

Contamination Event Diagnosis in Drinking Water Networks: A Review

Demetrios G. Eliades^a, Stelios G. Vrachimis^{a,b}, Alireza Moghaddam^{a,b}, Ioannis Tzortzis^{a,b}, Marios M. Polycarpou^{a,b}

^aKIOS Research and Innovation Center of Excellence, University of Cyprus, Cyprus

^bElectrical and Computer Engineering Dept., University of Cyprus, Cyprus

Abstract

Water distribution systems are susceptible to contamination events, which can occur due to naturally occurring events, accidents or even malicious attacks. When a contamination event occurs, dangerous substances infiltrating the network may be consumed and eventually have a significant impact on the consumers' health and the economy. Advances in information and telecommunication technologies, sensors and actuators, are enabling water networks to become smarter and more resilient to these types of events. In the last decades, water security research focused on contamination event diagnosis, and delivered important scientific results. This paper provides an extensive review of the state-of-the-art research in water security, focusing on water quality sensor placement, estimation and disinfection control, contamination detection and source isolation, as well as contamination management. Moreover, this paper attempts to formulate the results in a unified systems-theoretic mathematical framework.

Keywords: drinking water, water distribution, water quality, contamination event, critical infrastructure systems, water security, event preparedness, event diagnosis, emergency event management, early-warning systems

1. Introduction

1.1. Water Distribution Systems

Water Distribution Systems (WDS) are networks of pipes of different lengths and diameters, whose purpose is to connect water sources (such as water storage tanks or reservoirs) with a typically large number of outflow points, that are used for supplying water to consumers. The flow of water within the network is mainly driven by consumer demands. In general, WDS are pressurized and should be able to guarantee the delivery of a sufficient quantity of water to all consumers in the network, up to a certain elevation above ground. To achieve this, the design of the system can include actuators, such as pumping stations and valves, as well as dynamic elements, such as water tanks. At the same time, it's important to provide safe drinking water to consumers, typically through the use of disinfection agents or other technologies. Therefore, from a systems engineering perspective, there are two main components of interest in WDS: (i) the *hydraulic* part, which characterizes the water quantity; and (ii) the *quality* part, which characterizes the water quality and the presence of possible contamination events.

1.2. Water quality contamination events

Maintaining water quality within the regulations specified by the World Health Organization (WHO) (World Health Organization, 2008), the European Commission (The European Parliament and the Council of the European Union, 2020), or the U.S. Environmental Protection Agency (EPA) (U.S. Government, 2002), is an important challenge faced by water utilities which supply water to consumers through drinking water distribution networks. Guaranteeing a high level of water

quality, continuously, is a challenging task, since quality deterioration events, such as contamination due to pathogens or other substances, may occur in the system, affecting quality and posing a risk to human health. To safeguard a high level of water quality, various disinfection technologies are employed.

In many parts of the world, disinfectants are used in prescribed concentrations to prevent bacteria growth and neutralize certain chemical agents, and, in general, to maintain a high level of drinking water quality. For instance, according to the World Health Organization (WHO), a free-chlorine residual concentration should exist in drinking water distribution systems, with a minimum target concentration of $0.2 \frac{\text{mg}}{\text{L}}$ at the point of delivery (World Health Organization, 2008). However, when chlorine reacts with organic compounds in water, it creates disinfection by-products, such as Trihalomethanes (THMs) and Haloacetic Acids (HAA), which present health risks and are linked with certain bladder and rectal cancers (Morris, 1995; Richardson, 2009). For this reason, the objective is to maintain prescribed minimum concentrations at consumer locations, while ensuring the appropriate disinfection of water.

In general, the key challenge in water quality is to monitor and control the use of disinfectant substances as well as their disinfection by-products, in order to guarantee compliance with the relevant water quality regulations. It is important to note the distinction between monitoring and control of water disinfection under normal conditions and the severe case of contamination events, which may cause significant disruption to the water supply and, thus, needs to be detected and managed as soon as possible.

1.3. Water quality monitoring technologies

To monitor water quality, it is common for water authorities to employ laboratory-based methods for measuring the concentrations of different water quality parameters. This procedure, however, is usually conducted in a manual way and is time-consuming, without the ability to provide feedback in near-real-time. With the advancement of Information and Communication Technologies (ICT), it is possible to monitor and control the operation of water systems through the use of Supervisory Control and Data Acquisition (SCADA) systems, wireless networks and specialized sensors. Specifically, significant research and innovation has been invested in the past decades in relation to water quality monitoring, focusing on the design of reliable online sensors that are able to monitor various chemical and biological parameters within WDS (McKenna et al., 2010; Raich, 2013; Kruse, 2018). This includes color, conductivity, dissolved oxygen (DO), free chlorine (when used as a disinfectant), spectral absorbance/ light absorption at UV 254nm (UV254), oxidation-reduction potential (ORP), pH, temperature, total organic carbon (TOC), turbidity, (Hall et al., 2009). More advanced sensors provide the ability to also measure bacteria activity, toxicity and radiation. Currently, installing specialized sensors that monitor the system for all possible waterborne contaminants is considered too costly, and as a result, the decision on the number and type of sensors to be installed needs to be based on a risk-based analysis, considering different types of contamination events that may occur in the systems.

1.4. Examples of contamination events

Water contamination events can be caused due to the infiltration of various chemical, biological, or radioactive substances in the drinking water system. These may occur due to natural or accidental events, or even malicious actions. Some examples of contaminants include agriculture chemicals (such as pesticides), toxins, metals, bacteria, viruses, protozoa, helminths, etc (World Health Organization, 2017).

These events can occur typically as a result of raw-water contamination, treatment deficiencies, and distribution network failures (Nascetti et al., 2021). For instance, a contamination event can occur when wastewater with significant concentrations of pathogens, infiltrates the drinking water distribution system due to a hazard, such as a flood or a storm. Pathogens, in specific, can affect significantly the population's health. Typical examples of these waterborne disease agents include *Escherichia coli* (E. Coli) O157:H7, *Legionella*, *Toxoplasma*, and others (Hunter et al., 2001).

Various real waterborne disease outbreaks have been studied in the literature and the reader is referred to the relevant studies in which these water incidents are analyzed (Xin et al., 2017; Ford, 1999; Riera-Montes et al., 2011; Jacqueline et al., 2022; Jakopanec et al., 2008; Corso et al., 2003; Bopp et al., 2003). A brief overview of some notable contamination events is provided below.

In Nokia (Finland), in November 2007, 450 m³ of wastewater was injected into the town's drinking water distribution

system as a result of human error, causing an outbreak of *gastroenteritis*. Consequently, thousands were infected, hundreds were hospitalized, and the authorities imposed a complete ban on all water usage for 12 weeks (Lienemann et al., 2011). Similar events due to accidental wastewater infiltration have occurred during recent years, affecting the local communities and causing outbreaks of various diseases, e.g., in Finland (Kauppinen et al., 2019), in Italy (Giammanco et al., 2018), in Denmark (Kuhn et al., 2017), and in France (Beaudeau et al., 2008).

During 2011, one of the largest reported European waterborne outbreaks of *cryptosporidiosis* occurred in the municipality of Skellefteå (Sweden), in which 18,500 people were infected by the *Cryptosporidium* pathogen. The reason for this outbreak was the quality deterioration of the public water supplied by the water treatment plant (Bjellmar et al., 2017).

Cyber-attacks have also been linked with possible contamination events. For instance, a hacker was able to infiltrate remotely a water treatment plant in Florida in February 2021, and was able to change by 100 times the injection concentration of *sodium hydroxide*, an agent used for water purification. This action would have caused a significant impact if it was not detected on time (Davies et al., 2022)

Contamination events can also be caused due to cascading events during disasters. For example, after a fire in a town in Italy in 2020, a number of citizens started reporting symptoms attributed to various waterborne pathogens. The source of the contamination is suspected to be the firefighters' pressurized water tanks which were connected to fire hydrants on the water distribution network. The tanks were previously filled with water from different sources, including a sewage-polluted river (Nascetti et al., 2021).

1.5. Motivation and Key Contributions

Previous related survey papers on water contamination have considered different aspects of the problem. In (Islam et al., 2015), an extensive literature review of modelling and optimization methods for detecting and mitigating contamination events, and provides an integrated framework. Moreover, in (Kanakoudis & Tsitsifli, 2017), a literature review is provided, covering monitoring, modelling, optimization and early-warning-system aspects. Their paper proposes a "simheuristics" framework, i.e, using simulation models and meta-heuristics to handle uncertainties. In both aforementioned references, the authors do not provide a common mathematical problem formulation of the different methodologies presented, whereas this is one of the contributions of the present work.

This paper provides a review of the theoretical, modeling, and computational developments in the area of contamination event diagnosis for WDS, by analyzing previous work within a common systems and control framework. In addition, research has been segmented into three main phases, aligned with (Qiu et al., 2020) as well as the JRC Water Security guidelines (Teixeira et al., 2019, 2022), which can be summarized as the "Preparedness" (Phase 1), "Event Detection" (Phase 2) and "Emergency Event Management" (Phase

3). Note that we consider the recovery phase as beyond the scope of this review.

We outline the contributions of this work as follows:

- It provides a systematic review of state-of-art research performed in the area of contamination event diagnosis within the operations research and machine learning communities.
- It provides a general mathematical framework that represents, in a unified way, a broad range of problems and challenges associated with contamination event diagnosis.
- From the modeling and theoretical perspectives, we elaborate on the principles of hydraulics and water quality analysis, which provide a mathematical representation of WDS.
- From an algorithmic perspective, we review solution techniques for contamination event diagnosis.

This paper is structured as follows: in Section 2, we present an integrated mathematical modelling of water distribution systems, covering hydraulic, quality and impact dynamics. In Section 3, we present a review of technologies related to the “preparation” phase, focusing on sensor placement and disinfection control, and in Section 4 the “contamination diagnosis” phase with the detection and isolation challenges. Section 5 focuses on emergency contamination management methods. Finally, Section 6 concludes the paper and future work is discussed.

2. Modeling of Water Distribution Systems

A Water Distribution System (WDS) is characterized by a set of differential, algebraic and partial differential equations that describe its hydraulic and quality dynamics, given a topological and structural network graph, and considering certain input and disturbance signals. Note that, in normal operating conditions, hydraulic dynamics affect the water quality dynamics, but not vice versa. This concept is illustrated in Fig. 2, which provides an overview of the information flow between the different components of the water distribution system, including the modules related to *preparedness*, *event diagnosis* and *emergency event management* which will be described in the following sections.

In this section, we describe the fundamental mathematical models of a WDS, including the network graph, hydraulic and quality dynamics, as well as impact dynamics. Furthermore, we provide information regarding the computational tools for simulating the aforementioned models.

2.1. Network graph

The topology of a WDS is modeled by a directed graph denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Specifically, let \mathcal{V} be the set of vertices (nodes) defined as $\mathcal{V} \subset Z \times T_v \times \Theta_v \times D$. The set

$Z = \{1, \dots, n_v\}$ indicates the positive integers corresponding to the index of the i -th node, $v_i \in \mathcal{V}$, and $|\mathcal{V}| = n_v$. The set T_v indicates the type of node, and consists of junctions of pipes and consumer (water demand) locations, reservoirs (water sources), and tanks (storage), such that $T_v = \{\text{junction}, \text{reservoir}, \text{tank}\}$. The set Θ_v associates each node with parameters (real numbers) that describe the physical properties of the network that affect the hydraulics and water quality; e.g., the elevation of all nodes. The set $D \subset \mathbb{R}^{n_v}$ describes the consumer water demands at nodes, such that for node v_j the demand is denoted by d_j . We also define $\mathcal{V}_J \subset \mathcal{V}$ the set of nodes which correspond to ‘junction’ types, and similarly, $\mathcal{V}_T \subset \mathcal{V}$ the tank nodes ($n_t = |\mathcal{V}_T|$), and $\mathcal{V}_R \subset \mathcal{V}$ the reservoir nodes, such that, $\mathcal{V} = \mathcal{V}_J \cup \mathcal{V}_T \cup \mathcal{V}_R$.

Let \mathcal{E} be the set of edges defined as $\mathcal{E} \subset \mathcal{V} \times \mathcal{V} \times T_e \times \Theta_e$, where $e_{(i,j)} \in \mathcal{E}$ denotes an edge between nodes v_i and v_j with $i, j \in Z, i \neq j$, and $|\mathcal{E}| = n_e$ being the total number of links. Let the network *incidence matrix* $M \in \{-1, 0, +1\}^{n_e \times n_v}$ be a matrix indicating the connectivity of nodes with links, such that element $M_{(i,j)} = +1$ if the conventional direction of link i enters node j ; element $M_{(i,j)} = -1$ if the conventional direction of link i leaves from node j ; otherwise $M_{(i,j)} = 0$. The set T_e indicates the type of link, and consists of pipes, pumps, and valves, such that $T_e = \{\text{pipe}, \text{pump}, \text{valve}\}$. Pumps and valves are actuators and the main hydraulic control elements in a water network. The set Θ_e corresponds to edge parameters (real numbers) which depend on the edge type, e.g., length, diameter and roughness when the link is a pipe, or polynomial coefficients determining the characteristic curve in the case of a pump.

In this work, we consider that the parameters in the sets Θ_v and Θ_e are time-invariant. However, it is important to note that certain parameters are accurately known, and some may change slowly in the long-term, e.g., the pipe roughness which causes friction and energy losses.

2.2. Hydraulics modeling

The key hydraulic quantity associated with each node v_j is the *hydraulic head*, denoted by h_j . Hydraulic head is defined as the energy per unit weight, measured in meters of water column. In practice, the hydraulic head corresponds to the elevation of the node with respect to a geodesic reference (elevation head), a function of the velocity of water in the pipe (kinetic head), and a function of the pressure (pressure head). Typically, the kinetic head is neglected because it is insignificant compared to elevation and pressure head (Boulos et al., 2006). Water flows from higher to lower heads. The main hydraulic quantity associated with a link $e_{(i,j)}$ is the *water flow*, denoted by $q_{(i,j)}$.

The overall hydraulic state $x^h \in \mathbb{R}^{n_h}$ of a water network, is defined by the head at nodes and flow in links, thus $n_h = n_v + n_e$. These states are calculated using a hydraulic model of a WDS, which is a set of equations derived from the laws of: (i) conservation of mass; and (ii) conservation of energy in the network. These equations typically neglect the fast (transient) hydraulic dynamics, because of the nominally low resolution of measurements available in water networks, as well

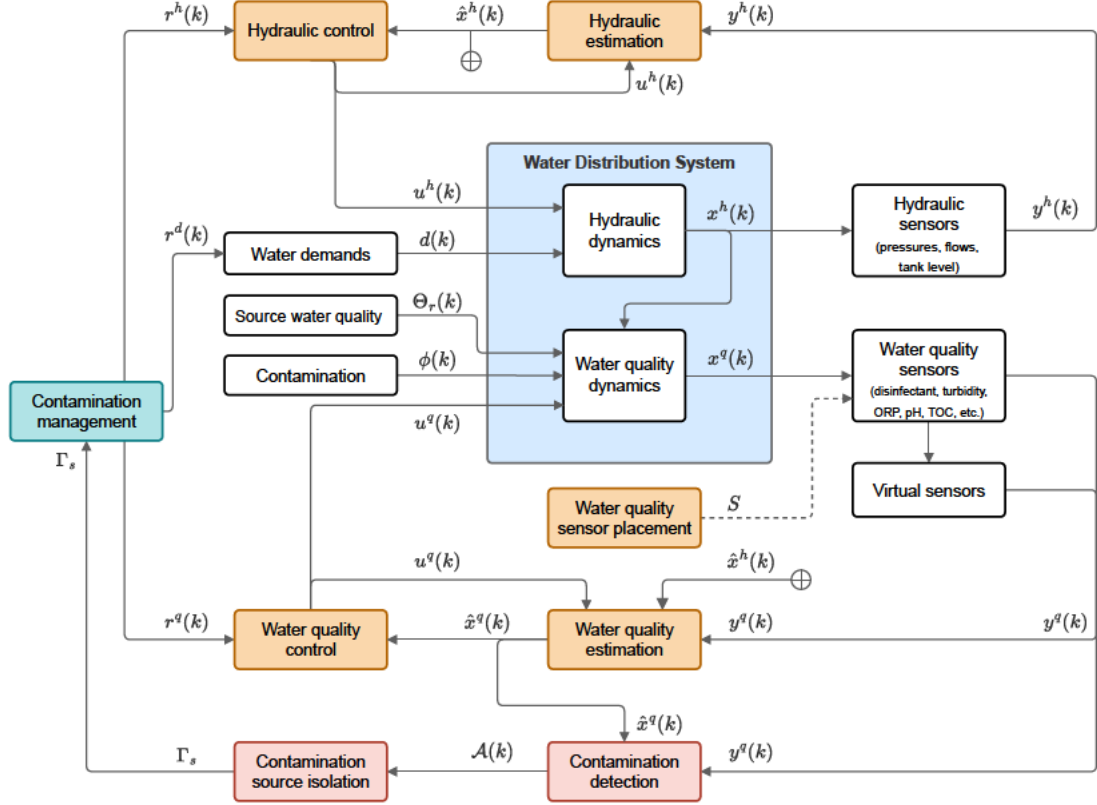


Figure 1: Information flow diagram for the water contamination event diagnosis components. The orange boxes correspond to Phase 1 “preparedness”, red boxes to Phase 2 “event diagnosis”, whereas the green box to Phase 3 “emergency event management”

as due to the fact that a steady-state analysis is adequate for most practical applications. In this work, we use the *pipe formulation* of these equations as used by (Todini & Pilati, 1988), which has been shown to be robust in computer simulations (Rossman, 2000).

For notational convenience, in the following, we consider the discrete-time representation of the system dynamics, indicated by k and corresponding to a hydraulic measurement time step of Δt (typically in the order of minutes or hours).

In accordance to the principle of mass conservation (similar to Kirchhoff’s junction rule), the sum of all inflows and outflows for a node $v_i \in \mathcal{V}_j$ at time k is equal to:

$$\sum_{j \in \mathcal{N}(v_i)} q_{(i,j)}(k) = f_d(h_i(k), d_i(k); \Theta_{v_i}), \quad (1)$$

where $\mathcal{N}(v_i)$ is the set of neighboring nodes of v_i , such that $\mathcal{N}(v_i) = \{v_j \in \mathcal{V} \mid e_{(i,j)} \vee e_{(j,i)} \in \mathcal{E}\}$. The function $f_d(\cdot)$ calculates the outflow water demand at node v_i , given the node head h_i , requested demand d_i (assuming a pressure-driven model), and the node parameters Θ_{v_i} (e.g., elevation). A special case of (1) is when the total requested demand is satisfied (pressure-sufficient conditions), thus $f_d(h_i(k), d_i(k); \Theta_{v_i}) = d_i(k)$ (demand-driven model). Note that in practice, $d_i(k)$ is unknown, except in cases when demand is measured in real-time by “smart metering” sensors. If $d_i(k)$ is not available, then it is possible to consider estimates, or pseudo-measurements, of the demand $\hat{d}_i(k)$ at each node v_i . Tanks

are modeled as nodes in the water distribution network and are dynamic elements of the system (Boulos et al., 2006). The head state of the tank node $v_i \in \mathcal{V}_T$, is given by

$$h_i(k+1) = h_i(k) + \frac{\Delta t}{\alpha_i} \sum_{j \in \mathcal{N}(v_i)} q_{(i,j)}(k), \quad (2)$$

where the tank head $h_i(k)$ corresponds to the relative water level plus the tank elevation, and α_i is the cross-sectional area of the tank (assuming uniform shape). The initial tank head is given by $h_i(0)$.

Furthermore, according to the principle of energy conservation, the relationship between flow and loss of hydraulic head due to friction (headloss) across each link in the network, must be balanced. Let $h_i(k)$ and $h_j(k)$ be the hydraulic heads at the i -th and j -th node respectively, for $v_i, v_j \in \mathcal{V}$. When $h_i(k) > h_j(k)$, water flows from node v_i (higher head) to node v_j (lower head) along the edge $e_{(i,j)}$ (e.g., the (i, j) -th pipe) with flow $q_{(i,j)}(k)$; the head-flow relationship is given by

$$h_i(k) - h_j(k) = f_h(q_{(i,j)}(k), u_{(i,j)}^h(k); \Theta_{e_{(i,j)}}), \quad (3)$$

where $f_h(\cdot)$ is a nonlinear (headloss) function that depends on link $e_{(i,j)}$ type and parameters $\Theta_{e_{(i,j)}}$, and $u_{(i,j)}^h(k)$ is a control input.

From a systems and control viewpoint, pumps and valves are used as actuators for hydraulic control in a WDS. Pumps

are actuators that add head across a link, while valves regulate the headloss in a desired way (e.g., pressure-reducing valves). Check-valves are a special group of actuators because they can only be opened or closed ($u_{(i,j)}^h \in \{0,1\}$), essentially changing the connectivity of the network. This affects the size of the incidence matrix M , and the number of states, as the flow state of closed pipes cannot be defined by the hydraulic equation of (3). For this reason, it is more suitable to model these inputs by reformulating the hydraulic equations (mode change). In the following, the modeling of control inputs corresponds to changing the settings of pumps that exist in the network by altering the pump head-flow curve or valves that alter the flow-headloss relationship of links (Boulos et al., 2006).

Following the work by Pasqualetti et al. (2015) and Vrachimis et al. (2018), the differential-algebraic hydraulic equations of a drinking water network (1)–(3), can be written in state-space form using the conservation of energy and mass equations, based on the following assumptions: 1) demand-driven modeling; 2) head-loss functions $f_h(\cdot)$ do not depend on previous flows; 3) control inputs do not change the network topology and are linear with respect to the head-loss function $f_h(\cdot)$:

$$E \begin{bmatrix} h(k+1) \\ q(k+1) \end{bmatrix} = \begin{bmatrix} E_t & M^\top \\ M & F_h(q(k)) \end{bmatrix} \begin{bmatrix} h(k) \\ q(k) \end{bmatrix} + Bu^h(k) + \begin{bmatrix} -d(k) \\ 0 \end{bmatrix}, \quad (4)$$

where $E \in \mathbb{R}^{n_h \times n_h}$ identifies the dynamic states corresponding to tank levels and multiplies them with the tank-base area (assuming cylindrical shape); $E_t \in \mathbb{R}^{n_v \times n_v}$ identifies the tank head states and multiplies them with the tank-base area; $M \in \mathbb{R}^{n_e \times n_v}$ is the network incidence matrix; $F_h : \mathbb{R}^{n_e} \mapsto \mathbb{R}^{n_e \times n_e}$, contains the nonlinear functions $f_h(q_{(i,j)}(k)) q_{(i,j)}^{-1}(k)$ at its diagonal; $B \in \mathbb{R}^{n_h \times n_e}$ identifies the additive control inputs in each equation; and $d(k) \in \mathbb{R}^{n_v}$ are the water demands at nodes, modeled as an uncontrolled (known or unknown) input to the system.

The matrix E is singular due to the presence of algebraic equations in (4). Note that the system can be transformed into a non-singular form by rearranging and separating the dynamic states from the algebraic states (Vrachimis et al., 2018) as follows:

$$\chi(k+1) = A^{(1)}\chi(k) + A^{(2)}\zeta(k) \quad (5a)$$

$$\zeta(k) = f_\zeta(\chi(k), u^h(k), d(k)) \quad (5b)$$

where $\chi(k)$ and $\zeta(k)$ are the dynamic and algebraic states respectively. The matrices $A^{(1)} \in \mathbb{R}^{n_t \times n_t}$, $A^{(2)} \in \mathbb{R}^{n_t \times (n_v - n_t)}$ are derived from (4), and $f_\zeta(\cdot)$ is a nonlinear function solving the algebraic hydraulic equations. Also, note that in this representation the dynamic (tank) equations (5a) are linear time-invariant with the algebraic state $\zeta(k)$ treated as an input. The hydraulic state of a WDS is defined by:

$$x^h(k) = [\chi(k)^\top \zeta(k)^\top]^\top \in \mathbb{R}^{n_h}, \quad (6)$$

where $n_h = n_v + n_e$ is the hydraulic state vector size. It is typical that pressure, flow, and tank-level sensors are installed at some limited number of locations within the network. The measured hydraulic states of the system are defined by:

$$y^h(k) = C^h x^h(k) + \eta^h(k) \quad (7)$$

where $y^h(k) \in \mathbb{R}^{n_s}$ denotes the measurement vector at time k , n_s is the number of sensors, and $\eta^h(k) \in \mathbb{R}^{n_s}$ is the measurement noise vector.

2.3. Water quality modeling

Water-quality is an abstract concept that aims to characterize the quality of water consumed by humans and other animals. In practice, it is represented by the concentration of certain key variables in the water. Water quality dynamics describe how the mass per unit volume (concentration) of one or more physical, chemical, and biological agents within the water distribution network change in space and time. In general, multiple agents may exist simultaneously within a WDS, and some may react to produce new chemical compounds, whereas some are non-reactive and may remain constant, grow or decay in time and space.

Let $w \in \mathbb{R}^{n_w}$ be a vector indicating the concentration of n_w substances which may exist within a water distribution system at a certain location and time, with $w^{(i)}$, $i \in \{1, \dots, n_w\}$ corresponding to the concentration of the i -th substance. We are typically interested only in a subset of substances that may exist in the water and need to be regulated, or that should not exist because they pose a threat to human health. For example, $w^{(1)}$ may correspond to the disinfectant concentration (e.g., Chlorine), $w^{(2)}$ may correspond to certain bacteria (e.g., E. coli), and so on. Next, we discuss the models describing the reactions in water, the advection and reactions in pipes, as well as a model for the quality of water in storage tanks.

2.3.1. Reaction dynamics

The reaction dynamics characterize how the concentration of one or more substances changes when reacting, decaying, or growing within a finite volume of water. Single-species reaction dynamics are widely used in research, to describe the rate of decay or growth of a substance (Rossman, 2000). This corresponds to the case of considering the concentration of a single substance, described by the state $w^{(i)}$, with all other substances neglected. The single-species representation is convenient since the reactions and even the existence of all species in a water volume are typically unknown.

The modeling of multiple-species reactions (Shang et al., 2008) involves coupled sets of differential and algebraic equations, which may be summarized by:

$$\frac{dw(t)}{dt} = f_r(w(t), \Theta_r), \quad (8)$$

where w is a vector of concentrations of n_w substances of interest, $f_r(\cdot)$ is a vector field describing the concentration change due to the decay/growth reactions between substances of interest, and Θ_r are the reaction kinetics (Shang

et al., 2008). In practice, f_r is derived from the stoichiometry of the reactions and may be composed of linear or bilinear terms. A special case of (8) is the first order decay model:

$$\frac{dw(t)}{dt} = \theta_r w(t), \quad (9)$$

where $\theta_r \in \Theta_r$ is a parameter vector indicating the decay/growth rates of the substances w .

2.3.2. Advection-Reaction dynamics

When substances enter a pipe corresponding to edge $e_{(i,j)}$, in which water flows under pressure, those substances move along with that flow and may react with other substances which are also floating in the bulk water or residing on the pipe walls. The change in space and time of the substance concentration in water, in continuous time t , as well as its reactions dynamics occurring at the bulk water and pipe surface (neglecting axial dispersion for simplicity), is described by a first-order hyperbolic partial differential equation:

$$\begin{aligned} \frac{\partial w_{(i,j)}(z, t)}{\partial t} + \frac{q_{(i,j)}(t)}{\alpha_{(i,j)}} \frac{\partial w_{(i,j)}(z, t)}{\partial z} = & f_r(w_{(i,j)}(z, t), \Theta_r) \\ & + u_{(i,j)}^q(z, t) \\ & + \phi_{(i,j)}(z, t), \end{aligned} \quad (10)$$

where $w_{(i,j)}(z, t)$ is the vector of substance concentrations in water at continuous time t and at distance z along a pipe corresponding to edge $e_{(i,j)}$, with water flow $q_{(i,j)}(t)$ and cross-sectional area of the pipe $\alpha_{(i,j)} \in \Theta_e$. The function $f_r(\cdot)$ corresponds to the change in the concentration of all the substances of interest due to reactions with other substances flowing in the bulk water or residing on the pipe walls, taking into account the pipe parameter vector Θ_r ; this includes also growth or decay of the agents without reactions, e.g., in the case of bacteria or radioactive material. The function $u_{(i,j)}^q$ corresponds to the controlled input of substances (e.g., the controlled addition of a disinfectant). The function $\phi_{(i,j)}(z, t)$ corresponds to the signal of uncontrolled or controlled injection of contaminants that can occur at any location in a water distribution network (e.g., accidental injection of contaminants or a malicious attack). Finally, note that the addition of a substance not included in w will result in the change of the function $f_r(\cdot)$, if the aforementioned substance reacts with any substances in w .

In general, it is difficult to solve the complete set of hyperbolic partial differential equations in (10) analytically. However, numerical/computational methods have been employed for solving this problem (Rossman & Boulos, 1996; Shang et al., 2008). Typically, Eulerian and Lagrangian approaches are considered for solving the advection dynamics. In the Eulerian approach, space is discretized, and therefore the states correspond to fixed locations, whereas in the Lagrangian approach, the state corresponds to a moving particle (LeVeque, 2002). In the following, we demonstrate the Eulerian approach which leads to a convenient discrete-time state-space formulation.

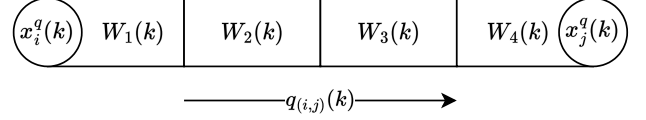


Figure 2: A single pipe is segmented into four finite volumes (rectangles), with concentration. The nodes (circles) correspond to the boundary conditions

We consider the “Finite Volumes Method”, which is a typical method for solving hyperbolic partial differential equations (LeVeque, 2002). According to the Finite Volume Method, the (i, j) -th pipe is segmented into a finite number of uniform volume cells, with $n_{z(i,j)}$ being the total number of cells. For an arbitrary volume cell $l \in \{1, \dots, n_{z(i,j)}\}$, we define $w_{(i,j,l)}(k)$ as the average concentration vector of each substance in that cell at discrete time k for a Δt quality time step, such that

$$w_{(i,j,l)}(k) = \frac{1}{\Delta z} \int_{z_l}^{z_l + \Delta z} w_{(i,j)}(z, k\Delta t) dz, \quad (11)$$

where Δt is the length of a hydraulic discrete time-step, which is a design parameter and may depend on the available sensors, Δz is the width of a single cell and z_l is the distance of the beginning of the l -th finite volume from the pipe inflow point. Both Δt and Δz are assumed to be fixed within a certain pipe. Note that the hydraulic and quality time steps may be different, however in this work we assume that both are equal and sufficiently small to describe both dynamics. For notational convenience, we can omit the (i, j) indicator, so that $w_l(k)$ is the concentration vector at the l -th finite volume of a certain pipe.

To numerically solve the advection and reaction equations in (10), we can use the *Lax-Wendroff* or the *Upwind* scheme to solve the advection dynamics (i.e., the propagation of a substance in a pipe), and the single finite-difference *Unsplit* method, or the *Fractional step* method for solving the reaction dynamics (LeVeque, 1998, 2002). Finally, the reaction Ordinary Differential Equations (ODEs), in the case of multiple substances, can be solved using a numerical integrator solution such as *Runge-Kutta* method Shang et al. (2008). A detailed comparison of different solution methods can be found in (Rossman & Boulos, 1996) and more recently in (Elsherif et al., 2022). In the following, we provide a high-level outline of the solution approach as described in (Eliades & Polycarpou, 2010).

Figure 2 illustrates a pipe segmented into four finite volumes. We assume a single contaminant that decays in time linearly. We demonstrate how to compute the l -th finite volume concentration using the *Upwind and Unsplit* method, for which

$$\begin{aligned} w_l(k+1) = & w_l(k) - f_\lambda(q_{(i,j)}(k)) [w_l(k) - w_{l-1}(k)] \\ & - f_r(w_l(k), \Theta_r), \end{aligned} \quad (12)$$

where Θ_r corresponds to the reaction kinetics and $f_\lambda(q_{(i,j)}(k)) = \frac{q_{(i,j)}(k)}{\alpha_{(i,j)}} \frac{\Delta t}{\Delta z}$ computes the Courant Number

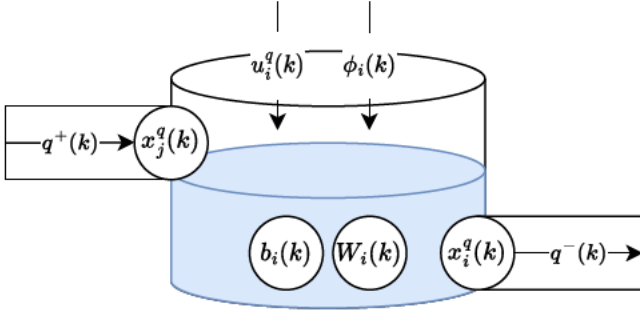


Figure 3: Tank schematic for water quality states.

for a certain water flow, where the Courant-Friedrichs-Lewy (CFL) condition needs to be satisfied, i.e., $0 < f_\lambda(\cdot) < 1$ for the stability of the solution (LeVeque, 1998).

The concentrations at the nodes are considered boundary conditions. Nodes can be considered as “ghost” or “virtual” cells, and their concentration can be computed as the weighted sum of the concentrations that inflow in that cell. For illustration, in the example of Fig. 2, the boundary conditions are $w_0(k) = x_i^q(k)$, and $x_j^q(k) = w_4(k)$.

In case of multiple pipes supplying water to a junction, the concentration at the j -th junction node ($v_j \in \mathcal{V}_j$) is therefore given by

$$x_j^q(k) = \left[\sum_{i \in \mathcal{P}(v_j, k)} q_{(i,j)}(k) w_{(i,j)}^+(k) \right] \cdot \left[\sum_{j \in \mathcal{P}(v_j, k)} q_{(i,j)}(k) \right]^{-1}, \quad (13)$$

where $\mathcal{P}(v_j, k)$ is the set of edges from where water flows into node v_j at time-step k , such that $\mathcal{P}(v_j, k) = \{e_{i,j} \in \mathcal{E} \mid q_{(i,j)}(k) > 0 \vee q_{(j,i)}(k) < 0\}$, and $w_{(i,j)}^+$ is the concentration of the finite volume adjacent to node j .

2.3.3. Tank water quality dynamics

Intuitively, water quality in tanks depends on the concentration of substances within the tank, the reactions between them, as well as the inflows and outflows. Typically, it is assumed that substances are perfectly mixed and uniformly spread within the tank. This model is referred to as a *Continuous Stirred-Tank Reactor* (CSTR). In contrast to finite volumes within pipes, the volume of water within a tank changes over time. For the i -th tank $v_i \in \mathcal{V}_T$, as in Fig. 3, the average concentration $x_i^q(k) = w_i(k)$ is computed as:

$$\begin{aligned} w_i(k+1) &= \frac{1}{b_i(k+1)} \left[(b_i(k) - q^-(k)\Delta t) w_i(k) \right. \\ &\quad + q^+(k)\Delta t x_j^q(k) + f_r(w_i(k), \Theta_r) \\ &\quad \left. + u_i^q(k) + \phi_i(k) \right], \\ b_i(k+1) &= b_i(k) + q^+(k)\Delta t - q^-(k)\Delta t, \end{aligned} \quad (14)$$

where $b_i(k)$ is the water volume in the i -th tank, $q^+(k)$, $q^-(k)$ are the tank inflow and outflow respectively, $f_r(\cdot)$ the function which describes the reactions, $u_i^q(k)$ the controlled inputs

(e.g., disinfection) and $\phi_i(k)$ the uncontrolled inputs (e.g., contamination), such that $\phi \in \Phi$, where Φ is the set of the contamination dynamics possible.

2.4. Water-quality state-space model

By combining equations (12)–(14) for all finite volumes in pipes and tanks, it is possible to express water quality dynamics in a state-space formulation:

$$\begin{aligned} w(k+1) &= A^q(x^h(k)) w(k) + B^q(x^h(k)) u^q(k) \\ &\quad + f_r(w(k), \Theta_r) + \Gamma_\phi \phi(k), \end{aligned} \quad (15)$$

where $w(k) \in \mathbb{R}^{n_w n_z}$ is the vector of concentrations across all volumes (including pipes and tanks), where n_w is the number of substances modeled and n_z is the total number of finite volumes in the network, $A^q(x^h(k))$ is a time-varying square matrix that depends on the numerical solution (discretization) used and the changes in the flows, $u^h(k) \in \mathbb{R}^{n_u}$ is the control input signal, $B^q(x^h(k)) \in \mathbb{R}^{n_w n_z \times n_u}$ is the input matrix which depends on the flows as well as the location of the inputs, $\phi(k) \in \mathbb{R}^{n_\phi}$ is an unknown signal corresponding to n_ϕ possible contaminant injections, $\Gamma_\phi \in \{0, 1\}^{n_w n_z \times n_\phi}$ is a matrix that identifies the source/location (discrete volume) of the contamination, and $f_r(x^q(k), \Theta_r)$ corresponds to the overall reaction dynamics.

The vector of concentrations across volumes $w(k)$ is a complete representation of the network water quality state, however, it is often more useful when analyzing the system to use only the water quality states at nodes, indicated by $x^q(k) \in \mathbb{R}^{n_v}$ (i.e. junctions, tanks and reservoirs). The advantage of this reduced representation is that it maintains the water quality information at the locations of interest (nodes) where consumers are located. The water quality state at nodes can be expressed as a linear combination of the states $w(k)$, derived using (13), as follows:

$$x^q(k) = L(x^h(k)) w(k), \quad (16)$$

where $L(x^h(k))$ is a time-varying output matrix that depends on the hydraulic states, as seen from (13). For a more detailed description of the state-space formulation for water quality see Wang et al. (2021) and Elsherif et al. (2022).

The locations where water-quality sensors are installed are modeled as network nodes. Different kinds of conventional water-quality sensors can be employed to monitor directly the concentration of various substances in the water network, such as chlorine sensors. When the concentration of the substance of interest is directly measured, the measurement vector $y^q(k)$ can be defined using the matrix $C^q \in \mathbb{R}^{n_s \times n_v}$ which identifies the location and state measured as follows:

$$y^q(k) = C^q x^q(k) + \eta^q(k), \quad (17)$$

where $\eta^q(k)$ is the sensor noise term.

Currently, certain online water-quality sensors measure several physio-chemical properties of drinking water such as pH, turbidity, and conductivity, which are indirectly linked

with the concentration of key substances that characterize the quality of water. In this case, a nonlinear function $f_c(\cdot)$ describes the relationship between the measurements and the concentration of different substances defined here as the water-quality states:

$$y^q(k) = C^q f_c(x^q(k)) + \eta^q(k). \quad (18)$$

The function $f_c(\cdot)$ can be considered as a *virtual sensor* (Albertos & Goodwin, 2002), that uses measurements of water properties to give an estimate of the concentration state $x^q(k)$ (Bourgeois et al., 2001; Paepae et al., 2021).

2.5. Impact modeling

In addition to hydraulic and water-quality dynamics, it is crucial for water authorities and decision-makers to be able to predict the public health effects and risks when trace elements, chemical, or biological contaminants enter the drinking water (Chaves et al., 2019). In general, human exposure to contaminants can occur through three major routes: direct ingestion, inhalation through the mouth and nose, and dermal absorption (Delmaar & van Engelen, 2007). Such contaminations can have a measurable impact on society (e.g., people infected, hospitalized, or diseased), as well as the economy (e.g., the loss of work hours, the cost of hospitalizations, reduction of GDP, etc.). A key concept is the “dose/concentration-response” (Xie et al., 2016; World Health Organization, 2017), which describes the likelihood and severity of the health effects, related to the concentration of a certain contaminant. This implies that there is a function that describes the relationship between the concentration of a contaminant with the likelihood that an average person will become infected.

Various studies have investigated the impact of different contaminants. For instance, disinfection by-products, such as *Trihalomethanes*, have been studied and shown to correlate with higher bladder cancer risks (Evlampidou et al., 2020). In general, the impact depends on the contaminant concentration, the volume of water an average person is exposed to via different routes, the average body weight, as well as the number of people who are exposed to the contaminant and are susceptible.

Impact dynamics can be described using a state-space formulation, such that

$$\xi_i(k+1) = \xi_i(k) + f_\xi(x_i^q(k)), \quad (19)$$

$$\psi(k) = f_\psi(\xi(k)), \quad (20)$$

where $\xi_i(k)$ is the “exposure” at node $v_i \in \mathcal{V}$, sampled at time k and $f_\xi(x_i^q(k))$ is a function which computes the exposure as a function of the concentration and other parameters in network graph G (such as water demands by human consumers). Exposure states may correspond to the volume of contaminated water consumed per person, or to the number of people receiving contaminated water above a certain threshold. To compute the total impact $\psi(k)$, an output function $f_\psi(\cdot)$ can be considered which utilizes the state mea-

surements, along with network graph parameters, to compute the total impact metric. This may correspond to an estimate of the number of people infected, hospitalized or diseased, as well as the total economic cost of the contamination event (Rasekh et al., 2014; Shafiee & Zechman, 2013; Rasekh & Brumbelow, 2013; Shahra et al., 2021; Shafiee et al., 2018; Janne et al., 2017).

It is important to note that these impact calculations aim to compute the impact in “real-time” or in simulation studies. In the literature, a significant volume of research has explored the short and long-term impacts of various contaminants, including trace elements (Aghlmand et al., 2021; Çiner et al., 2021; Kicinska & Wysowska, 2021; of Emergency & Response, 1989) and pathogens (Gerba, 2015; He & Huang, 2020; Organization, 2016). Specifically for pathogens, previous studies were used for estimating the risk of infection from cross-connection and backflow events (Viñas et al., 2022), pathogens contamination (Jamal et al., 2020), cryptosporidium, Giardia, and escherichia coli (*E. coli*) from faecal contamination in a drinking water source (Tolouei et al., 2019; Tinelli et al., 2018), faecal contamination (pathogens) events after repairs of drinking water mains (Batista et al., 2018), intentional pathogenic contamination in a WDS from consuming contaminated unboiled drinking water, and inhaling contaminated aerosol droplets when taking a shower (Schijven et al., 2016), potential treatment failures and unexpected variations in water quality and operating parameters of a water treatment plant (Hamouda et al., 2016), legionellae contamination in drinking WDNs (Hamilton & Haas, 2016; Buse et al., 2012), *Campylobacter jejuni* contamination (Van Abel et al., 2014) and intrusion of *E. Coli* bacteria during a sudden power shutdown associated with low or negative pressures (Farahat et al., 2019). In (Janne et al., 2017), models were used to track the effects of water sources on the consumers and a statistical methodology was employed to assess the public health risks of water consumption as well as the economic impacts of increased illnesses. Recently, a detailed review paper focusing on water system risks assessment methodologies was presented by Kombo Mpindou et al. (2022).

2.6. Simulation of water hydraulic, quality and contamination dynamics

One of the key developments in the last few decades is the design of software tools that simulate the hydraulic and quality dynamics of water distribution systems (Clark, 2015). A significant milestone in the field was the release by the U.S. Environmental Protection Agency (EPA) of the open-source hydraulic and quality simulation tool EPANET, which enabled computational water analytics research, worldwide (Rossman, 2000). The EPANET open-source software is currently being supported by the “Open Water Analytics” community¹. EPANET utilizes the *Global Gradient Algorithm* (GGA) for solving the hydraulic equations described in Section 2.2, at each time step (Todini & Pilati, 1988), assuming the consumer

¹<https://github.com/OpenWaterAnalytics/>

demands are known, while taking into account the impact of pressure on the water demands (Siew & Tanyimboh, 2012). For modeling quality, numerical solutions have been implemented to solve the hyperbolic partial differential equations, as well as the dynamics of disinfectants, and the production of disinfection by-products (Rossman et al., 1994; Vasconcelos et al., 1997; Biswas et al., 1993; Clark, 1998; Clark & Sivaganesan, 2002; Boccelli et al., 2003; Hua et al., 1999, 2015). These reaction models can be coupled with the hydraulic (transport) dynamics, to model water quality dynamics at different locations in a WDS (Rossman & Boulos, 1996; Shang et al., 2002; Islam & Chaudhry, 1998; Rossman et al., 1994). This typically involves the use of Eulerian numerical methods (such as the Finite Volume Method (LeVeque, 1998; Grayman et al., 1988)) or Lagrangian numerical methods (event-based (Liou & Kroon, 1987)) for solving the hyperbolic partial differential equations that characterize the system dynamics, as well as solving differential-algebraic equations in the case of multiple reacting substances. The EPANET-MSX software tool was developed to expand EPANET for handling multiple reactions, (Shang et al., 2008). In general, an ecosystem of open-source tools has evolved around EPANET and EPANET-MSX, e.g., the EPANET-MATLAB Toolkit (Eliades et al., 2016), that enables researchers to simulate different scenarios in an effective way by utilizing the large set of MATLAB simulation tools. Other tools include the EPANET-RTX which integrates with SCADA measurements, TEVA-SPOT for sensor placement, CANARY for contamination event detection and WST for responding actions to contamination events (Berglund et al., 2020).

From a modeling viewpoint, contamination dynamics can be described as differential equations to formulate the complete advection-reaction equations, which characterize water quality in distribution networks. In some cases, the contaminant may react with the disinfectant, causing a decrease in the disinfectant concentration which can be observed via measurements (e.g., bacteria with chlorine (Helbling & VanBriesen, 2009), arsenic with chlorine (Burkhardt et al., 2017; Abhijith & Ostfeld, 2021)). A notable effort to model and simulate water quality is with the EPANET-C toolkit (Abhijith & Ostfeld, 2022), where the authors present a collection of various quality dynamics (e.g., chlorine, bacteria, TOC, THMs and others).

3. Preparedness

In this section, we present two research areas that relate to the preparedness phase, before a contamination event has occurred. Typically, this phase includes measures to improve the monitoring capabilities of the system, by optimally selecting the location for installing water quality sensors within the network, as well as establishing a state estimation and water quality control scheme, for retaining the concentration of the disinfectant in the water network within certain bounds and thus reducing the possibility of a contamination event. In the following subsections, we introduce the two research

areas and present a literature review as well as a general problem formulation. The relation to the other cyber and physical components is depicted in Fig. 2 in orange color.

3.1. Water Quality Sensor Placement

The water quality sensor placement problem aims to identify at which locations (nodes) in the water distribution network to install a specific number of online sensors measuring certain water quality parameters, such that one or more objectives are optimized (e.g., minimizing the impact of a contamination event), assuming partially known network topology and water demands, as well as probabilities and dynamics of contamination events. The location of sensors in WDS must be selected in such a way that they detect contamination events as effectively and quickly as possible, so that immediate corrective actions are enabled to protect the population from consuming contaminated water, as much as possible. The water quality sensor placement problem may be formulated as the selection of sensor locations in a water network without any such sensors, however, a common situation is installing *additional* water quality sensors. Typically, it is not feasible to relocate existing sensors due to significant installation costs.

The sensor placement problem is typically posed as an optimization problem with the location and number of sensors to install as the decision variable. Let $\mathcal{V}_s \subseteq \mathcal{V}$ be the set of all nodes where it is possible to install a water quality sensor, and $s \in \{0, 1\}^{n_v}$ be the vector indicating the sensor locations, i.e., $s_i = 1$ if a sensor is located at node $v_i \in \mathcal{V}_s$, and $s_i = 0$ otherwise; $n_s = |s|$ is the total number of sensors. The output matrix in (17), is given by $C^q \in \{0, 1\}^{n_s \times n_v}$, which can be computed based on s , as follows: let $\sigma = \{s_j \in s | s_j = 1\}$; for each element in the set σ for which $\sigma_i = s_j$, the (i, j) -th element in C^q is '1', otherwise it is '0'. Let also $\mathcal{V}_0 \subseteq \mathcal{V}$ be the set of all nodes where quality sensors are already installed.

Specific sensor configurations are evaluated using single- and multi-objective approaches. Let $J(s) = [J_1(s), \dots, J_{n_o}(s)]^\top$, be a multi-objective function with n_o objectives, where $J_i(s) : \{0, 1\}^{n_v} \mapsto \mathbb{R}^+$ is a function which, based on the sensor placement defined by s , computes a value with respect to a possible contamination scenario. The objective functions may correspond to minimizing an impact metric as in (20), or other suitable metrics. An example of an impact metric is the number of people exposed to a contamination event prior to its detection, referred to as the *population exposed* (Berry et al., 2005). Another metric, *time-to-detection*, is the time interval between a contamination event and its detection (Kumar et al., 1999). Note that impact metrics are significantly affected by differences in flow patterns, by different contamination event locations, and by variations in sensor placements. *Network coverage* is also an important metric since certain contamination events might pass undetected due to the limited number of monitoring sensors. Other metrics include the amount of contaminated water consumed, the total length of the contaminated pipeline, etc. For an extensive review of the various sensor placement performance objectives see (Adedoja et al., 2018a). A critical review of the differ-

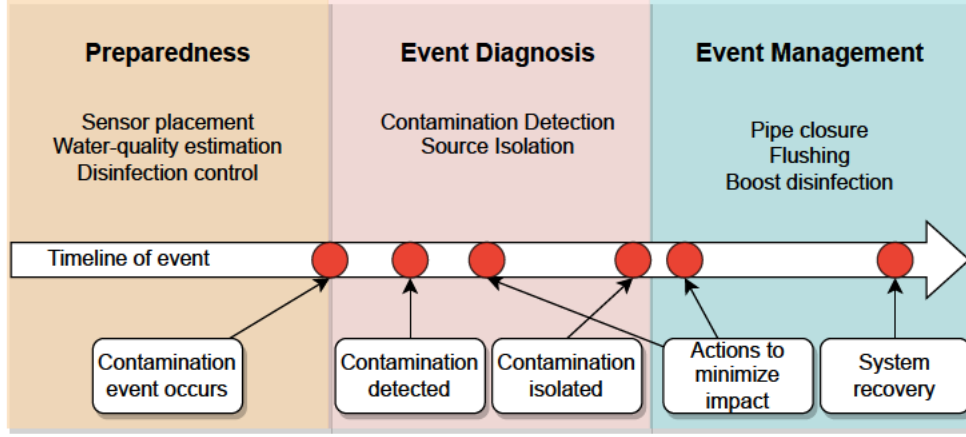


Figure 4: Contamination event timeline and management phases to improve water security.

ent objective approaches can also be found in (Rathi & Gupta, 2014).

Moreover, multiple constraints can be considered, such as the maximum number of sensors, the allowable locations for sensor installation, etc. Let then $g(s)$ be a vector of n_g constraint functions, such that $g(s) : \{0, 1\}^{n_v} \mapsto \mathbb{R}^{n_g}$,

Modeling contamination events in WDS is challenging due to the high uncertainty in determining the contamination type, location, concentration, occurrence time, duration, or a combination of the above. Moreover, uncertainty related to the spread of contaminants is present. In particular, contaminants in water networks are transported by water flows which vary continuously over time (i.e., changing flow rate and direction due to consumer demands). A typical way to hedge against uncertainty is within a stochastic optimization framework. More precisely, to model this uncertainty within a decision-making framework by assuming complete knowledge of the probability distribution of each of the possible contamination events, and by optimizing a function of the cost (e.g., minimizing the maximum detection time, and/or minimizing the number of people exposed to a contamination event prior its detection during the worst-case contamination scenarios). In practice, to identify a sensor placement that is optimal in a general sense, a finite number of contamination scenarios are evaluated.

Let $\gamma \in \{0, 1\}^{n_w n_z}$ be a binary vector indicating all possible locations where a certain contaminant is injected, such that the contamination source matrix Γ_ϕ in (15) is given by $\Gamma_\phi \in \{0, 1\}^{n_w n_z \times n_\phi}$, which can be computed based on γ , as follows: let $\kappa = \{\gamma_j \in \gamma | \gamma_j = 1\}$; for each element in the set κ for which $\kappa_i = \gamma_j$, the (i, j) -th element in Γ_ϕ is '1', otherwise it is '0'. For the computation of the optimal sensor placements, we need to consider different contamination scenarios, however, some events may be more likely than others (e.g., contaminations may occur at locations nearby sewerage networks or at cracks in the pipes). This can be considered in the formulation by assigning probabilities. We denote by $\mathbb{P}_\phi(i, j) \geq 0$ as the probability that the i -th contaminant

may enter the network at the j -th finite volume, such that

$$\sum_{i \in \{1, \dots, n_w\}} \sum_{j \in \{1, \dots, n_z\}} \mathbb{P}_\phi(i, j) = 1. \quad (21)$$

When multiple scenarios are evaluated, appropriate metrics are needed to aggregate the cost and uncertainty of each scenario to achieve an overall optimized sensor placement decision. Often in the literature the *average* function is used to analyze the outcomes of multiple scenarios; however, this may not be a suitable metric as it does not take well into account extreme cases. A more suitable metric is the coherent risk measure *Conditional Value-at-Risk* (CVaR) (e.g., the CVaR of the population exposed with 95% confidence), or the maximum (worst-case) time to detection.

A general sensor placement optimization problem is formulated here as follows:

$$S = \arg \min_{s \in \{0, 1\}^{n_v}} \{\mathcal{R}_\mathbb{P}[J(s)]\} \quad (22)$$

s.t.

$$s_i = 1, v_i \in \mathcal{V}_0$$

$$|s| = n_s$$

$$\mathcal{R}_\mathbb{P}[g(s)] \leq 0$$

where $\mathcal{R}_\mathbb{P} : \mathcal{Z} \mapsto \mathbb{R}$ denotes a (component-wise) real-valued functional under \mathbb{P}_ϕ , where \mathcal{Z} is a linear space of measurable functions on a measurable space. The functional $\mathcal{R}_\mathbb{P}$ is used to account for the uncertainty in the outcomes of the decision, given a fixed probability measure \mathbb{P}_ϕ .

Finding a solution to the sensor placement problem via different optimization approaches has been widely studied in the sensor placement literature. A comprehensive review of the existing optimization-based approaches can be found in (Hart & Murray, 2010) and (Adedola et al., 2018a). The sensor placement problem can be categorized into three different optimization approaches, namely, deterministic optimization, stochastic optimization, and robust optimization. A summary of indicative literature on the use of these approaches on the sensor placement problem is given in Table 1.

References	Optimization Method
Krause et al., 2008; Kumar et al., 1999	Deterministic
Berry et al., 2006; Shastri & Diwekar, 2006; Weickgenannt et al., 2010, Rico-Ramirez et al., 2007; Cozzolino et al., 2011; Comboul & Ghanem, 2013	Stochastic
Snyder, 2006; Carr et al., 2006; Sela & Amin, 2018; Langowski et al., 2012, Watson et al., 2009; Xu et al., 2010	Robust

Table 1: Sensor placement optimization-based approaches.

Deterministic optimization approaches are developed and applied to assume no uncertainty is present. For such approaches, it is possible to evaluate the merit of any given sensor placement just from a single simulation path of the water distribution network. In the case of stochastic optimization, however, this is not possible. Although stochastic optimization approaches have been developed significantly over the last years, their applicability is limited due to the requirement of having complete and accurate distributional information about the different contamination event scenarios. This problem becomes particularly severe in the case of multiple sources or contributors of uncertainty each one characterized by its own probability distribution, as described above. In robust optimization, on the other hand, it is assumed that no distributional information exists about the uncertainty of the different contamination events, and a worst-case cost, over an uncertainty set, is optimized.

3.2. Estimation and Disinfection Control

In most of the world, disinfectants such as chlorine or chloramines, are used in prescribed concentrations to maintain drinking water quality, by preventing bacteria growth and neutralizing chemical agents. According to the World Health Organization (World Health Organization, 2017), a disinfectant residual needs to be sustained throughout the drinking water network, such that it is sufficient to deactivate water-borne pathogens, with a minimum target concentration of $0.2 \frac{\text{mg}}{\text{L}}$ at the point of delivery and $0.5 \frac{\text{mg}}{\text{L}}$ for high-risk circumstances (World Health Organization, 2008). It is common practice to supply water with a few tenths of a milligram per liter of “free chlorine” residual.

In general, disinfectants react and destroy pathogens such as *Escherichia coli* (*E. coli*), as well as other organic matter. A critical trade-off of the disinfection process is the creation of Disinfection By-Products (DBPs), such as Trihalomethanes (THMs) and Haloacetic Acids (HAA), which are linked with various health issues when consuming water with high concentrations of these substances for a long period (Mouly et al., 2010). In addition to the DBPs, which are regulated, a large number of substances have been shown to be of concern, as they are associated with various health issues. These include emerging contaminants, pharmaceuticals, persistent and mobile contaminants, per- and polyfluoroalkyl substances (PFAS), and others (Richardson & Postigo, 2012).

A key challenge in water quality is to monitor via sampling or online sensors the disinfectant substances and their disinfection by-products to guarantee compliance with the rel-

evant water quality regulations (Eliades & Polycarpou, 2010). Disinfectant concentration is affected by chlorine injection at specific locations in the network. However, due to the dependence of water-quality dynamics on hydraulic dynamics, as seen in Fig. 2, hydraulic actuators may also be used to affect the disinfectant concentration.

The task of estimating the concentration of these substances throughout the network, in real-time, is challenging due to the highly uncertain available models and the small number of measurements (Vrachimis et al., 2021). Mathematical models that estimate disinfectant concentration in the water at different parts of the network are typically complex and inaccurate, especially in large-scale networks. Disinfectant reaction dynamics are usually only partially known and the reaction parameters may vary due to the use of water originating from diverse sources or due to fluctuations in temperature (Monteiro et al., 2017). The largest source of uncertainty is the water flows due to the unknown water demands (Pasha & Lansey, 2010).

Some approaches have exploited the approximate periodicity of water demands, and by extent, water flows over a 24-hour period, to create time-varying water-quality models using online learning (Polycarpou et al., 2002). These approaches omit the information provided by hydraulic sensors, which may lead to significant estimation errors. The most common approach to deal with modeling uncertainty due to water flows is to utilize them as an input to the water-quality model, by assuming that the hydraulics are constant (steady-state) and known over a period of time called the hydraulic step (Rossman, 2000).

A *Hydraulic State Estimator* (HSE) is a function that uses the hydraulic model in (5), the known hydraulic inputs $u^h(k)$, demand estimates $\hat{d}(k)$, as well as the current measurements $y^h(k)$ to calculate the most suitable hydraulic-state estimate $\hat{x}^h(k) \in \mathbb{R}^{n_v+n_e}$ as follows:

$$\hat{x}^h(k) = f_{HSE}\left(y^h(k), u^h(k), \hat{d}(k); \mathcal{G}\right) \quad (23)$$

However, hydraulic state estimates obtained when solving the flow network during this period may be inaccurate due to errors in water demand estimates and other model uncertainties, such as unknown pipe roughness coefficients (Hutton et al., 2014). Recent efforts in developing hydraulic-state estimation methodologies try to overcome the challenge of using a small number of measurements compared to the system hydraulic states to obtain reliable hydraulic-state estimates, while also calibrating model parameters and water de-

mands (Tshehla et al., 2017; Díaz et al., 2017). Methodologies have also been developed which are able to quantify the hydraulic-state estimation error using set-based (Vrachimis et al., 2019; Rego et al., 2021) and probabilistic approaches (Wang et al., 2022b). Assuming the existence of hydraulic state estimates, a disinfectant concentration estimator can be constructed using the hydraulic states as input to the water-quality model.

Water-quality models include a large number of dynamic states, as seen in (15), and initial conditions are needed for the calculation of the current state. However, the initial conditions throughout the network are typically unknown. A more convenient approach is to assume that enough time has passed such that a substance of interest can be estimated using only the known input of the system. Water-quality models exhibit large time delays from input to output, thus the effect of inputs $U^q(k, k-m) \triangleq \{u^q(k), u^q(k-1), \dots, u^q(k-m)\}$, $m \in \mathbb{Z}^+$ are needed for the calculation of the current water-quality state, where m is the memory of the system corresponding to the maximum time that a disinfectant remains in the network before decaying or consumed. Let the flow estimates for all the network links for the memory of the system m be defined by $\hat{Q}(k, k-m) = \{\hat{q}(k), \hat{q}(k-1), \dots, \hat{q}(k-m)\}$. A water Quality State Estimator (QSE) should be able to calculate the disinfectant concentration states $\hat{x}^q(k)$, given the available measurements $y^q(k)$, the sequence of past inputs $U^q(k, k-m)$ and the sequence of estimated water flows $\hat{Q}(k, k-m)$, as follows:

$$\hat{x}^q(k) = f_{QSE}(y^q(k), U^q(k, k-m), \hat{Q}(k, k-m); \mathcal{G}) \quad (24)$$

Note that typically, only a subset of water-quality states are observable given the available measurements. Moreover, an estimator intended for water-quality monitoring should be able to quantify the errors that arise from the combined hydraulic and water-quality parameter uncertainty. One approach to do this is proposed by Eliades et al. (2015), where chlorine concentration bounds are calculated through Monte-Carlo Simulations (MCS) using the Lagrangian model of EPANET software and used as contamination detection thresholds. A notable set-membership application to water quality was proposed by Łangowski & Brdys (2007, 2017), where a distributed interval-model of water quality is formulated following the Lagrangian modeling approach, based on known-bounded uncertainties on water flows and water-quality parameters. The researchers go a step further by proposing an interval observer structure for chlorine residual estimation. In Vrachimis et al. (2021) a methodology to calculate chlorine concentration bounds at nodes of WDS is proposed, using an input-output modeling approach, hydraulic-state estimates as input and incorporates water-quality measurements in a parameter estimation scheme.

Feedback control of disinfectant concentration states aims to keep them within pre-defined levels and reduce disinfection by-products health impacts in the long term (e.g., in the case of Trihalomethanes). This is achieved by designing control algorithms that are able to regulate the disinfectant concentration (e.g., free chlorine concentration) at a minimal

level that guarantees safety, at different locations in the network by controlling disinfectant injection at specific locations in the network, through the use of so called booster disinfection stations (Polycarpou et al., 2002; Zhong Wang et al., 2006; Kang & Lansey, 2010; Constans et al., 2003; Duzinkiewicz et al., 2005; Sakarya & Mays, 2000; Boccelli et al., 1998).

The booster disinfection station placement problem (Islam et al., 2017; Tsitsifli & Kanakoudis, 2021; Islam et al., 2013; Seth et al., 2017; Ohar & Ostfeld, 2014; Hernandez Cervantes et al., 2016) aims to deal with the fact that, commonly, large quantities of disinfectants are released at the sources of the network, making it infeasible to distribute the disinfectant evenly in the network. Uneven disinfectant distribution results in unpleasant taste and odor problems near the disinfectant injection locations, and excessive disinfectant by-product formations (DBPs) at the far network ends. To optimally place the disinfectant boosting stations, models of normal system operation are constructed using extended-period hydraulic simulation conditions and utilizing multi-species water-quality models. The optimization problem formulated for finding the location and settings of booster stations is solved using various optimization approaches such as genetic algorithms (Ohar & Ostfeld, 2014) and convex optimization (Pecci et al., 2022).

A significant challenge when attempting to regulate the concentration of a disinfectant in a WDS is the large and uncertain delay between the time that a disinfectant chemical is injected at the input node and the time that the concentration is measured at the monitored output nodes. One of the first studies addressing this challenge is (Polycarpou et al., 2002), where the proposed Input-Output (I-O) model of water quality is used to adaptively learn periodic parameters to characterize the time-delay. The input-output approach is a model-reduction technique that overcomes the challenge of using a small number of measurements compared to the system states, to design a water-quality control algorithm (Zhong Wang et al., 2006). Xie & Brdys (2015) used a simulation-based model to compute hydraulics and water-quality, while nonlinear Model Predictive Control (MPC) within a hierarchical control framework is then used to control disinfectant levels. Xie et al. (2018) use a Multi-Input Multi-Output modeling approach, which considers also the formation of DBPs. More recently, Wang et al. (2021), derived a state-space model of chlorine concentration in WDS, which includes all water-quality states; i.e., the concentration at all the links and nodes of the network. A drawback of this state-space formulation in practice is the lack of observability for estimating the complete state vector, and thus model-reduction techniques are needed to calculate a control law (Wang et al., 2022a).

Model predictive control is one of the key control approaches that has traditionally been applied to chemical processes with large time delays; this has also been investigated for the regulation of water-quality (Xie & Brdys, 2015; Wang et al., 2021; Brdys et al., 2001; Duzinkiewicz et al., 2005). Let us define the prediction horizon $N \in \mathbb{N}^+$ for which we want to estimate, for all time steps within the prediction horizon

$\tau = \{k+1, \dots, k+N\}$, the output $y(k) \forall k \in \tau$. Due to the large time delays in the system, the output is affected by the past inputs $U^q(k, k-m)$. Considering that future inputs $u^q(k) = 0, k \in \tau$, the water quality estimator of (24) is used to estimate the effect of past inputs to the future outputs, indicated here by $y^0(k) \forall k \in \tau$. Note that in (24) an estimate of the future water flows is used, calculated using (23). The effect of past inputs on future outputs $y^0(k)$ are then used to adjust a pre-defined tracking/reference signal $r(k)$ as follows:

$$r_y(k) = r(k) - y^0(k), \forall k \in \tau = \{k+1, \dots, k+N\}. \quad (25)$$

The optimization formulation for calculating the optimal control input sequence $U^*(\tau) \triangleq \{u^q(k+1), \dots, u^q(k+N)\}$ for all time steps in the prediction horizon can then be calculated using the following optimization problem:

$$U^*(\tau) = \arg \min_{u^q} \sum_{k \in \tau} (|\hat{y}(k) - r_y(k)| + |\Delta u^q(k)|) \quad (26)$$

s.t.

$$r_y^l \leq \hat{y}(k) \leq r_y^u, \quad \forall k \in \tau \quad (27)$$

$$u^l \leq u^q(k) \leq u^u, \quad \forall k \in \tau \quad (28)$$

$$\tau = \{k+1, \dots, k+N\} \quad (29)$$

where $\hat{y}(k)$ is calculated using (24), r_y^l and r_y^u are the reference lower and upper bounds respectively, u^l and u^u are the control input lower and upper bounds respectively. Typically, the rapid changes of disinfectant injection should be avoided to reduce equipment wear, thus $\Delta u^q(k) \triangleq u^q(k) - u^q(k-1), \forall k \in \tau$ is included in the objective function.

4. Event Detection and Diagnosis

This section focuses on the relevant tasks after a contamination event has occurred. Specifically, the problem of contamination detection is examined, as well as the problem of isolating the source of the contamination event. Typically, these two procedures must be effective in producing results as fast as possible after the occurrence of a contamination event, to allow first responders and water authorities to gain better situational awareness and to make informed decisions to manage the event. Figure 2 illustrates the relation between the relevant components in red color.

4.1. Contamination detection

Accidental or intentional contamination events threaten the quality of drinking water, and consequently, the public health. For these reasons, there is significant research activity on the problem of contamination event detection, which can be summarized as the problem of determining whether a contamination event has occurred based on certain water quality sensor measurements, taking into account hydraulic and water quality models, control inputs, and uncertainties. Various possible approaches have been identified in the literature (Zulkifli et al., 2018), including but not limited to:

(i) sample-based analyses, (ii) in-line sensor-based monitoring, and (iii) algorithmic model-based techniques. Choosing one over another depends on the particular type of contaminant that intruded into the water distribution network, and consequently, on the particular type of water analysis that is required to be performed (i.e., qualitative, quantitative, or both) for fast and reliable detection.

A key trade-off in contamination detection is to detect contamination events as soon as possible, while at the same time avoid false alarms associated with the estimation error exceeding the adaptive threshold due to modeling uncertainty, even though there is no contamination. A key issue in practice is the design of a suitable adaptive threshold signal $\epsilon(k)$ such that it is small enough to prevent false negatives, while at the same time it is large enough to avoid false positives (false alarms). In general, the problem can be defined mathematically as follows: let $e(k)$ be the estimation error, such that

$$e(k) = y^q(k) - C^q f_c(\hat{x}^q(k)), \quad (30)$$

where $y^q(k)$ are the sensor readings, $\hat{x}^q(k)$ are the water-quality state estimates as described in (24), assuming no contaminant intrusions, i.e., $\phi(k) = 0$. Moreover, let C^q be the output matrix, and $f_c(\cdot)$ is a function that relates the estimates $\hat{x}^q(k)$ with the sensor readings, if required. Let $\bar{\epsilon}(k)$ be an adaptive threshold that may be computed using historical data of normal water quality measurements, or based on the knowledge regarding the uncertainty in modeling, inputs, measurements, as well as the safety levels of waterborne substances. Finally, let $f_e: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a function that compares the estimation error $e(k)$ with the threshold $\bar{\epsilon}(k)$, such that, if it is positive, there exists an input signal $\phi(k)$ which can be considered a contamination event. In practice, this can be $f_e \equiv |e(k)| - \bar{\epsilon}$, however other functions could be more complex, to consider multiple thresholds, as in (Eliades et al., 2015). Therefore, an alert signal $\mathcal{A}(k)$ can be generated, as follows:

$$\mathcal{A}(k) = \begin{cases} \text{True} & f_e(e(k), \bar{\epsilon}(k)) > 0 \\ \text{False} & \text{otherwise.} \end{cases} \quad (31)$$

Currently, most of the literature is focused on developing techniques which allow monitoring of water quality and detecting contamination events in real-time, based on signals generated by low-cost sensors. Different kinds of sensors, possibly located at different locations of the water distribution network are employed for monitoring water quality parameters such as chlorine concentration, pH, turbidity, conductivity, total organic carbon, etc., (Lambrou et al., 2012). By processing and analyzing these signals, the aim is to detect potential contamination events and raise an alarm signal.

Contamination detection techniques relying on Artificial Intelligence (AI), and the classification of water quality data into anomalous categories, are based on (i) artificial neural networks and deep learning networks (Perelman et al., 2012; Olikier & Ostfeld, 2015; Arad et al., 2013; Qian et al., 2020), and (ii) support vector machines and regression trees (Fitore et al., 2019; Wang et al., 2018; Nguyen & Logofătu, 2018; Yan-jun &

Qian, 2012). An extensive review and comparison between different AI methods for water quality analysis can be found in (Kang et al., 2017). Another type of model-based event detection technique relies on Monte-Carlo simulations to obtain bounds on the chlorine concentration at various sensing locations (Eliades et al., 2014). Research on the periodic fluctuation of chlorine as a normal pattern for event detection, using a singular value decomposition based-method, was investigated in (Zhao et al., 2015). Further research on event detection using chlorine concentration as a surrogate parameter for contaminants can also be found in (Jonkergouw et al., 2004; Hua et al., 1999). A more recent model-based approach (Vrachimis et al., 2020) explores the concept of active fault detection to confirm the existence of a contaminant in the network by re-routing suspected contaminated water to the available sensors.

4.2. Source isolation

In some cases, a contamination event needs to be forensically investigated, to identify the source of the contamination and hypothesize the contamination cause (e.g., bacteria due to the use of cast-iron pipes, surface runoff infiltration after a flood, etc.) (Danon-Schaffer, 2001). In the literature, the problem is typically defined as an “inverse” problem, i.e., given one or more water quality measurements (or alerts) at some sensor or sampling locations, identify which are the most probable locations where one or more contamination events could have occurred at, as well as at which times and with what magnitudes (Adedola et al., 2018b).

At an abstract level, the problem can be described mathematically as finding the smallest set of the most probable source locations which are in agreement with the measurements, considering the uncertainties in the hydraulic and quality dynamics, as well as the sensor noise. Let $\Gamma' \subset \{0, 1\}^{n_w n_z}$ be the set of the possible contamination source locations, and $\mathbb{P}_\phi(i, j) \geq 0$ is the probability that the i -th contaminant enters the network at the j -th finite volume. Therefore,

$$\begin{aligned} \Gamma_s = \arg \max_{\gamma \in \Gamma'} & \left[\mathcal{R}_\mathbb{P} \left(\mu \left(\gamma, y^q(\tau), \hat{d}(\tau), u^h(\tau), u^q(\tau) \right) \right) \right] \quad (32) \\ \text{s.t.} & \\ & |\gamma| \leq n_\gamma \\ & \tau \in \mathcal{T}, \end{aligned}$$

where γ is the decision variable vector which indicates the location (and the type) of the contaminant and $\mu(\cdot)$ is a function that calculates a metric of how well the existence of a contamination event at the location described in γ satisfies the formulated problem, given demands \hat{d} , inputs u^q and u^h , and sensor measurements y^q , over a time period $\tau \in \mathcal{T}$. Moreover, n_γ is the maximum number of simultaneous contaminant sources. Finally, the function $\mathcal{R}_\mathbb{P}(\cdot)$ computes the probability of that event.

A key challenge is to relate output measurements with certain contaminant inputs. Usually, computational tools

such as EPANET are employed, which simulate contamination events forward in time. A computational approach, such as the Particle Backtracking Algorithm (PBA) can also be utilized (Shang et al., 2002). PBA can be considered an “inverse simulator” that computes the concentrations upstream of the location of interest, given a known model, assuming known hydraulic dynamics (water demands) and contaminant reaction models. Other ways have been proposed to model these relationships, such as model trees (Preis & Ostfeld, 2006).

During a source isolation investigation, it is possible to use sensors that are installed in the network, as well as manual water samplings taken at different parts of the network, at certain time instances. Multiple positive or negative (usually binary) signals can be considered for the isolation problem (Mann et al., 2012; Kumar et al., 2012; Costa et al., 2013). This can be extended to identifying at which locations to sample and when, to help in reducing the risk and search area (Eliades et al., 2011; Eliades & Polycarpou, 2012).

The problem of (32) can be formulated as a mathematical program, to minimize the error between measured measurements and recorded values. For example, Mixed-Integer Quadratic Program (MIQP) formulations have been considered (Laird et al., 2005, 2006), and similarly, in (Cristo & Leopardi, 2008) an optimization method was proposed that considered the use of a “pollution matrix” as described in (Kessler et al., 1998).

One of the most common solution methodologies when addressing the contamination isolation problem is to consider a simulation-optimization approach. For this, a simulation engine (e.g. EPANET) or other equivalents, can be used, to evaluate different potential scenarios using a suitable optimization methodology, e.g., evolutionary algorithms, logistic regression and others. (Guan et al., 2006; Preis & Ostfeld, 2007; Liu et al., 2012b; Yan et al., 2019). Specifically, computational intelligence methods were widely used in the literature to solve this problem (Vankayala et al., 2009; Zechman & Ranjithan, 2009; Liu et al., 2011; Xuesong et al., 2017; Lučin et al., 2021b). It is important to note that there may be multiple possible solutions, and a systematic way is required to handle this issue (Liu et al., 2008).

Two factors that affect the problem formulation, are the use of more information regarding sensor measurements, and the uncertainties. In (Preis & Ostfeld, 2008a), fuzzy variables are considered for the water quality sensors, and in addition, in (Preis & Ostfeld, 2011), uncertainties are considered through the use of Monte-Carlo simulations. Imperfect sensors, e.g., sensors that give a binary signal indicating the presence of a contamination or fuzzy sensors that measure concentration within a set range, were considered in (Wagner et al., 2015).

Another class of methodologies considers the problem of estimating, in a probabilistic framework, the possible contamination sources. Bayesian networks were first utilized by Perelman & Ostfeld (2012), to compute the most likely sources. This work has also been extended to consider in-pipe mobile sensors (Sankary & Ostfeld, 2019). Other approaches have also been proposed which describe, proba-

bilistically, some of the dynamics (De Sanctis et al., 2010; Neupauer et al., 2010; Wang & Zhou, 2017; Ortega et al., 2020; Cai & Ye, 2021). A comparison of the Bayesian and the optimization-based methods, was presented in (Seth et al., 2016). Other Bayesian methodologies also consider external information, e.g., the consumer complaints (Tao et al., 2012a).

Statistical and machine learning have also received significant attention in contamination isolation research, using data mining (Huang & McBean, 2009), pattern recognition (Tao et al., 2012b), entropy-based methods (Propato et al., 2010), greedy and local search algorithms (Liu et al., 2012a; Adedola et al., 2020), density-based clustering algorithms (Mandel et al., 2021), nearest neighbor and random forest algorithms (Barros et al., 2022).

Recently, researchers investigated the challenge as an estimation problem, e.g., using the restart Ensemble Kalman Filters (Butera et al., 2021). Moreover, various methods based on artificial neural network structures (convolutional neural networks, graph convolutional networks, deep neural networks etc.) have been employed, sometimes in conjunction with methods for regression such as random forests (Sun et al., 2019; Grbčić et al., 2020; Hu et al., 2020a; Lučin et al., 2021a; Zhou et al., 2021; Qian et al., 2021).

5. Emergency Event Management

An important concern of water utilities and policymakers is the management of emergency response in the case of intrusion of contaminants such as protozoans, bacteria, and toxic chemicals into the water. In this case, decision-makers in water boards need to come up with solutions for reducing and, finally, eliminating the contamination in order to ensure that it is consumed by as few people as possible.

Previous studies have listed a number of response actions that can be used during emergency management situations in contaminated WDS. These include: water network partitioning for the creation of isolated network sections (Ciaponi et al., 2018, 2019), pipe valve closure to change the network topology and isolate contaminated areas (Hu et al., 2022; Palleti et al., 2018), flushing water from hydrants to remove contaminated water from the network (Moghaddam et al., 2022, 2020; Fasaee et al., 2020; Hu et al., 2020b), warning consumers to change water use (Shafiee & Zechman, 2013; Kadinski et al., 2022), filtration of water (Cossali et al., 2016; Montenegro-Ayo et al., 2020; Silvestry-Rodriguez et al., 2007), boosting disinfection (Parks & VanBriesen, 2009; Karamouz et al., 2022; Van Bel et al., 2019; Helbling & VanBriesen, 2009), and slug-feed disinfection (Qiu et al., 2021). A combination of the aforementioned strategies can be applied to have a more effective response strategy.

In Fig. 2 the relation of the management component with the different control systems is depicted, in relation to the other elements of the system.

The problem can be described mathematically as an optimization problem of finding the reference signals for the

quality and hydraulic control, as well as the consumer demands, to minimize the risk of a severe impact. Let $r(k) = [r^q(k), r^h(k), r^d(k)]^\top$ be the reference signal vector, and $r(k) \in \mathbb{R}$ is the space of all feasible reference signals. Note that in the case of demands, the reference signal can be a linguistic instruction such as “do not consumer water”, “toilet use only” or “boil”; in the case of hydraulic reference signals this can correspond to manipulating valves or pumps, whereas, in the case of quality reference signals, it may refer to the disinfection concentration.

$$r(k) = \underset{r(\tau) \in R}{\operatorname{argmin}} \rho(r(\tau), \psi(\tau)) \quad (33)$$

s.t.

$$\begin{aligned} g(\hat{x}_j^q(k), \hat{x}_j^h(k)) &= 0 \\ \tau &\in \{k, k+1, \dots, k+T_f\} \end{aligned}$$

where $\psi(\cdot)$ is a function for evaluating the total impact of contamination as in (20), ρ is a risk function, $g(\cdot)$ are constraints which must be satisfied on hydraulic and water-quality states and T_f is the number of discrete steps in the future considered in the problem.

Contamination propagation in the network is a dynamic process that is affected by the hydraulics (water demands), the hydraulic actions by operators, and disinfection inputs. Typically the optimal strategy is difficult to be calculated due to unknown hydraulic and water-quality dynamics in large-scale WDS and the plethora of control actions that may need mixed-integer optimization approaches to calculate, e.g., in the case of closing valves to isolate contaminated areas. Typically, the water utilities prepare several response protocols *a priori*, which are evaluated using simulation and optimization models. Therefore, simulation-optimization models are used to evaluate the effectiveness of the strategies for emergency response events. A suitable intervention strategy is calculated using single or multiple objectives, using computational intelligence methods, including evolutionary computation (Afshar & Najafi, 2014; Rasekh & Brumbelow, 2014, 2015; Hu et al., 2020b; Fasaee et al., 2020; Qiu et al., 2021; Hu et al., 2022).

The main objective during the operational and emergency response to a contamination event is to minimize in some form the contamination impact, generally defined in (20). The impact metrics used are typically calculated for a predefined operational period and have the following form:

$$\psi = \sum_{j=1}^n \sum_{k \in \tau} f_\psi(\hat{x}_j^q(k), \hat{x}_j^h(k), \theta_{v,j}) \quad (34)$$

where j indicates the consumer node and n the total number of consumer nodes; $\tau = \{k, k+1, \dots, k+T_f\}$ is the time period of interest; \hat{x}_j^q is an estimation of the contaminant concentration at node j ; \hat{x}_j^h is an estimation of the hydraulic state at node j ; $\theta_{v,j}$ is a parameter vector corresponding to node j ; and $f_\psi(\cdot)$ is a function for evaluating the impact at node j at time k .

Different impact metrics used in the literature include: Minimization of the number of contaminated nodes, using a function that defines whether a node is contaminated or not based on a threshold applied to the contaminant concentration, or minimization of the total contaminant concentration at nodes (Afshar & Najafi, 2014; Bashi-Azghadi et al., 2017; Moghaddam et al., 2020); Minimization of the number of exposed individuals, using a function that calculates the number of people affected at a node based on the population at the node (Zechman, 2013; Moghaddam et al., 2022); Minimization of the volume of contaminated water consumed, using a function of the demand when the concentration at a node is above a threshold (Preis & Ostfeld, 2008b; Guidorzi et al., 2009; Rasekh & Brumbelow, 2014; Bashi-Azghadi et al., 2017; ?); Maximization of the flushed contamination mass by the hydrant set, using a function that calculates the contaminant mass that exits the network from hydrant flushing locations (Fasaee et al., 2020). A notable approach uses the objective to maximize the number of protected consumers when the control action is warning the consumers through “warning tours” about how to use water safely before they are exposed to the contaminant (Shafiee & Berglund, 2016; Kadin-ski et al., 2022).

Many of these approaches use secondary objectives which make the response more efficient. Minimizing the operational costs in contamination management strategies is mostly related to the cost of control actions, such as valve closure and hydrant opening. There are different factors for computing the operational costs, such as the type of valves and hydrant that need to be operated, the distance to the operational location, and the response crew salary (Hu et al., 2020b). The objective may be to minimize only the number of valves and hydrants which need to be operated in the contaminant network (Bashi-Azghadi et al., 2017; Guidorzi et al., 2009; Preis & Ostfeld, 2008b), or include also a distance-based model to actuation locations (Hu et al., 2020b). Minimization of the total time required to fill the network with the desired disinfectant concentration is another secondary objective, where a minimum disinfectant concentration required for network decontamination should be imposed at the contaminated locations. This objective reduces the impact of network closure on the public. Moreover, minimizing the total amount of disinfectant inserted in the network reduces the cost associated with the amount of disinfectant used. Finally, minimizing the total amount of water drained from a contaminated section of the network reduces the amount of water that needs to be disposed of or treated (Qiu et al., 2020, 2021).

Certain operational constraints, defined as $g(\hat{x}_j^q(k), \hat{x}_j^h(k))$ in (33), should be satisfied when trying to achieve the aforementioned objectives. The most important constraint is maintaining sufficient pressure conditions in the network such that the system remains operational and water demands are satisfied, i.e., $p_i(k) > p_{min}$ for all k , where p_i is the pressure at node i and p_{min} is the minimum sufficient pressure. Some works assume $p_{min} = 0$ during emergency response

to also limit consumption of contaminated water; however, negative pressures should be avoided since they may become sources of additional contaminants into the network. Additional operational constraints include flushing water only from pre-defined drainage locations, inserting disinfectants only at disinfection stations, and maintaining a minimum disinfectant concentration at each contaminated node of the network.

Real-time contamination response is still an open challenge. Automated emergency response requires real-time hydraulic and water-quality state estimation to cope with the changes in the system during contamination events (Lifshitz & Ostfeld, 2019; Berglund et al., 2020).

6. Conclusions and Future Challenges

In this paper, we have presented a comprehensive review of the state-of-the-art in contamination event diagnosis methodologies with emphasis on model-based monitoring and control aspects. A unified modeling framework was presented, integrating hydraulic, quality, contamination and impact dynamics. Next, different classes of methodologies were presented as part of three main phases in water security: preparedness, event detection and emergency event management. Specifically, overviews of the state-of-the-art and problem formulations are suggested for: the water quality sensor placement problem, the estimation and disinfection control problem, the contamination detection and source isolation problems, as well as the contamination management problem. The overall architecture, where all the physical and interconnected cyber components are mapped, is illustrated in Fig. 2.

One of the major future challenges of water contamination management is the need for a better understanding of contamination events in order to predict how they will spread and what their impact will be. This will require the development of more sophisticated mathematical (systems-theoretic) approaches for modeling, detecting, and responding to contamination events. Another challenge is the integration of large volumes of possibly heterogeneous sensors for monitoring multiple parameters, which will require more data-driven approaches and the use of machine-learning techniques to handle these situations effectively. Changes in technology will also continue to be a challenge, as both short-term and long-term changes in the system, including changes in demands, topology, and components, will need to be taken into account. Real-time automated responses to contamination events will be crucial in minimizing their impact, both in terms of the infrastructure, as well as human health, and will require the development of advanced technologies and strategies to enable this. Finally, the human component in decision-making is also an important consideration, as there is a need to consider factors such as the explainability of AI-based decision support systems, biases in decision-making, in order to make effective decisions.

7. Acknowledgments

This work was co-funded by the European Research Council (ERC) under the ERC Synergy grant agreement No. 951424 (Water Futures), the European Union's (EU) Horizon 2020 research and innovation programme under grant agreement No. 883484 (PathoCERT), and supported by the EU Horizon 2020 programme under grant agreement No. 739551 (KIOS CoE), and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

References

- Abhijith, G., & Ostfeld, A. (2021). Modeling the response of nonchlorinated, chlorinated, and chloraminated water distribution systems toward arsenic contamination. *Journal of Environmental Engineering*, 147, 04021045.
- Abhijith, G. R., & Ostfeld, A. (2022). Contaminant fate and transport modeling in distribution systems: EPANET-C. *Water*, 14, 1665.
- Adedoja, O. S., Hamam, Y., Khalaf, B., & Sadiku, E. R. (2020). Development of an algorithm for the estimation of contamination sources in a water distribution network. *IEEE Access*, 8, 200412–200419.
- Adedoja, O. S., Hamam, Y., Khalaf, B., & Sadiku, R. (2018a). A state-of-the-art review of an optimal sensor placement for contaminant warning system in a water distribution network. *Urban Water Journal*, 15, 985–1000.
- Adedoja, O. S., Hamam, Y., Khalaf, B., & Sadiku, R. (2018b). Towards development of an optimization model to identify contamination source in a water distribution network. *Water*, 10, 579.
- Afshar, A., & Najafi, E. (2014). Consequence management of chemical intrusion in water distribution networks under inexact scenarios. *Journal of Hydroinformatics*, 16, 178–188.
- Aghlmand, R., Rasi Nezami, S., & Abbasi, A. (2021). Evaluation of chemical parameters of urban drinking water quality along with health risk assessment: a case study of Ardabil province, Iran. *International Journal of Environmental Research and Public Health*, 18, 5179.
- Albertos, P., & Goodwin, G. C. (2002). Virtual sensors for control applications. *Annual Reviews in Control*, 26, 101–112.
- Arad, J., Housh, M., Perelman, L., & Ostfeld, A. (2013). A dynamic thresholds scheme for contaminant event detection in water distribution systems. *Water Research*, 47, 1899–1908.
- Barros, D. B., Cardoso, S. M., Oliveira, E., Brentan, B., & Ribeiro, L. (2022). Using data mining techniques to isolate chemical intrusion in water distribution systems. *Environmental Monitoring and Assessment*, 194, 203.
- Bashi-Azghadi, S. N., Afshar, M. H., & Afshar, A. (2017). Multi-objective optimization response modeling to contaminated water distribution networks: pressure driven versus demand driven analysis. *KSCE Journal of Civil Engineering*, 21, 2085–2096.
- Batista, A. M. M., Meynet, P., Garcia, G. P. P., Costa, S. A. V., Araujo, J. C., Davenport, R. J., Werner, D., & Mota Filho, C. R. (2018). Microbiological safety of a small water distribution system: evaluating potentially pathogenic bacteria using advanced sequencing techniques. *Water Science and Technology: Water Supply*, 18, 391–398.
- Beaudeau, P., de Valk, H., Vaillant, V., Mannschott, C., Tillier, C., Mouly, D., & Ledrans, M. (2008). Lessons learned from ten investigations of waterborne gastroenteritis outbreaks, France, 1998–2006. *Journal of Water and Health*, 6, 491–503.
- Berglund, E. Z., Pesantez, J. E., Rasekh, A., Shafiee, M. E., Sela, L., & Haxton, T. (2020). Review of modeling methodologies for managing water distribution security. *Journal of Water Resources Planning and Management*, 146, 03120001.
- Berry, J., Hart, W. E., Phillips, C. A., Uber, J. G., & Watson, J.-P. (2006). Sensor placement in municipal water networks with temporal integer programming models. *Journal of Water Resources Planning and Management*, 132, 218–224.
- Berry, J. W., Fleischer, L., Hart, W. E., Phillips, C. A., & Watson, J.-P. (2005). Sensor placement in municipal water networks. *Journal of Water Resources Planning and Management*, 131, 237–243.
- Biswas, P., Lu, C., & Clark, R. L. (1993). A model for chlorine concentration decay in pipes. *Water Research*, 27, 1715–1724.
- Bjelkmar, P., Hansen, A., Schonning, C., Bergstrom, J., Lofdahl, M., Lebbad, M., Wallensten, A., Allestam, G., Stenmark, S., & Lindh, J. (2017). Early outbreak detection by linking health advice line calls to water distribution areas retrospectively demonstrated in a large waterborne outbreak of cryptosporidiosis in Sweden. *BMC Public Health*, 17.
- Boccelli, D. L., Tryby, M. E., Uber, J. G., Rossman, L. A., Zierolf, M. L., & Polycarpou, M. M. (1998). Optimal scheduling of booster disinfection in water distribution systems. *Journal of Water Resources Planning and Management*, 124, 99–111.
- Boccelli, D. L., Tryby, M. E., Uber, J. G., & Summers, R. S. (2003). A reactive species model for chlorine decay and THM formation under rechlorination conditions. *Water Research*, 37, 2654–2666.
- Bopp, D. J., Sauters, B. D., Waring, A. L., Ackelsberg, J., Dumas, N., Braun-Howland, E., Dziewulski, D., Wallace, B. J., Kelly, M., Halse, T., Musser, K. A., Smith, P. F., Morse, D. L., & Limberger, R. J. (2003). Detection, isolation, and molecular subtyping of *Escherichia coli* O157:H7 and *Campylobacter jejuni* associated with a large waterborne outbreak. *Journal of Clinical Microbiology*, 41, 174–180.
- Boulos, P. F., Lansey, K. E., & Karney, B. W. (2006). *Comprehensive water distribution systems analysis handbook for engineers and planners*. MWH Soft, Incorporated.
- Bourgeois, W., Burgess, J. E., & Stuetz, R. M. (2001). On-line monitoring of wastewater quality: a review. *Journal of Chemical Technology & Biotechnology*, 76, 337–348.
- Brdys, M. A., Chang, T., & Duzinkiewicz, K. (2001). Intelligent model predictive control of chlorine residuals in water distribution systems. In D. Phelps, & G. Sehlke (Eds.), *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges* (pp. 1–11). ASCE.
- Burkhardt, J. B., Szabo, J., Klosterman, S., Hall, J., & Murray, R. (2017). Modeling fate and transport of arsenic in a chlorinated distribution system. *Environmental Modelling & Software*, 93, 322–331.
- Buse, H. Y., Schoen, M. E., & Ashbolt, N. J. (2012). Legionellae in engineered systems and use of quantitative microbial risk assessment to predict exposure. *Water Research*, 46, 921–933.
- Butera, L., Gómez-Hernández, J. J., & Nicotra, S. (2021). Contaminant-source detection in a water distribution system using the ensemble Kalman filter. *Journal of Water Resources Planning and Management*, 147, 04021029.
- Cai, J., & Ye, Z.-S. (2021). Contamination source identification: A Bayesian framework integrating physical and statistical models. *IEEE Transactions on Industrial Informatics*, 17, 8189–8197.
- Carr, R. D., Greenberg, H. J., Hart, W. E., Konjevod, G., Lauer, E., Lin, H., Morrison, T., & Phillips, C. A. (2006). Robust optimization of contaminant sensor placement for community water systems. *Mathematical Programming*, 107, 337–356.
- Chaves, R. S., Guerreiro, C. S., Cardoso, V. V., Benoliel, M. J., & Santos, M. M. (2019). Hazard and mode of action of disinfection by-products (DBPs) in water for human consumption: evidences and research priorities. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, 223, 53–61.
- Ciaponi, C., Creaco, E., Nardo, A. D., Natale, M. D., Giudicianni, C., Musmarra, D., & Santonastaso, G. F. (2018). Optimal sensor placement in a partitioned water distribution network for the water protection from contamination. *Proceedings*, 2, 670.
- Ciaponi, C., Creaco, E., Nardo, A. D., Natale, M. D., Giudicianni, C., Musmarra, D., & Santonastaso, G. F. (2019). Reducing impacts of contamination in water distribution networks: A combined strategy based on network partitioning and installation of water quality sensors. *Water (Switzerland)*, 11, 1315.
- Çiner, F., Sunkari, E. D., & Enba, B. A. (2021). Geochemical and multivariate statistical evaluation of trace elements in groundwater of Niğde municipality, south-central Turkey: implications for arsenic contamination and human health risks assessment. *Archives of Environmental Contamination and Toxicology*, 80, 164–182.
- Clark, R. M. (1998). Chlorine demand and TTHM formation kinetics: a second-order model. *Journal of Environmental Engineering*, 124, 16–24.
- Clark, R. M. (2015). The USEPA's distribution system water quality modelling program: a historical perspective. *Water and Environment Journal*, 29, 320–330.
- Clark, R. M., & Sivaganesan, M. (2002). Predicting chlorine residuals in drink-

- ing water: second order model. *Journal of Water Resources Planning and Management*, 128, 152–161.
- Comboul, M., & Ghanem, R. (2013). Value of information in the design of resilient water distribution sensor networks. *Journal of Water Resources Planning and Management*, 139, 449–455.
- Constans, S., Brémond, B., & Morel, P. (2003). Simulation and control of chlorine levels in water distribution networks. *Journal of Water Resources Planning and Management*, 129, 135–145.
- Corso, P. S., Kramer, M. H., Blair, K. A., Addiss, D. G., Davis, J. P., & Haddix, A. C. (2003). Costs of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee, Wisconsin. *Emerging Infectious Diseases*, 9, 426.
- Cossali, G., Routledge, E. J., Ratcliffe, M. S., Blakes, H., Fielder, J. E., & Karayiannis, T. G. (2016). Inactivation of *E. coli*, *Legionella*, and *Pseudomonas* in tap water using electrochemical disinfection. *Journal of Environmental Engineering*, 142, 04016063.
- Costa, D. M., Melo, L. F., & Martins, F. G. (2013). Localization of contamination sources in drinking water distribution systems: A method based on successive positive readings of sensors. *Water Resources Management*, 27, 4623–4635.
- Cozzolino, L., Della Morte, R., Palumbo, A., & Pianese, D. (2011). Stochastic approaches for sensors placement against intentional contaminations in water distribution systems. *Civil Engineering and Environmental Systems*, 28, 75–98.
- Cristo, C. D., & Leopardi, A. (2008). Pollution source identification of accidental contamination in water distribution networks. *Journal of Water Resources Planning and Management*, 134, 197–202.
- Danon-Schaffer, M. (2001). Walkerton's contaminated water supply system: a forensic approach to identifying the source. *Environmental Forensics*, 2, 197–200.
- Davies, A. W., Dubow, J. B., Collins, G., & Borky, J. M. (2022). Analyzing and mitigating water utility system vulnerabilities. *Journal-American Water Works Association*, 114, 58–66.
- De Sanctis, A. E., Shang, F., & Uber, J. G. (2010). Real-time identification of possible contamination sources using network backtracking methods. *Journal of Water Resources Planning and Management*, 136, 444–453.
- Delmaar, J., & van Engelen, J. (2007). *Aggregating human exposure to chemicals: an overview of tools and methodologies*. Technical Report 63070001 Rijksinstituut voor Volksgezondheid en Milieu RIVM.
- Díaz, S., Mínguez, R., & González, J. (2017). Calibration via multi-period state estimation in water distribution systems. *Water Resources Management*, 31, 4801–4819.
- Duzinkiewicz, K., Brdys, M. A. A., & Chang, T. (2005). Hierarchical model predictive control of integrated quality and quantity in drinking water distribution systems. *Urban Water Journal*, 2, 125–137.
- Eliades, D. G., Kyriakou, M., Vrachimis, S., & Polycarpou, M. M. (2016). EPANET-MATLAB Toolkit: An open-source software for interfacing EPANET with MATLAB. In *Proc. 14th International Conference on Computing and Control for the Water Industry (CCWI)* (p. 8).
- Eliades, D. G., Panayiotou, C., & Polycarpou, M. M. (2014). Contamination event detection in drinking water systems using a real-time learning approach. In *2014 International Joint Conference on Neural Networks (IJCNN)* (pp. 663–670). IEEE.
- Eliades, D. G., & Polycarpou, M. M. (2010). A fault diagnosis and security framework for water systems. *IEEE Transactions on Control Systems Technology*, 18, 1254–1265.
- Eliades, D. G., & Polycarpou, M. M. (2012). Water contamination impact evaluation and source-area isolation using decision trees. *Journal of Water Resources Planning and Management*, 138, 562–570.
- Eliades, D. G., Polycarpou, M. M., & Charalambous, B. (2011). A security-oriented manual quality sampling methodology for water systems. *Water Resources Management*, 25, 1219–1228.
- Eliades, D. G., Stavrou, D., Vrachimis, S. G., Panayiotou, C. G., & Polycarpou, M. M. (2015). Contamination event detection using multi-level thresholds. In *Computing and Control for the Water Industry CCWI2015* (pp. 1429–1438). volume 119.
- Elsherif, S. M., Wang, S., Taha, A. F., Sela, L., Giacomoni, M. H., & Abokifa, A. (2022). Control-theoretic modeling of multi-species water quality dynamics in drinking water networks: survey, methods, and test cases. *Annual Reviews in Control*, . (in press).
- of Emergency, U. S. E. P. A. O., & Response, R. (1989). *Risk assessment guidance for superfund*. Office of Emergency and Remedial Response, US Environmental Protection Agency.
- Evlampidou, I., Font-Ribera, L., Rojas-Rueda, D., Gracia-Lavedan, E., Costet, N., Pearce, N., Veneis, P., Jaakkola, J. J., Delloye, F., Makris, K. C., Stephanou, E. G., Kargaki, S., Kozisek, F., Sigsgaard, T., Hansen, B., Schullehner, J., Nahkur, R., Galey, C., Zwiener, C., Vargha, M., Righi, E., Aggazzotti, G., Kalnina, G., Grazuleviciene, R., Polanska, K., Gubkova, D., Bitenc, K., Goslan, E. H., Kogevinas, M., & Villanueva, C. M. (2020). Trihalomethanes in drinking water and bladder cancer burden in the European union. *Environmental Health Perspectives*, 128, 017001.
- Farahat, A., Mahmoud, M. T., & Khalil, A. (2019). Assessment of the risk associated with *E. coli* bacterial intrusion in drinking water distribution networks. *Arabian Journal for Science and Engineering*, 44, 4161–4168.
- Fasaee, M. A. K., Nikoo, M. R., Bakhtiari, P. H., Monghasemi, S., & Sadeh, M. (2020). A novel dynamic hydrant flushing framework facilitated by categorizing contamination events. *Urban Water Journal*, 17, 199–211.
- Fitore, M., Doina, L., & Florin, L. (2019). Machine learning approaches for anomaly detection of water quality on a real-world data set. Fourier analysis for demand forecasting in a fashion company. *Journal of Information and Telecommunication*, 3, 294–307.
- Ford, T. E. (1999). Microbiological safety of drinking water: United states and global perspectives. *Environmental Health Perspectives*, 107, 191–206.
- Gerba, C. P. (2015). *Chapter 24 - Risk Assessment*. (Third edition ed.). San Diego: Academic Press.
- Giammanco, G. M., Bonura, F., Urone, N., Purpari, G., Cuccia, M., Pepe, A., Li Muli, S., Cappa, V., Saglimbene, C., Mandolfo, G., Marino, A., Guercio, A., Di Bartolo, I., & De Grazia, S. (2018). Waterborne norovirus outbreak at a seaside resort likely originating from municipal water distribution system failure. *Epidemiology & Infection*, 146, 879–887.
- Grayman, W. M., Clark, R. M., & Males, R. M. (1988). Modeling distribution-system water quality; dynamic approach. *Journal of Water Resources Planning and Management*, 114, 295–312.
- Grbčić, L., Lučin, I., Kranjčević, L., & Družeta, S. (2020). A machine learning-based algorithm for water network contamination source localization. *Sensors*, 20, 2613.
- Guan, J., Aral, M. M., Maslia, M. L., & Grayman, W. M. (2006). Identification of contaminant sources in water distribution systems using simulation optimization method: case study. *Journal of Water Resources Planning and Management*, 132, 252–262.
- Guidorzi, M., Franchini, M., & Alvisi, S. (2009). A multi-objective approach for detecting and responding to accidental and intentional contamination events in water distribution systems. *Urban Water Journal*, 6, 115–135.
- Hall, J., Szabo, J. G., Panguluri, S., & Meiners, G. (2009). *Distribution system water quality monitoring: Sensor technology evaluation methodology and results*. Technical Report EPA/600/R-09/076 U.S. Environmental Protection Agency.
- Hamilton, K. A., & Haas, C. N. (2016). Critical review of mathematical approaches for quantitative microbial risk assessment (QMRA) of *Legionella* in engineered water systems: research gaps and a new framework. *Environmental Science: Water Research & Technology*, 2, 599–613.
- Hamouda, M. A., Anderson, W. B., Van Dyke, M. I., Douglas, I. P., McFadyen, S. D., & Huck, P. M. (2016). Scenario-based quantitative microbial risk assessment to evaluate the robustness of a drinking water treatment plant. *Water Quality Research Journal of Canada*, 51, 81–96.
- Hart, W. E., & Murray, R. (2010). Review of sensor placement strategies for contamination warning systems in drinking water distribution systems. *Journal of Water Resources Planning and Management*, 136, 611–619.
- He, X., & Huang, K. (2020). Chapter 7 - assessment technologies for hazards/risks of wastewater. In H. Ren, & X. Zhang (Eds.), *High-Risk Pollutants in Wastewater* (pp. 141–167). Elsevier.
- Helbling, D. E., & VanBriesen, J. M. (2009). Modeling residual chlorine response to a microbial contamination event in drinking water distribution systems. *Journal of Environmental Engineering*, 135, 918–927.
- Hernandez Cervantes, D., Mora Rodríguez, J., Delgado Galvan, X., Ortiz Medel, J., & Jimenez Magana, M. R. (2016). Optimal use of chlorine in water distribution networks based on specific locations of booster chlorination: analyzing conditions in Mexico. *Water Science and Technology: Water Supply*, 16, 493–505.
- Hu, C., Cai, J., Zeng, D., Yan, X., Gong, W., & Wang, L. (2020a). Deep reinforcement learning based valve scheduling for pollution isolation in water distribution network. *Mathematical Biosciences and Engineering*, 17, 105–121.

- Hu, C., Wang, Q., Gong, W., & Yan, X. (2022). Multi-objective deep reinforcement learning for emergency scheduling in a water distribution network. *Memetic Computing*, 14, 211–223.
- Hu, C., Yan, X., Gong, W., Liu, X., Wang, L., & Gao, L. (2020b). Multi-objective based scheduling algorithm for sudden drinking water contamination incident. *Swarm and Evolutionary Computation*, 55, 100674.
- Hua, E., West, J., Barker, R., & Forster, C. (1999). Modelling of chlorine decay in municipal water supplies. *Water Research*, 33, 2735–2746.
- Hua, P., Vasyukova, E., & Uhl, W. (2015). A variable reaction rate model for chlorine decay in drinking water due to the reaction with dissolved organic matter. *Water Research*, 75, 109–122.
- Huang, J. J., & McBean, E. A. (2009). Data mining to identify contaminant event locations in water distribution systems. *Journal of Water Resources Planning and Management*, 135, 466–474.
- Hunter, P. R., Colford, J. M., LeChevallier, M. W., Binder, S., & Berger, P. S. (2001). Waterborne diseases. *Emerging Infectious Diseases*, 7, 544.
- Hutton, C. J., Kapelan, Z., Vamvakieridou-Lyroudia, L., & Savić, D. A. (2014). Dealing with uncertainty in water distribution system models: A framework for real-time modeling and data assimilation. *Journal of Water Resources Planning and Management*, 140, 169–183.
- Islam, M. R., & Chaudhry, M. H. (1998). Modeling of constituent transport in unsteady flows in pipe networks. *Journal of Hydraulic Engineering*, 124, 1115–1124.
- Islam, N., Farahat, A., Al-Zahrani, M. A. M., Rodriguez, M. J., & Sadiq, R. (2015). Contaminant intrusion in water distribution networks: review and proposal of an integrated model for decision making. *Environmental Reviews*, 23, 337–352.
- Islam, N., Rodriguez, M. J., Farahat, A., & Sadiq, R. (2017). Minimizing the impacts of contaminant intrusion in small water distribution networks through booster chlorination optimization. *Stochastic Environmental Research and Risk Assessment*, 31, 1759–1775.
- Islam, N., Sadiq, R., & Rodriguez, M. J. (2013). Optimizing booster chlorination in water distribution networks: A water quality index approach. *Environmental Monitoring and Assessment*, 185, 8035–8050.
- Jacqueline, C., del Valle Arrojo, M., Bellver Moreira, P., Rodríguez Feijóo, M. A., Cabrerizo, M., & Fernandez-Garcia, M. D. (2022). Norovirus GI.3[p12] outbreak associated with the drinking water supply in a rural area in Galicia, Spain, 2021. *Microbiology Spectrum*, 10, e01048–22.
- Jakopanec, I., Borgen, K., Vold, L., Lund, H., Forseth, T., Hannula, R., & Nygard, K. (2008). A large waterborne outbreak of campylobacteriosis in norway: The need to focus on distribution system safety. *BMC Infectious Diseases*, 8, 1–11.
- Jamal, R., Mubarak, S., Sahulka, S. Q., Kori, J. A., Tajammul, A., Ahmed, J., Mahar, R. B., Olsen, M. S., Goel, R., & Weidhaas, J. (2020). Informing water distribution line rehabilitation through quantitative microbial risk assessment. *Science of The Total Environment*, 739, 140021.
- Janne, J., Paivi, M., & Antti, S. (2017). Public health and economic risk assessment of waterborne contaminants and pathogens in Finland. *Science of the Total Environment*, 599, 873–882.
- Jonkergouw, P., Khu, S., & Savić, D. (2004). Chlorine: A possible indicator of intentional chemical and biological contamination in a water distribution network. In *Proc. IWA Conference on Automation in Water Quality Monitoring*.
- Kadinski, L., Salcedo, C., Boccelli, D. L., Berglund, E., & Ostfeld, A. (2022). A hybrid data-driven-agent-based modelling framework for water distribution systems contamination response during covid-19. *Water*, 14.
- Kanakoudis, V., & Tsitsifli, S. (2017). Potable water security assessment - a review on monitoring, modelling and optimization techniques, applied to water distribution networks. *Desalination and Water Treatment*, 99, 18–26.
- Kang, D., & Lansey, K. (2010). Real-time optimal valve operation and booster disinfection for water quality in water distribution systems. *Journal of Water Resources Planning and Management*, 136, 463–473.
- Kang, G., Gao, J. Z., & Xie, G. (2017). Data-driven water quality analysis and prediction: A survey. In *2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService)* (pp. 224–232). IEEE.
- Karamouz, M., Yousefi, A., Zahmatkesh, Z., Mahmoodzadeh, D., & Pirooz, M. D. (2022). Simulation of chlorine injection in water distribution networks in response to contaminations. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH*, 16.
- Kauppinen, A., Pitkänen, T., Al-Hello, H., Maunula, L., Hokajärvi, A.-M., Rimhanen-Finne, R., & Miettinen, I. T. (2019). Two drinking water outbreaks caused by wastewater intrusion including sapovirus in Finland. *International Journal of Environmental Research and Public Health*, 16, 4376.
- Kessler, A., Ostfeld, A., & Sinai, G. (1998). Detecting accidental contaminations in municipal water networks. *Journal of Water Resources Planning and Management*, 124, 192–198.
- Kicinska, A., & Wysowska, E. (2021). Health risk related to the presence of metals in drinking water from different types of sources. *Water and Environment Journal*, 35, 27–40.
- Kombo Mpindou, G. O. M., Escuder Bueno, I., & Chordà Ramón, E. (2022). Risk analysis methods of water supply systems: comprehensive review from source to tap. *Applied Water Science*, 12, 1–20.
- Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., & Faloutsos, C. (2008). Efficient sensor placement optimization for securing large water distribution networks. *Journal of Water Resources Planning and Management*, 134, 516–526.
- Kruse, P. (2018). Review on water quality sensors. *Journal of Physics D: Applied Physics*, 51, 203002.
- Kuhn, K. G., Falkenhorst, G., Emborg, H. D., Ceper, T., Torpdahl, M., Krogfelt, K. A., Ethelberg, S., & Molbak, K. (2017). Epidemiological and serological investigation of a waterborne Campylobacter jejuni outbreak in a Danish town. *Epidemiology & Infection*, 145, 701–709.
- Kumar, A., Kansal, M. L., Arora, G., Ostfeld, A., & Kessler, A. (1999). Detecting accidental contaminations in municipal water networks. *Journal of Water Resources Planning and Management*, 125, 308–310.
- Kumar, J., Brill, E. D., Mahinthakumar, G., & Ranjithan, S. R. (2012). Contaminant source characterization in water distribution systems using binary signals. *Journal of Hydroinformatics*, 14, 585–602.
- Laird, C. D., Biegler, L. T., van Bloemen Waanders, B. G., & Bartlett, R. A. (2005). Contamination source determination for water networks. *Journal of Water Resources Planning and Management*, 131, 125–134.
- Laird, C. D., Biegler, L. T., & van Bloemen Waanders, B. G. (2006). Mixed-integer approach for obtaining unique solutions in source inversion of water networks. *Journal of Water Resources Planning and Management*, 132, 242–251.
- Lambrou, T. P., Panayiotou, C. G., & Anastasiou, C. C. (2012). A low-cost system for real time monitoring and assessment of potable water quality at consumer sites. In *Sensors, 2012 IEEE* (pp. 1–4). IEEE.
- Langowski, R., & Brdys, M. (2007). Monitoring of chlorine concentration in drinking water distribution systems using an interval estimator. *International Journal of Applied Mathematics and Computer Science*, 17, 199–216.
- Langowski, R., & Brdys, M. A. (2017). An interval estimator for chlorine monitoring in drinking water distribution systems under uncertain system dynamics, inputs and chlorine concentration measurement errors. *International Journal of Applied Mathematics and Computer Science*, 27, 309–322.
- Langowski, R., Brdys, M. A., & Qi, R. (2012). Optimised robust placement of hard quality sensors for robust monitoring of quality in drinking water distribution systems. In *Proceedings of the 10th World Congress on Intelligent Control and Automation* (pp. 1109–1114).
- LeVeque, R. J. (1998). Nonlinear conservation laws and finite volume methods. In O. Steiner, & A. Gautschi (Eds.), *Computational methods for astrophysical fluid flow* (pp. 1–159). Springer.
- LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*. Cambridge University Press.
- Lienemann, T., Pitkänen, T., Antikainen, J., Mölsä, E., Miettinen, I., Haukka, K., Vaara, M., & Siitonen, A. (2011). Shiga toxin-producing escherichia coli o100: H-: stx 2e in drinking water contaminated by waste water in Finland. *Current microbiology*, 62, 1239–1244.
- Lifshitz, R., & Ostfeld, A. (2019). Clustering for real-time response to water distribution system contamination event intrusions. *Journal of Water Resources Planning and Management*, 145, 04018091.
- Liou, C., & Kroon, J. (1987). Modeling the propagation of waterborne substances in distribution networks. *Journal - American Water Works Association*, 79, 54–58.
- Liu, L., Ranjithan, S. R., & Mahinthakumar, G. (2011). Contamination source identification in water distribution systems using an adaptive dynamic optimization procedure. *Journal of Water Resources Planning and Man-*

- agement, 137, 183–192.
- Liu, L., Zechman, E. M., Brill, E. D., Jr, Mahinthakumar, G., Ranjithan, S., & Uber, J. (2008). Adaptive contamination source identification in water distribution systems using an evolutionary algorithm-based dynamic optimization procedure. In *Water Distribution Systems Analysis Symposium 2006* (pp. 1–9).
- Liu, L., Zechman, E. M., Mahinthakumar, G., & Ranjithan, S. R. (2012a). Coupling of logistic regression analysis and local search methods for characterization of water distribution system contaminant source. *Engineering Applications of Artificial Intelligence*, 25, 309–316.
- Liu, L., Zechman, E. M., Mahinthakumar, G., & Ranjithan, S. R. (2012b). Identifying contaminant sources for water distribution systems using a hybrid method. *Civil Engineering and Environmental Systems*, 29, 123–136.
- Lučin, I., Grbčić, L., Čarija, Z., & Kranjčević, L. (2021a). Machine-Learning classification of a number of contaminant sources in an urban water network. *Sensors*, 21, 245.
- Lučin, I., Grbčić, L., Družeta, S., & Čarija, Z. (2021b). Source contamination detection using novel search space reduction coupled with optimization technique. *Journal of Water Resources Planning and Management*, 147, 04020100.
- Mandel, P., Wang, Y., Parre, A., Féliers, C., & Heim, V. (2021). Quality zones automatically identified in water distribution networks by applying data clustering methods to conductivity measurements. *Water Research*, 207, 117716.
- Mann, A. V., McKenna, S. A., Hart, W. E., & Laird, C. D. (2012). Real-time inversion in large-scale water networks using discrete measurements. *Computers & Chemical Engineering*, 37, 143–151.
- McKenna, S. A., Hart, D. B., Klise, K., Koch, M., Vugrin, E. D., Martin, S., Wilson, M., Cruz, V., & Cutler, L. (2010). *Water quality event detection systems for drinking water contamination warning systems*. Technical Report US Environmental Protection Agency.
- Moghaddam, A., Afsharnia, M., & Minaee, R. P. (2020). Preparing the optimal emergency response protocols by MOPSO for a real-world water distribution network. *Environmental Science and Pollution Research*, 27, 30625–30637.
- Moghaddam, A., Afsharnia, M., Mokhtari, M., & Peirovi-Minaee, R. (2022). Management and health risk assessment of chemical contamination events in water distribution systems using PSO. *Environmental Monitoring and Assessment*, 194.
- Monteiro, L., Figueiredo, D., Covas, D., & Menaia, J. (2017). Integrating water temperature in chlorine decay modelling: a case study. *Urban Water Journal*, 14, 1097–1101.
- Montenegro-Ayo, R., Barrios, A. C., Mondal, I., Bhagat, K., Carlos Morales-Gomero, J., Abbaszadegan, M., Westerhoff, P., Perreault, F., & Garcia-Segura, S. (2020). Portable point-of-use photoelectrocatalytic device provides rapid water disinfection. *Science of The Total Environment*, 737, 140044.
- Morris, R. D. (1995). Drinking water and cancer. *Environmental Health Perspectives*, 103, 225–231.
- Mouly, D., Joulin, E., Rosin, C., Beaudeau, P., Zeghnoun, A., Olszewski-Ortar, A., Munoz, J. F., Welté, B., Joyeux, M., Seux, R., Montiel, A., & Rodríguez, M. J. (2010). Variations in trihalomethane levels in three French water distribution systems and the development of a predictive model. *Water Research*, 44, 5168–5179.
- Nascetti, S., Busani, L., Bartoli, F., Orioli, R., Stenico, A., & Regele, D. (2021). Community waterborne outbreak linked to a firefighting response during the COVID-19 emergency. *Annali Dell'istituto Superiore di Sanita*, 57, 226–232.
- Neupauer, R. M., Records, M. K., & Ashwood, W. H. (2010). Backward probabilistic modeling to identify contaminant sources in water distribution systems. *Journal of Water Resources Planning and Management*, 136, 587–591.
- Nguyen, M., & Logofătu, D. (2018). Applying tree ensemble to detect anomalies in real-world water composition dataset. In *International Conference on Intelligent Data Engineering and Automated Learning* (pp. 429–438). Springer.
- Ohar, Z., & Ostfeld, A. (2014). Optimal design and operation of booster chlorination stations layout in water distribution systems. *Water Research*, 58, 209–220.
- Oliker, N., & Ostfeld, A. (2015). Comparison of two multivariate classification models for contamination event detection in water quality time series. *Journal of Water Supply: Research and Technology AQUA*, 64, 558–566.
- Organization, W. H. (2016). *Quantitative microbial risk assessment: application for water safety management*. Technical Report World Health Organization.
- Ortega, E., Braunstein, A., & Lage-Castellanos, A. (2020). Contamination source detection in water distribution networks using belief propagation. *Stochastic Environmental Research and Risk Assessment*, 34, 493–511.
- Paepae, T., Bokoro, P. N., & Kyamakya, K. (2021). From fully physical to virtual sensing for water quality assessment: A comprehensive review of the relevant state-of-the-art. *Sensors*, 21, 6971.
- Palleti, V. R., Kurian, V., Narasimhan, S., & Rengaswamy, R. (2018). Actuator network design to mitigate contamination effects in water distribution networks. *Computers and Chemical Engineering*, 108, 194–205.
- Parks, S. L. I., & VanBriesen, J. M. (2009). Booster disinfection for response to contamination in a drinking water distribution system. *Journal of Water Resources Planning and Management*, 135, 502–511.
- Pasha, M. F. K., & Lansey, K. (2010). Effect of parameter uncertainty on water quality predictions in distribution systems-case study. *Journal of Hydroinformatics*, 12, 1–21.
- Pasqualetti, E., Dorfler, F., & Bullo, F. (2015). Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems. *IEEE Control Systems Magazine*, 35, 110–127.
- Pecci, F., Stoianov, I., & Ostfeld, A. (2022). Optimal design-for-control of chlorine booster systems in water networks via convex optimization. In *2022 European Control Conference (ECC)* (pp. 1988–1993). IEEE.
- Perelman, L., Arad, J., Housh, M., & Ostfeld, A. (2012). Event detection in water distribution systems from multivariate water quality time series. *Environmental science & technology*, 46, 8212–8219.
- Perelman, L., & Ostfeld, A. (2012). Water-distribution systems simplifications through clustering. *Journal of Water Resources Planning and Management*, 138, 218–229.
- Polycarpou, M. M., Uber, J. G. J. G., Wang, Z. W. Z., Shang, F. S. F., & Brdys, M. (2002). Feedback control of water quality. *IEEE Control Systems*, 22, 68–87.
- Preis, A., & Ostfeld, A. (2006). Contamination source identification in water systems: A hybrid model trees linear programming scheme. *Journal of Water Resources Planning and Management*, 132, 263–273.
- Preis, A., & Ostfeld, A. (2007). A contamination source identification model for water distribution system security. *Engineering Optimization*, 39, 941–947.
- Preis, A., & Ostfeld, A. (2008a). Genetic algorithm for contaminant source characterization using imperfect sensors. *Civil Engineering and Environmental Systems*, 25, 29–39.
- Preis, A., & Ostfeld, A. (2008b). Multiobjective contaminant response modeling for water distribution systems security. *Journal of Hydroinformatics*, 10, 267–274.
- Preis, A., & Ostfeld, A. (2011). Hydraulic uncertainty inclusion in water distribution systems contamination source identification. *Urban Water Journal*, 8, 267–277.
- Propato, M., Sarrazzy, F., & Tryby, M. (2010). Linear algebra and minimum relative entropy to investigate contamination events in drinking water systems. *Journal of Water Resources Planning and Management*, 136, 483–492.
- Qian, K., Jiang, J., Ding, Y., & Yang, S. (2020). Deep learning based anomaly detection in water distribution systems. In *2020 IEEE International Conference on Networking, Sensing and Control (ICNSC)* (pp. 1–6). IEEE.
- Qian, K., Jiang, J., Ding, Y., & Yang, S.-H. (2021). DLGEA: a deep learning guided evolutionary algorithm for water contamination source identification. *Neural Computing and Applications*, 33, 11889–11903.
- Qiu, M., Salomons, E., & Ostfeld, A. (2020). A framework for real-time disinfection plan assembling for a contamination event in water distribution systems. *Water Research*, 174, 115625.
- Qiu, M., Salomons, E., & Ostfeld, A. (2021). An analytical model for the decontamination of water distribution systems using slug-feed method of disinfection. *Water Resources Research*, 57, e2020WR028277.
- Raich, J. (2013). *Review of sensors to monitor water quality*. Technical Report Publications Office of the European Union.
- Rasekh, A., & Brumbelow, K. (2013). Probabilistic analysis and optimization to characterize critical water distribution system contamination scenarios. *Journal of Water Resources Planning and Management*, 139, 191–199.
- Rasekh, A., & Brumbelow, K. (2014). Drinking water distribution systems

- contamination management to reduce public health impacts and system service interruptions. *Environmental Modelling & Software*, 51, 12–25.
- Rasekh, A., & Brumbelow, K. (2015). A dynamic simulation-optimization model for adaptive management of urban water distribution system contamination threats. *Applied Soft Computing Journal*, 32, 59–71.
- Rasekh, A., Shafiee, M. E., Zechman, E., & Brumbelow, K. (2014). Sociotechnical risk assessment for water distribution system contamination threats. *Journal of Hydroinformatics*, 16, 531–549.
- Rathi, S., & Gupta, R. (2014). Sensor placement methods for contamination detection in water distribution networks: A review. *Procedia Engineering*, 89, 181–188.
- Rego, B. S., Vrachimis, S. G., Polycarpou, M. M., Raffo, G. V., & Raimondo, D. M. (2021). State estimation and leakage detection in water distribution networks using constrained zonotopes. *IEEE Transactions on Control Systems Technology*, (pp. 1–14).
- Richardson, S., & Postigo, C. (2012). Drinking water disinfection by-products. In *Emerging Organic contaminants and human health* (pp. 93–137). Springer volume 20 of *Emerging Organic Contaminants and Human Health*.
- Richardson, S. D. (2009). Water analysis: emerging contaminants and current issues. *Analytical Chemistry*, 81, 4645–4677.
- Rico-Ramirez, V., Frausto-Hernandez, S., Diwekar, U. M., & Hernandez-Castro, S. (2007). Water networks security: A two-stage mixed-integer stochastic program for sensor placement under uncertainty. *Computers & chemical engineering*, 31, 565–573.
- Riera-Montes, M., Sjolander, K. B., Allestam, G., Hallin, E., Hedlund, K.-O., & Lofdahl, M. (2011). Waterborne norovirus outbreak in a municipal drinking-water supply in Sweden. *Epidemiology & Infection*, 139, 1928–1935.
- Rossman, L. A. (2000). *EPANET 2: users manual*. Technical Report US Environmental Protection Agency. Office of Research and Development. National Risk Management Research Laboratory.
- Rossman, L. A., & Boulos, P. F. (1996). Numerical methods for modeling water quality in distribution systems: a comparison. *Journal of Water Resources Planning and Management*, 122, 137–146.
- Rossman, L. A., Clark, R. M., & Grayman, W. M. (1994). Modeling chlorine residuals in drinking water distribution systems. *Journal of Environmental Engineering*, 120, 803–820.
- Sakarya, A. B. A., & Mays, L. W. (2000). Optimal operation of water distribution pumps considering water quality. *Journal of Water Resources Planning and Management*, 126, 210–220.
- Sankary, N., & Ostfeld, A. (2019). Bayesian localization of water distribution system contamination intrusion events using inline mobile sensor data. *Journal of Water Resources Planning and Management*, 145, 04019029.
- Schijven, J., Foret, J. M., Chardon, J., Teunis, P., Bouwknegt, M., & Tangena, B. (2016). Evaluation of exposure scenarios on intentional microbiological contamination in a drinking water distribution network. *Water Research*, 96, 148–154.
- Sela, L., & Amin, S. (2018). Robust sensor placement for pipeline monitoring: Mixed integer and greedy optimization. *Advanced Engineering Informatics*, 36, 55–63.
- Seth, A., Hackebeil, G. A., Klise, K. A., Haxton, T., Murray, R., & Laird, C. D. (2017). Efficient reduction of optimal disinfectant booster station placement formulations for security of large-scale water distribution networks. *Engineering Optimization*, 49, 1281–1298.
- Seth, A., Klise, K. A., Sirola, J. D., Haxton, T., & Laird, C. D. (2016). Testing contamination source identification methods for water distribution networks. *Journal of Water Resources Planning and Management*, 142, 04016001.
- Shafiee, M. E., & Berglund, E. Z. (2016). Agent-based modeling and evolutionary computation for disseminating public advisories about hazardous material emergencies. *Computers, Environment and Urban Systems*, 57, 12–25.
- Shafiee, M. E., Berglund, E. Z., & Lindell, M. K. (2018). An agent-based modeling framework for assessing the public health protection of water advisories. *Water Resources Management*, 32, 2033–2059.
- Shafiee, M. E., & Zechman, E. M. (2013). An agent-based modeling framework for sociotechnical simulation of water distribution contamination events. *Journal of Hydroinformatics*, 15, 862–880.
- Shahra, E. Q., Wu, W., & Gomez, R. (2021). Human health impact analysis of contaminant in IoT-enabled water distributed networks. *Applied Sciences* (Switzerland), 11, 3394.
- Shang, F., Uber, J. G., & Polycarpou, M. M. (2002). Particle backtracking algorithm for water distribution system analysis. *Journal of Environmental Engineering*, 128, 441–450.
- Shang, F., Uber, J. G., & Rossman, L. A. (2008). Modeling reaction and transport of multiple species in water distribution systems. *Environmental Science & Technology*, 42, 808–814.
- Shastri, Y., & Diwekar, U. (2006). Sensor placement in water networks: A stochastic programming approach. *Journal of Water Resources Planning and Management*, 132, 192–203.
- Siew, C., & Tanyimboh, T. T. (2012). Pressure-Dependent EPANET Extension. *Water Resources Management*, 26, 1477–1498.
- Silvestry-Rodriguez, N., Bright, K. R., Uhlmann, D. R., Slack, D. C., & Gerba, C. P. (2007). Inactivation of *Pseudomonas aeruginosa* and *Aeromonas hydrophila* by silver in tap water. *Journal of Environmental Science and Health, Part A*, 42, 1579–1584.
- Snyder, L. V. (2006). Facility location under uncertainty: A review. *IIE transactions*, 38, 547–564.
- Sun, L., Yan, H., Xin, K., & Tao, T. (2019). Contamination source identification in water distribution networks using convolutional neural network. *Environmental Science and Pollution Research*, 26, 36786–36797.
- Tao, T., Hai-dong, H., Kun-lun, X., & Shu-ming, L. (2012a). Identification of contamination source in water distribution network based on consumer complaints. *Journal of Central South University*, 19, 1600–1609.
- Tao, T., Lu, Y.-j., Fu, X., & Xin, K.-l. (2012b). Identification of sources of pollution and contamination in water distribution networks based on pattern recognition. *Journal of Zhejiang University-SCIENCE A*, 13, 559–570.
- Teixeira, R., Carmi, O., Gattinesi, P., & Hohenblum, P. (2022). *Water security plan: implementation manual for drinking water systems*. Technical Report Publications Office of the European Union.
- Teixeira, R., Carmi, O., Raich, J., Gattinesi, P., Hohenblum, P., Theocharidou, M., & Giannopoulos, G. (2019). *Guidance for production of a Water Security Plan in drinking water supply*. Technical Report KJ-NA-29846-EN-N (online), KJ-NA-29846-EN-E (ePub) Publications Office of the European Union Luxembourg (Luxembourg).
- The European Parliament and the Council of the European Union (2020). *Directive (EU) 2020/2184 of the European Parliament and of the Council*. Technical Report March 2019 European Union.
- Tinelli, S., Juran, I., & Cantos, W. P. (2018). Development of risk assessment tools for early detection of bio-contamination in water distribution systems. *Water Supply*, 18, 2151–2161.
- Todini, E., & Pilati, S. (1988). A gradient algorithm for the analysis of pipe networks. In *International Conference on Computer Applications for Water Supply and Distribution* (pp. 1–20). Leuicester: Research Studies Press Ltd.
- Tolouei, S., Dewey, R., Snodgrass, W. J., Edge, T. A., Andrews, R. C., Taghipour, M., Prevost, M., & Dorner, S. (2019). Assessing microbial risk through event-based pathogen loading and hydrodynamic modelling. *Science of The Total Environment*, 693, 133567.
- Tshehla, K. S., Hamam, Y., & Abu-Mahfouz, A. M. (2017). State estimation in water distribution network: A review. In *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)* (pp. 1247–1252). IEEE.
- Tsitsifli, S., & Kanakoudis, V. (2021). Assessing the impact of DMAs and the use of boosters on chlorination in a water distribution network in Greece. *Water*, 13.
- U.S. Government (2002). National primary drinking water regulations - title 40, code of federal regulations, part 141 - environmental protection agency (EPA).
- Van Abel, N., Blokker, E. J. M., Smeets, P. W. M. H., Meschke, J. S., & Medema, G. J. (2014). Sensitivity of quantitative microbial risk assessments to assumptions about exposure to multiple consumption events per day. *Journal of Water and Health*, 12, 727–735.
- Van Bel, N., Hornstra, L. M., Van Der Veen, A., & Medema, G. (2019). Efficacy of flushing and chlorination in removing microorganisms from a pilot drinking water distribution system. *Water*, 11, 903.
- Vankayala, P., Sankarasubramanian, A., Ranjithan, S. R., & Mahinthakumar, G. (2009). Contaminant source identification in water distribution networks under conditions of demand uncertainty. *Environmental Forensics*, 10, 253–263.
- Vasconcelos, J. J., Rossman, L. A., Grayman, W. M., Boulos, P. F., & Clark, R. M. (1997). Kinetics of chlorine decay. *Journal - American Water Works Association*

- ciation, 89, 54–65.
- Viñas, V., Sokolova, E., Malm, A., Bergstedt, O., & Pettersson, T. J. (2022). Cross-connections in drinking water distribution networks: Quantitative microbial risk assessment in combination with fault tree analysis and hydraulic modelling. *Science of the Total Environment*, 831, 154874.
- Vrachimis, S. G., Eliades, D. G., & Polycarpou, M. M. (2018). Leak detection in water distribution systems using hydraulic interval state estimation. In *Proceedings of IEEE Conference on Control Technology and Applications (CCTA)* (pp. 565–570).
- Vrachimis, S. G., Eliades, D. G., & Polycarpou, M. M. (2021). Calculating chlorine concentration bounds in water distribution networks: a backtracking uncertainty bounding approach. *Water Resources Research*, 57, e2020WR028684.
- Vrachimis, S. G., Lifshitz, R., Eliades, D. G., Polycarpou, M. M., & Ostfeld, A. (2020). Active contamination detection in water-distribution systems. *Journal of Water Resources Planning and Management*, 146, 04020014.
- Vrachimis, S. G., Timotheou, S., Eliades, D. G., & Polycarpou, M. M. (2019). Iterative hydraulic interval state estimation for water distribution networks. *Journal of Water Resources Planning and Management*, 145, 04018087.
- Wagner, D. E., Neupauer, R. M., & Cichowitz, C. (2015). Adjoint-based probabilistic source characterization in water-distribution systems with transient flows and imperfect sensors. *Journal of Water Resources Planning and Management*, 141, 04015003.
- Wang, C., & Zhou, S. (2017). Contamination source identification based on sequential bayesian approach for water distribution network with stochastic demands. *IIEE Transactions*, 49, 899–910.
- Wang, H., Hu, C., & Shi, B. (2021). The control of red water occurrence and opportunistic pathogens risks in drinking water distribution systems: A review. *Journal of Environmental Sciences*, 110, 92–98.
- Wang, K., Wen, X., Hou, D., Tu, D., Zhu, N., Huang, P., Zhang, G., & Zhang, H. (2018). Application of least-squares support vector machines for quantitative evaluation of known contaminant in water distribution system using online water quality parameters. *Sensors*, 18, 938.
- Wang, S., Taha, A. E., Chakrabarty, A., Sela, L., & Abokifa, A. A. (2022a). Model order reduction for water quality dynamics. *Water Resources Research*, 58, e2021WR029856.
- Wang, S., Taha, A. E., Gatsis, N., Sela, L., & Giacomoni, M. H. (2022b). Probabilistic state estimation in water networks. *IEEE Transactions on Control Systems Technology*, 30, 507–519.
- Watson, J.-P., Murray, R., & Hart, W. E. (2009). Formulation and optimization of robust sensor placement problems for drinking water contamination warning systems. *Journal of Infrastructure Systems*, 15, 330–339.
- Weickgenannt, M., Kapelan, Z., Blokker, M., & Savic, D. A. (2010). Risk-based sensor placement for contaminant detection in water distribution systems. *Journal of Water Resources Planning and Management*, 136, 629–636.
- World Health Organization (2008). *Guidelines for drinking-water quality, Vol. 1, 3rd edition incorporating 1st and 2nd addenda*. (3rd ed.). Geneva, Switzerland: World Health Organization.
- World Health Organization (2017). *Guidelines for drinking-water quality, 4th edition, incorporating the 1st addendum*. (4th ed.). Geneva, Switzerland: World Health Organization.
- Xie, G., Roiko, A., Stratton, H., Lemckert, C., Dunn, P. K., & Mengersen, K. (2016). A generalized QMRA beta-poisson dose-response model. *Risk Analysis*, 36, 1948–1958.
- Xie, M., & Brdys, M. (2015). Nonlinear model predictive control of water quality in drinking water distribution systems with DBPs objectives. *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, 9, 354–360.
- Xie, M., Wang, P., Zhang, X.-P., & Jayaweera, D. (2018). Robust parameter estimation and output prediction for nonlinear water quality control in water distribution systems. *Journal of Water Resources Planning and Management*, 144, 04017092.
- Xin, K.-l., Tao, T., Li, S., & Yan, H. (2017). Contamination accidents in China's drinking water distribution networks: status and countermeasures. *Water Policy*, 19, 13–27.
- Xu, J., Johnson, M. P., Fischbeck, P. S., Small, M. J., & VanBriesen, J. M. (2010). Robust placement of sensors in dynamic water distribution systems. *European Journal of Operational Research*, 202, 707–716.
- Xuesong, Y., Jie, S., & Chengyu, H. (2017). Research on contaminant sources identification of uncertainty water demand using genetic algorithm. *Cluster Computing*, 20, 1007–1016.
- Yan, X., Zhao, J., Hu, C., & Zeng, D. (2019). Multimodal optimization problem in contamination source determination of water supply networks. *Swarm and Evolutionary Computation*, 47, 66–71.
- Yan-jun, L., & Qian, M. (2012). AP-LSSVM modeling for water quality prediction. In *Proceedings of the 31st Chinese Control Conference* (pp. 6928–6932). IEEE.
- Zechman, E. M. (2013). Integrating evolution strategies and genetic algorithms with agent-based modeling for flushing a contaminated water distribution system. *Journal of Hydroinformatics*, 15, 798–812.
- Zechman, E. M., & Ranjithan, S. R. (2009). Evolutionary computation-based methods for characterizing contaminant sources in a water distribution system. *Journal of Water Resources Planning and Management*, 135, 334–343.
- Zhao, H., Hou, D., Huang, P., & Zhang, G. (2015). Periodic pattern extraction and anomaly detection for free chlorine in drinking water network. *Water Science and Technology: Water Supply*, 15, 541–551.
- Zhong Wang, Polycarpou, M., Uber, J., & Feng Shang (2006). Adaptive control of water quality in water distribution networks. *IEEE Transactions on Control Systems Technology*, 14, 149–156.
- Zhou, Y., Jiang, J., Qian, K., Ding, Y., Yang, S.-H., & He, L. (2021). Graph convolutional networks based contamination source identification across water distribution networks. *Process Safety and Environmental Protection*, 155, 317–324.
- Zulkifli, S. N., Rahim, H. A., & Lau, W.-J. (2018). Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications. *Sensors and Actuators B: Chemical*, 255, 2657–2689.