*I Simpósio Regional da Sociedade Brasileira de Geofisica*

São Paulo 2004

# Estimation of Errors in Spherical Harmonic Representation of Global Heat Flow

Cosme F. Ponte Neto, ON/MCT, (cosme@on.br) and Valiya M. Hamza, ON/MCT, (hamza@on.br)



## Abstract

A simple method of estimating errors in the calculation of harmonic coefficients is presented. The equations for spherical harmonic representation are expressed as a set of linear equations amenable for analysis using least square methods. This system of equations is recast in matrix form, which allows calculation of the uncertainties as diagonal elements of the inverse of the summation matrix. This method was applied for evaluating errors in the determination of harmonic coefficients of global heat flow data.

## Introduction

Advances in terrestrial heat flow measurements on land and sea floor over the last few decades have contributed considerably to improvements in the world geothermal database. However most of the measurements have been carried out in shallow boreholes and mines, reaching depths of less than a few hundred meters. Direct determination of deep thermal regime continues to remain a difficult task, in view of the technical and economic limitations in deep drilling. Difficulties also arise from the large variations in data density and the substantial differences in the quality of primary measurements. The problem is further complicated by the fact that in most cases experimental results refer only to the conductive component of the total heat flux, there being no easy means of determining the convective component. The controversy involved in interpretation of heat flow measurements near lithospheric spreading centers in oceanic areas is a classical example of the difficulties in estimating the deep heat flux.

A direct consequence of such difficulties in data processing and interpretation is that our understanding of global thermal field continues to be poor. Mapping heat flow fields on a regional scale is one form of minimizing problems arising from non-homogeneous distribution. However features revealed in maps are directly related to the density and distribution of the observational data used in the analysis. Spherical harmonic analysis is one of the convenient forms of examining characteristics of potential fields of the Earth on a global scale. In the earlier attempts for harmonic analyses (see for example, Lee and MacDonald, 1963, Lee and Uyeda, 1965; Horai and Simmons, 1969) the analytical procedure has been based on methods that create an overdetermined set of equations based on experimental data, which in turn is solved for the unknown coefficients. This method has the inherent weakness that the determination of coefficients is sensitive to the characteristics of data distribution. In the later works by Chapman and Pollack (1975) and Pollack et al (1993) problems arising from uneven data distribution are minimized using empirical predictors, (based on the heat flow-age relation proposed by Polyak and Smirnov (1968) and Hamza and Verma (1969)) of heat flow for unsurveyed areas. The practice of using empirical predictors in harmonic representation of heat flow is open to criticism as it is based on a priori knowledge of thermal processes at deeper levels in the crust, which is what we are trying to determine in the first place. In the work of Pollack et al (1993) substantial portions of the experimental data were set side in favor of synthetic data generated by empirical predictors, which is a matter of concern. Another problem with this second method is that it requires extensive pre-processing of related geological and geophysical information for selecting 'suitable' values of the empirical predictors for the different terrains with different geotectonic characteristics. Such procedures are cumbersome and prone to errors, especially when large data sets are involved. In this context, it is convenient to note that the set of coefficients calculated by Pollack et al (1993) are incorrect.

Another related problem is that uncertainties involved in the determination of the harmonic coefficients have not been estimated in these earlier works. In the present work we outline a simple procedure for estimating errors in the determination of the harmonic coefficients.

## Characteristics of the Global Heat Flow Database

The heat flow database was downloaded from the web site of the National Geophysical Data Center (NGDC). This compilation was completed in 1991 and includes 20201 records of heat flow measurements over the globe. Of these 10042 are on land, 9864 in oceanic regions and the remaining 295 in transition regions such as continental platform areas and shallow water bodies. The geographic distribution of the database, illustrated in Figure (1), indicates significant differences in data density in both continental and oceanic regions.

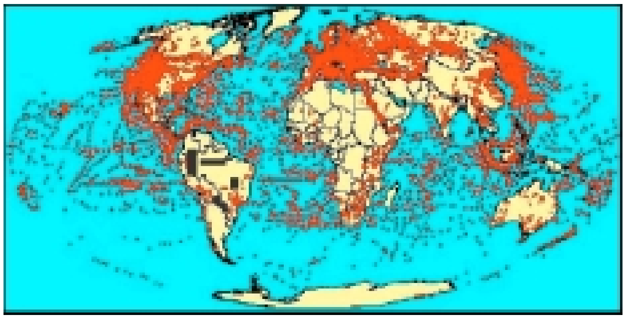

Figure (1) Global distribution of heat flow data. Shaded areas in South America refer to localities where new data was compiled by Hamza et al (2004).

In dealing with such database, with wide disparities in data density, it is common practice to adopt procedures that minimize problems arising from the non-homogeneous distribution. In the present case, the surface area of the globe was divided into a regular grid system composed of 5 x 5 degree squares and average values of heat flow in the grid elements were calculated. Experimental data are available for 12349 of these grid elements, there being 1353 elements without data. For those grid elements without data, the same criterion as that used by Chapman and Pollack (1975) was employed in assigning estimated heat flow values. The reduced data set at regular grid points was used in determination of harmonic coefficients.

**Spherical Harmonic Representation**

The harmonic representation of heat flow (q) may be represented as either:

$$q(\theta,\phi)=\sum_{n=0}^{N}\sum_{m=0}^{n}\left[A_{nm}\cos(m\phi)+B_{nm}\,\mathrm{sen}(m\phi)\right]P'_{nm}(\cos\theta) \qquad (1)$$

or

$$q(\theta,\phi)=\sum_{n=0}^{N}\sum_{m=0}^{n}\left[A_{nm}\cos(m\phi)P'_{nm}(\cos\theta)+B_{nm}\,\mathrm{sen}(m\phi)P'_{nm}(\cos\theta)\right] \qquad (2)$$

where $\phi$ is the longitude $\theta = 90 - \psi$, is the colatitude, $P'nm(\cos\theta)$ is the associated Legendre function that is fully normalized and $A_{nm}$ and $B_{nm}$ the coefficients of the harmonic expansion. The expression for evaluation of $P'_{nm}$ is:

$$P'_{nm}=\frac{P_{nm}}{\sqrt{K_n^m}} \qquad (3)$$

where $P_{nm}$ is the associated Legendre function given by:

$$P_{nm}(\cos\theta)=\frac{\mathrm{sen}^m\theta}{2^n}\sum_{t=0}^{Int\left(\frac{n-m}{2}\right)}\frac{(-1)^t(2n-2t)!}{t!(n-t)!(n-m-2t)!}\cos^{(n-m-2t)}\theta \quad (4)$$

and

$$K_n^m=\frac{1}{2(2n+1)}\frac{(n+m)!}{(n-m)!} \qquad (5)$$

In equation (4) $Int\,(n-m/2)$ is the largest integer that is lower than $(n-m)/2$.

Full normalization of associated Legendre functions ($P_{nm}$) requires that the following equations be satisfied:

$$\int_0^{2\pi}\int_0^{\pi}\left[P'_{nm}(\cos\theta)\mathrm{sen}(m\phi)\right]^2\mathrm{sen}\,\theta\;d\theta\;d\phi=4\pi \quad (6a)$$

$$\int_0^{2\pi}\int_0^{\pi}\left[P'_{nm}(\cos\theta)\cos(m\phi)\right]^2\mathrm{sen}\,\theta\;d\theta\;d\phi=4\pi \quad (6b)$$

**Least Squares Determination of the Coefficients**

We now turn our attention to the procedure adopted for estimating the coefficients $A_{nm}$ and $B_{nm}$. Note that if N is the degree of the harmonic expansion, the relations $((N^2+3N)/2+1)$ and $(N(1+N)/2)$ give the number of coefficients of $A_{nm}$ and $B_{nm}$ respectively. Thus, for example, in twelve-degree expansion there are 91 coefficients for $A_{nm}$ and 78 for $B_{nm}$, leading to a total of 169 coefficients.

The estimates of the coefficients may be obtained by fitting the harmonic expansion to the set of experimental data, which are the heat flow values (q) and their respective geographic coordinates ($\phi$ e $\theta$). The expression for q may be written as:

$$\begin{aligned}q &= A_{00}\cos(0.\phi)P'_{00}+A_{10}\cos(0.\phi)P'_{10}+A_{11}\cos(1.\phi)P'_{11}+\\ &B_{11}\,\mathrm{sen}(1.\phi)P'_{11}+A_{20}\cos(0.\phi)P'_{20}+A_{21}\cos(1.\phi)P'_{21}+ \\ &B_{21}\,sin(1\phi)\,P'_{21}+\ldots\ldots\ldots\ldots B_{1212}\,sin(12\phi)P'_{q1212}\end{aligned} \qquad (7)$$

where the expression for $\mathbf{P'_{nm}(\cos\theta)}$ is abbreviated as $P'_{nm}$. Though not necessary from the theoretical point of view, it is possible to impose the condition that the first coefficient ($A_{00}$) is equal to the average of all the data, in other words:

$$A_{00}=\sum_{i=1}^{w}q_i\,/\,w \qquad (8)$$

where w is the number of data. Equation (7) may be rewritten as:

$$q=A_{00}C_1+A_{10}C_2+A_{11}C_3+B_{11}C_4+\ldots B_{1212}C_{169} \quad (9)$$

$$c_1=\cos(0\phi)P'_{00},\;\; c_2=\cos(0\phi)P'_{10},\;\; c_3=\cos(1\phi)P'_{11},$$

$$c_4=sin(1\phi)P'_{11},\ldots\ldots\ldots\ldots,\;\; c_{169}=sin(12\phi)P'_{1212}$$

Equation (9) may now be recast in a form such that the right hand side contains only the coefficients to be determined in the least square adjustment:

$$\frac{q-A_{00}.C_1}{C_2}=A_{10}+A_{11}\frac{C_3}{C_2}+B_{11}\frac{C_4}{C_2}+A_{20}\frac{C_5}{C_2}+A_{21}\frac{C_6}{C_2}+\ldots\ldots+B_{1212}\frac{C_{169}}{C_2} \quad (10)$$

Equation (10) may now be written in a more compact form as:

$$y=a1+a2.x1+a3.x2+a4.x3+\ldots\ldots\ldots\ldots a168.x167 \qquad (11)$$

where $y=(q-A_{00}\,C_1)/C_2$, $a_j$ (j= 1 to 168) are the coefficients $A_{nm}$ and $B_{nm}$ and $x_{j-1}$, the ratios ($C_j/C_2$) with j varying from 2 to 168. Note that $C_1 = 2^{1/2}$ and $A_{00}$ are constants and independent of m, n and $\theta$. Associated with each value of $q_i$ is a set of values of Cij (1<j<168). Equations like (11) may be written down for each of the data $q_i$, forming thus a system of equations for the whole set of w data points:

$$\begin{cases} y_1=a1+a2.x1_1+a3.x2_1+a4.x3_1+\ldots\ldots\ldots a168.x167_1 \\ y_2=a1+a2.x1_2+a3.x2_2+a4.x3_2+\ldots\ldots\ldots a168.x167_2 \\ \;.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad. \\ y_w=a1+a2.x1_w+a3.x2_w+a4.x3_w+\ldots\ldots\ldots a168.x167_w \end{cases} \qquad (12)$$

For number of data greater than the number of coefficients this leads to an over determined system of equations, amenable for analysis using standard least square methods. Noting that for a given matrix $A_{MxN}$ of coefficients, with M > N, and a vector $\overline{y}$ of observations or data, the least square criterion provide estimates (say $\widehat{x}$ and $\widehat{y}$ ) which satisfy the model:

$$\widehat{y} = A\,\widehat{x} = A\left(A^T A\right)^{-1} A^T\,\overline{y}$$

Thus (12) may be reformulated in matrix form as:

$$\begin{bmatrix} \sum \frac{y_i}{\sigma_i^2} \\ \sum \frac{y_i.x1_i}{\sigma_i^2} \\ \sum \frac{y_i x2_i}{\sigma_i^2} \\ \\ \\ \\ \sum \frac{y_i x167_i}{\sigma_i^2} \end{bmatrix} = \begin{bmatrix} \sum \frac{1}{\sigma_i^2} & \sum \frac{x1_i}{\sigma_i^2} & \sum \frac{x2_i}{\sigma_i^2} & . & \sum \frac{x167_i}{\sigma_i^2} \\ \sum \frac{x1_i}{\sigma_i^2} & \sum \frac{x1_i^2}{\sigma_i^2} & \sum \frac{x1_i.x2_i}{\sigma_i^2} & \cdots & \sum \frac{x1_i.x167_i}{\sigma_i^2} \\ \sum \frac{x2_i}{\sigma_i^2} & \sum \frac{x2_i.x1_i}{\sigma_i^2} & \sum \frac{x2_i^2}{\sigma_i^2} & \cdots & . \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ \sum \frac{x167_i}{\sigma_i^2} & \sum \frac{x167_i.x1_i}{\sigma_i^2} & \sum \frac{x167_i.x2_i}{\sigma_i^2} & \cdots & \sum \frac{x167_i^2}{\sigma_i^2} \end{bmatrix} \begin{bmatrix} a1 \\ a2 \\ a3 \\ \\ \\ \\ a168 \end{bmatrix} \quad (13)$$

In (13) the summation runs over the range of data 1 to w and $\sigma_i$ the respective standard deviations. Designating the column matrix on the left hand side by *D* and the second member of the right hand side as the coefficient matrix *F* we may rewrite (13) in compact form as:

$$D = M\,.\,F \qquad (14)$$

where M is designated as the summation matrix, a square matrix of 168x168. It is fairly straightforward to show that coefficient matrix is given by the inverse of the summation matrix M and the column matrix D:

$$F = M^{-1}\,.\,D \qquad (15)$$

**Uncertainty in the Estimate of Heat Flow (q)**

Matrix formulation allows a fairly simple way of obtaining estimates of the uncertainties in the coefficients. The variances of the coefficients are the diagonal elements of the inverse of the summation matrix. Thus:

$$\sigma_{ak}^2 = \left(M^{-1}\right)_{kk} \qquad (16)$$

Since the uncertainties a1, a2, .......a168 are known from (16) we may estimate the uncertainty in y making use of the relation:

$$\sigma_y = \sqrt{\left(\frac{\partial y}{\partial a1}\right)^2 \sigma_{a1}{}^2 + \left(\frac{\partial y}{\partial a2}\right)^2 \sigma_{a2}{}^2 + .. \left(\frac{\partial y}{\partial a168}\right)^2 \sigma_{a168}{}^2} \quad (17)$$

If it is further assumed that the uncertainties in $A_{00}$, C1 and C2 are equal and zero, we have from equation (10):

$$\sigma_y = \sigma_q \qquad (18)$$

This leads to an estimation of the uncertainty in heat flow (q) as:

$$\sigma_q = \sqrt{\left(\frac{\partial y}{\partial a1}\right)^2 \sigma_{a1}{}^2 + \left(\frac{\partial y}{\partial a2}\right)^2 \sigma_{a2}{}^2 + ..... \left(\frac{\partial y}{\partial a168}\right)^2 \sigma_{a168}{}^2} \quad (19)$$

**Discussion and Conclusions**

The procedure outlined in the previous section was used in obtaining estimates of errors in the harmonic coefficients of the global heat flow data set. The results presented are Table (1) as percentage errors for the coefficients $A_{NM}$ and $B_{NM}$.

The value of $\sigma_q$ calculated using equation (19) might be compared with the error estimates in experimental data. Note that increasing the order of expansion leads to higher values of $\sigma_q$. Thus a limit can be set for selecting the order of expansion that is compatible with the inherent uncertainties in the experimental data.

It is convenient to point out that the set of harmonic coefficients published by Pollack et al (1993) are found to be incorrect. A new set of coefficients based on the available data on conductive heat flow has been calculated and will be published elsewhere (Ponte Neto and Hamza, 2004).

**Acknowledgements**

We would like to thank Jean Marie Flexor and Fernando Jose Soares Silva Dias for fruitful discussions on the methods used for analysis of potential fields on a global scale. Computational facilities for carrying out this work were acquired through funding provided by FAPERJ (Process No. E-26/150.954/99).

**References**

Chapman, D.S. and Pollack, H. N., 1975, Global heat flow: A new look. Earth Planet. Sci. Lett.,28(1975), 23-32.

Lee, W. H. K. and McDonald, G. J. F., 1963, The Global variation of terrestrial heat flow, J. Geophys. Res. 68, 6481-6492.

Lee, W. H. K. and Uyeda, S., 1965, Review of heat flow data, in Terrestrial Heat Flow, Geophysical Monograph Series No.8 (American Geophysical Union), 87-190.

Hamza, V.M. and Verma, R.K., 1969, Relationship of heat flow with the age of basement rocks, *Bull. Volcan*. **33**, 123-152.

Hamza, V.M., Silva Dias, F.J.S., Gomes, A.J.L. and Terceros, Z.G.D., 2004, Numerical and Functional representation of regional heat flow in South America, Submitted for publication in Tectonophysics.

Horai, K. and Simmons, G., 1969, Spherical harmonic analysis of terrestrial heat flow, *Earth Planet. Sci. Lett.* 6, 386-394.

Pollack, H.N., Hurter, S.J. and Johnson, J.R., 1993, Heat flow the Earth's interior: analysis of the global data set, Reviews of Geophysics, 31,3, 267-280.

Polyak, B.G. and Smirnov, Y.A., 1968, Relationship between terrestrial heat flow and tectonics of continents. *Geotectonics* (Eng. Transl.) **4**, 205-213.

Ponte Neto, C.F. and Hamza, V.M., 2004, Spherical Harmonic Analysis of Conductive Heat Flow: A Reappraisal of Global Data. (Submitted for publication).

Table –1. Estimates of percent errors (σ values) in the spherical harmonic coefficients $A_{NM}$ and $B_{NM}$ of Global Heat Flow Data.

| n | m | $\sigma_{Anm}$ (%) | $\sigma_{Bnm}$ (%) | n | m | $\sigma_{Anm}$ (%) | $\sigma_{Bnm}$ (%) | n | m | $\sigma_{Anm}$ (%) | $\sigma_{Bnm}$ (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 4,3E-11 | 0,0E+00 | 7 | 2 | 1,8E-08 | 8,8E-09 | 10 | 5 | 9,4E-09 | 6,6E-09 |
| 1 | 0 | 1,2E-09 | 0,0E+00 | 7 | 3 | 2,3E-07 | 3,3E-08 | 10 | 6 | 1,5E-08 | 9,6E-08 |
| 1 | 1 | 2,9E-09 | 2,4E-09 | 7 | 4 | 2,8E-09 | 7,7E-09 | 10 | 7 | 2,5E-09 | 1,2E-08 |
| 2 | 0 | 1,5E-10 | 0,0E+00 | 7 | 5 | 1,5E-08 | 1,3E-08 | 10 | 8 | 5,3E-08 | 3,2E-08 |
| 2 | 1 | 1,5E-09 | 4,3E-09 | 7 | 6 | 1,3E-08 | 5,8E-09 | 10 | 9 | 4,9E-09 | 3,3E-10 |
| 2 | 2 | 1,6E-09 | 6,1E-09 | 7 | 7 | 1,0E-08 | 4,0E-09 | 10 | 10 | 2,3E-08 | 1,2E-08 |
| 3 | 0 | 1,8E-08 | 0,0E+00 | 8 | 0 | 7,9E-09 | 0,0E+00 | 11 | 0 | 1,3E-08 | 0,0E+00 |
| 3 | 1 | 6,0E-09 | 3,6E-09 | 8 | 1 | 8,8E-09 | 2,9E-09 | 11 | 1 | 1,4E-08 | 8,9E-09 |
| 3 | 2 | 3,1E-09 | 2,0E-08 | 8 | 2 | 5,4E-09 | 5,4E-10 | 11 | 2 | 3,4E-09 | 9,2E-10 |
| 3 | 3 | 1,8E-09 | 4,8E-09 | 8 | 3 | 1,0E-08 | 4,6E-08 | 11 | 3 | 1,4E-08 | 6,9E-09 |
| 4 | 0 | 3,7E-10 | 0,0E+00 | 8 | 4 | 2,0E-08 | 1,5E-08 | 11 | 4 | 2,1E-09 | 1,4E-09 |
| 4 | 1 | 4,8E-09 | 1,7E-09 | 8 | 5 | 3,4E-08 | 5,0E-09 | 11 | 5 | 7,2E-09 | 5,6E-09 |
| 4 | 2 | 1,3E-09 | 8,7E-09 | 8 | 6 | 3,9E-09 | 3,6E-09 | 11 | 6 | 4,3E-09 | 3,4E-10 |
| 4 | 3 | 3,5E-08 | 4,1E-09 | 8 | 7 | 2,2E-08 | 5,6E-07 | 11 | 7 | 6,9E-09 | 4,0E-09 |
| 4 | 4 | 2,6E-08 | 1,2E-08 | 8 | 8 | 6,1E-08 | 4,1E-09 | 11 | 8 | 2,7E-09 | 3,2E-10 |
| 5 | 0 | 1,7E-09 | 0,0E+00 | 9 | 0 | 6,1E-09 | 0,0E+00 | 11 | 9 | 6,0E-08 | 6,9E-09 |
| 5 | 1 | 1,5E-09 | 3,1E-08 | 9 | 1 | 9,2E-09 | 1,8E-08 | 11 | 10 | 1,6E-09 | 3,8E-10 |
| 5 | 2 | 1,9E-09 | 2,1E-08 | 9 | 2 | 1,8E-08 | 6,7E-09 | 11 | 11 | 4,0E-08 | 6,6E-09 |
| 5 | 3 | 2,9E-09 | 1,7E-08 | 9 | 3 | 6,2E-09 | 1,1E-08 | 12 | 0 | 3,3E-08 | 0,0E+00 |
| 5 | 4 | 3,1E-09 | 1,8E-08 | 9 | 4 | 4,7E-08 | 9,1E-09 | 12 | 1 | 1,3E-08 | 1,2E-08 |
| 5 | 5 | 5,2E-09 | 1,8E-08 | 9 | 5 | 1,5E-08 | 8,5E-08 | 12 | 2 | 2,3E-09 | 1,7E-09 |
| 6 | 0 | 5,4E-10 | 0,0E+00 | 9 | 6 | 2,5E-09 | 9,1E-10 | 12 | 3 | 2,4E-09 | 2,0E-07 |
| 6 | 1 | 1,0E-09 | 7,5E-09 | 9 | 7 | 1,2E-08 | 1,2E-09 | 12 | 4 | 2,8E-09 | 1,9E-08 |
| 6 | 2 | 2,5E-09 | 4,0E-09 | 9 | 8 | 1,2E-08 | 2,6E-09 | 12 | 5 | 2,3E-09 | 1,5E-08 |
| 6 | 3 | 5,1E-09 | 1,2E-08 | 9 | 9 | 4,9E-08 | 5,2E-09 | 12 | 6 | 9,2E-09 | 1,2E-09 |
| 6 | 4 | 6,2E-09 | 8,1E-09 | 10 | 0 | 9,2E-09 | 0,0E+00 | 12 | 7 | 1,3E-09 | 6,5E-09 |
| 6 | 5 | 1,7E-09 | 3,4E-09 | 10 | 1 | 5,7E-09 | 4,7E-08 | 12 | 8 | 2,5E-08 | 6,2E-09 |
| 6 | 6 | 2,5E-08 | 3,4E-09 | 10 | 2 | 7,0E-09 | 4,7E-09 | 12 | 9 | 1,1E-09 | 2,0E-09 |
| 7 | 0 | 4,6E-09 | 0,0E+00 | 10 | 3 | 4,2E-08 | 3,6E-08 | 12 | 10 | 2,2E-08 | 2,7E-07 |
| 7 | 1 | 1,2E-07 | 8,9E-08 | 10 | 4 | 4,5E-09 | 2,6E-08 | 12 | 11 | 7,9E-10 | 2,0E-07 |
| | | | | | | | | 12 | 12 | 5,1E-11 | 1,7E+00 |