

# Iterative Feature Selection with Redundancy Accounting for the Neural Network Solution of Inverse Problems of Magnetotelluric Sounding



Nickolay Shchurov , Igor Isaev , Ivan Osbornev , Eugeny Osbornev,  
Eugeny Rodionov, Mikhail Shimelevich, and Sergey Dolenko 

**Abstract** In this paper, we consider a neural network solution of the inverse problem (IP) of magnetotelluric sounding (MTS) which consists in constructing the electrical conductivity distribution in the Earth's interior from the values of the electromagnetic field components measured on its surface. It has a high input dimension (thousands of features), so it is necessary to reduce the input data dimension to achieve a more accurate and stable solution while reducing computational complexity. Neighboring measurement points and neighboring frequencies carry similar information dictating the need to use a selection method that considers this feature. The present work is devoted to the study of a method based on the iterative selection of features with the highest correlation with respect to the target variable and the exclusion of features with high cross-correlation. This method was compared with the traditional selection method, the cross-correlation filter.

**Keywords** Magnetotelluric sounding · Feature selection · Multicollinearity · Dimensionality reduction

---

N. Shchurov

Faculty of Physics, M.V. Lomonosov Moscow State University, 1/2 Leninskiye Gory,  
Moscow 119991, Russia

N. Shchurov · I. Isaev · I. Osbornev · S. Dolenko (✉)

D.V. Skobel'syn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University,  
Moscow, Russia

e-mail: [dolenko@srd.sinp.msu.ru](mailto:dolenko@srd.sinp.msu.ru)

I. Isaev

Kotel'nikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences,  
Moscow, Russia

I. Osbornev · E. Osbornev · E. Rodionov · M. Shimelevich

S. Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia

## 1 Introduction

In the solution of many physical problems with neural networks (NN), it is necessary to reduce the dimension of the input data [1]. This usually allows one to obtain a more accurate and resilient solution, also reducing computational complexity. In addition, such data preprocessing improves the generalizing ability of the model.

Exploration geophysics uses methods based on the measurement of physical fields at the Earth's surface to determine the distribution of some physical quantity in the Earth's interior. Magnetotelluric sounding (MTS) is one of such methods reconstructing electrical conductivity distribution from the properties of electromagnetic fields measured on the surface. However, such reconstruction is an inverse problem (IP) which is often ill-posed or ill-conditioned, and this IP has high dimension both by input and by output.

Such problems may be efficiently solved using NN. Starting from more general pioneering studies in the beginning of 1990s [2] and the first investigations devoted to the NN solution of the IP of MTS [3], subsequent studies of the MTS IP solution within the approximation approach differed by increasing complexity of the parameterization scheme, increasing dimensionality of the IP, and by various improvements in the methods of NN approximation [4–8]. However, the increase in the number of input features hampers the efficiency of approximation. Therefore, reduction of the input dimensionality of the problem is an important part of building its solution.

The MTS data is often characterized by multicollinearity. Therefore, a method used to select significant input features should take into account the correlation between them. There are several methods to detect multicollinearity, and various approaches are used to solve this problem [9, 10].

Feature selection (FS) is a general approach that chooses the subset of features most important for the target variable by removing irrelevant and redundant features. The methods most often used for FS in the case of high-dimensional data are filtering methods [11]. The approach studied in this article considers a special method of filter type. It iteratively selects features with the highest Pearson correlation with the target variable and discards features with high mutual correlation.

In this study, we compare the quality of the NN solution to the MTS IP on the full set of input features and on its subsets. These subsets are created using the considered selection method, as well as using traditional FS methods, such as cross-correlation based selection of significant input features.

The primary objective of this study is to test the effectiveness of the novel method of selecting essential features, whose main contribution is taking into account feature multicollinearity, in solving the MTS IP. In addition, we consider determination of the optimal parameters of the algorithm, and we compare the results obtained using this method with the results of cross-correlation based FS, and with the results obtained when neural networks are trained on the full data set.

## 2 Problem Statement

### 2.1 Parameterization Scheme

The MTS IP model considered in the present study is an integral part of the general model designed for joint application of three physical methods: magnetometry, gravimetry, and magnetotelluric sounding. To ensure the possibility of simultaneous use, it is necessary that the formulation of the problem was similar for all the physical methods considered [12]. In this case, such a formulation would consist of determining the structural boundaries that separate geological layers with constant values of the parameters: magnetization in magnetometry, density in the gravimetry problem, electrical resistivity in MTS.

The considered parameterization scheme was a 4-layer 2D model corresponding to the section of the Norilsk region, relevant in the context of ore exploration. The first layer modeled a basalt layer, the second and fourth ones—terrigenous-carbonate deposits of the Tunguska series, and the third one—gabbro-dolerite massive copper-nickel-platinum ores. The medium parameterization scheme is shown in Fig. 1, and described in more detail in [13–15].

The values of the layers resistivity were fixed, i.e. the same for the entire data set. The determined parameters were the values of the depths of the boundaries of the layers  $h(y)$  along the section, the thickness of each layer for each  $y$  was greater than zero.

The depth values for each pattern of the training sample were set randomly in the range of layer boundaries considered. Next, the direct problem was solved by the finite difference method. In this case, the six components of the EM field were calculated: the real and imaginary components of the impedance tensor  $Z$  ( $Z_{YX}$ —H polarization and  $Z_{XY}$ —E polarization) and tipper  $W$  [16, 17]. The calculation was made for 13 frequencies ranging from 0.001 to 100 Hz.

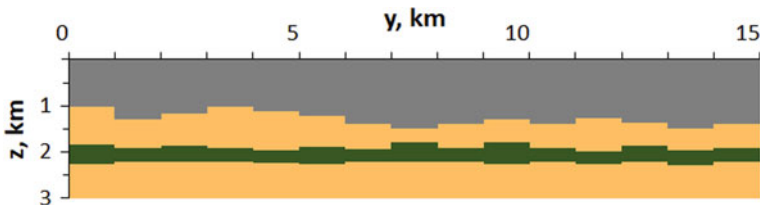


Fig. 1 Parameterization scheme

## 2.2 Data

The data array was obtained by repeatedly solving the direct problem, as stated above. For each pattern, the direct 2D problem was solved for random distribution of the depths of the boundaries of the layers [18].

Thus, the input dimension of the problem was:

6 field components (taking into account the complex – valued data presentation)  
× 13 frequencies × 31 pickets = 2418 attributes (features)

It should be noted that due to the geometry and physics of the problem, many features are correlated with each other, which is an additional argument in favor of FS.

The output dimension of the problem was: 3 layers × 15 depths = 45 parameters. A total of 30,000 patterns were calculated.

## 3 Methods of Solving the Inverse Problem

### 3.1 The Use of Neural Networks

In this study, to solve the IP, we use the type of NN called a multilayer perceptron (MLP), which is known to be a universal approximator [19–21].

Here we apply the approach of autonomous determination of parameters, when a separate single-output NN is used to determine each target parameter independently. The architecture used was an MLP with a single output and 32 neurons in the single hidden layer. To reduce the influence of weights initialization, three NN's were trained in each considered case; the statistical indicators of the solution quality of the 3 NN's were averaged. To prevent overtraining of the NN, training was stopped after 500 epochs with no improvement in the solution quality on the validation set.

The initial data set was divided into training, validation and test sets in the ratio of 70:20:10. The size of the sets was 21,000, 6,000, and 3,000 patterns, respectively.

### 3.2 Description of the Iterative Feature Selection Algorithm

Traditionally [22, 23], among the methods of selecting essential attributes based on supervised training, three groups are distinguished: filter methods, embedded methods and wrappers. Filter methods are highly computationally efficient; however, feature sets selected by them may be not optimal.

Here we consider a method based on iterative feature selection (IFS), which takes into account multicollinearity of the input features [24]. Hereafter, by “correlation” we mean the Pearson correlation.

As the first step, the algorithm selects the feature with the highest correlation with the target variable. As the second step, all features whose correlation with the one chosen at the first step was higher than some threshold value are excluded from the set.

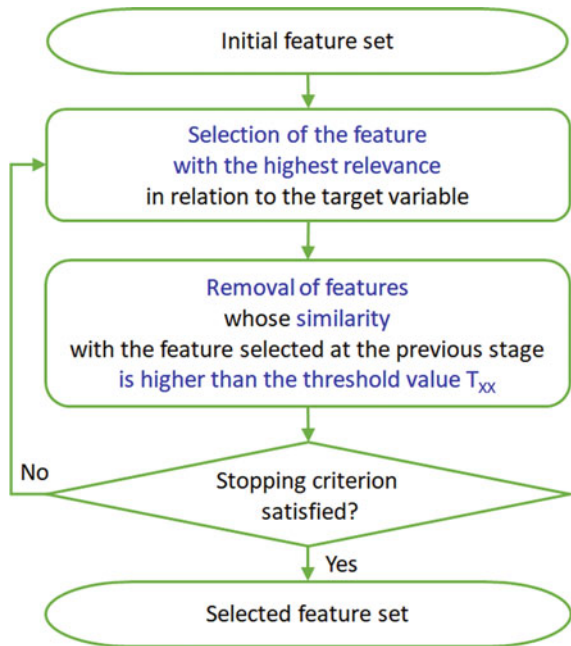
This process is repeated either until the features run out, or until there are no features left in the initial set whose correlation with the target variable is greater than a certain threshold (Fig. 2).

Thus, the described method has two parameters that need to be set:

1. The maximum allowable value of correlation with other input features  $T_{xx}$ .
2. The minimum allowable value of correlation with the target  $T_{xy}$ .

It should be noted that the studied algorithm has been already tested by the authors on an inverse problem in spectroscopy in their previous study [24], and proved its efficiency. A similar FS method is discussed in [25].

**Fig. 2** Scheme of the iterative feature selection algorithm



### 3.3 Application of the Iterative Feature Selection Algorithm

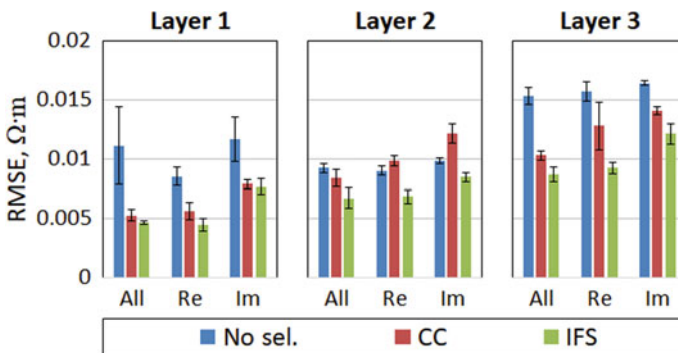
For efficient implementation of the described IFS algorithm, the absolute value of Pearson correlation coefficient between each pair of the input features was calculated. A large number of input features, especially neighboring ones, have indeed a high degree of mutual correlation. This may indicate redundancy of some features, which will be excluded from the feature set using the developed FS method.

As specified above, the method discussed in this article has two threshold values to be set ( $T_{xx}$  and  $T_{xy}$ ). We chose the values of the thresholds at which the number of selected features is 600 (~25% of all input parameters). There are several ways to achieve this. We used 4 different pairs of threshold values, and then the best ones for each of the methods were selected for further consideration.

## 4 Results of Solving the Inverse Problem

The quality of the NN solution of the inverse problem on the full feature set was compared with the quality obtained using the IFS algorithm and CC (cross-correlation based FS). The CC-based FS method calculates the correlation of each of the input features with the target variable and either takes the specified number of features with larger correlation values or takes all the features whose correlation with the target variable exceeds a pre-defined threshold. The results of comparison of the three approaches are presented in Fig. 3.

The results are provided separately for the real part of the fields, for the imaginary part of the fields, and for both parts of input data considered simultaneously. In each case, the number of selected features was set to 600 for both IFS and CC methods. It should be noted that the fourfold reduction in the number of input features with



**Fig. 3** The quality of the solution (root mean squared error) of the inverse problem on the test set on the full set of input features (No sel.), for cross-correlation feature selection (CC), and for the iterative feature selection algorithm (IFS), for 600 selected features

respect to the full feature set, when any of the two FS methods is used, causes also a significant reduction in the computational resources needed to train the NN. As for the FS procedure itself, both considered methods require computation of the correlation values for all the initial features, with their subsequent selection based on different types of comparison of the obtained values. This selection stage requires comparable numbers of comparison operations for the two methods, so the total amount of computational resources required for implementation of the two methods is nearly the same.

The presented results are for the three blocks of the central vertical line, placed above each other and equally spaced from the edges of the section. The quality of the solution of all methods decreases with increasing layer number (depth). This is due to the distortions introduced by the upper layers in the readings for the lower ones.

## 5 Conclusions

Solving the inverse problem of MTS using the IFS algorithm considered in the article gives better results than solving this IP on the full set of input features. Another popular FS method (by cross-correlation) does not show such good results due to the high multicollinearity of this problem. Thus, the IFS algorithm makes it possible to reduce the error of solving the IP of MTS with a decrease in the input dimension.

**Acknowledgements** This study has been performed at the expense of the grant of the Russian Science Foundation no. 19-11-00333, <https://rscf.ru/en/project/19-11-00333/>.

## References

1. Chang, L., Wang, J., Woodgate, W.: Analyzing spectroscopy data using two-step group penalized partial least squares regression. *Environ. Ecol. Stat.* 28(2), 445–467 (2021). DOI: <https://doi.org/10.1007/s10651-021-00496-2>
2. Raiche, A.: A pattern recognition approach to geophysical inversion using neural nets. *Geophys. J. Int.* 105(3), 629–648 (1991). DOI: <https://doi.org/10.1111/j.1365-246X.1991.tb00801.x>
3. Spichak, V., Popova, I.: Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macroparameters. *Geophys. J. Int.* 142(1), 15–26 (2000). DOI: <https://doi.org/10.1046/j.1365-246x.2000.00065.x>
4. Spichak, V., Fukuoka, K., Kobayashi, T., Mogi, T., Popova, I., Shima, H.: ANN reconstruction of geoelectrical parameters of the Minou fault zone by scalar CSAMT data. *J. Appl. Geophys.* 49(1–2), 75–90 (2002). DOI: [https://doi.org/10.1016/S0926-9851\(01\)00100-8](https://doi.org/10.1016/S0926-9851(01)00100-8)
5. Dolenko, S., Isaev, I., Osborne, E., Persiantsev, I., Shimelevich, M.: Study of Influence of Parameter Grouping on the Error of Neural Network Solution of the Inverse Problem of Electrical Prospecting. In: Iliadis, L., Papadopoulos, H., Jayne, C. (eds) *Engineering Applications of Neural Networks. EANN 2013. Communications in Computer and Information Science*, vol. 383, pp. 81–90. Springer, Berlin, Heidelberg (2013). DOI: [https://doi.org/10.1007/978-3-642-41013-0\\_9](https://doi.org/10.1007/978-3-642-41013-0_9)

6. Montahaei, M., Oskooi, B.: Magnetotelluric inversion for azimuthally anisotropic resistivities employing artificial neural networks. *Acta Geophys.* 62(1), 12–43 (2013). DOI: <https://doi.org/10.2478/s11600-013-0164-7>
7. Isaev, I., Osborne, E., Osborne, I., Shimelevich, M., Dolenko, S.: Neural Network Recognition of the Type of Parameterization Scheme for Magnetotelluric Data. In: Kryzhanovsky, B., Dunin-Barkowski, W., Redko, V., Tiumentsev, Y. (eds) *Advances in Neural Computation, Machine Learning, and Cognitive Research II. NEUROINFORMATICS 2018. Studies in Computational Intelligence*, vol. 799, pp. 176–183. Springer, Cham (2018). DOI: [https://doi.org/10.1007/978-3-030-01328-8\\_19](https://doi.org/10.1007/978-3-030-01328-8_19)
8. Wang, H., Liu, W., Xi, Z. Z., Fang, J. H.: Nonlinear inversion for magnetotelluric sounding based on deep belief network. *J. Cent. South Univ.* 26(9), 2482–2494 (2019). DOI: <https://doi.org/10.1007/s11771-019-4188-2>
9. Askin, R.G.: Multicollinearity in regression: Review and examples. *Journal of Forecasting* 1(3), 281–292 (1982). DOI: <https://doi.org/10.1002/for.3980010307>
10. Belsley, D.A., Kuh, E., Welsch, R.E.: *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons, Hoboken, New Jersey (2005)
11. Chandrashekar, G., Sahin, F.: A survey on feature selection methods. *Comput. Electr. Eng.* 40(1), 16–28 (2014). DOI: <https://doi.org/10.1016/j.compeleceng.2013.11.024>
12. I. Isaev, E. Osborne, I. Osborne, E. Rodionov, M. Shimelevich, V. Shirokiy, S. Dolenko.: Using Domain Knowledge for Feature Selection in Neural Network Solution of the Inverse Problem of Magnetotelluric Sounding. A. V. Samsonovich et al. (Eds.): *Brain-Inspired Cognitive Architectures for Artificial Intelligence: BICA\*AI 2020. Advances in Intelligent Systems and Computing*, 2021, V.1310, pp.115–126. DOI: [https://doi.org/10.1007/978-3-030-65596-9\\_15](https://doi.org/10.1007/978-3-030-65596-9_15)
13. Isaev, I., Osborne, I., Osborne, E., Rodionov, E., Shimelevich, M., Dolenko, S.: Neural Network Solution of Inverse Problems of Geological Prospecting with Discrete Output. *Proc. of Science* 410, 003 (2021). DOI: <https://doi.org/10.22323/1.410.0003>
14. Isaev, I., Osborne, I., Osborne, E., Rodionov, E., Shimelevich, M., Dolenko, S.: Comparison of data integration methods for neural network solution of the inverse problem of exploration geophysics. In: VIII International Conference on Information Technology and Nanotechnology (ITNT), pp. 1–4. IEEE (2022). DOI: <https://doi.org/10.1109/ITNT55410.2022.9848628>
15. Isaev, I., Osborne, I., Osborne, E., Rodionov, E., Shimelevich, M., Dolenko, S.: Integration of Geophysical Methods for Solving Inverse Problems of Exploration Geophysics Using Artificial Neural Networks. In: Kosterov, A., Bobrov, N., Gordeev, E., Kulakov, E., Lyskova, E., Mironova, I. (eds) *Problems of Geocosmos–2020. Springer Proceedings in Earth and Environmental Sciences*, pp.77–87. Springer, Cham, (2022). DOI: [https://doi.org/10.1007/978-3-030-91467-7\\_7](https://doi.org/10.1007/978-3-030-91467-7_7)
16. Berdichevsky, M., Dmitriev, V.: *Models and Methods of Magnetotellurics*. Springer, Heidelberg (2010)
17. Zhdanov, M.: *Geophysical Electromagnetic Theory and Methods. Methods in Geochemistry and Geophysics*. Elsevier, Amsterdam (2009)
18. Shimelevich, M., Osborne, E., Osborne, I., Rodionov, E.: An algorithm for solving inverse geoelectrics problems based on the neural network approximation. *Numer. Anal. Appl.* 11(4), 359–371 (2018). DOI: <https://doi.org/10.1134/S1995423918040080>
19. Hornik, K., Stinchcombe, M., and White, H.: Multilayer Feedforward Networks are Universal Approximators. *Neural Netw.* 2(5), 359–366 (1989). DOI: [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
20. Cybenko, G.: Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems* 2, 303–314 (1989). DOI: <https://doi.org/10.1007/BF02551274>
21. Funahashi, K.: On the Approximate Realization of Continuous Mappings by Neural Networks. *Neural Netw.* 2(3), 183–192 (1989). DOI: [https://doi.org/10.1016/0893-6080\(89\)90003-8](https://doi.org/10.1016/0893-6080(89)90003-8)
22. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182 (2003)
23. Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R.P., Tang, J., Liu, H.: Feature selection: a data perspective. *ACM Comput. Surv.* 50(6), 1–45 (2017). DOI: <https://doi.org/10.1145/3136625>



24. Shchurov, N.O., Isaev, I.V., Burikov, S.A., Dolenko, T.A., Laptinskiy, K.A., Dolenko, S.A.: Taking into Account Mutual Correlations during Selection of Significant Input Features in Neural Network Solution of Inverse Problems of Spectroscopy. *Proc. of Science* 429, 026 (2022). DOI: <https://doi.org/10.22323/1.429.0026>
25. Biesiada, J., Duch, W.: Feature selection for high-dimensional data — a Pearson redundancy based filter. In: Kurzynski, M., Puchala, E., Wozniak, M., Zolnierek, A. (eds) *Computer Recognition Systems 2. Advances in Soft Computing*, vol. 45, pp. 242–249. Springer, Berlin, Heidelberg (2007). DOI: [https://doi.org/10.1007/978-3-540-75175-5\\_30](https://doi.org/10.1007/978-3-540-75175-5_30)