JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 105, NO. A11, PAGES 25,305–25,315, NOVEMBER 1, 2000

Automatic identification of large-scale field-aligned current structures

Tomoyuki Higuchi

Institute of Statistical Mathematics, Tokyo, Japan

Shin-ichi Ohtani

Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

Abstract. The study of large-scale field-aligned currents (LSFACs) has depended on visual inspection of data plots, which is often subjective and time-consuming, and therefore limits the efficiency of analyses. The present paper reports a new procedure to automatically identify the spatial structure of LSFACs from satellite magnetic field measurements. The procedure is based on the concept of the first-order B spline fitting with variable node positions, which may be envisioned as fitting line segments to a line plot. The fitting is made for the maximum variance component of magnetic variations in the plane perpendicular to the background magnetic field. If the distribution of LSFACs can be approximated as an infinite sheet, each slope of the plot corresponds to the crossing of a FAC sheet. The number of node points, which determines the number of FAC sheets, is one of the fitting parameters and is optimized for each orbit so that the Akaike information criterion (AIC) is minimized. Whereas other methods, such as a spherical harmonic fitting, seek to obtain two-dimensional distributions of FACs from assembled data, the present method is basically the automation of the way we visually examine a plot of satellite magnetic field data, and it can be applied even to a single satellite pass. Therefore the procedure can be implemented to real-time data processing and now-casting. The procedure should also provide a powerful tool for data mining. Magnetic field data from the Defense Meteorological Satellite Program–F7 (DMSP–F7) were used for demonstration, and the present paper reports the initial results.

1. Introduction

The characteristics of field-aligned currents (FACs) have been examined for more than three decades since FACs were initially detected by a low-altitude satellite at the very early stage of satellite observation [e.g., *Zmuda et al.,* 1966; *Cummings and Dessler,* 1967]. Large-scale FACs (LSFACs) are classified into three systems, that is, region 2 (R2), region 1 (R1), and region 0 (R0) systems from equatorward to poleward [*Iijima and Potemra,* 1976]. Whereas the R0 system is distributed in the midday sector, the R2 and R1 currents encircle a magnetic pole but are off-center toward midnight. The R1 FAC tends to flow toward and away from the ionosphere in the morning and evening sectors, respectively, and at a given local time the polarity of a R2 current is the opposite to that of a R1 current. Whereas such average properties of large-scale FACs are well accepted, it is also known that the spatial structure and intensity of these current systems are far from stationary and depend on both external and internal conditions.

The FAC has been studied based mostly on the observation of magnetic perturbations by low-altitude spacecraft. Provided that LSFACs extend in the east-west direction, a satellite observes a changing azimuthal magnetic component as it crosses FAC sheets. Each slope of the azimuthal component corresponds to a crossing of a FAC sheet. The flowing polarity (upward or downward) and intensity of each FAC can be determined from the sense (increase or decrease) and magnitude of the slope, respectively. However, magnetic perturbations, and therefore FACs, are generally structured, and it is not always straightforward to identify LSFAC systems. Moreover, in general, a FAC sheet is inclined from the east-west direction, and its extent is finite, causing variations of the latitudinal magnetic component. Despite such a complex nature of FACs, the study of FACs has depended on visual inspection of data plots, which (1) limits the efficiency of the analysis of FACs since such visual inspection is both subjective and time-

Paper number 2000JA900073.
0148-0227/00/2000JA900073$09.00

consuming and (2) often lacks quantitative evaluation. The development of an automatic procedure to identify LSFACs with quantitative criteria is most desirable.

The present paper reports an automatic procedure we developed to determine the structure of LSFACs and its application to magnetic field data from the Defense Meteorological Satellite Program–F7 (DMSP–F7) satellite. Section 2 describes the DMSP-F7 satellite and magnetometer data acquired from this satellite. The procedure rotates the magnetic field perturbation in the horizontal plane so that one axis is aligned with FAC sheets. Section 3 explains this coordinate transformation. Section 4 describes a model for describing magnetic perturbations and explains how the procedure searches for the optimal fit. In section 5 we apply the procedure to the DMSP-F7 data. In section 6 we briefly compare the present procedure with other methods and comment on future applications.

2. Data

DMSP is a Sun-synchronous satellite with a nearly circular polar orbit at 835 km in altitude, with its ascending and descending nodes at 1030 and 2230 LT(local time), respectively. The orbital period is about 101 min. The satellite was operational from December 1983 to January 1988.

The magnetometer on board [*Rich et al.*, 1985] measures three magnetic components every second. Data were provided in the form of the difference between the measured and model magnetic fields, $\mathbf{B} = [B_X, B_Y, B_Z]'$ (prime denotes the transposition), in the spacecraft coordinate system. In this system the X axis is directed vertically downward, the Y axis is in the direction of the projection of the satellite orbital velocity onto the plane perpendicular to the X axis, and the Z axis completes a right-hand orthogonal system. The magnetic field is approximately vertical in the high-latitude region. For example, at a magnetic latitude of 70°, the inclination of the dipole magnetic field is 80°, and the transverse magnetic perturbation projected onto the Y–Z plane keeps 98% of the original amplitude; the ratio is even larger at higher latitudes. Therefore we can focus on the Y and Z magnetic components to examine magnetic perturbations, $\mathbf{B}$, caused by FACs.

We divide each polar crossing into two parts, dayside and nightside passes, based on the highest latitude satellite position. A total of 71,594 data files were created for the entire interval of the satellite operation. Each data file, $\mathbf{B}_n$ $(n = 1, \ldots, N)$, typically consists of 600-800 measurements and should include no more than one crossing of LSFACs.

3. Coordinate Transformation

3.1. Preparatory Determination of the Interval of the Satellite Crossing of LSFACs

If a LSFAC has a sheet structure, the associated magnetic perturbation is parallel to the current sheet; see Figure 1. We determine the orientation of LSFAC sheets by applying the principal component analysis (PCA) to $\mathbf{B}_n$; the sheet orientation is defined as the orientation of the principal axis. The PCA is applied only to the interval of the satellite crossing of LSFACs. In this subsection we explain how our procedure determines such an interval for each orbit.

In general, each data file includes measurements made equatorward (E), inside (F), and poleward (P) of LSFAC systems. We denote the corresponding intervals as $\mathcal{I}_Z^E$, $\mathcal{I}_Z^F$, and $\mathcal{I}_Z^P$, in this order. Here the subscript Z explicitly indicates that the classification is made on the basis of the Z component measurement. (In general, the east-west component should be used for the present purpose.) Since LSFACs extend in the east-west direction, which is close to the Z direction, the adoption of the Z component for the preparatory classification can be justified; we will later redefine each interval based on the principal axis component.

Now we have to introduce a parameter that quantifies the goodness of the classification. Although magnetic variations are expected to be small outside of LSFAC systems, $B_{Z,n}$ often changes gradually in $\mathcal{I}_Z^E$ and $\mathcal{I}_Z^P$. Such a slope is possibly an effect smearing out from LSFAC systems. Note that magnetic perturbations are completely confined in the region of LSFACs only if LSFACs have an infinitely extending sheet structure and the intensities of upward and downward flowing FACs are balanced. Therefore the smoothness of B_Z variations for $\mathcal{I}_Z^E$ and $\mathcal{I}_Z^P$ should be evaluated based on the variance of the linearly detrended signal, instead of the original signal. We introduce U_Z^E and U_Z^P, which are defined as

$$U_Z^E = \frac{1}{L_Z^E - 1} \sum_{n=1}^{L_Z^E - 1} \left[B_{Z,n} - \left(a^E n + b^E \right) \right]^2 \quad (1)$$

and

$$U_Z^P = \frac{1}{N - L_Z^P} \sum_{n=L_Z^P+1}^{N} \left[B_{Z,n} - \left(a^P n + b^P \right) \right]^2, \quad (2)$$

where a^E and b^E (a^P and b^P) are the results of the least squares fit for $\mathcal{I}_Z^E$ ($\mathcal{I}_Z^P$). As a parameter quantifying the smoothness of B_Z variations for $\mathcal{I}_Z^E$ and $\mathcal{I}_Z^P$, we adopt the average of U_Z^E and U_Z^P weighted by the number of data points of the corresponding interval,

$$U_Z^{E,P} = \frac{(L_Z^E - 1)U_Z^E + (N - L_Z^P)U_Z^P}{L_Z^E - 1 + N - L_Z^P}. \quad (3)$$

One may attempt to determine an optimal $\mathcal{I}_Z^F$ by maximizing the ratio of the variance of $B_{Z,n}$ in $\mathcal{I}_Z^F$ to $U_Z^{E,P}$. However, this approach sometimes misidentifies the boundaries of LSFACs because of smaller-scale magnetic variations embedded in large-scale ones associated with LSFACs; the boundaries of $\mathcal{I}_Z^F$ determined tend to be located far inside of the actual LSFAC region.

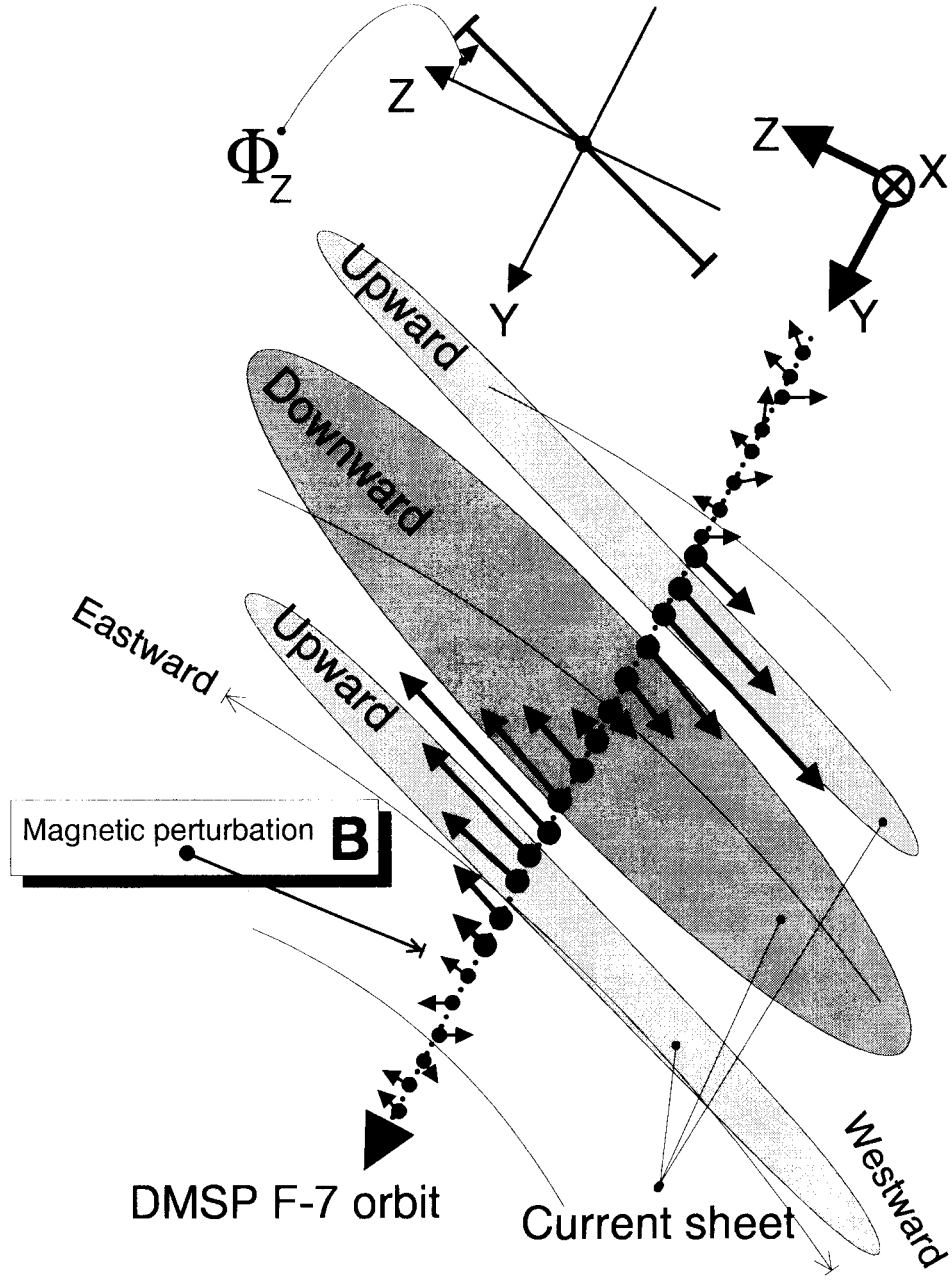

Figure 1. A schematic illustration of the large-scale field-aligned current (LSFAC) structure in the horizontal $(Y-Z)$ plane. The arrows represent magnetic perturbations expected to be observed along the satellite pass. An inclination of the orientation of the FAC sheets from the Z axis is denoted as Φ_Z.

Therefore we need to smooth the data for the interval of $[L_Z^E, L_Z^P]$.

The smoothing procedure we use in this study is called the Bayesian smoothing procedure [*Kitagawa,* 1981], which is based on a state space model with the Kalman filter and smoothing algorithms [*Anderson and Moore,* 1979]. We preliminarily examined 300 DMSP data files and found that the mean interval of the satellite crossing of LSFACs systems is 133 data points, which corresponds to 9° in latitude if the satellite orbit is along a meridian. We adopted the inverse of this length as the half bandwidth of the filter; for a different data set, the half bandwidth needs to be predetermined by converting this spatial scale (9°) to the number of data points. In the Bayesian smoothing procedure the smoothness of the result is controlled by one trade-off parameter that is usually specified by τ^2. The relationship between the half bandwidth and this trade-off parameter has been examined numerically by *Higuchi* [1991]. The smoothed data obtained by this procedure are denoted by $\widetilde{B}_{Z,n}$ $(n = 1, \ldots, N)$ henceforth.

Whereas the value of $\widetilde{B}_{Z,n}$ itself does not distinguish $\mathcal{I}_Z^F$ from its outside regions ($\mathcal{I}_Z^E$ and $\mathcal{I}_Z^P$), the charac-

teristics of the first-order difference of $B_{Z,n}$, especially its variance, are significantly different between the two regions. For $\mathcal{I}_Z^E$ and $\mathcal{I}_Z^P$, $\widetilde{B}_{Z,n}$ tends to change linearly, indicating that the first-order difference has similar values, and therefore its variance is small. In contrast, in $\mathcal{I}_Z^F$, the variance of the first-order difference should be large, because the variation of $\widetilde{B}_{Z,n}$ is structured. We use the variance of the first-order difference of $\widetilde{B}_{Z,n}$, $\widetilde{U}_Z^F$, as a quantity to characterize the disturbances observed in $\mathcal{I}_Z^F$.

The optimal interval is determined by maximizing the ratio of $\widetilde{U}_Z^F$ to $U_Z^{E,P}$,

$$Q(L_Z^E, L_Z^P) = \frac{\widetilde{U}_Z^F}{U_Z^{E,P}}, \tag{4}$$

where

$$\widetilde{U}_Z^F = \frac{1}{L_Z^P - L_Z^E} \sum_{n=L_Z^E+1}^{L_Z^P} \left[\left(\widetilde{B}_{Z,n} - \widetilde{B}_{Z,n-1} \right) - b^F \right]^2 \tag{5}$$

and $b^F = 1/(L_Z^P - L_Z^E) \sum_{n=L_Z^E+1}^{L_Z^P} \left(\widetilde{B}_{Z,n} - \widetilde{B}_{Z,n-1} \right)$.

A fine search for $[L_Z^E, L_Z^P]$ requires computational time, even though this step of the procedure is only for determining the interval for the PCA. For DMSP-F7 data, we decided to use data every 10 points at this step. Ten data points correspond to a latitudinal width of 0.6° for DMSP-F7 data. The resolution should be sufficient for determining $\mathcal{I}_Z^F$.

3.2. Maximum Variance Direction

We perform the PCA for $(B_{Y,n}, B_{Z,n})$ $(n = L_Z^E, \ldots, L_Z^P)$, for $\mathcal{I}_Z^F$. The square root of the ratio of the maximum to the minimum eigenvalues is denoted by α. Fluctuations are isotropic if $\alpha = 1$. Increasing in α, the observed magnetic fluctuations tend to lie in a certain direction, which corresponds to the orientation of LSFAC systems.

An ambiguity remains concerning the direction of the maximum variance orientation. We choose the direction of the new coordinate so that the angle from the positive Z direction Φ_Z is smaller (see Figure 1). Measured magnetic vectors are transformed from the satellite coordinate system to the maximum–minimum variance coordinate systems by rotating by Φ_Z in the Y–Z plane. In this new coordinate system the components in the maximum and minimum variance directions are denoted by B_A and B_L, respectively. B_A is the component parallel to FAC sheets and is usually close to the "A"zimuthal direction, whereas B_L can be regarded as the "L"atitudinal component. The geological directions of the "A" and "L" axes, that is, eastward/westward and poleward/equatorward, vary depending on the hemisphere (Northern/Southern) and the local time (dayside/nightside) just as those of the Z and Y axes do. If LSFACs systems have a sheet structure, the associated magnetic variation should be confined in B_A. Only $B_{A,n}$ $(n = 1, \ldots, N)$ will undergo a further procedure to identify the structure of LSFACs.

4. A Model for Magnetic Field Perturbation

4.1. Linear Spline With Variable Nodes

A measurement $B_{A,n}$ $(i = 1, \ldots, N)$ consists of a true value $B_{fit,n}$ and an observation noise e_n, that is, $B_{A,n} = B_{fit,n} + e_n$. The observation noise e_n is an independently and identically distributed (i.i.d.) white noise sequence that has a Gaussian (normal) distribution with mean 0 and unknown variance σ^2. Our procedure approximates $B_{fit,n}$ by a first-order B spline function, that is, a sequence of straight lines [*Williams*, 1978], which is sometimes called a polyline (J. L. Maryak et al., unpublished manuscript, 1997) or simply, a linear spline [*Ja-Yong*, 1997]; we call it polyline henceforth. A polyline that has J node points can be specified by $\mathbf{T}_J = \{t_j | j = 1, \ldots, J\}$ and g_j $(j = 1, \ldots, J)$ along with values at both ends, $g_0 = f_1$ and $g_{J+1} = f_N$, where g_j is the value at the node position t_j. A set of g_j $(j = 0, \ldots, J+1)$ for a polyline with J nodes is denoted by $\mathbf{G}_J$ henceforth.

The optimal set of g_j (for a given number of node points) can be determined so that it minimizes the residual sum of squares (RSS) defined by

$$\mathrm{RSS}_J = \sum_{n=1}^{N} \left[B_{A,n} - B_{fit,n} \right]^2. \tag{6}$$

The minimization of (6) can be reduced to a linear least squares problem after an appropriate linear transformation of g_j, which can then be solved by a normal numerical procedure [*Hiragi et al.*, 1985].

Although node points are fixed in most applications, the benefit of the spline fitting can be maximized when node points are allowed to move [*Hiragi et al.*, 1985]. We therefore treat $\mathbf{T}_J$ as variables to be optimized. The minimization of RSS_J is nonlinear with respect to $\mathbf{T}_J$. We initially tried to use the quasi–Newton method, specifically, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) updating formula [*Hiragi et al.*, 1985]. However, it turned out that this approach requires significant computation, therefore making it unfeasible to deal with a large data set. Furthermore, in that procedure, two node positions sometimes intersect in the course of iteration. We therefore developed a method to give a better initial (starting) point for $\mathbf{T}_J$ in the J-dimensional space, which we will describe in the next subsection.

4.2. Initial Guess on Node Position

It is reasonable to expect that optimal node points are located near local maximum and minimum points

of large-scale variations. We select the positions of local maxima and minima of smoothed data as candidates for optimal node points. There are two considerations to take into account with regard to the smoothing. First, the characteristic length of large-scale variations is different for different events, and therefore the half bandwidth of smoothing must be determined for each data file. Second, the characteristics of magnetic fluctuations are different inside and outside the field-aligned current region. It is therefore necessary to build different smoothing filters for different regions.

The procedure begins with dividing the interval of a data file into three subintervals: $\mathcal{I}_A^E = [1, L_A^E - 1]$, $\mathcal{I}_A^F = [L_A^E, L_A^P]$, and $\mathcal{I}_A^P = [L_A^P + 1, N]$. We determine these intervals in the same way we did in section 3.1, except that we use B_A instead of B_Z.

Figure 2a shows the B_A component measured on the dayside in the Northern Hemisphere for the eleventh orbit of August 26, 1986. The value of α was 6.3, and Φ_Z was 9.0° for this orbit. The two long vertical arrows drawn from the top to the bottom of the figure mark the interval of $[L_A^E, L_A^P]$.

The next step is to determine a characteristic spatial scale for each subinterval. For this purpose we calculate autocorrelation, which is defined, for the interval $\mathcal{I}_A^F$, as

$$R^F(k) = \sum_{n=L_A^E}^{L_A^P-k} \left(B_{A,n+k} - \overline{B}_A^F\right)\left(B_{A,n} - \overline{B}_A^F\right) / \left(L_A^P - L_A^E - 1\right) \quad \text{for } k = 0, \ldots, \tag{7}$$

where $\overline{B}_A^F$ is the mean of $B_{A,n}$ for $n \in \mathcal{I}_A^F$. The characteristic spatial scale λ^F is defined as $4k_c$, where k_c is the value at which $R^F(k_c)$ first turns negative. Similarly, λ^E and λ^P are obtained from $R^E(k)$ and $R^P(k)$, respectively.

The inverse of the obtained characteristic spatial scale is adopted as the half bandwidth for designing the low-pass filter based on the Bayesian smoothing procedure. The data for each interval are smoothed with the corresponding low-pass filter. The results, which hereafter will be presented as $\widetilde{B}_{A,n}^E$, $\widetilde{B}_{A,n}^F$, $\widetilde{B}_{A,n}^P$, are plotted by the smooth curves in Figure 2a. Note that the smoothness of these three curves is different for each curve.

As the candidates of node points, we select points of local maxima or local minima of the smoothed data:

$$\mathbf{T}^E = [t_1^E, \ldots, t_{M^E}^E], \tag{8}$$

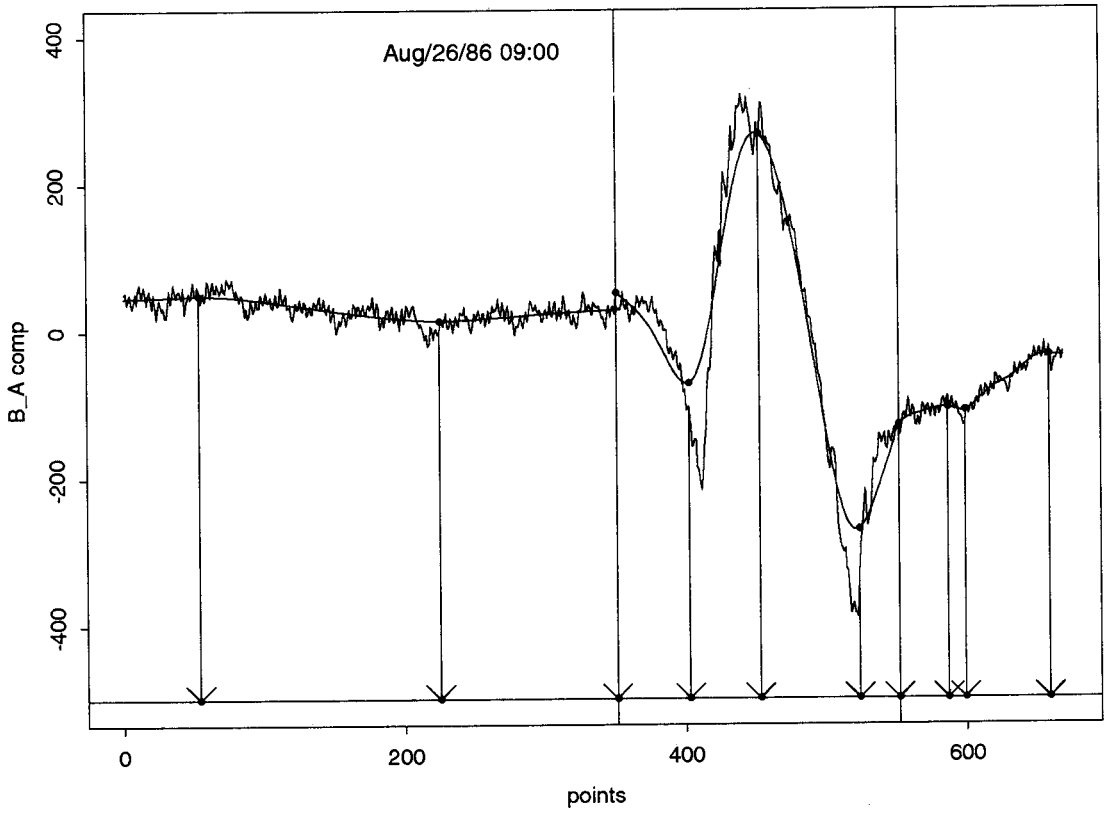

Figure 2a. The magnetic field perturbation of the B_A component observed in the northern dayside region for the eleventh polar pass of August 26, 1986. The two long vertical arrows separate the whole interval into three subintervals, $\mathcal{I}_Z^E$, $\mathcal{I}_Z^F$, and $\mathcal{I}_Z^P$. Data are smoothed by a low-pass filter defined for each subinterval, and the result is represented by the smoothed curves. The locations of the local maximum and minimum points of these smooth curves are marked by the shorter vertical arrows.

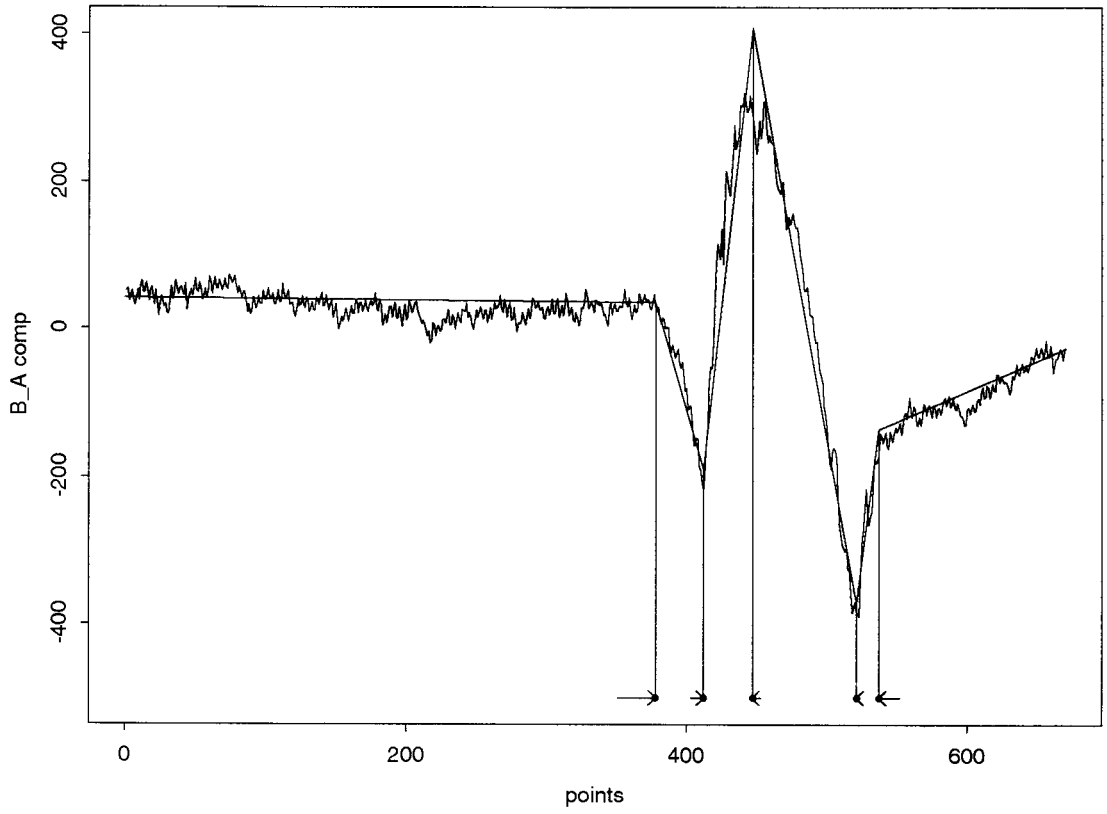

Figure 2b. The fitted polyline together with the original data shown in Figure 2a. The five solid circles at the bottom indicate the optimal node positions. The horizontal arrows represent the adjustments made by the quasi-Newton method.

$$\mathbf{T}^F = [t_1^F, \ldots, t_{M^F}^F], \text{and} \quad (9)$$

$$\mathbf{T}^P = [t_1^P, \ldots, t_{M^P}^P], \quad (10)$$

where M^E, M^F, and M^P are the sum of the numbers of local maximum and minimum points for each subinterval. In Figure 2a, the positions of node candidates obtained are marked by the arrows drawn downward from the plots of the smoothed data. For the present example, $M^E = 2$, $M^F = 3$, and $M^P = 3$. The edges of the LSFAC interval, L_A^E and L_A^P, are added to a set of candidates. We combine all candidates:

$$\mathbf{T}_M = [t_1^E, \ldots, t_{M^E}^E, L_A^E, t_1^F, \ldots, t_{M^F}^F, L_A^P, t_1^P, \ldots, t_{M^P}^P]. \quad (11)$$

Here $M = M^E + M^F + M^P + 2$; $M = 10$ for this example.

4.3. Determination of Node Positions

We now apply the polyline fit with J node points to $B_{A,n}$. We confine the value of J to no more than 7. There are $K_J = {}_M\mathrm{C}_J$ ways to choose J node points from $\mathbf{T}_M$. The procedure examines all possible combinations, $\mathbf{T}_J(k)$ $(k = 1, \ldots, K_J)$.

For a given $\mathbf{T}_J(k)$, the procedure calculates the residual sum of squares $\mathrm{RSS}_J(k)$ to evaluate the goodness of a designed spline fit. The polyline fit that has the minimum $\mathrm{RSS}_J(k)$, $\mathrm{RSS}_J^*(k)$, can be regarded as optimal for that value of J, and we denote the corresponding $\mathbf{T}_J(k)$ as $\mathbf{T}_J^*$.

The final step of the polyline fit is to determine the optimal number of node points, J^*. Here we adopt the Akaike information criterion (AIC) [*Akaike*, 1974; *Sakamoto et al.*, 1986], which is defined for J node points as [*Hiragi et al.*, 1985],

$$\mathrm{AIC}_J = N \log\left(\frac{\mathrm{RSS}_J^*}{N}\right) + 4J + \text{constant}, \quad (12)$$

where the constant is independent of the model selection. The AIC evaluates the efficiency of fitting in terms of the number of fitting parameters, J in the present case. RSS is a decreasing function with J; in other words, the measurements can be better fitted with more node points. Thus the value of RSS alone does not determine J. The second term on the right-hand side serves as a penalty against using more parameters. The optimum number of FAC sheets J^* is determined as the one that gives minimum AIC_J among AIC_J $(J = 1, \ldots, 7)$, and accordingly the optimal combination of node points is selected; we refer to the corresponding RSS as RSS_{J^*}.

Figure 2b shows the best model for the example of Figure 2a, for which J^* was found to be 5. The positions of $\mathbf{T}_J^*$ $(= \mathbf{T}_{J^*}^*)$ are marked by the solid circles.

The positions of the local maximum or minimum of the smoothed data do not always agree with those of the original data, and the horizontal arrows in Figure 2b indicate the results of this final adjustment of $\mathbf{T}_j^*$. In general, the adjustment is insignificant except for the edges of the LSFAC region.

5. Analysis of DMSP–F7 Data

5.1. Preparatory Check

Data files often include no magnetic signature that can be identified as LSFACs. There are at least two reasons for this. One is that LSFAC systems disappear or diminish in magnitude during geomagnetically quiet periods, and the other is that an orbit is not high enough in latitude to traverse the LSFAC region. The procedure requires significant computation to make a spline fit for such files because of the absence of coherent large-scale signatures, whereas the output does not include any useful information. To remove such files in advance, the procedure examines the following three quantities: (1) variance of the first-order difference of the smoothed data, $\widetilde{B}_{Z,n}$, within the field-aligned current region: $\widetilde{U}_Z^F$; (2) spatial scale between the equatorward and poleward edges of the LSFACs: $|L_Z^E - L_Z^P|$; and (3) range of $\widetilde{B}_{Z,n}$ within the field-aligned current region (i.e., $n \in \mathcal{I}_Z^F$). Here the region of LSFACs refers to the result of the step mentioned in section 3. The procedure regards that a data file does not include a LSFAC signature if any one of the following three conditions is satisfied: (1) $\widetilde{U}_Z^F \leq 1$ (nT/s); (2) a spatial length is larger than 600 data points ($\sim$ 4,000 km); (3) the range is smaller than 25 (nT).

The whole DMSP–F7 data set consists of a total of 71,594 data files. After removing data files that do not include any clear FAC signatures as well as data files that include significant data gap or artificial signatures, the procedure analyzed 45,587 data files.

5.2. Selection Criteria of LSFACs

The polyline fit represents the observed sequence of magnetic field variations with a set of (t_j, g_j) $(j = 0, \ldots, J^* + 1)$, where t_j and g_j represent a node position and a value at that position, respectively. Each segment, except for the most equatorward and most poleward ones, corresponds to a LSFAC; the jth current system is given by the segment for the jth interval, the duration of which is $I_j = t_{j+1} - t_j$ $(j = 0, \ldots, J^*)$. The intensity and density of a LSFAC are given by the level shift and the slope of a segment, respectively, which are defined as

$$H_j = g_{j+1} - g_j, \tag{13}$$

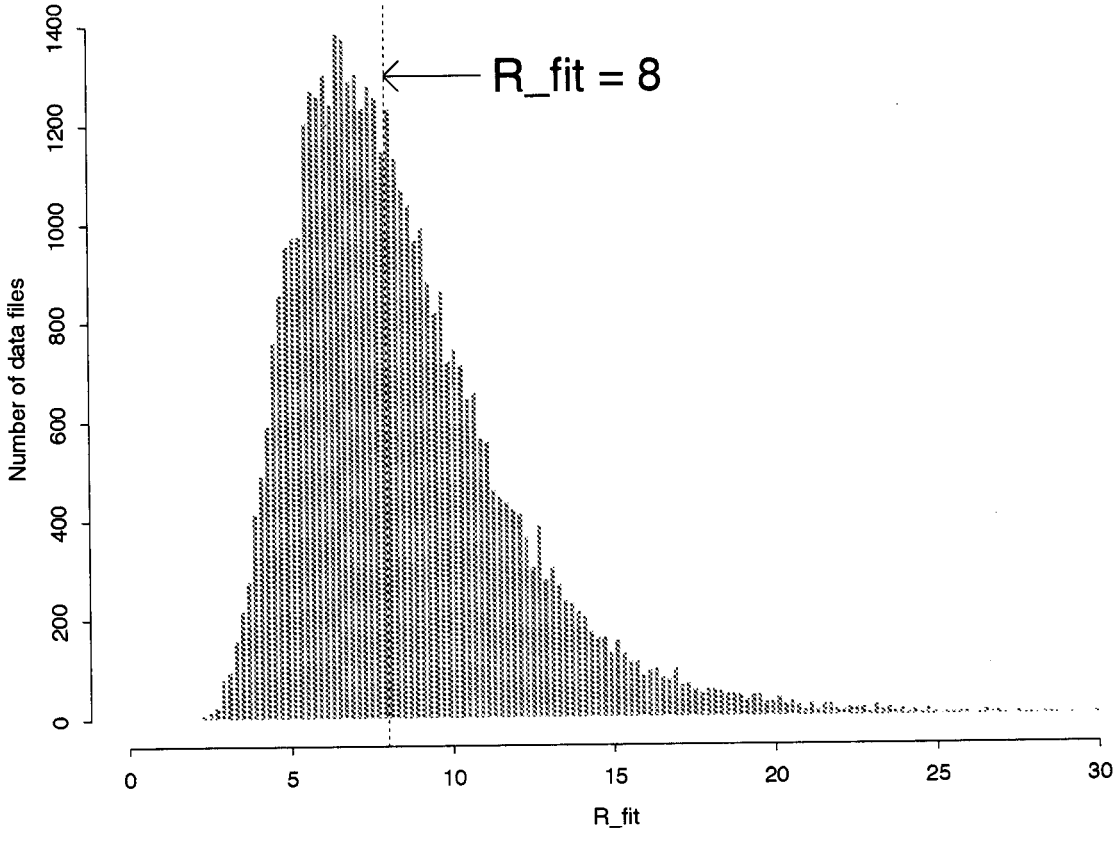

Figure 3. Histogram of R_{fit} for the 45,587 files.

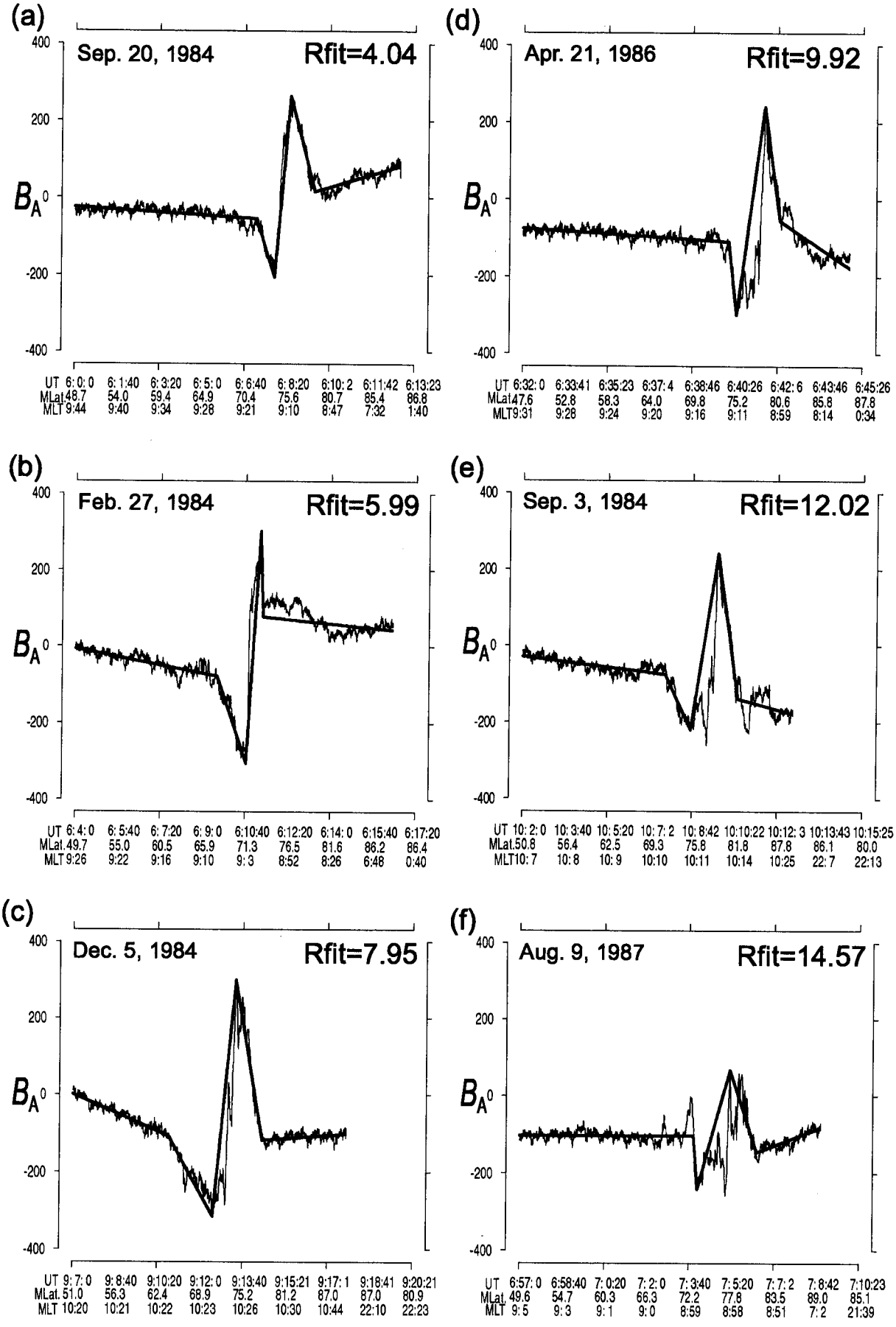

Figure 4. Examples of three-FAC-sheet structures with different R_{fit}.

and

$$\delta_j = \frac{g_{j+1} - g_j}{t_{j+1} - t_j} \quad \text{for} \quad j = 0, \ldots, J^*. \qquad (14)$$

We denote the maximum absolute value of H_j as H_{max}. In this study a segment satisfying all following conditions is regarded as a LSFAC: (1) $|H_j| \geq \gamma_{th} H_{max}$, (2) $|H_j| \geq H_{th}$, and (3) $|\delta_j| \geq \delta_{th}$, where γ_{th}, H_{th}, and δ_{th} are set to be $\gamma_{th} = 0.3$, $H_{th} = 50$ (nT), and $\delta_{th} = 1.2$ (nT/s) (0.128 (μA/m^2)). Both the most equatorward and most poleward segments are also reevaluated based on the above criteria; this is necessary especially for the most poleward interval since for a skimming orbit the boundary between dayside and nightside passes could be located in a LSFAC system.

The fitness of a polyline is often assessed by a root-mean-square error: $\sigma^* = \sqrt{\mathrm{RSS}_{J^*}/N}$. However, because the intensity of LSFACs is highly variable, we found that a certain normalization is necessary. We adopted H_{max} for this purpose. That is, we use $R_{fit} = \sigma^*/H_{max}$ to evaluate the fitness.

Figure 3 shows the histogram of R_{fit} of the 45,587 data files. Figure 4 shows six examples of three-sheet structures in the order of increasing R_{fit}. The result of the fitting is superposed to measurements in each plot. As R_{fit} decreases, the fitting gets closer to the visual inspection. Data files with $R_{fit} \leq 8$ underwent further investigation.

In addition, the square root of the ratio of the larger to smaller eigenvalue in the PCA, α, is important in evaluating the structure of LSFACs. Figure 5 shows the histogram of α for the 45,587 files; 80 data files with $\alpha \geq 15$ are excluded from this figure. For further investigation, α must be greater than 2. A total of 21,310 files (46.7%) satisfied both $R_{fit} \leq 8$ and $\alpha \geq 2$.

Figure 6a shows the histogram of the width of each LSFAC sheet, I_j, for the 21,310 data files. I_j is distributed from a few points up to hundreds of points. The histogram of the absolute value of the level shift of each LSFAC, $|H_j|$, is shown in Figure 6b. H_j is also widely distributed in the range from 50 (nT) to more than 1,000 (nT). The combination of the two panels elucidates a wide variety of LSFAC structures.

5.3. Determination of Number of Sheets

The number of LSFAC sheets for each file is one of the final outputs of this procedure. Special caution is required for a node point that is not located at either a local maximum or a local minimum of $\mathbf{G}_{J^*}$, although we found that this is not very common. Such a node makes a kink in a polyline, where the slope changes but $B_{fit,n}$ decreases or increases on both sides. In such a case we count two such segments as a single LSFAC sheet.

Table 1 lists the number of data files that satisfy $R_{fit} \leq 8$ and $\alpha \geq 2$ for different numbers of LSFAC sheets.

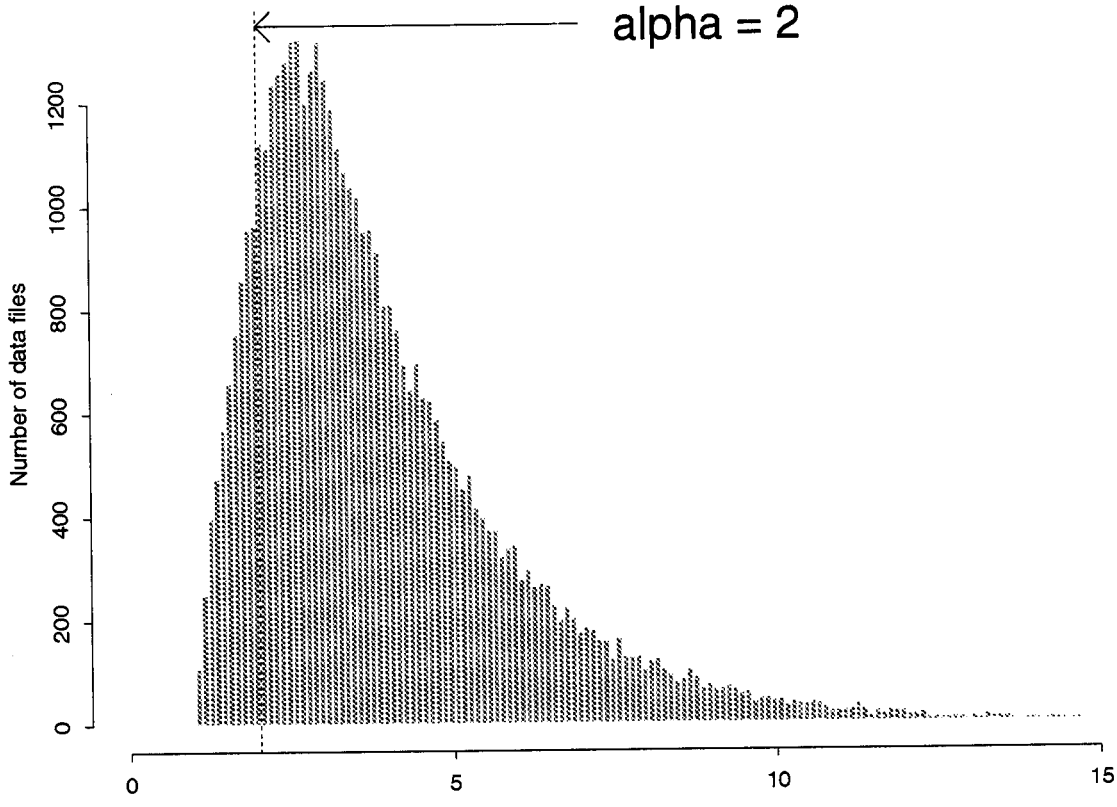

Figure 5. Histogram of α for the 45,587 files.

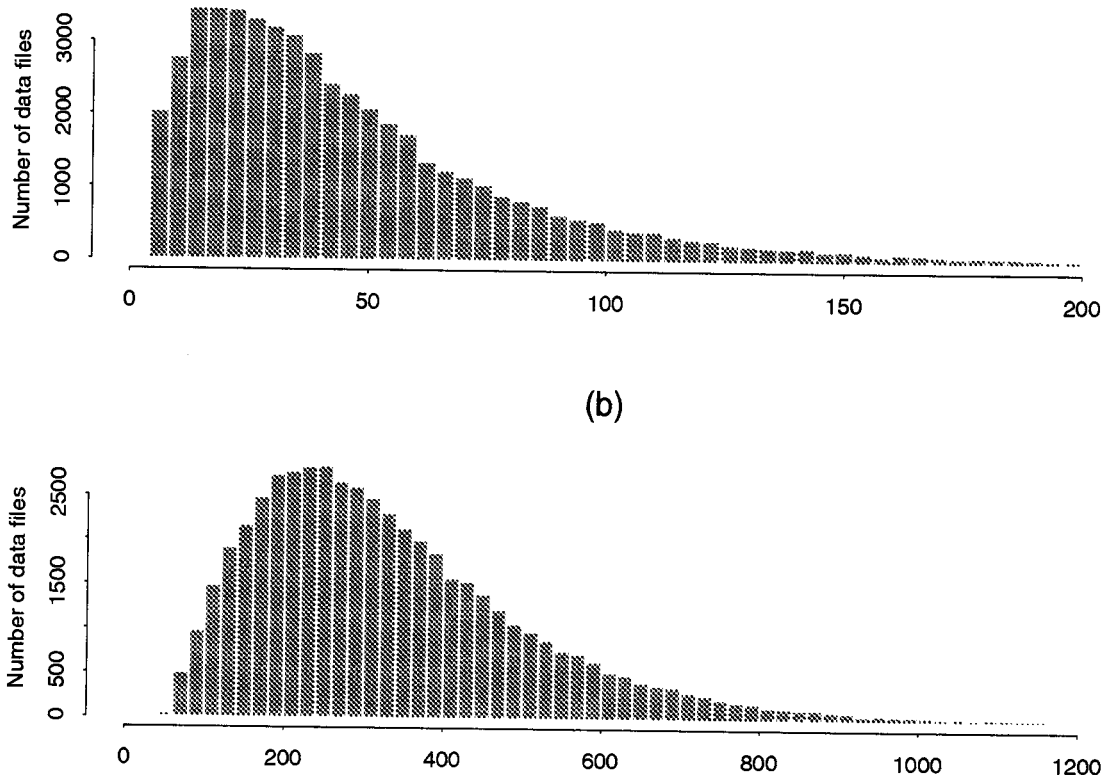

Figure 6. (a) Histogram of the width of the LSFAC region in units of data points for the 21,310 data files with $R_{fit} \leq 8$ and $\alpha \geq 2$. (b) Histogram of the FAC intensity in units of nanoteslas.

6. DISCUSSION AND SUMMARY

There are a few other techniques for calculating the spatial distribution of FACs. One popular example is the assimilative mapping of ionospheric electrodynamics (AMIE) technique [*Richmond and Kamide,* 1988]. Whereas the AMIE is most useful for examining the time development of the global FAC distribution, it requires magnetic field data from numerous ground stations, and it has to assume ionospheric conductance. The distribution of FACs can also be obtained from two-dimensional measurements of ionospheric convection, but again an assumption of ionospheric conductivity is required. A new technique developed by *Weimer* [2000] uses spherical harmonic functions for presenting the magnetic potential and does not require any assumption about the geometry of FACs. However, for obtaining the global distribution of FACs, it is necessary to assemble data from many satellite passes. In contrast to these techniques, the method we developed is basically the automation of the way we visually examine a plot of satellite magnetic field data. The procedure can be easily applied to any data set, from a single satellite pass to an entire data set of a certain satellite. The trade-off of our method is that it assumes that FACs extend infinitely, although the extent to which this assumption is justified is quantified by the PCA and is represented by α.

The output of the DMSP-F7 analysis was already used in a few statistical studies. *Higuchi and Ohtani* [2000] examined nightside region 1 FACs and reported empirical formulas for predicting the intensities of region 1 currents and westward electrojets from the latitudinal deviation of these current systems from the quiet time location. *Ohtani et al.* [2000] reported events

Table 1. Number of FAC Crossings

		Number of Sheets						
	Hemisphere	1	2	3	4	5	6	Subtotal
Dayside	North	27	3,812	1,839	517	143	21	6,359
	South	96	2,744	1,536	436	145	8	4,965
Nightside	North	53	3,381	1,561	483	110	2	5,590
	South	338	2,673	956	318	101	0	4,386
Column Totals		514	12,610	5,892	1,754	499	31	21,310

in which LSFAC systems disappear in the midday sector. They discussed the occurrence of such events from the viewpoint of the summer-winter asymmetry of ionospheric conductivity and its effect on the solar wind-magnetosphere interaction. The procedure selected 953 events in which DMSP-F7 traversed four current sheets on the dayside. *Ohtani and Higuchi* [this issue] examined the dependence of the event occurrence on the interplanetary magnetic field (IMF) and magnetic local time (MLT) and interpreted the four-sheet structure as an overlap of midday region 0 and region 1 currents and conventional region 2 and region 1 currents.

We believe that this procedure can make an important contribution to space weather as well as scientific research. For now-casting the characteristics of FACs such as location and intensity, we need to extract the characteristics of FACs from satellite observations as soon as data become available. It takes less than 5 s (typically 2-3 s) for the present procedure to process one pass with a 260-MHz Pentium II processor. For forecasting, we need empirical formulas for predicting FACs and associated quantities. Considering that the characteristics of FACs depend on many parameters of external and internal conditions, the development of such empirical formulas requires processing a huge database. As described previously, the present procedure allows us to analyze an entire set of the DMSP-F7 data, and it can be easily applied to other data sets. The robustness of our procedure will allow us to conduct a statistical study of an unprecedented size, which we believe will provide new insights into FACs.

Acknowledgments. The DMSP–F7 magnetometer data were provided by F. J. Rich. Thanks are also to K. B. Baker for providing the PACE coordinate software. A part of this work was done when T. H. was at JHU/APL as a visiting scientist. Work at the Institute of Statistical Mathematics was in part carried out under the ISM Cooperative Research Program (H9-ISM.CRP-B7 and H10-ISM.CRP-B12). Work at JHU/APL was supported by NASA and NSF.

Hiroshi Matsumoto thanks J. Burch and another referee for their assistance in evaluating this paper.

References

Akaike, H., A new look at the statistical model identification, *IEEE Trans. Autom. Control, AC–19*, 716, 1974.

Anderson, B. D. O, and J. B. Moore, *Optional Filtering*, Prentice–Hall, Englewood Cliff's, N.J., 1979.

Cummings, W. D., and A. J. Dessler, Field-aligned currents in the magnetosphere, *J. Geophys. Res., 72*, 1007, 1967.

Higuchi, T., Frequency domain characteristics of linear operator to decompose a time series into multi-components, *Ann. Inst. Stat. Math., 43*, 469–492, 1991.

Higuchi, T., and S. Ohtani, Automatic identification of large-scale field-aligned current structures and its application to night-side current systems, in *Geophys. Monogr. Ser., Vol. 118*, edited by S. Ohtani *et al.*, pp. 389–394, AGU, Washington, D.C., 2000.

Hiragi, Y., H. Urakawa, and K. Tanabe, Statistical procedure for deconvoluting experimental data, *J. Appl. Phys., 58*(1), 5, 1985.

Iijima, T., and T. A. Potemra, Field-aligned currents in the dayside cusp observed by Triad, *J. Geophys. Res., 81*, 5971, 1976.

Ja-Yong, K., Spline estimation of discontinuous regression functions, *J. Comput. Graph. Stat., 6*(3), 266, 1997.

Kitagawa, G., A nonstationary time series model and its fitting by a recursive filter, *J. Time Ser. Anal., 2*, 103, 1981.

Ohtani, S., and T. Higuchi, Four-sheet structure of dayside field-aligned currents: Statistical study, *J. Geophys. Res.*, this issue.

Ohtani, S., T. Higuchi, T. Sotirelis, and P. T. Newell, Disappearance of large-scale field-aligned current systems: Implications for the solar wind-magnetosphere coupling, in *Geophys. Monogr. Ser., Vol. 118*, edited by S. Ohtani *et al.*, pp. 253–259, AGU, Washington, D.C., 2000.

Rich, F. J., D. A. Hardy, and M. S. Gussenhoven, Enhanced ionosphere-magnetosphere data from the DMSP satellites, *Eos Trans. AGU, 66*, 513, 1985.

Richmond, A. D., and Y. Kamide, Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, *J. Geophys. Res., 93*, 5741, 1988.

Sakamoto, Y., M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics: Mathematics and Its Application*, D. Reidel, Norwell, Mass., 1986.

Weimer, D. R., A new technique for the mapping of ionospheric field-aligned currents from satellite magnetometer data, in *Geophys. Monogr. Ser., Vol. 118*, edited by S. Ohtani *et al.*, pp. 381–388, AGU, Washington, D.C., 2000.

Williams, C.M., An efficient algorithm for the piecewise linear approximation of planar curves, *Comput. Graph. Image Process., 8*, 286–293, 1978.

Zmuda, A. J., J. H. Martin, and F. T. Heuring, Transverse magnetic disturbances at 1100 km in the auroral region, *J. Geophys. Res., 71*, 5033, 1966.

T. Higuchi, The Institute of Statistical Mathematics, Tokyo 106-8569, Japan. (higuchi@ism.ac.jp)

S. Ohtani, The Johns Hopkins University, Applied Physics Laboratory, Johns Hopkins Road, Laurel, MD 20723-6099. (ohtani@jhuapl.edu)

(Received February 9, 2000; accepted May 18, 2000.)