Chlorine-age based booster chlorination optimization in water distribution network considering the uncertainty of residuals

Kunlun Xin, Xiao Zhou, Hao Qian, Hexiang Yan and Tao Tao

ABSTRACT

Booster chlorination has been applied by many utilities for better chlorine-residual maintenance. In this paper, a new water-age based method for optimal location and dosage of booster disinfection has been proposed, as well as an uncertainty analysis of chlorine residuals. Chlorine-age, a novel indicator of water quality, is firstly introduced based on water age. By minimizing the total chlorine-age of nodes in a water distribution network (WDN), a new model for optimal booster location is proposed. The chlorine-age based model is independent of chlorine-decay simulation, and therefore avoids the complexity of obtaining kinetic parameters and prevents misleading results caused by inaccurate simulation of chlorine residuals. The uncertainties of chlorine residuals increase along with the distance and time consumption for delivering water. In this study, chlorine-age is employed to measure the uncertainties of nodal residuals, and optimal chlorine dosage is calculated considering the uncertainties. The proposed method has been tested on an example network and a real-life network to illustrate its validity and applicability. The results have shown that the method is feasible and reliable in practical application.

Key words | booster chlorination, optimization, uncertainty analysis, water distribution system, water quality

Kunlun Xin
Xiao Zhou
Hao Qian
Hexiang Yan (corresponding author)
College of Environmental Science and Engineering,
Tongji University,
1239 Siping Road, Shanghai, 200092,
China
E-mail: *hxyan@tongji.edu.cn*

Tao Tao
UN Environment-Tongji Institute of Environment
 for Sustainable Development,
Tongji University,
1239 Siping Road, Shanghai, 200092,
China

INTRODUCTION

Maintaining the quality of finished water in water distribution networks (WDN) has widely concerned to water supply companies in the last few decades. Chlorine residuals in a WDN help to prevent the regrowth of pathogenic bacteria and to sustain the quality of water. However, chlorine decays easily in WDN (Tamminen *et al.* 2008) with the existence of organic and inorganic compounds. Hence, booster chlorination is applied for residual maintenance in WDN. Besides, booster chlorination has various advantages such as reduction of chlorine dosage, homogenizing of residual distribution (Tryby *et al.* 1999) and decreasing of disinfectant by-products (DBPs) (Boccelli *et al.* 1998). Two main aspects of booster disinfection are

of concern to researchers: (1) optimal number and locations of booster stations (Tryby *et al.* 2002; Lansey *et al.* 2007; Ohar & Ostfeld 2014) and (2) chlorine dosage (Munavalli & Kumar 2003; Prasad *et al.* 2004). The optimization objectives may include the homogeneity of residuals, economic efficiency, the percentage of qualified water and the concentration of DBPs, etc.

Many researchers have focused on the booster station optimization problem with the defining of single or several optimization objectives. Generally, the mathematical simulation of chlorine decay is necessary for these studies. For example, Harmant *et al.* (2000) employed the Lee–Deininger approach along with coverage matrix for booster station

selecting, of which the optimization objective is the maximum set covering (MSC) of satisfactory chlorine, and residuals are set as constraints and are calculated by kinetics parameters. The other widely applied MSC model is proposed by Uber *et al.* (1998), which aims to locate the minimum number of booster stations to cover a defined set of nodes and time periods. Due to the closer relevancy between water quality and objective functions, this model is more widely adopted by researchers (Kansal & Arora 2001; Tryby *et al.* 2002). After optimizing the booster locations, chlorine dosage is considered, which is mostly focused on minimizing dosage or homogenizing residual chlorine concentration. The dosage-optimizing problems are generally solved by optimization algorithms such as the genetic algorithm (GA). For instance, Nouiri (2017) proposed a two-step model and used a multi-objective genetic algorithm (MOGA) to solve the booster location and dosage-optimizing problems sequentially. These two problems can also be simultaneously solved. Tryby *et al.* (2002) set booster location as unknown and optimized location and chlorine dosage in one objective function. Prasad *et al.* (2004) introduced the concept of qualified water and proposed a multi-objective optimization model with two objectives: qualified water volume and chlorine dosage. Behzadian *et al.* (2012) introduced the formation of trihalomethane (THM) as a new optimization objective, and a two-phase multi-objective method was employed to determine optimal location and scheduling of booster disinfection. Estimation of chlorine residuals is also necessary for these studies.

Simulations of chlorine reaction and decay in a WDN (i.e., water-quality based models) provide better results than a hydraulic model, but these models require accurate kinetic parameters. However, obtaining accurate parameters needs plenty of labor, resources and time, which sometimes is unaffordable for decision-makers. Although these parameters can be assumed by experience, they will be error-containing and may cause uncertain or misleading results in booster chlorination optimization. In contrast, hydraulic models are more easily performed with fewer parameters, but they are irrelevant to water quality. Based on water age, chlorine-age is introduced in this paper. The chlorine-age can be calculated without chlorine reaction parameters and is closely relevant to water quality. By minimizing the total chlorine-age of

different nodes in WDN, an alternative method for booster location optimization is proposed, which is independent of the kinetic parameters and is more easily conducted in practical application.

Inaccurate kinetics of chlorine decay simulation may not only influence the results of booster location optimization but also cause uncertainties in optimal booster chlorine dosage. Although the uncertainties have been studied by some researchers (Rico-Ramirez *et al.* 2007; Koker & Altan-Sakarya 2015), a majority of these studies assumed that the variances of chlorine residuals in a WDN are homogeneous, and assigned the same probability distribution function (PDF) to all nodes. However, as is shown in this study, with the increase of distance and time consumption of water delivery, uncertainties of residuals increase as well, so the assumption of homogeneity will not hold anymore. Chlorine-age is adopted here to measure the degree of uncertainties, hence it is able to assign specialized PDFs to different nodes.

The remainder of this paper is organized as follows. Chlorine-age is introduced and explained firstly, followed by the optimization algorithms and objective functions for the booster location and chlorine dosage. After that, the uncertainties of nodal residuals are analyzed, and a method for measuring the uncertainties is presented. The proposed methods are tested on an example and a real-life network, and discussions and conclusions are presented.

METHODS

Nodal chlorine-age

Longer time consumption to deliver water means more chlorine decaying and lower residual concentration. The time spent by a parcel of water in the network is called water age (Rossman 1999), which provides a simple, non-specific measure of the overall quality of delivered drinking water, defined as:

$$T_i = \frac{\sum_{n=1}^{N} q_{0i}^{(n)} T_{0i}^{(n)}}{\sum_{n=1}^{N} q_{0i}^{(n)}}; N \in U^i \tag{1}$$

where $T_{0i}^{(n)}$ = time consumption to deliver water from source to node i through path n; $q_{0i}^{(n)}$ = volume of water which
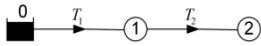
Figure 1 │ A simple network.

flows through path n; U^i = set of paths which are able to deliver water to node i.

Water age is a weighted average of time consumption along different paths. Larger nodal water age indicates more time consumption to deliver water from source to a node, as well as more chlorine residuals decaying.

As shown in Figure 1, T_1 and $T_1 + T_2$ represent the water age of node 1 and node 2, respectively. $T_1 + T_2 > T_1$, or in other words, more chlorine is consumed during water delivery from source to node 2, compared with node 1. When node 1 is set as a booster station, the rechlorination of water at node 1 refreshes the chlorine residual, and the time for residual decay to delivery of water to node 2 decreases from $T_1 + T_2$ to T_2. Here, we can deem that the water age of node 2 is reduced by rechlorination at node 1, and use chlorine-age to represent the impacts of booster stations on water age. Chlorine-age can be calculated as the weighted average time consumption to deliver water from different chlorine injectors to a specific node:

$$W_{ij} = T_{ij} - \sum_{k=1}^{n_b} T_{kj}\eta_{kij} \tag{2}$$

$$n_b < n \tag{3}$$

where W_{ij} = chlorine-age of node i at time j; T_{ij} = water age of node i at time j; n_b = number of booster stations; n = number of nodes; η_{kij} = percentage of water which comes from booster station k of node i at time j. η_{kij} integrates different paths to deliver water from station k to node i. It can be calculated with the EPANET Toolkit (Rossman 1999) by performing chemical analysis with reaction rates set to zero or by trace analysis. Although trace analysis is a water quality simulation method, it is independent of chlorine reaction parameters.

GA and NSGA-II

Inspired by natural selection, a self-adaptive global optimization probability search algorithm was initially proposed by

Holland (1975), which is called GA. In a GA, sets of solutions in the feasible region are deemed as individuals (also called creatures or phenotypes). The individuals are encoded into binary numbers as 0s and 1s, which are called the chromosomes or genotypes of individuals. At the beginning of the GA, initial individuals are randomly generated. In an iterative process, the individuals are stochastically selected based on their fitness, which is usually the value of the objective function in the optimization problem. Then, the genomes of the chosen individuals are randomly combined and mutated to generate new individuals. The iterative process terminates when any individual has satisfied the presupposed fitness level or a maximum number of iterations has been reached, and the individual which shows the best fitness is outputted as the final result.

The GA is generally used to solve single objective optimization problems. However, in many cases, two or more objective functions need to be considered, which cannot be transformed or compared together, and sometimes are contradictory. Therefore, it is impossible to find an optimum solution for all objective functions. The Pareto optimal solution set is widely used to solve multi-objective problems. In a Pareto optimal solution set, no solution can perform better than any other solution, and none of the objective functions can be improved in value without degrading some of the other objective values. Combining the multi-objective optimization problem and GA, Srinivas & Deb (1994) proposed the non-dominated sorting genetic algorithm (NSGA). After that, Deb *et al.* (2002) proposed the modified NSGA version named NSGA-II to improve some imperfections of NSGA. NSGA-II is similar to GA in selecting, combining and mutating, but the individuals are sorted according to non-domination (Deb *et al.* 2002). NSGA-II provides a Pareto optimal solution set for a multi-objective problem. Both the GA and NSGA-II are employed in this research for booster location and dosage optimization, respectively.

Notably, GA and NSGA-II are random searching algorithms, which sometimes suffer from premature convergence, resulting in local optimums rather than a global optimum. The premature convergence problem can be partially solved by some coding techniques such as mutation of the population, scaling the fitness (Whitley 1994) and elitist operation (Malik & Wadhwa 2014), and in this study, the improved coding techniques are used to avoid premature

convergence. Further, GA and NSGA-II are performed for a certain number of times and the optimum result is finally chosen.

Location of booster stations

Larger nodal chlorine-ages indicates more decaying of chlorine from different injectors to nodes. Hence, a minimal total nodal chlorine-age in a WDN represents the least diversity of chlorine residuals. To minimize the nodal chlorine-ages, an optimization model for booster station location is proposed as follows:

$$\min f_1 = \sum_{j=1}^{t} \sum_{i=1}^{n} \frac{W_{ij} q_{ij}}{Q} \tag{4}$$

subject to

$$n_b < N_b$$

where f_1 = average chlorine-age; q_{ij} = water demand of node i at time j; Q = total demand of the WDN during the simulation period; t = number of time steps; n_b = number of booster stations; N_b = maximum number of booster stations.

Equation (4) indicates the minimum total chlorine-ages in WDN, which are weighted by nodal demands. A node with higher demand has more influence over the average chlorine-age. Rather than a complicated water quality model, the chlorine-age-based optimal location model is simply based on hydraulic simulation of WDN and requires less labor and fewer resources to obtain accurate kinetic parameters. Despite the simplicity, it is closely related to water quality. Besides, for water-quality based models, inaccurate simulation of chlorine residuals may cause misleading results of optimal booster location. In case of uncertain kinetic parameters, the chlorine-age based method may provide more reliable results because the simulation of chlorine decaying is not required.

The GA and EPANET Toolkit are used to solve Equation (4). Individuals of GA represent different locations of booster stations, which are binary numbers and can be decoded into different node indexes. For each individual, W_{ij}, q_{ij} and Q can be evaluated by the EPANET toolkit.

Chlorine dosage optimization

Assuming (1) hydraulic states of a WDN are known and time-varying, (2) chlorine decay kinetics is first-order with respect to chlorine concentration, and (3) reaction rate coefficients are independent of the booster stations, the linear superposition (Boccelli *et al.* 1998) can be used for nodal residual calculation, given as

$$C_j(t) = \sum_{i=1}^{n} \sum_{k=1}^{t} a_{ij}^{kt} u_i^k \tag{5}$$

where $C_j(t)$ = residual concentration of downstream node j at time t; u_i^k = chlorine dosage of upstream injector i at time k; $a_{ij}^{kt} = \partial C_j(t)/\partial u_i^k$, describing the effect of u_i^k on the concentration at node j and time t. Equation (5) indicates that $C_j(t)$ is the linear summation of all chlorine injector effects. To simplify the calculation, the dosages of injectors are supposed as constants. Then Equation (5) can be rewritten as

$$C_j(t) = \sum_{i=1}^{n} a_{ij}^{t} u_i \tag{6}$$

EPANET is employed to calculate a_{ij}^t, and Flow Paced Booster (FPB) mode (Rossman 1999) is chosen as the source type. A detailed flowchart for a_{ij}^t calculation is shown in Figure 2.
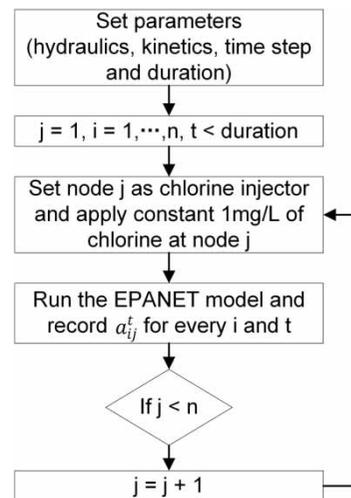

Figure 2 | Flowchart for a_{ij}^t calculation.

A two-objective optimization model is proposed to optimize the chlorination while minimizing the cost. As shown in Equations (7) and (8), the percentage of qualified water is used to indicate the effectiveness of chlorination, and total chlorine dosage is used to measure the cost. NSGA-II linked with the EPANET Toolkit is employed to solve the model:

$$\max \ \frac{\sum_i \sum_k d_i^k \varnothing(c_i^k)}{D} \tag{7}$$

$$\min \ \sum_{m=1}^{n_b} \sum_{k=1}^{t} u_m d_m^k \Delta t \tag{8}$$

where d_i^k = demand of node i at time k; $\varnothing(c_i^k)$ = qualified water multiplicator (Prasad *et al.* 2004), if $C_{\min} \le c_i^k \le C_{\max}$, $\varnothing(c_i^k) = 1$, otherwise $\varnothing(c_i^k) = 0$; D = total demand during the simulation; n_b = number of booster stations; u_m = booster chlorine dosage of station m; d_m^k = water inflow of station m at time k; Δt = time step of simulation model. The constraints of residuals, $C_{\min}$ and $C_{\max}$, are set as 0.2 mg/L and 4 mg/L respectively in this study.

Uncertainty analysis of nodal residuals

Generally, the simulation results of hydraulic and water quality models are certain. However, the model outputs can hardly be the same as the actual values due to the uncertainties of the model parameters. Many factors give rise to the uncertainties of simulated chlorine residuals, mainly including the following. (1) Uncertain hydraulic parameters, like nodal demands and pipe roughness coefficients. These parameters are necessary for a WDN hydraulic model to represent the real network accurately, but some of them cannot be obtained directly or need unaffordable resources to be sufficiently and precisely measured. The presumed uncertain parameters lead to inaccurate hydraulic simulation results like pipe flow and velocity, which will further result in inaccurate simulation of chlorine decay and distribution. (2) The uncertainty of chlorine decay. The reaction kinetic parameters are difficult to measure and are unsteady due to the altering pipe situation, hence the simulation of chlorine decay is uncertain. (3) Random events, such as leakage, pipe burst and changing of valve

status may occur at any time and influence the hydraulic and water quality conditions.

To sufficiently quantify the influence of the above factors could be impossible. Our methodology integrates all the uncertainties into the influence coefficient a_{ij}^{kt} (see Equations (5) and (6)). By assigning different PDFs to the influence coefficients, the uncertainties of residuals can be integrated into the proposed optimization method.

The uncertainties of residuals increase along with the distance and time consumption of water delivery. The positive relationship is illustrated with a simple network shown in Figure 3. The length, diameter, bulk coefficient, and wall coefficient of each pipe is set as 500 m, 200 mm, −0.1, −1 respectively. Demands of all nodes are set as 10 L/s, and the chlorine dosage of the water source is set as 2 mg/L with continuous injection for 24 hours. Water quality simulation shows that the simulated chlorine concentration values of the three nodes at the 8th hour are 1.68 mg/L, 1.44 mg/L, 1.16 mg/L, respectively.

To simulate the uncertainty of nodal demands, a zero-mean Gaussian noise with 1.0 standard deviation is added to the three nodal demands. Using the Monte Carlo simulation (MCS) method (Metropolis & Ulam 1949), water quality simulations are performed 10,000 times and the statistical data of residual concentration at the 8th hour are shown in Table 1.

The mean values agree with the results of simulation under certain demands. The standard deviation and difference between the maximum and the minimum values of the three nodes are increasing, showing that nodes farther from the source with more time consumption for water delivery have a larger uncertainty of residuals. Furthermore, if all the uncertain factors are considered, the difference of residual uncertainties would be more significant.

Chlorine-age is the time consumption for delivering water from different chlorine injectors to a node. As more time consumption refers to a larger uncertainty of nodal residual, chlorine-age is employed here to measure the uncertainty. For node α with the minimum chlorine-age
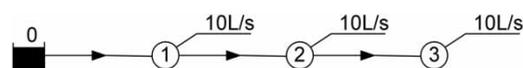

Figure 3 │ A simple network.

Table 1 | MCS results of residual concentration

	Max value (mg/L)	Min value (mg/L)	Mean value (mg/L)	Standard deviation (mg/L)
Node 1	1.697	1.666	1.680	0.0039
Node 2	1.460	1.409	1.437	0.0072
Node 3	1.197	1.096	1.161	0.0121

($W_{\min}$), its influence coefficient $a_{i\alpha}^t$ (see Equation (6)) is assumed to be subject to

$$\forall i, P\left(\left|a_{i\alpha}^t - \overline{a_{i\alpha}^t}\right| \leq \gamma \overline{a_{i\alpha}^t}\right) = \chi_\alpha \tag{9}$$

where $\overline{a_{i\alpha}^t}$ = mean value of $a_{i\alpha}^t$; γ = convergence coefficient, $\gamma \in (0, 1)$; χ_α = probability coefficient, $\chi_\alpha \in (0, 1)$. Equation (9) denotes that $a_{i\alpha}^t$ has a probability of χ_α to satisfy $a_{i\alpha}^t \in \left[\overline{a_{i\alpha}^t}(1 - \gamma), \overline{a_{i\alpha}^t}(1 + \gamma)\right]$.

Similarly, for node β with the maximum chlorine-age ($W_{\max}$), $a_{i\beta}^t$ is assumed to be subject to:

$$\forall i, P\left(\left|a_{i\beta}^t - \overline{a_{i\beta}^t}\right| \leq \gamma \overline{a_{i\beta}^t}\right) = \chi_\beta \tag{10}$$

As the chlorine-age of node α is smaller than node β, $\chi_\alpha > \chi_\beta$ with the same γ because node β has larger residual uncertainty. Referring to the actual operation data from water-supply companies, in this study, χ_α, χ_β and γ are set as 99%, 80% and 0.2 respectively.

Equations (9) and (10) can be rewritten as

$$P\left(\frac{-\gamma \overline{a_{i\alpha}^t}}{\sigma_{i a}} \leq \frac{a_{i\alpha}^t - \overline{a_{i\alpha}^t}}{\sigma_{i a}} \leq \frac{\gamma \overline{a_{i\alpha}^t}}{\sigma_{i a}}\right) = \chi_\alpha \tag{11}$$

and

$$P\left(\frac{-\gamma \overline{a_{i\beta}^t}}{\sigma_{i\beta}} \leq \frac{a_{i\beta}^t - \overline{a_{i\beta}^t}}{\sigma_{i\beta}} \leq \frac{\gamma \overline{a_{i\beta}^t}}{\sigma_{i\beta}}\right) = \chi_\beta \tag{12}$$

where $\sigma_{i a}$, $\sigma_{i\beta}$ are the standard deviations of $a_{i\alpha}^t$, $a_{i\beta}^t$, respectively. According to the theory of Gaussian distribution (Patel & Read 1995), $\sigma_{i a}$ and $\sigma_{i\beta}$ can be solved when χ_α, χ_β and γ are known.

For any node j with chlorine-age W, it is assumed that the standard deviation σ_{ij} varies linearly according to chlorine-age, i.e.

$$\sigma_{ij} = \sigma_{i a} + (\sigma_{i\beta} - \sigma_{i a}) \frac{W - W_{\min}}{W_{\max} - W_{\min}} \tag{13}$$

Therefore, if the distribution of the influence coefficients is Gaussian, for any node j, $a_{ij}^t \sim N\left(\overline{a_{ij}^t}, \sigma_{ij}\right)$.

MCS exactly converges to the uncertainty estimates but is computationally demanding. Kang *et al.* (2009) tested three approximate methods including MCS, Latin hypercube sampling (LHS) and first-order second-moment (FOSM) methods on WDN uncertainty analysis problems, and concluded that LHS provides accurate analysis with better stability (compared with FOSM) and lower computational requirements (compared with MCS). Therefore, the LHS method (Kang *et al.* 2009), which is a stratified sampling method, is adopted in this study to evaluate optimal chlorine dosage considering the uncertainty of residuals.

Based on the methods presented above, an overall flow-chart to optimize booster chlorination including booster station, chlorine dosage and uncertainty analysis is shown in Figure 4. The optimizations of booster station and chlorine dosage are separated into two sequential steps because: (1) the objective functions of booster station optimization and dosage optimization is different, which make it difficult to combine the two problems while keeping their characteristics; (2) considering the two problems together leads to a larger solution space, and results in difficulties for the GA to find a good solution. Although the sequential method may miss the global optimal solution, it more easily finds the optimum solution for each step, and ensures relatively good results.

RESULTS AND DISCUSSION

The booster chlorination optimization and uncertainty analysis method described above are applied to two cases to illustrate their performances and applications. Assuming periodicity of network hydraulic solutions with the dosages assumed as constants, the water quality simulation results will be periodic after running for some time (Boccelli *et al.* 1998). Therefore, the water ages and influence coefficients
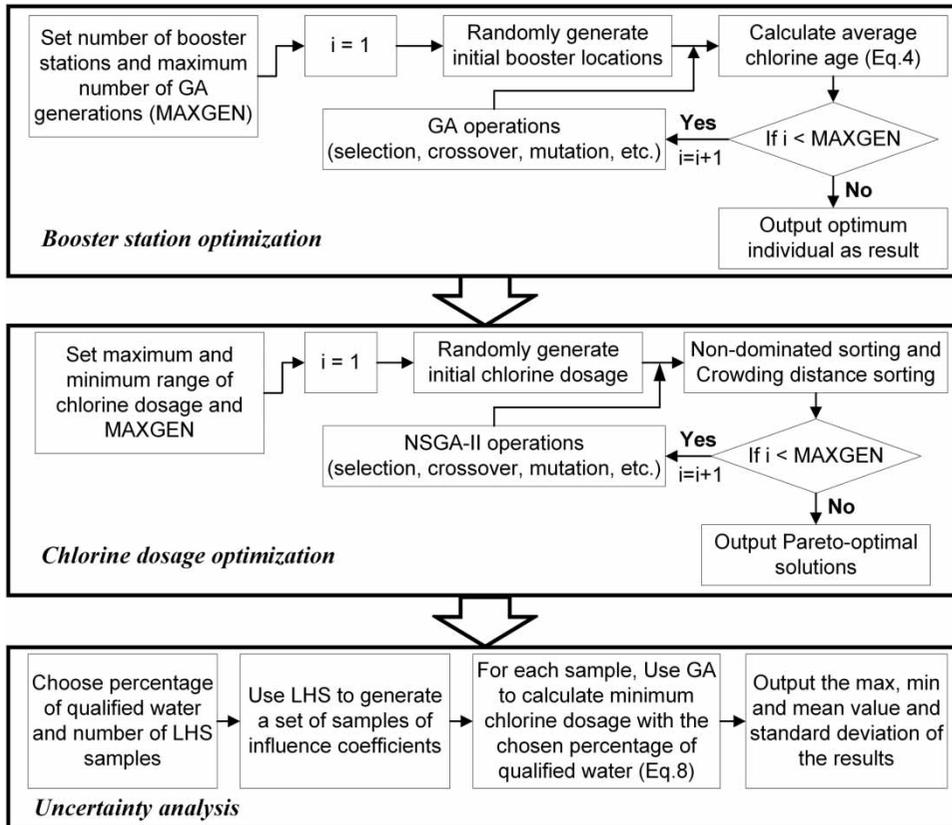
Figure 4 | An overall flowchart of booster chlorination optimization.

(a_{ij}^t) of one cycle can be collected to represent the long-running network for optimal chlorination analysis. They are obtained after the periodic changes of water quality simulation results reach a steady state. Specifically, the data of 49–72 h simulation of an example network and the data of 73–96 h simulation of a real-life network are collected.

CASE 1: EXAMPLE NETWORK

The proposed model is firstly tested with an example network which contains 16 pipes and 23 nodes. The bulk coefficient and wall coefficient of chlorine decay in this simulation are set as $-0.1 \, \mathrm{d}^{-1}$ and $-1.5 \, \mathrm{d}^{-1}$ respectively, and the time step of hydraulic and water quality simulation are set as 1 hour and 15 min respectively.

At first, the optimal locations of booster stations (see Equation (4)) with $N_b = 1, 2, 3$ are calculated respectively. The results are shown in Table 2 and Figure 5.

As shown in Table 2, when N_b changes from 1 to 2, the second location at node 5 is selected following the first one at node 17. This may indicate that booster chlorine for the nodes at the end of the network (i.e., nodes 18–23) can significantly improve water quality. The nodes in the bottom right of the network are far from the source (node 1) since they have relatively less residual concentration and greater chlorine-age. Even if node 17 is set as a booster station, nodes 10, 11, 15 and 16 are still far from the chlorine injectors and have greater chlorine-age than other nodes. Therefore, when $N_b = 3$, node 9 is selected as the new

Table 2 | Optimal location of booster stations and corresponding average chlorine-age

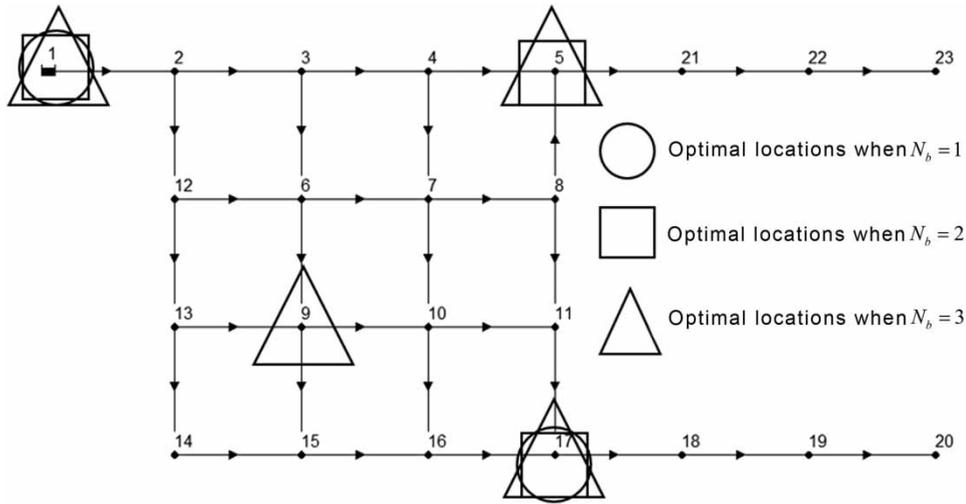| N_b	Optimal locations	Average chlorine-age (h)
0	Node 1	10.53
1	Node 1, node 17	7.46
2	Node 1, node 5, node 17	6.18
3	Node 1, node 5, node 9, node17	4.92

Figure 5 | Optimal locations of booster stations.

booster location to minimize total chlorine-age by reducing the chlorine-age of nodes 10, 11, 15 and 16. In that case, the average chlorine-age is significantly less than when $N_b = 2$, which indicates more uniform distribution of residuals.

After booster stations are chosen, the chlorine dosages of booster stations are optimized in the cases of different N_b (see Equations (7) and (8)). The result is shown in Figure 6: the more booster stations, the less total dosage. However, the decrease of total dosage is damped with the increase of N_b. For instance, to provide 95% qualified water, the decrease of total dosage is 118 kg/d, 35 kg/d,

12 kg/d respectively when N_b increases from 0 to 3. Considering the construction costs of booster stations, an appropriate number of booster stations is important for higher water quality as well as lower chlorine dosage. There are also other factors (e.g. complexity in management) that influence the decision of N_b, but are not discussed here.

To provide more qualified water, more chlorine is required when there are no or few booster stations (e.g. for $N_b = 0$ or 1 in Figure 6) compared with situations of more booster stations (e.g. for $N_b = 2$ or 3 in Figure 4). The residuals of nodes far from the sources or booster stations
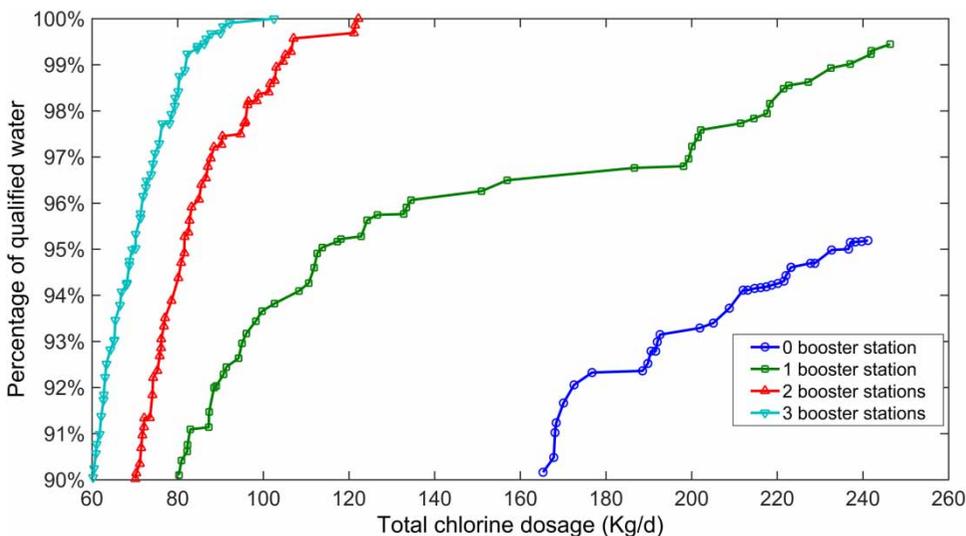

Figure 6 | Pareto curve for chlorine dosage – qualified water (example network).

Table 3 | Chlorine dosage of each injector

Node ID	1	5	17
Dosage (mg/L) (ignore uncertainty)	1.07	0.83	1.25
Dosage (mg/L) (consider uncertainty)	1.16	0.99	1.40

are critical factors which affect the ratio of qualified water. With more booster stations, the minimum residual constraints of those nodes are more easily met. It can also be seen from Figure 6 that the slope of the dosage-qualified water curve is larger with the increase of N_b.

Now consider the uncertainty of chlorine residuals. Two booster stations with a goal of 96% of qualified water (85 kg/d total chlorine) are chosen as a benchmark in the following discussion. The residuals are recalculated considering the uncertainties of influence coefficients according to Equation (13). LHS is employed to generate 500 samples of influence coefficients, and the minimum chlorine dosage of each sample is calculated respectively.

Considering the uncertainty of residuals, the optimal chlorine dosage varies. The maximum, minimum, mean value and standard deviation of the optimal chlorine dosage are 93.76 kg/d, 82.08 kg/d, 87.93 kg/d and 1.99 kg/d, respectively. The total dosage should be increased appropriately if uncertainties are considered. When the dosage is 87.93 kg/d, the ratio of qualified water just has 50% probability of reaching 96%, which is obviously insufficient. If 95% confidence is required, the total chlorine dosage should be 91.70 kg/d. The optimal dosage of each booster station is listed in Table 3.

CASE 2: REAL-LIFE NETWORK

To further illustrate the validity, the proposed method is applied to a real-life large-scale WDN. The network lies in a district in Shanghai, eastern China. As shown in Figure 7, the network is composed of 4,003 pipes and 4,878 nodes, and supplies 85,000 tons of water to 291,700 people per day. The water treatment plant supplies water to the whole network, and the booster pump station boosts pressure for the eastern areas which are far from the plant. Before booster chlorination optimization, the managers apply about 1.76 mg/L of chlorine at the plant and about 0.2 mg/L of

chlorine at the booster pump station. The distribution of chlorine residuals is shown in Figure 7(a). It can be seen that a majority of nodes satisfy the water qualify constraints, but the residuals in the eastern areas are relatively low and there is significant diversity between the western and eastern areas.

The average chlorine-age of the network without booster chlorination is 10.58 h. Considering the construction costs, the number of booster chlorine stations is set as 1 (i.e. $N_b = 1$) in this case. Equation (4) is used to calculate the optimal booster station location, and the result shows that the optimal location is at the original booster pump station. After the booster station is applied, the average chlorine-age decreases to 6.50 h, indicating that the booster station promotes the overall water quality obviously.

The chlorine dosage is optimized by NSGA-II subsequently, and the results are shown in Figure 8. When the percentage of qualified water is 90–94%, a slight increase of total chlorine dosage gains qualified water significantly. However, when the percentage is larger than 94%, the slope of the Pareto curve reduces sharply and the increase of chlorine dosage gains little. Considering both security and efficiency, 94% of qualified water is chosen as the acceptable level for this case. The chlorine dosages of the treatment plant and booster station are 0.85 mg/L and 1.36 mg/L, respectively, and the total dosage is 120.26 kg/d.

As tested in the example network, the total chlorine dosage increases when considering the uncertainty of residuals. Similarly, 500 samples are generated by LHS and the chlorine dosages of the samples are calculated respectively. The maximum, minimum, average value and standard deviation of the sampling results are 133.37 kg/d, 125.32 kg/d, 128.88 kg/d and 1.44 kg/d, respectively. It can be seen that the uncertainties of residuals will impact the overall water quality, which should be improved by applying more chlorine. If 95% confidence is required, the total dosage should be 131.71 kg/d and the dosages of treatment plant and booster station are 0.92 and 1.58 mg/L, respectively.

The overall comparison of before and after booster chlorination optimization is shown in Figure 7 and Table 4. After optimization, the total chlorine dosage decreases by 34.1%, indicating that the booster chlorination provides better economic efficiency. In addition, the distribution of residuals is
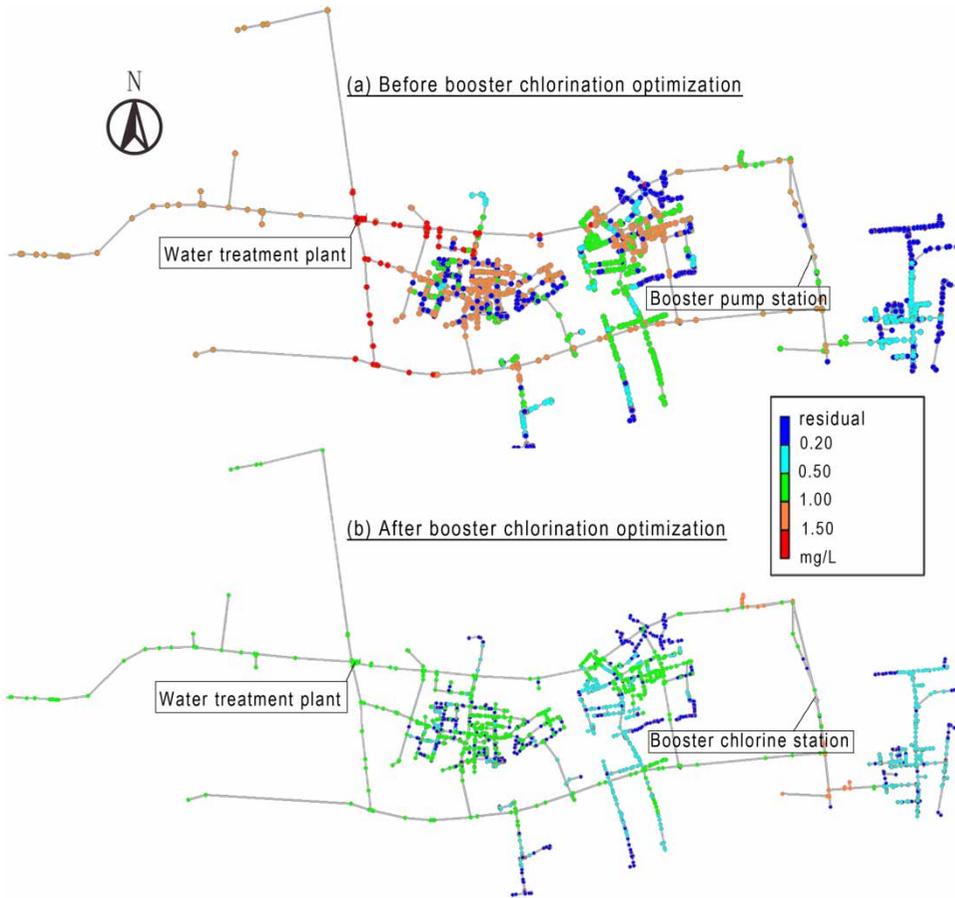
Figure 7 │ Residual concentration at 24:00 of the real-life network (a) before booster chlorination optimization and (b) after optimization.
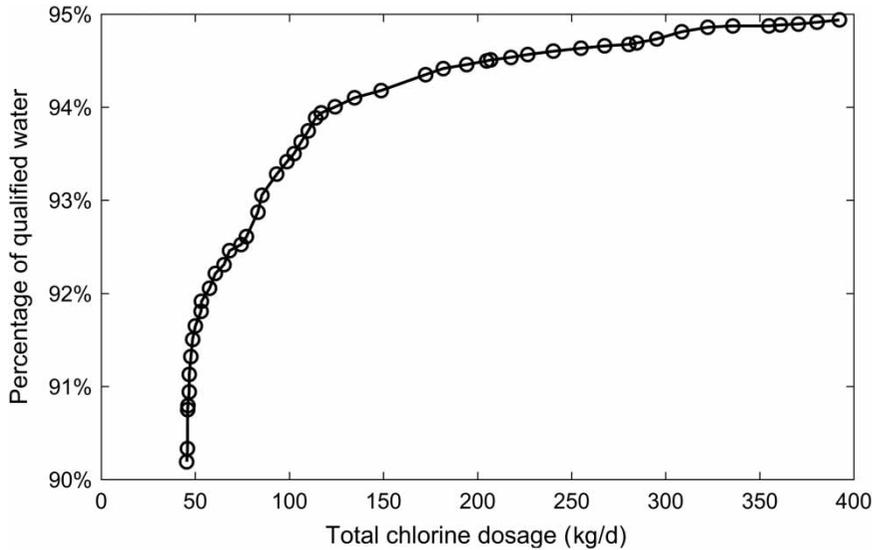

Figure 8 │ Pareto curve for chlorine dosage – qualified water (real-life network).

Table 4 | Comparison of chlorine dosage before and after booster chlorine optimization

	Dosage of treatment plant (mg/L)	Dosage of booster station (mg/L)	Total dosage (kg/d)	Percentage of qualified water	confidence
Before optimization	1.76	0.2	199.87	94%	–
After optimization (ignore uncertainty)	0.85	1.36	120.26	94%	–
After optimization (consider uncertainty)	0.92	1.58	131.71	94%	95%

more uniform and the overall residual concentration decreases, which may reduce DBP formation and prevent consumers from complaining of taste and odor. What is more, although the percentages of qualified water before and after optimization are all 94%, as the uncertainties of residuals are considered, there is better confidence of it satisfying the qualified water requirements.

CONCLUSIONS

Booster chlorination helps to decrease chlorine consumption, make uniform the distribution of residuals and reduce DBP formation. A majority of booster chlorine optimization methods are based on the simulations of chlorine reaction in the WDN, which require accurate kinetic parameters. As these parameters are hard to obtain and contain errors, an inaccurate water quality model may result in error-containing or misleading results in choosing optimal locations of booster stations and overconfident estimation of optimal chlorine dosage. In this study, a new water quality indicator, chlorine-age, is introduced. Methods to optimize booster chlorination based on chlorine-age are proposed, including optimal location, chlorine dosage optimization, and uncertainty analysis of residuals. The methods are able to find the optimal locations of booster stations in WDN without the requirement of complex chlorine reaction kinetic parameters. To minimize the chlorine dosage while maximizing the percentage of qualified water, a two-objective optimization model is built based on the existing research on optimal chlorination. This study tested the uncertainties of chlorine residuals and concluded that the uncertainties vary along with the distance and time consumption for delivering water. Hence, chlorine-age is employed to measure the uncertainties. A new linear relationship between the standard deviation of residuals

and nodal chlorine-age is presented to estimate the uncertainty of residuals. After that, optimal chlorine dosage considering the uncertainties is further studied. The results show that the optimum chlorine dosage should be increased to ensure confidence in water quality when the uncertainties of residuals are considered.

This study assumes that the chlorine decay kinetics is first-order with respect to chlorine concentration, and the uncertainties of chlorine residuals are Gaussian distributed. As many researchers verified that second-order or mixed first-order models have better accuracy, further research should focus on combining the proposed method with these models. What is more, the uncertainty of nodal residuals may not exactly follow the Gaussian distribution, hence further research on the distribution of residuals is recommended.

REFERENCES

Behzadian, K., Alimohammadnejad, M., Ardeshir, A., Jalilsani, F. & Vasheghani, H. 2012 A novel approach for water quality management in water distribution systems by multi-objective booster chlorination. *International Journal of Civil Engineering* **10** (1), 51–60.

Boccelli, D. L., Tryby, M. E., Uber, J. G., Rossman, L. A., Zierolf, M. L. & Polycarpou, M. M. 1998 Optimal scheduling of booster disinfection in water distribution systems. *Journal of Water Resources Planning and Management* **124** (2), 99–111.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6** (2), 182–197.

Harmant, P., Nace, A. & Kiene, L. 2000 An algorithm to optimize booster chlorination in water distribution network. In: *Building Partnerships: Proceedings of the 2000 Joint Conference on Water Resources Engineering and Water Resources Planning & Management* (R. H. Hotchkiss & M. Glade, eds), ASCE, Reston, VA, USA.

Holland, J. H. 1975 Adaptation in natural and artificial systems. *Quarterly Review of Biology* **6** (2), 126–137.

Kang, D. S., Pasha, M. F. K. & Lansey, K. 2009 Approximate methods for uncertainty analysis of water distribution systems. *Urban Water Journal* **6** (3), 233–249.

Kansal, M. & Arora, G. 2001 Reliability analysis of booster chlorination in urban water distribution networks. In: *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges – Proceedings of the World Water and Environmental Resources Congress.*

Koker, E. & Altan-Sakarya, A. B. 2015 Chance constrained optimization of booster chlorination in water distribution networks. *Clean: Soil Air Water* **43** (5), 717–723.

Lansey, K., Pasha, F., Pool, S., Elshorbagy, W. & Uber, J. 2007 Locating satellite booster disinfectant stations. *Journal of Water Resources Planning and Management* **133** (4), 372–376.

Malik, M. & Wadhwa, S. 2014 Preventing premature convergence in genetic algorithm using DGCA and elitist technique. *International Journal of Advanced Research in Computer Science and Software Engineering* **4** (6), 410–418.

Metropolis, N. & Ulam, S. 1949 The Monte Carlo method. *Journal of the American Statistical Association* **44** (247), 335–341.

Munavalli, G. R. & Kumar, M. S. M. 2003 Optimal scheduling of multiple chlorine sources in water distribution systems. *Journal of Water Resources Planning and Management* **129** (6), 493–504.

Nouiri, I. 2017 Optimal design and management of chlorination in drinking water networks: a multi-objective approach using genetic algorithms and the Pareto optimality concept. *Applied Water Science* **7**, 3527–3538.

Ohar, Z. & Ostfeld, A. 2014 Optimal design and operation of booster chlorination stations layout in water distribution systems. *Water Research* **58**, 209–220.

Patel, J. K. & Read, C. B. 1995 Handbook of the normal distribution. *Technometrics* **25** (1), 112–115.

Prasad, T. D., Walters, G. A. & Savic, D. A. 2004 Booster disinfection of water supply networks: multiobjective approach. *Journal of Water Resources Planning and Management* **130** (5), 367–376.

Rico-Ramirez, V., Iglesias-Silva, G. A., Gomez-De la Cruz, F. & Hernandez-Castro, S. 2007 Two-stage stochastic approach to the optimal location of booster disinfection stations. *Industrial & Engineering Chemistry Research* **46** (19), 6284–6292.

Rossman, L. A. 1999 The EPANET programmer's toolkit for analysis of water distribution systems. In: *29th Annual Water Resources Planning and Management Conference*, Tempe, AZ, USA.

Srinivas, N. & Deb, K. 1994 Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation* **2** (3), 221–248.

Tamminen, S., Ramos, H. & Covas, D. 2008 Water supply system performance for different pipe materials part I: water quality analysis. *Water Resources Management* **22** (11), 1579–1607.

Tryby, M. E., Boccelli, D. L., Koechling, M. T., Uber, J. G., Summers, R. S. & Rossman, L. A. 1999 Booster chlorination for managing disinfectant residuals. *Journal American Water Works Association* **91** (1), 95–108.

Tryby, M. E., Boccelli, D. L., Uber, J. G. & Rossman, L. A. 2002 Facility location model for booster disinfection of water supply networks. *Journal of Water Resources Planning and Management* **128** (5), 322–333.

Uber, J. G., Polycarpou, M. M. & Subramaniam, P. 1998 Optimal decoupling of booster disinfection systems in water distribution networks. In: *Water Resources and the Urban Environment: 25th Annual Conference on Water Resources Planning and Management* (E. D. Loucks, ed.), ASCE, Reston, VA, USA.

Whitley, D. 1994 A genetic algorithm tutorial. *Statistics and Computing* **4** (2), 65–85.