J Geophys (1987) 61:77–81

Journal of Geophysics

Stable inversions of MAGSAT data over the geomagnetic equator by means of ridge regression

C.J. Lötter
Magnetic Observatory, PO Box 32, Hermanus 7200, South Africa

Abstract. Several authors have reported on the instability problems experienced when inverting satellite magnetic anomaly data using the equivalent source technique. These instabilities particularly occur for inversions of data in regions near the geomagnetic equator and for certain dipole spacings in the equivalent source formulation. In this study the method of ridge regression is put forward as an alternative method, as it provides very stable solutions. In addition, it gives an indication of where the instabilities occur and can be utilized with smaller than usual dipole spacings. This method is demonstrated by applying it to the region containing the Bangui anomaly.

Key words: Stable solutions – Diagonal additive – Crustal magnetization/susceptibility – Equivalent source – Dipole spacing

Introduction

One of the reasons for inverting satellite magnetic anomaly data is to obtain more information on the spatial physical properties of the crust. The standard method is to invert the anomaly data to an equivalent layer magnetization model based on a latitude-longitude dipole source array at the earth's surface. Mayhew and Galliher (1982) reported that the inversion process becomes unstable for dipole spacings of less than 220 km. At these dipole spacings the parameter solutions exhibit an oscillatory nature; that is, adjacent sources generally take on alternately large positive and negative values and clearly have no physical meaning. For larger dipole spacings they found that the magnetization solutions vary systematically and have a physically meaningful interpretation. Their results were obtained for mid-latitude regions, utilizing the Moore-Penrose inverse:

$$\Delta\tilde{\mathbf{P}} = (\mathbf{A}^T\mathbf{A})^{-1}\,\mathbf{A}^T\Delta\mathbf{G} \qquad (1)$$

where $\Delta\tilde{\mathbf{P}}$ is as parameter correction vector, $\Delta\mathbf{G}$ is a residual vector (observed minus initial model) and $\mathbf{A}$ is the dipole source function matrix. Within $+20°$ of the geomagnetic equator, this technique yields unstable solutions for any reasonable dipole spacing.

The problem of inverting satellite magnetic anomaly data involves a large number of data points and of parameters to be resolved. Numerical manipulation of large matrices can lead to instabilities due to round-off error. In addition, the magnetic inverse problem is very similar to downward continuation right down to the surface of the source. This is a very unstable process and short wavelength variations in the residual total field, ΔF, are amplified out of proportion in the magnetization solution.

Another stability problem is associated with proximity to the geomagnetic equator. The presence of the asymmetric sine function in the dipole source function equation can lead to symmetric sign changes in values of the elements of the matrix $\mathbf{A}$ for this region. This is mainly due to the change in sign of the inclinations of the magnetization vector and main field direction, respectively, as the geomagnetic equator is crossed. A problem that now arises is that of linear dependence between the columns of the problem matrix $(\mathbf{A}^T\mathbf{A})$. If the condition of linear dependence is nearly satisfied, i.e. certain linear combinations of the columns are nearly zero, the determinant of $(\mathbf{A}^T\mathbf{A})$ is almost zero. The problem can then be classified as singular, ill-conditoned, non-orthogonal or ill-posed.

Von Frese et al. (1981) applied singular value decomposition to the problem matrix in order to obtain physically realistic parameter values for subsequent geological interpretation of the crust. They did an eigenvalue-eigenvector decomposition of $\mathbf{A}$ given by $\mathbf{A}=\mathbf{USV}^T$ to yield Lanczos' natural inverse:

$$\tilde{\mathbf{P}} = (\mathbf{VS}^{-1}\mathbf{U}^T)\mathbf{G}. \qquad (2)$$

$\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices whose columns are the eigenvectors associated with the columns and rows of $\mathbf{A}$, respectively, $\mathbf{S}$ is a diagonal matrix whose elements are the eigenvalues of $\mathbf{A}$. Here $\tilde{\mathbf{P}}$ and $\mathbf{G}$ are the estimated parameter and total field residual vectors, respectively. The authors do not explicitly mention any instability problems near the geomagnetic equator, or for certain dipole spacings for that matter.

Langel et al. (1984) suggested and demonstrated the use of principal component regression to overcome the instability problem. They also start off with a singular value decomposition and then force the system to be positive definite by truncating the smaller eigenvalues. This is equivalent to reducing the rank of the problem matrix. With this method one is restricted to a choice between two adjacent integer values for the rank, even though the optimum rank sometimes clearly lies between the two values (Marquardt, 1970). This normally happens when some of the eigenvalues are not precisely zero, as is the case with some dipole source function matrices.

Marquardt (1970) came to the conclusion that principal component regression (and the natural inverse) is preferable for problems containing some zero eigenvalues. Our problem involves some very small eigenvalues.

Other authors (Leite, 1983; Silva, 1986) have discussed the technique of adding a small value to the diagonal of the problem matrix, in order to force the system to be orthogonal. In the present study the method of ridge regression, initiated by Hoerl and Kennard (1970a, b) together with some ideas put forward by Marquardt (1970), is implemented to obtain stable inversions for regions near the geomagnetic equator and for smaller than usual dipole spacings.

Theory of the method

The theoretical magnetic field value at the j-th satellite observation point due to an array of k dipoles is given by:

$$g_j = \sum_{i=1}^{k} a_{ji} p_i + \varepsilon_j \quad j = 1, \ldots, n \tag{3}$$

with

n = number of observations,
a_{ij} = element of dipole source function matrix,
p_i = magnetization of i-th crustal block,
ε_j = measurement errors and noise.

The theoretical field arising from the source array is a linear function of the source parameters. The parameters may therefore be directly determined using a least-squares technique. The standard least-squares solution for an estimated parameter correction is given by Eq. (1).

Hoerl and Kennard (1970a) demonstrated (see Appendix) that when $\mathbf{A}^T\mathbf{A}$ is nearly singular the mean-square error between the unbiased estimates, $\tilde{\mathbf{P}}$, and the true parameters, $\mathbf{P}$, becomes very large when using Eq. (1). The same happens with the variance of the parameter. Instead of the least-squares estimator they suggest the use of the ridge regression estimate of $\mathbf{P}$, which is:

$$\mathbf{P}^* = (\mathbf{A}^T\mathbf{A} + \gamma\mathbf{I})^{-1}\,\mathbf{A}^T\mathbf{G} \tag{4}$$

where the relationship between the two estimators is given by:

$$\mathbf{P}^* = [\mathbf{I} + \gamma(\mathbf{A}^T\mathbf{A})^{-1}]^{-1}\,\tilde{\mathbf{P}}. \tag{5}$$

From (5) it is clear that $\mathbf{P}^*$ is a biased estimate of $\mathbf{P}$. Here $\mathbf{I}$ is the identity matrix and γ is a real positive value. According to the existence theorem of Hoerl and Kennard (1970a), a small value γ_c always exists such that the mean-squared difference between $\mathbf{P}^*$ and $\mathbf{P}$, $\bar{L}^2$, is smaller than that between $\mathbf{P}$ and the ordinary least-squares estimator $\tilde{\mathbf{P}}$. This relationship is graphically illustrated in Fig. 1.

Before adding the factor γ, the matrix $\mathbf{A}^T\mathbf{A}$ is standardized in such a way that it produces a matrix of correlation coefficients among the parameters. In other words, the diagonal values of $\mathbf{A}^T\mathbf{A}$ all became equal to unity. Therefore, the diagonal additive γ must have a value between 0 and 1. Equation (4) can now be rewritten in the form:

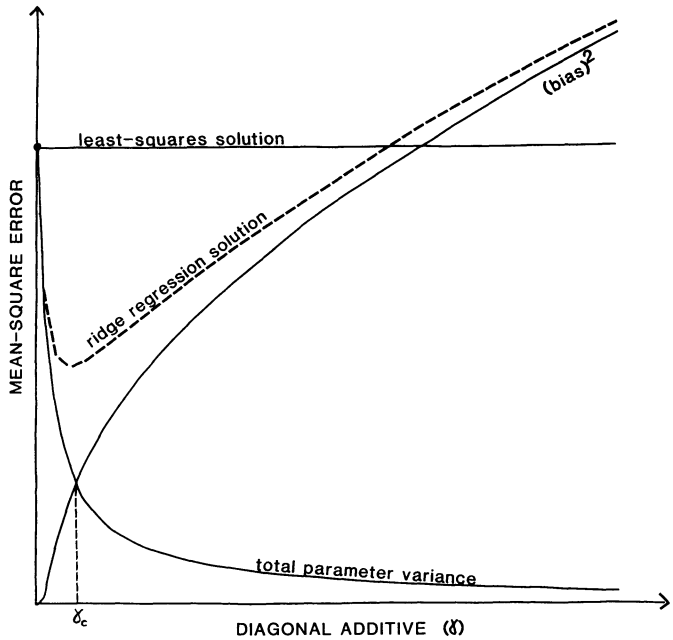

Fig. 1. Mean-square error functions for the ridge regression estimator (after Hoerl and Kennard, 1970a)

$$\mathbf{P}^* = \mathbf{D}(\mathbf{D}\mathbf{A}^T\mathbf{A}\mathbf{D} + \gamma\mathbf{I})^{-1}\,\mathbf{D}\mathbf{A}^T\mathbf{G} \tag{6}$$

where $\mathbf{D}$ is the diagonal scaling matrix with elements

$$d_{ii} = 1 \Big/ \left(\sum_{j=1}^{n} a_{ji}^2\right)^{1/2}.$$

Hoerl and Kennard (1970a) give criteria for the selection of an optimal γ. Leite (1983) provides another method to determine γ_c.

Off-diagonal coefficients in the correlation matrix $\mathbf{D}\mathbf{A}^T\mathbf{A}\mathbf{D}$ approximating unity indicate that the problem is nearly singular. This leads to very small eigenvalues which, in turn, render the problem sensitive to noise in the data and to computer round-off error. By adding a small value γ to the diagonal of $\mathbf{D}\mathbf{A}^T\mathbf{A}\mathbf{D}$, all small eigenvalues will be increased by an amount γ. This will lead to a more stable inversion.

The following is an algorithm to perform a stable inversion:

1. Let the variations in γ be defined by the exponential form $\gamma = 10^{(\tau - 2\lambda)}$ (Leite, 1983). Determine an optimum value for γ in the following way:

a) Set $\tau = -1$ and $\lambda = 1$ (These values are suitable for the data set used in this study).

b) Determine the solution $\mathbf{P}_1^* = \mathbf{D}(\mathbf{D}\mathbf{A}^T\mathbf{A}\mathbf{D}) + \gamma\mathbf{I})^{-1}\,\mathbf{D}\mathbf{A}^T\mathbf{G}$.

c) Determine $\mathbf{G}_1 = \mathbf{A}\mathbf{P}_1^*$, the first model.

d) Determine an estimate of the residual variance:

$$\hat{\sigma}_1^2 = \frac{\Delta\mathbf{G}_1^T\Delta\mathbf{G}_1}{n-k}, \text{ where } \quad \Delta\mathbf{G}_1 = \mathbf{G}(\text{observed}) - \mathbf{G}_1.$$

e) Determine the total parameter variance

$$\hat{\sigma}_p^2 = \sum_{i=1} [\mathrm{VAR}(\mathbf{P}_1^*)]_{ii}$$

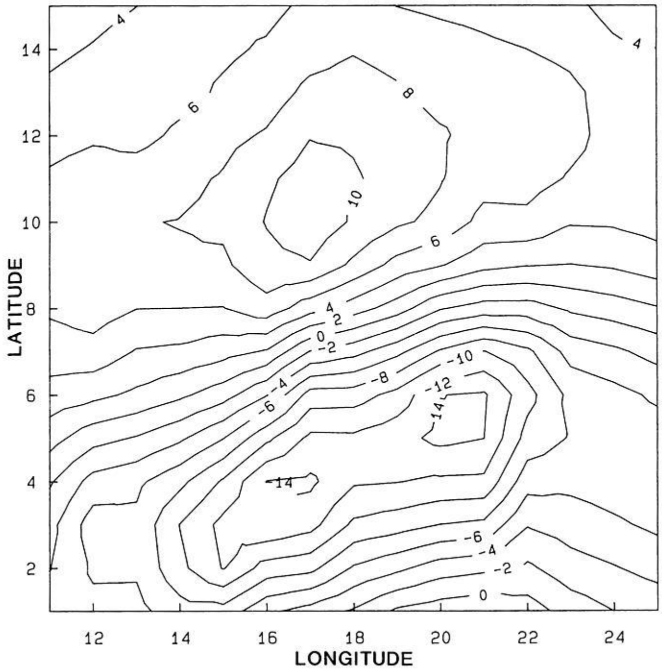

Fig. 2. Average scalar anomaly field from MAGSAT data. Contour interval is 2 nT

with $\mathrm{VAR}(\mathbf{P}_1^*) = \hat{\sigma}_1^2\ \mathbf{Z}(\mathbf{A}^T\mathbf{A})^{-1}\ \mathbf{Z}^T$ (Hoerl and Kennard, 1970a)

$$\mathbf{Z} = (\mathbf{A}^T\mathbf{A} + \gamma\mathbf{I})^{-1}(\mathbf{A}^T\mathbf{A}).$$

f) Determine a new value for γ by setting $\lambda = \lambda/_2$ and $\tau = \tau - \log_{10}(\hat{\sigma}_1^2/\hat{\sigma}_0^2)$

$$\text{where } \hat{\sigma}_0^2 = \frac{\mathbf{G}(\text{observed})^T\,\mathbf{G}(\text{observed})}{n-k}$$

g) Repeat steps (b)–(e) with $\mathbf{P}_2^*$, $\mathbf{G}_2$ and $\hat{\sigma}_2^2$.

h) If $\hat{\sigma}_2^2 < \hat{\sigma}_1^2$, repeat steps (f) and (g); if $\hat{\sigma}_2^2 > \hat{\sigma}_1^2$, set $\tau = \tau + \lambda$, leave λ unchanged and repeat step (g) with $\mathbf{P}_3^*$, $\mathbf{G}_3$ and $\hat{\sigma}_3^2$.

i) After three attempts, use Lagrange interpolation on the curve $\hat{\sigma}^2$ versus γ to obtain an optimal γ.

2. Solve Eq. (6) using the optimal γ value.

3. If not satisfied with the total variance of the parameters, increase the value γ as described above while closely monitoring the estimated residual variance $\hat{\sigma}^2$.

Results

The area in Central Africa where part of the Bangui anomaly is situated was chosen to demonstrate the application of the proposed method. A 1° latitude-longitude array of dipole sources was used for the equivalent source inversion. Each observation is an average of the MAGSAT residual total field values contained in a cylinder of radius 0.7°, height equal to the full elevation range of the satellite and centred at integer latitude-longitude intersections. Figure 2 shows a contour map of the data prior to inversion. The following corrections have been applied to this data set: removal of the main field, an improved ring current correction (Zaaiman and Kühn, 1986) and linear fits to individual passes to improve track-to-track consistency.

The optimal value of γ for this problem was found to be 0.005. Figure 3 shows a contour map of the synthesized field at a common altitude of 400 km, obtained by using $\gamma = 0.005$. From Fig. 1 it is clear that a $\gamma > \gamma_c$ can be used to obtain solutions with smaller parameter variance without increasing $\bar{L}^2$ above that obtained by the least-squares solution.

Leite (1983) used ridge regression to obtain depth to the top, width, direction and intensity of magnetization for basement blocks from aeromagnetic data. He reported that a relatively large value of γ (0.7) was required to determine

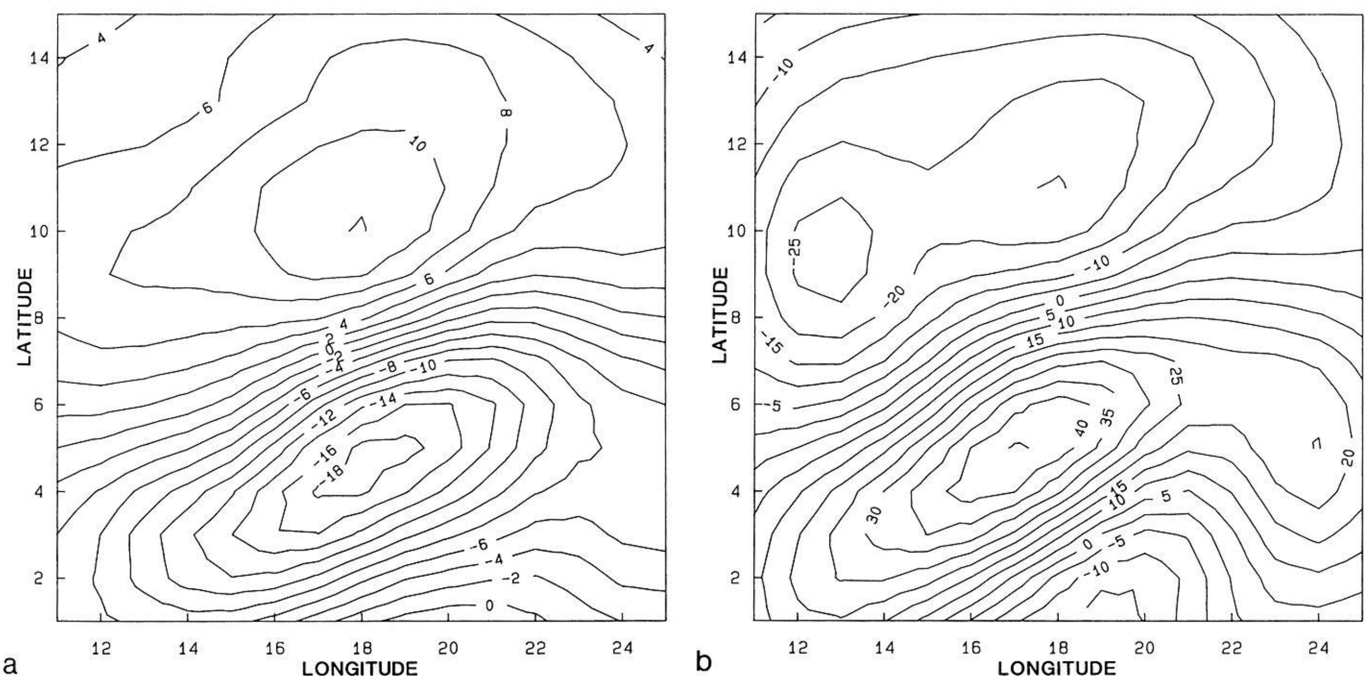

Fig. 3 a and b. a Anomaly field computed at 400 km altitude from equivalent source solution. Contour interval is 2 nT. **b** Anomaly field at 400 km altitude reduced to the pole. Contour interval is 5 nT

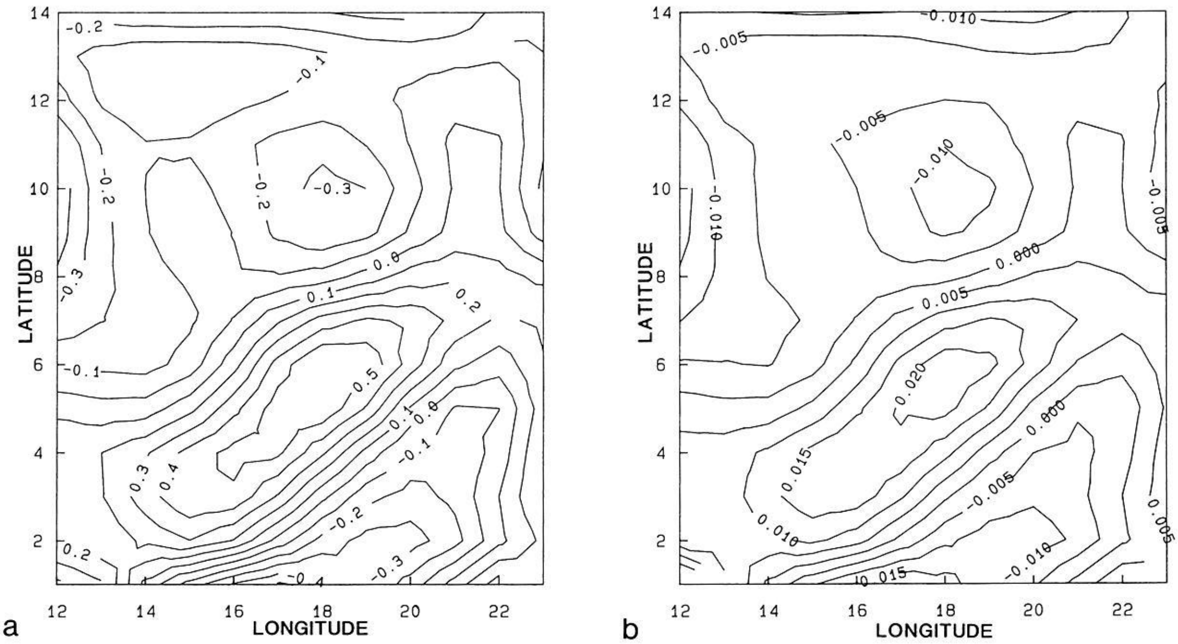

Fig. 4a and b. a Magnetization contours from equivalent source solutions. Contour interval is 0.1 A/m. **b** Susceptibility contours from equivalent source solutions. Contour interval is 0.005 SI units

physically meaningful magnetization and susceptibility values from aeromagnetic data. In the present study a value for γ which is orders of magnitude larger than the optimal value, γ_c, was also found to be necessary. In order to obtain smooth parameter solutions from satellite magnetic data, a value of 0.15 was used for γ. This is the largest value for which the inversion process is still stable and produces a smaller $\bar{L}^2$ than the least-squares technique. Figure 4a and b shows the apparent magnetization and apparent susceptibility contrast contour maps for the study area. There are only small differences between the two maps. The reason is the small variation of the main field intensity (H) in this region. It was assumed from the start that one has induced magnetization only, so that the relationship between the magnetization (P) and the susceptibility (K), which is

$$P = KH,$$

explains the similarity between the two maps.

The minor edge effects that are observed in Fig. 4a and b, can be avoided by using solutions for overlapping adjacent regions and combining them at a suitable boundary. From subsequent work on the Southern African region it was found that, with a 1° dipole spacing, an overlap of 6° for adjacent regions was sufficient to get satisfactory matching. The small perturbations which are superimposed on all the contour maps are due to the linear interpolation technique that was used prior to contouring the data.

The map reduced to the (south) pole was produced using the magnetization values of Fig. 4a.

Conclusions

Stable inversions cannot be obtained for regions near or over the geomagnetic equator when using standard inversion techniques. Ridge regression can be applied to obtain stable solutions and physically realistic parameter values for this region.

One of the most important advantages of the ridge regression method is that it allows much more freedom in selecting a suitable dipole spacing for the equivalent source grid to be used. This has been demonstrated by using a 1° × 1° dipole grid in this study. The minimum dipole spacing is only determined by the capabilities of the computer in use and the limitation that the system of equations must be overconstrained. It would be physically meaningless to decrease the dipole spacing beyond the resolution capabilities of the data set, but a higher sampling density of magnetization and/or susceptibility values facilitates effective contouring of the data. This obviates the shifting of the source grid and recalculation of parameters to obtain a satisfactory sampling density as described by Mayhew (1982).

Appendix

The following demonstrates why the least-squares estimate $\tilde{\mathbf{P}}$ of the true parameter correction vector $\mathbf{P}$ tends to be unreliable for an ill-conditioned $\mathbf{DA}^T\mathbf{AD}$ matrix.

Let the eigenvalues of $\mathbf{DA}^T\mathbf{AD}$ be denoted by

$$\lambda_{\max} = \lambda_1 \geqq \lambda_2 \ldots \geqq \lambda_p = \lambda_{\min} > 0.$$

From Hoerl and Kennard (1970a) the mean-square distance difference $\bar{L}^2$ between $\mathbf{P}$ and $\tilde{\mathbf{P}}$ is given by

$$\bar{L}^2 = \sigma^2 \,\mathrm{TRACE}(\mathbf{DA}^T\mathbf{AD})^{-1} = \sigma^2 \sum_{j=1}^{P} (1/\lambda_j)$$

and the variance is given by

$$\mathrm{VAR}[L^2] = 2\sigma^4 \sum_{j=1}^{P} (1/\lambda_j)^2$$

where σ^2 is the residual variance.

Lower limits for these two properties of L^2 are $\sigma^2/\lambda_{\min}$ and $2\sigma^2/\lambda_{\min}^2$. It is now clear that for a $\mathbf{DA}^T\mathbf{AD}$ with one or more very small eigenvalues, the distance between $\tilde{\mathbf{P}}$ and $\mathbf{P}$ will tend to be large. The variance of L^2 will also be large and this is responsible for the oscillatory behaviour of the magnetization solutions observed by Mayhew (1982) and Mayhew and Galliher (1982) for certain dipole spacings.

Acknowledgements. I thank Dr G.J. Kühn and Miss H. Zaaiman for numerous constructive discussions and for reviewing the manuscript. The author is grateful to NASA for making available the MAGSAT investigator-B data tapes.

References

Hoerl, A.E., Kennard, W.K.: Ridge regression: biased estimation for nonorthogonal problems. Technometrics **12**, 55–67, 1970a

Hoerl, A.E., Kennard, W.K.: Ridge regression: applications to nonorthogonal problems. Technometrics **12**, 69–82, 1970b

Langel, R.A., Slud, E.V., Smith, P.J.: Reduction of satellite magnetic anomaly data. J. Geophys. **54**, 207–212, 1984

Leite, L.W.B.: Application of optimization methods to the inversion of aeromagnetic data. Ph.D. dissertation, Faculty of the Graduate School of Saint Louis University, 1983

Marquardt, D.W.: Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. Technometrics **12**, 591–612, 1970

Mayhew, M.A.: An equivalent layer magnetization model for the United States derived from satellite altitude magnetic anomalies. J. Geophys. Res. **87**, 4837–4845, 1982

Mayhew, M.A., Galliher, S.C.: An equivalent layer magnetization model for the United States derived from MAGSAT data. Geophys. Res. Lett. **9**, 311–313, 1982

Silva, J.B.C.: Reduction to the pole as an inverse problem and its application to low-latitude anomalies. Geophysics **51**, 369–382, 1986

Von Frese, R.R.B., Hinze, W.J., Braile, L.W.: Spherical earth gravity and magnetic anomaly analysis by equivalent point source inversion. Earth Planet. Sci. Lett. **53**, 69–83, 1981

Zaaiman, H., Kühn, G.J.: The application of the ring current correction model to MAGSAT passes. J. Geophys. Res. **91**, 1986 (in press)

Received August 25, 1986; revised version December 31, 1986
Accepted January 20, 1987