

Izabela ZIMOCH¹ i Ewa ŁOBOS²

MODELOWANIE STĘŻENIA TRIHALOMETANÓW W SIECI WODOCIĄGOWEJ ZA POMOCĄ RÓŻNYCH RÓWNAŃ REGRESJI

THE MODELING OF TRIHALOMETHANES CONCENTRATION IN WATERPIPE NETWORK BASED ON DIFFERENT EQUATIONS OF REGRESSION

Abstrakt: Trihalometany (THM), jako uboczne produkty dezynfekcji wody chlorem, są związkami powstającymi w procesie dostarczania wody do konsumenta. Zarówno złożoność kinetyki reakcji, jak i duża liczba czynników wpływających na jej przebieg powodują trudności w szacowaniu poziomu stężenia trihalometanów w punktach kontroli jakości wody w systemie zaopatrzenia w wodę. W związku z rewizją w 2015 roku unijnej dyrektywy dotyczącej jakości wody przeznaczonej do spożycia zakres i częstotliwość monitoringu powinny uwzględniać ryzyko zdrowotne konsumenta w całym łańcuchu dostaw wody (ujęcie wody, uzdatnianie, magazynowanie, dystrybucja). Zatem rośnie znaczenie modeli szacowania stężenia trihalometanów w sieci wodociągowej, które uwzględniają parametry jakości wody oraz czynniki eksploatacyjne. Taki model może być istotnym narzędziem decyzyjnym w procesie eksploatacji systemów wodociągowych, ponieważ pozwoli na zarządzanie ryzykiem zdrowotnym przy jednoczesnym ograniczeniu kosztownych analiz laboratoryjnych. W pracy przedstawiono różne modele predykcji stężenia trihalometanów w sieci wodociągowej. Analizowano modele uzyskane za pomocą regresji wielorakiej dla różnych zbiorów danych oraz transformacji THM-ów i zmiennych niezależnych. Porównano wpływ różnych transformacji zmiennych na dokładność predykcji oraz ocenę statystyczną modelu. Badania zostały oparte na danych rzeczywistych, zebranych podczas siedmiu lat eksploatacji jednego z dużych polskich systemów zaopatrzenia w wodę.

Słowa kluczowe: trihalometany, system zaopatrzenia w wodę, modele matematyczne, regresja wieloraka

Wprowadzenie

W procesie eksploatacji systemów wodociągowych zasilanych wodami powierzchniowymi jednym z poważnych problemów jest kontrola generowania trihalometanów. Trihalometany powstają wskutek reakcji wydzielonej grupy substancji organicznych zawartych w wodzie z chlorem. Chlor jest powszechnie stosowanym środkiem ze względu na niskie koszty i wysoką skuteczność procesów dezynfekcji, zapewniającym stabilność mikrobiologiczną wody wodociągowej. Ponieważ trihalometany są związkami szkodliwymi dla zdrowia człowieka [1-3], a powstają nie tylko podczas uzdatniania wody, ale także podczas jej transportu rozległymi sieciami wodociągowymi [4], konieczne jest wdrożenie odpowiednich procedur związanych z zachowaniem wysokiej jakości wody, zgodnie z Dyrektywą Unii Europejskiej dotyczącej jakości wody przeznaczonej do spożycia przez ludzi [5] i jej rewizją z 2015 roku [6]. Monitoring stężeń THM-ów w systemach zaopatrzenia w wodę jest konieczny i istotny, lecz pociąga za sobą znaczne koszty, zatem ograniczany jest do niezbędnego minimum.

¹ Instytut Inżynierii Wody i Ścieków, Politechnika Śląska, ul. S. Konarskiego 18, 44-100 Gliwice, tel. 32 237 16 98, fax 32 237 10 47, email: Izabela.Zimoch@polsl.pl

² Instytut Matematyki, Politechnika Śląska, ul. Kaszubska 23, 44-100 Gliwice, tel. 32 237 12 03, fax 32 237 28 64, email: Ewa.Lobos@polsl.pl

Praca była prezentowana podczas konferencji ECOpole'16, Zakopane, 5-8.10.2016

Sam proces powstawania THM-ów nie jest ciągle dokładnie rozpoznany, mimo prowadzonych od wielu lat badań [7-9]. Ze względu na rosnące zanieczyszczenie antropogeniczne wody i dynamicznie zmieniające się warunki eksploatacji systemów wodociągowych rośnie znaczenie wykorzystania dobrych modeli matematycznych do symulacji stężeń THM-ów w sieci. Jedną z popularnych metod opracowywania modeli generowania THM-ów jest regresja wieloraka. Przegląd literatury wskazuje, że równania regresji ściśle zależą od indywidualnych cech i warunków eksploatacji systemów zaopatrzenia w wodę [10, 11]. Osiągnięcia współczesnej nauki jednoznacznie wskazują, że najlepsze modele (o wysokim współczynniku determinacji) uzyskuje się dla danych laboratoryjnych, zaś modele oparte na danych rzeczywistych zebranych z punktów monitoringu sieci dystrybucji są obciążone z reguły dużym błędem losowym [11].

Badania opisane w literaturze przedmiotu wskazują na różnorodne zmienne wpływające na końcowe stężenie THM w rzeczywistych systemach zaopatrzenia w wodę. Publikowane są modele wykorzystujące regresję wieloraką stosowaną bezpośrednio do danych surowych, jak również transformacji zmiennych niezależnych oraz zmiennej zależnej (najczęściej są to logarytmy tych zmiennych - naturalne lub dziesiętne). Czasem buduje się modele, w których wprowadza się dodatkowe zmienne (np. iloczyny bądź ilorazy pewnych zmiennych niezależnych). Modele o dobrych charakterystykach statystycznych mogą stanowić użyteczne, praktyczne i tanie narzędzie w zarządzaniu systemami wodociągowymi z uwzględnieniem analizy bezpieczeństwa [12] i oceny ryzyka zdrowotnego, którą wprowadziła ostatnia rewizja dyrektywy UE dotyczącej jakości wody przeznaczonej do spożycia [6].

W artykule przedstawiono najlepsze modele predykcji stężenia THM-ów w punktach monitoringu sieci wodociągowej. Uwzględniono wpływ różnych czynników na wielkość stężenia generowanych THM-ów, takich jak: odczyn wody (pH), absorbancja UV w 272 nm, utlenialność, OWO, przewodnictwo, temperatura wody, dawka chloru, stężenie chloru pozostałego. Porównano różne modele wynikające z regresji krokowej (postępującej oraz wstecznej) dla podanych zmiennych oraz ich logarytmów. Badania zostały oparte na danych zgromadzonych podczas badań prowadzonych w ciągu siedmiu lat w rzeczywistym systemie zaopatrzenia dostarczającym wodę mieszkańcom Krakowa.

Metoda regresji wielorakiej jako techniki badawczej

Metoda regresji wielorakiej jest metodą statystyki matematycznej [13], w której poszukuje się zależności między zmienną zależną (wyjaśnianą) a kombinacją liniową zmiennych niezależnych (wyjaśniających). Do modelowania stężenia trihalometanów w punktach sieci (THM_{PS}) najczęściej przyjmuje się następujące postacie równań:

$$THM_{PS} = a_0 + \sum_{i=1}^n a_i x_i, \quad (1)$$

$$\ln THM_{PS} = a_0 + \sum_{i=1}^n a_i x_i, \quad (2)$$

$$THM_{PS} = a_0 + \sum_{i=1}^n a_i \ln x_i, \quad (3)$$

$$\ln THM_{PS} = a_0 + \sum_{i=1}^n a_i \ln x_i, \quad (4)$$

gdzie $x_i, i = 1, 2, \dots, n$ są zmiennymi niezależnymi, zaś współczynniki $a_i, i = 0, 1, \dots, n$ są wyznaczone metodą najmniejszych kwadratów.

W badaniach prezentowanych w tej pracy uwzględniono wszystkie powyższe warianty oraz dodatkowo uwzględniono stężenie THM-ów w zakładzie uzdatniania wody (THM_{ZUW}) i przyjęto następujące transformacje zmiennej zależnej: $\Delta THM = THM_{PS} - THM_{ZUW}$, $\ln \Delta THM$, $\ln(1 + \Delta THM)$. Dla zmiennych niezależnych, których wartości były bliskie zera, rozważano zarówno zmienne postaci $\ln x_i$, jak i $\ln(1 + x_i)$. Do wyznaczenia równań regresji zastosowano regresję krokową postępującą i wsteczną. W obu przypadkach przyjęto wartości statystyki F równe 4 (do wprowadzenia zmiennej do równania) i 3 (do usunięcia zmiennej z równania). Obliczenia wykonano w pakiecie STATISTICA.

Przedmiot i zakres badań

Jakość wody dostarczanej mieszkańcom Krakowa ściśle związana jest z warunkami eksploatacji rozbudowanego podsystemu dystrybucji wody, do którego woda włączana jest z czterech niezależnych układów zasilania pracujących w oparciu o zasoby wód powierzchniowych (rzeki: Dłubnia, Rudawa, Sanka i zbiornik zaporowy Dobczyce na Rabie) i jednego układu zasilania Mistrzejowice, ujmującego wody podziemne. Kluczowy dla krakowskiego wodociągu jest układ zasilania Raba, który ze względu na jego dyspozycyjną zdolność produkcyjną (207 000 m³/d) jest w stanie zagwarantować średniodobowe zapotrzebowanie na wodę. Ponadto z układu tego woda dostarczana jest do ponad 50% mieszkańców Krakowa.

Tabela 1

Metody i urządzenia pomiarowe stosowane do wykonywania analiz

Table 1

Methods and measurement devices used to perform analyses

Wskaźnik	Metoda/urządzenie
odczyn (według [14])	laboratoryjny aparat wielofunkcyjny inoLab Multi Level 1
temperatura (według [15])	termometr elektryczny PT- 105 firmy Elmetron
chlor (według [16])	fotometr PC Compact Chlor
UV ₂₇₂ (według [17])	spektrofotometr HACH DR 4000 U
utlenialność (według [18])	metoda miareczkowa
OWO (według [19])	analizator TOC 5050 Shimadzu
przewodnictwo (według [20])	konduktometr
Σ THM (według [21])	chromatograf gazowy Agilent Technologies GC7890B z MSD5977A

Rozległy podsystem dystrybucji wody buduje złożona sieć wodociągowa o długości ponad 1800 km, współpracująca z 11 układami zbiorników sieciowych, o łącznej pojemności 276 200 m³. Infrastruktura liniowa krakowskiego wodociągu charakteryzuje się zarówno zróżnicowaniem wiekowym (przewody z początku ubiegłego wieku), jak i materiałowym z ponad 50% udziałem przewodów żeliwnych i stalowych (26% sieć żeliwna i 32% sieć stalowa). Stabilizację hydrauliczną pracy rozległego podsystemu dystrybucji wody zapewnia dyspozytornia Nastawnia Piaski Wielkie (NPW).

Pobór próbek odbywał się zgodnie z PN-ISO 5667-5:2003 [22] oraz PN-ISO 5667-6:2003 [23]. Wskaźniki jakości badanych próbek wody oznaczono zgodnie z procedurami polskich norm podanych w tabeli 1.

Wyniki badań

Podstawowe statystyki opisowe badanych zmiennych przedstawiono w tabeli 2. Większość parametrów w zakładzie uzdatniania wody (ZUW) była badana codziennie, wyjątkiem jest stężenie THM-ów, które monitorowane jest rzadziej ze względu na znaczne koszty analiz laboratoryjnych. W analizowanym zbiorze znalazły się tylko te przypadki, dla których znane było stężenie THM-ów w punktach monitoringu sieci.

Tabela 2

Statystyki opisowe badanych zmiennych

Table 2

Descriptive statistics of collected variables

Zmienna	Nr	Średnia	Mediana	Minimum	Maksimum	Odch. std.	Wsp. zmn.
odległość punktu od NPW [km]	394	8,08	8,40	2,13	13,70	3,13	38,71
czas przepływu ZUW-NPW [h]	394	11,84	11,70	8,73	16,87	1,40	11,82
chlor wolny w ZUW [mg/dm ³]	394	0,40	0,40	0,27	0,65	0,06	15,96
chlor wolny w punkcie [mg/dm ³]	394	0,11	0,10	0,00	0,50	0,07	60,87
całkowita dawka chloru [mg/dm ³]	394	1,22	1,13	0,73	2,33	0,30	24,56
pH ZUW	394	7,84	7,83	7,44	8,10	0,11	1,42
pH w punkcie	394	7,78	7,80	7,20	8,10	0,14	1,77
temperatura wody w ZUW [°C]	394	11,63	9,85	1,70	23,60	6,91	59,43
UV272 ZUW [cm ⁻¹]	384	0,02	0,02	0,01	0,06	0,01	28,64
UV272 w punkcie [cm ⁻¹]	11	0,01	0,01	0,01	0,02	0,00	18,61
utlenialność w ZUW [mg/dm ³]	394	1,07	1,00	0,60	1,70	0,25	23,25
OWO w ZUW [mg/dm ³]	185	1,90	1,82	1,00	3,25	0,46	24,34
OWO w punkcie [mg/dm ³]	367	1,57	1,57	0,01	5,30	0,66	42,42
przewodnictwo wł. w ZUW [μS/cm]	394	301,65	300,00	230,00	373,00	30,35	10,06
przewodnictwo wł. w punkcie [μS/cm]	394	322,76	311,5	215,00	612,00	57,07	17,68
ΣTHM w ZUW [μg/dm ³]	190	5,14	4,80	0,00	13,80	3,71	72,05
ΣTHM w punkcie [μg/dm ³]	394	18,61	16,75	0,50	48,60	9,34	50,22

Do dalszych analiz użyto 190 przypadków, w których nie występowały braki danych. Po zastosowaniu regresji krokowej otrzymano 33 równania (czasem regresja postępująca dawała ten sam wynik co regresja wsteczna). Odrzucono te równania, które miały co najmniej jeden współczynnik statystycznie nieistotny lub nie były zgodne z aktualną

wiedzą naukową dotyczącą powstawania THM-ów (np. odrzucono modele o ujemnym współczynniku przy dawce chloru). Ostatecznie uzyskano zaledwie pięć modeli:

model A

$$THM_{PS} = 7,10967 + 0,47939 T_{ZUW} + 1,2826 THM_{ZUW} \pm 6,1491,$$

model B

$$\ln THM_{PS} = 2,12830 + 0,02356 T_{ZUW} + 0,07977 THM_{ZUW} \pm 0,43084,$$

model C

$$THM_{PS} = 7,52077 + 8,42645 \ln(\Delta_1 Cl_2) + 7,91600 \ln(1 + THM_{ZUW}) \pm 6,3187,$$

model D

$$\ln(\Delta THM) = 2,05879 + 0,65770 \ln(\Delta_2 Cl_2) + 0,23708 \ln(1 + THM_{ZUW}) \pm 0,62236,$$

model E

$$\ln(1 + \Delta THM) = 2,20096 + 0,56836 \ln(\Delta_2 Cl_2) + 0,21814 \ln(1 + THM_{ZUW}) \pm 0,50856,$$

gdzie: $\Delta_1 Cl_2$ - zużycie chloru od zakładu uzdatniania wody do dyspozytorni Nastawnia Piaski Wielkie (ZUW-NPW), $\Delta_2 Cl_2$ - zużycie chloru na odcinku ZUW-punkt sieci.

Współczynniki determinacji dla uzyskanych modeli

Tabela 3

Determination coefficients for obtained models

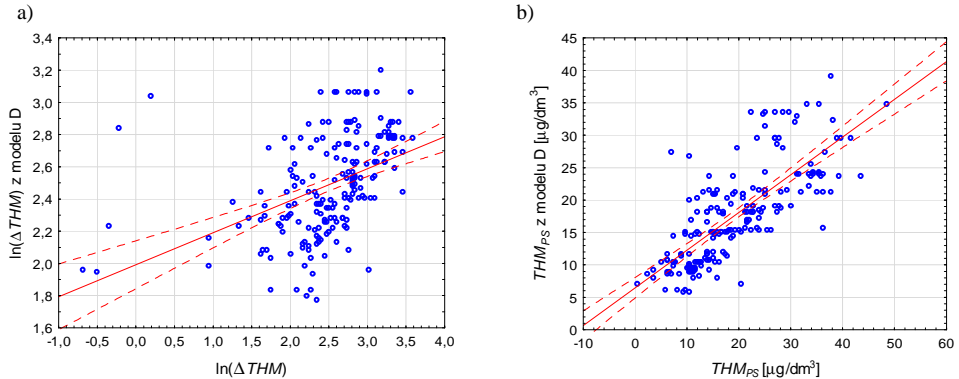
Table 3

Skorygowany współczynnik determinacji (R^2)	Model				
	A	B (B')	C	D (D')	E (E')
dla modelu w postaci uwikłanej	nie dotyczy	0,49	nie dotyczy	0,19	0,22
dla modelu w postaci jawnej	0,59	0,74 (0,72)	0,57	0,60 (0,57)	0,58 (0,56)

Dwa z tych modeli (A i C) są w postaci jawnej, pozostałe w postaci uwikłanej. Porównanie współczynników determinacji (tab. 3) wskazywałoby, że najlepszy jest model A, gdyż wyjaśnia on 59% zmienności zmiennej zależnej. Jednak po wyznaczeniu zmiennej THM_{PS} z modeli B, C, D i obliczeniu dla nich współczynników determinacji okazuje się, że model B wyjaśnia 74% zmienności zmiennej THM_{PS} . Warto zwrócić uwagę na znaczący wzrost współczynników determinacji dla modeli D i E w postaci jawnej w porównaniu z ich postaciami uwikłanymi. Przykładowo dla modelu D (rys. 1) w postaci jawnej punkty na wykresie rozrzutu są bardziej skupione wokół dwusiecznej pierwszej ćwiartki niż dla postaci uwikłanej.

Z modelami w postaci uwikłanej wiąże się jeszcze jeden problem. Metoda wyznaczania współczynników równania regresji gwarantuje, że średni błąd (różnica między wartościami obserwowanymi a estymowanymi z równania) jest równy zero. Jednak modele uzyskane w postaci uwikłanej po przekształceniu do postaci jawnej mogą mieć średni błąd różny od zera. Dokonano więc kolejnej transformacji - przesunięcia zmiennej

THM_{PS} o średnią wartość błędu (tzn. w modelu B o 1,04, D o 1,71, E o 1,46), otrzymując w ten sposób modele B', D', E'. Współczynniki determinacji nowych modeli zmniejszyły się o 2-3% (tab. 3).



Rys. 1. Wykresy rozrzutu wartości obserwowanych i estymowanych dla modelu D w postaci uwikłanej (a) i jawnej (b)

Fig. 1. Scatter plots for observed vs. estimated values for model D in the implicit form (a) and explicit form (b)

Na rysunku 2 przedstawiono histogramy reszt dla czterech modeli (model E' pominięto, ponieważ jest podobny do D') - model B' ma najwięcej reszt „małych”, jednak we wszystkich modelach zdarzają się niepożądane reszty większe. W zarządzaniu systemami zaopatrzenia w wodę ważną miarą oceny jakości modeli predykcji może być frakcja wartości, które w rzeczywistości przekroczyły normę, a model wskazuje na wartość dopuszczalną. Jednak tej oceny nie można dokonać dla podanych modeli, ponieważ zbiór danych nie zawierał wartości $THM_{PS} > 100\mu\text{g}/\text{dm}^3$.

Tabela 4

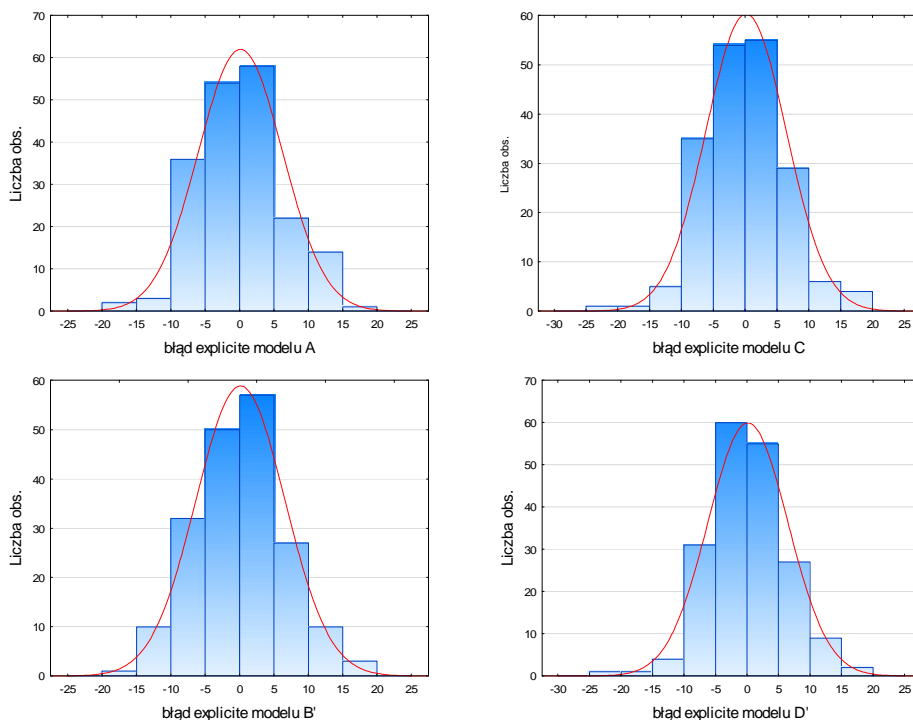
Statystyki błędu względnego dla uzyskanych modeli

Table 4

Statistics of relative error for obtained models

Model	Średnia	Mediana	Dolny kwartył	Górny kwartył	Wariancja
A	-24,53%	0,07%	-27,79%	18,54%	240,23%
C	-18,02%	-0,30%	-32,23%	19,38%	67,62%
B'	-24,54%	1,29%	-29,31%	17,40%	258,99%
D'	-23,32%	-0,96%	-32,13%	17,37%	171,13%
E'	-23,32%	-0,84%	-32,31%	17,07%	168,81%

Interesujące może być porównanie modeli ze względu na popełniane przy estymacji błędy względne (tab. 4). Wszystkie modele mają podobne charakterystyki błędu względnego (średnio przewidują o 20% wyższe stężenie THM, mediany są bliskie zeru, podobne dolne i górne kwartyły) z wyjątkiem wariancji. Z tego punktu widzenia najlepszy jest model C, który ma najmniejszą wariancję błędu względnego.



Rys. 2. Histogramy błędów bezwzględnych postaci jawnych modeli A, C, B' i D'

Fig. 2. Histograms of absolute errors for models A, C, B' and D' in their explicit forms

Podsumowanie i wnioski

Regresja wieloraka jest prostą i powszechnie stosowaną metodą w modelowaniu stężenia trihalometanów w systemach zaopatrzenia w wodę. Aby modele uzyskane tą metodą mogły być stosowane w praktyce (zarządzanie procesami technologicznymi, estymacja ryzyka zdrowotnego, ustalenie maksymalnych dopuszczalnych stężeń), muszą w miarę wiernie opisywać zjawisko generowania THM-ów oraz być statystycznie istotne.

W naukach fizycznych i technicznych przyjmuje się, że współczynnik determinacji powinien być większy niż 0,6, tzn. model powinien wyjaśniać co najmniej 60% zmienności zmiennej zależnej, natomiast modele o współczynniku determinacji mniejszym niż 0,25 uważa się za praktycznie nieużyteczne [13]. Jak pokazano w niniejszej pracy, nie należy odrzucać a priori modeli w postaci uwikłanej o niskim R^2 , ponieważ dla postaci jawnej współczynnik ten może być dużo wyższy (tak jest dla modeli D i E). Warto też zwrócić uwagę, że dla modelu B' uzyskano $R^2 = 0,72$, co jest wartością dużą dla modeli opartych na pomiarach w rzeczywistych systemach dystrybucji wody [11]. Jednakże wartość R^2 nie powinna być jedynym kryterium.

Z praktycznego punktu widzenia ważny jest odsetek predykcji z niedomiarem, zwłaszcza gdy wartość rzeczywista przekracza normę, a prognoza nie przewiduje

niebezpieczeństwa przekroczenia maksymalnego dopuszczalnego stężenia THM-ów. W tym aspekcie w ocenie modelu istotna jest analiza wielkości błędu względnego i prawdopodobnie zastosowanie - na etapie budowy modelu - wyższych wag dla przypadków o wyższym stężeniu THM w punktach monitoringu, co wymaga dalszych badań.

Literatura

- [1] Richardson S. *Trend Anal Chem.* 2003;22(10):666-684. DOI: 10.1016/S0165-9936(03)01003-3.
- [2] Richardson S, Plewa M, Wagner E, Schoeny R, DeMarini D. *Mutation Res.* 2007;636:178-242. DOI: 10.1016/j.mrrev.2007.09.001.
- [3] WHO. *Disinfectants and disinfectant by-products. Environmental Health Criteria 216.* Geneva: World Health Organization; 2000. http://www.who.int/ipcs/publications/ehc/ehc_216/en.
- [4] Włodyka-Bergier A, Bergier T, Kot M. *Desalin Water Treat.* 2014;52:3898-3907. DOI: 10.1080/19443994.2014.887490.
- [5] *Drinking Water Directive 98/83/EC. Council Directive of November 1998 on the quality of water intended for human consumption.* Off J EC. 1998; L 330.5.12.98:32-54. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31998L0083&from=EN>.
- [6] *Commission Directive (EU) 2015/1787 of 6 October 2015 on the quality of water intended for human consumption.* Off J EU. 2015; L 260: 6-17. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3A0J.L_.2015.260.01.0006.01.ENG.
- [7] Bellar TA, Lichtenberg RC, Kroner C. *J Am Water Works Assoc.* 1974;66(12):703-707. <https://www.awwa.org/publications/journal-awwa/abstract/articleid/8737.aspx>.
- [8] Zbieć E, Dojlido JR. *Ochr Środ.* 1999;3(74):37-44. http://www.os.not.pl/docs/czasopismo/1999/Zbiec_3-1999.pdf.
- [9] Zimoch I. *Zintegrowana metoda analizy niezawodności funkcjonowania i bezpieczeństwa systemów zaopatrzenia w wodę.* Gliwice: Wyd Politechniki Śląskiej; 2011.
- [10] Sadiq R, Rodriguez MJ. *Sci Tot Environ.* 2004;321:21-46. DOI: 10.1016/j.scitotenv.2003.05.001..
- [11] Zimoch I, Łobos E. Trihalomethanes in drinking water - the review of predictive models. In: Zimoch I, editor. *Current Issues in Water Treatment and Distribution*, vol. 5. Gliwice: Instytut Inżynierii Wody i Ścieków Politechnika Śląska; 2015;235-254.
- [12] Zimoch I, Łobos E, Zaba T. Drinking water quality aspect in ensuring the safety of water supply systems. In: Nowakowski T, et al, editors. *Safety and Reliability: Methodology and Applications.* London: Taylor & Francis Group; 2014;717-726. <http://www.crcnetbase.com/doi/abs/10.1201/b17399-100>.
- [13] Stanisław A. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny.* Kraków: StatSoft; 2007.
- [14] PN-EN ISO 10523:2012, Jakość wody - odczyn pH. Warszawa: PKN; 2012. <http://sklep.pkn.pl/pn-en-iso-7887-2012p.html>.
- [15] PN-77/C-04584, Woda i ścieki - pomiar temperatury. Warszawa: PKN; 1977. <http://sklep.pkn.pl/pn-c-04584-1977p.html>.
- [16] PN-ISO 7393-2:1997, Jakość wody - oznaczanie chloru wolnego i chloru ogólnego. Warszawa: PKN; 1997. <http://sklep.pkn.pl/pn-iso-7393-2-1997p.html>.
- [17] PN-84/C-05472, Woda i ścieki - Oznaczenia zawartości rozpuszczonych związków organicznych w wodzie metoda spektrofotometrii w nadfiolecie. Warszawa: PKN; 1984. <http://sklep.pkn.pl/pn-c-05472-1984p.html>.
- [18] PN-EN ISO 8467:2001, Jakość wody - oznaczanie indeksu nadmanganianowego. Warszawa: PKN; 2001. <http://sklep.pkn.pl/pn-en-iso-8467-2001p.html>.
- [19] PN-EN-1484-1999, Analiza wody - wytyczne oznaczenia ogólnego węgla organicznego (OWO) i rozpuszczonego węgla organicznego (RWO). Warszawa: PKN; 1999. <http://sklep.pkn.pl/pn-en-1484-1999p.html>.
- [20] PN-EN-27888-1999, Jakość wody - oznaczanie przewodności elektrycznej właściwej. Warszawa: PKN; 1999. <http://sklep.pkn.pl/pn-en-27888-1999p.html>.
- [21] PN-EN ISO 10301:2002, Jakość wody - Oznaczenie łatwo lotnych chlorowcowych pochodnych węglowodorów - Metody z zastosowaniem chromatografii gazowej. Warszawa: PKN; 2002. <http://sklep.pkn.pl/pn-en-iso-10301-2002p.html>.

- [22] PN-ISO 5667-5:2003, Jakość wody - Pobieranie próbek - Część 5: Wytyczne dotyczące pobierania próbek wody do picia i wody używanej do produkcji żywności i napojów. Warszawa: PKN; 2003. <http://sklep.pkn.pl/pn-iso-5667-5-2003p.html>.
- [23] PN-ISO 5667-6:2003, Jakość wody - Pobieranie próbek - Część 6: Wytyczne dotyczące pobierania próbek z rzek i strumieni. Warszawa: PKN; 2003. <http://sklep.pkn.pl/pn-iso-5667-6-2003p.html>.

THE MODELING OF TRIHALOMETHANES CONCENTRATION IN WATERPIPE NETWORK BASED ON DIFFERENT EQUATIONS OF REGRESSION

¹ Faculty of Energy and Environmental Engineering, Silesian University of Technology, Gliwice

² Faculty of Applied Mathematics, Silesian University of Technology, Gliwice

Abstract: Trihalomethanes, disinfection by-products resulting from the reaction between chlorine and organic matter contained in water, appear during delivery of water to consumers. Both the complexity of the reaction kinetics and a large number of factors determining the reaction cause difficulties in the estimation of concentration of THMs in monitoring points in the water supply system. In 2015 it was carried out a revision of the EU directive on the quality of water intended for human consumption. According to the current legal regulation, the scope and frequency of monitoring should take into account the risk of consumer health throughout the whole water supply chain from the catchment area through abstraction, treatment and storage to distribution. Thus there increases the role of prediction models of trihalomethanes concentration. Such models should contain water quality parameters as well as exploitation factors and they may be important decision tool in the process of water system operation. They allow to manage the human health risk, simultaneously they limit the number of expensive laboratory analyses. In this paper we present different prediction models for the concentration of trihalomethanes in waterpipe network. The multiple regression was applied to various data sets and transformations of THMs and independent variables. We establish how different transformations of variables influence on prediction accuracy and statistical properties of the model. Presented research was based on real data collected during seven years of exploitation of one of big Polish water supply systems.

Keywords: trihalomethanes, water supply system, mathematical models, multiple regression