Distributed network analytical GIS

V. G. Gitis,[1] A. P. Weinstock,[1] and A. N. Shogin[2]

Received 1 December 2007; accepted 15 December 2007; published 29 February 2008.

[1] The functionality of distributed analytical network systems GIS (Geographic Information Systems) GeoProcessor 2.0, GeoTime II and COMPASS V and elements of geoinformation technology of their use in scientific and applied research are analyzed. At the time of the analysis these systems can download data and plug-ins, distributed either in telecommunication systems or on a user's PC. The article is illustrated by examples of GIS applications for analysis of spatial and spatio-temporal data. The GIS discussed here represent a significant part of the analytical resources of geoinformation environment, elaborated in the framework of the Russian Academy of Sciences (RAS) Presidium Programme "Electronic Earth". *INDEX TERMS:* 0525 Computational Geophysics: Data management; 0530 Computational Geophysics: Data presentation and visualization; 0545 Computational Geophysics: Modeling; *KEYWORDS:* distributed analytical network systems, GIS technologies, "Electronic Earth" programme.

Citation: Gitis, V. G., A. P. Weinstock, and A. N. Shogin (2008), Distributed network analytical GIS, *Russ. J. Earth. Sci., 10,* ES2003, doi:10.2205/2007ES000236.

Introduction

[2] Analytical geoinformation technologies and GIS are developed for research of spatial and spatio-temporal data. Analytical GIS first appeared in the 1980s. Development of the Internet and the exponential growth of the volume of digital geographic information (GI) promoted the creation of network GIS. At the present time an increase in GI volumes entails their structuring and developing of thematic data storages, distributed over the Internet and in local networks. In this connection recently the research has focused on development of geoinformation environment of users, comprising distributed information as well as analytical and system resources. In particular, in 2006, ESRI company, a leader in GIS, has announced the development of distributed GeoWeb environment (ArcReview no. 1 (36) 2006). This direction of research in Russia was initiated in 2004 by the RAS Presidium programme "Electronic Earth: scientific data resources and information-communication technologies". In the framework of this programme in 2006 for the first time the basic version of geoinformation network environment "Electronic Earth" was elaborated [*Arsky et al.,* 2007a, 2007b]. Some of the most important analytical resources of this version are distributed network GIS: GeoProcessor, version 2.0 (http://www.geo.iitp.ru/GeoProcessor-2/new/index.htm), COMPASS, version V (http://www.geo.iitp.ru/) and GeoTime II, beta-version (http://www.geo.iitp.ru/geotime/), developed in the Institute of Information Transmission Problems (IITP RAS) and supported by the Russian Foundation for Basic Research (RFBR). These GIS are realized as Java applets. The systems support complex analysis of GI, ensure high interactivity of analysis, facilitate integration of data and plug-ins, distributed on network servers and on a user's PC. The connection of plug-ins to GIS kernels allows a user to carry out domain orientation of the system and local data networking ensures its confidentiality.

[1] A. Kharkevich Institute of Information Transmission Problems RAS, Moscow, Russia

[2] All-Russia Institute of Scientific and Technical Information RAS, Moscow, Russia

ISSN: 1681–1208 (online)

Methods of Geographic Information Research

[3] Functionality of analytical GIS is mainly focused on solving two types of problems: (1) exploration of multidisciplinary geographic information (GI) and evaluation of links between its components and (2) search of multivariate dependencies in GI, prediction, detection and identification of target stationary and dynamic parameters of examined environment. For solving these tasks three methods are widely used: Visual research, Analytical transformations and Plausible inference [*Gitis and Ermakov,* 2004].

[4] The goals of *visual research* are the determination of the spatial image of parameters and objects of an examined territory and cartographic measurements. For example, network GIS of IITP RAS allows dynamic operating of

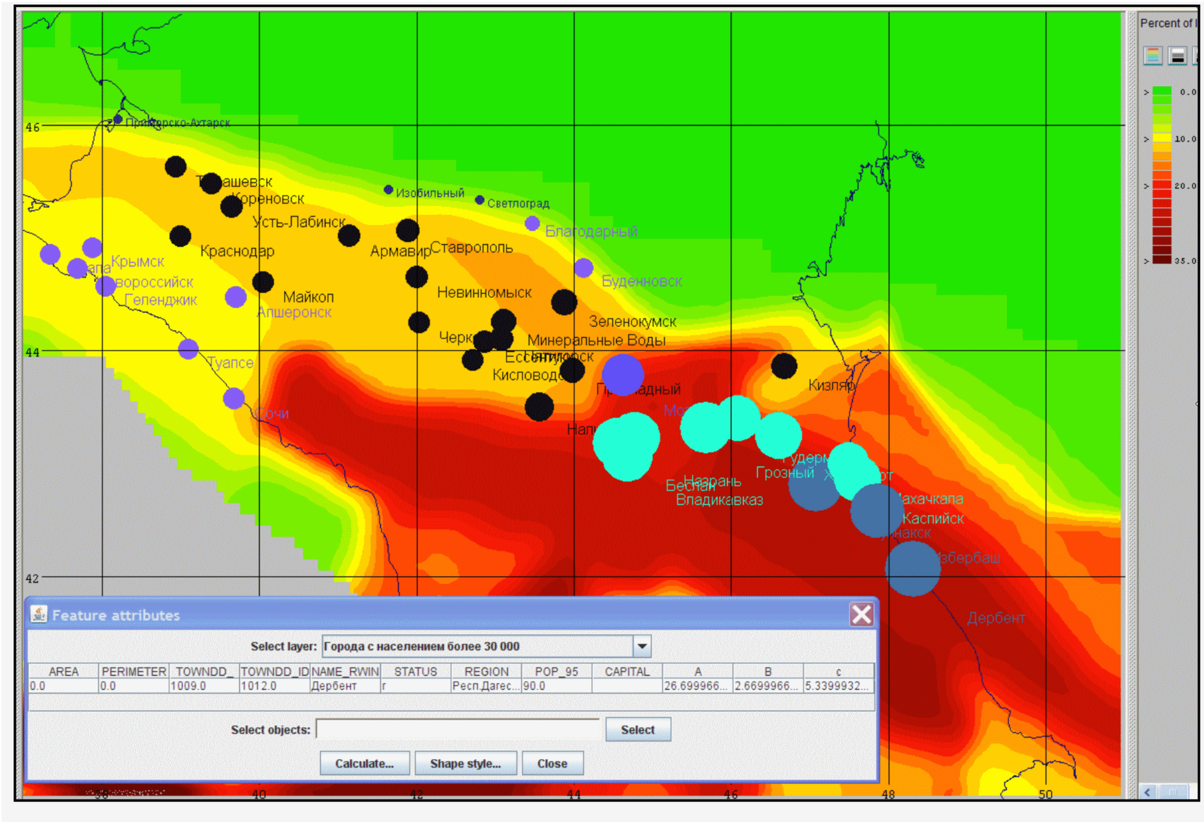

Figure 1. Destructions of area 7KP from maximal earthquakes in %.

cartographic layers, filling-in maps, size of pictograms, thickness of lines, parameters of 3D screen animation, acquisition of multilayered sections, measuring of grid-fields values and vector data attributes, dynamically display on a map groups of objects in a moving time interval (for example, earthquake epicenters), dynamically highlight areas on a map, identical by their parameters to standards, chosen by a user etc.

[5] *Analytical transformations* help to acquire new thematic and spatio-temporal GI parameters by predetermined operators. The most important are the following types of transformations: (1) *Grid layers* ⇒ *Grid layer,* (2) *Grid layers* and *Vector layers* ⇒ *Attributes of a vector layer,* (3) *Vector layer* ⇒ *Grid layer,* (4) *Vector layers* ⇒ *Attributes of a vector layer.*

[6] *Methods of plausible inference* allow determination of previously unknown operators of analytical transformation. Plausible inference tasks include: evaluation of dependencies and relations between an examined environment's parameters and its substances; prediction of target parameters of an environment; detection of target objects; prediction of spatio-temporal processes. Solving these tasks requires using of methods of multidimensional statistic analysis, pattern recognition, imitational modeling and artificial intelligence.

[7] All analytical operations are accompanied by interactive visualization. Close interaction of analytical and visual research methods lays the foundation of spatial and spatio-temporal cognitive modeling and significantly simplifies understanding of the researched material, accordingly increasing the efficiency of tasks solutions.

Examples of Tasks Solutions Technology

Types of Analytical Tasks

[8] The complexity of solving tasks of geographic information analysis depends essentially on the completeness of available data [*Gitis and Ermakov,* 2004].

[9] Tasks with complete information reveal the qualitative characteristics of GI by visualization, determine new GI parameters using previously known transformations and evaluate standard statistical GI parameters.

[10] Tasks with incomplete information emerge at solving problems of prediction which require a more profound investigation of geographic substances, their parameters and relations between them. Solution of such tasks is related

to a complex GI analysis. Such analysis is necessitated by interaction of researched processes, impossibility of direct measurements of their key parameters, lack of the volume of observations and impact of noise on the measurements' results.

Earthquakes Damage Evaluation (GeoProcessor 2.0)

[11] Let us examine the example of possible damage evaluation of a strong earthquake for the cities of the North Caucasus with a population of more than 100,000 people (see http://www.geo.iitp.ru/GeoProcessor-2/new/Caucasus2.htm). The data on peak acceleration was used [*Giardini et al.,* 2003] (this resource was obtained through the Central portal of the geographic information environment "Electronic Earth" http://eearth.viniti.ru).

[12] With the help of the transformations *Grid layers* ⇒ *Grid layer* for a grid layer of peak accelerations A the field of maximal magnitude of earthquakes $I = (\lg A - 0.014)/0.3$ was obtained [*Trifunac and Brady,* 1975]. For field I grid layer V of a proportion of destruction of buildings 7KP (this type of buildings was selected only for the illustration of the method): $V = 0$ at $I < 7$, $V = 3.5\%$ at $I = 7$, $V = 11.9\%$ at $I = 8$, $V = 37\%$ at $I = 9$. Then the transformation "Grid layers and Vector layers Attributes of a vector layer" was applied. With its help the proportions of destructions located at a distance of 5 kilometers in the vicinity of the cities were determined. Assuming that the area with buildings of 7KP type is homogenous for the selected size of the zone, the result can be accepted as an evaluation of damage of a maximal earthquake. Figure 1 shows the grid layer of destructions of the buildings of 7KP type in percentage terms, the size of circles showing the damage values for the cities. Below the destruction proportion value for the city of Derbent is shown, equal to 27%.

Seismic Danger Evaluation (GeoProcessor 2.0)

[13] Let us examine the example of detecting the zones of possible earthquake sources (PES) with magnitudes $M > 6.5$ for the Caucasus using the resource http://www.geo.iitp.ru/GeoProcessor-2/new/ARMEAST2-e.htm developed according to the data of *Gitis et al.,* [1993].

[14] According to [*Gitis and Ermakov,* 2004; *Gitis et al.,* 1993], it was assumed that the central zones of strongest earthquakes are timed to intersection of heterogeneous zones of the Earth crust with the zones of thrust and shear faults, active in the Cainozoic era.

[15] First with the help of analytical transformations and visual investigation an exploratory analysis of the benchmark and transformed data was made. To illustrate the method the most simple solution was chosen, using only a field of velocity gradient module of vertical motions in the post-Sarmatian time (characteristic X_1) and thrust faults, active in Cainosoe. By transformation *Vector layer* ⇒ *Grid layer* characteristic X_2 was obtained – the field of the distance to thrust faults. Further the method of inductive logical conclusion was applied [*Gitis and Ermakov,* 2004]. The obtained rule appears to be: IF velocity gradient of vertical tectonic motions in the post-Sarmatian time (X_1) exceeds 10 conventional units (cu) OR (X_1) > 6.4 cu AND distance to thrust faults ($X2$) < 6.75 km, THAN centers with $M > 6.0$ are possible. The PES zones, obtained according to this rule and epicenters with magnitudes $M > 6.0$ are shown in Figure 2. Prediction of sea zones and the southern and south-eastern zones hasn't been made due to the lack of geological and geophysical data.

Prediction of Oil and Gas Fields (GeoProcessor 2.0)

[16] Let us examine an example of selective regional prediction of oil and gas field in Western Siberia using the resource http://www.geo.iitp.ru/GeoProcessor-2/new/WestSiberia2.htm developed according to the data of *Gitis et al.,* [1994a].

[17] According to [*Gitis and Ermakov,* 2004; *Gitis et al.,* 1994a] it was assumed that a phase state of carbohydrates is determined by the history of tectonic development. Gas fields are usually characterized by deteriorated quality of primary organic matter and strong sedimentation, revealed in higher velocity of longitudinal seismic waves on the foundation surface. Oil deposits are characterized by a high quality of organic matter, forming a relatively small sedimentary deposit. Provinces with a thin sedimentary cover have low prospects related to oil and gas deposits. For the problem's solution the same parameters are chosen as in [*Gitis and Ermakov,* 2004; *Gitis et al.,* 1994a]. The transformations *Grid layers* ⇒ *Grid layer* were used to obtain them. Further a method of recognition was applied, according to the rule of the closest neighbor, affiliating a point to one or another class by similarity of the point parameters and reference objects of classes. The results of prediction are shown in Figure 3. Circles mark the known gas deposits, triangles – oil deposits, squares – unproductive areas. For the prediction of gas fields the following parameters are used: the depth of occurrence of the dogger's top and a half-sum of longitudinal seismic waves velocities on the surface of crystalline and folded foundations. For predicting oil deposits the following parameters are used: the depth of the top of Middle Jurassic sediments, the depth of the top of Upper Cretaceous sediments, and the thickness of upper layer of consolidated crust.

Analysis of Precursors According to Earthquakes Catalogue (GeoTime II)

[18] Let us examine the example of detecting precursors of the Susamyrsky earthquake: 19.08.1992, energy class $K = 17$, coordinates $\lambda = 73.63°$ longitude east and $\phi = 42.06°$ latitude north (see http://www.geo.iitp.ru/geotime/asia.html). The Central Asian earthquakes catalogue was used, cleared from aftershocks. In the catalogue 16329 events for 1980-2001 are presented at K from 7 to 17. The cata-

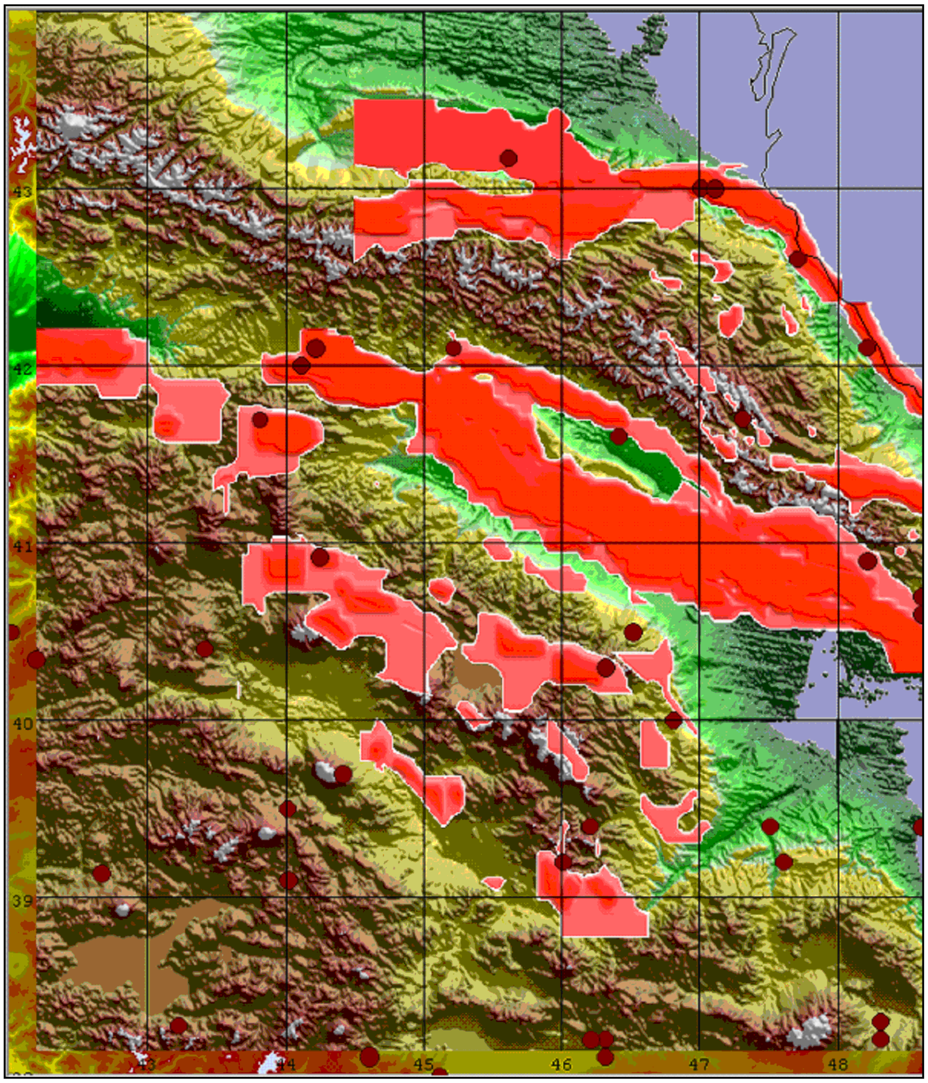

Figure 2. PES zones with $M > 6.0$ are marked with red color.

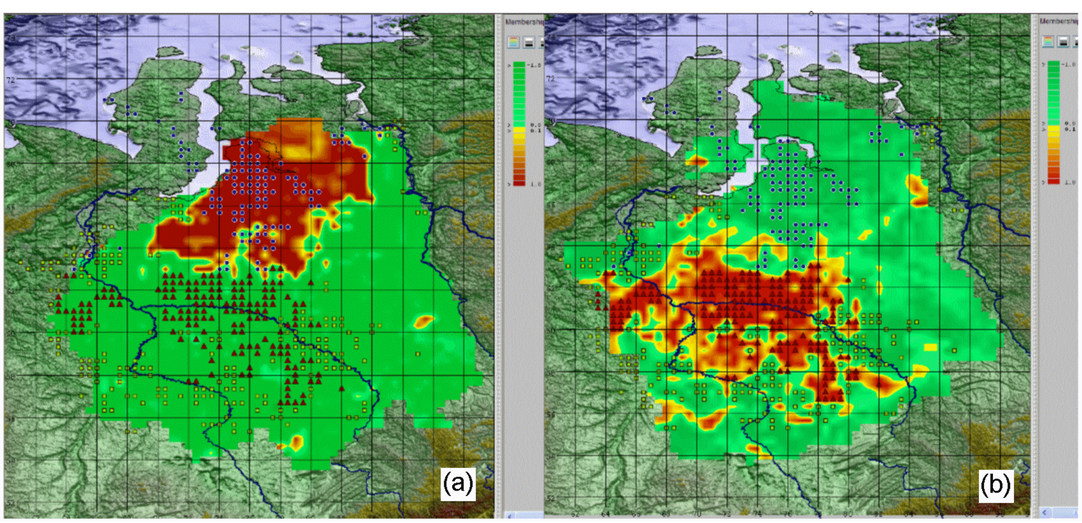

Figure 3. Predicted zones of gas (a) and oil (b) deposits (marked with red color).

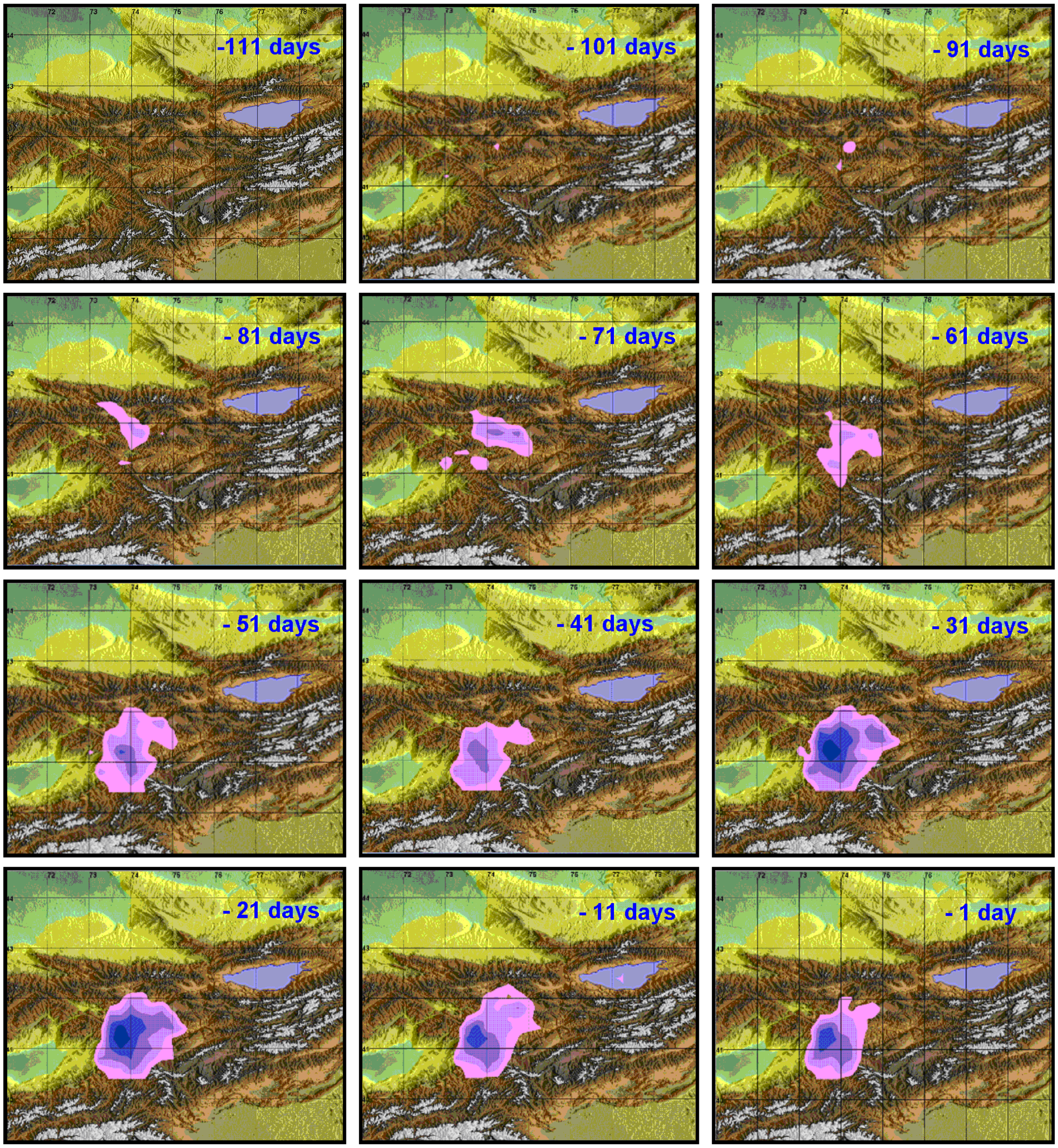

Figure 4. Spatio-temportal anomaly of earthquakes epicenters density, preceding the Susamyrsky earthquake.

logue's preliminary processing was implemented in IPE RAS by G. Sobolev.

[19] At the time of the analysis the method of detection of precursors was applied, elaborated in [*Gitis and Ermakov,* 2004; *Gitis et al.*, 1994b]. First by catalogue grid layer 3D of earthquakes centers density in a running cylindrical window with radius of 100 km and time interval of 10 days was obtained (transformation *Vector layer* ⇒ *Grid layer*). Further by transformation *Grid layers* ⇒ *Grid layer* grid layer 3D of anomalies was obtained. The anomalies detection algorithm is based on the method of statistical hypotheses verification. For each node of a spatial grid the statistics are prepared, equal to norm difference of averages $(m_2 - m_1)$ in two running windows: m_1 – the average in the first window 1440

days long for the estimation of the background value of density of earthquakes epicenters and m_2 – the average in the second 30-days window for current value estimation. Figure 4 shows the evolution of density anomaly of epicenters with high negative statistical values in the interval from –5.5 to –3.5. In the picture one can see 12 cuts of 3D anomaly from 111 to 1 day before the earthquake. The epicenter of Sysamyrsky earthquake is marked by a star. Negative values of the anomaly give evidence that average density of earthquakes is declining. It points at a lull, in many cases preceding a strong earthquake. The anomaly with a value less than –3.5 appears in the vicinity of the forthcoming earthquake 101 day before the earthquake. 61 day before the earthquake the anomaly transforms into a simply connected domain. The density of anomaly and numerical value increase monotonously and reach their maximum 31 day before the earthquake. Then the anomaly subsides.

Conclusion

[20] We have briefly examined some elements of the technology of network analytical GIS GeoProcessor 2.0, GeoTime II and COMPASS V. At the time of the analysis these systems can download for processing data and plug-ins, distributed either in telecommunication systems or on a user's PC. The described examples of spatial and spatio-temporal analysis provide the evidence of the efficiency of application of the above mentioned systems for solving relatively complicated research tasks in the Earth sciences.

[21] At the present time the number of new sources of digital GI is increasing, which has both the spatial and temporal components. This is related in the first place to the development of means of monitoring of natural and socio-economic processes. One of the most important initiatives in this field is the Project of setting up of the Global Earth Observation System of Systems (GEOSS), designed for forthcoming decades, which is being developed with the aim of better understanding and solving global problems of the environment and economics. These tendencies require modern geoinformation technologies of storing, transferring and processing of huge masses of data, representing researched processes in time and space. In the basic version of the geoinformation distributed environment "Electronic Earth" first steps towards this direction have been made.

[22] **Acknowledgment.** The work was fulfilled with the support of the RFBR grants 00-07-90100, 03-07-90114-v, 06-07-89139 and the RAS Presidium Programme "Development of the Fundamental Base of the Scientific Distributed Data-processing Environment on the basis of GRID Technologies", direction: "Electronic Earth: scientific information resources and information-communicative technologies".

References

Arsky, Yu., V. Gitis, and A. Shogin (2007a), Electronic Earth – the network environment of search, integration and analysis of geographical data, Smirnovsky Collected Works, PIK VINITI, Moscow.

Arsky, Yu., V. Gitis, A. Shogin, and A. Weinstock (2007b), Geoinformation Environment for Analysis of Spatial and Spatio-Temporal Data, IUGG XXIV General Assembly, IUGG, Perugia, Italy, 2–13 July 2007.

Giardini, D., G. Grünthal, K. Shedlock, and P. Zhang (2003), The GSHAP Global Seismic Hazard Map, in *International Handbook of Earthquake & Engineering Seismology, International Geophysics Series 81 B, edited by W. Lee et al.,* p. 1233, Academic Press, Amsterdam.

Gitis, V. G., B. V. Ermakov, L. V. Ivanovsnaia, B. V. Osher, D. P. Trofimov, V. Schenk, Ju. K. Shchukin, and E. F. Jurkov (1993), The Information Technology of the GEO system for Prediction Mmax of Earthquake, *Journal of Earthquake Prediction Research, 2*(2), 221.

Gitis, V. G., B. V. Ermakov, V. N. Semov, Ju. K. Shchukin, and E. F. Jurkov (1994), E.F.GEO Expert System Application to Oil and Gas Resource Forecasting by the Deep Criteria (Example for West-Siberian Platform), *Sci. de la Terre, Ser. Inf., 32,* 653.

Gitis, V. G., B. V. Osher, S. A. Pirogov, A. V. Ponomarev, G. A. Sobolev, and E. F. Jurkov (1994), A System for Analysis of Geological Catastrophe Precursors, *Journal of Earthquake Prediction Research, 3,* 540.

Gitis, V., and B. Ermakov (2004), *Basics of spatio-temporal prediction in geoinformatics,* 256 pp., Fizmatgiz, Moscow.

Trifunac, M. D., and A. G. Brady (1975), On the correlation of seismic intensity with peaks of recorded strong ground motion, *Bull. Seismol. Soc. Amer., 65*(1), 139.

V. G. Gitis and A. P. Weinstock, A. Kharkevich Institute of Information Transmission Problems RAS, Moscow, Russia (gitis@iitp.ru)

A. N. Shogin, All-Russia Institute of Scientific and Technical Information RAS, Moscow, Russia