

# The ARM-Approach Based Local Modelling of the Gravitational Field

Sultan Valeev and Konstantin Samokhvalov

Ulyanovsk State Technical University, 32 Sev Venets Str., 432027 Ulyanovsk, Russia  
{sgv,sam}@ulstu.ru

**Abstract.** Gravimetrical Measurements revealing particularities of the gravitational field, play an important role in a research after oil thanks to their capacity to reduce the cost of reconnaissance work. Graphic generalization of the field is given by gravimetrical maps of the field anomalies, drawn out with the corresponding software on the basis of one or other version of interpolation formula. Mathematical models of the trend forming regional reference-surface, can extend the capabilities of graphic generalization of measurements. With these models and corresponding contour maps, the local particularities of the gravitational field reveal themselves more clearly.

The use of models in the manner of truncated decompositions in Fourier series on spherical surfaces is offered for practical geophysical tasks. An algorithm of adaptive statistical modelling which allows optimum mathematical models to be built is considered.

In the last section we describe the developed software that allows us to build the statistical potential field models. The corresponding field maps of residual anomalies are described.

## 1 Introduction

Gravimetrical contour maps (isoanomaly contour maps) which are drawn according to the results of measurements and various reductions are used for geological object identifications which accompany oilfields and gas fields, or just hydro-carbon collectors. This information is used with (or without) some seismic and other types of measurements depending on a number of circumstances.

At present time, gravitational anomalies and, naturally, corresponding isolines on contour maps have mistakes because of incomplete account of global, regional and, in some cases, local components forming the gravitational field of a region. Besides, normal field of the Earth is fixed with sufficiently great mistakes, and mathematical presentation of the field of anomalies in points in which measurements are taken, can contain great casual mistakes because of the model inadequacy.

In practice of geodesic works, local mathematical models are not always used; more often one immediately goes from measurements on points to contouring of anomalies with the help of interpolations in graphic software packages. The usefulness of the local "background" models, averaging the influence of the rock

on "normal" Earth on the area under consideration is obvious. Forming some reference-surface they enable separate local disturbance from geological structures to reveal itself on average background.

So-called surface polynomials (usual two-dimensional algebraic polynomials) are often used to make a description of gravitational field of anomalies in the region. Their factors are defined by a standard method of least squares (MoLS), and dimensions of the model are hard fixed.

It is expected that the approach proposed in this paper will increase the accuracy of potential floor description by the mathematical model and, consequently, the accuracy of identification (in contrast with methods used in practice in Russia and other countries) from 10% to several times as large.

Using ARM-approach, based on adaptive regression modelling [1] the following is planned: 1) determination of global and regional components of the geopotential by a great multifunction array of measurements (overland and satellite data); 2) revision of normal field of the Earth; 3) building of regional and local mathematical models in the manner of optimum decompositions by Fourier series based on the overland gravimetric measurements.

## 2 Mathematical Models of Potential Fields

### 2.1 Global Models of Gravitational Field

Global models of potential fields (gravitational, magnetic and others) are necessary for solution of geophysical and other problems both for the whole Earth, its great regions, and for the building of local models. In the last case global models can be used as reference-models (relationship models).

Decompositions on spherical functions are more often used for the presentation of global models. Harmonic coefficients of those functions are defined by MoLS with lumpy distribution of data [2], [3]. Initial data for processing are the results of the Earth Satellite observing; average anomalies in free air in trapezoids are usually  $1^0 \times 1^0$ ; average one-degree altimetric heights of the geoid relative to ocean surface.

Decompositions of anomalies of gravity  $\Delta g$  are researched for the description of the global gravitational field in spherical drawing in the manner of:

$$\Delta g(z) = \frac{GM}{r^2} \sum_{n=2}^{\infty} (n-1) \frac{a^n}{r^n} \sum_{m=0}^n (\Delta \bar{C}_{nm} \cos m\lambda + \Delta \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\cos \theta) \quad (1)$$

where  $GM = G(M_0 + M_{\text{atm}})$  - is geocentric gravitational constant, taking into account the Earth atmosphere,  $a$  - big half-axis of the Earth ellipsoid;  $r, \theta=90^0 - \varphi$ ,  $\lambda$  - are spherical coordinates of the point - accordingly geocentric distance, arctic distance ( $\varphi$  - geocentric width), geographical longitude;  $\Delta \bar{C}_{nm}$ ,  $\Delta \bar{S}_{nm}$  - are differences of factors of normalizing spherical functions of real and normal fields;  $\bar{P}_{nm}(\cos \theta)$  - are normalized joined Legendre functions of  $n$  degree and  $m$  order.

Taking into account that model (1) has excluded the planet component, corresponding to normal Earth, it is possible to suppose that they include local and

regional components, generated respectively by anomalies of the gravitational field in upper mantle and in the Earth's crust. This circumstance ensures a possibility to use with a certain accuracy a mathematical model (1) to determine the identification of geological structures which generate anomalies.

Under standard approach the model (1) is considered to be determined for accepted order of decomposition  $m$ , if MoLS - evaluations of amendments to harmonica amplitudes are calculated on the surplus number of measurements of anomalies. Whereupon, we can draw a map of gravitational field (gravimetrical contour map), using expression (1) in the mode of forecast.

On the basis of the theory and computing experiments with great data arrays, we had proved that standard decompositions on spherical functions with MoLS-evaluations of amplitudes of harmonics do not fully correspond to real measurements, on which they are built [1]. The main reason of this fact is the redundancy of mathematical model, containing multiple (up to 60% and more) noise harmonics. There occurs the reduction of accuracy from 10% up to one order at the study of such models for the forecast (buildings isolines - isoanomalies or isolines in relief description), that causes corresponding displacing of isoanomalies and the loss of accuracy in the identification of geological structures.

We offered to use a statistical approach in the manner of adaptive regression modelling [1] for deciding a problem of adequacy of global models and accordingly gravimetrical contour maps of anomalies of power of gravity. This approach allows to raise the accuracy of identification of structures up to several times already at the stage of elimination of noises. Experiments on its using under comparatively small orders of decomposition are described in [1].

## 2.2 Local Models of Gravitational Field

Flat functional approximations in the manner of polynomials of low degrees are usually used for local models, but in global model decompositions on spherical functions are researched, using degree  $n$  at 360 order. However, spherical functions also can be used for regional and local models as trend functions. It is possible to use noted functional models for the forecast of regular values of anomalies.

The area of the local models application is sufficiently wide. They can assign regional or local, relation surfaces, formed by long geological structures, aside from the description of systematic behavior of a potential field. In the last case residue, for instance  $(\Delta g_i - \Delta \hat{g}_i) = e_i$  between observed and computed (forecasted) values of anomalies, can be interpreted as a manifestation of disturbances, generated by changing density of geological objects.

Accuracy of the end result, naturally, depends on the keeping of suggestions, mortgaged in the mathematical device of data processing for getting optimum models, MoLS in particular.

It is possible to offer the model

$$\Delta g(\lambda, \theta) = \sum_{n=0}^N \sum_{m=0}^n (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda) P_{nm}(\cos \theta), \quad (2)$$

as the local model, describing potential field on a segment of spherical surface. The formula (1) can be offered for the anomaly description.

### 2.3 Problems of Local Potential Fields Modelling and Ways of Their Solution

Researchers face the problems of non-observance of conditions of MoLS with all resulting consequences, when they're using approximate descriptions and MoLS. MoLS - evaluations are not the best single-line evaluations in practice.

In this paper the following problems are concerned:

1. Algorithm Development of statistical modelling of local potential fields on the basis of adaptive regression modelling (an ARM - approach);
2. The development of ARM - approach software with the use of the decompositions in Fourier series;
3. The research of ARM - approach efficiency on the example of real data array.

## 3 Methodology and Algorithms of Statistical (Regression) Modelling

### 3.1 Mathematical Model of Data Processing, to Forecast Potential Field Features

Clearly determined model of observed (measured) phenomena (objects, processes), in mathematical terms is defined as mathematical model of data processing. For instance, the relationship of anomaly of gravity power and measured coordinates of an object in some area can be described as the model

$$MY = \eta(X, \beta) \quad (3)$$

where  $Y$  - is dependent variable (anomalies of gravity power);  $X = (x_0, x_1, \dots, x_{p-1})$  - is matrix of independent variables (measured coordinates, their product or other combinations), which can change in certain areas of space  $R^0$ ;  $\beta = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$  - is the vector of unknown parameters defined according to the results of experiments;  $M$  - is an operator of math mean.

The analysis of standard strategy used in the geophysics for parametric model identification (3), has shown that for essential increasing of accuracy of presentation, it is necessary to solve a number of problems, connected with the use of MoLS.

Standard methodology can be subjected to criticism as methodology, from the point of mathematical statistics and the theory of function reconstruction.

It does not promote getting the whole possible statistical information about the model and its members, and does not help to get the identical structures and parameters according to the results of checking the hypotheses. These statements are considered in detail below.

1. Choice of measures of accuracy for the evaluation of a model quality, its fitness for the whole forecast is limited. Only one measure from the kit of internal measures exists in mathematical statistics - it is a remaining dispersion or its variety. Mixed measure of accuracy used in a number of works in the manner of forecast mistakes expects the observance of the assumption about the absence of systematic mistake in the model that is not always executed. At the same time external measures are not paid attention to, ways of their using are imperfect, and many approaches used are limited.

2. Mathematical model after postulating is considered to be hard given, regardless of its origins: no matter if it is a result of a project or it is approximated. In the last case the members of the model are not analyzed on value in the process of solution, i.e. the procedure of structured identifications is not executed.

3. In papers on modelling in geophysics interest is not shown to the circumstance that evaluations of parameters of the statistical model (3) are the best, i.e. consisted, not-shifted, efficient in the class of single-line not-shifted evaluations only in conditions of observance of a number of suggestions, considered in the following section.

### 3.2 Methodology of the ARM – Approach

**Regression Analysis (RA).** *Model of RA.* Data processing Model (1)-(3) can be presented in the manner of [1]

$$Y = X\beta + \varepsilon . \quad (4)$$

Usual mathematical presentation of direct and indirect observance resulted in the form (4) when passive experiment in geophysics is used.

- RA stipulates the performing of three stages: - model postulating,  
 - valuation of its parameters,  
 - solution of the problem of searching its optimum structure.

*RA-MoLS suggestions.* For correct using of RA it is necessary to observe a number of suggestions. One part of suggestions, given below, is generated by the statistical theory of valuation, in which condition for the sample of data and method of valuation (MoLS) are installed, the other part - is a theory of statistical criteria. Besides, we offer a suggestion (5.1) which concerns the structure identification method taking into consideration the dimensions of the problem.

In the respect of data samples  $\{y_i, x_{ii}\}$  it is supposed to be sufficiently representative, i.e.

- (1.1) - a size of observing is sufficient,
- (1.2) - casual selection of observing is ensured,
- (1.3) - the set of observing is uniform,

(1.4) - there are no rough misses of inwardly set.

Some suggestions about valued vector parameter  $\beta$  are formulated:

(2.1) - the model (4) is linear on vector  $\beta$ ,

(2.2) - no restrictions are imposed on the vector  $\beta$ ,

(2.3) - vector  $\beta$  contains the additive constant  $\beta_0$ ,

(2.4) - vector elements are calculated with a negligibly small computer inaccuracy.

Let us note suggestions about matrix  $X$ .

(3.1) - regressors  $x_0, x_1, \dots, x_{p-1}$  (matrix  $X$  columns) are linear-independent vectors of matrix  $X$  or of matrix:  $\text{rank } X = p$ ,

(3.2) - elements of matrix  $X$  are not random quantities.

The main suggestions concerning elements  $e_i$  of mistake vector  $e$  are:

(4.1) - they are additive casual mistakes,

(4.2) - they are distributed by normal law,

(4.3) - they do not contain systematic displacement,

(4.4) - they have constant dispersion,

(4.5) - they are not correlated and in the suggestion (4.2) are statistically independent.

In conclusion we will point out additional suggestions about vector  $Y$ :

(5.1) - the method of searching for optimum model or the identification of optimum kit of regressors  $\{x_i : j = \overline{1, p-1}, p-1 < p\}$  for vector  $y$  is exact,

(5.2) - MoLS can be used for each of regressions separately to solve a multi-responsible problem, containing two or more output parameters  $y_k$  ( $k < 2$ ).

Linear normal RA ensures the best (consisted, not-shifted, efficient) evaluations only if all RA conditions are observed. In practice, suggestions (1.1) - (5.2) are broken, so RA evaluations are not optimum. The degree of not-optimality is considered in detail in book [1].

*The main stages of RA.* There are three main RA stages:

(1) model postulating, (2) valuation of its MoLS parameters, (3) analysis and searching for optimum structure of the model. All these stages are considered in detail in [1].

**Adaptive regression modelling.** *Consequences of the RA-MoLS suggestions breaking.* Breaking of conditions of RA - MoLS application results in removable, insolvent and inefficient MoLS (for details see book [1]), and evaluations of both the parameters of the model processing  $\beta_j$  ( $j = \overline{0, p-1}$ ), and the values of  $\hat{Y}$  when using models in the mode of forecast.

*Revealing of breaches.* To reveal breaking of conditions of RA - MoLS use both corresponding statistics and various graphic procedures can be used [1].

*Methodology of the ARM approach.* As it was already noted, traditional methodology of the solution of problems concerning data processing is characterized by two moments: 1) the structure of model processing is considered to be given; 2) valuation of model parameters is performed by MoLS. A standard approach for given geophysical problem does not correspond to terms of accuracy. The more so because the simplicity of its use does not compensate forecast

losses in characteristics of best linear evaluations from dozens of percent to one order. It is possible to say, that a MoLS computing scheme application (without statistical analysis of the degree of breaking its usage conditions and corresponding adaptation) on the one hand makes the accuracy of the results higher and allows to consider statistically negligible factors. On the other hand (in conditions of some MoLS suggestions breaking) - it results in not-the-best evaluations of parameters and the model's prognoses characteristics become worse.

Certainly, the use of regression analysis is a significant step forward in contrast with MoLS application: 1) model analysis of a some number of criteria (R-, F - criteria) is made; 2) statistical value not only of a model as a whole, but each separate member of the model is analyzed. The last fact allows to develop to some extent the best model structure, using one or another method of structure identification.

Unfortunately, RA does not completely solve the problem of finding the best linear evaluations  $\hat{\beta}_j$  and  $\hat{Y}$ .

Additional stages for standard MoLS methodology (in system ARM - approach developed in [1]) are: 1) evaluation of model adequacy to observation and search for its optimum structure; 2) checking the observance of MoLS suggestions; 3) consequent adaptation of a processing scheme to breaking of MoLS conditions by using a kit of computing procedures (change of postulating models, methods of parametric valuation, structure identification and others. [1]); 4) application of a kit of measures (criterion of models quality, including multi-criteria concept).

### 3.3 The Algorithm of Building the Local Optimum Model of Field of Anomalies of Gravity Power

At present time, there are at least two rival scenarios of a solution of the geophysical problems of data processing, represented in the manner of models (1) - (4).

Let forecasting model be postulated in the manner of model (4).

One of the perspective scenarios of processing can be described in brief as follows:

1. Single-criteria search in the global criterion of optimum mathematical structure is realized on the base of initial description (4). Such search can be full, if computing possibilities allow, or incomplete. In the last case the method of incomplete change with the restriction can be one of pseudo-bul optimization methods. Casual or systematic mistakes, defined according to the objects being checked, which were not used in processing model building, or values of general F-criterion can serve as global (main) criterion of model quality.

2. The Second and further stages of the structured-parametric identification of optimum model are based on checking of all conditions of RA-MoLS application and consequent adaptation according to the degree of breach importance. In practice the "chains" of adaptation algorithms can be different, and it results in particular requirements to "intellectuality" automated data processing systems.

The Second rival scenario of data processing does not expect the use of global criterion in realization the last, one based on checkpoints is used at the end stage only. The base of this approach is a check of suggestion observance for initial models (4) and its consequent improvement by adaptation to the most serious breaking while the degree of characteristics distortion of best evaluations is decreasing.

### 3.4 Contouring Algorithm of the Anomalies Field

It is possible to use a method, based on the way from [3] to build isolines in the manner of isoanomalies on the equation (2). Adaptation of this way makes the following algorithm.

Let's calculate two-dimensional curves, defined by optimum mathematical description under initial model (2)

$$\sum_{n=0}^N \sum_{m=0}^n (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda) P_{nm}(\cos \theta) = 0.5k \quad (5)$$

for instance, under  $k=-4,-3,-2,-1,0,1,2,3,4,5,6$ , we will get isoanomalies, along which deflections of values  $\delta g$  from 0 will be -2, -1.5,-1 and etc to +3.

We rewrite equation (5) in the manner of

$$f(\lambda, \theta, k) = 0 \quad (6)$$

Let  $k$  be fixed, and free spot be chosen as initial  $(\lambda_0, \theta_0)$ . Then most likely that  $f(\lambda_0, \theta_0, k) \neq 0$ . To find the spot  $(\lambda_1, \theta_1)$ , for which

$$|f(\lambda_1, \theta_1, k)| \leq |f(\lambda_0, \theta_0, k)| \quad (7)$$

let us enter increments  $\delta\lambda$  and  $\delta\theta$  and require the equation to be satisfied

$$f(\lambda_0 + \delta\lambda, \theta_0 + \delta\theta, k) = 0 \quad (8)$$

In the first approximation we get

$$\left(\frac{\partial f}{\partial \lambda}\right)_{\lambda=\lambda_0} \delta\lambda - \left(\frac{\partial f}{\partial \theta}\right)_{\theta=\theta_0} \delta\theta = -f(\lambda_0, \theta_0, k) \quad (9)$$

Then, solving (9) with equation

$$\left(\frac{\partial f}{\partial \lambda}\right) \delta\lambda - \left(\frac{\partial f}{\partial \theta}\right) \delta\theta = 0 \quad (10)$$

we define

$$\delta\lambda = -\frac{f\left(\frac{\partial f}{\partial \lambda}\right)}{\left(\frac{\partial f}{\partial \lambda}\right)^2 + \left(\frac{\partial f}{\partial \theta}\right)^2} \quad (11)$$



$$\delta\theta = -\frac{f\left(\frac{\partial f}{\partial\theta}\right)}{\left(\frac{\partial f}{\partial\theta}\right)^2 + \left(\frac{\partial f}{\partial\lambda}\right)^2} . \quad (12)$$

If these values are added to  $\lambda_0$  and  $\theta_0$ , the condition (7) will be satisfied. Values  $\lambda_1$  and  $\theta_1$  which satisfy an equation (6) will be reached under following approximations. When we find one point on the curve (let it be  $(\lambda, \theta)$ ) we give it an increment along tangent, i.e. the increment must satisfy the condition

$$\left(\frac{\partial f}{\partial\lambda}\right)\delta\lambda + \left(\frac{\partial f}{\partial\theta}\right)\delta\theta = 0 . \quad (13)$$

To move forward the inner product of increments  $(\delta\lambda, \delta\theta)$  of two consequent points on the curve must be positive. Otherwise, we need to change the signs of both components. Equation (13) gives only the direction, so to move we need to choose a value of the step at each moment.

## 4 Software

### 4.1 The First Version of ASSI

The first version of developed automatic system [4] is a specialized system, realizing the strategy of the statistical (regression) modelling [1] for solution of a number of problems of mathematical relief description and gravitational fields of planets. The main purpose of the system is to obtain models of processes or phenomena with their further application for output features (responses) and realization of some functions of management in interactive (display) and packet modes of work. The need for this automatic system is due to great difficulties when performing similar work, requiring both multi-variant calculations, and using different methods of parameters valuation and structural identifications, as well as remainders analysis under chosen scenario of checking the observance of MoLS suggestions.

Module system realizing methods of regression analysis is considered to be base mathematical support. The given modules ensure the solution of normal and surplus (redefined) algebraic equation systems. Both original and modified programs of a well-known package of scientific programs in Fortran are used in this automatic system.

### 4.2 Modified Version of ASSI

The program package ASSI 1.0 [4] does not use computing facilities very much, the speed of calculations is not high enough because of the use of 16-bit code, so the process of modelling requires a lot of time. Possibilities of the package are limited by modelling when using the spherical functions of order  $N < 40$ .

New realization differs from initial one by using optimized 32-bit code, and by adding new procedures and modules and new multi-window interface. Methods

of mathematical modelling, unceasing and discrete optimization, numeric methods, the theory of probability methods are used in the program. Object-oriented programming methods were used when creating an algorithmic code, which resulted in less complicated package structure, new functions for realization of various methods of calculation have appeared.

In this realization of ASSI 2.0 for the operating system Windows 9x/2000/XP an efficient solution of the problem of relief models adequacy and gravitational fields of planets at the level of measurements accuracy and size of the information used is ensured. Most program code had been developed in Borland C++ Builder 5.0. Also Borland Delphi was used. ASSI 2.0 consists of interface, controlling part and calculation modules. Also a program for making contour maps and a program of sections building are included in the package.

## 5 Conclusion

The developed mathematical methods and software are used for getting local and regional trend models, as well as drawing maps of iso-anomalies and variations of the anomalies of gravity power for one of the regions at the Volga-river. The results prove the efficiency of the given technique.

The problems to be solved next are:

- adjustment of type models (2) within the framework of ARM-approach (reduction of the effect of interdependence of parameters of Fourier series, adaptation of MoLS use to other breaking of conditions);
- development of other type models, both on spherical and square coordinates (models with generally accepted normalizing harmonic factors in particular) which will provide more exact account in local models of global and regional components;
- expansion of type (2) models application (magnetic and other potential fields);
- practical identification of deposits (territorial expansions of investigated ranges, practical check of accuracy of geological structures localization, consideration of a problem of the depth of their presence bedding);
- improvement of software.

## References

1. Valeev, S.G.: Regression Modelling in Observations Treatment. Science, Moscow (1991) [*or* Valeev, S.G.: Regression Modelling in Data Processing. FEN, Kazan (2001)]
2. Torge, W.: Gravimetru. Walter de Gruyter, Berlin New York (1989)
3. Gudas, To.: Decomposition of Relief of the Moon on Spherical Functions / Figure of the Moon and Problems of Moon Topography. Science, Moscow (1968)
4. Valeev, S.G., Dyakov, V.I.: Automatic System for Modelling of the Megarelief and Gravitation Fields of Planets. Izvestiya VUZov. Series: Geodesy and Aerophotography **4-5** (1998) 45–49