2nd International Joint Conference on Water Distribution Systems Analysis & Computing and Control in the Water Industry
Valencia (Spain), 18-22 July 2022
doi: https://doi.org/10.4995/WDSA-CCWI2022.2022.14734

DISINFECTION BY PRODUCTS ESTIMATION IN A WATER DISTRIBUTION NETWORK

Laura Vinardell[1], Irene Jubany[1] and Ramon Pérez[2]

[1]Sustainability Area, Eurecat, Centre Tecnològic de Catalunya, Plaça de la Ciència, 2, 08243 Manresa, Spain;
[2]acs2ac, Universitat Politècnica de Catalunya, Rambla Sant Nebridi 10, 0822 Terrassa, Spain.

[1] laura.vinardell@eurecat.org ; irene.jubany@eurecat.org, [2] ramon.perez@upc.edu

Abstract

Even though disinfection is necessary to ensure water safety for human consumption, some disinfectants produce disinfection by-products (DBPs) that may be dangerous for human health. Current European legislation obligates water distributors to limit some DBPs concentration to final consumers. Then, water companies must control these compounds and are obligated to periodically monitor their network. DBPs modeling can be very useful for estimating online DBPs concentration throughout the network, increasing DBPs control and knowledge, but avoiding DBPs analytics time and resources consumption [1].

Trihalomethanes (THM), the first DBPs discovered, have long been the most studied and modeled. Previous studies have mostly used linear relations between variables and THM concentration, but also computational modelling, mechanistic and data driven models [2, 3]. Even though, there are still challenges to beat: most studies use a small database and laboratory-scale for model building, forgetting the impact of network pipelines and season. In addition, significant variables for DBPs' formation such as retention time are most of the time neglected due to its difficulty to measure. Finally, THMs are not the only DBPs generated from disinfection or even the most toxic: other DBPs must be studied, and their formation pathways along the network investigated.

In this study, data from a full-scale distribution network was used: online sensors and sampling campaigns. To include hydraulic conditions as retention time, EPANET software and R programming are used to simulate the network. Different models, mechanistic and data driven, have been used to estimate the chlorine decay and DBP formation within the network. Results of the calibration and validation of these models and the conclusions obtained are presented.

Keywords: disinfection by-products, modelling, distribution water networks, hydraulic simulation, EPANET, HAA, THM

1. INTRODUCTION

Disinfection by-products are an increasing concern for protecting human health. These compounds are formed due to a disinfection treatment: the disinfectant reacts with compounds present in water as natural organic matter and anthropogenic contaminants, forming halogenated molecules. This field was discovered in 1974 when chloroform and other trihalomethanes were found in drinking water supplies in the USA. Since then, more than 700 DBPs have been identified from all disinfection processes (chlorine dioxide, ozonation, chloramination, and chlorination) [4, 5]. Some of these compounds are carcinogenic and/or toxic at certain concentrations [6].

Since DBPs' discovery, some countries have included in their drinking water legislation a threshold for these compounds. Current European regulation about water intended for human consumption is (EU) 2020/2184 [7]. It considers microbiological, chemical, and

physical parameters. Regarding DBPs, it includes TTHM (Total THM: chloroform, bromodichloromethane, dibromochloromethane, and bromoform), five HAA (HAA_5: monochloroacetic acid, dichloroacetic acid, trichloroacetic acid, monobromoacetic acid, and dibromoacetic acid), chlorate, chlorite, and bromate (Table 1). This is relatively new legislation since it was implemented in December 2020. The previous one [8] is from 2003 and only included TTHM and bromate.

Table 1: European current DBPs legal thresholds for water for human consumption

	TTHM	**HAA_5**	**Bromate**	**Chlorite**	**Chlorate**
UE Legal threshold [µg/L]	100	60	10	250	250

To accomplish the European legislation, water distribution companies monitor water quality at different network locations. Water network operators are not forced to monitor their water quality online, and only need to analyze a few samples distributed along the network. For DBPs analysis it is required specific and expensive instrumentation: they are generally quantified with gas or liquid chromatography at laboratories, a time-consuming and expensive analytical technique.

Due to the high cost and maintenance of online sensors, only a few distribution networks have them installed. The alternative for DBPs online control is the use of mathematical models based on easily measurable data. Unfortunately, water quality modelling has some disadvantages: it is water source and treatment applied depending (highly specific), needs a large amount of data, and can require using complex mathematical tools. For all these reasons, DBPs online monitoring is a research area in development and needs to be deeply investigated.

This study aims to use a model from literature able to predict DBPs in a full-scale water distribution network, using water quality online sensors data, and hydraulic simulations from the pilot site. DBPs formation does not only depend on its precursors (organic matter, some inorganic compounds and type and dose of disinfectant), but also on the network's hydraulics (residence time, velocity) and environmental characteristics (temperature, pH, etc.). For that purpose, sampling and data from a full-scale network were used, and EPANET software tool for hydraulic simulations.

2. MATERIALS AND METHODS

The study site is in Tarragona (Catalonia, Spain). *Consorci d'Aigües de Tarragona* (CAT) is the entity managing water distribution in this province to the municipalities. Water comes from Ebro's River and is dispensed to the whole province after ozonation treatment at drinking water treatment plant (DWTP). Chlorination is applied in distribution network in different boosting stations.

Tarragona's network includes 1 DWTP, 23 pump stations (PS), 141 tanks, and online sensors in several locations (free chlorine, flowmeters, and other water quality parameters).

2.1. Sampling campaign definition

To decide the water network section for the study, historic data of the whole drinking water distribution network was analyzed. Data comprised TTHM, UV254, conductivity and total organic carbon (TOC) measurements from 2017 to 2019. In addition, distance and number of rechlorinations was also included in the study. The statistical analysis comprised

correlations between the different variables (TTHMs and water quality) and principal components analysis (PCA) to identify the more significant variables that explain TTHM values dispersion.

This analysis was necessary to decide CAT network's section where the sampling campaign should be performed, in other to include full-scale data to the research. Data available from the sampling campaign is summarized in Table 3.

Table 2: study parameters

Field measurements	Laboratory measurements
pH	THM_4 [µg/L]
Conductivity [µS/cm]	HAA_5 [µg/L]
Temperature [ºC]	Chlorate [µg/L]
Free Chlorine [mg/L]	Turbidity [NTU]
	TOC [mg/L]

Parameters measured in the field were measured using field sensors. THM were measured by HS / GC - µECD, HAA and chlorate with LC-MS/MS, turbidity with a turbidimeter and TOC with high temperature oxidation combustion equipment.

2.2. Water network hydraulic simulation

EPANET Software

EPANET is a free software for hydraulic network simulations, designed by the United States Environmental Protection Agency [9]. This software makes it possible to simulate hydraulic and quality behavior in a water distribution network: residence time, free chlorine concentration, storage tank pressure, and flow, among other parameters.

Figure 2 shows the complete hydraulic model of the study network. For this study, work was carried out on the A to C section corresponding to the water treatment plant effluent (A) and a pumping station (C) located 60 km away. Then the network was simplified preserving only the main pipeline. Pump curves, valve status, and pipeline diameters were revised and/or modified to balance flowmeter sensor measurements.

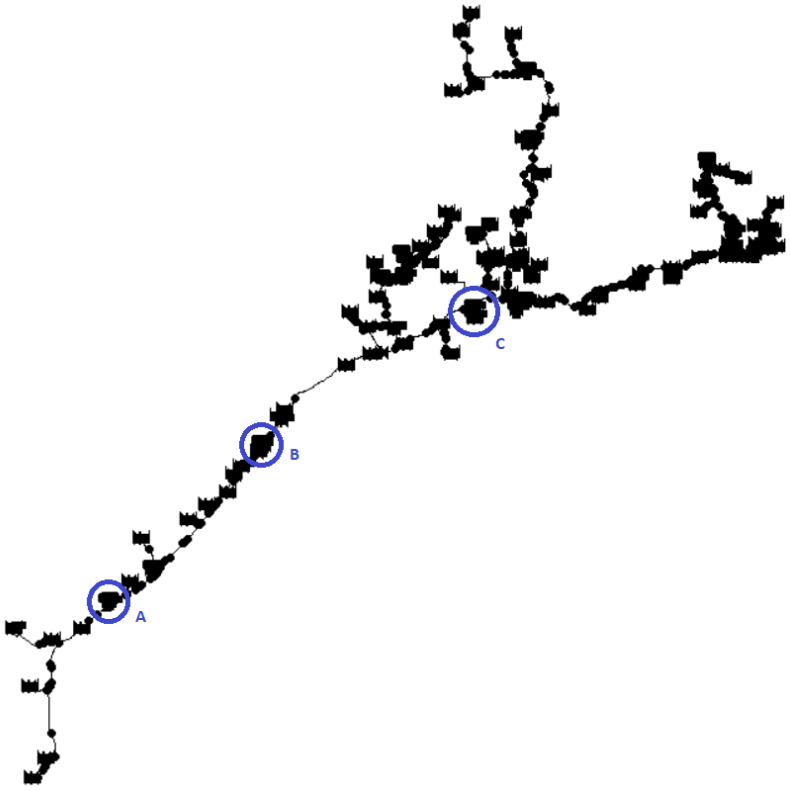

Figure 1: Case-study hydraulic model

The reason for this section's choice for hydraulic modeling was its simplicity for free chlorine modeling: there is only the chlorination at A (the following disinfection process is at C). Modeling this simple model is the first step in this research context.

For EPANET hydraulic simulation, flowmeter and free chlorine sensors data from October 2021 were used. The section chosen had different regimes depending on water demand: water can pass through intermediate tank storage (B in Figure 1) or can advance directly from A to C. Due to this complexity, clustering analysis of flowmeters data was done to discretize the different regimes at tank storage B, using PAM clustering function in RStudio language.

PAM clustering [10] is an algorithm that organizes data around medoids (cluster centers), to minimize the dissimilarities between data in the same cluster and maximize between different clusters. *NBClust* R package was used for the optimal number of clusters decision.

Hydraulic and quality simulations were done for 72h, in the regime where water by-pass B tank:

- First 24h for hydraulic model stabilization
- 24h for model identification
- Last 24h for model validation

Chlorine decay model

Free chlorine behavior in water distribution networks is commonly described as first-order kinetics [11]:

$$Cl = Cl_0 e^{-kt} \quad (1)$$

Where Cl is the final chlorine concentration, Cl_0 is the initial chlorine concentration, k is the decay constant and t the residence time. The same eq. (1) can be written as:

$$\ln\left(\frac{Cl_0}{Cl}\right) = kt \quad (2)$$

Using eq. (2), k was estimated by linear regression method using residence time (t) and free chlorine concentrations at A (initial concentration Cl_0) and C (final concentration Cl). k constant was estimated using identification data and then quality simulation was performed for 48h (identification and validation data)

The mean squared error (MSE) was calculated to compare predicted values and measured values (from free chlorine sensors). MSE must not worsen significantly with the validation data compared to the identification data to rule out overfitting.

TTHMs prediction model

Clark et al. [12] used a mathematical model for TTHM formation prediction using kinetics relationships:

$$TTHM = T\left\{C_{A0} - \left[\frac{C_{A0}(1-R)}{1-Re^{-ut}}\right]\right\} + M \quad (3)$$

Where M is the initial TTHMs, t is the residence time, and T, R and u are parameters depending on kinetic relations between free chlorine, chlorine demand matter and DBPs generated. This model does not include rechlorinations. Therefore, TTHM prediction could not be extended from C booster station.

In the same study, Clark et al. used different water samples with different T, R, u and M values. The samples were defined by location and quality parameters (pH, Temperature, TOC and residual free chlorine). For this research, the values T, R and u from the sample with quality parameters closer to the study case (data from A) were used (Table 2).

Table 3: water sample quality characteristics and model parameters of [12] reference study

	Quality conditions				Model parameters			
	pH	**T [ºC]**	**TOC [mg/L]**	**Free Cl [ppm]**	**T**	**R**	**u**	**M**
Reference sample from [12]	8.15	17.9	1.87	1.73	41.29	0.25	0.44	10.00
A	7.70	24.5	3.00	1.22	-	-	-	-

3. RESULTS

3.1. Sampling campaign definition

The statistical analysis showed that the highest correlations were found between TTHMs and the number of rechlorinations ("nº rechl.") (high correlations) and with conductivity (moderated correlation). Other moderated correlations were Temperature – TOC and year – TOC (Table 5).

Table 4: variables correlation

	Year	Temperature [ºC]	TTHM [µg/L]	Conductivity [µS/Cm]	TOC [mg/L]	UV_{254}	nº rechl.
Temperature [ºC]	-0.05021						
TTHM [µg/L]	-0.16646	0.17517					
Conductivity [µS/Cm]	-0.23006	-0.25826	0.30543				
TOC [mg/L]	-0.41349	-0.36898	-0.10633	0.17625			
UV_{254}	0.16498	-0.01802	-0.05994	-0.20935	0.07519		
nº rechl.	-0.22228	0.08059	0.56943	0.03430	-0.07588	-0.11523	
Distance [km]	0.00158	0.04319	0.02287	-0.02804	0.01283	-0.06475	0.09644

PCA study also showed a significative relation between TTHMs and the number of rechlorinations (Figure 2).

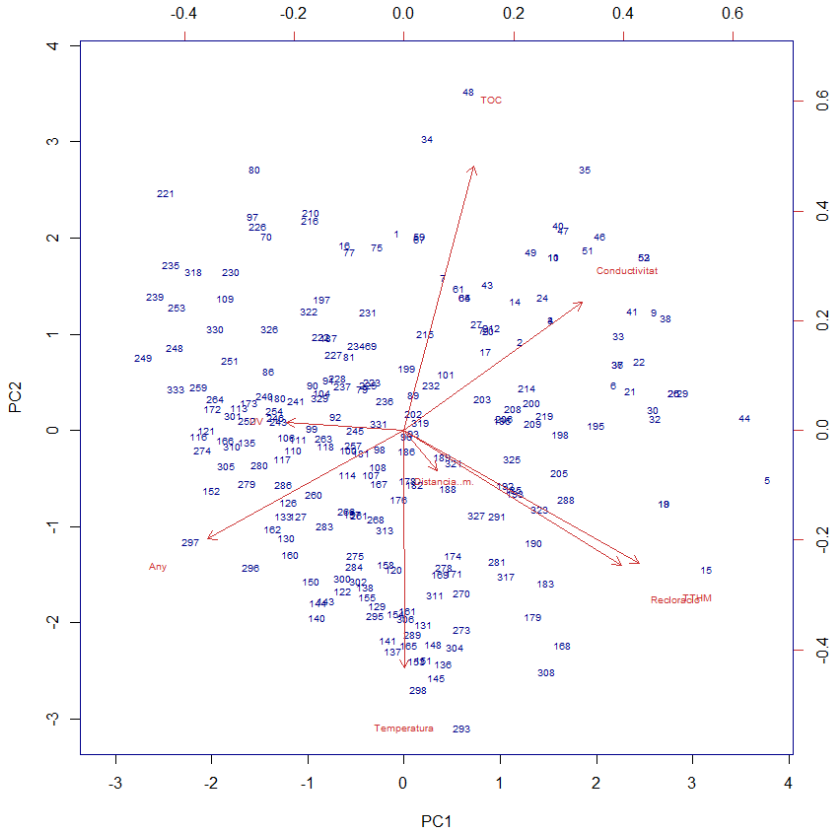

Figure 2: PCA results

These findings showed that rechlorinations may have a significant role in THMs formation in this network. For this reason, the main objective of the research is to study its effect on DBPs formation. Thus, the selected location for further studies (illustrated in red in Figure 3) is composed of the principal pipeline and two final branches (with three and four boosters stations respectively). To include seasoning in the study for predicting DBPs' formation, the field study will last a whole year (October 2021 – October 2022), and a sampling campaign will be done biweekly. Sampling campaigns started in September 2021.

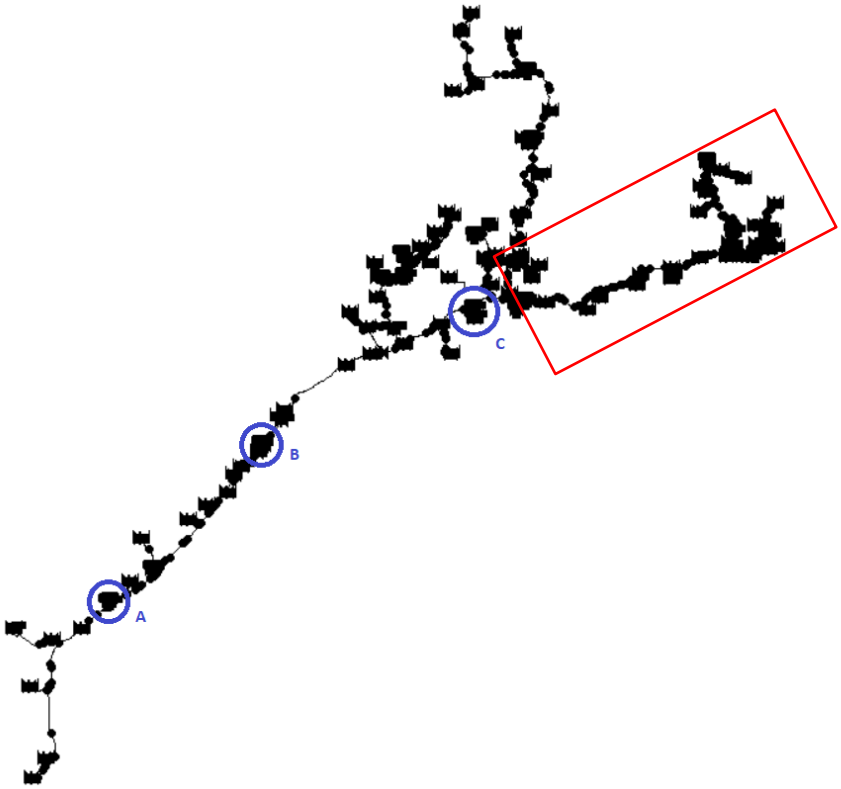

Figure 3: study section

The selected study section is posterior to the hydraulic modeled sector (A to C in Figure 3). Further work will include extending the hydraulic simulation to the studied section using a kinetic predictive model for DBPs that includes rechlorinations.

Currently, TTHMs predicted at C storage tank intake (before rechlorination) with Clark et al. model [12] were compared to the full-scale TTHM measured from the sampling campaign at the sampling location closer to C.

3.2. Chlorine decay model

The free chlorine decay constant obtained by linear regression was k = 1.001 [h^{-1}] (Figure 4). Then, free chlorine decay model for this site is:

$$Cl = Cl_0 e^{-1.001t} \quad (4)$$

Free chlorine decay constant *k* was included as *Global kbulk* and *Global kwall* in EPANET. Then, quality simulation was performed for the same 72h as the hydraulic model. Results are presented at Figure 5, where the difference between free chlorine simulation (using k decay obtained) and the measured from sensors can be seen.

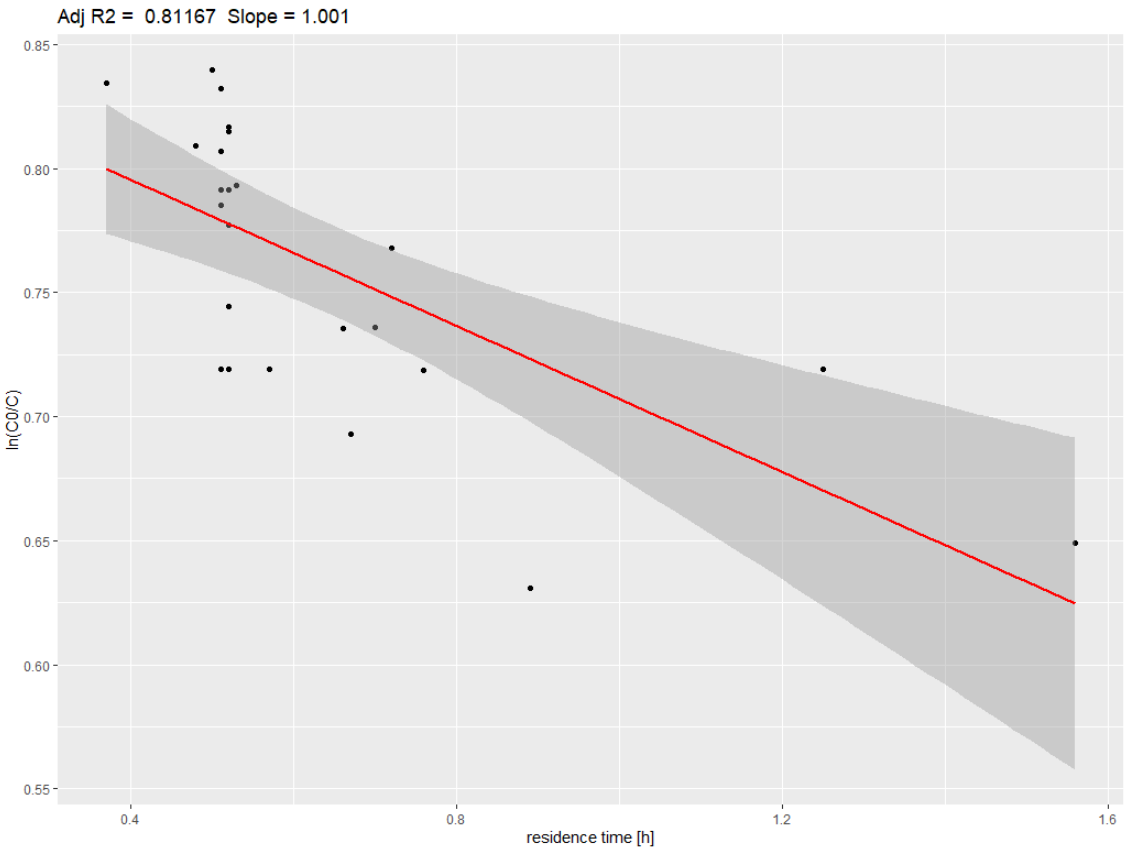

Figure 4: linear regression model for identification data

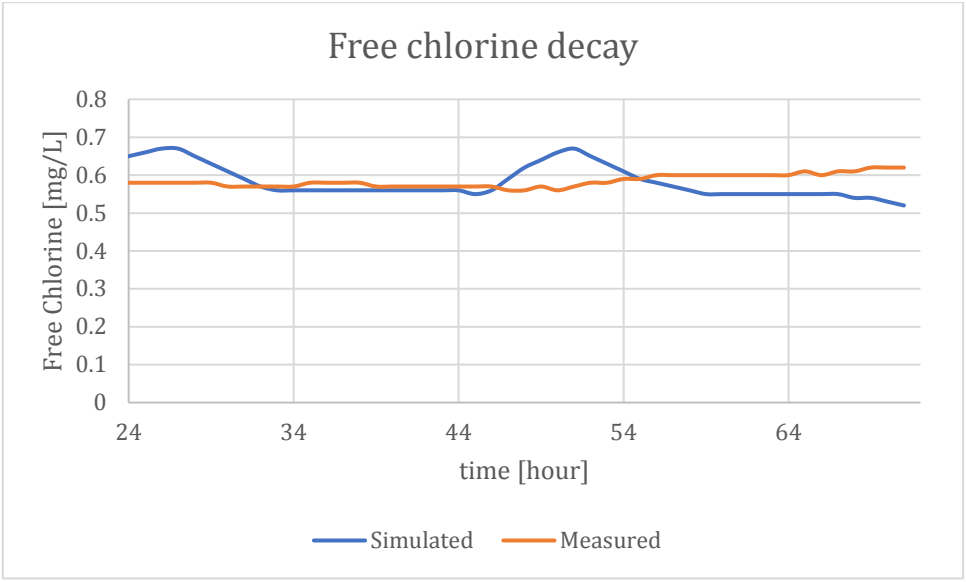

Figure 5: Comparison between measured and simulated free chlorine decay

MSE results (Table 4) indicate the error made by the model for identification and validation data. Though both values are very small, the result for validation duplicates the one for identification. The overfitting of the model should be treated including more data when the different hydraulic regimes are reliably simulated.

Table 5: MSE results

Identification data MSE	Validation data MSE
0.001696	0.003888

3.3. TTHMS predictive model

Residence time obtained from the hydraulic simulation combined with available data of free chlorine available were used together with a TTHM prediction model from the literature [12]. The intention was to study if TTHM prediction model estimates a coherent value at C tank intake with TTHM analytically measured in the sampled location closer to C tank.

TTHM estimated by the Clark et al. model was calculated using eq. (3). *T, R* and *u* parameters used were obtained from the sample reference (Table 2); residence time *t* is from EPANET hydraulic model calculus, and free chlorine initial concentration C_{A0} and initial TTHMs concentration *M* from A (drinking water treatment plant output) data.

Table 6: TTHMs predicted and measured

TTHMs predicted using Clark et al. model at C intake	TTHMs analytically measured
11.35 µg/L	58.6 µg/L

Table 6 shows the TTHMs calculated by Clark et al. model [12] and the full-scale measurement. The difference between both values could be explained because there are two booster stations between both locations. Residence time at C tank output is close to 5h, seven times more than in C tank intake. Contact time increment at storage tanks between free chlorine and organic matter may explain TTHMs increase.

4. CONCLUSIONS

This study shows that kinetic models and software for hydraulic simulations are useful tools for water quality predictions. Drinking water network simulations allow to include complex hydraulic calculations (water demand, pressure, residence time, etc.), improving distribution network management.

In this study, residence time was identified using full-scale data and EPANET software simulation. Free chlorine constant decay was calculated and validated for a first-order decay model. Finally, using a kinetic model from the bibliography, TTHMs were estimated, obtaining a realistic value.

Even though, this study is the first step to obtain a more accurate and specific model for TTHM prediction. In further research, the model will use only site parameters calibrated for the site and must include rechlorination data. The hydraulic simulation must be extended to the study section, including all the different regimes and chlorination conditions. Finally, other DBPs must be also considered: HAA and oxyhalides because they are also dangerous for human health and need deeper research.

5. ACKNOWLEDGEMENTS

The authors want to acknowledge *Consorci d'Aigües de Tarragona* for providing hydraulic and water quality data and providing access to the drinking water network for sampling. Laura Vinardell is a fellow of Eurecat's "Vicente López" PhD grant program.

6. REFERENCES

1. Ike IA, Karanfil T, Ray SK, Hur J (2020) A comprehensive review of mathematical models developed for the estimation of organic disinfection byproducts. Chemosphere 246:. https://doi.org/10.1016/j.chemosphere.2019.125797
2. Clark RM, Sivaganesan M (1998) Predicting Chlorine Residuals and Formation of TTHMs in Drinking Water. J Environ Eng 124:1203–1210. https://doi.org/10.1061/(asce)0733-9372(1998)124:12(1203)
3. Morrow CM, Minear RA (1987) Use of regression models to link raw water characteristics to trihalomethane concentrations in drinking water. Water Res 21:41–48. https://doi.org/10.1016/0043-1354(87)90097-2
4. Brown D, Bridgeman J, West JR (2011) Predicting chlorine decay and THM formation in water supply systems. Rev. Environ. Sci. Biotechnol. 10:79–99
5. John C. Crittenden, R.Rhodes Trussell, David W. Hand,Kerry J.Howe GT (2012) Water treatment principles and design
6. Richardson SD, Plewa MJ, Wagner ED, et al (2007) Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research. Mutat Res - Rev Mutat Res 636:178–242. https://doi.org/10.1016/j.mrrev.2007.09.001
7. European Commission (2020) Directive (EU) 2020/2184 of the European Parlliament and of the Council of 16 December 2020 on the quality of water intended for human consumption. Off J Eur Union 2019:1–61
8. Presidencia M De (2003) Real Decreto 140 / 2003 , de 7 de febrero , por el que se establecen los criterios sanitarios de la calidad del agua de consumo humano . 1–40
9. Rossman L a. (2000) Epanet 2. 104
10. Kaufman L, Rousseeuw PJ (1990) Partitioning Around Medoids (PAM). In: Finding Groups in Data: An Introduction to Cluster Analysis
11. Johnson J. (1978) Measurement and persistence of chlorine residuals in natural waters. In: Environmental Impact of Water Chlorination. pp 43–71
12. Clark RM (1998) Chlorine Demand and TTHM Formation Kinetics: A Second-Order Model. J Environ Eng 124:16–24. https://doi.org/10.1061/(asce)0733-9372(1998)124:1(16)