EURASIAN JOURNAL OF MATHEMATICAL
AND COMPUTER APPLICATIONS
ISSN 2306–6172
Volume 4 , Issue 3 (2016) 39 – 50

DEVELOPMENT OF THE INFORMATION-MODELING SYSTEM FOR FLOOD PREDICTION AT LARGE RIVERS IN SIBERIA

Zinoviev A.T. , Lovtskaya O.V, Baldakov N.A.

Abstract The paper gives the description of some components of the information-modeling system developed to predict floods. The scales and formats of the required cartographic information are analyzed. The composition and structure of the problem-oriented database for data storage and processing in the information-modeling systems is presented. The database designed as a "star" schema based on DBMS PostgreSQL is proposed. Tools for the construction of interactive maps of flood zones with the use of standard service of Google maps have been created. The maps allow to identify flood zones at various probability of exceedance, to get the information on areas and details of the flooded territories as well as to make science-based decisions on flood protection.

Key words: database, Web-GIS, PostgreSQL, Geoserver, CUAHSI

AMS Mathematics Subject Classification: 86A05, 68U35

1 Introduction

Inundations caused by spring and rainfall floods are among the most dangerous natural disasters in Russia. Prediction of such events is of great practical importance; it is a primary task the modern hydrology faces. On the one hand, being closely related with the protection of human life, health and well-being, the problem of river floods forecasting is currently central; on the other hand, it deals with operation of hydropower systems under release of large volumes of flood water [1]. The creation of methods for forecasting the extreme hydrological events (including floods) is one of major research areas of IWEP SB RAS [2].

When solving the urgent practical problems associated with prediction and evaluation of floodplains inundation, reliable results can be obtained on condition that rather complex mathematical models of flow are involved. Among feasible models for spatial calculation of kilometers-long streams, a planned 2DH model seems to be the most advanced [3].

To solve water-related problems, a large amount of spatially distributed hydrological and meteorological information is required, including GIS data on spatial characteristics of a catchment, data on the underlying surface, and a digital elevation model (DEM) of a river's bed and its valley. Thus, the development of information-modeling systems (IMS), which combine computational modules, databases and GIS, is topical nowadays. An integrated database provides the interaction of information-modeling systems with GIS; it contains the information on the creation of GIS geometric objects [4].

2 Structure of Information-Modeling System

To integrate the IMS into various databases DB and DBMS, probable standardization of major data flows in the IMS as well as between the IMS and the database was studied. The general structure of the IMS and related data flows are presented in Figure 1.

The data exchange between the IMS and the integrated database (DB) involving spatial-distributed hydrological, meteorological and GIS data is carried out via "database (DB) — converter (import) — unified data — IMS internal environment" and "unified data — IMS internal environment — converter (export) — database (DB)". The converters are designed as separate units; to match the types and identifiers of data in the database and the IMS, a link is given, i.e. each parameter in the IMS corresponds to a particular database resource. The structure of internal data in the IMS supports the creation and expansion of the library of modeling blocks for solving numerous hydrological problems. Databases for solving particular tasks are specific depending

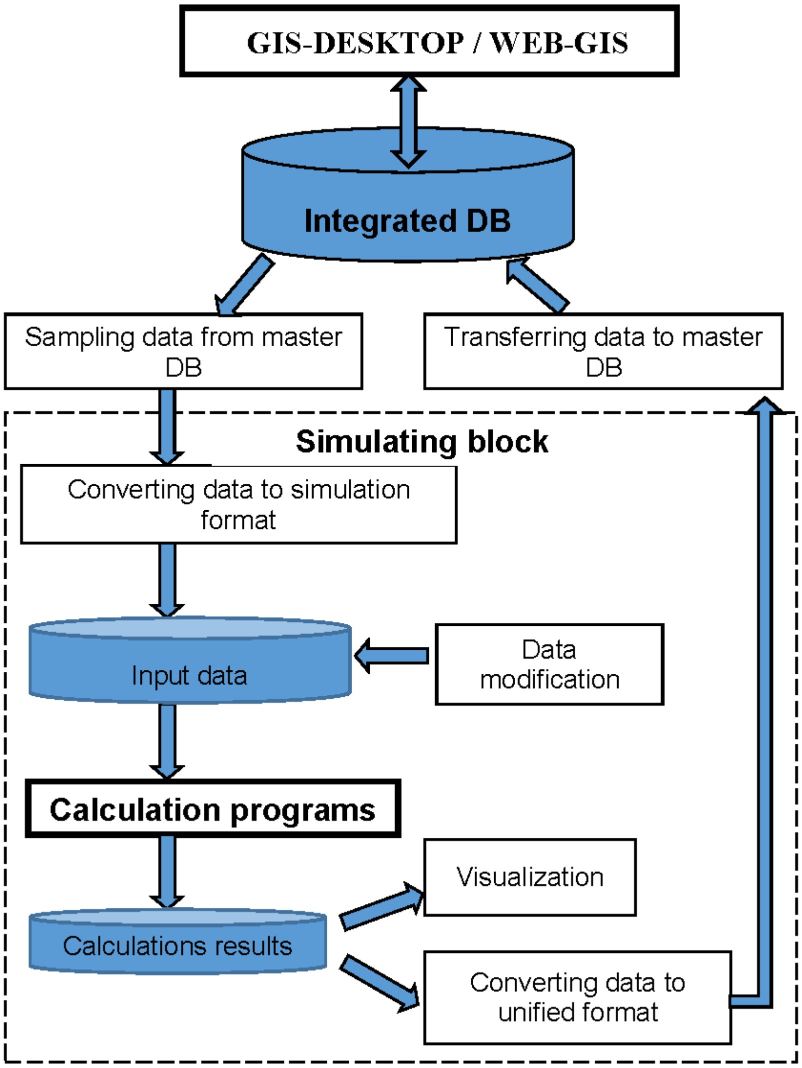


Figure 1: General structure of the IMS and related data flows

upon the subject of investigation. Here, the presence of a temporary component that has a marked impact on the data structure in the modeling complexes is critical. A large body of data is represented by time series of the measured and calculated values (discharge, water levels, meteorological parameters, etc.) assigned to certain spatial points (sites, weather stations, etc.).

Connecting GIS with modeling systems ensures the best solution to various hydrological tasks that is proved by successful national and international developments [5–9]. The concurrent use of GIS and databases is typical for expert support systems in water resources management [10]. The use of commercial GIS products with functional redundancy and high price as at the stage of data preparation as at calculation results visualization is a grave drawback of these systems. Moreover, the systems devotes much attention to the organization of observation data and calculation results.

The lack of standardized presentation of source and resultant data hinders the expansion of information-modeling complexes and the use of common data for solving some problems.

Under designing the IMS for flood prediction, the following tasks should be solved [11]:

- the creation of an integrated database with heterogeneous cartographic information, observations data on natural processes, including design and simulation data;
- the development of geoinformation systems providing the universal approach to the display of spatially distributed information within the IMS for floods prediction at large rivers in Siberia.

3 Development of integrated database to solve hydrological problems

The complexity and heterogeneity of tasks determine nontrivial approaches to the development of the DB structure. In the unified system, it is desirable to store the simulation and the in situ observation data, the online instrumentation data, the data on different periods of averaging, as well as the data associated with different natural environments. The universal schema of such a base with the unified extensible system of dictionaries allows us to collect and summarize the most important data on the subject area of research and to use them for further analysis, calculations, and presentation of results.

When developing the data model, we used a common international data format CUAHSI as a prototype [10]. A standardized data storage schema facilitates the analysis of information from disparate sources both within a single study area or various hydrologic objects.

Hydrologic observations are identified by the following fundamental characteristics (Figure 2):

- the location at which the observations were made (space);
- the date and time at which the observations were made (time);

- the type and value of a variable that was observed (ID and value).

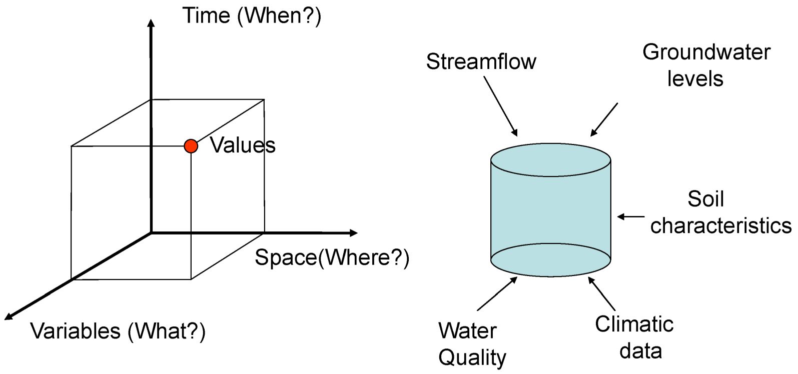


Figure 2: Schema of observational data description

In addition to these fundamental characteristics, there are many other distinctive attributes, which specify observational data. For example, the location of the observation, in addition to specifying the latitude and longitude coordinates, can be found in the local coordinate system, and can be accompanied by a text string. Other attributes can provide important context in interpreting the observational data, preserving the information about search conditions, methods of data processing as well as the organization that collected the data and data sources. Table 1 presents the general attributes associated with a point observation.

A conceptual model of the database (DB) presented in Figure 3 allows us to estimate the composition and quantity of the DB tables and their relationship as well.

In design of relational databases, such approach to data structuring is called a “star” schema. The data model consists of two types of tables: one fact table — the “star” center — and several dimension tables according to the number of dimensions in the data model — the “star” rays [12]. The main element in this schema is the fact table and numerous dimension tables. The fact table usually contains information about objects or events, the totality of which will be further addressed. Dimension tables contain persistent or seldom-changed data. They also contain at least one descriptive field (usually a variable name) and, as a rule, an integer key field to identify this variable. If the measurement corresponding to the table contains a hierarchy, then this table may also contain fields pointing to the variable “parent” in this hierarchy. Each dimension table should be in a “one–to many” relationship with a fact table.

In our database, the “data values” table acts as a “fact table”. Other tables presented in Figure 3, serve as the “dimension tables” or dictionaries.

The availability of ancillary dictionaries allows us to store the data of different kind (“raw”, that passed quality control, processed, etc.), the data of regular monitoring and single observations, the data on different periods of observation (from instant to average annual data), the data associated with any component of natural environment. A relational database with rows (records) that characterize a point observation provides maximum flexibility in the data analysis due to the possibility of their multicriteria selection and grouping.

Table 1: Attributes associated with an observation

Attribute	Definition
Variable Name	The name of the quantity that the data value represents
Data Value	The observation value itself
Accuracy	Quantification of the measurement accuracy
Date and Time	The date and time of observation
Location	The location of point (object) at which the observation was made
Units	The units and unit type
Interval	The interval over which each observation was made (averaged)
Offset	Distance from a reference point to the location at which the observation was made
Offset Type/ Reference Point	The reference point from which the offset is measured at the measurement location (e.g. annual average water level)
Data Type	The kind of quantity being measured (e.g. continuous, minimum, maximum, or cumulative measurement)
Censoring	An indication of censoring in sampling (e.g. lower than the detection limit, or <0.01)
Medium	The medium in which the sample was collected (water, air, soil, etc.)
Value Category	Characteristics of a value presented (an actual measurement, a calculated value, the result of a model simulation)
Analysis Method	An indication of what analysis method was used
Data Source	The organization that carried out measurements
Organization	The organization providing the data
Comments	Comments on the data quality that affect the way the data are used or interpreted

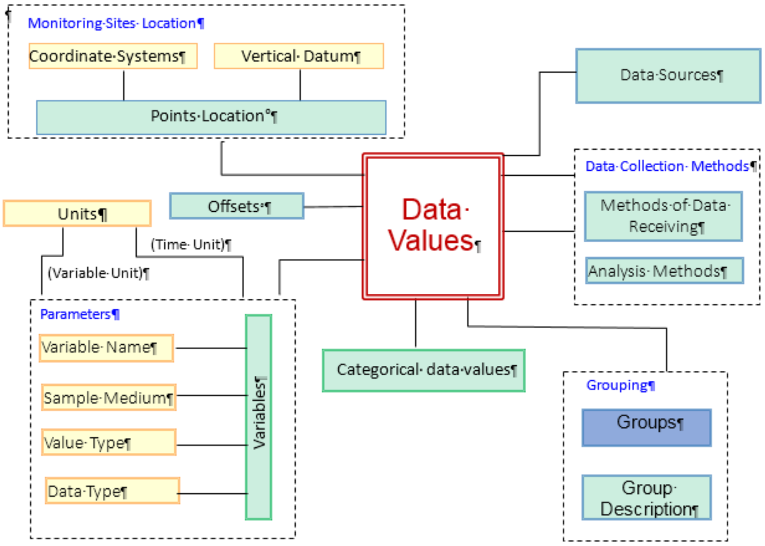


Figure 3: Conceptual database model

The availability of ancillary dictionaries allows us to store the data of different kind (“raw”, that passed quality control, processed, etc.), the data of regular monitoring and single observations, the data of different periods of observation (from instant to average annual data), the data associated with any component of natural environment. A relational database with rows (records) that characterize a point observation provides maximum flexibility in the data analysis due to the possibility of their multicriteria selection and grouping.

The observation data model is independent of the geographical representation of the site locations. The geographic location of sites is specified through the Latitude, Longitude and Elevation information stored in the Point Location table. The coordinates are specified in geographic or projection coordinate system, or by another method specific for the problem being solved. Each observation point has a unique identifier, which can be logically linked to one or more objects in the GIS data model. For example, a one–to one relationship between sites within the observation data model and meteo points on the vector layer. Such a relation between the observation sites and GIS objects is general in nature and carries no information about the structure and values of the GIS data. This architecture allows the observation data model to interact with any geographic data model that contains information about the location of observation sites (Figure 4).

Most hydrological processes vary in space and time. The choice of time and spatial scale of the information used plays an important role in building the IMS. Information with a scale larger than required, becomes “noisy”. Information with a scale smaller than required, is not representative or insignificant [9]. In [13], the scale of cartographic data is discussed.

The proposed structure of the cartographic database is as follows:

- a small-scale topographic base (vector), including hypsometry and hydrography (1:500000, 1:200000);

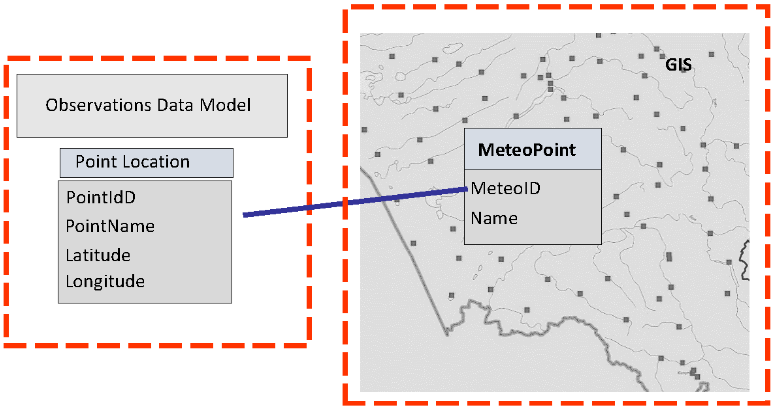


Figure 4: Schema of relationship between observations database and geospatial data

- a set of georeferenced raster coverages of 1:100000, 1:50000, 1:25000 scale and larger for the study area, covering it, if possible, completely;
- vector layers of hydrography and hypsometry of 1:25000 scale and larger for individual sites of the study area.

4 Development of geoinformation component of IMS

Geoinformation systems (GIS) is an integral component of modern systems of flood forecasting. GIS is used during the preparation of initial data for modeling and the analysis of the results of forecasting the development of hydrological situation on the rivers to determine the potential socio-economic effects due to the dangerous hydrological events.

Geoinformation component of IMS is intended for the following:

- collection, storage and graphical visualization of spatial data and related information about objects;
- formation of cartographic representation of assessment of flood hazard for the territory;
- graphical representation of results of calculations performed by modeling units;
- management of GIS objects and layers (publications, grouping, access, display of layers and attribute information).

In recent years there has been a trend of using web GIS, which allows to view and analyze spatial data with web browsers.

A typical web GIS includes three functional components: presentation service (client interface), application logic service (server application) and data service [14].

The architecture of this GIS application looks like the following (Figure 5):

- client;

- web-server of GIS applications;
- map-server;
- spatial data server.

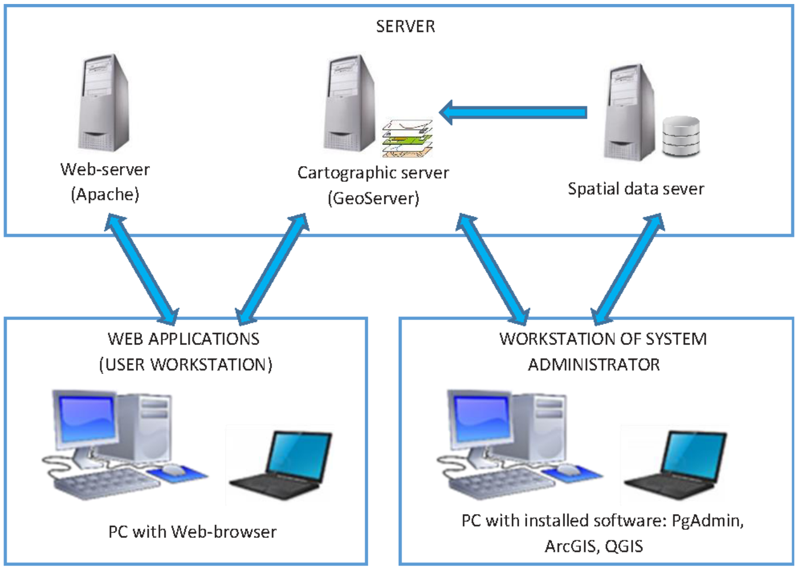


Figure 5: Web-GIS architecture

Presentation service (Web-server of GIS applications) is responsible for the interaction between application and user.

Map server implements a basic set of GIS functions, which are combined in software components that provide their own methods and properties to be used. The data server implemented through a database management system, enables the interaction between application and data.

This approach allows a client PC to use quite simple applications, including standard Internet browsers, and to perform most of calculations on the server. As far as the calculations are performed on the server, the amount of transmitted data reduces significantly since the user receives only the final product of the query processing but not all the data required for its execution.

To solve the problems of flood forecast on major rivers of Siberia, IWEP SB RAS creates a Web-server GIS application "Flood zones".

The main components of Web-GIS "Flood zones" are:

- web-server of GIS applications including "Flood zones" GIS;
- map-server GeoServer;
- DBMS PostgreSQL with spatial extension PostGIS.

Web-server of GIS applications uses the Apache HTTP-server.

"Flood zones" GIS, being a part of Web-server of GIS applications, is developed in Javascript using the OpenLayers 3 and ExtJS libraries [15, 16]. The interaction of application with GeoServer is implemented by AJAX-queries [17]. GeoServer provides the publication of spatial data as web services as well as caching and display style setting.

PostgreSQL is used for spatial data storage and management.

According to the scheme in Figure 1, the calculations to determine the flood area are performed by simulation unit that implements 1D - and 2D - hydrodynamic models [1–3].

For the parameterization and adaptation of hydrodynamic models, the following data from the integrated database are called for:

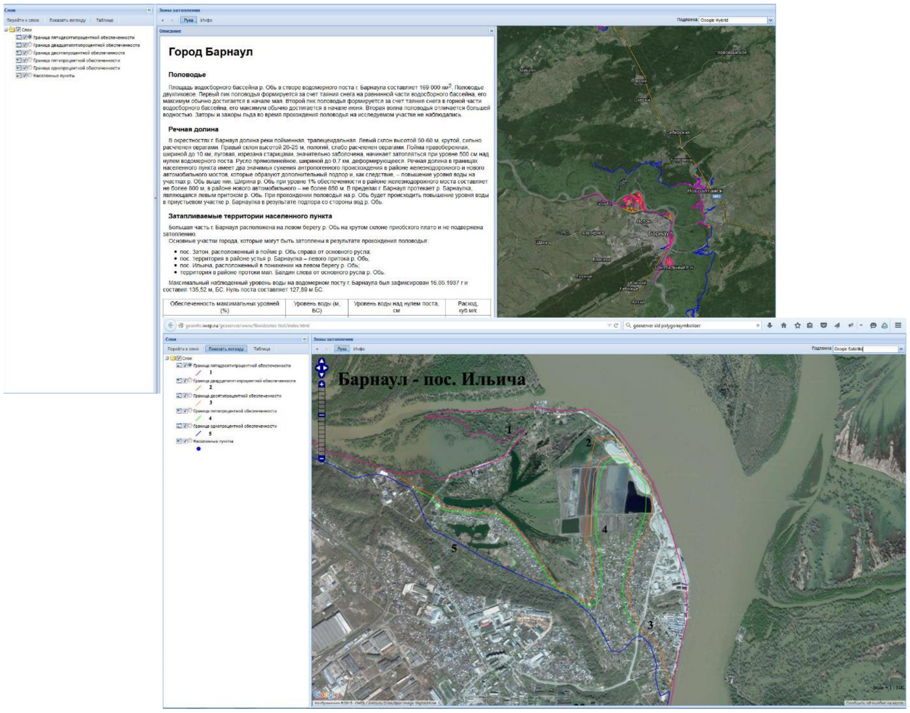
- river bed morphometry and floodplain relief;
- river bed and floodplain roughness;
- series of water level and discharge at hydrological sections of the studied river;
- data on meteorological conditions;
- data to optimize and verify model parameters (satellite images of overflowing, high-water marks, etc.).

The quality of the original data on river bed and floodplain relief is the most important factor influencing the final modeling of flood zones with the use of models of river hydraulics [18, 19].

The use of topographic maps of 1:25000 scale and larger to construct the floodplain relief as well as the data of the river bed instrumental survey allows the adequate description of processes of the floodplain inundation. A visual display of simulated results using Web browser enabled better forecast of water level in the rivers of the Upper Ob and Lena basins to reduce social and economic risks from dangerous floods (Figures 6,7).

Conclusions

- The structure of problem-oriented GIS system to forecast floods in the basins of major rivers of Siberia is developed.
- The structure of the database that contains real-time and historical hydrological and weather data on the hydrological situation is defined; a system of directories allowing to store in the database the data of different quality (“raw”, clean and processed data, etc.), regular monitoring data and the data of single observations, the data of the different periods of observation (from instant to average annual data), the data relating to any characteristics of the studied processes is defined.
- The tools for integrated processing of heterogeneous map data, observational data, calculations and simulation, which provide the possibility to solve the important hydrological problems of flood forecast and evaluation, are created.



Город Барнаул

Половодье

Площадь водосборного бассейна р. Обь в створе водомерного поста г. Барнаула составляет 169 000 км^2. Половодье двухпиковое. Первый пик половодья формируется за счет таяния снега на равнинной части водосборного бассейна, его максимум обычно достигается в начале мая. Второй пик половодья формируется за счет таяния снега в горной части водосборного бассейна, его максимум обычно достигается в начале июня. Вторая волна половодья отличается большей водностью. Заторы и зажоры льда во время прохождения половодья на исследуемом участке не наблюдались.

Речная долина

В окрестностях г. Барнаул долина реки пойменная, трапецеидальная. Левый склон высотой 50-60 м, крутой, сильно расчленен оврагами. Правый склон высотой 20-25 м, пологий, слабо расчленен оврагами. Пойма правобережная, шириной до 10 км, луговая, изрезана старицами, значительно заболочена, начинает затопляться при уровне 500 см над нулем водомерного поста. Русло прямолинейное, шириной до 0,7 км, деформирующееся. Речная долина в границах населенного пункта имеет два значимых сужения антропогенного происхождения в районе железнодорожного и нового автомобильного мостов, которые образуют дополнительный подпор и, как следствие, – повышение уровня воды на участках р. Обь выше них. Ширина р. Обь при уровне 1% обеспеченности в районе железнодорожного моста составляет не более 800 м, в районе нового автомобильного – не более 850 м. В пределах г. Барнаул протекает р. Барнаулка, являющаяся левым притоком р. Обь. При прохождении половодья на р. Обь будет происходить повышение уровня воды в приустьевом участке р. Барнаулка в результате подпора со стороны вод р. Обь.

Затапливаемые территории населенного пункта

Большая часть г. Барнаул расположена на левом берегу р. Обь на крутом склоне приобского плато и не подвержена затоплению.

Основные участки города, которые могут быть затоплены в результате прохождения половодья:

- пос. Затон, расположенный в пойме р. Обь справа от основного русла;
- пос. территория в районе устья р. Барнаулка – левого притока р. Обь;
- пос. Ильича, расположенный в понижении на левом берегу р. Обь;
- территория в районе протоки мал. Балдин слева от основного русла р. Обь.

Максимальный наблюденный уровень воды на водомерном посту г. Барнаула был зафиксирован 16.05.1937 г и составил 135,52 м, БС. Нуль поста составляет 127,89 м БС.

Обеспеченность максимальных уровней (%)	Уровень воды (м, БС)	Уровень воды над нулем поста, см	Расход, куб.м/с

Figure 6: Flood zones of 1%, 3%, 5% probability (city of Barnaul) http://geoinfo.iwep.ru/geoserver/www/floodzones-test/index.html

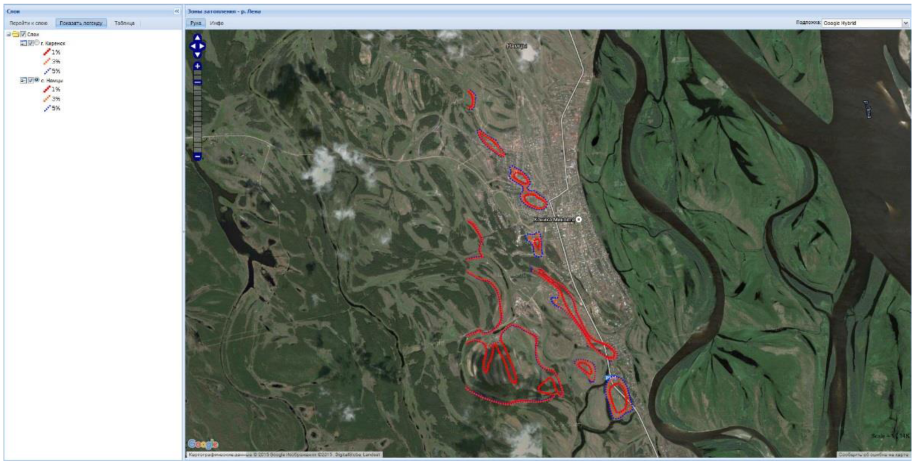


Figure 7: Flood zones of 1%, 3%, 5% probability (selo Namtsy, Respublika Sakha (Yakutiya)) http://geoinfo.iwep.ru/geoserver/www/floodzones-test/index.html

Acknowledgement

The work was supported by the projects of SB RAS 0383-2014-0006 “Development of information and analytical support for the study of water and environmental processes in water-bodies, flows and watersheds of Siberia” and 0383-2014-0001 “Study of river flow formation and development of information-modeling systems for operational forecasting of dangerous hydrological conditions for large river systems of Siberia” .

References

[1] Vasil'ev, O.F. *Creation of systems for real-time forecast of floods*, Bulletin of RAS, 2012, No.3, pp. 237-242 (in Russian).

[2] Zinov'ev, A.T., Koshelev, K.B., D'yachenko A.V., Kolomeitsev A.A. *Extreme rainfall flood of 2014 in the Upper Ob basin: causes, forecast and field observations*, Water economy of Russia. Problems, technologies, management, 2015, No. 6, pp. 93-104 (in Russian).

[3] Zinov'ev, A.T., Koshelev, K.B. *Modeling of floodplain flooding in large rivers with complex morphometry of the channel and floodplain*, Water economy of Russia. Problems, technologies, management, 2013, No. 6, pp. 17-31 (in Russian).

[4] Vasil'ev, O.F., Zinov'ev, A.T., Koshelev, K.B. Kudishin, A.V., Lovtskaya, O.V., Ovchinnikova T.E., and Semchukov A.N. *Development of information-modeling systems for operational forecasting of dangerous hydrological conditions for large river systems of Siberia (the Upper Ob as a case study)*, Water-related hazard: danger, forecasting, management and prevention of threats: Proceedings of the All-Russian scientific conference, October 7-12, 2013. - Novocherkassk: LIC, 2013, pp. 41-47 (in Russian).

[5] Ackerman, C.T., Evans, T.A., and Brunner, G.W. *HEC-GeoRAS: Linking GIS to Hydraulic Analysis using ARC/INFO and HEC-RAS*, Hydrologic and Hydraulic Modeling Support with Geographical Information Systems / In Maidment, D., and Djokic, D. (eds.), New York: ESRI Press, 2000, pp.155–176.

[6] *MIKE 21C-2D river hydraulics and morphology.* Available at: http://www.dhisoftware.com/Products/WaterResources/MIKE21C.aspx.

[7] *Delft3D-FLOW Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments*, *User Manual*, 2011 (Version 3.15).

[8] Motovilov Yu.G. *Current status and prospect of the hydrological modelling for the Russian river basins using ECOMAG package*, “Fundamental Problems of Water and Water Resources”: Proceedings of the 3rd All-Russian conference with international participation. Barnaul, 2010, pp. 530 – 532 (in Russian).

[9] Loucks Daniel P. and van Beek Eelco *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications.* Edited by M.V. Seliverstova; translation from English - A.V. Stepanov and others, Yustitsinform, Moskow, 2009 (in Russian).

[10] Maidment, D. R., ed. *CUAHSI Hydrologic Information System: Overview of Version 1.1.* Consortium of Universities for the Advancement of Hydrologic Science, Inc, 96 p. Available at: http://his.cuahsi.org/documents/HISOverview.pdf.

[11] Zinoviev A.T., Lovtskaya O.V., Baldakov N.A., Dyachenko A.V. *Geoinformation support of hydrologic problems*, Computational Technologies, Vol. 19, 2014, No. 3, pp. 14-26 (in Russian).

[12] Barsegjan A. A., Kuprijanov M. S, Stepanenko V. V., Holod I. I. *Methods and models for data analysis: OLAP u Data Mining*, BKhV-Peterburg, SPb., 2004 (in Russian).

[13] Lovtskaya O.V., Yakovchenko S.G., Zherelina I.V., Zhorov V.A. and Postnova I.S. *Providing water-management and hydrological calculations by GIS technology*, Sibirskiy ekologicheskiy zhurnal, 2005,No. 6, pp. 1013–1023 (in Russian).

[14] Chandra A.M., Ghosh S.K. *Remote sensing and geographical information system*, Tekhnosfera, Moskow, 2008 (in Russian).

[15] OpenLayers 3, (http://openlayers.org)

[16] *Sencha Ext JS* (https://www.sencha.com/products/extjs).

[17] *AJAX Tutorial* (http://www.w3schools.com/ajax/s).

[18] Borshch S. V., Samsonov T. E., Simonov T. E., L'vovskaya E. A. *Visualization of hydrological situation in the basins of major rivers by means of GIS-technologies*, Proceedings of the Hydrometeorological Research Center of the Russian Federation, Vol. 349, 2013, pp. 47-62 (in Russian).

[19] Werner M, Blazkova S, Petr J. *Spatially distributed observations in constraining inundation modelling uncertainties*, Hydrological Processes, Vol. 19, 2005, pp. 3081-3096.

Alexander Zinoviev,
Institute for Water and Environmental Problems, Siberian Branch of the Russian Academy of Sciences (IWEP SB RAS),
1, Molodezhnaya St., Barnaul, 656038, Altai Krai, Russian Federation,
Email: `zinoviev@iwep.ru`,
Olga Lovtskaya,
Institute for Water and Environmental Problems, Siberian Branch of the Russian Academy of Sciences (IWEP SB RAS),
1, Molodezhnaya St., Barnaul, 656038, Altai Krai, Russian Federation,
Email: `lov@iwep.ru`,
Nikita Baldakov,
Institute for Water and Environmental Problems, Siberian Branch of the Russian Academy of Sciences (IWEP SB RAS),
1, Molodezhnaya St., Barnaul, 656038, Altai Krai, Russian Federation,
Email: `nikita-baldakov@yandex.ru`

Received 12.09.2016, Accepted 15.09.2016