ISBN 978-602-344-251-5

Study of Bali rivers water quality via automated workflow using python and arcGIS tools

Pavel Kargashin, Andrew Karpachevskiy, Maria Kozlova, Nataliya Shchegolkova, I Gusti Ngurah Santosa, I Nyoman Sunarta, I Wayan Budiasa, I Ketut Suada, I Nyoman Rai
Moscow State University, Faculty of Geography, Department of Cartography and Geoinformatics, Moscow, Russia
**Coresponding author: p.e.kargashin@mail.ru*

ABSTRACT

The article shows the results of joint research of scientists from Indonesia (Udayana University) and Russia Federation (Lomonosov MSU and Water Problems Institute RAS). Within the frameworks of the research, the water quality of natural water bodies is explored in the Bali island. Initial data for investigation was taken from state environmental reports. The main goal was to systematize data, find optimal ways for data storage and representation and make an overview of water quality in Bali island, to explore the spatial and temporal aspects. Database construction and programming used Python to build GIS project to maintain the research. There are concepts of workflow in the article and some primary conclusions concerning the current monitoring system and spatial distribution of pollutants in river basins.

Keywords: *water quality, monitoring, geographic information system, geographical maps, pollutants*

INTRODUCTION

The system of monitoring the quality of natural surface water is a set of technologies for obtaining and processing data to describe the quality of natural water bodies according to the list of parameters. It is critical that monitoring routine ensure sufficient coverage and certain periodicity of observations. An important purpose of monitoring is to measure major parameters of water quality (Ministry of Technology UK, 1970), but also to identify the factors causing changes in the natural objects, as well as the reasons of generation of these factors. At the same time, in the field of natural resources management, the purpose of monitoring is considered to be the information support of decision making to decrease environmental risks and to mitigate impacts (S. Behmel et al., 2016). It is important to note that, besides the practical goals, monitoring allows to accumulate an array of data for scientific research.

ISBN 978-602-344-251-5

Monitoring can be classified according to many criteria. Well-known classification concerns the spatial coverage. Global monitoring involves the observation of processes or phenomena that have a global imprint. Regional monitoring means observations of the territory with similar geographical or socio-economic indicators. Local monitoring covers a relatively small area, which has a certain set of anthropogenic impacts. There are also monitoring systems covering the territories limited by administrative boundaries. Such system can accumulate data about neighboring countries, countries themselves, regions, provinces etc.

Water quality monitoring must be complex. Thus, monitoring of water quality of a large river involves not only investigation of the whole catchment area, but also inspecting quality of water, bottom sediments and precipitation. The prior task for such investigation is to find out the sources of impact and determine their features. Monitoring routine can include investigation of physical factors (radioactive, electromagnetic, thermal, acoustic, vibration, etc.), chemical factors (pollution by various substances), physical and chemical factors (for example, the impact of acid rain on soil and water) (Timothy, 1987; Takeshi, 1988).

Sometimes it is sensible to focus on monitoring of individual industries (for example, monitoring of ferrous or non-ferrous metallurgy, monitoring agriculture, etc.). These types of human activity are specific, so they have a certain row of pollutants and monitoring routine is aimed to detect substances from a limited list.

Monitoring technology uses different methods of observation and devices as well as various information technologies. Methods can be divided into those that are implemented by means of sensory organs (perception of smell, color, noise, comparative heat estimates, visual analysis of environmental objects directly in nature or in images, etc.), and methods implemented on the basis of technical devices which use different physical, chemical and physico-chemical principles.

Monitoring routines use a wide row of modern methods including analytical chemistry, as well as hydrological, geological, meteorological techniques and many others are applied. The corresponding devices can be placed on stationary and mobile technical means: on land they can be stationary or field laboratories, can be set on self-moving or transported vehicles; there are used vessels and floating devices for seas and also unmanned devices in the air and Space.

All methods can also be divided into two groups: contact ones, which implementation is carried out direct physical contact of the measuring device and the controlled natural object in the natural state or in the form of a sample, and remote methods. It should also be noted the current trends in the creation of

ISBN 978-602-344-251-5

devices that simultaneously measure a certain set of characteristics of the object, as well as current automation of the monitoring process.

Information technologies of monitoring, depending on the goals of monitoring, are used within the workflows of collection and primary processing of data, their accumulation in databases, in the construction of various spatio-temporal prediction models (M.Khadr and M. Elshemy, 2017; Siyu et al., 2018), when using geographic information systems (Zelhofer et al., 2007). They are also used to build predictive scenarios, which may include draft decisions to reduce damage to environment causing by current and future human activity. The monitoring system should be designed in conjunction with communication and data processing systems.

The essence of environmental monitoring is to establish cause-and-effect relationships between the anthropogenic load and the state of ecosystems, as well as the health of the population.

The creators and users of monitoring systems may be executive bodies authorized to protect the environment and human health on behalf of the state, municipal authorities, economic entities, scientific organizations, public associations, etc.

The optimal construction of the environmental protection system of the country should include the creation of a unified strategy for all monitoring systems, the development of common monitoring standards, the exchange of information, the creation of databases and their application in decision-making.

Unfortunately, often the construction of an optimal system of monitoring of water objects is almost an impossible task, especially for territories where environmental research is selective. For example, for the island of Bali, Indonesia, the authors could not find lots of scientific research relating to the assessment of the quality of the environment and, in particular, water bodies. This may indirectly indicate that in this region insufficient attention is paid to the problems of ecology and rational use of water resources. The similar issues noted Stroma Cole (Stroma, 2012) in the article. There is also some information in the book Ecology of Java & Bali (Anthony et al., 1996). Some authors pay attention to agriculture impacts (Machbub et al., 1998).

However, the government conducts annual monitoring of the quality of water. At the same time, it is quite difficult to make a comparison between the defined indicators, since the materials presented in the annual reports are not unified, often presented chaotically, they don’t cover each year for each analyzed indicator. Moreover, the number of locations and the frequency is not enough to get significant scientific results.

Therefore, an integrated approach to the processing of available data seems necessary for the analysis of the state of the water bodies of Bali and for making management decisions. This approach can be implemented in the use of

ISBN 978-602-344-251-5

GIS, allowing to demonstrate the results of the monitoring and indicate the missing layers of information, as well as the most disadvantaged areas of the study area. The suggested approach can be used even with insufficient data and will help to upgrade the system of monitoring.

MATERIALS AND METHODS

In this project we have set several tasks. First of all, we need to look through the source data and figure out the features of its processing in GIS. Also, we need to obtain methods of data verification for further automated processing of the environmental data. To define the best ways of data representation we need to construct the user-friendly workflow and program interface which allows to get map series based on data without learning theory of geomatics and ArcGIS or some other GIS software. Finally, using created map series we can analyze the environmental situation figuring out some spatial regularities and dynamics within 10 years. It could allow us to characterize the actual state of water bodies within Bali island and outline trends of human impact on the environment.

Our research was based on State environmental data reports for Bali island. The document includes the results of water sampling. This data refers to the rivers, lakes, underground water and sea water. Each sample was analyzed for a wide row of chemical and physical parameters. The total count of such parameters is nearly 20. In addition, there is some metadata, e.g. time, date, water body, longitude, latitude of the sample points.

The idea of the research is to get an automated workflow which can help users to create maps without deep working in GIS software. Besides, it is critical to use appropriate GIS-techniques within a workflow to get correct maps, so that we could analyze them and make some decisions. To provide the workflow with tools, which would preprocess and analyze data, we created some scripts using the Python programming language Python 2.7. Particularly we used several libraries for created scripts: arcpy, PyQt and some extra ones.

Arcpy is a python library which includes large set of ArcGIS tools. This library is usually installed together with ArcGIS but also it can be utilized in outstanding scripts. Arcpy is a key library for our project because it allows to automate spatial data (often called geodata) manipulation and designing maps. The tools from arcpy library were used to create geodatabase and datasets, to import attributes, to locate points due to its coordinates in source, to perform calculations. Other tools allow user to create map series with all cartographic attributes, including title, scale and legend and to export it to external raster file format.

PyQt is a library which was used to construct user interfaces for separate parts of workflow. We have divided our workflow into 3 blocks: 1) creation of

ISBN 978-602-344-251-5

the project, 2) Loading of the data, 3) Construction of the map or map series. Each block initially has at least one window by which user can select a file with source data or choose chemical parameter from dropdown list etc.

The user can create a map using ArcGIS capabilities even without opening it and that is very helpful for non-cartographers. Developed scripts also allow to input new environmental data and verify it. For example, verification finds the impossibility of entering some text instead of numerical value to describe concentration. Another aspect of verification refers to the sampling location. Each coordinate value compares with defined pattern and limits from the north to the south and from the west to the east. If the value doesn't suit, it must be re-entered. All the described solutions help to avoid errors caused by users and provide the automated data processing. Also, for source data we have developed some conventions of representation in Excel.

RESULTS AND DISCUSSION

The results of the research can be divided into 2 parts. The first one is an automated workflow. The screenshots below show the workflow (Figure 1). The project can be deployed on the PC within 10 minutes. The input data is base map which consists of appropriate shape files, and thematic data in Excel format.

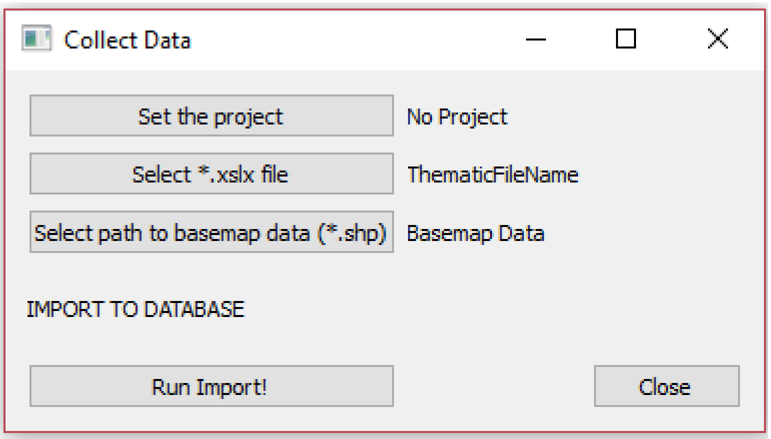

Figure 1. Window to select initial data for project.

Additionally, we can input data for a single sample and load it to the project. Mapping is presented in two modes: single or batch. For single map the user has to select a required data: year, units of territorial division (administrative division or river basins) and particular chemical substance or parameter. The result would be presented as a map and saved to the project folder. The example of window for mapping is shown below in the figure 2. The

ISBN 978-602-344-251-5

examples of the maps would be further in the article. The batch mode allows to create all possible maps (map series) via one click.

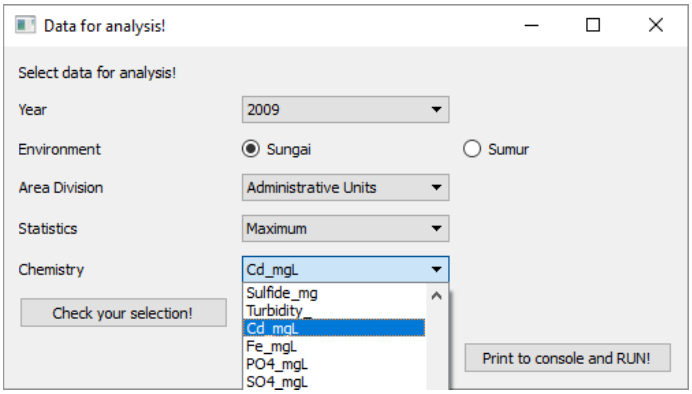

Figure 2.Window to select parameters for single map construction

With the help of developed tools, we have got a geodatabase containing the data about the quality of water bodies within the Bali island. The database includes quantitative indicators of water quality according to years, seasons and different types of water, such as rivers, lakes, groundwater, sea water for 7 years (from 2010 to 2017). By means of GIS-project we analyzed the maximum, minimum and average values of indicators such as BOD, COD, biogenic elements (nitrogen, phosphorus), dissolved oxygen and its range (min-max), bacteriological parameters (total coliform bacteria), iron, copper, calcium and other pollutants (zinc, chromium, cadmium, mercury, detergents, oil and other pollutants fat).

There are some notices before map analysis. First, we found out that some data can't be located due to the errors in coordinates or their absence. So not each sample from database was placed on the map. Another conclusion is that there were different programs of water sampling and there is not enough data to construct maps which illustrate the situation in the whole island, some areas has value "No data". In the following maps color of polygons (the river basins or regencies) show maximum concentration of chemical substances or maximum value of other parameters within the one year. Other maps devoted to the lakes of Bali island show dynamics of maximum concentrations via diagrams.

If analyzed such integral indicators as BOD, COD, DO, which characterize the overall levels of water pollution and the possibility of self-purification of the ecosystem, there can be seen a noticeable correlation (see Figure 3). This may confirm the representativeness of the initial data on these characteristics in the annual state report.

ISBN 978-602-344-251-5

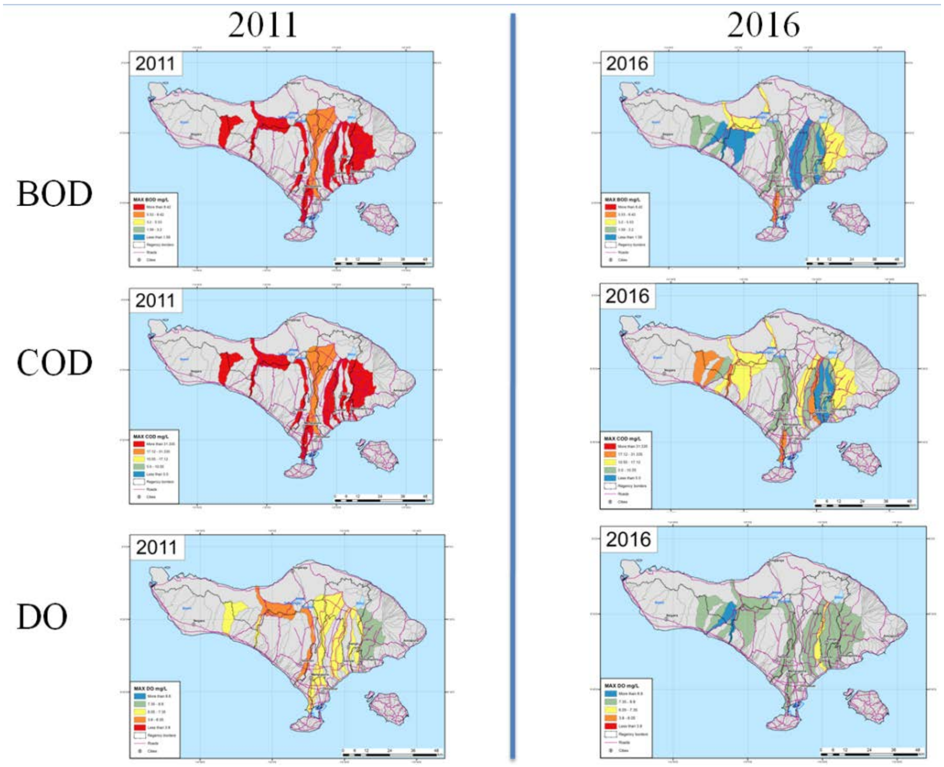

Figure 3.Correlation between BOD, COD and DO

Naturally, the maximum values of bacteriological parameters fall on densely populated areas of Kuta (Figure 4). Number of inhabitants and existing anthropogenic load is an initial explanation which must be clarified within further investigation. However, there is no data on Denpasar at all. Therefore, it should be recommended that sampling of the data should take into account possible sources of pollution and that maximum attention should be given to the densely populated areas of the island.

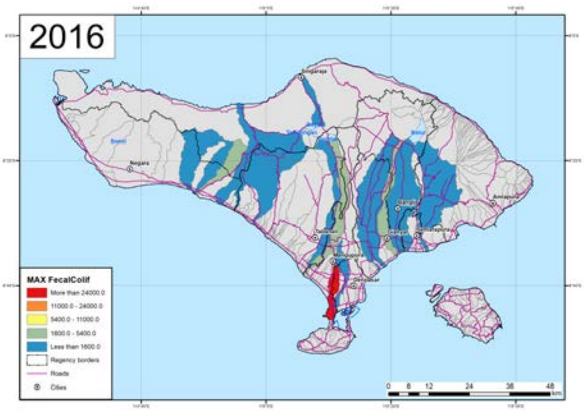

Figure 4. Identified heterogeneity in pollution level of the island.

ISBN 978-602-344-251-5

Parameter is fecal coliform bacteria

Some maps indicate a strong pollution of rivers by organic compounds. It can be noted that the high values of phosphorus in some areas in the eastern part of the island. The phosphorus spatial distribution is similar to distribution of various forms of nitrogen. As a result, conclusion about the most environmentally unfavorable areas and high anthropogenic load can be taken (Figure 5).

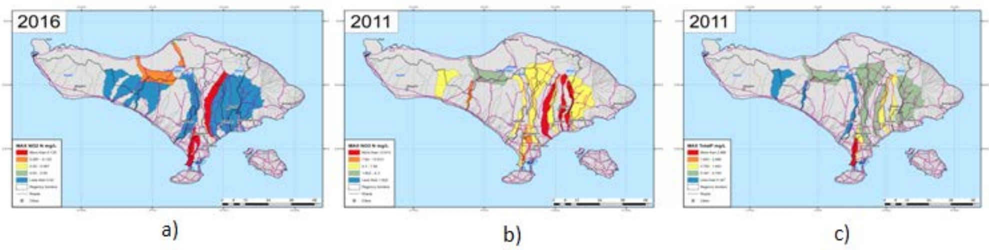

Figure 5. Identified heterogeneity in pollution level of the island. Parameters – a) NO_2-N, b) NO_3-N, c) Total P

In addition, the change in concentrations of pollutants within the river basin can be assessed by GIS application. It is easy to show on the map an increase in NO_3 concentration from the river source to the estuary. Thus, it is possible to allocate sites on which there is a strong anthropogenic impact. For example, having imposed on such map a layer with pollution sources.

By analyzing the maps obtained, we can distinguish four most anthropogenically loaded areas. They are Kuta, Denpasar, Gianyar and the western part of Tabanan. They have maximum concentrations of bacteriological indicators, and of certain metals oxides (e.g. iron, manganese, zinc). The quality of river water in these areas is worse comparing to neighboring plots and it is necessary to pay additional attention when planning monitoring program to this fact (Figure 6).

ISBN 978-602-344-251-5

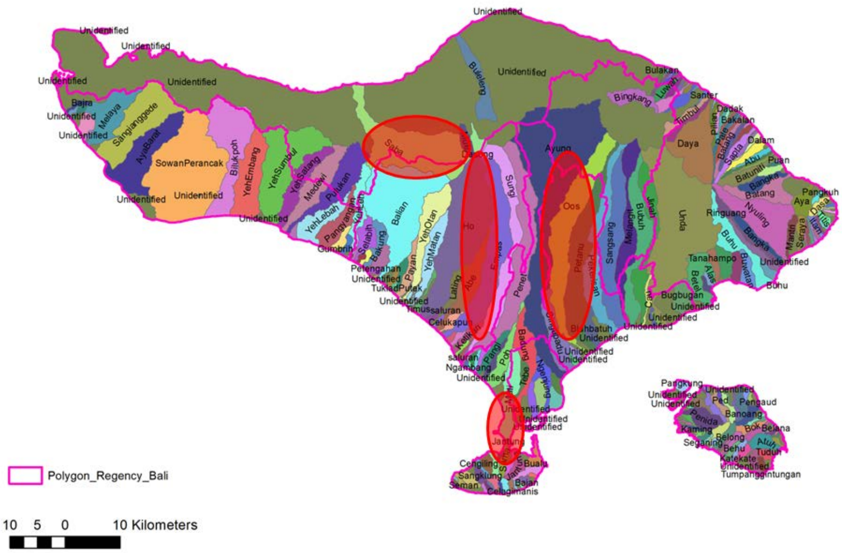

Figure 6. The most polluted areas of Bali

CONCLUSIONS

As it is shown that the utilization of GIS allows to highlight the type of pollution, to identify trends, to determine the basins of water bodies with maximum anthropogenic load, to determine the place of priority measures to improve the environmental situation.

Our GIS is now developing and in future we are planning to extend the capabilities of the system. Nevertheless, it is very noticeable the lack of the data and the lack of correct data. On the Figure 7 below it is shown that the same source data (river samples for the 2014 year) could be interpreted differently because of different units of territorial division – regencies and river basins. Territories colored grey have no data about water quality. If we take river basin division, it becomes obvious, that more than 75% of Bali island is unexplored.

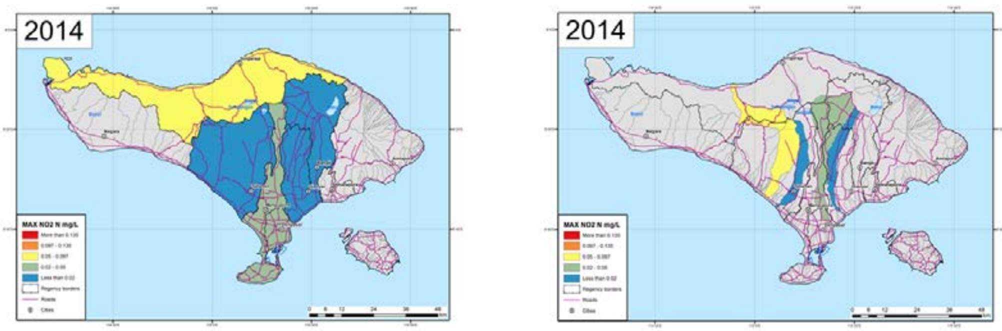

Figure 7. Data representation for a) regencies, b) river basins

ISBN 978-602-344-251-5

Future development of the project depends not only on the program capabilities and visualization, but also on the availability of representative data, so we propose our GIS for input of the new data.

REFERENCES

Anthony J. W., T. Whitten, R. S. Soeriaatmadja, R. E. Soeriaatmadja, and S. A. Afiff. 1996. Ecology of Java & Bali Periplus Editions (HK) Limited. p. 969.

Machbub, B., Ludwig, H.F. and Gunaratnam, D. 1998. Environ Monit Assess 11: 1. https://doi.org/10.1007/BF00394508

Ministry of Technology UK. 1970. Monitoring of water quality. *Water Research*.,**4**(11):765-768.

M. Khadr and M. Elshemy. 2017. Data-driven modeling for water quality prediction case study: the drains system associated with Manzala Lake. *Egypt Ain Shams Eng. J.*, **8**(4):549-557.

S. Behmel, M. Damour, R. Ludwig, and M.J. Rodriguez. 2016. Water quality monitoring strategies-a review and future perspectives Sci. *Total Environ.*, **571**:1312-1329

Siyu C. ,G. Fang , X. Huang and Y Zhang. 2018. Water quality prediction model of a water diversion project based on the improved artificial bee colony–backpropagation neural network. *Water*., **10**(6):80

Stroma C. 2012. A political ecology of water equity and tourism: A Case Study From Bali. *Annals of Tourism Research*., **39**(2):1221-1241.

Takeshi G. 1988. Monitoring and measurement of water quality parameters. *International Journal of Water Resources Development*.,**4**(4):270-275.

Timothy D. S. 1987 Water quality monitoring strategies. *Hydrological Sciences Journal*., **32**(2)207-213.

Zeilhofer, Peter, Zeilhofer, L. V. A. Corrêa, Hardoim, E. Lopes, Lima, Z. M. de Oliveira, and C. Silva. 2007. GIS applications for mapping and spatial modeling of urban-use water quality: a case study in District of Cuiabá, Mato Grosso, Brazil.*Cadernos de Saúde Pública*., **23**(4)875-884.