Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2018.22809

134 (2018) 135–142
December

Modeling of changes in trihalomethanes concentration in a wide water supply system: the case study of the Silesian agglomeration

Izabela Zimoch[a,*], Ewa Łobos[b]

[a]*Institute of Water and Sewage Engineering, Silesian University of Technology, Konarskiego 18, 44-100 Gliwice, Poland, Tel. 32 237 2884, 606 62 41 96; email: Izabela.Zimoch@polsl.pl*
[b]*Institute of Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland*

Received 25 March 2018; Accepted 4 July 2018

ABSTRACT

The presence of secondary micro-pollutants, including trihalomethanes, in water intended for human consumption is a common phenomenon in water supply systems, therefore tap water is subjected to constant quality monitoring. Relations between the large number of factors affecting water quality are complicated and it is difficult to form them in simple mathematical formulas which would define simulation models applicable in practice. This paper presents the detailed analysis of statistical properties of prediction models for trihalomethanes in tap water. All models were based on data collected in real water supply system during its standard exploitation from 2007 to 2017 and obtained by multiple regression. We propose a procedure which allows to select the model with the best predictive properties. It is shown also that the determination coefficient (R^2) for the model in implicit form does not determine whether the model may be useful in the decision-making process: it is possible that model in implicit form has low R^2 while the correlation between observed and estimated values of trihalomethane at endpoints is high, and vice versa. The best model depicted in the paper has correlation 0.76 between observed and predicted values which is a good result for data from real water supply systems.

Keywords: Trihalomethanes (THMs); Disinfection by-products; Water supply systems; Predictive models; Multiple regression

1. Introduction

Growing pollution of the natural environment, including surface water resources, is the result of dynamic economic development and increasing technological progress. Especially surface water is a resource exposed to various contaminants both natural and anthropogenic origins. Numerous annual reports of world environment state are pointing out that approximately one-third of organic compounds produced in the national economy goes into the environment. Therefore, the scarcity and poor quality of surface water resources force water supply operators to use not only coagulation and filtration but also highly effective processes that guarantee good water quality in terms of its physicochemical parameters [1–4]. Disinfection with chlorine application is still a commonly used process which guarantees the microbiological stability of water. However, unavoidable consequence of water chlorination is the formation of disinfection by-products (DBPs), including trihalomethanes (THMs). Numerous researchers confirm that THMs have mutagenic and carcinogenic properties, dangerous for human health [5–11], that is why the law relating to drinking water quality protects the consumer health. The directive of European Union 98/83/EC of 3 November 1998 which concerns drinking water quality, defines maximal permissible concentrations of DBPs (mainly THMs) in water delivered directly to consumers. Standard water treatment processes both coagulation and filtration do not guarantee

* Corresponding author.

Presented at the 13th Conference on Microcontaminants in Human Environment, 4–6 December 2017, Czestochowa, Poland.

1944-3994/1944-3986

the efficient removal of DBP precursors, consequently the risk of THMs formation in chlorinated water may be significantly increased. Therefore, the harmfulness of THM and other DBPs forces the necessity of complex technological processes of water treatment such as preliminary oxidation, ozonization, sorption on activated carbon, and double disinfection process using chlorine and chlorine dioxide [12–16]. All these processes together guarantee the production of high quality tap water and thus the risk of THM formation in tap water has been significantly reduced. The essential problem is that THMs arise not only in disinfection in water treatment plants (WTPs) but also during treated water transport by very large water pipe network. In big cities, the transport of water needs long times (e.g., a few days) due to oversized water distribution system resulting in low water flow rates (sometimes less than 0.01 m/s). This is a consequence of the extensive water pipes network and reduced water consumption due to consumers growing awareness. Therefore, to reduce the risk of THM formation and to ensure the safety of consumers' health, it is necessary to optimize the chlorine use in the whole water supply system by placing into service in water network additional points of chlorine doses. Thus the monitoring of THM concentration in water supply systems in both production and distribution subsystems is important and necessary. On the other hand, in water supply systems the precise monitoring of THM is expensive and time consuming so it is limited to indispensable minimal range.

The formation of THM is a complicated process occurring within the overall water supply system. The concentration of THM in drinking water in consumer's taps depends on many factors such as temperature, pH, concentration of natural organic matter, chlorine dose, reaction time, and other. This problem of THM formation in both water production and water distribution subsystems is intensively studied for several decades but still it is not explored in details [17–21]. Modern statistical tools allow their use for the mathematical description of the THM formation process. A lot of problems in statistical application of data and then the interpretation of obtained results explain the fact that there were only a few attempts made in order to define a mathematical model which describes changes of level of THM formation in time, in real extensive water pipe network [22–27]. Well-calibrated model, estimating the concentration of chlorination by-products in drinking water, can be used to predict the THMs concentration in water at consumer's tap. Such mathematical predicting models might be an important tool both in decision making in determining the technological parameters of the water treatment processes and in identifying the critical control points in the monitoring process.

The aim of this paper is to present a method of selecting the best predictive model from many models obtained by multiple regression. Two variants of multiple regression (backward elimination and forward selection) applied to different sets of both original and transformed data give usually many models. However, some of them have the similar statistical characteristics; therefore, the problem of choosing the best model arises. It will be shown that coefficient of determination (commonly used in such studies) is not a sufficient measure, especially if statistical model is obtained in implicit form. All considerations were conducted on the base of data collected in Silesian water supply system.

2. Materials and methods

2.1. Characteristics of research subject

The subject of research is the separated zone of Silesian water supply system. The system is one of the biggest systems not only in Poland but also in Europe and it spreads out over an area of 4,300 km^2 providing water to nearly 3.5 million inhabitants of the Silesian region. The origins of this system are dating back to the 80's of the 19th century and they are strongly associated with the development of the mining and metallurgical industries. Currently, the system is built of 11 WTPs which were fed with 87% of water from surface water resources (Fig. 1). The two greatest WTPs (Goczałkowice and Czaniec) take water from Wisła and Soła rivers, and then treat it with highly efficient technological processes. The current average daily water production of whole Silesian water supply system is over 381,500 m^3. This value of average production represents scarcely 45.5% of full available production capacity of analyzed water supply system.

From an operational point of view, the Silesian water supply system is divided into eight major networks. The distribution system uses the combination of pressure and gravity mains conveying the water to 42 local distribution subsystems and industrial bulk buyers. The total length of the water pipe network is 876.2 km and it consists mainly of the mainline in the large diameter range from 500 to 1,800 mm, characterized by considerable material diversity (Table 1).

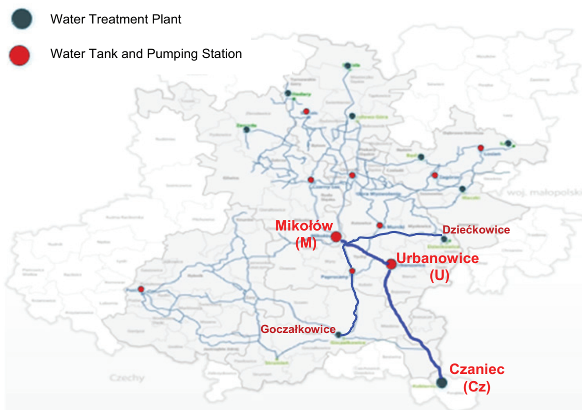

Fig. 1. Silesian water supply system – research subject.

Table 1
Materials structure of water pipe network in 2017[a]

Material	Length (km)	Percentage (%)
GRP	0.3	0.04
PE	102.3	11.70
Steel	620.7	70.77
Concrete	38.8	4.44
Grey iron	32.6	3.73
Ductile iron	81.5	9.32
Sum	876.2	100.00

[a]Archival data of Silesian Waterworks PLC.

Pipelines with diameters less than 500 mm cover only 12.8% of the total length. The average age of the pipelines is about 38 years (Fig. 2). Both materials and the age of water pipes affect the kinetics and concentration of THMs. Therefore, water company consequently invests in the modernization of water distribution network. Now only 20% of the pipelines (approximately 170 km) have been lined with cement-mortar, PU or PE to protect against corrosion and secondary water contamination.

An integral element of the water distribution system is nine complexes of network equalizing–storage tanks with the total capacity of 363,800 m^3 and five network pumping stations (Fig. 1).

A key role in the Silesian water supply system is played by the storage tanks in Mikołów, with the total capacity 96,000 m^3. Their convenient location allows gravity transport of water to the northern and western parts of the system. These tanks are supplied from three major WTPs: Goczałkowice, Czaniec, and Dziećkowice. WTP Czaniec, with average daily water production at the level of 66,000 m^3, delivers about 50% of water to storage tanks in Mikołów per day (via Urbanowice tank). This WTP intakes water from the Czaniec reservoir located on the Soła river, characterized by high seasonal variability of water quality. In particular, during spring rainfall, the amount of water turbidity significantly increases, often exceeding the level of 100 NTU. This causes difficulties in purification of water and leads directly on decreasing of the quality of tap water. The technological treatment system includes both contact filtration with an average dose of aluminium sulphate 1.02 mgAl/dm^3 (variation range from 0.37 to 5.52 mgAl/dm^3) and disinfection with chlorine (chlorine dose from 0.71 to 1.08 mg/dm^3). In order to protect microbiological stability of water chlorine is added both at WTP Czaniec and at Urbanowice tank (18,300 m^3). Chlorine use in the disinfection process and the long water transport to the Mikołów tanks (over 1 d) at the TOC concentration at level 1.44 mgC/dm^3 in treated water increases the risk of chlorine by-products formation, especially chloroform. The average concentration of chloroform in treated water at WTP Czaniec is 7.28 μg/dm^3, and it increases to 11.47 μg/dm^3 in tanks at Mikołów. Water from Mikołów tanks is delivered to consumer's tap, and this transport to farthest endpoints takes up to 6 d. Such long water transport results from low water flow with average speed 0.19 m/s (depending of water flow the speed changes from 0.02 to 0.33 m/s).

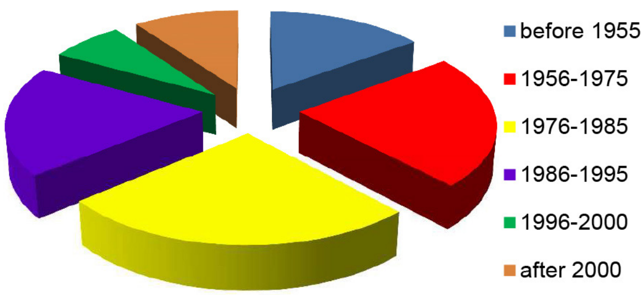

Fig. 2. Age structure of water pipe network in 2017 (based on archival data of Silesian Waterworks PLC).

From these reasons, quality of drinking water at Mikołów tanks should be as good as possible. WTP Czaniec plays an important role in supplying water to Mikołów tanks, thus the separated zone from Czaniec to Mikołów is the subject of presented study which initiates the complex research related to health risk assessment of Silesian inhabitants.

2.2. Research methods and data specification

The modeling of water quality, including changes of THMs concentration, may be based on statistical methods, especially multiple regression [18,21,26]. Statistical analysis and software give equations in which dependent variable (the concentration of THMs) is expressed as a function of selected independent variables. If there are nonzero correlations between inputs, then the equation may appear worthless for practical application or even it is possible that the equation does not exist. Research based on large sets of data may be examined for their suitability as reliable predictive models by adequate statistical tests at the assumed significance level (such as the Kolmogorov–Smirnov test, chi-square test, goodness of fit, correlation between observed and predicted values, analysis of statistical errors of estimation). For these reasons, data should be carefully selected due to following stages of research method. At the beginning, the correlations between all possible pairs of variables should be calculated. The most important in future prediction models will be these independent variables, which have significant correlation (positive or negative) with dependent variable. In the second stage, there are constructed sets of independent variables such that the correlations between them are close to zero or negligible. Above analyses are provided also for transformations of all variables (usually logarithms, powers, products or quotients of two variables). Then, using two variants of multiple regression (forward selection and backward selection), many different models with different statistical properties may be obtained. Therefore, the main research problem is the choice of the best predictive equation which would be applicable in management of water supply system. In this research method, a new approach to the solution of this problem is proposed. It is based on both determination coefficient and the analysis of goodness of fit for predictive equations in their explicit forms.

The study is based on data collected from January 2007 to September 2017 during normal exploitation of water supply system obtained from Silesian Waterworks PLC in Katowice. Technological parameters (chlorine dose and residual chlorine) were examined every day while quality parameters of tap water (temperature, pH, UV_{254}, UV_{272}, total organic carbon, THMs) were measured up to three times in each month according to monitoring plan. From this large set of data, the subset of records containing the concentration of THMs in Czaniec and Mikołów was excerpted. Finally, predictive models were built basing on 8,778 values of 17 water quality parameters and 2 operational parameters which were collected during 10 years by water company. All water samples were analyzed using standard methods for the examination of water. Table 2 presents basic descriptive statistics of these 19 variables which turned out to be significant in the generation of THMs during water transport from WTP Czaniec to Mikołów tanks.

The concentration of THMs at Mikołów is greater than at Czaniec and it is the result of an average 1-d reaction between chlorine and organic matter (water from WTP Czaniec to Urbanowice tank is transported during about 20 h, from Urbanowice tank to Mikołów tanks – about 9.5 h). It should be explained that the greater mean value of chlorine dose at Urbanowice is related to accidental failures in the system which happened a few times during examined time of exploitation.

The simulations by EPANET showed that water from Czaniec to Mikołów flows on average 1 d time so records at day N at Czaniec were paired with records at day $N + 1$ in Mikołów. This gave 142 records (of course with missing data because different parameters are tested according to different monitoring plans). Then the correlations between the concentration of THMs at Mikołów and variables at Czaniec and Urbanowice were calculated (Table 3). The highest correlations are noted between concentration of THMs at Mikołów and

- the concentration of THMs at Czaniec,
- the concentration of chloroform at Czaniec, because chloroform is the largest share of total THMs,
- temperature at Czaniec, which is similar to the temperature at Mikołów, because in higher temperature the chemical reaction is faster,
- residual chlorine at Urbanowice (where additional chlorine dose is applied), because the smaller residual chlorine, the greater chlorine usage per reaction of THMs forming.

Surprisingly, the correlation with pH is negative. Moreover, in further research from these data, a few models with negative coefficient in pH were obtained. Such situation is noted only in two papers [22,25].

Finally, for multiple regressions, there were considered following dependent variables:

- THM(M) – the concentration of THMs at Mikołów,
- ΔTHM = THM(M) – THM(Cz) – the increment of concentration of THMs from Czaniec to Mikołów,
- lnTHM(M),
- lnΔTHM,
- ln(1 + ΔTHM) – this variable was taken into consideration because it is possible that the increment of THMs is equal to zero or close to zero (which happens rarely).

Table 2
Basic descriptive statistics of considered variables

Variable	N	Mean	Minimum	Maximum	Standard deviation
Temperature (Cz) (°C)	410	10.67	0.70	24.00	6.39
Temperature (M) (°C)	378	10.90	1.00	22.40	6.41
pH (Cz)	425	7.22	6.39	7.76	0.16
pH (M)	437	7.31	7.00	7.63	0.13
UV_{254} (Cz) (cm^{-1})	420	0.12	0.05	0.21	0.03
UV_{254} (M) (cm^{-1})	131	0.10	0.03	0.19	0.03
UV_{272} (Cz) (cm^{-1})	420	0.09	0.01	0.17	0.02
Total organic carbon (Cz) (mg/L)	163	1.44	0.76	3.41	0.35
Total organic carbon (M) (mg/L)	145	1.40	0.76	2.31	0.29
Chlorine dose (Cz) (mg/L)	456	0.71	0.21	1.08	0.13
Chlorine dose (U) (mg/L)	467	1.68	0.00	5.42	0.83
Residual chlorine (Cz) (mg/L)	456	0.26	0.02	0.46	0.07
Residual chlorine (U) (mg/L)	467	0.18	0.14	0.33	0.04
Residual chlorine (M) (mg/L)	433	0.30	0.02	1.18	0.10
Bromates (Cz) (μg/L)	100	0.44	0.00	2.10	0.41
Chloroform (Cz) (μg/L)	414	7.28	1.30	19.80	3.69
Chloroform (M) (μg/L)	354	11.47	1.70	28.40	5.35
THM (Cz) (μg/L)	414	8.62	1.60	23.60	4.22
THM (M) (μg/L)	354	15.17	3.60	33.30	6.24

Table 3
Correlation coefficients between the concentration of trihalomethanes at Mikołów and variables at Czaniec and Urbanowice

Variable	Correlation coefficient	Variable	Correlation coefficient
Temperature (Cz)	0.6167	Chlorine dose (Cz)	0.1745
pH (Cz)	–0.3178	Chlorine dose (U)	0.1020
UV_{254} (Cz)	0.4306	Residual chlorine (U)	–0.6780
Total organic carbon (Cz)	0.4376	Chloroform (Cz)	0.7777
Residual chlorine (Cz)	–0.2956	THM (Cz)	0.7706

Independent variables considered in predictive models were:

- T(Cz) – temperature at Czaniec,
- pH(Cz) – pH at Czaniec,
- TOC(Cz) – total organic carbon at Czaniec,
- UV_{254}(Cz) – absorbance UV at 254 nm at Czaniec,
- UV_{272}(Cz) – absorbance UV at 272 nm at Czaniec,
- Br(Cz) – the concentration of bromates at Czaniec,
- THM(Cz) – the concentration of THMs at Czaniec,
- $CHCl_3$(Cz) – the concentration of chloroform at Czaniec,
- DCl_2(Cz) – chlorine dose applied at Czaniec,
- DCl_2(U) – chlorine dose applied at Urbanowice,
- RCl_2(Cz) – residual chlorine at Czaniec,
- RCl_2(U) – residual chlorine at Urbanowice,

and their logarithms. The six sets of independent variables were considered: pure variables with either THM(Cz) or $CHCl_3$(Cz), similarly natural logarithms but there was taken into account either lnpH or pH (because it is logarithm itself).

3. Results and discussion

The aim of presented study was to find the equation expressing the concentration of THMs at Mikołów as a function of operational and quality parameters at Czaniec and Urbanowice. Basing on 142 selected records (some with missing data), the statistical models were obtained by multiple regression. For all possible choices of dependent variable and sets of independent variables two variants of multiple regression were applied (the forward selection and backward elimination using STATISTICA software with $F_{\text{enter}} = 4$ and $F_{\text{delete}} = 3$). Although 60 equations were expected, finally 37 different models were obtained because in 5 cases backward eliminations gave no result and sometimes forward selection and backward elimination gave the same results. There were 22 models with at least one coefficient or constant term statistically insignificant or indicating process inconsistent with actual knowledge on THMs formation (e.g., such with negative coefficient in temperature or chlorine dose). At this stage, there were considered 15 predictive models with acceptable statistical characteristics. These models contain similar but independent variables as models reported in literature [17,18,20–22,24]. It should be stressed that coefficients in predictive equations are strongly related to operational conditions of whole water supply system as well as quality of drinking water.

From the set of 15 models with all terms statistically significant 6 models in explicit form with coefficient of determination less than 0.60 were removed (coefficients of determination in the case of models in implicit form were not considered because we are interested in the goodness of fit only for THMs concentration, not on its transformations).

In the end, nine models remained. The procedure of choosing the best one will be explained on the example on four of them. They were tested with respect to their residuals distributions, correlation between observed and predicted concentrations of THMs at Mikołów, and properties of relative errors of estimation. The four best models, based on 131 cases, are as follows:

model A:

$$\ln \mathrm{THM}(M) = 2,2798 + 0,0484\mathrm{CHCl}_3(Cz) + 0,0234T(Cz) - 2,0048RCl_2(U) + 2,0299UV_{254}(Cz) \pm 0,2286 \tag{1}$$

model B:

$$\Delta\mathrm{THM} = 3,3301 + 0,1626T(Cz) + 30,4509\mathrm{UV}_{254}(Cz) \pm 4,0518 \tag{2}$$

model C:

$$\ln(1 + \Delta\mathrm{THM}) = 1,6958 + 0,0141T(Cz) + 2,8140\mathrm{UV}_{254(Cz)} \pm 0,04193 \tag{3}$$

model D:

$$\ln \mathrm{THM}(M) = 1,6764 + 0,3958\ln \mathrm{CHCl}_3(Cz) + 0,1677\ln T(Cz) - 0,2928\ln \mathrm{RCl}_2(U) + 0,2342\ln \mathrm{UV}_{254}(Cz) \pm 0,1949 \tag{4}$$

All these models are in implicit form so each model above was rewritten in the explicit form and the correlation between observed and predicted values of THM(M) was calculated. The comparison of correlations for models in both implicit and explicit forms is shown in Table 4.

Although models B and C explain a small part of variability of dependent variable in its implicit form, they indicate quite good correlation between observed and predicted values of concentration of THMs at the endpoint Mikołów. All models have similar correlations so further analysis is necessary. In Fig. 3, the distributions of residuals (i.e., differences between observed and predicted values of dependent variables) are shown.

Table 4
Comparison of correlations for considered models in their implicit and explicit forms

Model	Standard estimation error for the model in its implicit form	R_a^2 for the model in its implicit form	R for the model in its implicit form	R for the model in its explicit form
A	0.2286	0.6598	0.8182	0.6920
B	4.0518	0.0983	0.3350	0.7601
C	0.4193	0.0712	0.2925	0.7571
D	0.1949	0.7527	0.8716	0.7583

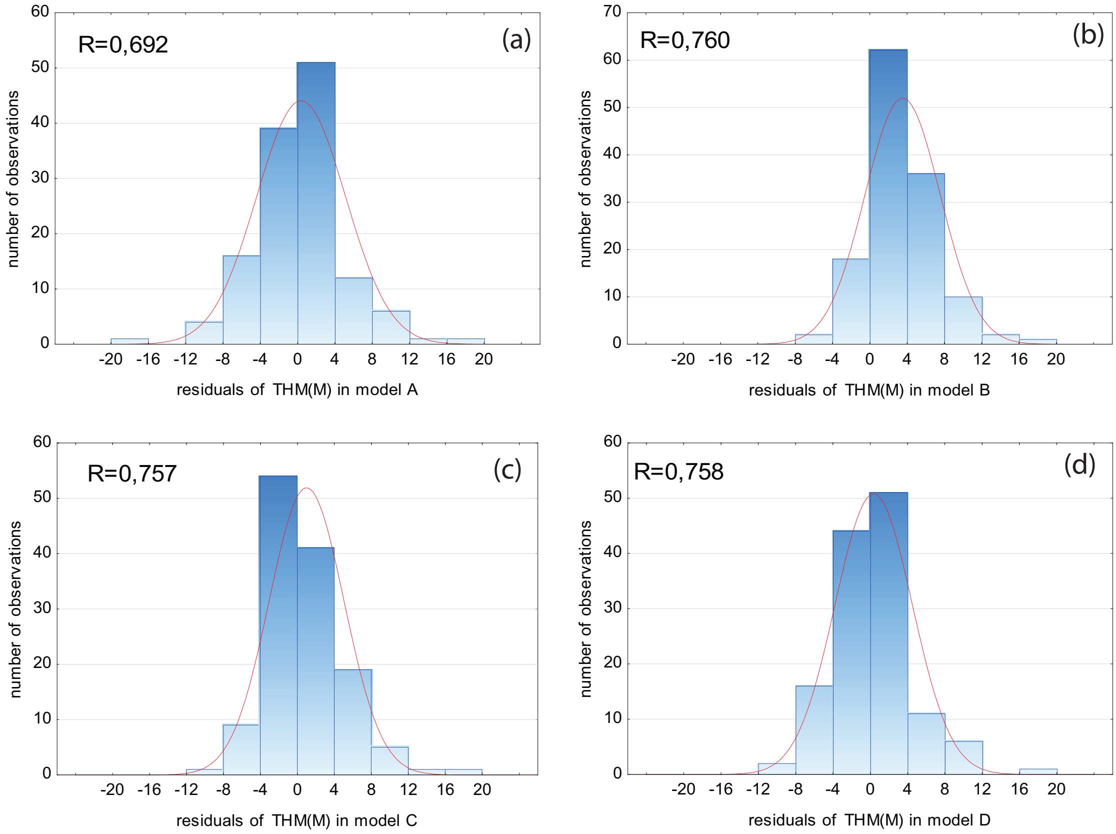

Fig. 3. Distributions of residuals of THM(M) in four models.

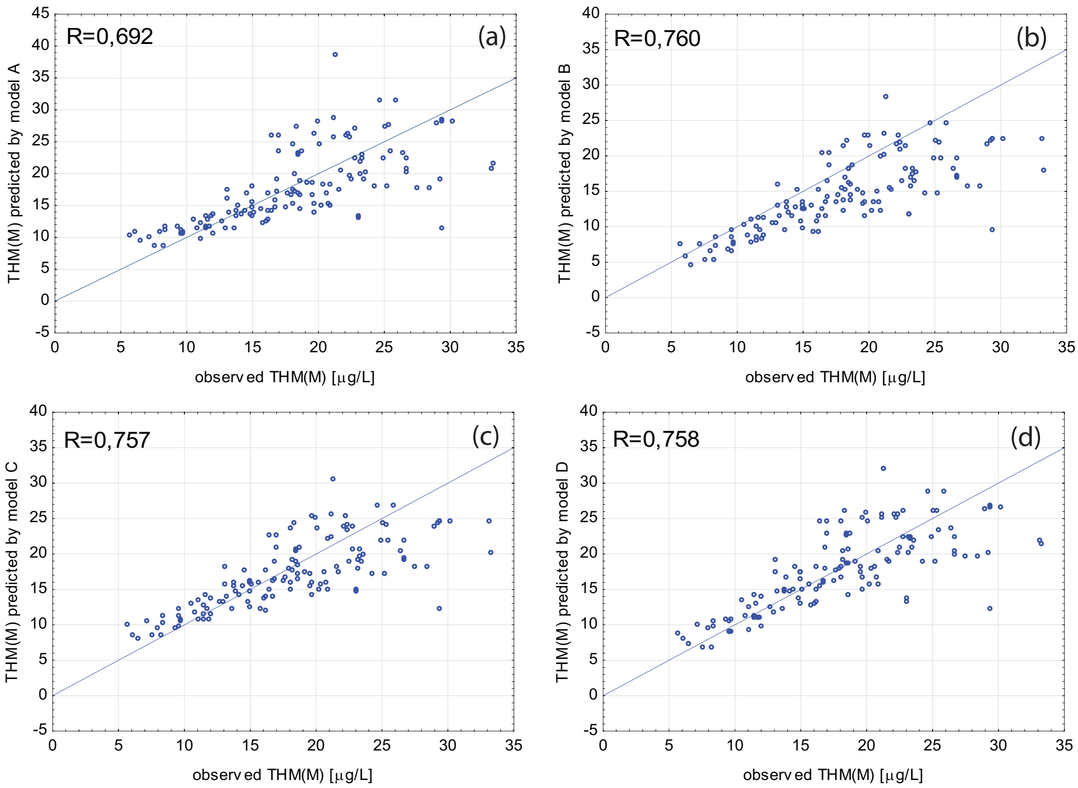

Fig. 4. Comparison of observed and predicted values of THM(M) in four models.

In all models residuals fit to normal distribution but only in the case of model B its mean is 3.38. In this model the majority of residuals is positive which means that the model usually gives values that are less than observed. If a model underestimates small values, it does not matter, but it would be worse if high values are underestimated. Indeed, model B is useless for the water company which may be seen in Fig. 4. Moreover, the best are models C and D, because in these two cases points on the graphs are much closer to line $y = x$.

The next step in the selection of the best model was to compare the number of correct predictions in the intervals: (3, 15), (15, 25), (25, 34). The intervals were chosen arbitrarily and the main idea was to find how models estimate low, medium and high concentrations of THMs. Note that although the maximal allowable concentration of THMs is 100 µg/L, concentrations close to 30 µg/L may be regarded high – from Mikołów water is transported further, up to 6 d, so the risk of exceeding the maximal allowable concentration of chloroform increases. The results are presented in Table 5.

Total number of correct estimations is: 86 for model A, 83 for B, 92 for C and D. Hence models C and D are comparable, but the decisive factor is the number of correct estimations in the interval with high THMs concentration. Model C predicts properly only one such case (per 18 possible) whereas model D – six cases. It is important to note that both C and D have a deficiency, that is, in one case (5.6%) they predict low THMs concentration while it is high. This must be borne in mind when models are used in decision-making processes.

Table 5
Number of observation with correct and incorrect estimations with respect to low, medium, and high concentrations of trihalomethanes at Mikołów

Model	Predicted values	Observed values		
		(3, 15)	(15, 25)	(25, 34)
A	(8, 15)	38	13	1
	(15, 25)	6	42	11
	(25,39)	0	14	6
B	(4, 15)	42	27	2
	(15, 25)	2	41	16
	(25, 29)	0	1	0
C	(8, 15)	36	8	1
	(15, 25)	8	55	16
	(25, 31)	0	6	1
D	(6, 15)	35	6	1
	(15, 25)	9	51	11
	(25, 33)	0	11	6

The analysis of relative errors of estimation (Table 6) gives helpful information if model will be applied in real water supply systems. The least dispersion is observed for model B; however, the mean disqualifies this model (values are underestimated by an average of 17%). Although model C has the best mean which is closest to zero, model D has slightly greater mean but less range of relative error of estimation. Thus, finally, model D is the best one.

4. Conclusions

The most important conclusion stemming from the research presented in the paper is that the coefficient of determination of the model in the implicit form does not settle the model usefulness (for example, model C with very low $R^2 = 0.0712$ is one of the four best predictive models). Of course mathematical characteristics of models are important but the purpose of building it should be taken into account. From the point of view of the water supply company, it is better to react to a false alarm (overestimated THMs concentration) than to put the consumer health to danger.

The presented model D is quite good as a mathematical model, especially considering the fact that it was based on data from the actual water supply system (and not on the laboratory data). The correlation between observed and predicted by model D concentrations of THM is 0.7583. Therefore it may be applied in water supply management. It should be stressed that, in this study, the contact time between chlorine and organic matter contained in water was approximately constant because the distance between initial point WTP Czaniec and endpoint Mikołów tanks is constant. Since the contact time is the key parameter to changes of THM concentration, the building of mathematical models predicting THMs concentration at different points on water pipe-network is more difficult because of varying contact time.

It is worth to continue study on mathematical models that can support the management and operation process. They are not expensive and simultaneously they could show how THMs concentration would change if some parameters were changed. The methodology of best model selection, presented in this case study for Silesian agglomeration, may be used in building THMs' prediction models in other water supply systems.

Acknowledgement

This work was supported by Ministry of Science and Higher Education Republic of Poland within statutory funds – project no BK-286/RIE-4/2017.

Table 6
Descriptive statistics of relative error of estimation

Model	N	Mean	Median	Minimum	Maximum	Standard deviation
A	131	–2.80%	1.95%	–82.21%	61.00%	25.03%
B	131	16.83%	19.12%	–32.84%	67.13%	17.72%
C	131	0.26%	0.26%	–78.24%	57.98%	20.98%
D	131	–1.92%	1.11%	–56.65%	58.00%	20.63%

References

[1] A. Matilainen, M. Vepsalainen, M. Sillanpaa, Natural organic matter removal by coagulation during drinking water treatment: a review, Adv. Colloid Interface Sci., 159 (2010) 189–197.

[2] M.J. Hedegaard, H.J. Albrechtsen, Microbial pesticide removal in rapid sand filters for drinking water treatment: potential and kinetics, Water Res., 48 (2014) 71–81.

[3] A. Matilainen, M. Sillanpaa, Removal of natural organic matter from drinking water by advanced oxidation processes, Chemosphere, 166 (2017) 497–510.

[4] A. Bhatnagar, M. Sillanpaa, Removal of natural organic matter (NOM) and its constituents from water by adsorption – a review, Chemosphere, 166 (2017) 497–510.

[5] K.P. Cantor, Ch.F. Lynch, M.E. Hildesheim, M. Dosemeci, J. Lubin, M. Alavanja, G. Craun, Drinking water source and chlorination by products in Iowa. III. Risk of brain cancer, Am. J. Epidemiol., 6 (1999) 552–560.

[6] M.J. Nieuwenhuijsen, R. Smith, S. Golfinopoulos, N. Best, J. Bennett, G. Aggazzotti, Health impacts of long-term exposure to disinfection by-products in drinking water in Europe: HIWATE, J. Water Health, 7 (2009) 185–207.

[7] N. Costet, C.M. Villanueva, J.J. Jaakkola, M. Kogevinas, K.P. Cantor, W.D. King, C.F. Lynch, M.J. Nieuwenhuijsen, S. Cordier, Water disinfection byproducts and bladder cancer: is there a European specificity? A pooled and meta-analysis of European case-control studies, Occup. Environ. Med., 68 (2011) 379–385.

[8] L. Zhang, Q. Cai, J. Lin, Y. Fang, Y. Zhan, A. Shen, L. Wei, L. Wang, J. Peng, Chloroform fraction of *Scutellaria barbata* D. Don promotes apoptosis and suppresses proliferation in human colon cancer cells, Mol. Med. Rep., 9 (2014) 701–706.

[9] T. Sasada, T. Hinoi, Y. Saito, T. Adachi, Y. Takakura, Y. Kawaguchi, Y. Sotomaru, K. Sentani, N. Oue, W. Yasui, H. Ohdan, Chlorinated water modulates the development of colorectal tumors with chromosomal instability and gut microbiota in *Apc*-deficient mice, PLoS One, 7 (2015) 1–15.

[10] S.E. Hrudey, L.C. Backer, A.R. Humpage, S.W. Krasner, D.S. Michaud, L.E. Moore, P.C. Singer, B.D. Stanford, Evaluating evidence for association of Human bladder cancer with drinking-water chlorination disinfection by-products, J. Toxicol. Environ. Health B Crit. Rev., 5 (2015) 213–241.

[11] I. Zimoch, E. Łobos, Evaluation of health risk caused by chloroform in drinking water, Desal. Wat. Treat., 57 (2016) 1027–1033.

[12] A. Matilainen, N. Vieno, T. Tuhkanen, Efficiency of the activated carbon filtration in the natural organic matter removal, Environ. Int., 32 (2006) 324–331.

[13] C. Ding, C. Shang, Mechanisms controlling adsorption of natural organic matter on surfactant-modified iron oxide-coated sand, Water Res., 44 (2010) 3651–3658.

[14] S.M. Korotta-Gamage, A. Sathasivan, A review: potential and challenges of biologically activated carbon to remove natural organic matter in drinking water purification process, Chemosphere, 167 (2017) 120–138.

[15] A. Chin, P.R. Berube, Removal of disinfection by-product precursors with ozone-UV advanced oxidation process, Water Res., 39 (2005) 2136–2144.

[16] S. Saraty, M. Mohseni, Effects of UV/H_2O_2 advanced oxidation on chemical characteristics and chlorine reactivity of surface water natural organic matter, Water Res., 44 (2010) 4087–4096.

[17] G.L. Amy, P.A. Chadik, Z.K. Chowdhury, Developing models for predicting trihalomethane formation-potential and kinetics, J. Am. Water Works Assoc., 79 (1987) 89–97.

[18] R. Sadiq, M.J. Rodriguez, Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review, Sci. Total Environ., 321 (2004) 21–46.

[19] S. Chowdhury, P. Champagne, P.J. McLellan, Factorial analysis of trihalomethanes formation in drinking water, Water Environ. Res., 82 (2010) 556–566.

[20] K. Ozdemir, I. Toroz, V. Uyak, Assessment of trihalomethane formation in chlorinated raw waters with differential UV spectroscopy approach, Sci. World J., 3 (2013) 1–7.

[21] I. Zimoch, E. Łobos, The optimization of chlorine dose in water treatment process in order to reduce the formation of disinfection by-products, Desal. Wat. Treat., 52 (2014) 3719–3724.

[22] S.K. Golfinopoulos, M. Xilourgidis, N. Kostopoulou, T.D. Lekkas, Use of multiple regression model for predicting trihalomethane formation, Water Res., 32 (1998) 2821–2829.

[23] L. Semerjian, J. Dennis, G. Ayoub, Modeling the formation of trihalomethanes in drinking waters of Lebanon, Environ. Monit. Assess., 149 (2009) 429–436.

[24] D. Mouly, E. Joulin, C. Rosin, P. Beaudeau, A. Zeghnoun, A. Olszewski-Ortar, J.F. Munoz, B. Welte, M. Joyeux, R. Seux, A. Montiel, M.J. Rodriguez, Variations in trihalomethane levels in three French water distribution systems and the development of a predictive model, Water Res., 44 (2010) 5168–5179.

[25] R. El-Shehawy, J. Awad, Monitoring and modeling of Fayoum distribution networks according to trihalomethanes formation, Civil Environ. Res., 2 (2012) 59–69.

[26] I. Zimoch, E. Łobos, T. Żaba, Drinking Water Quality Aspect in Ensuring the Safety of Water Supply Systems, T. Nowakowski, M. Młyńczak, A. Jodejko-Pietruczuk, S. Werbińska-Wojciechowska, Eds., Safety and Reliability: Methodology and Applications, Taylor & Francis Group, CRC Press/Balkema, 2014, pp. 717–726.

[27] P. Charisiadis, S.S. Andra, K.C. Makris, C.A. Christophi, D. Skarlatos, V. Vamvakousis, S. Kargaki, E.G. Stephanou, Spatial and seasonal variability of tap water disinfection by-products within distribution pipe networks, Sci. Total Environ., 506–507 (2015) 26–35.