



Disinfectant control in drinking water networks: Integrating advection–dispersion–reaction models and byproduct constraints[☆]

Salma M. Elsherif^{a,b,*}, Ahmad F. Taha^a, Ahmed A. Abokifa^c

^a Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, USA

^b Department of Irrigation and Hydraulics Engineering, Faculty of Engineering, Cairo University, Giza, Egypt

^c Department of Civil, Materials, and Environmental Engineering, The University of Illinois Chicago, Chicago, IL, USA

ARTICLE INFO

Keywords:

Disinfection byproducts
Water quality control
Chlorine
Advection–dispersion–reaction
Booster stations injections

ABSTRACT

Effective disinfection is essential for maintaining water quality standards in distribution networks. Chlorination, as the most used technique, ensures safe water by maintaining sufficient chlorine residuals but also leads to the formation of disinfection byproducts (DBPs). These DBPs pose health risks, highlighting the need for chlorine injection control (CIC) by booster stations to balance safety and DBPs formation. Prior studies have followed various approaches to address this research problem. However, most of these studies overlook the changing flow conditions and their influence on the evolution of the chlorine and DBPs concentrations by integrating simplified transport-reaction models into CIC. In contrast, this paper proposes a novel CIC method that: (i) integrates multi-species dynamics, (ii) allows for a more accurate representation of the reaction dynamics of chlorine, other substances, and the resulting DBPs formation, and (iii) optimizes for the regulation of chlorine concentrations subject to EPA mandates thereby mitigating network-wide DBPs formation. The novelty of this study lies in its incorporation of time-dependent controllability analysis that captures the control coverage of each booster station. The effectiveness of the proposed CIC method is demonstrated through its application and validation via numerical case studies on different water networks with varying scales, initial conditions, and parameters.

1. Introduction

Disinfectants play a pivotal role in water distribution networks (WDNs) to ensure compliance with water quality (WQ) safety standards. By maintaining the water pathogen-free, the occurrence of multiple waterborne diseases (e.g., cholera, typhoid) in the United States has significantly decreased since the utilization of disinfection process at the beginning of the 20th century (Constable and Somerville, 2003). Of the various disinfection methods available, chlorination is the most commonly used technique in WDNs (Grayman, 2018). While serving as a proxy for WQ monitoring, the objective is to have a sufficient chlorine residual all over the network to maintain safe water. However, chlorine actively reacts with various substances such as bacteria, organic matter, and microbial chemicals, leading to the formation of disinfection byproducts (DBPs) (Wang et al., 2013; Zhou et al., 2023; Shah et al., 2024). According to the Centers for Disease Control, exposure to these DBPs may increase the risk of cancer, liver damage, and decreased nervous system activity (DBP, 2022; Kalita et al., 2024). Therefore, it is crucial that chlorine injections into the

network by distributed booster stations are strategically controlled to achieve the primary goal of maintaining sufficient residuals and limiting the formation of DBPs.

Furthermore, accurate representation and detailed modeling are essential for effectively controlling chlorine concentrations in water distribution networks. The transport and reaction dynamics of chlorine are modeled using advection–dispersion–reaction partial differential equations (ADR-PDEs). While some studies neglect the dispersion process, assuming that advection dominates in zones with high velocities, this overlooks the importance of dispersive transport in low-flow velocity pipes, especially in networks with laminar flow conditions and numerous dead-ends. This is because exclusively advective transport models may either underestimate or overestimate actual concentrations within the low-velocity zones (Abokifa et al., 2016). Additionally, in contrast to the widely adopted single-species chlorine decay model with a constant decay rate, multi-species models offer a more accurate representation of chlorine dynamics in WDNs (Jonkergouw et al., 2009; Elsherif et al., 2023). These models consider scenarios where chlorine

[☆] This work is supported by the National Science Foundation, United States under Grants 2151392 and 2015603.

* Corresponding author.

E-mail addresses: salma.m.elsherif@vanderbilt.edu (S.M. Elsherif), ahmad.taha@vanderbilt.edu (A.F. Taha), abokifa@uic.edu (A.A. Abokifa).

interacts with other substances at rates different from the decay rate. Such scenarios include, but not limited to, contamination intrusion and substances derived from pipe materials. Furthermore, these multi-species models can be utilized to reflect the formation of DBPs, such as trihalomethanes (THMs), providing a comprehensive understanding of the WQ dynamics in WDNs (Moeini et al., 2023).

To that end, this paper introduces a chlorine control approach that incorporates multi-species dynamics to maintain chlorine concentrations within EPA-specified limits (Acrylamide, 2009) while limiting the formation of DBPs. This control approach is formulated based on a detailed transport model that accounts for the dispersion effect in low-velocity zones of the network providing precise simulations of chlorine and other substances. Furthermore, the control approach *strategically* allocates injections between booster stations based on a prior controllability analysis. This analysis implicitly prioritizes booster stations with broader coverage and assigns lower weights to zones with high existing DBPs concentrations. In the following section, we survey the literature on this topic and identify gaps that our study aims to address.

1.1. Literature review

The literature on chlorine modeling and control is broad and abundant, covering various reaction and decay dynamics, modeling techniques, and control approaches. Each of these aspects comes with its own set of underlying assumptions and limitations, contributing to the complexity of the topic. Next, we provide a brief summary of the literature on (i) modeling the transport, decay, and reaction dynamics of chlorine and the subsequent formation of the DBPs, and (ii) control techniques, objectives, constraints, and the specific scenarios under which they are applied.

Chlorine Modeling and DBPs Formation

Accurate chlorine modeling is essential for the development of effective WQ control frameworks. Chlorine evolution in WDNs is covered by the transport, decay, and reaction dynamics. The governing ADR-PDE divides these dynamics into three processes. The advection process accounts for the transport of chlorine along with the flowing water through the network, driven by the velocity of the water flow. The dispersion process captures the spreading of chlorine within the water flow caused by the irregularities and turbulence in the flow pattern and the diffusion of chlorine molecules across concentration gradients. Lastly, the reaction process includes chlorine decay, the formation of DBPs, and chlorine interaction with other substances in the water.

In WDNs, flow conditions are dependent on the consumers demand, connecting components, and pipes characteristics (i.e., diameter, length, material). Consequently, certain network zones exhibit turbulent flow conditions, while others demonstrate laminar flow conditions, particularly in dead-end sections. The transition between turbulent and laminar flow in various zones fluctuates throughout the day in response to changes in demand and the dynamics of storage components. In turbulent flow conditions and high flow velocities, advection has a greater influence than dispersion on the change of chlorine concentrations. Hence, in several studies (Grayman, 2018; Engineering Consultant et al., 2012; Elsherif et al., 2023) the dispersion process is neglected, with the focus placed on advection and reaction dynamics under the assumption of limited dead-end branches, higher velocities, and changing demands leading to a turbulent flow states. On the other hand, for WDNs with dead-ends comprising high percentage of the network, studies have proven that neglecting the dispersion effect results in inaccurate chlorine concentrations (Tzatchkov et al., 2002; Abokifa et al., 2016; Li et al., 2005). Furthermore, authors in Li et al. (2005) have linked the input patterns of chlorine injections to the dispersion effect, demonstrating the importance of including the dispersion in laminar flow zones when these injection patterns exhibit dramatic changes, a scenario common in many chlorine control

applications. To that end, to build a generalized model for chlorine evolution in WDNs, the model has to be adaptable to various flow conditions and network characteristics, including layout and topology. This adaptability allows for the accurate simulation of the physically-representative dynamics. This generalized model to be integrated into chlorine control framework is a gap to be filled in this paper.

In addition, a plethora of studies have investigated the modeling of the decay and reaction dynamics of chlorine in WDNs. These distinguish between these model vary in their order and the substances included (Helbling and VanBriesen, 2009). The most commonly used model in developing chlorine control frameworks is the first-order single-species decay model. This model assumes that chlorine is linearly decaying in time as a result of its reaction with the bacteria and microbial organics in the water (Jonkergouw et al., 2009). Nonetheless, this model fails to account for various scenarios where chlorine is actively reacting with other substances (e.g., contamination intrusion, substance from pipe material), as well as the formation of DBPs (Fisher et al., 2017). Conversely, the multi-species reaction model is an advanced approach capable of capturing these scenarios by introducing nonlinear reaction expressions that enable modeling the evolution of other substance(s) reacting with chlorine and the resultant formation of DBPs (Boccelli et al., 1998; Elsherif et al., 2023; Moeini et al., 2023; Li et al., 2019). In this paper, we develop a chlorine control approach based on multi-species reaction model. This model simulates chlorine decay and reaction with a fictitious reactant, which can represent various substances or chemicals in a generalized form, and the resultant formation of DBPs.

Lastly, by defining the processes and reaction dynamics to be modeled by the ADR-PDEs, the techniques used to solve these equations differ. Due to the interplay between the dynamics of different components and the complexity added by the unique layout of each network, analytical solutions do not exist for these equations. Therefore, spatio-temporal discretization schemes are utilized to solve these equations and obtain the concentrations of chemicals (Rossman and Boulos, 1996). These schemes are categorized based on how the spatio-temporal grid is constructed and the interdependency between the calculations of grid points' concentrations. They are classified as Eulerian, Lagrangian, or hybrid Eulerian–Lagrangian based according to the grid construction approach. Eulerian schemes divide the grid into fixed-size meshes over space and time, while Lagrangian schemes use variable-sized segments. Hybrid schemes combine aspects of both, with fixed-size segments over time and variable segments over space, or vice versa. Moreover, they are classified based on which neighboring points' concentrations are considered in the calculation of a segment's concentration and whether these concentrations are from the previous time-step or the current one (Basha and Malaeb, 2007; Rossman and Boulos, 1996).

Several studies (Munavalli and Mohan Kumar, 2005; Blokker et al., 2007) employ one or more of these schemes in their studies of modeling the chlorine decay under advection-dominant conditions. Study (Elsherif et al., 2023) surveys the applicability of a range of Eulerian and Lagrangian based schemes on the chlorine multi-species advection-reaction PDEs, validating their performance against EPANET and its extension EPANET-MSX, which utilizes a Lagrangian-based scheme (Rossman et al., 2020). Other studies (Abokifa et al., 2020, 2016; Tzatchkov et al., 2002; Ozdemir, 2023) apply different schemes to simulate the chlorine decay and transport, while accounting for the dispersion effect. However, none of these studies investigate these scheme's performance while considering the multi-species dynamics and the alternation between advection- and dispersion-dominant conditions.

Chlorine Control Approaches and Underlying Scenarios

The topic of chlorine control is addressed in many studies, each with different objective function(s), constraints, and approaches to solving the problem, while falling within specific scenario frameworks (Fisher et al., 2018; Ding et al., 2024). Several studies focus on minimizing

chlorine injections while maintaining specific chlorine residual levels throughout the network, utilizing various optimization techniques such as linear programming (Boccelli et al., 1998; Constans et al., 2012), nonlinear programming (Mala-Jetmarova et al., 2017), and nonlinear genetic algorithm (Munavalli and Kumar, 2003). Additionally, other studies couple the objective of minimizing chlorine injections with the booster stations allocation as a multi-objective programming problem (Tryby et al., 2002; Pineda Sandoval et al., 2021; Ayvaz and Kentel, 2015). Moreover, several studies have incorporated the DBPs formation issue into their optimization frameworks (Zhang and Lu, 2021). A straightforward approach involves implicitly minimizing chlorine injections to reduce DBP formation (Prasad et al., 2004). Furthermore, studies such as Ardesir et al. (2011) and Behzadian et al. (2012) tackle multi-objective optimization problems aiming to minimize both chlorine injections and THM formation, a specific byproduct of chlorine reactions. The authors in Cozzolino et al. (2005) investigate the optimal number and allocation of booster stations and chlorine injections to control DBPs formation under feasible operating scenarios. Moreover, Maheshwari et al. (2020) have addressed the issue in a multi-objective genetic optimization in an integrated MATLAB-EPANET-MSX platform, utilizing WQ model simulating chlorine residuals and DBPs levels.

We note a prevalent limitation across most of the aforementioned studies: they lack explicit articulation or formulation of WQ models that accurately represent the input/output relationship between booster stations injections and the concentrations at critical network components. Such WQ models facilitate the integration of state-of-the-art control algorithms within a control-theoretic framework. Efforts have been made to address this limitation by developing input/output WQ models and incorporating them into control frameworks (Shang et al., 2002; Wang et al., 2005; Duzinkiewicz et al., 2005). Furthermore, Wang et al. (2021) introduce a novel approach to model WQ dynamics using a linear state-space representation, enabling the modeling of system inputs, outputs, and chemical concentration states as a closed-form control-theoretic model. This study has combined this representation with a model predictive control (MPC) algorithm to determine chlorine injections into the system. However, this formulation is based solely on the first-order single-species chlorine decay and transport dynamics. Additionally, this study neglects the dispersion process and employs an Eulerian discretization scheme that does not accurately represent advection-dominant processes. The study in Elsherif et al. (2024) has expanded upon this control framework to address scenarios where substances interact with chlorine by utilizing the control-oriented nonlinear multi-species reaction model built on validated Eulerian discretization schemes (Elsherif et al., 2023) and has applied model order reduction techniques to manage the associated high dimensionality. However, this study does not consider varying flow conditions in the system or their impact on chemical transport dynamics. Furthermore, the control approach in Elsherif et al. (2024) focuses on setting chlorine residuals within standardized limits, without considering the formation of DBPs or how network characteristics and flow conditions influence this formation and the control actions. Our paper aims to address these gaps, as detailed in the next section on our paper's contributions.

1.2. Paper contributions and organization

This paper's main objective is to develop a comprehensive strategic WQ control approach that takes into account the different flow conditions across various zones within the network and their influence on the chemicals' evolution, the complex dynamics of chlorine reaction and decay involving multiple species, and the formation of DBPs, which pose potential health risks. The development of this approach relies on a detailed analysis of system controllability under all these aspects, strategically guiding the control algorithm by weighting the input injections by booster stations to achieve enhanced controllability tailored to the unique characteristics of the network and hydraulics. Moreover, the approach seeks to mitigate the excessive formation of

DBPs, thereby ensuring the maintenance of WQ standards and protecting public health. The corresponding detailed paper's contributions are as follows.

- This paper extends and utilizes a multi-species reaction model that simulates chlorine decay and reaction with various substances, including the formation of DBPs. Moreover, this model accounts for the different hydraulic dynamics and flow conditions, capturing the advection- and dispersion-dominated states across different zones within the network. It dynamically switches between these states in certain zones while maintaining a single dominant mode throughout the simulation period in others. Formulated as a control-oriented state-space representation, this model integrates numerical discretization schemes to solve the ADR-PDE. The educational and theoretical significance of this work lies in its exploration of these schemes when coupled with multi-species dynamics and their impact on simulation variables such as time-step, discretization grid size, and numerical stability. This approach facilitates the integration of the model into control strategies while accommodating variations in process dynamics.
- The paper introduces a controllability analysis technique, providing valuable foresight for shaping the control strategy and optimizing the allocation of chlorine injections among booster stations. By leveraging insights from system controllability, the approach enhances the efficacy of control algorithms, enabling proactive decision-making and resource allocation to mitigate DBPs formation while ensuring WQ standards are met.
- Development of a control approach that integrates detailed modeling insights with real-world operational considerations. This approach offers a generalized solution applicable to diverse network configurations and hydraulic settings, accommodating predetermined booster station locations and evolving network dynamics. Through thorough validation on different networks and case studies, the proposed control approach demonstrates its effectiveness in addressing the dynamic challenges of WQ regulation.

The remainder of this paper is structured as follows: Section 2 presents the multi-species disinfectant and byproducts dynamics model, detailing the governing equations governing the evolution of chemicals across network components. Based on this model, the formulation of the disinfectant control problem is derived in Section 3. In Section 4, we validate our proposed approach through several case studies, encompassing various scales, layouts, and scenarios. Finally, Section 5 offers conclusions, discusses the limitations of the study, and provides recommendations for future research directions.

Notation. Italicized, boldface upper and lowercase characters represent matrices and column vectors: a is a scalar, \mathbf{a} is a vector, and \mathbf{A} is a matrix. The notation \mathbb{R}^n denotes the sets of column vectors with n real numbers, while $\mathbb{R}^{n \times m}$ denotes the sets of matrices with n rows and m columns. Variables with uppercase characters J , R , TK , P , M , and V correspond to junctions, reservoirs, tanks, pipes, pumps, and valves, respectively. Additionally, B_{TK} and B_J represent variables associated with booster stations located at tanks and junctions, while D_j denotes the demand variables at the junctions.

2. Multi-species disinfectant and byproducts dynamics model

We model the WDN by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$. The set \mathcal{N} defines the nodes and is partitioned as $\mathcal{N} = \mathcal{J} \cup \mathcal{T} \cup \mathcal{R}$ where sets \mathcal{J} , \mathcal{T} , and \mathcal{R} are collections of junctions, tanks, and reservoirs. Let $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ be the set of links, and define the partition $\mathcal{L} = \mathcal{P} \cup \mathcal{M} \cup \mathcal{V}$, where sets \mathcal{P} , \mathcal{M} , and \mathcal{V} represent the collection of pipes, pumps, and valves. The total number of states is $n_x = n_L + n_N$, where n_L and n_N are numbers of links and nodes. The number of reservoirs, junctions, tanks, pumps, valves, and pipes are $n_R, n_J, n_{TK}, n_M, n_V$, and n_P . Each pipe i with length L_i is spatially discretized and split into s_{L_i} segments. Hence, the number of

links states is expressed as $n_L = n_M + n_V + \sum_{i=1}^{n_p} s_{L_i}$ while $n_N = n_R + n_J + n_{TK}$ represents the number of nodes states.

We apply the principles of conservation of mass, transport, decay, and reaction to model the evolution of chemicals and substances throughout the different components of the WDNs. This modeling assumes that the hydraulic variables and settings are pre-determined and operate under non-transient conditions for extended period simulations. This means that the network operates on a steady-state basis for each time step, assuming that the system reaches equilibrium within each interval before moving to the next. In the following sections, we provide the governing equations for the transport of any chemical by denoting the concentration by c . Additionally, for dynamics that depend on specific chemical properties, such as chlorine decay and mutual reactions with other substances, we specify which chemical's concentration we are calculating by adding a special notation to the symbol c .

2.1. Transport and reaction in pipes: Advection-dispersion-reaction model

In our paper, we simulate the transport and reaction in pipes by the ADR-PDE, which for Pipe i is expressed as

$$\partial_t c_i^P = -v_i(t) \partial_x c_i^P + D_i(t) \partial_{xx} c_i^P + R_{MS}^P(c_i^P(x, t)), \quad (1)$$

where $c_i^P(x, t)$ is concentration in pipe at location x along its length and time t ; $v_i(t)$ is the mean flow velocity; $D_i(t)$ is the effective longitudinal dispersion coefficient of the chemical; and $R_{MS}^P(c_i^P(x, t))$ is the decay and reaction expression in pipes (more explanation is given in Section 2.3).

In Eq. (1), the term $v_i(t) \partial_x c_i^P$ describes the change in concentration as an impact of the advection process. Advection causes translation of the chemical by moving it with the flow velocity. Whilst, the term $D_i(t) \partial_{xx} c_i^P$ represents the impact of chemical dispersion, where the effective longitudinal dispersion coefficient $D_i(t)$ encompasses both molecular diffusion and shear dispersion caused by the unevenness of the velocity profile. The calculation of $D_i(t)$ for Pipe i is dependent on the flow condition, which can be categorized as laminar, transitional, or turbulent flow based on the unitless Reynolds number Re_i

$$Re_i(t) = \frac{\rho L_{D_i}}{\mu} v_i(t), \quad (2)$$

where ρ represents the water density, which is typically 62.4 lb/ft³ (998.4 kg/m³), μ denotes the water dynamic viscosity, typically 2.42 × 10⁻⁵ lb ft⁻¹ s⁻¹ (3.6 × 10⁻⁵ kg m⁻¹ s⁻¹) at room temperature, and L_{D_i} stands for the characteristic length, which is the hydraulic diameter for water flowing pressurized in a pipe, equal to the pipe diameter d_{p_i} .

For laminar flow conditions ($Re < 2300$), $D_i(t)$ is calculated as an averaged value over the solute residence time (Lee, 2004):

$$D_i(t) = \frac{d_{p_i}^2 v_i^2(t)}{12 d_m} \left[1 - \left[\frac{1 - \exp(-\frac{4d_m t_r(t)}{d_{p_i}^2})}{\frac{4d_m t_r(t)}{d_{p_i}^2}} \right] \right], \quad (3)$$

where t_r is the pipe residence time, $t_r(t) = \frac{L_i}{v_i(t)}$; and d_m is the molecular diffusion coefficient—Refs. Leaist (1986), Lee (2004) and Rossman et al. (2020) provide values for this coefficient for chlorine and different by-products.

For turbulent and transitional flow conditions ($Re \geq 2300$), $D_i(t)$ does not depend on the molecular diffusion coefficient, rather on the flow condition through the pipe (Basha and Malaeb, 2007):

$$D_i(t) = \frac{d_{p_i} v_{\tau_i}(t)}{2} \left[10.1 + 577 \left(\frac{Re_i(t)}{100} \right)^{-2.2} \right], \quad (4)$$

where v_{τ} is the shear velocity taken as a percentage of the mean flow velocity.

The influence of dispersion may be minimal in certain sections or across the entire network when turbulent conditions prevail. In such cases, the advection process is the dominant process in the chemical transportation throughout these network segments. However, it

is crucial to account for dispersion when the flow velocity is low and the flow approaches the laminar state (e.g., dead-end segments). The significance of dispersion can be assessed quantitatively using the Peclet number. For Pipe i , the Peclet number Pe_i at a given time t can be calculated using Eq. (5). The Peclet number serves as a dimensionless indicator of the relative importance between advection and dispersion, with a high value suggesting a flow scenario primarily governed by advection, where dispersion can be neglected. Certain simulation software establish a threshold value for the Peclet number Pe^{th} beyond which the dispersion effect is neglected (e.g., EPANET sets this threshold at 1000 (Rossman et al., 2020)).

$$Pe_i(t) = \frac{v_i(t) L_i}{D_i(t)}. \quad (5)$$

Following the explanation of all parameters and variables in Eq. (1), the subsequent step is to discuss how it is actually solved. Notably, there is no analytical solution for Eq. (1) in many scenarios with complex dynamics and different components models. That is, numerical discretization schemes are employed. These schemes discretize the pipe over a spatio-temporal grid, and accordingly, concentrations are calculated at each segment of this grid. A variety of schemes have been developed and utilized in the literature (Rossman and Boulos, 1996). The choice of the scheme depends on the actual physics of the system and the processes involved in the simulation of chemicals evolution. For instance, if the dispersion influence is neglected, a different method should be chosen compared to scenarios where it is taken into consideration. To that end, we detail herein numerical schemes to be employed for both scenarios (i.e., ADR-PDE and AR-PDE). Afterwards, we explain how to determine the grid size and simulation time-step, while ensuring numerical stability. That is, for the first two sections (2.1.1 and 2.1.2), we use the notation of Δt and Δx for the time-step and segment size for each pipe. Following, we introduce how to determine these parameters in Section 2.1.3.

2.1.1. Advection-dominant PDE discretization schemes

In advection-dominant transport, the two main processes are the advection where the concentration at a certain location and time is affected by upstream concentrations, and reaction where chemicals decay and/or mutually react. That being said, *Upwind* discretization schemes are more descriptive of the actual physical process considered among other schemes (Hirsch, 1990). In this paper, we consider both *Explicit* and *Implicit* Upwind schemes to investigate their performance in the proposed control framework. The distinction between *explicit* and *implicit* notations is clarified in Definition 1.

Definition 1. A discretization scheme is referred to as explicit or implicit depending on whether the concentrations of neighboring segments/nodes, on which the concentration of the segment being calculated depends, are taken from the previous time-step (and thus determined); Explicit, or from the current time-step (and thus solved for simultaneously); Implicit.

For Pipe i divided into n_{s_i} segments, the concentrations at $t + \Delta t$ of the first, last, and any in-between segment (i.e., $c_i^P(1, t + \Delta t)$, $c_i^P(s, t + \Delta t)$, $c_i^P(s_{L_i}, t + \Delta t)$) is calculated as listed in Table 1 for both schemes. In these calculations, $\tilde{\lambda}_i(t)$ is the Courant number (CN) and expressed as

$$\tilde{\lambda}_i(t) = \frac{v_i(t) \Delta t}{\Delta x_i}. \quad (6)$$

2.1.2. Dispersion-dominant PDE discretization schemes

By considering the dispersion process in modeling the transport of chemicals, both upstream and downstream concentrations influence the concentration at a specific location. Thus, we employ two central Eulerian discretization schemes to solve the ADR-PDE: the explicit Lax-Wendroff (L-W) Scheme and the implicit Backward Euler Scheme. Refer to Definition 1 for differentiation between explicit and implicit

Table 1

Expressions for four discretization schemes; two applied for the AR-PDE: **Explicit Upwind** (EU) and **Implicit Upwind** (IU), and the other two are for solving the ADR-PDE: **Lax-Wendroff** (L-W) and **Backward Euler** (BE). The header row of the table includes variables representing the concentrations at the upstream Junction j , first, intermediary, last segments of Pipe i , and the downstream Junction k at times t and $t + \Delta t$. For each scheme, three expression can be formulated from this table depending on the variables in the header row: these expressions are for the first, any intermediary, and last segments of the pipe. To formulate these expressions, treat each row independently. Multiply the values in each cell by the corresponding header variable of the same column. Then, sum these resultant terms on each side of the equal sign to complete the expression. In the column labeled $R_{MS}^P(\cdot)\Delta t$, the cells specify the variable upon which this function is calculated for. Note that, for simplicity and a more compact table, we eliminate writing the time factor for all the parameters $\alpha, \tilde{\lambda}, \lambda, \underline{\lambda}$, and $\bar{\lambda}$ of Eqs. (6), (7), and (8), knowing that all of them are taken at time (t) .

	$c_j^P(t + \Delta t)$	$c_i^P(1, t + \Delta t)$	$c_i^P(s, t + \Delta t)$	$c_i^P(s_L, t + \Delta t)$	$c_i^J(t + \Delta t)$	$c_j^J(t)$	$c_i^P(1, t)$	$c_i^P(s, t)$	$c_i^P(s_L, t)$	$c_k^J(t)$	$R_{MS}^P(\cdot)\Delta t$	Eq.
Explicit Upwind	1					=	$\tilde{\lambda}_i$	$1 - \tilde{\lambda}_i$			$(c_i^P(1, t))$	(9)
		1			1	=	$\tilde{\lambda}_i$	$1 - \tilde{\lambda}_i$			$(c_i^P(s, t))$	(10)
						=	$\tilde{\lambda}_i$	$1 - \tilde{\lambda}_i$			$(c_i^P(s_L, t))$	(11)
Implicit Upwind	$-\tilde{\lambda}_i$	$1 + \tilde{\lambda}_i$				=		1			$(c_i^P(1, t))$	(12)
		$-\tilde{\lambda}_i$	$1 + \tilde{\lambda}_i$			=			1		$(c_i^P(s, t))$	(13)
			$-\tilde{\lambda}_i$	$1 + \tilde{\lambda}_i$		=				1	$(c_i^P(s_L, t))$	(14)
Lax-Wendroff	1					=	$\underline{\lambda}_i + \alpha_i$	$\lambda_i - 2\alpha_i$	$\bar{\lambda}_i + \alpha_i$		$(c_i^P(1, t))$	(15)
		1			1	=	$\underline{\lambda}_i + \alpha_i$	$\lambda_i - 2\alpha_i$	$\bar{\lambda}_i + \alpha_i$		$(c_i^P(s, t))$	(16)
						=	$\underline{\lambda}_i + \alpha_i$	$\lambda_i - 2\alpha_i$	$\bar{\lambda}_i + \alpha_i$		$(c_i^P(s_L, t))$	(17)
Backward Euler	$-0.5\tilde{\lambda}_i - \alpha_i$	$1 + 2\alpha_i$	$0.5\tilde{\lambda}_i - \alpha_i$			=		1			$(c_i^P(1, t))$	(18)
		$-0.5\tilde{\lambda}_i - \alpha_i$	$1 + 2\alpha_i$	$0.5\tilde{\lambda}_i - \alpha_i$		=			1		$(c_i^P(s, t))$	(19)
			$-0.5\tilde{\lambda}_i - \alpha_i$	$1 + 2\alpha_i$	$0.5\tilde{\lambda}_i - \alpha_i$	=				1	$(c_i^P(s_L, t))$	(20)

schemes. Similar to the previous schemes, Table 1 provides the formulations for the concentrations at $t + \Delta t$ of the first, last, and any in-between segment (i.e., $c_i^P(1, t + \Delta t)$, $c_i^P(s, t + \Delta t)$, $c_i^P(s_L, t + \Delta t)$) of Pipe i by utilizing these schemes. The coefficients used in these formulations are the dispersion number expressed as

$$\alpha_i(t) = D_i(t) \frac{\Delta t}{\Delta x_i^2}, \quad (7)$$

while $\underline{\lambda}_i(t)$, $\lambda_i(t)$, and $\bar{\lambda}_i(t)$ are the weighting coefficients calculated as follows

$$\begin{aligned} \underline{\lambda}_i(t) &= 0.5\tilde{\lambda}_i(t)(1 + \tilde{\lambda}_i(t)), \\ \lambda_i(t) &= 1 - \tilde{\lambda}_i^2(t), \\ \bar{\lambda}_i(t) &= -0.5\tilde{\lambda}_i(t)(1 - \tilde{\lambda}_i(t)). \end{aligned} \quad (8)$$

2.1.3. Pre-simulation preparations: Time-step, grid size, and numerical stability assurance

Applying explicit discretization methods requires satisfying numerical stability conditions. For the Explicit Upwind Scheme applied on AR-PDE, the Courant–Friedrichs–Lowy condition states that the CN number is to be maintained in the range of $0 < \tilde{\lambda}_i(t) \leq 1$ for Pipe i . Additionally, to solve the ADR-PDE using L-W scheme for Pipe i and ensure its numerical stability, the von Neumann condition states that $0 < \tilde{\lambda}_i^2(t) \leq 2\alpha_i(t) \leq 1$. In addition, for any central scheme applied for ADR-PDE, the ratio $\tilde{\lambda}_i(t)/\alpha_i(t)$ is set to be less than or equal to 2 to achieve high accuracy and avoid spatial oscillations (Durran, 2010).

That being said, different conditions need to be satisfied, according to which the time-step and grid size are determined. In Procedure 1, we list the steps needed to compute these parameters while following a numerically stable approach.

2.1.4. Dynamic modeling of the transport and reaction dynamic in pipes

As explained in the previous sections, our approach facilitates dynamic switching between discretization methods based on the prevailing processes in transport and reaction dynamics. To begin with, Procedure 1 ensures consistency in system dimensions (specifically, the WQ time-step and number of segments for each pipe) while maintaining numerical stability throughout the simulation period, regardless of whether dispersion effects are considered. To that end, the switch between the discretization methods is solely about the considered elements and the calculations of the dependency between the segments and nodes from one time-step to the next (see Table 1). With each hydraulic time-step for updating system matrices, an additional condition is introduced to assess the current Peclet number (Eq. (5)) against the threshold P_{th} , determining whether dispersion effects should be neglected. Accordingly, parameters and elements outlined in Table 1 are computed.

Procedure 1: Time-step and grid size determination and numerical stability assurance

```

Input: WDN topology, components' characteristics, hydraulics parameters, and  $P_{th}$ 
Output: Time-step  $\Delta t$  and number of segments and segment size for each pipe;  $s_{L_i}$  and  $\Delta x_i$ ,  $i = 1, \dots, n_p$ 
1 Initialize  $\Delta t_{temp} \leftarrow 0$ 
2 Initialize  $\Delta t \leftarrow 10^6$  // Start with a large number
3 for  $i = 1$  to  $n_p$  do
4   Initialize  $D_i \leftarrow \emptyset$  and  $P_{e_i} \leftarrow \emptyset$ 
5   for each hydraulic time-step do
6     Calculate  $R_{e_i}(t)$  via (2)
7     if  $R_{e_i}(t) < 2300$  then
8       | Calculate  $D_i(t)$  via (3)
9     else
10      | Calculate  $D_i(t)$  via (4)
11    Calculate  $P_{e_i}(t)$  via (5)
12    Append  $D_i(t) \rightarrow D_i$  and  $P_{e_i}(t) \rightarrow P_{e_i}$ 
13    if  $P_{e_i}(t) \leq P_{th}$  then
14      |  $\Delta t_{temp} = 2D_i(t)/v_i^2(t)$  // To avoid spatial oscillations
15      | if  $\Delta t_{temp} < \Delta t$  then
16        | |  $\Delta t \leftarrow \Delta t_{temp}$ 
17      end if
18    end if
19  return  $\Delta t$  // Final time-step
20 if Applying Explicit Discretization Methods then
21   for  $i = 1$  to  $n_p$  do
22     Initialize  $\Delta x_{temp_1} \leftarrow 0$  and  $\Delta x_{temp_2} \leftarrow 0$ 
23     Initialize  $\Delta x \leftarrow 0$ 
24     for each hydraulic time-step do
25       if  $P_{e_i}(t) \leq P_{th}$  then
26         | |  $\Delta x_{temp_1} = (2D_i(t)\Delta t)^{-\frac{1}{2}}$  // von Neumann condition
27       else
28         | |  $\Delta x_{temp_2} = v_i(t)\Delta t$  // Courant-Friedrichs-Lowy condition
29       |  $\Delta x_i \leftarrow \max\{\Delta x_{temp_1}, \Delta x_{temp_2}, \Delta x_i\}$ 
30      $s_{L_i} = \lfloor L_i/\Delta x_i \rfloor$ 
31      $\Delta x_i = L_i/s_{L_i}$ 
32 else if Applying Implicit Discretization Methods then
33   Set  $s_{L_i}$  as fixed arbitrary integer for all  $i = 1, \dots, n_p$ 
34    $\Delta x_i = L_i/s_{L_i}$ ,  $\forall i = 1, \dots, n_p$  // Final number of segments and segment size for each pipe

```

2.2. Conservation of mass

For nodes and links other than pipes, the conservation of mass principle is applied to formulate the governing equations for concentration calculations within water distribution networks.

2.2.1. Mass balance at reservoirs

For any Reservoir i , concentration of a chemical is assumed to remain constant over time, such that $c_i^R(t + \Delta t) = c_i^R(t)$.

2.2.2. Mass balance at pumps and valves

In our WQ model, we deal with pumps and valves as transmission links with negligible length. Accordingly, chemicals concentrations at these elements are taken equal to the concentration of the upstream node. That being said, for Pump i or Valve j installed after Reservoir/Tank/Junction k , concentrations are expressed as $c_i^M(t + \Delta t) = c_k^M(t + \Delta t)$, and $c_j^V(t + \Delta t) = c_k^V(t + \Delta t)$.

2.2.3. Mass balance at junctions

Water from all inflows into a junction is assumed to have complete and instantaneous mixing. That is, while assuming that there is no storage time at junctions, at a Junction i , all outflows have the same concentration for a specific chemical. This concentration is expressed as

$$c_i^J(t) = \frac{\sum_{j \in L_{in}} q_{in}^j(t) c_{in}^j(t) + q_i^{B_j}(t) c_i^{B_j}(t)}{q_i^{D_j}(t) + \sum_{k \in L_{out}} q_{out}^k(t)}, \quad (21)$$

where j and k represent the counters for the elements of the set L_{in} of links flowing into the junction and a set L_{out} of links withdrawing flow from the junction, respectively; $q_{in}^j(t)$ and $q_{out}^k(t)$ are the corresponding inflows and outflows from these links; $c_{in}^j(t)$ is the chemical concentration in each of the inflows; $q_i^{B_j}(t)$ is the flow injected into the junction with chemical concentration $c_i^{B_j}(t)$ by a booster station, if located; and $q_i^{D_j}(t)$ represents the consumer's demand.

2.2.4. Mass balance at tanks

In our model, we assume complete and instantaneous mixing of all inflows, outflows, and stored water in a tank, following the continuously stirred tank reactor (CSTR) model. Consequently,

$$\begin{aligned} V_i^{TK}(t + \Delta t) c_i^{TK}(t + \Delta t) &= V_i^{TK}(t) c_i^{TK}(t) + \sum_{j \in L_{in}} q_{in}^j(t) c_{in}^j(t) \Delta t \\ &+ V_i^{B_{TK}}(t + \Delta t) c_i^{B_{TK}}(t + \Delta t) \\ &- \sum_{k \in L_{out}} q_{out}^k(t) c_i^{TK}(t) \Delta t + R_{MS}^{TK}(c_i^{TK}(t)) V_i^{TK}(t) \Delta t, \end{aligned} \quad (22)$$

where $V_i^{TK}(t)$ is the tank volume and $V_i^{B_{TK}}(t + \Delta t)$ is the chemical solute volume injected to the tank by a booster station, if located, with a concentration $c_i^{B_{TK}}(t + \Delta t)$. Note that the effect of booster station injections is considered immediate according to the CSTR model. Therefore, their volume and concentration are accounted for at $t + \Delta t$ to calculate the tank volume concentration at the same time instant. $R_{MS}^{TK}(c_i^{TK}(t))$ is the multi-species dynamics in tanks expression (refer to Section 2.3).

2.3. Chlorine multi-species reaction and decay dynamics and DBPs formation models

In this paper, we utilize a first-order decay dynamics model for chlorine. Additionally, we model the mutual reaction between chlorine and a fictitious reactant using a second-order formulation, which also considers the formation of DBPs. Specifically, our focus is on the formation of one of the most common types of DBPs: trihalomethanes (THMs) (DBP, 2022). The chlorine decay reaction rates for Pipe i and Tank j are $k_i^P = k_b + \frac{2k_w k_f}{r_{Pi}(k_w + k_f)}$, $k_j^{TK} = k_b$, where k_b is the bulk reaction

Table 2

Chlorine multi-species reaction and decay dynamics and DBPs formation models. In these models, concentrations for chlorine, fictitious reactant, and THMs are denoted by c , \tilde{c} , and \hat{c} , respectively.

Segment s of pipe i	Tank j	
$R_{MS}(c(t))$	$-k_i^P c_i^P(s, t) - k_r c_i^P(s, t) \tilde{c}_i^P(s, t)$	$-k_j^{TK} c_j^{TK}(t) - k_r c_j^{TK}(t) \tilde{c}_j^{TK}(t)$
$R_{MS}(\tilde{c}(t))$	$-Y_{FR} k_r c_i^P(s, t) \tilde{c}_i^P(s, t)$	$-Y_{FR} k_r c_j^{TK}(t) \tilde{c}_j^{TK}(t)$
$R_{MS}(\hat{c}(t))$	$Y_{THMs} k_r c_i^P(s, t) \hat{c}_i^P(s, t)$	$Y_{THMs} k_r c_j^{TK}(t) \hat{c}_j^{TK}(t)$

rate constant; k_w is the wall reaction rate constant; k_f is the mass transfer coefficient between the bulk flow and the pipe wall; r_{Pi} is the pipe radius. Both decay rates are in $1/s$.

By donating the chlorine concentration to be c , \tilde{c} for fictitious reactant concentrations, and \hat{c} for THMs, the multi-species dynamics models (Moeini et al., 2023) in pipes and tanks are expressed as in Table 2. In this table, k_r ($L \text{ mg}^{-1} \text{ s}^{-1}$) denotes the mutual reaction coefficient, while Y_{FR} and Y_{THMs} represent the unitless ratios between the stoichiometric coefficients of the fictitious reactant and THMs, respectively, to that of chlorine.

2.4. Chlorine and byproducts multi-species dynamics in a form of state-space representation

The state-space representation of the WQ multi-species dynamics of chlorine and byproducts (WQMS-CLBP) is expressed in Eq. (23) as nonlinear difference equations (NLDE).

WQMS-CLBP	NLDE
$E(t)x(t + \Delta t) = A(t)x(t) + B(t)u(t) + f(x(t)),$	(23a)
$y(t) = C(t)x(t),$	(23b)

where $x(t)$ represents the state vector, which concatenates the concentrations of chlorine, fictitious reactant, and the byproducts at the various components within the network. The control input vector, $u(t)$, encompasses chlorine injections and can also accommodate unplanned and planned injections of contamination associated with the fictitious reactant. The nonlinear vector, $f(x(t))$, encapsulates the expressions for mutual reactions and byproduct formation. Vector $y(t)$ contains sensor measurements of the chemicals. Matrices $E(t)$, $A(t)$, $B(t)$, and $C(t)$ are time-varying and depend on factors such as network topology, hydraulic parameters, decay rates, coefficients for mutual reactions between chemicals, and the locations of booster stations and sensors.

Note that, matrix $E(t)$ equals identity under the condition of applying explicit discretization schemes. However, for the implementation of implicit discretization schemes, its construction depends on the system's hydraulics (refer to Table 1). The system's hydraulics are updated every hydraulic time-step, which is typically longer than the WQ one. Therefore, it is customary to denote the time instant when this matrix is taken as t .

As explained, the one source of nonlinearity in our model is the mutual reaction dynamics between the chemicals. To overcome the complexity associated with this nonlinearity, we employ a linearization technique, specifically utilizing Taylor series approximation as detailed in Elsherif et al. (2024). Herein, we showcase the linearization process for the mutual reaction expression, which applies uniformly across all such expressions listed in Table 2. The linearization is performed around operating points represented by c_o for chlorine and \tilde{c}_o for the fictitious reactant.

$$R_{MS}(\tilde{c}(t)) = -Y_{FR} k_r c(t) \tilde{c}(t) \Rightarrow -Y_{FR} k_r (c_o \tilde{c}(t) + \tilde{c}_o c(t) - c_o \tilde{c}_o). \quad (24)$$

For both chlorine and the fictitious reactant, the linearized mutual reaction breakdown consists of terms dependent on their respective

concentrations, terms dependent on the concentrations of the other chemical, and constant terms. While THMs concentrations depend on the other two chemicals concentrations and a constant term. Consequently, while the general state-space representation in Eq. (23) features a block-diagonal matrix of A matrices with no interdependency between the chemicals apart from within the f function, the application of linearization alters the state-space representation to linear difference equations (LDEs) with interdependencies among these chemical compounds. This transformation can be expressed as:

WQMS-CLBP	LDE
$E(t)x(t + \Delta t) = \tilde{A}(t)x(t) + B(t)u(t) + \Phi$,	(25a)
$y(t) = C(t)x(t)$,	(25b)

where $\tilde{A}(t)$ is the updated matrix to account for the interdependencies among the chemicals after linearization, while Φ is the vector gathering the constant terms.

The operating points around which the system is linearized are dynamically updated throughout the simulation period, typically every other time-step. The updating process is guided by the frequency with which the system dynamics evolve. In practice, the operating points are updated on a time scale wider than the WQ time-step. This approach ensures that the linearization captures the evolving behavior of the system adequately, allowing for accurate modeling of the system dynamics over time.

3. Disinfectant control problem formulation

The primary objective of this study is to regulate disinfectant levels within WDNs to comply with standard thresholds, while simultaneously addressing the formation of DBPs. Our proposed approach involves the implementation of a MPC algorithm, which is constructed based on the WQ multi-species model outlined in (25), and constrained by specified chlorine limits and DBPs cutoffs. As a preliminary step, we conduct a controllability analysis of the system to identify critical states (e.g., dead-ends) and determine control inputs from booster stations. Utilizing controllability metrics obtained from this analysis, we assign higher weights to control inputs that exhibit greater effectiveness in influencing critical states or achieving desired objectives.

First, we explain the formulation of the disinfectant control problem. The control problem is formulated over the simulation period $[0, T_s]$, with the objective of minimizing the cost of chlorine injections. The problem is constrained by maintaining the chlorine concentrations within the standard levels of 0.2 mg/L and 4 mg/L, and THMs levels lower than 0.08 mg/L. Additionally, the problem can be utilized to constrain the fictitious reactant concentrations to a specific level within the simulation period. The control inputs for chlorine are constrained to be non-negative and limited by the availability of chlorine and the capacity of booster stations. All these constraints must be satisfied while complying with the actual governing equations of the WQ dynamics. Combining all this information, the water quality multi-species control problem (WQMS-CP) is formulated as shown in (26).

WQMS-CP	
minimize $x(t), \hat{u}(t)$	$\mathcal{J}(\hat{u}(t)) = \epsilon \sum_{t=1}^{N_s} q^B(t)^\top \hat{u}(t)$
subject to	WQMS-CLBP (25), $x_{\min} \leq x(t) \leq x_{\max}$, $\hat{u}_{\min} \leq \hat{u}(t) \leq \hat{u}_{\max}$,

where problem variables $x(t)$ and $\hat{u}(t)$ are chemicals concentrations network-wide and chlorine injections through booster stations, $q^B(t)$ is the flow rates at the nodes corresponding to the locations of the booster stations, ϵ is the unit cost of chlorine in \$/mg, and WQMS-CLBP is the WQ model we are simulating and controlling following the representation in (25). Finally, N_s is the number of time-step in the simulation period, $N_s = \frac{T_s}{\Delta t}$.

The control problem described in (26) is formulated as a linear program (LP) due to its linear objective function and constraints. The subsequent step involves reformulating this problem into a quadratic program (QP) with a quadratic objective function in a MPC framework. This reformulation aims to minimize chlorine injections while ensuring the smoothness of control inputs while ensuring falling within the states boundaries. The study (Wang et al., 2021) has provided a comprehensive derivation of this problem for single-species WQ dynamics based on AR models. Despite focusing on a different system, the derivation is applicable to our study since both systems are linear, and the form of the objective functions and constraints remains consistent. The one distinction lies in the constant terms concatenated in Φ of (25a). However, these terms solely depend on the operating point around which the system is linearized. In accordance with the approach outlined in Wang et al. (2021), the full derivation and formulation of the control problem are provided therein. For brevity, we direct readers to this study for a detailed understanding of the derivation and the final formulation of the control problem.

In this final formulation of the control problem as presented in Wang et al. (2021, Eq. (34)), two weight matrices are introduced: $Q = Q^\top$ and $R = R^\top$. These matrices specify the relative importance of measurement deviations and the smoothness of control inputs, respectively. In addition, we are building these matrices to reflect on prioritizing each of control input according to the controllability analysis we perform beforehand. Detailed information on the controllability analysis and the construction of these matrices is provided in the next section.

3.1. Preliminary controllability analysis

In this section, we present the procedure for conducting the prior controllability analysis and deriving the corresponding weight matrices of the MPC algorithm. To do so, we start by introducing the notion of controllability for dynamic systems, focusing specifically on the WQ dynamic system in our study. We then introduce metrics aimed at quantifying the influence of each booster station within the control framework. We note that the controllability notion introduced in this section is based on the linearized form of the WQMS-CLBP model as a simplified approximation, yet, enough to provide the needed insights for the analysis.

From a control-theoretic perspective, controllability refers to the capability of guiding a system from its initial states $x_o := x_{t(0)}$ to a desired state $x_p := x_{T_p}$ by some input $u(t)$ over a specific time window of T_p (Kalman, 1963). In the context of WQ control, we want to measure the ability to regulate chlorine concentrations by adjusting chlorine injections from booster stations, ensuring chlorine levels remain within predefined bounds.

The dynamic linear system (25), where $\tilde{A} \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times n_u}$, is said to be controllable if and only if the controllability matrix for $N_p = \frac{T_p}{\Delta t}$ time-steps given as

$$C_{N_p} := \{B, \tilde{A}B, \tilde{A}^2B, \dots, \tilde{A}^{N_p-1}B\} \in \mathbb{R}^{n_x \times N_p n_u}, \quad (27)$$

is full row rank, i.e., $\text{rank}(C_{N_p}) = n_x$, without loss of generality as we assume that $N_p n_u > n_x$. This is known as Kalman's rank condition (Kalman, 1963).

For our analysis objective, simply assessing the rank property is not sufficient to evaluate the impact of each booster station's injections within the control framework. The concept of control energy is important as well. In the WQ control context, control energy relates to the

amount of chlorine injections to reach the desired chlorine levels at the networks' components. Metrics concerning the control energy are derived from the controllability Gramian $\mathbf{W}_c(\tilde{\mathbf{A}}, \mathbf{B}, N_p) := \mathbf{W}_c \in \mathbb{R}^{n_x \times n_x}$, defined for N_p pairs of matrices $\tilde{\mathbf{A}}$ and \mathbf{B} as

$$\mathbf{W}_c := \sum_{\tau=0}^{N_p-1} \tilde{\mathbf{A}}^\tau \mathbf{B} \mathbf{B}^\top (\tilde{\mathbf{A}}^\top)^\tau = C_{N_p} C_{N_p}^\top, \quad (28)$$

where the controllability Gramian \mathbf{W}_c , that is a positive semidefinite matrix. In this study, we employ the trace(\mathbf{W}_c) metric, which inversely correlates with the average controllability energy across all state-space directions. In the context of WQ control, a higher controllability energy indicated a greater potential for chlorine injections to impact various system states within the specified time interval. In addition, rather than utilizing the rank metric for the controllability matrix in higher dimensions, we apply it for the controllability Gramian—a symmetric matrix with lower dimensions.

However, there are two aspects that demand attention during our controllability analysis: (i) not all the states of the system dynamics (25) are critical or controllable by booster stations, and (ii) calculating the trace for an uncontrollable subspace can be misleading due to the averaging of energy across uncontrollable directions. The first is addressed by adopting the concept of *target controllability* (Gokhale et al., 2021). Target controllability enables us to specify the desired target nodes, thereby mitigating the challenge of high dimensionality associated with the WQ representation. In this scenario, the metrics are applied to the targeted controllability Gramian $\mathbf{W}_T = \mathbf{C}_T \mathbf{W}_c \mathbf{C}_T^\top$, where the output matrix \mathbf{C}_T identifies the set of target nodes \mathcal{T} of size n_T . The notions of controllability and control energy apply to the subspace of the target nodes. Regarding the second aspect, to quantify the control energy for the controllable subspace within the space under consideration – where the rank of the controllability Gramian is $k < n$, with n denoting the number of states in the original space – a decomposition approach is utilized (Datta, 2004). This approach begins by defining a nonsingular matrix $T \in \mathbb{R}^{n \times n}$ such that

$$\tilde{\mathbf{A}} = T \tilde{\mathbf{A}} T^{-1} = \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \mathbf{0} & \tilde{\mathbf{A}}_{22} \end{bmatrix}, \quad \tilde{\mathbf{B}} = T \mathbf{B} = \begin{bmatrix} \tilde{\mathbf{B}}_1 \\ \mathbf{0} \end{bmatrix}, \quad (29)$$

where $\tilde{\mathbf{A}}_{11}$, $\tilde{\mathbf{A}}_{12}$ and $\tilde{\mathbf{A}}_{22}$ have dimensions of $k \times k$, $k \times (n - k)$ and $(n - k) \times (n - k)$, and $\tilde{\mathbf{B}}_1$ has k rows. Matrices $\tilde{\mathbf{A}}_{11}$ and $\tilde{\mathbf{B}}_1$ define a controllable subspace.

The key question to address now is: *over which time interval should the WQ controllability metrics be evaluated, and correspondingly, how often should the MPC weight matrices be updated?* To answer this question, we emphasize that this time interval should coincide with the frequency at which booster station injections' effectiveness varies. This effectiveness is dependent on the changes in system dynamics, particularly system's hydraulics. Hence, the time interval is taken to align with the hydraulic time-step, during which the WQ time-step is updated due to their different scales.

After establishing the controllability metrics to be utilized, herein we outline the analysis conducted to determine the weights for the MPC matrices. Initially, for each hydraulic time-step, the matrix \mathbf{B} is constructed for each booster station individually as the sole station in the system, followed by the computation of the corresponding controllability matrix and Gramian. Subsequently, for each critical set of target nodes \mathcal{T} and each booster station, the rank and trace metrics of the Gramians are calculated, leveraging the proposed methodologies addressing target controllability and uncontrollable subspace decomposition. Next, based on the importance index assigned to each critical set, each booster station is allocated a score relative to others according to its WQ controllability metrics. These scores are then aggregated for each booster station and weighted together relatively. Thus, the weights to construct the matrix $\mathbf{R}(t)$ are computed for each hydraulic time-step.

On the other hand, matrix \mathbf{Q} is constructed relative to matrix \mathbf{R} . Essentially, if prioritizing the smoothness of control inputs and the

Procedure 2: Offline prior controllability analysis and MPC weight matrices determination

Input: WDN topology, components' characteristics, hydraulics parameters, the linearized WQMS-CLBP system matrix $\tilde{\mathbf{A}}$, and WQ initial conditions

Output: MPC weight matrices $\mathbf{R}(t)$ and $\mathbf{Q}(t)$ at each hydraulic time-step

1 Initialization

2 Define critical n_T sets of target nodes $\mathcal{T}_i, i = 1, \dots, n_T$

3 Assign index $\eta_i, i = 1, \dots, n_T$ to each target set based on their criticality/importance

4 Obtain Δt , s_{L_i} and $\Delta x_i, \forall i = 1, \dots, n_p$ via following Procedure 1

5 for each hydraulic time-step do

6 **for** $j = 1$ **to** n_u **do**

7 Initialize $w_j \leftarrow 0$

8 Construct \mathbf{B}_j assuming only booster station j is allocated // n_u is booster stations count

9 Construct \mathbf{C}_{N_p} via (27)

10 Construct \mathbf{W}_c via (28)

11 **for** $k = 1$ **to** n_T **do**

12 Calculate $\mathbf{W}_{\mathcal{T}_k} = \mathbf{C}_{\mathcal{T}_k} \mathbf{W}_c \mathbf{C}_{\mathcal{T}_k}^\top$

13 Calculate $\text{rank}(\mathbf{W}_{\mathcal{T}_k})_j$

14 **if** $\text{rank}(\mathbf{W}_{\mathcal{T}_k})_j = n_k$ **then**

15 Calculate $T_r = \text{trace}(\mathbf{W}_{\mathcal{T}_k})_j$

16 Update $w_j \leftarrow w_j + \eta_k \cdot n_k \cdot T_r$

17 **else**

18 $r = \text{rank}(\mathbf{W}_{\mathcal{T}_k})_j$

19 Apply subspace decomposition as in (29)

20 Update $\tilde{\mathbf{C}}_{\mathcal{T}_k}$ to the first r of $T \mathbf{C}_{\mathcal{T}_k} T^{-1}$

21 $\tilde{\mathbf{W}}_{\mathcal{T}_k} \leftarrow \tilde{\mathbf{C}}_{\mathcal{T}_k} \mathbf{W}_c \tilde{\mathbf{C}}_{\mathcal{T}_k}^\top$

22 Calculate $T_r = \text{trace}(\tilde{\mathbf{W}}_{\mathcal{T}_k})$

23 Update $w_j \leftarrow w_j + \eta_k \cdot r \cdot T_r$

24 Calculate diagonal elements $a_{jj}, j = 1, \dots, n_u$ of $\mathbf{R}(t)$,
 $a_{jj} = w_j / \sum_{j=1}^{n_u} w_j$

25 Assign diagonal weights to $\mathbf{Q}(t)$ // According to critical states and importance in comparison to $\mathbf{R}(t)$

distinction between booster station injections is of utmost importance, the weights in \mathbf{R} are proportionally higher. Conversely, if achieving desired chemical levels promptly is more important, the weights in \mathbf{Q} are elevated. Such decisions are guided by the deviation of current concentrations from the desired setpoints and the how effective the booster stations injections are to cover the network. Typically, tuning these matrices is performed individually for each network to achieve suitable settings.

All these steps are summarized in Procedure 2. Note that, this approach is flexible, allowing for consideration of either one metric or both in the assessment and weighting procedure, albeit with different weights. Additionally, the critical sets of target nodes are categorized to be one of the following: dead-ends, zones with low initial chlorine conditions compared to the rest of the network, zones with high contaminant concentrations, or zones with pre-existing elevated levels of byproducts. For the latter scenario, the assigned index to this set is the lowest.

4. Case studies

In this section, we validate the proposed chlorine control approach through numerical case studies. These studies include several networks characterized by different scales, layouts, hydraulic settings and scenarios, and initial WQ conditions. The hydraulic settings are obtained by running different scenarios using the EPANET toolkit on MATLAB (Rossman et al., 2020). We apply our approach on four benchmark

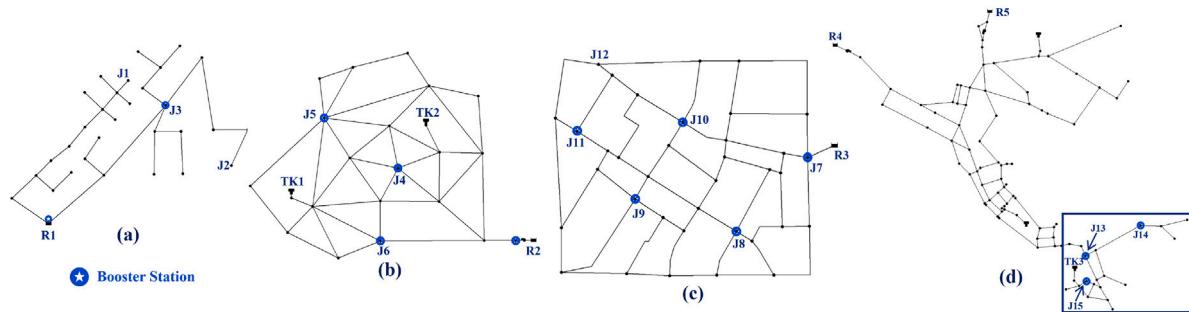


Fig. 1. Networks under study and their layouts: (a) BLA-M, (b) Anytown, (c) FOS, and (d) Net3 networks.

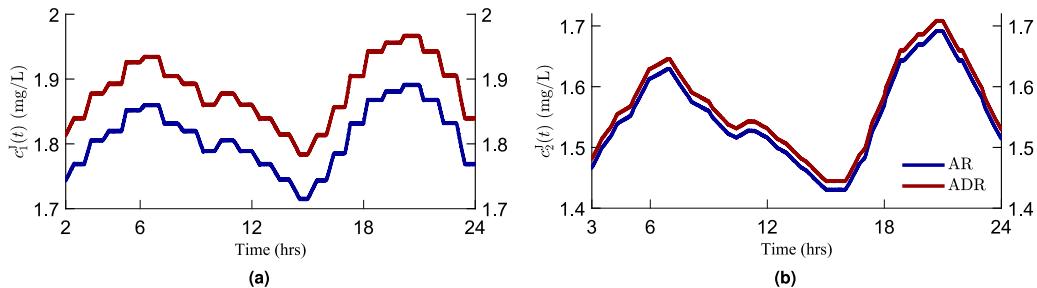


Fig. 2. Chlorine concentrations at Junctions (a) J1 and (b) J2 of BLA-M network without (AR) and with (ADR) the consideration of the dispersion process effect.

Table 3
Components count for each of the test networks.

Network	Junctions	Reservoirs	Tanks	Pipes	Pumps
BLA-M	30	1	0	30	0
Anytown	22	1	2	43	3
FOS	36	1	0	58	0
Net3	90	2	3	114	2

networks, each has a different number of components and layouts, including looped networks and those with a relatively higher number of dead-ends. These networks are: a modified version of the Blacksburg network (BLA-M), Anytown, Fossolo network (FOS), and Net3 (Wang et al., 2015; Rossman et al., 2020). Fig. 1 illustrates these networks layouts while Table 3 lists the count for each component.

First, we demonstrate the effect of considering dispersion in the simulation of the chemicals transport and evolution. We simulate the three chemical compounds with and without the inclusion of dispersion on the BLA-M and Anytown networks for comparison. For both networks, the simulation period is 24 h with a hydraulic time-step of 1 h. In addition, fixed sources of chemicals are maintained at Reservoirs R1 and R2, with concentrations of 2 mg/L for chlorine, 0.3 mg/L for the fictitious reactant, and 0.01 mg/L for the THMs. For the BLA-M network, Fig. 2 shows the results for chlorine concentrations at Junctions J1 and J2 with and without dispersion. As shown in the figure, the inclusion of dispersion has a greater impact on Junction J1 than J2. Neglecting dispersion leads to an underestimation of chlorine concentrations, particularly at J1, with an underestimation of around 8%, while for J2, it is 3%. This is due to the lower velocity in the pipe leading to Junction J1 compared to J2, resulting in a more dispersion-dominant process.

For the Anytown network, Fig. 3 illustrates chlorine and THMs concentrations at Tank TK1 with and without dispersion. The reason for showing these results is to highlight that although the underestimation of chlorine concentrations at Tank TK1 may seem negligible, the difference in THMs concentrations can be substantial, potentially leading to overlooking it reaching its upper bound, thus compromising water safety. This situation arises from how the mutual reactions are handled in our model. These reactions are expressed as second-order

nonlinear formulations. Consequently, the dispersion effect on chlorine and the fictitious reactant implicitly affects THMs formation while explicitly impacting its evolution in the transport expression. This effect becomes more noticeable with higher concentrations of chlorine and the fictitious reactant, as well as at slower velocities.

It is worth mentioning that the results in Fig. 2 are obtained by applying the explicit discretization schemes listed in Table 1, while the results in Fig. 3 are obtained using the implicit discretization schemes. Additionally, Procedure 1 is employed to determine the WQ time-step and system dimensions for both networks under different scenarios to test its applicability and numerical stability. Although the discretization procedure has proven its applicability, in some scenarios, the resulting WQ time-step is restrictively small to satisfy the condition listed in step 14 of Procedure 1. Subsequently, applying explicit discretization schemes leads to a high number of segments required for each pipe, resulting in high system dimensionality. This high system dimensionality demands significant computational time for the simulation of chemical evolution, considering that we simulate for three chemicals. Thus, for each network component and pipe segment, we have three states. Conversely, using an implicit discretization scheme for either the AR-PDE or ADR-PDE avoids this issue but requires performing matrix inversion. This problem can be mitigated by constructing these matrices as sparse matrices and utilizing fast inversion commands according to the coding language used. To that end, the results obtained in the remainder of this section are based on models developed using implicit discretization schemes.

Next, we present results from quantifying the controllability of each booster station located on the FOS network individually, aimed at steering chlorine concentration to a critical target node: Junction J12. This junction could represent a high-demand area or an area with initially low chlorine concentrations. We calculate the rank metric for each booster station and present the results in Fig. 4 for two hydraulic settings of the system over a 24-h simulation period with a hydraulic time-step of 1 h. The WQ time-step is updated within the hydraulic time-step on a scale of 1 min. These metrics are measured on a system linearized around operating points of 0.5 mg/L for chlorine, 0.1 mg/L for the fictitious reactant, and 0.01 mg/L for THMs concentrations—the initial WQ concentrations of the system.

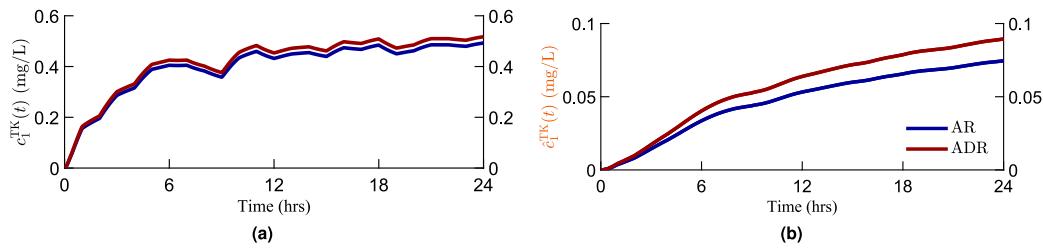


Fig. 3. (a) Chlorine and (b) THMs concentrations at Tank TK1 of the Anytown network without (AR) and with (ADR) the consideration of the dispersion process effect.

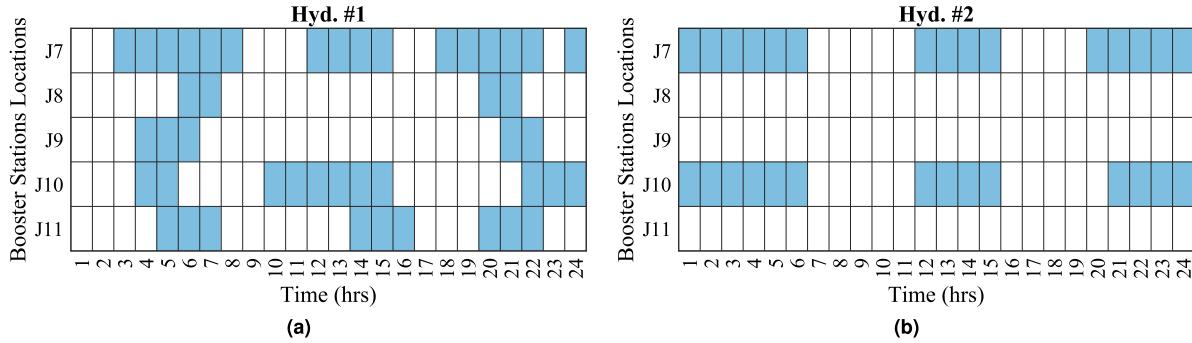


Fig. 4. Controllability of chlorine injections by each booster stations allocated over the FOS network if worked solely to steer the concentrations at the target junction J12, for two hydraulic settings: (a) Hyd. #1 and (b) Hyd. #2. Colored tiles indicate full rank metric. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

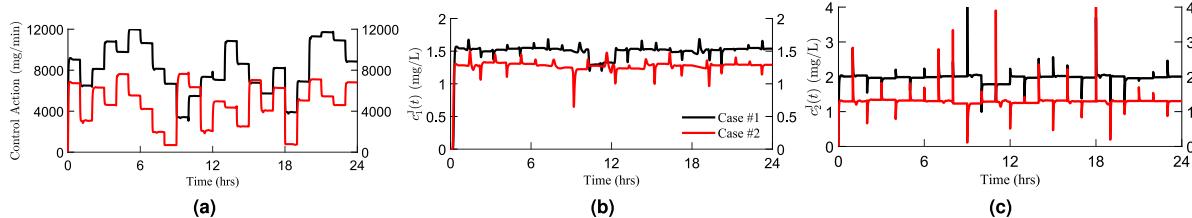


Fig. 5. (a) Summation of MPC control actions at Reservoir R1 and Junction J3 of the BLA-M network, and the corresponding chlorine concentrations at (b) Junction J1 and (c) Junction J2 for two cases: Case #1, applying MPC without prior controllability analysis and without DBPs constraints.

In Fig. 4, achieving full rank is highlighted by the colored blocks within the specific hydraulic 1-h time-step, demonstrating the achieved target controllability for Junction J12. As shown, Junction J12 is not always controllable by booster stations due to changes in flow directions and actual flow rates, which can make it difficult to reach the junction within the specified time interval. For the first hydraulics scenario (Hyd. #1), Junction J12 can be influenced by different booster stations at various time-steps, with varying control energy depending on the flow paths between them. Conversely, for the second hydraulics scenario (Hyd. #2), higher demands exist on the opposite side of the network, causing flow directions to be opposite to those leading to J12, with low flow rates to J12 except during the controllable windows shown in Fig. 4(b). This suggests that the other side of the network might also represent critical target nodes that demand higher priority indices.

By applying the proposed MPC control approach based on the results from the controllability analysis and constrained by THMs formation (Case #2), the results for the BLA-M network are demonstrated in Fig. 5. Additionally, this figure shows results for Case #1, where the MPC control algorithm is applied without prior controllability analysis, and without constraining the problem with the maximum THMs concentrations. It is important to note that the results for Case #2 are obtained by tuning the matrices R and Q to balance control

input smoothness, achieve the desired concentrations, and weight the inputs appropriately. As illustrated in Fig. 5(a), the combined control actions from R1 and J3 are lower for the second case, while maintaining chlorine concentrations between the lower and upper bounds across the network—see Figs. 5(b) and 5(c) for examples. Two significant factors explain these results. The first factor is the values and scaling of R and Q compared to the cost of chlorine injections, allowing the latter to dominate the objective functions with higher trade-offs for the first two. The second factor is the constraint on chlorine injections to prevent THMs formation from exceeding its maximum bound of 0.08 mg/L. Furthermore, in Case #1, the control actions are primarily assigned to the booster station at R1, while in Case #2, they are distributed between the stations at R1 and J3. This distribution is due to the higher indices assigned to the dead-ends downstream of J3, where initial chlorine concentrations are zero and the fictitious reactant has relatively high initial concentrations of 0.5 mg/L, representing a contamination intrusion event in this zone. Additionally, fluctuations in chlorine concentrations at J2 shown in Fig. 5(c) are due to the need to recover chlorine levels after their consumption by the fictitious reactant and the withdrawn water with high demands.

For the Net3 case study, we focus on the zone framed in Fig. 1. In this network, Tank TK3 faces periods of filling and periods of serving parts of the network throughout the simulation period. Thus,

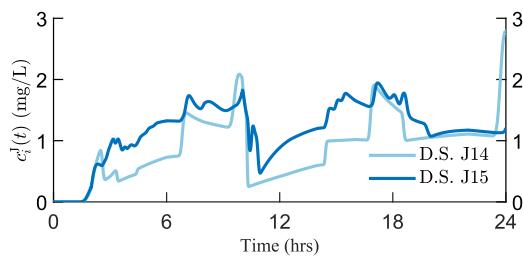


Fig. 6. Chlorine concentrations at the dead-ends downstream (D.S.) Junctions J14 and J15 after applying the MPC algorithm for Net3.

it is considered a critical target node during the filling windows to ensure the stored water has sufficient chlorine concentrations for safe distribution. As shown in Fig. 1, the nearest booster stations to this tank are located at Junctions J13 and J15. Notably, water reaching J15 must first pass through J13, which serves more dead-ends than J15. To accommodate these scenarios, a higher priority index is assigned to the target set leading to the tank during its filling windows. When the tank is emptying, it is considered uncontrollable, and its weight is set to zero in Procedure 2. This approach demonstrates the scalability and flexibility of the proposed control strategy, allowing it to be easily customized for each network and each unique scenario or special consideration that arises.

By applying the proposed MPC approach based on the aforementioned scenarios for Net3, Fig. 6 illustrates chlorine concentrations at the dead-ends downstream Junctions J14 and J15. Results show how the water coming from Junction J13 with the same chlorine concentration, but due to the different weights and interactive dynamics, the chlorine injections by booster stations at J14 and J15 are different. We also want to highlight that although Net3 is a relatively large-scale network, the switching between the two discretization schemes for the cases with and without dispersion has been smooth, and the computational time for a simulation period of 24 h has only increased by 3% to run through the if condition that decides which scheme is used.

5. Conclusion, paper's limitations, and recommendations for future work

This study introduces a novel disinfectant control approach that ensures WQ safety through chlorine injections while mitigating the hazardous formation of DBPs. This approach integrates an advanced multi-species WQ model that is based on the ADR dynamics, dynamically switching between discretization schemes to accurately simulate chemical transport according to the dominant process. Accounting for the dispersion process in chemical evolution simulation can significantly impact results, particularly in low-velocity zones such as branched networks with dead-ends and stagnant water. Additionally, this study demonstrates the utilization of prior controllability analysis, enabling the storage of offline weight matrices that are updated every hydraulic time-step, reflecting the subsequent changes in the WQ controllability metrics. These matrices inform the MPC algorithm for the chlorine control problem, strategically prioritizing booster stations based on their influence on critical states and desired objectives. Consequently, optimal chlorine injections are allocated among booster stations based on the insights from the controllability analysis. The proposed control approach is validated through different numerical case studies with various scales, layouts, and hydraulic conditions.

In our study, we exclusively focus on chlorine as the disinfectant and accordingly refer to its dynamics as the WQ dynamics, given that chlorine is one of the most widely used disinfectants. However, there are

other disinfection methods that differ in their composition, effectiveness, dynamics in WDNs, byproducts, and usage limitations. For a detailed comparison of these disinfection methods and their broader impacts, we refer readers to the study in Tsitsifli and Kanakoudis (2018). Furthermore, our study acknowledges several limitations within its structure, including the high dimensionality associated with the developed approach, potentially leading to demanding computational times. However, model order reduction approaches are available to be utilized and integrated into the control framework as presented in Elsherif et al. (2024). Other limitations are the predetermined locations of booster stations and the prerequisite of detailed hydraulic settings before the WQ control approach implementation. To address these limitations, future research direction for our group is to explore incorporating uncertainty in the hydraulic settings, modeling transient conditions to better reflect real-world scenarios, and adopting an integrated approach that considers optimizing both hydraulic and WQ dynamics. Additionally, it is worth noting that our controllability analysis is performed for the linearized system representing an approximation of the actual system dynamics.

CRediT authorship contribution statement

Salma M. Elsherif: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Ahmad F. Taha:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Ahmed A. Abokifa:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ahmad F. Taha reports financial support was provided by National Science Foundation. Ahmed A. Abokifa reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The .inp files of the test networks are available in the online repository (Elsherif, 2024). Any other data, models, and codes that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Abokifa, A.A., Xing, L., Sela, L., 2020. Investigating the impacts of water conservation on water quality in distribution networks using an advection-dispersion transport model. *Water* 12 (4), 1033.
- Abokifa, A.A., Yang, Y.J., Lo, C.S., Biswas, P., 2016. Water quality modeling in the dead end sections of drinking water distribution networks. *Water Res.* 89, 107–117.
- Acrylamide, O., 2009. National Primary Drinking Water Regulations. Vol. 2, pp. 0–07.
- Ardeshir, A., Alimohammadnezhad, M., Behzadian, K., Jalilzani, F., 2011. Control of THM formation in multi-objective booster chlorination for water distribution systems. In: Computing and Control for the Water Industry, CCWI Conference.
- Ayvaz, M.T., Kentel, E., 2015. Identification of the best booster station network for a water distribution system. *J. Water Resour. Plan. Manag.* 141 (5), 04014076.
- Basha, H.A., Malaeb, L.N., 2007. Eulerian-Lagrangian method for constituent transport in water distribution networks. *J. Hydraul. Eng.* 133 (10), 1155–1166.
- Behzadian, K., Alimohammadnejad, M., Ardeshir, A., Jalilzani, F., Vasheghani, H., 2012. A novel approach for water quality management in water distribution systems by multi-objective booster chlorination. pp. 51–60.
- Blokker, M., Vreeburg, J.H.G., Buchnberger, S.G., van Dijk, J.C., 2007. Importance of demand modelling in network water quality models: A review. In: Conference High Quality Drinking Water 2007, Delft. Delft University of Technology, pp. 1–10.
- 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. *J. Water Resour. Plan. Manag.* 124 (2), 99–111.

Constable, G., Somerville, B., 2003. *A Century of Innovation: Twenty Engineering Achievements that Transformed Our Lives*. Joseph Henry Press.

Constans, S., Brémond, B., Morel, P., 2012. Using linear programs to optimize the chlorine concentrations in water distribution networks. pp. 1–12. [http://dx.doi.org/10.1061/40517\(2000\)208](http://dx.doi.org/10.1061/40517(2000)208).

Cozzolino, L., Pianese, D., Pirozzi, F., 2005. Control of DBPs in water distribution systems through optimal chlorine dosage and disinfection station allocation. *Desalination* 176 (1–3), 113–125.

Datta, B., 2004. *Numerical Methods for Linear Control Systems*. Academic Press.

2022. Disinfection by-products (DBPs) factsheet | national biomonitoring program | CDC. https://www.cdc.gov/biomonitoring/THM-DBP_FactSheet.html.

Ding, Y., Sun, Q., Lin, Y., Ping, Q., Peng, N., Wang, L., Li, Y., 2024. Application of artificial intelligence in (waste) water disinfection: Emphasizing the regulation of disinfection by-products formation and residues prediction. *Water Res.* 121267.

Durran, D.R., 2010. *Numerical Methods for Fluid Dynamics: with Applications to Geophysics*. Vol. 32, Springer Science & Business Media.

Duzinkiewicz, K., Brdys, M., Chang, T., 2005. Hierarchical model predictive control of integrated quality and quantity in drinking water distribution systems. *Urban Water J.* 2 (2), 125–137.

Elsherif, S.M., 2024. Github repository with inp. files of the test netowrks. https://github.com/salmamelsherif/BenchmarkNetworks_DBPsPaper.git.

Elsherif, S.M., Taha, A.F., Abokifa, A.A., Sela, L., 2024. Comprehensive framework for controlling nonlinear multispecies water quality dynamics. *J. Water Resour. Plan. Manag.* 150 (2), 04023077.

Elsherif, S.M., Wang, S., Taha, A.F., Sela, L., Giacomoni, M.H., Abokifa, A.A., 2023. Control-theoretic modeling of multi-species water quality dynamics in drinking water networks: Survey, methods, and test cases. *Annu. Rev. Control* 55, 466–485.

Engineering Consultant, Grayman, W.M., Kshirsagar, S., Global Quality Corp., Rivera-Sustache, M., USACE, Ginsberg, M., USACE, 2012. An Improved Water Distribution System Chlorine Decay Model Using EPANET MSX. *J. Water Manag. Model.* 20, R245–21.

Fisher, I., Kastl, G., Sathasivan, A., 2017. A comprehensive bulk chlorine decay model for simulating residuals in water distribution systems. *Urban Water J.* 14 (4), 361–368.

Fisher, I., Kastl, G., Shang, F., Sathasivan, A., 2018. Framework for optimizing chlorine and byproduct concentrations in drinking water distribution systems. *J. - Am. Water Works Assoc.* 110 (11), 38–49.

Gokhale, A., Valli, S.M., Kalaimani, R., Pasumarthy, R., 2021. Optimizing controllability metrics for target controllability. In: 2021 Seventh Indian Control Conference. ICC, IEEE, Mumbai, India, pp. 141–146.

Grayman, W.M., 2018. History of water quality modeling in distribution systems. In: *WDSA / CCWI Joint Conference Proceedings*. Vol. 1.

Helbling, D.E., VanBriesen, J.M., 2009. Modeling residual chlorine response to a microbial contamination event in drinking water distribution systems. *J. Environ. Eng.* 135 (10), 918–927.

Hirsch, C., 1990. *Numerical Computation of Internal and External Flows*. Vol. 2- Computational Methods for Inviscid and Viscous Flows. Vol. 708, John Wiley & Sons, Chichester, England and New York, 1990.

Jonkergouw, P.M.R., Khu, S.-T., Savic, D.A., Zhong, D., Hou, X.Q., Zhao, H.-B., 2009. A variable rate coefficient chlorine decay model. *Environ. Sci. Technol.* 43 (2), 408–414.

Kalita, I., Kamarilas, A., Havinga, P., Reva, I., 2024. Assessing the health impact of disinfection byproducts in drinking water. *ACS Est Water* 4 (4), 1564–1578.

Kalman, R.E., 1963. Mathematical description of linear dynamical systems. *J. Soc. Ind. Appl. Math. Ser. A* 1 (2), 152–192.

Leaist, D.G., 1986. Absorption of chlorine into water. *J. Solut. Chem.* 15, 827–838.

Lee, Y., 2004. *Mass Dispersion in Intermittent Laminar Flow* (Ph.D. dissertation). University of Cincinnati, United States – Ohio.

Li, Z., Buchberger, S.G., Tzatchkov, V., 2005. Importance of dispersion in network water quality modeling. In: *Impacts of Global Climate Change*. American Society of Civil Engineers, Anchorage, Alaska, United States, pp. 1–12.

Li, R.A., McDonald, J.A., Sathasivan, A., Khan, S.J., 2019. Disinfectant residual stability leading to disinfectant decay and by-product formation in drinking water distribution systems: A systematic review. *Water Res.* 153, 335–348.

Maheshwari, A., Abokifa, A., Gudi, R.D., Biswas, P., 2020. Optimization of disinfectant dosage for simultaneous control of lead and disinfection-byproducts in water distribution networks. *J. Environ. Manag.* 276, 111186.

Mala-Jetmarova, H., Sultanova, N., Savic, D., 2017. Lost in optimisation of water distribution systems? A literature review of system operation. *Environ. Model. Softw.* 93, 209–254.

Moeini, M., Sela, L., Taha, A.F., Abokifa, A.A., 2023. Bayesian optimization of booster disinfection scheduling in water distribution networks. *Water Res.* 242, 120117.

Munavalli, G.R., Kumar, M.S.M., 2003. Optimal scheduling of multiple chlorine sources in water distribution systems. *J. Water Resour. Plan. Manag.* 129 (6), 493–504.

Munavalli, G.R., Mohan Kumar, M.S., 2005. Multi-step Eulerian method for multicomponent transport in water networks. *ISH J. Hydraul. Eng.* 11 (3), 103–115.

Ozdemir, O.N., 2023. Discussion of “Lagrangian method to model advection-dispersion-reaction transport in drinking water pipe networks” by Feng Shang, Hyoungmin Woo, Jonathan B. Burkhardt, and Regan Murray. *J. Water Resour. Plan. Manag.* 149 (1), 07022004.

Pineda Sandoval, J.D., Brentan, B.M., Lima, G.M., Cervantes, D.H., García Cervantes, D.A., Ramos, H.M., Delgado Galván, X., Mora Rodríguez, J.D.J., 2021. Optimal placement and operation of chlorine booster stations: a multi-level optimization approach. *Renewable Energy* 14 (18), 5806.

Prasad, T.D., Walters, G.A., Savic, D.A., 2004. Booster disinfection of water supply networks: multiobjective approach. *J. Water Resour. Plan. Manag.* 130 (5), 367–376.

Rossman, L.A., Boulos, P.F., 1996. Numerical methods for modeling water quality in distribution systems: a comparison. *J. Water Resour. Plan. Manag.* 122 (2), 137–146.

Rossman, L.A., Woo, H., Tryby, M., Shang, F., Janke, R., Haxton, T., 2020. EPANET 2.2 user manual water infrastructure division, center for environmental solutions and emergency response.

Shah, M., Mazhar, M.A., Ahmed, S., Lew, B., Khalil, N., 2024. Recent trends in controlling the disinfection by-products before their formation in drinking water: A review. In: *Drinking Water Disinfection by-Products: Sources, Fate and Remediation*. Springer, pp. 177–192.

Shang, F., Uber, J.G., Polycarpou, M.M., 2002. Particle backtracking algorithm for water distribution system analysis. *J. Environ. Eng.* 128 (5), 441–450.

Tryby, M.E., Boccelli, D.L., Uber, J.G., Rossman, L.A., 2002. Facility location model for booster disinfection of water supply networks. *J. Water Resour. Plan. Manag.* 128 (5), 322–333.

Tsitsifli, S., Kanakoudis, V., 2018. Disinfection impacts to drinking water safety—a review. *Proceedings* 2 (11), 603.

Tzatchkov, V.G., Aldama, A.A., Arreguin, F.I., 2002. Advection-dispersion-reaction modeling in water distribution networks. *J. Water Resour. Plan. Manag.* 128 (5), 334–342.

Wang, Q., Guidolin, M., Savic, D., Kapelan, Z., 2015. Two-objective design of benchmark problems of a water distribution system via MOEAs: towards the best-known approximation of the true Pareto front. *J. Water Resour. Plan. Manag.* 141 (3), 04014060.

Wang, J.-J., Liu, X., Ng, T.W., Xiao, J.-W., Chow, A.T., Wong, P.K., 2013. Disinfection byproduct formation from chlorination of pure bacterial cells and pipeline biofilms. *Water Res.* 47 (8), 2701–2709.

Wang, Z., Polycarpou, M.M., Uber, J.G., Shang, F., 2005. Adaptive control of water quality in water distribution networks. *IEEE Trans. Control Syst. Technol.* 14 (1), 149–156.

Wang, S., Taha, A.F., Abokifa, A.A., 2021. How effective is model predictive control in real-time water quality regulation? state-space modeling and scalable control. *Water Resour. Res.* 57 (5), e2020WR027771.

Zhang, C., Lu, J., 2021. Optimizing disinfectant residual dosage in engineered water systems to minimize the overall health risks of opportunistic pathogens and disinfection by-products. *Sci. Total Environ.* 770, 145356.

Zhou, Q., Bian, Z., Yang, D., Fu, L., 2023. Stability of drinking water distribution systems and control of disinfection by-products. *Toxics* 11 (7), 606.