ISSN 0145-8752, Moscow University Geology Bulletin, 2016, Vol. 71, No. 5, pp. 347–354. © Allerton Press, Inc., 2016.
Original Russian Text © D.V. Epishkin, 2016, published in Vestnik Moskovskogo Universiteta, Seriya 4: Geologiya, 2016, No. 4, pp. 40–46.

Improving Magnetotelluric Data-Processing Methods

D. V. Epishkin

Faculty of Geology, Moscow State University, Moscow, Russia

e-mail: dmitri_epishkin@mail.ru

Received March 25, 2016

Abstract—A magnetotelluric (MT) data processing algorithm that demonstrates high robustness to intense electromagnetic noise that occurs in measured MT data has been developed. The key features of the algorithm are a specific approach to estimating different transfer functions and the ability to utilize all four channels acquired at remote reference stations. The code utilizes various techniques to reduce the estimation errors, including the robust Huber estimator, jack-knife approach, improved remote reference technique, and compensating for overestimation of power spectra. The proposed algorithm has high efficiency in processing data with a low signal-to-noise ratio.

Keywords: magnetotelluric data processing, response function estimation, robust estimator

DOI: 10.3103/S0145875216050057

INTRODUCTION

Most Russian organizations utilize foreign hardware and software during industrial-scale magnetotelluric (MT) studies. Considering technological development, including the development of new Russian hardware for magnetotelluric sounding (MTS), the respective software is also necessary. At the same time, enhancing the processing of procedures compared to those that are widely used, especially under the conditions of a low signal-to-noise ratio, is an important problem.

The mentioned points determine the topicality of the development of new computational algorithms and software tools for MT data processing, i.e., for the estimation of magnetotelluric transfer operators. In the present work an algorithm for MT data processing that takes the peculiarities of different transfer operators into consideration is proposed. Various approaches to reducing overestimation are tested.

THE AIMS, PRINCIPLES, AND PROBLEMS OF MT DATA PROCESSING

MTS data processing consists of finding transfer functions (the impedance tensor and Vise-Parkinson matrix, as well as the telluric and horizontal magnetic tensors) from the records of MT field components. The main stage of processing is the separation of MT field records into spectral components, on whose basis components of the sought transfer functions are then found for a set range of frequencies (Semenov, 1985; Larsen, 1989).

The basic approach to the analysis of records during sounding has been considered in detail, for example, in (Sims et al., 1971; Semenov, 1985). A search for some transfer function underlies this approach; for example, we need to find a transfer function of the impedance tensor

$$Z = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix},$$

that would satisfy (in the mean square or any other sense) the entire set of the observed MT field:

$$E_x = Z_{xx}H_x + Z_{xy}H_y, \quad (1a)$$

$$E_y = Z_{yx}H_x + Z_{yy}H_y. \quad (1b)$$

Here, E_x and E_y are the spectra of the horizontal components of the electrical field, while H_x and H_y are the spectra of the horizontal components of the magnetic field. All of the values in Eqs. (1a) and (1b) are functions of the frequency. Equations (1a) and (1b) are traditionally solved by separating a record into segments (windows); these segments then undergo the Fourier transform and the residual error between the left and right parts of each equation in the system (1) is minimized, for example, by the least-squares method (LSM). One shortcoming of the standard LSM is that the result considerably depends on the presence and characteristics of noise, which can always be found in the measured components of the field. Some enhancements of this method have been proposed, in particular, different statistical robust procedures (Egbert and Booker, 1986; Chave et al., 1987) and the method of the remote reference point (Gamble et al., 1979). Nevertheless, under the conditions of a low signal-to-noise ratio their results still suffer from

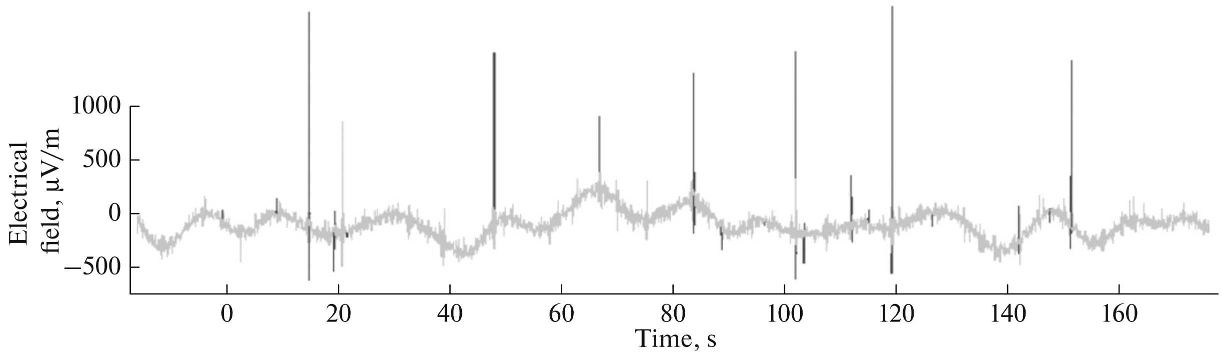

Fig. 1. Removal of peaks in records of the electrical field. The peaks are marked in black; the gray oscillating curve is the corrected time series.

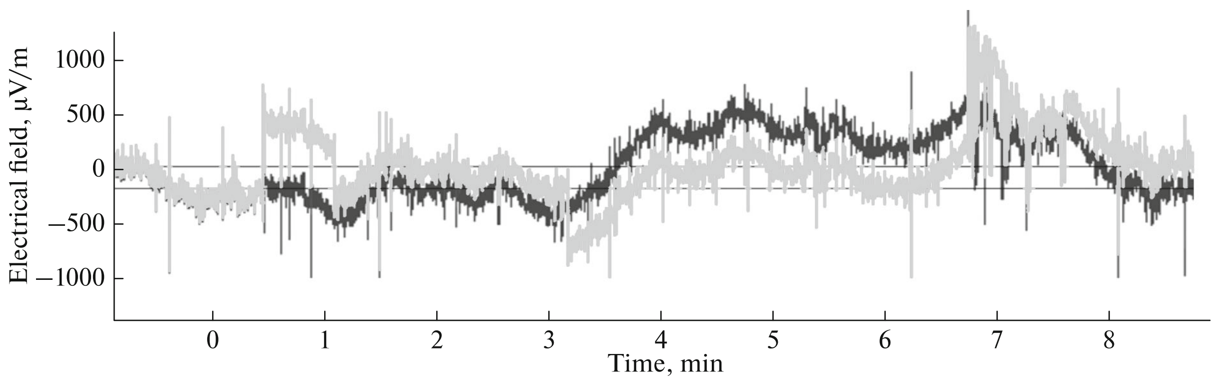

Fig. 2. Correction of steps in the records of the electrical field. The gray oscillating curve denotes the time series before correction; the black one denotes the corrected time series.

distortion (overestimation), and often very significant distortion.

The proposed algorithm provides accuracy of the estimation of MT operators and therefore the quality of MT data for subsequent interpretation. This algorithm consists of a set of sequentially performed procedures. They are considered in more detail below.

PRELIMINARY PROCESSING OF TIME SERIES

At the first stage, we transform the initial records of the measured field in order to remove two specific types of noise that are often present in records of the components of electrical fields: intensive "peaks" that significantly affect the spectrum of the signal (Fig. 1) and "steps" (Fig. 2).

To search for peaks, we choose a window (a fragment of the time series) of some width and then find a median and robust estimation of the dispersion within it (Huber, 1981). If the signal count in the window center differs from the median by more than the set number of dispersions the count is considered anomalous. Then, by sequential verification of adjacent counts, we determine the width of the anomalous zone and the signal counts that fall into it are replaced with interpolated values.

The search for steps is performed on the basis of two criteria: a greater (relative to the other points) value of the modulus of the derivative and a significant increase in signal dispersion in the vicinity of a step. The algorithm for searching for anomalies in the derivative is analogous to the algorithm for searching for peaks in the signal. The beginning of a step is found by the maximum value of the derivative (averaged in a window of several points in width) and by the maximum dispersion.

After the transformations in the time region described above, we proceed to the frequency region using the fast Fourier transform and calculate the averaged spectral characteristics of the components of the MT field on the basis of statistical procedures.

CALCULATION OF SPECTRA AND WEIGHTING FUNCTIONS

Spectra are calculated in a window of fixed width. The window moves along the time series with a set step and the Fourier transform is performed for each position. Via the application of such a scheme we can obtain a certain number of estimates for the spectra of each component of the field for the set of frequencies. We then find the quality of each spectral estimate and, in accordance with this quality, set the weights of all estimates in the range from 0 to 1.

The algorithm utilizes two robust procedures for calculating weights:

coherence- and residual error-based sorting using the jack-knife method (Jones et al., 1989) to search for and remove spectral estimates whose presence considerably worsens the result (this reduces the coherence and increases the confidence intervals);

robust *M*-estimator or Huber estimator (Huber, 1981) (see the discussion of the application of this estimate to the processing of MT data in (Egbert and Booker, 1986; Chave et al., 1987; Egbert and Livelybrooks, 1996)); in addition to the original Huber weighting function, the described algorithm also utilizes the Thompson function (Chave et al., 1987).

Sorting based on coherence is the first step; this is an auxiliary procedure that provides more effective use of the *M*-estimator at the next stage. Sorting based on coherence is a useful tool even when the signal-to-noise ratio is quite small. At the same time, the robust *M*-estimator appears to be effective in the case of weak signal (so-called dead ranges) and in the low-frequency region where samplings are small. Sorting of estimates using the jack-knife method yields better results at high frequencies, where samplings are larger.

THE SPECIFICITY OF PROCEDURES FOR ESTIMATING DIFFERENT TRANSFER FUNCTIONS

In the considered algorithm, different transfer operators are found by solving the respective optimization problems, including those using different formulations.

As an example, for the impedance tensor we use two estimators that correspond to the solutions by the method of weighted least squares for two formulations of the problem on the minimization of the residual error for Eqs. (1) (Sims et al., 1971):

$$\begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} = \begin{pmatrix} \overline{E_x H_x^*} & \overline{E_x H_y^*} \\ \overline{E_y H_x^*} & \overline{E_y H_y^*} \end{pmatrix} \begin{pmatrix} \overline{H_x H_x^*} & \overline{H_x H_y^*} \\ \overline{H_y H_x^*} & \overline{H_y H_y^*} \end{pmatrix}^{-1}, \quad (2a)$$

$$\begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} = \begin{pmatrix} \overline{E_x E_x^*} & \overline{E_x E_y^*} \\ \overline{E_y E_x^*} & \overline{E_y E_y^*} \end{pmatrix} \begin{pmatrix} \overline{H_x E_x^*} & \overline{H_x E_y^*} \\ \overline{H_y E_x^*} & \overline{H_y E_y^*} \end{pmatrix}^{-1}, \quad (2b)$$

where the overbar indicates the calculation of spectral densities by averaging on the set of windows with weight coefficients.

The weights are found while performing the above-described robust procedures. Finding weights is an iteration process, with independent calculation for the first and second estimators at each iteration and subsequent multiplication of the obtained weighting functions:

$$\overline{AB^*} = \sum_{i=1}^{M} w_i^a w_i^b A_i B_i^*. \quad (3)$$

After all of the iterations of this multiplicative weighting, we finally calculate the spectral densities on the basis of the values of the weights and then find the transfer functions.

The priority of the first estimator is higher because it minimizes the noise in electrical (which are usually more noisy) channels relative to magnetic ones (less noisy); therefore, the final calculation of the transfer functions is made with its use.

The algorithm can be applied to calculation of all of the main transfer functions: the impedance tensor and the Vise vector (tipper), as well as the telluric and magnetic tensors. In each particular case, the weighting functions are adjusted in a way to provide the maximum suppression of noise in the respective components of the MT field. In addition, each transfer function is processed with certain peculiarities.

Let us consider the algorithm for estimating the telluric tensor. In the proposed algorithm, we involved time series for eight components of the MT field. In addition to the electrical components of the field, at the reference point and the ordinary one we used the magnetic components of these points.

In order to estimate the telluric tensor using the algorithm we apply the iteration method of weighted least squares (analogous to that used when estimating the impedance). However, the weights in this case are calculated in a special way. The weight of each estimate is calculated as a product of the weights, which were preliminarily obtained via the four procedures described below:

(1) calculation of the telluric tensor using the basic algorithm described above for impedance, but under the assumption that only the field in an ordinary point is affected by noise;

(2) calculation of the telluric tensor for the case where an ordinary point is assumed to be a remote reference, and vice versa, with the respective assumption that only the field at the remote reference is noise affected;

(3) calculation of the impedance tensor at an ordinary point, so that the weights found this way reduce the influence of noise in electrical channels of an ordinary point;

(4) calculation of the impedance tensor at the remote reference to reduce the noise effects in electrical channels of the remote reference.

It was found when conducting experimental works in Kazakhstan in 2014 that inclusion of magnetic channels into processing significantly improves the quality of the estimation. Figure 3 shows the determi-

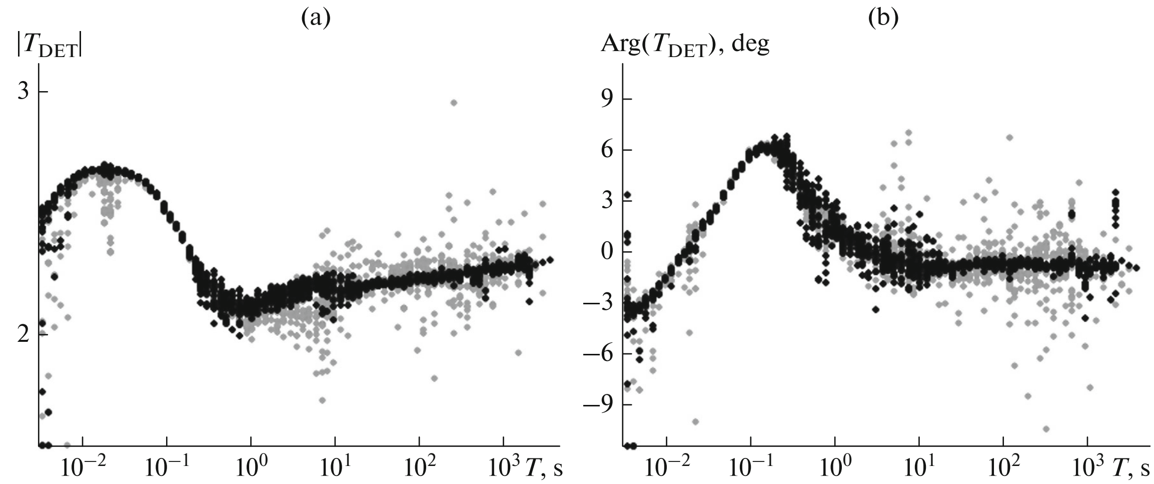

Fig. 3. The determinant of the telluric tensor as a function of the frequency: amplitude and phase. Comparison between two types of processing: (a) standard, marked with grey dots; (b) using magnetic channels, marked with black dots.

nant of the telluric tensor as a function of the frequency, comparing the results of two types of processing: basic (weighting involves procedures 1 and 2 described above) and modified (with the addition of procedures 3 and 4). At each frequency, 20 solutions were obtained (the entire time series was preliminarily partitioned into 20 segments, with independent processing being performed by two versions for each of them). Obviously, the dispersion of the solutions is less in the case of the magnetic channels involved, especially in the time interval from 10 to 100 s.

This algorithm for the estimation of T was initially developed to solve the concrete problem related to acquiring data on rock polarizability (Zorin et al., 2015). It was necessary to obtain the phase of a telluric tensor within an accuracy of 0.2° (based on estimation of the dispersion). The basic algorithm failed to solve this problem, whereas the developed algorithm provided the desired accuracy.

With the parameters of robust statistics being optimized to find the telluric tensor and with the magnetic channels included in the processing we increased the quality of the frequency dependences.

CORRECTIONS ON OVERESTIMATION

In the case of processing without a remote reference, the described algorithm utilizes the procedure based on the ideas by A. Müller (2000). A number of points are assumed when introducing corrections. In particular, it is assumed that additional components of the tensor are small in comparison with the main ones and there is a linear relationship between the value of overestimation and the noise level in magnetic channels:

$$Z^{b}_{xyi} \approx Z^{0}_{xy} - Z^{0}_{xy}\alpha_{xy}\left(\frac{1 - Coh^{2}_{0}(H_y, H_y{}^{p})_i}{1 - Coh^{2}_{0}(H_x, H_y)_i}\right), \qquad (4)$$

where Z^{b}_{xyi} is the overestimated modulus of impedance for the time interval i; Z^{0}_{xy} is the true value of the impedance modulus; α_{xy} is the linear coefficient; $Coh^{2}_{0}(H_y, H^{p}_{y})_i$ is the coherence between component H_y and its predicted value (as reconstructed from the electrical field); and $Coh^{2}_{0}(H_x, H_y)_i$ is the coherence between components H_x and H_y.

This equation contains two unknown values, Z^{0}_{xy} and α_{xy}. In order to find them, we estimate Z^{b}_{xy} for different time intervals using the robust procedures described above and calculate the coherences for the obtained solution. We then get an overdetermined system of equations and find the overestimation values using it. As was shown by the testing, this procedure can be successfully used to take the overestimation into account in the absence of a remote reference. The introduction of corrections is made after the transfer of the result to the logarithmic scale of frequency. Corrections are introduced only for the main components of the impedance tensor.

USE OF ALL CHANNELS OF A REMOTE REFERENCE

As is known, estimates of the considered transfer functions are overestimated because of the unavoidable presence of noise in the observed fields (Goubau et al., 1978). Quite a simple solution of this problem was proposed in the United States by T.D. Gamble et al. in 1978 and in the Soviet Union by I.A. Bezruk and V.O. Lakhtinov in 1979 (Zhdanov, 1986). The influence of noise can be reduced by the introduction of a field measured at some distance from the observation point (at a remote reference) in the processing.

This method to reduce noise has proved to be effective (Jones et al., 1989). The remote reference is located in such a way that the linear dependence between the desired signal at the main point and at the remote reference remained sufficient and, at the same time, there was no correlation between noise at these two points.

The solution for the impedance tensor using a remote reference is of the following form (Gamble et al., 1979):

$$\begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} = \begin{pmatrix} \overline{E_x R_x^*} & \overline{E_x R_y^*} \\ \overline{E_y R_x^*} & \overline{E_y R_y^*} \end{pmatrix} \begin{pmatrix} \overline{H_x R_x^*} & \overline{H_x R_y^*} \\ \overline{H_y R_x^*} & \overline{H_y R_y^*} \end{pmatrix}^{-1}, \quad (5)$$

where R_x and R_y are the horizontal components of magnetic field at the remote reference.

With such an approach, robust procedures are used only once, when calculating the respective spectral densities. The noise in the magnetic channels has a weaker effect on the result in this case, because its contribution to the solution considerably decreases under the condition that it weakly correlates with noise at the basic point.

In the framework of the remote reference method, the components of the electrical field are often measured simultaneously with those of magnetic field. However, it appears in most cases that electrical channels are noisy to a higher degree compared to magnetic ones and their use in conventional remote reference schemes analogous to Eq. (5) lead to even greater overestimation.

The Remote E + H algorithm proposed by the author allows remote electrical channels to be used, preventing overestimation of results.

Let there be M spectral estimates for each component of the field; E_{xi} and E_{yi} are components of the magnetic field at the observation point; H_{xi} and H_{yi} are electric field components; R_{xi} and R_{yi} are the components of the magnetic field at the basic point; E_{xi}^{ref} and E_{yi}^{ref} are the components of the electrical field at the basic point; and $i = \overline{1, M}$.

First, we perform classical processing with a remote reference:

$$\begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} = \begin{pmatrix} \overline{E_x R_x^*} & \overline{E_x R_y^*} \\ \overline{E_y R_x^*} & \overline{E_y R_y^*} \end{pmatrix} \begin{pmatrix} \overline{H_x R_x^*} & \overline{H_x R_y^*} \\ \overline{H_y R_x^*} & \overline{H_y R_y^*} \end{pmatrix}^{-1}$$

where Z_{ij} are the components of the impedance tensor and values of the form $\overline{AB^*}$ are the auto-spectra and cross-spectra, i.e., pairwise products of the spectral estimates averaged using weights obtained via robust procedures:

$$\overline{AB^*} = \sum_{i=1}^{M} w_i^a A_i B_i^*. \quad (6)$$

At the next stage, we use the electrical channels of a remote reference in order to suppress noise in the respective channels of the local point. For this purpose, we first find the robust estimate of the telluric tensor, which determines the relationship between electrical fields at the local and reference points:

$$\begin{cases} E_x^{\text{loc}} = E_x^{\text{ref}} T_{xx} + E_y^{\text{ref}} T_{xy} \\ E_y^{\text{loc}} = E_x^{\text{ref}} T_{yx} + E_y^{\text{ref}} T_{yy} \end{cases}. \quad (7)$$

Tensor T is calculated by analogy with Z, as follows

$$\begin{pmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{pmatrix} = \begin{pmatrix} \overline{E_x^{\text{loc}} E_x^{\text{ref}*}} & \overline{E_x^{\text{loc}} E_y^{\text{ref}*}} \\ \overline{E_y^{\text{loc}} E_x^{\text{ref}*}} & \overline{E_y^{\text{loc}} E_y^{\text{ref}*}} \end{pmatrix} \times \begin{pmatrix} \overline{E_x^{\text{ref}} E_x^{\text{ref}*}} & \overline{E_x^{\text{ref}} E_y^{\text{ref}*}} \\ \overline{E_y^{\text{ref}} E_x^{\text{ref}*}} & \overline{E_y^{\text{ref}} E_y^{\text{ref}*}} \end{pmatrix}^{-1}. \quad (8)$$

We then calculate the residual errors for each ith estimate of the spectrum

$$\begin{aligned} r_{xi} &= E_{xi}^{\text{loc}} - E_{xi}^{\text{ref}} T_{xx} - E_{yi}^{\text{ref}} T_{xy} \\ r_{yi} &= E_{yi}^{\text{loc}} - E_{xi}^{\text{ref}} T_{yx} - E_{yi}^{\text{ref}} T_{yy} \end{aligned}. \quad (9)$$

We then calculate the weights for each estimate according to the scheme proposed by P. Huber (1981). The weights are set to be inversely proportional to the residual errors:

$$w_i^b = \begin{cases} 1 & |r| < r_0 \\ \dfrac{r_0}{|r|} & |r| \ge r_0 \end{cases}, \quad (10)$$

where $|r|$ is the modulus of the residual error and r_0 is the robust estimate of the scale.

Finally, the auto-spectra and cross-spectra in Eq. (5) are recalculated with the new weights

$$\overline{AB^*} = \sum_{i=1}^{M} w_i^a w_i^b A_i B_i^*, \quad (11)$$

and new estimates of the impedance are found on this basis. This procedure is repeated several times up to convergence, with of weights w_i^a and w_i^b being specified after each iteration.

This approach demonstrates its effectiveness due to the local character of the electrical noise: it can be observed simultaneously at the reference and local points, but it rarely correlates, especially with a con-

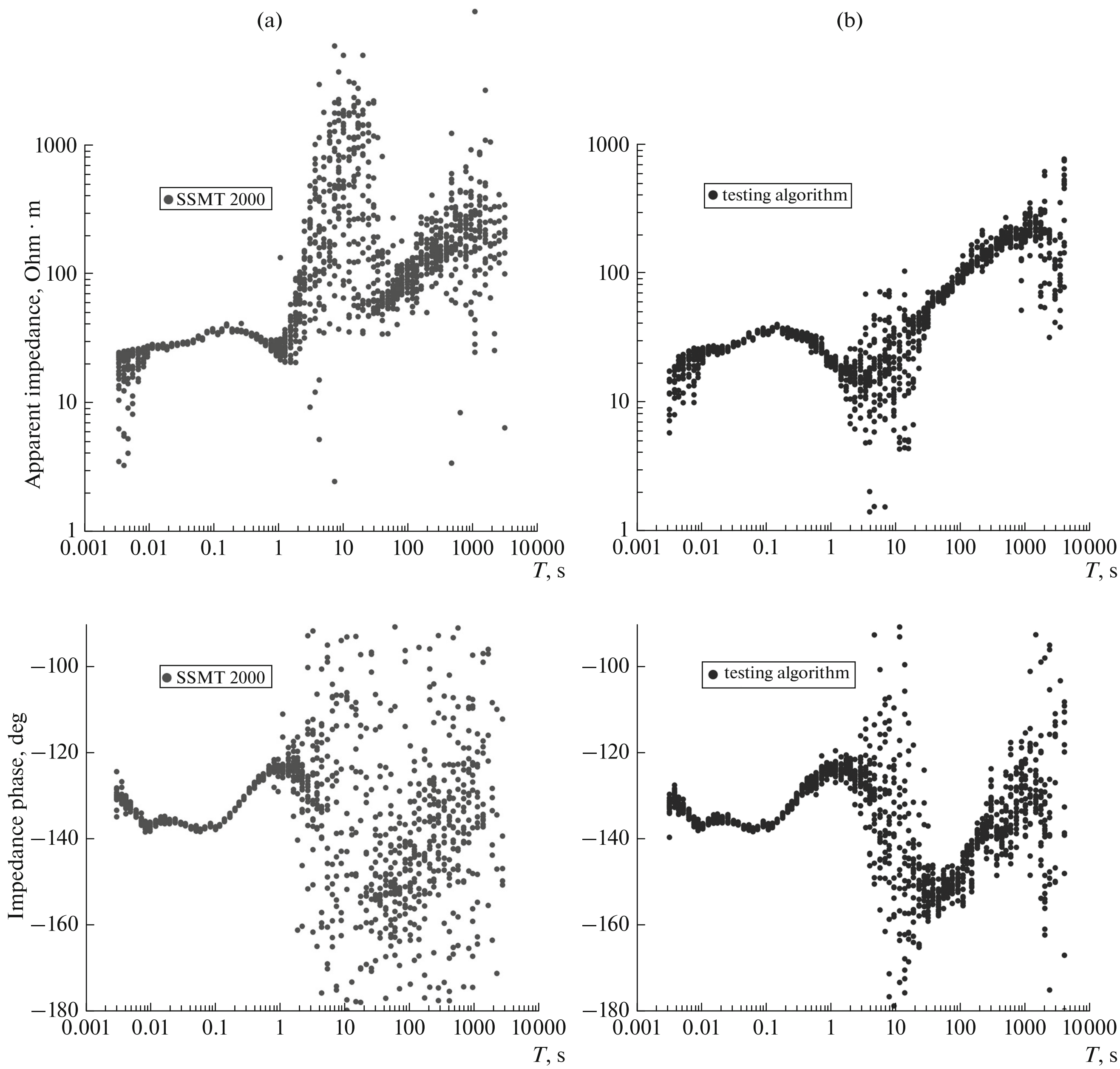

Fig. 4. Comparison between the results of robust processing using SSMT2000 software (a) and the proposed algorithm (b).

siderable distance between these points. Thus, electrical noise can be suppressed by setting small weights w_i^b that correspond to noisy runs of the record.

TESTING THE ALGORITHM

Testing of the efficiency of the algorithm was performed on a set of MT data of different qualities. Comparison was made mainly with the results obtained using SSMT 2000 software by Phoenix Geophysics (Canada).

In the case where the primary records of MT field are characterized by quite a high signal-to-noise ratio, the results from modes of simple one-point processing and processing using remote reference magnetic channels with the SSMT 2000 algorithm and the proposed one are quite similar. However, when the signal-to-noise ratio is low, the proposed algorithm is often more effective than that in SSMT 2000.

As an example, we compared the results of processing for MT data with a low signal-to-noise ratio in the range of periods of more than 1 s. It is seen in Fig. 4 that the results using the described processing algo-

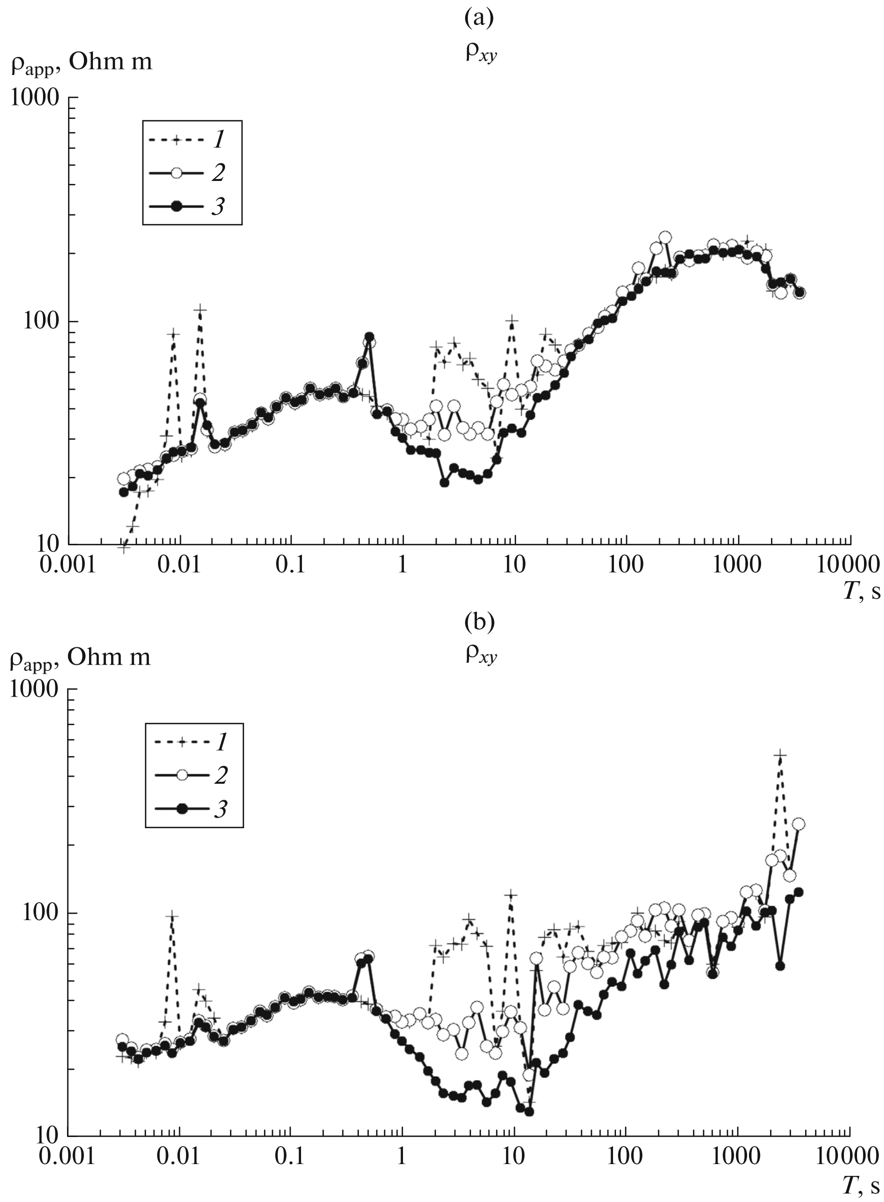

Fig. 5. Comparison between the results with the use of three processing modes: (*1*) standard robust processing (Single); (*2*) processing using remote reference magnetic channels (Remote H); (*3*) processing using four channels of a remote reference (Remote H + E).

rithm (system) are characterized by a significantly smaller dispersion of the runs and a smaller level of overestimation.

The efficiency of the Remote E + H algorithm is shown in Fig. 5, where the results of different types of processing are shown for a record with very noisy electrical channels. Processing by the Remote E + H algorithm in such a case allows us to increase the accuracy of the MTS curves in the form of smaller confidence intervals and more accurate dispersion ratios.

CONCLUSIONS

An algorithm (system) for MT data processing that is characterized by a high robustness to intensive electromagnetic noise, which is often presented in MT field records, has been developed. The distinctive feature of this algorithm is the application of special schemes of multiplicative weighting of data when calculating different transfer functions, as well as the possibility of the simultaneous use of both electrical and magnetic channels of a remote reference. The algorithm utilizes different methods to avoid overestimation of spectra, including the robust *M*-estimator, jack-knife method, remote reference, and introduction of corrections on spectral overestimation. The proposed algorithm is highly effective when processing MT data with a low signal-to-noise ratio.

REFERENCES

Chave, A.D., Thomson, D.J., and Ander, M.E., On the robust estimation of power spectra, coherences, and transfer functions, *J. Geophys. Res.*, 1987, no. 92, pp. 633–648.

Egbert, G.D. and Booker, J.R., Robust estimation of geomagnetic transfer functions, *Geophys. J. Int.*, 1986, no. 87, pp. 173–194.

Egbert, G.D. and Livelybrooks, D., Single station magnetotelluric impedance estimation: coherency weighting and the regression M-estimate, *Geophysics*, 1996, no. 61, pp. 964–970.

Gamble, T.D., Goubau, W.M., and Clarke, J., Magnetotellurics with a remote magnetic reference, *Geophysics*, 1979, no. 44, pp. 53–68.

Goubau, W.M., Gamble, T.D., and Clarke, J., Magnetotelluric data analysis: removal of bias, *Geophysics*, 1978, no. 43, pp. 1157–1166.

Huber, P.J., *Robust Statistics*, New York: John Wiley, 1981.

Jones, A.G., Chave, A.D., Egbert, G., Auld, D., and Bahr, K., A comparison of techniques for magnetotelluric response function estimation, *J. Geophys. Res.*, 1989, no. 94, pp. 14201–14213.

Larsen, J.C., Transfer functions: smooth robust estimates by least squares and remote reference methods, *Geophys. J.*, 1989, no. 99, pp. 655–663.

Müller, A., A new method to compensate for bias in magnetotellurics, *Geophys. J. Int.,* 2000, no. 142, pp. 257–269.

Semenov, V.Yu., *Obrabotka dannykh magnitotelluricheskogo zondirovaniya* (Processing of Magnetotelluric Sounding Data), Moscow: Nedra, 1985.

Sims, W.E., Bostick, F.X., and Smith, H.W., The estimation of magnetotelluric impedance tensor elements from measured data, *Geophysics*, 1971, no. 36, pp. 938–942.

Zhdanov, M.S., *Elektrorazvedka (Electrical Exploration)*, Moscow: Nedra, 1986.

Zorin, N., Epishkin, D., and Yakovlev, A., A telluric method for natural field induced polarization studies, *J. Appl. Geophys.*, 2015, no. 23.

Translated by N. Astafiev