Quantifying Failure for Critical Water Infrastructures under Cyber-Physical Threats

Georgios Moraitis[1]; Dionysios Nikolopoulos[2]; Dimitrios Bouziotas[3]; Archontia Lykou[4]; George Karavokiros[5]; and Christos Makropoulos[6]

Abstract: This paper presents a failure quantification methodology to assess the impact of cyber-physical attacks (CPAs) on critical water infrastructures, such as water distribution networks, by mapping simulation-derived data onto metrics. The approach sets out a three-step profiling architecture to interpret the consequences of failures resulting from CPAs against several dimensions of integrity, adjusted through user-defined service levels. Failure is examined in terms of its magnitude, propagation, severity, and crest factor, while rapidity is used to infer available time slots to react. The methodology is operationalized through a dedicated tool designed to assist water-sector critical infrastructures gauge and assess CPAs. The approach is demonstrated on a benchmark water distribution system, and results and insights from the metrics are presented and discussed. It is argued that the approach and the tool that operationalizes its application can be useful to water companies that need to assess and compare cyber-physical threats and prioritize mitigation actions based on quantitative metrics.
DOI: 10.1061/(ASCE)EE.1943-7870.0001765.

Author keywords: Cyber-physical attacks; Cyber-physical systems; Consequences; Risk assessment; Risk management.

Introduction

Water systems are essential for health, safety, and well-being. As such, they are considered critical infrastructures (CIs) (McPherson and Burian 2005) whose disruption of service can have significant debilitating impacts. Contemporary water CIs are moving toward cyber and physical integration, merging processes with computational systems to form cyber-physical systems (CPSs) (Lee 2008). Water CIs inherit larger attack surfaces (Howard et al. 2005) from the entanglement of cyber and physical layers, and additional pressure is introduced to their strategic and tactical planning. Advanced cyberattacks are designed to infringe upon the physical domain through communication and/or computational infrastructures, thereby evolving into cyber-physical threats. Increased functionalities related to the autonomous operation of subsystems, real-time monitoring, and remote-control capabilities, designed to increase efficiency, are becoming risk sources, exploited by adversaries to disturb or even weaponize water supplies (Janke et al. 2014). In this era, enhancing data-driven emergency preparedness and planning, to better comprehend and manage emerging risks, helps ensure safe and resilient water systems for communities (Ugarelli et al. 2018).

Cyber-Physical Vectors in Risk Management

The intertwining of cyber and physical layers arguably increases efficiency and accuracy by offering capabilities like remote real-time control (RRTC) for pressure management (Giustolisi et al. 2017; Page et al. 2017), mitigation of combined sewer overflow (CSO) (Garofalo et al. 2017) or extension of actuators' lifespan (Lund et al. 2018), and the detection of contamination (Wang et al. 2015) or leakages at the household level (Kossieris et al. 2014). On the other hand, just as water CIs benefit from shifting to more integrated CPSs, so do potential adversaries by constantly adjusting their tactics, techniques, and procedures (TTPs) (Johnson et al. 2016) to exploit the new cyber-physical domain. The European Union's (EU) Agency for Network and Information Security (ENISA) has reported a shift in the threat landscape from individuals to companies (ENISA 2019), while for the same period the annual strategic report of the European Cybercrime Center (EC3) identifies the convergence of cyber and terrorism (Europol 2018). Access to a range of malwares and anonymization and encryption tools or services through the Darknet enables even inexperienced threat actors to exploit vulnerabilities and perform cyber-physical attacks (CPAs) that go well beyond their actual know-how and skills. Common misguided security perceptions over industrial

[1]Civil Engineer and Ph.D. Candidate, Dept. of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical Univ. of Athens, Heroon Polytechneiou 5, Zografou GR-15780, Greece (corresponding author). Email: georgemoraitis@central.ntua.gr

[2]Civil Engineer and Ph.D. Candidate, Dept. of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical Univ. of Athens, Heroon Polytechneiou 5, Zografou GR-15780, Greece. Email: nikolopoulosdio@central.ntua.gr

[3]Civil Engineer and Researcher, KWR Water Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, Netherlands. Email: Dimitrios.Bouziotas@kwrwater.nl

[4]Civil Engineer and Ph.D. Candidate, Dept. of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical Univ. of Athens, Heroon Polytechneiou 5, Zografou GR-15780, Greece. Email: alykou@central.ntua.gr

[5]Computer Scientist, Dept. of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical Univ. of Athens, Heroon Polytechneiou 5, Zografou GR-15780, Greece. Email: gkaravo@itia.ntua.gr

[6]Associate Professor, Dept. of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical Univ. of Athens, Heroon Polytechneiou 5, Zografou GR-15780, Greece; Principal Scientist, KWR Water Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, Netherlands. Email: cmakro@mail.ntua.gr

Note. This manuscript was submitted on January 15, 2020; approved on March 23, 2020; published online on July 13, 2020. Discussion period open until December 13, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Environmental Engineering***, © ASCE, ISSN 0733-9372.

control systems (Loukas 2015), broad geographical expansion of CPSs (Konstantinou et al. 2015), and a rise in sophistication (Rasekh et al. 2016) of malicious codes allow for a range of manipulation and deception attacks.

According to the latest Verizon Data Breach Investigations Report (DBIR) (Verizon 2019), 23% of the reported breaches involved nation state- or state-affiliated actors and 28% leveraged malware to establish or advance attacks. Relevant to those findings, a technical alert by US-CERT (US-CERT 2018) revealed that since at least March 2016 multiple US CIs, including water-sector CIs, were strategically targeted by foreign government cyber actors who, inter alia, gained access to industrial control systems (ICSs). Compromised ICSs and unauthorized access over such systems can go undetected for long periods, exposing water CIs and society to significant risks. Such is the case of a company responsible for supplying a number of neighboring counties, anonymized under the pseudonym "Kemuri Water Company" (Verizon 2016), with an unusual operation of remotely controlled assets lasting nearly 2 months. Attackers accessed the platform that supervised hundreds of programmable logic controllers (PLCs), gaining control over and altering the dosing of chemicals used for water treatment and the water flow per se, compromising supply services. Another example from the recent water-sector CPA history is the 2013 near-miss after a successful supervisory control and data acquisition (SCADA) hack of Bowman Dam in New York. As made public by the relevant indictment, the attacker, who had repeatedly obtained data on water levels, temperature and sluice status, had gained access to the sluice remote control system as well. Fortunately, he was unable to escalate his threat only because the dam's sluice was disconnected for maintenance. A recent review on the sector's incidents by Hassanzadeh et al. (2020) reveals the diversity of attackers' TTPs and resulting consequences. Officially disclosed or otherwise, recent incidents or near-misses in water CIs have raised a caution flag that should not be overlooked.

Proactive risk management calls for prevention and preparedness, through structured multidisciplinary approaches for stress testing (Galbusera et al. 2014; Licák 2006), against current and future threats. Both the EU [CD 2008/114/EC, article 2(c)] and the USA (US Department Homeland Security 2009) critical infrastructure protection (CIP) frameworks recommend risk assessments that follow the threat scenarios approach. The latter, considered the drivers of emergency simulations (Grance et al. 2006), act as catalysts that trigger the exploration of infrastructure exposure to risk and inspire actions to protect against potential threats, up to weaponization of supply (ASME-ITI 2009). To meet those objectives, cyber-physical threat scenarios must address key threat characteristics and explore multiple durations and escalations of events under existing operating plans and available alternatives (Bouchon et al. 2008), while rendering the adversary's TTPs.

Toward Realistic Stress Testing

Stress tests are risk and safety assessment approaches designed to associate the severity of a threat scenario with its impact on the system or society, performing the core analysis required for the prevention of risks and the preparedness of the CI (Galbusera et al. 2014). Physical, cyber, geographical, and logical interdependencies within a system, as defined by Rinaldi et al. (2001), allow for cascading effects to occur. For water CIs to prepare against events that may cascade from the cyber to the physical layer and vice versa, appropriate stress-testing environments are required that can model those dynamics (Nikolopoulos et al. 2018). This has triggered cyber-oriented research on developing virtual SCADA environments and test vulnerabilities (Almalawi et al. 2013; Chen et al. 2015; Davis et al. 2006; Fovino et al. 2010; Queiroz et al. 2011; Siaterlis et al. 2013). In the water CI domain, the widely accepted EPANET model (Rossman 2000), used to simulate hydraulic systems, has recently started transforming toward bridging that cyber-physical gap. Eliades et al. (2016) provided a programming interaction for a simulator through MATLAB in an effort to assist research in the field of smart water networks, used by Taormina et al. (2018) to deploy the EpanetCPA toolbox and link monitoring and control device interactions to traditional network hydraulics. The latter, inter alia, provides a structured way of importing CPA scenarios and pass them, in a certain level of modeling abstraction, through a hydraulic solver. In a similar manner, Klise et al. (2017) developed an open-source Python software version 0.2.2 package, the Water Network Tool for Resilience (WNTR), which employs both EPANET and a purpose-built EPANET-based simulator to allow for the modeling and simulation of water distribution networks (WDNs), focused on network resilience in physical emergency states (e.g., earthquake, power outage). WNTR has been recently used in the work of Nikolopoulos et al. (2019a) in an early prototype of a cyber-physical stress-testing platform called RISKNOUGHT.

Stress testing introduces distributed or point loads that cause performance to drift outside normal boundaries and lead to nonideal conditions of service. Events, like an attack or power outage at a pumping station, can lead to pressure deficiency conditions, in which demand-driven analysis (DDA) solvers, such as the original EPANET solver, pose limitations (Chmielewski et al. 2016). DDA solvers continuously supply nodes regardless of the pressure, yielding unrealistic demand satisfaction and hydraulic behavior. Quality of generated data is directly linked to the simulation approaches and methods chosen (Wand and Wang 1996), and as such, DDA is not suitable for the purposes of a stress test. On the other hand, linking pressure to nodal outflow allows for pressure-driven analysis (PDA) through nodal head-flow relationship (NHFR) formulas (e.g., Fujiwara and Li 1998; Germanopoulos 1985; Wagner et al. 1988). Maximum demand satisfaction is met under optimal pressure conditions and decreases as pressure drops, down to a minimum operating value. Because water supply is based on available operating pressure at each node, PDA-based stress-testing platforms are indeed able to represent pressure deficiency effects more realistically (Todini 2003).

Utilizing state-of-the-art tools and methodologies that best fit the purposes of the analysis can form a realistic cyber-physical stress-testing approach. Subsequently, this produces data deemed to be of high quality that need to be mined to express failure. This work provides a defined structure to interpret a system's predicament by translating simulation data to failure information, aiming to enhance risk-informed decisions and prioritization of actions.

From Data to Performance Information

Translating model-derived data to meaningful aggregates and keeping an overview of the simulated behavior is a difficult task because of the large volume of raw data. Real water network models contain thousands of nodes and assets, dynamically operated over a simulation period. Even skeletonized network models, with known limitations and shortcomings (Davis and Janke 2018), produce large sets of data, while fine-time-resolution simulation adds to both the detail and volume of results. Thus, making sense of stress-test results in a structured and efficient way becomes of paramount importance for facilitating risk-informed decision-making (Hansson and Aven 2014) and can be achieved by mapping results to suitable indicators.

Water-sector experts often keep track of network service performance and communicate company goals through sets of measures designed for management which are found to be very similar across countries and companies (Vilanova et al. 2015). The similarity originates from the common fundamental processes, assets, and overall goals of water companies, with metrics usually focused on five main categories of management interest:

- Quality of service, which includes both quantity and quality delivered to customers;
- Asset, which includes the physical performance of the infrastructure;
- Operational, which relates to daily system monitoring and maintenance;
- Personnel, which focuses on human resources management; and
- Financial, which keeps track of the financial soundness and economic prosperity of the company.

Performance measures are metrics that quantify the efficiency or effectiveness of an action (Neely et al. 2005). These are categorized into result indicators (RIs), answering the question of what has been achieved so far, and performance indicators (PIs), indicating what needs to be done to increase performance (Parmenter 2015). Adding the word *key* indicates the importance of those factors in achieving the defined goals, revealing the critical success factors. RIs capture the results of operational actions and show whether the organization is "travelling at the right direction at the right speed" (Parmenter 2015), which is important for the governance of the organization. The difference can be seen through the definition provided by Alegre et al. (2016), where PIs are efficiency and effectiveness measures for the delivery of services with respect to target values. Such values are benchmarks used for comparison (Cable and Davis 2004) or as reference points of improved performance (Malano et al. 2004) and the establishment of policies (Walter et al. 2009). Thus, metrics like those presented by Alegre et al. (2016), Danilenko et al. (2014), Kanakoudis et al. (2011), Bouziotas et al. (2019), and others, serving trend monitoring (Andersen and Fagerhaug 2002) and long-term benchmarking objectives (Berg 2013), do not reveal the dynamics or inner characteristics of a system failing under stress.

An emerging concept related to the performance of water systems under stress is that of resilience. Being a relatively recent term in the water industry, it has received many definitions in the scholarly literature (Francis and Bekera 2014). The variations are mostly subtle (Butler et al. 2017), while a stress-testing-oriented approach regarding water system resilience is given by Makropoulos et al. (2018), who define resilience as "the degree to which an urban water system continues to perform under progressively increasing disturbance." As an expansion of stability (Holling 1996), resilience is linked to the ability of a system to react to stress conditions (Todini 2000) and reduce the magnitude or duration of disruptive events (NIAC 2009), to retain a level of functionality. Several studies have tried to capture and indicate resilience against a failure (mainly in network design), as a generic inverse function of failure time (Hashimoto et al. 1982; Kjeldsen and Rosbjerg 2005), quantified via resilience profile graphing tools (Makropoulos et al. 2018), through a demand satisfaction ratio (Mehran et al. 2015; Zhuang et al. 2012), available energy (Creaco et al. 2016; Todini 2000), or graph theory metrics (Herrera et al. 2016), while resilience through operational and financial dimensions is proposed in the Risk Analysis and Management for Critical Asset Protection (RAMCAP) approach (AWWA 2010). Though it is used in system design optimization, linked to recovery plans (Chmielewski et al. 2016), and becoming increasingly recognized as essential to rethinking contemporary water CPSs (Nikolopoulos et al. 2019b), it has been argued that no resilience measure proposed to date can adequately describe cascading failures (Shin et al. 2018) or provide adequate context on a system's complex integrity predicament during stress.

Recognizing a need to summarize, interpret, and communicate data derived from stress testing, a quantification framework is proposed in the following sections that is deployable through a purpose-built tool. It is designed specifically for water CIs under threat and is adjustable to any internal and external operating environment of a water utility.

Quantifying Consequences

Setting the Scene

Mapping data is essential in accurate information delivery for emergency preparedness (Zoppi et al. 2016) and data-driven risk management (Niesen et al. 2016) that aims to preserve water system integrity. According to the European standard on drinking water security [CEN-EN 15975-2 (CEN 2013)], referenced by Commission Directive (EU) 2015/1787, integrity means meeting "specified quality, quantity, continuity and pressure targets in accordance with legal/regulatory requirements and the drinking water supplier's objectives." These targets, regulatory restrictions, and acceptable levels of performance vary according to each company's operational environment and risk attitude. Results must be presented in a way that assists water companies to comprehend the nature and level of risk by linking potential consequences to a company's objectives.

Both events and their consequences are used to characterize risk (ISO 2018), but to obtain a balanced information flow, especially in a multistakeholder environment such as the water sector, it may be argued that shared information needs to be

- *Compatible* with the purposes of its use,
- *Concise* to avoid losing focus,
- *Comprehensive* to thoroughly cover the subject of interest,
- *Consistent* with the physical and logical environment it describes, and
- *Comparable* in order to discover differences or similarities between reference points.

The preceding "5C" checklist is used as a guide to the approach proposed in this work: the information produced through the proposed approach is failure-oriented (*compatible*) focused on the center of gravity of the CI (*concise*). Risk information can derive from different families of metrics in multiple dimensions (*comprehensive*) adjustable to the internal and external operating environment (*consistent*). Metrics are selected to be quantitative, dimensionful, and related to a reference state of the system (*comparable*). Each of these concepts is explained in one of the following sections, while the last section briefly presents a basic flowchart and functionalities of an early version of a tool designed to operationalize the proposed framework for interested water companies.

Failure-Oriented Approach

Risk analysis is defined as a necessary step to comprehend the nature and profile of a risk and define its level as a function of an event's consequences for a system. Consequences, either positive or negative, are the outcomes of an event that affect system objectives (ISO 2018). Performance improvements, measured through PIs, come from utility management actions, while negative effects on the system's objectives come as the result of malfunctions, accidents, natural disasters, malevolent actions, or hybrids of these. In the light of the cyber-physical era, threat actors can be categorized,

based on their intentions, as terrorists, state actors, cyber-criminals, hacktivists, hobbyists, and insiders (Nicholson et al. 2012; Rasekh et al. 2016). Evidently, the causality relationship between an actor's intentions and impact is affected by existing system vulnerabilities, required capabilities, and available resources of the actor. Whether seeking to harm people or an organization's reputation, the impact of an attack is measured by the decay of the system's performance. For CI stakeholders to acquire situational awareness in that perspective, the adversary's perspective should be adopted (Schnaubelt et al. 2014) and address system behavior under stress in a failure-oriented manner. This provides an information context that is more compatible with the purposes of risk analysis. Such a frame of reference also better serves the decision-making process: A positive tone risk approach can be misleading, biasing decision-making (Damasio et al. 1996; Sanfey et al. 2003), while in fact the prospect and experience of losses lead to safer and higher reward decisions (Bechara et al. 1999).

A water system's state can be defined, at any point or period, through sets of multidimensional metrics, which provide information on performance relative to specified reference values. As the operating environment changes, the system reacts to the received stimulus and alters its state. Ex ante risk analysis focuses on the prediction of the system's potential state, if a threat event were to occur, under specific internal and external operating conditions. Stress testing the system results in a snapshot of its state, allowing for a comparison between that and the system's state without the stimuli. The difference between the two is a measure of an event's consequences. This conceptualization of consequence is at the core of the approach proposed herein.

A system's performance, under normal operating conditions (undistorted state), is depicted here as a steady state. Imposing a stress would cause a change of state, shifting the system to lower performance levels (stress state), as in Fig. 1. As described in previous section, the retained functionality of the system under stress is related to the system's resilience. To provide a risk-compatible context, the failure-oriented approach proposed is focused on profiling the area between the performance curves, which is complementary, at least at an abstract level, to resilience.

To ensure a resilient and efficient system, companies orchestrate assets and operations throughout the urban water cycle. From the wide range of systems and subsystems included in the urban water cycle, here the focus is on the WDN, which is suggested is the so-called center of gravity of a water CI. The concept of center of gravity is borrowed here from the general security literature (Schnaubelt et al. 2014) and defined as the entity processing the key capabilities

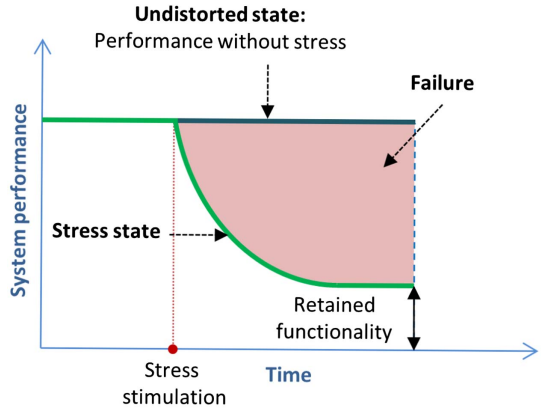

Fig. 1. Idealized system performance curves in undistorted state and under stress.

to achieve a CI's objectives. The purpose of identifying a center of gravity is to assist stakeholders in maintaining their focus on crucial risk information (Schnaubelt et al. 2014). For a resilient water infrastructure, arguably, the distribution network is the key indispensable resource utilized to provide services and support the objectives of safe and adequate water supply to the community. Destroying, weakening, or influencing a cyber-physical WDN's processes hinders the fulfillment of objectives, making it the system's center of gravity.

Failure Profile

A failure profile is defined as a collection of factual aspects of consequences that can be used to interpret a system's predicament. Plain as it may sound, reconstructing the failure information from simulation-derived data requires multiple viewpoints through metrics that allow an unbiased and opinion-free information flow to interested stakeholders. To assess a system's failure, especially in the case of deliberate CPAs, stakeholders need to be aware of the type, level, and characteristics of an event's aftermath. For this reason, this paper proposes a three-step approach for the main profiling structure. The first step is categorizing the services provided by the supplier to the community. The second step is to identify pivotal levels of failure for each service category. The third and last step is setting the dimensions on which failure, for each level, is mapped.

First Step: Service Categories

As mentioned earlier, water service integrity refers to the quality and quantity delivered within a given operational range of pressure. In an ex ante approach, simulation models transpose real-world systems and threats into the virtual domain and encapsulate them through sets of rules that define future system state changes at each time step (Borshchev and Filippov 2004). Stimulated by the threat scenario each time, a model reacts upon the predefined set of rules and system state changes by dynamically addressing system behavior. Such is the reaction of a hydraulic model in the presence of low-pressure conditions, affecting node supply and adjusting network flows based on the chosen NHFR formula. Thus, using the more realistic PDA approach reduces the required information of service integrity by one dimension because pressure and supply are addressed in a direct cause-effect modeling dependence. The proposed profiling approach needs to assess failure only in terms of supply, creating two sets of metrics: those related to sufficiency of quantity and those related to safety of quality.

Second Step: Service Levels

Urban water systems are operated under an ensemble of regulations and policies defined by their internal and external environments. Analogously, a system under failure is also called to balance under the weight of strategic guidelines, perspectives, and plans of the internal stakeholders and the ever-present accountability to regulatory, legislative, social, and political system expectations. To capture the importance of such strains and limitations, this paper proposes the use of *service levels*. These levels, confined within adjustable ranges per stakeholder, represent a state between complete failure and optimal performance of services. Though numerous intermediate levels of service can be defined, there are two generic, definable but not fixed, pivotal levels. The first level is the maximum allowed interruption of services (based on regulations, standards, and operational policies, for example) and the other can be a tolerable level below undistorted services. The latter can be associated with mild discomfort of customers and, possibly, reputational damage to a company. Hence, the key system service levels are defined as disrupted, degraded, and normal states [Fig. 2(c)].

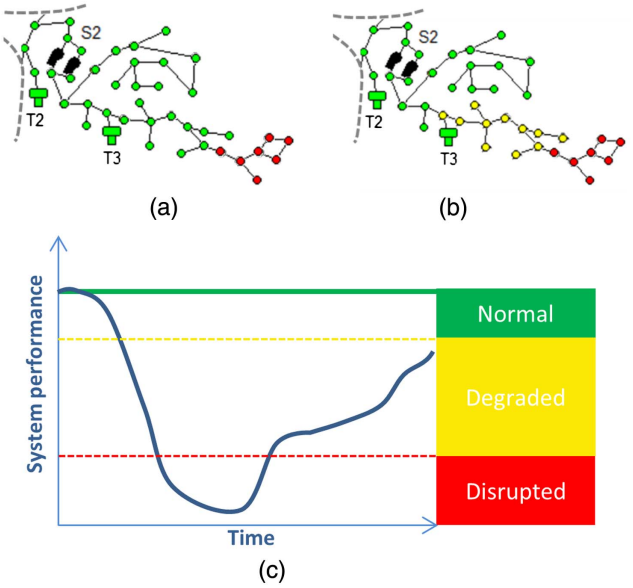

Fig. 2. Performance predicament for part of a network: (a) without and (b) with the proposed intermediate level; and (c) system performance and relative service levels.

The proposed service levels can be viewed as a "traffic lights for risks" analogy. Complete failure levels (polluted or interrupted supply) are a potentially harmful and operationally critical state of the system, to which special attention should be given (red light). Degraded services (insufficient or substandard supply) are a non-harmful and more reputation-oriented failure level, raising a caution flag (yellow light). Any service that meets the requirements above inadequate level thresholds, up to optimal service performance, is considered normal (green light). To understand the need for such a classification, one can assume a stress scenario that leads to pressure deficiency conditions. If only complete service interruption were considered a failure, the spatial image of failure could be represented by Fig. 2(a), with seven nodes being affected. But in reality, failure in this part of the system could be closer to Fig. 2(b), since pressure drops and partial supply of demand would not be acceptable to customers (11 nodes). This applies to both types of service recognized in Step 1 of the process. It is an intermediate classification of failure, important to the separation between critical and less critical levels per stakeholder, that makes it possible to apply the metrics to each.

For a cyber-physical stress-testing simulation of T hours for a network that contains N number of demand nodes, the parameters of time $t \in (0, T]$ and node $n \in [1, N]$ are defined. Denoting supply to node n of the network at time t by $S_{n,t}$ and demand by $D_{n,t}$, the thresholds and range for the disrupted (L_1), degraded (L_2), and normal (L_3) levels of service in terms of quantity are found in Table 1. In this paper, l denotes the lowest percentage of demand below which (even in PDA configurations) supply is perceived as zero-equivalent in terms of satisfaction of needs. Thus, complete failure is assumed to occur for a range of low demand satisfaction (L_1). For the degraded state, the upper boundary, denoted by h, is the lowest acceptable service provided to customers in a state of emergency before causing inconvenience or disturbance. For the creation of any other service level, the generic definition of its range is found in the last line of Table 1, for any new level L_i within the thresholds p_i and p_{i+1}.

With respect to quality of supply, quality standards and legal frameworks provide guidance in setting the minimum acceptable

Table 1. Service level ranges

Service level	Alias	Quantity Range	Quality Range
L_1	Disrupted	$0 \le S_{n,t} < l \times D_{n,t}$	$c_{s_e} \le c_{s_{n,t}}$
L_2	Degraded	$l \times D_{n,t} \le S_{n,t} < h \times D_{n,t}$	$c_{s_p} \le c_{s_{n,t}} < c_{s_e}$
L_3	Normal	$h \times D_{n,t} \le S_{n,t} < D_{n,t}$	$c_{s_{n,t}} < c_{s_p}$
L_i	Generic	$p_{i+1} \times D_{n,t} \le S_{n,t} < p_i \times D_{n,t}$	$c_{s_i} \le c_{s_{n,t}} < c_{s_{i+1}}$

Note: Service levels are for (1) quantity based on supply $S_{n,t}$ and higher and lower acceptable (h and l for proposed service levels) or generic (p_{i+1} and p_i for any level L_i) percentages of demand $D_{n,t}$ being met; and (2) quality of supply based on concentration $c_{s_{n,t}}$ of any substance s and excessive and permitted by legislation (c_{s_e} and c_{s_p} for proposed service levels) or generic ($c_{s_{i+1}}$ and c_{s_i} for any level L_i) concentrations of substance supplied.

quality, in view of customer safety. Both biological and chemical contaminations directly affect human health, so, based on their toxicity, different minimum levels are applied. Assuming an excessive concentration can be identified, a state of emergency with potential life losses could be triggered. Although such concentrations are rare, a low-probability, high-impact scenario should be properly profiled. Such excessive concentration (e.g., a substance's lethal concentration LC_{50}) is the lowest threshold for L_1 and is denoted by c_{s_e}. It is worth noting that such is usually the threshold that many regulations demand supply interruption or use restriction. Failure state L_2 is less critical but can still affect customers' well-being, with a lower concentration threshold c_{s_p}. In technical details, it can never exceed the concentration at which 0% of the population is expected to die (LC_0) or the Maximum Contaminant Level Goal (MCLG) for which there is no known or expected risk to customers' health. Any supply with a concentration below the maximum permissible value c_{s_p} is considered safe to consume (L_3). As previously, the generic quality service level L_i can be found in Table 1.

Third Step: Failure Dimensions

The very role and interests of each stakeholder shape and orient their perception of the system. Different viewpoints of system service failures can also be recognized by observing the more technical and network-oriented view by the company and the behavior of external stakeholders during emergencies. Though motives may differ, rationales converge on the overarching objective of restoring integrity. Therefore, service integrity dimensions, as found in CEN-EN 15975-2 (CEN 2013), can provide a suitable metric spectrum. Under optimal conditions, a system is expected to provide sufficient quantity and quality of water, continuously meeting customer needs and regulatory expectations for the entire network. Reversing this definition of ideal performance, the failure metric spectrum can be divided and categorized into four dimensions related to *service*, *spatial*, *social*, and *continuity* aspects. Those dimensions were chosen because they sequentially answer four key questions: How much service is lost? At what spatial extent? How many customers are affected? For how long?

Service metrics refer to the physical dimension of failure for supply ($S_{L_{i,n,t}}$) found within the range of any previously defined service level (L_i). Thus, $S_{L_{i,n,t}}$ is the notation for every supply that meets the criteria $S_{n,t} \in L_i | c_{s_{n,t}} \in L_i$. In terms of quantity, failure at each level of service is seen through unmet demand ($UD_{L_i,n,t}$), which is the difference between demand and supply as in Eq. (1). In quality-related problems, the physical dimension of failure $PS_{L_{i,n,t}}$ is identical to the supply $S_{L_{i,n,t}}$ found in the corresponding level:

$$UD_{L_i,n,t} = D_{n,t} - S_{L_{i,n,t}} \quad (1)$$

Spatial metrics demonstrate the extent of a scenario's impacts in terms of affected nodes. Each node can represent a single connection (high resolution) or an entire district with a number of blocks (skeletonized network). Regardless of the degree of skeletonization, the spatial extent of failure is related to the connectivity and dynamics of the system. Supply nodes that are affected by an event can be described by a logical index ($N_{L_i,t}$) for both services (1 affected, 0 not affected) as in Eq. (2):

$$N_{L_i,n,t} = \begin{cases} 1 & \text{if } S_{n,t} \in L_i || c_{s_{n,t}} \in L_i \\ 0 & \text{if } S_{n,t} \notin L_i | c_{s_{n,t}} \notin L_i \end{cases} \quad (2)$$

The spatial dimension of failure is critical in estimating the loss of service coverage, identifying cascading behavior, and examining potential risk reduction measures designed to isolate or decrease the failure's propagation.

Social metrics attempt to represent the social impact of an event by addressing failure in terms of people affected. Spatial metrics are unable to address this aspect since population density significantly varies from one area to another. In addition, a real community is an ever-shifting system, dynamic in its internal flows, with people working, living, or entertaining in different areas within the community web. The WDN, with its fixed nodes, must change accordingly. Exact knowledge of the number of people served in each node is impossible, so this temporal change at the local level is unveiled via demand curves. A rough estimation of customers ($C_{n,t}$) can be exported by assuming a *per capita consumption* D_{pc}, using Eq. (3):

$$C_{n,t} = \frac{D_{n,t}}{D_{pc}} \quad (3)$$

The spatiotemporal distribution of customers in the model directly affects multiple hydraulic characteristics, e.g., tank available storage, refill time and pump speed. Such values, which change over time, can assist in composing a more representative failure profile during different critical hours, e.g., peak demand hours (more people affected) and night hours (fewer people affected). The number of customers affected per service level is found using Eq. (4):

$$C_{L_i,n,t} = \begin{cases} C_{n,t} & \text{if } S_{n,t} \in L_i || c_{s_{n,t}} \in L_i \\ 0 & \text{if } S_{n,t} \notin L_i | c_{s_{n,t}} \notin L_i \end{cases} \quad (4)$$

Continuity metrics constitute the last dimension and relate to the duration of failure. The time dimension is crucial in risk management processes because it indirectly defines a level of importance in terms of exposure. Inarguably, spatial expansion, for instance, of high microbial load, to N nodes for 1 h becomes more severe, to the same extent, when it occurs for 2, 4, or 8 h. Time of failure for any level can be defined using a logical index as in Eq. (5):

$$T_{L_i,n,t} = \begin{cases} 1 & \text{if } S_{n,t} \in L_i || c_{s_{n,t}} \in L_i \\ 0 & \text{if } S_{n,t} \notin L_i | c_{s_{n,t}} \notin L_i \end{cases} \quad (5)$$

Since simulation time steps can vary from seconds to minutes to hours and so forth, the foregoing logical index is multiplied by the simulation time step Δt_t, at time t, to produce the physical dimension of failure duration. Such a dimension of information can be used as a guide to select actions that allow the system to recover faster from critical service failures by minimizing the duration of disrupted services (L_1). An overall schematic representation of the failure profiling structure described in this section can be seen in Fig. 3.

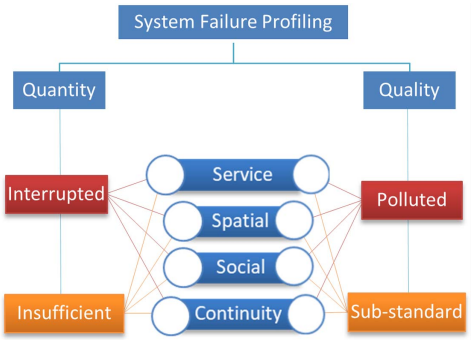

Fig. 3. Schematic representation of the three-step profiling approach.

Quantifying Failure

The previously defined four categories of metrics are the dimensions of failure manifestation for which failure needs to be quantified. Let us assume a CPA occurring in a system affecting it at time t_e. The attack affects performance, with loss of performance building up to a maximum point (t_p), while recovery actions taking place at time t_d start restoring performance up to a recovered state. Such a generic failure curve, inspired by EPA (2015), is represented by Fig. 4, which strongly resembles the shape of a flood hydrograph, with the CPA taking the place of the rain event. Just as the shape and size of the flood hydrograph are affected by rainfall and basin characteristics, so the CPA and water system characteristics affect the form of the failure curve.

With this analogy in mind, one observes that key information found in flood hydrographs can also be defined for the case of a CPA on a water system. The rising limb of the hydrograph, where runoff gradually increases up to peak flow, is seen as the failure from the undistorted state up to a peak value (t_e, t_p), while the receding limb, where flood discharge decreases down to basic flow again, is represented by the system recovering (t_r). It is possible to adopt a new, intermediate state prior to full restoration, similarly to the new base flow. Following this analogy, the failure is explored under a series of lenses to interpret the system's predicament, as presented in the following subsections.

Magnitude

Following the flood analogy, the first lens through which failure in every dimension (i.e., *service*, *spatial*, *social*, and *continuity*) can

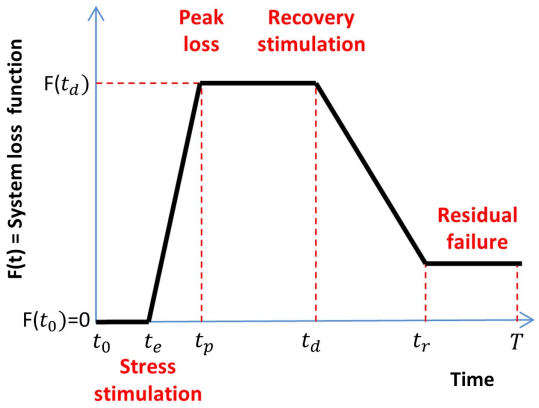

Fig. 4. Generic loss curve under the effect of stress events and resilience actions.

be viewed is magnitude, as total volume equivalent. Magnitude directly conveys the size of failure as total unmet demand (UD_{L_i}) or poor quality supply (PS_{L_i}), nodes (N_{L_i}) and customers affected (C_{L_i}), and total duration of failure (T_{L_i}).

Service metrics magnitude, for any level (L_i), as an absolute number or percentage against expected optimum performance, can be calculated using Eqs. (6)–(9) and quantify failure in terms of its overall size:

$$UD_{L_i} = \int_{t_0}^{T} \int_{n=1}^{N} UD_{L_i,n,t} \tag{6}$$

$$UD_{\%_{L_i}} = \frac{\int_{t_0}^{T} \int_{n=1}^{N} UD_{L_i,n,t}}{\int_{t_0}^{T} \int_{n=1}^{N} D_{n,t}} \times 100\% \tag{7}$$

$$PS_{L_i} = \int_{t_0}^{T} \int_{n=1}^{N} S_{L_i,n,t} \tag{8}$$

$$PS_{\%_{L_i}} = \frac{\int_{t_0}^{T} \int_{n=1}^{N} S_{L_i,n,t}}{\int_{t_0}^{T} \int_{n=1}^{N} D_{n,t}} \times 100\% \tag{9}$$

In terms of the magnitude of spatial manifestation, it is recognized that failure cannot be expressed as the integral of the nodes' loss function [Eq. (2)] because service coverage does not present cumulative behavior over time. The total number of nodes affected is in fact the number of nodes that experience service disturbance, according to each level, within the given timeframe. Thus, one can define a vector of affected nodes for the entire simulation duration T through

$$N_{L_i,n} = \begin{cases} 1 & \text{if } \exists t{:}N_{L_i,n,t} = 1 \\ 0 & \text{if } \nexists t{:}N_{L_i,n,t} = 1 \end{cases} \tag{10}$$

Following Eq. (10), one can now calculate the total spatial expansion of failure and the percentage of system compromised:

$$N_{L_i} = \sum_{n} N_{L_i,n} \tag{11}$$

$$N_{\%_{L_i}} = \frac{\sum_n N_{L_i,n}}{N} \times 100\% \tag{12}$$

The same principle holds for the customer dimension. Affected population for each node is defined as the maximum number of customers that experienced failure in that node [Eq. (13)] at any point in time. One can argue that customers in time t_1 may not be the same as those in t_2 for a node. But this practical assumption is unavoidable, because in practice it is impossible, or at least impractical, to track individual customers for the duration of an event:

$$C_{L_i,n} = \begin{cases} \max\limits_{t_0 \le t \le T} C_{L_i,n,t} & \text{if } \exists t{:}C_{L_i,n,t} = C_{n,t} \\ 0 & \text{if } \nexists t{:}C_{L_i,n,t} = C_{n,t} \end{cases} \tag{13}$$

Having estimated the affected population per node $\in [1, N]$, the total number of customers affected can be determined next. Note that while referring to the affected population, the authors refrain and strongly advise against the use of percentages. Such a metric could, unintentionally, be misunderstood if not coupled with its absolute size, especially in cases of microbial or chemical incidents:

$$C_{L_i} = \sum_{n} C_{L_i,n} \tag{14}$$

For continuity, duration of failure is the total time the system services below expectations even 1 node. Thus the continuity vector per time step can be defined as follows:

$$T_{L_i,t} = \begin{cases} 1 & \text{if } \exists n{:}N_{L_i,n,t} = 1 \\ 0 & \text{if } \nexists n{:}N_{L_i,n,t} = 1 \end{cases} \tag{15}$$

As mentioned, reducing the duration of a critical level failure (L_1) can positively reflect on the system's integrity. Failure duration is calculated through Eqs. (16) and (17) as a percentage of service hours:

$$T_{L_i} = \sum_{t} T_{L_i,t} \times \Delta t_t \tag{16}$$

$$T_{\%_{L_i}} = \frac{\sum_t T_{L_i,t} \times \Delta t_t}{T} \times 100\% \tag{17}$$

To the magnitude-related metrics is added one more, which is widely used in the sector—customer minutes lost (CML). This is the cumulative sum of customers experiencing 0-supply conditions (L_1) times the duration of failure in minutes. Note that this metric is the only metric that depends on the use of specific time units, so proper attention should be paid to unit conversions. Incorporating this metric into the proposed approach requires the introduction of the metric to multiple service levels. This paper proposes a new variation of that metric (CM_{L_i}), presented in the following equation, where $\Delta t_{\min_t}$ is the time step at time t in minutes:

$$CM_{L_i} = \sum_{n,t} C_{L_i,n,t} \times T_{L_i,n,t} \times \Delta t_{\min_t} \tag{18}$$

For disrupted services (L_1), Eq. (18) refers to the known CML, while at the degraded level (L_2) it represents the cumulative exposure to disturbing conditions, as improperly serviced customer minutes. The latter can be a metric concerning the exposure of the company to "reputational damage" for noncritical levels of failure. At this point recall that this metric also applies to quality-related incidents, revealing the cumulative exposure at certain levels of concentration of a chemical or microbial load.

Average Propagation

The second lens through which quantification is proposed is that of average propagation. The effects of a CPA propagate through the system over time, spreading and amplifying the overall failure. For a given dimension, the propagation over time can be seen through the failure curves. Failure curves can have shapes that are more complex than generic curves (Fig. 4), with multiple crests and valleys, demonstrating the dynamic evolution of failure. With magnitude metrics demonstrating the final size of failure, it is useful to also create a snapshot to represent the failure's average propagation. To do this, the arithmetic average is proposed, under the condition of nonzero values. Such a condition ensures the calculation of the metrics only for the duration of failure. It protects the accuracy and comparability of metrics, since identical events under different simulation durations would otherwise yield different results. For each of the service dimensions, the average propagation of failure is determined via the following equations:

$$\overline{UD_{L_i}} = \frac{\sum_{n,t} UD_{L_i,n,t}}{\sum_{n,t} T_{L_i,n,t}} \tag{19}$$

$$\overline{PS_{L_i}} = \frac{\sum_{n,t} S_{L_i,n,t}}{\sum_{n,t} T_{L_i,n,t}} \tag{20}$$

$$\overline{N_{L_i}} = \frac{\sum_{n,t} N_{L_i,n,t}}{\sum_t T_{L_i,t}} \tag{21}$$

$$\overline{C_{L_i}} = \frac{\sum_{n,t} C_{L_i,n,t}}{\sum_t T_{L_i,t}} \tag{22}$$

In terms of continuity, the average propagation of failure over time should be reflected in a manner that weights the importance of the average time of exposure. Average exposure time per customer is a social and continuity combinatory metric that is used, for example, to assess potential health hazards from exposure to a chemical. The average per-customer exposure is estimated using Eq. (23):

$$\overline{CT_{L_i}} = \frac{\sum_{n,t} C_{L_i,n,t} \times T_{L_i,n,t} \times \Delta t_t}{C_{L_i}} \tag{23}$$

Equivalently, to address the need for network continuity–related metrics that useful for water company stakeholders, the average failure per node can also be used. Note that the units of the time step are the units of the combinatory metrics:

$$\overline{NT_{L_i}} = \frac{\sum_{n,t} T_{L_i,n,t} \times \Delta t_t}{N_{L_i}} \tag{24}$$

Such metrics may look simple but can prove effective tools, not only for exploring different scenario configurations and mitigation plans but also for comparing effects at different parts of the system.

Severity

However, magnitude and average propagation do not uniquely define an event. As in the case of two flood events where the total runoff volume and average per time runoff can be similar, while the actual events could differ greatly in terms of severity when one of the two has a much higher peak discharge. In the same analogy, severity in the case of a CPA on a water system can be identified through peak effects. Peak temporal effects can be identified for three of the service dimensions, albeit not for continuity, as seen in Eqs. (25)–(28):

$$UD_{peak_{L_i}} = \max_{t_0:T} \sum_{n=1}^{N} UD_{L_i,n,t} \tag{25}$$

$$PS_{peak_{L_i}} = \max_{t_0:T} \sum_{n=1}^{N} S_{L_i,n,t} \tag{26}$$

$$N_{peak_{L_i}} = \max_{t_0:T} \sum_{n=1}^{N} N_{L_i,n,t} \tag{27}$$

$$C_{peak_{L_i}} = \max_{t_0:T} \sum_{n=1}^{N} C_{L_i,n,t} \tag{28}$$

Peak temporal failure can be considered a measure of an event's severity. Besides allowing a comparison between attack scenarios, these metrics can also capture a measure's ability to blunt the peak effect and create a ceiling on failure propagation.

Crest Factor

As failure propagates through a system (Fig. 4), it rises to its climax before moving to lower states. Though the climax itself is important, the rate of change in performance is also of interest. To continue with the flood analogy, it is useful to capture the difference between a gradual flood and a rapidly occurring flash flood. This can be assessed using a peak-to-average ratio (PAR) metric, which helps detect whether performance changes occur abruptly or gradually. PAR is a crest factor defined as the ratio between the peak effect and the average propagation (Rouphael 2009). PAR indicates how extreme peaks are, as failure escalation is translated to higher ratio values. Representing Eqs. (25)–(28) of peak failure as $F_{peak_{L_i}}$ and propagation Eqs. (19)–(22) as $\overline{F_{L_i}}$, PAR equations for any level and dimension are derived as follows:

$$PAR_{F_{L_i}} = \frac{F_{peak_{L_i}}}{\overline{F_{L_i}}} \tag{29}$$

PAR metrics are magnitude-independent. As PAR approaches 1, failure propagation exhibits a more uniform profile, while larger values imply a spikelike failure, indicative of disastrous major events that tend to abruptly affect the system.

Rapidity

Another important metric, in view of crisis-management resource allocation and mobilization, is one that quantifies a failure's rapidity. To capture this important aspect, a "time to peak" metric is proposed, specifically the interval from the beginning of an event in t_e to its climax in t_p, which will be referred to henceforth as the time from event to peak (TEP). The TEP is equivalent to the lag time in the flood hydrograph analogy, demonstrating the required time for peak discharge to occur after precipitation. The TEP is defined through Eq. (30), where $t_{p_{F_{L_i}}}$ is the occurrence of peak effect in the F_{L_i} specified failure dimension:

$$TEP_{F_{L_i}} = t_{p_{F_{L_i}}} - t_e \tag{30}$$

TEP quantifies the available time for stakeholders to react before the peak effect of an attack occurs and as such is important for emergency planning. A similarly defined metric associated with some user-defined "critical condition" can be used to act as a trigger for crisis management plans to be set in motion. This critical condition occurs at time t_{cr} in the generic loss curve (Fig. 4) after the event's occurrence and prior to t_d. This is defined here as the time from event to critical (TEC), which is calculated as follows, where $t_{Cr_{F_{L_i}}}$ is the occurrence time of the user-defined critical condition $Cr_{F_{L_i}}$ in the F_{L_i} specified failure dimension:

$$TEC_{F_{L_i}} = t_{Cr_{F_{L_i}}} - t_e \tag{31}$$

To summarize, the proposed methodology quantifies a water system's failure in terms of its magnitude, average propagation, severity, and peak-to-average ratio and identifies available reaction times. The approach and its equations are designed to work with different definitions of failure based on (user-defined) service levels.

Approach Deployment

Risk assessment for the emerging cyber-physical water domain is expected to become an integral component in water companies' workflow (Makropoulos and Savíc 2019). Nonetheless, operationalizing and integrating new (potentially disruptive) approaches can be challenging. To assist in this transition, the methodology presented in this paper is encapsulated as a standalone MATLAB compiled tool. The tool processes simulation data, of a specific format, regardless of the stress-test model used, and generates the failure profile for the given scenario based on user-defined criteria. Additional functionalities are also introduced and briefly presented next.

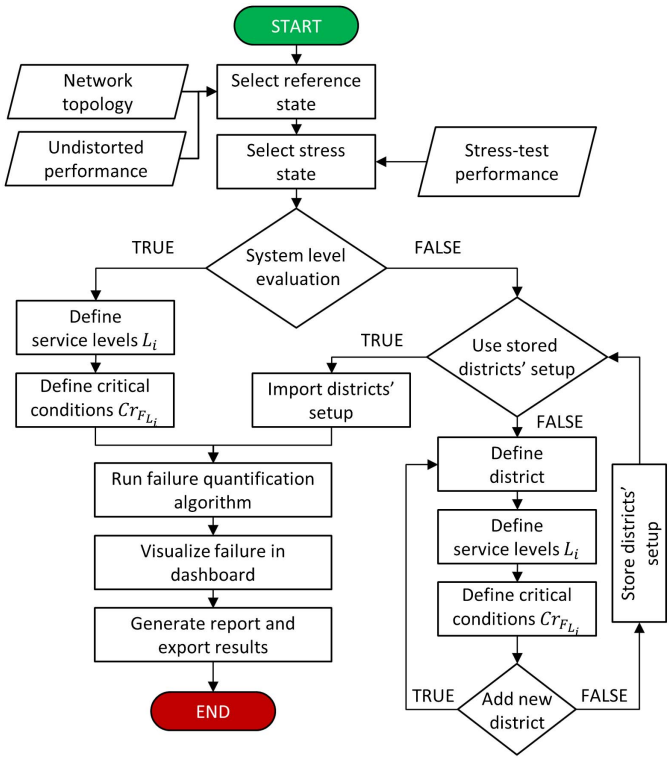

Fig. 5. Failure quantification tool simplified flowchart.

Consistent with the scope described in the section "Failure Oriented Approach," the tool focuses on profiling the failure area appearing between the ideal and the under-stress performance of the system, based on user-defined service levels. This process can be seen in the simplified flowchart of the tool in Fig. 5. The hydraulic network topology (EPANET.inp file) and data that compose the undistorted, attack-free, performance are used as the reference state. The network file provides information on units used and reports layout information for validation purposes, to ensure that performance data refer to the same topology. The stress-test results of the CPA scenario to be profiled are the next required input, reporting performance data per asset. Currently the tool supports.csv formats performance data tables that provide variable names and corresponding asset IDs. For example, the expected demand of Node N1 under normal conditions, for each time step, is found in the undistorted performance file column labeled "Demand_N1." User-defined service levels L_1 and L_2 and critical conditions $Cr_{F_{L_i}}$ are specified using the tool's interactive interface.

WDNs are typically zoned into districts with fixed boundaries, known as district metered areas (DMAs), primarily to control leakages and regulate supply pressure and quality (Butler et al. 2005). Within such districts, service is provided to various customers, some of which can be considered critical based on the (societal) impact a disruption of service to them can cause (e.g., hospitals, government, and military buildings). In this spirit, the tool provides network sectorization capabilities to quantify failure and evaluate CPAs at the district level. Different thresholds can be assigned per district to signify the importance or specific performance requirements for each DMA. Such capability allows for additional adjustability with spatial customization of failure quantification, to account for critical parts of the serviced community and support a more realistic assessment of cyber-physical consequences. After defining the service levels and critical conditions through the tool's interface, at either the system or district level, the core algorithm can be deployed and results can be visually explored through the tool's visualization capabilities. An example can be seen in Fig. 6, which shows the main dashboard, where key failure information can be explored. Peak effects and available time slots to react are also visualized in the tool. The tool profiles and represents failure using preferred simulation parameters; thus, both SI and customary US units are supported.

To assist the usability and communication of risk-related information, which is a very important aspect of rapid situational awareness, the tool supports two export capabilities. The first export capability produces a JSON file, which contains all failure profile data for both system and district levels and a set of metadata. This file can then be stored in a database or used as input in other software. The second export capability comes in the form of a report. The report-generating module of the tool includes human readable files, organizing numerical information into structured word templates with supporting figures to increase comprehension and provide a faster, easier, and more standardized way to communicate information. The modular design of the report generator allows for further customization, including customizing the reports' language to facilitate local operators.

Case Study

Case Study Configuration

To demonstrate the proposed failure quantification approach implemented through the purpose-built tool, a case study was conducted where a WDN was subjected to a CPA. The network selected was the well-known C-Town benchmark network, an EPANET network introduced by Ostfeld et al. (2012) that is composed of 388 demand nodes and 7 tanks linked with 429 pipes, 11 pumps, and 4 valves to a single source (1 reservoir) divided into 5 DMAs (Fig. 7). The cyber network is composed of 9 PLCs, linked to 7 tank stage sensors, and 12 actuators across the WDN, based on a configuration proposed by Taormina et al. (2017). To generate the data on the effects of CPA needed to demonstrate the quantification approach that was used, one of the CPA scenarios introduced in Taormina et al. (2017) was adopted and simulated using the same tool: the EpanetCPA toolbox. This CPA scenario is explained in what follows.

C-town actuators are operated by the main SCADA based on the water levels reported by the sensors found in each of the seven tanks. For example, tank T2 and control valve V2 are operationally linked through such a control logic. Based on the reported T2 level sensor readings, the SCADA updates the status of valve V2 and remotely opens it when a low level (<0.5 m) is detected. Or, when the tank sensor reports high water levels (>5.5 m), the SCADA closes the valve. For this case study, it is assumed that a threat actor, after breaching the system and repeatedly obtaining sensor and actuator status data, has leveraged operational knowledge over this refill protocol. This allows him to perform an attack that alters the T2 stage readings, introducing bogus input data (Zhu et al. 2011) that lead the assigned PLC to (erroneously) believe the tank is full. This information is then reported back to the SCADA, deceiving it and preventing it from activating the refill protocol by requesting V2 status to be "closed." This CPA scenario is expected to cause supply problems to the districts supplied by tank T2. It is assumed that the false sensor reading attack begins at time $t_e = 5$ h and has a duration of 15 h. To produce a more realistic behavior in the network, a PDA approach was selected using the NHFR formula of Wagner et al. (1988). For the purposes of estimating the number of customers based on Eq. (3), a per-capita consumption D_{pc} of

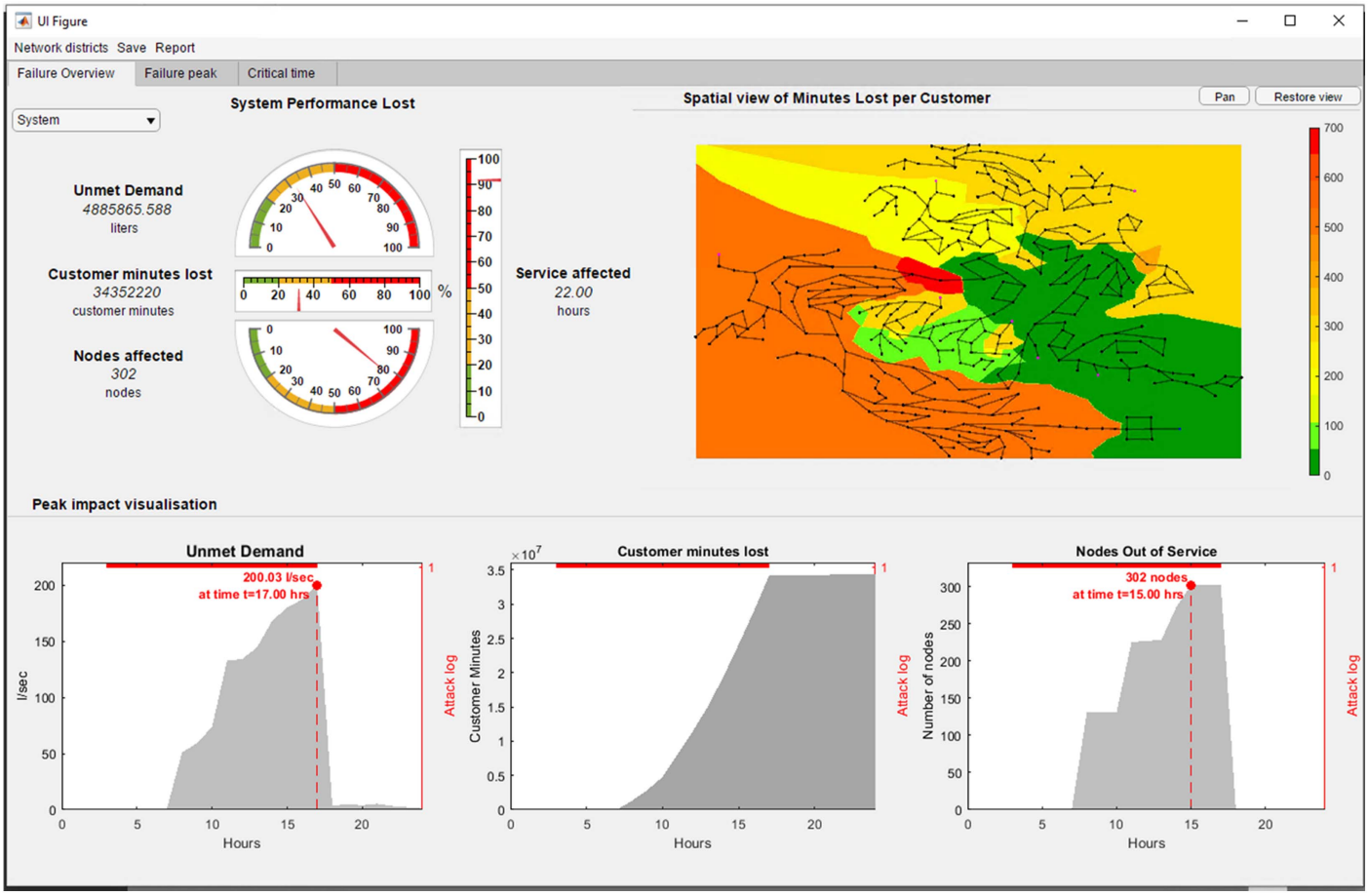

Fig. 6. Main dashboard tab for a CPA affecting quantity of supply.

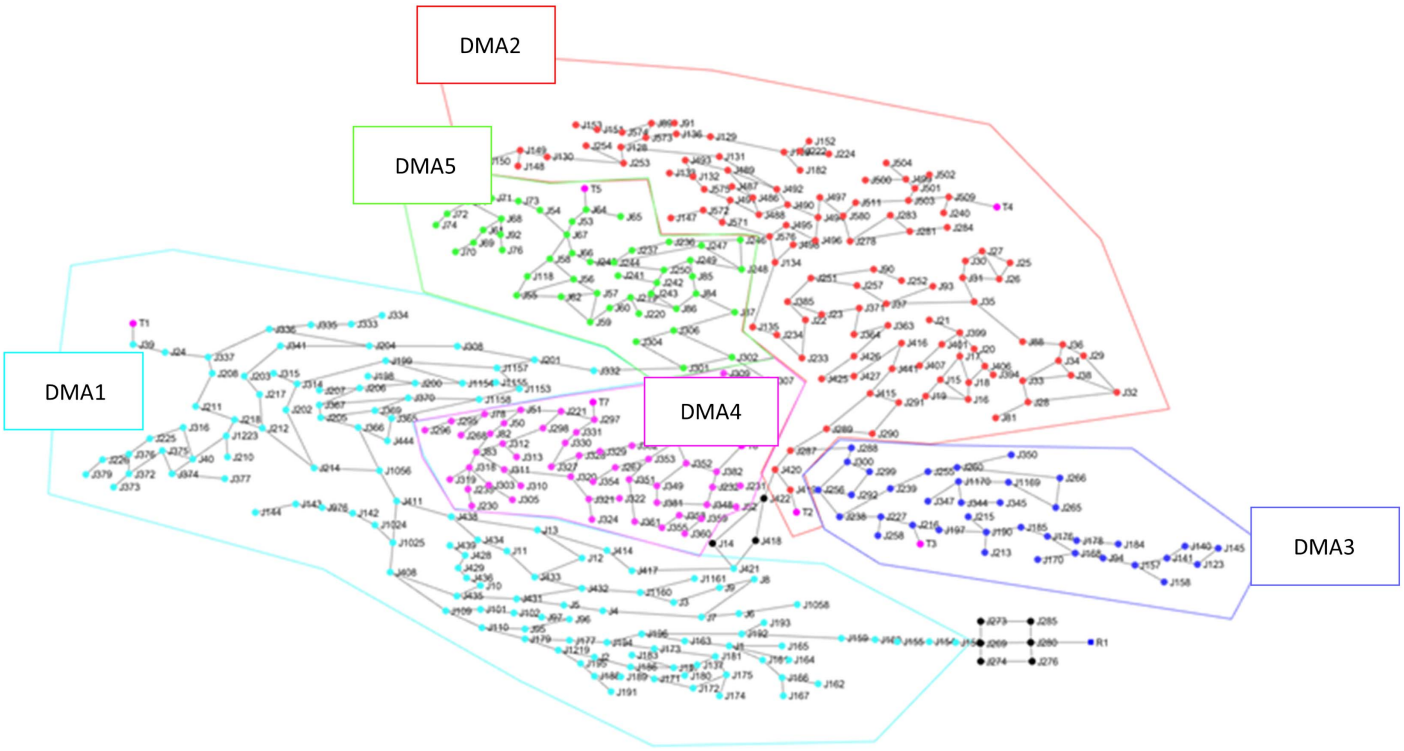

Fig. 7. C-Town DMAs compartmentalization as created by the tool and used for failure quantification.

Table 2. Service level thresholds and critical condition at system and DMA scales for C-Town

Analysis scale	Service level		Critical conditions		
	$L1$	$L2$	cr_{UD} (%)	cr_N (%)	$cr_{C_{L1}}$
System	$0 \le S_{n,t} < 15\% \times D_{n,t}$	$15\% \times D_{n,t} \le S_{n,t} < 90\% \times D_{n,t}$	10	10	1,250
DMA1	$0 \le S_{n,t} < 20\% \times D_{n,t}$	$20\% \times D_{n,t} \le S_{n,t} < 90\% \times D_{n,t}$	7.5	10	5,000
DMA2	$0 \le S_{n,t} < 22.5\% \times D_{n,t}$	$22.5\% \times D_{n,t} \le S_{n,t} < 95\% \times D_{n,t}$	5	5	5,000
DMA3	$0 \le S_{n,t} < 15\% \times D_{n,t}$	$15\% \times D_{n,t} \le S_{n,t} < 90\% \times D_{n,t}$	7.5	5	5,000
DMA4	$0 \le S_{n,t} < 20\% \times D_{n,t}$	$20\% \times D_{n,t} \le S_{n,t} < 98\% \times D_{n,t}$	7.5	5	5,000
DMA5	$0 \le S_{n,t} < 15\% \times D_{n,t}$	$15\% \times D_{n,t} \le S_{n,t} < 90\% \times D_{n,t}$	7.5	5	5,000

200 L/day is assumed. Failure quantification using the proposed approach was conducted both on the entire system and at the DMA scale. Service level thresholds [disrupted (L_1) and degraded (L_2)], defined for the system and per DMA, are shown in Table 2. Conditions critical for the examination of failure rapidity and available reaction times are also defined in the same table.

Results

The attack led to low pressure conditions and supply deficiency to parts of C-Town. Failure in terms of magnitude for all four dimensions (service, spatial, social, and continuity) can be found in Table 3.

The results suggest that C-Town was unable to meet nearly 12% of the total daily demand to 28.6% of the network due to interruptions. This resulted in 31,571 customers being cut off during the 11 h of failure. Note that, although the system faces degraded supply conditions for a longer period of time, its spatial expansion is limited, while high values for CM_{L_1} indicate a significant service outage. The main impact was felt in DMA2, which saw 47.5% of its supply lost owing to complete supply interruption (L_1). Some damage also spread to DMA3, but only partially, while the rest of the system remained intact. In fact, one can clearly see in Fig. 8(b) that DMA3 started to fail shortly before the attack ended, allowing the PLC to detect the low stage and react. The refill process caused pump activation and a rise in pressure, which restored supply services.

In terms of the failure's average propagation, the results are presented in Table 4. In an average hour of disrupted services, 82 nodes were cut off, unable to meet the demands of 22,814 customers for 52.71 L/s. On average each affected customer of the system experienced 6.48 h of disrupted service. The nearly identical down time per affected node in DMA2 ($\overline{NT_{L_1}} = 9.49$ h) and total disruption duration ($T_{L_1} = 9.5$ h) indicate that the attack affected a fixed area throughout the failure duration, which can be seen in Fig. 9(a).

Table 3. Magnitude of service failure due to CP attack on tank sensor

Failure magnitude	Unit	System	DMA1	DMA2	DMA3	DMA4	DMA5
UD_{L_1}	Liters	1.80×10^6	0	1.74×10^6	62.43×10^3	0	0
UD_{L_2}	Liters	4.76×10^3	0	142.76	0	5.61×10^3	0
N_{L_1}	Nodes	111	0	79	32	0	0
N_{L_2}	Nodes	4	0	2	0	2	0
C_{L_1}	Individuals	31,571	0	24,063	7,508	0	0
C_{L_2}	Individuals	114	0	21	0	93	0
T_{L_1}	Hours	9.5	0	9.5	1	0	0
T_{L_2}	Hours	11	0	11	0	13.5	0
CM_{L_1}	Customer minutes	13,004,400	0	12,553,920	450,480	0	0
CM_{L_2}	Customer minutes	34,800	0	1,080	0	41,130	0

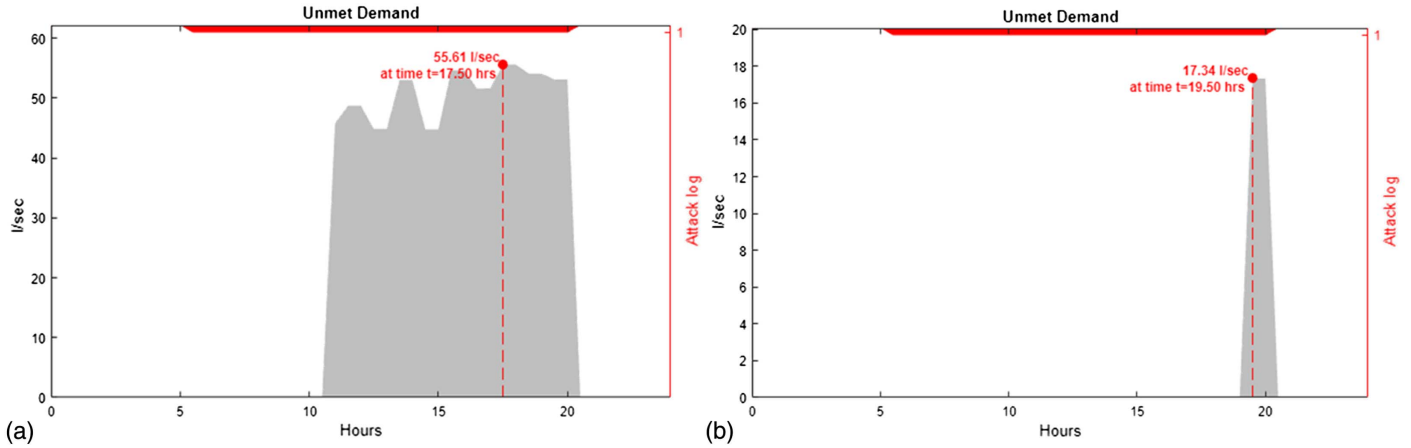

Fig. 8. Unmet demand time series and peak effect for (a) DMA2; and (b) DMA3 and 15-h attack log (top).

J. Environ. Eng., 2020, 146(9): 04020108

Table 4. Propagation of service failure due to CP attack on tank sensor

Failure propagation	Unit	System	DMA1	DMA2	DMA3	DMA4	DMA5
$\overline{UD_{L_1}}$	L/s	52.71	0	50.89	17.34	0	0
$\overline{UD_{L_2}}$	L/s	0.14	0	0.02	0	0.11	0
$\overline{N_{L_1}}$	Nodes	82.31	0	78.94	32	0	0
$\overline{N_{L_2}}$	Nodes	2	0	1	0	2	0
$\overline{C_{L_1}}$	Individuals	22,814	0	22,204	7,508	0	0
$\overline{C_{L_2}}$	Individuals	64.44	0	9	0	50.78	0
$\overline{CT_{L_1}}$	Hours	6.48	0	8.7	1	0	0
$\overline{CT_{L_2}}$	Hours	3.23	0	0.68	0	7.09	0
$\overline{NT_{L_1}}$	Hours	7.04	0	9.49	1	0	0
$\overline{NT_{L_2}}$	Hours	4.37	0	1	0	13.5	0

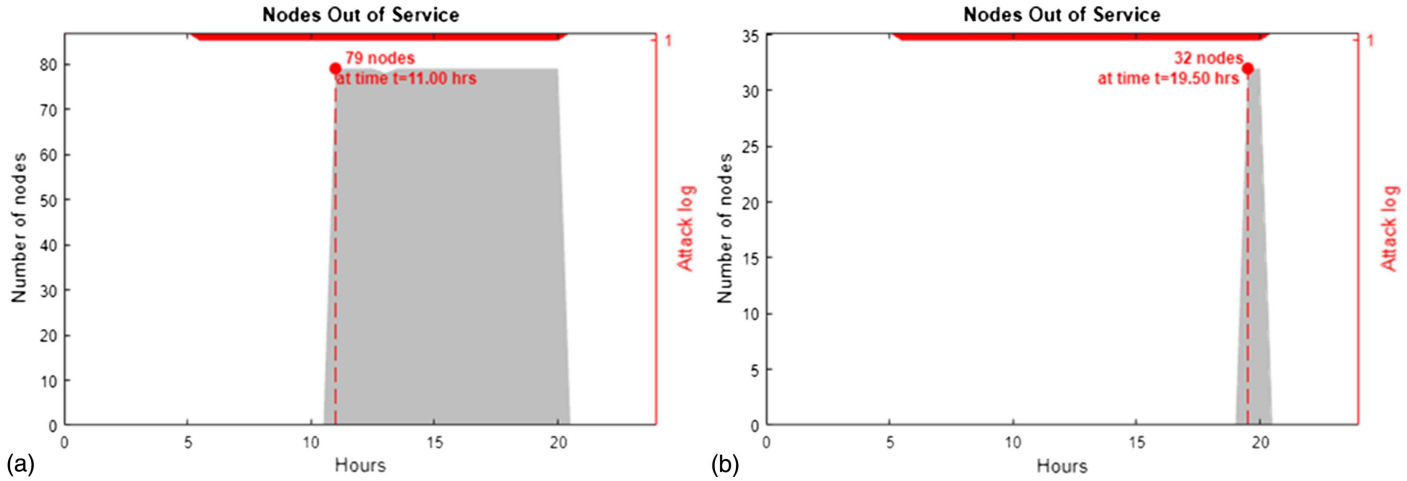

Fig. 9. Nodes out of service time series and peak effect for (a) DMA2; and (b) DMA3 and 15-h attack log (top).

Table 5. Peak of service failure and peak-to-average ratio due to CP attack on tank sensor

Failure climax	Unit	System	DMA1	DMA2	DMA3	DMA4	DMA5
$UD_{peak_{L_1}}$	L/s	70.66	0	55.61	17.34	0	0
$UD_{peak_{L_2}}$	L/s	0.21	0	0.04	0	0.21	0
$N_{peak_{L_1}}$	Nodes	111	0	79	32	0	0
$N_{peak_{L_2}}$	Nodes	3	0	1	0	2	0
$C_{peak_{L_1}}$	Individuals	30,504	0	24,063	7,508	0	0
$C_{peak_{L_2}}$	Individuals	93	0	20	0	93	0
Peak-to-average							
$PAR_{UD_{L_1}}$	—	1.34	—	1.09	1	—	—
$PAR_{UD_{L_2}}$	—	1.5	—	2	—	1.9	—
$PAR_{N_{L_1}}$	—	1.34	—	1	1	—	—
$PAR_{N_{L_2}}$	—	1.54	—	1	—	1	—
$PAR_{C_{L_1}}$	—	1.34	—	1.09	1	—	—
$PAR_{C_{L_2}}$	—	1.44	—	2.22	—	1.83	—

Severity was then calculated as a snapshot of the system at its highest stress point (Table 5). In addition to the absolute number of the peak effect for each dimension, the PAR metrics are also calculated and shown in the second part of Table 5. At the system level one can see a rather uniform profile, revealing a steadily evolving failure. The failure seems to gradually (as opposed to abruptly) add load to the system after its manifestation, as seen through the PAR metrics. This can be seen in Fig. 10, where the time series of unmet demand for the system are plotted.

At the district level, DMA2 PAR metrics ($PAR_{UD_{L_1}} = 1.09$, $PAR_{N_{L_1}} = 1$, $PAR_{C_{L_1}} = 1.09$) indicate a smoothly propagating failure, with no extreme peaks, as seen in Figs. 8(a) and 9(a).

The rapidity of propagation is calculated from the attack's occurrence ($t_e = 5$ h) until the peak and critical marks, respectively. Defining the available time until peak effect of the attack occurs creates the first and largest possible time window for mitigation actions to become effective. In the absence of effective actions, disruption of supply seems to gradually rise and spread, as $TEP_{UD_{L_1}}$

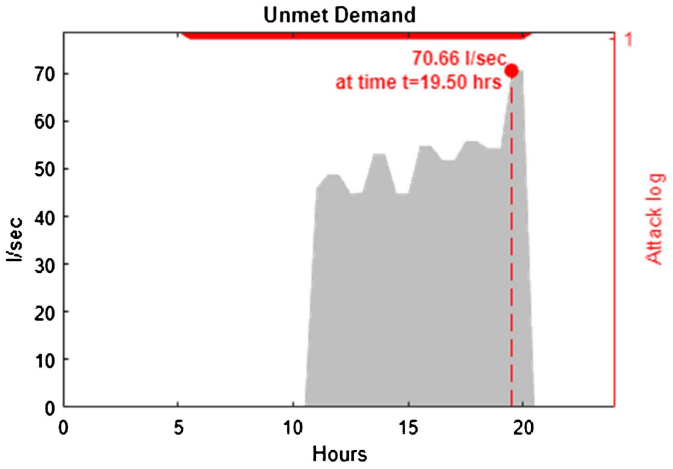

Fig. 10. Unmet demand time series and peak effect at system level and 15-h attack log (top).

Table 6. Time to peak and time to critical conditions due to CP attack on tank sensor

Time to peak	Unit	System	DMA1	DMA2	DMA3	DMA4	DMA5
$TEP_{UD_{L_1}}$	Hours	14.5	—	12.5	14.5	—	—
$TEP_{UD_{L_2}}$	Hours	8	—	5.5	—	8	—
$TEP_{N_{L_1}}$	Hours	14	—	6	14.5	—	—
TEP_{N_2}	Hours	7.5	—	5.5	—	5.5	—
$TEP_{C_{L_1}}$	Hours	14.5	—	12.5	14.5	—	—
$TEP_{C_{L_2}}$	Hours	8	—	5.5	—	8	—
Time to critical							
$TEC_{UD_{L_1}}$	Hours	5.5	—	5.5	14	—	—
$TEC_{N_{L_1}}$	Hours	5.5	—	5.5	14	—	—
$TEC_{C_{L_1}}$	Hours	5.5	—	5.5	14	—	—

and $TEP_{N_{L_1}}$ approach attack duration. This was seen previously, with failure cascading to DMA3 shortly before the end of the attack. With respect to the entire system, the water company has 8 h before maximum service disruption. The maximum disruption of service to DMA2 occurs 12 h after the attack, but the company only has 5.5 h to deploy the necessary resources before the maximum spatial spread to the subsystem. This is almost 2.5 times faster than the failure occurrence to DMA3, raising a priority flag in terms of available time.

The identical times to critical (TEC) found in Table 6 for all three conditions at the system and DMA levels are the result of lower critical condition thresholds determined and the C-Town dependence on tanks to regulate nodal pressure, which allows for immediate disruption once the tank is empty. Prioritization of actions for DMA2 is highlighted in terms of criticality, available time to react, and size of failure, as a result of the proposed methodology implementation.

Conclusions

In this article, a failure quantification methodology is presented that can capture the impact of a water CI under cyber-physical threats and communicate this information to relevant stakeholders. The method is designed to allow for the exploration of different dimensions of a failure's manifestation under user-defined levels of service. Two levels of service are identified, representing the critical levels specified by the regulatory, legislative, or internal operating environment. Since more intermediate service levels may be desired, generic equations to calculate the proposed metrics for any set of service levels are also provided. The method also allows for the identification of available response times as a basis for emergency planning. To support the operationalization of the method, a dedicated tool was developed within the STOP-IT project. The tool allows users to define thresholds, select DMAs, and implement the methodology at any scale (system or DMA), and it provides an additional spatial overview and possesses metric export capabilities. Based on this work, comparison of the effects of different attack types can be showcased for a network, including scenarios for denial of service, bad data injection, or replay attacks, for example. To consider the effect of user-defined thresholds per se, selected either based on experience or following adopted standards, legislation, and so forth, a sensitivity analysis can be conducted. Although both the tool and the methodology are hydraulic solver-agnostic, a purpose-built stress-testing platform is also developed and presented in a companion paper by Nikolopoulos et al. (2020). The intention behind the companion papers is to node towards potential future interactions between the two methods and tools to further advance research on the cyber-physical security of water systems.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

Acknowledgments

This paper was supported by the STOP-IT project. STOP-IT has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement 740610. The publication reflects only the authors' views, and the European Union is not liable for any use that may be made of the information contained therein.

References

Alegre, H., J. M. Baptista, E. C. Jr, F. Cubillo, P. Duarte, W. Hirner, W. Merkel, and R. Parena. 2016. *Performance indicators for water supply services*. London: IWA Publishing.

Almalawi, A., Z. Tari, I. Khalil, and A. Fahad. 2013. "SCADAVT-A framework for SCADA security testbed based on virtualization technology." In *Proc., Conf. on Local Computer Networks*, 639–646. Montreal, QC: LCN.

Andersen, B., and T. Fagerhaug (2002). *Performance measurement explained: Implementing your state-of-the-art system*. Milwaukee: ASQ Quality Press.

ASME-ITI. 2009. *All-hazards risk and resilience: Prioritizing critical infrastructures using the RAMCAP plus approach*. New York: ASME Press.

AWWA (American Water Works Association). 2010. *Risk and resilience management of water and wastewater systems*. AWWA J100-10 (R13). Denver: AWWA.

Bechara, A., H. Damasio, A. R. Damasio, and G. P. Lee. 1999. "Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making." *J. Neurosci.* 19 (13): 5473–5481. https://doi.org/10.1523/JNEUROSCI.19-13-05473.1999.

Berg, S. 2013. "Advances in benchmarking to improve water utility operations: A review of six IWA books." *Water Policy* 15 (2): 325. https://doi.org/10.2166/wp.2012.089.

Borshchev, A., and A. Filippov. 2004. "From system dynamics and discrete event to practical agent based modeling: Reasons, techniques, tools." In *Proc., 22nd Int. Conf. of the System Dynamics Society*, 25–29. Harvard, MA: Ventana Systems.

Bouchon, S., C. Di Mauro, C. Logtmeijer, J. Nordvik, R. Pride, B. Schupp, and M. Thornton. 2008. *Non-binding guidelines for application of the council directive on the identification and designation of European critical infrastructure and the assessment of the need to improve their protection*. Luxembourg: Office for Official Publications of the European Communities.

Bouziotas, D., D. van Duuren, H.-J. van Alphen, J. Frijns, D. Nikolopoulos, and C. Makropoulos. 2019. "Towards circular water neighborhoods: Simulation-based decision support for integrated decentralized urban water systems." *Water* 11 (6): 1227. https://doi.org/10.3390/w11061227.

Butler, D., F. Memon, and L. Seattle. 2005. *Water demand management*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Butler, D., S. Ward, C. Sweetapple, M. Astaraie-Imani, K. Diao, R. Farmani, and G. Fu. 2017. "Reliable, resilient and sustainable water management: The Safe & SuRe approach." *Global Challenges* 1 (1): 63–77. https://doi.org/10.1002/gch2.1010.

Cable, J. H., and J. S. Davis. 2004. *Key performance indicators for federal facilities portfolios: Federal facilities council technical report number 147*. Washington, DC: National Academies Press.

CEN (European Committee for Standardization). 2013. *Security of drinking water supply—Guidelines for risk and crisis management. Part 2: Risk management*. CEN-EN 15975-2. Brussels, Belgium: CEN.

Chen, B., N. Pattanaik, A. Goulart, K. L. Butler-purry, and D. Kundur. 2015. "Implementing attacks for modbus/TCP protocol in a real-time cyber physical system test bed." In *Proc., 2015 IEEE Int. Workshop Technical Committee on Communications Quality and Reliability (CQR)*, 1–6. New York: IEEE.

Chmielewski, H., R. Guidotti, T. McAllister, and P. Gardoni. 2016. "Response of water systems under extreme events: A comprehensive approach to modeling water system resilience." In *World Environmental and Water Resources Congress 2016*, 475–486. Reston, VA: ASCE.

Creaco, E., M. Franchini, and E. Todini. 2016. "The combined use of resilience and loop diameter uniformity as a good indirect measure of network reliability." *Urban Water J.* 13 (2): 167–181. https://doi.org/10.1080/1573062X.2014.949799.

Damasio, A. R., B. J. Everitt, D. Bishop, and A. R. Damasio. 1996. "The somatic marker hypothesis and the possible functions of the prefrontal cortex." *Philos. Trans. R. Soc. London, Ser. B* 351 (1346): 1413–1420. https://doi.org/10.1098/rstb.1996.0125.

Danilenko, A., C. Van der Berg, B. Macheve, and J. Moffitt. 2014. *The IBNET water supply and sanitation blue book*. Washington, DC: World Bank.

Davis, C. M., J. E. Tate, H. Okhravi, C. Grier, T. J. Overbye, and D. Nicol. 2006. "SCADA cyber security testbed development." In *Proc., 2006 38th Annual North American Power Symp., NAPS-2006*, 483–488. Piscataway, NJ: IEEE.

Davis, M. J., and R. Janke. 2018. "The effect of a loss of model structural detail due to network skeletonization on contamination warning system design: case studies." *Drinking Water Eng. Sci.* 11 (1): 49–65. https://doi.org/10.5194/dwes-2017-39.

Eliades, D. G., M. Kyriakou, S. G. Vrachimis, and M. M. Polycarpou. 2016. "EPANET-MATLAB Toolkit: An open-source software for interfacing EPANET with MATLAB." In *Proc., 14th Computer Control for Water Industry Conf., CCWI 2016*, 1–8. Exeter, UK: CCWI.

ENISA (European Union Agency for Cybersecurity). 2019. *ENISA threat landscape report 2018: 15 top cyberthreats and trends*. Heraklion, Greece: ENISA.

EPA. 2015. *System measures of water distribution resilience*. Washington, DC: EPA.

Europol. 2018. *Internet organised crime threat assessment (IOCTA 2018)*. Hague, Netherlands: Europol.

Fovino, I. N., M. Masera, L. Guidi, and G. Carpi. 2010. "An experimental platform for assessing SCADA vulnerabilities and countermeasures in power plants." In *Proc., 3rd Int. Conf. on Human System Interaction (HSI'2010)*, 679–686. New York: IEEE.

Francis, R., and B. Bekera. 2014. "A metric and frameworks for resilience analysis of engineered and infrastructure systems." In *Reliability engineering & system safety*, 90–103. Amsterdam, Netherlands: Elsevier.

Fujiwara, O., and J. Li. 1998. "Reliability analysis of water distribution networks in consideration of equity, redistribution, and pressure-dependent demand." *Water Resour. Res.* 34 (7): 1843–1850. https://doi.org/10.1029/98WR00908.

Galbusera, L., G. Giannopoulos, and D. Ward. 2014. *Developing stress tests to improve the resilience of critical infrastructures: A feasibility analysis*. Luxembourg: Publications Office.

Garofalo, G., A. Giordano, P. Piro, G. Spezzano, and A. Vinci. 2017. "A distributed real-time approach for mitigating CSO and flooding in urban drainage systems." *J. Network Comput. Appl.* 78 (Sep): 30–42. https://doi.org/10.1016/j.jnca.2016.11.004.

Germanopoulos, G. 1985. "A technical note on the inclusion of pressure dependent demand and leakage terms in water supply network models." *Civ. Eng. Syst.* 2 (3): 171–179. https://doi.org/10.1080/02630258508970401.

Giustolisi, O., R. Ugarelli, L. Berardi, D. Laucelli, and A. Simone. 2017. "Strategies for the electric regulation of pressure control valves." *J. Hydroinf.* 19 (5): 621–639. https://doi.org/10.2166/hydro.2017.101.

Grance, T., T. Nolan, K. Burke, R. Dudley, G. White, and T. Good. 2006. "Guide to test, training, and exercise programs for IT plans and capabilities." *No. Special Publication (NIST SP)-800-84*. Gaithersburg, MD: National Institute of Standards and Technology.

Hansson, S. O., and T. Aven. 2014. "Is risk analysis scientific?" *Risk Anal.* 34 (7): 1173–1183. https://doi.org/10.1111/risa.12230.

Hashimoto, T., J. R. Stedinger, and D. P. Loucks. 1982. "Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation." *Water Resour. Res.* 18 (1): 14–20. https://doi.org/10.1029/WR018i001p00014.

Hassanzadeh, A., A. Rasekh, S. Galelli, M. Aghashahi, R. Taormina, A. Ostfeld, and M. K. Banks. 2020. "A review of cybersecurity incidents in the water sector." *J. Environ. Eng.* 146 (5): 03120003. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001686.

Herrera, M., E. Abraham, and I. Stoianov. 2016. "A graph-theoretic framework for assessing the resilience of sectorised water distribution networks." *Water Resour. Manage.* 30 (5): 1685–1699. https://doi.org/10.1007/s11269-016-1245-6.

Holling, C. S. 1996. "Engineering resilience versus ecological resilience." *Eng. Ecol. Constraints* 31 (1996): 32.

Howard, M., J. Pincus, and J. Wing. 2005. *Measuring relative attack surfaces*, 109–137. Pittsburgh: Carnegie Mellon Univ.

ISO. 2018. *Risk management—Principles and guidelines*. ISO 31000. London: ISO.

Janke, R., M. E. Tryby, and R. M. Clark. 2014. *Securing water and wastewater systems*, edited by R. M. Clark and S. Hakim. Berlin: Springer.

Johnson, C. S., M. L. Badger, D. A. Waltermire, J. Snyder, and C. Skorupka. 2016. *Guide to cyber threat information sharing*. Gaithersburg, MD: NIST.

Kanakoudis, V., S. Tsitsifli, P. Samaras, A. Zouboulis, and G. Demetriou. 2011. "Developing appropriate performance indicators for urban water distribution systems evaluation at Mediterranean countries." *Water Util. J.* 1 (1): 31–40.

Kjeldsen, T. R., and D. Rosbjerg. 2005. "Choice of reliability, resilience and vulnerability estimators for risk assessments of water resources systems/Choix d'estimateurs de fiabilité, de résilience et de vulnérabilité pour les analyses de risque de systèmes de ressources en eau." *Hydrol. Sci. J.* 49 (5): 755–767. https://doi.org/10.1623/hysj.49.5.755.55136.

Klise, K. A., M. Bynum, D. Moriarty, and R. Murray. 2017. "A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study." In *Environmental modelling and software*, 420–431. Amsterdam, Netherlands: Elsevier.

Konstantinou, C., M. Maniatakos, F. Saqib, S. Hu, J. Plusquellic, and Y. Jin. 2015. "Cyber-physical systems: A security perspective." In *Proc., 2015 20th IEEE European Test Symp. (ETS)* 1–8. New York: IEEE.

Kossieris, P., S. Kozanis, A. Hashmi, E. Katsiri, L. S. Vamvakeridou-Lyroudia, R. Farmani, C. Makropoulos, and D. Savic. 2014. "A web-based platform for water efficient households." *Procedia Eng.* 89 (Jan): 1128–1135. https://doi.org/10.1016/j.proeng.2014.11.234.

Lee, E. A. 2008. "Cyber physical systems: Design challenges." In *Proc., 2008 11th IEEE Int. Symp. on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, 363–369. New York: IEEE.

Licák, M. 2006. "Stress testing as a risk management method." *BIATEC* 14 (3): 3–5.

Loukas, G. 2015. "Cyber-physical attacks on industrial control systems." In *Cyber-physical attacks*, 105–144. Amsterdam, Netherlands: Elsevier.

Lund, N. S. V., A. K. V. Falk, M. Borup, H. Madsen, and P. Steen Mikkelsen. 2018. "Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management." *Critical Rev. Environ. Sci. Technol.* 48 (3): 279–339. https://doi.org/10.1080/10643389.2018.1455484.

Makropoulos, C., D. Nikolopoulos, L. Palmen, S. Kools, A. Segrave, D. Vries, S. Koop, H. J. van Alphen, E. Vonk, P. van Thienen, E. Rozos, and G. Medema. 2018. "A resilience assessment method for urban water systems." *Urban Water J.* 15 (4): 316–328. https://doi.org/10.1080/1573062X.2018.1457166.

Makropoulos, C., and D. A. Savíc. 2019. "Urban hydroinformatics: Past, present and future." *Water (Switzerland)* 11 (10): 1959. https://doi.org/10.3390/w11101959.

Malano, H., M. G. Bos, W. Vlotman, and D. Molden. 2004. *Benchmarking of irrigation and drainage projects*. Bridgewater, NJ: John Wiley & Sons.

McPherson, T. N., and S. J. Burian. 2005. "The water infrastructure simulation environment (WISE) project." In *Proc., World Water Congress 2005: Impacts of Global Climate Change—Proc., 2005 World Water and Environmental Resources Congress*, 1–8. Reston, VA: ASCE. https://doi.org/10.1061/40792(173)58.

Mehran, A., O. Mazdiyasni, and A. AghaKouchak. 2015. "A hybrid framework for assessing socioeconomic drought: Linking climate variability, local resilience, and demand." *J. Geophys. Res.* 120 (15): 7520–7533. https://doi.org/10.1002/2015JD023147.

Neely, A., M. Gregory, and K. Platts. 2005. "Performance measurement system design: A literature review and research agenda." *Int. J. Oper. Prod. Manage.* 25 (12): 1264–1277. https://doi.org/10.1108/01443570510633648.

NIAC (National Infrastructure Advisory Council). 2009. *Critical infrastructure resilience final report and recommendations*, 1–43. Washington, DC: NIAC.

Nicholson, A., S. Webber, S. Dyer, T. Patel, and H. Janicke. 2012. "SCADA security in the light of cyber-warfare." *Comput. Security* 31 (4): 418–436. https://doi.org/10.1016/j.cose.2012.02.009.

Niesen, T., C. Houy, P. Fettke, and P. Loos. 2016. "Towards an integrative big data analysis framework for data-driven risk management in industry 4.0." In *Proc., Annual Hawaii Int. Conf. on System Sciences*, 5065–5074. New York: IEEE.

Nikolopoulos, D., C. Makropoulos, D. Kalogeras, K. Monokrousou, and I. Tsoukalas. 2018. "Developing a stress-testing platform for cyber-physical water infrastructure." In *Proc., 2018 Int. Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*, 9–11. New York: IEEE.

Nikolopoulos, D., G. Moraitis, D. Bouziotas, A. Lykou, G. Karavokiros, and C. Makropoulos. 2019a. "RISKNOUGHT: A cyber-physical stress-testing platform for water distribution networks." In *Proc., 11th World Congress on Water Resources and Environment (EWRA 2019)*, 25–29. Madrid, Spain: EWRA.

Nikolopoulos, D., G. Moraitis, D. Bouziotas, A. Lykou, G. Karavokiros, and C. Makropoulos. 2020. "Cyber-physical stress-testing platform for water distribution networks." *J. Environ. Eng.* 146 (7): 04020061. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001722.

Nikolopoulos, D., H. J. van Alphen, D. Vries, L. Palmen, S. Koop, P. van Thienen, G. Medema, and C. Makropoulos. 2019b. "Tackling the 'new normal': A resilience assessment method applied to real-world urban water systems." *Water (Switzerland)* 11 (2): 330. https://doi.org/10.3390/w11020330.

Ostfeld, A., et al. 2012. "Battle of the water calibration networks." *J. Water Resour. Plann. Manage.* 138 (5): 523–532. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000191.

Page, P. R., A. M. Abu-Mahfouz, and M. L. Mothetha. 2017. "Pressure management of water distribution systems via the remote real-time control of variable speed pumps." *J. Water Resour. Plann. Manage.* 143 (8): 04017045. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000807.

Parmenter, D. 2015. *Key performance indicators: Developing, implementing, and using winning KPIs*. Hoboken, NJ: Wiley.

Queiroz, C., A. Mahmood, and Z. Tari. 2011. "SCADASimA framework for building SCADA simulations." *IEEE Trans. Smart Grid* 2 (4): 589–597. https://doi.org/10.1109/TSG.2011.2162432.

Rasekh, A., A. Hassanzadeh, S. Mulchandani, S. Modi, and M. K. Banks. 2016. "Smart water networks and cyber security." *J. Water Resour. Plann. Manage.* 142 (7): 01816004. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000646.

Rinaldi, S. M., J. P. Peerenboom, and T. K. Kelly. 2001. "Identifying, understanding, and analyzing critical infrastructure interdependencies." *IEEE Control Syst. Mag.* 21 (6): 11–25. https://doi.org/10.1109/37.969131.

Rossman, L. A. 2000. *EPANET programmer's toolkit*, 1–74. Cincinnati: EPANET.

Rouphael, T. J. 2009. "*Chapter 7—Uniform sampling of signals and automatic gain control*, 199–234. Burlington, VT: Newnes.

Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E. Nystrom, and D. Cohen. 2003. "The neural basis of economic decision-making in the ultimatum game." *Science* 300 (5626): 1755–1758. https://doi.org/10.1126/science.1082976.

Schnaubelt, C. M., E. V. Larson, and M. E. Boye. 2014. *Vulnerability assessment method pocket guide: A tool for center of gravity analysis*. Arlington, VA: RAND Arroyo Center.

Shin, S., S. Lee, D. Judi, M. Parvania, E. Goharian, T. McPherson, and S. Burian. 2018. "A systematic review of quantitative resilience measures for water infrastructure systems." *Water* 10 (2): 164. https://doi.org/10.3390/w10020164.

Siaterlis, C., B. Genge, and M. Hohenadel. 2013. "EPIC: A testbed for scientifically rigorous cyber-physical security experimentation." *IEEE Trans. Emerging Top. Comput.* 1 (2): 319–330. https://doi.org/10.1109/TETC.2013.2287188.

Taormina, R., S. Galelli, H. C. Douglas, N. O. Tippenhauer, E. Salomons, and A. Ostfeld. 2018. "Modeling cyber-physical attacks on water networks with epanetCPA overview of epanetCPA toolbox." In *Proc., WDSA/CCWI Joint Conf.* Kingston, ON: Water Distribution Systems Analysis. https://ojs.library.queensu.ca/index.php/wdsa-ccw/about.

Taormina, R., S. Galelli, N. O. Tippenhauer, E. Salomons, and A. Ostfeld. 2017. "Characterizing cyber-physical attacks on water distribution systems." *J. Water Resour. Plann. Manage.* 143 (5): 04017009. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000749.

Todini, E. 2000. "Looped water distribution networks design using a resilience index based heuristic approach." *Urban Water* 2 (2): 115–122. https://doi.org/10.1016/S1462-0758(00)00049-2.

Todini, E. 2003. "A more realistic approach to the 'extended period simulation' of water distribution networks." In *Advances in water supply management*. Oxfordshire, UK: Taylor & Francis.

Ugarelli, R., J. Koti, E. Bonet, C. Makropoulos, J. Caubet, S. Camarinopoulos, M. Bimpas, M. Ahmadi, L. Zimmermann, and M. G. Jaatun. 2018. *STOP-IT–Strategic, tactical, operational protection of water infrastructure against cyber-physical threats*, 2112–2119. Manchester, UK: EasyChair.

US-CERT (US Computer Emergency Readiness Team). 2018. "Alert (TA18-074A) Russian government cyber activity targeting energy and other critical infrastructure sectors." Accessed August 4, 2019. https://www.us-cert.gov/ncas/alerts/TA18-074A.

US Department Homeland Security. 2009. *National Infrastructure Protection Plan: Partnering to Enhance Protection and Resiliency*. Washington, DC: Dept. of Homeland Security.

Verizon. 2016. *Data breach digest: Scenarios from the field*. New York: Verizon.

Verizon. 2019. *2019 Data breach investigations report*. New York: Verizon.

Vilanova, M. R. N., P. Magalhães Filho, and J. A. P. Balestieri. 2015. "Performance measurement and indicators for water supply management:

Review and international cases." In *Renewable and sustainable energy reviews.*, 1–12. Amsterdam, Netherlands: Elsevier.
Wagner, J. M., U. Shamir, and D. H. Marks. 1988. "Water distribution reliability: Simulation methods." *J. Water Resour. Plann. Manage.* 114 (3): 276–294. https://doi.org/10.1061/(ASCE)0733-9496(1988)114:3(276).
Walter, M., A. Cullmann, C. von Hirschhausen, R. Wand, and M. Zschille. 2009. "Quo vadis efficiency analysis of water distribution? A comparative literature review." *Util. Policy* 17 (3): 225–232. https://doi.org/10.1016/j.jup.2009.05.002.
Wand, Y., and R. Y. Wang. 1996. "Anchoring data quality dimensions in ontological foundations." *Commun. ACM* 39 (11): 86–95. https://doi.org/10.1145/240455.240479.
Wang, Z., H. Song, D. W. Watkins, K. G. Ong, P. Xue, Q. Yang, and X. Shi. 2015. "Cyber-physical systems for water sustainability: Challenges and opportunities." *IEEE Commun. Mag.* 53 (5): 216–222. https://doi.org/10.1109/MCOM.2015.7105668.
Zhu, B., A. Joseph, and S. Sastry. 2011. "A taxonomy of cyber attacks on SCADA systems." In *Proc., 2011 Int. Conf. on Internet of Things and 4th Int. Conf. on Cyber, Physical and Social Computing*, 380–388. New York: IEEE.
Zhuang, B., K. Lansey, and D. Kang. 2012. "Resilience/availability analysis of municipal water distribution system incorporating adaptive pump operation." *J. Hydraul. Eng.* 139 (5): 527–537. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000676.
Zoppi, T., A. Ceccarelli, P. Lollini, A. Bondavalli, F. Lo Piccolo, G. Giunta, and V. Morreale. 2016. "Presenting the proper data to the crisis management operator: A relevance labelling strategy." In *Proc., IEEE Int. Symp. on High Assurance Systems Engineering*. New York: IEEE.