Geoinformation system with algorithmic shell as a new tool for Earth sciences

Alexander Beriozko,[1] Alexei Lebedev,[1] Anatoly Soloviev,[1] Roman Krasnoperov,[1] and Alena Rybkina[1]

Received 26 April 2011; accepted 29 April 2011; published 4 May 2011.

A new technology was elaborated, which combines geographical information system (GIS) technologies with GIS-oriented algorithmic methods of pattern recognition (PR). Numerous thematic data layers for geosciences, provided by Russian and international scientific institutions and data sources were imported into the GIS. Technology and software for the integration of PR methods within within the integrated geoinformation environment was developed in the form of the Centralized Catalogue of Geodata Processing Algorithms (CCGPA). A GIS visualization subsystem was created to provide interaction between the system and its users. It includes geodata layers visualization, map operations, geodata set management and representation of CCGPA-stored algorithm application results.
KEYWORDS: Geoinformation system; geodata; Earth sciences; pattern recognition.

Citation: Beriozko, Alexander, Alexei Lebedev, Anatoly Soloviev, Roman Krasnoperov, and Alena Rybkina (2011), Geoinformation system with algorithmic shell as a new tool for Earth sciences, *Russ. J. Earth. Sci., 12,* ES1001, doi:10.2205/2011ES000501.

Introduction

The fundamental problem of complex analysis of georeferenced data (geodata) is the combination of a geographic information system (GIS) and pattern recognition algorithmic methods for the creation of an integrated geoinformation environment for thematic data retrieval, storage, and processing. The present research is aimed at the development of new mathematical approaches and adaptation of existing methods of pattern recognition based on fuzzy logic and fuzzy mathematics approach, and their integration with data on geography, geology, geophysics, geoecology, and other Earth sciences in the comprehensive problem-oriented GIS including the intellectual environment for geodata analysis [*Berezko et al.,* 2008].

Research and Results

The main goals of the present research are:

- The development of an intellectual GIS for Earth sciences;
- Creation and development of PR methods and algorithms for recognition, classification, and cauterization and their integration with the GIS environment and geodata bases.

[1]Institution of the Russian Academy of Sciences Geophysical Centre of RAS (GC RAS), Moscow, Russia

ArcGIS 9.3.1 software package developed by Environmental Systems Research Institute (ESRI) was used for the creation of the basic GIS environment. The information content of the GIS was significantly extended and improved with data, provided by Russian and international scientific institutions and data sources. The main information sources were:

- RAS institutions;
- World data centers (WDC) within the World data system (WDS);
- United States Geological Survey (USGS);
- British Geological Survey (BGS);
- German Research Centre for Geosciences (GFZ);
- Institute of Earth Physics of Paris (IPGP);
- International Association of Geomagnetism and Aeronomy (IAGA);
- International Institute for Applied Systems Analysis (IIASA);
- Institute for Environment and Sustainability (IES), supported by the European Commission Joint Research Centre (EC JRC).

Currently the geodata base of GIS includes more than 200 thematic layers arranged in 19 data categories including Geodesy and Cartography, Geology, Geophysics, Glaciology,

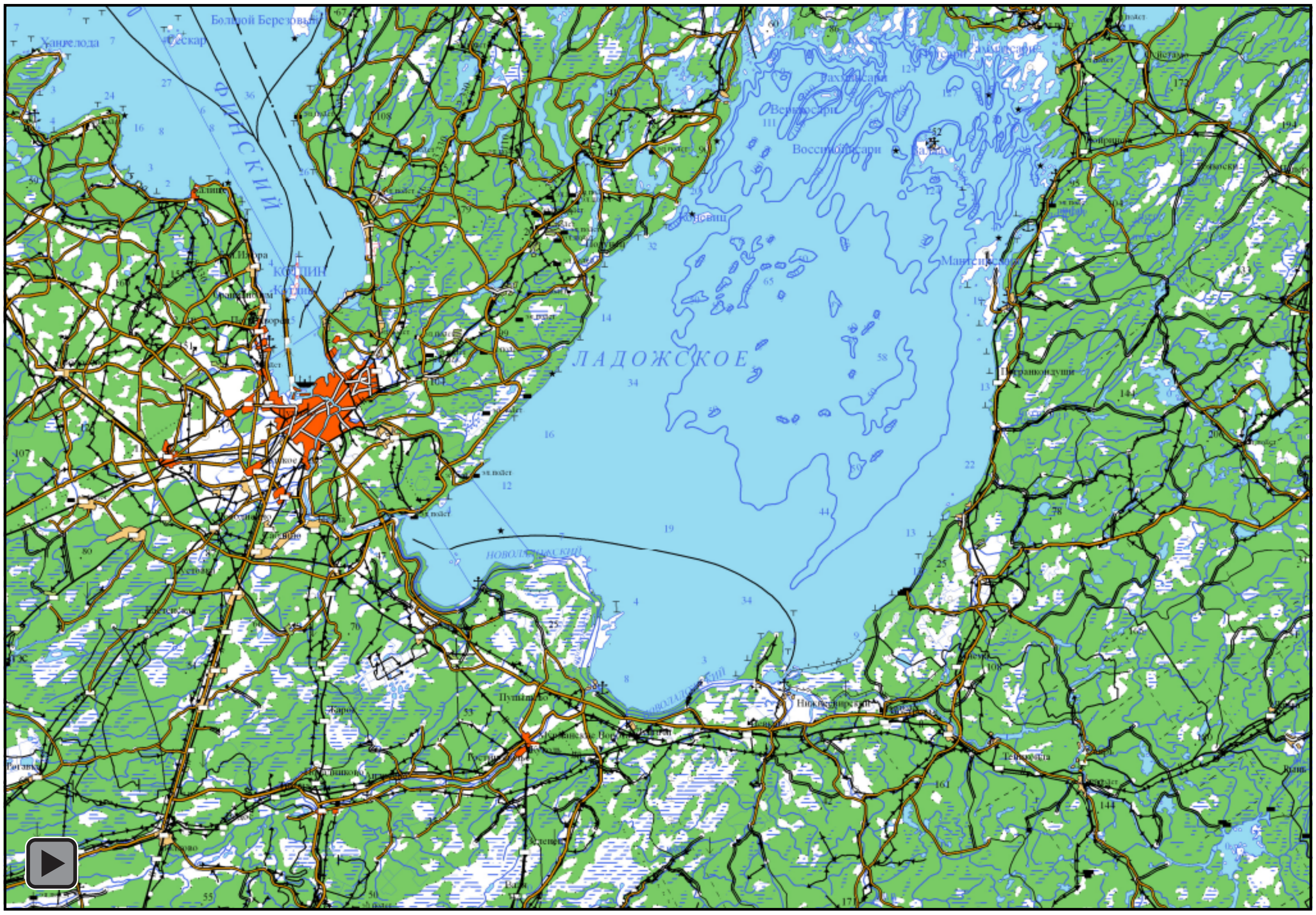

Figure 1. Visualization of thematic geodata layers in GIS: **a**) topographic map (fragment, $1:1,000,000$).

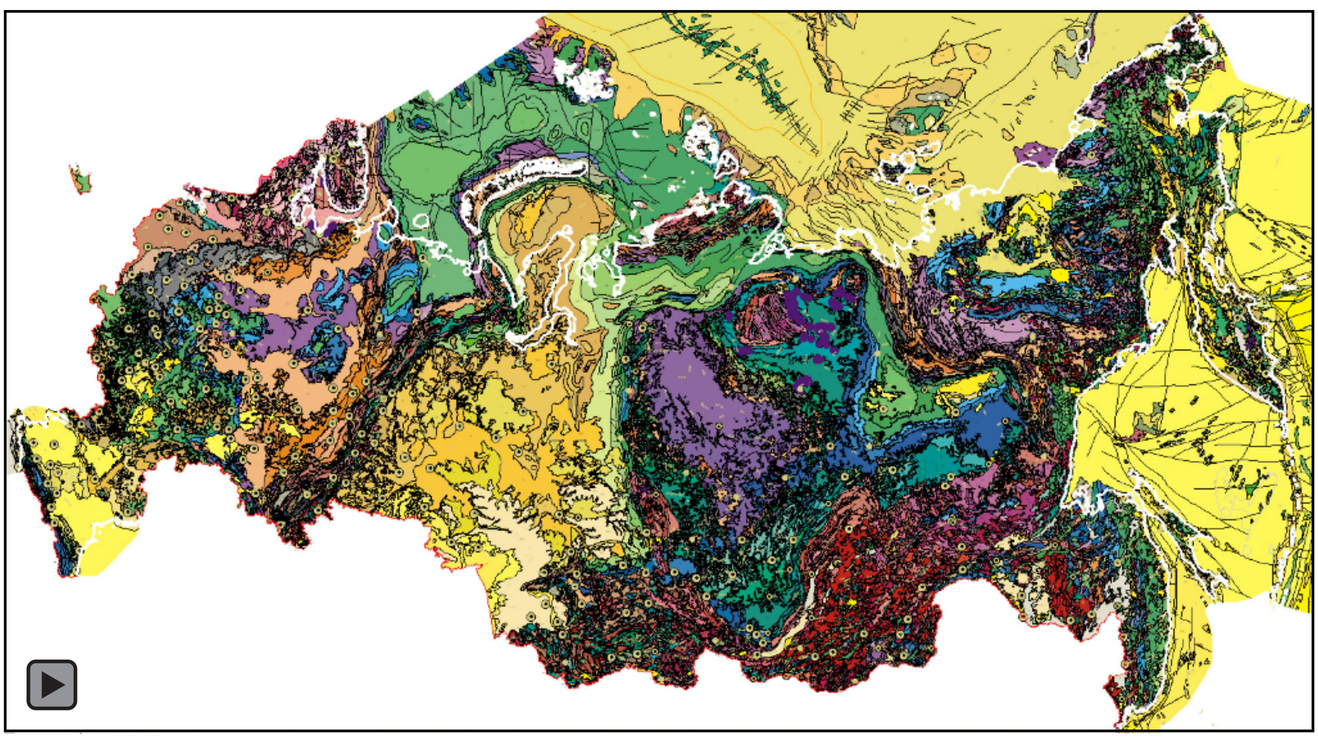

Figure 1. Visualization of thematic geodata layers in GIS: **b**) geological map ($1:5,000,000$). [Panels **a** and **b** are flash objects. Use Adobe Reader, version 8 or newer to see these maps in detail. Click on mouse right button and select *Zoom In*, then you can move the map inside the figure area while pressing mouse left button.]

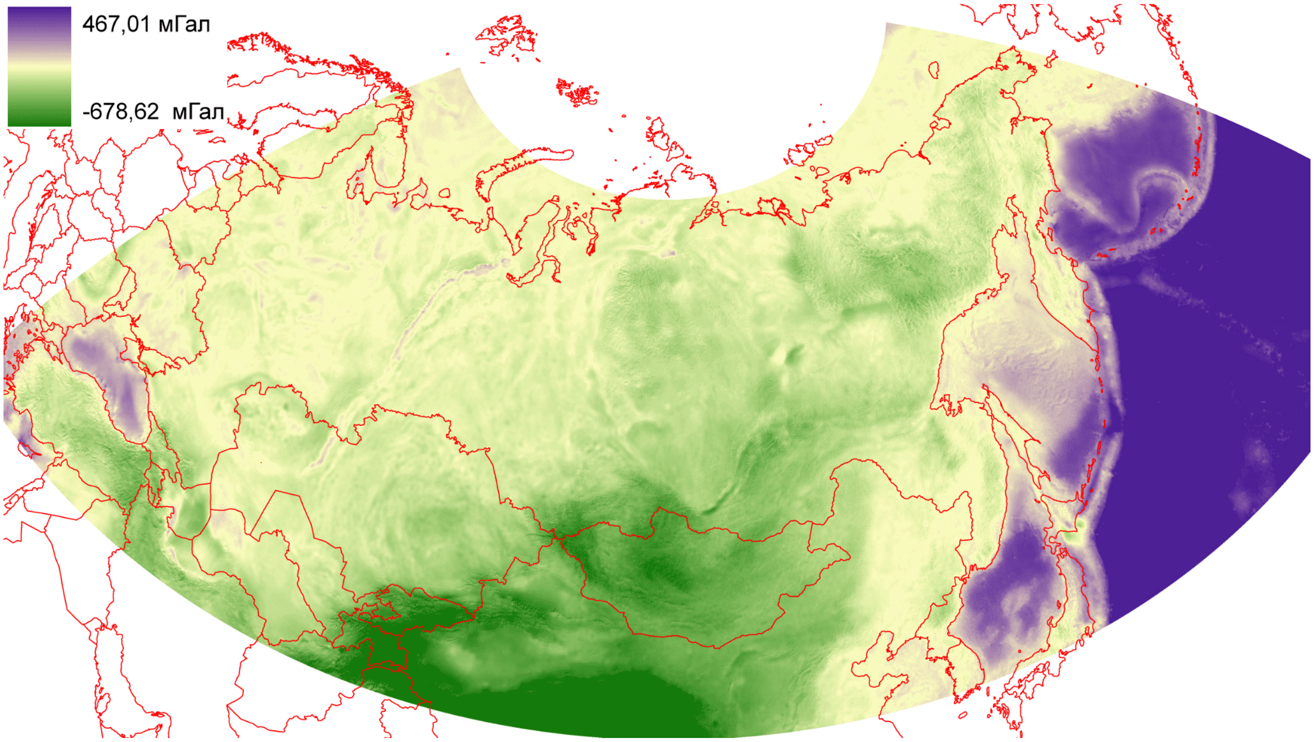

Figure 1. Visualization of thematic geodata layers in GIS: **c**) Bouguer gravity disturbances (5′).

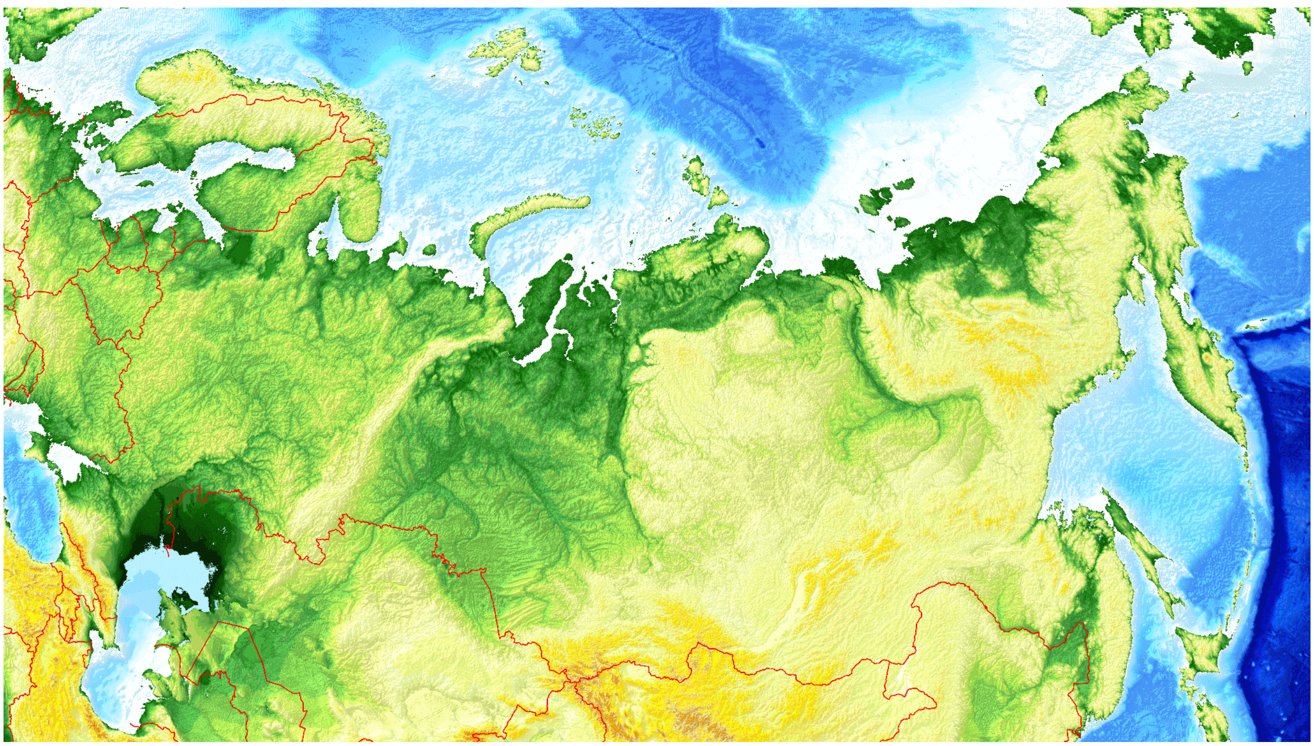

Figure 1. Visualization of thematic geodata layers in GIS: **d**) digital elevation model ETOPO1 (1′).

Hydrology, Remote Sensing, Mineral Deposits, Meteorology and Climatology, Pedology, Political Geography, Industry, Agriculture, Biogeography, Ecology. The cartographic base is formed with digital terrain maps at $1 : 1,000,000$ scale. All thematic geodata layers are stored as GIS-projects on database servers providing remote access via Internet. Examples of visualization of different thematic geodata layers in the GIS are shown in Figure 1.

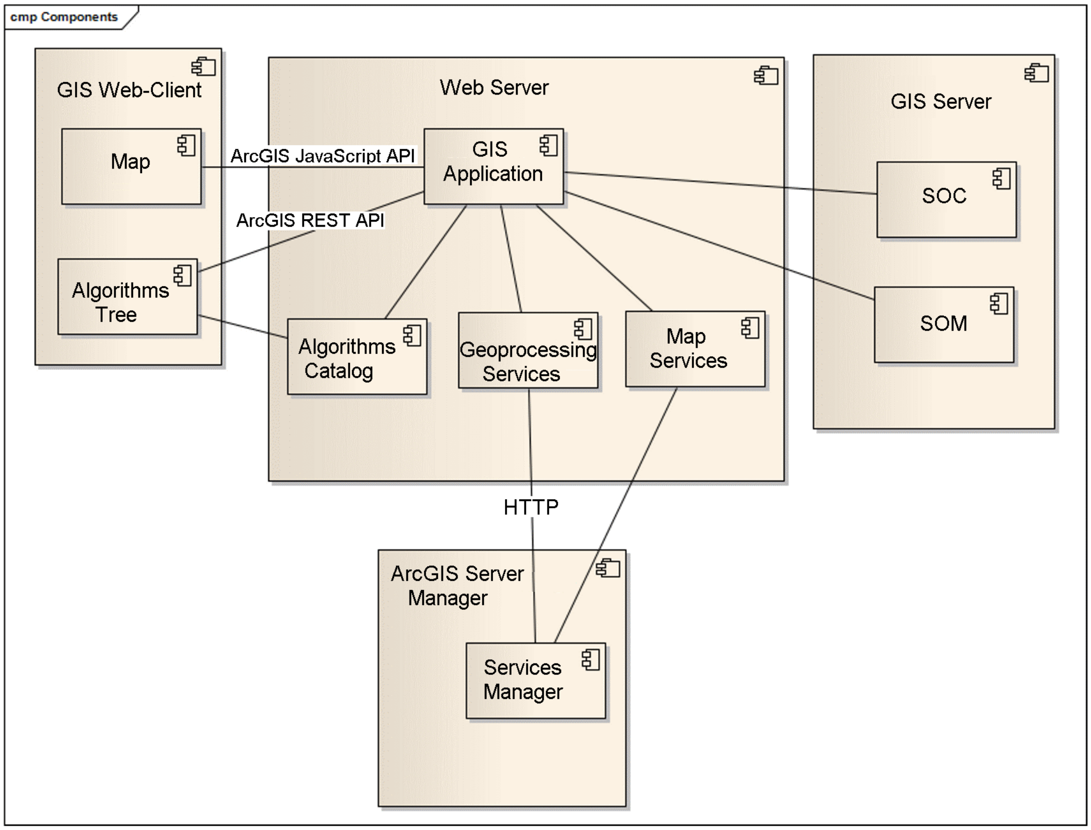

Figure 2. Architecture of CCGPA.

The developed theory and methods of pattern recognition must become not only an integral, but the main core of a modern GIS [*Beriozko et al.*, 2009a, 2009b]. At the present time GIS provide only limited opportunities for general analysis of geodata handled. At the same time, among the scientific community, dealing with the Earth sciences data, the requirement of more profound and comprehensive data analyzing and processing is constantly growing. Application and complex integration of PR methods and algorithms within the unified geoinformation environment, based on modeling of discrete analogs of the fundamental notions of mathematical analysis, is oriented at solving geophysical and geological-forecasting tasks and not only brings the analysis of initial geodata to the higher scientific and practical level, but provides scientifically based recommendations on interpreting the results obtained. The evaluation of natural environment and risk is a task of identification of a complex character of potentially dangerous situation on the basis of huge data volumes of environmental monitoring. The PR methods, developed by the GC RAS, are presently applied to seismology, volcanic activity monitoring, search and interpretation of anomalies in geophysical fields, detecting of signals on various types of times series records, to solving geodynamic problems, etc.

In order to integrate PR methods and algorithms within the GIS a new technology was elaborated. This advanced approach is implemented in the form of the Centralized Catalogue of Geodata Processing Algorithms (CCGPA) [*Lebedev and Beriozko*, 2009]:

- The CCGPA is a GIS-subsystem, which provides access to specific geodata processing algorithms centrally executed on a GIS-server;
- Users can choose available algorithms from the CCGPA or upload new algorithms to the CCGPA themselves;
- Each user-uploaded algorithm is to be verified and published by the system administrator and becomes available for the user community.

The advantages of this technology are:

- Possibility to upload new algorithms to the CCGPA;
- Possibility to execute a sequence of algorithms on several geodata sets;
- Minimal requirements to user workstation: all computations are performed by the server;
- Access to the CCGPA from any Internet node.

The CCGPA is implemented in three-level architecture as the GIS subsystem [*Lebedev and Berezko*, 2010] on the base of ArcGIS Server technology (Figure 2).

The CCGPA is managed by a GIS administrator using ArcGIS Server Manager application. The administrator publishes algorithms as Geoprocessing Services, which become available to GIS users.

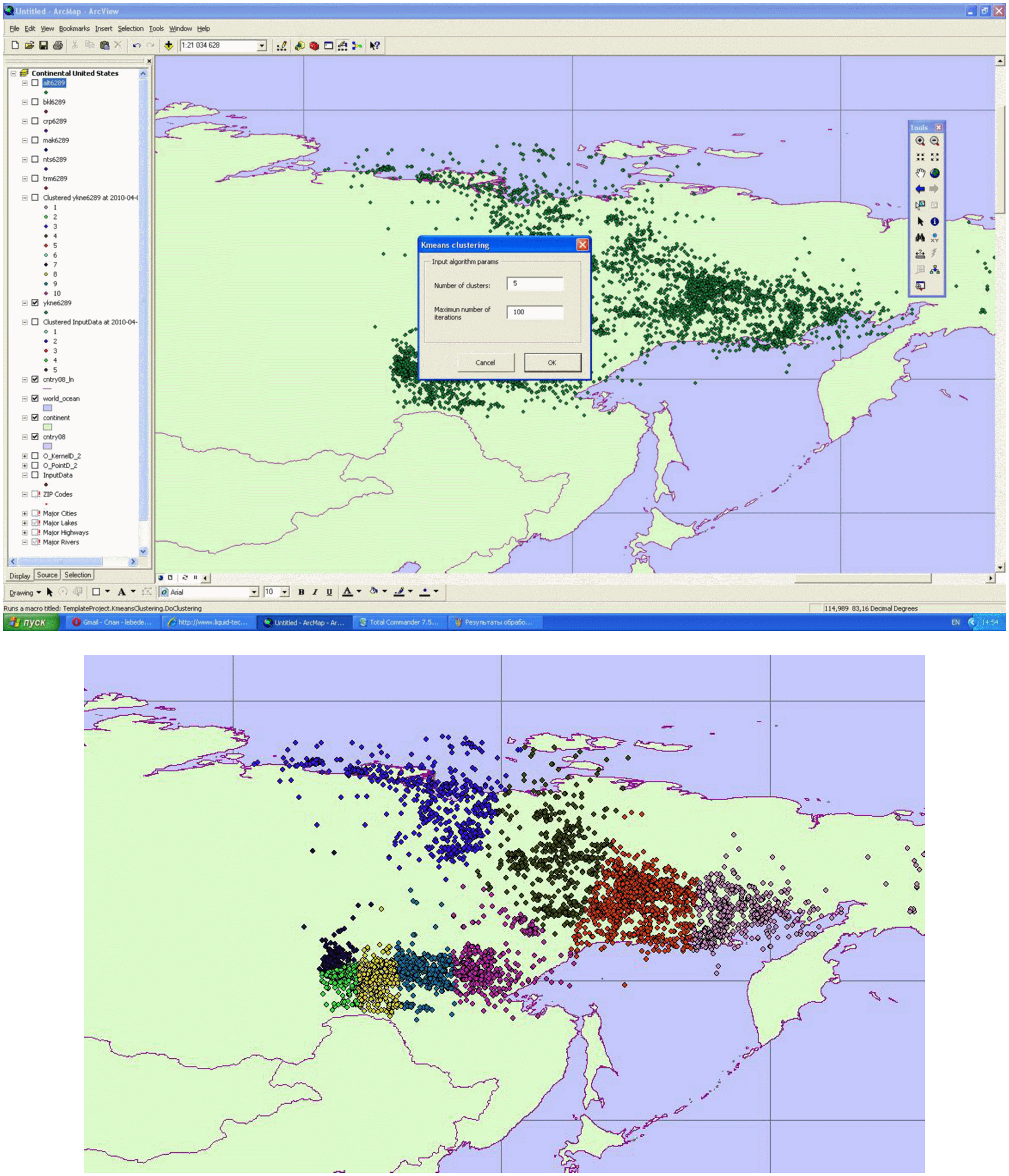

Figure 3. Application of CCPGA-stored K-means clustering algorithm: (top) input data (earthquake epicenters, Eastern Siberia, 1962–1989) and algorithm parameters window; (bottom) results (clusters are marked in colors).

The component Algorithms Catalog is responsible for uploading by users of new algorithms and their storage before inclusion in CCGPA. For uploading algorithms a client application opens a special form.

The client application is implemented via ArcGIS JavaScript API and includes a component for operations with algorithms (Algorithms Tree). When a user launches GIS Application via ESRI REST API, this component re-

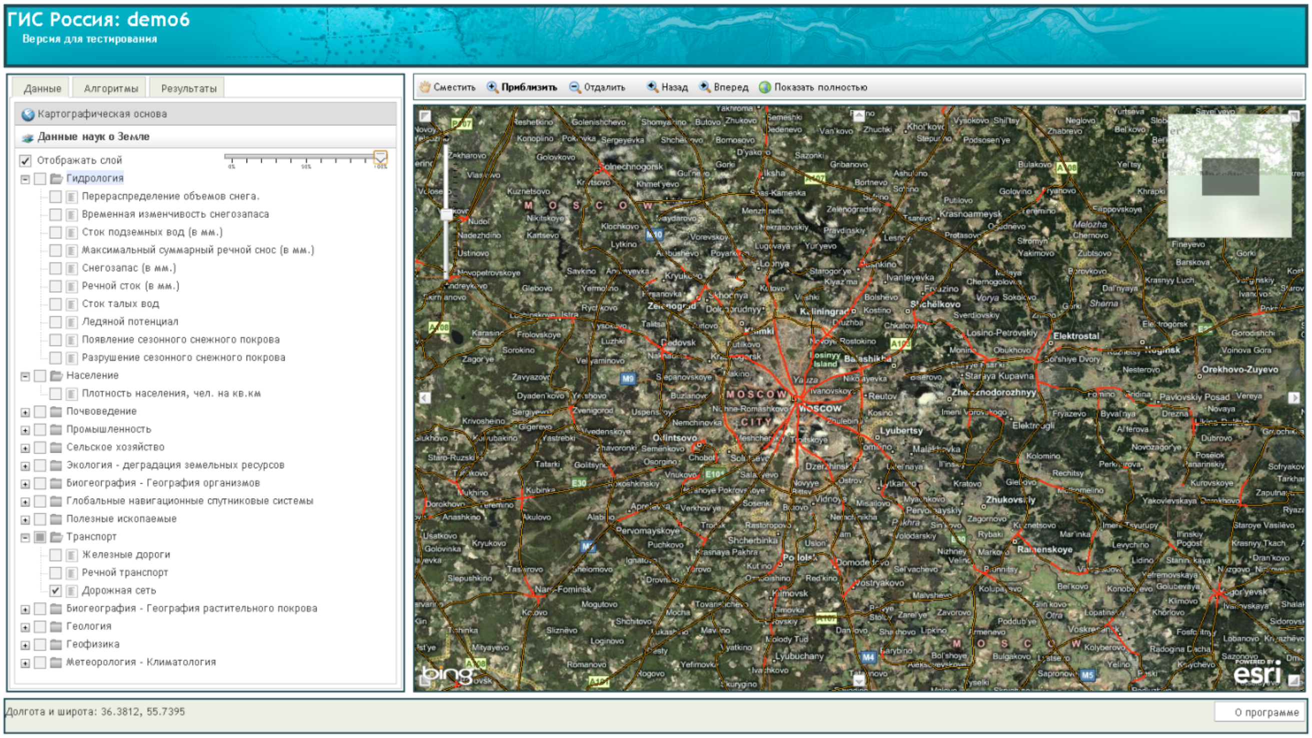

Figure 4. User interface of GIS web-application.

ceives a list of Geoprocessing Services with the information about them and fills a user's Algorithms Tree.

ArcGIS JavaScript API allows executing the algorithms with the user-specified parameters and returns the results to the client application for visualization.

Geodata processing is performed by the components Server Object Container (SOC) and Server Object Manager (SOM) of the GIS-server. On the base of the CCGPA technology an algorithmic shell was constructed in the frame of the GIS, ensuring the intellectual capacity of GIS environment. The shell comprises a set of algorithms of universal character bound by the common formal framework. It contains fundamental solutions of classic tasks of data analysis: clustering and search of linear structures in stationary multidimensional arrays, construction of smooth skeletons, search of signals and morphological analysis in time series. Figure 3 demonstrates the results of application of the K-means clustering algorithm stored in the CCGPA to geophysical data sets [*Berezko et al.*, 2010].

To provide interaction between the GIS and its users a GIS visualization subsystem was created. It performs geodata layers visualization, map operations, geodata set management, execution of CCGPA-stored algorithms and representation of application results.

The basic part of the GIS visualization subsystem includes a GIS application, implemented as a web-application. This GIS web-application can be launched at a user's Internet enabled workplace without installation of any additional software. It supports an interface between the GIS, its geodata base, the CCGPA, and remote users. The user interface of the GIS web-application is represented in a browser window by the following elements (Figure 4):

1. Map window, with minimap subwindow to ease navigation;
2. Panel of Instruments for map operations;
3. Data, Algorithms, Results tabs;
4. Title and Information line.

The Data tab contains a tree of thematic geodata layers available for visualization, including conventional signs and legend. A user can switch the layers on and off by placing or removing a tick in the appropriate box.

Geodata layers chosen by a user in the Data are visualized in the Map window. The Algorithms tab contains a tree of algorithms available in the CCGPA. The Results tab contains a list of layers created as results of application of chosen algorithms to chosen thematic geodata layers.

It is possible to choose either digital terrain maps at $1:1,000,000$ scale stored in the GIS geodata base, or web mapping service applications (Google Maps, Bing Maps, Imagery, etc.) as a cartographic base (Figure 5).

A user can apply to thematic geodata layers visualized in the Map window all instruments and operations contained in the Panel of Instruments (Figure 6).

The main disadvantage of cartographic editions published in the form of paper maps is their lack of visual representation. Modern GIS representation technologies

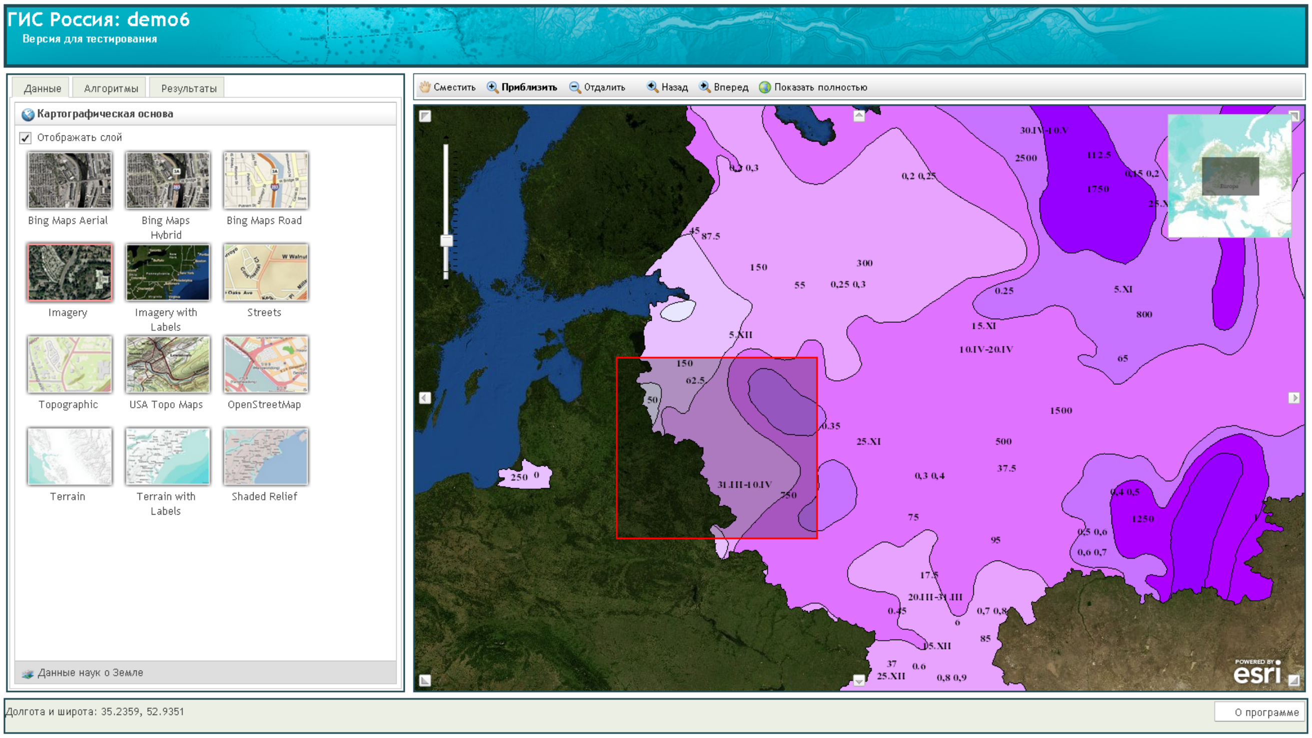

Figure 5. Choosing a cartographic base.

are highly efficient for the visual analysis of cartographic documents. GIS-technologies are common in demonstration systems based on up-to-date visualization means. One of the most advanced approaches for visualization and representation of georeferenced data is implementation of digital projection systems with a sphere shaped screen. This projection technology is a revolutionary instrument which brings the geoscience data visualization up to a new level. The demonstration complex consists of a spherical display, a tabletop digital projector in a metal chassis with a

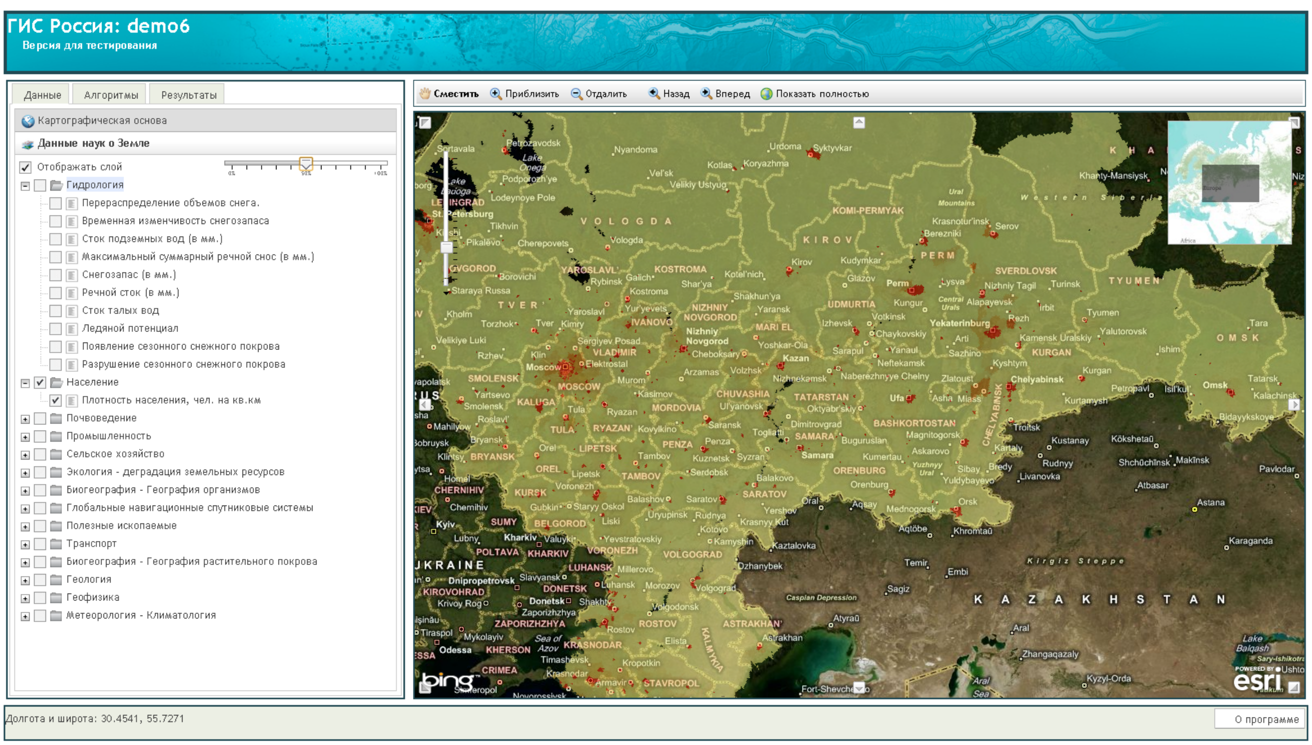

Figure 6. Overlaying several geodata sets and changing transparency.

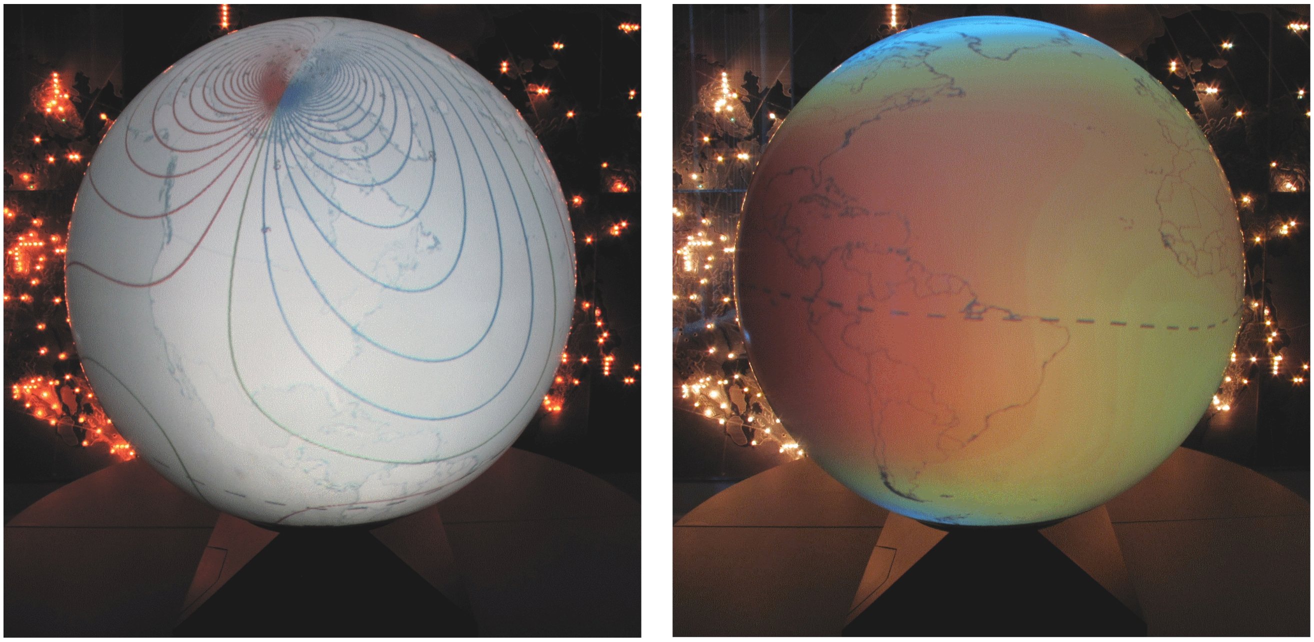

Figure 7. General view of the digital projection systems with a sphere shaped screen. Visualization of the maps of Earth Main Magnetic Field.

complicated catadioptric optical system (lens block and mirror), and a PC workstation with special software. The general view of the device is shown in Figure 7.

Conclusions

Application and complex integration of methods and algorithms of pattern recognition (PR) within the unified geoinformation environment are the key features of the discussed intellectual GIS. GIS-technologies, based on generalization and complex processing of geodata, in combination with appropriately adjusted PR algorithms stored in the CCPGA, provide efficient automation of Earth sciences data analysis and forecasting for fundamental and applied scientific research.

The interaction between the GIS and remote users is provided by the GIS visualization subsystem. It includes the GIS application, implemented as the web-application, which can be launched at a user's Internet enabled workplace without installation of any additional software. The GIS visualization subsystem supports an interface between the GIS, its geodatabase, the CCGPA, and remote users. Thus the problem of geodatabase publication and interface between data, CCGP and remote users was solved.

One of the most advanced approaches for visualization and representation of georeferenced data is implementation of digital projection systems with a sphere shaped screen. This demonstration complex was implemented in GC RAS; and currently it is the only device of this type in Russia.

References

Berezko, A. E., A. A. Soloviev, A. D. Gvishiani, E. A. Jalkovski, R. I. Krasnoperov, S. A. Smagin, E. S. Bolotsky (2008), Intellectualnaya geoinformacionnaya sistema "Dannye nauk o Zemle po territorii Rossii", *Ingenernaya Ekologiya* (in Russian), 5, 32–40.

Beriozko, A., A. Soloviev, R. Krasnoperov, A. Rybkina, E. Kedrov, E. Bolotsky (2009a), Intellectual analytical geoinformation system "Earth sciences data for the territory of Russia", *Materials of the International Conference "Electronic Geophysical Year: State of the Art and Results", 2009, edited by V. Nechitailenko,* 117, GC RAS, Moscow, doi:10.2205/2009-REGYconf.

Berezko, A., A. Soloviev, R. Krasnoperov, A. Rybkina (2009b), Intellectual analytical geoinformation system "Earth Science Data for the Territory of Russia", *Proceedings of the 7th International Scientific and Practical Conference "Environment. Technology. Resources", 2009*, Rezeknes Augstskola, Rezekne, RA Izdevnieciba, 1, 215–221,

Berezko, A. E., A. D. Gvishiani, A. A. Soloviev, R. I. Krasnoperov, A. I. Rybkina, A. Yu. Lebedev (2010), Intellectualnaya GIS "Dannye nauk o Zemle po territorii Rossii", *Materialy XV mezhdunarodnoi nauchno-prakticheskoi konferencii "Problemy zaschity naseleniya i territorii ot chrezvychainyh situacii", 2010* (in Russian), FGU VNII GOChS (FC), 512, 210–218.

Lebedev, A. Yu., A. E. Beriozko (2009), Development of centralized catalog of geophysical data processing algorithms, *Russ. J. Earth Sci., 11*, RE2002, doi:10.2205/2009ES000399.

Lebedev, A. Yu., A. E. Berezko (2010), Sozdanie centralizovannogo cataloga algoritmov obrabotki geodahhyh, *Geoinformatika* (in Russian), 2, 67–70.

A. Beriozko, R. Krasnoperov, A. Lebedev, A. Rybkina and A. Soloviev, Institution of the Russian Academy of Sciences Geophysical Centre of RAS (GC RAS), 3 Molodezhnaya st., 119296 Moscow, Russia. (a.berezko@gcras.ru)