Effect of data uncertainty on trihalomethanes prediction in water supply systems using kinetic models

C. Di Cristo*, A. Leopardi*, G. de Marinis*

*Università di Cassino e del Lazio Meridionale, Via Di Biasio 43, Cassino 03043, Italy

Abstract

The present work compares the performances of first and second order kinetic models for predicting trihalomethanes (THMs) formation in a real case-study. The kinetic parameters are evaluated through an automatic calibration procedure, in which a least-squared objective function relating measured and computed residual chlorine and THMs concentrations is adopted. The effect of measurements uncertainty on the calibrated parameters and on THMs concentrations predictions are quantified in terms of confidence limits using the First Order Second Moment approach. The study reveals that the performances of the two models are quite similar, but the second order one results less influenced by uncertainty.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of the CCWI2013 Committee

Keywords: Water quality; Chlorine; Trihalomethanes; Uncertainty; Calibration;

1. Introduction

The primary rule of water supply systems is to provide the water resource to the users with prefixed quality requirements. For this reasons it is crucial to elaborate strategies and methodologies for improving the drinking water security by deliberate (Di Cristo et al. 2011) or accidental (Di Cristo and Leopardi, 2008) contaminations and by phenomena which can determinate a quality deterioration through the distribution systems. In this framework, for making drinking water safe from pathogenic microorganisms, chlorine is commonly used as disinfectant. Chlorine, typically applied as solution of either sodium or calcium hypochlorite, is a strong oxidizing

* Corresponding author. Tel. +39 (0)776 2994340, Fax +39 (0)776 2993939
E-mail address: dicristo@unicas.it
agent and it decays over time. A minimum level of chlorine residual must be maintained in the distribution system to preserve the quality of treated water respect to microbial re-grow. While the use of chlorine is beneficial for this purpose, unwanted disinfection by-products (DBPs) are usually formed from its reaction with the Natural Organic Matter (NOM) and other chemical species, such as bromide, naturally present in water. Epidemiological and toxicological studies have suggested that some of these DBPs, such as the trihalomethanes (THMs) have carcinogenic effects to human health (Weinberg et al. 2002). For this reasons there are regulatory limits on total THMs concentration in tap water: the US Environmental Protection Agency in the Stage 2 DBP regulations has fixed the maximum level to 80 μg/l; the Canadian and the Australian – New Zealand guidelines have maxima of 100 μg/l and 250 μg/l, respectively, while in Italy a very restrictive regulation has fixed the limit value to 30 μg/l. Regulations require that those standards have to be met at remote distribution system locations, where the water residence time and THMs concentrations are higher.

For identifying management strategies in order to respect the quality standards, it is essential to have models able to reproduce THMs formation in water supply systems. Many proposed predictive models are based on empirical relationships among the total THMs concentration and some parameters, which influence their formation, developed using laboratory or field data (Chowdhury et al. 2009). Di Cristo et al. (2013), comparing different empirical models, investigated on their applicability to a water supply systems and on their ability in reproducing THMs concentrations in real situations, different from those considered for their elaboration, individuating the more suitable ones. The alternative is to use models based on kinetics involved during chlorine reactions (Clark and Sivaganesan 1998; Gang and at. 2002), but in this case the kinetic parameters have to be quantified. Only for few models (Clark and Sivaganesan 1998, Clark and Sivaganesan 2001) there are empirical relationships for estimating the THMs formation kinetic rate, relating its value with parameters influencing the transformation. However their applicability to real water distribution systems is still an open question, which requires further investigations. Another possibility is to estimate the kinetic parameters through a mathematical calibration procedure using measurements, which is the option considered in the analysis performed in this paper. While the calibration of chlorine decay parameters was widely investigated (Powell et al. 2000, Al-Omari and Chaudhry 2001), not much has been done for the THMs production rate (Courtis et al. 2009). However errors in measuring concentrations result in uncertainty in the estimated parameters. Such uncertainty, along with its effect on model predictions, has to be quantified for a reliable use of a water quality simulation.

In the present paper, the performances of the two more used kinetic models for predicting THMs formation are compared. In both cases, the chlorine decay and the THMs formation kinetic parameters are calibrated through an automatic calibration procedure, in which a least-squared objective function relating measured and computed residual chlorine and THMs concentrations is adopted. The used field measurements were realized on the main trunk of the Aurunci-Valcanneto Water Supply System (AVWSS), adopted as case-study, where water disinfection is performed using sodium hypochlorite. Moreover, the effect of measurements uncertainty is investigated evaluating its assessment on the calibrated parameters and on THMs concentration predictions though a First Order Second Moment (FOSM) analysis.

Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>time (h)</td>
</tr>
<tr>
<td>([Cl]_0)</td>
<td>initial chlorine concentration (mg/l)</td>
</tr>
<tr>
<td>([Cl])</td>
<td>chlorine concentration at time ( t ) (mg/l)</td>
</tr>
<tr>
<td>([THM]_0)</td>
<td>initial total THMs concentration (μg/l),</td>
</tr>
<tr>
<td>([THM])</td>
<td>total THMs concentration at time ( t ) (μg/l)</td>
</tr>
<tr>
<td>(D)</td>
<td>THMs formation kinetic coefficient (μg of THMs formed /mg chlorine consumed)</td>
</tr>
<tr>
<td>(K_1)</td>
<td>first order chlorine decay kinetic coefficient (1/h)</td>
</tr>
<tr>
<td>(K_c)</td>
<td>dimensionless constant</td>
</tr>
<tr>
<td>(\mu)</td>
<td>rate constant (1/h)</td>
</tr>
<tr>
<td>(f)</td>
<td>dimensionless objective function (-)</td>
</tr>
<tr>
<td>(f_C)</td>
<td>objective function respect to chlorine measurements (mg²/l²)</td>
</tr>
</tbody>
</table>
f_T \quad \text{objective function respect THMs measurements (μg²/l²)}

2. THMs formation prediction

2.1. THMs kinetic models

In the framework of kinetic models, previous researches have shown that THMs formation can be modeled as function of chlorine consumption (Clark and Sivaganesan 1998, Bocelli et al. 2003, Brown et al. 2011). The relationship between THMs formation and chlorine consumption is assumed linear:

\[
[THM] = [THM]_0 + D([CI]_0 - [CI])
\] (1)

The kinetic parameter \( D \) is an indicator of the THMs productivity in water. In order to predict the THMs concentrations through Eq.(1), the following first and second order (Clark and Sivaganesan, 1998) models are considered for estimating the actual chlorine concentration:

\[
[CI] = [CI]_0 e^{-K_1 t}
\] (2)

\[
[CI] = [CI]_0 \left( \frac{1 - K_c}{1 - K_c e^{-u_1 t}} \right)
\] (3)

The parameters which have to be estimated are \( K_1, D \) for the first order model (Eqs.1 and 2) and \( K_C, u, D \) for the second order one (Eqs.1 and 3). In the following, they are estimated through the calibration procedure described in the next paragraph.

2.2. Parameters calibration procedure

The values of the kinetic parameters are obtained minimizing the following objective function:

\[
f = \sum_{i=1}^{NC} \left( \frac{[CI]^c_i - [CI]^m_i}{[CI]^m_i} \right)^2 + \sum_{i=1}^{NT} \left( \frac{[THM]^c_i - [THM]^m_i}{[THM]^m_i} \right)^2
\] (4)

where \( NC \) and \( NT \) are the numbers of the measurements of \([CI]\) and \([THM]\), respectively; the apexes \( c \) and \( m \) indicate the calculated and measured concentrations in the measurement point \( i \), respectively. It is worth of note that in applying a THMs formation model to a water system, the kinetic parameters, depending on water quality, operational and environmental conditions, should theoretically be considered different in different locations. However, in the presented case-study, unique global values are assumed for the whole system, since the conditions are not very variable.

For solving the optimization problem (4) a gradient based approaches is used. In particular, the EPANET-MSX software (Shang et al. 2008), adopted for the multi-species quality simulation, is coupled with a gradient based optimization algorithm. The use of a gradient based optimization method is well suited for the first order model since the objective function is a quadratic form. For the second order model the objective function is more complex, but it has been tested that the gradient method is rapidly converging to the optimal solution.

2.3. Uncertainty analysis

Given the errors in data measurements, the calibrated values represent the best estimate in terms of objective
function, but they include also uncertainty. Assuming the first-order approximation (Pasha and Lansey 2012), the variances of the calibrated parameters are computed based on the variances of the measured concentrations. In particular, as shown in the following relations, the variances of the parameters related to chlorine decay, $P_c$, are estimated considering errors in chlorine measurements, $\sigma_c^2$, while the variance of $D$ is evaluated based on errors in THMs measures, $\sigma_T^2$:

$$Var(P_c) = \left[ \left( \frac{\partial[C]}{\partial P_c} \right)_{NC \times 1} \right]^{-1} \sigma_c^2$$  \hspace{1cm} (5)

$$Var(D) = \left[ \left( \frac{\partial[THM]}{\partial D} \right)_{NT \times 1} \right]^{-1} \sigma_T^2$$  \hspace{1cm} (6)

where

$$\sigma_c^2 = \frac{f_c}{NC - 1} \hspace{1cm} \sigma_T^2 = \frac{f_T}{NT - 1}$$  \hspace{1cm} (7)

being $f_c$ and $f_T$ the sums of the squared differences between the measured and simulated values after the calibration of chlorine and THMs concentrations, respectively. The parameters related to chlorine decay $P_c$ are $K_1 (c=1)$ for the first order model and $K_c$ and $u (c=2)$ for the second order one. Then, the covariance matrixes of THMs concentrations are estimated from the variances of the parameters using the First Order Second Moment (FOSM) approach, which has been efficiently applied to quantify uncertainty in calibration of the hydraulic parameters (Kang and Lansey 2011) and of the wall reaction coefficient (Pasha and Lansey 2012). One covariance matrix of THMs concentrations is computed for each parameter as:

$$Cov([THM])_p = \left[ \left( \frac{\partial[THM]}{\partial P} \right)_{N \times 1} \right] Var(P) \left( \frac{\partial[THM]}{\partial P} \right)_{N \times 1}$$  \hspace{1cm} (8)

where $N$ is number of the measurement points and $P$ is the considered parameter. The covariance matrixes are $c+1$, since the parameters are $P_c$ and $D$. The diagonal elements of the covariance matrixes are the variances of THMs concentrations respect to the considered parameter. The total variance of the THMs concentration prediction in each measurement point is computed as the sum of variances estimated separately respect each parameter. Since, the FOSM has been successfully applied based on a Gaussian assumption (Pasha and Lansey 2012), the confidence limits (CLs) are computed as the predicted value minus (plus) 1.96 time the corresponding standard deviation, which corresponds to the 2.5% (97.5%) percentile of the Gaussian distribution.

3. Case study

The considered case-study is the main trunk of the AVWSS, located in the Lazio region in Southern Italy. It serves about 110’000 inhabitants, divided in 25 towns. The AVWSS scheme, reported in Fig. 1, is composed by 1 reservoir (Madonna di Canneto Spring), 70 junction nodes, 51 pipes, 28 regulation valves and 2 tanks (ValCanneto tanks), which furnishes the water to the internal distribution networks. The water comes into the system from the primary source of the Madonna di Canneto Spring (Node 1), connected to the annex reservoir, and from two secondary inputs in the junctions Forestelle (Node 13, $Q_{13}=22$ l/s) and Cippone-Collelungo (Node 12, $Q_{12}=24$ l/s), as shown in Fig. 1. The 9 output junctions, along with the relative demands, are also reported in Fig. 1. Water disinfection is made using sodium hypochlorite ($NaOCl$) before the water entrance into the system.
A campaign of water quality measurements was realized during a period in which the hydraulic behaviour of the system can be considered steady. All realized measurements are detailed described in Di Cristo et al. (2013), while Table 1 reports only the data used for the calibration, represented by the residual $\text{HOCl}$, indicated as chlorine $[\text{Cl}]$, and the total THMs concentrations, measured in the 9 nodes indicated in Fig. 1. For each measurement node is also reported the water age, evaluating a 48 hours hydraulic simulation using EPANET with a time step of 300 s. The hydraulic parameters were previously calibrated using tank levels and flow demands measurements, with the Darcy-Weisbach as resistance formula (Russo Spena and Vacca, 1998).

![Fig.1 Scheme of the AVWSS.](image)

### 4. Results

As shown in Fig. 1, since the nodes 1, 2 and 12 are just downstream the water inflow points, the measurements in those nodes are used for assigning the initial concentrations of chlorine $[\text{Cl}]_0$ and trihalomethanes $[\text{THM}]_0$. In fact, even if in Eq.(1) it is usually assumed $[\text{THM}]_0=0$, this is not the case of the presented study, since not zero THMs concentrations have been measured in the inlet nodes 1 and 2 (Table 1). The other values are used for the calibration, which is realized with $NC=4$ and $NT=6$. During the calibration process water quality simulations are realized using EPANET-MSX with a time step of 60 s.

<table>
<thead>
<tr>
<th>Node</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[\text{Cl}]$ (mg/l)</td>
<td>0.30</td>
<td>0.27</td>
<td>0.27</td>
<td>0.26</td>
<td>-</td>
<td>-</td>
<td>0.24</td>
<td>0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>$[\text{THM}]$ (µg/l)</td>
<td>0.64</td>
<td>1.97</td>
<td>1.49</td>
<td>1.32</td>
<td>2.00</td>
<td>2.45</td>
<td>3.02</td>
<td>1.77</td>
<td>0.00</td>
</tr>
<tr>
<td>Age (h)</td>
<td>0.05</td>
<td>1.97</td>
<td>0.96</td>
<td>6.49</td>
<td>5.78</td>
<td>8.00</td>
<td>21.95</td>
<td>7.54</td>
<td>0.00</td>
</tr>
</tbody>
</table>

For solving the optimization problem expressed by Eq.(4), joint and two steps sequential approaches are tested. In the joint calibration all kinetic parameters are simultaneously estimated considering both chlorine and THMs measurements. In the two steps approach, first the chlorine decay parameters $P_c$ are calibrated using only the residual chlorine measurements, then, fixed the $P_c$ values obtained in the former step, a second calibration is performed for estimating $D$ using only the THMs measurements. In Table 2 the calibration results of both
procedures are presented for the first order model, revealing that they have the same performance in terms of objective function. For the second order, the joint approach does not represent a good option, because the model results only sensitive to the chlorine decay parameters $K_c$ and $u$ and then the simultaneous calibration it is not able of modifying the initial $D$ values. For this reason, for the second order model the results of the calibration with only the two steps approach are reported in Table 3. Comparing the performances of the first and second order models, obtained with the two steps calibration, the latter has a smaller objective function value, even if the difference is marginal. The calibration furnishes analogous $D$ values for the two models, which are very similar to the value 41 (μg/mg) obtained by Courtis et al. (2009).

<table>
<thead>
<tr>
<th>Calibration</th>
<th>Optimal values</th>
<th>$K_1$ (1/h)</th>
<th>$D$ (μg/mg)</th>
<th>$f$ (–)</th>
<th>$f_c$ (mg/l²)</th>
<th>$f_t$ (μg²/l²)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint</td>
<td></td>
<td>0.014</td>
<td>35</td>
<td>0.247039</td>
<td>0.001250</td>
<td>0.9734</td>
<td></td>
</tr>
<tr>
<td>Two-steps</td>
<td></td>
<td>0.012</td>
<td>40</td>
<td>0.248706</td>
<td>0.001193</td>
<td>1.0001</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Table 2. Calibration results of the first order model.

<table>
<thead>
<tr>
<th>Calibration</th>
<th>Optimal values</th>
<th>$K_C$ (–)</th>
<th>$u$ (1/h)</th>
<th>$D$ (–)</th>
<th>$f$ (–)</th>
<th>$f_c$ (mg/l²)</th>
<th>$f_t$ (μg²/l²)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td></td>
<td>1.0195</td>
<td>-0.0003</td>
<td>38</td>
<td>0.22629</td>
<td>0.00101</td>
<td>0.8353</td>
<td>6.8 $10^{-5}$</td>
</tr>
</tbody>
</table>

Table 3. Calibration results of the second order model with the two-steps approach.

The uncertainty analysis is performed for evaluating the standard deviations of the calibrated parameters and the confidence limits of THMs concentrations prediction for both the considered models. The standard deviations of all kinetic parameters, computed through Eqs.(5)-(6), are reported in Table 2 and 3 for the first and the second order models, respectively. The standard deviations of $K_1$, $K_C$ and $D$ are of the same order of magnitude, while it is much smaller for the $u$ parameter of the second order model. The difference between the $D$ values obtained for the different models and the different calibration procedures is within the uncertainty range.

![Fig. 2 – First Order Model. CLs for THMs concentration in different nodes. Boxes represent CLs considering $K_1$, Lines represent CLs for $K_C$ and $D$. Crosses represent the measured values.](image)
The CLs of the THMs concentrations are evaluated for the nodes in which THMs measurements are available ($N = NT = 6$). For the first order model Fig. 2 reports the CLs of the THMs concentration considering the $K_1$ variance only and both the $K_1$ and $D$ effects. The CLs size increases with the node water age (Table 1), with a linear law, since both the transport-decay model and the FOSM analysis are linear. The largest CLs value of about 5 µg/l is reached in node 9, when both parameters are considered.

For the second order model, the CLs of the THMs concentrations are reported considering the variances of the chlorine decay parameters only, along with the ones obtained taking into account the effect of the all three coefficients $K_C$, $u$ and $D$. Also in this case the CLs amplitude increases with the node water age, but not with a linear law, with widest interval at node 9 of about 3 µg/l. Moreover, the smaller CLs values of the second order model respect to the ones of the first order, indicate its major robustness respect to measurements uncertainty. For both models the measured THMs values also reported in Fig. 2 and Fig. 3 are contained in the CLs, except for node 3. This fact could be explained by a possible non-gaussian behavior in such node, but this conjecture should be verified through a Montecarlo analysis.

![Fig. 3 – Second Order Model. CLs for THMs concentration in different nodes. Boxes represent CLs for $K_C$ and $u$, Lines represent CLs for $K_C$, $u$ and $D$. Crosses represent the measured values.](image)

5. Conclusion

The comparison of the performances of the considered first and second order models for predicting THMs formation in water supply systems has been performed calibrating the kinetic parameters using real measured data. A two-steps calibration, modifying first the chlorine decay parameters and then the $D$ coefficient, has been adopted, since the joint procedure does not represent a fruitful option for the second order model.

The comparison among the results of the two models after the calibration shows similar performances in terms of predicted concentration. These results suggest of using the simpler first order model with only two parameters. On the other side the FOSM analysis reveals that the second order model results less affected by measurements uncertainty. The values of the THMs formation kinetic coefficient obtained for both models are quite similar, indicating a $D$ value of about 40 ± 5 µg/mg, coherent with previous literature estimates.

References


