V International Workshop on Deep Learning in Computational Physics

June 29, 2021. Moscow.

NEURAL NETWORK SOLUTION OF INVERSE PROBLEMS OF GEOLOGICAL PROSPECTING WITH DISCRETE OUTPUT*

Igor Isaev[1,2], Ivan Obornev[1], Eugeny Obornev[3], Eugeny Rodionov[3], Mikhail Shimelevich[3] and Sergey Dolenko[1].

[1] D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia.

[2] Kotelnikov Institute of Radio Engineering and Electronics, RAS, Moscow, Russia.

[3] S.Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia.

This study has been performed at the expense of the grant of the Russian Science Foundation (project no. 19-11-00333).

Inverse problems of exploration geophysics

Inverse problems are the type of problems
when the system parameters are determined
from the observed data describing the system state.

The inverse problem of exploration geophysics
consists in reconstructing the spatial distribution
of the properties of the medium in the Earth's interior
from measurements on its surface.

Features of inverse problem of exploration geophysics:

- ✓ Nonlinear
- ✓ Multi-parametrical
- ✓ High-dimensional
- ✓ ill-posed
- ✓ ill-conditioned
- ✓ In general case do not have a direct numerical solution

Introduction

Solution methods

Traditional solution methods:

- Optimization methods based on
  the multiple solution of the direct problem
  with the minimization of residuals in the space of the observed fields
  - ✓ High computational cost and low speed of work
  - ✓ Need for a good first approximation
  - ✓ Need to have a correct model for solving the direct problem
  - ✓ Small residual in the space of the observed quantities does not guarantee a small residual in the space of the determined parameters
- Matrix methods based on regularization
  - ✓ Need to choose the regularization parameter.
  - ✓ Linear method.
    It is necessary to perform nonlinear data preprocessing.

Introduction

Solution methods

Neural network solution is considered as an alternative.

- ❑ Neural network solution
  - ✓ Free from the disadvantages of traditional methods
  - ✓ High computational cost when using machine learning methods are shifted from the stage of application of the computing system to the stage of its development, which increases the convenience of practical use of such a system.

To improve the quality of the solution,
the integration of geophysical methods is considered –
the simultaneous use of data from several geophysical methods

Neural network application scheme

The scheme of application of neural networks
for solving inverse problems of exploration geophysics:

- ❑ Define a parameterization scheme with a finite number of parameters
- ❑ Create a training data set:
  - For each training pattern
    - ✓ Set a random distribution of parameters on macrogrid
    - ✓ Calculate distribution of parameters on microgrid
    - ✓ Calculate field values by solving the direct problem using the finite difference method
- ❑ Train neural networks on a training dataset
- ❑ Apply neural networks to the studied data

To use the integration of geophysical methods, it is necessary that the determined parameters of each method are the same.

Problem statement

Parameterization scheme

Description:

- 2D model (section)
- 4 layers
  - ✓ The physical characteristics of the 2-nd and 4-th layers were the same
- The physical properties of the layers are fixed
  - ✓ Do not change within the section
  - ✓ Do not change in the entire dataset
- Variable (determined) parameters
  - ✓ Depths of the lower boundaries of layers
- Calculated physical fields
  - ✓ Gravimetry
  - ✓ Magnetometry
  - ✓ Magnetotelluric sounding

Problem statement

Properties of the layers

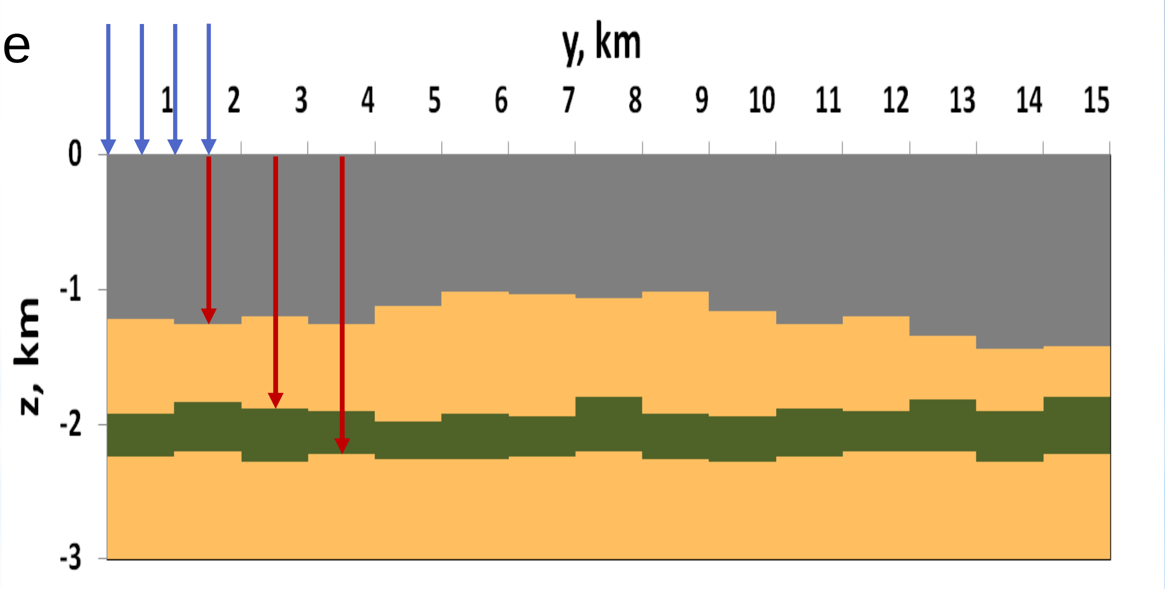
Layer	**Description**	**Physical properties**			**Spatial properties**		
		Density **σ**, kg/m³	Magnetization **μ**, A/m	Resistivity **ρ**, Ω·m	Upper bound, min-max, km	Lower bound, min-max, km	Thickness, min-max, km
1	Basalt	2 800	3	2 000	0	1 – 1.48	1 – 1.48
2	Terrigenous carbonate deposits of the Tunguska series	2 550	0.5	100	1 - 1.48	1.8 – 1.98	0.32 — 0.98
3	Gabbro-dolerites massive copper-nickel-platinum ores	3 000	0.9	1 000	1.8 - 1.98	2.2 — 2.28	0.22 — 0.48
4	Terrigenous carbonate deposits of the Tunguska series	2 550	0.5	100	2.2 — 2.28	—	—

Problem statement

Parameterization scheme

Properties:

- ❑ Geological section size
  - ✓ Depth- 3 km
  - ✓ Width -15 km
- ❑ Physical field measurement step
  - ✓ 0.5 km
  - ✓ 31 measurement points along the profile
- ❑ Step of changing the boundaries of geological layers
  - ✓ 1 km
  - ✓ 15 depth values for each layer



- ❑ The discreteness of changing the values of depths
  - ✓ 0.02 km

Purpose of the study

- Comparison of the neural network solution obtained by directly solving the regression problem to the approach based on “one-hot encoding” of discrete outputs and the subsequent solution of the classification problem.
- The study of the effect of the integration of geophysical methods using regression and classification approaches.

Computational experiment

Dataset

- Dataset
  - Were obtained by numerical solution of the direct problem
  - Number of patterns 30 000 patterns
  - Split into sets:

✓ Training set	70%	21 000 patterns
✓ Validation set	20%	6 000 patterns
✓ Test set	10%	3 000 patterns

Computational experiment

Data

- Data dimensionality
  - Output dimensionality
    - ✓ *45 parameters = 3 layers * 15 values of layer boundary depth*
  - Input dimensionality
    - ✓ Gravimetry:

      *31 features = 1 field component * 31 measurement point (picket)*
    - ✓ Magnetometry:

      *31 features = 1 field component * 31 picket*
    - ✓ Magnetotelluric Sounding:

      *62 features = 2 field components * 1 frequency * 31 picket*
    - ✓ Integration of geophysical methods:

      62, 93 or 124 features

Computational experiment

Use of neural networks

- Architecture:
  - Multilayer perceptron
  - 1 hidden layer - 32 neurons
  - Activation function:
    - ✓ hidden layer – sigmoid
    - ✓ output layer: linear – for regression approach
sigmoid – for classification approach
- Prevent overfitting - early stopping method
  - Stop training after 500 epochs
with no improvement on the validation set
- Weights initialization
  - Each neural network was trained 5 times
with various initial weights values.
  - The statistic indexes of the results of application of the 5 networks were averaged

Computational experiment

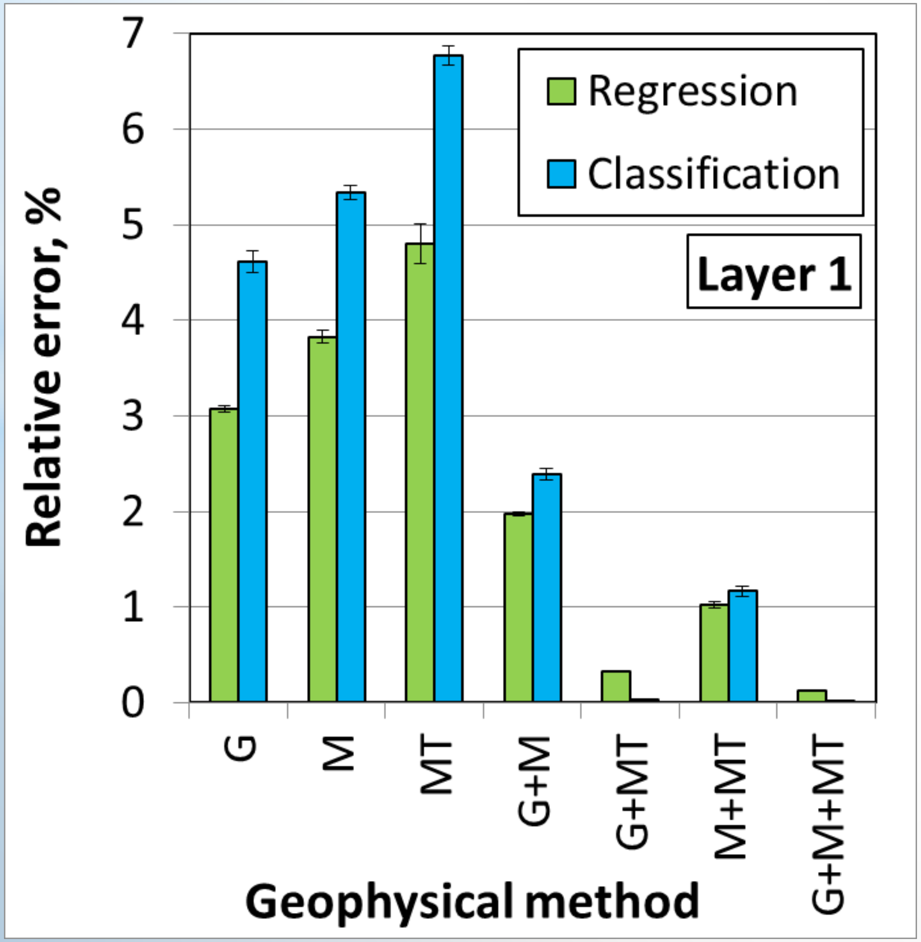
Use of neural networks

- ❑ Regression approach
  - Autonomous determination
    - ✓ Individual determination of each parameter
    - ✓ Training a separate single-output neural network
- ❑ Classification approach
  - Autonomous determination
  - "One-hot encoding"
    - ✓ The number of outputs is equal
      to the number of discrete values of determined parameter.
    - ✓ 25, 10, 5 discrete values for 1-st, 2-nd, 3-rd layer respectively
    - ✓ All outputs are "0" except for one
    - ✓ The position corresponding to the parameter value is set to "1"
  - Classification
    - ✓ The solution is the output with the maximum amplitude

Results

Layer 1

Dependence of the quality of the solution on input data
for regression and classification approach

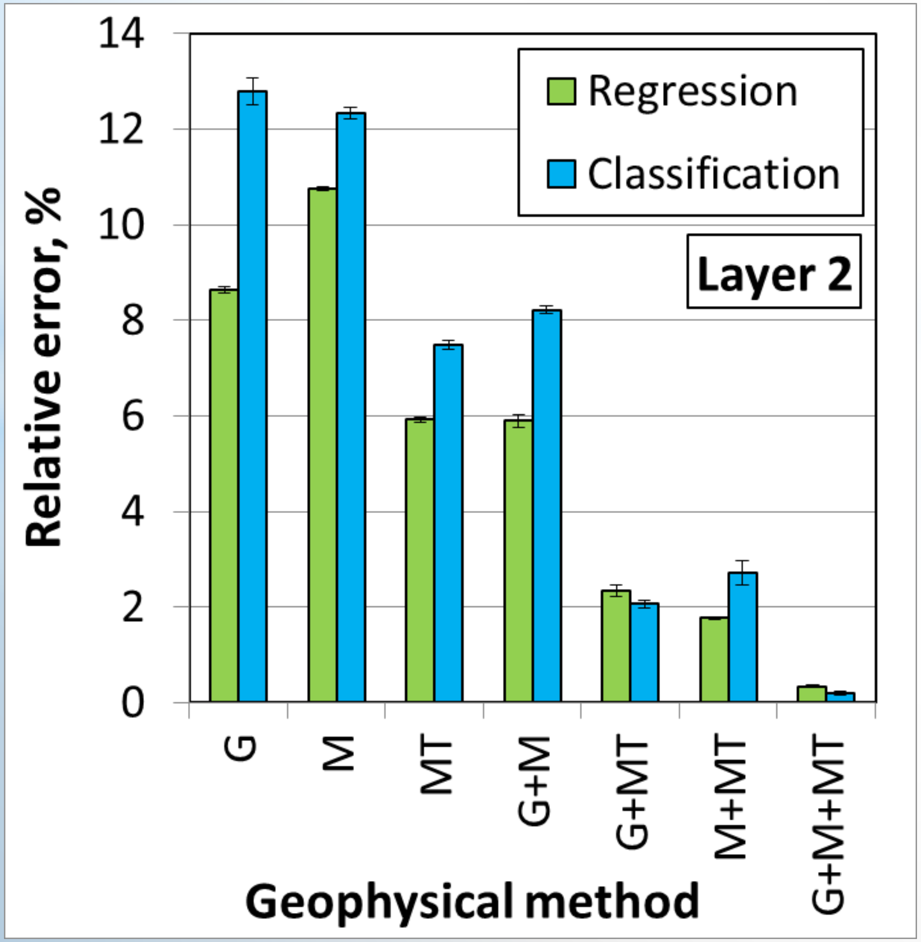


- Simultaneous use of data from any two geophysical methods reduces the error compared to the individual use of data from any of them.
- The best result is shown by the simultaneous use of data from all the three geophysical methods.
- The classification approach gives a better result than the regression approach for simultaneous use of gravimetry and magnetotellurics, and for simultaneous use of all the three geophysical methods.

Results

Layer 2

Dependence of the quality of the solution on input data for regression and classification approach

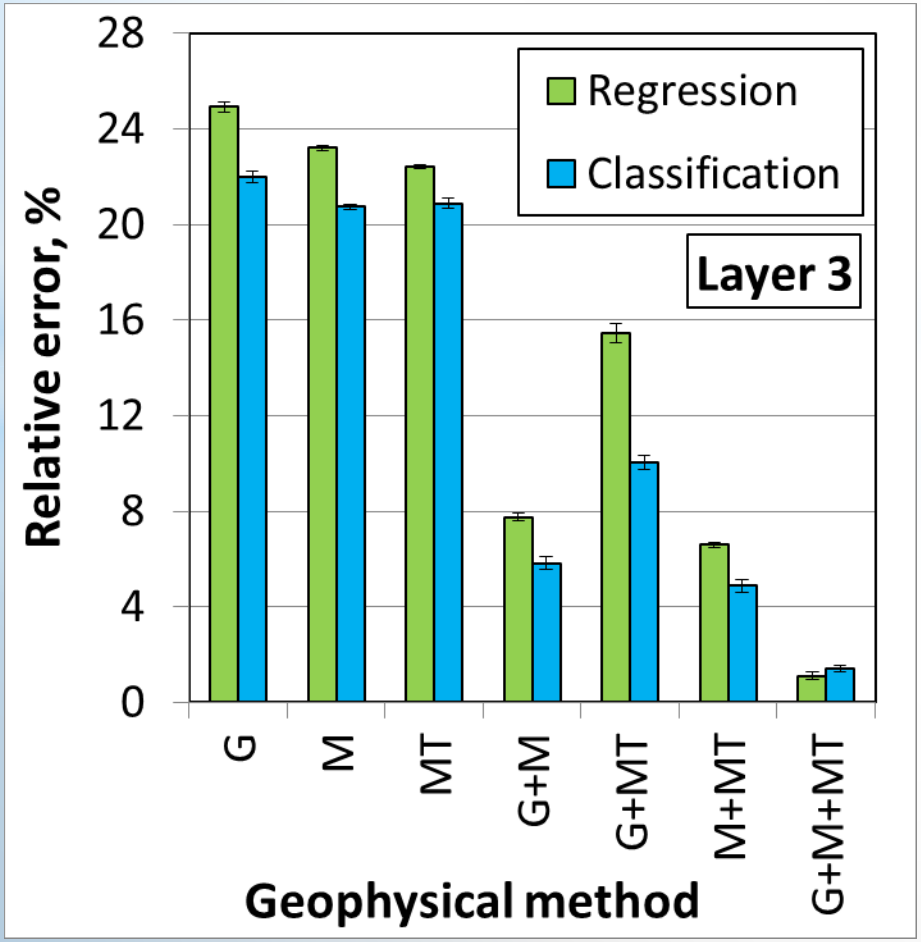


- Simultaneous use of data from any two geophysical methods reduces the error compared to the individual use of data from any of them.
- The best result is shown by the simultaneous use of data from all the three geophysical methods.
- The classification approach gives a better result than the regression approach for simultaneous use of gravimetry and magnetotellurics, and for simultaneous use of all the three geophysical methods.

Results

Layer 3

Dependence of the quality of the solution on input data for regression and classification approach



- Simultaneous use of data from any two geophysical methods reduces the error compared to the individual use of data from any of them.
- The best result is shown by the simultaneous use of data from all the three geophysical methods.
- The classification approach gives a better result than the regression approach in most cases.

Conclusions

Integration of geophysical methods.

- For all layers and for all approaches to solving the problem:
  - ✓ Simultaneous use of data from any two geophysical methods reduces the error compared to the individual use of data from any of them.
  - ✓ The best result is shown by the simultaneous use of data from all the three geophysical methods.

Classification approach and regression approach for solving the problem.

- In some cases the classification approach gives a better result than the regression approach:
  - ✓ For 1-st and 2-nd layer with simultaneous use of gravimetry and magnetotellurics and with simultaneous use of all the three geophysical methods.
  - ✓ For 3-rd layer in most cases.

Conclusions

Further work

Classification approach and regression approach for solving the problem.

- The study of the use of continuous,
  rather than discrete random values
  of the determined parameters
  to improve the quality of the solution of the regression approach.

Thank you

for your attention!