Control-Theoretic Modeling of Multi-Species Water Quality Dynamics in Drinking Water Networks: Survey, Methods, and Test Cases

Salma M. Elsherif[†,¶], Shen Wang[*], Ahmad F. Taha[†,**], Lina Sela[♠], Marcio H. Giacomoni[‡], and Ahmed Abokifa[◇]

Abstract

Chlorine is a widely used disinfectant and proxy for water quality (WQ) monitoring in water distribution networks (WDN). Chlorine-based WQ regulation and control aims to maintain pathogen-free water. Chlorine residual evolution within WDN is commonly modeled using the typical single-species decay and reaction dynamics that account for network-wide, spatiotemporal chlorine concentrations only. Prior studies have proposed more advanced and accurate descriptions via multi-species dynamics. This paper presents a host of novel state-space, control-theoretic representations of multi-species water quality dynamics. These representations describe decay, reaction, and transport of chlorine and a fictitious reactive substance to reflect realistic complex scenarios in WDN. Such dynamics are simulated over space- and time-discretized grids of the transport partial differential equation and the nonlinear reaction ordinary differential equation. To that end, this paper *(i)* provides a full description on how to formulate a high fidelity model-driven state-space representation of the multi-species water quality dynamics and *(ii)* investigates the applicability and performance of different Eulerian-based schemes (Lax-Wendroff, backward Euler, and Crank-Nicolson) and Lagrangian-based schemes (method of characteristics) in contrast with EPANET and its EPANET-MSX extension. Numerical case studies reveal that the Lax-Wendroff scheme and method of characteristics outperform other schemes with reliable results under reasonable assumptions and limitations.

Index Terms

Multi-species reaction, chlorine decay and reaction, water quality, state-space representation.

I. Introduction and Literature Review

AS water travels through water distribution networks (WDN), water quality deteriorates due to the physical-chemical processes (e.g., growing water age and microbial load increasing), malicious attacks, natural hazards, or lack of maintenance leading to potential contaminant intrusion through leaks. To that end, disinfection has been introduced as an important treatment process for drinking water to ensure meeting the public health requirements (e.g., waterborne disease prevention). To point out the great impact of disinfection, studies [1], [2] state that the occurrence of multiple waterborne diseases (e.g., cholera, typhoid) in the U.S. has decreased remarkably since its application at the start of the 20th century [3]. The National Academy of Engineering (NAE) listed *Safe and Abundant Water* as one of the greatest engineering achievements. Yet, the disinfection process is not straightforward as there are challenges to overcome. The objective is to have a sufficient disinfectant residual all over the network to maintain pathogen-free water while limiting the formation of the health-threatening disinfection by-products (DBPs), while not causing undesirable odor, taste, or color [4], [5].

There are different disinfection methods which are listed and compared in [6] according to their composition (chemical/non-chemical), effectiveness, and limitations. Among these disinfectants, chlorine is the most widely used one in drinking water distribution networks. Modeling the fate of chlorine concentrations within the WDN has been an effective proxy for monitoring real-time water quality. Consequently, various studies have focused on modeling chlorine as it leaves the treatment plant, and travels through the network, until it reaches water users [7].

[†]Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, USA. Emails: salma.m.elsherif@vanderbilt.edu, ahmad.taha@vanderbilt.edu

[¶]Secondary: Department of Irrigation and Hydraulics Engineering, Faculty of Engineering, Cairo University, Giza, Egypt.

[*]School of Cyberspace Security, Beijing University of Posts and Telecommunications. Email: shen.wang@bupt.edu.cn

[♠]Department of Civil, Architecture, and Environmental Engineering, The University of Texas at Austin. Email: linasela@utexas.edu

[‡]School of Civil & Environmental Engineering, and Construction Management, The University of Texas at San Antonio. Email: marcio.giacomoni@utsa.edu

[◇]Department of Civil, Materials, and Environmental Engineering, The University of Illinois Chicago. Email: abokifa@uic.edu

[**]Corresponding author.

To model chlorine evolution in WDN, a full understanding of its behavior is required. As chlorine travels within the water, it decays and reacts with other substances through the processes of Reaction-Transport-Diffusion. However, chlorine evolution is widely modeled by the one-dimension advection-reaction (1-D AR) partial differential equation (PDE). Several numerical methods are applied to solve the PDE, that are Eulerian, Lagrangian, and hybrid Eulerian-Lagrangian based schemes [8]–[11], which discretize the PDE in both time and space over a specific domain. The main distinction between these approaches is the formulation of either fixed or variable-sized segments in the pipes to form the numerical grid over which the equations are solved. Additionally, chlorine decay and reaction processes are characterized by nonlinear functions, which include interactions between multiple species. Yet, due to the modeling complexity the most frequently used model is the single-species first-order equations [12]. This simplified model depicts the chlorine decay and reaction as a function of the chlorine concentration with a constant decay rate without accounting for reactions with other species. Whilst, the decay and reaction models in realistic WDN are much more complicated and the simplified single-species model fails to capture the actual performance, especially under high variability in hydraulics and water quality states. On the other hand, a more advanced and complex description is provided by the multi-species second-order model [13]–[15], which produces a more accurate description. The multi-species model can be used to model chlorine interactions with other species (e.g., microbial, chemical) in the bulk flow, attached to the pipe walls, or contamination events [16]. The simulation can be limited to consider one specific substance (in addition to organic matter) or it can be upgraded to include more substances that consume chlorine interdependently or in form of a fictitious reactant [12], [13].

Building reasonably good models to predict the evolution of water quality states (e.g., concentrations of chemicals) in WDN is the key to perform model-driven feedback control—a mainstream control approach in various infrastructure. The aim is to control and maintain the desired chlorine residual that meets water quality standards. This has been expressed as an interesting research and applied problem in several studies [17], [18]:

> *Is it more effective to apply primary chlorination with a high dose at the treatment plant, the very start of the water distribution networks, or distribute the injection locations over the network with smaller doses? Would the answer to this previous question change under abnormal conditions (e.g., contamination intrusion)?*

Yet, to the best of our knowledge, there are virtually no model-, system-, and network-theoretic studies on the control of water quality in WDN in the existence of other reactive substances (i.e., multi-species dynamics), let alone, to solve the aforementioned research problem. Therefore, this paper's main objective is to investigate and build a model that traces chlorine concentrations in presence of such events. This model is built to be easily integrated in model-based control frameworks (e.g., model-predictive control, state-feedback control, output-feedback control, Lyapunov-based control) while integrating measurement models from WQ sensors. The proposed models can also be used to study control-theoretic and optimal WQ sensors and booster station placement formulations.

Paper's objectives. this paper's objectives are *(i)* to formulate high fidelity water quality model depicting multi-species reaction dynamics of chlorine in the entire network, and *(ii)* to investigate the performance, scalability, and complexity of various discretization-based techniques to simulate the multi-species water quality dynamics. This paper's results pave the way to build novel control algorithms of multi-species dynamics—the algorithms are outside the scope of this paper. Henceforward, we survey the literature to highlight the gap bridged by this paper.

A. Literature review

The literature review covers two main aspects of water quality dynamics that are chlorine transport modeling and decay and reaction modeling. Study [15] states that the system composition and how diverse and case-oriented the reactants with which chlorine reacts establish the existence of huge uncertainty and assumptions. Thus, simulating chlorine transport and reaction has been covered by several studies with different approaches with a wide range of applicability and limitation domains.

Chemicals transport model is derived using the advection-reaction equation in several studies [9], [19], [20]. Study [8] applies different numerical schemes to discretize the advection-reaction equation. Two of these schemes are Eulerian-based: finite-difference and discrete-volume methods, and the other two are Lagrangian-based: time-driven and event-driven methods. Eulerian methods solve the equation on a fixed grid system while Lagrangian methods employ a moving coordinate system using particle tracking algorithms. Each scheme's stability is sensitive to one parameter or another including water quality time-step, segments sizes, and allowed simulation tolerance. Studies [11], [21] propose Lagrangian based-solution to the advection and reaction using the explicit method of characteristics to solve the advection and reaction terms. Study [22] uses a multi-step Eulerian scheme that solves the transport equations by an explicit Eulerian scheme, the MacCormack scheme, while solving the kinetic dynamics using the explicit Runge-Kutta method. On the other hand, the latest version of EPANET, a standard software widely used to model hydraulics and water quality of water distribution systems [23], applies the Lagrangian time-driven method where it works within fixed time intervals and variable-sized segments.

Moreover, the effect of dispersion in chemicals transport has been investigated in several studies [21], [24], [25]. Results have shown that dispersion dominates over advection in dead-end branches where the flow velocities are low and approaching a laminar state. That is, it is an acceptable assumption to rely on the AR dynamics neglecting the dispersion term in networks with limited dead-end branches, higher velocities, and changing demands leading to a turbulent flow state.

In addition to the transport modeling, decay and reaction modeling plays an important role in water quality dynamics. The most widely used model for chlorine decay is the first-order decay model which is based on the assumption that chlorine reacts at a constant rate with excess chlorine-demanding reactants [12], [26]. Moreover, this first-order decay model has been used in EPANET for water quality modeling in water networks [23]. However, the study [27] verifies that the first-order decay model that accounts for only chlorine is not accurate under high variability in concentrations and rechlorination scenarios. Not only the reaction dynamics are affected by the chlorine-consuming reactant concentration in many cases, but also depend on the contact time needed for the reaction to be completed. Thus, other studies [28]–[30] categorize the reactants to slow and fast dynamics and accordingly model the chlorine decay with two different second-order reaction models simultaneously. These models outperform the first-order model in the rechlorination scenarios. Studies [12], [13] introduce a multi-species model of chlorine with every reactive component that presents in the water modeled with its specific reaction rate. Yet, this model requires pre-categorization for each substance according to their reaction rates and to know their initial concentrations at all the network components [16].

Furthermore, in [31], a multicomponent (organic carbon, biomass, and chlorine) reaction transport model is developed. Although this model does not include the reaction dynamics with the microbial contaminant itself, it is proposed to be an alternative approach by tracking the organic carbon and bacterial growth. Yet, it has various parameters to be predetermined/measured through a complex approach. However, this study uses one explicit Eulerian discretization scheme and the time-driven Lagrangian method with no reference to compare to and with no brief description of the implementation of the method on different networks. In addition, the authors in [14], [15] investigate the performance of a reaction model between chlorine and fictitious reactive species in a form of a second-order model to test its advantages over the typically used first-order model. In comparison, their models show a more accurate description, particularly for scenarios with chlorine injected into the network by chlorine booster stations. Note that both decay and reaction of chlorine are compromised in this fictitious reactant. That is, the parameters of their models have to be predetermined for the specific water characteristics of the network to reflect organic matter and substances reacting with chlorine. Yet, the inclusion of this fictitious reactant in the chlorine reaction and decay can play an important role to represent a specific substance/contaminant reacting with chlorine in addition to chlorine decay due to organic matter and bacterial growth. As the study [32] states, along with tracing chlorine concentrations throughout the network to maintain the desired levels under the normal conditions, it is important to model the chlorine with different special cases. These cases include contaminated water due to several reasons (e.g., microbial/chemical contamination intrusion, poorly treated water, substances derived from pipe materials) and due to the fact that WDN are considered vulnerable to contamination and health violation attacks which affect millions of people annually [33], [34].

Contamination sources that can affect water quality parameters vary from microbial contaminants, heavy metals, natural organic matter, etc. [35]. Study [32] investigates chlorine residual modeling's response to microbial contamination events. Results from this study show how chlorine residual is affected by the contamination species, event location, and network topology and characteristics. Study [36] addresses uncontrolled microbial contamination caused by sewage intrusion and how chlorine is consumed as a result. Study [37] states that arsenic (As) is a naturally occurring toxic substance from pipe materials that is highly soluble and often found in water sources, while study [38] considers it the biggest mass poisoning case in the world. However, the authors in [39] show that maintaining residual chlorine is recognized as an effective strategy to control levels of soluble arsenic. These aforementioned contaminates/substances are examples that react with chlorine at various rates leading to different scenarios, hence, controlling chlorine throughout the network follows a case-oriented approach.

On the other hand, several studies [17], [18], [40] investigate the water quality control (WQC) problem and apply different optimization algorithms to reach the optimum scheduling/locations of the injection boosters while minimizing the system cost (i.e., minimizing the cost of chlorine injections and maintaining minimal deviations from chlorine setpoint concentrations) and minimizing the formation of the undesired DBPs. A framework for optimization chlorine and by-product concentrations is presented in [41]. However, such studies do not include a closed-form, network- and control-theoretic representation of all system inputs, variables, and output measurements under normal and abnormal operation conditions thereby limiting their ability to perform network-wide WQC.

Recently, a novel state-space water quality control model based on the conversation of mass and a single-species reaction model has been developed in [42]. This state-space model captures the spatio-temporal evolution of chlorine concentrations throughout all the components of any network (e.g., reservoirs, tanks, junctions, pipes, pumps, valves, and booster stations). However, this state-space model does not simulate or include multi-species interactions between chlorine and any other reactants that are critical in water quality modeling as mentioned—and only uses the Lax-Wendroff scheme without further discussion on other discretization schemes.

B. Paper Contributions and Organization

The paper contributions are three-fold:

- We present a state-space water quality model to include the decay and reaction of chlorine with a fictitious reactant resulting in a multi-species dynamic model. This model is a step forward to depict the existence of a specific reactant(s) with the chlorine in the system and to reflect various real-time water quality scenarios (e.g., water characteristics-, pipes material-, hazard-related). With which, a closed-form and network-theoretic representation is included of all system inputs and variables.

- Although the Lagrangian method has been used within EPANET for decades, other methods are less utilized, understood, or derived in the context of the nonlinear multi-species water quality dynamics. In this context, this paper investigates the performance, scalability, and complexity of various discretization-based techniques used in building this state-space water quality model. These techniques are Lax-Wendroff, Backward Euler, Crank-Nicolson, and Method of Characteristics. The derivation of these techniques is provided through a detailed example and applied on various scaled case studies. This adds a novel educational and theoretical value to the paper as it fills the gap of how different discrete methods perform for multi-species dynamics, how to apply them, and to add a comprehensive framework of the dynamics in a system theoretic way that empowers applying feedback control.
- To compare techniques' validity and simulation results with EPANET and its multi-species extension EPANET-MSX which allows the modeling of multi-species dynamics. EPANET-MSX extension gives the user free control to define the chemical reactions to be included in his model [23], [43]. Comparison is considered reliable as the covering laws and equations are the same for all network components in both models.
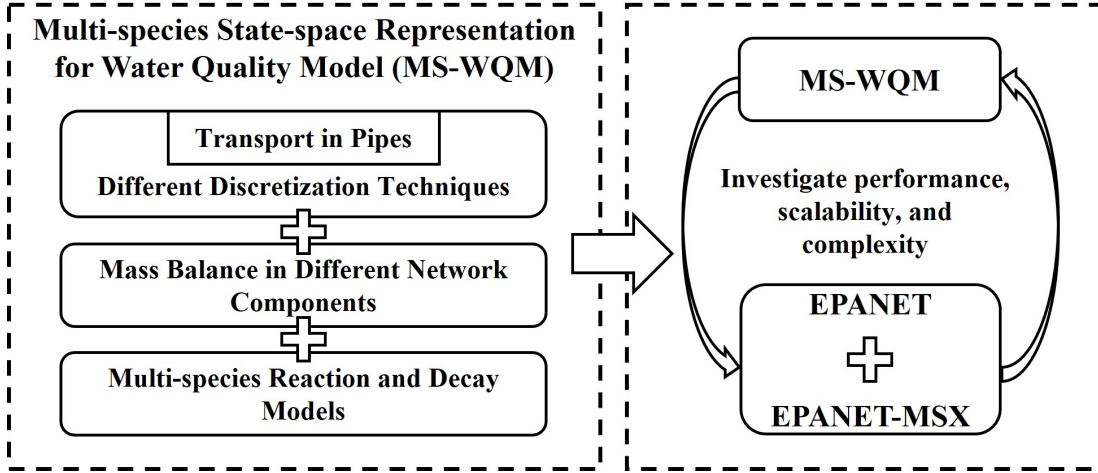

Fig. 1: Conceptual framework of the paper.

The conceptual framework of the paper is illustrated in Fig. 1. The formulation of the state-space representation for multi-species water quality model (MS-WQM) is based on transport in pipes, mass balance at the rest of the network components, and multi-species reaction and decay dynamics. In this paper, we adopt different discretization methods to simulate the transport and reaction in pipes. These methods are compared with EPANET and its extension EPANET-MSX to test their performance, scalability, and complexity.

The rest of the paper is organized as follows. Section II describes the problem formulation based on the principles of water quality in WDN. Section III delineates the transport and reaction model in pipes in detail and how pipes can be discretized using different schemes. Section IV lists the governing equations for different network components by applying mass balance. Section V breaks down the reaction models into decay and mutual reaction dynamics and their implementations. Section VI provides the final formulation of the state-space representation of the system. Section VII showcases the model implementation results on different networks. Based on these results, conclusion, paper's limitations, and recommendations for future work are all given in Section VIII.

II. Water Network Fundamentals and Problem Formulation

The WDN is modeled 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 the collections of junctions, tanks, and reservoirs, respectively. 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.

We simulate the interaction and fate of two chemical substances in our model, thus, the governing equations may differ in one component according to their characteristics. Therefore, general statements of laws and equations for each network component are briefly formulated followed by their application for different substances. In particular, herein we showcase a high-level state-space representation of the water quality model which captures the transport and evolution of the chemicals' concentrations in every component of the network (e.g., reservoirs, junctions, tanks, pipes, pumps, and valves). This representation is a generalized multi-species model that traces the evolution of chlorine and a fictitious reactant that can represent one specific chemical substance or multiples which react at a relatively close rate with chlorine. Furthermore, we showcase how this model can be appended to include more than two distinct chemicals by introducing different reactions dynamics representations.

In general, for each component the water quality model is a function of time-dependent parameters and variables such as velocities, and flow rates, and might depend on the model of another component. The problem formulation in this paper is

to obtain a state-space representation for both chlorine and the fictitious reactant concentrations evolution and their mutual reaction as a nonlinear difference equation (NDE). This representation is formulated for various discretization methods of the transport advection-reaction PDE. Eventually, we formulate a system representation of two-species that is able to capture chemicals evolution, booster stations injections, and sensors measurements, is given by an NDE as

$$
\underbrace{\begin{bmatrix} E_{11}(t) & 0 \\ 0 & E_{22}(t) \end{bmatrix}}_{E(t)} \underbrace{\begin{bmatrix} x_1(t+\Delta t) \\ x_2(t+\Delta t) \end{bmatrix}}_{x(t+\Delta t)} = \underbrace{\begin{bmatrix} A_{11}(t) & 0 \\ 0 & A_{22}(t) \end{bmatrix}}_{A(t)} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} B_{11}(t) & 0 \\ 0 & B_{22}(t) \end{bmatrix}}_{B(t)} \underbrace{\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}}_{u(t)} + f(x_1, x_2, t), \quad \text{(1a)}
$$

$$
\underbrace{\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}}_{y(t)} = \underbrace{\begin{bmatrix} C_{11}(t) & 0 \\ 0 & C_{22}(t) \end{bmatrix}}_{C(t)} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}}_{x(t)} \quad \text{(1b)}
$$

where variable t represents specific time index in a simulation period $[0, T_s]$; Δt is the time-step or sampling time; vectors $x_1(t)$ and $x_2(t) \in \mathbb{R}^{n_x}$ depict the concentrations of chlorine and the fictitious reactant (two species model) in the entire network; vector $u_1(t) \in \mathbb{R}^{n_{u_1}}$ represents the dosages of injected chlorine; vector $u_2(t) \in \mathbb{R}^{n_{u_2}}$ accounts for planned or unplanned injection of the fictitious component; vector $f(x_1, x_2, t)$ encapsulates the nonlinear part of the equations representing the mutual nonlinear reaction between the two chemicals; vector $y_1(t) \in \mathbb{R}^{n_{y_1}}$ denotes the sensor measurements of chlorine concentrations at specific locations in the network while $y_2(t) \in \mathbb{R}^{n_{y_2}}$ captures the fictitious reactant measurements by sensors in the network if they exist. The state-space matrices $E_{11}, E_{22}, A_{11}, A_{22}, B_{11}, B_{22}, C_{11}$ and C_{22} are all time-varying matrices that depend on the network topology, hydraulic parameters, and disinfectant decay rate coefficients. It is customary to assume that these matrices evolve at a slower pace than the states $x(t)$ and control inputs $u(t)$.

The model, built on the conservation of mass law, transport, decay, and reaction models of the substances, captures the concentrations in all network components (junctions, tanks, pipes, valves, pumps, and reservoirs). Some of these models are in a form of ordinary differential equations (ODEs) or PDEs and converted to difference algebraic equations by applying different schemes and approaches (e.g., Eulerian-based and Lagrangian-based schemes for pipes, continuously stirred tank reactor model for tanks). On the other hand, the decay and reaction models are combined in a form of a multi-species model. Notice that the A matrices are placed in a form of a block-diagonal big matrix, which indicates that the state-space modeling is mainly decoupled for the two chemicals and only the nonlinear part f in (1a) reflects their reaction and relation. To that end, the objective of the next sections is to derive the matrices for a generic water distribution network.

As a preparation for the next sections, we list the symbols for different quantities in Tab. I. For example, $c(t)$ represents chemical concentration; $q(t)$ depicts the flow rate for a specific component depending on the superscript it has; $q^D(t)$ is the demand drawn from a node; $q^B(t)$ is the booster flow rate and in this case, the subscript represents the node with the injected flow; V is the volume of the tank. As for the network components, P, M, V, J, TK, R, and B are super/sub-scripts representing pipes, pumps. valves, junctions, tanks, reservoirs, and booster stations respectively; while super/sub-scripts L and N represent links and nodes. Note that links represent pipes, pumps, and valves, and nodes include reservoirs, tanks, and junctions.

It is worthwhile to note that all hydraulic variables and parameters (e.g., flow rates, velocities, water levels, pump setting, etc.) are assumed to be predetermined. This assumption is widely accepted due to the difference between the water quality modeling time-step and the one for the hydraulic simulation [42], [44]. Hydraulic modeling time-step is acceptable to be within an hourly scale to be aligned with the assumption of steady system simulation and to reflect the change in demand, while the range is between minutes and seconds for water quality modeling to allow a stable numerical simulation that captures chemicals reaction and evolution [43], [45].

III. TRANSPORT AND REACTION IN PIPES: DISCRETIZATION TECHNIQUES

Conservation of mass during transport and reaction in pipes is modeled by the one-dimension advection-reaction (1-D AR) equation. That is, we assume that the dispersion is neglected during our simulation (i.e., medium to high velocity ranges, limited number of dead-end nodes) [21]. To that end, in any Pipe i of the network, the 1-D AR equation is expressed by a PDE as:

$$
\frac{\partial c_i^P}{\partial t} = -v_i(t)\frac{\partial c_i^P}{\partial x} + R_{MS}^P(c_i^P(x,t)), \quad \text{(2)}
$$

where $c_i^P(x,t)$ is the concentration in Pipe i at location x and time t; $v_i(t)$ is the mean flow velocity which equals $\frac{4q_i^P(t)}{\pi r_{P_i}^2}$; $q_i^P(t)$ is flow; r_{P_i} is the pipe radius; and $R_{MS}^P(c_i^P(x,t))$ is the multiple-species reaction rate expression, which is explained briefly in Section V.

The PDE (2) can be solved using different numerical schemes that are Eulerian, Lagrangian, and mixed Eulerian-Lagrangian schemes [8]–[11]. The approach of the Eulerian Finite-Difference (EFD) schemes relies on defining a fixed numerical grid that

Tab. I: Paper variables and parameters notation

Symbol	Description	Dimensions
$c^{\mathrm{N}} := \{c^{\mathrm{R}}, c^{\mathrm{J}}, c^{\mathrm{TK}}\}$	Concentrations at nodes (reservoirs, junctions, and tanks)	$\mathbb{R}^{n_{\mathrm{N}}}$
$c^{\mathrm{L}} := \{c^{\mathrm{M}}, c^{\mathrm{V}}, c^{\mathrm{P}}\}$	Concentrations at links (pumps, valves, and pipes)	$\mathbb{R}^{n_{\mathrm{L}}}$
$x := \{c^{\mathrm{N}}, c^{\mathrm{L}}\}$	System states (i.e., chemical concentrations)	$\mathbb{R}^{n_x}$
$u := \{c^{\mathrm{BJ}}, c^{\mathrm{BTK}}\}$	System inputs (i.e., chemical injections at nodes)	$\mathbb{R}^{n_u}$
$q^{\mathrm{M}}, q^{\mathrm{V}}, q^{\mathrm{P}}$	Flow rates in pumps, valves, and pipes	$\mathbb{R}^{n_{\mathrm{M}}}, \mathbb{R}^{n_{\mathrm{V}}}, \mathbb{R}^{n_{\mathrm{P}}}$
q^{DJ}	Demands from junctions	$\mathbb{R}^{n_{\mathrm{J}}}$
$q_{\mathrm{in}}, q_{\mathrm{out}}$	Inflows and outflows at nodes	—
q^{BJ}	Flow injected to junctions by booster stations	$\mathbb{R}^{n_{\mathrm{J}}}$
V^{TK}	Tank volume	$\mathbb{R}^{n_{\mathrm{TK}}}$
V^{BTK}	Volume injected to tank by booster station	$\mathbb{R}^{n_{\mathrm{TK}}}$
v	Flow velocity in pipes	$\mathbb{R}^{n_{\mathrm{P}}}$
L, r_{P}, s	Pipes length, radius, and number of segments to be divided into	$\mathbb{R}^{n_{\mathrm{P}}}$
$\tilde{\lambda}$	Courant number	$\mathbb{R}^{n_{\mathrm{P}s}}$
$R_{\mathrm{MS}}, R_{\mathrm{D}}, R_{\mathrm{M}}$	Multi-species, decay, and mutual reaction expressions	—
k_b, k_w, k_f, k_r	Chlorine bulk and wall reaction rate constants, mass transfer coefficient between bulk flow and pipe wall, and mutual reaction rate constant between two chemicals	—
$k^{\mathrm{P}}, k^{\mathrm{TK}}$	Chlorine decay-reaction rates for pipes and tanks	$\mathbb{R}^{n_{\mathrm{P}s}}, \mathbb{R}^{n_{\mathrm{TK}}}$
T_s	Simulation period	—

discretizes the pipes in time and space resulting in approximate algebraic equation over the numerical grid. In Equation (3), Pipe i with length L_i is split into a number of segments s_i of length Δx_i.

$$s_i = \left\lfloor \frac{L_i}{v_i(t)\Delta t} \right\rfloor, \quad \Delta x = \frac{L_i}{s_i} \tag{3}$$

As such, the EFD schemes calculates the concentration at any segment s depending on the concentrations of the upstream and downstream nodes/segments. The scheme is considered *implicit* when the upstream and downstream nodes/segments concentrations are taken at the current time-step; it is considered *explicit* when they are taken at the previous time-step; *explicit-implicit* when averaging the current and previous time-steps concentrations; see Fig. 2.
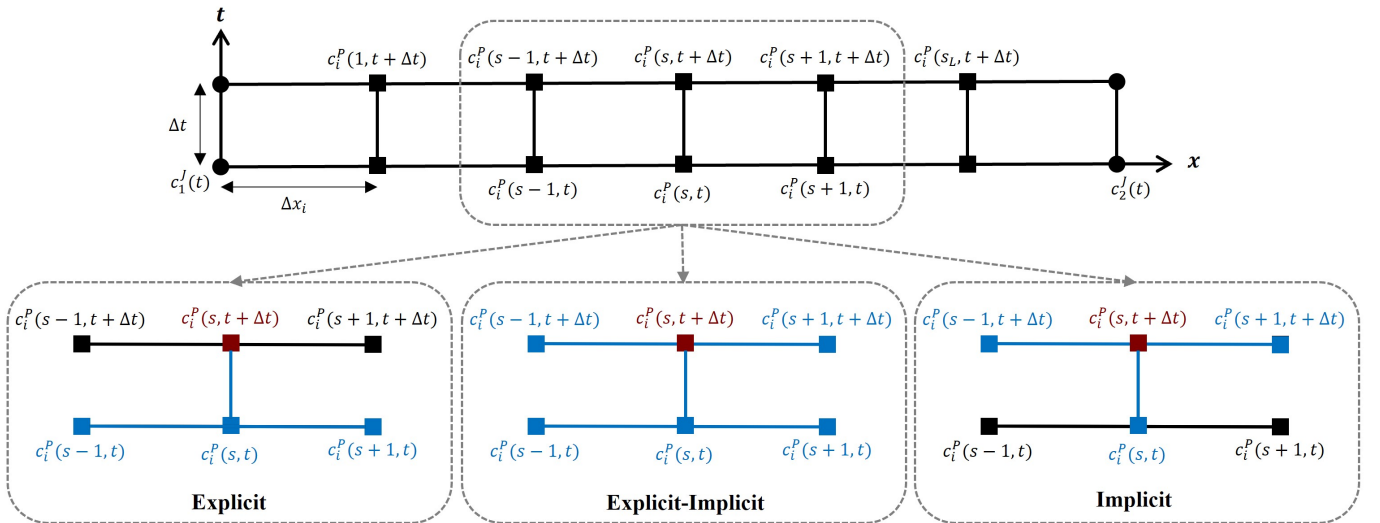
Fig. 2: Eulerian Finite-Difference discretization schemes for Pipe i connecting Junctions 1 and 2. Schemes calculate concentration $c_i^{\mathrm{P}}(s, t + \Delta t)$ at segment s (colored in maroon) depending on concentrations at the other segments/nodes colored in blue

However, EPANET and its extension EPANET-MSX use the Lagrangian time-driven method (LTD) [10], [23]. We give a full description of EPANET's algorithm in Appendix A. The main difference between EFD and LTD schemes is that LTD divides the pipe into non-overlapping changing-sized segments by allowing the most upstream segment to change its size at the expense of the most downstream segment. The change in the segment's size depends on the mass and flow entering the pipe every time step. However, EFD schemes work within a fixed grid that produces a state-space description of finite-dimension. This description allows us to formulate a control problem and solve an optimization problem to identify optimum solutions.

In this paper, we use a Lagrangian-based method that uses an explicit method of characteristics (MOCs) to project concentrations forward in time which is achieved through integration and interpolation [11], [21], [46]. This method allows for the state-space formulation of the concentrations profile with higher complexity than EFD schemes, more details on that are explained in Section III-D.

We also investigate three EFD schemes: *explicit, explicit-implicit,* and *implicit* schemes in addition to a Lagrangian-based method. When applying these schemes, the pipes are divided into a number of segments, thus, the size of vector $c^{\mathrm{P}}(t + \Delta t)$ is the summation of segments count for all pipes in the network. Hence, the size of the vector is $n_{\mathrm{P}_s} = \sum_{i=1}^{n_{\mathrm{P}}} s_i$ with n_{P} pipes in the network. For these techniques, the concentration at any segment $c_i^{\mathrm{P}}(s, t + \Delta t)$ depends on the concentrations at the previous time-step $c_i^{\mathrm{P}}(s, t)$ and the adjacent segments/nodes at the previous and/or current time-steps (i.e., $c_i^{\mathrm{P}}(s-1, t), c_i^{\mathrm{P}}(s+1, t), c_i^{\mathrm{P}}(s-1, t+\Delta t), c_i^{\mathrm{P}}(s+1, t+\Delta t)$) depending on the applied technique; see Fig. 2 and Fig. 3. This dependency is formulated using the Courant number $\tilde{\lambda}_i(t) = v_i(t)\frac{\Delta t}{\Delta x_i}$, where $v_i(t)$ is the mean flow velocity in Pipe i; Δt is the water quality modeling time-step; and Δx_i is the chosen segment length. In the following sections, a brief description of each technique is provided.

A. Technique 1: Explicit Eulerian scheme — Lax-Wendroff scheme

One of the EFD schemes is the Lax-Wendroff (L-W) explicit scheme [8], [47] which is widely used for solving hyperbolic differential equations with second-order accuracy [48]–[50]. In water quality modeling research, a water quality state-space representation of a single-species model is previously developed for the first time in [42] using the Lax-Wendroff scheme to discretize the 1-D AR PDE. In this scheme, the chemical concentration at a segment s and time $t + \Delta t$ depends on the concentrations at time t of segment s itself, upstream segment $s - 1$, and downstream segment $s + 1$. Hence, by applying the Lax-Wendroff scheme for segment s of Pipe i, except for the first and last segments, the concentration is calculated as

$$c_i^{\mathrm{P}}(s, t + \Delta t) = \underline{\lambda}_i(t)c_i^{\mathrm{P}}(s-1, t) + \lambda_i(t)c_i^{\mathrm{P}}(s, t) + \overline{\lambda}_i(t)c_i^{\mathrm{P}}(s+1, t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s, t))\Delta t, \tag{4}$$

where $\underline{\lambda}_i(t), \lambda_i(t),$ and $\overline{\lambda}_i(t)$ are the weighting coefficients calculated as follow

$$\begin{aligned}
\underline{\lambda}_i(t) &= 0.5\tilde{\lambda}_i(t)\left(1 + \tilde{\lambda}_i(t)\right), \\
\lambda_i(t) &= 1 - \tilde{\lambda}_i^2(t), \\
\overline{\lambda}_i(t) &= -0.5\tilde{\lambda}_i(t)\left(1 - \tilde{\lambda}_i(t)\right).
\end{aligned} \tag{5}$$

Note that the L-W scheme is conditionally stable by applying the Courant-Friedrichs-Lewy condition (CFL) [47]. The CFL condition puts the Courant number in the range of $0 < \tilde{\lambda}_i(t) \le 1$. Thus, the choice of the number of segments and the water quality time-step must be in accordance to satisfy the CFL condition. The velocities are predetermined during the hydraulic simulation, subsequently, the first step for water quality simulation, when using this technique, is choosing the time-step or number of segments of all pipes in the network to ensure fulfilling the stability condition.

Remark 1. *For the first segment $s = 1$ in a pipe, there is no upstream segment and the concentration depends on the concentration of its upstream node. Likewise, the concentration of the last segment $s = s_{\mathrm{L}}$ depends on the downstream node's concentration. The concentrations in other network components connected to a pipe are affected by the properties and parameters of the pipe's first or last segment according to the flow direction.*

According to Remark 1, the concentrations in the first and last segment are expressed in Equation (6) assuming that the connecting nodes are two junctions that are Junctions j and k — a full description of how to calculate concentrations at junctions in the next section is given.
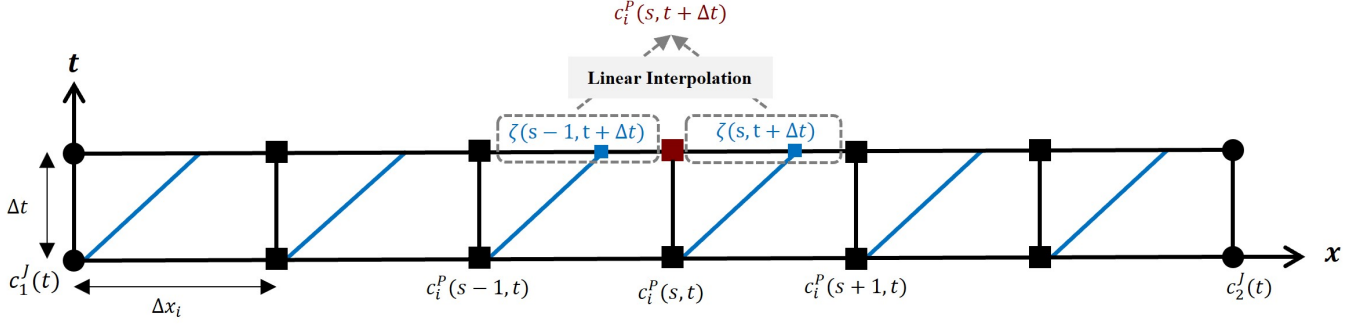
$$c_i^{\mathrm{P}}(1, t + \Delta t) = \underline{\lambda}_i(t)c^{\mathrm{J}}(t) + \lambda_i(t)c_i^{\mathrm{P}}(1, t) + \overline{\lambda}_i(t)c_i^{\mathrm{P}}(2, t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(1, t))\Delta t, \tag{6a}$$

$$c_i^{\mathrm{P}}(s_{\mathrm{L}}, t + \Delta t) = \underline{\lambda}_i(t)c_i^{\mathrm{P}}(s_{\mathrm{L}} - 1, t) + \lambda_i(t)c_i^{\mathrm{P}}(s_{\mathrm{L}}, t) + \overline{\lambda}_i(t)c^{\mathrm{J}}(t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s_{L_i}, t))\Delta t. \tag{6b}$$

B. Technique 2: Implicit Eulerian scheme — Backward Euler scheme

Backward Euler scheme follows an implicit EFD approach [51]. The approach is for chemical concentration at segment s of Pipe i is expressed as:

$$0.5\tilde{\lambda}_i(t + \Delta t)c_i^{\mathrm{P}}(s+1, t + \Delta t) + c_i^{\mathrm{P}}(s, t + \Delta t) - 0.5\tilde{\lambda}_i(t + \Delta t)c_i^{\mathrm{P}}(s-1, t + \Delta t) = c_i^{\mathrm{P}}(s, t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s, t))\Delta t. \tag{7}$$

Fig. 3: Lagrangian-MoCs for Pipe i connecting Junctions 1 and 2.

Assumption 1. *Water quality time-step is much smaller than the hydraulic simulation time-step. That is, $\tilde{\lambda}_i$ is constant within the same hydraulic simulation time-step.*

By applying Assumption 1 and Remark 1, Equation (7) is updated. To that end, for any, first, and last segment (assuming connected to Junctions j and k) of the pipe, concentrations are expressed as Equations (8a), (8b), and (8c).

$$0.5\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(2,t+\Delta t) + c_i^{\mathrm{P}}(1,t+\Delta t) - 0.5\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t+\Delta t) = c_i^{\mathrm{P}}(1,t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(1,t))\Delta t, \tag{8a}$$

$$0.5\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s+1,t+\Delta t) + c_i^{\mathrm{P}}(s,t+\Delta t) - 0.5\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s-1,t+\Delta t) = c_i^{\mathrm{P}}(s,t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s,t))\Delta t, \tag{8b}$$

$$0.5\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t+\Delta t) + c_i^{\mathrm{P}}(s_{\mathrm{L}},t+\Delta t) - 0.5\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s_{\mathrm{L}}-1,t+\Delta t) = c_i^{\mathrm{P}}(s_{\mathrm{L}},t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s_{\mathrm{L}},t))\Delta t. \tag{8c}$$

Note that the Backward Euler scheme is different from the Lax-Wendroff scheme, and the stability check is not needed as it is unconditionally stable. Therefore, the choice of the number of segments does not depend on fulfilling any stability condition and can be chosen as an arbitrary number that preserves the desired accuracy.

C. Technique 3: Explicit-Implicit Eulerian scheme — Crank-Nicolson scheme

Crank-Nicolson scheme [52] is as a hybrid method between the two previous ones and takes an average of the explicit and implicit schemes. And by considering Assumption 1and Remark 1, concentrations at any segment s, the first segment and the last segment are expressed through Equations (9a), (9b), and (9c).

$$0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(2,t+\Delta t) + c_i^{\mathrm{P}}(1,t+\Delta t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t+\Delta t)$$
$$= 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t) + c_i^{\mathrm{P}}(1,t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(2,t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(1,t))\Delta t, \tag{9a}$$

$$0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s+1,t+\Delta t) + c_i^{\mathrm{P}}(s,t+\Delta t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s-1,t+\Delta t)$$
$$= 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s-1,t) + c_i^{\mathrm{P}}(s,t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s+1,t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s,t))\Delta t, \tag{9b}$$

$$0.25\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t+\Delta t) + c_i^{\mathrm{P}}(s_{\mathrm{L}},t+\Delta t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s_{\mathrm{L}}-1,t+\Delta t)$$
$$= 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s_{\mathrm{L}}-1,t) + c_i^{\mathrm{P}}(s_{\mathrm{L}},t) - 0.25\tilde{\lambda}_i(t)c_i^{\mathrm{J}}(t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s_{\mathrm{L}},t))\Delta t. \tag{9c}$$

Moreover, this method is unconditionally stable as the Backward Euler scheme where the choice of the number of segments does not depend on fulfilling the CFL condition. Yet, temporal numerical dispersion and oscillation may have a non-neglectable effect on the results [53].

D. Technique 4 — Method of Characteristics

Method of Characteristics (MoCs) reduces Equation (2) to an ODE along the advection characteristics line whose slope is $\dfrac{dt}{dx_i}$; see the blue line in Fig. 3. By the definition of advection, we find that $\frac{dx_i}{dt} = v_i(t)$ for that line. However, to guarantee the line falls within the same segment in the next time-step the number of segments and segment size are calculated by (3).

The forward projection of concentrations from time t through this line specifies new locations at $t + \Delta t$, shifted by a $v_i(t)\Delta t$ distance. The new location for concentrations are donated by ζ as illustrated in Fig. 3 where the concentration $c^{\mathrm{P}}(s,t)$ is projected to $\zeta(s,t+\Delta t)$. However, to calculate concentrations profile for different ζ at time-step $t + \Delta t$ according to the new reduced ODE, the reaction term should be added to the advection profile. That is, the derivation of the concentration expression is included in Section V after introducing the reaction and decay models. Henceforward, segments' concentrations profile at time-step $t + \Delta t$ is calculated by applying linear interpolation from the ζ profile at same time-step [11], [46], [54].

IV. Mass balance at Different Network's Components

For components other than pipes (i.e., reservoirs, pumps, valves, junctions, and tanks), the conservation of mass is applied to obtain concentrations at different time steps. Different descriptions are provided for components depending on their characteristics and connected nodes/links.

A. Mass balance at reservoirs

Reservoirs are assumed to have constant concentrations where they can present a continuous source of a specific chemical. In vector $c^{\mathrm{R}}(t + \Delta t) \in \mathbb{R}^{n_{\mathrm{R}}}$, for each Reservoir i concentration is expressed as

$$c_i^{\mathrm{R}}(t + \Delta t) = c_i^{\mathrm{R}}(t). \tag{10}$$

B. Mass balance at pumps and valves

Links other than pipes which are represented in pumps and valves are assumed to have no defined length or storage. Subsequently, the model deals with them as transmission links with concentration equals to the concentration of the node upstream while heir functions are reflected in the hydraulic model. That being said, the concentrations for Pump i or Valve j, in vectors $c^{\mathrm{M}}(t + \Delta t) \in \mathbb{R}^{n_{\mathrm{M}}}$ and $c^{\mathrm{V}}(t + \Delta t) \in \mathbb{R}^{n_{\mathrm{V}}}$, installed after Reservoir k (as an example), are expressed as

$$c_i^{\mathrm{M}}(t + \Delta t) = c_k^{\mathrm{R}}(t + \Delta t), \tag{11}$$

$$c_j^{\mathrm{V}}(t + \Delta t) = c_k^{\mathrm{R}}(t + \Delta t). \tag{12}$$

C. Mass balance at junctions

Chemicals are assumed to have a complete and instantaneous mixing in the junctions with no storage volume [8], [43], [55]. Thus, the concentration at each Junction i in vector $c^{\mathrm{J}}(t) \in \mathbb{R}^{n_{\mathrm{J}}}$ is expressed as

$$c_i^{\mathrm{J}}(t) = \frac{\sum_{j \in L_{\mathrm{in}}} q_{\mathrm{in}}^j(t) c_{\mathrm{in}}^j(t) + q_i^{\mathrm{B_J}}(t) c_i^{\mathrm{B_J}}(t)}{q_i^{\mathrm{D_J}}(t) + \sum_{k \in L_{\mathrm{out}}} q_{\mathrm{out}}^k(t)}, \tag{13}$$

where j and k are the counters for total L_{in} links flowing into the junction and L_{out} links extracting flow from the junction; $q_{\mathrm{in}}^j(t)$ and $q_{\mathrm{out}}^k(t)$ are the inflows and outflows from these links connected to the junction; $c_{\mathrm{in}}^j(t)$ is the concentration in the inflow solute; $q_i^{\mathrm{B_J}}(t)$ is the flow injected to the junction with concentration $c_i^{\mathrm{B_J}}(t)$ by booster station if located; and $q_i^{\mathrm{D_J}}(t)$ is demand.

Equation (13) can be applied at any arbitrary modeling time $t + \Delta t$ and according to Remark 1, with inflow entering the junction from a pipe, $c_{\mathrm{in}}^j(t)$ is taken as the concentration of the last segment of the pipe calculated in Equation (6b).

D. Mass balance at tanks

Tanks are used as storage facilities being filled during low demands to cover the high demands. Mass conservation in tanks assumes complete instantaneous mixing of all inflows, outflows, and stored water following the continuously stirred tank reactor (CSTR) model [43], [56].

Remark 2. *The CSTR model assumes that the stored water gets affected by any inflow or injected disinfection instantaneously. It also assumes that all outflows and stored volume have the same solute concentrations.*

The change in solute concentration at tanks is expressed by an ODE [57] as

$$\frac{\mathrm{d}(c^{\mathrm{TK}} V^{\mathrm{TK}})}{\mathrm{d}t} = \sum_{j \in L_{\mathrm{in}}} q_{\mathrm{in}}^j(t) c_{\mathrm{in}}^j(t) - \sum_{k \in L_{\mathrm{out}}} q_{\mathrm{out}}^k(t) c^{\mathrm{TK}}(t) + R_{\mathrm{MS}}^{\mathrm{TK}}(c^{\mathrm{TK}}(t)) V^{\mathrm{TK}}(t). \tag{14}$$

where V^{TK} is the stored water volume at time t; j and k are the counters for total L_{in} links flowing into the tank and L_{out} links extracting flow from the tank, these links are either part of the network or not; $q_{\mathrm{in}}^j(t)$ is inflow either from links or injected; $c_{\mathrm{in}}^j(t)$ is substance concentration in the inflow solute; $q_{\mathrm{out}}^k(t)$ is outflow by links connected to the tank; $c^{\mathrm{TK}}(t)$ is concentration at tank and any flow drawn from it according to Remark 2; and $R_{\mathrm{MS}}^{\mathrm{TK}}(c_i^{\mathrm{TK}})$ is the multi-species reaction rate expression explained briefly in Section V.

Applying Forward Euler discretization method [58] in Equation (14), the concentration at each Tank i in $c^{\mathrm{TK}}(t + \Delta t) \in \mathbb{R}^{n_{\mathrm{TK}}}$ is expressed as

$$V_i^{\mathrm{TK}}(t + \Delta t) c_i^{\mathrm{TK}}(t + \Delta t) = V_i^{\mathrm{TK}}(t) c_i^{\mathrm{TK}}(t) + \sum_{j \in L_{\mathrm{in}}} q_{\mathrm{in}}^j(t) c_{\mathrm{in}}^j(t) \Delta t - \sum_{k \in L_{\mathrm{out}}} q_{\mathrm{out}}^k(t) c_i^{\mathrm{TK}}(t) \Delta t + R_{\mathrm{MS}}^{\mathrm{TK}}(c_i^{\mathrm{TK}}(t)) V_i^{\mathrm{TK}}(t) \Delta t. \tag{15}$$

Yet, according to Remark 2, the effect of the injections from are instantaneous. resulting in an immediate effect [50]. Updating Equation (15) to be

$$V_i^{\text{TK}}(t + \Delta t)c_i^{\text{TK}}(t + \Delta t) = V_i^{\text{TK}}(t)c_i^{\text{TK}}(t) + \sum_{j \in L_{\text{in}}} q_{\text{in}}^j(t)c_{\text{in}}^j(t)\Delta t + V_i^{\text{B}_{\text{TK}}}(t + \Delta t)c_i^{\text{B}_{\text{TK}}}(t + \Delta t)$$
$$- \sum_{k \in L_{\text{out}}} q_{\text{out}}^k(t)c_i^{\text{TK}}(t)\Delta t + R_{\text{MS}}^{\text{TK}}(c_i^{\text{TK}}(t))V_i^{\text{TK}}(t)\Delta t, \quad (16)$$

where $V_i^{\text{B}_{\text{TK}}}(t + \Delta t)$ is the volume injected to the tank with concentration $c_i^{\text{B}_{\text{TK}}}(t + \Delta t)$ by booster station if located.

Similarly to junctions, $c_{\text{in}}^j(t)$ is taken as the concentration of the last segment of the pipe calculated in Equation (6b) for inflows entering the tank.

V. MULTI-SPECIES REACTION AND DECAY MODEL

Having modeled mass balance in various components as well as discretization of pipes, herein we showcase the chlorine decay and reaction models. Chlorine decay and reaction models have been broadly studied as stated in Section I. In Tab. II, we provide a list of different, widely used and studied, bulk decay and reaction models [32], [59].

Tab. II: Chlorine bulk decay and reaction models

Model	Model formulation	Model description	Citation
First-order	$\frac{dc}{dt} = -kc(t)$	the simplest model with a first-order decay rate	[12], [23], [60]
First-order with stable component	$\frac{dc}{dt} = -k(c(t) - c_{\text{L}})$	this model is built on the first-order one and adding a stable component that represents the limiting concentration for chlorine c_{L}	[60], [61]
Parallel first-order	$\frac{dc_1}{dt}\Big\|_{\text{fast}} = -k_{\text{fast}}c_1(t)$ $\frac{dc_2}{dt}\Big\|_{\text{slow}} = -k_{\text{slow}}c_2(t)$ $c_t(t) = c_1(t) + c_2(t)$	this model divides chlorine decay into two stages, and each stage decays independently, according to a first-order reaction with its own individual decay rate. The two stages are a slow and a fast decay dynamics	[60], [61]
Parallel second-order	$\frac{dc_{\text{F}}}{dt}\Big\|_{\text{fast}} = -k_{\text{fast}}c(t)c_{\text{F}}(t)$ $\frac{dc_{\text{S}}}{dt}\Big\|_{\text{slow}} = -k_{\text{slow}}c(t)c_{\text{S}}(t)$ $\frac{dc}{dt} = \frac{dc_{\text{F}}}{dt} + \frac{dc_{\text{S}}}{dt}$	In comparison to the previous model, the decay rate of chlorine is divided into fast and slow second-order dynamics where c_{F} and c_{s} are representing the concentrations of fast and slow reducing agents.	[28]–[30]
n^{th}-order	$\frac{dc}{dt} = -kc^n(t)$	the rate of reaction can generally be described as a power function of concentration	[60], [61]
n^{th}-order with stable component	$\frac{dc}{dt} = -k(c(t) - c_{\text{L}})c^{(n-1)}$	c_{L} puts a cut-off for the regular n^{th}-order model	[60], [62], [63]
Second-order with fictitious component	$\frac{dc}{dt} = -kc(t)\tilde{c}(t)$ $\frac{d\tilde{c}}{dt} = -kc(t)\tilde{c}(t)$	a fictitious reactant with concentration $\tilde{c}(t)$ is introduced in the system and accounts for all components reacting with chlorine	[14], [15], [64]
Second-order with multiple components	$\frac{dc_i}{dt} = -k_i c(t)\tilde{c}_i(t)$ $\frac{d\tilde{c}_i}{dt} = -k_i c(t)\tilde{c}_i(t)$ $\frac{dc}{dt} = \sum_i \frac{dc_i}{dt}$	separate rate is considered for every component reacting with chlorine, where $\tilde{c}_i$ is the concentration of the ith reacting substance	[12], [13]

In this paper, we consider a hybrid model that accounts for both chlorine decay with a constant rate and a reaction dynamic with a fictitious constituent. This model description has been used to detect different contamination events and their effect on chlorine residual in [65]. Dividing the model into decay and mutual reaction dynamics allows it to consider a substance with relatively different reaction rates than the decay rate, and allows the model to be less sensitive to the other reactants' concentrations. Hence, the model formulation is

$$R_{\text{MS}}(c(t)) = R_{\text{D}}(c(t)) + R_{\text{M}}(c(t)), \quad (17)$$

where $R_{\text{D}}(c(t))$ is the decay reaction expression; and $R_{\text{M}}(c(t))$ is the mutual reaction expression.

Note that the decay and reaction model (17) is assumed to occur in pipes and tanks only. We assume that there is no storage nor enough contact time in the other network components (e.g., pumps, valves, junctions). Likewise, the reservoir is considered a fixed source of chemicals with no decay or reactions [23], [43].

A. Decay reaction model

This decay model is a first-order model that depends only on chlorine concentration and constant decay rate. In addition to chlorine bulk decay models listed in Tab. II, chlorine decay also takes place at the pipe's wall. While disinfectant bulk fluid reactions occur in pipes and tanks, wall reactions only considered to take place in pipes. The bulk decay is a result of the reaction with natural organic matter, while pipe wall decay is due to the reaction with the materials released from its boundary layer [8], [11], [43]. Different wall reaction models have been developed/used varying between the zero-order and first-order models available in EPANET and its multi-species extension [?] and the developed and validated EXPBIO model by [66]. The wall decay zero-order model assumes constant decay while in the first-order model the decay rate reduces with the disinfactant's concentration, on the other hand, they have negative relation in the EXPBIO model. In this paper the focus is on the different bulk reaction models and we consider the first-order wall decay model in our simulation and when comparing our results with EPANET-MSX.

Hence, the chlorine decay reaction rates [26] for Pipe i and Tank j are

$$k_i^{\mathrm{P}} = k_b + \frac{2k_w k_f}{r_{\mathrm{P}_i}(k_w + k_f)}, \quad k_j^{\mathrm{TK}} = k_b, \tag{18}$$

where k_b is the bulk reaction 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_{P_i} is the pipe radius.

Eventually, the chlorine decay reaction expressions for segment s of Pipe i and Tank j are

$$R_{\mathrm{D}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s,t)) = -k_i^{\mathrm{P}} c_i^{\mathrm{P}}(s,t), \quad R_{\mathrm{D}}^{\mathrm{TK}}(c_j^{\mathrm{TK}}(t)) = -k_j^{\mathrm{TK}} c_j^{\mathrm{TK}}(t). \tag{19}$$

B. Mutual reaction model

The reaction model is introduced in a form two-species reaction dynamics between chlorine and a fictitious reactant. The model is covered by the general reaction model of any two substances [13], [32], [67]). Therefore, the mutual reaction model is expressed by a second-order nonlinear ODEs as

$$\frac{\mathrm{d}c}{\mathrm{d}t} = -k_r c(t)\tilde{c}(t), \quad \frac{\mathrm{d}\tilde{c}}{\mathrm{d}t} = -k_r c(t)\tilde{c}(t), \tag{20}$$

where $c(t), \tilde{c}(t)$ are the concentrations for chlorine and fictitious reactant respectively; and k_r is the mutual reaction rate between them.

The different techniques described in Section III to solve the advection-reaction PDE (2) for pipes need different approaches to include Equation (20). First, when adopting one of the EFD discritization techniques (see Sections III-A, III-B, and III-C), the mutual reaction model in Equation (20) (i.e., ODEs) is discretized applying the forward Euler discrtezation method resulting in

$$\begin{aligned} c(t+\Delta t) - c(t) &= -k_r \Delta t(c(t)\tilde{c}(t)), \\ \tilde{c}(t+\Delta t) - \tilde{c}(t) &= -k_r \Delta t(c(t)\tilde{c}(t)). \end{aligned} \tag{21}$$

Similarly, discretized reaction model in Equation (21) is used to formulate the reaction expression in tank Equation (16) along with Equation (19). Note that, for the techniques considered for pipes and for tanks they are discretized through the Eulerian discretization schemes in Sections III and IV-D. Thus, to avoid double discretizing, the left-hand-side of Equation (21) is not reconsidered, and we have

$$R_{\mathrm{M}}^{\mathrm{P}}(c_i^{\mathrm{P}}(s,t)) = -k_r c_i^{\mathrm{P}}(s,t)\tilde{c}_i^{\mathrm{P}}(s,t), \quad R_{\mathrm{M}}^{\mathrm{TK}}(c_j^{\mathrm{TK}}(t)) = -k_r \tilde{c}_j^{\mathrm{TK}}(t)c_j^{\mathrm{TK}}(t), \tag{22a}$$

$$R_{\mathrm{M}}^{\mathrm{P}}(\tilde{c}_i^{\mathrm{P}}(s,t)) = -k_r c_i^{\mathrm{P}}(s,t)\tilde{c}_i^{\mathrm{P}}(s,t), \quad R_{\mathrm{M}}^{\mathrm{TK}}(\tilde{c}_j^{\mathrm{TK}}(t)) = -k_r \tilde{c}_j^{\mathrm{TK}}(t)c_j^{\mathrm{TK}}(t). \tag{22b}$$

Second, for pipes when applying the Lagrangian-MoCs technique explained in Section III-D the following approach is followed. We start by showing that when the MoCs method is applied on a first-order decay model (Equation (19)) only, we get

$$\zeta(s, t+\Delta t) = c_i^{\mathrm{P}}(s,t)\exp(-k_i^{\mathrm{P}}\Delta t). \tag{23}$$

However, the system of equations in (20) can be turned into a pseudo-first-order reaction model by assuming that the other chemical's concentration in the previous time-step is not variable. The limitation of this approach is to have one chemical's initial concentration much greater than the other's concentration. Yet, the desired difference between initial concentrations depends directly on the mutual reaction rate [53], [59]. To that end, similar to Equation (23) we obtain

$$\begin{aligned} \zeta(s, t+\Delta t) &= c_i^{\mathrm{P}}(s,t)\exp(-(k_i^{\mathrm{P}} + k_r\tilde{c}_i^{\mathrm{P}}(s,t))\Delta t), \\ \tilde{\zeta}(s, t+\Delta t) &= \tilde{c}_i^{\mathrm{P}}(s,t)\exp(-k_r c_i^{\mathrm{P}}(s,t))\Delta t). \end{aligned} \tag{24}$$

By applying linear interpolation at time-step $t + \Delta t$; see Fig. 3, concentrations at segment s of Pipe i at time-step $t + \Delta t$ are expressed as

$$
\begin{aligned}
c_i^{\mathrm{P}}(s, t + \Delta t) &= \frac{v_i(t)\Delta t}{\Delta x}\zeta(s - 1, t + \Delta t) + (1 - \frac{v_i(t)\Delta t}{\Delta x})\zeta(s, t + \Delta t), \\
\tilde{c}_i^{\mathrm{P}}(s, t + \Delta t) &= \frac{v_i(t)\Delta t}{\Delta x}\tilde{\zeta}(s - 1, t + \Delta t) + (1 - \frac{v_i(t)\Delta t}{\Delta x})\tilde{\zeta}(s, t + \Delta t).
\end{aligned}
\tag{25}
$$

Since Courant number $\tilde{\lambda}_i(t)$ equals $v_i(t)\frac{\Delta t}{\Delta x_i}$, and after substituting Equation (24) to Equation (25), we have

$$
\begin{aligned}
c_i^{\mathrm{P}}(s, t + \Delta t) &= \tilde{\lambda}_i(t)c_i^{\mathrm{P}}(s - 1, t)\exp(-(k_i^{\mathrm{P}} + k_r\tilde{c}_i^{\mathrm{P}}(s - 1, t))\Delta t) + (1 - \tilde{\lambda}_i(t))c_i^{\mathrm{P}}(s, t)\exp(-(k_i^{\mathrm{P}} + k_r\tilde{c}_i^{\mathrm{P}}(s, t))\Delta t), \\
\tilde{c}_i^{\mathrm{P}}(s, t + \Delta t) &= \tilde{\lambda}_i(t)\tilde{c}_i^{\mathrm{P}}(s - 1, t)\exp(-k_r c_i^{\mathrm{P}}(s - 1, t))\Delta t) + (1 - \tilde{\lambda}_i(t))\tilde{c}_i^{\mathrm{P}}(s, t)\exp(-k_r c_i^{\mathrm{P}}(s, t))\Delta t)).
\end{aligned}
\tag{26}
$$

Note that in the Method of Characteristics, the first segment's concentration is taken as equal to the concentration of the upstream node. While the last segment's concentration is calculated depending on a virtual forward projection of its concentration at the previous time-step and the projection of the previous segment.

VI. STATE-SPACE FORMULATION

The multi-species water quality models have been analyzed in Sections III, IV, and V-B and the equations for WDN components have been derived. In this section, the detailed derivation of a state-space representation of the system based on these equations is shown in this section. We note that each technique listed in Section III for pipes leads to a different formulation. That is, the general formulation is expressed in Equation (1) with $E_{11}(t) = E_{22}(t) = E(t)$. Moreover, $E(t)$ matrix is a non-singular matrix and invertible allowing $x(t + \Delta t)$ to be calculated for every time-step according to Equation (27). As such, Tab. III lists the formulation for each of the techniques which applies to any of the two chemicals. Differences in the techniques' approaches are reflected in some of the representation matrices — highlighted in a different color for each technique.

$$
\begin{bmatrix} x_1(t + \Delta t) \\ x_2(t + \Delta t) \end{bmatrix} = \begin{bmatrix} E'(t)A_{11}(t) & 0 \\ 0 & E'(t)A_{22}(t) \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} E'(t)B_{11}(t) & 0 \\ 0 & E'(t)B_{22}(t) \end{bmatrix}\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \begin{bmatrix} E'(t) & 0 \\ 0 & E'(t) \end{bmatrix}f(x_1, x_2, t).
\tag{27}
$$

In these formulations, each of the chemicals $x_i(t)$ collects the concentrations in nodes and links of the network:

$$
x_i(t) := \{c_i^{\mathrm{N}}(t), c_i^{\mathrm{L}}(t)\} = \{c_i^{\mathrm{R}}(t), c_i^{\mathrm{J}}(t), c_i^{\mathrm{TK}}(t), c_i^{\mathrm{M}}(t), c_i^{\mathrm{P}}(1, t), \ldots, c_i^{\mathrm{P}}(s_{\mathrm{L}}, t), c_i^{\mathrm{V}}(t)\},
$$

where $x(t) \in \mathbb{R}^{n_x}$; n_x is the total number of nodes and links including pipes segments: $n_x = n_{\mathrm{R}} + n_{\mathrm{J}} + n_{\mathrm{TK}} + n_{\mathrm{M}} + n_{\mathrm{P_s}} + n_{\mathrm{V}}$.

This state-space formulation is updated every water quality time-step within a hydraulic time-step to reflect parameters and variables change. However, the system dimensions depend on how many segments each pipe is divided into, the number of segments depends on the water quality time-step and velocity (refer to Equation (3)) and the water quality time-step is fixed throughout the whole simulation. On the other hand, velocities change from hydraulic time-step to another for a system in a dynamic state (i.e., changing demands and flows). Consequently, in a dynamic system, with every hydraulic time-step applying (3) gives a different number of segments for each pipe which leads to different system dimensions.

Remark 3. *For a specific network, model dimensions change with changing the hydraulic parameters and variables (e.g., demands, levels, diameters, etc.) leading to different concentrations evolution throughout the network.*

Therefore, to be able to apply our model for an extended hydraulic period, we fix the number of segments for each pipe i according to the following approach:

$$
s_i = \left\lfloor \frac{L_i}{\max_{t \in [0, T_s]}(v_i(t))\Delta t} \right\rfloor, \quad \Delta x = \frac{L_i}{s_i},
\tag{28a}
$$

for $v_i(t) \leq \max_{t \in [0, T_s]}(v_i(t))$,

$$
0 \leq \tilde{\lambda}_i(t) = v_i(t)\frac{\Delta t}{\Delta x_i} \leq 1.
\tag{28b}
$$

To that end, the state-space formulation process for multi-species water quality dynamics in WDN is described in Algorithm 1. WDN topology and hydraulics throughout the simulation period are the algorithm inputs with an initialization step that constructs the numerical grid in time and space. Afterwards, within every hydraulic time-step parameters are updated and the state-space representation is formulated for each water quality time-step according to Tab. III and the governing equations in Sections III, IV, and V.

Algorithm 1: State-space formulation for multi-species water quality dynamics in WDN

Input: WDN topology, components' charachtristics, and hydraulics parameters (velocities, flow rates, and flow directions)
Output: State-space representation for the WDN for Techniques 1, 2, 3, and 4

1 Initialization
2   Choose Δt
3   For each pipe i in set $\mathcal{P}$, calculate s_i and Δx_i by applying (28)
4   Obtain initial concentrations for both chemicals and reaction rates
5 Proceed
6 foreach *Hydraulic time-step (t_H)* **do**
7   Run hydraulic simulation and update velocities and flow rates
8   Update flow directions for each pipe and order segments accordingly
9   Update Courant number $\tilde{\lambda}_i(t)$ (28b) for set $\mathcal{P}$
10   **foreach** $t \leq t_\mathrm{H}$ **do**
11   **for** *both chemicals* **do**
12   Construct $A_\mathrm{M}^\mathrm{R}, A_\mathrm{M}^\mathrm{J}$ by applying (11) for set $\mathcal{M}$
13   Construct $A_\mathrm{V}^\mathrm{J}, A_\mathrm{V}^\mathrm{TK}$ by applying (12) for set $\mathcal{V}$
14   Construct $A_\mathrm{J}^\mathrm{J}, A_\mathrm{J}^\mathrm{M}, A_\mathrm{J}^\mathrm{V}, A_\mathrm{J}^\mathrm{P}$ and B_J by applying (13) for set $\mathcal{J}$
15   Obtain reaction trems reflecting decay R_D^TK and R_D^P via (18) for chlorine and no decay for the other reactant
16   Construct $A_\mathrm{TK}^\mathrm{TK}, A_\mathrm{TK}^\mathrm{P}$, and B_TK by applying (16) for set $\mathcal{T}$
17   Construct $f(c^\mathrm{TK}(t), \tilde{c}^\mathrm{TK}(t))$ via (22) for set $\mathcal{T}$
18   **if** *Applying Lax-Wendroff scheme (Tech. 1)* **then**
19   Construct $A_\mathrm{P}^\mathrm{R}, A_\mathrm{P}^\mathrm{J}, A_\mathrm{P}^\mathrm{TK}$, and A_P^P by applying (4) and (6) for set $\mathcal{P}$
20   Construct $f(c^\mathrm{P}(t), \tilde{c}^\mathrm{P}(t))$ via (22) for set $\mathcal{P}$
21   **else if** *Applying Backward Euler scheme (Tech. 2)* **then**
22   Construct $E_\mathrm{P}^\mathrm{R}, E_\mathrm{P}^\mathrm{J}, E_\mathrm{P}^\mathrm{TK}, E_\mathrm{P}^\mathrm{P}$ and A_P^P by applying (8) for set $\mathcal{P}$
23   Construct $f(c^\mathrm{P}(t), \tilde{c}^\mathrm{P}(t))$ via (22) for set $\mathcal{P}$
24   **else if** *Applying Crank-Nicolson scheme (Tech. 3)* **then**
25   Construct $E_\mathrm{P}^\mathrm{R}, E_\mathrm{P}^\mathrm{J}, E_\mathrm{P}^\mathrm{TK}, E_\mathrm{P}^\mathrm{P}, A_\mathrm{P}^\mathrm{R}, A_\mathrm{P}^\mathrm{J}, A_\mathrm{P}^\mathrm{TK}$, and A_P^P by applying (9) for set $\mathcal{P}$
26   Construct $f(c^\mathrm{P}(t), \tilde{c}^\mathrm{P}(t))$ via (22) for set $\mathcal{P}$
27   **else**
28   *Applying Method of Characteristics (Tech. 4)*:
29   Construct $A_\mathrm{P}^\mathrm{R}, A_\mathrm{P}^\mathrm{J}, A_\mathrm{P}^\mathrm{TK}$, and A_P^P by applying (26) for set $\mathcal{P}$
30   **end if**
31   **end for**
32   Formulate the state-space representation according to Tab. III
33   Calculate concentrations for both chemicals by applying state-space representation (27)
34 $t = t + \Delta t$
**35**   **end foreach**
36 end foreach

Tab. III: State-space representation with different advection-reaction discretization techniques[a]

All four techniques are expressed in the general state-space form:

$$E(t)\,x(t+\Delta t) = A(t)\,x(t) + B(t)\,u(t) + f(x_1,x_2,t)$$

Common vectors and blocks (techniques 1–4):

$$x(t) = \begin{bmatrix} c^{R} \\ c^{J} \\ c^{TK} \\ \hline c^{M} \\ c^{V} \\ c^{P} \end{bmatrix}, \qquad
u(t) = \begin{bmatrix} c^{B_R} \\ c^{B_J} \\ c^{B_{TK}} \end{bmatrix}, \qquad
B(t) = \begin{bmatrix} O_{n_R} & B_J & B_{TK} \\ \hline & O_{n_L\times n_N} & \end{bmatrix}$$

Tech. 1 — Lax–Wendroff scheme

$$E(t) = I_{n_x}$$

$$A(t) = \begin{bmatrix} I_{n_R} & & & & & \\ & A_J^{J} & & & & \\ & & A_{TK}^{TK} & & & \\ \hline A_M^{R} & A_M^{J} & & A_M^{M} & & \\ & A_V^{J} & A_V^{TK} & & A_V^{V} & \\ A_P^{R} & A_P^{J} & A_P^{TK} & A_P^{M} & A_P^{V} & A_P^{P} \end{bmatrix}$$

$$f(x_1,x_2,t) = \begin{bmatrix} O_{n_R} \\ O_{n_J} \\ f(c^{TK}(t), \tilde{c}^{TK}(t)) \\ \hline O_{n_M} \\ O_{n_V} \\ f(c^{P}(t), \tilde{c}^{P}(t)) \end{bmatrix}$$

Tech. 2 — Backward Euler scheme

$$E(t) = \begin{bmatrix} I_{n_L\times n_L} & & & O_{n_L\times n_N} \\ & I_{n_M} & & \\ & & I_{n_V} & \\ E_P^{R} \;\; E_P^{J} \;\; E_P^{TK} & & & E_P^{P} \end{bmatrix}$$

$$A(t) = \begin{bmatrix} I_{n_R} & & & & & \\ & A_J^{J} & & & & \\ & & A_{TK}^{TK} & & & \\ \hline A_M^{R} & A_M^{J} & & A_M^{M} & & \\ & A_V^{J} & A_V^{TK} & & A_V^{V} & \\ A_P^{R} & A_P^{J} & A_P^{TK} & A_P^{M} & A_P^{V} & A_P^{P} \end{bmatrix}$$

$$f(x_1,x_2,t) = \begin{bmatrix} O_{n_R} \\ O_{n_J} \\ f(c^{TK}(t), \tilde{c}^{TK}(t)) \\ \hline O_{n_M} \\ O_{n_V} \\ f(c^{P}(t), \tilde{c}^{P}(t)) \end{bmatrix}$$

Tech. 3 — Crank–Nicolson scheme

$$E(t) = \begin{bmatrix} I_{n_L\times n_L} & & & O_{n_L\times n_N} \\ & I_{n_M} & & \\ & & I_{n_V} & \\ E_P^{R} \;\; E_P^{J} \;\; E_P^{TK} & & & E_P^{P} \end{bmatrix}$$

$$A(t) = \begin{bmatrix} I_{n_R} & & & & & \\ & A_J^{J} & & & & \\ & & A_{TK}^{TK} & & & \\ \hline A_M^{R} & A_M^{J} & & A_M^{M} & & \\ & A_V^{J} & A_V^{TK} & & A_V^{V} & \\ A_P^{R} & A_P^{J} & A_P^{TK} & A_P^{M} & A_P^{V} & A_P^{P} \end{bmatrix}$$

$$f(x_1,x_2,t) = \begin{bmatrix} O_{n_R} \\ O_{n_J} \\ f(c^{TK}(t), \tilde{c}^{TK}(t)) \\ \hline O_{n_M} \\ O_{n_V} \\ f(c^{P}(t), \tilde{c}^{P}(t)) \end{bmatrix}$$

Tech. 4 — Method of Characteristics

$$E(t) = I_{n_x}$$

$$A(t) = \begin{bmatrix} I_{n_R} & & & & & \\ & A_J^{J} & & & & \\ & & A_{TK}^{TK} & & & \\ \hline A_M^{R} & A_M^{J} & & A_M^{M} & & \\ & A_V^{J} & A_V^{TK} & & A_V^{V} & \\ A_P^{R} & A_P^{J} & A_P^{TK} & A_P^{M} & A_P^{V} & A_P^{P} \end{bmatrix}$$

$$f(x_1,x_2,t) = \begin{bmatrix} O_{n_R} \\ O_{n_J} \\ f(c^{TK}(t), \tilde{c}^{TK}(t)) \\ \hline O_{n_M} \\ O_{n_V} \\ O_{n_P} \end{bmatrix}$$

[a] Tech. (1) Lax-Wendroff scheme, (2) Backward Euler scheme, (3) Crank-Nicolson scheme, and (4) Method of Characteristics.
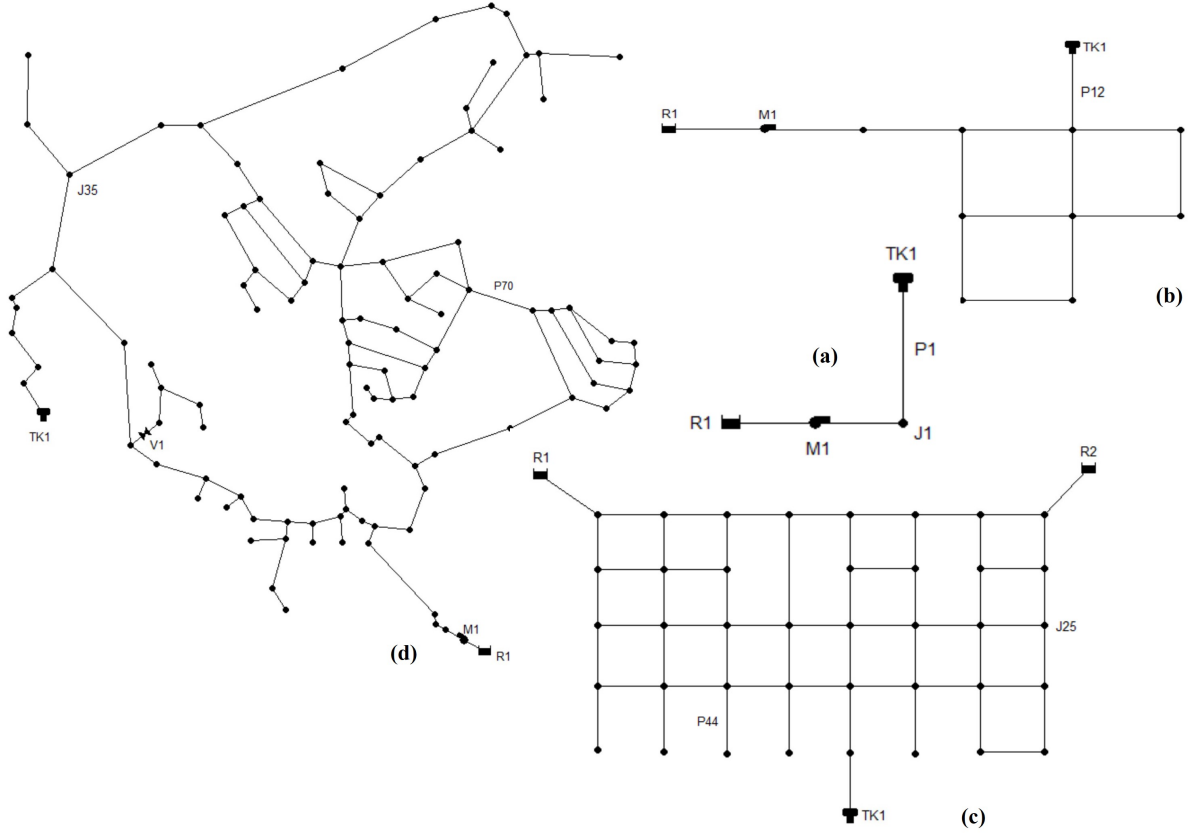
Fig. 4: Case studies' layouts: (a) Three-node network, (b) Net1, (c) Net2, and (d) FFCL-1.

VII. CASE STUDIES

In this section, several networks are used to validate the accuracy of the proposed two-species water quality dynamics and compare the results of different discretization schemes and their applicability. Four networks are considered, three-node, Net1, Net2, and FFCL-1 [23]. Note that the FFCL-1 network is modified accordingly based on the topology from EPANET to include a reservoir, a pump, and a valve. The topologies of tested networks are illustrated in Fig. 4 and the details of their components are listed in Tab. IV. The Three-node network is used as an example of how to build and apply our state-space formulation for all the techniques due to the simple and clear topology it has. Net1 and Net2 are used to show how the inclusion of a fictitious reactant affects chlorine concentrations and compare results from applying all techniques, and the scalability of the proposed method is tested on the FFCL-1 network.

Tab. IV: Test networks components

Network	Components Count					
	Reservoir	Tank	Junction	Pipe	Pump	Valve
Three-node	1	1	1	1	1	0
Net1	1	1	9	12	1	0
Net2	2	1	39	60	0	0
FFCL-1	1	1	108	121	1	1

A. *State-space formulation illustration using Three-node network*

In this section, the details of how the state-space formulation is derived are given with the Three-node network that has a simple topology consisting of four components (i.e., a reservoir, a pump, a junction, a pipe, and a tank) as listed in Tab. IV. We formulate the state-space representation using the Lax-Wendroff scheme (Tech. #1) explained in Section III-A and Appendix B covers the formulation for the rest of the techniques. We follow the approach of Algorithm 1 assuming that the initialization step has been carried out.

We assume that Pipe P1 of Three-node network shown in Figure 4 is divided into three segments $s_1 = 3$ to help readers understand our method; flow $q_1^{\mathrm{M}}(t)$ is pumped from Reservoir R1 to Junction J1 by Pump M1; flow in P1 is $q_1^{\mathrm{P}}(t)$; and demand at J1 is $q_1^{\mathrm{D_J}}(t)$. One chlorine booster station is located at TK1 injecting volume of $V^{\mathrm{B}}(t+\Delta t)$ with concentration $c^{\mathrm{B_{TK}}}(t+\Delta t)$. In this water quality model, we consider $u_1(t) = c^{\mathrm{B_{TK}}}(t+\Delta t)$. This can be explained physically as the amount of chlorine to be injected in the time step we are calculating at.

Tab. V summarizes the governing equations for two species that are chlorine and fictitious reactant. That is, the following state-space representation is formulated for x_1 as the vector of chlorine concentrations and x_2 as the vector of fictitious component concentrations.

$$
x_1(t+\Delta t) =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a_{24} & 0 & 0 & 0 \\
0 & 0 & a_{33} & 0 & 0 & 0 & a_{37} \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a_{52} & 0 & 0 & a_{55} & a_{56} & 0 \\
0 & 0 & 0 & 0 & a_{65} & a_{66} & a_{67} \\
0 & 0 & a_{73} & 0 & 0 & a_{76} & a_{77}
\end{bmatrix}
\begin{bmatrix}
c_1^{\mathrm{R}}(t) \\
c_1^{\mathrm{J}}(t) \\
c_1^{\mathrm{TK}}(t) \\
c_1^{\mathrm{M}}(t) \\
c_1^{\mathrm{P}}(1,t) \\
c_1^{\mathrm{P}}(2,t) \\
c_1^{\mathrm{P}}(3,t)
\end{bmatrix}
+
\begin{bmatrix}
0 \\
0 \\
b_3 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\left[c^{\mathrm{B_{TK}}}(t+\Delta t) \right] - k_r
\begin{bmatrix}
0 \\
0 \\
\tilde{c}_1^{\mathrm{TK}}(t) c_1^{\mathrm{TK}}(t) \\
0 \\
\tilde{c}_1^{\mathrm{P}}(1,t) c_1^{\mathrm{P}}(1,t) \\
\tilde{c}_1^{\mathrm{P}}(2,t) c_1^{\mathrm{P}}(2,t) \\
\tilde{c}_1^{\mathrm{P}}(3,t) c_1^{\mathrm{P}}(3,t)
\end{bmatrix},
$$

$$
x_2(t+\Delta t) =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \tilde{a}_{24} & 0 & 0 & 0 \\
0 & 0 & \tilde{a}_{33} & 0 & 0 & 0 & \tilde{a}_{37} \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \tilde{a}_{52} & 0 & 0 & \tilde{a}_{55} & \tilde{a}_{56} & 0 \\
0 & 0 & 0 & 0 & \tilde{a}_{65} & \tilde{a}_{66} & \tilde{a}_{67} \\
0 & 0 & \tilde{a}_{73} & 0 & 0 & \tilde{a}_{76} & \tilde{a}_{77}
\end{bmatrix}
\begin{bmatrix}
\tilde{c}_1^{\mathrm{R}}(t) \\
\tilde{c}_1^{\mathrm{J}}(t) \\
\tilde{c}_1^{\mathrm{TK}}(t) \\
\tilde{c}_1^{\mathrm{M}}(t) \\
\tilde{c}_1^{\mathrm{P}}(1,t) \\
\tilde{c}_1^{\mathrm{P}}(2,t) \\
\tilde{c}_1^{\mathrm{P}}(3,t)
\end{bmatrix}
- k_r
\begin{bmatrix}
0 \\
0 \\
\tilde{c}_1^{\mathrm{TK}}(t) c_1^{\mathrm{TK}}(t) \\
0 \\
\tilde{c}_1^{\mathrm{P}}(1,t) c_1^{\mathrm{P}}(1,t) \\
\tilde{c}_1^{\mathrm{P}}(2,t) c_1^{\mathrm{P}}(2,t) \\
\tilde{c}_1^{\mathrm{P}}(3,t) c_1^{\mathrm{P}}(3,t)
\end{bmatrix},
$$

where

$$
a_{24} = \tilde{a}_{24} = \frac{q_1^{\mathrm{M}}(t+\Delta t)}{q_1^{\mathrm{D_J}}(t+\Delta t) + q_1^{\mathrm{P}}(t+\Delta t)},
$$

$$
a_{33} = \frac{(1 - k_1^{\mathrm{TK}}\Delta t)V_1^{\mathrm{TK}}(t) - q_1^{\mathrm{D_{TK}}}(t)\Delta t}{V_1^{\mathrm{TK}}(t+\Delta t)},
$$

$$
\tilde{a}_{33} = \frac{V_1^{\mathrm{TK}}(t) - q_1^{\mathrm{D_{TK}}}(t)\Delta t}{V_1^{\mathrm{TK}}(t+\Delta t)},
$$

$$
a_{37} = \tilde{a}_{37} = \frac{q_1^{\mathrm{P}}(t)\Delta t}{V_1^{\mathrm{TK}}(t+\Delta t)}
$$

$$
a_{52} = \tilde{a}_{52} = a_{65} = \tilde{a}_{65} = a_{76} = \tilde{a}_{76} = \underline{\lambda}_1(t),
$$

$$
a_{55} = a_{66} = a_{77} = \lambda_1(t) - k_1^{\mathrm{P}}\Delta t,
$$

$$
\tilde{a}_{55} = \tilde{a}_{66} = \tilde{a}_{77} = \lambda_1(t),
$$

$$
a_{56} = \tilde{a}_{56} = a_{67} = \tilde{a}_{67} = a_{73} = \tilde{a}_{73} = \overline{\lambda}_1(t),
$$

$$
b_3 = \frac{V^{\mathrm{B}}(t+\Delta t)}{V_1^{\mathrm{TK}}(t+\Delta t)}.
$$

Tab. V: Water quality model equations for the three-node example[a]

Network Component	Chlorine	Fictitious Reactive Component
R1	$c_1^{\mathrm{R}}(t + \Delta t) = c_1^{\mathrm{R}}(t)$	
M1	$c_1^{\mathrm{M}}(t + \Delta t) = c_1^{\mathrm{R}}(t + \Delta t) = c_1^{\mathrm{R}}(t)$	
J1	$c_1^{\mathrm{J}}(t + \Delta t) = \dfrac{q_1^{\mathrm{M}}(t+\Delta t)c_1^{\mathrm{M}}(t+\Delta t)}{q_1^{\mathrm{D_J}}(t+\Delta t)+q_1^{\mathrm{P}}(t+\Delta t)} = \dfrac{q_1^{\mathrm{M}}(t+\Delta t)c_1^{\mathrm{R}}(t)}{q_1^{\mathrm{D_J}}(t+\Delta t)+q_1^{\mathrm{P}}(t+\Delta t)}$	
P1[b]	$c_1^{\mathrm{P}}(1, t + \Delta t) = \underline{\lambda}_1(t)c_1^{\mathrm{J}}(t) + \lambda_1(t)c_1^{\mathrm{P}}(1, t) + \overline{\lambda}_1(t)c_i^{\mathrm{P}}(2, t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_i^{\mathrm{P}}(1, t))\Delta t$ $c_1^{\mathrm{P}}(2, t + \Delta t) = \underline{\lambda}_1(t)c_1^{\mathrm{P}}(1, t) + \lambda_1(t)c_1^{\mathrm{P}}(2, t) + \overline{\lambda}_1(t)c_1^{\mathrm{P}}(3, t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_1^{\mathrm{P}}(2, t))\Delta t$ $c_1^{\mathrm{P}}(3, t + \Delta t) = \underline{\lambda}_1(t)c_1^{\mathrm{P}}(2, t) + \lambda_1(t)c_1^{\mathrm{P}}(3, t) + \overline{\lambda}_1(t)c_1^{\mathrm{TK}}(t) + R_{\mathrm{MS}}^{\mathrm{P}}(c_1^{\mathrm{P}}(3, t))\Delta t$	
TK1[b]	$V_1^{\mathrm{TK}}(t + \Delta t)c_1^{\mathrm{TK}}(t + \Delta t) = V_1^{\mathrm{TK}}(t)c_1^{\mathrm{TK}}(t)$ $+ q_1^{\mathrm{P}}(t)c_1^{\mathrm{P}}(3, t)\Delta t + V^{\mathrm{B}}(t + \Delta t)c^{\mathrm{B_{TK}}}(t + \Delta t)$ $+ R_{\mathrm{MS}}^{\mathrm{TK}}(c_1^{\mathrm{TK}}(t))V_1^{\mathrm{TK}}(t)\Delta t$	$V_1^{\mathrm{TK}}(t + \Delta t)\tilde{c}_1^{\mathrm{TK}}(t + \Delta t) = V_1^{\mathrm{TK}}(t)\tilde{c}_1^{\mathrm{TK}}(t)$ $+ q_1^{\mathrm{P}}(t)\tilde{c}_1^{\mathrm{P}}(3, t)\Delta t + R_{\mathrm{MS}}^{\mathrm{TK}}(\tilde{c}_1^{\mathrm{TK}}(t))V_1^{\mathrm{TK}}(t)\Delta t$

[a] For equations applicable for both chemicals, c is the concentration for any of the two species. Otherwise, c is chlorine concentration while $\tilde{c}$ is fictitious reactant concentration.
[b] Reaction expressions are according to Equations (19) and (22).
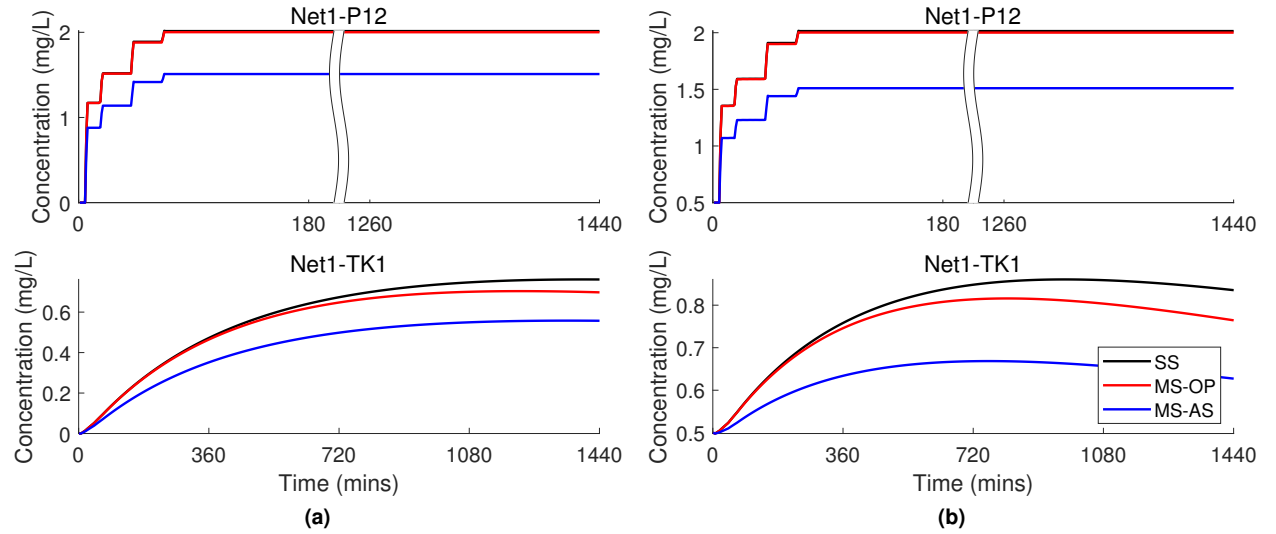
Fig. 5: Chlorine concentrations at Tank TK1 and Pipe P12 in Net1 in case of (a) zero initial concentrations for both chemicals (b) 0.5 mg/L chlorine concentration, in all components except for Reservoir R1 with source concentrations.

*SS: Single-species model, MS-OP: Multi-species model with a fictitious reactant with slow reaction rate (e.g., pro-oxonic organo-phosphate), MS-AS: Multi-species model with a fictitious reactant with relatively rapid reaction rate (e.g., sodium arsenite). Multi-species model is simulated by applying Tech. #1 - Lax-Wendroff scheme

B. NDE model implementation with different discretization techniques, validity, and complexity

In this section, we test the state-space representation formulated using the considered discretization techniques on different case studies. Then, we compare the results with the ones obtained from EPANET and its multi-species simulation extension EPANET-MSX. All parameters are matched for our NDE and EPANET to have a reliable comparison. Codes are run on MATLAB and EPANET results are obtained using its toolkit. Bare in mind, EPANET-MSX is not yet inherent in EPANET software. One way to run it is on the Command Prompt in Windows, which results in a form of .rpt-extended file. This file extension can be opened by any text editor. Yet, to do any further analysis/modeling, the procedure to read the results needs to follow a certain complex format. Another way to run it is through the toolkit on MATLAB. Both approaches take seconds to run the model. However, using the toolkit to get these results takes a longer time. To that end, the run time results listed in the following subsections are the actual total time needed to run and extract the results of the simulation. On the other hand, according to Remark 3 the dimension of the model depends on the hydraulics within the system for the same network. Thus, we list the number of states in the model for each network simulation.

Additionally, pipes are divided into a number of segments with a fixed length per pipe, thus, an average concentration over pipe segments represents the simulation result for each pipe. The same approach is followed in EPANET-MSX but with variable-sized segments per pipe.

1) Multi-species model vs. single-species reaction and transport models: Firstly, we showcase the impact of the inclusion of the fictitious reactant by simulating three cases using Lax-Wendroff scheme (Tech. #1): no fictitious reactant included, a
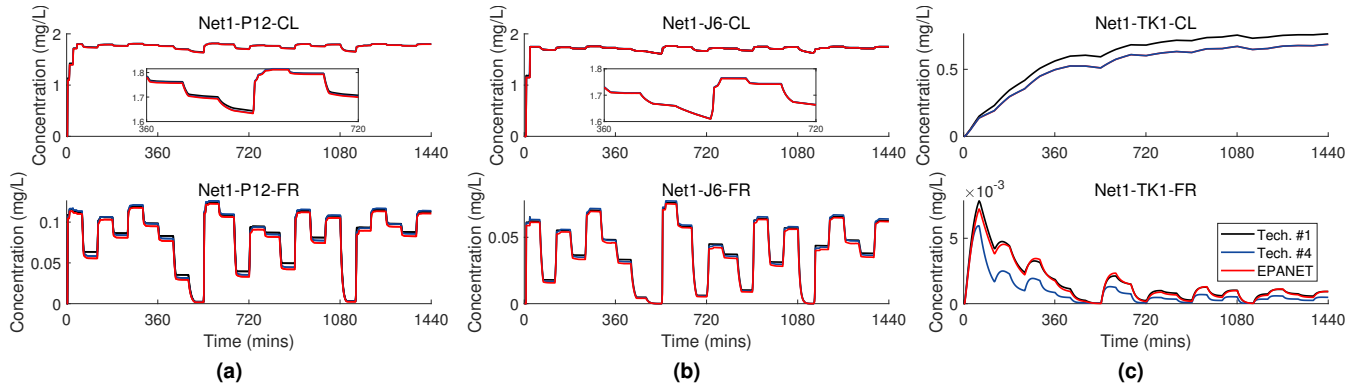
Fig. 6: Simulation Results at (a) Pipe P12 and (b) Junction J6 (c) Tank TK1 in Net1 with reservoir source concentrations of 2 and 0.3 mg/L, and zero initial concentrations at the rest of the components. Simulation is performed over 24hrs period with demand pattern changing every 1hr by applying Tech. #1 - Lax-Wendroff scheme and Tech. # 4 - Method of Characteristics.

fictitious reactant with a slow reaction rate (e.g., pro-oxonic organo-phosphate), and a fictitious reactant with relatively rapid reaction rate (e.g., sodium arsenite) [65]. Simulations are performed on Net1 a constant source of both chemicals at Reservoir R1 of 2 mg/L and 0.3 mg/L with chlorine and fictitious reactant, respectively. The water quality time-step is considered to be $\Delta t = 1$ sec and pipes are divided into a different number of segments satisfying the stability condition. Moreover, we consider two scenarios for initial chlorine concentrations in networks components other than Reservoir R1, zeros (Fig. 5a) and 0.5 mg/L (Fig. 5b). Results from both scenarios confirm how chlorine concentrations can be affected by the presence of another substance. Furthermore, depending on the substance and its reaction rate with chlorine, it can consume chlorine slowly or rapidly.

2) Discretization schemes/techniques performance: Each of the four discretization schemes follows a different approach as explained. Henceforward, we investigate the behavior of each and compare the results with EPANET-MSX. We apply each of the schemes on Net1 with a fixed chlorine concentration of 2 mg/L and 0.3 mg/L for the fictitious reactive substance at Reservoir R1. A dynamic hydraulic model is considered that resulted in a number of segments of pipes ranging between 44 to 301 segments. Fig. 6 depicts the results of the L-W scheme (Tech. #1), MoCs (Tech. #4), and EPANET with patterned demand and zero initial concentrations for chlorine in the first row and the fictitious reactant in the second one. MoCs gives almost exact chlorine concentrations for all components as EPANET-MSX but relatively different for fictitious reactant with cases of low concentrations (e.g., TK1). On the contrary, results of the L-W scheme are closer in case of low concentrations of the fictitious reactant and chemicals concentrations in pipes and junctions, while maintaining a maximum relative difference of 9.8% for chlorine in TK1. In general, we can see that both methods give a good representation of both chemicals' evolution, follow similar behavior and give close results to EPANET-MSX.
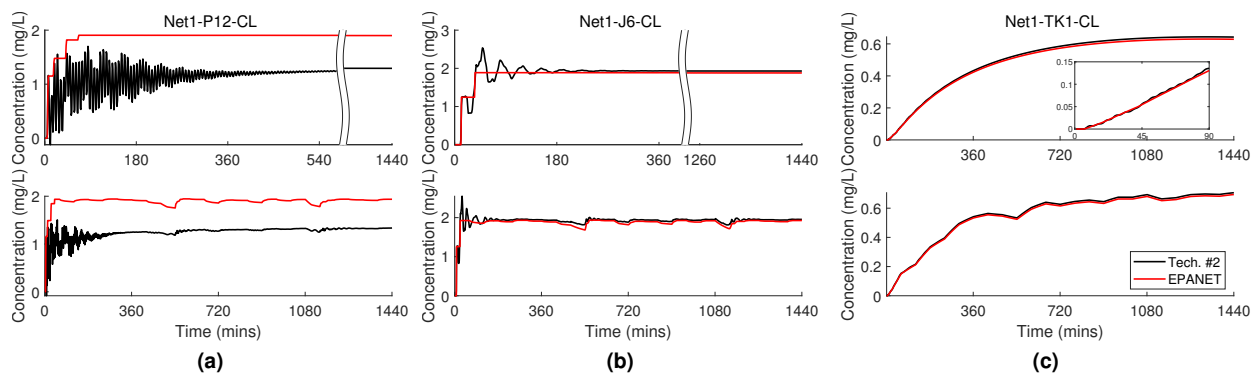

Fig. 7: Chlorine concentrations at (a) Pipe P12, (b) Junction J6, and (c) Tank TK1 in Net1 under static- and dynamic-states by applying Tech. #2 - Backward Euler scheme.

On the other hand, Fig. 7 shows results from simulating chlorine using the Backward Euler scheme (Tech #2) with the same setting on Net1 within static-state in the first row and dynamic hydraulics in the second one for Pipe P12, Junction J6, and Tank TK1. It is obvious that it generates numerical dispersion. Although Tech. #2 is unconditionally stable, a relatively small time step is required to achieve accurate solutions. Otherwise, it generates time- step-dependent numerical dispersion. The numerical dispersion drops out as the static condition is reached. Moreover, with dynamic hydraulics it starts to show at the transition

phase between one hydraulic step to the next one. Furthermore, this dispersion affects both chemicals simultaneously and each of them directly affects the other one implicitly through their mutual reaction which effect appears in Pipe P12. On the other hand, Tank TK1 is affected by the last segment in Pipe P12 which does not participate in its concentration/volume leading to small dispersion (see Fig. 7b). However, This dispersion leads to misleading unrealistic overestimated/underestimated results that controllers act depends on. Whilst, to avoid dispersion a significantly smaller time-step compared to other methods is needed which increases the run time exponentially.

As for Crank-Nicolson scheme (Tech. #3), Fig. 8 shows chemicals concentrations evolution in Net1 with the aforementioned fixed setting. Results exhibit high oscillation reaching illogical concentrations values for both chemicals. That is, this scheme which follows an explicit-implicit discretization approach is sensitive to sharp initial transients that cause an oscillatory response. This oscillatory affects both chemicals separately and they depend on each other which results in a dramatically large error. To test this sensitivity, the scheme is applied on Net1 with only single-species chlorine first-order decay dynamics and with initial conditions of 0.3 mg/L for all components other than the source. Results from this scenario are shown in Fig. 9. We can see that the oscillation effect is reduced but not avoided completely. In fact, sharp transitions in concentrations take place repeatedly during dynamic hydraulic and water quality simulation with demands fluctuation and chemicals injections/intrusion. Comprehensively, it is concluded that the Backward Euler scheme (Tech. #2) and Crank-Nicolson scheme (Tech. #3) put stiff limitations to be used in water quality modeling including extremely small time-step to avoid numerical dispersion and no sharp transitions in chemical concentrations to eliminate the oscillation action.
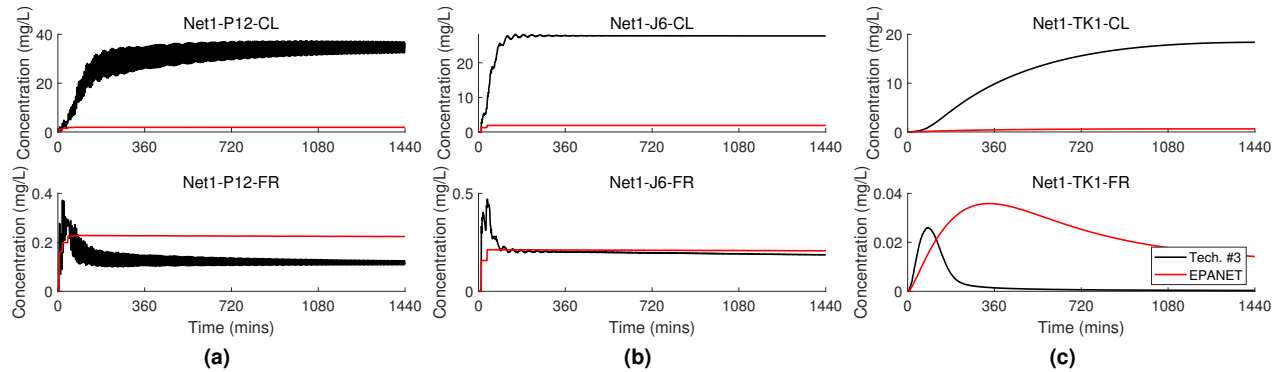

Fig. 8: Simulation results for chlorine and the fictitious reactant on Net1 by applying Tech. #3 - Crank-Nicolson scheme at (a) Pipe P12, (b) Junction J6, and (c) Tank TK1.
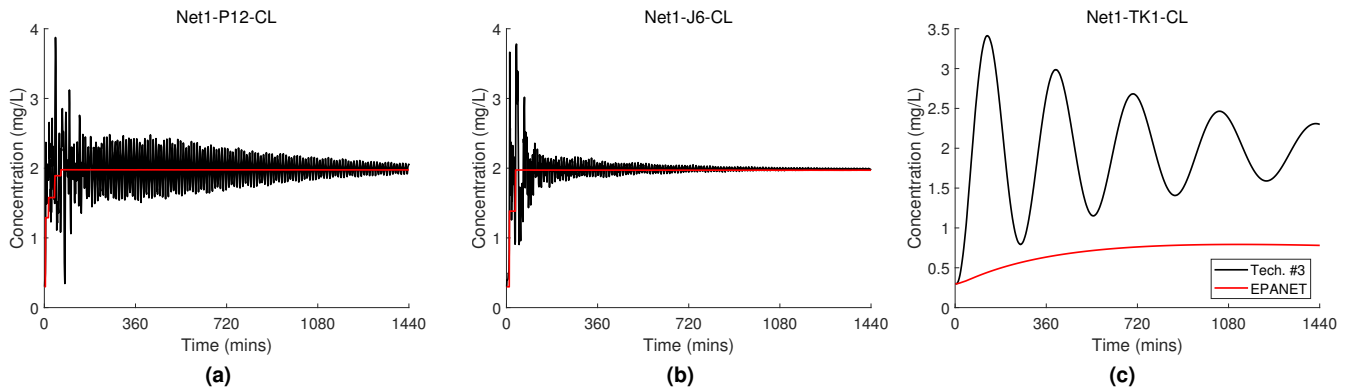
Fig. 9: Simulation results for chlorine on Net1 by applying Tech. #3 - Crank-Nicolson scheme on single-species first-order model with initial chlorine concentration of 0.3 mg/L and source concentration of 2 mg/L at (a) Pipe P12, (b) Junction J6, and (c) Tank TK1.

L-W scheme and MoCs are applied on networks with different scales under various hydraulic states to test their performance. First, we apply both techniques on Net2 with chlorine and fictitious reactant concentrations of 2 and 0.3 mg/L at both Reservoirs R1 and R2. Results in Fig. 10 imply the reliable performance of both techniques. Yet, the L-W scheme results in numerical dispersion at sharp fronts in Junction J25 as shown in Fig. 10b. This dispersion is likely to take place when the Courant number $\tilde{\lambda}$ is small close to zero. Such a scenario occurs when fixing the number of segments via Equation (28) for a system with a high range of velocities resulted from extremely high or low hydraulic states (see Fig. 10c). However, this dispersion can be eliminated by dividing the simulation period to separate these extreme states and maintain the Courant number as close as possible to 1. First, the system dimensions are defined by velocities in the first phase then the system is resized according

to velocities in the following phase. Average concentration over pipe at the end of the first phase is used as the initial value for the next one.
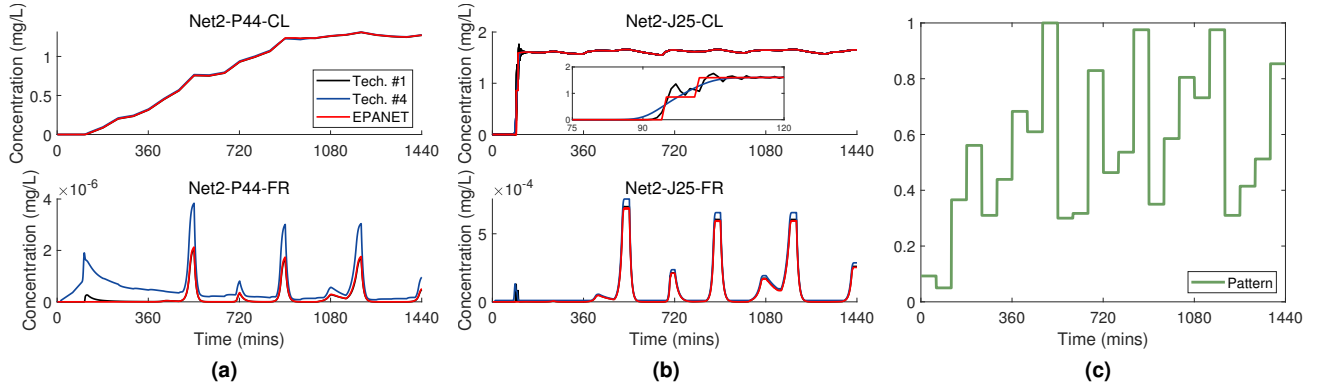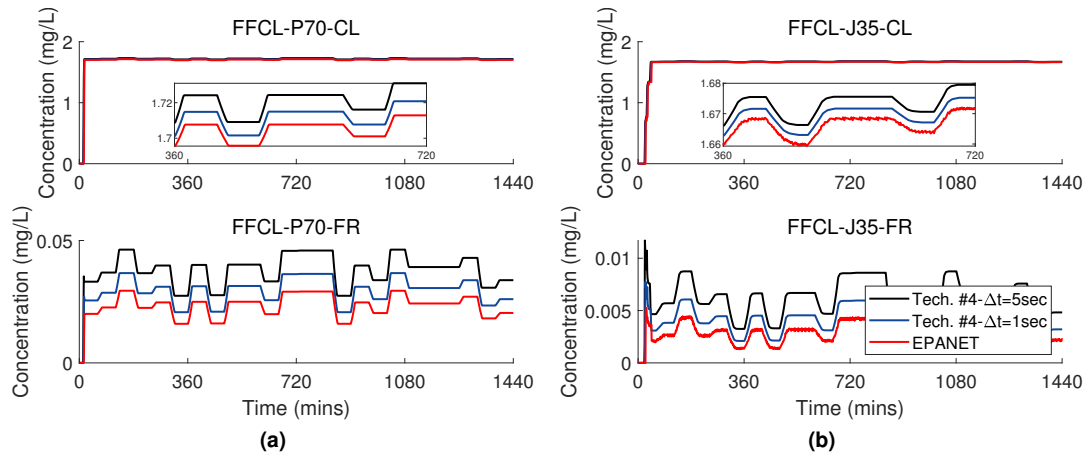

Fig. 10: Simulation results on Net2 by applying Tech. #1 - L-W scheme and Tech. #4 - MoCs at (a) Pipe P44 and (b) Junction J25 with demand pattern shown in (c).

Moreover, both techniques are applied on FFCL-1 with fixed concentrations of 2 and 0.3 mg/L at Reservoir R1 to test their scalability. To compare results, we calculate relative difference ratio between the applied technique and EPANET-MSX to be

$$Relative\ Difference\ (RD) = \frac{c_{\text{EPANET-MSX}} - c_{\text{model}}}{c_{EPANET\text{-}MSX}}.$$

The maximum relative difference ratio for the L-W scheme (Tech. #1) for both chlorine and fictitious reactant within these case studies ranges between 5-23%. However, the maximum relative difference ratio for the MoCs scheme (Tech. #4) varies between 1-11% for chlorine and between 7-33% for fictitious reactants. This difference is caused by the assumption made to transform the second-order dynamics into pseudo-first-order which affects the substance with smaller concentrations. Nonetheless, the error decreases with decreasing the water quality time-step as shown in Fig. 11.

Fig. 11: Simulation results for chlorine and the fictitious reactant on FFCL-1 by applying Tech. #4 - Method of Characteristics at (a) Pipe P70, and (b) Junction J35 with water quality time-steps of 5 and 1 sec.

Moreover, the effect of hydraulic parameters variability and its direct impact on system dimensions as stated in Remark 3 is tested. For each network, demands values and patterns are changed leading to different velocities and different number of states. Simulation run-time for each case study is listed in Tab. VI. Notice that the hydraulic simulation output almost does not affect EPANET AND EPANET-MSX run-time due to the fact that the time needed is for extracting the results. On the other hand, increasing number of states increases computational time, especially for Method of Characteristics (Tech. #4).

Tab. VI: Simulations run-time (secs)

Tab. VII: Water quality model equations for the three-node example[a]

Network	# States[a]	Tech. #1 - Lax-Wendroff scheme	Tech. #4 - Method of Characteristics	EPANET-MSX
Net1	4586	81.9	96.5	615
	6360	113	116	643
Net2	23680	466	483	745
	57186	1207.5	2180	763
FFCL-1	35370	611	1044	2092
	52454	1225	3929	2201

[a] States count is the summation of the number of states (i.e., $x_i(t)$) for both chemicals.

Tab. VIII: Techniques/methods applicability and limitations in multi-species water quality modeling

Technique/method	Applicability	Limitation
EPANET and EPANET-MSX	Widely used and EPANET-MSX allows full control to specify the reaction model to be considered	High run-time to obtain results even for small networks. Complexity to couple a control model with it
Tech. #1 - Lax-Wendroff Scheme	Gives a reliable results in a form of nonlinear state-space representation and traces the evolution of both chemicals	Conditionally stable. Stability should be checked throughout whole simulation period. According to networks topology and hydraulics parameters, may results in high numbers of state-space variables, hence, high run-time.
Tech. #2 - Backward Euler Scheme	Sensitive to simulation time-step causing dispersion that results in non-realistic results especially at the start of the simulation	To avoid numerical dispersion extremely small time-step is needed even for small- and mid-sized network leading to high run-time
Tech. #3 - Crank-Nicolson Scheme	Results in high oscillatory response to transients	These transitions take place repetitively in water quality modeling and control
Tech. #4 - Method of Characteristics	Depicts chemical evolution in WDN while overcoming the nonlinearity complexity	Presents the second-order reaction model in a form of pseudo-first order one resulting in approximation error that increases with the high difference between chemical concentrations, network's scale and hydraulic parameters change

VIII. CONCLUSION, PAPER'S LIMITATIONS, AND RECOMMENDATIONS FOR FUTURE WORK

This paper presents a full description of the formulation of multi-species water quality state-space representation. This description is provided and tested over a fixed time and space grid discretized using distinguishably based techniques. Three Eulerian Finite Difference schemes, Lax-Wendroff, Backward Euler, and Crank-Nicolson schemes, and one Lagrangian time-driven method, Method of Characteristics are considered for implementation and testing. Case studies show that Backward Euler and Crank-Nicolson schemes are strictly limited to avoid numerical dispersion and oscillation. On the contrary, the Lax-Wendroff scheme and Method of Characteristics reliably trace chlorine contractions within different networks with various scales. Yet, the Lax-Wendroff scheme's performance worsens with highly variable hydraulic dynamics, however, it is proposed to divide the simulation period to separate extremely low and high hydraulic states as mitigation to this problem. Tab. VIII lists techniques' applicability and limitations concluded from the research output.

An educational and theoretical value with a broader impact of this paper is providing a comprehensive framework and detailed example of how the state-space representation for water quality dynamics is formulated. The algorithm provided is easily editable to consider different scenarios and different chlorine reaction models.

However, the presented research has its limitations and recommended extensions. First, this paper tests numerically-driven discretization methods, however, more methods varying from Eulerian-, Lagrangian-, hybrid-based are available and can lead to a more accurate/limitation-free simulation of the model. Second, the multi-species model is valid for substances actively reacting with chlorine only. Also, in this paper we consider the first order wall reaction model, while applying other models including the zero-order and the validated EXPBIO models in the closed-form state space representation is worth investigating. Additionally, discretization methods result in large model dimensions that vary with the considered hydraulic parameter even for mid-sized networks. Subsequently, it is proposed for the authors' future work to perform a model order reduction taking

into consideration the nonlinearity associated with applying the Lax-Wendroff scheme for the multi-species dynamics. And to investigate and apply different control algorithms to maintain chlorine residuals within the desired standard range under multi-species water quality dynamics.

REFERENCES

[1] E. Ohanian, C. Mullin, and J. Orme, "Health effects of disinfectants and disinfection by-products: A regulatory perspective," *Water Chlorination Chem. Environ. Impact Health Effects*, vol. 6, no. 6, pp. 75–86, 1990.

[2] S. D. Richardson, M. J. Plewa, E. D. Wagner, R. Schoeny, and D. M. DeMarini, "Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research," *Mutation Research/Reviews in Mutation Research*, vol. 636, no. 1-3, pp. 178–242, 2007.

[3] G. Constable and B. Somerville, *A century of innovation: Twenty engineering achievements that transformed our lives*. Joseph Henry Press, 2003.

[4] D. M. DeMarini, "A review on the 40th anniversary of the first regulation of drinking water disinfection by-products," *Environmental and molecular mutagenesis*, vol. 61, no. 6, pp. 588–601, 2020.

[5] N. Ngwenya, E. J. Ncube, and J. Parsons, "Recent advances in drinking water disinfection: successes and challenges," *Reviews of environmental contamination and toxicology*, pp. 111–170, 2013.

[6] S. Tsitsifli and V. Kanakoudis, "Disinfection impacts to drinking water safety—a review," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 2, no. 11, 2018, p. 603.

[7] W. M. Grayman, "History of water quality modeling in distribution systems," in *WDSA/CCWI Joint Conference Proceedings*, vol. 1, 2018.

[8] L. A. Rossman and P. F. Boulos, "Numerical methods for modeling water quality in distribution systems: A comparison," *Journal of Water Resources planning and management*, vol. 122, no. 2, pp. 137–146, 1996.

[9] L. A. Rossman, P. F. Boulos, and T. Altman, "Discrete volume-element method for network water-quality models," *Journal of Water Resources Planning and Management*, vol. 119, no. 5, pp. 505–517, 1993.

[10] C. Liou and J. Kroon, "Modeling the propagation of waterborne substances in distribution networks," *Journal-American Water Works Association*, vol. 79, no. 11, pp. 54–58, 1987.

[11] H. Basha and L. Malaeb, "Eulerian–lagrangian method for constituent transport in water distribution networks," *Journal of Hydraulic Engineering*, vol. 133, no. 10, pp. 1155–1166, 2007.

[12] P. M. Jonkergouw, S.-T. Khu, D. A. Savic, D. Zhong, X. Q. Hou, and H.-B. Zhao, "A variable rate coefficient chlorine decay model," *Environmental science & technology*, vol. 43, no. 2, pp. 408–414, 2009.

[13] A. Jadas-Hécart, A. El Morer, M. Stitou, P. Bouillot, and B. Legube, "Chlorine demand of a treated water[modelisation de la demande en chlore d'une eau traitee]," *Water Research*, vol. 26, no. 8, pp. 1073–1084, 1992.

[14] R. M. Clark, "Chlorine demand and tthm formation kinetics: a second-order model," *Journal of environmental engineering*, vol. 124, no. 1, pp. 16–24, 1998.

[15] D. L. Boccelli, M. E. Tryby, J. G. Uber, and R. S. Summers, "A reactive species model for chlorine decay and thm formation under rechlorination conditions," *Water research*, vol. 37, no. 11, pp. 2654–2666, 2003.

[16] L. Monteiro, D. Figueiredo, S. Dias, R. Freitas, D. Covas, J. Menaia, and S. Coelho, "Modeling of chlorine decay in drinking water supply systems using epanet msx," *Procedia Engineering*, vol. 70, pp. 1192–1200, 2014.

[17] Z. Ohar and A. Ostfeld, "Optimal design and operation of booster chlorination stations layout in water distribution systems," *Water research*, vol. 58, pp. 209–220, 2014.

[18] A. Ostfeld and E. Salomons, "Conjunctive optimal scheduling of pumping and booster chlorine injections in water distribution systems," *Engineering optimization*, vol. 38, no. 03, pp. 337–352, 2006.

[19] P. F. Boulos, T. Altman, P.-A. Jarrige, and F. Collevati, "An event-driven method for modelling contaminant propagation in water networks," *Applied mathematical modelling*, vol. 18, no. 2, pp. 84–92, 1994.

[20] E. Blokker, J. Vreeburg, S. Buchberger, and J. Van Dijk, "Importance of demand modelling in network water quality models: a review," *Drinking Water Engineering and Science*, vol. 1, no. 1, pp. 27–38, 2008.

[21] V. G. Tzatchkov, A. A. Aldama, and F. I. Arreguin, "Advection-dispersion-reaction modeling in water distribution networks," *Journal of Water Resources Planning and Management*, vol. 128, no. 5, pp. 334–342, 2002.

[22] G. Munavalli and M. Mohan Kumar, "Multi-step eulerian method for multicomponent transport in water networks," *ISH Journal of Hydraulic Engineering*, vol. 11, no. 3, pp. 103–115, 2005.

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

[24] A. A. Abokifa, L. Xing, and L. Sela, "Investigating the impacts of water conservation on water quality in distribution networks using an advection-dispersion transport model," *Water*, vol. 12, no. 4, p. 1033, 2020.

[25] F. Shang, H. Woo, J. B. Burkhardt, and R. Murray, "Lagrangian method to model advection-dispersion-reaction transport in drinking water pipe networks," *Journal of Water Resources Planning and Management*, vol. 147, no. 9, p. 04021057, 2021.

[26] J. J. Vasconcelos and P. F. Boulos, *Characterization and modeling of chlorine decay in distribution systems*. American Water Works Association, 1996.

[27] I. Fisher, G. Kastl, and A. Sathasivan, "Evaluation of suitable chlorine bulk-decay models for water distribution systems," *Water research*, vol. 45, no. 16, pp. 4896–4908, 2011.

[28] ——, "A suitable model of combined effects of temperature and initial condition on chlorine bulk decay in water distribution systems," *Water research*, vol. 46, no. 10, pp. 3293–3303, 2012.

[29] Y.-H. Wang, Y.-H. Wu, Y. Du, Q. Li, Y. Cong, Z.-Y. Huo, Z. Chen, H.-W. Yang, S.-M. Liu, and H.-Y. Hu, "Quantifying chlorine-reactive substances to establish a chlorine decay model of reclaimed water using chemical chlorine demands," *Chemical Engineering Journal*, vol. 356, pp. 791–798, 2019.

[30] I. Fisher, G. Kastl, and A. Sathasivan, "A comprehensive bulk chlorine decay model for simulating residuals in water distribution systems," *Urban Water Journal*, vol. 14, no. 4, pp. 361–368, 2017.

[31] G. Munavalli and M. M. Kumar, "Dynamic simulation of multicomponent reaction transport in water distribution systems," *Water research*, vol. 38, no. 8, pp. 1971–1988, 2004.

[32] D. E. Helbling and J. M. VanBriesen, "Modeling residual chlorine response to a microbial contamination event in drinking water distribution systems," *Journal of Environmental Engineering*, vol. 135, no. 10, pp. 918–927, 2009.

[33] M. Allaire, H. Wu, and U. Lall, "National trends in drinking water quality violations," *Proceedings of the National Academy of Sciences*, vol. 115, no. 9, pp. 2078–2083, 2018.

[34] K. E. Furst, B. M. Pecson, B. D. Webber, and W. A. Mitch, "Tradeoffs between pathogen inactivation and disinfection byproduct formation during sequential chlorine and chloramine disinfection for wastewater reuse," *Water research*, vol. 143, pp. 579–588, 2018.

[35] K. N. Palansooriya, Y. Yang, Y. F. Tsang, B. Sarkar, D. Hou, X. Cao, E. Meers, J. Rinklebe, K.-H. Kim, and Y. S. Ok, "Occurrence of contaminants in drinking water sources and the potential of biochar for water quality improvement: A review," *Critical Reviews in Environmental Science and Technology*, vol. 50, no. 6, pp. 549–611, 2020.

[36] S. Mohan, G. Abhijith, and B. Aneesh, "Modeling chlorine response to uncontrolled contamination events in drinking water distribution systems," *Journal of Water Supply: Research and Technology-Aqua*, vol. 67, no. 8, pp. 834–845, 2018.

[37] J. B. Burkhardt, J. Szabo, S. Klosterman, J. Hall, and R. Murray, "Modeling fate and transport of arsenic in a chlorinated distribution system," *Environmental Modelling & Software*, vol. 93, pp. 322–331, 2017.

[38] S. Sharma and A. Bhattacharya, "Drinking water contamination and treatment techniques," *Applied Water Science*, vol. 7, no. 3, pp. 1043–1067, 2017.

[39] G. Abhijith and A. Ostfeld, "Modeling the response of nonchlorinated, chlorinated, and chloraminated water distribution systems toward arsenic contamination," *Journal of Environmental Engineering*, vol. 147, no. 10, p. 04021045, 2021.

[40] G. Munavalli and M. M. Kumar, "Optimal scheduling of multiple chlorine sources in water distribution systems," *Journal of water resources planning and management*, vol. 129, no. 6, pp. 493–504, 2003.

[41] I. Fisher, G. Kastl, F. Shang, and A. Sathasivan, "Framework for optimizing chlorine and byproduct concentrations in drinking water distribution systems." *Journal: American Water Works Association*, vol. 110, no. 11, 2018.

[42] S. Wang, A. F. Taha, and A. A. Abokifa, "How effective is model predictive control in real-time water quality regulation? state-space modeling and scalable control," *Water Resources Research*, vol. 57, no. 5, p. e2020WR027771, 2021.

[43] F. Shang, J. G. Uber, L. A. Rossman, and R. Janke, "Epanet multi-species extension user's manual," *Risk Reduction Engineering Laboratory, US Environmental Protection Agency, Cincinnati, Ohio*, 2008.

[44] Z. Wang, M. M. Polycarpou, J. G. Uber, and F. Shang, "Adaptive control of water quality in water distribution networks," *IEEE transactions on control systems technology*, vol. 14, no. 1, pp. 149–156, 2005.

[45] A. G. Seyoum and T. T. Tanyimboh, "Integration of hydraulic and water quality modelling in distribution networks: Epanet-pmx," *Water Resources Management*, vol. 31, no. 14, pp. 4485–4503, 2017.

[46] A. A. Abokifa, *Developing Multi-Scale Models for Water Quality Management in Drinking Water Distribution Systems*. Washington University in St. Louis, 2018.

[47] P. D. Lax and B. Wendroff, "Difference schemes for hyperbolic equations with high order of accuracy," *Communications on pure and applied mathematics*, vol. 17, no. 3, pp. 381–398, 1964.

[48] G. D. Smith, *Numerical Solution of Partial Equations: Finite Difference Methods*. Oxford University Press, 1978.

[49] K. S. Tshehla, Y. Hamam, and A. M. Abu-Mahfouz, "State estimation in water distribution network: A review," in *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*. IEEE, 2017, pp. 1247–1252.

[50] S. Wang, A. F. Taha, L. Sela, N. Gatsis, and M. H. Giacomoni, "State estimation in water distribution networks through a new successive linear approximation," in *IEEE 58th Conference on Decision and Control, Nice, France*, 2019, pp. 5474–5479. https://arxiv.org/pdf/1909.03182.pdf

[51] J. Mackenzie and W. Mekwi, "An analysis of stability and convergence of a finite-difference discretization of a model parabolic pde in 1d using a moving mesh," *IMA journal of Numerical analysis*, vol. 27, no. 3, pp. 507–528, 2007.

[52] J. Crank and P. Nicolson, "A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 43, no. 1. Cambridge University Press, 1947, pp. 50–67.

[53] S. C. Chapra, *Surface water-quality modeling*. Waveland press, 2008.

[54] A. Salih, "Method of characteristics," *Indian Institute of Space Science and Technology, Thiruvananthapuram*, 2016.

[55] P. F. Boulos, K. E. Lansey, and B. W. Karney, *Comprehensive water distribution systems analysis handbook for engineers and planners*. American Water Works Association, 2006.

[56] J. C. Butcher, *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[57] L. D. Schmidt *et al.*, *The engineering of chemical reactions*. Oxford University Press, USA, 1998.

[58] S. P. Nørsett and G. Wanner, *Solving ordinary differential equations: Stiff and differential-algebraic problems*. Springer-Verlag, 1993.

[59] I. Fisher, G. Kastl, A. Sathasivan, and V. Jegatheesan, "Suitability of chlorine bulk decay models for planning and management of water distribution systems," *Critical Reviews in Environmental Science and Technology*, vol. 41, no. 20, pp. 1843–1882, 2011.

[60] C. N. Haas and S. Karra, "Kinetics of wastewater chlorine demand exertion," *Journal (Water Pollution Control Federation)*, pp. 170–173, 1984.

[61] J. C. Powell, J. R. West, N. B. Hallam, C. F. Forster, and J. Simms, "Performance of various kinetic models for chlorine decay," *Journal of Water Resources Planning and Management*, vol. 126, no. 1, pp. 13–20, 2000.

[62] J. D. Johnson, "Measurement and persistence of chlorine residuals in natural waters," in *Proceedings of the Conference on the Environmental Impact of Water Chlorination*, 1975.

[63] D. Feben, M. J. Taras, H. Faber, and L. Hedgepeth, "Studies on chlorine demand constants [with discussion]," *Journal (American Water Works Association)*, vol. 43, no. 11, pp. 922–932, 1951.

[64] R. M. Clark and M. Sivaganesan, "Predicting chlorine residuals in drinking water: Second order model," *Journal of water resources planning and management*, vol. 128, no. 2, pp. 152–161, 2002.

[65] K. Umberg, J. G. Uber, and R. Murray, "Performance evaluation of real-time event detection algorithms," in *Water Distribution Systems Analysis Symposium 2006*, 2008, pp. 1–14.

[66] I. Fisher, G. Kastl, and A. Sathasivan, "New model of chlorine-wall reaction for simulating chlorine concentration in drinking water distribution systems," *Water research*, vol. 125, pp. 427–437, 2017.

[67] J. G. Uber, R. Murray, M. Magnuson, and K. Umberg, "Evaluating real-time event detection algorithms using synthetic data," in *World Environmental and Water Resources Congress 2007: Restoring Our Natural Habitat*, 2007, pp. 1–11.

APPENDIX A
EPANET'S AR DISCRETIZATION ALGORITHM

EPANET and its extension, EPANET-MSX, use the Lagrangian time-driven method to solve the advection-reaction PDE in pipes. Hereinafter, Algorithm 2 gives a full description of how the method is applied for each hydraulic time-step within the simulation period $[0, T_s]$.

Algorithm 2: EPANET's Lagrangian time-driven method to solve advection-reaction PDE for each hydraulic time-step

---

    Input: Hydraulic parameters (e.g., velocities, flow rates), pipe's length, and adjustment tolerance (τ_{adj})
    Output: Chemical concentrations in each pipe
1 **if** *It is the first hydraulic time-step* **then**
2   Pipe is considered as one segment with initial concentration equal to the upstream node's concentration
3 **else**
4   Segments are ordered in the direction of the flow
5 **end if**
6 **while** $t \leq T_s$ **do**
7   Apply reaction to each segment
8   Calculate volume entering each downstream node, which equal pipe's flow rate times time-step
9   **if** *volume is larger than last segment's volume* **then**
10   Last segment is distroyed
11   Segment before the last one contributes its volume
12   **else**
13   Last segment's volume is decreased by the volume contributed
14   **end if**
15   Update concentrations at nodes according to entering volumes and demands
16   **if** *difference between concentrations of the upstream node and the first segment* $\leq \tau_{adj}$ **then**
17   Update first segment's size by the inflowing volume
18   **else**
19   Create new segment with the upstream node's concentration
20   **end if**
21   Obtain the average of the segments' concentrations to represent the whole pipe's concentration at time t
22     $t = t + \Delta t$
23 **end while**

APPENDIX B
STATE-SPACE FORMULATION EXAMPLE - DIFFERENT TECHNIQUES

In this appendix we apply Backward Euler, Crank-Nicolson schemes, and Method of Characteristics (Tech. 2,3, and 4) on the same example explained in Section VII-A by highlighting the differences from the Lax-Wendroff scheme (Tech. 1). As noticed in Tab. III, the main differences between the techniques are in $E(t), A(t)$, and $f(x_1, x_2, t)$. That is, we give a general form for them (Tab. IX) for the example then list the specific application for each of the techniques (Tab. X). We only list elements for the pipe's segments in Tab. X) as the remaining elements are the same as explained in Section VII-A. Notice that, the listed values are for chlorine. As for the fictitious reactant, k_1^{P} is considered equal to zero.

Tab. IX: General form for the state-space matrices for the three-node network example

Tab. X: Different techniques implementation on the three-node network example

Element	Tech. #2 - Backward Euler Scheme	Tech. #3 - Crank-Nicolson Scheme	Tech. #4 - Method of Characteristics
e_{52} = e_{65} = e_{76}	$-0.5\tilde{\lambda}_1$	$-0.25\tilde{\lambda}_1$	0
e_{55} = e_{66} = e_{77}	1	1	1
e_{56} = e_{67} = e_{73}	$0.5\tilde{\lambda}_1$	$0.25\tilde{\lambda}_1$	0
a_{52}	0	$0.25\tilde{\lambda}_1$	1
a_{65}	0	$0.25\tilde{\lambda}_1$	$\tilde{\lambda}_i(t)\exp(-(k_i^{\mathrm{P}}+k_r\tilde{c}_1^{\mathrm{P}}(1,t))\Delta t)$
a_{76}	0	$0.25\tilde{\lambda}_1$	$\tilde{\lambda}_i(t)\exp(-(k_i^{\mathrm{P}}+k_r\tilde{c}_1^{\mathrm{P}}(2,t))\Delta t)$
a_{55}	$1-k_1^{\mathrm{P}}$	$1-k_1^{\mathrm{P}}$	0
a_{66}	$1-k_1^{\mathrm{P}}$	$1-k_1^{\mathrm{P}}$	$(1-\tilde{\lambda}_i(t))\exp(-(k_i^{\mathrm{P}}+k_r\tilde{c}_1^{\mathrm{P}}(2,t))\Delta t)$
a_{77}	$1-k_1^{\mathrm{P}}$	$1-k_1^{\mathrm{P}}$	$(1-\tilde{\lambda}_i(t))\exp(-(k_i^{\mathrm{P}}+k_r\tilde{c}_1^{\mathrm{P}}(3,t))\Delta t)$
a_{56} = a_{67} = a_{73}	0	$-0.25\tilde{\lambda}_1$	0
$f_1^{\mathrm{P}}(1)$	$\tilde{c}_1^{\mathrm{P}}(1,t)c_1^{\mathrm{P}}(1,t)$	$\tilde{c}_1^{\mathrm{P}}(1,t)c_1^{\mathrm{P}}(1,t)$	0
$f_1^{\mathrm{P}}(2)$	$\tilde{c}_1^{\mathrm{P}}(2,t)c_1^{\mathrm{P}}(2,t)$	$\tilde{c}_1^{\mathrm{P}}(2,t)c_1^{\mathrm{P}}(2,t)$	0
$f_1^{\mathrm{P}}(3)$	$\tilde{c}_1^{\mathrm{P}}(3,t)c_1^{\mathrm{P}}(3,t)$	$\tilde{c}_1^{\mathrm{P}}(3,t)c_1^{\mathrm{P}}(3,t)$	0