ELABORATION OF A COMPLEX ALGORITHM OF NEURAL NETWORK SOLUTION OF THE INVERSE PROBLEM OF ELECTRICAL PROSPECTING BASED ON DATA CLASSIFICATION

S.A. Dolenko¹, I.V. Isaev¹, I.E. Obornev², E.A. Obornev², I.G. Persiantsev¹, M.I. Shimelevich²

¹D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia, e-mail: dolenko@srd.sinp.msu.ru;
²S. Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia

Abstract. The inverse problem (IP) of electrical prospecting is a complicated high-dimensional ill-posed problem with a well-known instability. To describe the sought distribution of the electrical conductivity, different parameterization schemes are used. The most general scheme uses the values of conductivity in the nodes of a pre-defined grid, with further interpolation between nodes. More specific schemes may assume presence of certain geological structures. Transfer from the solution of the IP within general scheme to its much more stable solution within one of specific schemes in a narrower class of geoelectric sections causes the necessity of prior classification of the studied data pattern, resulting in the selection of the most appropriate parameterization scheme. In their previous studies, the authors considered the solution of the IP of magnetotelluric sounding (MTS) using artificial neural networks (ANN) (perceptrons). Also, it was demonstrated that the described classification problem can be successfully solved by ANN with average rate of correct determination of the parameterization scheme exceeding 97%. Since then, the authors have elaborated a novel method of ANN-based solution of the MTS IP within scheme G₀, based on simultaneous determination of a group of several parameters at once. In this study, the developed method has been extended to other parameterization schemes. It is demonstrated that group determination of parameters is an effective method for ANN solution of the MTS IP for any parameterization scheme. In future studies, it is planned to test the complex algorithm combining classification and IP solution within partial classes of geoelectrical sections against the general approach within the most general parameterization scheme G₀.

INTRODUCTION

The inverse problem (IP) of electrical prospecting is the problem of determining the distribution of electrical conductivity (EC) in the thickness of earth by the values of electromagnetic field observed on its surface. The distribution of the EC is described by a finite number of parameters using some parameterization scheme. Usually the parameters represent the EC in a number of discrete points of the area, with subsequent interpolation between the points providing the continuous EC distribution. Some parameterization schemes also assume presence of certain geological structures within the studied area; then some of the parameters may describe the geometrical dimensions of these structures.

It is well known that the IP of electrical prospecting is a complicated high-dimensional ill-posed problem (Zhdanov, 2002). The number of observed values of electromagnetic fields is about $10^3-10^4$, and the distribution is usually described by several hundred parameters even in the 2D case. Transfer from the most general parameterization scheme using only a pre-defined spatial grid to specific parameterization schemes assuming presence of certain structures allows reducing the number of parameters; it may increase the stability of the IP and convert the general problem incorrect by Hadamard to several partial problems correct by Tikhonov. However, a solution of the IP may be only as good as close is the selected parameterization scheme to actual structure of the area. To apply this approach, the following tasks should be accomplished:

1) Elaboration of general parameterization scheme and several specific parameterization schemes;
2) Elaboration of a classification method capable to choose the most appropriate parameterization scheme by the observed values, without solving the IP for all the schemes;
3) Implementation of a technology of IP solution working within the general scheme;
4) Testing adaptability of the IP solution technology to specific parameterization schemes.

* This study was supported by the grant of Russian Scientific Foundation (project no.14-11-00579).
If all these tasks are accomplished successfully, then it would be possible to claim elaboration of a complex IP solution algorithm based on data classification. Within this algorithm, first the classification method would be applied to select the most appropriate parameterization scheme, and then the IP solution technology would be used based on this scheme.

In this paper, we report accomplishment of the four listed tasks for a specific type of electrical prospecting – magnetotelluric sounding (MTS) with the help of artificial neural networks (ANN). For MTS, the electromagnetic waves used for sounding are induced by ionosphere sources and may be assumed to be plane waves. The details of the studied case, the obtained results and the perspective of elaboration of the complex algorithm are discussed.

PARAMETERIZATION SCHEMES

The considered case is a two-dimensional case, for which the MTS IP is a coefficient IP for the Helmholtz equation (Zhdanov, p. 214) with 6500+ observed field values.

One general parameterization scheme and six specific schemes assuming presence of layers with different types of conductivity were considered (Fig. 1, 2).

Scheme $G_0$ (macro-grid, MG) is a structure present in all models. The parameters in $G_0$ are the values of electrical conductivity in fixed nodes of MG, which may change in the wide range $10^{-4} < \sigma_{i,j} < 1$ Sm/m (Fig. 1, a, blue dots). The MG is interpolated by a 2D spline for numeric solution of the direct problem. This parameterization scheme is described in more detail in (Dolenko et al, 2013).

The rest six schemes assume explicit presence of one, two or three conducting or dielectric layers with alternating properties ($G_1^C$, $G_2^{CI}$, $G_3^{CIC}$, $G_1^I$, $G_2^{IC}$, $G_3^{ICI}$). In all the models, index $C$ denotes conductor with the range of specific electrical conductivity $10^{-2} < \sigma_{i,j}^{(L)} < 1$ Sm/m, while index $I$ denotes isolator with the values of specific electrical conductivity in the range $10^{-4} < \sigma_{i,j}^{(L)} < 10^{-2}$ Sm/m. Fig. 1, b illustrates the principle of parameter setting for schemes $G_1^C$ and $G_1^I$. The MG constructed similar to MG of scheme $G_0$, is partially overlapped in its upper part by the layer. The blocks of this layer are not rectangle; the EC distribution is set by its values in the centers of blocks (red circles) and by the values of layer thickness at the edges of the grid (red arrows).
It should be separately noted that the dimensionality of this problem is very high. The number of input features of the IP (observed field values) was $N_I = 4$ field components $\times$ 13 frequencies $\times$ 126 pickets $= 6552$, and the number of output features was $N_O = 336$ parameters for scheme $G_0$ and $N_O = 233$ parameters for all the other six schemes.

CLASSIFICATION

In their previous study (Dolenko et al., 2008), the authors solved the classification problem by artificial neural networks (ANN). Within this approach, the vector of the observed data is fed to the inputs of an ANN. The ANN has several outputs, each corresponding to one of the considered classes (parameterization schemes). The distribution of amplitudes at the outputs of the classifier can be used to estimate the certainty of the classifier answer, and to make the decision about applicability of specific models for the IP solution. The decision is made using a pre-set threshold. If the highest amplitude among the outputs of the classifier exceeds the threshold value, then the classified sample is assigned to the class to which this output corresponds (this may be a correct or a wrong answer). If the highest amplitude is less than the threshold, then it is accepted that the classifier refused to recognize this sample. This is a standard technique for applying NN to solve classification problems.

As an alternative classification method to NN classification, the method of K nearest neighbors (K-NN) has been considered in the following implementation. For a studied sample, the belonging of which should be determined, K samples of the training data set are found that are nearest to it in the space of input features (K nearest neighbors). If it turns out that more than half of the nearest neighbors belong to the same class, then the studied sample is considered to be attributed by the method to the same class; otherwise it is considered that the method failed to give an answer (recognition refusal).

The Table presents the results obtained by different classification methods – single perceptron, hierarchical neural classifiers with fixed and with adaptively defined structure, probabilistic neural network, and k-NN with one nearest neighbor. The presented statistics are correct recognition rates (green), false recognition rates (pink) and refusal rates (yellow), averaged over all seven classes, and for the class with the worst result. The detailed description of these results can be found in (Dolenko et al., 2008).

<table>
<thead>
<tr>
<th>Method</th>
<th>Perceptron</th>
<th>Fixed HNC</th>
<th>Adaptive HNC</th>
<th>PNN</th>
<th>1-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold</td>
<td>0.7</td>
<td>0.1</td>
<td>0.7</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>% correct average</td>
<td>96.78</td>
<td>97.49</td>
<td>97.43</td>
<td>78.94</td>
<td>87.78</td>
</tr>
<tr>
<td>% wrong average</td>
<td>2.71</td>
<td>2.48</td>
<td>2.03</td>
<td>5.52</td>
<td>12.22</td>
</tr>
<tr>
<td>% refusals average</td>
<td>0.50</td>
<td>0.03</td>
<td>0.54</td>
<td>15.54</td>
<td>-</td>
</tr>
<tr>
<td>% correct worst</td>
<td>95.20</td>
<td>95.84</td>
<td>92.68</td>
<td>18.80</td>
<td>42.68</td>
</tr>
<tr>
<td>% wrong worst</td>
<td>3.96</td>
<td>4.16</td>
<td>6.52</td>
<td>33.36</td>
<td>57.32</td>
</tr>
<tr>
<td>% refusals worst</td>
<td>0.84</td>
<td>0.00</td>
<td>0.80</td>
<td>47.84</td>
<td>-</td>
</tr>
</tbody>
</table>

One should stress a very important peculiarity of the considered classification problem. It is mediated by the IP, i.e. the classes are determined by the distribution of the conductivity in the space of parameters, while the data for the solution of the classification problem are the values of the observed fields. Due to the incorrectness of the IP mentioned above, solvability of the classification problem with required precision was not guaranteed, and the complexity of recognition could turn out to be different for different classes. However, the results of classification turned out to be very good in these circumstances, reaching nearly 97.5% on the average for the hierarchical classifiers, and more than 95% for the worst class. So we can consider the classification task to be accomplished to an extent sufficient to move to the next task.

IP SOLUTION WITH PARAMETER GROUPING FOR GENERAL SCHEME $G_0$

There are several possible approaches to the ANN solution of multi-parameter inverse problems:

1) Solution of a separate single-output IP with construction of a separate ANN for each of the determined parameters (autonomous determination). This is the most universal approach, used most often.
2) Solution of a single IP with simultaneous determination of all the sought-for parameters. This approach requires construction of a single ANN with the number of outputs \( N_O \) equal to the full number of the determined parameters. The efficiency of this approach rapidly degrades with increase of \( N_O \).

3) Aggregation of parameters into groups with simultaneous determination of the parameters (by construction of a single ANN) within each group (group determination). The method of aggregation is governed by the physical sense of the determined parameters and by their known interrelations. This approach was investigated in (Dolenko et al, 2013), where an improvement in the IP solution quality has been demonstrated.

4) Sequential determination of parameters. Within this approach, at the first stage the values of those parameters for which this problem is solved with acceptable precision, are determined independently of each other or simultaneously. At the subsequent stages, the values of these parameters obtained by applying the ANN of the first stage are fed to the inputs of the ANN of the next stage, together with the values of the input features. This approach was investigated in (Isaev and Dolenko, 2014), where it has been demonstrated that in some cases it may also have positive effect on the IP solution precision.

Among all these approaches, the best results for scheme \( G_0 \) were demonstrated by group determination. So it was selected as the main IP solution algorithm.

As the initial input data dimensionality is very large (\( N_I = 6552 \)), prior two-step selection of significant input features was performed for each of the parameters using ANN weight analysis. Note that if such selection is performed properly, the computational cost of the ANN solution is reduced, and the quality of the IP solution is increased (Dolenko et al., 2009). The number of significant input features for various parameters ranged from 16 to 94. For group determination, each ANN was fed with all the input features significant at least for one of the determined output parameters. Depending on the size of the group and on the intersection of the sets of significant input features, the total number of input features for group determination of parameters ranged from 32 to 940.

In all experiments (except some special cases) the problem was solved by a perceptron with three hidden layers having 24, 16, and 8 neurons, linear activation function in the output layer, and logistic activation function in all the other layers. Training was stopped 500 epochs after the minimal value of the mean squared error on the test set of data has been reached.

To assess the quality of problem solution, the multiple determination coefficient \( R^2 \) (R squared), equal to 1 for exact approximation of the studied dependence with a model, and 0 for the trivial model-average, was used in this study. To reduce the dependence of the results on random factors, each ANN was trained five times, with different sets of initial weights, and the results were averaged.

Note that group determination turned out to be efficient in reducing IP solution error only then the aggregated parameters (whose values were determined together) corresponded to the blocks lying just one above another, in a single vertical, for the reasons discussed in detail in (Dolenko et al, 2013). So the considered parameters correspond to EC in the points along vertical grid edge (blue dots in Fig. 1, a).

![Fig. 3. IP solution quality, scheme \( G_0 \): a – autonomous determination (open circles) and simultaneous determination (line), b-f – group determination for grid layers 1-5 respectively, for various group positions, versus group size. Layers 1-5 describe upper 3.5 km of the section.](image-url)
Fig. 3 displays the quality of IP solution for autonomous and simultaneous determination (a) and for group determination with vertical grouping (b-f) for the five uppermost layers (total thickness 3.5 km). Different points at the diagrams for the same group size correspond to various positions of the parameter grouping window. Specially marked (with open signs and curves) are the values for window shifted up or down to maximum possible extent. Points for Group size = 1 correspond to autonomous determination.

The following facts attract attention.

1) For the uppermost block (Fig. 3, b), the quality of the group solution of the IP is increased as compared with that of autonomous determination, if the group includes up to 4 output features. However, the best result is achieved for group size $S_g=2$.

2) For the second layer block (Fig. 3, c), for which the problem solution quality in autonomous mode is the best one (Fig. 3, a), group determination always gives worse results.

3) For deeper occurring blocks, the situation depends on the position of parameter grouping window. If the window is shifted up (towards parameters that are better determined in autonomous mode), group determination improves the results; if the window is shifted down, the group determination usually makes the results worse.

In the whole, we see that group determination of parameters, when properly used, can improve the IP solution quality for scheme $G_0$. Now we need to check the applicability of this approach for other schemes of parameterization.

**IP SOLUTION WITH PARAMETER GROUPING FOR SCHEMES $G_1^C$ AND $G_1^I$**

Fig. 4 displays the quality of IP solution for scheme $G_1^C$, for autonomous determination (a) and for group determination with vertical grouping (b-f) for the layers 3-7. Points for Group size = 1 correspond to autonomous determination. Fig. 4 is similar to Fig. 3 except for the following.

As described above, scheme $G_1^C$ assumes presence of a sub-surface layer with variable thickness (Fig. 1, b). So the leftmost parameter in Fig. 4, a describes thickness of the sub-surface layer (red arrow in Fig. 1, b) just above the vertical grid edge corresponding to the grouped parameters describing EC in the points along vertical grid edge (blue dots in Fig. 1, b). The second and third parameters from the left in Fig. 4, a describe EC in the centers of the two adjacent sub-surface blocks (red circles in Fig. 1, b) lying just above this vertical grid edge. Now, the sub-surface layer fully overlaps Layer 1 of the grid (that is why the value of $R^2$ for this layer is close to zero) and partially overlaps Layer 2. So the grouping effect is studied and demonstrated for Layers 3-7.

**Fig. 4. IP solution quality, scheme $G_1^C$: a – autonomous determination, b-f – group determination for grid layers 3-7 respectively, for various group positions, versus group size. Three leftmost parameters (a) describe: Thck – sub-surface layer thickness, $SSB1$ and $SSB2$ – EC of two blocks in the sub-surface layer.**

Fig. 5 displays the quality of IP solution for scheme $G_1^I$, for autonomous determination (a) and for group determination with vertical grouping (b-f) for the layers 3-7. Points for Group size = 1 correspond to autonomous determination. Fig. 5 is similar to Fig. 4 starting from Layer 3.
Fig. 5. IP solution quality, scheme $G_1^1$: $a$ – autonomous determination, $b$-$f$ – group determination for grid layers 3-7 respectively, for various group positions, versus group size.

It can be seen from Fig. 4 and Fig. 5 that the positive effect of group determination of parameters is confirmed. One can conclude that this effect is a general property of ANN solving multi-parameter inverse problems, and therefore it can be used for other schemes of parameterization.

CONCLUSION

In this study it has been demonstrated that all the main parts needed to create a complex algorithm for ANN solution of MTS IP based on data classification have been elaborated and tested. Namely, (a) various parameterization schemes, (b) a classification algorithm able to distinguish the most appropriate scheme by the observed field values, and (c) an advanced algorithm of IP solution based on ANN with group determination of parameters, efficient with various parameterization schemes, are available.

Future studies should include putting all these parts together and testing of the complex algorithm against the approach with no classification within the most general parameterization scheme $G_0$.

ACKNOWLEDGEMENT

The authors thank A.G. Guzhva for writing the ANN software which was used to conduct this study.

REFERENCES


