Geographic Information System and Geoportal «River basins of the European Russia»

O P Yermolaev, S S Mukharamova, K A Maltsev, M A Ivanov, P O Ermolaeva, A I Gayazov, V V Mozzherin, S V Kharchenko, O A Marinina, F N Lisetskii

Institute of Environmental Sciences, Kazan Federal University, Kremlyovskaya str., 18, 420008, Kazan, Russia

E-mail: oyermol@gmail.com

Abstract. Geographic Information System (GIS) and Geoportal with open access «River basins of the European Russia» were implemented. GIS and Geoportal are based on the map of basins of small rivers of the European Russia with information about natural and anthropogenic characteristics, namely geomorphometry of basins relief; climatic parameters, representing averages, variation, seasonal variation, extreme values of temperature and precipitation; land cover types; soil characteristics; type and subtype of landscape; population density. The GIS includes results of spatial analysis and modelling, in particular, assessment of anthropogenic impact on river basins; evaluation of water runoff and sediment runoff; climatic, geomorphological and landscape zoning for the European part of Russia.

1. Introduction

Currently, in Russia, there is no integrated geospatial database or geographic information system tied to the basins of small rivers. The urgency of creating such a complete coverage GIS that is capable to accumulate large volumes of spatial information on natural systems and integrated information on the state of small river basins is in relevancy of studying the results of increasing anthropogenic impact, as well as climatic changes observed in different landscape areas on basin geographic systems.

The large region of Russia – its European part which is some 4 million square kilometres with the most part of population and industrial and agricultural potential of the country – is being fundamentally studied in terms of geographical analysis of minor river basins and river runoff while creating designated geographic information system (GIS) "River basins of the European Russia". The designed GIS is considered to be a base for modern data and knowledge on geographic, hydroclimatic, geoecological and other characteristics of natural resource potential of tens thousands river basins. The thematic information sources are the data from long-term monitoring's, Earth's remote sensing, and accumulated corpus of cartographic materials from state surveys. The GIS includes not only actual information, but also the results of its comprehensive spatial analysis and modelling, assessments of anthropogenic load on river basins, the results of studying the patterns of water runoff formation depending on the landscape and geographical conditions in the European part of Russia.

The information accumulated in the GIS is located on the geoportal "River basins of the European Russia" with public (open) access, which will enable a wide range of representatives of a scientific community and experts in the field of environmental management and protection to obtain thematic and geographic information on natural and anthropogenic characteristics of river basins, as well as it will serve to meet the information needs of the population.

2. Material and methods

Study area is the European part of Russia, except for mountainous areas (figure 1). Scale studies (level of spatial detail) is regional (corresponding to a scale 1: 1 000 000).

Figure 1. Study area: The European part of Russia.

2.1. Building the layer of river basins – input data and methodology

The basis for making the GIS is an electronic map (vector layer) of river basins in the study area. The river basin acts here as a basic operational and territorial unit for the generalisation of diverse natural and anthropogenic information, as well as for conducting spatial analysis. The basin approach provides, it appears, the most convenient unit of spatial analysis for the Earth's temperate zone, where there is a dense river network, and the main relief-forming role is played by the permanent and temporary streams. An important feature for the use of river basins as territorial units for assessment and analysis is the fact that they are geosystem formation with all inherent properties. Pools also assure the demand for representativeness to the fullest extent. They are easily constructed when the level of generalisation of work changes. A large number of elementary basins covering the study area provide a representative sampling for statistical analysis and modelling.

Currently, there are a number of products obtained through various projects that are models of catchment basins of a particular territorial coverage. Such products differ in their intended purpose, in terms of spatial detail, in the degree of data openness, and etc. Speaking of open products, such projects as HydroSHEDS/HydroBASINS, CCM, and Ecrins [1-3], which cover the territory of the European Russia, should be mentioned first of all. However, the distinctive feature of these global models of basin geosystems is that they have been created for streamflows constructed using digital elevation model (DEM) and are models of catchments of temporary and permanent hydrographic network that is far from the river basins presented on topographic maps. As far as we know, a detailed network of river basins for the entire territory of Russia or at least for large parts of it has not been created yet. Regional geodatabases created by individual researchers or groups of researchers usually within a single constituent entity of the Russian Federation, are characterised by the disunity and lack of a unified methodology for constructing model boundaries. All of this has determined the task of creating our own model of river basins on the territory of the European Russia.

The process of creating a model of river basins (a vector layer of river basins) comprises the following stages [4]:

- initial data selection;
- pre-processing of the relief model;
- construction of pools' boundaries in automatic mode;
- post-processing – automatic and manual correction of pools' boundaries.

As initial data, we have selected DEM GMTED2010 (Global Multi-resolution Terrain Elevation Data 2010) [5] of 250 m spatial resolution and vector map of water objects of the Russian Federation, scale of 1: 1 000 000 (refined DCW). Pre-processing included creation of a single picture from 9 fragments of GMTED2010, correction of DEM using the hydrographic network map [6, 7], elimination of local "swallow holes" that were present in the original model [8, 9]. For the entire study area, river basins' boundaries were allocated in automatic mode according to the algorithm implemented in the Whitebox GAT programme [10].

The vector geometry of the boundaries after the automatic selection with accuracy up to a linear raster cell size was corrected by smoothing the polylines. The topological consistency of the geometry of rivers and basins' boundaries in river mouths (the "convergence" of the basin boundary to the final node of the polyline representing the river) was improved with procedures implemented in ArcGIS and partially (about 5%) in manual way (figure 2).

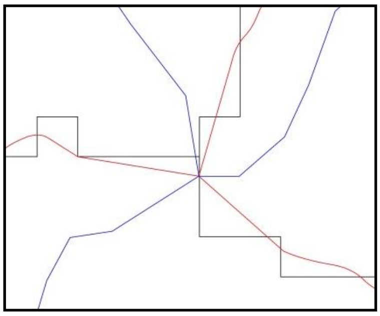

Figure 2. The example of river basins' boundaries before (black lines) and after post-processing (red lines); blue lines are rivers.

The allocation of basins was carried out planarly, that is, not only the basins of small rivers (on the accepted scale they are streamflows of the first order), but also inter-basin (or inter-flow) spaces. In total, layer of river basins comprises 53865 objects.

2.2. Sources of thematic geodata

For the GIS' content, thematic geodata (mainly from open sources), which characterise the natural and resource potential of the territory and correspond to the level of spatial detail of the studies have been processed.

Geomorphometry of the relief. On the basis of the relief model GMTED2010, the derived layers of the morphometric characteristics of study area's relief have been calculated. They are: steepness of slopes, exposure of slopes, erosivity of relief, length of water thread, profile and planned curvature, and order of streamflows. To minimise distortion in the calculation of morphometric indicators, equidistant projection was applied. The steepness (angle of inclination), exposure, curvature, and lengths of water threads were calculated using the ArcGIS software complex, which applies techniques for calculations [11, 12]. The erosivity of the relief and the order of the streamflows were calculated in the WhiteBox GAT software [10]. When constructing the model of water runoff orders, the traditional scheme of the Philosof-Strahler was applied.

Climate. In order to determine the climatic indicators, data of meteorological stations of Roshydromet were used [13]. Sixteen climatic parameters, representing averages, variation, seasonal variation, extreme values of temperature and precipitation, were estimated on the basis of long-term observations of daily air temperature and precipitation at meteorological stations for the period from 1960 to 2014.

Spatial interpolation of these 16 climatic parameters was constructed for the entire territory of the European part of Russia (spatial interpolation method is Multilevel B-spline Approximation).

Land Cover. The source of information on land cover types of the territory was spatial database, gathered based on set of maps of terrestrial ecosystems of Russia *TerraNorte RLC*, elaborated in Space Research Institute (IKI) Russian Academy of Sciences [14]. The main source of the initial information is satellite observation data by spectroradiometer MODIS (MOD09 standard products). The detection method involves multi-annual time-series of spectral reflectance satellite measurements preliminarily cleaned from impact of cloud and other contaminating factors, along with automated land cover types recognition based on a locally-adaptive classification algorithm [15]. The map legend consists of thematic classes, including classes of vegetation cover (forest: evergreen dark needle-leaf, evergreen light needle-leaf, broadleaf, mixed with needle-leaf majority, mixed, mixed with broadleaf majority, deciduous needle-leaf, sparse deciduous needle-leaf; grasslands and shrub lands; tundra; wetlands; recent burns; croplands and non-vegetated areas).

Soils and landscapes. GIS includes coverage of the study area with electronic soil and landscape maps. The electronic soil map of Russia was developed in the Soil Science Institute named after V.V. Dokuchaev and is available as the Unified State Register of Soil Resources of Russia [16, 17]. The polygonal objects of the layer are the digitised contours of the RSFSR soil map at a scale of 1: 2,500,000. Each polygon contains information on the soil cover and soil-forming rock material. The map's explanatory notes include 205 soil varieties, 70 soil complexes, 6 non-soil formations, and 30 variants of soil texture and petrographic composition of soil-forming rock material.

An electronic landscape map of the USSR is available through the link [18] on the website of the International Institute for Applied Systems Analysis (IIASA). This is a digital version of the landscape map of the USSR at a scale of 1: 2,500,000. The section, group, class, type, subtype, sort, sort types, and landscape types are presented as attributes of landscapes. It should be noted that the coordinate system indicated in the map description does not allow us to obtain an exact location (closure error: 15 km). We carried out a refinement of the digital map georeferencing, and selected projection's parameters ensuring its spatial linking with the rest of the geoinformation.

Hydrological geodata. The main source of data on water runoff and runoff of suspended load of rivers in the European Russia is published materials of long-term monitoring observations at hydrological posts – "Surface water resources of the USSR". On the basis of these materials, a database has been created. It presents information on the average annual water discharge and solid discharge for all the years of observation. A part of the data is obtained at the open Global Runoff Data Center (http://www.bafg.de/GRDC/). A part of the information for posts of the Volga-Kama basin has been updated according to data up to 1985 obtained from open sources (http://caspi.ru/), and according to data up to 2013 for posts in the Upper Volga, Kama, and Lower Volga basins, obtained within the framework of work under the contract with FGBU VNIIGMI-MTSD (RIHMI-WDC, http://meteo.ru/data/). The initial information on the localisation of gauging stations is FGBU VNIIGMI-MTSD's open data (http://meteo.ru/data/). To correct the location of posts, information from the volumes of hydrological directories (a name of a river and a settlement, the length of a river from the mouth to the post), layers of topographical map at a scale of 1: 1,000,000, Google Maps, and Yandex.Maps were used. The creation of point objects at gauging stations' layer was carried out in compliance with topological relations with the objects of the hydrographic network map (figure 3).

To study the dependencies of a river runoff on the landscape and geographical conditions of its formation, we have created a vector layer of catchments related to hydrological stations. In order to solve this issue, we used the processed model of the relief GMTED2010 and a layer of point objects representing hydrological stations. The construction of the watersheds' boundaries was performed in an automatic mode using the WhiteBox GAT program [10]. In order to assess the quality of the process of boundary setting, a comparison between the polygon areas of the obtained layer and the values of areas of the gauging stations' basins taken from the hydrological reference books was made.

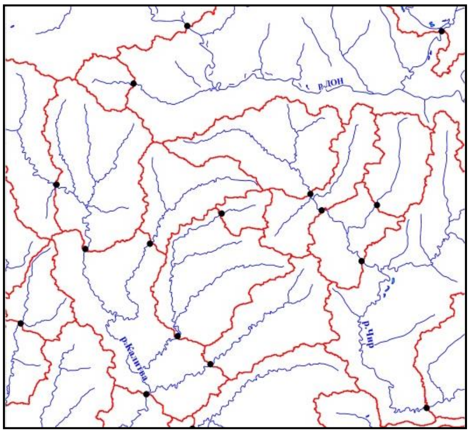

Figure 3. Boundaries of catchments related to hydrological stations: blue lines are rivers, black points are hydrological stations, red lines are boundaries of catchments.

2.3. Generalization of geodata
A generalisation of the thematic information included in the GIS into the operational and territorial units of analysis was done, i.e. river basins and basins related to hydrological stations. For each basin, generalised morphometric characteristics of the relief, average climatic parameters, prevailing soil characteristics, predominant type and subtype of landscape, and percentage of different land cover types were determined. All operations with geodata, processing and analysis (spatial interpolation, generalisation of data pools from thematic raster and vector layers) were performed using programs written in R language (rgdal, raster, MBA packages) [19-24], and with the help of developed SQL-requests in the DBMS PostgreSQL/PostGIS.

3. Results and discussion

3.1. GIS composition
The GIS "River basins of the European Russia" currently includes the following vector and raster cartographic layers: river basins; hydrological stations; catchments related to hydrological stations; digital elevation model; terrain slope; terrain aspect; plan curvature; profile curvature; erosional potential of relief; lengths of the flows; meteorological stations; average annual air temperature; average air temperature in January; average air temperature in July; average maximum of year temperature; average minimum of year temperature; average annual air temperature amplitude; temperature standard deviation for the year; average number of days with temperatures below 8 degrees; frequency of severe frosts (-30 deg. and below); sum of active temperatures; average annual precipitation; average precipitation in May-August; average precipitation for the cold season; average precipitation for the warm season; coefficient of variation of annual precipitation (%); hydrothermal coefficient; bioclimatic layers from WorldClim database; geological map (classes of pre-Quaternary deposits); geomorphological map; soil map; land cover types (vegetation); landscape map; layers of topographic map scale of 1: 1,000,000; boundary of the study area.

Attributes of objects of main layers (in particular layer of river basins (53865 basins) and layer of basins related to hydrological stations (1892 basins) are the following: basin identifier, basin area, predominant soil type in basin, predominant type of soil-forming rock, predominant class of pre-Quaternary deposits, predominant type and subtype of landscape in basin, percentage of woodland, cropland, grassland, shrubland, wetland, lakes in basin, average, minimum and maximum elevation in basin, average slope, average aspect (from "cold" -1 to "warm" 1), erosional potential of relief and flows length in basin, average annual air temperature, average air temperature in January, average air temperature in July, average maximum of year temperature, average minimum of year temperature, average annual air temperature amplitude, mean square temperature deviation for the year, average number of days with temperatures below 8 degrees, frequency of severe frosts, sum of active temperatures, average annual precipitation, average precipitation in May-August, average precipitation for the cold season, average precipitation for the warm season, coefficient of variation of annual precipitation, hydrothermal coefficient.

3.2. Spatial analysis based on GIS

The geodata accumulated in the GIS make it possible to study the features of forming and functioning of basin geosystems in the landscape and geographical conditions of the European part of Russia, to conduct thematic mapping of basins with a sufficiently high level of spatial detail, and to solve tasks of comprehensive spatial analysis and modelling. As examples, we present the results of solving two tasks (the solution itself is not discussed in this paper).

Climatic zoning. The method of automated zoning of the study area was developed. It is based on Kohonen Self Organizing Map (SOM) method [25] and uses the authors' software. On the basis of 16 climatic parameters we obtained climatic zoning for the European part of Russia. Basins of small rivers were used as spatial units of analysis. Figure 4 shows the result of climatic zoning and the class ordination in the attribute space. Classes are ordinated by the temperature gradient in the vertical direction on the ordinate plane, and by humidity gradient in the horizontal direction.

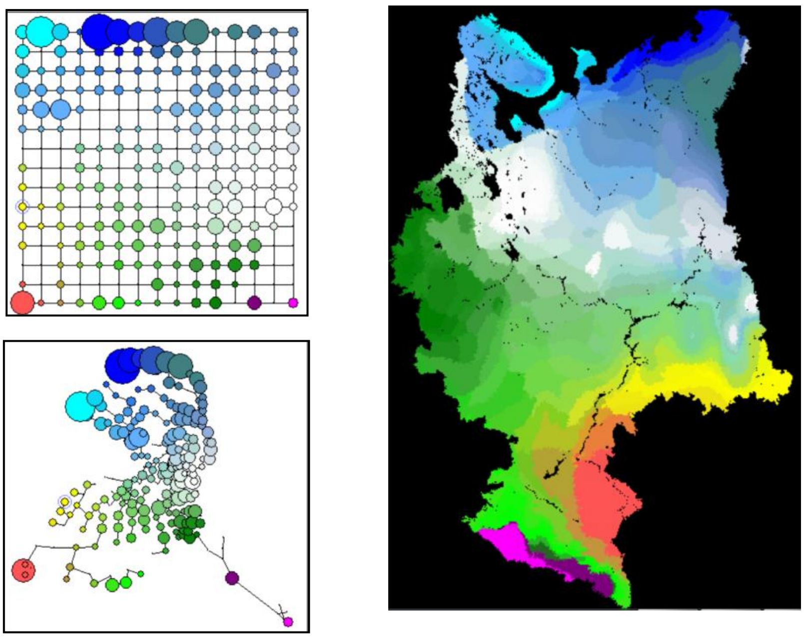

Figure 4. Climatic zoning for the European part of Russia: left are classes in geographical space, right up are neurons (classes) on ordination plane, right bottom is Sammon map of the minimal skeleton of classes (with a thematic palette, the size of the icons is proportional to the number of basins assigned to the corresponding class).

The class of semi-desert dry climate basins (the red color) is clearly distinguished; classes of steppe basins with unstable moisturisation, lower temperature class (the yellow color) and class of steppes with insufficient moisturisation (the bright green color) are close to the previous one in terms of their properties. Then, there are classes of the basins of subboreal and boreal temperate continental climate. They are: forest-steppe with sufficient moisturisation and mixed forests (shades of green, transition to white), and taiga with excessive moisturisation (shades of light-grey and grey-blue). And then, there are basins in the subarctic climate (the blue color). The classes of high mountain climate basins (the purple color) and even the subtropical climate (the pink color) are clearly distinguished. In geographical space (on the map), it can be observed that the obtained classes demonstrate not only the latitudinal (zonal); but also the longitude and sector gradient. In general, the climatic zoning carried out is quite adequate to the published climatic maps of Russia. However, unlike classical climate maps with their clear boundaries between climatic regions or zones, it reflects continual changes in climatic conditions.

Evaluation of water runoff. Dependence of the characteristics of water runoff on the landscape and geographical conditions and anthropogenic stress are studied by statistical methods (GLM, GAM). Statistical models of these relations were constructed. The most significant predictors of specific water discharge model are the following: sum of active temperatures, annual precipitation, average slope, average elevation (for mountain areas), percentage of woodland in basin, percentage of cropland, soil category.

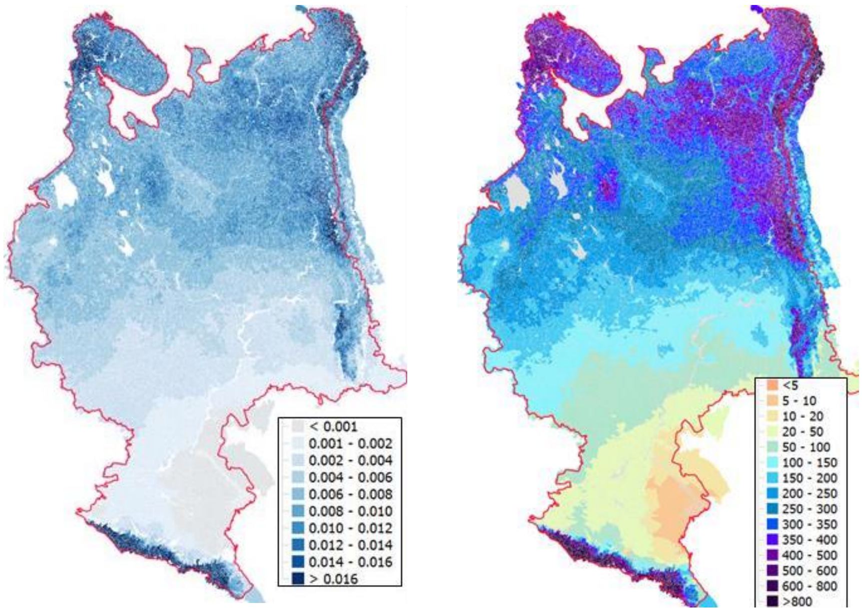

Figure 5. Map of specific water discharge (left) and map of annual water runoff layer (right).

Models explain about 80% of data variability, provide acceptable prediction accuracy and represent the main dependencies between water runoff and its formation conditions at the scale of study.

Spatial extrapolation of the values of water runoff was produced in hydrological unexplored areas of the European Russia on the basis of constructed models. Figure 5 shows the results as a map of specific water discharge and map of annual water runoff layer.

3.3. Geoportal "River basins of the European Russia"

The geoportal "River basins of the European Russia" has been developed and put on the Internet: http://kpfu.ru/ecology/nauchno-issledovatelskaya-rabota/rechnye-bassejny-evropejskoj-rossii; direct link: http://bassepr.kpfu.ru/ .

On the geoportal, with the help of cartographic web services, the major part of the geodata that has been accumulated in the GIS is displayed. It is, first of all, the layer "Pools" that is basins of small rivers and their inter-tributory spaces. On the geoportal, there are also such layers, as "Gauging stations" layer which is hydrological posts of Rosgidromet (Russian Federal Service for Hydrometeorology and Environmental Monitoring) provided with data on average annual water consumption; "Study area" layer that shows the boundary of the studied area, and two layers in the group "Hydrography" that are hydrographic network objects represented on the map M 1:1,000,000. The geoportal enables its visitors to use maps, such as a map of Russia, OpenStreetMaps, Yandex-map, or Google-map, and satellite images (mosaic of synthesised images from Google, Yandex, Roskosmos, ScanEx) as a “background layer” (Figure 6).

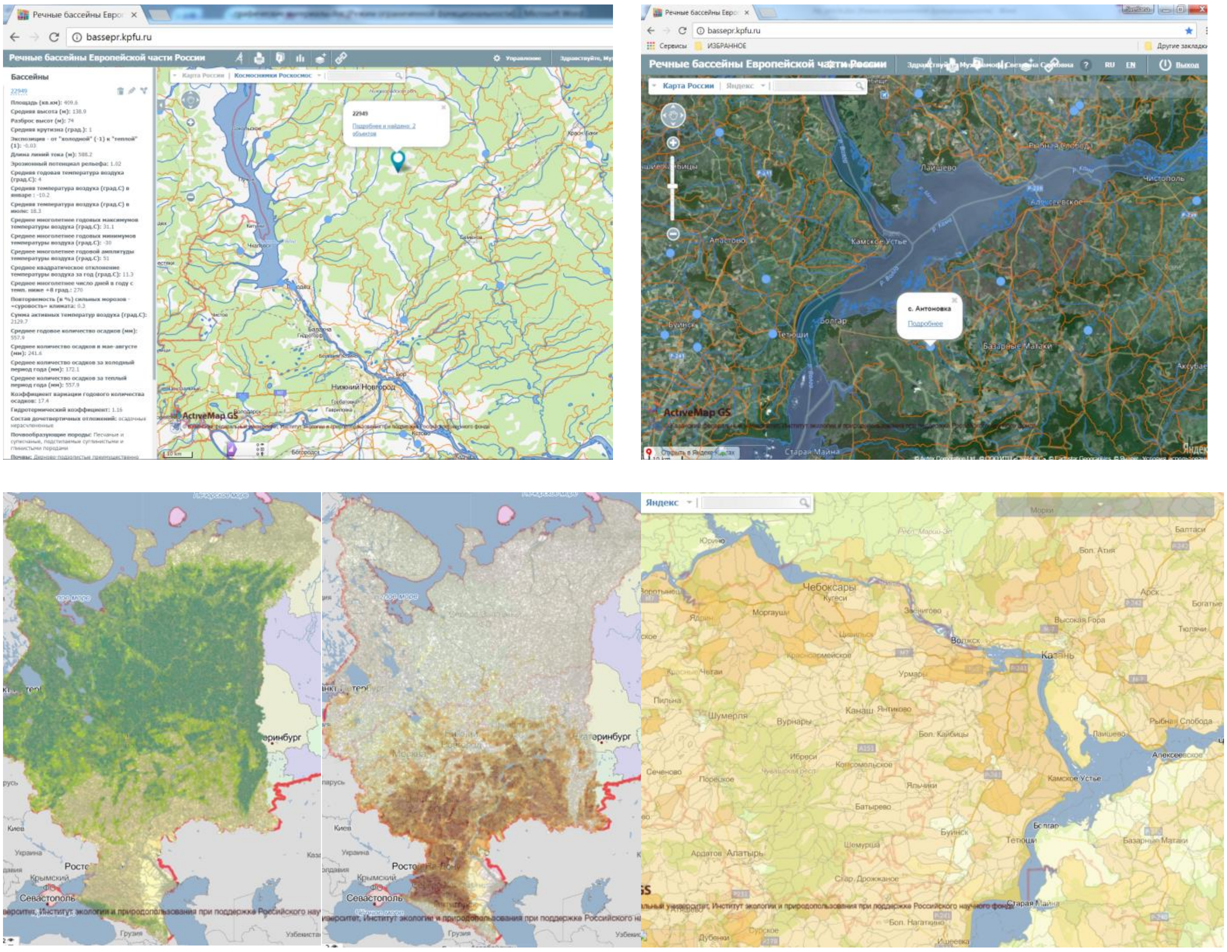

Figure 6. Examples of pages from the geoportal “River basins of the European Russia” on the Internet.

We have provided public (open) access to the developed resource. Any user who visits the geoportal gets the opportunity to familiarise themselves with the map of basins, information on the natural and anthropogenic characteristics of basins, and to save the information they are interested in (in tabular form). One can see the locations of the hydrological observation posts and acquire information on the duration of observations of the water runoff and the flow of suspended load at posts.

We view the developed geoportal as the first version. Expert interviews (via online platform *Testograf*) with the faculty, research scientists (ecologists, geographers, social ecologists, sociologists, etc.), representatives of project and environmental organisations, representatives of the public sector, mass media, and environmental NPOs from different regions of Russia have been conducted to optimise the content of the geoportal. The follow-on revision of the geoportal is carried out to expand the presented geoinformation, improve the functionality, display data on the geoportal in the form of thematic, analytical, and integrated maps, and implement a “download geodata” option.

Acknowledgement
This study was supported by the Russian Science Foundation (project 15-17-10008).

IOP Conf. Series: Earth and Environmental Science **107** (2017) 012108 doi:10.1088/1755-1315/107/1/012108

References

[1] Lehner B, Verdin K, Jarvis A 2008 *Eos, Transactions* **89**(10) 93–4
[2] Vogt J et al. 2007 *Report EUR 22920, European Commission – Joint Research Centre* 120p
[3] 2012 *EEA Catchments and Rivers Network System ECRINS v1.1* 111p
[4] Ermolaev O et al. 2017 *Geography and Natural Resources* **38**(2) 131-8
[5] Danielson J, Gesch D 2011 Global multi-resolution terrain elevation data 2010 (GMTED2010) *Open-File Report 2011–1073* (Reston: U.S. Geological Survey) 26p
[6] Mal'tsev K, Yermolayev O 2014 *Geomorfologiya* **1** 45-53 (In Russian)
[7] Yermolaev O, Mal'cev K, Ivanov M 2014 *Geogr. i prirodnye resursy* **3** 32-9 (In Russian)
[8] Pogorelov A, Dumit Zh 2009 *Kuban river basins relief: morphologic analy*sis (M.: GEOS) 218p (In Russian)
[9] O'Callaghan J, Mark D 1984 *Comput. Vis. Graph. Image Process.* **28**(3) 323–344
[10] Lindsay J 2014 *Proceedings of the GIS Research* (UK 22nd Annual Conference, The University of Glasgow, UK) 16–18
[11] Burrough P, McDonell R 1998 *Principles of Geographical Information Systems* (Oxford University Press, New York) 190 p
[12] Zeverbergen L, Thorne C 1987 *Earth Surface Processes and Landforms* **12** 47–56
[13] Buligina O, Razuvayev V, Aleksandrova T *Daily Temperature and Precipitation Data for Russia and U.S.S.R. Stations (TTTR)* Patent № 2014620942
[14] Bartalev S et al. 2011 *Current problems in remote sensing of the earth from space* (CPRSES) **8**(4) 285-302 (In Russian)
[15] Bartalev S et al. 2012 *Current problems in remote sensing of the earth from space* (CPRSES) **9**(2) 9-27 (In Russian)
[16] 2014 *Edinyy gosudarstvennyy reestr pochvennyh resursov Rossii* [Unified state register of soil resources of Russia] V. 1.0 Collective monography (M.: V.V. Dokuchaev Soil Inst.) 768 p (In Russian)
[17] Rukhovich D et al. 2013 *Pochvovedenie* **3** 251-67 (In Russian)
[18] *Forests and forestry in Russia. Data and analysis. Landscapes* (http://webarchive.iiasa.ac.at/Research/FOR/forest_cdrom/russian/for_cond_ru.html#landscapes) (Accessed 18.05.2017) (In Russian)
[19] Lee S, Wolberg G, Shin S 1997 *IEEE Transactions on Visualization and Computer Graphics* **3**(3) 229-44
[20] Finley A, Banerjee S 2014 *MBA: Multilevel B-spline Approximation* (R package version 0.0-8)
[21] Wood S 2003 *Journal of the Royal Statistical Society* **65**(1) 95-114
[22] Hijmans R 2014 *raster: Geographic data analysis and modeling* (R package version 2.2-31)
[23] Bivand R, Keitt T, Rowlingson B 2014 *rgdal: Bindings for the Geospatial Data Abstraction Library* (R package version 0.8-16)
[24] R Core Team 2014 R: *A language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria)
[25] Kohonen T 1997 *Self-organization and Associative Memory* (New York: Springer-Verlag) 428 p