Sociotechnical risk assessment for water distribution system contamination threats

Amin Rasekh, M. Ehsan Shafiee, Emily Zechman and Kelly Brumbelow

ABSTRACT

Water distribution systems (WDS) are vulnerable to contaminants, and systematic risk assessment can provide valuable information for assisting threat management. Contamination events are sociotechnical systems, in which the interactions among consumers and water infrastructure may generate unpredicted public health consequences. This research develops a sociotechnical risk assessment framework that simulates the dynamics of a contamination event by coupling an agent-based modeling (ABM) framework with Monte Carlo simulation (MCS), genetic algorithm (GA) optimization, and a multi-objective GA. The ABM framework couples WDS simulation with agents to represent consumers in a virtual city. MCS is applied to estimate the uncertainty in human exposure, based on probabilistic models of event attributes. A GA approach is used to identify critical contamination events by maximizing risk, and a multi-objective approach explores the trade-off between consequence and occurrence probabilities. Results that are obtained using the sociotechnical approach are compared with results obtained using a conventional engineering model. The sociotechnical approach removes assumptions that have been used in engineering analysis about the static, homogeneous, and stationary behaviors of consumers, and results demonstrate new insight about the impacts of these actions and interactions on the public health consequences of contamination events.

Key words | complex adaptive system, contamination, evolutionary algorithm, risk assessment, sociotechnical, water distribution system

Amin Rasekh (corresponding author)
Kelly Brumbelow
Zachry Department of Civil Engineering,
Texas A&M University,
3136 TAMU, College Station,
TX 77843-3136,
USA
E-mail: *aminrasekh@tamu.edu*

M. Ehsan Shafiee
Emily Zechman
Department of Civil, Construction,
and Environmental Engineering,
North Carolina State University,
C.B. 7908,
Raleigh, NC 27695,
USA

INTRODUCTION

Water distribution systems (WDS) are critical infrastructure systems that are vulnerable to contamination events, which occur when a chemical contaminant or pathogen is accidentally or intentionally introduced to a pipe system. Contamination events endanger public health, erode community trust in WDS serviceability, and interrupt non-consumptive uses, such as firefighting. Water utilities are responsible for managing WDS contamination threats and are required to allocate resources to protect a community from probable events under the Public Health Security and Bioterrorism Response Act of 2002 (United States Congress 2002). Threat management activities include hardening infrastructure components, responding to events, and restoring infrastructure service after an event has occurred (Lindell *et al.* 2006; Haimes 2009). Risk assessment is an important initial step that should be conducted to guide analysis and decision-making required for threat management. WDS contamination risk assessment is a systematic process to identify critical contamination scenarios and estimate the expected risk, which is the product of the consequences and occurrence probability associated with potential scenarios (Kaplan & Garrik 1981; Rose *et al.* 1991; Rasekh & Brumbelow 2013).

Risk can be assessed through a number of methodologies that are based on simulation and optimization. Monte Carlo simulation (MCS) generates a large number of possible intrusion events through randomly sampling contamination scenario attributes and estimating the values for

doi: 10.2166/hydro.2013.023

health consequences through hydraulic simulation (Nilsson et al. 2005; Khanal et al. 2006; Torres et al. 2009). Critical contamination scenarios may be identified through extreme-consequence sampling (Perelman & Ostfeld 2010) and through systematic optimization of the components of risk, which are likelihood and consequences (Rasekh & Brumbelow 2013).

During a contamination event, the reactions and behaviors of a population of consumers may generate public health outcomes that are not predicted by an engineering modeling approach. Consumers may change their water demands from typical usage patterns, as they become aware of a contaminant through symptoms of exposure, warnings from peers, and alerts from a local utility. As consumers obtain more information about an event, they may reduce their water activities and influence other consumers to also reduce their consumption. Significant shifts in consumer demand can feed back into the network hydraulics and alter the predicted public health consequences of an event. New hydraulic conditions alter the spatial and temporal characteristics of the contaminant plume, exposing a different set of consumers than originally predicted. Feedback between the infrastructure and social systems may significantly influence the emergent public health consequences of a water contamination event, creating a sociotechnical system and requiring new analytical methods (Glouberman 2001; Woo & Vicente 2003; Vicente & Christoffersen 2006; Rasekh et al. 2010; Zechman et al. 2011).

A new sociotechnical risk assessment methodology is developed and demonstrated here. Sociotechnical risk assessment accounts for the influence of feedback mechanisms among consumers and the WDS during an emergency to identify critical contamination event characteristics. Existing methods for risk assessment are based on assumptions that the demands exerted during an event are consistent with those exerted during normal operating conditions. The influence of changes in consumer behaviors may introduce error in the predictions of an engineering model (Shang et al. 2006), and as a consequence, the most vulnerable regions in the network may differ from those identified through engineering analysis. Sociotechnical risk assessment uses a new simulation methodology that was developed to simulate the feedbacks among the adaptive and responsive actions of consumers and the WDS in a contamination event (Zechman 2011). This methodology is based on Complex Adaptive System (CAS) simulation, which is a technique for simulating a large number of interacting actors to holistically evaluate the emergent properties of a system (Holland 1995; Miller & Page 2007). The CAS approach couples agent-based modeling (ABM) with a water distribution model to evaluate the number of exposed consumers when the changes in water activities are considered. The sociotechnical risk assessment methodology couples MCS and evolutionary computation-based optimization methods with the ABM framework to evaluate public health impacts and characterize extreme-risk accidental contamination events. The integrated framework is evaluated and demonstrated for a virtual mid-sized city. Results demonstrate that using an engineering model for risk assessment may focus mitigation efforts on high consequence, low-probability events. By using an ABM framework that incorporates consumer dynamics, values for exposures are predicted at lower levels than the engineering model predicts, and critical contamination events that are identified have a higher relative probability of occurrence. These insights can be used to direct the attention of a utility manager in selecting critical contamination events to determine resource allocation for emergency response.

SOCIOTECHNICAL RISK ASSESSMENT FOR WDS MANAGEMENT

The sociotechnical risk assessment methodology that is described here integrates recent developments in sociotechnical WDS simulation and the state-of-the-art methods for characterizing risk and critical contamination events. A sociotechnical modeling methodology was developed by coupling ABM with a hydraulic simulation model for assessing the impacts of consumer behaviors on public health in a contamination event (Zechman 2011). During a contamination event, consumers may move among contaminated and clean nodes, adapt their demands based on warnings, and communicate about an event. Readily available water distribution models do not include capabilities to simulate these interactions and, instead, assume that consumer demands are homogeneous, static, and stationary. Engineering assumptions do not reflect reality, but they are adequate for selecting engineering planning decisions in typical operating conditions. In a contamination event, however, these assumptions may not be adequate and may lead to a misapproximation of consumer health impacts.

To create an ABM framework of the contamination event, each individual is represented as an agent; each agent receives messages from a node in the water network and other agents; agents reduce demands by a percentage of their total use; and demand reductions are translated as new input to the water distribution network model. The ABM framework was applied as a proof-of-concept for two simple networks described as tutorials in the EPANET User's Guide (Rossman 2000), including Net3 (Zechman 2007) and Net1 (Shafiee & Zechman 2010). Zechman (2011) developed the framework using realistic estimations of the timing and volume of ingested water for individuals. The framework was applied for the city of Micropolis (Brumbelow *et al.* 2007), which is an all-pipe network model for a virtual community of approximately 5000 residents. The numbers of exposed consumers were compared for variations in the timeliness of reducing water demands, and the work demonstrated that the model could be used to evaluate a utility's response actions, including flushing water through hydrants and broadcasting warnings. Shafiee & Zechman (2013) further developed the framework for Mesopolis (Johnson & Brumbelow 2008), which is a skeletonized network model for a virtual city of 150,000 residents. For Mesopolis, the ABM framework was extended to more realistically represent the volume and likelihood of demand reduction using data from surveys about the responses of individuals to a water emergency. This work analyzed the extent to which hydraulics were changed in the network due to consumer response. For an event that introduces a large dose of a potent chemical (arsenic), the predictions of the location and timing of a contaminant plume differ significantly from engineering predictions when adaptive demands are included in the model, and hydraulic conditions are altered beyond normal operating conditions. The ABM framework was extended to include an optimization module for effectively selecting and manipulating hydrants that should be flushed (Zechman 2010, 2013; Shafiee & Zechman 2012) and for routing emergency vehicles to efficiently warn consumers (Shafiee & Zechman 2011). These studies compared the performance of management strategies predicted by models that include consumer behaviors with models that neglect consumer behaviors.

The previous studies analyzed the difference in public health and hydraulics as predicted using a sociotechnical approach and an engineering approach for a small set of potential contamination events; however, the contamination events were created and selected using engineering judgment and limited preliminary analysis. A sociotechnical risk assessment provides a new methodology to provide better guidance on selecting contamination events for developing management strategies, by exploring how the actions and interactions of consumers influence the vulnerability of the WDS and the characteristics of high-risk contamination events. Advanced risk assessment methodologies, including MCS, genetic algorithm (GA)-based optimization, and multi-objective optimization, are coupled with the ABM framework in this research. While previous studies explored the impacts of a small number of events on consumer health, the application of the sociotechnical risk assessment approach simulates thousands of events to determine: (1) the range of consequences that could be expected due to a contamination event; (2) the timing, location, and contaminant type of the worst-case events that could occur; and (3) the trade-off in consequences and likelihood for the worst-case events. This research explores and quantifies the difference between risk assessment results generated using a sociotechnical approach with those calculated using a pipe network model alone. This research can assist decision-makers in determining the importance of insight provided by sociotechnical approaches in selecting events for planning response strategies.

SOCIOTECHNICAL RISK ASSESSMENT METHODOLOGY

This study develops a methodology for sociotechnical risk assessment by coupling an ABM framework that simulates water distribution contamination events with risk assessment algorithms, including MCS, GA-based single-objective optimization, and GA-based multi-objective optimization (Figure 1). The ABM evaluates the exposure for water contamination events that are generated by risk

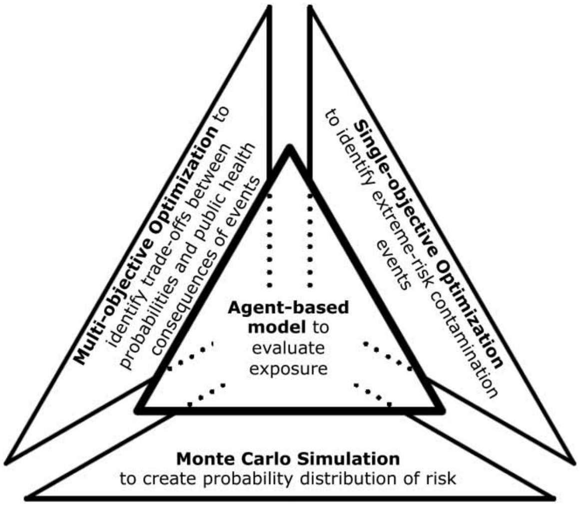

Figure 1 | Sociotechnical risk assessment framework. The framework is applied to estimate public health consequences and characterize critical contamination scenarios in a WDS.

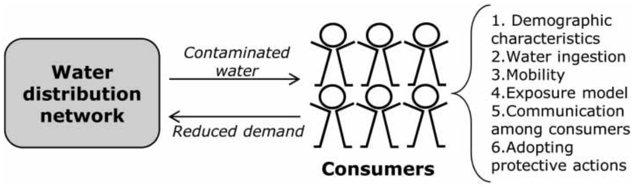

Figure 2 | ABM framework for a WDS contamination event. Each consumer is simulated as an agent with attributes and behaviors as shown.

assessment algorithms. MCS generates a probability distribution for the public health impacts of contamination events. The single-objective optimization identifies critical contamination events with the highest values for risk, while the multi-objective optimization identifies and explores the trade-offs between the likelihood of events and the public health consequences of events. Abbreviated descriptions are given below of each module, and the reader is referred to related documents in the following sections for complete descriptions.

ABM for simulation of WDS contamination events

ABM is a computational method that is used for simulating a CAS (Axelrod 1997). ABM simulates a system as a set of aggregated autonomous agents that receives information from other agents and its environment and selects actions based on a set of rules. ABM provides the analysis capabilities to assess the emergence of system-level properties that result from complex nonlinear interactions among decentralized agents.

An ABM framework models a sociotechnical water distribution contamination event and evaluates public health consequences measured as the number of exposed consumers (shown in Figure 2). The framework couples consumer agent models with a WDS model. Consumer agents are modeled using a set of attributes and rules to specify behaviors (listed in Figure 2 and described below). At each simulation time step, the WDS model passes to consumer agents water quality data, which are used to calculate each agent's level of exposure to the contaminant based on the concentration value and the volume of water that is ingested. Consumers exert demands at different locations within a city as they travel to and from business and residential districts. A consumer agent ingests water at five ingestion events during the day and changes its water use when it consumes a critical mass of contaminant, which is calculated by the exposure model. A consumer agent sends a message to other agents once it reduces water usage. Consumers receive messages from other agents and reduce their water demand along with those who are exposed. New demand patterns feed back to the WDS model and alter hydraulic conditions. The ABM is implemented by coupling a dynamic modeling system, AnyLogic (XJ Technology 2010) with a hydraulic simulator, EPANET (Rossman 2000). Rules and attributes of the consumer agents are described briefly in the following sections, and complete descriptions are provided by Zechman (2011) and Shafiee & Zechman (2013).

Demographic characteristics

Each consumer agent is assigned a value for weight, age, gender, and employment status. Demographic attributes are generated probabilistically, based on available data (United States Environmental Protection Agency 2000), to represent the statistical distribution of the characteristics of the national population. The Bureau of Labor Statics of the US Department of Labor (2011) provides employment rates across age groups.

Water ingestion

Each consumer agent is encoded with a rule that specifies the timing for five ingestion events each day. Davis & Janke (2008, 2009) developed a model to approximate the amount of tap water a consumer ingests per day and the timing of five ingestion events. The timing model uses probability distributions to generate the times at which three major meals are taken, along with the timing of two minor meals, which occur mid-way between major meals. The amount of water a consumer agent consumes each day is based on gender and weight (United States Environmental Protection Agency 2000). The total volume of water is divided evenly among the five ingestion events.

Mobility

Each consumer agent is assigned a residential node, a non-residential node, and timing for leaving residential and non-residential nodes. Agents that are tagged as 'employed' are assigned a non-residential node at commercial and industrial nodes. The diurnal demand pattern that is provided as input for the hydraulic network model is used to determine the number of agents who should be located at each node and time step. Mobility can affect the exposure of an agent, as it travels among different sections of the city and visits nodes that deliver contaminated water.

Exposure model

An exposure model is used to estimate the exposure of each agent, based on the volume of ingested water, the concentration of the contaminant, and the body weight (Ayotte *et al.* 2006). The amount of contaminant ingested by each consumer agent is an accumulative value, which is calculated as

$$M_c = \sum_{i=1}^{n} V_i \times Q_i \tag{1}$$

where M_c = mass of contaminant consumed, n = number of ingestion events, V_i = volume of water consumed at ingestion event i in liters, and Q_i = concentration of contaminant in water at ingestion event i in mg/liter. The agent is flagged as 'exposed' once the mass in the agent's body exceeds a specified dose. The critical dose can be represented as an infectious dose for pathogens (White 1999), to represent that the agent would experience symptoms of illness. The analysis presented here explores risk assessment for pathogens, and the approximate number of pathogen cells in an infectious dose for each pathogen is shown in Table 1.

Communication among consumers

During a WDS contamination event, a consumer communicates with peers and colleagues based on actual and perceived risks. Lindell & Perry (2004) developed a communication model that is based loosely on the small world network model (Watts 1999) to simulate the unidirectional flow of information within a cluster of peers via word-of-mouth during an emergency. Each consumer agent is assigned a cluster, which represents the collection of peers and family members. Consumer agents are labeled as original source, intermediate receivers, ultimate receivers, or information isolates. Ultimate receivers can receive warning messages from the original source or intermediate receivers, and intermediate receivers can receive messages from the original source. Once an agent receives a message or becomes exposed, it passes a message to receiving agents after one time step has passed.

Table 1 | Statistical data for pathogens (Rasekh & Brumbelow 2013)

Pathogen type	Occurrence probability (%)	Infectious dose (cells)	Expected IDPC [a]
Giardia	20.6	10 to 100[b]	195
E. coli	15.9	1.0 E + 08[c]	131
C. jejuni	20.6	500[c]	197
Norwalk-like virus	15.9	100[d]	239
Cryptosporidium	27.0	10[c]	IDPC quartiles: 2%: 0.29, 25%: 1.0, 50%: 2.3, 75%: 87, 98%: 412

[a]IDPC = infectious doses per capita.
[b]Roxström-Lindquist *et al.* (2006).
[c]Kothary & Babu (2001).
[d]Centers for Disease Control and Prevention (2001).

Adopting protective actions

Agents become aware of contamination events as they are alerted by peers (described above in 'Communication among consumers') or once they have consumed enough contaminant to experience symptoms (described above in 'Exposure'). At the same time step that an agent is alerted via either pathway, it takes protective actions to safeguard its health. The effect of these protective actions is modeled as a reduction in the hourly demand of consumer agents, based on typical demands exerted through the use of water appliances. Mayer *et al.* (1999) categorized indoor end-uses for US households, including washing clothes (22% of total water demand), toilet (26%), shower (17%), faucet (16%), leakages (14%) and other miscellaneous uses (5%). Lindell *et al.* (2010) fielded a survey to determine how respondents would reduce water activities due to the potential contamination of tap water. Based on survey results, each consumer agent is assigned a set of values to represent the probability of suspending a water activity (e.g., washing clothes, washing kitchen, taking a shower, washing hands, rinsing mouth, washing dishes, brewing coffee, cooking, rinsing fresh vegetables), which leads to a reduction in the demand of a consumer agent. An agent selects to suspend water activities once it receives a message from another agent or once the agent becomes exposed. For example, a consumer that is aware of an event has a 38.4% probability of suspending its use of tap water for taking showers, which reduces its demand by 11.7%. Probabilities for all water activities and analysis of survey results are provided by Shafiee & Zechman (2013).

MCS for characterizing consequences

MCS is a numerical procedure for generating a probability distribution of system response based on estimated uncertainties in system inputs. MCS executes a large number of random realizations of input variables, which are contamination event characteristics, and simulates the response, or consumer exposure, using a system model. MCS and sensitivity studies have been conducted to estimate the likely health effects from WDS contamination events, which provides insight to the variability of exposure levels across various values for event characteristics (Uber *et al.* 2004; Nilsson *et al.* 2005; Khanal *et al.* 2006; Torres *et al.* 2009; Pasha & Lansey 2010; Davis & Janke 2011). In simulating a WDS contamination scenario, a realization of random variables for the pathogen type, the pathogen load, the intrusion location, the water demand multiplier, the time of day, and the intrusion duration, are generated using the distributions described below. The system response of interest is the number of exposed consumers, which is the output of the ABM simulation.

Statistical inference to characterize contamination event attributes

Documentation and meta-analysis of past contamination events is needed to determine the likelihood of contamination event characteristics. Data about historic accidental contamination outbreaks have been collected by agencies (e.g., the Centers for Disease Control and Prevention) and scholars (e.g., Hrudey & Hrudey 2004, 2007), and these data have been analyzed to provide quantitative information on expected contamination event patterns and characteristics (Reynolds *et al.* 2008; Rasekh & Brumbelow 2013). Statistical inference was conducted to determine the probability distributions that describe contamination event characteristics. Rasekh & Brumbelow (2013) analyzed 70 historic accidental contamination outbreaks collected by Hrudey & Hrudey (2004, 2007) and used statistical inference to construct a set of probability distributions for scenario attributes. Each contamination scenario is defined by a set of attributes, including the contaminant intrusion location (L), the time of day the contamination begins (T), and the intrusion duration (ΔT), contaminant type (C), contaminant loading (M), and the time of year contamination occurs (represented by a surrogate WDS-wide demand multiplier (D)).

Analysis of historic events reveals that 89% of documented events originated at water treatment plants. To generate a value for L, the location is selected uniformly from all water treatment plants in a network. The time of day (T) that a contaminant is first introduced to the network and the duration of the intrusion (ΔT) are subject to high uncertainty (Hrudey & Hrudey 2004; Bristow & Brumbelow 2006), and a uniform distribution is used to generate these attributes. Five pathogens appeared most frequently in the historic dataset: *Giardia lamblia*, *Escherichia coli*, *Campylobacter jejuni*, *Cryptosporidium*, and Norwalk-like virus. Occurrence probabilities for the five pathogens (C) were calculated based on their occurrence frequency in the compiled data. An exponential distribution was fit to describe the amount of contaminant (M) introduced into the WDS, which is measured as the number of infectious doses per capita (IDPC). Frequency and expected values of IDPC for the five pathogens are reported in Table 1.

Contamination events can create diverse public health consequences based on the time of year at which they occur. This is because water demands fluctuate among seasons, as consumers use more water in summer months, and less water in winter months. The time of year that a contamination occurs is represented in the modeling as seasonal variations in demand through a WDS-wide demand multiplier. A shifted gamma distribution was fit to describe the demand multiplier, using real data obtained from New York City for 1982 (Protopapas *et al.* 2000) and an unpublished data set for a water utility in Texas. Probability distributions for all attributes are summarized in Table 2.

Optimization-based approaches to identify critical events

Optimization methods can be applied to identify a small set of critical contamination scenarios, which occur at maximum values of likelihood and consequences (Kaplan & Garrik 1981). Perelman & Ostfeld (2010) proposed a method to sample and identify a set of events with small probabilities of occurrence and extreme impacts. Rasekh & Brumbelow (2013) identified critical events by maximizing risk, defined as the product of likelihood and consequences, using a GA-based approach. Rasekh & Brumbelow (2013) further demonstrated the utility of applying multi-objective optimization for risk assessment. While events of maximum risk (the product of consequences and probability of occurrence) can be considered for placing sensors and planning emergency response, decision-makers may be interested in a range of events that could occur, including rare events with extreme consequences and more likely events with low consequences. By considering a diverse ensemble of critical events, instead of one event alone, decision-makers can explore and enhance the robustness and reliability of risk mitigation and emergency response plans. Multi-objective methods can be applied to generate a trade-off relationship between likelihood and consequences and provide additional information for decision-making.

Table 2 | Probability distributions for contamination scenario attributes

Scenario attributes	Probability distribution	Possible values
Intrusion location	Uniform	{West WTP[a], East WTP}
Contaminant type	Experimental	5 pathogens[b]
Contaminant amount	Exponential	{0.00, 0.01, …, 1.00}[c]
Demand multiplier	Shifted gamma	{0.600, 0.625, …, 2.300}
Intrusion start time	Uniform	{00:00, 01:00, …, 23:00}
Intrusion duration	Uniform	{24, 25, …, 96}

[a]WTP: water treatment plant.
[b]Listed in Table 1.
[c]Normalized by dividing by the 98th percentile.

Single-objective optimization

The identification of a critical event is posed as an optimization model, where the product of the likelihood and consequences defines risk, R_p, which should be maximized (Kaplan & Garrik 1981). Multiplication of likelihood and consequences to construct a single measure of expected damages is a conventional approach and has been extensively applied for risk assessment in water resources engineering and many other disciplines (Tung *et al.* 2006). The scenario likelihood, p, is defined here as a joint probability of multiple random variables, which represent event attributes (Ang & Tang 2007):

$$p = P(L = l, C = c, M = m, D = d, T = t, \Delta T = \delta t) \tag{2}$$

where l, c, m, d, t, and δt represent specific values for random scenario attributes, including the contaminant intrusion location, contaminant type, contaminant loading, the WDS-wide demand multiplier, the time of day the contamination begins, and the intrusion duration, respectively. Statistical independence is assumed among all scenario attributes, except for the contaminant type and amount, and Equation (2) is reformulated as

$$p = P(L = l) \times P(C = c) \times P(M|C = m|c) \times P(D = d) \\ \times P(T = t) \times P(\Delta T = \delta t) \quad (3)$$

where $P(M|C = m|c)$ denotes the probability that $M = m$ given $C = c$. Because some scenario attributes are continuous random variables (e.g., intrusion duration), the probability functions are discretized using discrete variable intervals (probability functions are described in Table 2). The number of all possible scenarios and, consequently, the scenario likelihood, p, depends on the degree of discretization, as each random variable is rounded to the upper or lower bound of a range, to approximate its likelihood. Choosing smaller ranges for discretization can estimate the true probability distribution more accurately. Probability approximations based on the probability functions are used in this study as the occurrence probability of events.

The consequence of an event is quantified here as the number of people that become exposed (N_s) due to a contamination event. As described above, an agent is considered exposed if the amount of ingested contaminant mass exceeds the infectious dose, and the number of exposed consumers is calculated using the ABM framework. Critical contamination events can be identified by solving the following equation:

$$\text{Maximize } R_p = p \times N_s \quad (4)$$

The decision variables are the contamination event characteristics (l, c, m, d, t, and δt, as defined above) that are used to calculate p, the probability of an event, and the number of exposed consumers. Equation (4) represents a highly nonlinear and discrete objective function due to the dynamics of the ABM approach and the complexity of the combination of multiple probability distributions for characterizing scenario attributes. Evolutionary algorithms (Holland 1975) have been used effectively to solve water resources planning and management problems (Nicklow *et al.* 2010). A GA-based approach (Goldberg 1989) is used to identify characteristics of the critical contamination event. The GA mimics natural evolution to converge to a near-optimal solution by generating and selecting new solutions through recombination (crossover), mutation, and selection.

Multi-objective optimization

A multi-objective problem is formulated to maximize the components of likelihood and consequences as separate objectives, as follows:

$$\begin{aligned} &\text{Maximize } Z_1 = p \\ &\text{Maximize } Z_2 = N_s \end{aligned} \quad (5)$$

where Z_i represents the ith optimization function, and the decision variables are the contamination event characteristics, l, c, m, d, t, and δt. Solution of Equation (5) through the use of a multi-objective optimization algorithm will yield a set of non-dominated solutions, which is called the 'maximum-risk frontier'. The maximum-risk frontier is a set of diverse critical scenarios with varying levels of likelihood and consequences (Rasekh and Brumbelow 2013). The optimization problem represented by Equation (5) is solved here using an evolutionary algorithm for multi-objective problems, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb *et al.* 2002). NSGA-II uses a non-dominated sorting strategy, which sorts and ranks solutions based on non-dominance, and a crowding distance, which measures the degree to which solutions are spread uniformly across the non-dominated front. The selection operator chooses solutions based on Pareto-dominance and values for the crowding distance. Crossover and mutation operators as implemented within NSGA-II are similar to archetypical operators used for a GA.

ILLUSTRATIVE CASE STUDY: MESOPOLIS

The risk assessment framework is applied to generate a distribution function of consequences, critical contamination events, and a maximum-risk frontier for Mesopolis, a virtual city. The Mesopolis dataset was developed as a case study for urban infrastructure research (Johnson & Brumbelow 2008). Mesopolis is simulated with a population of 146,716. Land use in Mesopolis is comprised of residential, commercial, and industrial areas, and within the city limits, there is a naval base, an airport, and a university (Figure 3). Water is withdrawn at an intake located south of the city, from a river that runs north through the center of Mesopolis. A branched pipe delivers raw water to two water treatment plants (WTP), located on opposite sides of the river. The West WTP supplies water to the older sections of Mesopolis, located on the western side of the river, and the East WTP distributes water to the eastern section and, during peak demand periods, to a large portion of the central and western districts. The network is modeled as a skeletonized water network with one reservoir, 1588 nodes, 2058 pipes, 13 tanks, and 65 pumps. Four demand patterns are applied for different nodes based on residential, commercial, industrial, and naval land uses.

Engineering and sociotechnical modeling frameworks

This work compares the sociotechnical model, constructed as the ABM framework, with a conventional engineering model. The engineering model, which is a hydraulic simulator (EPANET), follows a typical engineering approach and does not account for sociotechnical interactions among the consumers and the WDS. Changes in consumer water demands during a contamination event are not considered, and, instead, consumer demands are simulated as static, homogeneous, and aggregated only at residential nodes. The number of consumers at each residential node is calculated by dividing the base demand value at that node by the daily water consumption per consumer, which is assumed as 105 gallons per person per day. Consumers are not simulated at non-residential nodes and do not travel to other nodes during the day. Each consumer ingests 0.93 liters of water per day, distributed uniformly over five ingestion events at 07:00, 09:30, 12:00, 15:00, and 18:00. Both the engineering model and the sociotechnical model are simulated for a 10-day period, and both models are coupled with the MCS and optimization methodologies.

RESULTS

MCS results

MCS is executed for 20,000 and 5,000 random event realizations for the engineering model and the sociotechnical model, respectively. Simulation of a single event requires 15 seconds and 600 seconds for the engineering model and the sociotechnical model, respectively, on a personal desktop computer. The increased simulation time for the ABM framework is due to the extra computation required to simulate the interactions among agents and the WDS.

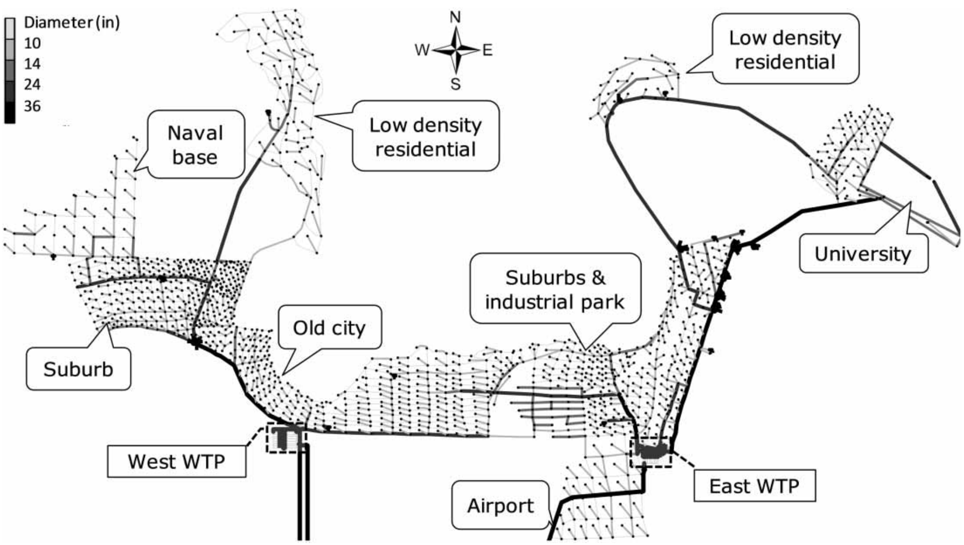

Figure 3 | Mesopolis WDS. Different land uses have distinct water demand patterns in the hydraulic model.

MCS results are used to generate the cumulative distribution function (CDF) for the number of exposed consumers. The CDF obtained using the engineering modeling has a discontinuous step-wise shape, while the CDF estimated using the ABM framework increases monotonically (Figure 4). There is a 'hydraulic barrier', which cuts through the center of the city, and during most hours of the day, the West WTP supplies water to western nodes, with a population of approximately 38,000, and the East WTP supplies water to eastern nodes. During peak demand hours, the East WTP also provides water to western nodes. All events are simulated at one of the two WTPs. As a result of the hydraulic barrier, events at both WTPs can generate consequences of less than 38,000 exposures, while only events at the East WTP can generate greater consequences when contaminated water propagates to the western residential sections during peak demand periods. This discontinuity does not appear in the CDF for the sociotechnical framework, because agents travel across the hydraulic barrier, and events at the West WTP can reach more consumers and create more significant consequences than predicted by the engineering model.

Figure 5 shows that the average exposure for events that originate at the West WTP is higher than the average exposure for events at the East WTP. These results are counter intuitive when compared to the location of the maximum-risk events (East WTP), but again result from dynamics caused by the hydraulic barrier. For events at the West WTP, the contaminant remains in the western region and occurs at high concentrations, even for moderate loads. As a result, a significant portion of the consumers becomes sick. For many events at the East WTP that are initiated with moderate loads, the contaminant disperses throughout the entire network and the contaminant concentrations become diluted, so that only a few consumers are exposed in these events.

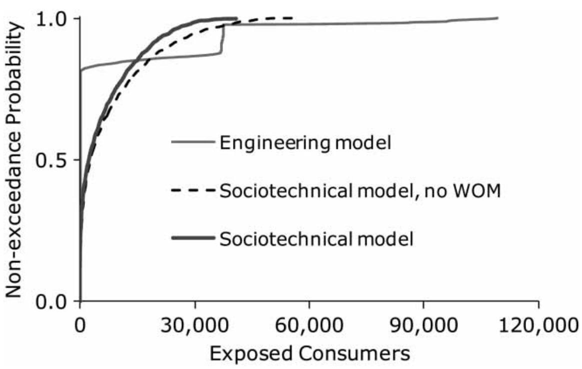

Figure 4 | MCS results for engineering and sociotechnical models. WOM indicates word-of-mouth mechanism.

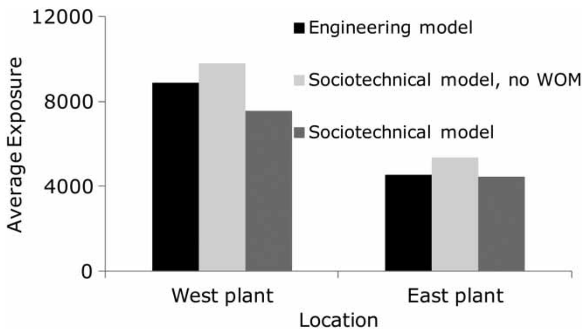

Figure 5 | Average exposure over all MCS realizations separated by the contaminant intrusion site. WOM indicates word-of-mouth mechanism.

The sociotechnical framework predicts a lower number of zero-exposure events than the engineering model predicts. For the engineering model, 81% of the realizations result in zero exposures, while approximately 10% of the realizations results in zero exposures using the sociotechnical model. Because the engineering model does not simulate that consumers travel in the city, consumers are safe when their fixed residential locations are not contaminated. In reality, however, members of a community move within a city during a contamination event, and they may be exposed to contaminated areas even when their places of residence remain uncontaminated. This dynamic is included in the ABM framework through simulation of mobility, which leads to a higher number of non-zero-exposure contamination events predicted by the ABM framework.

Word-of-mouth mechanism

The effect of the word-of-mouth mechanism on the exposure CDF was assessed through further simulations. The word-of-mouth mechanism was disabled, and MCS was again executed using the sociotechnical model for the same 5000 event simulations. By excluding word-of-mouth, the exposure CDF is shifted to reflect an increase in the total number of exposed consumers (Figure 4). The exposure CDF has a continuously increasing shape when the word-of-mouth mechanism is excluded, which indicates that the communication among consumers does not drive the distribution of impacts; instead, it is the heterogeneous drinking patterns and mobility of agents that drive the differences between the CDFs generated by the engineering and sociotechnical models. As shown in Figure 5, average exposure across all events is higher when the word-of-mouth is excluded, compared with the sociotechnical simulation that includes it.

Optimization results

A GA is used to solve the single-objective problem (Equation (4)) for a set of contamination events, including events at the two WTPs for two pathogens, which are simulated using the ABM framework and engineering model separately, for a total of eight optimization scenarios. *Cryptosporidium* and *Giardia* were chosen as the pathogens of study because they produced, on average, higher risk values than the other pathogens, based on the MCS results. NSGA-II is used to solve the multi-objective problem (Equation (5)) for the same set of eight scenarios. Solutions are encoded using a real-valued representation, and the GA is implemented using roulette wheel selection, simulated binary crossover (Deb & Agrawal 1995), and polynomial mutation (Deb 2001). The algorithmic parameter settings for the optimization models are shown in Table 3, and the number of function evaluations for one execution of an optimization algorithm is approximately 1,400 and 4,100 for the GA and NSGA-II approaches, respectively. Because an evolutionary algorithm uses random operators in the search, there is stochasticity in the results, and separate runs of an algorithm may produce solutions that vary in fitness and decision variable values. For each scenario and model, five random trials were run for the single-objective GA to provide some assessment of the variability in solution quality. Because of the computational time required for solution evaluation, the multi-objective algorithm was executed only once for each model and contamination event.

Table 3 | Algorithmic parameter settings for GA and NSGA-II

Parameter	GA	NSGA-II
Population size	40	100
Number of generations	50	50
Tournament size	N/A	3
Crossover type	Simulated binary crossover	Simulated binary crossover
Crossover rate	0.7	0.8
Mutation type	Polynomial	Polynomial
Mutation rate	0.05	0.10

Single-objective optimization results

The objective function value, or risk, for each of the five random GA trials is shown in Table 4. The similarity among results for each contamination event indicates that the GA reliably identifies critical contamination events with similarly high risk. There is an exception when *Cryptosporidium* is introduced into the East WTP. For this contamination scenario, three out of five random optimization trials fail to find a non-zero-risk contamination event using the engineering model. The IDPC for *Cryptosporidium* is low and the dilution effect is significant when the East WTP is the injection site; as a result, the search domain for this contamination event is dominated mostly by zero-exposure contamination events.

Tables 4 and 5 show the decision variables and risk for each of the five random trials. The maximum-risk events that occur at the East WTP are initialized with larger loads of contaminant mass than the maximum-risk events at the West WTP. Though the probability of occurrence is smaller for larger contaminant loads, the consequences are higher, producing a higher value for risk for the large contaminant loads at the East WTP. The potential area that can be contaminated by the East WTP is larger, and the dilution effect is more significant than for events that occur at the West WTP. The value of the normalized pathogen load is larger for *Cryptosporidium* events than *Giardia* events, as the IDPC for *Cryptosporidium* is much smaller than that of *Giardia*.

The demand multipliers associated with different maximum risk scenarios are consistently lower than the average demand multiplier (1.00), because lower demand multipliers correspond to lower dilution effects and higher exposure. The minimum value for the demand multiplier is used in the final solution for critical contamination events, however, because while it increases the exposure level, it significantly decreases the occurrence probability.

Table 4 | Value of risk for five solutions identified through random trials of a GA for the single-objective problem

Location	West WTP				East WTP			
Pathogen	Giardia		Crypto.		Giardia		Crypto.	
Model	Engr.	ABM	Engr.	ABM	Engr.	ABM	Engr.	ABM
Risk values for 5 solutions	$\mathbf{1.39\times10^{-3}}$	$\mathbf{6.10\times10^{-4}}$	$\mathbf{3.43\times10^{-4}}$	$\mathbf{2.57\times10^{-4}}$	$\mathbf{6.60\times10^{-4}}$	$\mathbf{3.32\times10^{-4}}$	$\mathbf{4.30\times10^{-5}}$	$\mathbf{8.67\times10^{-5}}$
	1.36×10^{-3}	6.06×10^{-4}	3.42×10^{-4}	2.55×10^{-4}	6.08×10^{-4}	3.31×10^{-4}	2.77×10^{-5}	8.57×10^{-5}
	1.36×10^{-3}	5.98×10^{-4}	3.29×10^{-4}	2.54×10^{-4}	6.03×10^{-4}	3.30×10^{-4}	0.00×10^{-5}	8.56×10^{-5}
	1.31×10^{-3}	5.98×10^{-4}	3.19×10^{-4}	2.54×10^{-4}	5.87×10^{-4}	3.28×10^{-4}	0.00×10^{-5}	8.55×10^{-5}
	1.29×10^{-3}	5.96×10^{-4}	3.17×10^{-4}	2.47×10^{-4}	5.87×10^{-4}	3.23×10^{-4}	0.00×10^{-5}	8.52×10^{-5}

Bold indicates the maximum risk solution.

Table 5 | Decision variable values for five solutions for each optimization scenario. Results are obtained through execution of a GA for the single-objective problem

Location	West WTP				East WTP			
Pathogen	Giardia		Crypto.		Giardia		Crypto.	
Model	Engr.	ABM	Engr.	ABM	Engr.	ABM	Engr.	ABM
Start time	{**8**, 7, 10, 1, 6}	{**15**, 16, 12, 11, 1}	{**6**, 7, 8, 5, 5}	{**15**, 15, 15, 16, 13}	{**20**, 16, 1, 20, 14}	{**3**, 1, 4, 2, 1}	{**20**, 18, 8, 14, 0}	{**1**, 2, 1, 2, 1}
Duration (hr)	{**36**, 82, 56, 91, 38}	{**32**, 60, 39, 66, 51}	{**38**, 82, 37, 86, 66}	{**34**, 39, 30, 32, 62}	{**37**, 72, 95, 60, 65}	{**38**, 24, 29, 38, 24}	{**26**, 24, 77, 73, 36}	{**25**, 24, 24, 38, 24}
Demand multiplier	{**0.925**, 0.9, 0.95, 0.9, 0.925}	{**0.875**, 0.925, 0.875, 0.9, 0.95}	{**0.8**, 0.85, 0.775, 0.875, 0.775}	{**0.9**, 0.875, 0.875, 0.875, 0.9}	{**0.95**, 0.85, 0.825, 0.925, 0.75}	{**0.875**, 0.825, 0.875, 0.875, 0.825}	{**0.75**, 0.7, 0.975, 0.975, 0.975}	{**0.825**, 0.825, 0.8, 0.825, 0.8}
Normalized pathogen load	{**0.35**, 0.35, 0.36, 0.36, 0.37}	{**0.3**, 0.31, 0.3, 0.3, 0.31}	{**0.71**, 0.75, 0.7, 0.78, 0.71}	{**0.44**, 0.41, 0.42, 0.41, 0.43}	{**0.74**, 0.76, 0.74, 0.76, 0.67}	{**0.42**, 0.42, 0.42, 0.41, 0.4}	{**0.86**, 0.86, 0.02, 0.02, 0.02}	{**0.6**, 0.62, 0.57, 0.57, 0.62}

The normalized pathogen amount represents the number of infectious doses, normalized by dividing by the 98th percentile of the exponential distribution for each pathogen. Solution order corresponds to the list of risk values in Table 4, of decreasing risk. Bold indicates the maximum-risk solution.

The critical injection start times vary based on the location and the simulation model. The contaminant injection at the West WTP occurs in the evening as consumers return from unaffected central commercial and industrial districts to contaminated residential nodes in the western part of the network. When the engineering framework is used, critical contamination events at the West WTP occur in the morning because the population does not move to the unaffected central or eastern areas during the day, and consumers drink water at regular intervals in the day, beginning at 7 a.m. When the East WTP is contaminated, injection is more critical in the morning for the sociotechnical model, due to the influence of mobility. Early injection exposes the population members who reside in the west, but travel to the contaminated central and eastern commercial and industrial districts in the morning. The western population segments are not at risk when they are simulated as stationary at western areas, using the engineering model.

The worst event across all events, as predicted by both the engineering model and the sociotechnical model, occurs when *Giardia* is introduced at the West WTP (Table 6). Both models predict similar injection durations, season of occurrence (based on the demand multiplier, approximately June or October), and loading mass (Table 5). The ABM framework predicts the number of exposed customers to be approximately 38% of the number predicted by the engineering model, for the worst-case event. Both models predict the same ranking of maximum-risk events as: (1) *Giardia* at the West WTP, (2) *Giardia* at the East WTP, (3) *Cryptosporidium* at the West WTP, and (4) *Cryptosporidium* at the East WTP. The

Table 6 | Objective function values for one representative solution for each optimization scenario

Location	West WTP				East WTP			
Pathogen	Giardia		Crypto.		Giardia		Crypto.	
Model	Engr.	ABM	Engr.	ABM	Engr.	ABM	Engr.	ABM
Occurrence probability	3.79×10^{-8}	4.34×10^{-8}	9.30×10^{-9}	3.41×10^{-8}	8.40×10^{-9}	2.71×10^{-8}	4.30×10^{-9}	1.58×10^{-8}
Exposure	36,682	14,057	36,718	7,538	78,879	12,244	10,091	5,499
Risk	1.39×10^{-3}	6.10×10^{-4}	3.42×10^{-4}	2.57×10^{-4}	6.60×10^{-4}	3.32×10^{-4}	4.30×10^{-5}	8.67×10^{-5}

Results are obtained through execution of a GA for single-objective problem.

major difference between the characteristics of maximum-risk events reported by the two modeling approaches is in the values for pathogen loading. The engineering model, compared to the sociotechnical model, consistently reports a higher load for the pathogens, which causes higher consequences, but corresponds to consistently lower likelihoods of occurrence, by one order of magnitude for three of four events. This difference highlights the utility of the sociotechnical model. Because the ABM framework simulates more population dynamics to calculate the number of exposures, the discontinuities that are predicted through an engineering approach are smoothed out by population behavior, and the ABM framework does not predict extreme events with high exposures. By maximizing the number of exposed consumers (as a part of risk), the optimization algorithm takes advantage of the artifacts of the engineering simulation to identify what may be unrealistically high exposure values. For example, simulating that all consumers drink water at five specific time steps leads to the prediction of higher impacts than simulation of a heterogeneous population that drinks at different times throughout a day.

The contamination events identified through use of the sociotechnical model are more probable than corresponding events identified by the engineering model. In using the ABM framework, decision-makers can plan for more likely events, compared to the engineering approach. The trade-off between occurrence likelihood and number of exposed consumers, however, can cause difficulties for decision-makers, because it may be unclear which component should be weighted more heavily in planning for events. The issue of this trade-off is addressed in the following section, which uses multi-objective optimization to explore a non-dominated set of contamination events.

Multi-objective optimization results

NSGA-II is applied to solve the multi-objective problem (Equation (5)) for eight optimization scenarios. The performance of NSGA-II is evaluated based on the hypervolume (Zitzler & Thiele 1998), which is the non-overlapping volume in objective space covered by members of a non-dominated set of solutions with respect to a reference point, set as the worst value for each objective. The hypervolume is used to assess the convergence of the maximum-risk frontier as it evolves. Increasing values of the hypervolume necessarily represent an increase in the uniformity and proximity to a true Pareto optimality, which are characteristics that represent the quality of a non-dominated set of solutions. The convergence of the hypervolume for one trial is shown in Figure 6. Increasing values of the hypervolume for increasing generations shows that the maximum-risk frontier is migrating away from the worst point at the origin (where probability and consequences are equal to 0.0) and spreading uniformly across both axes. The search converges and stabilizes after approximately 30 generations.

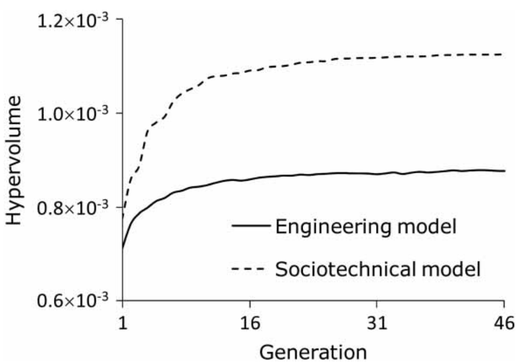

Figure 6 | Multi-objective optimization model convergence history for *Giardia* introduced at the East WTP.

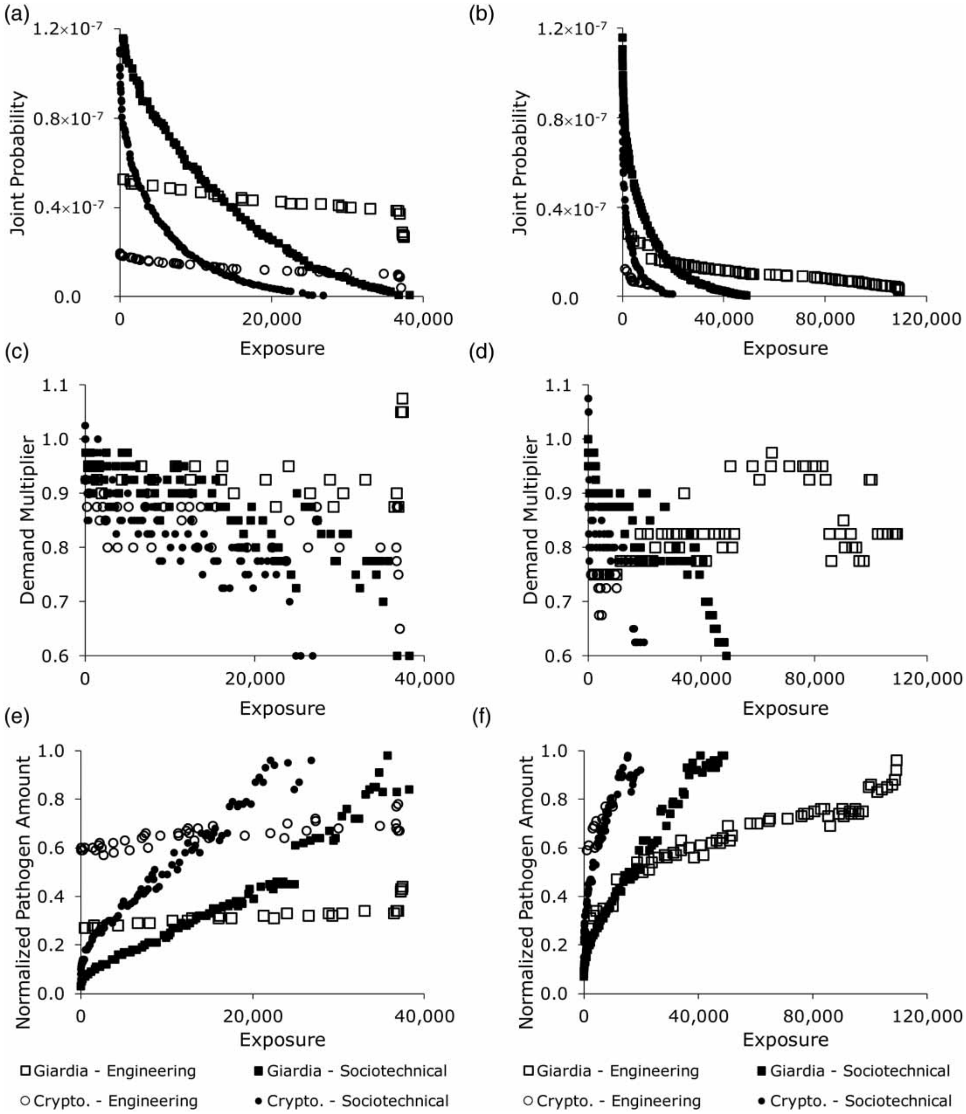

Figure 7 | Multi-objective optimization results for contamination of West WTP (left column) and East WTP (right column); (a) and (b) illustrate maximum-risk frontiers, (c) and (d) show demand multiplier associated with scenarios in maximum-risk frontiers, and (e) and (f) indicate injected pathogen amount associated with scenarios in maximum-risk frontiers.

The results of the multi-objective optimization are eight maximum-risk frontiers, for *Giardia* and *Cryptosporidium* injected at the East and West WTPs for the engineering model and ABM framework (Figure 7). As shown by the non-dominated fronts in Figures 7(a) and 7(b), the engineering model produces lower-risk, higher-consequence frontiers, compared to the sociotechnical model. The engineering model predicts that the events occurring at the West WTP have similar probabilities of likelihood, though they vary in exposure from 0.0 to almost 38,000 exposed consumers.

Values for scenario attributes, including the demand multiplier and pathogen loading, as they correspond to the non-dominated front, are shown in Figure 7. Maximum-risk scenarios with higher exposures, and thus lower occurrence probabilities, occur when the aggregate water demand (demand multiplier) is lower, as simulated using the sociotechnical model (Figures 7(c) and 7(d)). This trend in the demand multiplier is less apparent for the engineering model. The values for the contaminant loading vary across the entire range (0.0–1.0) for the sociotechnical model, while the contaminant loading for the engineering model

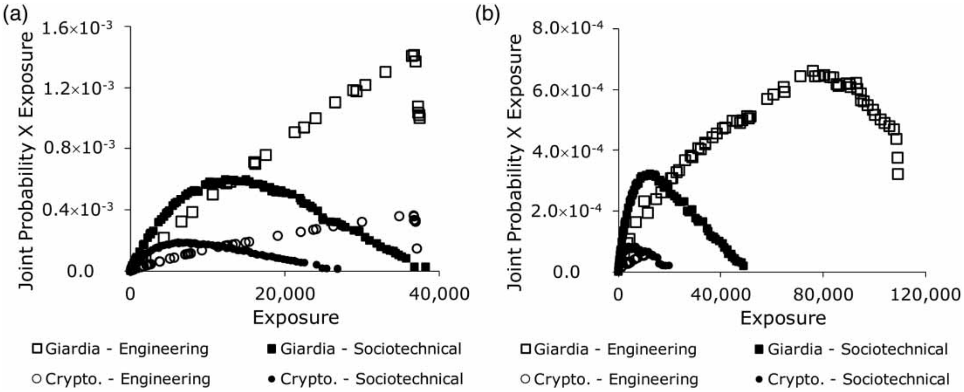

Figure 8 | Value of risk associated with scenarios in maximum-risk frontiers for contamination of (a) West WTP and (b) East WTP.

remain in smaller windows of values, particularly for events at the West WTP (Figure 7(e)). For example, the range of exposures for *Cryptosporidium* at the West WTP is simulated with the engineering model using normalized loading rates between 0.6 and 0.8. The timing of events drives the changes in population exposure across the set of non-dominated solutions for the engineering model; for the sociotechnical model, the loading rate drives the variation in exposure.

A decision-maker, such as a water utility manager, can use the non-dominated sets of solutions for planning for contamination events and determine if high probability or high consequences should be weighted more heavily in planning. Because of the discontinuity of *Giardia* at the West WTP event trade-off curve (Figure 7(a)), a decision-maker may tend to focus on the 'knee', where there is a large increase in occurrence probability with little decrease in the number of exposed consumers. This part of the trade-off curve, however, may be only an artifact of the discrete simulation of consumer behaviors, and should not be considered at the cost of neglecting other events. Specifically, the same trade-off curve for the sociotechnical model shows a smoother relationship. These trade-off curves can be used to assist a decision-maker in selecting the number of exposures that should be sustained and a level of occurrence probability that is appropriate.

Figure 8 shows the magnitude of risk for the non-dominated solutions identified using the multi-objective optimization. For all curves, the peak value of the risk is approximately equal to the maximum value of the risk found through single-objective optimization for each optimization scenario (shown for five solutions in Table 4 and for the best solution identified in Table 6). The peak values for risk in Figure 8 should be multiplied by 0.5, the location probability, to match the values reported in Tables 4 and 6.

CONCLUSIONS

A new framework is developed in this study for conducting risk assessment of water distribution contamination events. This research coupled an ABM with MCS and evolutionary algorithms to assess the risk of contamination events occurring in a water network of a mid-sized city, and a virtual city is used as a case study to demonstrate the framework. The ABM framework is the basis of a sociotechnical approach that simulates the dynamics of a contamination event by simulating the influence of the interactions among consumers and the WDS on the outcomes of the event. Conventional engineering approaches typically neglect the dynamic interactions among social and technical elements of a water contamination event and the adaptive behaviors of consumers. The risk assessment that is generated through the sociotechnical approach is compared with the results generated by an engineering model alone, in which a hydraulic model simulates the consequences of contamination events. The engineering model and the ABM framework calculate the number of exposed consumers, and the risk of an event is calculated as the product of the exposure and the occurrence likelihood. This study found that estimates of risk that are obtained using a sociotechnical approach may differ significantly from estimates obtained using an engineering approach. The sociotechnical model removes some of the unnatural components and discontinuities that are introduced through an engineering approach that assumes static, homogeneous, and stationary behaviors of consumers in drinking water.

The probability distribution of public health consequences, measured as the total number of exposed consumers, is estimated using MCS. The CDFs are smooth and continuous when the sociotechnical framework is used. Results for the engineering model show that hydraulic discontinuities and barriers cause the distribution curves to follow a step-wise discontinuous trend. In addition, the engineering model predicted a higher percentage of zero-exposure events than predicted by the ABM framework.

A GA-based approach is applied to identify critical contamination scenarios by maximizing risk for both the engineering model and sociotechnical model. Both models predicted that the maximum risk event would occur due to *Giardia* injection at the West WTP. The engineering model assumes static, homogeneous, and stationary behaviors for consumers, and the artifact of this simulation is that the exposure of consumers is predicted as high. Engineering model results generate high-consequence events and low likelihoods. When the sociotechnical aspects of an event are included, high exposures are not predicted, and the critical contamination events that are identified have a higher probability of occurrence. A multi-objective GA-based approach is used to identify a maximum-risk frontier, which is a set of contamination events that represent the trade-off between the occurrence likelihood and the consequences of contamination events. The maximum-risk frontier can guide the development of risk mitigation and emergency response plans, as decision-makers can explore the number of exposed consumers for events that vary in likelihood. Frontiers of the ABM are continuous, compared to those of the engineering models, which are discontinuous and discrete. Due to the discrete behaviors of consumers in drinking water that are simulated using the engineering model, many solutions in the engineering frontier are located at similar points, creating discontinuous fronts.

A set of assumptions about consumer behavior has been adopted in the ABM framework. The word-of-mouth simulation that is used in the current framework assumes a structured cluster of a small group of consumers. Due to social media, microblogging, and online social networking, the propagation of information among some consumers may vary distinctly from traditional word-of-mouth mechanisms. New network models that become available can be integrated within the ABM framework to explore the effects of a well-connected community on public health outcomes. The current model also simulates that an agent that is exposed to a contaminant immediately reduces its water consumption. The modeling is based on the assumption that consumers recognize that the contaminant originates in the water supply. More realistically, consumers may blame food sources or contact with exposed peers, rather than tap water, creating a delay in any reduction in water demands.

The sociotechnical risk assessment framework described here is designed to assess critical contamination events, which should be identified so that water utilities and public officials can plan approaches for mitigating WDS for these events. While there are no currently available data to validate the sociotechnical approach or engineering modeling approach for a realistic contamination event, the sociotechnical approach provides new insight about the WDS in contamination events that has not been available for analysis before. The ABM approach creates a flexible model, so that new analysis from surveys can be used to improve simulation of individual behaviors. Data about social networks and models for the cognitive process that individuals take to diagnose symptoms can be included to improve the accuracy of the modeling framework. Utility managers can be simulated as they interact during an event and take actions to alter WDS hydraulics through opening hydrants to flush a contaminant and to warn consumers about an event. Public health officials and pharmacies may also play an important role in detecting an event; data about the communication and networking among health care providers and water utilities can be used to simulate the confirmation of a water quality threat as an increasing number of consumers seek medical attention.

Further exploration of critical contamination events may reveal that a set of events exist that cause similar levels of risk. These sets of solutions should be identified to allow planning for a diversity of occurrence characteristics that may cause public health consequences. Further risk assessment approaches for exploring alternative solutions (e.g., Zechman & Ranjithan 2007; Zechman et al. 2013) can be coupled with the ABM framework to identify sets of critical contamination events and multiple maximum-risk frontiers. Ongoing explorations can create new frameworks to identify discrete sets of solutions and frontiers that provide better understanding of water events for effectively developing consumer protection strategies.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant Number CMMI-0927739. Any opinions, findings, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Ang, A. H. & Tang, W. H. 2007 *Probability Concepts in Engineering*. John Wiley and Sons, Hoboken, NJ.

Axelrod, R. 1997 *The Complexity of Cooperation*. Princeton University Press, Princeton, NJ.

Ayotte, J. D., Nolan, B. T., Nuckols, J. R., Cantor, K. P., Robinson, G. R., Baris, D., Hayes, L., Karagas, M., Bress, W., Silverman, D. T. & Lubin, J. H. 2006 Modeling the probability of Arsenic in groundwater in New England as a tool for exposure assessment. *Environ. Sci. Technol*. **40** (11), 3578–3585.

Bristow, C. B. & Brumbelow, K. 2006 Delay between sensing and response in water contamination events. *J. Infrastruct. Syst*. **12** (2), 87–95.

Brumbelow, K., Torres, J., Guikema, S., Bristow, E. & Kanta, L. 2007 Virtual cities for water distribution and infrastructure system research. In: *Proceedings of the World Environmental and Water Resources Congress* (K. C. Kabbes, ed.). Tampa, FL, May 15–18, 2007. ASCE, Weston, VA.

Centers for Disease Control and Prevention 2001 *Norwalk-like Viruses: Public Health Consequences and Outbreak Management. MMWR Recommendations and Reports, 50 (No. RR-9)*.

Davis, M. & Janke, R. 2008 Importance of exposure model in estimating impacts when a water distribution system is contaminated. *J. Water Resour. Plan. Manage*. **134** (5), 449–456.

Davis, M. & Janke, R. 2009 Development of a probabilistic timing model for the ingestion of tap water. *J. Water Resour. Plan. Manage*. **135** (5), 397–405.

Davis, M. J. & Janke, R. 2011 Patterns in potential impacts associated with contamination events in water distribution systems. *J. Water Resour. Plan. Manage*. **137** (1), 1–9.

Deb, K. 2001 *Multiobjective Optimization using Evolutionary Algorithms*. Wiley, Chichester.

Deb, K. & Agrawal, R. B. 1995 Simulated binary crossover for continuous search space. *Complex Syst*. **9** (2), 115–148.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm NSGA-II. *IEEE Trans. Evol. Comput*. **6** (2), 182–197.

Glouberman, S. 2001 Walkerton water and complex adaptive systems. *Hosp. Q*. **4** (4), 28–31.

Goldberg, D. E. 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.

Haimes, Y. Y. 2009 *Risk Modeling, Assessment, and Management*. John Wiley & Sons, Hoboken, NJ.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Holland, J. H. 1995 *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley, Reading, MA.

Hrudey, S. & Hrudey, E. 2004 *Safe Drinking Water: Lessons from Recent Outbreaks in Affluent Nations*. IWA Publishing, London.

Hrudey, S. & Hrudey, E. 2007 Published case studies of waterborne disease outbreaks: evidence of a recurrent threat. *Water Environ. Res*. **79** (3), 233–245.

Johnson, G. & Brumbelow, J. K. 2008 Developing Mesopolis- A Virtual City for Research in Water Distribution System and Interdependent Infrastructures. Available online at https://ceprofs.civil.tamu.edu/kbrumbelow/Mesopolis/index.htm.

Kaplan, S. & Garrik, B. J. 1981 On the quantitative definition of risk. *Risk Anal*. **1** (1), 11–27.

Khanal, N., Buchberger, S. G. & McKenna, S. A. 2006 Distribution system contamination events: Exposure, influence, and sensitivity. *J. Water Resour. Plan. Manage*. **132** (4), 283–292.

Kothary, M. H. & Babu, U. S. 2001 Infective dose of foodborne pathogens in volunteers: a review. *J. Food Saf*. **21** (1), 49–68.

Lindell, M. K., Mumpower, J., Wu, H-C. & Hwang, S-K. 2010 *Perceptions and Expected Responses to a Water Contamination Emergency*. Texas A&M University Hazard Reduction & Recovery Center, College Station, TX.

Lindell, M. K. & Perry, R. W. 2004 *Communicating Environmental Risk in Multiethnic Communities*. Sage, Thousand Oaks, CA.

Lindell, M. K., Prater, C. S. & Perry, R. W. 2006 *Fundamentals of Emergency Management*. Federal Emergency Management Agency, New York, NY.

Mayer, P., Tawler, E. & DeOreo, W. B. 1999 *Residential End Uses of Water*. American Water Works Association, Denver, CO.

Miller, J. & Page, S. 2007 *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press, Princeton, NJ.

Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Hilton, A. C., Karamouz, K., Minsker, B., Ostfeld, A., Singh, A. &

Zechman, E. 2010 State of the art for genetic algorithms and beyond in water resources planning and management. *J. Water Resour. Plan. Manage.* **136** (4), 412–432.

Nilsson, K. A., Buchberger, S. G. & Clark, R. M. 2005 Simulating exposures to deliberate intrusions into water distribution systems. *J. Water Resour. Plan. Manage.* **131** (3), 228–236.

Pasha, M. F. & Lansey, K. 2010 Effect of parameter uncertainty on water quality predictions in distribution system-case study. *J. Hydroinf.* **12** (1), 1–21.

Perelman, L. & Ostfeld, A. 2010 Extreme impact contamination events sampling for water distribution systems security. *J. Water Resour. Plan. Manage.* **136** (1), 80–87.

Protopapas, A. L., Katchamart, S. & Platonova, A. 2000 Weather effects on daily water use in New York City. *J. Hydrol. Eng.* **5** (3), 332–338.

Rasekh, A. & Brumbelow, K. 2013 Probabilistic analysis and optimization to characterize critical water distribution system contamination scenarios. *J. Water Resour. Plan. Manage.* **139** (2), 191–199.

Rasekh, A., Brumbelow, K. & Zechman, E. M. 2010 WDS vulnerability analysis: focusing on random factors, consumer behavior, and system dynamics in contamination events. In: *Proceedings of the ASCE/EWRI World Environmental and Water Resources Congress*, Providence, RI, May 2010.

Reynolds, K. A., Mena, K. D. & Gerba, C. P. 2008 Risk of waterborne illness via drinking water in the United States. *Rev. Environ. Contam. Toxicol.* **192**, 117–158.

Rose, J. B., Haas, C. N. & Regli, S. 1991 Risk assessment and control of waterborne giardiasis. *J. Public Health* **81** (6), 709–713.

Rossman, L. A. 2000 *EPANET 2.0, User's Manual*. National Risk Management Research Laboratory, US EPA, Cincinnati.

Roxström-Lindquist, K., Palm, D., Reiner, D., Ringqvist, E. & Svärd, S. G. 2006 *Giardia* immunity – an update. *Trends Parasitol.* **22** (1), 26–31.

Shafiee, M. & Zechman, E. M. 2010 An agent-based modeling approach for simulating contamination events applied to the Mesopolis water distribution system. In: *Proceedings of the ASCE/EWRI World Environmental and Water Resources Congress*, Providence, RI, May 2010.

Shafiee, M. E. & Zechman, E. M. 2011 Sociotechnical simulation and evolutionary algorithm optimization for routing siren vehicles in a water distribution contamination event. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2011 (GECCO)*, Evolutionary Computation and Multi-agent Systems and Simulation (ECoMASS) Workshop, Dublin, Ireland, July 2011.

Shafiee, E. & Zechman, E. M. 2012 Integrating evolutionary computation and sociotechnical simulation for flushing contaminated water distribution systems. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2012*, Evolutionary Computation and Multi-agent Systems and Simulation (ECoMASS) Workshop, Philadelphia, PA, July 2012.

Shafiee, M. & Zechman, E. M. 2013 An agent-based modeling framework for sociotechnical simulation of water distribution contamination events. *J. Hydroinf.* **15** (3), 862–880.

Shang, F., Uber, J., van Bloememn, B., Boccelli, D. & Janke, R. 2006 Real time water demand estimation in water distribution system. In: 8th Annual Water Distribution Systems Analysis Symposium, Ohio, USA, pp. 95–103.

Torres, J. M., Brumbelow, K. & Guikema, S. 2009 Risk classification and uncertainty propagation for virtual water distribution systems. *Reliab. Eng. Syst. Saf.* **94** (8), 1259–1273.

Tung, Y. K., Yen, B. C. & Melching, C. S. 2006 *Hydrosystems Engineering Reliability Assessments and Risk Analysis*. McGraw-Hill, USA.

Uber, J., Regan, M. & Janke, R. 2004 Use of system analysis to assess and minimize water security risks. *J. Contemp. Water Res. Educ.* **129** (1), 34–40.

United States Congress 2002 Public health security and bioterrorism preparedness and response act of 2002. In: *Public Law No. 107-188, H.R. 3448, 107th Congress*. Government Printing Office, Washington, DC.

United States Department of Labor, Bureau of Labor Statistics 2011 Employment Situation Summary. Available at http://www.bls.gov/news.release/empsit.nr0.htm, Table A-1 to A-16.

United States Environmental Protection Agency 2000 *Estimated per capita water ingestion in the United States*. EPA-822-R-00-008, Office of Water, Washington, DC.

Vicente, K. J. & Christoffersen, K. 2006 The Walkerton *E. coli* outbreak: A test of Rasmussen's framework for risk management in a dynamic society. *Theor. Issues Ergon. Sci.* **7** (2), 93–112.

Watts, D. J. 1999 *Small Worlds*. Princeton University Press, Princeton, NJ.

White, J. W. 1999 *Hazards of Short-Term Exposure to Arsenic-Contaminated Soil*. Office of Environmental Health Assessment Services, Olympia, WA.

Woo, D. M. & Vicente, K. J. 2003 Sociotechnical systems, risk management, and public health: comparing the North Battleford and Walkerton outbreaks. *Reliab. Eng. Syst. Saf.* **80** (3), 253–269.

XJ Technologies 2010 *Anylogic 6 User's Manual*. St. Petersburg, Russia.

Zechman, E. M. 2007 Agent-based modeling to simulate contamination events and to analyze threat management strategies in water distribution systems. In: *Proceedings of the ASCE/EWRI World Environmental and Water Resources Congress*, Tampa, FL, May 2007.

Zechman, E. M. 2010 Integrating complex adaptive system simulation and evolutionary computation to support water infrastructure threat management. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2010 (GECCO)*, Evolutionary Computation and Multi-agent Systems and Simulation (ECoMASS) Workshop, Portland, OR, July 2010.

Zechman, E. M. 2011 Agent-based modeling to simulate contamination events and evaluate threat management strategies in water distribution systems. *Risk Anal*. **31** (5), 758–772.

Zechman, E. M. 2013 Integrating evolution strategies and genetic algorithms with agent-based modeling for flushing a contaminated water distribution system. *J. Hydroinf*. **15** (3), 798–812.

Zechman, E. M., Giacomoni, M. & Shafiee, M. 2013 An evolutionary algorithm approach to generate distinct sets of non-dominated solutions for wicked problems. *Eng. Appl. Artif. Intell*. **26** (5), 1442–1457.

Zechman, E. M., Lindell, M., Brumbelow, K., Mumpower, J., Rasekh, A. & Shafiee, M. 2011 Including consumer behavior and system dynamics in vulnerability analysis for water distribution system contamination events. In: Proceedings of the NSF Engineering Research and Innovation Conference, Atlanta, GA.

Zechman, E. M. & Ranjithan, S. 2007 Generating alternatives using evolutionary algorithms for water resources and environmental management problems. *J. Water Resour. Plan. Manage*. **133** (2), 156–165.

Zitzler, E. & Thiele, L. 1998 *Multiobjective Optimization using Evolutionary Algorithms – A Comparative Case Study. Parallel Problem Solving from Nature – PPSN V*. Springer, Berlin/Heidelberg.

First received 19 February 2013; accepted in revised form 17 July 2013. Available online 29 August 2013