Manuscript
Click here to view linked References

Groundwater Pollution Transport Modelling using Weighted Cellular Automata approach

Miloš Milašinović[a,*], Anja Ranđelović[a], Nenad Jaćimović[a] and Dušan Prodanović[a]

[a] Faculty of Civil Engineering, Department of Hydraulic and Environmental Engineering, Belgrade, Serbia, Bulevar Kralja Aleksandra 73

***Corresponding author**: Miloš Milašinović, e-mail address: mmilasinovic@grf.bg.ac.rs

E-mail adresses: mmilasinovic@grf.bg.ac.rs, arandjelovic@grf.bg.ac.rs, njacimovic@grf.bg.ac.rs, dprodanovic@grf.bg.ac.rs

Abstract: High urbanization puts many groundwater resources at risk of quality deterioration. Analysing all viable potential groundwater contamination scenarios for good decision making requires reliable tool. Coupling several complex models in integrated modelling can often fail to perform in reasonable time. Possible solution in that case could be usage of simplified models in order to speed up long-term continuous calculations and simulations. The paper presents the application of the Cellular automata (CA) approach in modelling of the contaminant transport in unsteady groundwater conditions. It compares the results obtained using coupled CA models with well-known analytical solutions and standard methods used for pollution transport modelling in groundwater conditions, such as coupled MODFLOW and MT3DMS. Results obtained in this paper show that CA approach can be satisfactorily used for simulations of unsteady groundwater conditions, caused by surface-groundwater interaction, and pollution transport, especially in diffusion dominant cases.

Keywords: Integrated modelling; Cellular Automata; contaminant transport, surface-groundwater interaction

NOTATIONS

Table 1. Groundwater flow and pollutant transport model parameters and variables

Groundwater flow model parameters and variables			Pollutant transport model parameters and variables		
ΔH_{ij}	[L]	Hydraulic gradient	m_i^t	[M]	current pollutant mass in central cell ***i***
H_i^t	[L]	current central cell ***i*** state (head)	m_j^t	[M]	current pollutant mass in adjacent cell ***j***
H_j^t	[L]	current adjacent cell ***i*** state (j=1,2,3,4)	C_i^t	[ML^{-3}]	current pollutant concentration in

ΔV_{ij}	[L^3]	potential intercellular volume that can leave central cell ***i*** to the adjacent cell ***j***
A_i	[L^2]	cell basis area
ΔV_{min}	[L^3]	minimum value of the potential intercellular volume
ΔV_{max}	[L^3]	maximum value of the potential intercellular volume
w_j	[-]	cell ***j*** weight in the process of intercellular volume exchange
v_M	[LT^{-1}]	max allowed velocity (*M* is the index of max weighted cell)
K_i	[LT^{-1}]	hydraulic conductivity of the central cell
K_M	[LT^{-1}]	hydraulic conductivity of an adjacent cell with max weight
H_M^t	[L]	water level (head) in the max weighted adjacent cell
d_{iM}	[L]	boundary length between two cells
ΔV_M	[L^3]	max allowed intercellular volume
Δt	[T]	time step
ΔV_{tot}	[L^3]	total intercellular volume
ΔV_j^{real}	[L^3]	real intercellular volume that leaves central cell ***i*** to the each adjacent cell ***j*** (***j***=1,2,3,4)
$H_i^{t+\Delta t}$	[L]	updated cell state (head)
S_y	[-]	specific yield of the cell
n	[-]	porosity
V_{out}	[L^3]	cellular volume that leaves the central cell ***i*** (e.g. pumping rate)
V_{inp}	[L^3]	volume that enters the central cell I (e.g. infiltration)
Q_{ij}	[L^3T^{-1}]	intercellular discharge
T_i	[L^2T^{-1}]	transmissivity of the central cell
T_j	[L^2T^{-1}]	transmissivity of the adjacent cell
Q_{out}	[L^3T^{-1}]	discharge that leaves the central cell ***i*** (e.g. pumping)
Q_{inp}	[L^3T^{-1}]	flow (discharge) that enters the central cell ***i*** (e.g. infiltration)

		central cell ***i***
C_j^t	[ML^{-3}]	current pollutant concentration in central cell ***j***
Δm_{ij}^{adv*}	[MLT^{-1}]	pollutant mass gradient in intercellular interaction by the advection process (can have all real values; positive value when mass leaves the central cell)
v_{ij}^t	[LT^{-1}]	Darcy's velocity at time *t*
Δs	[L]	spatial resolution
$\Delta m_{ij}^{adv,pot}$	[MLT^{-1}]	Indicator of potential pollutant mass that leaves the central cell ***i*** and enters the adjacent cell ***j***
Δm_{min}^{adv}	[M]	Minimum gradient of pollutant mass that leaves central cell *i* during one time step
w_j^{adv}	[-]	cellular weight of the adjacent cell *j* in pollutant mass delivering from central cell *i* through advection mechanism
$\Delta m_{iM}^{adv,pot}$	[MLT^{-1}]	Gradient of the potential pollutant mass delivered from central cell ***i*** to the max weighted cell
w_M^{adv}	[-]	Max cellular weight in advection process
Δm_{tot}^{adv}	[M]	total pollutant mass delivered from the central cell in one time step through advection process
Δm_{ij}^{adv}	[M]	amount of the contaminant (mass) delivered from central cell to each adjacent cell
D_j	[L^2T^{-1}]	hydrodynamic dispersion coefficient
α_l	[L]	dispersivity
D_{mol}	[L^2T^{-1}]	molecular diffusion (Bear 1972)
Δm_{ij}^{disp*}	[M]	pollutant mass gradient between central cell ***i*** and the adjacent cell ***j*** (it can all real values; positive value when mass leaves the central cell)
$\Delta m_{ij}^{disp,pot}$	[M]	gradient of the potential pollutant mass that leaves the central cell towards the adjacent cell only
Δm_{min}^{disp}	[M]	Minimum gradient of the pollutant mass delivered from central cell in dispersion process
w_j^{disp}	[-]	cellular weight of the adjacent cell ***j*** in pollutant mass delivering from central cell ***i*** through dispersion mechanism
$\Delta m_{iM}^{disp,pot}$	[M]	Gradient of the potential pollutant mass delivered from central cell ***i*** to the max weighted cell
w_M^{disp}	[-]	Max cellular weight in dispersion process
Δm_{tot}^{disp}	[M]	total pollutant mass delivered from the central cell in one time step through dispersion process
Δm_{ij}^{disp}	[M]	amount of the contaminant (mass) delivered from central cell to each adjacent cell by the dispersion process
$m_i^{t+\Delta t}$	[M]	pollutant mass in the central cell *i* at the next time
Δm_{inp}	[M]	external pollutant mass input to the central cell (e. g. from some other cells)
Δm_{out}	[M]	represents external output of the pollutant mass from the cell (e. g. pollutant mass extracted from the

			cell by pumping rate)
	$C_i^{t+\Delta t}$	$[ML^{-3}]$	pollutant concentration at next time

1. INTRODUCTION

Water quality in urban catchments is one of the crucial problems that need to be addressed to assure high quality of life in cities. Particularly vulnerable to quality deterioration is the groundwater, especially if the aquifer lies near or beneath the city (>50% global population uses groundwater as potable water, WWAP 2015). This type of catchments is not uncommon, and put additional pressure on decision makers to find a fine balance between city expansion and groundwater catchment protection (e.g. Dimkić et al. 2013; Petrović Pantić, Mandić, and Samolov 2016). Contaminants affecting groundwater resources belong to a wide group of biological, chemical, inorganic and organic pollutants. Sources of the groundwater pollution also cover a wide range of locations, such as on-site sanitations system (septic systems), effluent from wastewater treatment plants, gas and petrol filling stations, landfills which are defined as point sources and can be well identified on the field. Beside point sources there are, also, nonpoint sources that cover larger area than point sources, which makes them more difficult to identify.

Urban stormwater runoff is being categorized as one of the major pollution sources to receiving waters, especially for containing significant amount of sediments and heavy metals (Deletic and Orr 2005; Djukić et al. 2016; Duong and Lee 2011; Revitt et al. 2014). Although build-up and wash-off mechanisms of pollution in highly urbanized areas have been subjects of numerous studies (Barbé, Cruise, and Mo 1996; Deletic, Maksimovic, and Ivetic 1997), interaction between groundwater and sewer systems should be further investigated. Old combined sewer systems (and even inadequately constructed separate ones) may leak underground. Sewer leakage combined with absence of a wastewater treatment system (Belgrade being such an extreme example), puts groundwater resources at a high risk of contamination. To assess various "*what if*" scenarios, particularly when there are water quality monitoring points, it is of outmost importance to develop a usable integrated model that provides a sound basis for decision making.

Due to the complexity of the commonly used physically based models, computation cost occurs as one of the main problems, especially in long-term simulations. Combination of several numerical models in order to solve physically based models for interactions in a chain of models can, often, create an unsolvable problem. Hence, simplified models, such as the Cellular Automata (CA) (Wolfram 1998) based ones have been a topic for many researchers in the last two decades, especially since they can speed up the computation by exploiting the explicit nature of CA. Additionally, with the development of parallel computing, CA models became highly exploited methods in different areas (Bandini, Mauri, and Serra 2001), including the water cycle modelling.

CA found its application in different areas of research, such as simulation of wildfire spreads (Ghisu et al. 2015), lava motion during volcano eruptions (Vicari et al. 2007), urban growth and land-use change (Barredo and Kasanko 2003; Feng and Tong 2018).

Dottori and Todini (2010, 2011) developed flood inundation models (diffusion dominated problem) based on the CA approach. The authors employ the Manning's equation for computation of interaction/discharge between computational cells, and therefore solving a two-dimensional problem as four one-dimensional problems. Ghimire et al., 2013 analysed Manning's formula-based CA model proposed by Dottori & Todini, 2010 and found that exponential nature of Manning's equation creates a time and computational expensive problem. In order to reduce this problem, they proposed a CA-based methodology with cell ranking to determine the direction of flow between cells. It should be noted that cell ranking also creates a somewhat computational time expensive problem. Hence, Guidolin et al., 2016 developed weighted cellular automata model for rapid flood inundation analysis. Guidolin et al. model calculates weights between cells based on water level differences in order to calculate discharge between them. However, Guidolin's model lacks the ability to be used for inertial dominated problems.

Efficiency of the CA models combined with parallel computing are demonstrated in Gibson, Keedwell, & Savić, 2015. In addition to flood inundation modelling, CA approach has been successfully applied in surface runoff problems (Shao et al. 2015, Cai et al. 2014), and in rainfall simulation, water evaporation and groundwater flow in three-dimensional satellite images (Espínola et al. 2016). Some researchers improved surface runoff CA model outputs by setting it as a data-driven model (Li et al., 2015). Ravazzani, Rametta, & Mancini, 2011 used CA for transient groundwater state modelling caused by constant pumping rate from a well, showing good results. Transient groundwater state is modelled using a simple Darcy's law for intercellular discharge calculation, combined with a water balance equation for cell state updating.

In addition to water quantity, CA approach is used to model pollution (contaminant) and other transport processes in atmosphere, surface water and groundwater. However, CA based models for pollution transport have been mostly used in *steady state conditions*: air pollution (Guariso & Maniezzo, 1992, Marín et al., 2000, Lauret et al. 2016), river pollution transport (Rui et al. 2013), probabilistic two dimensional contaminant transport in groundwater (Palanichamy, Schüttrumpf, & Palani, 2008), etc.

Researchers who have used CA approach with contaminant transport modelling in the air (e.g. Guariso and Maniezzo 1992; Guariso, Maniezzo, and Salomoni 1996), or water (e.g. Gług and Wąs 2018; Lin and Yao 2018) have not coupled the transport part with airflow or hydrodynamic modelling, limiting the type of problems that can be solved using the proposed methodology. Additionaly, these types of CA algorithms mostly use linear, empirical, functions for intercellular interaction rules that introduce new transport parameters. These new transport parameters are sometimes not easily related to physical phenomena, and their estimation is therefore difficult. On the other hand, Lauret et al. (2016) used artificial neural networks (ANNs) approach to model air pollution transport. However, due to lack of physics, ANNs may encounter problems when the input data differs substantially from the training set, producing poor results.

If there is a requirement for fast evaluation of "what-if" scenarios in integrated models, such as the problem of pollution transport in both urban surface and groundwater catchment, than there is a necessity for development of a modelling strategy able to cope with: (1) physically based pollution transport, (2) hydrodynamic modelling and (3) unsteady boundary conditions.

This paper presents a further development of a weighted CA approach for pollutant transport that complies with all three previously stated requirements, firstly presented by Milasinovic et al. (2019). Further testing of the initial methodology showed its inability to satisfy pollutant mass conservation, therefore requiring some structural modifications, mainly in the cell state variables, along with additional physical constraints. Here presented is a research that analyses and compares the behaviour of the two CA-based hydrodynamic models coupled with CA-based transport model for contaminant transport by an advection-dispersion mechanism. The methodology is tested against steady analytical 1D and 2D solutions, and a linear pollutant source under unsteady groundwater conditions.

2. MATERIALS AND METHODS

2.1 Overview of the methodology

Cellular Automata based methodology used for groundwater flow/contaminant transport phenomena simulation is developed as a two-layer (two-stage) model (Figure 1). First layer (Layer 1) is used for hydrodynamic modelling of the unsteady groundwater conditions. Therefore, the following values are defined as cell's *static* parameters (constant during simulation period): porosity ***n*** ***(dimensionless)***, Darcy's coefficient of permeability ***K*** **(L**enght**/T**ime**)**, specific yield S_y ***(dimensionless)*** and spatial resolution (Δs). Cell's *dynamic* parameters are: water level (head) ***H*** **(L)** , intercellular discharge Q_j (L^3/T) (j=1,2,3,4) and intercellular Darcy's velocity V_j **(L/T)** (j=1,2,3,4). Von Neumann's approach is used for neighbourhood representation (Wolfram 1998). Intercellular interaction rules are based on the Darcy's law and a water balance equation.

Layer 1 uses two CA approaches: a modification of WCA2D, proposed by Guidolin et al., 2016, and a simple Darcy's law – method, MACCA-GW proposed by Ravazzani et al., 2011 (Section 2.3) Since WCA2D model is developed for urban flood modelling, it is modified here to include physical constraints suitable for groundwater conditions forming a new model named WCAGW (Section 2.2). MACCA-GW is used in its original format.

Layer 2 is used for modelling of pollution transport using the velocity-field and head-field from Layer 1 (hydrodynamic layer). The velocity – field and head - field, therefore, are the link between the two simulated physical processes, groundwater flow and mass transport. Layer 2 also has static and dynamic parameters, cell neighbourhood and set of intercellular interaction rules. The following set of values is defined as static parameters significant for pollutant transport simulation: porosity $\boldsymbol{n}$ (/), dispersivity $\boldsymbol{\alpha_L}$ (**L**), molecular diffusion $\boldsymbol{D_{mol}}$ ($\mathbf{L^2/T}$). Values defined as dynamic parameters are: pollutant concentration $\boldsymbol{C}$ (**M**ass/$\mathbf{L^3}$) averaged over the entire cell, pollutant mass in a cell $\boldsymbol{m}$ (**M**) and mass exchange rate between cells $\boldsymbol{\Delta m_j}$ (**M**). Cell neighbourhood in Layer 2 is also Von Neumann's type. Intercellular interaction rules in Layer 2 are originally developed in this research by using simplified advection-dispersion equation (Section 2.4).

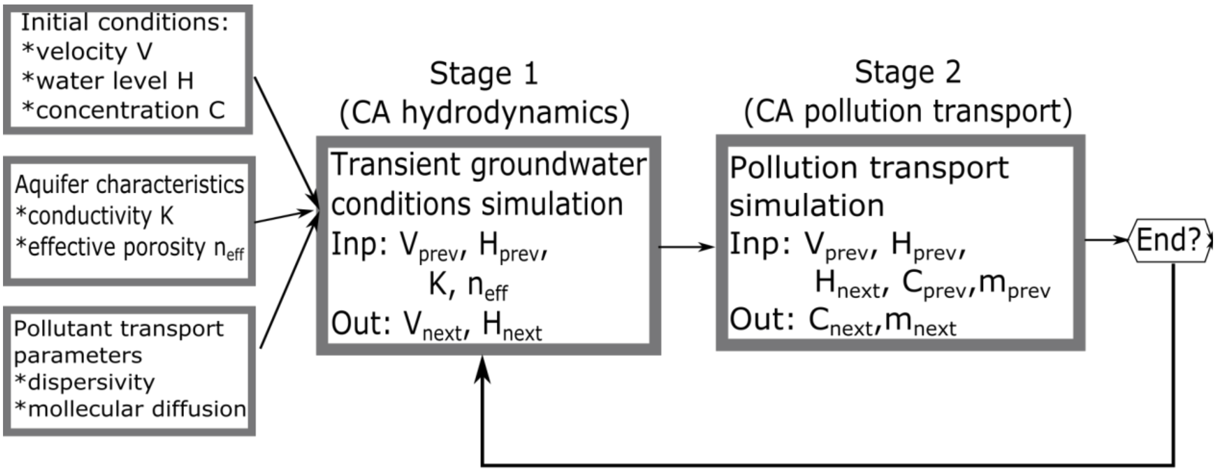


Figure 1. Schematic view of the CA two-layer model: Layer 1 is computed either by WCAGW or by MACCA-GW and Layer 2 by CAPT model

2.2 Weighted Cellular Automata for Unsteady Groundwater Flow Modelling – WCAGW

Weighted Cellular Automata (WCAGW) approach implemented in this paper (fig. 2a), for Layer 1 simulations, uses the same principle of cellular weights calculation as the original WCA2D method (Guidolin et al., 2016). However, differences exist in physical limitations that are used for maximum intercellular velocity (eq. 6), and the update step for cell state (water level) (eq. 10), both accounting for intrinsic nature of the porous media.

Generally, process of the hydrodynamic simulations using Cellular Automata can be described with the following operations which are applied to each cell representing the domain:

a. Obtain current cells state value, both the central cell (cell being currently considered) and adjacent cells (cells representing the neighbourhood)
b. Calculate the cellular weights – eqs. 1, 2, 3, 4 and 5 from table 2
c. Calculate the intercellular volume according to the cellular weights – eqs. 6, 7, 8 and 9 from table 2
d. Update the cell state; calculate next cell state – eq. 10 from table 2
e. If the simulation time reaches the limit, stop simulation, otherwise repeat steps a., b., c. and d.

For better display all equation related to Cellular Automata algorithms are grouped in appropriate tables.

Table 2. Groundwater flow equations - WCAGW hydrodynamic model

Groundwater flow equation	Eq. No.
Hydraulic gradient	
$\Delta H_{ij} = H_i^t - H_j^t$	(1)
Potential intercellular water volume exchange	
$\Delta V_{ij} = A_i \cdot \max\left(0, \Delta H_{ij}\right)$	(2)
$\Delta V_{\min} = \min(\Delta V_{ij})$	(3)
$\Delta V_{\max} = \max(\Delta V_{ij})$	(4)
Weightening the cells in the water volume exchange process	
$w_j = \frac{\Delta V_{ij}}{\sum_{j=1}^{4} \Delta V_{ij} + \Delta V_{\min}}$	(5)
Embeding the physicall limitation in water flow equation	
$v_M = \frac{2 \cdot K_i \cdot H_i^t \cdot K_M \cdot H_M^t}{\left(K_i \cdot H_i^t + K_M \cdot H_M^t\right) \cdot \frac{\left(H_i^t + H_M^t\right)}{2} \cdot d_{iM}} \left(H_i^t - H_M^t\right)$	(6)
$\Delta V_M = v_M \cdot H_i^t \cdot d_{iM} \cdot \Delta t$	(7)
Total water volume transfer	
$\Delta V_{tot} = \min\left(\Delta V_M / w_M ; H_i^t \cdot A_i \cdot n\right)$	(8)
Real intercellular water volume transfer	
$\Delta V_j^{real} = w_j \cdot \Delta V_{tot}$	(9)
Cell state (head) update	

$$H_i^{t+\Delta t} = H_i^t - \frac{1}{S_y} \cdot \frac{\sum_{j=1}^{4} \Delta V_j^{real} + V_{out} - V_{inp}}{A_i}$$	(10)

2.3 Darcy’s law-based Cellular Automata for Unsteady Groundwater Flow Modelling – MACCA-GW

Darcy’s law-based Cellular Automata for Unsteady Groundwater Modelling (Ravazzani et al. 2011), originally named as **MAC**roscopic **C**ellular **A**utomata for **G**round**W**ater modelling (MACCA - GW) (fig. 2b), is developed for simulation of unsteady groundwater conditions by using same dynamic and static parameters defined in section 2.1 for Layer 1. Similarly with WCAGW, the two dimensional problem is solved as a set of four one-dimensional problems, using a simplified Darcy’s law to calculate intercellular discharges. MACCA-GW model implementation has one step less than process described with WCAGW model. The following set of operations shows general algorithm that’s being applied to each cell of the domain:

a. Obtain current cells state value, both the central cell (cell being currently considered) and adjacent cells (cells representing the neighbourhood)
b. Calculate the rate of intercellular interaction – eqs. 11 and 12 from table 3
c. Update each cell state – eq. 13 from table 3
d. If the simulation time reaches the limit stop the simulation, otherwise repeat steps a., b. and c.

Table 3. Groundwater flow equations – MACCA-GW hydrodynamic model

Intercellular discharge	Eq.no.
$$Q_{ij} = \frac{2 \cdot K_i \cdot H_i^t \cdot K_j \cdot H_j^t}{K_i \cdot H_i^t + K_j \cdot H_j^t}\left(H_j^t - H_i^t\right)$$	(11)
$$Q_{ij} = \frac{2 \cdot T_i \cdot T_j}{T_i + T_j}\left(H_j^t - H_i^t\right)$$	(12)
Cell state (head) update	
$$H_i^{t+\Delta t} = H_i^t - \frac{1}{S_y} \cdot \frac{\sum_{j=1}^{4} Q_{ij} + Q_{out} - Q_{inp}}{A_i} \cdot \Delta t$$	(13)

2.4 Weighted Cellular Automata for Pollution Transport Modelling – CAPT

Initial methodology (Milasinovic et al. 2019) used pollutant concentration as a primary cell state variable, with intercellular interaction rules implemented by four 1D transport equations. Further testing of the initial methodology encountered problems with pollutant mass conservation. Hence, the methodology is improved by using mass of the pollutant as a primary cell state variable (pollutant concentration can be derived from pollutant mass). Additionally, weighted CA approach is used to keep the mass conserved. Using this approach, CA algorithm transfers only the available pollutant mass between adjacent cells, following the appropriate physical constraints. This combination of CA weighting algorithm with embedded physical constraints based on simplified transport equation (instead of empirical relations) accounts for the nature of the actual transport mechanisms.

Pollution transport model (CAPT) (fig. 2c) uses velocity and head fields calculated either by WCAGW or MACCA-GW models as input. Contaminant transport is represented by two transport mechanisms, advection and dispersion, with reasonable simplifications. Therefore, whole algorithm is divided in two parts, one for the advection mechanism and the other for the dispersion mechanism. Pseudo code which describes this methodology can be written as:

a. Obtain cell state values, pollutant concentration ***C*** and pollutant mass ***m*** – eqs. 14 and 15
b. Calculate intercellular pollutant mass exchange by advection
    i. Calculate cellular weights in advection process based on the velocity and head field and pollutant mass gradient – eqs. 16, 17, 18 and 19 from table 4a
    ii. Calculate advection forced intercellular pollutant mass exchange using the obtained cellular weights – eqs. 20, 21 and 22 from table 4a
c. Calculate intercellular pollutant mass exchange by dispersion
    i. Calculate cellular weights in dispersion process – eqs. 23, 24, 25, 26, 27 from table 4b
    ii. Calculate dispersion forced intercellular pollutant mass exchange using the obtained cellular weights – eqs. 28, 29 and 30 from table 4b

d. Calculate total intercellular mass exchange, advection + dispersion

e. Update cell state, pollutant mass and concentration in cell – eqs. 31 and 32 from table 4b

f. If the simulation time reaches the limit stop the simulation, otherwise repeat steps a., b., c., d. and e.

Table 4a. Contaminant transport equations – CAPT transport model (advection transport)

Cell pollutant mass	Eq. no.
$m_i^t = C_i^t \cdot H_i^t \cdot A_i \cdot n_i$	(14)
$m_j^t = C_j^t \cdot H_j^t \cdot A_j \cdot n_j$	(15)
Advection process Pollutant mass gradient in advection mechanism	
$\Delta m_{ij}^{adv*} = \frac{v_{ij}^t}{n} \cdot \left\| m_i^t - m_j^t \right\|$	(16)
$\Delta m_{ij}^{adv,pot} = \max\left(0; \Delta m_{ij}^{adv,*}\right)$	(17)
$\Delta m_{\min}^{adv} = \min\left(\Delta m_{ij}^{adv,pot}\right)$	(18)
Weightening the cells in mass transfer through advection process	
$w_j^{adv} = \frac{\Delta m_{ij}^{adv,pot}}{\sum_{j=1}^{4} \Delta m_{ij}^{adv,pot} + \Delta m_{\min}^{adv}}$	(19)
Embeding the physical limitation in pollutant mass transfer – advection mechanism	
$\Delta m_{iM}^{adv,pot} = \frac{v_{iM}^t}{n} \cdot \frac{\left\| m_i^t - m_M^t \right\|}{\Delta s} \cdot \Delta t$	(20)
Total pollutant mass transfer – advection mechanism	
$\Delta m_{tot}^{adv} = \min\left(\frac{m_i^t}{2}; \frac{\Delta m_{iM}^{adv,pot}}{w_M^{adv}}\right)$	(21)
Real intercellular pollutant mass transfer – advection mechanism	
$\Delta m_{ij}^{adv} = w_j^{adv} \cdot \Delta m_{tot}^{adv}$	(22)

Table 4b. Contaminant transport equations – CAPT transport model (dispersion transport + cellular pollutant mass update)

Dispersion process Pollutant mass gradient in dispersion mechanism	Eq. no.
$D_j = \alpha_l \cdot \frac{\left\| v_{ij}^t \right\|}{n} + D_{mol}$	(23)
$\Delta m_{ij}^{disp*} = D_j \cdot \left(m_i^t - m_j^t\right)$	(24)

$$\Delta m_{ij}^{disp,pot} = \max\left(0; \Delta m_{ij}^{disp*}\right) \tag{25}$$

$$\Delta m_{\min}^{disp} = \min\left(\Delta m_{ij}^{disp,pot}\right) \tag{26}$$

Weightening the cells in mass transfer through dispersion process

$$w_j^{disp} = \frac{\Delta m_{ij}^{disp,pot}}{\sum_{j=1}^{4} \Delta m_{ij}^{disp,pot} + \Delta m_{\min}^{disp}} \tag{27}$$

Embeding the physical limitation in pollutant mass transfer – dispersion mechanism

$$\Delta m_{iM}^{disp,pot} = D_M \cdot \frac{m_i^t - m_M^t}{\Delta s^2} \cdot \Delta t \tag{28}$$

Total pollutant mass transfer – dispersion mechanism

$$\Delta m_{tot}^{disp} = \min\left(\frac{m_i^t}{2}; \frac{\Delta m_{iM}^{disp,pot}}{w_M^{disp}}\right) \tag{29}$$

Real intercellular pollutant mass transfer – dispersion mechanism

$$\Delta m_{ij}^{disp} = w_j^{disp} \cdot \Delta m_{tot}^{disp} \tag{30}$$

Cell state (pollutant mass) update – advection + dispersion

$$m_i^{t+\Delta t} = m_i^t - \sum_{j=1}^{4}\left(\Delta m_{ij}^{adv} + \Delta m_{ij}^{disp}\right) + \Delta m_{inp} + \Delta m_{out} \tag{31}$$

$$C_i^{t+\Delta t} = \frac{m_i^{t+\Delta t}}{H_i^{t+\Delta t} \cdot A_i \cdot n} \tag{32}$$

Considering the explicit nature of MACCA-GW, WCAGW and CAPT models, time step limit has to be determined in order to provide numerical stability. Time step limit can be determined by using the eq. 33. derived by unifying different time step limitations (Ravazzani et al. 2011; Zheng and Wang 1999).

$$\Delta t \le \min\left(\frac{\Delta s^2 \cdot S_y}{4 \cdot T}; \frac{\Delta s}{v_{ij}^{\max}}; \frac{\Delta s^2}{2D_j}; \frac{2D_j}{\left(v_{ij}^{\max}\right)^2}\right) \tag{33}$$

Where Δs is spatial resolution in meters, S_y specific yield of the cell, T transmisivity of the cell, v_{ij}^{max} max real intercellular velocity, D_j hydrodynamic dispersion and Δt time step.

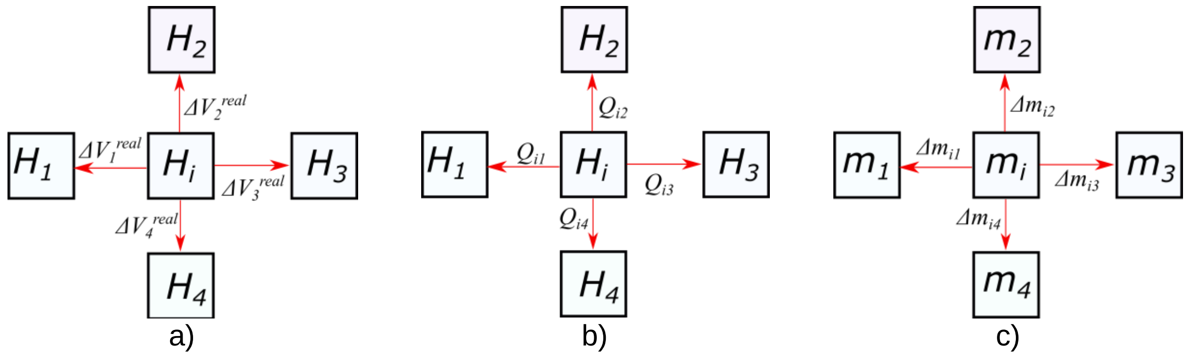


Figure 2. CA Von Neumann's neighbourhood and cell state variables and variables representing intercellular interaction for: a) WCAGW; b) MACCA-GW; c) CAPT

2.5 ANALYTICAL SOLUTIONS

The CA based pollution transport model, CAPT, is tested with two analytical solutions to transport problems: (1) two-dimensional analytical solution for point source contaminants transport in a semi-infinite homogeneous porous medium (Bear 1972) (eq 34) and (2) one-dimensional transport equation with continuous constant pollution source (Ogata and Banks 1961) (eq. 35).

$$C(x,y,t)=\frac{M/h}{4\pi t\sqrt{D_x\cdot D_y}}\exp\left(-\frac{(x-x_0-v_x\cdot t)^2}{4D_x\cdot t}-\frac{(y-y_0-v_y\cdot t)^2}{4D_y\cdot t}\right) \quad (34)$$

$$C(x,t)=\frac{C_0}{2}\cdot\left[erfc\left(\frac{x-x_0-v_x\cdot t}{2\sqrt{D_x\cdot t}}\right)+e^{\frac{v_x\cdot(x-x_0)}{D_x}}\cdot erfc\left(\frac{x-x_0+v_x\cdot t}{2\sqrt{D_x\cdot t}}\right)\right] \quad (35)$$

Eq. 34 shows contaminant concentration at the point with coordinates (*x,y*) at time *t*, when mass over depth *M/h* of the contaminant was injected instantly at $t=t_0$ at the point (x_0,y_0). This analytical solution considers steady hydrodynamic conditions, represented by two velocity components v_x and v_y in directions *x* and *y*, respectively. D_x and D_y are hydrodynamic dispersion coefficients in *x* and *y* direction, respectively.

Eq. 35 shoes contaminant concentration in 1D problem, at the point with coordinates (*x,y*) at time *t*, when constant pollution concentration C_0 is given at the pollution source with coordinates (x_0,y_0).

2.6 MODFLOW & MT3DMS

Additionally, WCAGW/CAPT and MACCA-GW/CAPT models are compared to MODFLOW/MT3DMS modules incorporated in ModelMuse 3.9.0.0 (Winston 2009). MODFLOW (Langevin et al. 2017) is a widely used code for simulating groundwater flow using control-volume finite difference method (Panday et al. 2013) to solve a three dimensional groundwater equation. MT3DMS (Zheng and Wang 1999) is a three – dimensional transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems. Depending on the problem solved, it can use multiple numerical solvers such as the standard finite-difference method and the third-order TVD method (Niswonger, Panday, and Ibaraki 2011; Panday et al. 2013).

2.7 Numerical test cases – problem description

Numerical test case 1, POINT-1D, represents a comparison between 1D CA solution and 1D analytical solution (Ogata and Banks 1961) for continuous, constant source. This test is conducted on a domain with the following characteristics: length L=500 m, spatial resolution Δx=10 m, temporal resolution Δt=4 h, seepage velocity in x direction $v=10^{-6}$ m/s, longitudinal dispersivity α_L=1000 m and solute concentration at the source $C(0,t)=C_0$=5mg/l. Water levels doesn't have any significance in 1D analytical solution (eq. 35), but uniform head field (15 m) is applied in order to calculate pollutant mass which is necessary in CAPT model.

Numerical test case 2, POINT-2D, compares 2D CA model and 2D analytical solution in eq. 36 (Bear 1972). The domain has the following characteristics: dimensions 500x500 m, with spatial resolution set at *10x10* m, time step Δt=4 h, hydrodynamic dispersion in longitudinal direction $D_x=3*10^{-4}$ m^2/s, hydrodynamic dispersion in transversal direction $D_y=3*10^{-5}$ m^2/s and without seepage velocity, so that dispersion is the only transport mechanism considered. The plume point source is presented with an initial pollutant mass released into the aquifer. The amount of the pollutant mass is $M_0=7.5*10^6$ mg, and it is released over the 15 m of water depth, h=15 m. Thus, the initial pollutant concentration at the source equals to C_0=50 mg/l. Same as 1D analytical solution, 2D solution (eq. 34) doesn't require head field. Uniform (15 m) head field is applied in this situation in order to calculate pollutant mass in CAPT model.

Dispersion parameters are selected in order to prevent pollution from reaching the boundaries of the domain. In that case semi-infinite domain assumption used for analytical solution deriving can be partially applied in CA model.

Numerical test case 3, LINEAR, is a linear pollution source in transient conditions caused by a constant infiltration rate (fig. 3). It is set in a hypothetical square domain which represents an unconfined aquifer. Boundary conditions include constant head H=15 m on western and northern boundaries, zero flow and zero mass flux on eastern and southern boundaries and pollution concentration 0 *mg/l* on northern and western boundaries. The domain is assumed to be homogeneous with the following set of static parameters of the CA hydrodynamic layer (Layer 1 in Fig.1): hydraulic conductivity $K=1.25*10^{-5}$ m/s and effective porosity $n_{eff} = 0.26$. Specific yield S_y is assumed equal to the effective porosity. Static parameters of the CA pollution transport layer (Layer 2 in Fig.1) have the following values: porosity $n=n_{eff}=0.26$, and transversal dispersivity $\alpha_T= 0$ m. Since longitudinal dispersivity takes a wide range of values, depending on the scale of the problem in question eg. 0.01 – 14 970 m (Schulze-Makuch 2005), a range of longitudinal dispersivities is tested: $\alpha_L=\{10, 100, 1000, 5000\}$ m, to assess its impact on the CA based transport model. Molecular diffusion coefficient is tested over a range of values: $D_{mol}=\{10^{-4},10^{-5},10^{-7}\}$ m^2/s. The domain is discretized in a *10x10* m cellular grid. Pollution is set as a linear source pollution in order to represent potential groundwater quality deterioration when sewage system leakage is present. Transient groundwater conditions are forced by a constant infiltration rate of i=0.1 mm/h spread homogeneously over the entire domain, during 1000 time steps using time steps Δt=4 h and Δt=6 h. The initial conditions include a constant head of H=15 m with no flow and a concentration of C=0 mg/l over the entire domain. Test case specifications are given in Table 5.

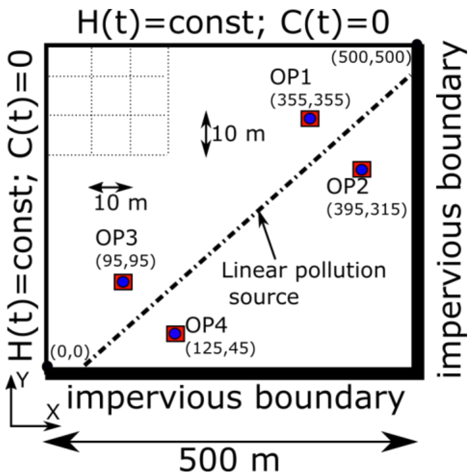


Figure 3. Schematic overview of the numerical case 1. Blue dots represent water level observation point (OP) and red squares represent pollutant concentration observation points

Table 5. Test cases specifications

Case	α_L [m]	D_{mol} [m^2 / s]	Δt [h]	Duration of pollution injection at the source
1	10	10^{-7}	6	1000 Δt
2	100	10^{-7}	6	1000 Δt
3	1000	10^{-7}	4	1000 Δt
4	5000	10^{-7}	4	1000 Δt
5	1000	10^{-4}	4	1000 Δt
6	1000	10^{-5}	4	1000 Δt
7	10	10^{-7}	6	1 Δt
8	100	10^{-7}	6	1 Δt
9	1000	10^{-7}	4	1 Δt
10	5000	10^{-7}	4	1 Δt
11	1000	10^{-4}	4	1 Δt
12	1000	10^{-5}	4	1 Δt

First six test cases (cases 1 - 6) from table 5 are conducted in order to analyse transport parameters impact on CA-based transport model in scenarios when pollution source is given as a constant pollution concentration during the simulation. In other words, it is assumed that pollutant is constantly loaded from the source in order to maintain constant concentration. Other six cases (7 - 12) are conducted in order to represent scenarios when pollutant is injected into the aquifer from the source over one time step (e.g. potential hazard situations caused by sewage system or oil pipeline failure).

2.8 Model assessment methodology

For the assessment of the CA transport modelling, three statistical indicators are calculated: (1) coefficient of determination R^2 (eqs. 36 and 39), (2) root mean square error *RMSE* (eqs.37 and 40) and (3) normalised root mean square error *NRMSE* (eqs. 38 and 41). These indicators are calculated as both spatial and temporal measures, to assess agreement between CA based and MODFLOW/MT3DMS model results for both water levels and pollution concentrations.

Table 6. Statistical parameters used for consistency assessment

Statistical paramaters for spatial consistency at the end of the simulation	
$R^2_{spatial} = \left(\frac{\sum_{i=1}^{N_{cell}} \left(X_{CA,i} - \overline{X_{CA,i}}\right)\left(X_{MF,i} - \overline{X_{MF,i}}\right)}{\sqrt{\sum_{i=1}^{N_{cell}} \left(X_{CA,i} - \overline{X_{CA,i}}\right)^2 \sum_{i=1}^{N_{cell}} \left(X_{MF,i} - \overline{X_{MF,i}}\right)^2}} \right)^2$	(36)

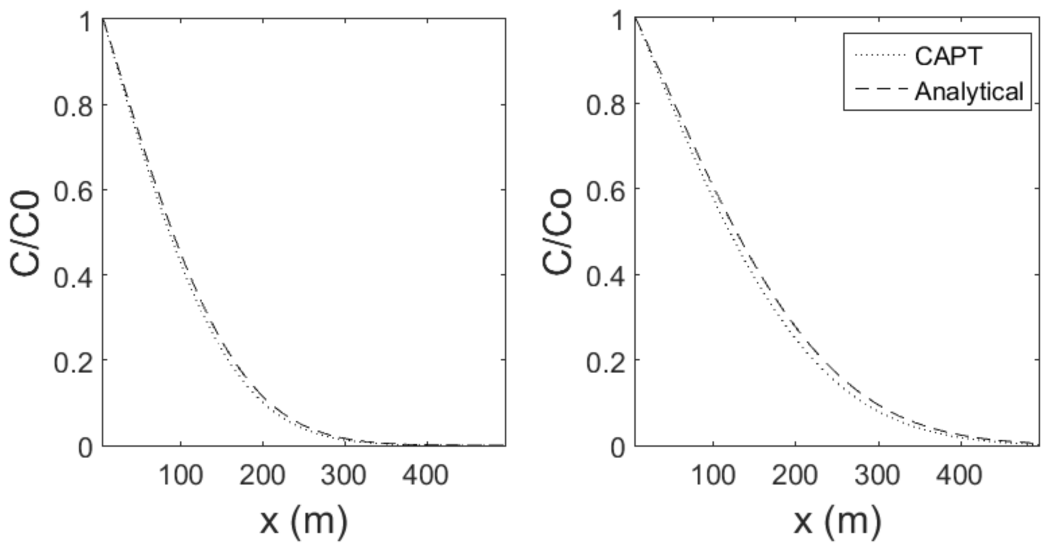
$RMSE_{spatial}=\sqrt{\frac{\sum_{i=1}^{N_{cell}}\left(X_{CA,i}-X_{MF,i}\right)^2}{N_{cell}}}$	(37)
$NRMSE_{spatial}=\frac{1}{X_0}\sqrt{\frac{\sum_{i=1}^{N_{cell}}\left(X_{CA,i}-X_{MF,i}\right)^2}{N_{cell}}}\cdot 100$	(38)
Statistical parameters for time series (temporal) consistency at observation points	
$R_{temporal}^2=\left(\frac{\sum_{i=1}^{N_{\Delta t}}\left(X_{CA,i}-\overline{X_{CA,i}}\right)\left(X_{MF,i}-\overline{X_{MF,i}}\right)}{\sqrt{\sum_{i=1}^{N_{\Delta t}}\left(X_{CA,i}-\overline{X_{CA,i}}\right)^2\sum_{i=1}^{N_{\Delta t}}\left(X_{MF,i}-\overline{X_{MF,i}}\right)^2}}\right)^2$	(39)
$RMSE_{temporal}=\sqrt{\frac{\sum_{i=1}^{N_{\Delta t}}\left(X_{CA,i}-X_{MF,i}\right)^2}{N_{\Delta t}}}$	(40)
$NRMSE_{temporal}=\frac{1}{X_0}\sqrt{\frac{\sum_{i=1}^{N_{\Delta t}}\left(X_{CA,i}-X_{MF,i}\right)^2}{N_{\Delta t}}}\cdot 100$	(41)
Total pollutant mass in the aquifer at specific time t	
$TotalMass^t=\sum_{i=1}^{N_{CA}}C_i^t\cdot H_i^t\cdot n\cdot A_i=\sum_{i=1}^{N_{CA}}m_i^t$	(42)
Normalised discrepancy in pollution concentration spatial distribution	
$NDD_j^t=\frac{\left\lvert X_{CA,j}^t-X_{analytical,j}^t\right\rvert}{X_0}\cdot 100\quad\left(j=1,\ldots,N_{CA}\right)$	(43)

In table 6. $X_{CA,i}$ represents head H or pollutant concentration C obtained by CA-based models, $X_{MF,i}$ represents head H or pollutant concentration C obtained by MODFLOW/MT3DMS model. N_{CA} is the number of cells used for domain spatial discretization and $N_{\Delta t}$ is number of time steps. X_0 in eq. 41 represent initial value of the water level H_0 or contaminant concentration C_0. Eqs. 39, 40 and 41 are also used for total pollutant mass discrepancy assessment. On that case $X_{CA,i}$ represents total pollutant mass in the aquifer at specific time obtained by CA-based models and $X_{MF,i}$ represents total pollutant mass in the aquifer at specific time obtained by MODFLOW/MT3DMS model. Total pollutant mass in the aquifer at specific time is calculated by eq. 42, which uses cellular pollutant mass obtained by eq. 14. Eq. 43 is used for model consistency assessment when CAPT transport model is compared to analytical solution. NDD_j^t represent normalised discrepancy between pollutant concentration (or pollutant mass) in cell indexed with j, at time t, and pollutant concentration (pollutant mass) at the point with same coordinates as center of the cell j, at the same time.

3. RESULTS and DISCUSSION

3.1 Test Cases 1 and 2: Comparison with 1D and 2D analytical solution for point source contaminants transport

When CA-based pollution transport model CAPT is compared with one-dimensional and two-dimensional analytical solution, relatively good consistency is achieved, especially in one-dimensional case where total mass discrepancy (eq. 38) is less than 1%, 0.7% to be more precise. Fig. 4 shows a comparison of the pollutant transport model results across the domain at simulation half-time (fig. 4a) and at the end of the simulation (fig. 4b). When two-dimensional problem is considered, figure 5 shows pollutant concentration spatial distribution at the end of the simulation for CAPT model (fig. 5a), analytical solution (fig. 5b) and normalised discrepancy between these two solutions calculated by eq. 43 (fig. 5c). Total mass discrepancy (eqs. 42 and 43) between CAPT model and analytical solution is bigger, with 1% approximately, but this discrepancy can be caused by inability to create adequate test case for Cellular Automata modelling which will consider all of the assumptions used when analytical solution was derived.



a) b)

Figure 4. Comparison of the 1D transport results: a) After 500 time steps; b) After 1000 time steps

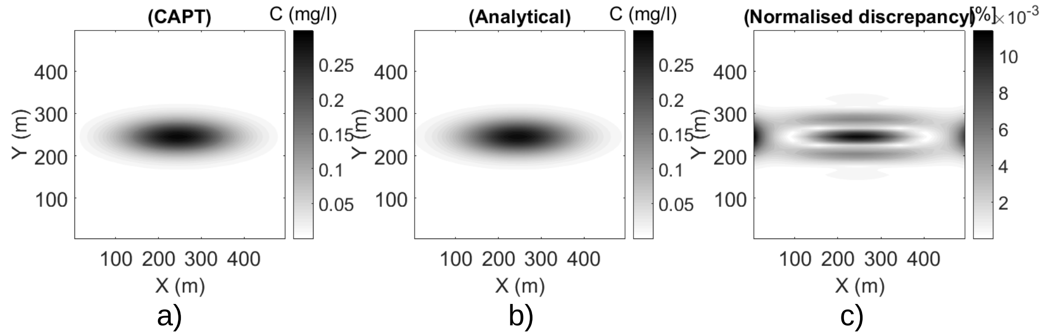
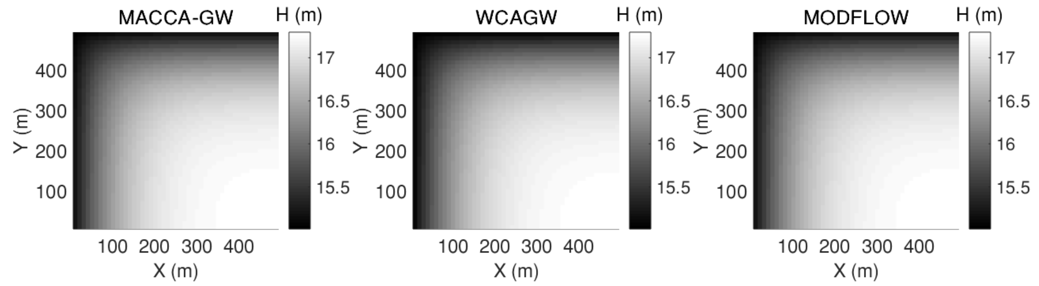


Figure 5. Comparison of the results for 2D test; spatial distribution of the pollution concentration at the end of the simulation (1000 time steps): a) Cellular Automata; b) Analytical solution; c) Normalised discrepancy distribution (Discrepancy is normalised to initial value of the pollutant concentration at the source by eq. 43)

3.2 Test case 3: Linear pollution source in transient groundwater conditions caused by constant infiltration rate

Water quantity model results for MACCA-GW, WCAGW, and MODFLOW are presented in Figure 6 as water level spatial distribution over the entire domain (end of simulation) and water level time series at four observation points. Location of these observation points can be seen in fig. 3. Figure 6 shows visually good agreement between results of all three models, with WCAGW somewhat struggling at Observation point 2 (*RMSE* = 0.037 m). It is hypothesized, that these slight disagreements in the time series are caused by weighting process in WCAGW model (eq. 5). This process in WCAGW model always leaves small amount of water in the cell (ΔV_{min} in eqs. 3 and 5) in order to prevent potential numerical instabilities. This was suggested by the original WCA2D model authors (Guidolin et al. 2016). MACCA-GW doesn't have that limitation. This could be a reason for slight differences in intercellular water volume exchanged by MACCA-GW and WCAGW models. Hydrodynamic conditions shown in fig. 5 are the same for all transport scenarios described in table 5.



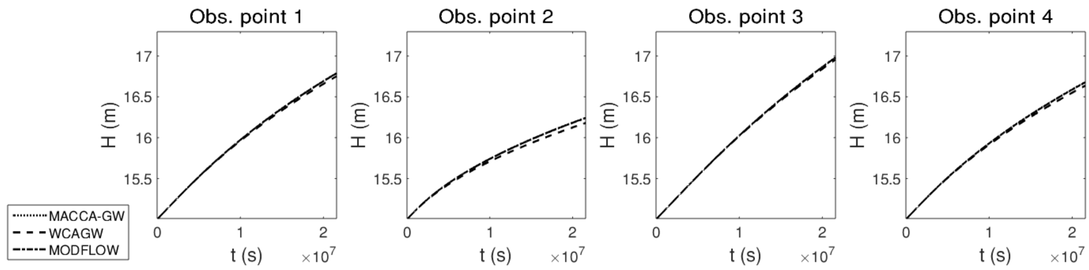


Figure 6. Water level H spatial distribution (X,Y) over the domain at the end of the simulation (top row) and water level time series at four observation points (bottom row) for MACCA-GW, WCAGW, MODFLOW models

Spatial distributions of pollutant concentration for cases 3, 4 and 5, are presented in the fig. 7. Fig. 8 shows spatial distribution of the pollutant concentration for cases 9, 10 and 11. Both figures show visually good agreements between MACCA-GW/CAPT, WCAGW/CAPT and MODFLOW/MT3DMS model results. These good agreements are quantified by RMSE and NRMSE values shown in table 9. For scenarios with constant pollutant concentration at the source during the simulation (cases 1-6), RMSE values for spatial distribution vary from 0.03 (case 3) mg to 0.17 mg (case 1) and NRMSE from 0.6 to 3.39 % when MACCA-GW/CAPT model is applied. When WCAGW/CAPT coupled model is applied, these values vary in the range 0.07 – 0.17 mg for RMSE and in the range 1.49 – 3.31 % for NRMSE. For the scenarios when pollution is injected into the aquifer over one time step (cases 7-12) RMSE values vary in the following ranges: 0.004 mg (case 12) to 0.09 mg (case 7) when MACCA-GW/CAPT model is applied, 0.004 mg (case 10) to 0.09 mg (case 7) when WCAGW/CAPT model is applied. Generally, discrepancy between MACCA-GW/CAPT and WCAGW/CAPT models results from seepage velocity differences. This difference is propagated from hydrodynamic modelling results which are then used as the input for transport modelling. In some cases (1 and 7) statistical parameters show better consistency between WCAGW/CAPT and MODFLOW/MT3DMS models than consistency between MACCA-GW/CAPT and MODFLOW/MT3DMS. It is assumed that this is the result of the longitudinal dispersivity and molecular diffusion relatively low values, 10 m and 10^{-7} m^2/s, respectively. When these low values of transport parameters are used, pollutant mass that is transferred by some of the defined transport mechanisms is also low, which can lead to the better consistency between WCAGW/CAPT and MODFLOW/MT3DMS models.

Table 7.	Statistical indicators of consistency between CA-based models and MODFLOW– comparison of the water level time series in observation points				

	R^2 [-]		RMSE [m]		NRMSE [%]	
	MACCA-GW	WCAGW	MACCA-GW	WCAGW	MACCA-GW	WCAGW
Obs. Point 1	0.99	0.99	0.0015	0.019	0.01	0.125
Obs. Point 2	0.99	0.99	0.0011	0.037	0.008	0.244
Obs. Point 3	0.99	0.99	0.0015	0.012	0.01	0.080
Obs. Point 4	0.99	0.99	0.0015	0.022	0.01	0.149

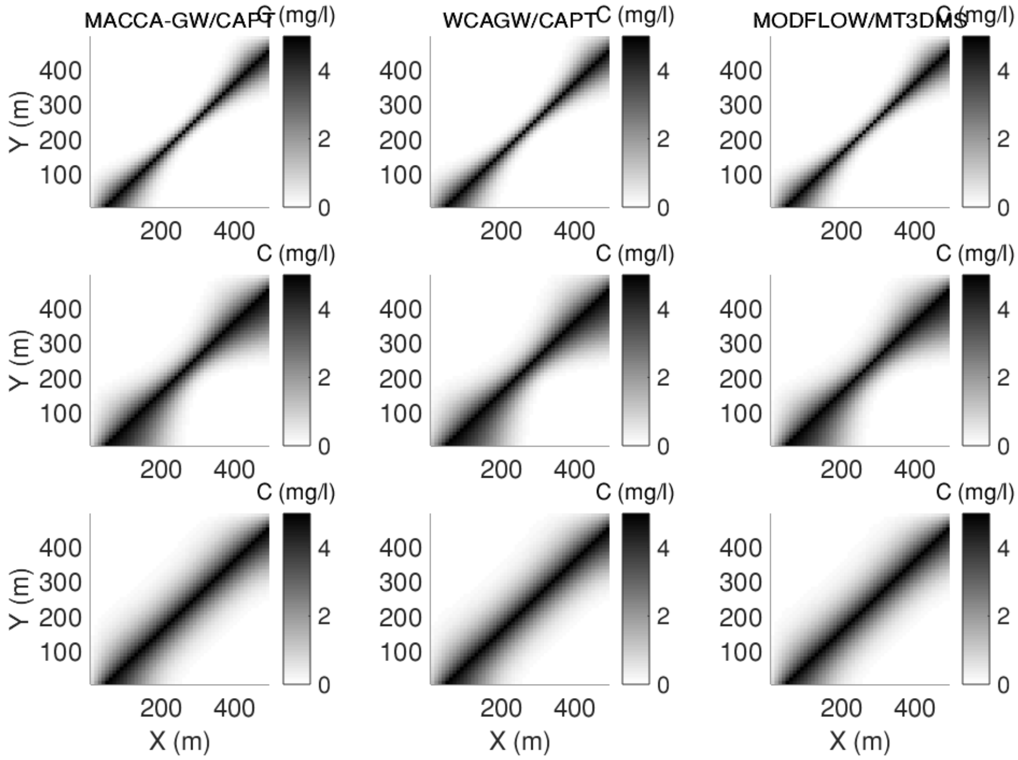


Figure 7. Spatial distribution (X,Y) of the pollutant concentration at the end of the simulation for case 3 (top row), case 4 (middle row) and case 5 (bottom row) for MACCA-GW/CAPT, WCAGW/CAPT and MODFLOW/MT3DMS– constant pollutant concentration at the source

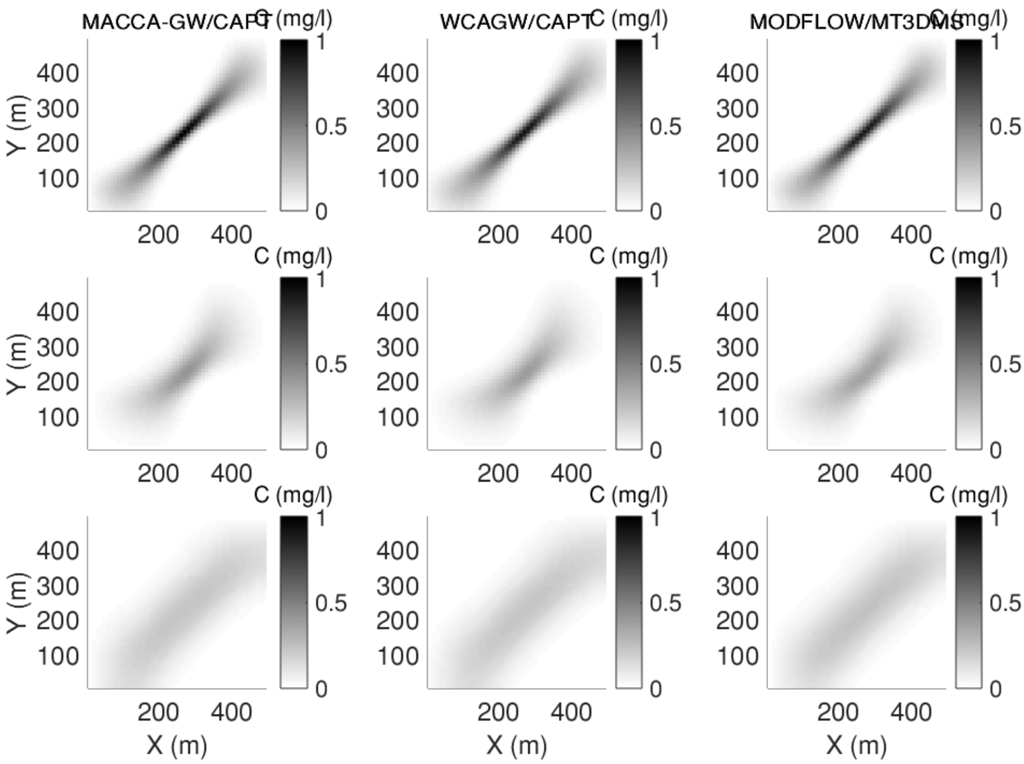


Figure 8. Spatial distribution of the pollutant concentration at the end of the simulation for case 9 (top row), case 10 (middle row) and case 11 (bottom row) – pollutant instantaneously injected at the source

Figs. 9 and 10 represent pollutant concentration time series in four observation points. Fig. 9 represent pollutant concentration time series for cases 3, 4 and 5, when pollution source is given as constant pollution concentration during the simulation. Fig. 10 represents pollutant concentration time series for cases 9, 10 and 11, when pollution is injected into the aquifer over one time step. Pollutant concentration time series at observation points (figs. 9 and 10) show, visually, good agreement between all applied models, except in observation point 4. In obs. point 4 CA-based models struggle to achieve max concentration values obtained by MODFLOW/MT3DMS model in cases 9, 10 and 11. It is hypothesized that this is caused by the proximity (see OP4 coordinates in fig. 3) of the impervious boundary condition to the obs. point 4. It is assumed that slight differences in boundary conditions setting between CA-based models and MODFLOW/MTDMS can cause slight disagreements in pollutant concentration time series. CA based transport model uses zero pollutant mass gradient to implement impervious boundary, while ModelMuse 3.9.0.0 doesn't have that option (no mass flux boundary is set by using zero flow boundary by default, without further explanation).

Statistical indicators of model results agreement for time series for four observation points are given in table 8. Considering *RMSE* and *NRMSE* as more suitable parameters than R^2 for model consistency assessment, it can be seen that these values are mostly well below 10 mg (2.5 % for NRMSE), except for obs. point 4, where highest *RMSE*/*NRMSE* is 0.3mg/5.97% (MACCA-GW/CAPT) and 0.27mg/5.42% (WCAGW/CAPT) in case 5.

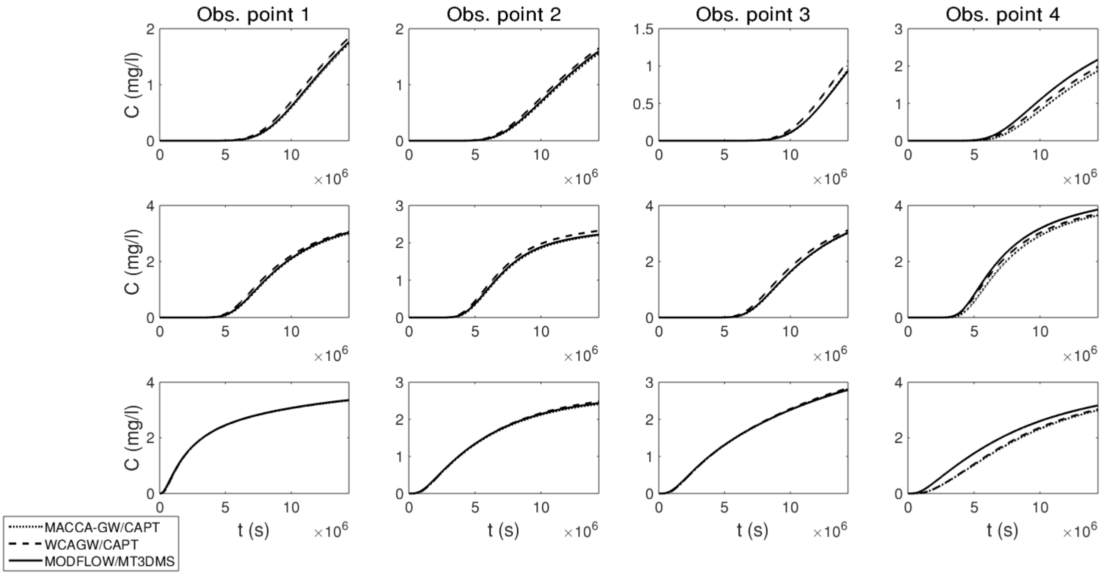


Figure 9. Time series of the pollutant concentration at four observation points for case 3 (top row), case 4 (middle row) and case 5 (bottom row) – pollutant constantly injected at the source (constant concentration)

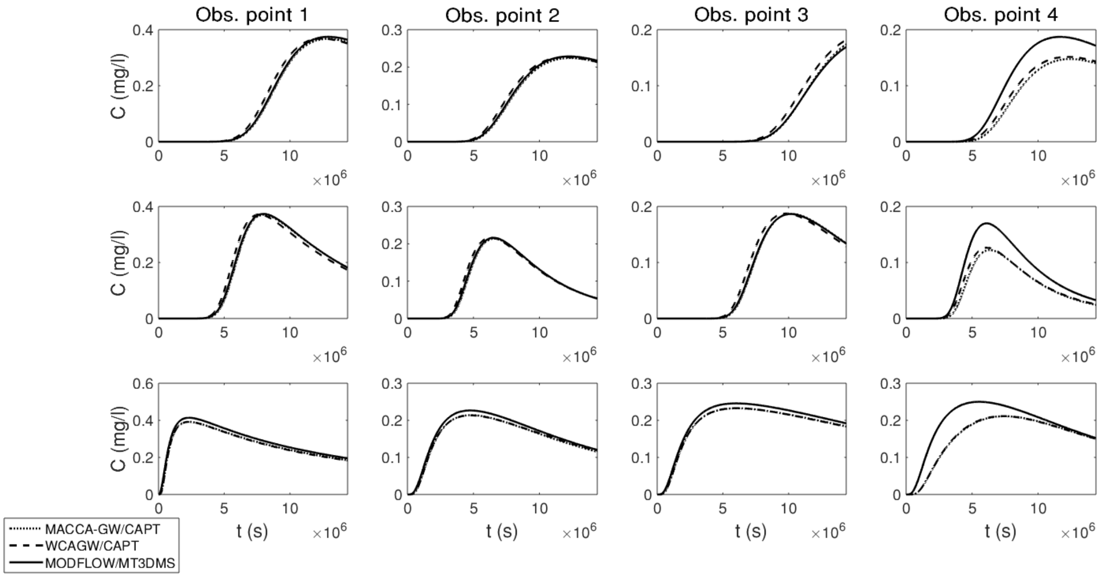


Figure 10. Time series of the pollutant concentration at four observation points for case 9 (top row), case 10 (middle row) and case 11 (bottom row) – pollutant instantaneously injected at the source

Figures 11 and 12 represent total pollutant mass in the aquifer during the simulation. Fig. 11 shows total pollutant mass time series for scenarios with constant concentration at the pollution source (cases 1 - 6). Total pollutant mass in the aquifer is growing (cases 1-6) because pollutant mass is constantly loaded at the source in order to keep constant pollutant concentration. Fig. 12 shows total pollutant mass time series for scenarios with pollutant injected over one time step (cases 7 - 12). Total pollutant mass for cases 7-12 decreases during the simulation due to mass leakage through the western and northern boundary. Figs. 11 and 12 show both visually and statistically good agreement between MACCA-GW/CAPT, WCAGW/CAPT and MODFLOW/MT3DMS results. Using *NRMSE* as best indicator for quantification of the models consistency, it can be seen that *NRMSE* for total pollutant mass has lower values than 0.82 %, which is the max NRMSE value in case 5 when MACCA-GW/CAPT model is used. When WCAGW/CAPT model is used, max *NRMSE* value is 1.56 %.

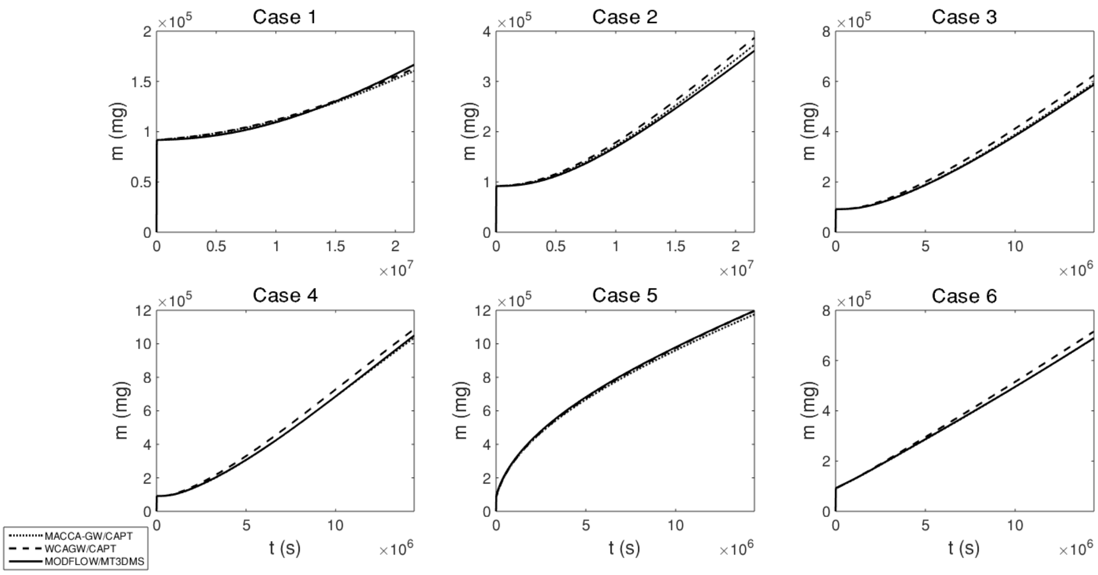


Figure 11. Total pollutant mass in the aquifer during the simulation – cases 1 – 6 (constant pollution concentration at the source)

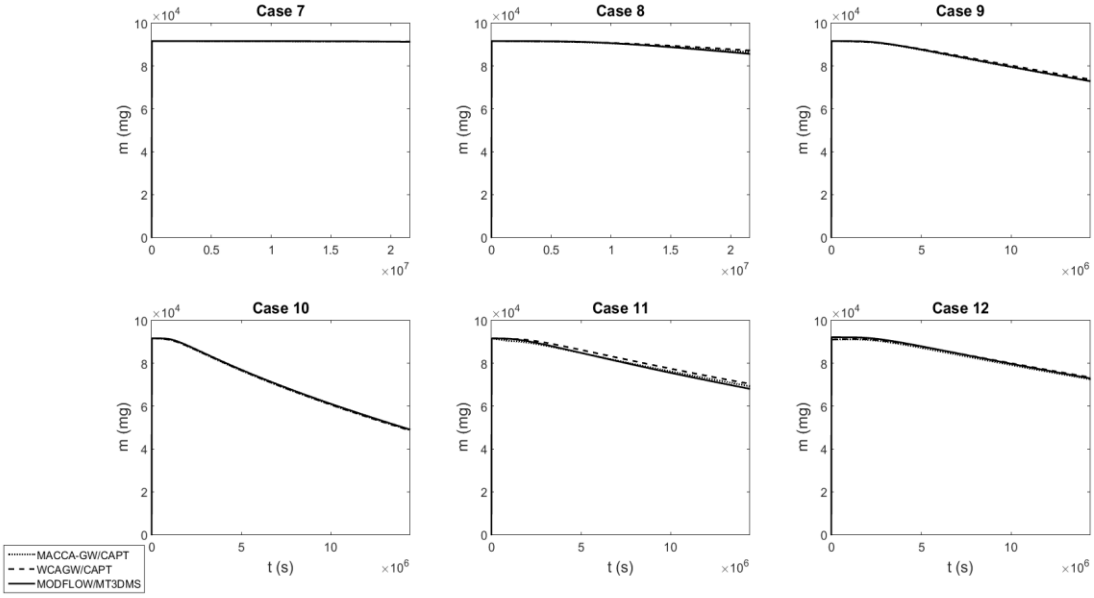


Figure 12. Total pollutant mass in the aquifer during the simulation – cases 7 – 12 (instantaneously injected pollution at the source)

Table 8. Statistical indicators of consistency between CA-based models and MODFLOW/MT3DMS – comparison of the pollutant concentration time series in observation points

	Obs. Point 1 - R^2 / RMSE / NRMSE		Obs. Point 2 - R^2 / RMSE / NRMSE		Obs. Point 3 - R^2 / RMSE / NRMSE		Obs. Point 4 - R^2 / RMSE / NRMSE	
Case	MACCA-GW/CAPT	WCAGW/CAPT	MACCA-GW/CAPT	WCAGW/CAPT	MACCA-GW/CAPT	WCAGW/CAPT	MACCA-GW/CAPT	WCAGW/CAPT
1	0.99 / 0.002 / 0.04	0.99 / 0.001 / 0.03	0.99 / 10^{-4} / 0.002	0.99 / 10^{-4} / 0.002	0.99 / $3*10^{-6}$ / $7*10^{-5}$	0.99 / $5*10^{-6}$ / 10^{-4}	0.98 / $5*10^{-6}$ / 10^{-4}	0.99 / $2*10^{-6}$ / $5*10^{-5}$
2	0.99 / 0.03 / 0.61	0.99 / 0.007 / 0.13	0.99 / 0.03 / 0.63	0.99 / 0.02 / 0.31	0.99 / 0.004 / 0.09	0.99 / 0.01 / 0.20	0.99 / 0.05 / 0.92	0.99 / 0.03 / 0.65
3	0.99 / 0.02 / 0.34	0.99 / 0.05 / 1.09	0.99 / 0.02 / 0.45	0.99 / 0.04 / 0.82	0.99 / 0.004 / 0.08	0.99 / 0.05 / 1.07	0.99 / 0.2 / 3.95	0.99 / 0.12 / 2.45
4	0.99 / 0.03 / 0.53	0.99 / 0.07 / 1.43	0.99 / 0.02 / 0.47	0.99 / 0.08 / 1.52	0.99 / 0.01 / 0.28	0.99 / 0.1 / 1.96	0.99 / 0.22 / 4.37	0.99 / 0.12 / 2.34
5	0.99 / 0.01 / 0.29	0.99 / 0.01 / 0.26	0.99 / 0.02 / 0.4	0.99 / 0.02 / 0.42	0.99 / 0.01 / 0.23	0.99 / 0.02 / 0.45	0.99 / 0.3 / 5.97	0.99 / 0.27 / 5.42
6	0.99 / 0.02 / 0.36	0.99 / 0.03 / 0.61	0.99 / 0.02 / 0.46	0.99 / 0.03 / 0.67	0.99 / 0.006 / 0.12	0.99 / 0.05 / 1.03	0.99 / 0.23 / 4.59	0.99 / 0.16 / 3.19
7	0.99 / 0.002 / 0.03	0.99 / 0.001 / 0.02	0.99 / 10^{-4} / 0.002	0.99 / $8.5*10^{-4}$ / 0.002	0.99 / $2.8*10^{-6}$ / $5.7*10^{-5}$	0.99 / $4.1*10^{-6}$ / $8.3*10^{-5}$	0.98 / $3.7*10^{-6}$ / $7.4*10^{-5}$	0.99 / $1.5*10^{-6}$ / $3*10^{-5}$
8	0.99 / 0.02 / 0.32	0.99 / 0.007 / 0.15	0.99 / 0.01 / 0.26	0.99 / 0.008 / 0.16	0.99 / 0.002 / 0.05	0.99 / 0.005 / 0.09	0.98 / 0.01 / 0.27	0.99 / 0.01 / 0.20
9	0.99 / 0.005 / 0.10	0.99 / 0.01 / 0.24	0.99 / 0.004 / 0.07	0.99 / 0.006 / 0.12	0.99 / 0.002 / 0.03	0.99 / 0.009 / 0.18	0.98 / 0.03 / 0.64	0.99 / 0.03 / 0.52
10	0.99 / 0.004 / 0.09	0.99 / 0.01 / 0.3	0.99 / 0.003 / 0.06	0.99 / 0.006 / 0.13	0.99 / 0.001 / 0.03	0.99 / 0.007 / 0.15	0.98 / 0.03 / 0.55	0.99 / 0.02 / 0.48
11	0.99 / 0.02 / 0.34	0.99 / 0.02 / 0.36	0.99 / 0.01 / 0.19	0.99 / 0.009 / 0.19	0.99 / 0.01 / 0.23	0.99 / 0.01 / 0.23	0.8 / 0.04 / 0.89	0.81 / 0.04 / 0.88
12	0.99 / 0.005 / 0.09	0.99 / 0.005 / 0.1	0.99 / 0.004 / 0.07	0.99 / 0.003 / 0.067	0.99 / 0.001 / 0.03	0.99 / 0.006 / 0.11	0.98 / 0.04 / 0.78	0.98 / 0.03 / 0.68

Table 9. Statistical indicators of consistency between CA-based models and MODFLOW/MT3DMS – comparison of the pollutant concentration spatial distribution at the end of the simulation

Case	Pollution concentration spatial R^2 / RMSE / NRMSE		Total pollution mass R^2 / RMSE / NRMSE	
	MACCA-GW / CAPT	WCAGW / CAPT	MACCA-GW / CAPT	WCAGW / CAPT
1	0.95 / 0.17 / 3.39	0.95 / 0.17 / 3.31	0.98 / 3792.45 / 0.19	0.98 / 3466.08 / 0.18
2	0.99 / 0.08 / 1.50	0.99 / 0.08 / 1.63	0.99 / 6056.19 / 0.31	0.99 / 13662.36 / 0.7
3	0.99 / 0.03 / 0.6	0.99 / 0.07 / 1.49	0.99 / 4438.73 / 0.23	0.99 / 22501.6 / 1.16
4	0.99 / 0.04 / 0.89	0.99 / 0.08 / 1.51	0.99 / 5571.98 / 0.29	0.99 / 30515.5 / 1.56
5	0.99 / 0.05 / 1.02	0.99 / 0.08 / 1.55	0.99 / 15899.5 / 0.82	0.99 / 5857.96 / 0.3
6	0.86 / 0.04 / 0.83	0.86 / 0.08 / 1.54	0.99 / 820.54 / 0.04	0.99 / 15729.1 / 0.81
7	0.94 / 0.09 / 1.87	0.95 / 0.09 / 1.79	0.99 / 79.53 / 0.004	0.99 / 68.87 / 0.004
8	0.99 / 0.02 / 0.39	0.98 / 0.02 / 0.47	0.99 / 304.54 / 0.02	0.98 / 682.58 / 0.04
9	0.99 / 0.008 / 0.16	0.99 / 0.007 / 0.15	0.99 / 152.11 / 0.008	0.99 / 481.74 / 0.02
10	0.99 / 0.007 / 0.14	0.99 / 0.004 / 0.08	0.99 / 354.4 / 0.02	0.99 / 371.66 / 0.02
11	0.99 / 0.005 / 0.10	0.93 /0.007 / 0.14	0.99 / 5062.90 / 0.26	0.99 / 3890.31 / 0.2
12	0.72 / 0.004 / 0.09	0.73 / 0.006 / 0.13	0.99 / 686.62 / 0.04	0.99 / 419.86 / 0.02

CONCLUSIONS

This paper presents a novel approach in modelling groundwater pollution transport by cellular automata based model CAPT, coupled with hydrodynamic cellular automata models MACCA-GW (Ravazzani et al. 2011) and modified weighted cellular automata WCAGW, derived by modifying WCA2D (Guidolin et al. 2016) model for groundwater problems. This modelling tool is compared with analytical solutions for one-dimensional (Ogata and Banks 1961) and two-dimensional (Bear 1972) problems , showing excellent results: *NRMSE* for total mass is in the order of 1%, both for 1D and 2D tests. These good results were guidance for additional comparison of the proposed CA-based methodology to widely used numerical model MODFLOW/MT3DMS. Accordingly, hypothetical test case was created on rectangular domain, with constant infiltration rate causing unsteady hydrodynamic conditions. Different test cases were created by analysing the impact of different transport parameters (longitudinal dispersivity and molecular diffusion coefficient) values and different type of pollution source. Results (Water level and pollution concentration) are presented in the way of spatial distribution at the end of the simulation and time series at selected points.

Analysing the results through the statistical indicators the following conclusions can be derived:

- Both hydrodynamic Cellular Automata models, MACCA-GW and WCAGW, show very good agreement with MODFLOW obtained results, considering MACCA-GW being better model based on statistical and visual indicators.
- Consistency in pollution concentration and pollutant mass is lower than consistency in hydrodynamic modelling, which is expected. Seepage velocity field and head field, obtained by CA-based hydrodynamic models, are propagated in the CA transport model. That causes bigger discrepancies between CA-based transport model CAPT and MODFLOW/MT3DMS model.
- Boundary conditions setting can affect bigger discrepancy in some regions of the domain. Thus, boundary conditions impact should be further investigated.
- Statistical indicators show good agreement between CA-based coupled hydrodynamic/transport models (MACCA-GW/CAPT and WCAGW/CAPT) and MODFLOW/MT3DMS model. In most cases, coupled MACCA-GW/CAPT model shows better agreement with coupled MODFLOW/MT3DMS model. In some cases, especially when longitudinal dispersivity is low (and seepage velocity impact accordingly), coupled WCAGW/CAPT shows better agreement with MODFLOW/MT3DMS model.
- CAPT model, compared to 1D and 2D analytical solutions, shows excellent results, considering 2D analytical solution being dispersion dominant (advection transport mechanism is neglected by setting the zero seepage velocity). Thus, it can be said that CA-based models show good results for dispersion dominant cases (dispersion is dominant transport mechanism).
- When total pollutant mass is analysed, CA based models are mass conservative due to the limitation implemented in eqs. 21 and 29. These equations limit the mass being transferred by advection and dispersion from a cell to the value of the cell pollutant

mass in previous step. In other words, pollutant mass available for transport cannot be greater than current mass in the cell. By applying these limitations, potential instabilities are avoided.

Based on the results analysis and previous specific conclusions, general conclusion can be derived. Simplified method for solving transport problems in groundwater, based on significant equations complexity reduction, shows high usage potential, without significant sacrifice of the accuracy, when compared to standard methods. Hence, CA-based models for groundwater and pollution transport modelling show usage justification, especially in integrated hydrological models where model simplifications are inevitable when long-term simulations are performed. Therefore, further analysis and investigations are necessary. Following future research steps should be investigated:

- Cell grid shape impact analysis (e.g. hexagonal cell grid instead of the rectangular cell grid)
- Implementing the CA model simulations on multi-core CPU (Central Processing Unit) or GPU (Graphics Processing Unit) in order to speed up Cellular Automata simulations by exploiting full CA potential through parallel computing
- Testing the proposed methodology in different test cases, especially on real scenarios with field collected hydrodynamic and pollution data
- Upgrade current CA model for pollution transport modelling by implementing additional transport mechanisms, such as sorption, volatilisation, biodegradation and radioactive decay
- Model uncertainty analysis.

ACKNOWLEDGEMENT

The authors are grateful to the Serbian Ministry of Education, Science and Technological Development for its financial support, project No. TR37010.

References

Bandini, S., G. Mauri, and R. Serra. 2001. “Cellular Automata: From a Theoretical Parallel Computational Model to Its Application to Complex Systems.” *Parallel Computing* 27(5):539–53.

Barbé, D. E., J. F. Cruise, and X. Mo. 1996. “Modeling the Buildup and Washoff of Pollutants on Urban Watersheds.” *Journal of the American Water Resources Association* 32(3):511–19. Retrieved (http://doi.wiley.com/10.1111/j.1752-1688.1996.tb04049.x).

Barredo, Ji and Marjo Kasanko. 2003. “Modelling Dynamic Spatial Processes: Simulation of Urban Future Scenarios through Cellular Automata.” *Landscape and Urban* … 64:145–60. Retrieved (http://www.sciencedirect.com/science/article/pii/S0169204602002189).

Batty, M., Yichun Xie, and Zhanli Sun. 1999. “Modeling Urban Dynamics through GIS-Based Cellular Automata.” *Computers, Environment and Urban Systems* 23(3):205–33.

Bear, Jacob. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York.

Cai, Xin, Yi Li, Xing-Wen Guo, and Wei Wu. 2014. “Mathematical Model for Flood Routing Based on Cellular Automaton.” *Water Science and Engineering* 7(2):133–42. Retrieved (http://www.waterjournal.cn).

Deletic, Ana, [Cbreve]edo Maksimovic, and Marko Ivetic. 1997. “Modelling of Storm Wash-off of Suspended Solids from Impervious Surfaces.” *Journal of Hydraulic Research* 35(1):99–118. Retrieved (http://www.tandfonline.com/doi/abs/10.1080/00221689709498646).

Deletic, Ana and David W. Orr. 2005. “Pollution Buildup on Road Surfaces.” *Journal of Environmental Engineering* 131(1):49–59.

Dimkić, Milan, Milenko Pušić, Dragan Vidović, Dušan Djurić, and Đulija Boreli-Zdravković. 2013. “Pollution Transport Analysis in Defining the Sanitary Protection Zones of Groundwater Sources in Alluvial Areas.” *Vodoprivreda* 45(1):203–18. Retrieved (http://www.vodoprivreda.net/analiza-transporta-zagadenja-kod-odredivanja-zona-sanitarne-zastite-izvorista-podzemnih-voda-u-aluvijalnim-sredinama/).

Djukić, Aleksandar et al. 2016. “Further Insight into the Mechanism of Heavy Metals Partitioning in Stormwater Runoff.” *Journal of Environmental Management* 168:104–10.

Dottori, F. and E. Todini. 2010. “A 2D Flood Inundation Model Based on Cellular Automata Approach.” *XVIII International Conference on Water Resources* (2):1–8.

Duong, Trang T. T. and Byeong Kyu Lee. 2011. “Determining Contamination Level of Heavy Metals in Road Dust from Busy Traffic Areas with Different Characteristics.” *Journal of Environmental Management* 92(3):554–62. Retrieved (http://dx.doi.org/10.1016/j.jenvman.2010.09.010).

Espínola, Moisés et al. 2016. “Simulating Rainfall, Water Evaporation and Groundwater Flow in Three-Dimensional Satellite Images with Cellular Automata.” *Simulation Modelling Practice and Theory* 67:89–99.

Feng, Yongjiu and Xiaohua Tong. 2018. “Dynamic Land Use Change Simulation Using Cellular Automata with Spatially Nonstationary Transition Rules.” *GIScience and Remote Sensing* 55(5):678–98. Retrieved (http://dx.doi.org/10.1080/15481603.2018.1426262).

Ghimire, Bidur et al. 2013. "Formulation of a Fast 2D Urban Pluvial Flood Model Using a Cellular Automata Approach." *Journal of Hydroinformatics* 15(3):676. Retrieved (http://jh.iwaponline.com/cgi/doi/10.2166/hydro.2012.245).

Ghisu, Tiziano, Bachisio Arca, Grazia Pellizzaro, and Pierpaolo Duce. 2015. "An Optimal Cellular Automata Algorithm for Simulating Wildfire Spread." *Environmental Modelling and Software* 71:1–14. Retrieved (http://dx.doi.org/10.1016/j.envsoft.2015.05.001).

Gibson, Michael J., Edward C. Keedwell, and Dragan A. Savić. 2015. "An Investigation of the Efficient Implementation of Cellular Automata on Multi-Core CPU and GPU Hardware." *Journal of Parallel and Distributed Computing* 77:11–25.

Gług, Maciej and Jarosław Wąs. 2018. "Modeling of Oil Spill Spreading Disasters Using Combination of Langrangian Discrete Particle Algorithm with Cellular Automata Approach." *Ocean Engineering* 156(March):396–405.

Guariso, Giorgio and Vittorio Maniezzo. 1992. "Air Quality Simulation through Cellular Automata." *Environmental Software* 7(3):131–41.

Guariso, Giorgio, Vittorio Maniezzo, and Paola Salomoni. 1996. "Parallel Simulation of a Cellular Pollution Model." *Applied Mathematics and Computation* 79(1):27–41. Retrieved (http://www.elsevier.com/cgi-bin/cas/tree/store/amc/cas_sub/browse/browse.cgi?year=1996&volume=79&issue=1&aid=9500214).

Guidolin, Michele et al. 2016. "A Weighted Cellular Automata 2D Inundation Model for Rapid Flood Analysis." *Environmental Modelling and Software* 84:378–94.

Langevin, Christian D. et al. 2017. "Documentation for the MODFLOW 6 Groundwater Flow Model." *Book 6: Modeling Techniques*. Retrieved (https://pubs.usgs.gov/tm/06/a55/tm6a55.pdf).

Lauret, Pierre, Frédéric Heymes, Laurent Aprin, and Anne Johannet. 2016. "Atmospheric Dispersion Modeling Using Artificial Neural Network Based Cellular Automata." *Environmental Modelling and Software* 85:56–69.

Li, Yi et al. 2015. "Real-Time Flood Simulations Using CA Model Driven by Dynamic Observation Data." *International Journal of Geographical Information Science* 29(4):523–35.

Lin, Menglong and Yiping Yao. 2018. "Simulation of Water Pollution Accident Based on Cellular Automata." *ICMSS 2018 - Proceedings of the 2018 2nd International Conference on Management Engineering, Software Engineering and Service Sciences* 270–74.

Marín, M. et al. 2000. "Cellular Automata Simulation of Dispersion of Pollutants." *Computational Materials Science* 18(2):132–40. Retrieved (http://linkinghub.elsevier.com/retrieve/pii/S0927025600000975).

Milasinovic, Milos, Anja Randelovic, Nenad Jacimovic, and Dusan Prodanovic. 2019. "Cellular Automata Approach for 2D Pollution Transport Modelling in Urban Groundwater." *In: Mannina G. (Eds) New Trends in Urban Drainage Modelling. UDM 2018. Green Energy and Technology. Springer, Cham* 765–70. Retrieved (http://link.springer.com/10.1007/978-3-319-99867-1).

Niswonger, Richard G., Sorab Panday, and Motomu Ibaraki. 2011. *MODFLOW-NWT, A Newton Formulation for MODFLOW-2005*. Retrieved (http://pubs.er.usgs.gov/publication/tm6A37).

Ogata, Akio and Robert Blackburn Banks. 1961. “A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media.” *Geological Survey (U.S.).; Professional Paper* A1–7. Retrieved (http://pubs.er.usgs.gov/publication/pp411A).

Palanichamy, Jegathambal, Holger Schüttrumpf, and Sundarambal Palani. 2008. “A Probabilistic Cellular Automaton for Two Dimensional Contaminant Transport Simulation in Ground Water.” *Water Science and Technology* 58(11):2083–92.

Panday, Sorab, Christian D. Langevin, Richard G. Niswonger, M. Ibaraki, and Joseph D. Hughes. 2013. “MODFLOW-USG Version 1: An Unstructured Grid Version of MODFLOW for Simulating Flow and Tightly Coupled Processes Using a Control Volume Finte-Difference Formulation.” 66.

Pantić, Tanja Petrović, Ţarko Veljković, Milan Tomić, and Katarina Samolov. 2017. “Hydrogeology and Vulnerability of Groundwater to Polluting in the Area of National Parka Fruška Gora.” *Vodoprivreda* 49(288):287–95. Retrieved (http://www.vodoprivreda.net/hidrogeologija-i-ranjivost-podzemnih-voda-u-podrucju-nacionalnog-parka-fruska-gora/).

Petrović Pantić, Tanja, Mihajlo Mandić, and Katarina Samolov. 2016. “Water Supply in the Area of Kosmaj, Mladenovac, Smederevo and Smederevska Palanka.” *Vodoprivreda* 48(282):267–75. Retrieved (http://www.vodoprivreda.net/hidrogeologija-i-vodoprivredna-problematika-na-podrucju-kosmaja-mladenovca-smedereva-i-smederevske-palanke/).

Ravazzani, G., D. Rametta, and M. Mancini. 2011. “Macroscopic Cellular Automata for Groundwater Modelling: A First Approach.” *Environmental Modelling and Software* 26(5):634–43. Retrieved (http://dx.doi.org/10.1016/j.envsoft.2010.11.011).

Revitt, D. Michael, Lian Lundy, Frédéric Coulon, and Martin Fairley. 2014. “The Sources, Impact and Management of Car Park Runoff Pollution: A Review.” *Journal of Environmental Management* 146:552–67.

Rui, Xiao Ping, Ding Zhen, Jin Lu, and Xian Feng Song. 2013. “Simulation of Point Source Pollution Diffusion Using a Velocity Field-Cellular Automata Coupled Method.” *Information Technology Journal* 12(20):5424–31.

Schulze-Makuch, Dirk. 2005. “Longitudinal Dispersivity Data and Implications for Scaling Behavior.” *Ground Water* 43(3):443–56.

Shao, Qi, Dion Weatherley, Longbin Huang, and Thomas Baumgartl. 2015. “RunCA: A Cellular Automata Model for Simulating Surface Runoff at Different Scales.” *Journal of Hydrology* 529(P3):816–29. Retrieved (http://dx.doi.org/10.1016/j.jhydrol.2015.09.003).

Vicari, Annamaria et al. 2007. “Modeling of the 2001 Lava Flow at Etna Volcano by a Cellular Automata Approach.” *Environmental Modelling and Software* 22(10):1465–71.

Winston, Richard B. 2009. “ModelMuse: A Graphical User Interface for MODFLOW-2005 and PHAST.” *U.S. Geological Survey Techniques and Methods 6-A29* 52 p. Retrieved (http://pubs.usgs.gov/tm/tm6A29/tm6A29.pdf).

Wolfram, Stephen. 1998. “Cellular Automata as Models of Complexity.” *Nonlinear Physics for Beginners: Fractals, Chaos, Solitons, Pattern Formation, Cellular Automata, Complex Systems* 311(5985):197. Retrieved (http://www.stephenwolfram.com/publications/academic/cellular-automata-models-complexity.pdf).

WWAP (United Nations World Water Assessment Programme). 2015. *The United Nations World Water Development Report 2015: Water for a Sustainable World.* Paris, UNESCO.

Zheng, Chunmiao and P. Wang. 1999. “MT3DMS: A Modular Three-Dimensional Multispeces Transport Model for Simulation of Advection, Dispersion, and Chemical Reactions of Contaminants in Groundwater Systems. Technical Report, Waterways Experiment Station, US Army Corps of Engineers.” *A Modular Three-Dimensional Multi-Species …* (June):239. Retrieved (http://www.geo.tu-freiberg.de/hydro/vorl_portal/gw-modellierung/MT3DMS Ref Manual.pdf).