Geoinformation Technologies and Software for Operational Assessment of Air Pollution

Alexey A. Kadochnikov[1,2], Alexey V. Tokarev[1,2], Oleg E. Yakubailik[1,2,3]

[1] Federal Research Center Krasnoyarsk Science Center of the SB RAS, Krasnoyarsk, Russia
[2] Institute of Computational Modelling SB RAS, Krasnoyarsk, Russia
[3] Siberian Federal University, Krasnoyarsk, Russia
scorant@icm.krasn.ru, tav@icm.krasn.ru, oleg@icm.krasn.ru

Abstract. The features of the development of an environmental monitoring system based on the integration of data from various sources is considered. Operational data collection from automated observation posts for the state of the atmospheric air and meteorological information has been implemented. Technology platform of the geoportal, intended for the collection, processing and presentation of data from various observations, is the system basis of completed development.

Keywords: web mapping, GIS, the natural environment, spatial data, monitoring.

1 Introduction

Air quality is the most important environmental factor that determines the health of the population and the state of ecosystems. Environmental and air monitoring systems particular are being created and implemented in many cities of the world both abroad and in Russia [1, 2]. A number of federal and regional organizations in Krasnoyarsk monitor the state of air pollution on the basis of their own methods and technologies of data processing [3]. The variety of solutions used, along with interdepartmental and organizational fragmentation, leads to the fact that a comprehensive analysis and rapid assessment of the entire array of recorded information is currently technically difficult and practically not carried out. The data collected are often published late and are not widely understood.

The use of international standards would greatly facilitate the collection and analysis of observations from different sources. However, today the situation in Krasnoyarsk and Russia as a whole is far from ideal. In most cases, access to operational observations is limited and there are no specialized tools for obtaining such data.

It should be noted that today the regional authorities, together with the Federal services, are trying to create a unified system for monitoring air quality. This example is a "Unified Information System of Monitoring of Pollution of Atmospheric Air of Krasnoyarsk" (http://www.feerc.ru/uisem/krasnoyarsk/). However, it is not possible to analyze operational data with such a system, as the system provides only daily average air pollution data.

2 Methods and tools

In the global community the initiative for the establishment of a network of sensors Sensor Web Enablement (SWE) of Open Geospatial Consortium (OGC) has gained prominence due to its maturity and broad support from academics and industry. SWE is committed to providing open standards and protocols to improve communication between different platforms and data providers. The SWE initiative aims to facilitate search, improve accessibility and provide real-time monitoring of various sensors via the Internet [4–7]. Currently, the SWE family consists of seven standards: Sensor Model Language (SensorML), Observations & Measurements (O&M), Transducer Model Language (TML), Sensor Observation Service (SOS), Sensor Planning Service (SPS), Sensor Alert Service (SAS), Web Notification Service (WNS).

The most interesting and useful tool for analyzing the environmental situation in Krasnoyarsk would be a tool based on the standard Sensor Observation Service (SOS). This standard describes a service that provides access to measurement results from the sensor or sensor network, provides the software information about the capabilities of the sensor and the quality of the measurements. It also provides access to the parameters of the sensor. In other words, SOS groups a collection of possible heterogeneous sensors and provides their measurements through a standardized service interface. SOS specification defines the operations offered by a specific sensor. The minimum set of methods includes GetCapabilites, DescribeSensor and GetObservation that return information about the observations and measurements supported by SOS. The types of data provided by the sensor and the sensor types themselves can be obtained from the sensor registry.

3 Data collection technology

As a result of the analysis of the current situation in Krasnoyarsk, a list of several sources of operational information on the state of atmospheric air, which can be obtained and analyzed, was formed. However, these data have to be extracted by special tools. It should be noted that in most cases the collection of data from external sources can be divided into several groups:

1. Data is presented in the form of web services that transmit information in a structured form, for example, in CSV, JSON, XML formats. Such data is easy to process and check for errors. This block includes data provided within the OGC SOS standard.

2. There are no special web services for receiving data, but there are service services of the resource necessary for its operation. For example, services for plotting or output tabular data in a web interface, tools, data export, etc. Such services can also be used to retrieve the data, but it is necessary to develop a number of tools for data analysis, including the inspection of changes in the structure of the data, changing internal IDs of the format of the output data, the dimensionality of the data, etc.

3. The data is presented in the form of HTML pages of the Internet resource. To obtain such data, you need to analyze the page in order to find some reference text blocks, with which you can always find the necessary information on the page. In this case, a specific text block is extracted from the page and processed further. In some cases, these page fragments can be treated as xml, making it easier to extract information.

In some cases, a combination of these groups can be used, because the services may require service information contained in the page itself. Special attention should be paid to structural changes in the received data, developing a series of verification units, including changes in the number of sensors and monitoring stations, the integrity and completeness of the transmitted data, etc. For the operational refinement of the software modules of the collection, the means of logging text information and alerts, for example, by e-mail, should be created. Such tools allow you to receive information about changes on the remote resource and its status constantly. Promptly making corrections to the program code can ensure the integrity and correctness of the information received.

There are a number of systems in the world for accessing observational data. For example, the smart Emission platform (https://data.smartemission.nl/data), the OpenSensorHub project (https://opensensorhub.org/), project 52°North (https://52north.org/), istSOS OSGeo project (http://istsos.org/), et al.

4 Results and discussion

On the basis of the geoportal of the Institute of Computational Modeling of SB RAS, a unit of scientific and research monitoring was developed for the collection, processing and presentation of data from various observations [8]. Indicators from external data sources are collected on the SensorCollector collection server (http://gis.krasn.ru/sc/). Data sources can be either individual sensors or external databases and information systems through additional adapters. The data collected is recorded in a fact table and stored in multidimensional form with specific dimensions such as date-time, measure, and site.

On the basis of primary data, automatic aggregation with different time intervals is performed, calculation of derived indicators is supported. A software interface (API) was implemented to populate and retrieve data, as well as a basic web interface for metadata management and storage navigation. Implemented support for differentiation of access rights to objects of the system depending on the role of the user. The service is implemented in PHP 5 scripting language using Yii framework, PostgreSQL DBMS with PostGIS extension module is used for data storage.

The organization of access to observation data is carried out by various applications and services, including viewing tabular data, export, viewing data on maps with the ability to select time intervals and access using generally accepted standards (OGC WMS, SOS). The General scheme of data collection is shown in Figure 1.

At the moment, the portal has implemented the collection of operational data from several sources and created the following sections:

- Monitoring system of air quality in the city of Krasnoyarsk.
- Subsystem and database of the official regional ecological system of municipal authorities.
- Subsystem and database of territorial bodies of hydro meteorological service of Russia in Krasnoyarsk region.
- Environmental data of independent ecologists and public organizations.
- Data of hydrological observations in the Krasnoyarsk region.
- et al.

Data are also collected from the devices of the network of scientific and research environmental monitoring of the Federal Research Center "Krasnoyarsk Science Center SB RAS". Data is collected at 1 minute intervals from 24 posts using the data access API. Monitoring stations are developed by CityAir company (https://cityair.io/) and allow to obtain data of measurements of particulate matter concentrations in atmospheric air (PM2.5, PM10) and basic meteorological parameters such as temperature, air humidity and pressure. An archive of data obtained from these stations has been accumulated since December 2018.

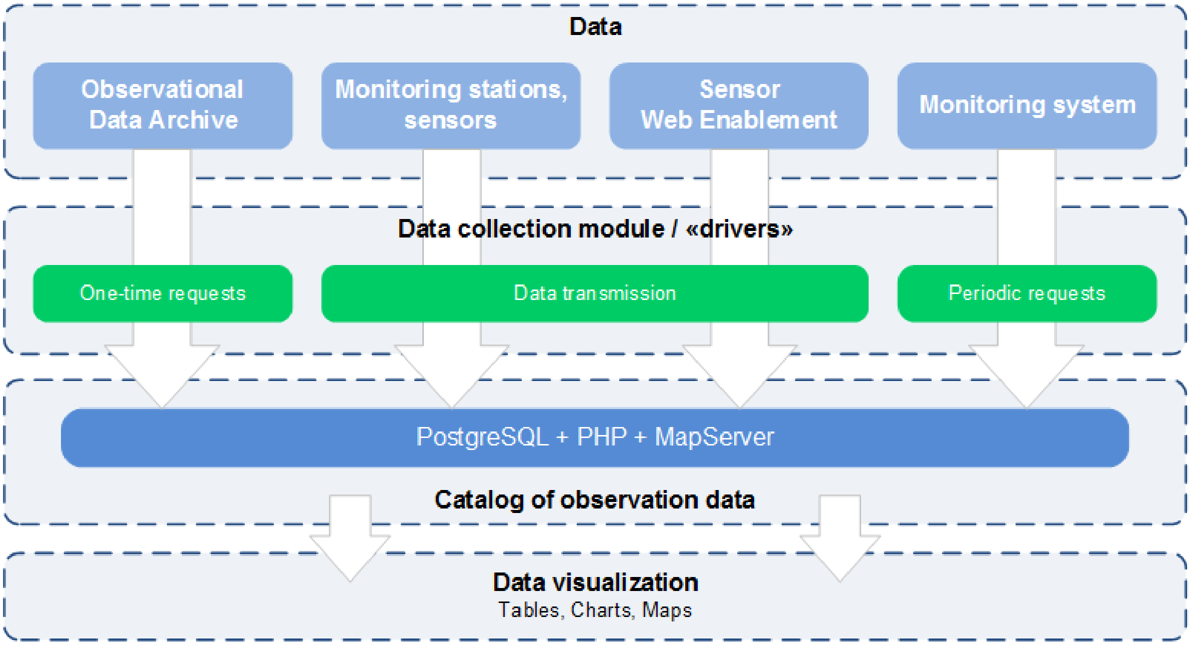

Figure 1. Scheme of data collection observations.

The data flow on atmospheric air from the Regional Departmental Information and Analytical System on the State of the Environment of the Krasnoyarsk territory (RDIAS) is set up (http://krasecology.ru/Air/). Observations of air quality are carried out at 11 automated observation posts, 9 of which are located in Krasnoyarsk. To download data from the RDIAS site, a software module with an appropriate "driver" for processing and converting the input data has been prepared, which provides periodic loading of observation data from this system through a web service. At the moment, most of the downloaded data on the fixed observation posts on the remote server is generated every 20 minutes. Data from mobile laboratories appear with a frequency of once a day, after processing them by employees of the regional system.

Meteorological data are automatically downloaded from the website of the Federal State Institution "Middle Siberian Department of Hydrometeorology and Environmental Monitoring" (http://meteo.krasnoyarsk.ru/). Meteorological information comes in real time from automatic monitoring posts, it is published on a special website and is updated every 3 hours. Data are collected on the following indicators: atmospheric pressure, humidity, wind direction, wind speed, air temperature. The archive has been maintained since the beginning of 2013. Previously, the site was available daily data on a number of gases (NO2, NO, SO2, CO, hydrofluoride, formaldehyde) and particulate matter PM10. These data were downloaded daily, but since June 2016 free access to these data has been closed.

Another source of information is data from the Nebo community project (https://nebo.live/), created by environmental activists from Krasnoyarsk. The project collects data from about 20-30 instruments measuring the particulate matter concentrations in the air (PM2.5) and the main meteorological parameters: temperature and relative humidity. At the end of last year, the owners of the service began to upload archives of data for the past period. However, these archives have gaps in time, but sometimes they allow you to fill in the data missed by the import service. Currently (summer 2019) access to the archives is closed.

Special attention was paid to the fact that the data of the Nebo community project on the charts with the average values for the day and month were compared with the maximum one-time maximum permissible concentration of harmful substances, which is incorrect. The site contains a link to the documentation for working with the service API (https://docs.nebo.live/api), but it is not currently available. The software API would make it easier to download data from the site and reduce the number of missing data due to problems with the project site.

An adapter has been developed to download daily reports of global coverage (GSOD) from the National Oceanic and Atmospheric Administration (NOAA) of the USA from weather stations of the Krasnoyarsk territory. These data represent records that contain information about temperature, humidity, wind speed and direction, atmospheric pressure, types of precipitation, etc. A data archive has been formed since 2003, services for automatic loading of new information have been created.

Special attention was paid to the subsystem of analysis of data received from third-party services. The subsystem analyzes the following events:

1. Unavailability of a web service or site via http protocol or lack of communication.
2. Correctness of the received data (negative or zero values for the data series, coincidence of time in the received data with the requested period, data from the "future", etc.).
3. The unavailability of individual monitoring stations or their indicators.

4. The appearance or disappearance of monitoring stations or indicators on a remote service or website (allows you to quickly add new monitoring stations to the system).

The system analyzes the above problems and records all events in a text log on the server, as well as informs the system administrator by e-mail. This subsystem is currently being refined to limit the number of e-mail messages and to send summary information about non-critical errors once in a given period (by default, once a day). This development of the subsystem is due to the fact that part of the services collects data every 10 minutes and each time for the entire service or for each monitoring post separately formed an error message and the number of such messages per hour can be measured in tens of pieces. This is especially true for the stations of CityAir company, because their software is developing and often there are various errors: the data does not come sequentially, some packages have a date with a shift of a week in the future, there are new "virtual" devices on existing stations, etc.

To date, in cooperation with the system of operational monitoring data, which is a part of geoportal of Institute of Computational Modeling SB RAS, several web applications have been developed for the publication and analysis of the collected observations that are part of the emerging regional spatial data infrastructure [9]. Figure 2 shows a web application with operational data on all the posts of the CityAir company of the Federal Research Center KSC SB RAS in Krasnoyarsk. The current weather conditions and the level of air pollution (PM2.5) are displayed. Depending on the level of contamination, the indicators are colored. The data is updated on the screen every 20 minutes.

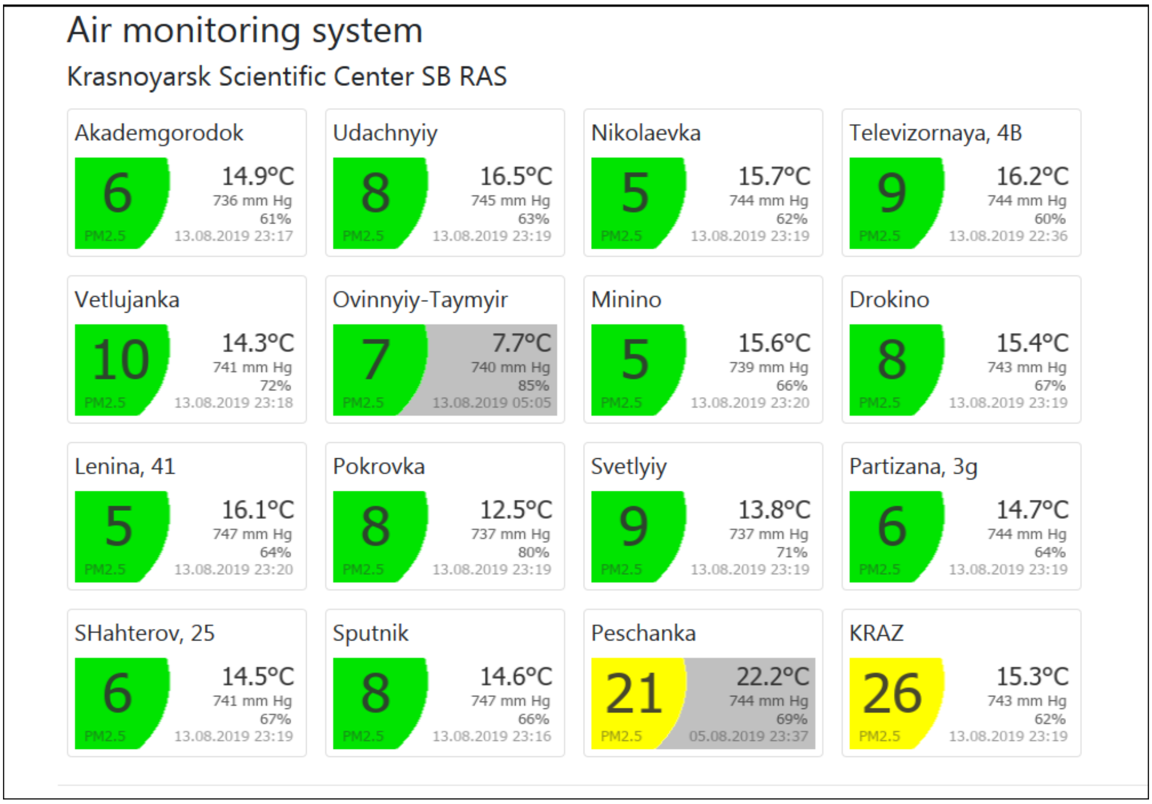

Figure 2. Real-time observation data in Krasnoyarsk from CityAir stations.

When developing web applications, the analysis of existing solutions related to the presentation of operational information on the state of atmospheric air in the world and Russia was carried out. Among the considered systems there are several examples. Air Quality System based on the data of the United States Environmental Protection Agency (https://gispub.epa.gov/airnow/) provides access to data in the form of Air Quality Index (AQI) [10]. The data on air quality and weather conditions of the AirVisual international system (now part of the IQAir, https://www.airvisual.com/air-quality-map) obtained on the basis of the data from small home stations AirVisual Pro. Non-profit World Air Quality Index project (http://aqicn.org/map/world/) based in China provides access to data in the form of an air quality index for many countries of the world according to the environmental protection Agency methodology.

Examples of several regional systems were also considered. One of them is the European system The Copernicus Atmosphere Monitoring Service (https://www.regional.atmosphere.copernicus.eu). Another successful project is the system of the Republican Center for Hydrometeorology, Radioactive Pollution Control and Environmental Monitoring of the Ministry of Natural Resources of the Republic of Belarus (http://rad.org.by/monitoring/air.html). It is also worth mentioning the system "air Quality of Moscow" (http://mosecom.ru/vozdux/) state environmental budget institution "Mosecomonitoring"; system "Environmental portal of St. Petersburg" (http://www.infoeco.ru/index.php?id=53), data published by the Committee on Nature Management, Environmental Protection and Environmental Safety, Air Quality Monitoring in Mumbai, India [11]. Most of these examples provide access only through a web browser or mobile application. Access using standard services (e.g., OGC SOS) is not available.

An example of a custom web application built on the components and API of our Institute's geoportal is shown in Figure 3. This development is considered as a prototype of the created regional environmental monitoring system [12]. The map shows data on particulate matter concentrations (PM2.5) in the atmosphere of Krasnoyarsk, obtained from various observation systems. The shape of the icons on the map characterizes the belonging to the data source, and the color reflects the impurity concentration. In addition, the figure shows an example of a graph with Air Quality Index from the stations of the regional system in the city of Krasnoyarsk in general for the first months of 2019.

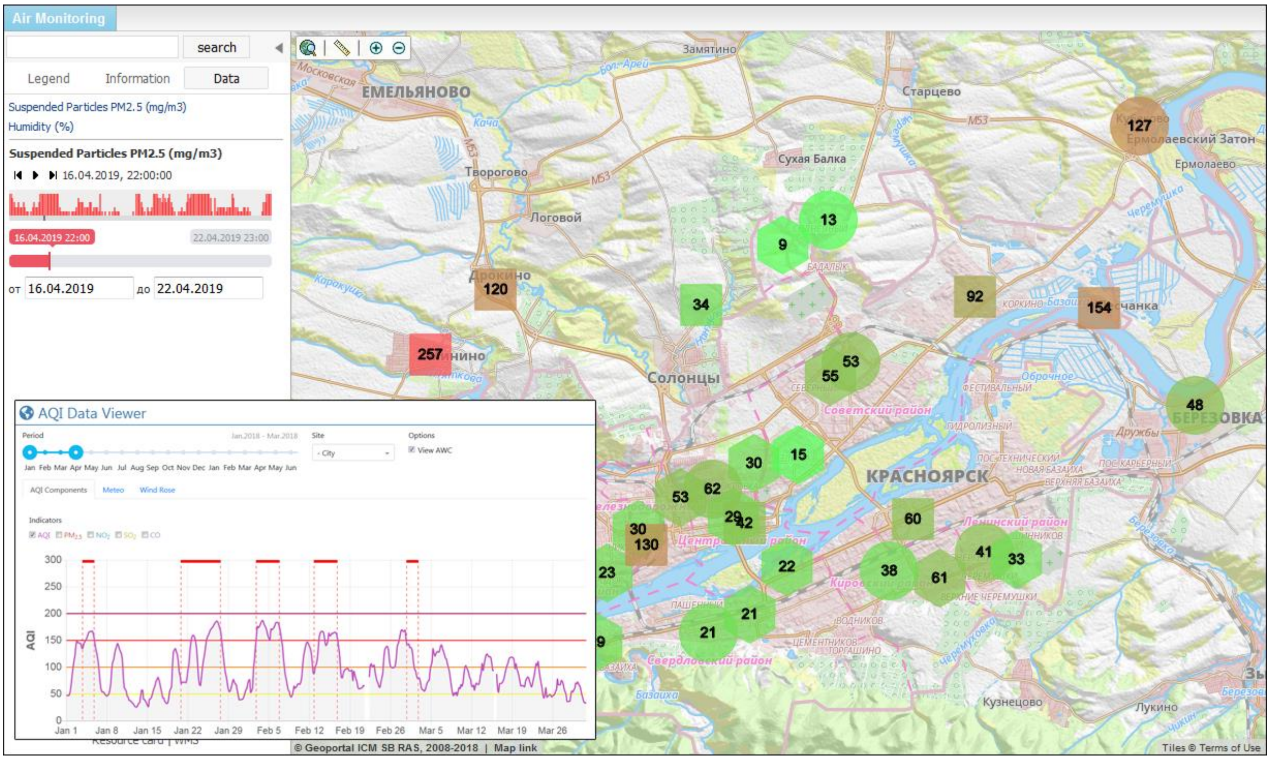

Figure 3. Map with observational data for PM2.5 in Krasnoyarsk on 04.16.2019.

5 Conclusions

For operational monitoring and analysis of the environmental situation in Krasnoyarsk, data collection from different sources was organized into a single system "Operational Monitoring Data" of the Institute of Computational Modeling SB RAS.

The developed software for importing observations from various resources allows automatic loading of data, regardless of how the access to these data is organized. In case of any changes and errors in the import, e-mail messages are sent, allowing you to quickly make changes to the import program or data on sensors and their composition. In the future, the possibility of organizing access to observational data in Krasnoyarsk with the help of the international standard OGC SOS is being considered. Unfortunately, in Russia this standard is not well developed, the suppliers of monitoring data.

References

[1] Manuel A. Regueiro, José R.R. Viqueira, Christoph Stasch, José A. Taboada. Semantic mediation of observation datasets through Sensor Observation Services // Future Generation Computer Systems. 2017. 67, pp. 47–56.

[2] Azemov D. Saint-Petersburg air pollution monitoring system // Environment of Saint- Petersburg, 2016. 2 (2), pp 8–14.

[3] Yakubailik O.E., Kadochnikov A.A., Tokarev A.V. Development of computational technologies and software for assessment of air pollution in Krasnoyarsk // IOP Conference Series: Earth and Environmental Science, 2019, Vol. 211 (1), 012080.

[4] Botts M, Percivall G, Reed C, Davidson J 2008 OGC® Sensor Web Enablement: Overview and High Level Architecture. In: Nittel S, Labrinidis A, Stefanidis A (eds) GeoSensor Networks // GSN 2006. Lecture Notes in Computer Science. Vol 4540. Springer, Berlin, Heidelberg.

[5] Kamel Boulos et al. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples // International Journal of Health Geographics. 2011 10:67.

[6] Schmitt R.H. and Voigtmann C. Sensor information as a service – component of networked production // J. Sens. Sens. Syst., 7, 389–402, 2018.

[7] Kotsev A, Schade S, Craglia M, Gerboles M, Spinelle L, Signorini M. Next Generation Air Quality Platform: Openness and Interoperability for the Internet of Things // Sensors (Basel, Switzerland) vol. 16,3 403. 2016.

[8] Yakubailik O.E., Kadochnikov A.A., Tokarev A.V. Web geographic information system and the hardware and software ensuring rapid assessment of air pollution // Optoelectronics, Instrumentation and Data Processing. 2018. 54 (3), pp 243–249.

[9] Kadochnikov A.A., Tokarev A.V., Zavoruev V.V., Yakubailik O.E. Prototype of city environmental monitoring system based on geoportal technologies // IOP Conference Series: Materials Science and Engineering, 2019, Vol. 537, 062052.

[10] Technical Assistance Document for the Reporting of Daily Air Quality – the Air Quality Index (AQI) // U.S. Environmental Protection Agency. 2018. https://www3.epa.gov/airnow/aqi-technical-assistance-document-sept2018.pdf

[11] Awkash Kumar, Indrani Gupta, Jørgen Brandt, Rakesh Kumar, Anil Kumar Dikshit & Rashmi S. Patil. Air quality mapping using GIS and economic evaluation of health impact for Mumbai City, India // Journal of the Air & Waste Management Association, 2016, 66:5, pp 470-481.

[12] Yakubailik O.E. The formation of regional spatial data infrastructure // Journal of Siberian Federal University. Engineering & Technologies, 2016, Vol. 9 (7), pp. 979-986.