550. 838

# PERFORMANCE COMPARISON OF NON-LINEAR OPTIMIZATION METHODS APPLIED TO INTERPRETATION IN MAGNETIC PROSPECTING

by

SVEN-ERIK HJELT*)

Finnish Academy,
Research Council for Tehnical Sciences.

Abstract

The performance of the Davidon, Marquardt, Powell and Spiral algorithms of non-linear optimization have been tested on the profile interpretation problem of magnetic prospecting. The performance has been measured by the number of function evaluations needed to calculate the theoretical profile of the interpretation model. The model parameters are found most reliably by minimizing the sum of squares of the deviations between the measured and the theoretical profiles.

For models consisting of one or two two-dimensional plates, with six parameters each, the Powell method has the most reliable performance, but the Marquardt method often is faster. In the interpretation of long, complicated profiles a reduction of computing time is essential. It was obtained by partitioning the measured profile and optimizing the parameters of each plate by direct search with parabolic fit to a limited number of points. The dip and depth extent of the plates should then preferably be kept constant during optimization.

## 1. *Introduction*

The interpretation problem of applied geophysics consists of finding geologically feasible variations of various physical properties of the

*) Present address of the author: Lab. of Applied Geophysics, Helsinki Univ. of Technology, SF-02150 Otaniemi, Finland.

ground from measurements of local variations (anomalies) of physical fields caused by the physical property in question. The problem, however, has no unique solution. Several alternative structures may give rise to same anomalies within accuracy of measurement.

The degree of non-uniqueness can be decreased by specifying the approximate form of the structures generating the measured field. Simple geometrical forms, such as cylinders, spheres, sheets and prisms are usually chosen, since the theoretical anomalies of such bodies are often calculable in closed form. The problem of interpretation then consists of comparing the measured values with the theoretical field values of one or several of these simple structures and finding the best values for the structure parameters, the dimensions and the physical properties of the model bodies.

Before the general use of computers, the interpretation of magnetic anomalies was based on one or several characteristic points of the measured anomalies, such as points of maximum and minimum value, inflection points, half-widths etc. [22]. Soon it became appreciated, that the whole anomaly was to be used in order to improve the solution. Standard curves were published to aid the interpretation by curve comparison [5]..

After the use of computers became more common, the possibilities of iterative interpretation has been discussed by many authors. Most papers consider the interpretation of a single anomaly [1], [2], [10], [16] or the finding of one parameter (e.g. magnetic susceptibility or depth) of a more complicated structure [3], [6].

The theoretical field values are mostly non-linear functions of the structure parameters. Thus the use of optimization methods can be used in solving the interpretation problem. This article describes performance comparisons between various standard methods of optimization used in the interpretation of magnetic prospecting measurements. The speed of performance has been the ultimate goal of these investigations.

## *2. The two-dimensional magnetic interpretation problem*

Magnetic prospecting measurements are performed along straight lines, profiles, either on the ground or from an aeroplane. The vertical or the horizontal component of the anomalous magnetic field is usually measured in ground surveys. In airborne measurements the anomalous

total field variations — usually assumed to be small compared with the main geomagnetic field, $T_0$ — are registered.

Since many structures of interest in prospecting are long in one horizontal direction [12], the interpretation model is built up as a combination of thick plates (Fig. 1). Neglecting effects of anisotropy and remant magnetization, the magnetic field of a single plate is given by

$$F_i(x_i, y) = b \cdot \Phi_1 + c \cdot \Phi_2 \tag{1}$$

$$\Phi_1 = \arctan Q + \delta$$

$$Q = \frac{Q_1 - Q_2}{1 + Q_1 \cdot Q_2}$$

$$Q_1 = \frac{z_0^2 \cdot d}{z_0^2 + a_i \cdot (a_i - d)} \quad Q_2 = \frac{z_2 \cdot d}{z_2^2 + a_{2i} \cdot (a_{2i} - d)}$$

$$\delta = \begin{cases} 0 & Q_1 \cdot Q_2 > 0 \\ \operatorname{sign}(Q_2) \cdot \pi & -1 < Q_1 \cdot Q_2 < 0 \end{cases}$$

$$\Phi_2 = \frac{1}{2} \cdot \ln R$$

$$R = \frac{z_0^2 + a_i^2}{z_0^2 + (a_i - d)^2} \cdot \frac{z_2^2 + (a_{2i} - d)^2}{z_2^2 + a_{2i}^2}$$

$$a_i = x_i - x_0 + d/2$$

$$a_{2i} = a_i - h \cdot \cot \varphi$$

$$z_2 = z_0 + h$$

The parameter vector of the plate is

$$y = (d, x_0, z_0, h, \varphi, k)$$

where (cf. Fig. 1)

$d$ = thickness

$x_0$ = horizontal position

$z_0$ = depth to the top

$h$ = depth extent

$\varphi$ = dip

$k$ = apparent susceptibility

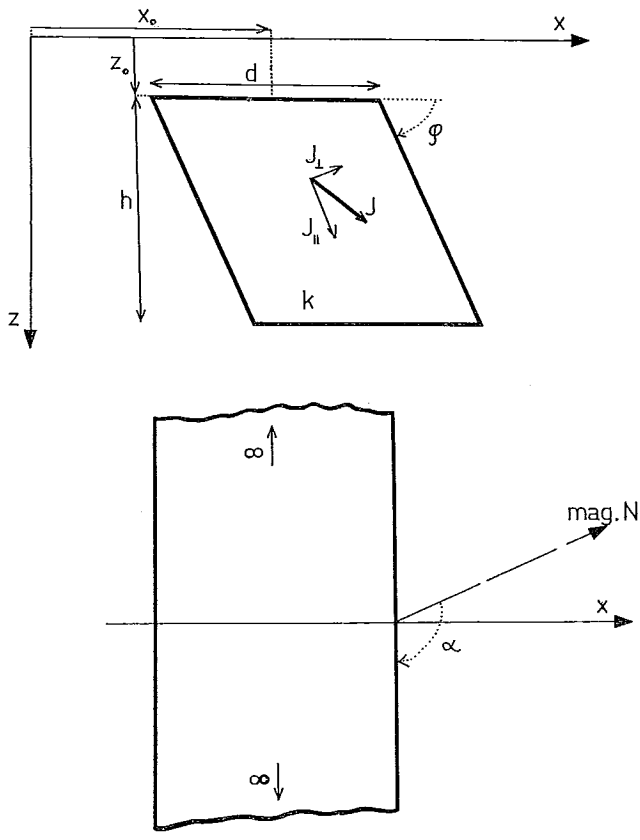


Fig. 1. The magnetic thick plate, definition of parameters.

The expressions $b$ and $c$ differ for vertical (Z), horizontal (H) and total (T) fields

$$\begin{cases} b_Z = 2kT_0 \cdot \sin\varphi \cdot (\sin I_0 \cdot \sin\varphi + \cos I_0 \cdot \sin\alpha \cdot \cos\varphi) \\ c_Z = 2kT_0 \cdot \sin\varphi \cdot (\sin I_0 \cdot \cos\varphi - \cos I_0 \cdot \sin\alpha \cdot \sin\varphi) \,/\, (1 + Nk) \end{cases}$$

$$\begin{cases} b_H = \sin\alpha \cdot c_Z \\ c_H = -\sin\alpha \cdot b_Z \end{cases} \qquad (2)$$

$$\begin{cases} b_T = \sin I_0 \cdot b_Z + \cos I_0 \cdot \sin\alpha \cdot c_Z \\ c_T = \sin I_0 \cdot c_Z - \cos I_0 \cdot \sin\alpha \cdot b_Z \end{cases}$$

where $T_0$ and $I_0$ are the amplitude and the inclination of the geomagnetic field, $\alpha$ the strike of the plate and $N$ the approximate demagnetization factor (having values between $2\pi$ and $4\pi$ cgs units depending on the thickness of the plate [5]).

Proper values of the plate parameters are found by minimizing the objective function, which is a measure of the deviation between the measurements and the theoretical field. Three different objective functions have been considered in this investigation:

$$\begin{cases} S_1 = \sum_{i=1}^{N_h} \Delta_i^2 \\ S_2 = \sum_{i=1}^{N_h} |F_{ti}| \cdot \Delta_i^2 \\ S_3 = \max_i |\Delta_i| \end{cases} \tag{3}$$

where

$$\Delta_i = F_h(x_i) - F_{ti}$$

$$F_{ti} = \sum_{l=1}^{N_h} F_t(x_i, \bar{y}_1)$$

and $y_l$ is the parameter vector of the $l$:th plate, $N_h$ the number of measuring points and $N_l$ the number of plates in the total model. Certain important properties of these funtions are briefly discussed in chapter 4. Correlation coefficients can also been used to measure the goodness of fit between measured and theoretical fields [17].

The geological feasibility of the final solution causes some restrictions on the parameter values allowed. These constraints were implemented by defining an analytic continuation of the objective function

$$S = \begin{cases} S_0 & y_{jl}^+ \geq y_{jl} \geq y_{jl}^- \\ S_0 \cdot \left[1 + \left(\frac{y_{jl}}{y_M} - 1\right)^2\right] & y_{jl} \geq y_l^+ = y_M \\ & y_{jl} \leq y_{jl} = y_M \end{cases} \Bigg\} \; l = 1 \ldots N_l \tag{4}$$

where $S_0$ is the basic objective function defined by equation (3) and

$y_{jl}^+$ = greatest allowed value of the $j$:th parameter

$y_{jl}^-$ = smallest allowed value of the $j$:th parameter

The chosen values of $y_{ji}^{\pm}$ are given in Table 1. It is necessary to make further restrictions on the horizontal position of the plates, if neighbouring plates are not allowed to cross. This is not, however, necessary, since the model satisfies the superposition principle.

When the depth extent, $h$, of the plate is great (infinite), the functions $\Phi_1$ and $\Phi_2$ in equation (1) do not longer depend on the dip, $\varphi$. The coefficients $b$ and $c$ can be used as help variables since they are the only factors depending on dip and susceptibility in this case [8].

This process, called *dip-linearization*, has been used in some of the examples to be discussed. The number of points associated with a measured anomaly profile and the starting values of the most parameters are found using a special algorithm described by HJELT [8]. (*Cf.* table 1).

## 3. *Description of the optimization methods*

### 3.1 The Marquardt algorithm

In the Newton method of finding the minimum of an objective function, $S$, the theoretical anomaly $F_t$ is approximated by the first term of its Taylor series (*cf.* definitions of Eq. (3))

$$\begin{aligned} F_{ti} &= F_{ti}(\boldsymbol{y}) + \sum_j J_{ij} \cdot \delta y_j \\ J_{ij} &= \frac{\partial F_t}{\partial y_j}(x_i, \boldsymbol{y}) \end{aligned} \tag{5}$$

An iterative refinement of $\boldsymbol{y}$ is given by

$$\begin{aligned} \boldsymbol{y}^{(k+1)} &= \boldsymbol{y}^{(k)} + \delta \boldsymbol{y} \\ \delta \boldsymbol{y} &= \boldsymbol{A}^{-1} \cdot \boldsymbol{b} \\ \boldsymbol{A} &= \{\sum_{i=1}^{N_h} (J_{ij} \cdot J_{ik})\} \\ \boldsymbol{b} &= \{\sum_i (\Delta_i \cdot J_{ij}\} \end{aligned}$$

where $\boldsymbol{A}$ and $\boldsymbol{b}$ should be calculated with the latest values, $\boldsymbol{y}^{(k)}$, of the parameters, which are available.

The method is ultimately unstable, and several methods to ensure reliable convergence have been put forward [19]. The suggestion of

Table 1. Thick plate parameters, their limits, starting values and steps in direct search iteration.

| Parameter name | $j$ | Parameter values of interest in practice | limits of parameter minimum $y_{jl}^{+}$ | limits of parameter maximum $y_{jl}^{-}$ | starting values $y_j^{(0)}$ | step $\Delta y_j^{(0)}$ |
|---|---|---|---|---|---|---|
| $d$ | 1 | 10 . . . 400 m | 5 m | 1 000 m | from anomaly half width | $0{,}1 \cdot d$ |
| $x_0$ | 2 | $x_1 \ldots x_{N_h}$ | $x_1 - 200$ m | $x_{N_h} + 200$ m | from anomaly maximum | $dx$ |
| $z_0$ | 3 | $>0{,}2$ m*) | 0,2 m*) | 100 m*) | given or 10 m*) | 5 m |
| $h$ | 4 | 100 . . . 2 500 m | 10 m | 10 000 m | given or 5 000 m | $-0.4 \cdot d$ |
| $\varphi$ | 5 | 45° . . . 135° | 0,1 rad | 3 rad | given | 10° |
| $k$ | 6 | 0,0005 . . . 0,05 cgs | 0,0001 cgs | 0,1 cgs | from anomaly maximum | $0.25 \cdot k$ [$0.8 \cdot k$, when $k < 2.\ 10^{-4}$] |

*) below the earth's surface

MARQUADT [15], LEVENBERG [14] and MORRISON [13] is the addition of a constant $\lambda$ to the diagonal of the multiplying matrix $A$. Too great values of $\lambda$ make the decrease of $S$ slow, whence a compromise between speed and stability has to be made. In this investigation the diagonal elements of $A$ were multiplied by $1 + \lambda$,

$$\lambda = \lambda_0 \cdot k^{-n}$$

where $k$ is the iteration cycle. The choise $\lambda_0 = 0.1$ and $n = 3$ was found to be appropriate, although it leads to greater values of $\lambda$ than those considered to be optimal by SHANNO [20]. If the value of $S$ still happens to increase during an iteration cycle, $\lambda$ was successively increased by multiplying with 5 until a decrease in $S$ again was found.

### 3.2 The Davidon method

When the method of DAVIDON [4], [13] is used to minimize the objective function, new values of the parameters are searched for in the conjugate directions

$$\boldsymbol{d}^{(k)} = \boldsymbol{H}^{(k)} \cdot \boldsymbol{b}^{(k)} \tag{7}$$

where $\boldsymbol{b}$ is the negative of the gradient of $S$ and $\boldsymbol{H}$ is an approximation of the inverse of the matrix formed by the second derivatives of $S$.

The objective function is minimized as an one-dimensional search in the direction defined by $\boldsymbol{d}$. The new parameter values will be

$$\boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \mu^{(k)} \cdot \boldsymbol{d}^{(k)} \tag{8}$$

where $\mu$ is a value giving $S_{\min}$. The matrix $\boldsymbol{H}$ is updated as a function tion of $\mu^{(k)}$, $\bar{\boldsymbol{d}}^{(k)}$, $-\bar{\boldsymbol{b}}^{(k)}$, $-\bar{\boldsymbol{b}}^{(k+1)}$ and $\boldsymbol{H}^{(k)}$.

### 3.3 The Powell algorithm

The optimization method of POWELL [18] does not need the calculation of the derivatives of the objective function. The search goes along $N_p$ conjugate directions $\boldsymbol{d}$, where $N_p$ is the total number of parameters to be optimized. New values of the parameters are given by

$$\begin{aligned} \boldsymbol{y}^{(k+1)} &= \boldsymbol{y}^{(k)} + \boldsymbol{\delta} \\ \boldsymbol{\delta} &= \sum_{i=1}^{N_p} \lambda_i \, \boldsymbol{d}_i^{(k)} \end{aligned} \tag{9}$$

where the coefficients $\lambda_i$ are the values minimizing

$$S(\boldsymbol{y}^{(k)} + \sum_{j=1} \lambda_i \boldsymbol{d}_j) \qquad i = 1, 2, \ldots, N_p \tag{10}$$

and the directions $\boldsymbol{d}_i$ are defined as

$$\begin{aligned} \boldsymbol{d}_i^{(k+1)} &= \boldsymbol{d}_{i+1}^{(k)} \qquad i = 1, 2, \ldots, n-1 \\ \boldsymbol{d}_n^{(k+1)} &= \boldsymbol{\delta} \end{aligned} \tag{11}$$

For $k = 1$, the $\boldsymbol{d}_i$:s are the directions defined by the components of the parameter vector $\boldsymbol{y}$.

When the number of parameters increase, restarting of the algorithm is recommended as a means to prevent the stagnation of the optimization process [13]. Tests with one-plate anomalies showed, that frequent restarting (even after every iteration) was efficient only at earlier stages of optimization, when the parameter values were still far from their optimal values. For many-plate anomalies with great number of parameters, the restarting process was, contrary to expectations, very ineffective. A closer look at the situation showed, that actually very few parameters changed during each iteration. The conjugate directions of the Powell method thus did not come to efficient use. This fact lead to the conclusion, that a simple direct search for each parameter in turn might be computationally simpler and faster to use.

### 3.4 Direct search with partial anomalies

When the length of the anomaly profile to be interpreted and the number of plates in the interpretation model increase, the computing time of the optimization process soon increases beyond the limits of feasibility. The concept of partial anomalies was then introduced and connected with the direct search method of parameter iteration. Considerable savings in computing time were noticed, when only a suitable number of points around the maximum anomaly corresponding to each plate (= the partial anomaly) is used in interpretation.

The maximum of each anomaly is identified by means of finite difference approximations to the horizontal derivate of the measured profile. Each partial anomaly consists of $NP_l$ points around the maxima so, that two neighbouring anomalies have no common points (*i.e.* if $X_l$

is the set of points of each partial anomaly, then $X_l \cap X_k = 0$ whenever $k \neq l$). At later stages of iteration, the number of points, $NP_l$, of the partial anomalies are increased, so that combined anomalies will be better interpreted.

Investigations of the objective function of a single plate indicated, that $S$ is quadratic over a wide range of parameter values. Each parameter is therefore refined using parabolic fit at three points. The partial theoretical anomaly is thus computed four times for each parameter value to be improved. The new parameter value chosen corresponds to the smallest value of $S$. The best steps in parameter iteration are those, which encounter the minimum of the objective function, and still are so small, that the parabolic nature of $S$ prevails. Too small steps seemed to be less dangerous in encountering the minimum than too large a step. Twodimensional cuts of the objective function surface were used to obtain information about suitable steps [8].

The steps decrease from one iteration cycle to another in proportion to $1/k$, where $k$ is the iteration cycle under consideration. Thus

$$\Delta y_j^{(k)} = \Delta y_j^{(k-1)}/k$$

The steps of the first iteration $\Delta \boldsymbol{y}^{(0)}$ are given in Table 1.

There is also strong correlation between some of the parameters of the model (most notably $d - z_o$, $d - h$ and $x_o - \varphi$), which appeared in the contour maps as inclined contour ellipses (or in case of strong non-linearity, distorted ellipses). These parameters should preferable be changed together, but no satisfactory general method has hitherto been found. The Marquardt method might offer a solution, but it was not tried in connection with partial anomalies.

The suitable number of points associated with each partial anomaly is a a matter of compromise. A small value of $NP_l$ should be preferred for a minimal computing time. When two or more anomalies are strongly interrelated so that the peaks of at least one anomaly disappears, correct positions of the plates can be obtained only by including points common to both anomalies. A good compromise is to start with $NP_l$ -values satisfying the criteria

$$6 \leq NP_l \leq 20 \qquad \sum_{l=1}^{N_l} NP_l \leq N_h$$

and to increase the number of anomaly points ($NP_l$) succesively at later stages of interpretation until a number depending on the anomaly

maximum amplitude (or $k$ -value of the corresponding plate) is reached. Figure 2A shows the effects of this type of procedure (curve $c$). The decrease of the objective function, $S$, soon stops, when the partial anomalies with original number of points only are used (curves $b$ and $d$). In curve $a$ $NP_l = NP_h$ for all plates. The decrease of $S$ is still quite favourable with respect to computing time, since the number of plates is small as well as the disturbing influence of neighbouring anomalies (*cf.* Fig. 7). The effect of improving the model by adding a fourth plate (curves $c$ and $d$) is clearly noticed both from the starting value and the final value of $S$.

Figure 2 B shows a detailed description of the performance of the direct search optimization, with partial anomalies $[\sum_l NP_l < N_h)$ and Fig. 2 C the same for $NP_l = N_h$. In both cases the parameters of each plate were iterated twice without decreasing the parameter step, before the parameters of the next plate were considered. The figures show clearly, that the second iteration cycle of each plate brings a singificant decrease of the objective function in very rare cases (plate no 1, position $x_o$).

### 3.5 The Spiral algorithm

Finally a recent method, the Spiral algorithm [11], was briefly tested on some interpretation examples. The search for new parameter values is done along spirals in the plane defined by the Taylor and the steepest descent directions of $S$. The Taylor step is halved whenever no decrease of the objective function can be noted along a spiral. If a maximum number of spirals have been searched without finding parameter values decreasing $S$, the search is continued in the direction of steepest descent, which is known always to lead to a decrease of $S$. Searches in the Taylor direction are performed by the Marquardt algorithm.

### 3.6 Minimum computing times for the methods

Some general information about the optimization programs used is found in table 2. Some ideas about the computational efficiency of the algorithms can be obtained by counting the arithmetic operations included in the function $F_t$ (table 3). The results have been used to construct table 4, giving the minimum computing time and number

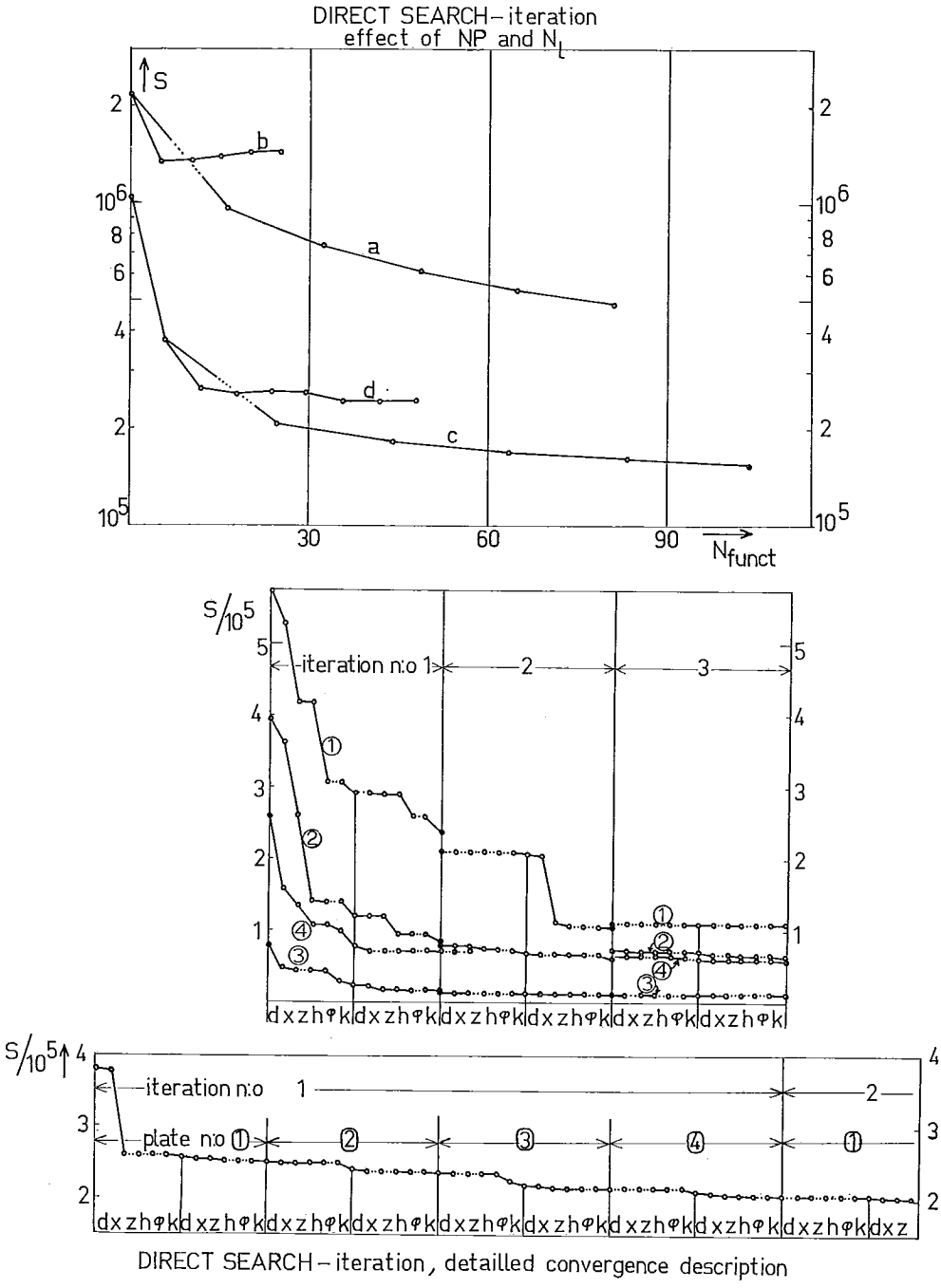


Fig. 2. Properties of direct search. A. The influence of $NP_l$, a. $NP_l = N_h$ for all plates ($N_l = 3$), b. $\Sigma NP_l = N_h$ ($N_l = 3$), c. $\Sigma NP_l = N_h$ for the first iteration, then $NP_l = N_h$ ($N_l = 4$), d. $\Sigma NP_l = N_h$ ($N_l = 4$); B. The decrease of the objective function of the partial anomalies ($N_l = 4$); C. The decrease of the total objective function ($N_l = 4$).

Table 2. Methods and programs used.

| Method | Description in Ref. no | Source of computer program used | Uses | |
|---|---|---|---|---|
| | | | Function values | Derivatives |
| Davidon | 4, 13, 19 | Reference 23 | Yes | Yes |
| Direct search with partial anomalies | 8 | Programmed by author | Yes | No |
| Marquardt | (10), 13, 14 15, 19 | Programmed by author | Yes | Yes |
| Powell | 13, 18 | Powell/Outokumpu Comp. Dept. (Reference 21) | Yes | No |
| Spiral | 11 | Outokumpu Comp.Dept. | Yes | Yes |

Table 3. Number of arithmetic operations in computing $F_t$ and its derivatives.

| | | ± | x | / | trig | arctan | ln |
|---|---|---|---|---|---|---|---|
| $F_t$ | $N_l$ times | 10 | 19 | 1 | 2 | 0 | 0 |
| | $N_l \cdot N_h$ times | 16 | 9 | 7 | 0 | 1 | 1 |
| Derivatives (in addition to $F_t$) | $N_l \cdot N_h$ times | 32 | 25 | 12 | 0 | 0 | 0 |

Table 4. Minimum computing time needed for aniteration cycle (arithmetic operations of $F_t$ and derivative calculations only).

| $N_l$ | $N_h = 10 \cdot N_l$ | $F_t$ | time/sec*) | | | $N_{funct}$ | | Powell Dir.s. |
|---|---|---|---|---|---|---|---|---|
| | | | Marq. | Dir.s. | Powell | Dir.s. | Powell | |
| 1 | 10 | 0,06 | 0,12 | 1,45 | 0,70 | 24,2 | 12 | 0,5 |
| 2 | 20 | 0,22 | 0,47 | 3,0 | 5,33 | 13,7 | 24 | 1,8 |
| 3 | 30 | 0,49 | 1,05 | 4,67 | 17,7 | 9,5 | 36 | 3,8 |
| 5 | 50 | 1,35 | 2,88 | 8,31 | 81 | 6,2 | 60 | 9,7 |
| 10 | 100 | 5,4 | 11,5 | 19,3 | 642 | 3,6 | 120 | 33,3 |
| 20 | 200 | 21,3 | 45,7 | 49,1 | 5110 | 2,3 | 240 | 104 |

*) FORTRAN IV, IBM 360/40

of function evaluations ($N_{\text{funct}}$) for one iteration cycle of the optimization methods. The operations required by the algorithms themselves are not included. It is further assumed, that the profiles contain 10 points per anomaly and that all six plate parameters are optimized. The computing times of the Fortran function statements are estimated from the machine language translation of the statements.

The last column shows the merits of the partial anomaly concept when the number of plates increases. In practice, the computational efficiency of the direct search with partial anomalies is still greater for many reasons. The direct search program is less complicated, non-anomalous parts of long profiles shorten the optimization of partial anomalies etc. In all results shown later, when the methods are compared, true computing times have been used.

## 4. *Properties of the objective functions*

The nature of the objective functions $S_1$, $S_2$ and $S_3$ is demonstrated by two-dimensional contours, when the thickness ($d$) and dip ($\varphi$) of a plate are changing (Fig. 3 A). The function $S_1$ has the smoothest behaviour when the parameters differ considerably from their correct values. $S_2$, and still more, $S_3$, has a tendency towards concavity, which is a highly undesired feature of the objective function in any optimization method.

The behaviour of the objective functions during optimization was also tested. The measured profile consisted of a theoretical anomaly and the Powell algorithm was used in all tests. Figure 3 B shows the decrease of the function $S_1$, when different objective functions were used during optimization (labels 1, 2 and 3 for $S_1$, $S_2$ and $S_3$ respectively). Figure 3 C shows the decrease of the objective functions themselves.

The results of Fig. 3 B confirm, that a linear objective function ($S_3$) does not give a very good overall fit in the interpretation ($S_1$ is related to the mean deviation of fit, $S_3$ is the maximum deviation). The objective function $S_3$ was abandoned because of its non-convexity and its slow convergence.

The objective function $S_2$ showed a better convergence as measured by $S_1$ than $S_1$ itself (Fig. 3 B). The convergence of $S_2$ turned out to be oscillatory, when the dip of the plate was optimized (dashed parts, Fig. 3 C). This objective function is useless for parameters, which

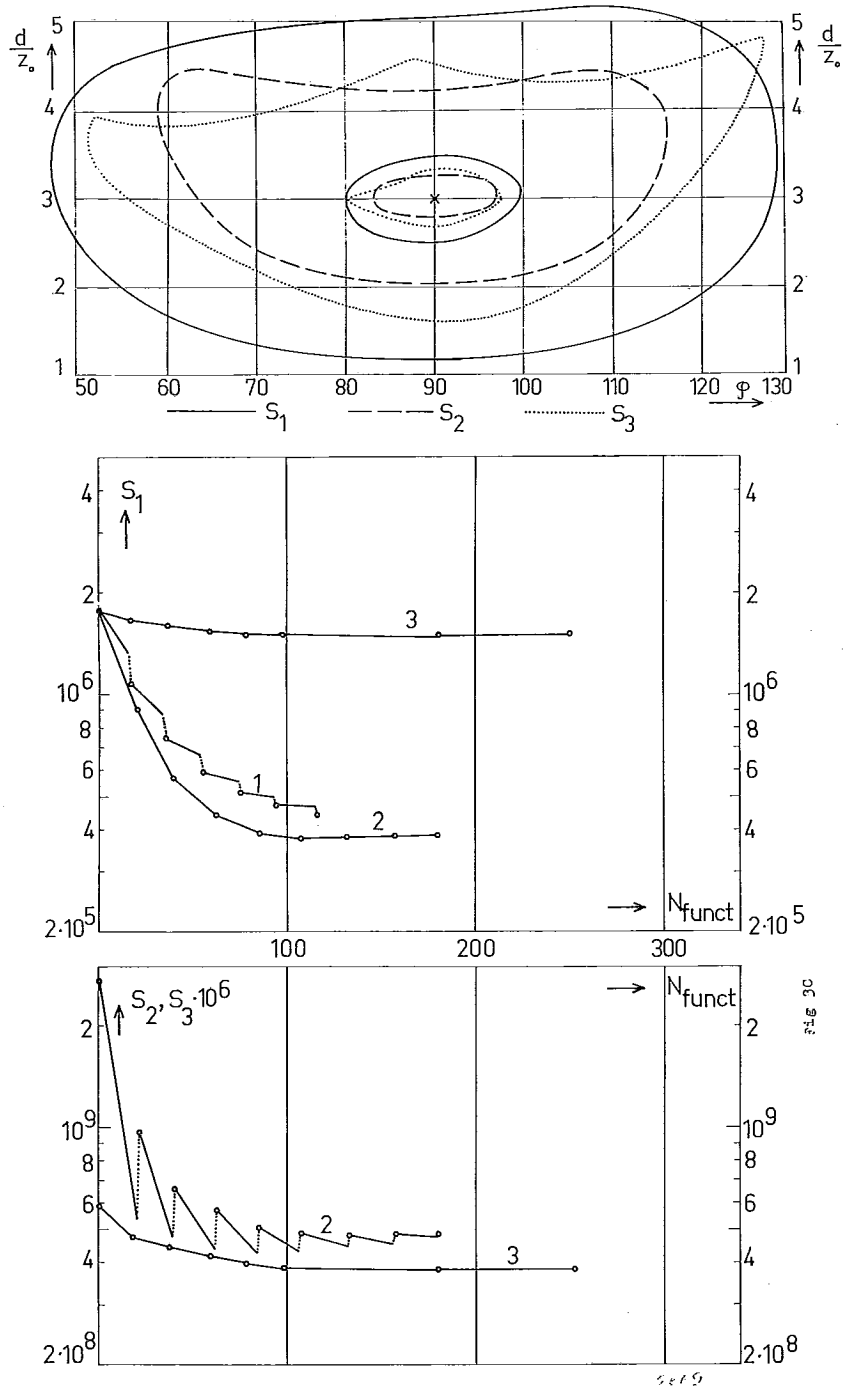


Fig. 3. Properties of three objective functions (equation 3). A. Equicontours of $S_1$, $S_2$ and $S_3$ when $d$ and $\varphi$ change.; B. The decrease of $S_1$ for different objective functions minimized; C. The decrease of the objective functions $S_2$ and $S_3$.

affect strongly the flanks of an anomaly. The conventional sum of squares, $S_1$, had the best overall properties for the purpose of optimization of the plate parameters.

## 5. *Comparison of optimization methods*

The experiments with the various optimization methods are described by means of 5 examples, shown in Figures 4—8. The first part of the Figures show the measured profiles (circles), the interpretation (line) and the interpretation model obtained. The parameters which are kept constant during optimization are indicated together with the mean deviation of fit, $\sigma$. The second parts of the Figures show the convergence of the various optimization methods. The speed of convergence is measured by the number of function evaluations, one function evaluation being equal to calculating the whole theoretical profile once. The number of interpretation points and the effect of the parameters not optimized are taken into account when estimating the true performance times.

The first profile is a theoretical one-plate anomaly without measurement errors (Fig. 4). Because the plate had infinite depth extent, the

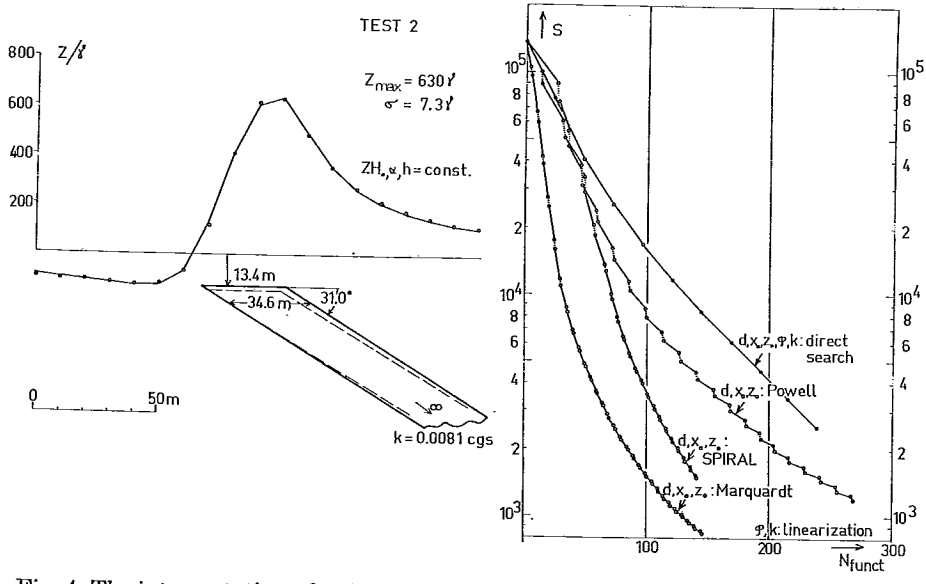


Fig. 4. The interpretation of a theoretical magnetic anomaly (dashed curve: original plate) and the decrease of the objective function for various methods of optimization (dotted parts indicate diplinearization).

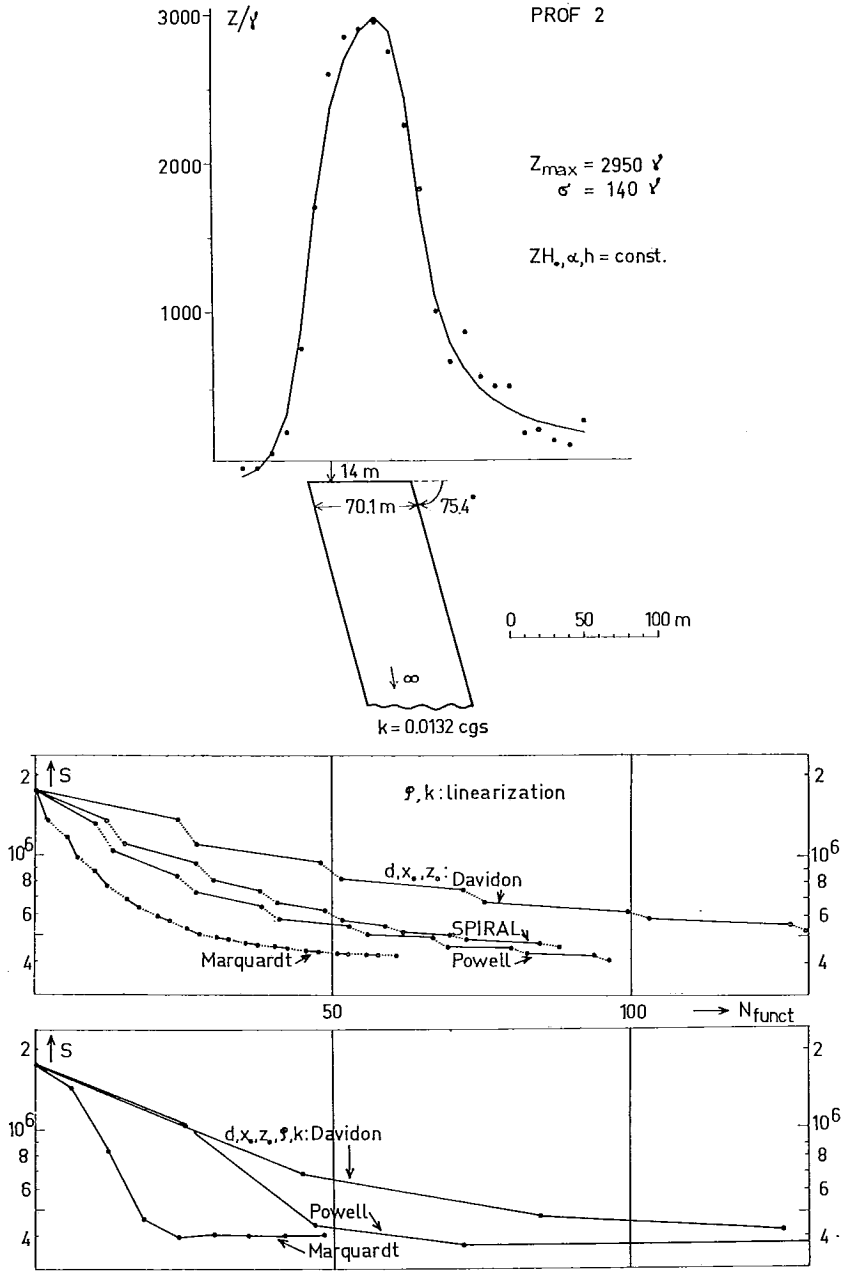


Fig. 5. The interpretation of a measured vertical component anomaly (upper convergence curves: dotted parts indicate dip-linearization).

parameter $h$ was kept constant and dip-linearization was used (dotted parts of the convergence curves) in connection with the standard optimization methods. The Marquardt method performed best and the direct search method was slow in this example. The Spiral method performed better than the Powell method since pure Taylor searches were effective during the most of the time. The greater time required by the method compared with the Marquardt method is due to more numerous program operations required by the Spiral algorithm.

The second example (Fig. 5) is the measured anomaly of a plate with infinite depth extent. The methods performed in a manner similar to the first example. The direct search was not applied. There were no essential differences between dip-linearization (upper convergence curves) and optimizing dip and susceptibility together with the other parameters (convergence curves at bottom).

The third example (Fig. 6) repeats the features of the two first examples. The Spiral method failed for some reason, not found. The Davidon method was comparable with the Marquardt method at later stages of optimization, but it failed when used with dip-linearization. The efficiency of the direct search method was equal to that of the Powell method, although not shown in the figure.

The fourth example (Fig. 7) shows a multiplate anomaly, where the comparison between the methods is done for a three plate interpretation. The Davidon method performed in this example equally to the Powell method when used together with dip-linearization, whereas the Davidon method failed when dip and susceptibility were optimized by the program itself. The efficiency of the direct search algorithm is evident. When a fourth plate (the plate most to the right) was added, the starting value of $S$ decreased significantly and the convergence of the direct search was still improved. The Marquardt method failed altogether on this interpretation example.

The last exmaple (Fig. 8) demonstrates merely some of the unambiguity properties of the interpretation. It is an aeromagnetic anomaly with an associated gradient caused by neighbouring anomalies. The gradient also compensates for the lack of exact two-dimensionality of the problem. When the method of dip-linearization was used in connection with the gradient term, the solution oscillated and the optimization did not converge after the first iteration cycle. When dip, susceptibility and the gradient were treated by the Powell algorithm, the convergence was normal and the fit given in the upper part of Fig. 8

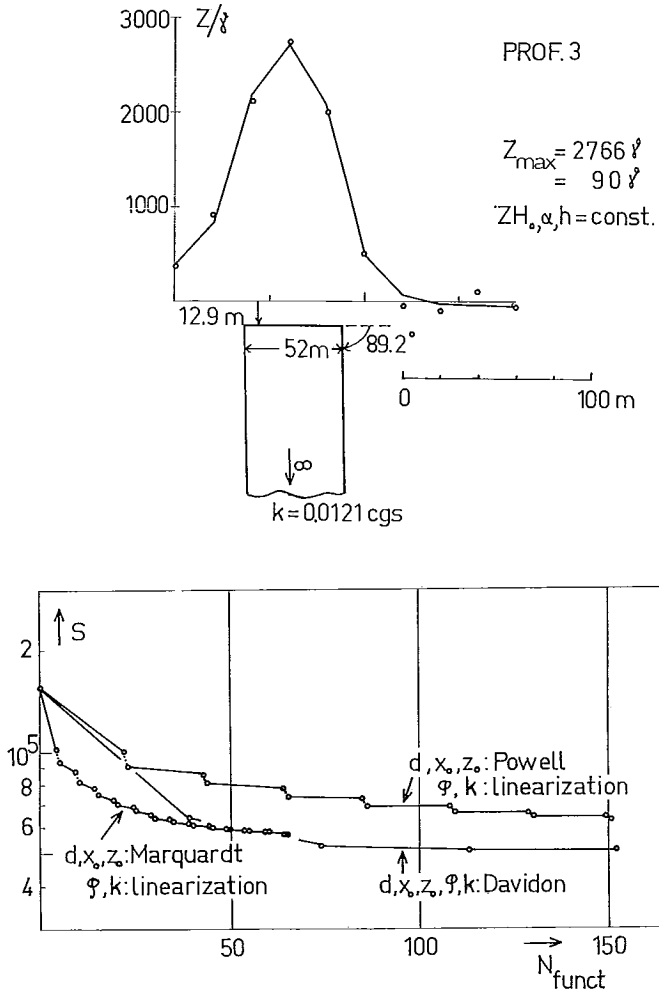


Fig. 6. Convergence in the interpretation of a measured vertical component anomaly.

was obtained. The Davidon method performed poorly even with constant gradient term and failed altogether when a variable gradient term was used. The Marquardt method and direct search were applied on a model consisting of three plates and their performance can not be directly compared with the other methods (the third plate was substituted for the gradient term). The decrease rate of the Marquardt method seems

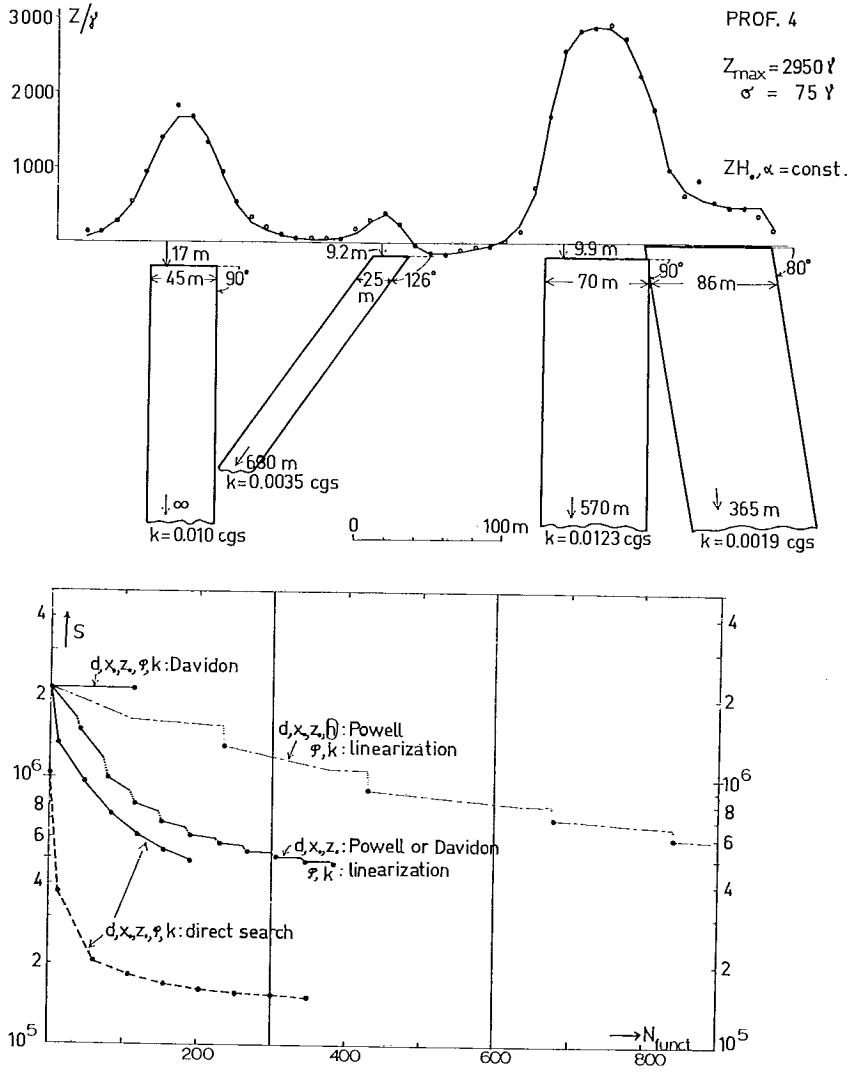


Fig. 7. Convergence in the interpretation of a measured multiplate vertical component anomaly.

to be comparable to that of the Powell method, whereas the direct search method did not converge after the first iteration cycle due to inappropriate number of points used for the partial anomalies.

According to these examples and other similar experiments [7], [8] the following conclusions can be made:

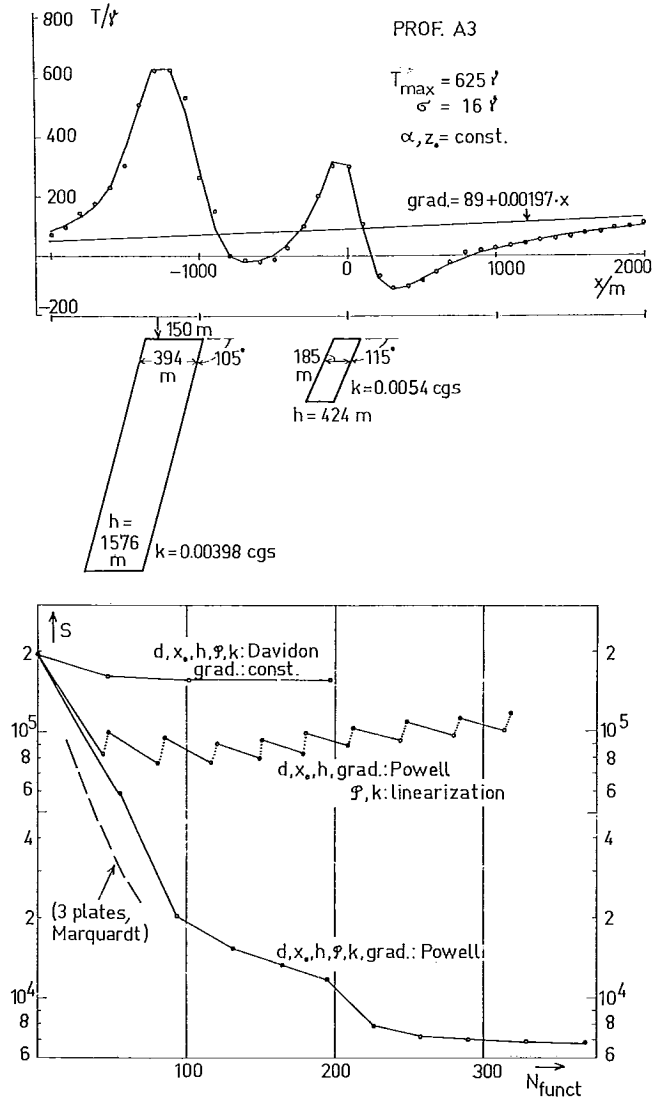


Fig. 8. Convergence in the interpretation of a measured multiplate, total field anomaly with linear gradient (dotted parts indicate dip-linearization; dashed curve refers to a three-plate model without gradient term).

1. The Powell method is capable of reliable performance for the greatest number of different parameter combinations. The method becomes slow when more than 3 plates are encountered. Dip-linearization is not to be used with this method, since the conjugate direction properties of the Powell algorithm are lost at the restart after each dip-linearization.

2. The direct search method in combination with partial anomalies is effective when two or more plates are used. The parabolic fit used for adjusting each parameter is effective, since the objective function for one plate is quadratic over a reasonably great region of parameter values.

3. The Marquardt method performs well for one and occasionally two-plate anomalies, but it fails when the number of plates is increased. Dip-linearization was succesful and lead to improved computing times in connection with the Marquardt algorithm.

4. The Davidon method performs well only occasionally. The starting values of the parameters are obviously quite far from their optimum values and the method then fails. AL-CHALABI [1] has found in connection with magnetic interpretation with polygonal cylinders, that the Davidon method is effective only at a final adjustment of the parameter values.

5. The Spiral algorithm had the poorest performance, obviously because of the success of the Marquardt principle in the searches along the Taylor direction.

6. A combination of partial anomalies and the Marquardt method might give further advantages and turn out to be the most effective combination in the interpretation of magnetic thick plate anomalies by optimization.

7. The sum of squares of the deviations between measured and theoretical anomalies, is the best objective function for the optimization process. Maximum deviation lead to slow convergence and weigthing the deviations squared with the absolute value of the anomaly lead to oscillatory behaviour of the objective function during optimization.

8. For very complicated anomalies (succesful interpretation has been obtained for models with up to 20 plates), dip and depth extent should be kept constant during optimization. The partial anomalies are not efficient for their interpretation, since the smallest errors of dip and depth extent is not obtained, when points around the anomaly maxima are used [9].

*Acknowledgements:* This paper contains results from an interpretation research project of the Finnish Academy, Research Council for Technical Sciences. The profiles of Figures 5—8 have been measured by the Exploration Department of Outokumpu Co. and the tests performed on the IBM 360/40 computer of the same company. The Spiral algorithm was programmed by Mr. E. KOIVULA. The author thanks Dr. M. KETOLA for discussions on the interpretation examples and Messrs V. SUOKONAUTIO, T. KATAJARINNE and O. RIIHIJÄRVI for their assistance in various computer operations.

## REFERENCES

1. AL-CHALABI, M., 1970: Interpretation of two-dimensional magnetic profiles by non-linear optimization. *Boll. Geof. Teor. Appl.*, **12**, 3—20.
2. BOSUM, W., 1968: Ein automatisches Verfahren zur Interpretation magnetischer Anomalien nach der Methode der kleinsten Quadrate. *Geophys. Prosp.*, **16**, 107—126.
3. BOTT, M. H. P., and M. A. HUTTON, 1970: A matrix method for interpreting oceanic magnetic anomalies. *Geoph. J. Royal Astr. Soc.*, **20**, 149—157.
4. FLETCHER, R. and M. J. D. POWELL, 1963: A rapidly convergent descent method for minimization. *Comp. J.*, **6**, 163—168.
5. GAY, S. P. Jr., 1963: Standard curves for interpretation of magnetic anomalies over long tabular bodies. *Ibid*, **28**, 161—200.
6. HARTMANN, R. H., TESKEY, D. J. and J. L. FRIEDBERG, 1971: A system for rapid digital aeromagnetic interpretation. *Ibid* **36**, 891—819.
7. HJELT, S. E., 1972: The interpretation of magnetic anomalies by computer (in Finnish). *Proc. of the Geophysics Conf. of the Finnish Geological Society, Helsinki 11.—12.11.1971.* Report HTKK-V-GF-2, Helsinki Univ. of Technology, Dept. of Mining and Metallurgy, Otaniemi, Finland.
8. —»— 1973 a: Experiences with automatic magnetic interpretation using the thick plate model. *Geophys. Prosp.*, **21**, 243—265.
9. —»— 1973 b: A contribution to the quantitative error analysis in magnetic and gravimetric interpretation. *35th E.A.E.G. Meeting*, Brighton 1973-06-05 . . . 08. (to appear in *Geoph. Prosp.*, **22**)
10. JOHNSON, W. W., 1969: A least-squares method of interpreting magnetic anomalies caused by two-dimensional structures. *Geophysics*, **34**, 65—74.
11. JONES, A., 1970: Spiral — a new algorithm for non-linear parameter estimation using least squares. *Comp. J.*, **13**, 301—308.
12. KETOLA, M., 1972: Some points of view on interpretation of magnetic and electromagnetic anomalies (in Finnish). *Proc. of the Geophysics Conf. of the Finnish Geological Society, Helsinki 11.—12.11.1971.* Report HTKK-V-GF-2, Helsinki Univ. of Technology, Dept. of Mining and Metallurgy, Otaniemi, Finland.

13. KOWALIK, J. and M. R. OSBORNE, 1968: *Methods for unconstrained optimization problems*, Elsevier Publ. Co, New York, 148 pp.
14. LEVENBERG, K., 1944: A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.*, **2**, 164—168.
15. MARQUARDT, D. W., 1963: An algorithm for least-squares estimation of non-linear parameters. *J. SIAM 11*, 431—441.
16. MCGRATH, P. M. and P. J. HOOD, 1970: The dipping dike case: a computer curve-matching method of magnetic interpretation. *Geophysics*, **35**, 831—848.
17. NAUDY, H., 1971: Automatic determination of depth on aeromagnetic profiles. *Ibid*, **36**, 717—722.
18. POWELL, M. J. D., 1964: An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comp. J.*, **6**, 155—162.
19. —»— 1970: A survey of numerical methods for unconstrained optimization. *SIAM Rev. 12*, 79—97.
20. SHANNO, D. F., 1970: Parameter selection for modified Newton methods for function minimization. *SIAM J. Numer. Anal.*, **7**, 366—372.
21. WINTER, M.: A description of the Powell program package VAO4A. (in Finnish) *Outokumpu Comp. Dept. Internal Memo.*
22. ÅM, K., 1972: The arbitrarily magnetized dyke: interpretation by characteristics. *Geoexploration*, **10**, 63—90.
23. H20-0205-3, 1968: IBM System/360 Scientific Subroutine Package. *Programmer's Manual.*