



Novel Eulerian Approach with Cellular Automata Modelling to Estimate Water Quality in a Drinking Water Network

M. A. Bonora¹ · G. Capano² · A. De Rango³ · Mario Maiolo⁴ 

Received: 8 July 2022 / Accepted: 26 September 2022 / Published online: 10 October 2022
© The Author(s) 2022

Abstract

The forecast analysis of the exposure to the contamination risk in a water distribution network requires increasing the quality of the applied input/outputs modeling. This need involves using non-traditional models responding to the increasingly high computation requirements. In this scenario, the Cellular Automata paradigm represents a new frontier with considerable potential. Specifically, this paper describes the Eulerian Water quality Modeling—Cellular Automata (EWAM-CA) model, aimed at simulating the sodium hypochlorite (chlorine) injection, transport, and reaction phase in a medium-sized drinking water network. The EWAM-CA accuracy was compared with the Epanet software on a Fossolo water network, in Bologna town (Italy), considering a constant and an impulsive input respectively. Due to CA's intrinsic aptitude for parallel computing, a parallel version of EWAM-CA was developed. Moreover, using the capability of the cellular automata to manage the modeling asynchronously, improving the computational efficiency, we propose a novel approach based on activation/deactivation asynchronous rules, avoiding unnecessary calculations in nodes or pipes where no pollution occurs. The different EWAM-CA versions were compared for the case study, and the parallel EWAM-CA approach coupled with asynchronous functionality significantly improved computational performance.

Keywords Water network modelling · Water quality · Forecasting contamination event · Cellular Automata · Extended Cellular Automata

✉ Mario Maiolo
mario.maiolo@unical.it

¹ Virtual Laboratory of Smart Water Management, NTT DATA Italia S.p.A., via Spagna 50, 87036 Rende, Italy

² Virtual Laboratory of Smart Water Management, Green HoMe S.C.Ar.L, via Ponte P. Bucci, 87036 Rende, Italy

³ Department of Environmental Engineering (DIAM), University of Calabria, via Ponte P. Bucci, 87036 Rende, Italy

⁴ Virtual Laboratory of Smart Water Management, University of Calabria, via Ponte P. Bucci, 87036 Rende, Italy

1 Introduction

The problems related to managing drinking water distribution networks (WDN), together with their interdependence with other technological infrastructures, require in-depth analysis, mainly when sensitive aspects affect their functioning. Furthermore, the interest related to the management of WDNs depends on the status of Critical Infrastructures (Adedoja et al. 2018) because the alteration of the correct functioning determines considerable impacts on the population's well-being (Johansen and Tien 2018), not only in terms of distributed resource but mainly on the quality of the same. Therefore, it is a priority to have tools capable of estimating the level of exposure to the risks associated with their optimal management to make it sustainable.

In this context, sustainable management models must be based on accurate planning of the entire life cycle of the network, with a short and long-term vision and considering the possible objectives to be achieved (Maiolo et al. 2020). These aspects, linked to quantitative and qualitative assessments, represent the input data for the management models. Expressly, the quality of the water resource assumes a priority role, ensuring compliance with internationally regulated qualitative parameters (e.g., 2000/60/EC) and sustainability requirements and protecting against the risk of contamination.

Identifying sources of contamination in a WDN represents a literary trend that has developed over time. The literature offers different models for estimating quality, prediction, and monitoring (Capano et al. 2019). Generally, it is impossible to disregard the analysis of the context related to modeling the movement and reactions of contaminants within a WDN. The studies of particle transport in computational fluid dynamics mainly show two approaches, the Lagrangian and the Eulerian, respectively (Zhang and Chen 2007). The Lagrangian approach describes the phenomena focusing on the particles that are chased in motion. The Eulerian approach describes phenomena from a specific point in space through which the particles transit. Both methods presuppose the previous knowledge of the WDN hydraulic operating (pipe flow direction and value) throughout the simulated period.

In general, modeling the quality of the resource in a drinking water network requires the calculation of three primary parameters: the position of the source, the concentration, and the intrusion time of the contaminant. The literature identifies different modeling and problem-solving approaches. Sun et al. (2019) analyzes the spatial-temporal characteristics of user alerts during a contaminant intrusion event, proposing a new analysis methodology aimed at reducing the arrival times of reports based on the Convolution Neural Network (CNN). Capano et al. (2019) presented a quick procedure for identifying one or more sensitive nodes in a network using a mainly topological approach called Identification of Contamination Potential Source (ICPS). Benamar et al. (2020) presented a study to obtain detailed information on global changes in Physico-chemical parameters at different sampling points. Ortega et al. (2020) use a Bayesian approach for the source of contamination determination, using an algorithm that reports the probability that each node is the source to explain the correlations between the sampling positions, defining a classification. Grbčić et al. (2021) present a complete framework for identifying the source of contamination (with a machine learning algorithm), the contamination times (with a stochastic fireworks optimization algorithm), and the injected contaminant concentration (through optimization and Random Forest algorithm).

However, the modeling and technological advancement required using forecast analyses point out that traditional models appear unsuitable for the scope. For example, in many

cases present in the literature, the analysis of water systems is based only on optimization models of hydraulic properties (velocity, pressure, length of pipelines) (Hassanvand et al. 2021). These represent the input of models with different analytical characteristics, including the Neural Network and the Fuzzy approach, which, through Artificial Intelligence (Salimi et al. 2020), increase the complexity of traditional modeling approaches. This aspect is a priority for the frequent use of management tools, such as the Decision Support System (DSS) (Grimaldi et al. 2020; Pagano et al. 2021), which require accurate and reliable estimates of the parameters.

In his attempt to understand the fundamental mechanisms behind self-reproduction, the Cellular Automata (CA) (Neumann and Burks 1966) are considered a new approach to treating some complex systems whose behavior may be expressed in terms of local laws. The system's complexity emerges from the interactions of its elementary (cellular) units by applying relatively simple local rules. Hence, it can be considered an alternative approach to differential equations and a general formalism that can represent and solve numerical methods. CA has been used to simulate many natural complex phenomena. In particular, Extended Cellular Automata (XCA), which is an extension of the original CA computational paradigm, have proven to be suitable for simulating complex natural phenomena like unsaturated water flow (Mendicino et al. 2006; De Rango et al. 2021), lava flows (Spataro et al. 2017), forest fires (Avolio et al. 2014).

Moreover, another type of CA exists, the Graph-based CA (Wu and Rosenfeld 1979), which has been widely used (Małeckı et al. 2019). The principal difference from the classical CA is the definition of the neighborhood modeled as a directed graph.

CA-based models provide other opportunities besides their natural predisposition to parallel computing, which can exploit to improve further computational speed. Specifically, asynchronous functionality optimization can be developed, allowing additional computation speed.

In this paper, the definition of asynchronous functionality is similar to that of Furnari et al. (2021), where a rule is imposed to allow only a part of the domain (called active cells or automata) to evolve step by step to the next state and keep quiescent (i.e., remaining in the same state) the residual part (called non-active cells or automata). From a computational point of view, this functionality avoids unnecessary calculations related to non-active parts of the domain which does not change their state during the simulation execution.

For WDNs the use of classical CA-based models is limited. This aspect is mainly related to the not fixed neighborhood since a WDN is not a classical regular grid. In Keedwell and Khu (2006), one of the first applications of CA to a WDN is proposed. This study introduces the CA approach in the context of optimal design in WDN, developing a direct performance comparison with genetic algorithms (GA). Guo et al. (2007), following the comparative approach between CA and GA, present a structured model with CA, which guarantees better performance than multi-objective genetic algorithms in terms of optimization efficiency. From a qualitative point of view, interest in the combination of CA and water quality modeling has been developing in recent years. Afshar and Hajiabadi (2018) propose a CA approach based on Parallel Cellular Automata (PCA) for multi-objective reservoir operation optimization. In Abhijith and Mohan (2020), the CA approach is used to predict chlorine's temporal and spatial variations in a WDN, using stochastic Lagrangian techniques to represent advection and dispersion processes. Abhijith and Mohan (2021) use a CA approach to model microbial growth and trihalomethane formation in chlorinated water distribution systems.

In this paper, we propose an Eulerian Water quality Modeling—Cellular Automata (EWAM-CA), based on a finite volume forecasting model formalized in the Graph-based

XCA approach in the water resource quality modeling proposed in Rossman et al. (1993). EWAM-CA enables the modeling of the transport and reaction of a chemical substance mixed with water within pipelines of a drinking water network. The model's potential is evaluated in the application by simulating chlorine injection in a medium-sized network with a steady-state hydraulic simulation. The model is structured to adapt to unsteady-state hydraulic simulations and model different reaction mechanisms of the transported substances. In detail, the EWAM-CA model performs a discretization of volumes, according to Rossman et al. (1994). Each pipeline is divided into a defined number of sub-volumes compatible with the hydraulic operating regime of the network. The chemical substance in each sub-volume is routed downstream, updating the states of the automata representing the network nodes.

The parallel implementation of the EWAM-CA model, together with the application of a new asynchronism rule, is applied on a Fossolo network (Bragalli et al. 2008; Meirelles et al. 2018). This network, single source, includes 36 nodes, 58 pipes that extend for about 8.4 km. For quality simulation a sodium hypochlorite is chosen as input substance, with a constant and an impulsive input, useful to comparing the result compared with the Epanet software. The main aim is to evaluate the potential of the EWAM-CA parallel implementation coupled with the asynchronous functionality in terms of computational efficiency.

Then, in Sect. 2 the criteria relating to the hydraulic modeling of the water network and the modeling with CA are detailed, including a case study presentation. In the Sect. 3 the results of the EWAM-CA application and the relative computational performances are presented. The conclusions are explained in Sect. 4.

2 Materials and Methods

2.1 Water Quality Modeling in an Urban Network

The hydraulic simulation aims to model the WDN operating regime starting from known data. On the other hand, the quality simulation goal is the knowledge of how chemical substances move and react in the network. Quality models use as input the hydraulic operating regime of and the characteristics of the introduced substances (diffusivity, solubility, concentration). Rossman et al. (1993) present a solution approach using a finite volume Eulerian scheme. The method assumes that the pipes are divided into equally sized volume segments once a temporal discretization is defined. The volumes composing each pipe are considered thoroughly mixed, limiting the numerical diffusion.

The authors, in this paper, propose a computation method based on the Eulerian finite volume method presented in Rossman et al. (1993). The model uses the operator splitting technique on the advective flow equation.

The tubes are divided into sub-volumes depending on the quality time step chosen. For each quality time interval, each volume element undergoes the reaction and dispersion (Eq. (1)) and is subsequently transported (Eq. (2)) following the hydraulic operating regime.

$$\frac{\partial C}{\partial t} - D \frac{\partial^2 C}{\partial x^2} + k_b C = 0 \quad (1)$$

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = 0 \tag{2}$$

In which: C is the concentration of the chemical substance along the pipe, u is the flow average velocity, x represent the position in the i -st pipe ($x = 0$ pipe start, $x = L_i$ pipe end, L_i is the i -st pipe length), D is the diffusion, k_b represent the reaction rate coefficient.

Rossman et al. (1993) use a criterion for dividing pipes into elements based on the pipe travel time:

$$\tau = \min\left(\frac{V_i}{Q_i}\right) \tag{3}$$

$$\eta_i^* = \left(\frac{V_i}{\tau Q_i}\right) \tag{4}$$

$$v_i = \frac{V_i}{\text{nearest integer}(\eta_i^*)} \tag{5}$$

where τ is the quality simulation time interval, i is the index pipes, V_i is the total volume, Q_i is the flow rate, η_i^* is the non-integer number of finite sub-elements, v_i is the volume of the η_i sub-elements. To assess the effectiveness of the EWAM-CA to model the introduction and transport of a chemical substance model, we compare the results obtained with the Epanet software, characterized by a Lagrangian solution approach (Liou and Kroon 1987). Epanet (Rossman 2000) is a software developed by the USEPA (the United States Environmental Protection Agency). It is one of the most known software for WDN hydraulic simulation. It allows performing hydraulic simulations of chemicals within a WDN.

Herein we chose sodium hypochlorite as the input chemical substance. This substance is commonly used for water disinfection because the reaction kinetics are well known. This substance, used for water disinfection, undergoes a decay process in the network. The sodium hypochlorite decay process releases free chlorine, which functions well in water quality control. In additions, its systematic determination allows recognizing degradative processes in progress in the network, providing important information on the disinfection process.

There are many chlorine decay analysis models. Most of these models refer to the one proposed by Rossman et al. (1994), which laid the content foundations for the quality simulation criteria setting in Epanet 2 (Rossman 2000). EWAM-CA simulates chlorine concentration decay using the one-dimensional conservation of mass equation for a dilute concentration of total free chlorine in water flowing through a section, proposed in Rossman et al. (1994) to use the Eq. (1):

$$\frac{\partial C}{\partial t} = -u \frac{\partial C}{\partial x} - k_b C - \frac{k_w(C - C_w)}{r} \tag{6}$$

where C is the chlorine concentration in the bulk flow, u is the flow average velocity, k_b is the bulk flow decay rate, C_w chlorine concentration at pipe wall and r is the hydraulic radius.

Equation (6) excludes the dispersion term, considered negligible compared to the others. The decision to neglect k_w depends on the variability of this parameter, linked to the temperature of the water, and the actual conditions of the pipe material. It is usually determined

during the calibration phase of the model. According to Eq. (6) the concentration depends on three terms. The first term considers the advection due to the water flow, while the second and third represent the decay in the bulk flow and on the pipe wall, respectively. In the often recurring hypothesis of considering the total chlorine decay mechanism within the bulk flow, Eqs. (1) and (2) can be rewritten as follows:

$$\frac{\partial C}{\partial t} = -k_B C \quad (7)$$

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = 0 \quad (8)$$

in which

$$C(t) = C(t - \Delta t)e^{-k_B \Delta t} \quad (9)$$

where k_B represents the total bulk decay rate.

2.2 Cellular Automata Water Quality Modeling

CA can be considered a time/space discrete model where space is represented by a d -dimensional (i.e., 1D, 2D, 3D, ...) structured grid of cells. At the beginning of the simulation, cells are in an arbitrary state, and the CA evolves by changing the states of the cells in discrete time steps by applying the same evolution law to each of them simultaneously. The evolution law can also consider the states of neighboring cells, defined by neighborhood conditions, a geometrical pattern invariant in time and space. Despite their simple definition, CA may give rise to highly complex behavior at a macroscopic level. Even if the local laws that regulate the system's dynamics are known, the system's global behavior can be tough to predict. In general, CA is formally identified by a quadruple which describes its fundamental characteristics. The quadruple is composed of: cellular space, in which the automata can evolve (e.g., one-, two-, three-dimensional); the pattern of the neighborhood, which identifies the spatial correlation between the cells in the computational domain and cannot change over time (e.g., Von Neumann or Moore neighborhood); the finite set of the cells states; the deterministic transition function, which allows evolving the cellular automata over time. It is simultaneously applied to each cell of the CA, and according to the neighborhood pattern, the cell changes its state. Moreover, the keys properties of the CA are: parallelism the state of each cell is updated independently from the others; locality, the new state of the cell depends only on its old state and the state of the neighborhood; homogeneity, the same deterministic transition function is applied to each cell of the domain.

Many different CA-based methods can be found in the literature to simulate many natural phenomena. XCA (Di Gregorio and Serra 1999) represent an extension of the original CA computational paradigm. The XCA formalism introduces some extensions concerning the classical CA. Major novelties regard the possibility of decomposing the state of the cells in substates (i.e., chemical concentration, total water volume, etc.) related to the phenomenon to be modeled. The cell state thus belongs to the Cartesian product of the considered substates. Phenomenon parameter behaviors, such as total bulk decay rate, can also be considered. In addition, the state transition function is divided into elementary processes, describing particular aspects of the supposed phenomenon.

A Graph-based XCA has been here considered as an explicit solver. The Graph-based XCA preserves all fundamental characteristics of the XCA. Nevertheless, the condition

of the neighborhood has been relaxed to be considered as a directed graph. The choice of the Graph-based XCA was dictated by the context of WDN in which a topological position describes the relationship between the various nodes of the network.

The formal definition of the Graph-based XCA is described by the following formulation:

$$EWAM-CA = \langle D, G, S, P, \sigma \rangle \tag{10}$$

where:

$D = [0, n - 1] \subset \mathbb{Z}$ is the one-dimensional discrete computational domain, with n representing the number of cells.

G is a directed graph $G = (V, E)$, which is defined by a set of nodes $V(G)$, set of edges $E(G)$. Each vertex v_i corresponds to one automaton cell and identifies a set of elements such as tanks, reservoirs, and junctions, while the edge represents the longitudinal elements such as pipes and valves. The neighborhood of the vertex $v_i \in V(G)$ is the set of vertices $X(v_i) = X_{in}(v_i) \cup X_{out}(v_i)$, in which $X_{in}(v_i) = \{v_j | v_j v_i \in E(G)\}$ and $X_{out}(v_i) = \{v_j | v_i v_j \in E(G)\}$.

In particular, the edge between a generic node v_i and node v_j is modeled to be subdivided into η_{ij} equal size volume elements, which identify the discrete mass value over the pipe, and are calculated by Eqs. (3) and (4).

S is the set of finite states of the cell that is defined as:

$$S_{v_i} = S_c \times S_f^{|X(v_i)|} \times S_m^{n_1} \times \dots \times S_m^{n_{|X(v_i)|}} \times S_c^{n_1} \times \dots \times S_c^{n_{|X(v_i)|}} \tag{11}$$

where $S_c[Mg/l]$ identifies the chemical concentration of the node, while the states $S_f[l/s]$ represent the neighborhood relationship with the other automata, $S_f^{|X(v_i)|}$ is the flow referring to the pipes connecting to the vertex v_i , $S_m^{n_1} \times \dots \times S_m^{n_{|X(v_i)|}}$ and $S_c^{n_1} \times \dots \times S_c^{n_{|X(v_i)|}}$ are the mass of the chemical and the concentration referring to pipes connecting the vertex v_i and the vertex v_j and subdivided into η_{ij} sections.

$P = \{k_b\}$ is the set of parameters identified, where k_b is the total bulk decay rate.

$\sigma = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ is the deterministic transition function. It is composed of the following elementary processes, listed below in the same order as they are applied.

It allows to define the state of the automaton at time $t + 1$ and depends on the state of its neighborhood at time t :

$$\sigma_1 : S_m^{n_1} \times \dots \times S_m^{n_{|X(v_i)|}} \rightarrow S_m^{n_1} \times \dots \times S_m^{n_{|X(v_i)|}} \tag{12}$$

σ_1 allows calculating the mass of chemicals by considering the elements inside the edge between node v_i and node v_j by applying the following equation:

$$m'_{ij}{}^k = m_{ij}{}^k e^{k_b \tau} \quad \forall k \in \{1, \dots, \eta_{ij}\} \tag{13}$$

where $m_{ij}{}^k$ is the chemical mass in the k -st element in the pipe between a generic node v_i and node v_j before reaction and $m'_{ij}{}^k$ is the new value after the reaction.

$$\sigma_2 : S_m^{n_1} \times \dots \times S_m^{n_{|X(v_i)|}} \times S_f^{|X(v_i)|} \rightarrow S_c \tag{14}$$

σ_2 updates the chemical concentration within the cell as:

$$M_{v_i} = \sum_j^{|X_{in}(v_i)|} m_{ji}^{n_{ji}-1} \quad (15)$$

$$V_{v_i} = \sum_j^{|X_{in}(v_i)|} \tau Q_{v_j} \quad (16)$$

$$C'_{v_i} = \frac{M_{v_i}}{V_{v_i}} \quad (17)$$

where M_{v_i} is the total mass of the chemical entering the i -st vertex, $m_{ji}^{n_{ji}}$ is the mass of the chemical in the last sub-element (n_{ji}) in the pipe between a generic node v_i and node v_j , V_{v_i} is the total water volume entering the i -st node, and C_j^{in} is the chemical concentration in the i -st node.

$$\sigma_3 : S_m^{n_{i1}} \times \dots \times S_m^{n_{i|X(v_i)|}} \rightarrow S_m^{n_{i1}} \times \dots \times S_m^{n_{i|X(v_i)|}} \quad (18)$$

σ_3 defines the update the mass and the concentration using the following equation:

$$m_{ij}^{k+1} = m_{ij}^k \quad \forall k \in \{1, \dots, n_{ij} - 1\} \quad (19)$$

where m_{ij}^k identifies the k^{th} mass element in the pipe between the generic node v_i and v_j .

$$\sigma_4 : S_m^{n_{i1}} \times \dots \times S_m^{n_{i|X(v_i)|}} \times S_f^{|X(v_i)|} \rightarrow S_m^{n_{i1}} \times \dots \times S_m^{n_{i|X(v_i)|}} \quad (20)$$

σ_4 allows to calculate the mass by using the following equation:

$$m_{ij}^1 = C_{v_i} Q_{v_i} \tau \quad (21)$$

Where m_{ij}^1 identifies the first mass element in the pipe between a generic node v_i and node v_j and C_{v_i} , Q_{v_i} and τ correspond to the concentration, the discharge, and the time interval, respectively.

Figure 1 illustrates, as pseudo-code, the primary simulation process of the EWAM-CA. The EWAM-CA was developed in C++. At the beginning of the process, the `init()` function allows the loading the computational domain by defining the relation between the cells of the Graph-based XCA. The main cycle, lines 7–14, permits the evolution of the EWAM-CA model by applying at each cell the set of the transition function. Finally, lines 11–12 allow updating the set of the state of the cellular space.

2.3 Making Cellular Automata Parallel and Asynchronous

The OpenMP API was considered to exploit the CA intrinsic parallelism property, supporting multi-platform shared-memory parallel programming in C/C++. This choice is mainly related to OpenMP portability, a simple and flexible interface that can exploit parallel computation from a workstation to supercomputer architectures.

The asynchronous functionality permits narrowing the computation only to the set of active nodes, allowing to skip the not active nodes, meaning that not only the transition function is not computed, but also that the not active cells are entirely skipped. At the


```

1 init()
2 forall  $q \in S$  do
3   update ( $q$ ) // Update the substate  $q$ 
4    $step \leftarrow initial\_step$ 
5    $tau \leftarrow time\_step$ 
6    $delta\_t\_simulation \leftarrow 0$ 
7   while ( $step \leq final\_step \parallel delta\_t\_simulation \leq final\_total\_time$ ) do
8     forall  $e \in \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$  do
9       forall  $i \in D$  do
10         $e(i)$  // Transition function  $e$  to the cell  $i$ 
11        forall  $q \in S$  do
12         update ( $q$ ) // Update the substate  $q$ 
13         $delta\_t\_simulation \leftarrow delta\_t\_simulation + tau$ 
14         $step \leftarrow step + 1$ 
15 return

```

Fig. 1 EWAM-CA pseudo-code to describe the main simulation process

beginning of the computation, all the node state is set to non-active unless for the source node. During the simulation, a node can set to active one of the neighborhood nodes by considering the following rule:

$$\text{active if } \exists i \in [1, \dots, |X_{in}(v_c)|] | m_{ic}^{n_{ic}-1} > 0 \quad (22)$$

A node can change its state to non-active using the following rule:

$$\begin{aligned} \text{nonactive if } C_{v_c} = 0 \wedge \forall i [1, \dots, |X_{out}(v_c)|] | \sum_{k=0}^{n_{ci}} m_{ci}^k \\ = 0 \wedge \forall i [1, \dots, |X_{in}(v_c)|] | \sum_{k=0}^{n_{ic}} m_{ic}^k = 0 \end{aligned} \quad (23)$$

More in detail, a node activates a neighborhood node if the chemical mass of the last element of the pipe is greater than zero. Instead, a node deactivates itself if its concentration is zero, the sum of all chemical mass in its outgoing pipe is zero, and the sum of all chemical mass in its incoming pipe is zero.

Applying the above rules allows skipping the non-active node's $\sigma_1, \sigma_2, \sigma_3$, and σ_4 deterministic transition functions computation altogether. Since the transport of a chemical substance does not necessarily and immediately involve all the network nodes, an asynchronous implementation deactivates the automata that do not participate in the chemical transport, leading to an increase in computational performance.

2.4 Case Study

The chosen case study is a network presented in Bragalli et al. (2008) and Meirelles et al. (2018) (Fig. 2). The network supplies water to a District Metered Area (DMA) in the city of Bologna. The DMA extends over 0.5 square kilometers.

The WDN counts 36 nodes, an average altitude of 64.3 a.s.l., and 58 pipes extending for about 8.4 km. The network has a single source modelled by a reservoir (tank of infinite volume) with a fixed hydraulic head of 121 a.s.l., which provides a flow of 33.9 l/s to the served area.

3 Results

The comparison between the Epanet solution model and the EWAM-CA concerns a steady flow hydraulic simulation. The simulation duration covers a total period of 1 h. The limited simulation duration is due to the small size of the network. The topologically farthest node from the reservoir has a water age of 18.6 min. The quality simulation models two scenarios. The first considers as input a constant amount of chlorine for the whole simulation period. The second one considers as inputs the same chemical concentration for a shorter period (10 min).

The chemical input affects node "1", located immediately downstream of the reservoir the chemical input is modeled by setting the concentration of each flow, leaving the node to the chosen value (500 $\mu\text{g/l}$). This chemical input corresponds to the setpoint booster source in the Epanet quality simulation module.



Fig. 2 Fossolo plan view. Google “Satellite” background

The limited simulation duration imposes short time-steps. Epanet hydraulic simulation is set to 5 min. Fossolo network is characterized by variable travel times ranging from 18.5 s to more than 40 min. According to what is reported in Rossman et al. (1993), the quality time-step must be at most equal to the shortest pipe time travel in the network (Eqs. (3), (4), and (5)) to ensure that water is not transported beyond its downstream boundary node. Therefore, for both the presented scenarios a time-steps of 10 s is used.

Figures 3 and 4 show a comparison between the results of the models for both scenarios. As seen from the figures, the behavior of the applied model is very similar to that

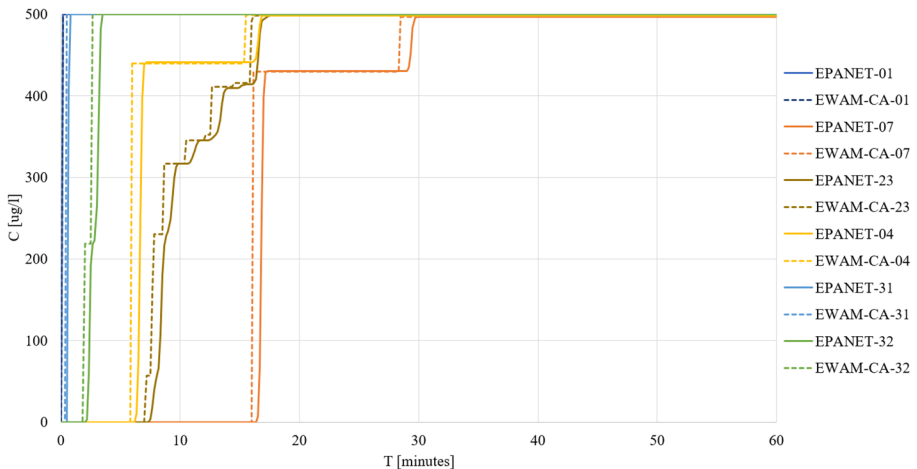


Fig. 3 Comparison between the simulation results with the $\Delta t = 10$ s and a constant input

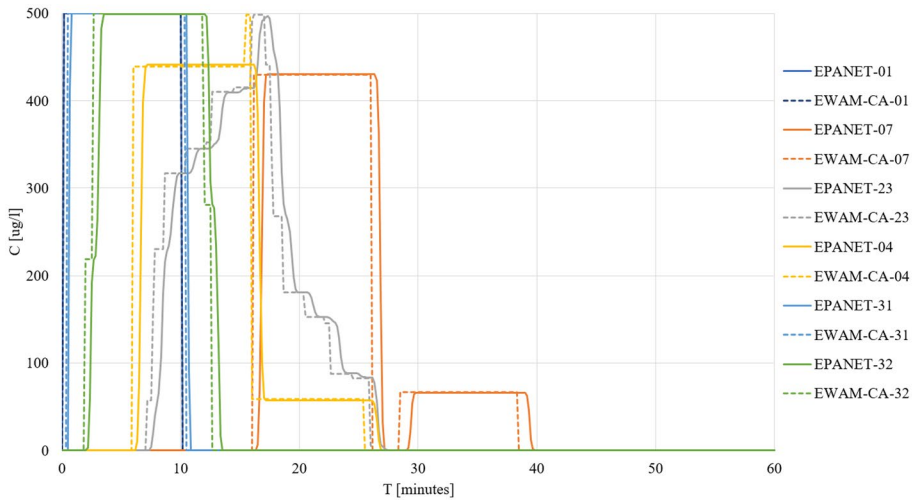


Fig. 4 Comparison between the simulation results with the $\Delta t = 10$ s and an impulsive input

of Epanet. The EWAM-CA model is characterized by the arrival times of the contaminant slightly lower than those of Epanet. This is due to the time shift error that characterizes EWAM-CA. As shown, the model is susceptible to choosing the quality timestep.

3.1 Computational Performance

The performance of EWAM-CA implementation is evaluated by considering the execution on a cluster node equipped with Intel(R) Xeon(R) Gold 6128 CPU @ 3.40 GHz with 64 GB of RAM. Two variants of the algorithm were tested for the above mentioned scenarios. The first former is based only on the Parallel Implementation, called hereafter PI. The second one is based on the Parallel and the Asynchronism Implementation, called hereafter PAI. The automata in the PAI are initially inactive and activated when the movement involves the automata connected to them. The automata are kept deactivated if they do not participate in the transport.

To evaluate the performance improvement of the PA and PAI with respect to the serial algorithm (i.e., Not parallelized), the use of the speedup is introduced. The speedup allows to characterize the performance improvement of a parallelized algorithm. The speedup is the ratio of the running time of the serial algorithm of its parallelized variant. The higher the speedup value, the better the performance gain due to parallelization.

The first series of experiments are shown in Fig. 5. The PAI version achieved the best speedup, approximately 3.2, with eight cores. The PAI demonstrated higher computational performance in all experiments than the PI implementation. Figure 5b shows the number of active nodes during the simulation. Since a constant injection rate is considered, the number of active nodes never decreases and reaches the total number of nodes approximately after 1000 s.

The second series of experiments are shown in Fig. 6. Again, the PAI version achieved the best speedup, approximately 4.2, with eight cores. The PAI demonstrated higher computational performance in all experiments than the PI implementation.

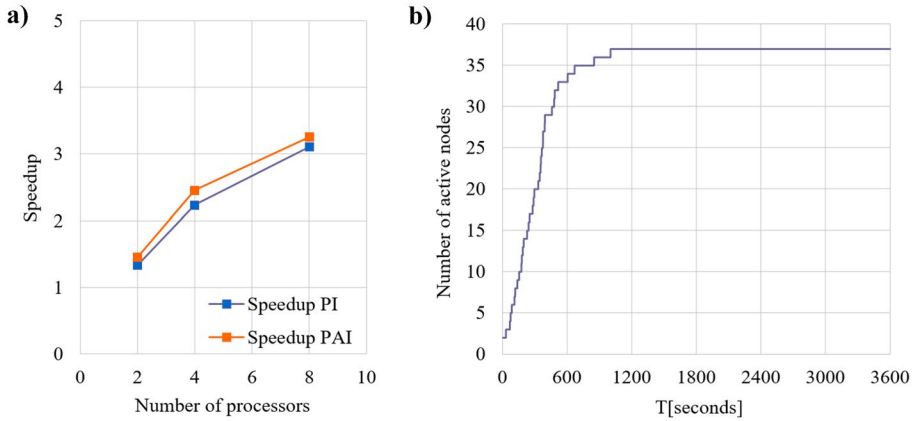


Fig. 5 **a** shows the speedup achieved by the PI and PAI versions. **b** illustrates the number of the active nodes in the system during the simulation for the PAI implementation

Figure 5b shows the number of active nodes during the simulation. Since an impulsive input is considered, the number of active nodes increases and decreases, reaching a peak approximately at the 500 s. Moreover, the gap between the two implementations is more noticeable in the first series of experiments. This trend is related to the asynchronous functionality, which can significantly improve the performance when only some of the nodes are involved in the contaminant propagation.

The PAI version achieved better speedup in all experiments, mainly concerning the impulsive injection scenario. While the performance gain is reduced, considering a constant input. This latter scenario is configured as the worst case since the input reaches all the network nodes, and being constant prevents the cells from changing their state to non-active. Otherwise, considering an impulsive injection, the performance gain is mainly due to the deactivation of the cells, which are not involved in the propagation process, as seen in Fig. 6.

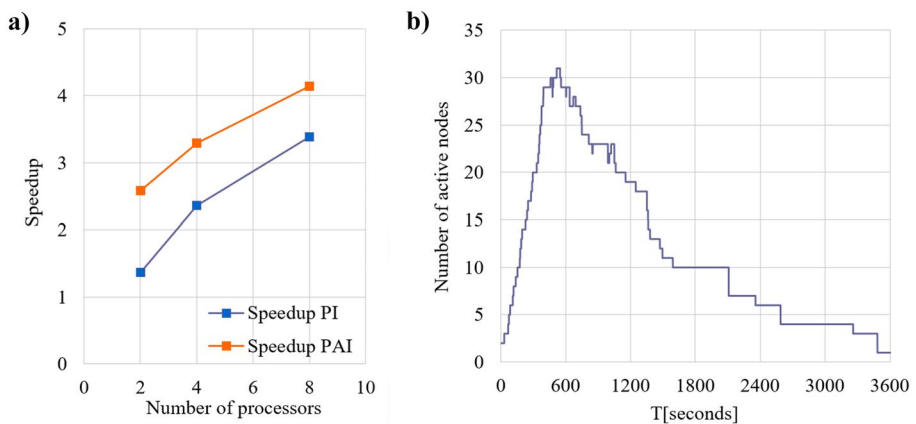


Fig. 6 **a** shows the speedup achieved by the PI and PAI versions. **b** illustrates the number of the active nodes in the system during the simulation for the PAI implementation

Further remarkable optimization can be achieved considering a constant source rate injection not located in the upper part of the network. The pollutant propagation process will probably not affect some network nodes that will never be activated by the model. Injecting a substance in a network-internal node reduces the maximum number of cells activated during a simulation.

4 Conclusions

Identifying a criterion to define the level of contamination risk in a water network requires increasingly sophisticated models and novel and performing computational tools. The proposed EWAM-CA interprets this need and declines it by integrating the CA paradigm in a non-continuous domain, such as a drinking water network, requiring the formalization of a Graph-based XCA approach.

The EWAM-CA simulation results have shown very similar behavior to the Epanet ones, confirming the reliability of the proposed model. Furthermore, the EWAM-CA computational efficiency has been improved by considering parallelization and asynchronous functionality. In particular, the introduction of the new asynchronous rules, which allows for locally managing the activation/deactivation of the automata, together with an efficient parallel implementation of the EWAM-CA, increased the computational simulation performance significantly. This aspect is crucial, especially when hundreds or thousands of simulations are in demand for real-time problems, such as contaminant-source detection.

Future developments are oriented to employ EWAM-CA model as an efficient simulation tool for a contaminant-source identification system. This latter will use a large batch of simulations over the domain by changing the source input for each network node with different source types and characteristics. Furthermore, the use of CA allows for building local rules that can be used to describe specific phenomena regarding the chemical transport and reactions within the networks, which is also suitable for emergency management.

Author Contribution All authors contributed to the study conception and design. Formal analysis and Validation, Mario Maiolo and Alessio De Rango; Methodology (definition of the AC modeling), Alessio De Rango; Methodology (definition of the hydraulic modeling), Mario Maiolo, Marco Amos Bonora, Gilda Capano; Software, Alessio De Rango and Marco Amos Bonora; Writing-original draft, Gilda Capano, Marco Amos Bonora and Alessio De Rango; Supervision, Mario Maiolo.

Funding Open access funding provided by Università della Calabria within the CRUI-CARE Agreement.

Data Availability The code used for this research is published at the following link: <https://github.com/alessioderango/EWAM-CA>

Declarations

Competing Interests The authors declare no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abhijith GR, Mohan S (2020) Random walk particle tracking embedded cellular automata model for predicting temporospatial variations of chlorine in water distribution systems. *Environ Proces* 7(1):271–296
- Abhijith GR, Mohan S (2021) Cellular automata-based mechanistic model for analyzing microbial regrowth and trihalomethanes formation in water distribution systems. *Int J Environ Eng* 147(1):04020145
- Adedoja OS, Hamam Y, Khalaf B, Sadiku R (2018) Towards development of an optimization model to identify contamination source in a water distribution network. *Water* 10(5):579
- Afshar MH, Hajibadi R (2018) A novel parallel cellular automata algorithm for multi-objective reservoir operation optimization. *Water Resour Manag* 32(2):785–803
- Avolio MV, Di Gregorio S, Trunfio GA (2014) A randomized approach to improve the accuracy of wildfire simulations using cellular automata. *J Cell Autom* 9(2–3):209–223
- Benamar A, Mahjoubi FZ, Ali GA, Kzaiber F, Oussama A (2020) A chemometric method for contamination sources identification along the Oum Er Rbia river (Morocco). *Bulg Chem Commun* 52:159–171
- Bragalli C, D'Ambrosio C, Lee J, Lodi A, Toth P (2008) Water network design by MINLP. IBM Research, Yorktown Heights
- Capano G, Bonora MA, Carini M, Maiolo M (2019) Identification of Contamination Potential Source (ICPS): a topological approach for the optimal recognition of sensitive nodes in a water distribution network. *International Conference on Numerical Computations: Theory and Algorithms*. Springer, Cham, pp 525–536
- De Rango A, Furnari L, Giordano A, Senatore A, D'Ambrosio D, Spataro W, Straface S, Mendicino G (2021) OpenCAL system extension and application to the three-dimensional Richards equation for unsaturated flow. *Comput Math Appl* 81:133–158. <https://doi.org/10.1016/j.camwa.2020.05.017>
- Di Gregorio S, Serra R (1999) An empirical method for modelling and simulating some complex macroscopic phenomena by cellular automata. *Future Gener Comput Syst* 16(2–3):259–271
- Furnari L, Senatore A, De Rango A, De Biase M, Straface S, Mendicino G (2021) Asynchronous cellular automata subsurface flow simulations in two- and three-dimensional heterogeneous soils. *Adv Water Resour* 153:103952. <https://doi.org/10.1016/j.advwatres.2021.103952>
- Grbčić L, Kranjčević L, Družeta S (2021) Machine learning and simulation-optimization coupling for water distribution network contamination source detection. *Sensors* 21(4):1157
- Grimaldi M, Sebillo M, Vitiello G, Pellecchia V (2020) Planning and managing the integrated water system: a spatial decision support system to analyze the infrastructure performances. *Sustainability* 12(16):6432
- Guo Y, Keedwell EC, Walters GA, Khu ST (2007) Hybridizing cellular automata principles and NSGAI for multi-objective design of urban water networks. *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, Berlin, Heidelberg, pp 546–559
- Hassanvand MR, Salimi AH, Kisi O, Omidvar Mohammadi H, Abouzari N (2021) Investigating application of adaptive neuro fuzzy inference systems method and Epanet software for modeling green space water distribution network. *Iran J Sci Technol Trans Civ Eng* 45(4):2765–2777
- Johansen C, Tien I (2018) Probabilistic multi-scale modeling of interdependencies between critical infrastructure systems for resilience. *Sustain Resilient Infrastruct* 3(1):1–15
- Keedwell E, Khu ST (2006) Novel cellular automata approach to optimal water distribution network design. *J Comput Civ Eng* 20(1):49–56
- Liou CP, Kroon JR (1987) Modeling the propagation of waterborne substances in distribution networks. *J Am Water Work Assoc* 79(11):54–58
- Maiolo M, Capano G, De Cicco R (2020) Metabolic approach for estimating the environmental loads associated with water distribution network of Rende: Life cycle assessment application with impact 2002+. *J Sustain Dev Energy Water Environ Syst* 9(4):1080323. <https://doi.org/10.13044/j.sdewes.d8.0323>
- Małecki K, Jankowski J, Szkwardowski M (2019) Modelling the impact of transit media on information spreading in an urban space using cellular automata. *Symmetry* 11:428
- Meirelles G, Brentan B, Izquierdo J, Ramos H, Luvizotto E (2018) Trunk network rehabilitation for resilience improvement and energy recovery in water distribution networks. *Water* 10(6):693
- Mendicino G, Senatore A, Spezzano G, Straface S (2006) Three-dimensional unsaturated flow modeling using cellular automata. *Water Resour Res* 42:W11419. <https://doi.org/10.1029/2005WR004472>
- Neumann J, Burks AW (1966) *Theory of self-reproducing automata*. University of Illinois Press, Illinois, Urbana and London
- Ortega E, Braunstein A, Lage-Castellanos A (2020) Contamination source detection in water distribution networks using belief propagation. *Stoch Environ Res Risk Assess* 34(3):493–511
- Pagano A, Giordano R, Vurro M (2021) A decision support system based on AHP for ranking strategies to manage emergencies on drinking water supply systems. *Water Resour Manag* 35(2):613–628

- Rossman LA (2000) EPANET 2: users manual
- Rossman LA, Boulos PF, Altman T (1993) Discrete volume-element method for network water-quality models. *J Water Resour Plan Manag* 119(5):505–517
- Rossman LA, Clark RM, Grayman WM (1994) Modeling chlorine residuals in drinking-water distribution systems. *Int J Environ Eng* 120(4):803–820
- Salimi A, Karami H, Farzin S, Hassanvand M, Azad A, Kisi O (2020) Design of water supply system from rivers using artificial intelligence to model water hammer. *ISH J Hydraul* 26(2):153–162
- Spataro D, D'Ambrosio D, Filippone G, Rongo R, Spataro W, Marocco D (2017) The new SCIARA-fv3 numerical model and acceleration by GPGPU strategies. *IJHPCA* 31(2):163–176
- Sun L, Yan H, Xin K, Tao T (2019) Contamination source identification in water distribution networks using convolutional neural network. *Environ Sci Pollut Res* 26(36):36786–36797
- Wu A, Rosenfeld A (1979) Cellular graph automata. I. Basic concepts, graph property measurement, closure properties. *Inf Control* 42(3):305–329
- Zhang Z, Chen Q (2007) Comparison of the Eulerian and Lagrangian methods for predicting particle transport in enclosed spaces. *Atmos Environ* 41(25):5236–5248

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.