Short Note

A method to minimize edge effects in two-dimensional discrete Fourier transforms

Yanick Ricard* and Richard J. Blakely‡

INTRODUCTION

Fourier transforms are widely used in analysis of two-dimensional (2-D) earth-science data, such as gravity and magnetic surveys, topographic models, and remote-sensing images. For example, manipulations of gridded magnetic or gravity data, such as upward and downward continuation, reduction to the pole, wavelength filters, pseudogravity transformation, and vertical derivatives (Hildenbrand, 1983), are greatly simplified with Fourier transforms, as are certain forward and inverse calculations (Parker, 1973; Parker and Huestis, 1974). Power spectra computed from 2-D Fourier transforms are used to estimate depth to the top and bottom of magnetic sources from gridded magnetic data (Spector and Grant, 1970; Connard et al., 1983) and to estimate lithospheric strength and Moho depth from gridded gravity data (Dorman and Lewis, 1970; Louden and Forsyth, 1982; McNutt, 1983).

Despite its utility, application of the 2-D Fourier transform to measured data presents three practical problems. (1) Measured data are necessarily sampled at discrete intervals, which places an upper limit on knowable wavenumbers. (2) Necessary restriction of measured data to finite areas may truncate important shapes in an arbitrary way and restrict resolution in the Fourier domain. (3) The fact that most observed quantities, including gravity and magnetic data, are not periodic functions (except at global scales) violates an implicit assumption of discrete Fourier analysis (Cordell and Grauch, 1982). These three problems, which are all dependent to some degree upon the orientation of the coordinate system, combine to produce edge effects that tend to obscure spectral patterns of geologic origin in calculated power-density spectra.

In this paper, we suggest a simple technique to reduce edge effects caused by axis orientation in power-density spectra. The method replaces the rectangular array of discrete data with a circular window of data and makes use of a simple fact:

rotation of a circular window of continuous data only affects the Fourier transform by an identical rotation. Data are first rotated about the origin by some angle, then Fourier transformed, and finally rotated back by the same angle. By stacking the results of this three-step procedure for various angles, the dependence on axis orientation is largely eliminated. We have found this procedure useful for studies that examine shapes of power-density spectra (e.g., Connard et al., 1983; Simpson et al., 1986).

THE EDGE EFFECT

In the following, we denote spatial and wavenumber coordinates by 2-D vectors $\mathbf{s} = (x, y)$ and $\mathbf{k} = (k_x, k_y)$, respectively, and write the 2-D Fourier transform $F(\mathbf{k})$ of a continuous function $f(x, y)$ as

$$F(\mathbf{k}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{s}) e^{-i\mathbf{k}\cdot\mathbf{s}} \, dx \, dy \qquad (1)$$

and the convolution of two functions $f_1(x, y)$ and $f_2(x, y)$ as

$$f_1(\mathbf{s}) * f_2(\mathbf{s}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(\mathbf{s}_0) f_2(\mathbf{s} - \mathbf{s}_0) \, dx_0 \, dy_0. \qquad (2)$$

Wavenumbers are inversely proportional to wavelengths λ_x and λ_y in the x and y directions, respectively; i.e., $k_x = 2\pi/\lambda_x$ and $k_y = 2\pi/\lambda_y$.

The cause of the edge effect is most easily seen by the construction of a discrete sampling function. Let $f(\mathbf{s})$ represent a 2-D, continuous set of data. We shall assume that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(\mathbf{s})| \, dx \, dy$$

exists and is finite. The 2-D Fourier transform exists under this condition and is given by equation (1). Real-world

measurements of $f(\mathbf{s})$, however, are often discrete and always finite. The sampled version of $f(\mathbf{s})$ can be represented as the multiplication of $f(\mathbf{s})$ by a 2-D, rectangular array of impulses (Bracewell, 1965). For example, if the data are measured on (or are interpolated to) a rectangular grid of $2N + 1$ columns and $2M + 1$ rows spaced Δx and Δy apart, respectively, then the measured survey can be represented as

$$f_D(\mathbf{s}) = f(\mathbf{s}) \times d(\mathbf{s}), \tag{3}$$

where

$$d(\mathbf{s}) = \sum_{n=-N}^{N} \sum_{m=-M}^{M} \delta(x - n\Delta x)\delta(y - m\Delta y). \tag{4}$$

The Fourier transform of $f_D(\mathbf{s})$ is given by the convolution of the Fourier transforms of $f(\mathbf{s})$ and $d(\mathbf{s})$:

$$F_D(\mathbf{k}) = \frac{1}{4\pi^2} F(\mathbf{k}) * D(\mathbf{k}), \tag{5}$$

where

$$D(\mathbf{k}) = \frac{\sin\left[(N + \frac{1}{2})k_x \Delta x\right] \sin\left[(M + \frac{1}{2})k_y \Delta y\right]}{\sin(\frac{1}{2}k_x \Delta x) \sin(\frac{1}{2}k_y \Delta y)}. \tag{6}$$

Because of the periodicity of $D(\mathbf{k})$, only wavenumbers between $\pm \pi/\Delta x$ or $\pm \pi/\Delta y$ (the Nyquist wavenumbers) need be considered in the following discussion. For large N and M in equation (6), $D(\mathbf{k})$ approaches a single impulse at the origin, and $f_D(\mathbf{s})$ approaches a perfect representation of $f(\mathbf{s})$ for wavenumbers less than the Nyquist wavenumbers. For finite N and M, $D(\mathbf{k})$ has part of its energy distributed in side lobes that are not isotropically distributed about the origin. The side lobes are rectangular in shape and the sides are parallel to the k_x and k_y axes; most energy in fact lies along the k_x and k_y axes. Consequently, finite N and M cause $D(\mathbf{k})$ to smooth and distort $F(\mathbf{k})$ in ways that depend upon the orientation of the x, y coordinate system.

The cause of the edge effect can be seen in another way. Discrete Fourier analysis implicitly assumes that the rectangular window of data is repeated infinitely in both the x and y directions, like a checkerboard pattern. Because $f(\mathbf{s})$ is not usually periodic in this way, discontinuities exist along the boundaries of each of the infinite set of repeated rectangles. The attempt to model these abrupt discontinuities with discrete samples causes additional rectangular distortions of $F_D(\mathbf{k})$.

Several approaches are used to minimize the discontinuities along the boundary of $f_D(\mathbf{s})$. First, the outer rows and columns of $f_D(\mathbf{s})$ can be gradually attenuated to zero or to some other constant. Second, extra rows and columns can be added to $f_D(\mathbf{s})$. For example, rows and columns can be added so that the discontinuity between the first and last point of *each* original row and column is smoothed across the added margin by using a straight line or a smoothly curving function (Hildenbrand, 1983; Blakely, 1977). Finally, rows and columns can be added to $f_D(\mathbf{s})$ by repeating the rows and columns of $f_D(\mathbf{s})$ in some sequence, perhaps as mirror images (Mayhew, 1985).

IMPROVED SPECTRAL ESTIMATES BY SUCCESSIVE ROTATIONS OF THE DATA GRID

Our method attempts to remove edge effects by rotating edges relative to the data; real features and edge artifacts

should thereby be separable. The method makes use of the fact that rotation of $f(x, y)$ about its origin only affects $F(\mathbf{k})$ by an identical rotation. This can be seen by rewriting equation (1) in polar coordinates. Let $\mathbf{s} = (s, \theta)$ and $\mathbf{k} = (k, \phi)$. Then equation (1) becomes

$$F(k, \phi) = \int_0^{2\pi} \int_0^\infty f(s, \theta)e^{-isk\cos(\theta-\phi)}s \, ds \, d\theta.$$

Clearly, the Fourier transform of $f(s, \theta + \varepsilon)$ is $F(k, \phi + \varepsilon)$. Thus, if $f(\mathbf{s})$ is rotated about its origin by an angle ε, Fourier transformed, and rotated back by an angle $-\varepsilon$, the result is simply $F(\mathbf{k})$. We denote this three-step procedure by $R_\varepsilon[f(\mathbf{s})]$ where the subscript indicates the amount of rotation. Note that

$$R_\varepsilon\left[f(\mathbf{s})\right] = F(\mathbf{k}) \tag{7}$$

for any ε.

Rotation by angle ε and subsequent regridding of a sampled function $f_D(\mathbf{s})$ can be seen as a rotation of $f(\mathbf{s})$, while $d(\mathbf{s})$ remains parallel to the x and y axes. Consequently, the rectangular side lobes caused by $D(\mathbf{k})$ are always parallel to the k_x and k_y axes, whereas $F(\mathbf{k})$ is rotated by an angle ε. Hence, application of the three-step procedure described above to $d(\mathbf{s})$ [rotation of $d(\mathbf{s})$ by ε, Fourier transformation, and rotation back by $-\varepsilon$] causes rotation of $D(\mathbf{k})$ by an angle $-\varepsilon$; i.e.,

$$R_\varepsilon\left[d(\mathbf{s})\right] \neq D(\mathbf{k}). \tag{8}$$

Our method uses stacking to exploit relations (7) and (8) so that contributions to $F_D(\mathbf{k})$ due to $F(\mathbf{k})$ are amplified at the expense of $D(\mathbf{k})$: We make K calculations of $R_\varepsilon[f_D(\mathbf{s})]$ for K unique values of ε, add the amplitudes of the K spectra, and divide the sum by K. Hence, we define a new "rotational" amplitude spectrum of $f_D(\mathbf{s})$:

$$\mathscr{F}\left[f_D(\mathbf{s})\right] = \frac{1}{K} \sum_{k=1}^{K} \left| R_{k\Delta\varepsilon}\left[f_D(\mathbf{s})\right] \right|. \tag{9}$$

Because amplitude spectra of real functions are symmetrical through the origin, it is only necessary that ε range between 0 and $\pi/2$. We have found in most applications that ten to twenty rotations are sufficient to eliminate most of the edge effect.

Prior to each three-step calculation of $R_\varepsilon[f_D(\mathbf{s})]$, we discard all data that lie outside of the largest circle that fits within the rectangular grid of $f_D(\mathbf{s})$. Data outside the circle are replaced by the average of all values along the perimeter of the circle. Although this step reduces the quantity of data by a factor of $(4 - \pi)/4$ for square-shaped windows, or by about 21 percent, it helps eliminate the distortion caused by the rectangular window of data.

In summary, our method consists of the following steps:

(1) Find the largest circle which fits within the boundaries of $f_D(\mathbf{s})$ and replace with a constant all $f_D(\mathbf{s})$ that lie outside the circle.

(2) Calculate $R_\varepsilon[f_D(\mathbf{s})]$:

(a) rotate the coordinate system by an angle ε and regrid $f_D(\mathbf{s})$ using spline interpolation with respect to the new coordinate system;

(b) Fourier transform the rotated grid;

(c) calculate the amplitude grid of the Fourier transform grid; and

(d) regrid the amplitude spectrum with respect to a new coordinate system rotated by angle $-\varepsilon$.

(3) Repeat step (2) using K various values of ε between 0 and $\pi/2$ and add the results of each iteration to previous results.

(4) Normalize the sum by dividing by K.

TEST CASE

A digital representation of a function having known spectral properties is required to test the method and to compare the method with other techniques. We compose our test function from a sum of L sinusoids,

$$f(\mathbf{s}) = \sum_{\ell=1}^{L} A_\ell \cos (\mathbf{k}_\ell \cdot \mathbf{s}), \qquad (10)$$

which has a Fourier transform given by

$$F(\mathbf{k}) = 2\pi^2 \sum_{\ell=1}^{L} A_\ell \left[\delta(\mathbf{k} + \mathbf{k}_\ell) + \delta(\mathbf{k} - \mathbf{k}_\ell) \right]$$

(Bracewell, 1965).

In the following example, we let $L = 10$ and $A_\ell = 1$ for all ℓ. The digital version $f(\mathbf{s})$ (Figure 1) is an array of 128 by 128 points. The band of knowable wavenumbers ranges from the fundamental wavenumbers,

$$|k_{xf}| = 2\pi/N \, \Delta x$$

$$|k_{yf}| = 2\pi/M \, \Delta y$$

to the Nyquist wavenumbers,

$$|k_{xN}| = \pi/\Delta x$$

$$|k_{yN}| = \pi/\Delta y.$$

We purposely chose the wavelengths of two of the ten sinusoids to lie outside of this band. One sinusoid that has wavenumbers greater than the Nyquist wavenumbers contributes aliasing errors to $F_D(\mathbf{k})$. Another sinusoid, which has wavenumbers less than the fundamental wavenumbers, contributes most of its energy near the origin, but also contributes to the edge discontinuity. The theoretical continuous spectrum is represented in Figure 2 where, for display purposes, the Dirac impulses are shown by small squares.

Analyses of the test case by a conventional method and by the new method are shown in Figures 3 and 4, respectively. Figure 3 shows the power spectrum of $f_D(\mathbf{s})$ calculated by (1) adding 22 rows and columns to $f_D(\mathbf{s})$ in order to minimize the edge discontinuity, (2) calculating the Fourier transform of this expanded $f_D(\mathbf{s})$, and (3) squaring the amplitude of the Fourier coefficients. The result is substantially different from the theoretical continuous spectrum (Figure 2); it is difficult to discern the various peaks due to a rectangular distortion; and several spurious peaks appear.

Figure 4 shows the power spectrum computed with the discrete rotational Fourier transform using 30 rotations. Spurious peaks are not evident, and the shape in Figure 4 more closely resembles the theoretical continuous spectrum (Figure 2). Note that, because of the first step of the procedure, Figure 4 was computed with 21 percent fewer original data than Figure 3, yet produces a more accurate estimate of the continuous spectrum.

DISCUSSION

The new technique has its greatest use in those studies that examine shapes and trends of Fourier transformed data (e.g., Connard et al., 1983; Simpson et al., 1986). Figure 5 shows an example of the benefits of the discrete rotational transform to interpretation of aeromagnetic data. Connard et al. (1983) extended the method of Spector and Grant (1970) to estimate depth to the top and bottom of magnetic sources in part of the Cascade volcanic province in Oregon. Their technique was to (1) divide the magnetic survey into rectangular cells, (2) calculate the 2-D Fourier transform of each cell, (3) calculate averages of spectral values within rings concentric about the origin, and (4) analyze the shape of these averaged spectral values as a function of radial distance from the origin. Step (3) is a convenient way to render a 2-D data set into a 1-D curve, but its success depends upon accurate calculation of mean values in each concentric ring.

To compare our technique with standard Fourier analysis, we applied both techniques to a 64 by 64 km subarea of the aeromagnetic survey shown in Figure 5 of Connard et al. (1983). Logarithms of squared amplitudes are shown as a function of radial distance from the origin. Hence, calculating averages within concentric rings about the origin of the 2-D transform is identical to calculating averages within a sliding window in Figure 5. The scatter of spectral estimates is considerably reduced with the new technique; averages calculated within concentric rings will have a higher level of confidence. The method of Connard et al. (1983) assumes that magnetic sources are distributed so that 2-D transforms have no azimuthal dependence. The discrete rotational transform ensures that orientation of the coordinate system does not detract from this assumption.

CONCLUSIONS

We have demonstrated in Figures 3 and 4 that, for this particular test case, the discrete rotational transform produces a more accurate representation of the continuous spectrum than does a Fourier transform of a rectangular window of data. We compared but did not show, the new method with three other techniques designed to reduce the edge effect: (1) addition of rows and columns which smooth the discontinuity, (2) addition of rows and columns which are mirror images of the measured data, and (3) smooth attenuation of outer rows and columns to a common value. The new method was superior in all cases. The new method has a great advantage over existing methods in that the data are not modified and surveys are not artificially expanded.

We rely on empirical rather than mathematical arguments to justify our method. This method obviously suppresses distortions that depend upon the azimuthal orientation of the coordinate system, but the method may have little or no effect on isotropic noise (i.e., noise that is a function of $|\mathbf{k}|$).

We expect that the discrete rotational transform will most benefit those studies that require examination of patterns of 2-D spectra (e.g., Connard et al., 1983; Simpson et al., 1986). It eliminates artifacts associated with the orientation of the rectangular sampling window, an effect that may tend to obscure spectral patterns related to geologic sources.
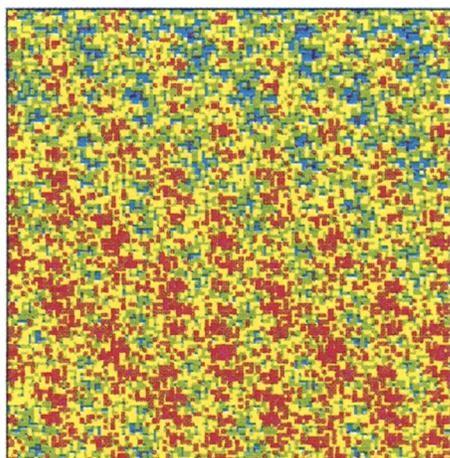
Test case

FIG. 1. Sum of 10 sinusoids used as a test case for Figures 3 and 4. Grid is composed of 128 rows and 128 columns. Contour interval is unity.
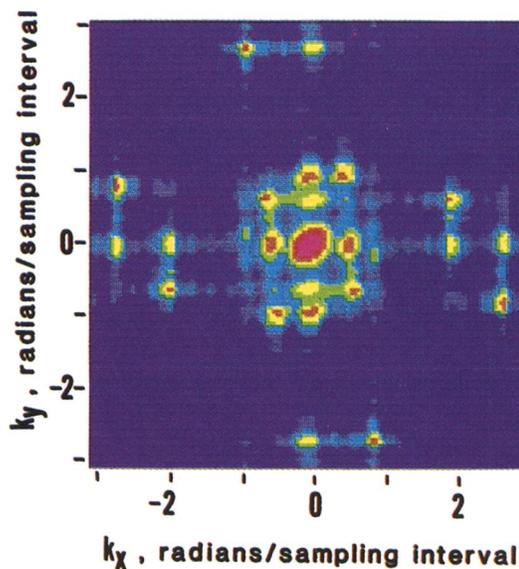
FIG. 3. Logarithm of squared amplitudes calculated from the data shown in Figure 1. Twenty-two rows and columns were added to the data prior to calculation of the Fourier transform so as to replace the discontinuity between the first and last elements of each row and column with a straight line. Contour interval is unity.
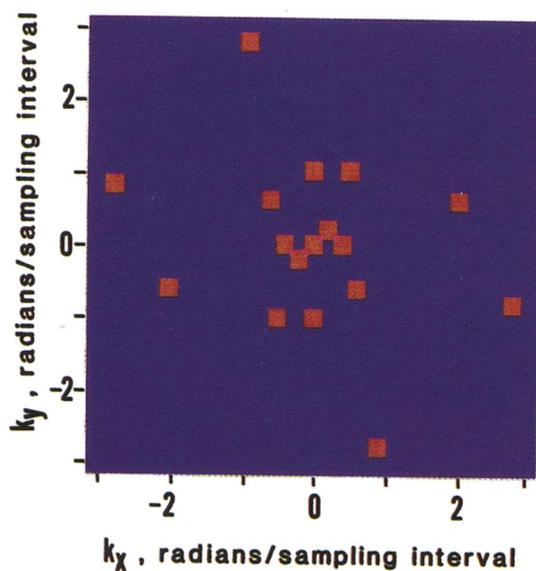
FIG. 2. Theoretical continuous spectrum of the function given by equation (10) and shown in Figure 1. Squares represent locations of Dirac impulses.
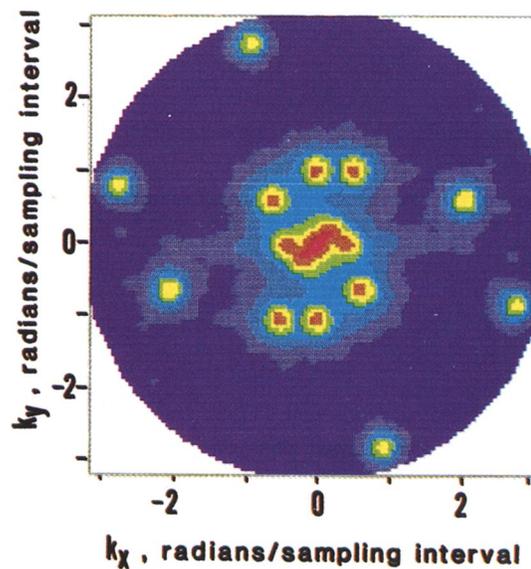
FIG. 4. Logarithm of squared amplitudes of the data shown in Figure 1 calculated by the discrete rotational transform method. Thirty rotations were used. Contour interval is unity.

k, (radians/km)
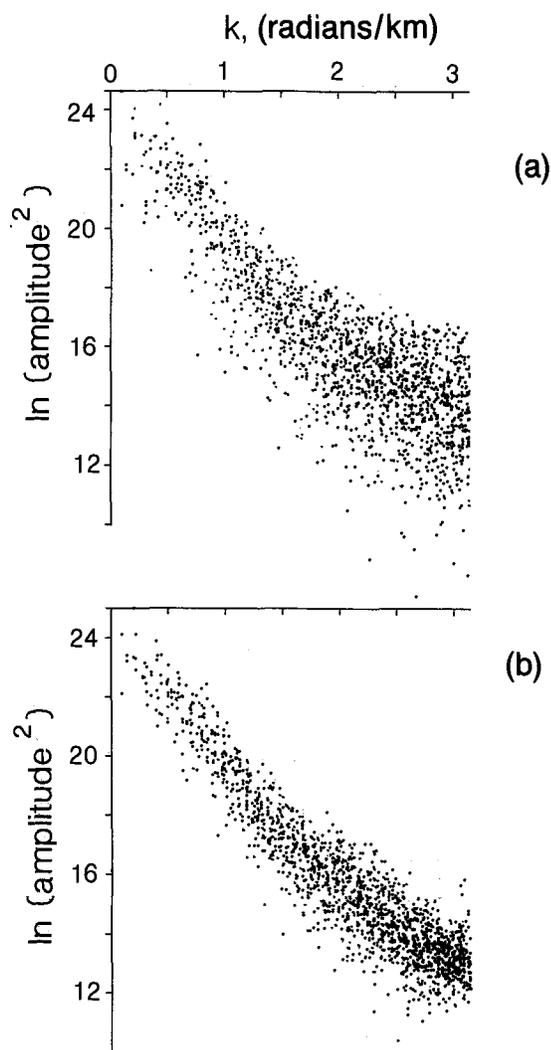
FIG. 5. Logarithm of squared amplitudes of aeromagnetic data from the Cascade volcanic province of Oregon. Values are plotted as a function of radial distance from the origin. (a) Calculations made with standard Fourier transform. (b) Calculations made with discrete rotational transform using 20 rotations.

REFERENCES

Bracewell, R., 1965, The Fourier transform and its applications: McGraw-Hill Book Co.

Blakely, R. J., 1977, Documentation for subroutine REDUC3, an algorithm for the linear filtering of gridded magnetic data: U.S. Geol. Surv. open-file rep. 77-784.

Connard, G., Couch, R. W., and Gemperle, M., 1983, Analysis of aeromagnetic measurements from the Cascade Range in central Oregon: Geophysics, 48, 376–390.

Cordell, L., and Grauch, V. J. S., 1982, Reconciliation of the discrete and integral Fourier transforms: Geophysics, 47, 237–243.

Dorman, L. M., and Lewis, B. T. R., 1970, Experimental isostasy, 1, theory of the determination of the Earth's isostatic response to a concentrated load: J. Geophys. Res., 75, 3357–3365.

Hildenbrand, T. G., 1983, FFTFIL: A filtering program based on two-dimensional Fourier analysis: U.S. Geol. Surv. open-file rep. 83-237.

Louden, K. E., and Forsyth, D. W., 1982, Crustal structure and isostatic compensation near the Kane fracture zone from topography and gravity measurements. I. Spectral analysis approach: Geophys. J. Roy. Astron. Soc., 68, 725–750.

Mayhew, M. A., 1985, Curie isotherm surfaces inferred from high-altitude magnetic anomaly data: J. Geophys. Res., 90, 2647–2654.

McNutt, M., 1983, Influence of plate subduction on isostatic compensation in northern California: Tectonics, 2, 399–415.

Parker, R. L., 1973, The rapid calculation of potential anomalies: Geophys. J. Royal Astron. Soc., 31, 447–455.

Parker, R. L., and Huestis, S. P., 1974, The inversion of magnetic anomalies in the presence of topography: J. Geophys. Res., 79, 1587–1593.

Simpson, R. W., Jachens, R. C., Blakely, R. J., and Saltus, R. W., 1986, A new isostatic residual gravity map of the conterminous United States with a discussion on the significance of isostatic residual anomalies: J. Geophys. Res., 91, 8348–8372.

Spector, A., and Grant, F. S., 1970, Statistical models for interpreting aeromagnetic data: Geophysics, 35, 293–302.