewall and the minor gradients caused by rotation in the surrounding cloud cap. The ring score is at a maximum in the center of the eye, and is not adversely affected by the clouds in the southwest portion of the eye. The final result is in excellent agreement with the best-track position (within 0.05°, or 5 km).

The ARCHER performance is also good when applied to major hurricanes with no discernable eye patterns. In Fig. 10, the first-guess position is 0.7° (77 km) to the west of the best-track position, yet the spiral score is still guided to the correct center. The strong spiral score in this scenario clearly dominates the combined score, leading to a center estimate that is 0.16° (18 km) away from the best-track center.

b. The 37-GHz brightness temperature

The imagery from 37-GHz microwave radiometers, particularly at the high resolution of the AMSR-E instrument, has a number feature-tracking characteristics found in both 85–92-GHz imagery and IR imagery. This channel is sensitive primarily to microwave emission by water droplets, so the image shows the distribution of low-level precipitation in the TC. Accordingly, the image...
is unobstructed by high clouds as in the 85–92-GHz imagery. However, the 37-GHz image also shows a wider distribution of convective events than the 85–92-GHz image and the convective bands are similar in appearance to those in an IR image. To apply ARCHER to the horizontal polarization of 37-GHz imagery, two simple changes are necessary because the gradients are generally of the opposite sign: first, reversing the weights on positive and negative gradient cross products of the guided fine spiral score [Eq. (3)]; and, second, reversing the sign in the ring score summation term [Eq. (4)]. These changes would not be necessary on 37-GHz polarization corrected temperature (PCT) imagery, but we apply the algorithm to the horizontal polarization instead because of its better texture.

Figures 11 and 12 show the ARCHER algorithm applied to two AMSR-E 37-GHz horizontal polarization images of Hurricane Katrina (2005) in the Gulf of Mexico. The first-guess position is located 0.78 (77 km) west of the best-track position. Both cases show ample distributions of spiral-oriented edges in the convective bands to guide the spiral scores to the rotational center. The eye is also evident in both images even though it is not apparent in the corresponding IR images. This leads the algorithm toward an accurate ring score center as well. In the final result, the ARCHER positions are

<table>
<thead>
<tr>
<th>TC intensity (kt)</th>
<th>Statistic</th>
<th>0.1°</th>
<th>0.4°</th>
<th>0.7°</th>
<th>1.0°</th>
<th>NATL basin</th>
<th>Other basins</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 ≤ ( V_{\text{max}} &lt; 65 ) (tropical storm)</td>
<td>No. cases</td>
<td>73.0</td>
<td>75.0</td>
<td>75.0</td>
<td>76.0</td>
<td>73.4</td>
<td>73.6</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>1.4</td>
<td>2.7</td>
<td>2.7</td>
<td>2.6</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.390</td>
<td>0.229</td>
<td>0.253</td>
<td>0.226</td>
<td>0.362</td>
<td>0.353</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.109</td>
<td>0.396</td>
<td>0.692</td>
<td>0.987</td>
<td>0.227</td>
<td>0.258</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.349</td>
<td>0.150</td>
<td>0.184</td>
<td>0.145</td>
<td>0.318</td>
<td>0.308</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.107</td>
<td>0.395</td>
<td>0.691</td>
<td>0.987</td>
<td>0.228</td>
<td>0.259</td>
</tr>
<tr>
<td>65 ≤ ( V_{\text{max}} &lt; 84 ) (category 1)</td>
<td>No. cases</td>
<td>70.0</td>
<td>71.0</td>
<td>69.0</td>
<td>70.0</td>
<td>70.2</td>
<td>70.2</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>55.7</td>
<td>54.9</td>
<td>50.7</td>
<td>44.3</td>
<td>55.4</td>
<td>55.3</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>10.0</td>
<td>12.7</td>
<td>10.1</td>
<td>8.6</td>
<td>10.5</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.529</td>
<td>0.506</td>
<td>0.542</td>
<td>0.529</td>
<td>0.525</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.400</td>
<td>0.461</td>
<td>0.625</td>
<td>0.825</td>
<td>0.420</td>
<td>0.427</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.499</td>
<td>0.475</td>
<td>0.513</td>
<td>0.499</td>
<td>0.495</td>
<td>0.494</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.379</td>
<td>0.443</td>
<td>0.612</td>
<td>0.817</td>
<td>0.398</td>
<td>0.406</td>
</tr>
<tr>
<td>( V_{\text{max}} \geq 84 ) (category 2–5)</td>
<td>No. cases</td>
<td>105.0</td>
<td>102.0</td>
<td>101.0</td>
<td>99.0</td>
<td>104.3</td>
<td>104.1</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>100.0</td>
<td>99.0</td>
<td>99.0</td>
<td>98.0</td>
<td>99.8</td>
<td>99.7</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>0.0</td>
<td>1.0</td>
<td>2.0</td>
<td>7.1</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.256</td>
<td>0.262</td>
<td>0.307</td>
<td>0.435</td>
<td>0.259</td>
<td>0.260</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.256</td>
<td>0.264</td>
<td>0.314</td>
<td>0.453</td>
<td>0.259</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.188</td>
<td>0.197</td>
<td>0.254</td>
<td>0.399</td>
<td>0.192</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.188</td>
<td>0.200</td>
<td>0.262</td>
<td>0.420</td>
<td>0.192</td>
<td>0.194</td>
</tr>
<tr>
<td>( V_{\text{max}} \geq 65 ) (category 1–5)</td>
<td>No. cases</td>
<td>175.0</td>
<td>173.0</td>
<td>170.0</td>
<td>169.0</td>
<td>174.5</td>
<td>174.3</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>82.3</td>
<td>80.9</td>
<td>79.4</td>
<td>75.7</td>
<td>82.0</td>
<td>81.8</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>4.0</td>
<td>5.8</td>
<td>5.3</td>
<td>7.7</td>
<td>4.4</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.351</td>
<td>0.348</td>
<td>0.382</td>
<td>0.459</td>
<td>0.351</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.321</td>
<td>0.358</td>
<td>0.466</td>
<td>0.634</td>
<td>0.333</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.306</td>
<td>0.301</td>
<td>0.341</td>
<td>0.425</td>
<td>0.306</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.280</td>
<td>0.323</td>
<td>0.439</td>
<td>0.616</td>
<td>0.293</td>
<td>0.298</td>
</tr>
<tr>
<td>( V_{\text{max}} \geq 34 ) (all)</td>
<td>No. cases</td>
<td>248.0</td>
<td>248.0</td>
<td>245.0</td>
<td>245.0</td>
<td>247.9</td>
<td>247.9</td>
</tr>
<tr>
<td></td>
<td>Percentage applied</td>
<td>58.5</td>
<td>57.3</td>
<td>55.9</td>
<td>53.1</td>
<td>58.2</td>
<td>58.1</td>
</tr>
<tr>
<td></td>
<td>Percentage worsened</td>
<td>2.8</td>
<td>4.0</td>
<td>3.7</td>
<td>5.3</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>RMS (applied)</td>
<td>0.352</td>
<td>0.346</td>
<td>0.381</td>
<td>0.457</td>
<td>0.351</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>RMS (all)</td>
<td>0.276</td>
<td>0.370</td>
<td>0.545</td>
<td>0.762</td>
<td>0.306</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td>Img RMS (applied)</td>
<td>0.306</td>
<td>0.300</td>
<td>0.339</td>
<td>0.422</td>
<td>0.306</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>Img RMS (all)</td>
<td>0.243</td>
<td>0.346</td>
<td>0.529</td>
<td>0.751</td>
<td>0.275</td>
<td>0.287</td>
</tr>
</tbody>
</table>

a Abbreviations explained in Table 4.
accurate to 0.11° (11 km) in Fig. 11 and 0.05° (5 km) in Fig. 12. Part of the reason for such close agreement is likely the fewer opportunities for intrinsic error because it represents an estimate that is much closer to the surface than that of the 85–92-GHz retrieval.

6. Conclusions

ARCHER is a widely applicable objective TC center-finding algorithm with an ability to distinguish spiral and eye patterns in satellite images and estimate the rotational center with high accuracy. Its ancillary inputs require only the information that is freely available in real time so that it can run in a fully automated mode. The algorithm shows considerable skill for center-fixing most TCs, and accuracy increases with TC intensity. For less organized TCs, the algorithm is not as accurate because of the high proportion of convective patterns that do not have coherent eyewalls and corresponding spiral bands. When using this algorithm operationally, it is important to interpret the result as the estimated rotational center derived from the retrieval height and navigation of the image signal rather than at the surface, although some corrections for parallax are possible. Tables 4 and 5 provide measures of the algorithm accuracy that will aid in this interpretation. When applied to less organized storms, the interpretation of the result depends heavily on the type of convective structure. In particular, eye-spiral structures lead to the most accurate ARCHER target positions, while axially asymmetric hook structures lead to the least accurate ARCHER position estimates.

Overall, the skill of this algorithm is limited by the fraction of TCs that do not conform, even partially, to the spiral-eye mode of convective organization. In such
cases the outstanding problem is that the structures in the image do not align with the actual surface rotational center (i.e., highly sheared or tilted vortices). A possible remedy would be to develop an ancillary technique based on other information that identifies these vortex structures in order to apply the ARCHER algorithm with customized parameters and thresholds. Beyond that, a significantly greater improvement in center-fixing accuracy would probably require the management of multisatellite information and perhaps the inclusion of numerical model information as well.

Acknowledgments. This work was sponsored by the Oceanographer of the Navy through the PEO C4I PMW-150 program office and the Naval Research Laboratory (Jeff Hawkins). We thank Derrick Herndon of CIMSS for collating the best-track and flight reconnaissance data, and we thank the three reviewers for their constructive comments.

APPENDIX A

Characterization of Errors in Operational Tropical Cyclone Position Forecasts

The development of a PMW center-fixing technique requires that the algorithm be calibrated using a first-guess position estimate with an error profile matching that of the forecast positions typical in real-time applications. The most widely available and reliable first-guess TC position estimates are from the official forecast track updated regularly by the National Hurricane Center for TCs in the North Atlantic and east Pacific basins, and by the Joint Tropical Warning Center for TCs in

Fig. 11. ARCHER diagnostic images for a 37-GHz image: 1842 UTC 26 Aug 2005 case of Hurricane Katrina. Background images are from the AMSR-E conical scanner 37-GHz channel (horizontal polarization). The symbols, contours, and scores are as in Fig. 2.

Fig. 12. As in Fig. 11, but for 1925 UTC 27 Aug 2005.
the west Pacific, Indian Ocean, and Australia–Fiji basins. (Another good alternative would be track forecasts issued by Japan’s Regional Specialized Meteorological Center.) Applying this to incoming microwave satellite imagery in real time requires an interpolation between positions at forecast times from roughly 3 to 12 h, depending on the latency of the forecast information and the latency of the image processing and/or transmission.

To characterize the accuracy of this forecast position, we have collected several years of NHC and JTWC forecast positions used as real-time first-guess centers for PMW images in the MIMIC product (Wimmers and Velden 2007) since mid-2004. These forecast positions can be easily evaluated against the final best-track TC positions issued by the respective agencies. The forecasts in the North Atlantic basin must be treated separately from those of the other basins because the high number of regular aircraft observations there leads to an improvement in forecast accuracy.

The 2007 and 2008 North Atlantic storm seasons were selected to estimate the North Atlantic real-time first-guess center accuracy, and the 2008 west Pacific storm

![Fig. A1.](image1)

Fig. A1. (a) First-guess TC center error as the difference between real-time available NHC forecast TC center during the 2007–08 North Atlantic seasons and the best track; (b) distribution of error distance. Asterisks in (b) mark average values binned to 0.1°, 0.4°, and 0.7° (described in text).

![Fig. A2.](image2)

Fig. A2. As in Fig. A1, but (for the difference between real-time available JTWC forecast TC center during the 2008 west Pacific season and the best track.
season was selected to estimate the first-guess center accuracy in the remaining basins. The following filtering criteria were applied to the data: the maximum sustained winds must exceed 34 kt and the best-track center latitude must be south of 40°N. These two criteria reduce the effect of ambiguous centers of rotation during early storm development and extratropical transition, respectively, on the development sample distribution. Also, the data for North Atlantic TCs Chantal (2007), Noel (2007), Dolly (2008), and Nana (2008) were lost from the forecast position archive, and so are not included.

Figures A1 and A2 show the distribution of forecast error in the 2007–08 North Atlantic storm seasons and 2008 west Pacific storm season, respectively. Both figures show a roughly even azimuthal distribution of error, and are bounded within less than one degree. The average error for the North Atlantic (NHC forecast) is lower than the average error for the west Pacific (JTWC forecast), which is to be expected because of the higher density of observations (including aircraft reconnaissance) in the North Atlantic. However, the tail of the distribution for the North Atlantic is greater, which may result from the fact that the precise North Atlantic observations afford more opportunities to correct the forecasts. Three asterisks on the histograms of Figs. A1 and A2 mark the binned averages of the distances at 0° to 0.25°, 0.25° to 0.55°, and >0.55°. These binned values (Table A1) are used in sections 3 and 4 to weight the results at simulated first-guess errors of 0.1°, 0.4°, and 0.7°, respectively.

### APPENDIX B

#### Derivation of Image-Relative Error

The best possible statistic of algorithm accuracy is the error of the algorithm position with respect to the center position at the retrieval height of the image, because this is the center indicated by the information in the image. However, the only available “truth” dataset for the calibration and validation is the NHC best track, which is a position for the TC center of rotation at the surface. The difference between the best track and the image center can be treated as an additional, independent source of error in the validation, shown as follows.

Because error in the north–south direction is independent from error in the east–west direction, we can derive the relationship in one direction (denoted by $x$) and it will apply as well to the orthogonal direction ($y$):

$$e(x_{\text{ARCHER}}|x_{\text{Best Track}}) = e(x_{\text{ARCHER}}|x_{\text{image}}) + e(x_{\text{Best Track}}|x_{\text{image}}),$$  \hspace{2em} (B1)

where $e(x_A|x_B)$ is the error of the set of $x$ values in set $A$ with respect to the values in reference set $B$. The commonly accepted relationship for the error contribution of independent (uncorrelated) quantities to the error of the sum is the following (Brown and Berthouex 2002):

$$z = a + b + c + \cdots \text{ (for each measurement)} \hspace{2em} (B2a)$$

$$e(z) = e(a) + e(b) + e(c) + \cdots \text{ (for each measurement)} \hspace{2em} (B2b)$$

$$\text{Var}(z) = \text{Var}(a) + \text{Var}(b) + \text{Var}(c) + \cdots \text{ (for the entire set of measurements)} \hspace{2em} (B2c)$$

where $\text{Var}(\cdot)$ is the variance of the set. Therefore,

$$\text{Var}(x_{\text{ARCHER}}|x_{\text{Best Track}}) = \text{Var}(x_{\text{ARCHER}}|x_{\text{image}}) + \text{Var}(x_{\text{Best Track}}|x_{\text{image}}),$$  \hspace{2em} (B3)

The solution for $\text{Var}(x_{\text{ARCHER}}|x_{\text{image}})$ is simply

$$\text{Var}(x_{\text{ARCHER}}|x_{\text{image}}) = \text{Var}(x_{\text{ARCHER}}|x_{\text{Best Track}}) - \text{Var}(x_{\text{Best Track}}|x_{\text{image}}).$$  \hspace{2em} (B4)

In addition, the variance of the distance in general is simply the sum of the variances of the errors in each direction:

$$\text{Var}(x) = N^{-1} \sum_N (x - x_0)^2,$$  \hspace{2em} (B5a)

$$\text{Var}(y) = N^{-1} \sum_N (y - y_0)^2,$$  \hspace{2em} (B5b)

$$\text{Var}(\text{dist}) = N^{-1} \sum_N (x - x_0)^2 + (y - y_0)^2.$$  \hspace{2em} (B5c)

Therefore when Eq. (B4) is applied to both dimensions, the terms add to an equation of the variance of distance:

$$\text{Var}(\text{ARCHER}|\text{image}) = \text{Var}(\text{ARCHER}|\text{Best Track}) - \text{Var}(\text{Best Track}|\text{image}),$$  \hspace{2em} (B6)
where the notation $\text{Var}(A|B)$ is described to denote the variance of the distance between the positions for set $A$ relative to set $B$. The square root of the term $\sqrt{\text{Var}(\text{ARCHER}|\text{image})}$ is what we call the image-relative error and the square root of $\sqrt{\text{Var}(\text{Best Track}|\text{image})}$ is what we call the intrinsic error.

We have employed the following method for semi-manually estimating $\text{Var}(\text{Best Track}|\text{image})$ using the ARCHER results. The goal is to find the error (in distance) between the best-track position for a given observation and the rotational center indicated by the image. Although no objective, independent source is available for finding the image-relative rotational center, we can approximate this position by choosing a subset of microwave images in which a clear and rounded eye pattern indicates the image-based center with little or no ambiguity, and in which the ARCHER result matches the position to within less than approximately 0.02$^\circ$ (~2 km). This value is considerably lower than the scale of the expected intrinsic error. The ARCHER result is then used as an approximation of the image-based center of rotation, and the variance of the best track with respect to the image is

$$\text{Var}(\text{Best Track}|\text{image}) \approx N^{-1} \sum_{\text{best cases}} \text{dist}(\text{Best Track}|\text{ARCHER})^2. \quad (B7)$$

The “best cases” are taken from cases with first-guess displacements of 0.1$^\circ$ to the north only, because the results from 0.1$^\circ$ displacements are the most accurate, and effectively the same in each direction. For 79 possible cases of group A category 2–5 hurricanes, 59 revealed an unambiguous image-based center. When applied to 44 possible group A category 1 cases, 10 cases identified unambiguous image-based centers. However, the group A tropical-storm-strength cases did not yield a center of rotation with an equally high level of certainty. The decreasing fraction of “best cases” with decreasing TC

<table>
<thead>
<tr>
<th>TC intensity (kt)</th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$34 \leq V_{\text{max}} &lt; 65$</td>
<td>0.154$^a$</td>
<td>0.173$^a$</td>
</tr>
<tr>
<td>$65 \leq V_{\text{max}} &lt; 84$</td>
<td>0.154</td>
<td>0.173$^a$</td>
</tr>
<tr>
<td>$V_{\text{max}} \geq 84$</td>
<td>0.010</td>
<td>0.173</td>
</tr>
</tbody>
</table>

where $a$ is the low-end estimate in a group with even less organization and resistance to vortex slant.

Using the same method for a smaller set of group B cases, we calculate an intrinsic error RMS of 0.17$^b$, which is significantly larger than the intrinsic error for group A cases, as expected. Because of the lack of best cases for estimating intrinsic error for category 1 hurricane and tropical storm groupings, we use the category 2–5 hurricane intrinsic error as a conservative substitute for the same reasons as before.

REFERENCES


