

APPLYING THE METHOD OF MAXIMUM CONTRIBUTIONS TO THE MAGNETOGRAM INVERSION TECHNIQUE

Yu.V. Penskikh

*Institute of Solar-Terrestrial Physics SB RAS,
Irkutsk, Russia, penskikh@iszf.irk.ru*

Abstract. Fundamentals of the spherical harmonic analysis (SHA) of the geomagnetic field were created by Gauss. They acquired the classical Chapman-Schmidt form in the first half of the XXth century. The SHA method was actively developed for domestic geomagnetology by IZMIRAN, and then, since the start of the space age, by ISTP SB RAS, where SHA became the basis for a comprehensive method of MIT (magnetogram inversion technique). SHA solves the inverse problem of potential theory and calculates sources of geomagnetic field variations (GFV) — internal and external electric currents. The SHA algorithm forms a system of linear equations (SLE), which consists of $3K$ equations (three components of the geomagnetic field, K is the number of ground magnetic stations). Small changes in the left and (or) right side of such SLE can lead to a significant change in unknown variables. As a result, two consecutive instants of time with almost identical GFV are approximated by significantly different SHA coefficients. This contradicts both logic and

real observations of the geomagnetic field. The inherent error of magnetometers, as well as the method for determining GFV, also entails the instability of SLE solution. To solve such SLEs optimally, the method of maximum contribution (MMC) was developed at ISTP SB RAS half a century ago. This paper presents basics of the original method and proposes a number of its modifications that increase the accuracy and (or) speed of solving the SLEs. The advantage of MMC over other popular methods is shown, especially for the Southern Hemisphere of Earth.

Keywords: equivalent current function, magnetogram inversion technique, spherical harmonic analysis, system of linear equations.

INTRODUCTION

Spherical harmonic functions were first used for the analytical representation of the geomagnetic field almost two centuries ago [Gauss, 1839]. The Spherical Harmonic Analysis (SHA) created by Carl Gauss was developed in [Schuster, Lamb, 1889; Schmidt, 1935; Chapman, Bartels, 1940] and became the primary method for modeling the main magnetic field and for calculating internal and external sources (equivalent currents) of quiet geomagnetic variations (diurnal and seasonal) at middle and low latitudes within the two-dimensional ionospheric dynamo theory [Benkova, 1941; Fougere, 1963; Mishin, 1976; Yanovsky, 1978; Haines, Torta, 1994; Backus et al., 1996; Manda, Korte, 2010; Olsen et al., 2010].

At ISTP SB RAS, the SHA method for the variable geomagnetic field in the dynamo region, in particular in the high-latitude region of the polar ionosphere, began to develop rapidly under the leadership of V.M. Mishin since the foundation of the Institute in 1960 [Mishin Bazarzhapov, 1966; Bazarzhapov et al., 1966, 1979; Shpynev et al., 1974; Mishin, 1976; Mishin et al., 1982, 1984; Shirapov et al., 2000]. Currently, SHA remains the main block of the software package of the magnetogram inversion technique (MIT) [Bazarzhapov et al., 1979; Mishin, 1990; Shirapov, Mishin, 2009] — an innovative method for calculating ionospheric sources (electric fields and currents) of ground geomagnetic

variations from their measurements at the global network of magnetic stations.

The SHA method solves the inverse problem of the potential theory for the quasistationary harmonic geomagnetic field in the surface layer of the non-conducting atmosphere [Chapman, Bartels, 1940; Yanovsky, 1978]. The Laplace equation for the scalar magnetic potential and for the geomagnetic field potentiality yields a system of linear algebraic equations (SLAE), which in MIT [Bazarzhapov et al., 1979] takes the form

$$\begin{cases} \sum_{n=1}^N \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) \frac{\partial P_n^m(\cos \theta)}{\partial \theta} = X, \\ \sum_{n=1}^N \sum_{m=0}^n (g_n^m \sin m\lambda - h_n^m \cos m\lambda) \frac{m P_n^m(\cos \theta)}{\sin \theta} = Y, \\ \sum_{n=1}^N \sum_{m=0}^n (j_n^m \cos m\lambda + k_n^m \sin m\lambda) P_n^m(\cos \theta) = Z, \end{cases} \quad (1)$$

where X, Y, Z are components of geomagnetic variations in a dipole coordinate system, measured at magnetometer stations; θ is the dipole colatitude; λ is the local geomagnetic time; $g_n^m = E_n^m + I_n^m$; $h_n^m = e_n^m + i_n^m$;

$j_n^m = nE_n^m - (n+1)I_n^m$; $k_n^m = ne_n^m - (n+1)i_n^m$; $E_n^m, e_n^m, I_n^m, i_n^m$ are spherical coefficients; $P_n^m(\cos \theta)$ and $\partial P_n^m(\cos \theta)/\partial \theta$ are associated Legendre polynomials of

degree n and order m , specified in Schmidt normalization, and their derivatives [Schmidt, 1935].

In (1), the number of columns is equal to the number of unknown variables and depends on N , and the number of rows is $3K$, where K is the number of ground magnetometers. On the one hand, when selecting N , we should take into account K so that the number of unknowns does not exceed the number of equations. On the other hand, when using only part of the first columns, we cut off the rest of the spectrum, which in general could better approximate the original field specified at a nonuniform network. Thus, a problem arises about choosing a finite subset such that it can approximate GFV with acceptable accuracy from a number of spherical functions — the problem of choosing the optimal spectrum of approximating functions.

The SHA coefficients obtained for (1) are used to calculate external and internal current functions responsible for GFV observed on Earth's surface. The formula for calculating the external equivalent current function from the SHA coefficients $J(\theta, \lambda)$ has the form [Chapman, Bartels, 1940; Bazarzhapov et al., 1979; Haines, Torta, 1994]

$$J(\theta, \lambda) = -\frac{R_E}{\mu_0} \sum_{n=1}^N \frac{2n+1}{n+1} \left(\frac{r}{R_E} \right)^n \times \sum_{m=0}^n (E_n^m \cos m\lambda + e_n^m \sin m\lambda) P_n^m(\cos \theta), \quad (2)$$

where $\mu_0 = 4\pi \cdot 10^{-7}$ H/m is the magnetic constant; $r = R_E + h$; $R_E = 6371$ km is the Earth radius; $h = 115$ km is the reduced height of the ionospheric current-carrying layer.

In the MIT software package, the spherical coefficients of (1) and external equivalent current function (2) form the basis of calculating spatial distributions of major electrodynamic parameters of the polar ionosphere and magnetosphere in two hemispheres (boundaries of auroral ovals, integral conductivities, horizontal electric fields and currents, field-aligned currents, and other parameters) [Bazarzhapov et al., 1979; Mishin, 1990; Lunyushkin, Penskikh, 2019].

The number of rows of system (1) depends on the number of operating magnetometer stations; and the number of columns, on the spectrum of degree n and order m in use [Bazarzhapov et al., 1966; Mishin et al., 1984]. Thus, the system can be both overdetermined and underdetermined. When $m=0$, zero columns appear, and for the Y component in (1) zero substrings appear as well. If the system is ill-conditioned, small changes in the X , Y , Z components lead to significant changes in desired coefficients [Tikhonov, Arsenin, 1979]. As a result, sources of geomagnetic field variations (GFV) for two consecutive moments of time with almost the same variations are approximated by significantly different coefficients. This contradicts both logic and actual observations. The presence of the inevitable error of magnetometers that measure the geomagnetic field and the selection of a method for determining the GFV base-

line [Gjerloev, 2012] also cause instability in SLAE solution (1) [Mishin et al., 1982]. There are a number of methods for solving systems of linear equations [Faddeev, Faddeeva 1963]. Based on features of the global distribution of magnetic data in MIT, to solve SLAE a method of maximum contributions (MMC) has been developed which has demonstrated high reliability and efficiency [Shpynev et al., 1974; Bazarzhapov et al., 1979]. MMC relies on instantaneous data on GFV, specified at a non-uniformly distributed network of ground magnetometers. This complicates the solution of the inverse problem [Sneeuw, 1994], but provides 1-min instantaneous equivalent current functions without time averaging. Most other existing SHA methods unlike MMC require uniform distribution of raw data over a sphere [Barraclough, 1978].

Ground magnetometers are generally located in the Northern Hemisphere, therefore for a few decades the MIT problem was successfully solved using MMC only for one hemisphere. Due to the project SuperMAG, the amount of available data from the world's ground magnetometers, including the Southern Hemisphere, has increased [Gjerloev, 2012]. This made it possible to perform calculations not only for the Northern but also for the Southern Hemisphere, and was one of the reasons for creating a new MIT software package.

Purposes of this work are to describe the current state of MMC as one of the important elements of MIT, which has already provided first noncontroversial scientific results for the two hemispheres of Earth at a time [Lunyushkin et al., 2019]; to compare MMC with two other popular methods utilized to solve systems of equations of similar geophysical problems.

1. BASICS OF THE METHOD OF MAXIMUM CONTRIBUTIONS

In the case of a uniform network of raw data, GFV is expanded by a system of orthogonal spherical functions, which can be considered as a system of orthogonal vectors, i.e. orthogonal basis. In the case of a nonuniform network, the vector system is nonorthogonal.

In the matrix form, SLAE is expressed through:

$$\mathbf{Ax} = \mathbf{b}, \quad (3)$$

where \mathbf{A} is the matrix of coefficients, \mathbf{b} is the column vector of constant terms, \mathbf{x} is the column vector of unknown variables. In the general case, the matrix \mathbf{A} may be non-square.

Introduce the notations: $C = \cos(m\lambda)$, $S = \sin(m\lambda)$. For clarity, use a magnetometer station and fixate n and m , then SLAE (1) can be represented as (3), where

$$\mathbf{A} = \begin{bmatrix} C \frac{\partial P_n^m}{\partial \theta} & S \frac{\partial P_n^m}{\partial \theta} & C \frac{\partial P_n^m}{\partial \theta} & S \frac{\partial P_n^m}{\partial \theta} \\ S \frac{mP_n^m}{\sin \theta} & -C \frac{mP_n^m}{\sin \theta} & S \frac{mP_n^m}{\sin \theta} & -C \frac{mP_n^m}{\sin \theta} \\ CnP_n^m & SnP_n^m & -C(n+1)P_n^m & -S(n+1)P_n^m \end{bmatrix},$$

$$\mathbf{x} = \begin{bmatrix} E_n^m \\ e_n^m \\ I_n^m \\ l_n^m \end{bmatrix}, \mathbf{b} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}.$$

If the number of observation stations is equal to K , system (1) consists of $3K$ of such subsystems. With increasing n and m , the number of columns of the matrix \mathbf{A} and the number of rows of the matrix \mathbf{x} increase. Thus, in the general case, system (1) may be underdetermined and overdetermined. By performing successive iterations, MMC provides an approximate solution of such a system regardless of its dimension.

Consider the MMC basics by the simplified example of SLAE with two equations and two unknowns

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

The matrix can be regarded as a set of columns or a set of rows [Beklemishev, 1998]. If we separate the matrix \mathbf{A} into individual column vectors, SLAE can be represented as

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

This system can be written as

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 = \mathbf{b}, \quad (4)$$

where $\mathbf{a}_1 = (a_{11}, a_{21})$, $\mathbf{a}_2 = (a_{12}, a_{22})$ are new base vectors in the two-dimensional space, x_1, x_2 are scalars at the base vectors, $\mathbf{b} = (b_1, b_2)$ is the vector-result that is obtained by summing up the two base vectors multiplied by respective scalars. The vectors $\mathbf{b}, \mathbf{a}_1, \mathbf{a}_2$ are specified in an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2$.

Thus, the SLAE solution is equivalent to decomposition of the given vector \mathbf{b} in the given basis $\mathbf{a}_1, \mathbf{a}_2$, i.e. to the search for a vector \mathbf{x} such that its components are coordinates of the vector \mathbf{b} in the basis $\mathbf{a}_1, \mathbf{a}_2$.

Equation (4) can be written as

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2. \quad (5)$$

Thus, (4) yields: (b_1, b_2) are coordinates of the vector \mathbf{b} in the initial orthonormal basis $\mathbf{e}_1, \mathbf{e}_2$; (x_1, x_2) are coordinates of the vector \mathbf{b} in the new basis $\mathbf{a}_1, \mathbf{a}_2$; $(a_{11}, a_{21}), (a_{12}, a_{22})$ are coordinates of the vectors $\mathbf{a}_1, \mathbf{a}_2$ in the basis $\mathbf{e}_1, \mathbf{e}_2$.

From linear algebra we know that decomposition of a vector in a new basis requires solving SLAE [Beklemishev, 1998]. Nonetheless, we can pose an inverse problem — to solve SLAE by decomposing the vector in a new basis.

Denote the scalar projection of the vector on the vector by $proj_{\mathbf{a}} \mathbf{b}$; the vector projection of the vector on the vector, by $proj_{\mathbf{a}} \mathbf{b}$.

$$proj_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{b} \cdot \mathbf{a}}{\|\mathbf{a}\|},$$

$$proj_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a},$$

where $\mathbf{b} \cdot \mathbf{a}$ is the scalar product of \mathbf{b} and \mathbf{a} , $\|\mathbf{a}\|$ is the

Euclidean norm of \mathbf{a} . The value $\frac{\mathbf{b} \cdot \mathbf{a}}{\|\mathbf{a}\|}$ reflects the projection length in units of \mathbf{e} vector lengths (original base vector).

The value $\frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}$ reflects the projection length in units of vector \mathbf{a} lengths (new base vector).

It is obvious that $\|\mathbf{b} - proj_{\mathbf{a}} \mathbf{b}\| \leq \|\mathbf{b}\|$.

Consider the problem of finding an approximate solution of SLAE. This problem arises in the absence of an accurate solution, ill-conditioned SLAE, a time limit for SLAE solution, a need for finding the variables that have the greatest impact on the SLAE solution, etc. It is this problem that faces MIT: we have to select such columns from the matrix \mathbf{A} , such a spectrum of approximating functions that best approximates GFV.

SLAE solution (3) can be found by sequentially subtracting from the vector \mathbf{b} its projections on the base vectors included in the matrix \mathbf{A} . However, to provide the most rapid convergence we should subtract the projection that can maximally reduce the norm. To do this, we should subtract the projection not on an arbitrary vector, but on the vector the projection modulus on which is maximum because the greater the subtracted projection, the lower the residual. Following this principle at each stage, we can obtain the following algorithm for solving SLAE (3):

1. $t = 0, \mathbf{b}_0 = \mathbf{b}, \mathbf{x}_0 = \mathbf{0}$
2. do
 - (a) $m = \operatorname{argmax}(\|proj_{\mathbf{a}_j} \mathbf{b}\|)$
 - (b) $\mathbf{x}_{t+1} = \mathbf{x}_t + \frac{proj_{\mathbf{a}_m} \mathbf{b}_t}{\|\mathbf{a}_m\|} \mathbf{e}_m$
 - (c) $\mathbf{b}_{t+1} = \mathbf{b}_t - proj_{\mathbf{a}_m} \mathbf{b}_t$
 - (d) $t = t + 1$
3. while ($t < MAX_T$)
 - and ($\|\mathbf{b}_t\| > NORM$)

Here, t is the iteration number; m is the number of the base column vector on which the projection is made; \mathbf{b}_t is the residual vector at the t -th iteration; \mathbf{x}_t is the vector of SLAE roots at the t -th iteration; MAX_T is the maximum permissible number of iterations; $NORM$ is the norm at which the solution should be stopped.

The maximum projection of the vector on the base vector is defined as

$$maxProj = \max(\|proj_{\mathbf{a}_j} \mathbf{b}\|),$$

$$m = \operatorname{argmax}(\|proj_{\mathbf{a}_j} \mathbf{b}\|),$$

where $\operatorname{argmax}()$ is the function returning the index of the maximum element in the array; $\max()$ is the function returning the maximum element in the array.

As a function argument in $\operatorname{argmax}()$ and $\max()$ the array of projection moduli \mathbf{b} on each \mathbf{a}_j -th base vector is used.

In the first iteration of the algorithm, the vector \mathbf{x} contains only zero elements, by the end of the algorithm the vector \mathbf{x} includes both zero and nonzero elements. The base vectors of the matrix \mathbf{A} at nonzero \mathbf{x} elements comprise the resulting basis since the base vectors at zero elements of the vector \mathbf{x} are not actually involved in the formation of the vector \mathbf{b} .

The algorithm provides such a subset of base vectors of the matrix \mathbf{A} that optimally approximates the vector \mathbf{b} . In (1) of MIT, the matrix \mathbf{A} contains spherical functions; and the vector \mathbf{b} , GFV specified at a nonuniform or uniform network of magnetic stations. Thus, the essence of MMC is in the field approximation (in our case, GFV) with the use of the optimal subset of spherical harmonics.

2. RELAXATION PARAMETER IN THE METHOD OF MAXIMUM CONTRIBUTIONS

The need has been mentioned to subtract the maximum projection, though in some cases it is impossible to identify one such projection (Figure 1). The moduli of projections on base vectors in Figure 1 are equal; the first of the maximum projections is chosen.

The extremely simplified example of SLAE solution

$$\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 5 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3/7 \\ 3/7 \end{bmatrix}$$

shows (Figure 2) that when subtracting the entire projection ($\xi=1.0$) the solution converges nonuniformly; and if we interrupt the solution upon the limit to the number of iterations, it emerges that $x_1 \neq x_2$. For a more uniform convergence, we can subtract not the entire projection, but only its part ($\xi \neq 1.0$).

Let us demonstrate the validity of this approach. Consider the arbitrary vector \mathbf{v} as the sum of two vectors

$$\mathbf{v} = \xi \mathbf{v} + (1 - \xi) \mathbf{v} \quad (6)$$

where $0 < \xi \leq 1$.

Consider the item (c) of the algorithm:

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \text{proj}_{\mathbf{a}_m} \mathbf{b}_t$$

Since $\text{proj}_{\mathbf{a}_m} \mathbf{b}_t$ is a vector, using (6) we get

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \xi \text{proj}_{\mathbf{a}_m} \mathbf{b}_t - (1 - \xi) \text{proj}_{\mathbf{a}_m} \mathbf{b}_t,$$

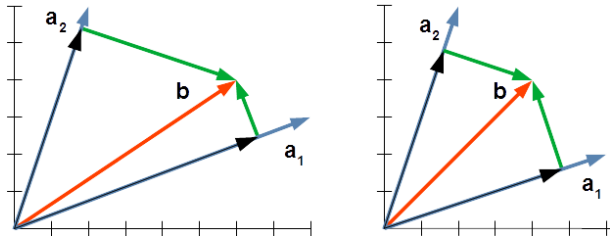


Figure 1. Selection of a base vector with maximum projection. To the left is a simple choice, the maximum projection of the vector \mathbf{b} is on the vector \mathbf{a}_1 ; to the right is an ambiguous choice because projections of \mathbf{b} on \mathbf{a}_1 , \mathbf{a}_2 are equal in modulus; \mathbf{b} (red) is the initial vector; \mathbf{a}_1 , \mathbf{a}_2 (blue) are the base vectors on which the projection is made; black marks the projections of \mathbf{b} on \mathbf{a}_1 , \mathbf{a}_2 respectively

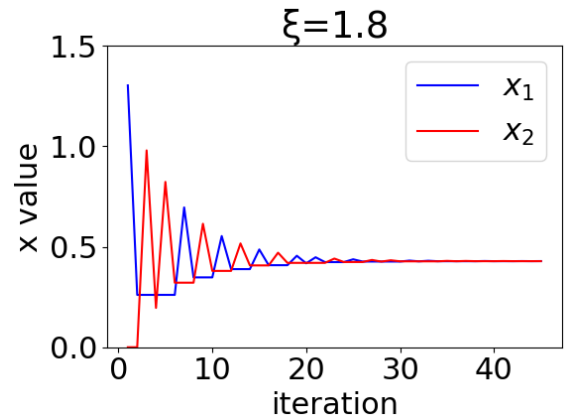
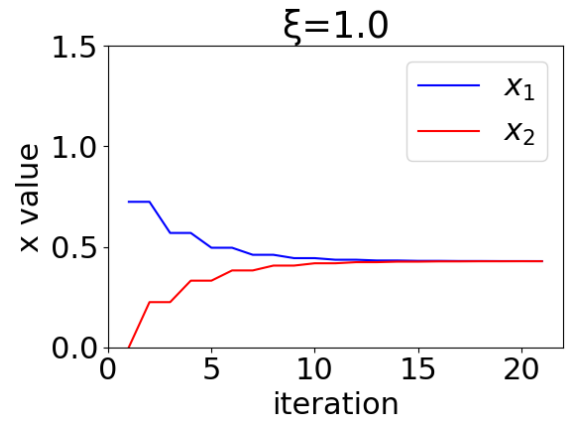
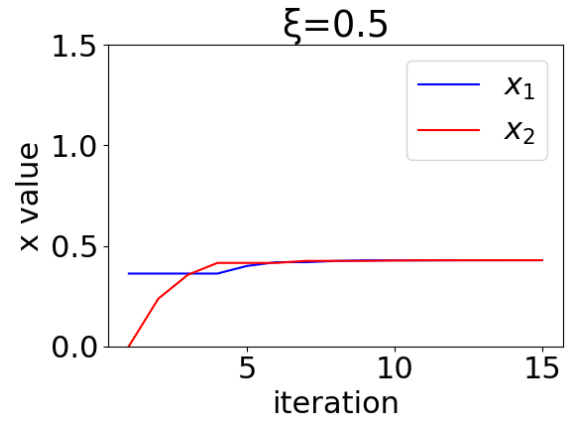
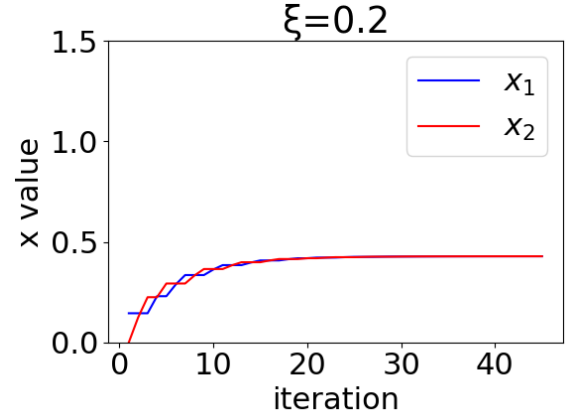


Figure 2. Course of solution at different relaxation parameters. The horizontal axis shows the iteration number, the vertical axis indicates a variable, blue is the variable x_1 , red is the variable x_2

$$\mathbf{b}_{t+1} + (1-\xi)\text{proj}_{\mathbf{a}_m} \mathbf{b}_t = \mathbf{b}_t - \xi\text{proj}_{\mathbf{a}_m} \mathbf{b}_t.$$

Thus, in the solution, not all the projection is subtracted, but only its part (Figure 2). Part of the projection remains in the residual vector. Nevertheless, the residual continues to decrease at each iteration. It is easy to see (Figures 2, 3) that the residual steadily decreases at $0 < \xi < 2$.

Then, the items (b), (c) of the algorithm can be rewritten as

$$(b) \mathbf{x}_{t+1} = \mathbf{x}_t + \xi \frac{\text{proj}_{\mathbf{a}_m} \mathbf{b}_t}{\|\mathbf{a}_m\|} \mathbf{e}_m,$$

$$(c) \mathbf{b}_{t+1} = \mathbf{b}_t - \xi\text{proj}_{\mathbf{a}_m} \mathbf{b}_t.$$

The algorithm with the above changes in terms of the MIT problem is the original MMC. In the source code of the previously developed original MMC, a constant parameter $\xi=0.7$ was specified. Furthermore, as an additional criterion for stopping the algorithm not only the residual norm is estimated but a change in the residual norm as well. MMC is based on incomplete residual vector length relaxation [Faddeev, Faddeeva 1963].

It is worth noting that introducing ξ for solving SLAE with the orthogonal basis leads only to an increase in the number of iterations. The use of $0 < \xi \leq 1$ may be justified for a nonorthogonal and/or excess basis. In this case, the other base vectors have a higher chance to fall into the final basis. Using $1 < \xi < 2$ may be justified when solving ill-conditioned SLAE. The relaxation parameter $\xi \neq 1.0$ can reduce the number of iterations sufficient for solving SLAE. Note that very small values of ξ can, on the contrary, lead to an increase in the number of iterations. The influence of the parameter ξ can be compared with the learning rate, which determines the step length in the gradient-descent method, $\mathbf{x}_{t+1} = \mathbf{x}_t - \xi \nabla F$.

Many researchers have studied optimal learning rate algorithms, but the problem is not completely solved yet. The learning rate in the gradient-descent method and the relaxation parameter in MMC have a similar mathematical sense, therefore for MMC we can use known convergence optimization techniques [Jacobs, 1988].

Consider the items (b), (c) of the algorithm. Mathematically, at any iteration the equality

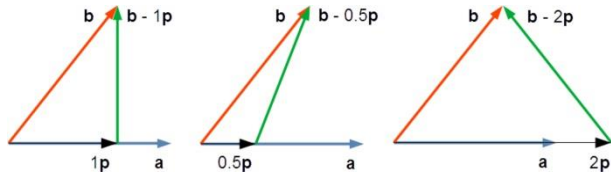


Figure 3. On the left ($\xi=1.0$) is the vector projection subtraction; in the center ($\xi=0.5$) is the half vector projection subtraction; on the right ($\xi=2.0$) is the subtraction of two vector projections: \mathbf{b} (red) is the initial vector, \mathbf{a} (blue) is the vector on which the projection is made; $\xi\mathbf{p}$ (black) is the projection of the vector on the vector multiplied by the corresponding relaxation parameter; green is a deviation from the projection of the vector on the vector multiplied by the corresponding relaxation parameter

$$\mathbf{b}_{t+1} \equiv \mathbf{b}_0 - \mathbf{A}\mathbf{x}_{t+1} \quad (7)$$

should hold. However, the implementation of the item (c) owing to the computational errors caused by floating point arithmetic [Goldberg, 1991] results in a violation of equality (7), which can give significant rounding errors at a large number of iterations. This problem can be solved in two ways.

The first way is to change the item (c) and to explicitly recalculate residual at each iteration:

$$\mathbf{b}_{t+1} = \mathbf{b}_0 - \mathbf{A}\mathbf{x}_{t+1}. \quad (8)$$

The second way is to use compensated summation algorithms [Kahan, 1965; Klein, 2006].

Note that both ways improve the numerical stability of the method, but reduces its speed. Formulas (7), (8) are shown for $\xi=1$, but it is obvious that they are easily modified for other values of ξ too.

3. COMPARISON BETWEEN METHODS OF SOLVING THE MAJOR SYSTEM OF EQUATIONS IN THE MAGNETOGRAM INVERSION TECHNIQUE

Indeterminate SLAE are solved using various methods, each having advantages and disadvantages. The best known and frequently applied method is the ordinary least squares (OLS), which is also widely used in geophysics [Rangarajan, Rao, 1975; Weimer, 1995].

The SLAE solution with OLS can be represented as

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

$$\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T \mathbf{b},$$

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}.$$

The MIT matrix $\mathbf{A}^T \mathbf{A}$ is ill-conditioned (the condition number is $\sim 10^{20}$), therefore the mathematically exact solution of SLAE is incorrect; hence the need for a less exact but more stable solution. For this purpose, to solve SLAE, along with OLS we adopt the biconjugate gradient stabilized method (BiCGStab) because it allows us to interrupt calculation at any iteration, without waiting for an exact solution.

Another method of solving indetermined SLAE popular in geophysics is based on the matrix \mathbf{A} singular value decomposition (SVD) [Pulkkinen et al., 2003]. This method calculates the pseudo-inverse matrix \mathbf{A}^+ used for solving SLAE.

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T,$$

$$\mathbf{x} = \mathbf{A}^+ \mathbf{b}.$$

One iteration performed with BiCGStab changes all elements of the vector \mathbf{x} , so it is impossible to separate individual columns (optimal spectrum). This also applies to SVD — there are no zero columns in the matrix \mathbf{A} even if we null all singular values except for the first one. Consequently, there are no zero values of the \mathbf{x} vector coordinates. Thus, SVD and BiCGStab do not allow us to select a subset of SLAE columns (optimal spectrum of approximation), as is done with MMC.

Compare methods of solving the inverse problem, using the well-studied isolated substorm that occurred on August 27, 2001 as an example [Mishin et al., 2017]. Solve system (1) by MMC and the above two methods for 03:55 UT of the substorm growth phase. Since during the substorm growth phase a smooth gain of the stationary magnetospheric convection takes place, the equivalent Hall current (ionospheric convection analog) [Lunyushkin, Pensikh, 2019] should reflect the classical plane two-vortex system of magnetospheric-ionospheric convection.

Figure 4 shows equivalent current functions for the Northern and Southern hemispheres calculated for the selected event by three different methods. It is seen that in the Northern Hemisphere, all methods generally give structurally similar equivalent current functions, but in the polar region OLS and SVD show small-scale changes in the two-vortex convection system. In the Southern Hemisphere, the expected (symmetric) two-vortex system of ionospheric convection was obtained only through MMC.

One of the current function characteristics is the transpolar current reflecting the ionospheric convection intensity through the polar cap [Bazarzhapov et al., 1979; Lunyushkin et al., 2019]. Table based of Figure 4 for each method presents current function values in two main foci, the transpolar current, and the interhemispheric ratio of transpolar currents.

Magnetospheric and auroral substorms occur in two hemispheres simultaneously [Akasofu, 1977]. The two-vortex system of ionospheric currents, GFV for two hem-

ispheres should generally be the same since they equally depend on external parameters of the solar wind and IMF, as well as on internal magnetospheric processes. The interhemispheric asymmetry is primarily affected by the tilt angle of the geomagnetic dipole. At a large angle, one of the hemispheres is more sunlit which increases the ionospheric conductivity and hence the ionospheric current.

The dipole angle for August 27, 2001, 03:55 UT is 0.24° , therefore the luminosity and related wave conductivity in the Northern and Southern hemispheres in the dipole coordinate system are virtually identical. Accordingly, the equivalent current functions and the transpolar current in the Northern and Southern hemispheres should also be almost identical. Table shows that this condition holds for MMC, but does not hold for the other two methods. The equivalent current functions of the two hemispheres obtained by MMC are generally similar, and the ratio of transpolar currents is 0.94, i.e. is close to 1. On the contrary, for the other two methods the equivalent current functions of the Northern and Southern hemispheres are markedly different, and the transpolar current in the Northern Hemisphere exceeds that in the Southern Hemisphere more than two times.

Due to the nearly zero tilt angle of the dipole, the interhemispheric similarity should also be observed in field-aligned currents and their boundaries. Figure 5 depicts field-aligned currents for uniform conductivity with boundaries obtained by an automatic method [Lunyushkin, Pensikh, 2019].

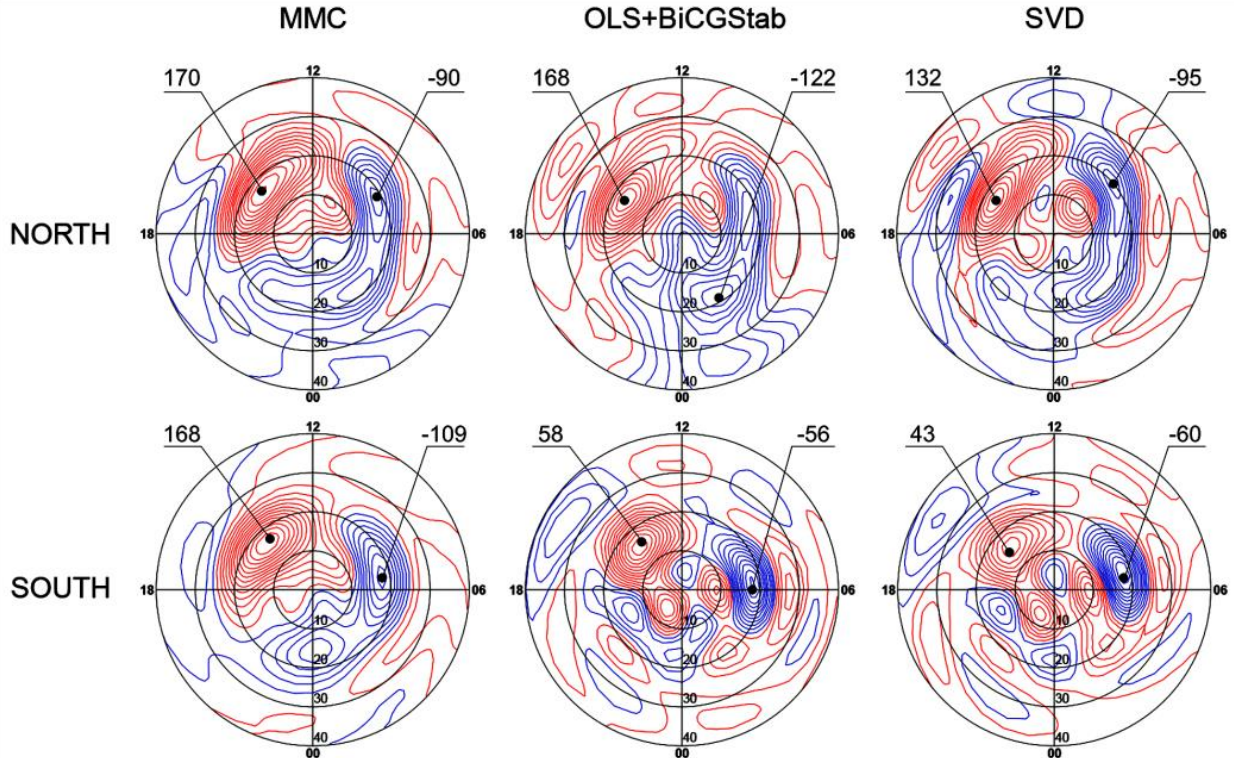


Figure 4. Comparison between results of numerical methods for solving equation (1) by the example of the equivalent current functions calculated for 03:55 UT of the August 27, 2001 isolated substorm. The top row shows current functions in the Northern Hemisphere; the bottom row, in the Southern Hemisphere. The first column presents MMC results; the second, OLS and BiCGStab results; the third, SVD results. Extension lines show values of the current function in the main foci (in kA)

Comparison between transpolar currents of two hemispheres

	Northern, kA			Southern, kA			$I_{tr,N}/I_{tr,S}$
	min	max	I_{tr}	min	max	I_{tr}	
MMC	-90	170	260	-109	168	277	0.94
OLS+BiCGStab	-122	168	290	-56	58	114	2.54
SVD	-95	132	227	-60	43	103	2.20

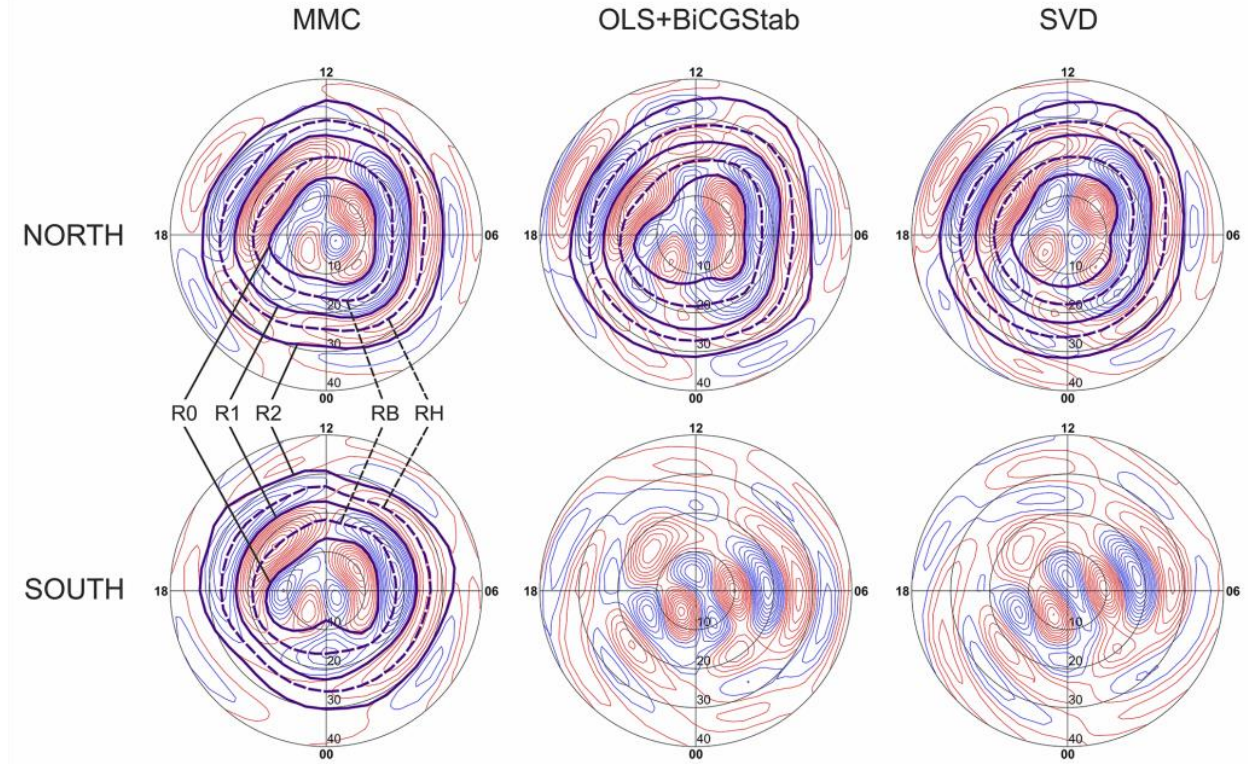


Figure 5. Comparison between results of numerical methods for solving equation (1) by the example of field-aligned currents at uniform conductivity, calculated for 03:55 UT of the August 27, 2001 isolated substorm. The top row is field-aligned currents in the Northern Hemisphere; the bottom row, in the Southern Hemisphere. The first column presents MMC results; the second, OLS and BiCGStab results; the third, SVD results. R0 is the polar cap boundary, RB is the convection reversal boundary (line of maximum field-aligned currents of zone 1), R1 is the boundary between zone 1 and zone 2 of field-aligned currents, RH is the line of maximum field-aligned currents of zone 2, R2 is the equatorial boundary of the auroral oval

The field-aligned currents of the two hemispheres obtained by MMC are generally similar, their boundaries are also similar. This was not obtained by OLS and SVD because distributions of field-aligned currents of the Northern and Southern hemispheres differ greatly, and it is practically impossible to identify their zones.

Thus, we can conclude that the solution derived with MMC proves to be the best of the three methods considered.

CONCLUSION

The paper gives details of the original MMC used in MIT. The paper has demonstrated the work of the relaxation parameter ξ in MMC and has shown that the algorithm stably converges at $0 < \xi < 2$. The influence of the relaxation parameter ξ in MMC is similar to that of the learning rate in gradient-descent methods, thus making it possible to use known algorithms for optimization. Modifications of the original method of maximum contribu-

tions have been proposed which increase the accuracy and/or the speed of solving SLAE.

For the MIT problem, MMC, OLS+BiCGStab, and SVD have been compared. MMC showed the most reasonable result, especially for the Southern Hemisphere. Calculations based on MMC also confirm the expected interhemispheric symmetry in equivalent current functions transpolar currents, distributions of field-aligned currents and their boundaries.

I am grateful to the developers of the original method of maximum contributions V.M. Mishin, B.G. Shpynev, A.D. Bazarzhapov, and D.Sh. Shirapov. I thank S.B. Lunyushkin for the goal and fruitful discussions, V.V. Mishin and A.V. Tashchilin for helpful comments.

This work was performed with financial support from RFBR under scientific project No. 19-35-90046.

For the ground magnetometer data I gratefully acknowledge: INTERMAGNET, Alan Thomson; CARISMA, PI Ian Mann; CANMOS, Geomagnetism Unit of the Geological Survey of Canada; The S-RAMP Da-

tabase, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; IMAGE, PI Liisa Juusola; Finnish Meteorological Institute, PI Liisa Juusola; Sodankylä Geophysical Observatory, PI Tero Raita; UiT the Arctic University of Norway, Tromsø Geophysical Observatory, PI Magnar G. Johnsen; GFZ German Research Centre For Geosciences, PI Jürgen Matzka; Institute of Geophysics, Polish Academy of Sciences, PI Anne Neska and Jan Reda; Polar Geophysical Institute, PI Alexander Yahnin and Yarolav Sakharov; Geological Survey of Sweden, PI Gerhard Schwarz; Swedish Institute of Space Physics, PI Masatoshi Yamauchi; AUTUMN, PI Martin Connors; DTU Space, Thom Edwards and PI Anna Willer; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; British Antarctic Survey; MacMac, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); MFGI, PI B. Heilig; Institute of Geophysics, Polish Academy of Sciences, PI Anne Neska and Jan Reda; University of L'Aquila, PI M. Vellante; BCMT, V. Lesur and A. Chambodut; Data obtained in cooperation with Geoscience Australia, PI Andrew Lewis; AALPIP, co-PIs Bob Clauer and Michael Hartinger; SuperMAG, PI Jesper W. Gjerloev; Data obtained in cooperation with the Australian Bureau of Meteorology, PI Richard Marshall. The data used in this study are available on the website of SuperMAG [<http://supermag.jhuapl.edu>].

REFERENCES

- Akasofu S.-I. *Physics of Magnetospheric Substorms*. Dordrecht, Holland, Springer, 1977, 619 p. DOI: [10.1007/978-94-010-1164-8](https://doi.org/10.1007/978-94-010-1164-8).
- Backus G., Parker R.L., Constable C. *Foundations of Geomagnetism*. Cambridge, UK, Cambridge University Press, 1996, 369 p.
- Barraclough D.R. Spherical harmonic models of the geomagnetic field. *Geomagn. Bull. Inst. Geol. Sci.* 1978, vol. 8, pp. 1–68.
- Bazarzhapov A.D., Matveev M.I., Mishin V.M. *Geomagnitnye variatsii i buri* [Geomagnetic variations and storms]. Novosibirsk, Nauka Publ., 1979, 248 p. (In Russian).
- Bazarzhapov A.D., Mishin V.M., Nemtsova E.I., Platonov M.L. The method of analytical presentation of “instant” fields of magnetic variations. *Geomagnitnye issledovaniya* [Geomagnetic Res.]. 1966, no. 8, pp. 5–22. (In Russian).
- Beklemishev D.V. *Kurs analiticheskoi geometrii i lineinoi algebry* [Course of Analytic Geometry and Linear Algebra]. Moscow, Vysshaya shkola, 1998, 320 p. (In Russian).
- Benkova N.P. *Spokoinye solnechno-sutochnye variatsii zemnogo magnetizma* [Quiet Solar Diurnal Variations of Terrestrial Magnetism]. Leningrad, Hydrometeoizdat, 1941. 76 p. (In Russian).
- Chapman S., Bartels J. *Geomagnetism, Vol. I-II*. London, Great Britain, Oxford University Press, 1940, 1049 p.
- Faddeev D.K., Faddeeva V.N. *Vychislitel'nye metody lineinoi algebry* [Computational methods of linear algebra]. 2nd Ed. Moscow, Fizmatgiz, 1963, 734 p. (In Russian).
- Fougere P.F. Spherical harmonic analysis: 1. A new method and its verification. *J. Geophys. Res.* 1963, vol. 68, no. 4, pp. 1131–1139. DOI: [10.1029/JZ068i004p01131](https://doi.org/10.1029/JZ068i004p01131).
- Gauss J.C.F. *Allgemeine Theorie des Erdmagnetismus. Resultate aus den Beobachtungen des Magnetischen Verein im Jahre 1838*. Ed. by V.B. Werke, Leipzig, Göttinger Magnetischer Verein, 1839, pp. 119–175.
- Gjerloev J.W. The SuperMAG data processing technique. *J. Geophys. Res.: Space Phys.* 2012, vol. 117, no. A9, p. A09213. DOI: [10.1029/2012ja017683](https://doi.org/10.1029/2012ja017683).
- Goldberg D. What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys*. 1991, vol. 23, no. 1, pp. 5–48. DOI: [10.1145/103162.103163](https://doi.org/10.1145/103162.103163).
- Haines G.V., Torta J.M. Determination of equivalent current sources from spherical cap harmonic models of geomagnetic field variations. *Geophys. J. Intern.* 1994, vol. 118, no. 3, pp. 499–514. DOI: [10.1111/j.1365-246X.1994.tb03981.x](https://doi.org/10.1111/j.1365-246X.1994.tb03981.x).
- Jacobs R.A. Increased rates of convergence through learning rate adaptation. *Neural Networks*. 1988, vol. 1, no. 4, pp. 295–307. DOI: [10.1016/0893-6080\(88\)90003-2](https://doi.org/10.1016/0893-6080(88)90003-2).
- Kahan W. Further remarks on reducing truncation errors. *Communications of the ACM*. 1965, vol. 8, no. 1. DOI: [10.1145/363707.363723](https://doi.org/10.1145/363707.363723).
- Klein A. A Generalized Kahan-Babuška-Summation-Algorithm. *Computing*. 2006, vol. 76, pp. 279–293. DOI: [10.1007/s00607-005-0139-x](https://doi.org/10.1007/s00607-005-0139-x).
- Lunyushkin S.B., Mishin V.V., Karavaev Y.A., Pensikh Y.V., Kapustin V.E. Studying the dynamics of electric currents and polar caps in ionospheres of two hemispheres during the August 17, 2001 geomagnetic storm. *Solar-Terr. Phys.* 2019, vol. 5, no. 2, pp. 15–27. DOI: [10.12737/stp-52201903](https://doi.org/10.12737/stp-52201903).
- Lunyushkin S.B., Pensikh Y.V. Diagnostics of the auroral oval boundaries on the basis of the magnetogram inversion technique. *Solar-Terr. Phys.* 2019, vol. 5, no. 2, pp. 88–100. DOI: [10.12737/stp-52201913](https://doi.org/10.12737/stp-52201913).
- Mandea M., Korte M. (Eds.). *Geomagnetic Observations and Models*, Dordrecht, Holland, Springer, 2010, 360 p. DOI: [10.1007/978-90-481-9858-0](https://doi.org/10.1007/978-90-481-9858-0).
- Mishin V.M. *Spokoinye geomagnitnye variatsii i toki v magnitosfere* [Quiet geomagnetic variations and currents in the magnetosphere]. Novosibirsk, Nauka Publ., 1976, 208 p. (In Russian).
- Mishin V.M. The magnetogram inversion technique and some applications. *Space Sci. Rev.* 1990, vol. 53, no. 1-2, pp. 83–163. DOI: [10.1007/bf00217429](https://doi.org/10.1007/bf00217429).
- Mishin V.M., Bazarzhapov A.D. Selection of the spectrum of Legendre polynomials approximating the observed Sq-field. *Geomagnitnye issledovaniya* [Geomagnetic Res.]. 1966, no. 8, pp. 23–30. (In Russian).
- Mishin V.M., Shpynev G.B., Bazarzhapov A.D. Computation of electric field and currents in the Earth's magnetosphere on a continuous basis by means of ground geomagnetic measurements. *Issledovaniya po geomagnetizmu, aeronomii i fizike Solntsa* [Res. on Geomagnetism, Aeronomy and Solar Phys.]. 1982, no. 58, pp. 178–186. (In Russian).
- Mishin V.M., Bazarzhapov A.D., Shpynev G.B. Mathematical analysis of the field of geomagnetic variations. *Geomagnetizm i aeronomiya* [Geomagnetism and Aeronomy]. 1984, vol. 24, no. 1, pp. 160–162. (In Russian).
- Mishin V.M., Mishin V.V., Lunyushkin S.B., Wang J.Y., Moiseev A.V. 27 August 2001 substorm: Preonset phenomena, two main onsets, field-aligned current systems, and plasma flow channels in the ionosphere and in the magnetosphere. *J. Geophys. Res.: Space Phys.* 2017, vol. 122, no. 5, pp. 4988–5007. DOI: [10.1002/2017ja023915](https://doi.org/10.1002/2017ja023915).
- Olsen N., Glassmeier K.H., Jia X. Separation of the magnetic field into external and internal parts. *Space Sci. Rev.*

2010, vol. 152, no. 1-4, pp. 135–157. DOI: [10.1007/s11214-009-9563-0](https://doi.org/10.1007/s11214-009-9563-0).

Pulkkinen A., Amm O., Viljanen A., BEAR working group. Separation of the geomagnetic variation field on the ground into external and internal parts using the spherical elementary current system method. *Earth, Planets and Space*. 2003, vol. 55, no. 3, pp. 117–129. DOI: [10.1186/BF03351739](https://doi.org/10.1186/BF03351739).

Rangarajan G.K., Rao D.R.K. A Fortran computer programme for spherical harmonic analysis of geomagnetic field by numerical integration. *Proc. Indian Acad. Sci.* 1975, vol. 82, no. 6, pp. 236–244. DOI: [10.1007/bf03046733](https://doi.org/10.1007/bf03046733).

Schmidt A. *Tafeln der Normierten Kugelfunktionen*. Gotha, Engelhard-Reyher, 1935. 52 p.

Schuster A., Lamb H. The diurnal variation of terrestrial magnetism. *Phil. Trans. R. Soc. Lond. A*. 1889, vol. 180, pp. 467–518. DOI: [10.1098/rsta.1889.0015](https://doi.org/10.1098/rsta.1889.0015).

Shirapov D.S., Mishin V.M. *Modeling of the global electrodynamic processes in the geomagnetosphere*. Ulan-Ude, East Siberian State Technological University, 2009, 217 p. (In Russian).

Shirapov D.S., Mishin V.M., Bazarshapov A.D., Saifudinova T.I. An improved version of the magnetogram inversion technique, and its application to the problem of the open magnetic flux dynamics in the geomagnetospheric tail. *Issledovaniya po geomagnetizmu, aeronomii i fizike Solntsa* [Res. on Geomagnetism, Aeronomy and Solar Phys.]. 2000, no. 111, pp. 154–172. (In Russian).

Shpynev G.B., Bazarshapov A.D., Mishin V.M. Selection of the optimal spectrum for approximating functions at analytical representation of experimental data. *Issledovaniya po geomagnetizmu, aeronomii i fizike Solntsa* [Res. on Geomagnetism, Aeronomy and Solar Phys.]. 1974, no. 32, pp. 60–65. (In Russian).

Sneeuw N. Global spherical harmonic analysis by least-squares and numerical quadrature methods in historical perspective. *Geophys. J. Intern.* 1994, vol. 118, no. 3, pp. 707–716. DOI: [10.1111/j.1365-246X.1994.tb03995.x](https://doi.org/10.1111/j.1365-246X.1994.tb03995.x).

Tikhonov A.N., Arsenin V.Y. *Solution methods for incorrect problems*. Moscow, Nauka Publ, 1979, 284 p. (In Russian).

Weimer D.R. Models of high-latitude electric potentials derived with a least error fit of spherical harmonic coefficients. *J. Geophys. Res.: Space Phys.* 1995, vol. 100, no. A10, pp. 19595–19607. DOI: [10.1029/95ja01755](https://doi.org/10.1029/95ja01755).

Yanovsky B.M. *Zemnoi magnetism* [Terrestrial Magnetism]. Leningrad, 1978, 592 p. (In Russian).

How to cite this article

Penskikh Yu.V. Applying the method of maximum contributions to the magnetogram inversion technique. *Solar-Terrestrial Physics*. 2020. Vol. 6. Iss. 4. P. 57–65. DOI: [10.12737/stp-64202009](https://doi.org/10.12737/stp-64202009).