BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, September 2015. Vol. **12**(Spl. Edn. 2), p. 651-656

Publicly Available Sources of Geographical Data About River networks of European Russia[1]

Sergey Vladimirovich Kharchenko[1], Oleg Petrovich Yermolaev[2], Kirill Aleksandrovich Maltsev[3], Svetlana Sayasovna Muharamova[4] and Maxim Andreevich Ivanov[5]

[1]Candidate of Geography. KFU, Researcher, [2]Doctor of Geography. KFU, Professor, [3]Candidate of Geography. KFU, Assistant Professor, [4]Candidate of Biology. KFU, Assistant Professor, [5]KFU, Researcher Kazan Federal University, 18 Kremlevskaya street, Kazan, Russia, 420000

doi: http://dx.doi.org/10.13005/bbra/2245

(Received: 11 June 2015; accepted: 21 August 2015)

The study provides results of a comparative analysis of several publicly available sources of data about the hydrographic network of the European part of Russia. The comparison was made by a number of quantitative and qualitative characteristics such as: detailed data (detailed network), accuracy (the degree of coincidence with real situation), initial materials used in the creation of a source, presence or absence of the attribute information. These parameters determine the applicability of a variety of sources to solve scientific and applied problems.

Key words: River network, publicly available data, European part of Russia, comparative analysis.

The European part of Russia covers an area of a little less than 4 million square kilometers. Drainage of such a large area is made through the complex and extensive network of tens of thousands of rivers, small rivers and streams. In support of this we suggest an example. According to the Water Registry of Russia, the basin of Volga alone includes 8,875 rivers and 688 streams[1]. It is natural to assume that a small number of streams are only due to generalization of data, and the real proportion is in streams' favor.

When solving a variety of problems (for example, hydrological and geomorphological modeling, runoff forecasting, identifying geological structures, etc.), most geographers face the challenge of selecting the correct information on the hydrographic network. It is clear that none of the sources – regardless of whether it is publicly available or "restricted" – is totally reliable. Any electronic or analog (including paper maps) image of the hydrographic network is just its model characterized by one or another level of abstraction. The task of the researcher in the selection of a source of information about the "drainage system" to solve some of its task is to determine the applicability of each of the available sources for this particular task. This study presents the results of a comparison of various publicly available sources of information about the structure of the river network that can be used by geographers in the work of the European part of Russia. Some of these sources cover a substantially larger area, which will be indicated in the text of this study further.

* To whom all correspondence should be addressed.

MATERIALS AND METHODS

The study analyzes 6 sources of publicly available information about the structure of the hydrographic network. These sources are: Russian DTM[2], VMap0 [3], Open Street Map[4], CCM 2.1 [5], Ecrins[6] and HydroSheds (with HydroBasins)[7]. They have different coverage, and in Russia they cover only its European part (not further north than 60° N). The Asian part is covered only by 4 out of 6 sources (VMap0, Russian DTM, OpenStreetMap and HydroSheds). The European part of Russia is planar covered by only 3 sources (VMap0, Russian DTM, OpenStreetMap). In most cases (and, most importantly, in the "primary sources"), these data are provided in the form of vector files used in GIS applications (in particular, files in *.shp format).

The comparison was made in the GIS MapInfo. Obviously more precise "standards", with which different versions of the submission of the hydrographic network were compared, were maps of 1:25,000 scale, as well as GoogleMaps satellite images uploaded to GIS (shoots of DigitalGlobe and GeoEye satellites).

Given the possible discrepancies in delimitation of the borders of European part of Russia (hereinafter – ER), we will separately mention our understanding of the borders of this large region. We consider the border of the "European Russia" as being drawn along the Arctic coast of the mainland (excluding the Arctic islands), then on the state border of the country (excluding the peninsula of Crimea, the islands in the Baltic Sea, the Sea of Azov, the Black Sea and the Caspian Sea-Lake) until the crossing of the border between Russia and Kazakhstan and the main watershed of the Ural Mountains – the Ural-Tau chain. Then along this chain and the main watershed of Pai-Khoi chain to the Arctic coast of Russia.

RESULTS AND DISCUSSION

Russian DTM

Within the ER borders, the hydrographic network of Russian DTM is represented with linear (streams) and polygonal (basins) layers. The first includes nearly 62,000 network segments. Most of the network is divided into elementary segments, i.e. sections of watercourses between the mouths of the tributaries flowing into it. In other words, only streams of the 1 order (by any of the common classifications [8-11], except for the classifications by Gravelius [12] and Horton [13]) are represented by 1 segment, other watercourses are represented by 2 or more segments. However, this approach is not sustained by the creators of Russian DTM everywhere – in rare cases, the confluence of inflow into the main river for some reason does not serve as a basis for splitting it at the confluence. This creates difficulties in automated classification of watercourses in the structure of river network (calculation of their orders). The great benefit of the database is presence of attribute information and Cyrillic names of the rivers. In total, more than 24,200 segments have such marks. In addition, it provides information on the navigable waterways. A layer of water bodies includes more than 27,200 objects. 1,160 of them have Cyrillic names – but they are not only lakes and reservoirs, but also lake-like expansions of the river beds of large rivers expressed on a map scale. The inventive mapping scale in Russian DTM is 1:1,000,000.

VMap0

The VMap0 hydrographic network includes two layers with linear objects (rivers and "canals") and one layer of polygonal objects (lakes and reservoirs). Within the ER borders, the river network includes 24,890 segments, the canal network – 108 segments. A layer of lakes and reservoirs is represented by a total of more than 12,000 objects. However, many water bodies are divided into parts not expressed under real conditions. This division increases the formal number of water bodies 3-4 times. In the attribute information, natural water bodies and reservoirs are separated – they correspond to the "InlandWater" (lakes, river openings) and "LandSubjecttoInundation" classification (reservoirs). More than 4,400 streams and 2,300 water bodies (!not geographical objects but namely objects in the GIS layer) are marked in Latin script. In addition, the information is given on whether it is a permanent water body/watercourse or sporadic. The inventive scale (and therefore, the detail and accuracy of the data) is 1:1,000,000. In contrast to Russian DTM, the VMap0 hydrographic network is divided into segments correctly.

Open Street Map

Data of the noncommercial portal

openstreetmap.org – free map of the world – is also available in the file format for geographic information systems[4]. Very highly detailed OSM "drainage system" is explained with the fact that the map itself and by its content largely corresponds to topographic maps of the largest scale – 1:10,000 or 1:25,000. The map is created by enthusiasts from satellite images, photos, GPS tracks and other spatial information. As in the two previous cases, data on river networks represent a set of data about streams and reservoirs. The number of watercourse objects within the ER borders is 224,600, i.e. 3.5-4 times more than in Russian DTM and almost 10 times more than on VMap0. The watercourses are classified into two types: "stream" (122,100 objects) and "river" (102,500 objects). Despite the less number of objects of the "river" type, their average length is 3 times longer than the length of the objects of the "stream" type. The value of a maximum length differs 6 times already. These differences, as well as the position of objects of the "stream" type in the structure of the network suggests that the creators predominantly include streams and small rivers of the upper parts of the river systems in them. Nearly 59,700 objects are marked. As is the case with Russian DTM, many of the watercourses are not divided into elementary segments, which may hinder the task of automated classification of flows by ordinal position in the network.

A layer of OSM water bodies on the ER territory includes more than 352,800 objects – they are lakes and reservoirs ("water"), bogs ("wetland"), expansions of riverbeds expressed on a scale ("riverbank"). More than 24,900 objects have names, both the names of their own and – less frequently – common nouns (like "pond", etc.).

All three described data sources have a topological feature that distinguish them from the other three sources – the existence of the "circular" formations, bifurcations (the river branches into arms and then back), which also doesn't allow automated classification, for example, by A. Strahler [8]. A potential solution could be a principle of conditional orders proposed by N.I. Alekseevskiy, applied by him for the determination of the ordinal position (corresponding to the water content of the flow) of arms in branched riverbeds and deltas[14]. However, this principle is not yet widely used, even in Russia. At least one of the sources – Russian DTM – has "circular" formations that do not exist in the real river network. There are situations where, for example, two adjacent tributaries take start at a single point, which does not correspond to reality.

Three other sources – CCM 2.1, Ecrins and HydroSheds – are in fact not networks of rivers, but rather networks of thalwegs of any linear drainage depressions. In humid climate conditions, these concepts are often identical, but in the arid and semi-arid landscapes "dry lands" are often found, which are at best the "canals" for temporary water flows during periods of snow melt or rain. These networks are built according to the morphometric analysis of global digital elevation models SRTM and GTOPO.

CCM 2.1 and *Ecrins* are almost identical networks. Their territorial coverage is all Foreign Europe and much of the ER, excluding the basin of the Pechora river. Reservoirs in these data sets are allocated separately, but the main linear layer – streams – also is not broken by the presence of lakes or reservoirs on the rivers, but is included by their axis lines into the network as a kind of thalwegs. Each river system has a clear tree structure without "circular" formations.

Within the ER territory, the CCM data is broken down into a series of "macro-basins" called "Volga", "Don", "Baltics", "Dnieper" and "Ural" and having their numeric codes. Four of the five names correspond to major rivers of the four major ER river systems, and one would think that the CCM data is divided by sections of watersheds of the Volga, the Don, the Dnieper and the Ural. However, in this case, obviously, the vast areas of Ciscaucasia, Baltic crystalline shield and continental European Arctic are not covered. I.e. the names of these "macro-basins" are conditional and only very roughly characterize their position and the borders.

The CCM data includes layers with so-called major rivers ("mainrivers"), segments of the river network ("riversegments"), river confluence nodes – the mouths of tributaries ("rivernodes"), the mouths of major rivers ("seaoutlets") and the catchment areas modeled for each of the network segments ("catchements"). "Mainrivers" layer contains the largest network of waterways, not divided into elementary segments.

For a small part of the objects, the attribute contains the name of the river in the Latin script. Layer «riversegments» contains much more detailed spatial and attribute information; the orders of segments by A. Strahler, the drained catchment area, the length of the segment and the cumulative length from the source, etc. may appear useful to researchers. The total number of segments of the ER river network, according to CCM data, slightly exceeds 400,000.

The Ecrins data for the ER territory is not broken up into five sections, as in the case of CCM (Ecrins is the product of CCM development and refinement). The attribute data of the layer of streams contains all the information that is available in the CCM database. The total number of segments of the river network is more than 352,000. In addition, the layers of watersheds, lakes and reservoirs are available.

The claimed scale of the two sources of data is 1:250,000. The detailed "drawing" of the drainage system fits this scale. Both CCM and Ecrins are built using the SRTM (south of 60° N) and GTOPO (north of 60° N) data adjusted by the interpolation methods to the SRTM resolution.

HydroSHEDS is the network of thalwegs built over a total area of land covered by the SRTM data. The digital terrain model before watercourses rendering was roughened up to the resolution of 15 arcseconds per pixel and 30 arcseconds per pixel. The HydroSHEDS data is respectively spread over two levels of details. In the first case (15 arcsec/ pixel), the network includes just under 106,400 segments. Splitting into segments is correct, so HydroSHEDS can be used for the calculation of thalwegs/watercourses orders. However, this data has another disadvantage – orthogonality of lines showing the thalwegs. The individual segments that simulate the planned configuration of the channel are located only at 0, 45, 90 and 135° angles. This feature greatly limits the possibility of, for example, river network lineament analysis to identify fault tectonics. The attribute data for each of the network segments includes the number of cells of the digital terrain model upstream of the segment, i.e. a parameter that with some degree of freedom can correlate with the distance from the source.

The minimum length of the broken line segment that approximates thalwegs in HydroSHEDS is about 500m. Accordingly, the hydrographic network structure in this data is reflected at least no worse than on paper maps at a scale 1:1,000,000.

Figure 1 shows the relative spatial relationship of various options of hydrographic network in the upper reaches of the Tim river (the Don basin) in Tim area of the Kursk region, where in the conditions of dissected, erosional relief of the Central Russian Upland, one should expect a good likeness of the planned morphology of the river network reflected by different sources.

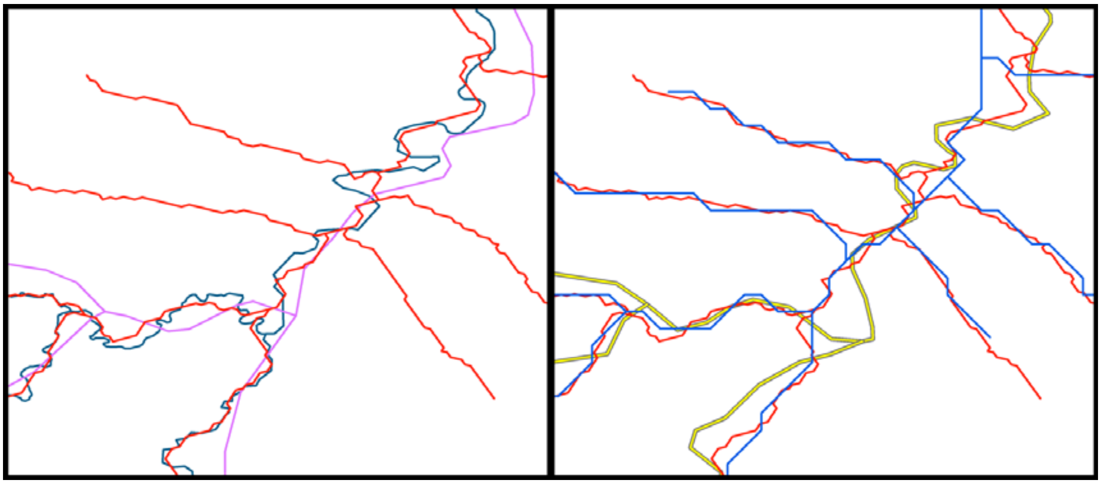

Fig. 1. To the left – fragments of hydrographic networks according to VMap0 (purple), Open Street Map (blue) and CCM 2.1 (red) data. To the right – fragments of hydrographic networks according to Russian DTM (yellow), HydroSHEDS 15s (blue) and Ecrins (red) data. Land area from west to east is 9 km

There is a relative "primitiveness" of the morphology of VMap0 networks, to a lesser extent – of those of Russian DTM and HydroSHEDS. The above-described orthogonality of HydroSHEDS network is also evident, when individual segments of curves take only discrete values of the azimuth orientation. Only CCM and Ecrins have complex structures, shown in red on the left and right images, respectively. It is easy to note the identity of their structure – these networks differ only in some parts of the upper parts of the river network, on the poorly dismembered vertex surfaces of interstream areas. The OpenStreetMap data most closely matches the actual morphology of the river network – this is proved by the imposition of the OSM river network on the topographic map of 1:25,000 scale and highly detailed satellite images. It should be noted that the OSM network – the most comprehensive of all the above – lacks 4 large right- and left-bank segments (2 on each side) available in all networks modeled on the basis of the digital elevation models. These segments are thalweg beams rather than permanent streams. I.e. in the forest steppe and steppe zones with characteristic ratios of precipitation and water runoff by DTM, a lot of "false" watercourses are allocated.

CONCLUSIONS

Various sources of data on the structure of the hydrographic network of the European part of Russia are characterized by a number of parameters, which together determine the applicability of a source for each specific task. Thalweg networks modeled on a digital elevation model are conveniently applicable and sufficiently reliable for the tasks of small and medium-sized part of the hydrological modeling (based on the topological structure of the river network). However, their use requires consideration of climatic conditions and the "dryland" drop-out. At the same time, they quite accurately reflect thalwegs and wellhead points where the accuracy of the spatial reference is essential to delineate watersheds. Without taking into account the spatial accuracy, the data of VMap0, Russian DTM and, especially, OSM have greater topological correctness. This data, however, has a number of other disadvantages. They include a significant shift of network elements with respect to their actual position of the slopes to the riverbeds, and sometimes to interstream areas (mostly – at the VMap0 and Russian DTM); details that are unnecessary and even harmful for some types of modeling (circuits, bifurcations).

ACKNOWLEDGEMENTS

The study was funded through the grant from the Russian Science Foundation (project No.15-17-10008)

REFERENCES

1. *Russian Federal Water Register*. textual.ru Retrieved 20.07.2015 from http://textual.ru/gvr.
2. *Digital maps from Roscartography in the ArcInfo format. DATA+: geographical informational system for business and society*. Retrieved 20.07.2015 from http://www.dataplus.ru/news/arcreview/detail.php?ID=2717&SECTION_ID=67&sphrase_id=2894291.
3. *Vmap0 – data description. GIS-Lab: geographical informational system and remote sensing*. Retrieved 20.07.2015 from http://gis-lab.info/qa/vmap0-about.html.
4. *OpenStreetMap data in the shape files format. GIS-Lab: geographical informational system and remote sensing*. Retrieved 20.07.2015 from http://gis-lab.info/qa/osmshp.html.
5. *Catchment Characterisation and Modelling (CCM). JOINT RESEARCH CENTRE. The European Commission's in-house science service*. Retrieved 20.07.2015 from http://ccm.jrc.ec.europa.eu/php/index.php?action=view&id=23.
6. *European catchments and Rivers network system (Ecrins). European Environment Agency (EEA)*. Retrieved 20.07.2015 from http://www.eea.europa.eu/data-and-maps/data/european-catchments-and-rivers-network.
7. *HydroSHEDS. Overview. HydroSHEDS*. Retrieved 20.07.2015 from http://hydrosheds.org/page/overview.
8. Strahler, A.N. (1957). Dynamic basis of geomorphology. *Geological Society of America Bulletin, 63*, 923–938.
9. Filosofov, V.P. (1959). *Valley's orders and their applications for geological surveying. Research yearbook 1955*. Saratov: Saratov University Press.
10. Rzhanitsin, N.A., *Morphological and hydrological regularities of river network*

organization (pp. 238). Leningrad: Hydrometeorological Publishing, 1960.

11. Schreve, R.L. (1967). Infinite topologically random channel networks. *J. Geol., Vol. 75, 2,* 178–186.
12. Gravelius, H. (1914). Rivers (pp. 179). Berlin: G.J. göschen Publishing.
13. Horton, R.E. (1948). Erosional development of streams and their drainage basins (pp. 158). Ìoscow: Publishing House of Foreign Fiction.
14. Alekseevskiy, N.I., & Sokolova, Yu.V. (1999). Organization of streams network in the bifurcations of river channels and deltas and approaches of its formalization. *Research's News of the Moscow University. Series 5. Geography, 2,* 13–19.