Detection of anomalies of the aquatorium ecosystems of Sevastopol on the basis of the markov model

A. Skatkov[1], A. Bryukhovetskiy[1], and D. Moiseev[1,2,*]

[1] Federal State Autonomous Educational Institution of Higher Education «Sevastopol State University», Russian Federation, 299053, Sevastopol
[2] Black Sea Higher Navy Order of the Red Star Academy by P.S. Nakhimov, Federal State Owned Military Educational Institution of the Higher Professional Education, Russian Federation, 299028, Sevastopol

Abstract. This paper is devoted to the development of the Markov model for the detection of anomalies of ecosystems of the waters of the city of Sevastopol. The states of the objects under study, the algorithm for constructing the matrix of transition probabilities and determining the final probabilities are presented. Some of the important applications of anomaly detection based on the Markov model are the prediction and recognition of anomalies.

Environmental monitoring and control is a set of measures to identify and assess the sources and levels of contamination of natural objects with harmful substances as a result of discharges or emissions of these substances into the environment by nature users, as well as due to natural formation and accumulation in environmental objects, including through chemical and biochemical transformation of natural and man-made substances into compounds with harmful properties [1].

The lack of comprehensive monitoring studies of the background state of the marine environment on the seashore of Sevastopol, including oil and phenolic pollution, control of which is especially needed in the waters adjacent to the infrastructure of the cargo-passenger and military fleets, does not allow an adequate assessment of the state of ecosystems and the future development of environmental situations [2]. Therefore, the development of methods for monitoring the detection of ecosystems anomalies in the waters of the city of Sevastopol is an urgent task.

For modern information systems for environmental monitoring, one of the main features is that their parameters and structures at various stages of the life cycle change under the influence of objective and subjective reasons. When choosing control methods, it is necessary to determine the probability of a reliable measurement of the characteristics of an object and how to solve the problem at the lowest financial and other costs.

The approach proposed in the article corresponds to the concept of the priority regional program "Smart City" in accordance with the strategy of socio-economic development of the city of Sevastopol up to 2030. The implementation of this concept includes, inter alia, the creation and maintenance of continuous monitoring of key environmental indicators.

---

[*] Corresponding author: dmitriymoiseev@mail.ru

Therefore, the solution of tasks related to the development of intelligent technologies as a key component of the monitoring system, quickly adapting to changes in the internal states of the critical ecosystem of the water area of the Sevastopol region and detecting critical events in order to promptly take urgent measures to eliminate them, is a necessity.

In [3], an approach to developing an intelligent monitoring system for solving large-scale scientific problems in cloud computing environments in the context of a lack of a priori information about the state of the environment is considered. A set of models for the classification of information situations is presented, allowing cloud brokers to support the decision-making process for choosing the optimal strategy for solving monitoring tasks. Development of decision-making methods in conditions of uncertainty, including methods for detecting abnormal values of monitoring data, will eliminate the shortcomings and limitations inherent in classical approaches and will allow to solve new scientific problems, characterized by variable intensity data flows, non-uniform information flows in conditions of low-informative, missing, noisy data and incomplete information.

One of the methods [4] of identifying the ecological state of the water area is a visualization method based on the modern concept of visual analytics systems Visual Analytics - VA or Visual Data Discovery - VDD. The most important element of the user interface of visual analytics systems are Dashboards - user panels that contain multiple visual indicators on one page. As visualization tools for the multidimensional space of quality indicators, the authors suggest using the methods of parallel coordinates and radar charts. These methods provide the ability to effectively visualize multidimensional information, measured in different scales.

In [5], a method of restructuring information systems for environmental monitoring based on the analysis of options for their implementation is considered. The method of solving the problem is based on a complex of interrelated stages. At the first stage, options are generated based on expert assessments and the minimum necessary functions of the system that implement the i-th task. For thinning the set of admissible solutions, the effectiveness of the obtained variants is determined on the basis of the vector approach, in which the object is estimated by the results of a comparison of its individual properties. Thus, it becomes possible to formalize the decision-making process.

The paper [6] is devoted to methods for modeling the detection of anomalies in complexly structured monitoring data. The article discusses the problem of detecting anomalies in the complexly structured data of environmental monitoring and critical-use systems. An approach is proposed to solve the problem of detecting changes in the state of the monitoring data stream using normal distribution models based on Spearman's nonparametric statistics criterion. The solution of the problem of estimating the influence of values of generation of requests, intensity of service of requests, system load, sample size, time of measurement of characteristics, time intervals of measurement of characteristics and levels of significance on the change of state of the monitoring object is given.

In this article, a method of monitoring the state of the ecosystem of the water area of the city of Sevastopol is proposed based on a Markov model [7]. The method will allow estimating such values of water areas as pollution level, content of hazardous substances, transparency, salinity, temperature, water conductivity, current and wind speed, and others. It is assumed that the state of the aquatic ecosystem is observed first in one time interval, and then in another time interval. Depending on the goals and duration of the observation period, a forecast is possible, for example, for a week, month, year, etc. The main task in the construction of Markov models is to collect data for calculating transition probabilities and constructing transition probabilities matrices, which requires information about changes occurring at certain time intervals and responses to various types of disturbances.

To assess the quality of the monitoring objects, we will use three states in which the objects can be located:

S1 – normal state, parameter value is within normal range,

S2 - precritical state, parameter value is different from normal,

S3 – critical state, anomalous, parameter value significantly differs from normal.

In general, the simulation algorithm contains the following sequence of actions:

1. The vector of initial probabilities is formed $p(0) = \{p(0)(1), p(0)(2), p(0)(3)\}$.

2. A set containing N states of Si is generated and a sequence is constructed Si,…Sj,…Sl,   где i, j, l ϵ {1,2,3}.

3. Calculate the number of |Si,j| pairs of Si,j, where i, j ϵ {1,2,3}, Si,j are the transition from the current state of Si to the next state Sj.

4. Find the probabilities of state transitions Pij =|Si,j| /N and build a matrix of transition probabilities |Pij|, in which the probability of the system's transition to any possible state at each moment of time is determined only by its state at the previous moment of time and does not depend on the earlier history:

$$
\left| P_{ij} \right| = \begin{vmatrix}
P_{11} & P_{12} & \cdots & P_{1j} & \cdots & P_{1n} \\
P_{21} & P_{22} & \cdots & P_{2j} & \cdots & P_{2n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
P_{i1} & P_{i2} & \cdots & P_{1ij} & \cdots & P_{in} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
P_{n1} & P_{1n2} & \cdots & P_{nj} & \cdots & P_{mn}
\end{vmatrix}
$$

The element pij of the matrix means the probability of the transition of the system from the state si to the state sj at discrete points in time t1 , t2 …. Obviously, the sum of the elements of any row in the matrix is 1.

5. Calculate the vector of final probabilities $p(n) = p(0) \times P\,n$.

When n grows, the p(n) vectors stabilize in some cases — they converge to a certain probability vector p, which can be called the stationary distribution of the chain. Stationarity manifests itself in the fact that by taking p(0)= p, we get p(n)= p for any n. The simplest criterion that guarantees convergence to a stationary distribution is as follows: if all elements of the transition probabilities matrix P are positive, then as n tends to infinity, the vector p(n) tends to the vector ρ, which is the only solution of the system p × P = p.

The advantages of using the Markov model include are the following:

1. Model is easy to build on the basis of data on the current state of the ecosystem of the water area.

2. The main matrix of transition probabilities displays the main parameters of dynamic changes in the system.

3. Analysis results are easy to visualize.

4. Computational needs in the study of Markov models of low-cost, especially with a small number of states.

The disadvantages of using the Markov model are the following:

1. Deviation from stationarity, under the assumption of which first-order Markov chains were obtained.

2. In some cases, the available data may not be sufficient to reliably estimate the probabilities and rates of transition.

3. As in other models, determining the adequacy of the model depends on the ability to predict the behavior of the system.

Thus, the main objective of the proposed approach is to estimate the values of the specified parameters and further, based on the data obtained, to predict changes in the ecological state of the water area. The development of intelligent technology for detecting anomalies of ecosystems in the water area of the city of Sevastopol, based on the use of new

approaches and methods, will lead to an increase in the validity, reliability and efficiency of decision support processes for assessing the ecological state of the environment.

References

1. Ecological bases of management of natural-technical systems / Ed. Mn Fedorov. SPb .: Publishing house Polytech. University, 2007. - 506 p.

2. Ovsyany EI, Romanov A.S., Minkovskaya R.Ya., Krasnovid I.I., Ozyumenko B.A., Tsymbal I.M. The main sources of pollution of the marine environment of the Sevastopol region // Environmental safety of the coastal and shelf zones and the integrated use of shelf resources. Sevastopol: ECOSY-Hydrophysics, 2001.– P. 138-152.

3. Skatkov A.V., Bryukhovetsky A.A., Moiseev D.V. Intellectual monitoring system for solving large-scale scientific problems in cloud computing environments. Information Management Systems. - 2017. № 2 (87). Pp. 19-25.

4. E.N. Mashchenko, Yu.P. Nyukina. Approach to the virtualization of data monitoring systems of multidimensional objects // Environmental Control Systems - 2016 / Abstracts of the International Scientific and Technical Conference. - Sevastopol, October 24 - 27, 2016 - Sevastopol: IPTS, 2016. - 234 p.

5. Y.V. Doronin, V.O. Ryabovaya. Development of the method of restructuring information systems for environmental monitoring based on the analysis of options for their implementation // Environmental Control Systems - 2016 / Abstracts of the International Scientific and Technical Conference. - Sevastopol, October 24 - 27, 2016 - Sevastopol: IPTS, 2016. - 234 p.

6. Bryukhovetsky A.A., Skatkov A.V., Shishkin Yu.E. Simulation of anomaly detection processes in complexly structured monitoring data // Environmental control systems. Sevastopol: IPTS, 2017. Vol. 9 (29). - p. 45–49.

7. Truhaev R.I. Models of decision making under uncertainty. - M .: "Science", 1981. - 258 p.