ISSN 0097-8078, Water Resources, 2018, Vol. 45, Suppl. 1, pp. S59–S66. © Pleiades Publishing, Ltd., 2018.

Information System to Support Regional Hydrological Monitoring and Forecasting

A. N. Bugaets[a, b, c, *], L. V. Gonchukov[a, c], O. V. Sokolov[a, b], B. I. Gartsman[b, c], and S. M. Krasnopeyev[c, d]

[a] Far Eastern Regional Hydrometeorological Research Institute, Vladivostok, 690091 Russia

[b] Water Problems Institute, Russian Academy of Sciences, Moscow, 119333 Russia

[c] Pacific Geographical Institute, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, 690041 Russia

[d] Far Eastern Federal University, Vladivostok, 690950 Russia

*e-mail: andreybugaets@yandex.ru

Received June 27, 2018

Abstract—This paper describes the conceptual framework and software components of the automated hydrological monitoring system (AHMS), developed as a part of the project aimed at recovery, modernization, and development of the hydrometeorological network and hydrological forecasting system in the Amur basin in Russia. AHMS information technology platform provides sustainable functioning of the observation network, data exchange (within regional hydrometeorological state agencies), and interaction with external information systems.

Keywords: information system, hydrological monitoring

DOI: 10.1134/S0097807818050329

INTRODUCTION

Efficient hydrological information monitoring systems require standardization of making observations with heterogeneous measurement equipment, unification of methods of storage of the ever-increasing body of information, processing and publication, and data exchange and dissemination. The UN World Meteorological Organization (WMO) formulated standards of making meteorological observations, their description, and international exchange of data quite a long time ago. In hydrology, the situation seems somewhat more difficult. The basic principles of describing hydrological processes, reflecting the link between the spatiotemporal dimension of instrumental observations and the interpretation of data received as a result of these observations, have been formulated only recently [7].

These principles found their practical application in the structure of a relational database (DB) developed by The Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI, www.cuahsi.org). It enables the description and storage of observation data from related sciences: meteorology, hydrology, hydrochemistry, etc. The infrastructure and data exchange format (OGC) WaterML developed by the Consortium made it possible to merge the academic and state components of the US hydrological and ecological monitoring services into one information network [3].

The rapid development in this field have been promoted by interdisciplinary scientific collaborations based on the common open information space and the adaptation of the information and modelling platforms to the standards of high-performance computing. One of the most successful examples is an open fully integrated platform to predict and manage the risks of extreme hydrometeorological phenomena created by the European Project of the Distributed Research Infrastructure for Hydro-Meteorology—DRIHM (www.drihm.eu) [2]. Based on the DRIHM framework, a wide range of hydrometeorological processes from microscale ones, happening on the slopes, up to mesoscale processes of interaction between the atmosphere and the underlying surface, have been modelled.

The development of Russian information technologies in terms of hydro-ecological monitoring is still lagging behind dramatically. Mostly the issues of a technical character were elaborated, connected with optimizing the composition and maintaining the efficiency of the observation network. The actual inaccessibility of the operational hydrometeorological data led to the technology development being dependent mostly on the commercial equipment and software vendors, rather than on the interaction with the scientific community. The implementation of the open distributed information systems (IS) for monitoring and prediction support will lead to economic benefits through optimizing the measurement network,

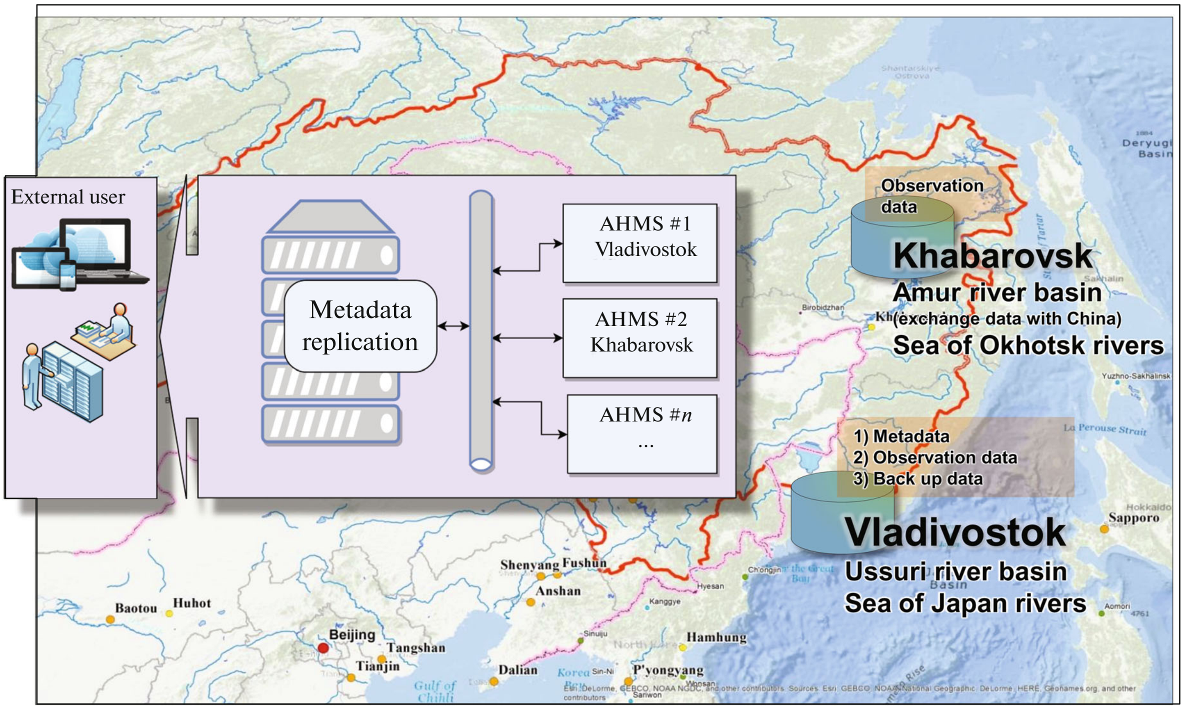

Fig. 1. General scheme of the distributed ASGM network in the Russian Far East region.

increasing the stuff effectiveness, and reducing the costs of research and engineering.

This paper describes the general principles and software components of the automated hydrological monitoring system (AHMS). It was designed within the project of restoration, modernization and development of the hydrometeorological network and the system of hydrological prediction in the Amur River basin for optimizing the operational work of the regional departments of the Russian Hydrometeorological Service (Roshydromet). The considered example includes the Primorskiy and Far East Administrations for Hydrometeorology and Environmental Monitoring located in the cities of Vladivostok and Khabarovsk.

THE GENERAL PRINCIPLES

AHMS general principles are based on the hydrological community experience and the open standards for creating software for managing operational hydrometeorological data of different types and aggregation levels. The main objectives of AHMS are providing information technology for secure functioning of the observation network of regional Roshydromet administration; the independence of individual instance of AHMS from the measuring equipment set-up; the standardization of storage and data exchange protocols; the scalability, i.e., the consolidation of local instances into a network; data fusion and interoperability with up-to-date developments in the field of monitoring and managing water resources. Special attention is given to the measurement quality control, as an essential condition of the reliability of predictive model operation results.

AHMS underlying structure has been implemented as a network of data sources separated on the local (low) and management (high) level nodes, integrated via web services (Fig. 1). The local nodes were organized on the territorial (or river-basin) principle. The high level (servers) controls metadata consistency on the low level (local network nodes), supports data fusion and information backup, as well as the internal and external communication (between users and systems).

IT infrastructure at a local node is responsible for data receiving, storage, and processing; quality control; visualization (GUI); GIS component, etc. All datasets (generated by sensor networks, different kind of simulation, samples, etc.) are organized and stored within a central database (DB) implemented on open-source object-relational database system PostgreSQL. The local software complex of AHMS is divided into the client and server parts. All software units cannot directly address each other; the interaction is realized only through addressing the DB by means of program access services (Fig. 2).

The server part consists of the database, processing services, and modules of automatic data control. The client part of AHMS, relying on the graphic interface of a user, provides viewing, editing, and critical control of the data, generating reports and bulletins of differ-

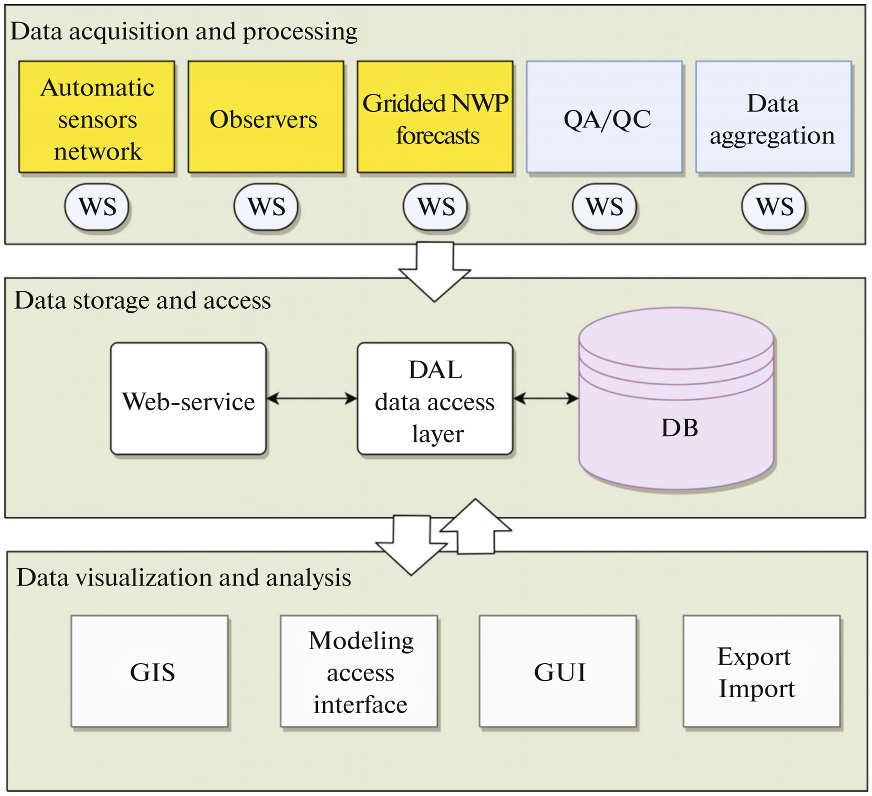

Fig. 2. Data exchange and interaction in a distributed ASGM network.

ent kinds, and import and export of the data into the formats of third-party developers (MS Office, ASCII, PDF, etc.). The web service of the local server provides the integrity, the consistency of the data, and the interaction between platforms. The coordination of metadata is achieved by replicating meta-database dictionaries of managing servers.

AHMS distributed data network, as an application of open Postgres RDMS, employs the well-known web and OGC standards and exchange formats to unify data access for most of the hydrological application. In particular, OpenMI (http://openmi.org) [6] wrapper for AHMS web-services create a framework where data and models could be executed in a single workflow [1, 4, 5].

The data exchange is organized in two levels. The first level provides the external interaction between all components of the modelling system using the OpenMI. The second level of integration, through OpenMI-wrapped SOAP web-services and the DAL (data access layer) library, provides the cross-platform interaction for operation with the distributed ASHM data-storage system. On the OpenMI level, the description of the variables uses international SI system only. When accessing the database web-service, the SI units were reduced to the terms of the ASHM database dictionaries.

THE DATABASE STRUCTURE

The database is intended to store the results of all types of measurements realized on the observation network, estimated characteristics and metadata vocabularies in the relevant, unambiguous, and universal representation. AHMS DB schema was adopted from CUAHSI Observation Data Model (http://his.cuahsi.org). It allows all types of hydrometeorological data to be unambiguously interpreted by sufficient auxiliary information (metadata) and provides traceable heritage from raw measurements to useable information. The tests conducted at the base of the Primorskoe Administration for Hydrometeorology and Environmental Monitoring in the period between 2010 and 2013 [3], demonstrated that such database structure corresponds to the accepted norms of storing all kinds of data and information received from the observation network.

AHMS extends CUAHSI ODM for saving the historical changes in observation site definition, warning events, multiple “qualifiers” for each observation, the original texts of messages and files, and the history of data values correction, and includes other improvement needed to meet the requirements of the operational mode and every-day practical work.

Additional tables for observation site (gauging or meteorological stations, etc.) contain historical extremes, typical values, acceptable intensities of change in measured values, etc. An important feature of such data is their binding to a certain period of time, which allows all the history of their changes to be saved and used for a wide range of tasks, including quality control, analysis, calculations (for instance, with the help of rating curves $Q = f(H)$ or coefficients), and providing information.

The history of data availability and processing is supported by the DB scheme through entering the term "a data entity," which allows one to choose the only relevant value from a number of records associated with the same place and time of observation, variable, measurement method and QCL that were received as a result of systematically getting new specifying messages from observation platforms along carrying out the procedures of automatic and manual control.

In the current scheme, there is a significant increase in the role and scope of application of the so-called qualifiers (in the ODM descriptions accompanying the suspected data). First, it has become possible to group the necessary amount of qualifiers from the corresponding DB table with every single change in the value. This enables describing every possible issue, emerging in the operational work and subsequent post-processing.

To accelerate data selection query, qualifiers are divided into two types: critical and descriptive. The critical qualifiers include the cases of exceeding the threshold of sensitivity established for the sensor or the limits of changes in the given variable, going beyond the permissible limits of changes in a value over some time interval, etc. The descriptive qualifiers are associated with the data related to actions not directly related to carrying out the measurements, for instance, data that came in with a delay (behind schedule), the availability of duplicate values, etc. Additionally, every record in the data value table is marked by a flag, pointing to the possibility of using the given value in calculations (including derivatives of variables, expenses, different aggregated characteristics, etc.) and predictions. By default, the permission of usage is granted in case of absence of critical qualifiers, while descriptive ones can be present.

AHMS, as well as ODM, is a point data observation schema, but AHMS metadata can describe external data sources of gridded data like a result of numerical weather-prediction models (GFS, WRF, etc.). These datasets are usually stored as remote file structures with http or ftp access. A special designed web-service allows getting this data in particular grid nodes, interpolated into fixed points (meteostation) or inside an extent frame. In the operative forecast mode in the case of absence of the required meteorological forecast data, ASHM OpenMI component can be tuned to try to import these data directly from the grid files of near-surface meteorological characteristics.

As a result of the above modifications, the DB scheme of AHMS makes it possible to store and effectively manage all types of data received from the observation network of Roshydromet and to follow the history of changes in the characteristics of observation points, and applies the methods of measurements and calculations, criteria of data control, visualization and editing.

DATA RECEIVING AND DECODING

The block of receiving and decoding data gathers the information from various sources and casts it for the DB update. This is a set of Microsoft Windows-services (WS), each one associated with a certain type of data source: the automatic hydrological or meteorological complex, hydrological model, external DB, etc. The output of all decoding WS connects to the input of the main AHMS data access library. Thus, any measuring equipment or data source is AHMS-compliant. Besides numerical data, the modules can save in DB texts, messages, files, etc. Every decoding Win-service keeps event log, which gets updated while receiving and analyzing data.

Using the service allowing for data access to the central DB, all modules can conduct a preliminary quality control to determine errors generated by measuring equipment or malfunctions in data communication network, because of manual information encoding, and other human-related factors. At this stage, such control is implemented as search for the typical data errors, e.g., using the criteria of the data value falling within a specified interval in accordance with the variable or the measurement device. The occurrence of input data outside the limits is marked by a qualifier, signaling suspicion that the measurement may be erroneous. All limits and parameters that are necessary for these procedures are stored in the DB, which allow the system to respond flexibly to the changes in equipment and system software settings.

DATA QUALITY CONTROL

One of the main objectives of AHMS is data control. The general principle adopted here is the background automatic data control (ADC). ADC provides the operational staff the diagnostic information about devices and observation network sensors in the form of QCL values and accompanying qualifiers. These facilitate finding and correcting errors and omissions in data on the second stage, where following manual control can effectively eliminate all problems identified in the data.

The list of QCLs and their semantic content was borrowed from the CUAHSI ODM; however, the rules for their designation were updated. In the operational workflow, there are three levels of control used:

—"0"—"raw" data;

—“1”—data reviewed by critical control and/or confirmed by an operator;

—“2” —derived values.

Beginning from the level “3” and higher, QCLs must reflect the degree of involvement the data into scientific research and engineering projects. While designating control levels, the ADC procedure uses the following rules:

—all data measured on the network are recorded in the DB with QCL = “0” and are saved without changes;

—derived data inherit the control level of their original data (for example, the daily average level, calculated from hourly values);

—the main task of ADC is to transfer raw data from QCL = “0” to the level “1,” i.e., to create new records with identifying problematic values on the basis of established criteria and attributing corresponding qualifiers to them. Each criterion is allied with a site, a variable, and a period of time when this criterion can be valid. The criterion has two thresholds for a data value and the way of its applying depends on the type of the criterion. So far, there are three types of criteria for value control range;

—maximum value change rate;

—difference between sensor and manual measurement.

All criteria of control available for the given site and variable are applied for every value. Each rating curve $Q = f(H)$ is validated for water stage and discharge. In order to optimize control procedures, apart from the value of QCL, the flag field of ADC was added to the data table. It can take values: “0”—without control, “1”—successful, “2”—there are errors, “3”—corrected or confirmed manually by an authorized staff member. The last value means that ADC procedures cannot be applied again to such data.

The records that have not passed automatic control (QCL “1,” ADC flag “2”), need to be analyzed manually, and they can be subsequently accepted (confirmed), edited, or rejected. To facilitate the efficiency of operational data control and correction, such records are specially highlighted in all AHMS tools of visualization and analysis (graphs, tables, reports). Editing tools allow getting all information and metadata from the DB for the indicated records, retrace and associate the derived values to the sources.

The standard editing functionality comprises the following set of instruments:

—highlighting, adding, and deleting records;

—copying and pasting;

—gap-filling;

—correcting (to add/multiply by the constant value);

—adding qualifiers.

The simple editing tools are self-explanatory and hardly require any further explanations. Manual adding of qualifiers is necessary only if automatically controlled data are ambiguous or completely lack in credibility. By default, linear interpolation is applied as the function of gap-filling; however, it is not always suitable method. A more reliable way is the use of imitation models available for the given object or reliable regression relations, which can be connected using OpenMI standard.

DATA AGGREGATION AND THE CALCULATION OF DERIVED VALUES

The final stage of preparing usable information from the raw measurement is data aggregation. Raw data are converted to temporal intervals applied in practice. Daily, decade and monthly values are being derived, which could be both average and total characteristics obtained in points or by basin averaging, etc. Variable from DB metadata vocabulary with corresponding time units of derived values is used. The spatial support of the derived characteristic is implicitly defined in the description of aggregation algorithm or measurement method.

DB scheme allowed storing groups of unique identifiers for binding each derived value with the used original information or measurements. While computing aggregated values (sums, average, etc.), the initial data are variables of the same type. For instance, for daily discharges, the initial data must be discharges over a shorter period of discretization—hourly, semidiurnal, etc., which, in turn, can be received after measuring the water level using a rating curve.

The functions of the module of data aggregation are called automatically, once the critical control procedures are finished, or it is done manually via the graphical user interface tools. By default, only data with maximal QCL are used. The hourly time interval is the basic for data aggregation of automatic sensors, and the diurnal interval is such for standard measurement. If an observation has critical qualifiers, the corresponding record is marked by the sign of rejection for the following use. These data will be ignored by aggregation algorithms until the problem is resolved, e.g., as a result of an operator intervention.

THE GRAPHICAL USER INTERFACE

AHMS graphical user interface (GUI) is designed for managing local server, data editing and publication using standard forms of tables and graphs used in the operational practice of Roshydromet. The GUI generally follows the design of Windows Explorer application (Fig. 3). The working space of the main window is split into navigational, working, and instrumental panels.

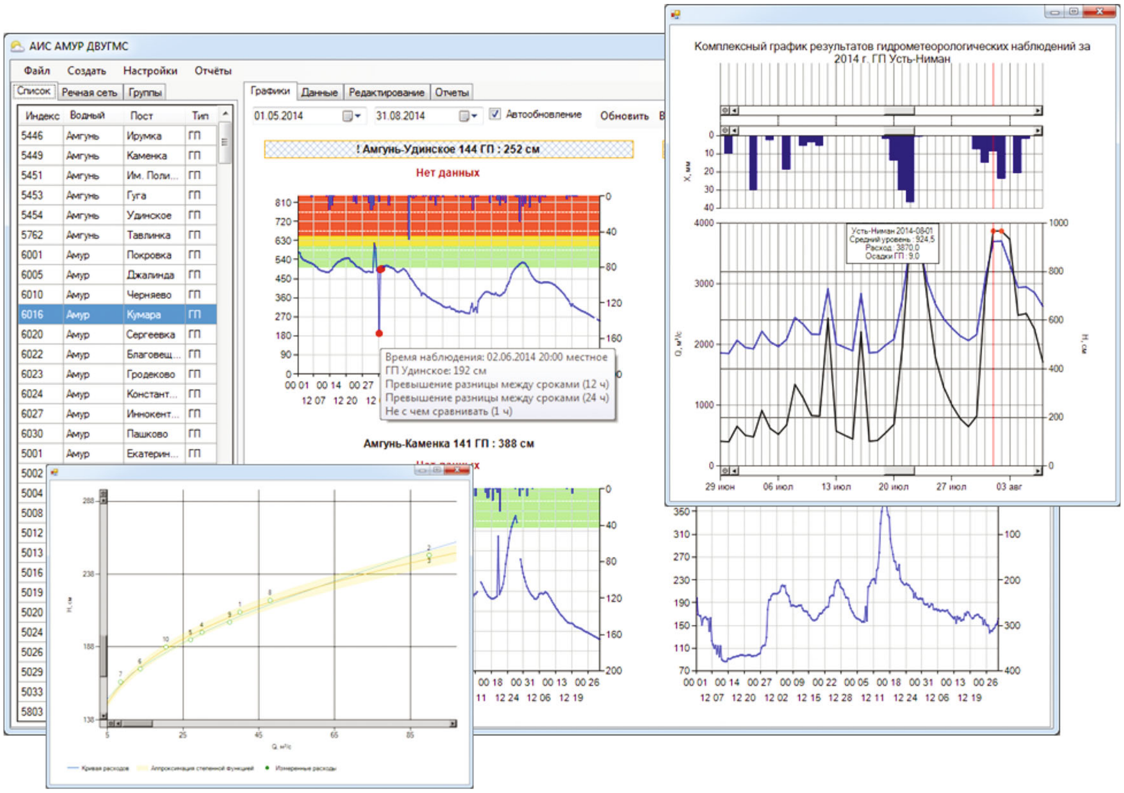

Fig. 3. Mosaic panel of the graphic user interface—tables and diagrams (Russian is used).

DB navigation panel elements of the observation network are linked to the hydrographic network in a tree-like or tabular form. The functionality of the application allows editing attributes of the observation network sites, adding new ones, and georeferencing and linking their location to the hydrographic network. The module includes instruments for viewing and editing observation data: different variants of the graphs combining series of water levels and discharges, air and water temperature, precipitation, ice thickness, and snow depth on the ice, etc. In addition, there is a possibility to display the long-term graphs of level and discharge, hyetographs, and the rating curves $Q = f(H)$. The graphs reflect the results of measurements regarding their norms and warning levels (if any). In the monitoring mode, GUI allows the operator to control hydrometeorological situation of several sites simultaneously via the multipanel mosaic.

The main form of tabular data representation is the modified existing standard operational journal, including the results of observing water levels, ice phenomena, ice thickness and snow depth on the ice, discharges, and precipitation. There are tabular forms of summary tables, similar to the widely accepted standards: for water equivalent of snow cover, daily average temperatures, results of the analysis of high water and floods (levels, discharges, precipitation, the dates of exceeding critical levels, and the probability of exceeding the hydrological value). To facilitate the work with historical data, there are exchange tools (import/export) with the software approved for use by Roshydromet provided.

The basic scheme of processing operational information consists of 5 main points, describing the sequence of everyday work of the hydrological forecasts department:

(1) integrity control of data coming from the observation network;

(2) critical data control;

(3) data analysis;

(4) forecast;

(5) dissemination of information to the end-users.

When the AHMS application starts, the implementation of the measuring programme is checked for every gauging station and measurement platform registered on the system. All detected failures are reported as tables specifying the site index and a possible reason (equipment malfunctions, message time delay, offline, low battery, etc.). The functionality of graphical and tabular forms allow the most problematic sites to be effectively discovered, fixing potentially incorrect information (with critical qualifiers), and also, in case of necessity, grant consistent access for editing records, indicating metadata and providing access to the initial measured values.

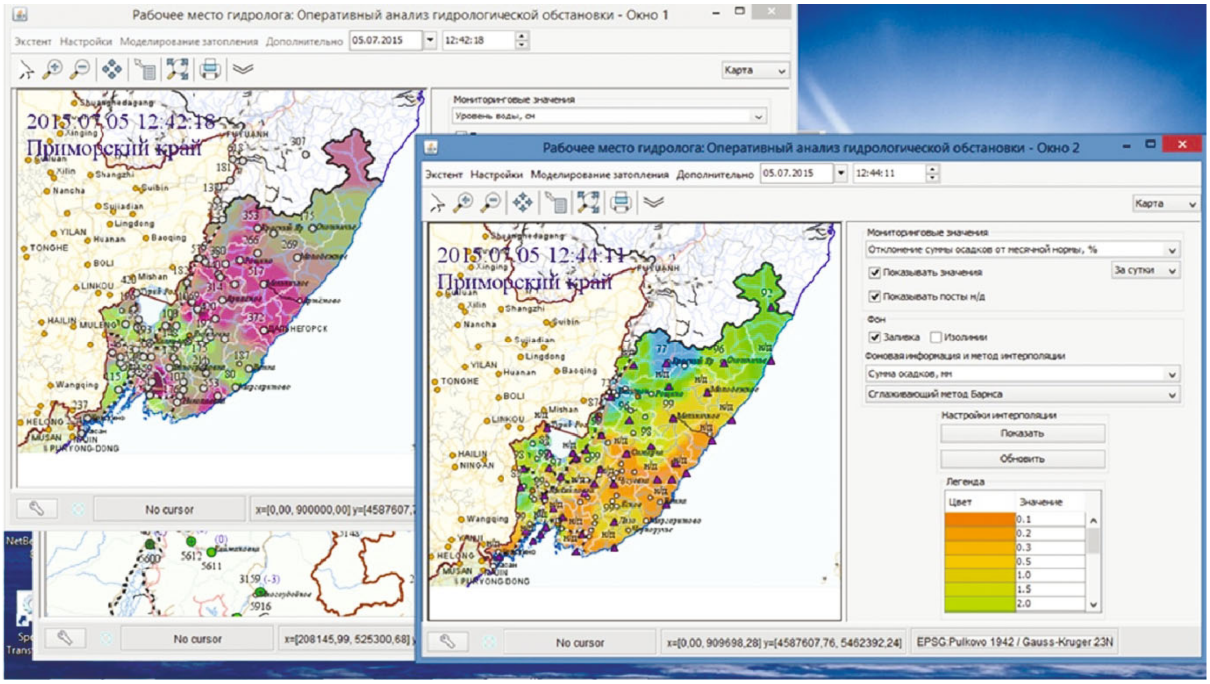

Fig. 4. Mosaic panel of the graphic user interface—the maps of GIS component (Russian is used).

Creating of finished products (reports, forecasts, maps, etc.) and their dissemination according to the predetermined schedules marks the end of the of operational data processing phase. The outputs include standard and specialized hydrological bulletins, special cartographic materials, observation data summaries for transferring to users, databases (including internal use).

To enhance the efficiency of AHMS operational work, a module was developed that carries out spatial interpretation of hydrological and thematic information, prognostic results of hydrological and hydrodynamic modelling with inundation area and water velocity maps (Fig. 4). The module includes the following technological blocks:

—visual monitoring of the current hydrometeorological situation in real time with the possibility to manage a number of hydrometeorological characteristics on the basis of terrestrial network data, information coming from the world data centers, numerical weather prediction models, and satellite monitoring data;

—generating a predetermined set of cartographic representations of hydrometeorological characteristics on auxiliary panels, as well as forecasting fields of meteorological elements, prognostic information of hydrological models and warning about the risk of emergency;

—generating presentational cartographic materials for preparing reports, providing bulletins, profiles and other consumer products in a semiautomatic mode;

—generating large-scale cartographic representations of predetermined areas on auxiliary panels and delivering full operational and prognostic information about a hydrometeorological situation, as well as the results of modelling relevant and expected areas of inundation.

The specialized web-service generates hourly automatic publication of cartographic layers with actual operative hydrometeorological information coming from observation network: water level and discharge, precipitation, air temperature, ice thickness, snow depth, and water equivalent of snow cover.

CONCLUSIONS

This paper discusses the motivation, the main idea, and the structure and functionality of the AHMS as an IT-based supporting tool for hydrological monitoring, data management, and modelling. The modernization of hydrometeorological monitoring systems requires standardization of observations on heterogeneous measurement platforms, as well as unification of methods of data storage, processing, and provision, without which it is impossible to use an ever-increasing amount of information effectively.

The presented project is the result of summarizing the best international experience and uses the open standards for creating the system that makes it possible to effectively manage the complex hydrometeorological data of various types. The AHMS distributed data network, application of open Postgres RDMS, applying of well-known web and OGC standards and exchange formats, unifies data access for most of the hydrological application and allows the data and models to be executed as a single workflow.

The general scientific and technical concept of the project facilitates the extensibility and scalability of the proposed solution, as well as its compatibility with the main functionally-interrelated developments of the top research institutes and commercial vendors in the field of monitoring, modelling, and managing water resources, which is particularly relevant for transboundary territories. The informational and technological foundation of AHMS allows the current needs of monitoring and forecasting services to be fully met, and also provides sustainable functioning and development of the observation network and informational and technological base of territorial administrations for an extended period of time.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation, project nos. 17-77-30006 and 16-17-00105, and the Russian Foundation for Basic Research, project no. 16-05-00864.

REFERENCES

1. Bugaets, A.N., Gartsman, B.I., Gelfan, A.N., Motovilov, Yu.G., Sokolov, O.V., Gonchukov, L.V., Kalugin, A.S., Moreido, V.M., Suchilina, Z.A., and Fingert, E.A., The Integrated System of Hydrological Forecasting in the Ussuri River basin based on the ECOMAG model, *Geosciences*, 2018, vol. 8, no. 1, p. 5.
2. Clematis, A., D'Agostino, D., Danovaro, E., Galizia, A., Quarati, A., Parodi, A., Rebora, N., Bedrina, T., Kranzlmüller, D., Schiffers, M., Jagers, B., Harpham, Q., and Cros, P.H., DRIHM: Distributed Research Infrastructure for Hydro-Meteorology, in *7th International Conference on System of Systems Engineering (SOSE2012) Proceedings*, 2012, pp. 149–154.
3. Bugaets, A.N., Gartsman, B.I., Krasnopeyev, S.A., and Bugaets, N.D., An experience of updated hydrological network data processing using the CUAHSI HIS ODM data management system, *Russ. Meteorol. Hydrol.*, 2013, vol. 38, no. 5, p. 359–366.
4. Bugaets, A.N., and Gonchukov, L.V., OpenMI 2.0 based WRF-SWAT models integration, in *1st European Fully Coupled Atmospheric-Hydrological Modeling and WRF-Hydro Users Workshop. Book of Abstracts*, University of Calabria, Italy, June 11–13, 2014.
5. Bugaets, A.N., and Gonchukov, L.V., Application of WRF-SWAT OpenMI 2.0 based models integration for real time hydrological modelling and forecasting, in *Geophys. Res. Abstr.*, 2014, vol. 16, EGU2014-4712.
6. Bugaets, A.N., Using the OpenMI standard for developing integrated systems of hydrological modeling, *Russ. Meteorol. Hydrol.*, 2014, vol. 39, no. 7, pp. 498–506.
7. Blöschl, G., and Sivapalan, M., Scale issues in hydrological modelling: A review, *Hydrol. Processes*, 1995, vol. 9, nos. 3–4, pp. 251–290.