# Principal Component Analysis of Doppler Radar Data. Part I: Geometric Connections between Eigenvectors and the Core Region of Atmospheric Vortices

PAUL R. HARASTI* AND ROLAND LIST

*Department of Physics, University of Toronto, Toronto, Ontario, Canada*

(Manuscript received 28 October 2004, in final form 27 May 2005)

ABSTRACT

This is the first in a three-part series of papers that present the first applications of principal component analysis (PCA) to Doppler radar data. Although this novel approach has potential applications to many types of atmospheric phenomena, the specific goal of this series is to describe and verify a methodology that establishes the position and radial extent of the core region of atmospheric vortices. The underlying assumption in the current application is that the streamlines of the nondivergent component of the horizontal wind are predominantly circular, which is a characteristic often observed in intense vortices such as tropical cyclones.

The method employs an $S^2$-mode PCA on the Doppler velocity data taken from a single surveillance scan and arranged sequentially in a matrix according to the range and azimuth coordinates. Part I begins the series by examining the eigenvectors obtained from such a PCA applied to a Doppler velocity model for a modified, Rankine-combined vortex, where the ratio of the radius of maximum wind to the range from the radar to the circulation center is varied over a wide range of values typically encountered in the field. Results show that the first two eigenvectors within the eigenspace of range coordinates represent over 99% of the total variance in the data. It is also demonstrated that the coordinates of particular cusps in the curves of the eigenvector coefficients plotted against their indices are geometrically related to both the position of circulation center and the radius of maximum wind.

## 1. Introduction

Over the last 100 yr, principal component analysis (PCA) has been employed in many diverse fields of study to identify the dominant variables and mechanisms that describe and control the structure and processes underlying a specific dataset. In meteorology and oceanography, the dataset entails observations of one or more physical fields at one or more times and locations. Either the physical field, time, or location is considered to be the observation while one, or both, of the other two are used to index a series of such observations. This results in eight possible modes of PCA analysis that are referred to as O, P, Q, R, S, T, $S^2$, or $T^2$ mode; details are given in Preisendorfer (1988). The values of the observations are placed in a 2-dimensional data matrix where the rows represent one index and the columns represent the other. The eigenvalues and eigenvectors are then computed from the scatter matrix of the data matrix. Fewer eigenvectors than the number of original data vectors are required to sufficiently represent the data thus allowing for data compression. Moreover, it is often possible to find physical interpretations of the eigenvectors and principal components.

This paper presents the first application of PCA to data taken from a Doppler radar that is utilized for weather surveillance. There were two previous applications of PCA to other Doppler-type data. Williams (1997) performed an S-mode PCA on 50-MHz wind profiler data to study the hourly, daily, annual, and interannual variability of the wind above Christmas Island, Kiribati. S-mode in this case indicated that the indices of time and location (profiler altitudes) of the zonal wind estimates (observations) were arranged along the rows and columns, respectively, of the data

* Current affiliation: Visiting Scientist Program, University Corporation for Atmospheric Research, Boulder, Colorado.

*Corresponding author address:* Dr. Paul R. Harasti, UCAR Project Scientist, Naval Research Laboratory, Marine Meteorology Division, 7 Grace Hopper Avenue, MS-2, Monterey, CA 93943-5502.
E-mail: paul.harasti@nrlmry.navy.mil

matrix. Bradley (1997) performed a P-mode PCA on size spectra of raindrops between 20 and 220 m altitude deduced from acoustic radar data obtained at South Island, New Zealand. P-mode in this case signified that the indices of time and drop diameter (physical field) were arranged along the rows and columns, respectively, while using the location (profile altitude) as the observation in the data matrix. The eigenvectors provided a means to characterize the temporal departures of the drop size distribution from its theoretical exponential distribution. For some other examples of previous uses of PCA in meteorology, the reader is referred to Stidd (1967), Phillips (1972), Horel (1981), Richman (1986), Barnston and Livezey (1987), Nigam and Shen (1993), Johnson and Hanson (1995), Huang and Antonelli (2001), and Hillger and Ellrod (2003).

The objective of this paper is to present the main principles behind a new Doppler radar analysis method for atmospheric vortices whose further refinements in Part II of this series will offer fewer limitations than the current methods that are used to estimate the position of the circulation center and the radius of maximum wind. In particular, PCA is employed in a new technique that extracts the essential information from a single Doppler radar scan with a higher degree of simplicity and accuracy. Although the current application utilizes an analytic model for the Doppler velocity data of tropical cyclones by way of example, the same procedure may be applied to radar data of smaller-scale vortices and other atmospheric phenomena.

The general procedure for applying PCA to Doppler radar data is described in section 2. Section 3 describes the details of the analytical model for Doppler velocity used in this study. PCA is then applied to a wide range of model configurations and parameters in section 4, which reveals geometric connections between the eigenvectors and the position and radial extent of the core region of atmospheric vortices. Section 5 provides a discussion on the significance of these connections. Finally, a summary and conclusions are presented in section 6.

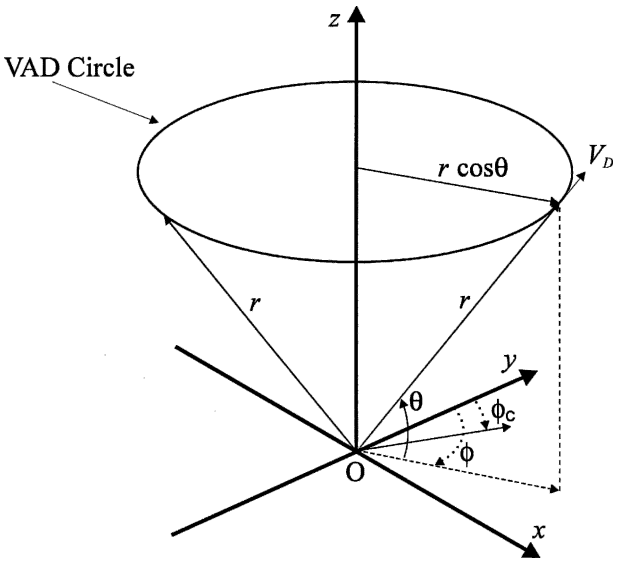


FIG. 1. Bird's-eye view of a PPI scan. The radar scans through a complete horizontal circle at a fixed elevation angle ($\theta$). The Doppler velocity of the precipitation particles ($V_D$) is estimated at fixed positions of range ($r$) and azimuth ($\phi$) along the surface of the cone defined by the complete PPI scan. There is a VAD circle located at every $r$ position, each having a different altitude $z$ above the earth-centered spherical surface that intersects the radar. The circulation center of the tropical cyclone is located in the direction of the azimuth position $\phi_C$. The coordinates of $V_D$ are expressed in terms of the spherical coordinate system ($r$, $\phi$, $\theta$) whose origin is at the radar (O); $y$ and $x$ point to true north and east, respectively.

## 2. PCA of Doppler radar data

This application of PCA to Doppler radar data utilizes the Doppler velocity data ($V_D$) taken from a plan-position-indicator (PPI) scan. Figure 1 shows a bird's-eye view of the geometry associated with a PPI scan. The Doppler radar scans through a complete circle at a fixed elevation angle, $\theta$. The current approach seeks properties of vortices in the region of maximum winds that, for intense vortices such as tropical cyclones, occur in the lower troposphere (e.g., Jorgensen 1984; Liu et al. 1999). If the radar's beamwidth is ~1°, it is recommended that the PPI be executed at $\theta \approx 0.5°$ so as to maximize the horizontal cross section through the storm and minimize the altitude spanned by the surface of the cone defined by the PPI scan. If it is not desirable or possible (e.g., because of a wider beamwidth or terrain blockage of the radar's view of the vortex) to execute a PPI at $\theta \approx 0.5°$, it is recommended that a constant altitude PPI (CAPPI) be utilized instead (e.g., Mohr et al. 1986), which is a representation of the Doppler velocity at constant altitude derived from the interpolation of data from several PPI scans of different elevation angles executed during a radar volume scan. However, if a CAPPI is used, one should correct the Doppler velocity data for the terminal fall speed of hydrometeors estimated from a standard reflectivity factor and fall speed relationship.

Depending on the Nyquist velocity dictated by the pulse repetition frequency of the radar, it might be necessary to de-alias the Doppler velocity data via an algorithm such as that of Bargen and Brown (1980). The de-aliased data are arranged by their range and azimuth coordinates ($r$, $\phi$) into an $N \times M$ matrix designated by

**D**, where $N$ is the number of range gates and $M$ is the number of azimuth positions. The $D_{ij}$ element of **D** is the Doppler velocity datum at the $i$th range gate and $j$th azimuth. Each row of **D** represents one of the $N$ velocity azimuth display (VAD) circles (Lhermite and Atlas 1961) in the PPI scan. It is an arbitrary, though logical, choice to place the range index along the rows of **D**.

Most of the modes of PCA deal with the one scatter matrix that defines the mode. However, the current application is concerned with the dual-variation of the Doppler velocity in both the azimuthal and range directions. Therefore, a dual $S^2$-mode PCA is performed on **D**. The $S^2$ mode indicates that the physical field (Doppler velocity, in this case) is the observation, indexed by its coordinates in two dimensions (range and azimuth relative to the radar) in the matrix **D**. The data are centered separately in both the range and azimuth direction so that the dual scatter matrices become covariance matrices. Centering with respect to range is accomplished by subtracting the average $V_D$ of each column (radial ray) of **D** from every datum of that column; let this centered matrix be denoted by $\mathbf{D}_r$. Centering with respect to azimuth is accomplished by subtracting the average $V_D$ of each row (VAD circle) of **D** from every datum of that row; let this centered matrix be denoted by $\mathbf{D}_\phi$. If there are missing $V_D$ values in **D** then these are set to the average of their particular column or row depending on whether $\mathbf{D}_r$ or $\mathbf{D}_\phi$ are calculated, respectively. Such a procedure is common practice in PCA (Jolliffe 1986).

The variance from the mean of the data is then expressed by formulating the two covariance matrices

$$\mathbf{C}_r = \frac{1}{N-1}\mathbf{D}_r^{\mathrm{T}}\mathbf{D}_r \tag{1}$$

$$\text{and} \quad \mathbf{C}_\phi = \frac{1}{M-1}\mathbf{D}_\phi\mathbf{D}_\phi^{\mathrm{T}}, \tag{2}$$

where T designates the transpose operation, and $\mathbf{C}_r$ and $\mathbf{C}_\phi$ are $M \times M$ and $N \times N$ matrices, respectively. Let the $j$ nonzero eigenvalues and eigenvectors of $\mathbf{C}_r$ be denoted by $\mu_j$ and $\mathbf{a}_j$, respectively, and let the $k$ nonzero eigenvalues and eigenvectors of $\mathbf{C}_\phi$ be denoted by $\lambda_k$ and $\mathbf{f}_k$, respectively. The eigenstructures of the covariance matrices are then determined by solving the dual eigenproblems

$$\mathbf{C}_r\mathbf{a}_\mathrm{j} = \mu_\mathrm{j}\mathbf{a}_\mathrm{j} \tag{3}$$

$$\text{and} \quad \mathbf{C}_\phi\mathbf{f}_\mathrm{k} = \lambda_\mathrm{k}\mathbf{f}_\mathrm{k}, \tag{4}$$

$$\text{where} \quad \mathbf{a}_j = [a_{1j}, a_{2j}, a_{3j}, \ldots, a_{Mj}]^{\mathrm{T}}, \tag{5}$$

$$j \le \min(N-1, M), \tag{6}$$

$$\mathbf{f}_k = [f_{1k}, f_{2k}, f_{3k}, \ldots, f_{Nk}]^{\mathrm{T}}, \tag{7}$$

$$\text{and} \quad k \le \min(N, M-1). \tag{8}$$

The projection of the original data vectors onto the eigenvectors $\mathbf{a}_j$ and $\mathbf{f}_k$ yields their respective principal components,

$$\mathbf{b}_j = \frac{\mathbf{D}_r\mathbf{a}_j}{|\mathbf{D}_r\mathbf{a}_j|} \tag{9}$$

$$\text{and} \quad \mathbf{e}_k = \frac{\mathbf{D}_\phi^{\mathrm{T}}\mathbf{f}_k}{|\mathbf{D}_\phi^{\mathrm{T}}\mathbf{f}_k|}, \tag{10}$$

$$\text{where} \quad \mathbf{b}_j = [b_{1j}, b_{2j}, b_{3j}, \ldots, b_{Nj}]^{\mathrm{T}} \tag{11}$$

$$\text{and} \quad \mathbf{e}_k = [e_{1k}, e_{2k}, e_{3k}, \ldots, e_{Mk}]^{\mathrm{T}}. \tag{12}$$

As one can see, the principal components are also vectors whose magnitudes have been normalized to unity in (9) and (10); the normalization is optional for the current approach and is usually not done in PCA applications. The vectors $\mathbf{a}_j$ and $\mathbf{e}_k$ have $M$ coefficients; one for each azimuth coordinate, $\phi$. Similarly, the vectors $\mathbf{b}_j$ and $\mathbf{f}_k$ have $N$ coefficients; one for each range coordinate, $r$. Thus, the coefficients of each of these vectors may be represented as functions of the coordinates as follows:

$$a_{\phi j} = a_j(\phi), \tag{13}$$

$$b_{rj} = b_j(r), \tag{14}$$

$$e_{\phi k} = e_k(\phi), \tag{15}$$

$$\text{and} \quad f_{rk} = f_k(r). \tag{16}$$

Using modern PCA jargon, the coefficients (13) and (16) are referred to as the eigenvector loadings whereas the coefficients (14) and (15) are referred to as the principal component scores (normalized to unity). As an alternative, the coefficients (13)–(16) may also be obtained from $\mathbf{D}_r$ or $\mathbf{D}_\phi$ in one step from the related multivariate technique of singular value decomposition (Jolliffe 1986).

The current application is only concerned about the physical interpretation of the eigenvectors and principal components. That is, the shape of the functions (13)–(16) is analyzed for particular cusps whose abscissas are related to the position and radial extent of the vortex core. Therefore, the scaling of the vectors $\mathbf{a}_j$, $\mathbf{b}_j$, $\mathbf{e}_k$, and $\mathbf{f}_k$ is arbitrary. The reason for scaling the principal components to unity in (9) and (10) is to underscore the fact that they may also be considered eigenvectors. As Table 1 demonstrates, if one were to reverse the order of the matrix products on the right-hand sides of (1) and (2), the respective eigenvectors would be identical to (9) and (10), and the respective normalized principal components would be identical to (5) and (7). In other words, the roles of the eigenvectors and principal components would become interchanged (Preisendorfer 1988). This duality property has two

TABLE 1. The duality of the eigenvectors and principal components is revealed when one formulates the scatter matrix in both azimuth space ($M \times M$ matrix) and range space ($N \times N$ matrix). One can perform the eigenanalysis in the smaller of these two spaces for computational efficiency and retrieve the same information that one would obtain from the larger space; the eigenvectors of the former are equivalent to the normalized principal components of the latter and vice versa.

| | $M \times M$ scatter matrix | | $N \times N$ scatter matrix | |
|---|---|---|---|---|
| | $\mathbf{D}_r^{\mathrm{T}}\mathbf{D}_r$ | $\mathbf{D}_\phi^{\mathrm{T}}\mathbf{D}_\phi$ | $\mathbf{D}_r\mathbf{D}_r^{\mathrm{T}}$ | $\mathbf{D}_\phi\mathbf{D}_\phi^{\mathrm{T}}$ |
| Eigenvector | $\mathbf{a}_j$ | $\mathbf{e}_k$ | $\mathbf{b}_j$ | $\mathbf{f}_k$ |
| Normalized principal component | $\mathbf{b}_j$ | $\mathbf{f}_k$ | $\mathbf{a}_j$ | $\mathbf{e}_k$ |

practical consequences for the current method. First, one may formulate (1) and (2) into two $L \times L$ matrices, where $L = \min(N, M)$ thus creating the smallest possible matrices for computational efficiency while providing the same information that one would obtain from (1) and (2) left unchanged. Second, the duality property allows one to refer to all of the vectors $\mathbf{a}_j$, $\mathbf{b}_j$, $\mathbf{e}_k$, and $\mathbf{f}_k$ as eigenvectors, which will simplify the classification of eigenvectors into common types as revealed by the shapes of the functions (13)–(16).

The eigenvalues $\mu_j$ and $\lambda_k$ represent the variance from the mean of the data accounted for by each corresponding eigenvector. As a fraction of the total variance in the data, the eigenvectors $\mathbf{a}_j$ and $\mathbf{b}_j$ account for $\mu_j/\mathrm{trace}(\mathbf{C}_r)$, whereas the eigenvectors $\mathbf{e}_k$ and $\mathbf{f}_k$ account for $\lambda_k/\mathrm{trace}(\mathbf{C}_\phi)$. PCA arranges the eigenvectors in descending order of the variance reduction that each affords. In other words, the first eigenvectors ($j, k = 1$) give the dominant modes of variability in the data, and all of the remaining eigenvector modes give the maximum amount of variability in the data not explained by any of those preceding them. Unlike the original data vectors, the eigenvectors are uncorrelated and orthonormal. The higher-indexed eigenvectors tend to represent the noise in the dataset, although determining the cutoff index that separates the signal from the noise can be a subjective procedure.

It should be noted that a subset of the eigenvectors and principal components may be used to represent the original data in an approximate form. For example, the $N \times M$ matrix $\mathbf{D}_r$ can be approximated via $\mathbf{D}_r \approx \mathbf{P}\mathbf{A}^{\mathrm{T}}$, where $\mathbf{A}$ is the $M \times S$ matrix whose columns are the first $S$ eigenvectors of the set $\mathbf{a}_j$, $\mathbf{P}$ is the $N \times S$ matrix whose columns are the associated, nonnormalized principal components $\mathbf{b}_j$ $|\mathbf{D}_r\mathbf{a}_j|$, and $S \ll \min(N - 1, M)$. Jolliffe (1986) describes several techniques to determine the best value for $S$. An approximate version of the original matrix $\mathbf{D}$ can then be constructed by reversing the centering procedure on the S-vector approximation of $\mathbf{D}_r$. Such PCA syntheses are beyond the scope of this paper but merit further investigation since there is a great potential to significantly compress all types of Doppler radar data while eliminating unwanted noise, reducing storage space, and facilitating electronic transfer.

## 3. A Doppler velocity model for axisymmetric rotation

By way of example, the parameters used in the analytic model for this study are chosen to approximate the observed structure of tropical cyclones. The primary circulation (tangential wind component) of most mature tropical cyclones is strongly axisymmetric in the lower troposphere with an asymmetric and radial wind component of less than 10% (Gray 1982; Emanuel 1991; Willoughby 1988, 1995). Thus, it is necessary and appropriate to develop wind models and Doppler radar analysis methods specifically for atmospheric vortices that are dominated by axisymmetric rotation. The current model describes the variation of the horizontal wind components with respect to a cylindrical coordinate system whose origin is the tropical cyclone's circulation center (the point C in Fig. 2), which is located at the azimuth position $\phi_c$ and radar range $r_c$. The tropical cyclone's horizontal wind is resolved into its tangential ($V_T$) and radial ($V_R$) components along every cylindrical surface of constant radius, $R$. As a start-

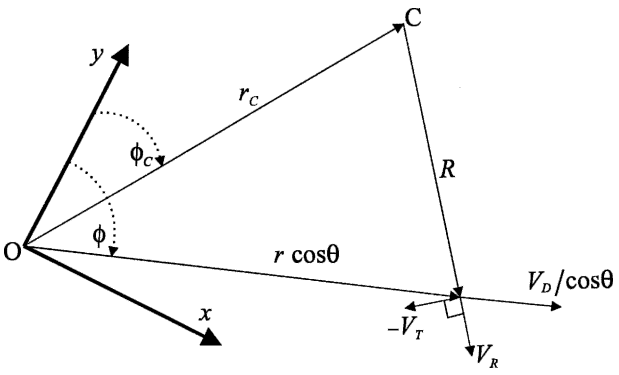


FIG. 2. Plan view of Fig. 1 through the plane of the VAD circle. The circulation center of the tropical cyclone is located at the point $C$, whose coordinates are $(r, \phi,) = (r_C, \phi_C,)$ The horizontal component of the Doppler velocity ($V_D/\cos\theta$) is shown at the same horizontal position as the Doppler velocity ($V_D$) shown in Fig. 1; $V_D/\cos\theta$ is the sum of the projection of the horizontal wind components onto the direction of $V_D$: the tangential wind ($V_T$), which is positive counterclockwise with respect to $C$, and the radial wind ($V_R$), which is positive outward from $C$. Also, $R$ is the radial distance from $C$ to the center of the sampling volume within which $V_D$ is measured.

ing point, this paper presents PCA results from an analytic model for $V_D$ containing a horizontal component resulting from axisymmetric rotation alone, as governed by an analytic equation for $V_T$, to be described in the following.

The majority of tropical cyclones have a single tangential wind maximum (Willoughby et al. 1982; Samsury and Zipser 1995). For this situation, the magnitude of $V_T$ is a maximum at the radius of maximum wind (RMW = $R_m$) and is adequately described by the modified, Rankine-combined vortex flow model (Rankine 1901):

$$V_T(R) = \begin{cases} V_T(R_m)\dfrac{R}{R_m} & R \le R_m \\ V_T(R_m)\sqrt{\dfrac{R_m}{R}} & R > R_m \end{cases}, \qquad (17)$$

and $V_T$ ($R_m$) is the maximum tangential wind at $R_m$, which occurs at an altitude $z_m$ such that 0.5 km $\le z_m \le$ 2 km (Powell et al. 1991) and $R_m < 55$ km in most instances (Shea and Gray 1973; Jorgensen 1984; Roux and Viltard 1995). For simplicity, the vertical variations in $V_T$ are not modeled. The square root law for $R$ in the region $R > R_m$ is based on the average-observed wind profile calculated along many different $R$ radii of hundreds of tropical cyclones (e.g., Gray and Shea 1973; Samsury and Zipser 1995), and its theoretical basis was expounded by Riehl (1963). The model assumes a moderately intense tropical cyclone with $V_T(R_m) = 50$ m s$^{-1}$ and 5 km $< R_m <$ 65 km.

Here, $V_D$ is defined with respect to a spherical ($r$, $\phi$, $\theta$) coordinate system whose origin is at the location of the radar (Fig. 1). Equations that relate $V_D$ to $V_T$ and $V_R$ are shown in appendix A of Wood and Brown (1992). As is customary in many Doppler radar data studies, it is assumed that the radar emits radiation in a straight beam with an infinitesimal width, and that the earth has an effective radius equal to four-thirds its actual radius (Doviak and Zrnic 1984). The radial ($r$), and azimuthal ($\phi$) resolutions used in the model are chosen to match those of Weather Surveillance Radar-1988 Doppler (WSR-88D) Archive Level II data; namely, 250 m and 1°, respectively. The maximum, unambiguous range ($r_{\max}$) for WSR-88D radars operating in the Doppler mode varies between 150 km $< r_{\max} <$ 230 km; however, for simplicity, $r_{\max} = 150$ km is assumed. In addition, the vertical component of the Doppler velocity is ignored since it is small compared to the horizontal component when $\theta \approx 0.5°$ for $r_{\max} \approx$ 150 km.

As an example, Figs. 3e,f illustrate two of the simulated PPIs of Doppler velocity that are examined in the next section, where $r_C = 70$ km, and the two cases $R_m = 20$ km and $R_m = 60$ km are shown. Figures 3a,b show the tangential wind profile as per (17) in each case, and Figs. 3c,d show the corresponding wind fields whose radar-radial components result in Figs. 3e,f. The Doppler velocity data reveal a primary couplet of maximum (positive) and minimum (negative) velocities at the points where the radar beam is tangent to the ring of maximum winds centered on the circulation center. The contours that are concentric with each component of the couplet turn inward toward the radar as the RMW becomes larger (cf. Figs. 3e,f), or more precisely, as the RMW approaches the radar.

Referring to Figs. 3c–f, if one were to change $\phi_C$ from its value of 180° to $\phi_C = 180° + \beta$, where $\beta$ is any positive or negative angle, the azimuth coordinates of the vector wind field points, along with the azimuth coordinates of their corresponding Doppler velocity data would simply change by $\beta$. Similarly, the azimuth coordinates of the eigenvector coefficients $a_j(\phi)$ and $e_k(\phi)$ would also change by $\beta$. Therefore, only the starting position of the trace (phase), not the shape, of the eigenvector curves of $a_j(\phi)$ and $e_k(\phi)$ versus $\phi$ is dependent of the value of $\phi_C$ chosen in the model for $V_D$. As will be demonstrated, the shapes of these particular curves for $j, k = 1$ and $j, k = 2$ are largely symmetrical about $\phi_C$, which is a point of interest. Since it is desirable to have this symmetry displayed with respect to the center of the 360° range of azimuth, only the value $\phi_C = 180°$ is used for demonstration purposes in section 4. In practical applications, however, the user will need to identify the same patterns, shifted by $\beta°$ of azimuth along the abscissa of the curves.

It is also useful to define the dimensionless aspect ratio

$$\rho = \frac{R_{\mathrm{m}}}{r_C}. \qquad (18)$$

Wood and Brown (1992) show that, regardless of the scale of the core diameter of axisymmetric rotation, the Doppler velocity pattern is governed by (18). Therefore, the current methodology may also be applied to the Doppler velocity data of smaller-scale intense vortices such as extratropical mesocylones and tornadoes observed both by WSR-88D-like radars and by Doppler radars of much smaller scanning range, such as the mobile X- and W-band radars described in Wurman et al. (1997) and Bluestein et al. (2003).

Only the region where $\rho < 1$ is considered in this study; that is, the radar is located outside the RMW. This is the most common configuration during the landfall of a tropical cyclone, and smaller-scale vortices are

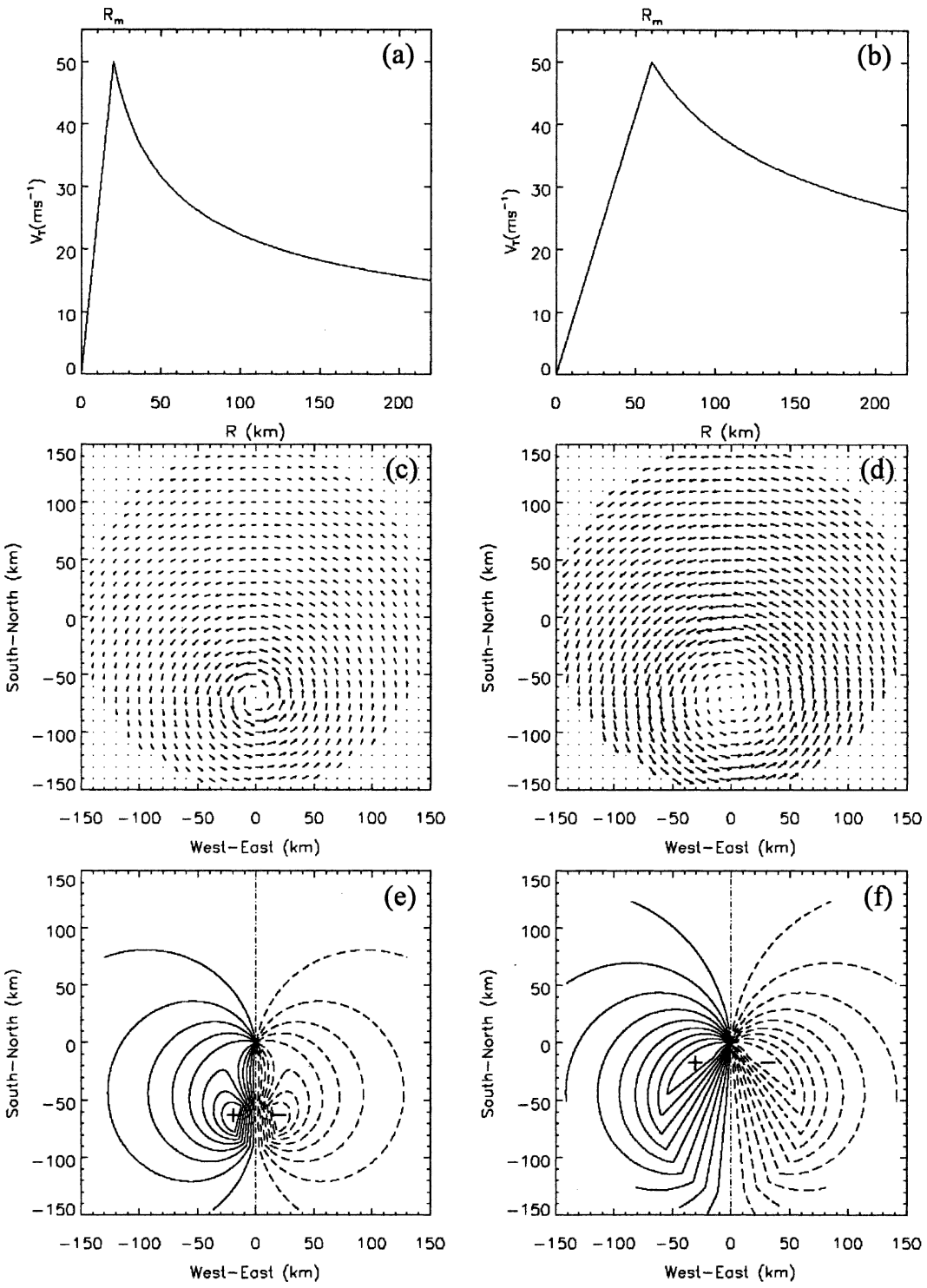


FIG. 3. Examples of the analytic models for the cases (left) $R_m$ = 20 km and (right) $R_m$ = 60 km. (a), (b) The modified, Rankine-combined vortex profiles of the tangential wind; (c), (d) the corresponding vector wind fields; (e), (f) the corresponding Doppler velocity contours spaced in 5 m s$^{-1}$ intervals with positive (away) and negative (toward) values delineated by solid and dashed contours, respectively. The zero isodop (vertical dash–dotted line), and the extrema positions labeled "+" and "−" are also shown. In (c)–(f), the radar is located at the origin, and the circulation center is located at ($r$, $\phi$) = (70 km, 180°).

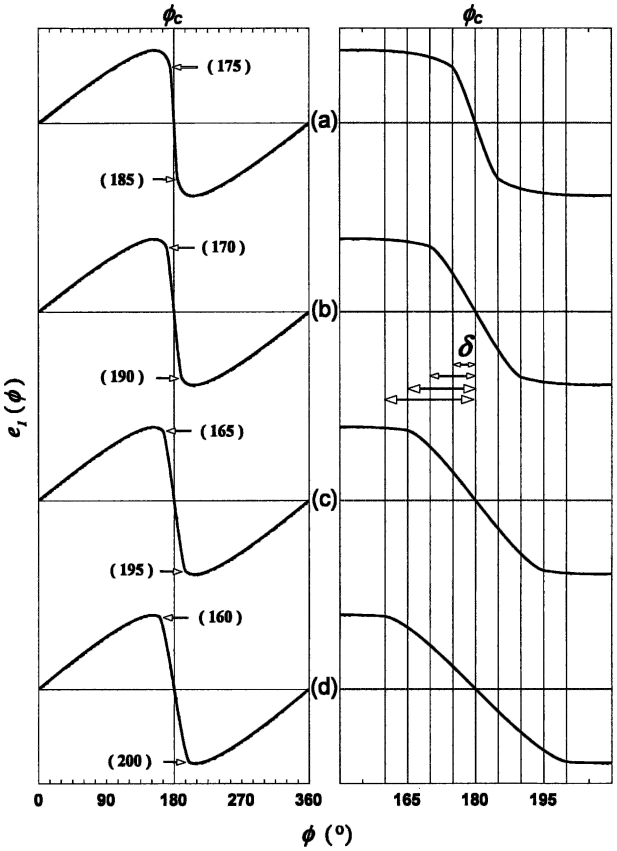


FIG. 4. Eigenvector coefficients $e_1(\phi)$ vs $\phi$ derived from simulated PPI scans of the analytic vortex model (17) using (a) $\rho$ = 0.0872, (b) $\rho$ = 0.1736, (c) $\rho$ = 0.2588, and (d) $\rho$ = 0.342 with $r_C$ = 70 km and $\phi_C$ = 180°. The right frame shows the same curves as those in the adjacent left frame except over a smaller range of azimuth centered on $\phi_C$. The cusp points of interest are shown by the left- and right-pointing arrows; their abscissas, $\phi_{\delta-}$ and $\phi_{\delta+}$, respectively, are the azimuth angles shown in parentheses in the left frame and are also delineated by the vertical gridlines in the right frame. Half of the difference between each of these paired abscissas is the angle δ indicated by the double arrows. The vertical scale is arbitrary. The horizontal gridlines $e_1(\phi) = 0$ are shown for each curve.

even less likely to traverse directly over the radar. Although a PCA may be performed during the infrequent case of $\rho \geq 1$, there are additional effects and considerations regarding the retrieval of the position and radial extent of the vortex core that are beyond the scope of this paper.

## 4. Geometric connections

### *a. Azimuth eigenvectors*

This section presents the eigenstructure of $\mathbf{C}_\phi$ derived from a wide range of simulated PPI scans whose Doppler velocity data correspond to the wind field given by (17). The eigenstructure of $\mathbf{C}_r$ does not reveal any new physics, however, it offers a wider range of practical uses that will be discussed in Part II of this work. Nevertheless, the fundamental physical interpretations of the eigenvectors are most clearly revealed from the eigenstructure of $\mathbf{C}_\phi$ and is therefore an appropriate starting point.

Using the analytic vortex model (17) and fixing $\phi_C$ = 180° and $r_C$ = 70 km, PCAs were performed on simulated PPIs for increasing values of $\rho$ (by varying $R_m$) in the region $\rho < 1$. Figures 4 and 5 show $e_1(\phi)$ versus $\phi$ for eight different values of $\rho$ in the range $0.0872 \leq \rho \leq 0.9063$. Similarly, Figs. 6 and 7 show $e_2(\phi)$ versus $\phi$ for eight different values of $\rho$ in the range $0.0875 \leq \rho \leq 0.9004$. The reasons for the specific values of $\rho$ are related to, and therefore discussed after, the geometric relations revealed in this section.

The left frames of Figs. 4–7 show the eigenvector coefficients through the full 360° range of azimuth. For display purposes, the magnitudes of the coefficients are arbitrary; only the zero-magnitude horizontal gridline is shown for each plot. The $e_1(\phi)$ curves display a single maximum and a single minimum, both of which are

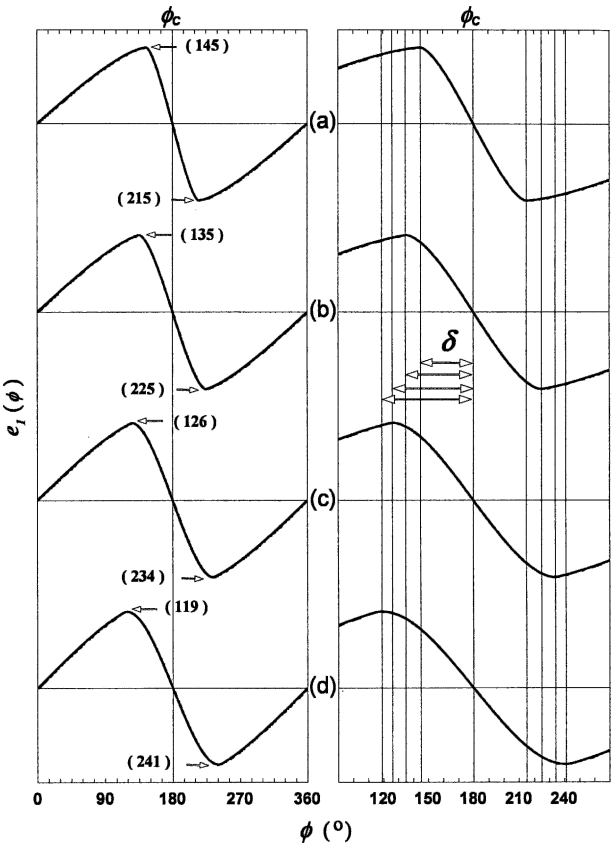


FIG. 5. Same as in Fig. 4 except for (a) $\rho$ = 0.5736, (b) $\rho$ = 0.7071, (c) $\rho$ = 0.8192, and (d) $\rho$ = 0.9063.

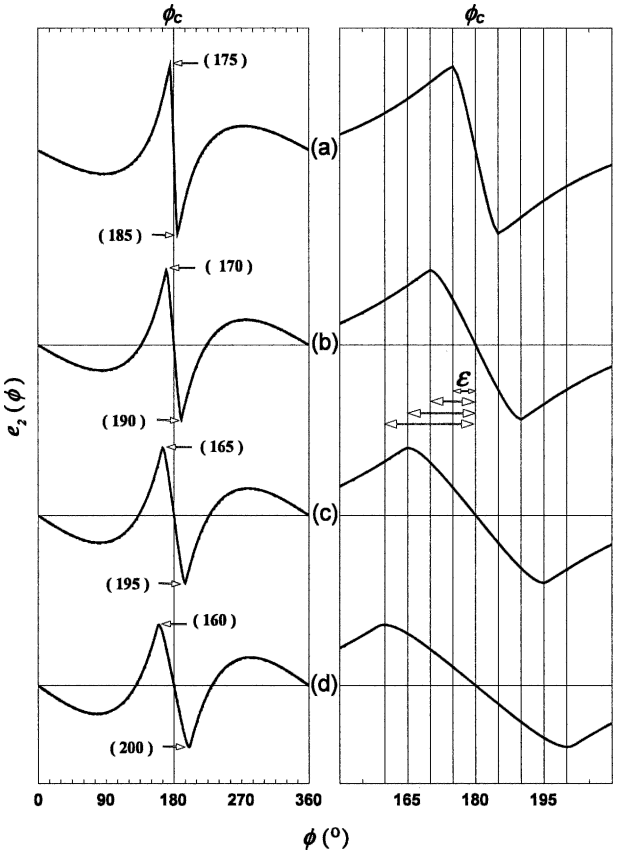


FIG. 6. Eigenvector coefficients $e_2(\phi)$ vs $\phi$ derived from simulated PPI scans of the analytic vortex model (17) using (a) $\rho$ = 0.0875, (b) $\rho$ = 0.1763, (c) $\rho$ = 0.2679, and (d) $\rho$ = 0.364 with $r_C$ = 70 km and $\phi_C$ = 180°. The right frame shows the same curves as those in the adjacent left frame except over a smaller range of azimuth centered on $\phi_C$. The cusp points of interest are shown by the left- and right-pointing arrows; their abscissas, $\phi_{\epsilon-}$ and $\phi_{\epsilon+}$ respectively, are the azimuth angles shown in parentheses in the left frame and are also delineated by the vertical gridlines in the right frame. Half of the difference between each of these paired abscissas is the angle $\epsilon$ indicated by the double arrows. The vertical scale is arbitrary. The horizontal gridlines $e_2(\phi)$ = 0 are shown for each curve.

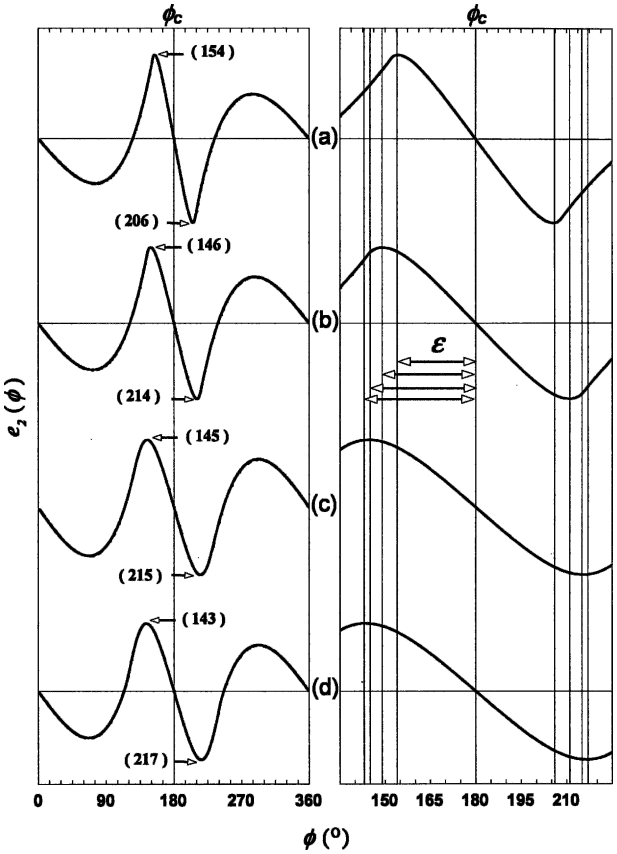


FIG. 7. Same as in Fig. 6 except for (a) $\rho$ = 0.4663, (b) $\rho$ = 0.7002, (c) $\rho$ = 0.8098, and (d) $\rho$ = 0.9004.

vertically displaced from the zero gridlines by equal amounts, and have an azimuthal spacing that increases with increasing $\rho$. Similarly, the $e_2(\phi)$ curves show a global maximum and minimum that are vertically displaced from the zero gridlines by equal amounts, and have an azimuthal spacing that increases with increasing $\rho$; the smaller local extrema are not of primary importance. There are two cusp points on every curve of $e_1(\phi)$ and $e_2(\phi)$ that are physically related to the geometry of the relative positions of the radar, the circulation center, and the RMW.

The cusp points of interest for $e_1(\phi)$ are indicated by the left- and right-pointing arrows in Figs. 4–5; their abscissas are shown in parentheses and are equal to the azimuth positions

$$\phi_{\delta-} = \phi_C - \delta \tag{19}$$

$$\text{and} \quad \phi_{\delta+} = \phi_C + \delta, \tag{20}$$

respectively, where $\delta$ is a function of $\rho$. Therefore,

$$\phi_C = \frac{[\phi_{\delta-} + \phi_{\delta+}]}{2} \tag{21}$$

$$\text{and} \quad \delta = \frac{[\phi_{\delta+} - \phi_{\delta-}]}{2}. \tag{22}$$

The left frames of Figs. 4 and 5 show that $\phi_{\delta-}$ ($\phi_{\delta+}$) tends to move toward the abscissa of the maximum (minimum) in $e_1(\phi)$ as $\rho$ increases. Note that $\delta$ increases with $\rho$.

Figure 8 shows the sine of each $\delta$ angle calculated from (22) as a function of its corresponding $\rho$ value. One can see that

$$\rho = \frac{R_m}{r_C} = \sin\delta \quad \rho < 0.75, \tag{23}$$

and the same relationship is approximately valid for larger values of $\rho$. This physical relationship is shown in

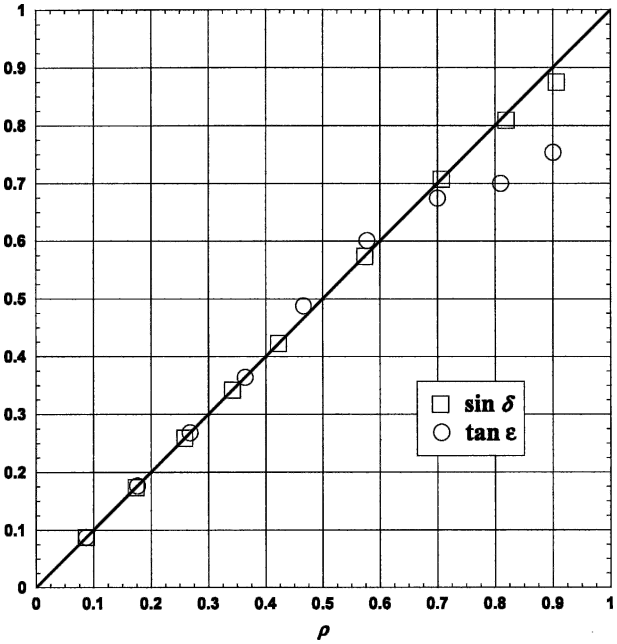


FIG. 8. Sinδ and tanε vs ρ for the values of δ, ε, and ρ shown in Figs. 4–7 including two extra points derived from graphical analyses not shown (ρ = 0.4226 for sinδ and ρ = 0.5774 for tanε). The solid diagonal line delineates ρ = sinδ and ρ = tanε.

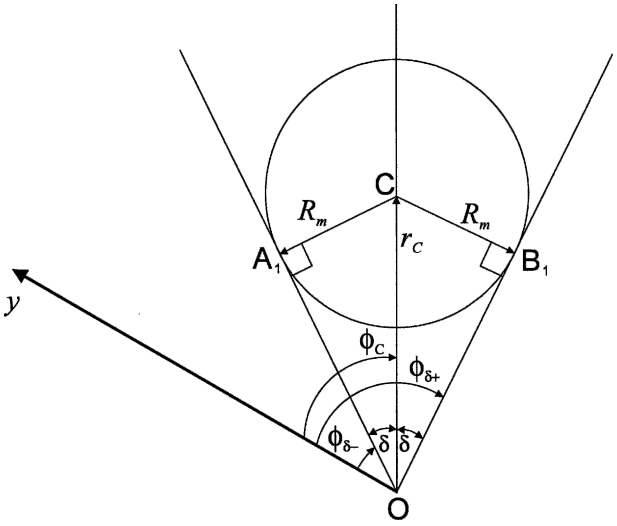


FIG. 9. The physical interpretation for the cusp points of $e_1(\phi)$ shown in Figs. 4 and 5 is described by the two right-angled triangles $OA_1C$ and $OB_1C$. The radar is located at O and the radar beam is tangent to the circle of radius $R_m$ at the points $A_1$ and $B_1$ which have the azimuth coordinates $\phi_{\delta-}$ and $\phi_{\delta+}$, respectively. The circulation center of the tropical cyclone is located at C, which has the radar-range coordinate $r_C$ and azimuth coordinate $\phi_C = \phi_{\delta+} - \delta = \phi_{\delta-} + \delta$. Simple trigonometry yields $R_m/r_C = \rho = \sin\delta$ from the triangles $OA_1C$ and $OB_1C$. The largest Doppler velocities in the PPI are located at the points $A_1$ and $B_1$ where the strongest tangential winds at the RMW are parallel to the radar beam. Given the results of Fig. 8, the abscissas of the cusp points shown in Figs. 4 and 5a,b ($\rho < 0.75$) therefore delineate the azimuthal coordinates of these points, and also give a measure of the azimuthal extent of the RMW.

Fig. 9 via the right-angled triangles $OA_1C$ and $OB_1C$. The largest Doppler velocities in the PPI are located where the radar beam is tangent to the circle of radius $R_m$ at the points $A_1$ and $B_1$, which have the azimuth coordinates $\phi_{\delta-}$ and $\phi_{\delta+}$, respectively. Therefore, $\phi_{\delta-}$ and $\phi_{\delta+}$ shown in Figs. 4 and 5a,b ($\rho < 0.75$) delineate the azimuthal coordinates of these points in physical space, and also give a measure of the azimuthal extent of the RMW. Evaluating (21) with the values of $\phi_{\delta-}$ and $\phi_{\delta+}$ shown parenthetically in Figs. 4 and 5, one can readily conclude that the value $\phi_C = 180°$ is retrieved exactly, even for those values of ρ for which the relationship $\rho = \sin\delta$ is not exact.

The cusp points of interest for $e_2(\phi)$ are the global extrema shown by the left- and right-pointing arrows in Figs. 6 and 7; their abscissas are shown in parentheses and are equal to the azimuth positions

$$\phi_{\varepsilon-} = \phi_C - \varepsilon \tag{24}$$

$$\text{and} \quad \phi_{\varepsilon+} = \phi_C + \varepsilon, \tag{25}$$

respectively, where ε is a function of ρ. Therefore,

$$\phi_C = \frac{[\phi_{\varepsilon-} + \phi_{\varepsilon+}]}{2} \tag{26}$$

$$\text{and} \quad \varepsilon = \frac{[\phi_{\varepsilon+} - \phi_{\varepsilon-}]}{2}. \tag{27}$$

The left frames of Figs. 6 and 7 show that the shape of the curve $e_2(\phi)$ near its global extrema tends to become more rounded as ρ increases. Note that ε increases with ρ.

Figure 8 shows the tangent of each ε angle calculated from (27) as a function of its corresponding ρ value. Clearly,

$$\rho = \frac{R_m}{r_C} = \tan\varepsilon \quad \rho < 0.45, \tag{28}$$

and the same relationship is approximately valid for $0.45 < \rho < 0.75$. This physical relationship is depicted in Fig. 10 by the right-angled triangles $OA_2C$ and $OB_2C$. The triangles reveal that $\phi_{\epsilon-}$ and $\phi_{\epsilon+}$ shown in Fig. 6 ($\rho < 0.45$) delineate the azimuth coordinates of the particular points $A_2$ and $B_2$, which are the points where the orthogonal to the radar beam passing through C intersects the circle whose radius is the RMW. The evaluation of (26), via visual inspection of the values of $\phi_{\epsilon-}$ and $\phi_{\epsilon+}$ shown parenthetically in Figs. 6 and 7, reveals that the value $\phi = 180°$ is retrieved exactly,

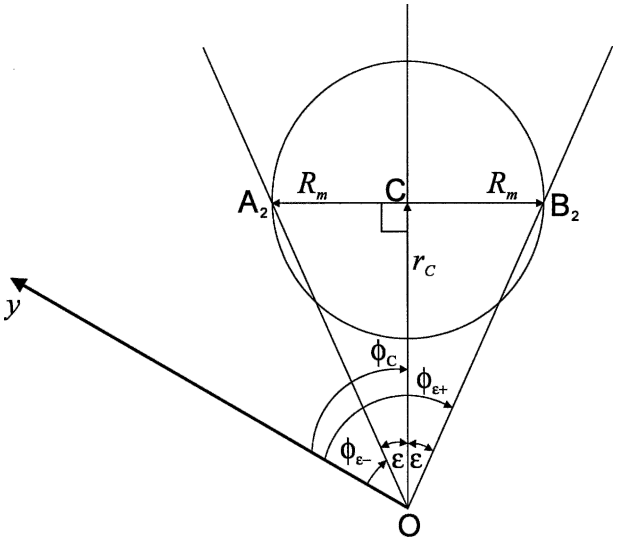


FIG. 10. The physical interpretation for the cusp points of $e_2(\phi)$ shown in Figs. 6–7 is described by the two right-angled triangles $OA_2C$ and $OB_2C$. The radar is located at O and the circulation center of the tropical cyclone is located at C, which has the radar-range coordinate $r_C$ and the azimuth coordinate $\phi_C = \phi_{\epsilon+} - \epsilon = \phi_{\epsilon-} + \epsilon$. The orthogonal to the radar beam passing through C intersects the circle of radius $R_m$ at the points $A_2$ and $B_2$, which have the azimuth coordinates $\phi_{\epsilon-}$ and $\phi_{\epsilon+}$ respectively. The radar beam also passes through these two points thus forming the right-angled triangles $OA_2C$ and $OB_2C$ from which it is clear that $R_m/r_C = \rho = \tan\epsilon$. Given the results of Fig. 8, the abscissas of the cusp points shown in Fig. 6 ($\rho < 0.45$) therefore delineate the azimuth coordinates of the particular points $A_2$ and $B_2$.

even for those values of $\rho$ for which the relationship $\rho = \tan\varepsilon$ is not exact. Note that the cusp points of $\mathbf{e}_2$ are more easily located compared to those of $\mathbf{e}_1$, thus it is the preferred eigenvector for the measurement of $\phi_C$.

As will be shown in Part II of this work, the shape and properties of the eigenvector $\mathbf{e}_2$ are also duplicated in either the eigenvector $\mathbf{a}_1$ or $\mathbf{a}_2$ when one performs a PCA on $\mathbf{C}_r$. Because this eigenvector is common to the azimuth spaces of both $\mathbf{C}_\phi$ and $\mathbf{C}_r$, and since it is the preferred standard for measurement of $\phi_C$, it shall be given a special name for future reference: the standard azimuth eigenvector (SAE). Similarly, the corresponding eigenvectors of the range spaces of both $\mathbf{C}_\phi$ and $\mathbf{C}_r$ ($\mathbf{f}_2$ and $\mathbf{b}_1$ or $\mathbf{b}_2$) shall be referred to as the standard range eigenvector (SRE). Once the SAE is identified via its unique pattern, the SRE is immediately identified; for example, if the eigenvector $\mathbf{a}_1$ or $\mathbf{a}_2$ is identified as the SAE, the SRE is automatically its corresponding normalized principal component $\mathbf{b}_1$ or $\mathbf{b}_2$.

Given a radar azimuthal resolution of $\Delta\phi$, the radar beam does not necessarily pass through the points $A_1$ and $B_1$ in Fig. 9, and through $A_2$ and $B_2$ in Fig. 10. To demonstrate the relationships of Fig. 8 independent of any $\Delta\phi$ effects, $\delta$ and $\varepsilon$ were set to integral multiples of $\triangle\phi = 1°$, then Eqs. (23) and (28) were used to calculate the specific $R_m$ values for those angles with $r_C = 70$ km fixed; thus, the reason for the specific values of $\rho$ in Figs. 4–7. Results of the simulations involving nonintegral multiples of $\delta$ and $\varepsilon$ to be shown in Part II indicate that (23) and (28) are correct to within $\pm\triangle\phi/2$ of the arguments of their trigonometric functions for those values of $\rho$ where the functions are valid. However, $\triangle\phi$ does not affect the accuracy of the measurement of $\phi_C$ since $\delta$ and $\varepsilon$ mathematically cancel in (21) and (26) regardless of any bias that they may contain. Moreover, the choice of $r_C$ has no effect on the accuracy of $\phi_C$ provided that $r_C \leq r_{\max}$ and $\rho < 1$.

### b. Range eigenvectors

The utility of the range eigenvectors $\mathbf{f}_k$ corresponding to the azimuth eigenvectors $\mathbf{e}_k$ shown in Figs. 4–7 is most clearly shown graphically if the values of $R_m$ are rounded to their nearest whole numbers and the PCA redone with these rounded values; all other parameters left unchanged. For an adequate demonstration, it is sufficient to choose the $R_m$ values of 10, 20, 30, 40, 50, and 60 km, which represent the range used for Figs. 4–7 well. Figures 11–14 show $f_k(r)$ versus $r$ for $k = 1$ and 2 constructed from these values of $R_m$. For each value of $k$, the eigenvectors are displayed in sequence from the smallest to largest value of $\rho$. Starting at $\rho = 0.1429$, $f_1(r)$, and $f_2(r)$ each display a maximum near $r = r_C$, which becomes progressively less pronounced and drifts away from $r = r_C$ as $R_m$ ($\rho$) increases; the maximum of $f_1(r)$ drifts toward smaller $r$ values, whereas the maximum of $f_2(r)$ drifts toward larger $r$ values. None of these extrema in the $f_k(r)$ curves have any exact or reliable relationship with the physical properties of the tropical cyclone's wind field. However, there are two inflection points on every curve of $f_1(r)$ and $f_2(r)$ that are physically related to the radial positions of $R_m$ and the tropical cyclone center.

The inflection points of interest for $f_k(r)$, $k = 1$ and 2, are shown by the left- and right-pointing arrows in Figs. 11–14; their abscissas are indicated by the vertical grid-lines and are equal to the range positions

$$r_{C-} = r_C - R_{\mathrm{m}} \tag{29}$$

$$\text{and} \quad r_{C+} = r_C + R_{\mathrm{m}}, \tag{30}$$

respectively. Therefore,

$$r_C = \frac{[r_{C-} + r_{C+}]}{2} \tag{31}$$

$$\text{and} \quad R_{\mathrm{m}} = \frac{[r_{C+} - r_{C-}]}{2}. \tag{32}$$

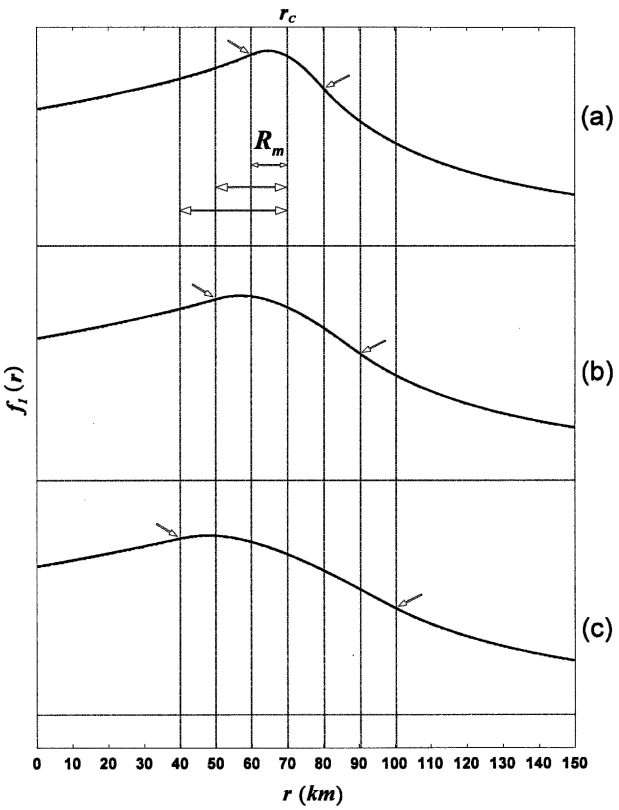


FIG. 11. Eigenvector coefficients $f_1(r)$ vs $r$ derived from simulated PPI scans of the analytic vortex model (17) using (a) $\rho = 0.1429$ ($R_m = 10$ km), (b) $\rho = 0.2857$ ($R_m = 20$ km), and (c) $\rho = 0.4286$ ($R_m = 30$ km). The circulation center of the tropical cyclone is located at $r = r_C = 70$ km. The inflection points of interest are shown by the left- and right-pointing arrows; their abscissas are delineated by the vertical gridlines: the first and last gridlines pertain to the curve (c); the second and second last gridlines pertain to the curve (b); and so on for curve (a). Half of the difference between each of these paired abscissas is $R_m$, indicated by the double arrows. The vertical scale is arbitrary. The horizontal gridlines $f_1(r) = 0$ are shown for each curve.

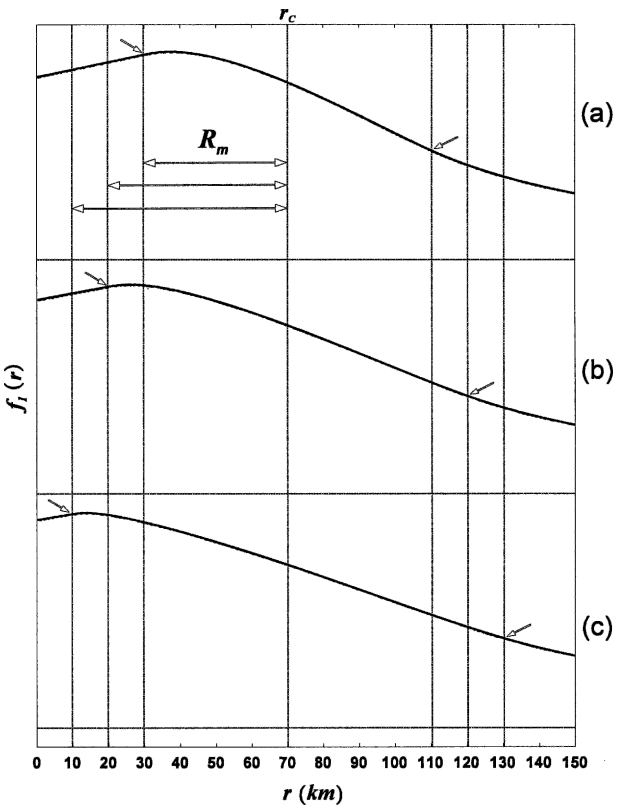


FIG. 12. Same as in Fig. 11 except for (a) $\rho = 0.5714$ ($R_m = 40$ km), (b) $\rho = 0.7143$ ($R_m = 50$ km), and (c) $\rho = 0.8571$ ($R_m = 60$ km).

Figure 15 shows the physical relationship between these parameters. The radar beam passes through the point C and intersects the circle of radius $R_m$ at the points $A_3$ and $B_3$, which have the range coordinates $r_{C-}$ and $r_{C+}$, respectively.

Visual identification of the inflection points shown in Figs. 11–14 is difficult in some cases; simply recording the computer mouse cursor position of the visually estimated locations may result in a reading error of greater than 2 km. To alleviate this problem, it is necessary to accentuate the inflection points by calculating the first derivative of $f_k(r)$, $df_k/dr$. The basis of such an approach is that, by definition, inflection points in $f_k(r)$ occur where $d^2f_k/dr^2 = 0$ or $d^2f_k/dr^2 = \infty$; thus $df_k/dr$ has a local extremum or an inflection point at the abscissas of the inflection points of $f_k(r)$. Of course, the same is also true for $b_k(r)$. Figures 16 and 17 show $df_1/dr$ and $df_2/dr$ corresponding to $f_1(r)$ and $f_2(r)$ shown in Figs. 11–14. There is a clear minimum in both $df_1/dr$ and $df_2/dr$ near $r_{C-}$, which is indicated by the up-pointing arrows. The maximum shown in both $df_1/dr$ and $df_2/$ is often greater than 2 km away from $r_{C+}$, particularly for large values of $\rho$. However, the down-pointing arrows indicate the first inflection point immediately to the right of this maximum that is much nearer to $r_{C+}$. Note that the cusp points of interest indicated by the arrows in Figs. 16 and 17 are much easier to locate along the curve of the SRE ($\mathbf{f}_2$) compared to $\mathbf{f}_1$. Table 2 shows the results of the evaluation of (31) and (32) using the range coordinates corresponding to these cusp points, expressed as biases from the true values of $r_C$ and $R_m$ used in the model. These results suggest that the SRE also provides the most accurate results with mean biases of $-0.062$ and 0.187 km for $r_C$ and $R_m$, respectively, therefore, the SRE is indeed the preferred eigenvector to use.

Figure 18 shows the percentage of the total variance of the modeled data represented by the eigenvectors shown in Figs. 11–14. As $\rho$ increases, the eigenvector $\mathbf{f}_2$ accounts for a much smaller percentage (~7% decreas-

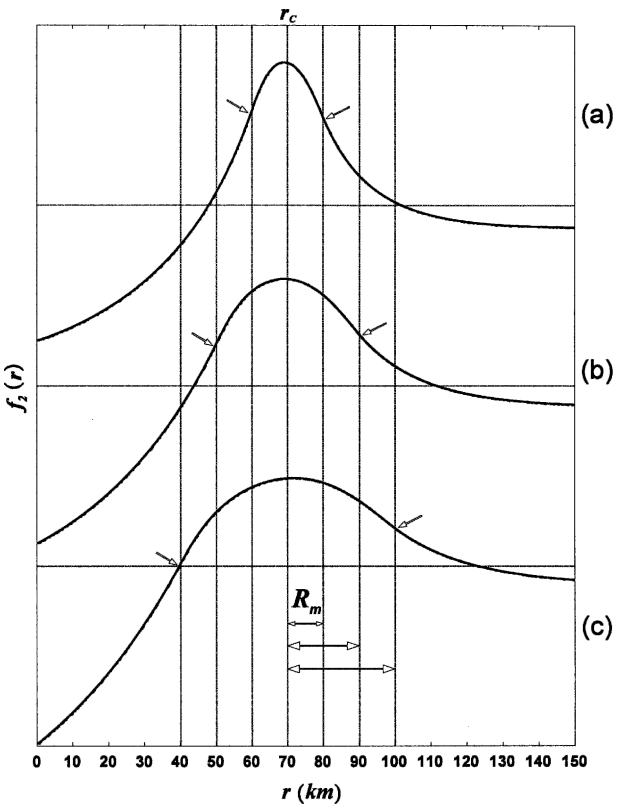


FIG. 13. Eigenvector coefficients $f_2(r)$ vs $r$ derived from simulated PPI scans of the analytic vortex model (17) using (a) $\rho = 0.1429$ ($R_m = 10$ km), (b) $\rho = 0.2857$ ($R_m = 20$ km), and (c) $\rho = 0.4286$ ($R_m = 30$ km). The circulation center of the tropical cyclone is located at $r = r_C = 70$ km. The inflection points of interest are shown by the left- and right-pointing arrows; their abscissas are delineated by the vertical gridlines: the first and last gridlines pertain to the curve (c); the second and second last gridlines pertain to the curve (b); and so on for curve (a). Half of the difference between each of these paired abscissas is $R_m$, indicated by the double arrows. The vertical scale is arbitrary. The horizontal gridlines $f_2(r) = 0$ are shown for each curve.

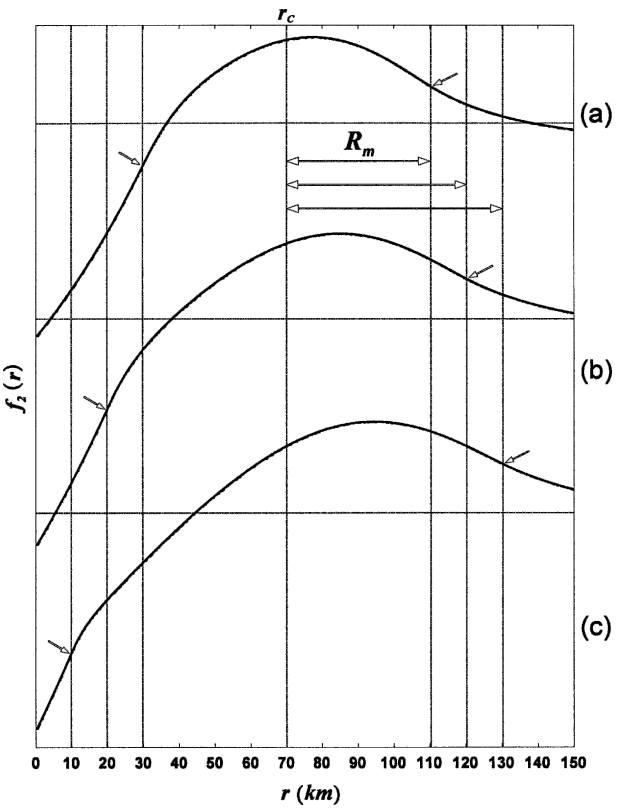


FIG. 14. Same as Fig. 13 except for (a) $\rho = 0.5714$ ($R_m = 40$ km), (b) $\rho = 0.7143$ ($R_m = 50$ km), and (c) $\rho = 0.8571$ ($R_m = 60$ km).

ing to ~2%) as compared to the eigenvector $\mathbf{f}_1$ (~92% increasing to ~98%). Therefore, summing the results for $\mathbf{f}_1$ and $\mathbf{f}_2$ shown in Fig. 18, we arrive at the conclusion that the first two eigenvectors of the $\mathbf{C}_\phi$ eigenspace represent over 99% of the total variance in the data from the analytic model of axisymmetric rotation.

When the tropical cyclone is near the outer reaches of the PPI scan, the cusp point whose abscissa is $r_{C+}$ may not be apparent for some time, or at all, depending on the trajectory of the tropical cyclone. Also, either one of $r_{C-}$ or $r_{C+}$ could be uncertain by an unacceptable amount (say > 1 km, depending on the user's required accuracy) owing to the distortions of the eigenvectors caused by excessive missing data and/or noise. Although such situations may not occur very often, it would be useful to device an alternative means of estimating $r_C$ and $R_m$ that only involves either $r_{C-}$ or $r_{C+}$. The alternative formulation is accomplished by combining (23) and (28)–(30), which yields

$$r_C = r_{C-}/(1 - \sin\delta), \tag{33}$$

$$r_C = r_{C+}/(1 + \sin\delta), \tag{34}$$

$$r_C = r_{C-}/(1 - \tan\varepsilon), \tag{35}$$

$$\text{and} \quad r_C = r_{C+}/(1 + \tan\varepsilon). \tag{36}$$

Then $R_m$ is determined from (29) or (30). Of course the accuracy of (33)–(36) will depend upon the restrictions on $\rho$ given by (23) and (28). Part II of this study will give examples of PCA results from additional analytical models where the cusp point whose abscissa is $r_{C+}$ is not visible.

## 5. Discussion

The methodology for estimating the position of the circulation center and RMW presented above has many advantages over the methods of Wood and Brown (1992), Jou et al. (1996), and Lee and Marks (2000). The essence of the contribution that the PCA method

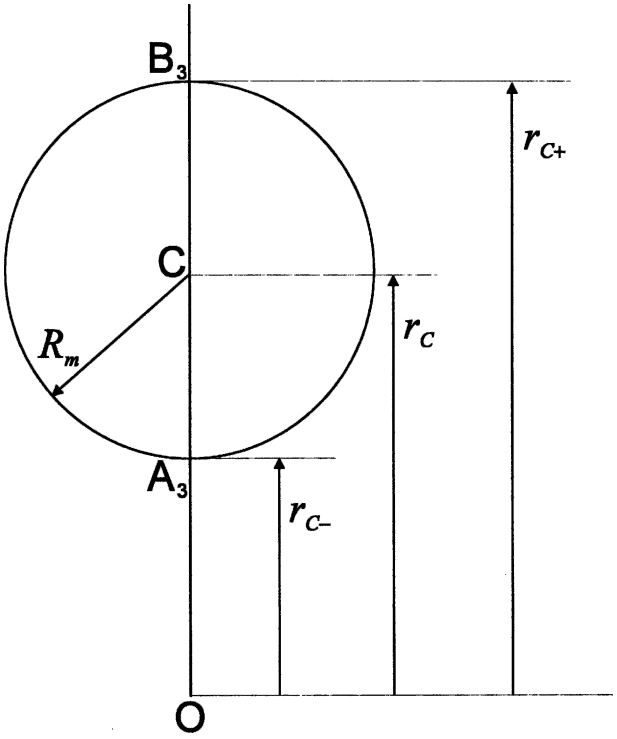


FIG. 15. The physical interpretation for the inflection points of $f_k(r)$ for $k = 1$ and 2 shown in Figs. 11–14. The radar beam passes through the points $A_3$, C, and $B_3$, which have the radar-range coordinates $r_{C-}$, $r_C$, and $r_{C+}$, respectively. The circulation center of the tangential wind is located at the point C, $R_m$ is the RMW.

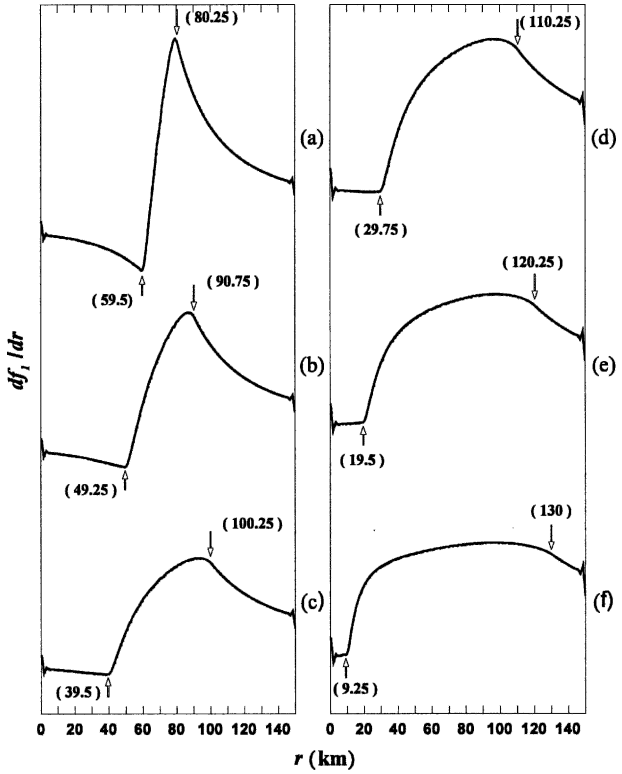


FIG. 16. Here $df_1/dr$ is derived from the curves shown in Figs. 11 and 12; results are ordered in the same sequence. The up- and down-pointing arrows indicate cusps whose range coordinates are shown in parentheses; their values closely approximate the abscissas of the inflection points shown in Figs. 11 and 12.

offers can be understood immediately by a comparison of the general underlying principles of each methodology. Specifically, in contrast to the search for extrema or cusp points within the one-dimensional eigenvector plots that the current PCA application offers, the underlying principle of the other methods is a search for extrema in the data, or extrema in parameters estimated from the data, within the two dimensions of a PPI or CAPPI. Therefore the PCA method offers a reduction in degrees of freedom and complexity involved, and is thus likely to be more robust and accurate when applied to real data that often contain asymmetries and noise in the vicinity of the extrema in the data.

In addition, the geometrically based methods of Wood and Brown (1992) and Jou et al. (1996) seek to find the position of the extrema similar to those shown in Figs. 3e,f; that is, only two Doppler velocity points are utilized in the calculations. Similarly, the algorithm-based method of Lee and Marks (2000) seeks to find the position of an assumed circular ring of maximum winds (whose center is the circulation center) from a Fourier analysis of the Doppler velocity in the vicinity of all candidate rings, via iteration, until the RMW is found; that is, again, the full Doppler velocity field is not utilized. The point here is that, whereas other methods provide a core-local estimate of the position of the circulation center and RMW, the PCA method provides a global estimate of the circulation center and RMW in the sense that it considers all the data in the PPI domain when going from (1) and (2) through to the calculation of the eigenvector coefficients (13)–(16). The significance of this is as follows.

All methods, including the PCA method, will suffer some degree of error in their estimates when noncircular asymmetries are present in the data that violate the underlying assumption of circularity in each method. This is especially true after the tropical cyclone makes landfall, when the tropical cyclone interacts with the terrain in a complex fashion, usually resulting in the production of additional wind asymmetries below 500 m altitude (Powell 1982). For the case of tropical cyclones prior to landfall, the strongest asymmetries at low altitudes are usually found in the core region and on the outskirts where the cyclonic winds blend into the surrounding environment, and the circulation in between these two regions tends to be the most circular

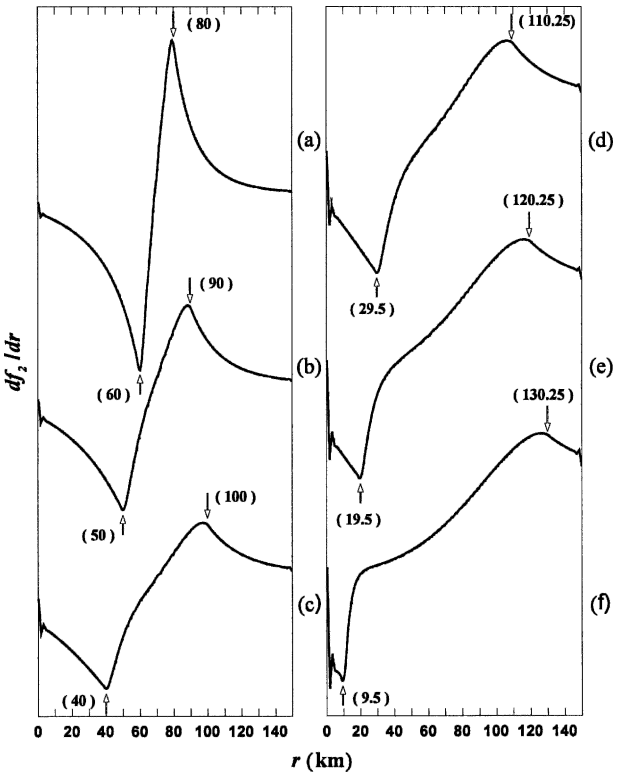


FIG. 17. Same as Fig. 16 except $df_2/dr$ is derived from the curves shown in Figs. 13 and 14.

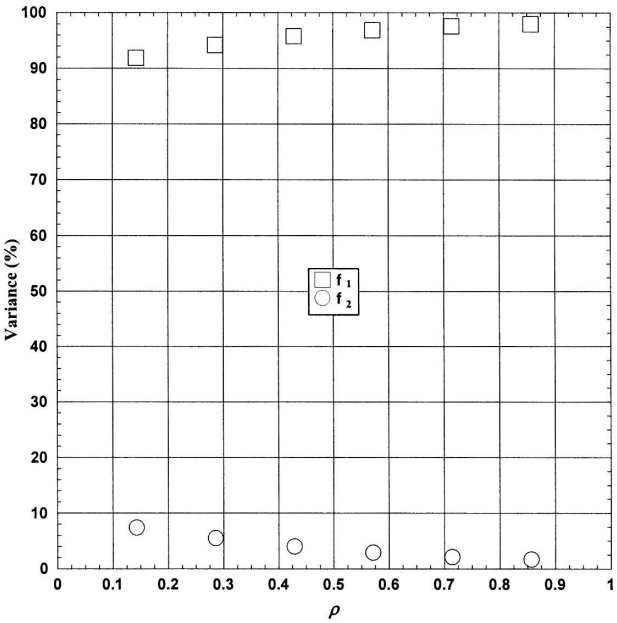


FIG. 18. Percentage of the total variance represented by the eigenvectors shown in Figs. 11–14.

(F. D. Marks, NOAA/AOML/Hurricane Research Division, 1997, and S. Stewart, NOAA/NWS/NCEP/National Hurricane Center, 2000, personal communications). Since there are few or no radar echoes typically found on the outskirts of the cyclonic winds, or in regions of terrain blockage in the landward direction of the coastline as well, the radar primarily observes the asymmetries in the core region. Although the PCA method will include these asymmetries in its calculations, it will also equally weight all the Doppler velocity data found in the region where the circulation is circular. In contrast, the other methods will only include data from the core region where the most significant small-scale asymmetries exist. Thus, the estimated parameters from the PCA method are likely to be less biased by the asymmetries in the core region since it includes more information from the circular region of the storm, and therefore should provide a more robust, global estimate of the position of the circulation center and the RMW.

Harasti et al. (2004) have tested the accuracy of the PCA technique using radar data of very coarse resolution in a study of two tropical cyclones that displayed substantial missing data. The results provided a circulation center estimate for a single-Doppler radar wind analysis method whose results are sensitive to the accuracy of the center that it uses. The winds constructed using the PCA center agreed very well with the available ground truth.

TABLE 2. The measured biases (Δ) in the modeled parameters $r_C$ and $R_m$ as a function of $\rho$, estimated from Figs. 16 and 17 according to (31) and (32) and expressed in units of km. The curves of the same value of $\rho$ in Figs. 16 and 17 are group together by their figure labels and separated by a vertical dashed line with the calculated parameter biases arranged similarly and correspondingly.

| $\rho$ | Figure | | $\triangle r_C$ | | $\triangle R_m$ | |
|---|---|---|---|---|---|---|
| 0.1429 | 16a | 17a | −0.125 | 0.000 | 0.375 | 0.000 |
| 0.2857 | 16b | 17b | 0.000 | 0.000 | 0.750 | 0.000 |
| 0.4286 | 16c | 17c | −0.125 | 0.000 | 0.375 | 0.000 |
| 0.5714 | 16d | 17d | 0.000 | −0.125 | 0.250 | 0.375 |
| 0.7143 | 16e | 17e | −0.125 | −0.125 | 0.375 | 0.375 |
| 0.8571 | 16f | 17f | −0.375 | −0.125 | 0.375 | 0.375 |
| | | Mean | −0.125 | −0.062 | 0.417 | 0.187 |

## 6. Summary and conclusions

For the first time, PCA has been used to analyze single-Doppler radar data. The method employs an $S^2$-mode PCA on the Doppler velocity data taken from a single PPI scan and arranged sequentially in a matrix according to their range and azimuth coordinates. The

data are centered with respect to either their range or azimuth coordinates then the first two eigenvectors of the covariance matrix are calculated in both azimuth and range space. PCA is applied to an analytic model for Doppler velocity corresponding to atmospheric vortices whose wind field is described by a modified Rankine-combined vortex. The first two eigenvectors calculated from the PCA of this model represent over 99% of the total variance in the data and may be used to either analyze or synthesize this data; the current study has focused on the utility of the former approach. When the calculations are performed within azimuth space, the second of these eigenvectors is referred to as the standard range eigenvector, and its corresponding normalized principle component is given the name standard azimuth eigenvector. These eigenvectors are given specific names because they also appear when the calculations are performed within range space instead, and also because they are the preferred standard for measurement purposes owing to their ease of analysis.

Results from the model simulations have established links between the eigenvectors and the position and radial extent of the core region of atmospheric vortices. Specifically, estimates of the position of the circulation center and the radius of maximum wind are obtained by identifying particular cusps in the curves of the eigenvectors' coefficients or their derivatives. The method retrieves the azimuth coordinate of the circulation center exactly from the cusp analyses of the first two azimuth eignenvectors calculated from the model. It is also demonstrated that the standard range eigenvector provides the most accurate estimates of the range to the circulation center and the radius of maximum wind with mean biases of $-0.062$ and 0.187 km, respectively. These model results apply to all atmospheric vortices dominated by axisymmetric rotation with similar values of the aspect ratio $\rho$, in the region where $\rho < 1$ (the radar is located outside the RMW).

The potential advantages of the PCA method compared to other methods that estimate the position of the circulation center and RMW of atmospheric vortices is given in general terms. The two main advantages are 1) the one-dimensional nature of the PCA analysis makes it potentially faster and easier to use with greater robustness and accuracy in the presence of asymmetries and noise in real data, and 2) the PCA method provides a global estimate of the position of the circulation center and the RMW that is likely to be less adversely affected by the local, small-scale asymmetries that impact the other methods in a more direct and all-encompassing way.

A forthcoming study (Part II) will present the PCA results from several combined $V_T$ and $V_R$ analytical models to demonstrate the robustness of the method for various symmetric and asymmetric flow fields commonly observed in atmospheric vortices while another study (Part III) will apply PCA to three tropical cyclones. Part III will summarize and expand upon the application of the PCA technique by Harasti (2000) and Harasti and List (2001) to full-resolution radar data of three different tropical cyclones, thereby further demonstrating the general robustness of the method in the presence of typical missing data and wind asymmetries of real tropical cyclones.

*Acknowledgments.* This paper is based on the Ph.D. dissertation of the first author (Harasti 2000), and was carried out within the research program supported by the Meteorological Service of Canada and the Natural Sciences and Engineering Research Council of Canada, with additional support provided by the University of Toronto Open Doctoral Fellowship and the Ontario Graduate Scholarship programs. Fond memories and appreciation are expressed for the late Professor Reuben Gabriel of the Department of Statistics, University of Rochester for some very enlightening discussions on the subject of PCA. Special thanks go to Mr. Colin McAdie of the NOAA/NWS/NCEP/Tropical Prediction Center/National Hurricane Center (TPC/NHC) who supported the completion of the final chapters of Harasti (2000) simultaneously with the testing of the PCA method on operational WSR-88D data of tropical cyclones, during the first six months of a UCAR postdoctoral fellowship awarded to the first author at TPC/NHC in 2000–01. The support and counsel of Dr. Frank Marks of NOAA/AOML/Hurricane Research Division and Dr. Wen-Chau Lee of NCAR/Earth Observing Laboratory during this period, and for many years previous and since on this subject, are also greatly appreciated. We also thank the Naval Research Laboratory at Monterey for supporting the completion of this manuscript.

## REFERENCES

Bargen, D. W., and R. C. Brown, 1980: Interactive radar velocity unfolding. Preprints, *19th Conf. on Radar Meteorology*, Miami Beach, FL, Amer. Meteor. Soc., 278–285.

Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.,* **115,** 1083–1126.

Bluestein, H. B., W.-C. Lee, M. Bell, C. C. Weiss, and A. L. Pazmany, 2003: Mobile Doppler radar observations of a tornado in a supercell near Bassett, Nebraska, on 5 June 1999. Part II: Tornado-vortex structure. *Mon. Wea. Rev.,* **131,** 2968–2984.

Bradley, S. G., 1997: Acoustic radar studies of rain microphysics. *J. Atmos. Oceanic Technol.,* **14,** 547–553.

Doviak, R. J., and D. S. Zrnic, 1984: *Doppler Radar and Weather Observations*. Academic Press, 458 pp.

Emanuel, K. A., 1991: The theory of hurricanes. *Annu. Rev. Fluid Mech.,* **23,** 179–196.

Gray, W. M., 1982: Tropical cyclone genesis and intensification. *Topics in Atmospheric and Oceanographic Sciences—Intense Atmospheric Vortices, Proceedings of the Joint Symposium (IUTAM/IUGG)*, Reading, UK, L. Bengtsson and J. Lighthill, Eds., Springer Verlag, 3–20.

——, and D. J. Shea, 1973: The hurricane's inner core region. II. Thermal stability and dynamic characteristics. *J. Atmos. Sci.,* **30,** 1565–1576.

Harasti, P. R., 2000: Hurricane properties by principal component analysis of Doppler radar data. Ph.D. dissertation, University of Toronto, 162 pp.

——, and R. List, 2001: Nowcasting hurricane properties by principal component analysis of Doppler velocity data. Preprints, *30th Conf. on Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc., 255–258.

——, C. J. McAdie, P. P. Dodge, W.-C. Lee, S. T. Murillo, and F. D. Marks Jr., 2004: Real-time implementation of single-Doppler radar analysis methods for tropical cyclones: Algorithm improvements and use with WSR-88D display data. *Wea. Forecasting,* **19,** 219–239.

Hillger, D. W., and G. P. Ellrod, 2003: Detection of important atmospheric and surface features by employing principal component image transformation of GOES imagery. *J. Appl. Meteor.,* **42,** 611–629.

Horel, J. D., 1981: A rotated Principal Component Analysis of the interannual variability of the Northern Hemisphere 500-mb height field. *Mon. Wea. Rev.*, **109,** 2080–2092.

Huang, H.-L., and P. Antonelli, 2001: Application of principal component analysis to high-resolution infrared measurement compression and retrieval. *J. Appl. Meteor.,* **40,** 365–388.

Johnson, G. L., and C. L. Hanson, 1995: Topographic and atmospheric influences on precipitation variability over a mountainous watershed. *J. Appl. Meteor.,* **34,** 68–87.

Jolliffe, I. T., 1986: *Principal Component Analysis*. Springer-Verlag, 271 pp.

Jorgensen, D. P., 1984: Mesoscale and convective-scale characteristics of mature hurricanes. Part II: Inner core structure of Hurricane Allen (1980). *J. Atmos. Sci.,* **41,** 1287–1311.

Jou, B. J.-D., S.-M. Deng, and B.-L. Chang, 1996: Determination of typhoon center and radius of maximum wind by using Doppler radar. *Atmos. Sci.,* **24,** 1–23.

Lee, W.-C., and F. D. Marks Jr., 2000: Tropical cyclone kinematic structure retrieved from single Doppler radar observations. Part II: The GBVTD-simplex center finding algorithm. *Mon. Wea. Rev.,* **128,** 1925–1936.

Lhermite, R. M., and D. Atlas, 1961: Precipitation motion by pulse-Doppler radar. Preprints, *Ninth Conf. on Radar Meteorology*, Kansas City, MO, Amer. Meteor. Soc., 218–223.

Liu, Y., D.-L. Zhang, and M. K. Yau, 1999: A multiscale numerical study of Hurricane Andrew (1992). Part II: Kinematics and inner-core structures. *Mon. Wea. Rev.,* **127,** 2597–2616.

Mohr, C. G., L. J. Miller, R. L. Vaughan, and H. W. Frank, 1986: The merger of mesoscale datasets into a common Cartesian format for efficient and systematic analyses. *J. Atmos. Oceanic Technol.,* **3,** 143–161.

Nigam, S., and H.-S. Shen, 1993: Structure of oceanic and atmospheric low-frequency variability over the tropical Pacific and Indian Oceans. Part I: COADS observations. *J. Climate,* **6,** 657–676.

Phillips, D. W., 1972: Modification of surface air over Lake Ontario in winter. *Mon. Wea. Rev.,* **100,** 662–670.

Powell, M. D., 1982: The transition of the Hurricane Frederic boundary-layer wind field from the open Gulf of Mexico to landfall. *Mon. Wea. Rev.*, **110,** 1912–1932.

——, P. P. Dodge, and M. L. Black, 1991: The landfall of Hurricane Hugo in the Carolinas: Surface wind distribution. *Wea. Forecasting,* **6,** 379–399.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier Science, 425 pp.

Rankine, W. J. M., 1901: Motions of fluids. *A Manual of Applied Mechanics*, 16th ed., Charles Griff and Company, 574–578.

Richman, M. B., 1986: Rotation of principal components. *J. Climatol.,* **6,** 293–335.

Riehl, H., 1963: Some relations between wind and thermal structure of steady state hurricanes. *J. Atmos. Sci.,* **20,** 276–287.

Roux, F., and N. Viltard, 1995: Structure and evolution of Hurricane Claudette on 7 September 1991 from airborne Doppler radar observations. Part I: Kinematics. *Mon. Wea. Rev.,* **123,** 2611–2639.

Samsury, C. E., and E. J. Zipser, 1995: Secondary wind maxima in hurricanes: Airflow and relationship to rainbands. *Mon. Wea. Rev.,* **123,** 3502–3517.

Shea, D. J., and W. M. Gray, 1973: The hurricane's inner core region. I. Symmetric and asymmetric structure. *J. Atmos. Sci.,* **30,** 1544–1564.

Stidd, C. K., 1967: The use of eigenvectors for climatic estimates. *J. Appl. Meteor.,* **6,** 255–264.

Williams, C. R., 1997: Principal Component Analysis of wind profiler observations. *J. Atmos. Oceanic Technol.,* **14,** 386–395.

Willoughby, H. E., 1988: The dynamics of the tropical cyclone core. *Aust. Meteor. Mag.,* **36,** 183–191.

——, 1995: Mature structure and evolution. *Global Perspectives on Tropical Cyclones,* R. L. Elsberry, Ed., WMO/TD 693, 21–62.

——, J. A. Clos, and M. G. Shoreibah, 1982: Concentric eye walls, secondary wind maxima, and the evolution of the hurricane vortex. *J. Atmos. Sci.,* **39,** 395–411.

Wood, V. T., and R. A. Brown, 1992: Effects of radar proximity on single-Doppler velocity signatures of axisymmetric rotation and divergence. *Mon. Wea. Rev.,* **120,** 2798–2807.

Wurman, J., J. Straka, E. Rasmussen, M. Randall, and A. Zahrai, 1997: Design and deployment of a portable, pencil-beam, pulsed, 3-cm Doppler radar. *J. Atmos. Oceanic Technol.,* **14,** 1502–1512.