Development and identification of a multi-species water quality model for reclaimed water distribution systems

Shun Li, Fu Sun, Siyu Zeng, Xin Dong and Pengfei Du

ABSTRACT

With the rapid development of a centralized wastewater reuse scheme in China, water quality concerns arise considering the long-distance transport of reclaimed water in distribution systems from wastewater treatment plants to points of use. To this end, a multi-species water quality model for reclaimed water distribution systems (RWDSs) was developed and validated against the data from part of a full-scale RWDS in Beijing. The model could simulate organics, ammonia nitrogen, residual chlorine, inert particles, and six microbial species, i.e. fecal coliforms, *Enterococcus* spp., *Salmonella* spp., *Mycobacterium* spp., and other heterotrophic and autotrophic bacteria, in both the bulk liquid and the biofilm. Altogether, 56 reaction processes were involved, and 37 model parameters and seven initial values were identified. Despite the limited monitoring data and the associated gross uncertainty, the model could simulate the reclaimed water quality in the RWDS with acceptable accuracy. Regional sensitivity analysis suggested that the model had a balanced structure with a large proportion of sensitive parameters, and the sensitivity of model parameters could be reasonably interpreted by current knowledge or observation. Furthermore, the most sensitive model parameters could generally be well identified with uncertainties significantly reduced, which also favored the trustworthiness of the model. Finally, future plans to improve and apply the model were also discussed.

Key words | distribution system, model, multi-species, parameter identification, reclaimed water

Shun Li
Fu Sun
Siyu Zeng
Xin Dong
Pengfei Du (corresponding author)
School of Environment, Tsinghua University,
Beijing 100084,
China
E-mail: *dupf@tsinghua.edu.cn*

INTRODUCTION

Wastewater reclamation has been promoted in China to meet the increasing water demand of the growing population and economy. In the *National 12th Five-Year Plan for the Development of Urban Wastewater Treatment and Reuse Infrastructure*, China aims to reuse more than 15% of its urban wastewater by 2015, while some water-scarce cities have set more ambitious local goals, e.g. 75% for Beijing. According to the national plan, centralized recycling based on existing wastewater treatment plants (WWTPs) will be the dominant scheme, and therefore long-distance transfer and distribution of reclaimed water would be unavoidable at this stage. In Beijing, for example, the *Three-Year Action Plan for the Development of Urban Wastewater Treatment and Reuse Infrastructure (2013–2015)* expects to lay 158 km pipelines for reclaimed water and facilitate its transportation and allocation among the service areas of different WWTPs.

Many countries including China have established water quality standards for wastewater reuse. These standards are usually applied by the water industries to the effluent of the wastewater reclamation facilities rather than the point of use where human or aquatic life is directly exposed. However, the quality of treated reclaimed water may deteriorate in distribution systems, e.g. disinfectant decay, microbial regrowth, and disinfection by-products (DBPs) formation, similar to what happens to drinking water distribution systems (DWDSs) (Arminski *et al.* 2013). To ensure the safety of human or ecosystem health at the point of use, it is imperative to understand how the quality of treated reclaimed water changes in the distribution systems and

doi: 10.2166/wrd.2015.115

thus design preventative or remedial measures to maintain or improve the point-of-use water quality.

Modeling is a widely recognized tool to study the temporal and spatial variations of water quality in distribution systems. However, water quality models for reclaimed water distribution systems (RWDSs) have only been developed since the 2000s plausibly due to the late start of large scale wastewater reclamation practice. Diaper *et al.* (2001) simulated the infection probability of *Salmonella* spp. due to disinfection failure as well as the variations in dissolved oxygen and chemical oxygen demand (COD) due to pump failure in a small scale gray water recycling system. Narasimhan *et al.* (2005) applied a H_2ONET model to simulate the water age and residual chlorine in a RWDS and evaluate several water quality improvement options, e.g. changing the operation of tanks and pumps and increasing chlorine doses. Narasimhan *et al.* (2005) also modeled residual chlorine in another two RWDSs with the EPANET and WaterCAD software, respectively. Xin *et al.* (2007) studied the planned RWDS in Beijing, China and simulated the water age and residual chlorine in the RWDS in different seasons. Chi (2010) studied an RWDS in another Chinese city and simulated the residual chlorine. Instead of these previously mentioned models based on hydraulic and kinetic principles, Takahashi (2012) developed a system dynamics model to design optimal water reclamation scheduling in terms of flow, quality and energy, and the water quality module could simulate biochemical oxygen demand (BOD), total nitrogen, and total phosphorus.

As compared with water quality models for DWDSs, which have been studied since the 1980s (Walski *et al.* 2003), research into water quality modeling for RWDSs is still in its infancy. Only a few existing models were validated against monitoring data collected from real RWDSs, e.g. those presented by Narasimhan *et al.* (2005), while others were based on literature data and scenario analysis. Furthermore, existing models only focused on limited water quality indicators, mainly residual chlorine, and simulated them in a separate rather than coupled way. Therefore, these single-species models could not characterize the interactions among the different constituents in reclaimed water. To this end, this paper proposed a multi-species water quality model for RWDSs and validated it against field monitoring data from a full-scale RWDS in Beijing, China, referring to the increasing multi-species water quality modeling studies on DWDSs especially since the release of the EPANET multi-species extension (EPANET-MSX) software (Shang *et al.* 2008) in 2008, e.g. Sun *et al.* (2008), Helbling & VanBriesen (2009), Shu (2011), Seyoum *et al.* (2013), and Schwartz *et al.* (2014). In the next section, the studied RWDS and the data were first introduced, and then details about model development and identification were followed. After that, the results of model performance, and the sensitivity, identifiability and uncertainty of the model parameters were presented, and future plans to improve and apply the model were also discussed before concluding the paper.

THE RWDS FOR THE QINGHE RECLAIMED WATER TREATMENT PLANT

The multi-species water quality model was developed for the RWDS for the Qinghe reclaimed water treatment plant (RWTP) in Beijing, China. The monitored RWDS received treated water from an 80,000 m^3/d section of the Qinghe RWTP, which adopted a treatment train including screening with a 300 μm strainer, ultrafiltration, ozone oxidation, and chlorine disinfection. The treated water was transported through an 8 km long, pressurized main and used to augment the flow of the Qing River, a major urban stream in Beijing.

The field monitoring program was conducted between June 2012 and March 2013, and water samples were collected from both ends of this main on a monthly basis. However, the field monitoring program was not specifically designed for this modeling study, so the data were actually not adequate for model development and identification. On the one hand, the observed water quality indicators only included total organic carbon (TOC), ammonia nitrogen, residual chlorine, fecal coliforms, and three species of micro-organisms, i.e. *Enterococcus* spp., *Salmonella* spp., and *Mycobacterium* spp. TOC, ammonia nitrogen, and residual chlorine were found decreased during the transport, while significant regrowth of the four microbial species were observed in this process. These results are consistent with those in previous studies (Ryu *et al.* 2005; Jjemba *et al.*

2010; Thayanukul *et al.* 2013). Details about the monitoring program in the Qinghe reclaimed water systems and the results could be found in Lin (2013). On the other hand, the flow rate or the travel time of the treated water in the main was not recorded or measured, so the average flow rate was used to estimate the travel time. Despite the data limitation, this study developed a multi-species water quality for the Qinghe RWDS and preliminarily validated it against the available monitoring data.

THE MULTI-SPECIES WATER QUALITY MODEL FOR RWDSS

Model conceptualization

The proposed multi-species water quality model for RWDSs divided a pipe segment into four compartments, including pipe wall, biofilm, boundary layer, and bulk liquid, as shown in Figure 1. Six types of water quality species were involved in the model, i.e. organics represented as TOC, ammonia nitrogen, residual chlorine, heterotrophic bacteria, autotrophic bacteria, and inert particles. Considering data availability, the proposed model further broke down heterotrophic bacteria into five species, including fecal coliforms, *Enterococcus* spp., *Salmonella* spp., *Mycobacterium* spp., and other heterotrophic bacteria. This conceptualization could facilitate future microbial risk assessment. Therefore, altogether 10 species were simulated in both the bulk liquid and the biofilm by the proposed model. These species interacted with each other through a series of physical, chemical, and biological processes, as shown in Figure 1, including microbial growth, inactivation, and decay, decay of residue chlorine, mass transfer of soluble components through the boundary layer, and deposition and detachment of particular components. Regarding these processes, the major assumptions and mechanisms for

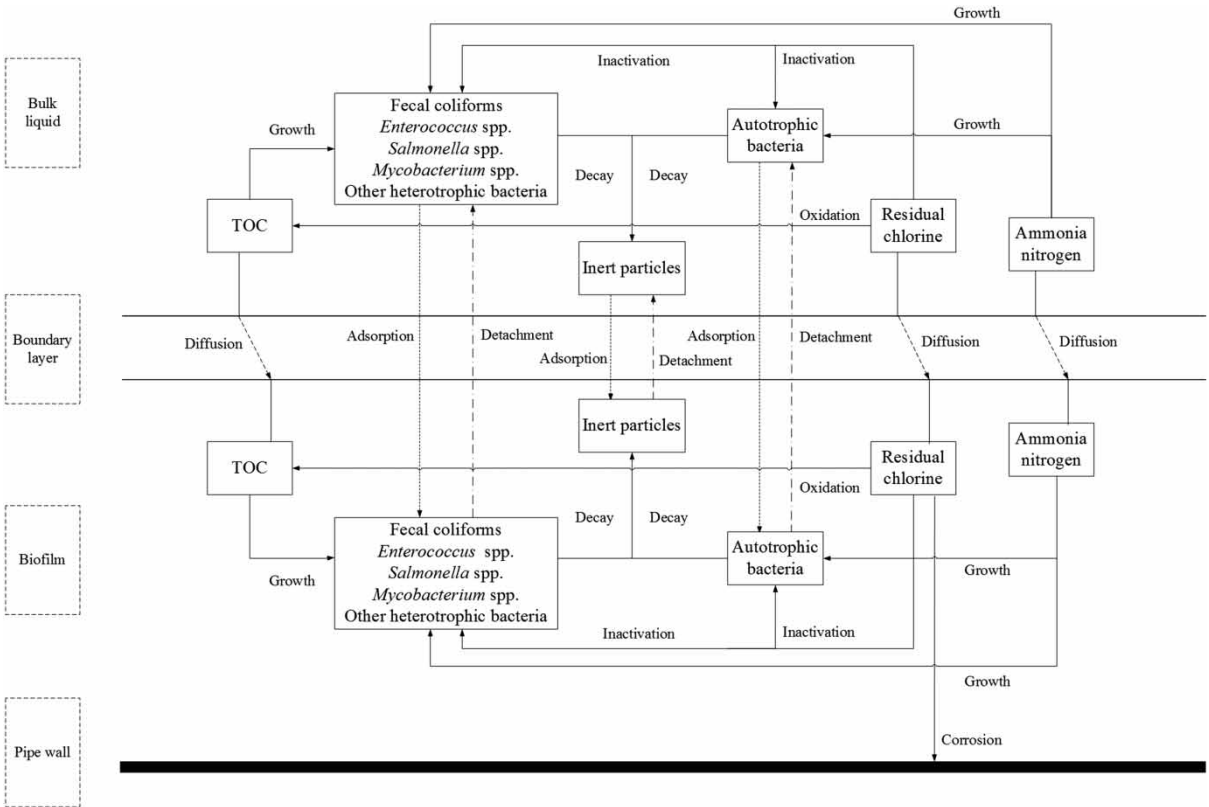

Figure 1 | Conceptualization of the multi-species water quality model for RWDSs.

these processes considered in model conceptualization were as follows:

1. One-dimensional advective transport along the pipes was assumed for all the species in the bulk liquid, and axial and radial dispersion in pipes was negligible, and thus the bulk liquid in each pipe segment was completely mixed (Lu *et al.* 1995; Jegatheesan *et al.* 2000; Walski *et al.* 2003; Munavalli & Mohan Kumar 2004; Sun *et al.* 2008).
2. The biofilm on the pipe wall was assumed to be homogeneous in each pipe segment and all the species were also completely mixed with the biofilm (Bois *et al.* 1997; Jegatheesan *et al.* 2000; Munavalli & Mohan Kumar 2004; Sun *et al.* 2008). Therefore, all the microbial species could have access to organics and nutrients for growth as well as be inactivated by disinfectants.
3. Mass transfer from the bulk liquid to the biofilm and vice versa only took place in the direction perpendicular to the biofilm, and no advective transport or dispersion was assumed within the biofilm (Lu *et al.* 1995; Bois *et al.* 1997; Jegatheesan *et al.* 2000; Munavalli & Mohan Kumar 2004; Sun *et al.* 2008).
4. Soluble species from the bulk liquid entered the biofilm through the boundary layer, where diffusion was assumed to occur (Munavalli & Mohan Kumar 2004; Sun *et al.* 2008). Furthermore, for simplicity, the diffusion coefficients were assumed to be the same for all soluble species.
5. Particular species in the bulk liquid were assumed to be deposited to the biofilm following a first-order reaction with respect to their contents in the bulk liquid (Bois *et al.* 1997; Munavalli & Mohan Kumar 2004; Zhang *et al.* 2004; Sun *et al.* 2008). Again, for simplicity, the deposition coefficients were assumed to be the same for all particular species.
6. Particular species were assumed to detach from the biofilm into the bulk liquid following a first-order reaction with respect to their contents in the biofilm, water velocities in the pipes, and biofilm thickness, respectively (Bois *et al.* 1997; Munavalli & Mohan Kumar 2004; Zhang *et al.* 2004; Sun *et al.* 2008). For simplicity, the detachment coefficients were assumed to be the same for all particular species.
7. Micro-organisms in the bulk liquid and the biofilm were assumed to have the same metabolisms (Munavalli & Mohan Kumar 2004; Sun *et al.* 2008). Heterotrophic species grew upon the utilization of available organics and nutrients following the Monod equation, while for autotrophic species only nutrient utilization was considered in the Monod equation (Henze *et al.* 2000; Sun *et al.* 2008). Both heterotrophic and autotrophic species decayed naturally following a first-order reaction with respect to their contents, and they were also inactivated by disinfectants following a first-order reaction with respect to the contents of micro-organisms and disinfectants, respectively (Bois *et al.* 1997; Munavalli & Mohan Kumar 2004; Zhang *et al.* 2004; Sun *et al.* 2008). Dead and inactivated micro-organisms became part of the inert particles (Henze *et al.* 2000; Munavalli & Mohan Kumar 2004; Sun *et al.* 2008).
8. Organic matter continued to react with residual disinfectants in the distribution systems following a first-order reaction with respect to their contents, respectively (Bois *et al.* 1997; Sun *et al.* 2008).

Model formulation

Based on the first assumption discussed above, the following mass equation could be established for the bulk liquid in each pipe segment. In the equation, $[C_j]$ denoted the concentration of the species C_j ($j = 1, 2, \ldots, 10$) in the bulk liquid, U represented the flow velocity in the pipe, $R(C_j)$ indicated the overall reaction rate of C_j in the pipe segment, t was time, and x was the distance along the pipe

$$\frac{\partial[C_j]}{\partial t} = -U\frac{\partial[C_j]}{\partial x} + R(C_j) \quad (1)$$

To give a simple, straightforward overview of the simulated species, processes and reaction kinetics of the proposed model, the Petersen matrix (Petersen 1965) was adopted herein to illustrate the linkages among the variables, reaction rates, and stoichiometric coefficients for each process as shown in Figure 2. The columns labeled from 'C1' to 'C20' represented the 20 model variables, i.e. 10 species in the bulk liquid and another 10 in the biofilm. Variables NH_3, Cl_2, and TOC, i.e. 'C1' to 'C3', represented ammonia nitrogen, residual chlorine, and TOC, respectively,

	Reaction components																				Reaction rates
	NH_3	Cl_2	TOC	FC	EN	SA	MY	OH	AU	IP	NH_3@	Cl_2@	TOC@	FC@	EN@	SA@	MY@	OH@	AU@	IP@	
No.	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	r_i
1	$-\frac{ea}{YH}$		$-\frac{1}{YH}$	1																	$ua\frac{[TOC]}{ka+[TOC]}\frac{[NH_3]}{knh+[NH_3]}[FC]$
2	$-\frac{eb}{YH}$		$-\frac{1}{YH}$		1																$ub\frac{[TOC]}{ka+[TOC]}\frac{[NH_3]}{knh+[NH_3]}[EN]$
3	$-\frac{ec}{YH}$		$-\frac{1}{YH}$			1															$uc\frac{[TOC]}{ka+[TOC]}\frac{[NH_3]}{knh+[NH_3]}[SA]$
4	$-\frac{ed}{YH}$		$-\frac{1}{YH}$				1														$ud\frac{[TOC]}{ka+[TOC]}\frac{[NH_3]}{knh+[NH_3]}[MY]$
5	$-\frac{es}{YH}$		$-\frac{1}{YH}$					1													$us\frac{[TOC]}{ka+[TOC]}\frac{[NH_3]}{knh+[NH_3]}[OH]$
6	$-\frac{1}{YA}$								1												$ux\frac{[NH_3]}{knh+[NH_3]}[AU]$
7											$\frac{ea}{YH}$		$-\frac{1}{YH}$	1							$ua\frac{[TOC@]}{ka+[TOC@]}\frac{[NH_3@]}{knh+[NH_3@]}[FC@]$
8											$\frac{eb}{YH}$		$-\frac{1}{YH}$		1						$ub\frac{[TOC@]}{ka+[TOC@]}\frac{[NH_3@]}{knh+[NH_3@]}[EN@]$
9											$\frac{ec}{YH}$		$-\frac{1}{YH}$			1					$uc\frac{[TOC@]}{ka+[TOC@]}\frac{[NH_3@]}{knh+[NH_3@]}[SA@]$
10											$-\frac{ed}{YH}$		$-\frac{1}{YH}$				1				$ud\frac{[TOC@]}{ka+[TOC@]}\frac{[NH_3]}{knh+[NH_3@]}[MY@]$
11											$-\frac{es}{YH}$		$-\frac{1}{YH}$					1			$us\frac{[TOC@]}{ka+[TOC@]}\frac{[NH_3@]}{knh+[NH_3@]}[OH@]$
12											$-\frac{1}{YA}$								1		$ux\frac{[NH_3@]}{knh+[NH_3]}[AU@]$
13				-1						em											$ba[FC]$
14					-1					em											$bb[EN]$
15						-1				em											$bc[SA]$
16							-1			em											$bd[MY]$
17								-1		em											$bs[OH]$
18									-1	em											$bx[AU]$
19														-1						em	$ba[FC@]$
20															-1					em	$bb[EN@]$
21																-1				em	$bc[SA@]$
22																	-1			em	$bd[MY@]$
23																		-1		em	$bs[OH@]$
24																			-1	em	$bx[AU@]$
25		$-ecl$		-1						em											$kina[Cl_2][FC]$
26		$-ecl$			-1					em											$kinb[Cl_2][EN]$
27		$-ecl$				-1				em											$kinc[Cl_2][SA]$
28		$-ecl$					-1			em											$kind[Cl_2][MY]$
29		$-ecl$						-1		em											$kint[Cl_2][OH]$
30		$-ecl$							-1	em											$kint[Cl_2][AU]$
31												$-ecl$		-1						em	$kina[Cl_2@][FC@]$
32												$-ecl$			-1					em	$kinb[Cl_2@][EN@]$
33												$-ecl$				-1				em	$kinc[Cl_2@][SA@]$
34												$-ecl$					-1			em	$kind[Cl_2@][MY@]$
35												$-ecl$						-1		em	$kint[Cl_2@][OH@]$
36												$-ecl$							-1	em	$kint[Cl_2@][AU@]$
37	$-Av$										$\frac{1}{Lf}$										$D_pU([NH_3]-[NH_3@])$
38			$-Av$										$\frac{1}{Lf}$								$D_pU([TOC]-[TOC@])$
39		$-Av$										$\frac{1}{Lf}$									$D_pU([Cl_2]-[Cl_2@])$
40				$-Av$										$\frac{1}{Lf}$							$kadp[FC]$
41					$-Av$										$\frac{1}{Lf}$						$kadp[EN]$
42						$-Av$										$\frac{1}{Lf}$					$kadp[SA]$
43							$-Av$										$\frac{1}{Lf}$				$kadp[MY]$
44								$-Av$										$\frac{1}{Lf}$			$kadp[OH]$
45									$-Av$										$\frac{1}{Lf}$		$kadp[AU]$
46										$-Av$										$\frac{1}{Lf}$	$kadp[IP]$
47			$Av\cdot Lf$											-1							$kdet[FC@]U$
48				$Av\cdot Lf$											-1						$kdet[EN@]U$
49					$Av\cdot Lf$											-1					$kdet[SA@]U$
50						$Av\cdot Lf$											-1				$kdet[MY@]U$
51							$Av\cdot Lf$											-1			$kdet[OH@]U$
52								$Av\cdot Lf$											-1		$kdet[AU@]U$
53									$Av\cdot Lf$											-1	$kdet[IP@]U$
54		-1	$-eccl$																		$kchr[TOC][Cl_2]$
55											-1	$-eccl$									$kchr[TOC@][Cl_2@]$
56												-1									$kt[Cl_2@]$

Figure 2 | The Petersen matrix of the multi-species water quality model for RWDSs.

while variables FC, EN, SA, MY, OH, AU, and IP, i.e. 'C4' to 'C10', represented fecal coliforms, *Enterococcus* spp., *Salmonella* spp., *Mycobacterium* spp., other heterotrophic bacteria, autotrophic bacteria, and inert particles, respectively. The 10 species in the biofilm used the same variables as their counterparts in the bulk liquid with a symbol '@' added correspondingly, i.e. 'C11' to 'C20' in Figure 2. The 56 rows characterized the reactions as shown in Figure 1 and discussed previously, and their rates were listed in the last column labeled 'r_i'. Rows 1–6, 13–18, and 25–30 quantified the growth, natural decay, and disinfectant inactivation processes of the six microbial species in the bulk liquid, while the counterparts in the biofilm were characterized in rows 7–12, 19–24, and 31–36, respectively. Rows 37–39 characterized the diffusion processes of ammonia nitrogen, residual chlorine, and TOC, respectively. Rows 40–46 and rows 47–53 represented the deposition and detachment processes of the seven particular species, respectively. Rows 54 and 55 characterized the reactions between organics and residual chlorine in the bulk liquid and the biofilm, respectively, while the last row characterized the chlorine demand on the pipe wall. The elements in the 56×20 matrix, except the last column 'r_i' in Figure 2, namely $M(i,j)$, represented the stoichiometric coefficient of species C_j in the ith reaction. So the overall reaction rate of species C_j, i.e. $R(C_j)$ in Equation (1), was given by the following equation:

$$R(C_j) = \sum_{i=1}^{56} M(i,j) \cdot r_i \tag{2}$$

Substituting Equation (2) into Equation (1), the mass balance equation of each species could be formulated. Herein, the equation for FC in the bulk liquid is given below as an example

$$\begin{aligned} \frac{\partial [\mathrm{FC}]}{\partial t} = & - U \frac{\partial [\mathrm{FC}]}{\partial x} + \mathrm{ua} \cdot \frac{[\mathrm{TOC}]}{\mathrm{ka} + [\mathrm{TOC}]} \frac{[\mathrm{NH_3}]}{\mathrm{knh} + [\mathrm{NH_3}]} [\mathrm{FC}] \\ & - \mathrm{ba} \cdot [\mathrm{FC}] - \mathrm{kina} \cdot [\mathrm{Cl_2}][\mathrm{FC}] - \mathrm{Av} \cdot \mathrm{kadp} \cdot [\mathrm{FC}] \\ & + \mathrm{Av} \cdot \mathrm{Lf} \cdot \mathrm{kdet} \cdot U \cdot [\mathrm{FC@}] \end{aligned} \tag{3}$$

All the parameters involved in the 'r_i' column in Figure 2 are described in detail in Table 1. The model was realized in the EPANET-MSX software, an extension to the original EPANET, which allowed it to model water distribution systems involving multiple, interacting chemical species (Shang *et al.* 2008).

Model identification

The 37 model parameters, together with seven initial values with respect to the contents and thickness of the biofilm on the pipe wall as shown in Table 1, were identified with the Hornberger–Spear–Young (HSY) algorithm, the procedure of which has been detailed in Beck (1987) and Sun *et al.* (2010), based on a Latin Hypercube Sampling (LHS) method. It is worth noting that the initial concentrations of different species of bacteria in biofilm in Table 1 were given in mass unit, i.e. mg/L, instead of colony-forming unit (CFU), i.e. CFU/L, to be compatible with the equations describing bacteria metabolisms, e.g. growth, decay, and inactivation, as shown in Figure 2. So the observed bacteria concentrations in CFU/L were converted to mass concentrations in the model by considering the sizes and mass densities of these microbial species, and vice versa. For simplicity, however, a same size of 1.0 μm and a same density of 1.0 g/cm^3 were assumed for all the modeled species in this study.

To apply the HSY algorithm for model identification, the initial ranges of these parameters and initial values were first determined according to the reported values in the literature (Henze *et al.* 2000; Sun 2007; Sun *et al.* 2008) as well as through some trial-and-error model simulations, which are given in Table 1. For simplicity, their initial probability distributions were assumed to be uniform distributions. Then, system behavior was defined to set a criterion to discriminate an acceptable model simulation from an unacceptable one. As a result, the acceptable simulations were categorized as behavior-giving ones, while those unacceptable were non-behavior-giving ones. For ammonia nitrogen, residual chlorine, and TOC, in spite of the effort to minimize the difference between observation and simulation, a behavior-giving simulation was practically defined as one with at least seven out of the nine data points that had absolute values of the relative errors less than 50%. Considering the gross uncertainty and randomness in water sampling in the full-scale RWDS and the errors in the plate count method in laboratory tests, the behavior-giving simulation for micro-organisms was

Table 1 | Description, initial ranges, and typical identified values of the model parameters

	Description	Units	Initial ranges	Typical identified values
ua	Maximum specific growth rate of fecal coliforms	min^{-1}	$0,1.0\times10^{-1}$	3.52×10^{-2}
ub	Maximum specific growth rate of *Enterococcus*	min^{-1}	$0,1.0\times10^{-1}$	2.58×10^{-2}
uc	Maximum specific growth rate of *Salmonella*	min^{-1}	$0,1.0\times10^{-1}$	3.86×10^{-2}
ud	Maximum specific growth rate of *Mycobacterium*	min^{-1}	$(0,1.0\times10^{-1})$	2.82×10^{-2}
us	Maximum specific growth rate of autotrophic bacteria	min^{-1}	$0,1.0\times10^{-2}$	2.22×10^{-3}
ux	Maximum specific growth rate of heterotrophic bacteria	min^{-1}	$0,1.0\times10^{-2}$	2.14×10^{-3}
ka	Half-saturation constant of TOC	Mg TOC/L	0,10	9.21
knh	Half-saturation constant of NH_3	Mg N/L	(0,1.0)	2.5×10^{-1}
ba	Natural attenuation rate of fecal coliforms	min^{-1}	$(0,1.0\times10^{-4})$	3.05×10^{-5}
bb	Natural attenuation rate of *Enterococcus*	min^{-1}	$0,1.0\times10^{-4}$	2.17×10^{-5}
Bc	Natural attenuation rate of *Salmonella*	min^{-1}	$0,1.0\times10^{-4}$	8.42×10^{-5}
Bd	Natural attenuation rate of *Mycobacterium*	min^{-1}	$0,1.0\times10^{-4}$	7.50×10^{-5}
Bs	Natural attenuation rate of autotrophic bacteria	min^{-1}	$0,1.0\times10^{-4}$	1.81×10^{-6}
Bx	Natural attenuation rate of heterotrophic bacteria	min^{-1}	$0,1.0\times10^{-4}$	3.73×10^{-5}
kint	Rate of residual chlorine inactivate micro-organisms	L $(mg\ min)^{-1}$	$0,1.0\times10^{-5}$	9.93×10^{-6}
kina	Rate of residual chlorine inactivate fecal coliforms	L $(mg\ min)^{-1}$	$0,1.0\times10^{-5}$	6.84×10^{-7}
Kinb	Rate of residual chlorine inactivate *Enterococcus*	L $(mg\ min)^{-1}$	$0,1.0\times10^{-5}$	2.45×10^{-7}
Kinc	Rate of residual chlorine inactivate *Salmonella*	L $(mg\ min)^{-1}$	$0,1.0\times10^{-5}$	6.57×10^{-7}
Kind	Rate of residual chlorine inactivate *Mycobacterium*	L $(mg\ min)^{-1}$	$0,1.0\times10^{-5}$	2.44×10^{-6}
D_F	Mass transfer coefficient of liquid-film	1	$0,1.0\times10^{-4}$	1.51×10^{-5}
Kadp	Adsorption rate of insoluble components	m^3/min	$0,1.0\times10^{-6}$	3.29×10^{-7}
Kdet	Desorption rate of insoluble components	m^{-1}	$0,1.0\times10^{-4}$	8.33×10^{-5}
Kchr	Oxidation rate of organic matter by residual chlorine	L $(mg\ min)^{-1}$	$0,1.0\times10^{-3}$	1.68×10^{-4}
kt	Rate of chlorine consumption by pipe wall	min^{-1}	0,1.0	9.76×10^{-1}
YH	Yield coefficient of heterotrophic bacteria	mg/mg	0,0.6	5.44×10^{-1}
YA	Yield coefficient of autotrophic bacteria	mg/mg	0,0.4	8.19×10^{-2}
ea	Nitrogen content of fecal coliforms	1	$0,1.0\times10^{-1}$	1.76×10^{-2}
eb	Nitrogen content of *Enterococcus*	1	$0,1.0\times10^{-1}$	6.89×10^{-2}
ec	Nitrogen content of *Salmonella*	1	$0,1.0\times10^{-1}$	1.09×10^{-3}
ed	Nitrogen content of *Mycobacterium*	1	$0,1.0\times10^{-1}$	1.78×10^{-2}
es	Nitrogen content of autotrophic bacteria	1	$0,1.0\times10^{-1}$	3.05×10^{-2}
ex	Nitrogen content of heterotrophic bacteria	1	$0,1.0\times10^{-1}$	4.66×10^{-3}
em	Conversion coefficient between microbes and inert particles	1	0,6	6.58×10^{-1}
ecl	Consumption coefficient of residual chlorine inactivate micro-organisms	1	0,10	5.29
eccl	Stoichiometric coefficient of TOC oxidation reactions by residual chlorine	1	0,10	4.31
Pf	Density of biofilm	mg/L	1.0×10^6,1.2×10^6	1.13×10^6
MK	Percentage of autotrophic bacteria in the inflow	1	0, 1.0	4.44×10^{-1}
A0	Initial concentration of fecal coliforms in the biofilm	mg/L	$0, 1.0\times10^4$	5.44×10^3
B0	Initial concentration of *Enterococcus* in the biofilm	mg/L	$0, 1.0\times10^4$	3.91×10^2
C0	Initial concentration of *Salmonella* in the biofilm	mg/L	$0, 1.0\times10^4$	6.64×10^3
D0	Initial concentration of *Mycobacterium* in the biofilm	mg/L	$0, 1.0\times10^4$	8.55×10^3
S0	Initial concentration of autotrophic bacteria in the biofilm	mg/L	$0, 1.0\times10^6$	7.91×10^4
X0	Initial concentration of heterotrophic bacteria in the biofilm	mg/L	$0, 1.0\times10^6$	9.98×10^4
Lf	Thickness of the biofilm	m	$0, 1.0\times10^{-4}$	7.94×10^{-5}

defined by the goodness-of-fit of the probability distributions between observed and simulated data, which was quantified by the Kolmogorov–Smirnov (K–S) test at a 0.05 significance level. However, it was still difficult to obtain a behavior-giving simulation where all the four species of micro-organisms could satisfy this criterion. Therefore, a more lenient criterion was adopted that three out of the four species should have no statistically significant difference in the two probability distributions between observation and simulation.

After making the above definition, all the parameters and initial values in Table 1 were randomly and independently sampled, with an LHS approach, according to their designated ranges and probability distributions and used for model simulation. The simulation results, together with the parameters and initial values, were then classified into a behavior-giving set and a non-behavior-giving set. The simulation continued until enough behavior-giving simulations were obtained. Finally, following the methodology proposed by a previous study (Sun *et al.* 2010), the simulation results from the behavior-giving set and the parameters and initial values from both the behavior-giving set and the non-behavior-giving one were analyzed to evaluate model performance and the sensitivity, identifiability and uncertainty of the model parameters.

RESULTS AND DISCUSSION

Model performance

Figure 3 compares the results of a typical behavior-giving simulation and the observed data for residual chlorine, ammonia nitrogen, and TOC. As shown in the figure, the developed model could simulate these three species quite well, especially ammonia nitrogen and TOC. For these two species, eight out of the nine data points had absolute values of the relative errors lower than 50%, and most of them were actually lower than 20%. For residual chlorine, the model could not simulate the first two data points well, whereas all the other data points had absolute values of the relative errors below 35%. Furthermore, as shown in Figure 3, the model could generally capture the unusual surges in the concentrations of residual chlorine in September 2012, ammonia nitrogen in November 2012, and TOC in March 2013.

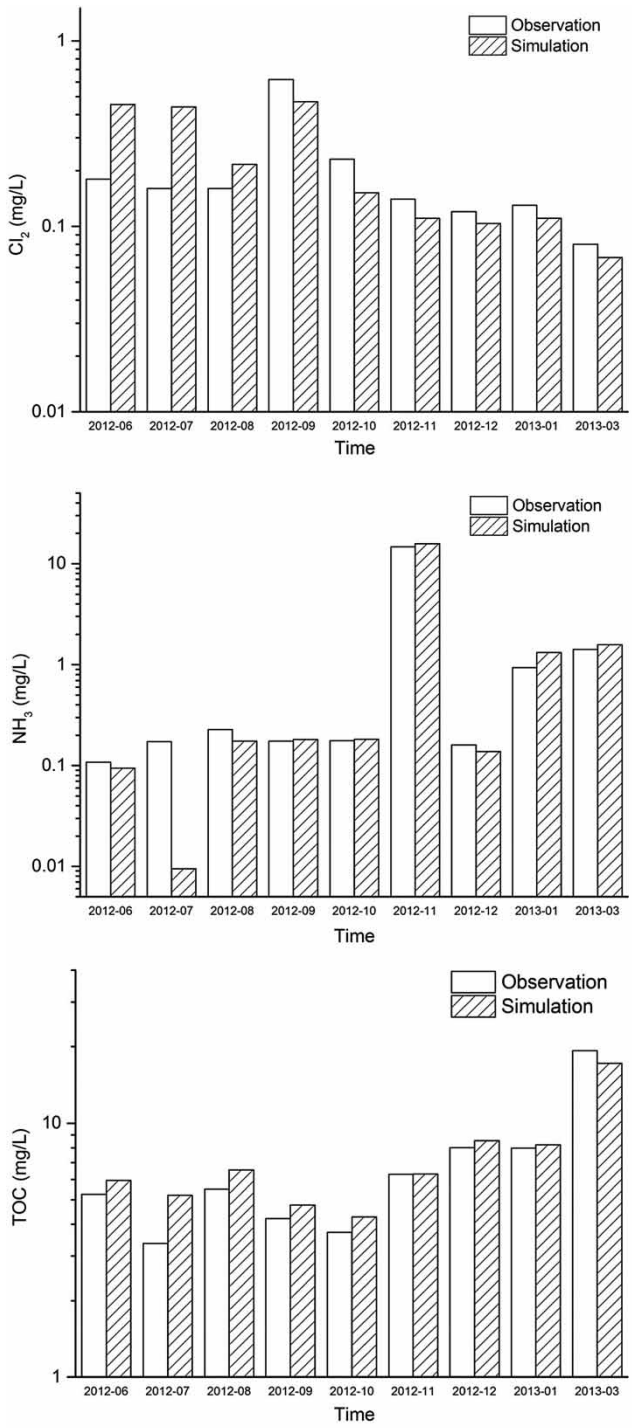

Figure 3 | Simulation results of residual chlorine, ammonia nitrogen, and TOC.

As discussed previously, the K–S test was applied to examine the similarity of the two probability distributions between observed and simulated data. Therefore, in Figure 4,

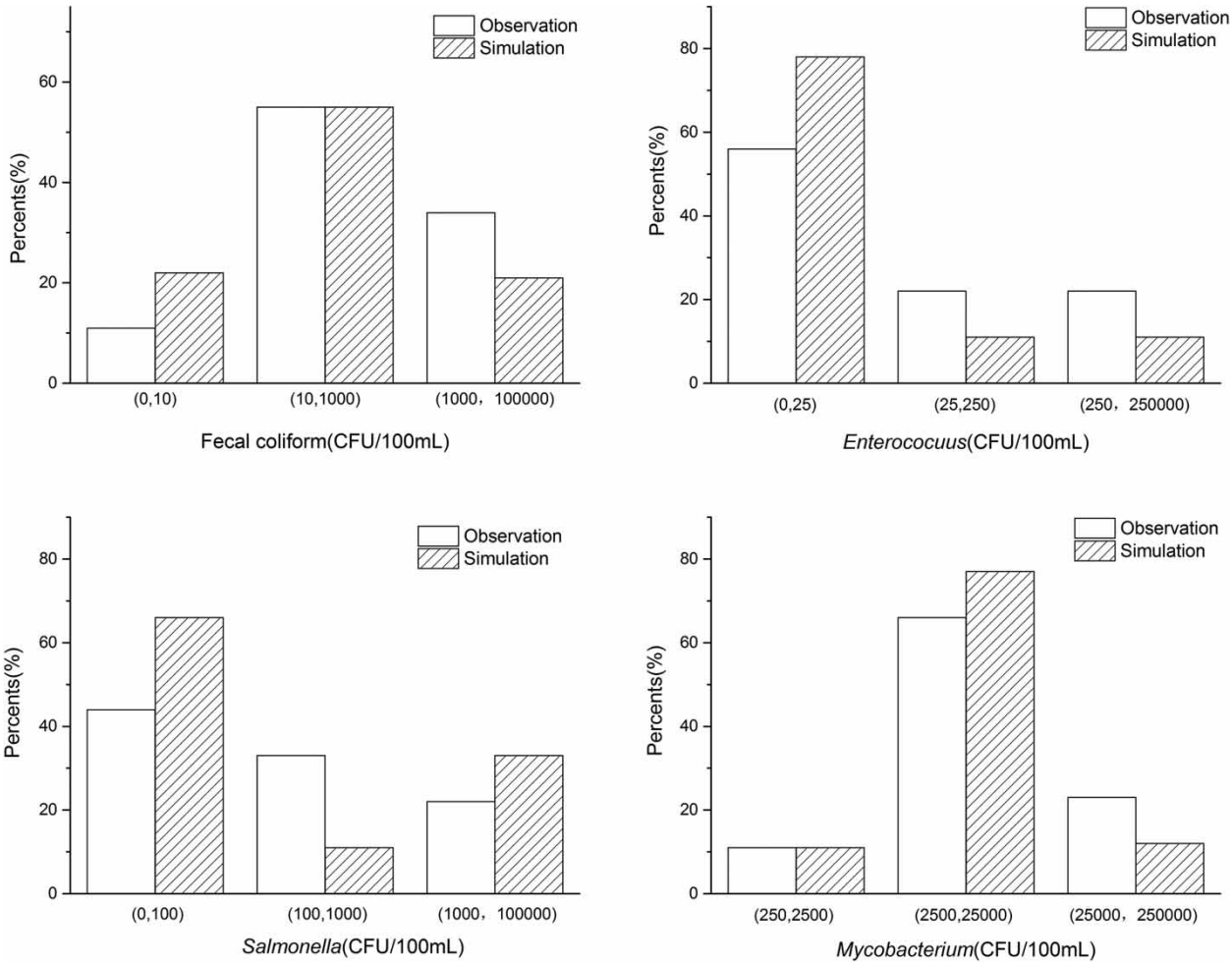

Figure 4 | Simulation results of fecal coliforms, *Enterococcus* spp., *Salmonella* spp., and *Mycobacterium* spp.

the observed and simulated histograms are compared for each of the four microbial species, i.e. fecal coliforms, *Enterococcus* spp., *Salmonella* spp., and *Mycobacterium* spp. Given the small number of water samples, the observed and simulated histograms generally fitted well, and in particular the most frequently occurring intervals in their histograms were exactly reproduced by the model.

Model parameters

The last column in Table 1 shows the identified values for the model parameters that gave the simulation results in Figures 3 and 4. For each parameter and initial value in Table 1, the probability distributions of the behavior-giving set and the non-behavior-giving one were also compared with the K–S test at a 0.01 significance level to analyze its regional sensitivity (Beck 1987; Sun *et al.* 2010). Figure 5 illustrates the results of the regional sensitivity analysis, in which the bars represented the K–S test statistic ($d_{\text{statistic}}$) of each parameter and initial value and the line across the bars represented the critical K–S test value of 1.63 at a significance level of 0.01. Those with $d_{\text{statistic}}$ values greater than 1.63 were therefore judged to be sensitive. The greater the $d_{\text{statistic}}$ value, the more sensitive and identifiable the parameter was. As shown in Figure 5, 33 out of the 44 model parameters and initial values were sensitive in this multi-species water quality model for RWDSs, and the high proportion of sensitive parameters indicated a balanced structure of the developed model (Chen & Beck 1999; Sun *et al.* 2010).

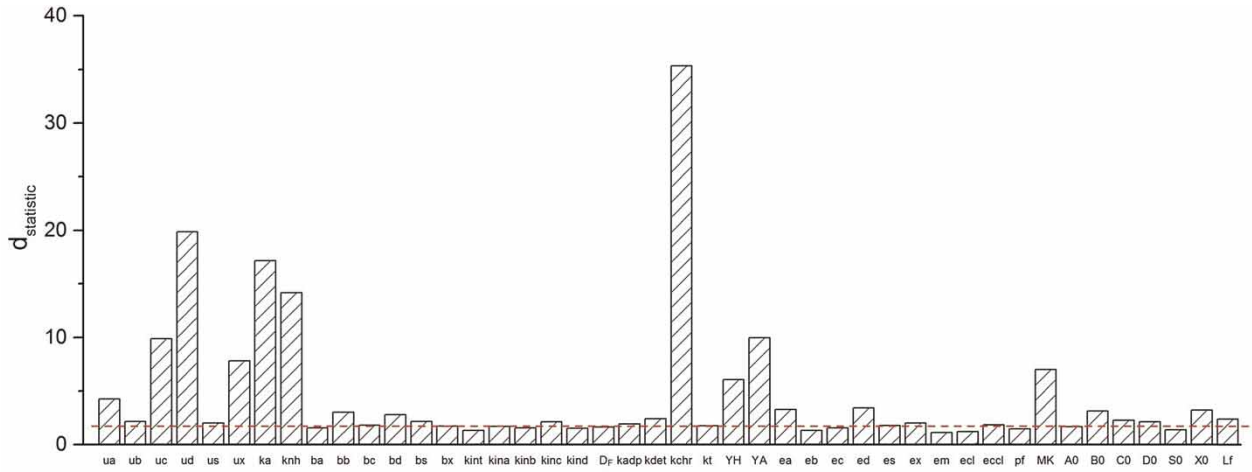

Figure 5 | Results of regional sensitivity analysis of the multi-species model.

Among all the sensitive parameters, kchr was the most sensitive. This implied that the reaction between organic matter and residual chlorine contribute significantly to their decreasing concentrations in the Qinghe RWDS. Except kchr, the other nine most sensitive parameters were all related to microbial growth, which agreed with the observation that all the four species of micro-organisms grew noticeably in the RWDS. For example, ud and uc represented the maximum specific growth rate of *Mycobacterium* spp. and *Salmonella* spp., and were ranked as the second and sixth most sensitive parameters, respectively. This could result from the continuous detection and remarkable growth of these two microbial species in the RWDS. So the results from the regional sensitivity analysis of model parameters could be reasonably interpreted by the current knowledge or observed facts about the system.

Finally, model parameter uncertainty was examined by the standard deviation of the behavior-giving set. The uncertainties of the 10 most sensitive parameters were reduced by 5.6–72.6% as compared with their initial ranges listed in Table 1. Furthermore, the more sensitive the parameter, the more the parameter uncertainty reduced. For example, the uncertainties associated with the first and second most sensitive parameters, i.e. kchr and ud, were reduced by 72.6 and 30.9% respectively. Significant reduction in the uncertainties of the most sensitive parameters could increase the trustworthiness of the model.

DISCUSSION

Given the limited monitoring data and the gross uncertainty involved, the proposed multi-species water quality model for RWDSs could still provide acceptable simulation results in the full-scale Qinghe RWDS, and furthermore the results regarding the sensitivity, identifiability, and uncertainty of model parameters all favored the trustworthiness of the model. However, opportunities exist to improve the performance of the model if a well-designed monitoring program could be implemented on a full-scale RWDS. On the one hand, other relevant water quality indicators need to be monitored, e.g. temperature, dissolved oxygen, suspended solids, and DBPs, which may not only contribute to better model performance, but also improve the identifiability of model parameters. On the other hand, the hydraulics of RWDS, e.g. flow rate and node pressure, should be monitored simultaneously with water quality. Owing to the lack of measurement of the flow rate, the travel time of reclaimed water in this study was estimated with the average flow rate, and this could cause significant errors when the actual flow rates during water sampling deviated from the average. In addition, the simulated part of the Qinghe RWDS in this study only consisted of a main pipe, and the simple hydraulic conditions facilitated sole focus on the multi-species water quality model during parameter identification. However, for application purposes, the model should also be validated against looped

networks where hydraulic modeling and water quality modeling need to be coupled.

The proposed multi-species water quality model for RWDSs could be used to assess and manage human health risk associated with wastewater reuse. As discussed in the introduction section, current water quality standards for reclaimed water usually regulate the effluent of RWTPs rather than the point of use, where water quality could have deteriorated during transport in the RWDSs. With the developed model, human health risk or the violation risk of water quality against relevant standards at the point of use could be predicted. Furthermore, water quality requirements on the effluent of RWTPs could be inversely derived based on the water safety at the point of use. These requirements could be used to develop a comprehensive Water Safety Plan or a hazard analysis and critical control point system for the whole reclaimed water treatment and distribution systems to monitor and manage the sources of health risk along the whole process.

CONCLUSION

A multi-species water quality model for RWDSs was proposed in this paper. Organics, ammonia nitrogen, residual chlorine, inert particles, and six microbial species, i.e. fecal coliforms, *Enterococcus* spp., *Salmonella* spp., *Mycobacterium* spp., other heterotrophic bacteria, and autotrophic bacteria, in both the bulk liquid and the biofilm were simulated. Altogether, 56 physical, chemical, and biological reaction processes were simulated in the model, and 37 model parameters and seven initial values needed to be identified. The developed model was identified against the field monitoring data from part of the full-scale Qinghe RWDS with the HSY algorithm based on a LHS method.

Despite the limitation of the available data from the Qinghe RWDS, the proposed multi-species water quality model could simulate ammonia nitrogen, residual chlorine, TOC, and four microbial species, i.e. fecal coliforms, *Enterococcus* spp., *Salmonella* spp., *Mycobacterium* spp., with acceptable accuracy. Regional sensitivity analysis suggested that the model has a balanced structure with a large proportion of sensitive parameters, and the sensitivity of model parameters could be reasonably interpreted by the current knowledge or observation. Furthermore, the most sensitive model parameters could generally be well identified with uncertainties significantly reduced, which also favored the trustworthiness of the model. However, the model could be improved by implementing a well-designed field monitoring program on full-scale RWDSs, and the validated model could be used for human health risk assessment at the point of wastewater reuse as well as risk management over the whole reclaimed water treatment and distribution systems.

REFERENCES

Arminski, K., Zubowicz, T. & Brdys, M. 2013 A biochemical multi-species quality model of a drinking water distribution system for simulation and design. *Int. J. Appl. Math. Comp. Sci.* **23** (3), 571–585.

Beck, M. B. 1987 Water quality modeling: a review of the analysis of uncertainty. *Water Resour. Res.* **23** (8), 1393–1442.

Bois, F. Y., Fahmy, T., Block, J. & Gatel, D. 1997 Dynamic modeling of bacteria in a pilot drinking-water distribution system. *Water Res.* **31** (12), 3146–3156.

Chen, J. & Beck, M. B. 1999 Quality assurance of multi-media model for predictive screening tasks. Report EPA/600/R-98/106. Office of Research and Development, EPA Washington, DC, USA.

Chi, H. 2010 *Study on Water Quality Model of Water Distribution Network of Non-traditional Water Source*. PhD Thesis. Tianjin University, Tianjin (in Chinese).

Diaper, C., Dixon, A., Butler, D., Fewkes, A., Parsons, S. A., Strathern, M., Stephenson, T. & Strutt, J. 2001 Small scale water recycling systems – risk assessment and modelling. *Water Sci. Technol.* **43** (10), 83–90.

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

Henze, M., Gujer, W., Mino, T. & van Loosdrecht, M. 2000 *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*. IWA Publishing, London.

Jegatheesan, V., Kastl, G., Fisher, I., Angles, M. & Chandy, J. 2000 Modelling biofilm growth and disinfectant decay in drinking water. *Water Sci. Technol.* **41** (4), 339–345.

Jjemba, P. K., Weinrich, L. A., Cheng, W., Giraldo, E. & LeChevallier, M. W. 2010 Regrowth of potential opportunistic pathogens and algae in reclaimed-water distribution systems. *Appl. Environ. Microbiol.* **76** (13), 4169–4178.

Lin, Y. 2013 *Regrowth of Viable but Nonculturable Pathogens and Risk Assessment in Reclaimed Water Systems*. Master Thesis. Tsinghua University, Beijing.

Lu, C., Biswas, P. & Clark, R. M. 1995 Simultaneous transport of substrates, disinfectants and microorganisms in water pipes. *Water Res.* **29** (3), 881–894.

Munavalli, G. R. & Mohan Kumar, M. S. 2004 Dynamic simulation of multicomponent reaction transport in water distribution systems. *Water Res.* **38** (8), 1971–1988.

Narasimhan, R., Brereton, J., Abbaszadegan, M., Ryu, H., Butterfield, P., Thompson, K. & Werth, H. 2005 *Characterizing Microbial Water Quality in Reclaimed Water Distribution Systems*. AWWA Research Foundation, Denver.

Petersen, E. E. 1965 *Chemical Reaction Analysis*. Prentice-Hall, Englewood Cliffs.

Ryu, H., Alum, A. & Abbaszadegan, M. 2005 Microbial characterization and population changes in nonpotable reclaimed water distribution systems. *Environ. Sci. Technol.* **39** (22), 8600–8605.

Schwartz, R., Lahav, O. & Ostfeld, A. 2014 Integrated hydraulic and organophosphate pesticide injection simulations for enhancing event detection in water distribution systems. *Water Res.* **63**, 271–284.

Seyoum, A. G., Tanyimboh, T. T. & Siew, C. 2013 Assessment of water quality modelling capabilities of EPANET multiple species and pressure-dependent extension models. *Water Sci. Technol. Water Suppl.* **13** (4), 1161–1166.

Shang, F., Uber, J. G. & Rossman, L. 2008 *EPANET Multi-species extension software and user's manual*. EPA/600/S-07/021. U.S. Environmental Protection Agency, Washington, DC, USA.

Shu, S. 2011 Direct and indirect modeling of bacterial regrowth in water distribution system. *J. Qingdao Technol. Univ.* **32** (2), 11–17 (in Chinese).

Sun, F. 2007 *Simulation and Management of Water Quality Risk in a Water Supply System*. PhD Thesis. Tsinghua University Beijing (in Chinese).

Sun, F., Chen, J., Tong, Q. & Zeng, S. 2010 Structure validation of an integrated waterworks model for trihalomethanes simulation by applying regional sensitivity analysis. *Sci. Total Environ.* **408** (8), 1992–2001.

Sun, F., Chen, J. & Zeng, S. 2008 Development and application of a multi-species water quality model for water distribution systems with EPANET-MSX. *Environ. Sci.* **29** (12), 3360–3367 (in Chinese).

Takahashi, S. 2012 Water cycle analysis system dynamics model for designing optimal water reclamation scheduling. In *Proceedings of the 30th International Conference of the System Dynamics Society, St. Gallen, Switzerland, 22–26 July*.

Thayanukul, P., Kurisu, F., Kasuga, I. & Furumai, H. 2013 Evaluation of microbial regrowth potential by assimilable organic carbon in various reclaimed water and distribution systems. *Water Res.* **47** (1), 225–232.

Walski, T. M., Chase, D. V., Savic, D. A., Grayman, W., Beckwith, S. & Koelle, E. 2003 *Advanced Water Distribution Modeling and Management*. Haestead Press, Waterbury.

Xin, K., Jia, H., Wei, W., Wang, J. & Liu, J. 2007 GIS-based reclaimed water network planning: a case study in Beijing urban area. *China Water Wastewater* **23** (5), 69–72 (in Chinese).

Zhang, W., Miller, C. T. & DiGiano, F. A. 2004 Bacterial regrowth model for water distribution systems incorporating alternating split-operator solution technique. *J. Environ. Eng. ASCE* **130** (9), 932–941.

First received 23 November 2014; accepted in revised form 4 February 2015. Available online 18 March 2015