Modeling and system analysis of drinking water parameters in urban water supply systems

V G Isakov, S V Vologdin, D S Ponomarev and M Y Dyagelev

1 Kalashnikov Izhevsk State Technical University, Studencheskaya st., 7, 426069, Izhevsk, Russia
2 Research Institute of the Federal Penitentiary Service of Russia (Izhevsk branch), Kommunarov st., 216, 426004, Izhevsk, Russia

E-mail: ponomarev.dmitry1990@mail.ru

Abstract. The article discusses the use of mathematical modeling and system analysis methods to solve problems of water preparation. Models for predicting indicators of drinking water based on the quality of the source water are proposed. The technological process of deodorization of drinking water is considered. The following system analysis methods were used in the models development: correlation analysis of data, the principal component method, regression modeling, and the least squares method. Hidden relationships between indicators of drinking and source water were identified. The adequacy of the obtained results was proved by comparing them with the actual values. The obtained results approximate well the actual values. The results can be used in water supply systems to improve the quality of drinking water.

1. Introduction

One of the pressing problems of our time is to obtain drinking water of adequate quality. Particularly relevant is the use of water resources, the use of drinking water, as well as the development of models in this area [1–3]. Because of the deteriorating state of drinking water supply sources and the impact of anthropogenic factors on the environment, these issues become more and more urgent. It is known that the parameters of the source water directly affect the quality of drinking water. Many factors need to be taken into account, including the distribution of algae, bacteria, microbiological indicators, environmental factors, seasonal changes in water supply sources [4–6]. It should be noted that deodorization of water is significantly important. One of the most common methods of deodorizing water is the use of sorbents and activated carbons. Today, there are many developments in this area [7, 8]. The research in this area at a theoretical level is of great interest as it would constitute resources for experiments and reduce the risk of disruption of water supply.

The use of systems analysis and mathematical modeling in various fields of science seems to be promising in recent years. The use of these methods in medicine and ecology is of special interest [9, 10]. Enterprises engaged in the preparation of drinking water can be considered as an open, complex system that can be divided into subsystems and interacting elements. The most popular methods, such as correlation data analysis, regression modeling, the principal component method, and the least squares method can be used as tools for analyzing and subsequent model development. It would help to identify hidden relationships between the parameters of the source and drinking water, develop
models that can be used to obtain a forecast of the drinking water quality based on the values of the source water parameters.

2. Selection of the conditions for the models development

First of all, it is necessary to analyze and identify the principal causes of the deterioration of the organoleptic properties of drinking water. The Municipal unitary enterprise Izhvodokanal and Izhevsk reservoir were chosen as an example. The problem of deodorizing drinking water was particularly acute in the enterprise in 2002-2004. In particular, consumers of drinking water noted the appearance of an earthy smell in the water, which corresponds to such compounds as geosmin, methyl isobenzene. In addition, there were problems with chlorides and chloroform as a result of reagent water treatment. Analysis of deodorization methods in the drinking water supply system has shown that one of the promising solutions is the use of sorbents, such as activated carbon.

The company introduced a deodorization system for drinking water using powdered activated carbon. The principal advantages of this sorbent are: low cost, the possibility of seasonal use, high adsorption activity, the possibility of use in any node of the technological scheme. Apart from that, there is no need to reconstruct the existing water treatment facilities. It should be noted that water treatment facilities should be considered as a complex system that has an internal structure, interaction of elements, possibility of division into subsystems, and influence of external factors from the environment. The system has the following properties: static, dynamic, synthetic.

The principal parameters of source and drinking water were selected to develop the model. The record of these parameters was kept monthly (from 2002 to 2014) in the enterprise. The most significant parameters that affect the organoleptic properties of drinking water were selected out. The sample consists of 62 points. The values of random variables were obtained in the interval \([0; 1]\).

The principal component method (PCA) \([11–14]\) was applied to eliminate multicollinearity between the parameters \((L_i)\). The reduction of the random variables values to a dimensionless form was carried out by formula (1), the values of random variables were obtained in the interval \([0; 1]\).

\[
\bar{x}_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}
\]

A matrix of correlation coefficients \((R)\) was calculated for indicators of source water \((L_i)\) (table 1).

| Parameter | \(L_1\) | \(L_2\) | \(L_3\) | \(L_4\) | \(L_5\) | \(L_6\) | \(L_7\) | \(L_8\) | \(L_9\) |
|-----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| \(L_1\)   | 1       | 0.44    | 0.32    | 0.01    | -0.14   | 0.51    | 0.43    | 0.49    | 0.36    |
| \(L_2\)   | 0.44    | 1       | 0.94    | 0.01    | -0.06   | 0.18    | -0.01   | 0.19    | 0.14    |
| \(L_3\)   | 0.32    | 0.94    | 1       | -0.06   | -0.10   | 0.07    | -0.11   | 0.11    | 0.09    |
| \(L_4\)   | 0.01    | 0.01    | -0.06   | 1       | -0.17   | -0.15   | 0.12    | 0.03    | 0.22    |
| \(L_5\)   | -0.14   | -0.06   | -0.10   | -0.17   | 1       | -0.05   | -0.38   | -0.37   | -0.49   |
| \(L_6\)   | 0.51    | 0.18    | 0.07    | -0.15   | -0.05   | 1       | 0.48    | 0.45    | 0.15    |
| \(L_7\)   | 0.43    | -0.01   | -0.11   | 0.12    | -0.38   | 0.48    | 1       | 0.70    | 0.53    |
| \(L_8\)   | 0.49    | 0.19    | 0.11    | 0.03    | -0.37   | 0.45    | 0.70    | 1       | 0.49    |
| \(L_9\)   | 0.36    | 0.14    | 0.09    | 0.22    | -0.49   | 0.15    | 0.53    | 0.49    | 1       |
The eigenvectors and eigenvalues of the matrix \( R \) are determined. The Jacobi rotation method was used to find eigenvectors and eigenvalues. There are three principal components that form the main data set \((P_1 - 36\%, P_2 - 22\%, P_3 - 15\%)\). The transition matrix into the selected principal components were determined (table 2).

**Table 2. The calculation of the principal components coefficients.**

| Parameter | \( P_1 (36\%) \) | \( P_2 (22\%) \) | \( P_3 (15\%) \) |
|-----------|------------------|------------------|------------------|
| \( L_1 \) | 0.41             | 0.14             | -0.20            |
| \( L_2 \) | 0.26             | 0.60             | 0.14             |
| \( L_3 \) | 0.21             | 0.62             | 0.19             |
| \( L_4 \) | 0.06             | -0.16            | 0.56             |
| \( L_5 \) | -0.28            | 0.19             | -0.42            |
| \( L_6 \) | 0.33             | -0.01            | -0.53            |
| \( L_7 \) | 0.42             | -0.33            | -0.11            |
| \( L_8 \) | 0.45             | -0.16            | -0.10            |
| \( L_9 \) | 0.38             | -0.20            | 0.32             |

Correlation was established between principal components \((P)\) and the concentration of geosmin in drinking water, mg/l \((K_2)\); chloroform concentration, mg/l \((K_3)\); chloride concentration, mg/l \((K_4)\). The results are presented in table 3.

**Table 3. The results of the correlation analysis for the principal components and parameters of drinking water.**

| Parameter | \( P_1 \) | \( P_2 \) | \( P_3 \) | \( K_2 \) | \( K_3 \) | \( K_4 \) |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| \( P_1 \) | 1.00      | -0.05     | -0.47     | -0.06     | -0.48     | 0.51      |
| \( P_2 \) | -0.05     | 1.00      | -0.13     | 0.62      | 0.12      | 0.16      |
| \( P_3 \) | -0.47     | -0.13     | 1.00      | -0.22     | 0.05      | -0.97     |
| \( K_2 \) | -0.06     | 0.62      | -0.22     | 1.00      | 0.09      | 0.24      |
| \( K_3 \) | -0.48     | 0.12      | 0.05      | 0.09      | 1.00      | 0.02      |
| \( K_4 \) | 0.51      | 0.16      | -0.97     | 0.24      | 0.02      | 1.00      |

The least squares method \((LSM)\) [15–17] was applied, equations for the concentration of odorizing substances in drinking water were developed. An equation (3) for the concentration of geosmin \((K_2)\) was developed:

\[ K_2 = 10^{-3}(-1.8P_1^3 + 4.2P_1 + 2.3P_2^3 - 0.03P_1 + 0.2P_2 + 0.2) , \]  

(3)

The concentration of chloroform \((K_3)\) has correlation with the first principal component \((P_1)\), Its equation was developed (4):

\[ K_3 = 0.4P_1^6 - 2.1P_1^5 + 4.4P_1^4 - 4.2P_1^3 + 1.7P_1^2 - 0.2P_1 + 0.1 , \]  

(4)

Correlation between the concentration of chlorides \((K_4)\) and the first \((P_1)\) and the third principal component \((P_3)\) was found with- A regression equation was developed (correlation coefficients \(r = 0.5\) and \(r = -0.9\)) (5):

\[ K_4 = 3.9P_1 - 5.4P_3 + 8.6 \]  

(5)

The developed models should be verified.

4. Verification of the results

The adequacy verification of the obtained results was carried out by using correlation coefficients \((r)\) and the Fisher criterion \((F)\) [18-21]. The actual and acquired values were compared by using the model.
In order to calculate the degree of statistical dependence between two numerical variables (model and actual values), the correlation coefficient (6) was calculated.

\[ r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}, \]  

where: \( n \) is the number of observations; \( x \) is the input variable; \( y \) is the output variable, \( \sigma \) - dispersion.

In this case, the correlation coefficients are: \( r_{K2} = 0.81; r_{K4} = 0.97; r_{K3} = 0.76 \).

The calculations are reliable if the correlation coefficient is higher than |0.5|. For all three parameters, this condition is met.

Formula (7) was used to calculate the Fisher criterion.

\[ F = \frac{\sigma_1^2}{\sigma_2^2} \]  

Where \( \sigma_1^2 > \sigma_2^2 \). In this case, the Fisher criterion: \( F_{K2} = 1.30; F_{K3} = 1.72; F_{K4} = 1.05 \). The calculations are reliable if the Fisher criterion coefficient in our case is no higher than 2.97 with \( P = 0.95 \).

5. Conclusions
1. Conducted correlation analysis of data allowed us to reveal hidden relationships between the parameters of the source (\( L_n \)) and drinking water (\( K_n \)). The use of this tool helps to study the processes of deodorization and the relationship between the indicators of water in water treatment facilities.

2. The models, that have been developed, can be used to obtain the values of drinking water parameters (geosmin (\( K_2 \)), chloroform (\( K_3 \)) and chlorides (\( K_4 \)) concentrations) based on the quality of the source water (parameters \( L_{1-9} \)). Application of the obtained results allows to predict and reduce the risk of deterioration in the quality of drinking water.

3. The method based on correlation data analysis, regression modeling, the principal component method, and the least squares method was developed. Thanks to this method the cost of conducting experimental studies aimed at improving the quality of water treatment systems can be reduced.

Acknowledgments
This work was supported by grant 27.06.01/18BCB of Kalashnikov ISTU.

References
[1] Alqadi K and Kumar L 2014 Water policy in jordan *International Journal of Water Resources Development* 10 322-34
[2] Droogers P and Bouma J 2014 Simulation modelling for water governance in basins *International Journal of Water Resources Development* 30 475-94
[3] Falkenmark M 2017 Water and human livelihood resilience: a regional-to-global outlook *International Journal of Water Resources Development* 33 181-97
[4] Siebel E, Wang Y and Egli T 2008 Correlations between total cell concentration, total adenosine tri-phosphate concentration and heterotrophic plate counts during microbial monitoring of drinking water *Drinking Water Engineering and Science* 1 1-6
[5] Bautista-De los Santos Q, Schroeder J, Sunghong R, Ijaz U, Sloan W, Sevillano-Rivera M and Pinto A 2016 Emerging investigators series: microbial communities in full-scale drinking water distribution systems-a meta-analysis *Environmental Science: Water Research and Technology* 2 631-44
[6] Gomez-Alvarez V, Pfaller S, Pressman J, Wahman D and Revetta R 2016 Resilience of microbial communities in a simulated drinking water distribution system subjected to
disturbances: role of conditionally rare taxa and potential implications for antibiotic-resistant bacteria. *Environmental Science: Water Research and Technology* 2 645-657

[7] Van Der Aa L, Rietveld L and Van Dijk J 2011 Effects of ozonation and temperature on the biodegradation of natural organic matter in biological granulated activated carbon filters. *Drinking Water Engineering and Science* 4 25-35

[8] Arefieva O, Zemukhova L, Morgun N, Tsvetnov M, Kovshun A, Panasenko A and Rybin V 2015 Removal of (2,4-dichlorophenoxo) acetic acid from aqueous solutions using low-cost sorbents. *Air Soil and Water Research* 8 59-65

[9] Gorokhov M, Ponomarev S, Vologdin S, Blagodatsky G and Ponomarev D 2017 Informatization of the quality of medical care in the management of the medical service of the federal penitentiary service of Russia. *Advances in Computer Science Research (ACSR)* 72 pp 481-4

[10] Isakov V and Ponomarev D 2015 Regression modeling of geosmin concentration in drinking water based on the data of MUE «Izhvodokanal» *Bulletin of the Izhavsk State Technical University* 26 107-8

[11] Vilanovaa M, Genishevab Z, Masaa A and Oliveira J 2010 Correlation between volatile composition and sensory properties in Spanish Albariño wines. *Microchemical Journal* 95 240-6

[12] Dougherty C 2011 *Introduction to econometrics* (Oxford University Press) p 573 p

[13] Harry T and Hildegarde H 2010 *Data relationships and multivariate applications food science text series* (New York: Springer New York) pp 433–49

[14] Pearson K 2011 On Lines and Planes of Closest Fit to Systems of Points in Space *Philosophical Magazine* 11 559–72

[15] Aldrich J 1998 Doing least squares: perspectives from Gauss and Yule *International Statistical Review* 66 61-81

[16] Charnes A, Frome E and Yu P 1976 The Equivalence of Generalized Least Squares and Maximum Likelihood Estimates in the Exponential Family. *Journal of the American Statistical Association* pp 169-71

[17] Stephen M. Stigler 1981 Gauss and the Invention of Least Squares *Ann. Stat.* 9 465-74

[18] Boddy R and Smith G 2009 *Statistical methods in practice: for scientists and technologists* (Chichester, U.K.: Wiley) pp 95–6

[19] Fisher R 1936 The Use of Multiple Measurements in Taxonomic Problems *Annals of Eugenics* pp 179-88

[20] Taylor R 1997 *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* Sausalito (CA: University Science Books) pp 217

[21] Venables W and Ripley B 2002 Modern Applied Statistics with S *Springer Verlag* 4 15-7