STOP-IT

D4.1 Asset Vulnerability Assessment to Risk Events

Supporting document for the Asset Vulnerability Assessment Tool (AVAT)

Technion

November, 2018

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 740610.

Supporting document for AVAT

D4.1 Asset Vulnerability Assessment to Risk Events

SUMMARY

This report describes the background and implementation of AVAT (Asset Vulnerability Assessment Tool) within the STOP-IT project. It elaborates the processes undertaken for the AVAT construction including a literature review and methodology development, and its potential future utilization and expansion. The AVAT is an online tool acting as a procedural "step-by-step" guide for the assessment of vulnerability of water distribution system assets taking into consideration the specific characteristics of the assets (i.e., geophysical, structural, dependence on other infrastructures), and the importance of the components for water supply (criticality of assets) and their "attractiveness" to be attacked. Within AVAT, vulnerability metrics are calculated for water distribution system assets (nodes and links). AVAT was developed in MATLAB® and compiled as a standalone application as well as a web application. As such, it mainly relies on MATLAB's Runtime shared libraries.

DELIVERABLE NUMBER	WORK PACKAGE
D4.1	WP4

LEAD BENEFICIARY	DELIVERABLE AUTHOR(S)
TECHNION	Avi Ostfeld (TECH) Elad Salomons (TECH) Rebecca Roth (TECH) Gil Zeevi (TECH) Hanoch Weiss (TECH) Jørn Vatn (NTNU/SINTEF) Eivind Okstad (SINTEF)

QUALITY ASSURANCE

Christos Makropoulos (KWR)
Dionysis Nikolopoulos (ICCS)

PLANNED DELIVERY DATE	ACTUAL DELIVERY DATE
30/11/2018	31/12/2018

DISSEMINATION LEVEL	
	☑ PU = Public ☐ PP = Restricted to other programme participants ☐ RE = Restricted to a group specified by the consortium. Please specify: _____ ☐ CO = Confidential, only for members of the consortium

Table of contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AVAT = asset vulnerability assessment tool,
CI = connectivity index,
d_j = demand at node j,
E (P) = the set of system elements,
G [N, E (P)] = the graph of the system,
h_{aj} = required minimum hydraulic head at node j,
h_i = hydraulic head at reservoir i
h_j = hydraulic head at node j,
itermax = maximum number of iterations,
IN = set of links entering node I,
LCI = link critical index,
nn = number of nodes in the system,
n_0 = number of reservoirs in the systems,
n_n = number of nodes in the network,
n_p = number of pumps in the network,
N = the set of system nodes,
Ni = set of direct upstream nodes j connected to node I,
NCF = number of connectivity failures,
P = vector of probabilities for all links,
P_k = power of pump k,
q_i = outflow from reservoir I, and demand at node i,
q_{ij} = flow in link from node i to node j,
Q_i = total flow into node i,
RI = reachability index
S = Entropy,
T = total network inflow from reservoir/tanks,
T_i = the total flow reaching node i;
TI = Todini index,
V = set of nodes,
WDS = water distribution system.
WP = work package, and
γ_w = water specific weight.

Executive summary

Vulnerability analysis of water distribution systems is a complex task. A review of the literature reveals that there is currently no universally acceptable definition or metric for the vulnerability of water distribution systems. Different definitions are proposed in the literature and some of the most relevant ones are summarized here, looking at both the quantification of vulnerability metrics and criteria as well as the degree to which these are meaningful and appropriate for water distribution systems, while still computationally feasible.

AVAT calculates two vulnerability indices at the system level, one at the node level and one at the link/element level.

The system vulnerability indices are: the Todini Index (TI) – which is a system relative aggregated measure defining how close a water distribution network operates compared to its minimum required level, and the Connectivity Index (CI) - which is the probability that all nodes in the system are connected to at least one source. The node vulnerability index is the Reachability Index (RI) – which is the probability that a given node in the system is connected to at least one source, and the Link Critical Index (LCI) which identifies the number of disconnected nodes resulted from an element outage.

The required data for AVAT consists of two parts: (1) a steady state hydraulic simulation EPANET file which runs without any errors, and (2) a MS-Excel file with probability failure data. The output of AVAT consists of tabular data exported to MS-Excel and color-bar figures.

AVAT was developed in MATLAB® and compiled as a standalone application and as a web application. As such, it mainly relies on MATLAB's Runtime libraries.

AVAT is designed as a standalone tool to be further integrated within the next tasks of WP4 and within WP6 and WP7 of the STOP-IT project, helping water utilities to make an initial screening evaluation of the vulnerability of their systems, to focus more detailed assessments at the most vulnerable parts of the system. As such the tool was explicitly designed to only require limited data on the water distribution system, including the layout of the system and one loading (demand) condition.

There are three levels of risk assessment supported by the STOP-IT tools. The Asset Vulnerability Assessment Tool (AVAT) belongs mainly to Level 2 which incorporates a single scenario assessment with known site-related data, giving a closer estimate on how the utility water distribution system performs under specific events/threats. It can however be used as the first step of both Level 1 (expert evaluation based) and Level 3 (multiple scenarios) assessments to support the relevant processes.

1. AVAT within STOP-IT

Task 4.1 is part of the STOP-IT project which works towards the development, demonstration, evaluation and preparation of scalable, adaptable and flexible solutions to support strategic/tactical planning, real-time/operational decision making and post-action assessment for the key parts of the water infrastructure.

Within this context, Task 4.1 is one of the modular components of the STOP-IT risk management platform of WP4 entitled “The Risk Assessment and Treatment Framework”.

This platform includes the Risk Identification Database (RIDB) (T3.2), a step-by-step guide for vulnerability assessment implemented through the **Asset Vulnerability Assessment Tool (AVAT) (reported here as D4.1)**, the Risk Analysis and Evaluation Toolkit (RAET) which houses state of the art models and tools for the analysis and evaluation of risk (from physical, cyber and combined events perspective) to the water systems (T4.2) integrated with a Scenario Planner (SP) and a Probabilistic Safety Assessment tool i.e. Fault Trees Explorer (PSA Explorer) and, a Risk Reduction Measure Database (RRMD)(T4.3) recommending actions to avoid or mitigate the occurrence and consequences of risk events for water critical infrastructures. Different tools are linked up into a Stress-Testing Platform (STP) able to conduct simulation but also to evaluate the effectiveness of risk reduction measures (T4.4) against Key Performance Indicators (KPIs) (T4.2). Finally, a decision support framework guides the user through the different components and tools (T4.5).

Risk assessment in STOP-IT is partitioned into three levels:

(1) Level 1 - Generic Analysis
This is the lowest level of risk analysis which requires no specific data and modelling skillsets through which users may have a first assessment of vulnerability and risks of their infrastructure and identify potential risk reduction measures based only on what is known about the type of infrastructure of their interest and high-level knowledge about the site from experts.

(2) Level 2 - Single Scenario Assessment
This level incorporates a single scenario assessment with known site-related data, giving a closer estimate on how the utility water distribution system performs on given events/threats. This second level utilizes data on the water distribution system including the layout of the system. **Vulnerability assessment for this level is implemented herein and reported through D4.1 in this accompanying report and the AVAT software.**

(3) Level 3 – Multiple Scenario Assessment
This is the most detailed vulnerability analysis level in which multiple scenarios are imposed for an all-hazard approach. Herein, a large number of various threats are considered with different characteristics and magnitude of consequences.

Fig. 1 is a schematic visualization description of AVAT within WP4 and its interconnections within STOP-IT. The figure shows the interrelationships among the different work packages starting from WP2 on the Communities of Practice for water systems protection and WP3

on the identification of risks in water distribution systems which serve as inputs to AVAT and to the preceding WP4 tasks, and WP6-WP8 which serve as the outputs of the entire STOP-IT project.

The remainder of this report is organized as follows:

- **Chapter 2** details the background for this delivery: background on water distribution systems vulnerability (2.1); vulnerability assessment methods (2.2) partitioned into three parts: Indirect/surrogate vulnerability assessment methods (2.2.1); topological vulnerability assessment methods (2.2.2), and stochastic simulations vulnerability assessment methods (2.2.3).
- **Chapter 3** describes the measures and methodologies implemented within AVAT: the **system vulnerability measures** (3.1) which include the Todini index (TI) (3.1.1), and the Connectivity index (CI) (3.1.2); and the **node and link vulnerability measures:** the Reachability Index (RI) (3.2.1), and the Link Critical Index (LCI) (3.2.2).
- **Chapter 4** incorporates the AVAT technical software description (4.1), and a case study demonstration (4.2).

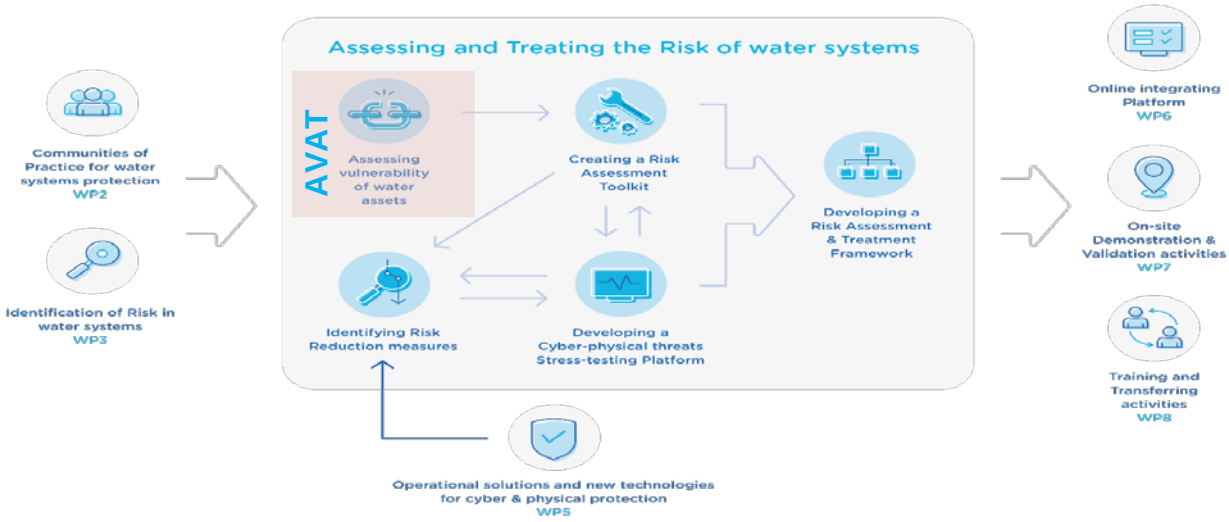

Figure 1: AVAT within WP4 and STOP-IT

2. Background

A water distribution system (WDS) is an interconnected collection of sources, pipes, and hydraulic control elements (e.g., pumps, valves, regulators, tanks) aimed at delivering water to consumers at prescribed quantities, desired pressures, and water qualities.

Water distribution systems are often described as a graph $G(N, E)$, with the set of nodes N representing connections between pipes, consumers, and sources, and the set of links E representing the pipes and hydraulic control elements such as pumps or valves. The behavior of a WDS is governed by: (1) the physical laws which describe the flow relationships in the pipes and its hydraulic control elements, (2) the consumer demands, and (3) the system layout (topology).

Most water supply networks are looped. The advantage of a looped layout resides in the possibility of obtaining a modified flow regime in case of a component failure, without disrupting the consumers supply. However, there is a significant difference between the ability of various designs to overcome a failure. Systems with the same layout and demand requirements, but with different designs might create systems with diverse reliability levels.

Vulnerability in general, and that of a water distribution system in particular, is a measure of performance. The opposite of vulnerability (henceforth termed 'invulnerability'[1]) can be defined as the ability of a system to function properly for a specified time interval under prescribed environmental conditions. This performance is relatively ease to quantify, with the help of metrics for reliability, robustness and resilience. Defining vulnerability however is less straightforward as it requires both the quantification and calculation of vulnerability measures. In the literature various definitions of vulnerability exist, and different aspects of vulnerability will be discussed later on.

Vulnerability considerations for water distribution systems are an integral part of all decisions regarding the planning, design, and operation phases. A major problem in vulnerability analysis of water distribution systems is defining vulnerability measures which are meaningful and appropriate, while still being computationally feasible.

Traditionally, improving the invulnerability of a water distribution system is achieved by following heuristic guidelines, such as ensuring two alternate paths to each demand node from at least one source, or selecting pipe diameters greater than a minimum prescribed value. By adopting these guidelines, it is implicitly assumed the system will be less vulnerable, but the resulting (reduced) vulnerability level is not quantified or measured.

2.1 Water distribution systems vulnerability

Rausand (2011) defines vulnerability as a possible weakness of an asset or group of assets that can be exploited by one or more threat agents, for example, to gain access to the asset and subsequent destruction, modification, theft, and so on, of the asset or part of the asset. On the other hand a system or a unit is said to be *invulnerable* if it functions properly for a

[1] It is clearly understood that no system is perfectly "invulnerable". In every system, undesirable events/failures can cause a decline or an interruption in system performance. Failures are of a stochastic nature and are the result of unpredictable events that occur in the system itself and/or in its environs.

specified time interval under prescribed environmental conditions. For both vulnerability and invulnerability we have to specify three different 'types' of vulnerability:

1. **System Vulnerability**: is the overall system vulnerable due to risk of component failures, external threats etc.?
2. **Component vulnerability contribution**: which components contribute to the system's vulnerability?
3. **Inherent Vulnerability**: to what extent a specific component is exposed to threats?

In this report a methodology for assessing all three "types" of vulnerability is presented. The AVAT tool is able to calculate the main aspects of a vulnerability analysis, although not all aspects discussed in this report were implemented in the tool.

Quantitatively, the **"invulnerability" of a water distribution system can be defined as the counterpart of the probability that the system will fail, where a failure is defined as the system's inability to supply consumer demands for water quantity, water pressure, and water quality**.

Vulnerability analysis of water distribution systems involves three interconnected stages: (1) identification of measures and criteria to assess system vulnerability, (2) quantification of the probabilistic nature of the behavior of the system components and its consumer demands, and (3) definition of the environmental conditions under which the system is designed to operate.

Two distinct types of events can cause a water distribution system failure: (1) system components going out of service (e.g., pipes and/or hydraulic control elements), and/or (2) consumers demands, such as flow rates in case of a fire or drought, exceeding design values.

Three interconnected issues are involved in assessing the vulnerability of a water distribution system:

(1) Measures. Vulnerability measures should be quantified from the consumers point of view, defining a required level of service (e.g., maximum duration and frequency of supply interruptions at a given probability, expected unserved demand, damage incurred when failure occurs).

(2) Failures. Failure is an event in which a vulnerability measure is impaired. Failure can occur if a system component fails (e.g., a pipe, valve, pump, tank), in case of consumer demand exceeding design values, or as a consequence of both. When analysing the vulnerability of a WDS, these two types of events and their interdependencies should be explored.

(3) Assembly. Construction of a mathematical model for assessing the system vulnerability, subject to the measures defined in (1), and the failures in (2).

2.2 Vulnerability assessment methods

Vulnerability assessment methods for water distribution systems can be classified into three categories: (1) indirect/surrogate, (2) topological, and (3) probabilistic. Among probabilistic

approaches they may be classified into (3a) stochastic Monte Carlo simulation and (3b) analytical.

The ability to implement any vulnerability assessment method is foremost subject to the system data availability: from knowledge of only the system layout/connectivity to complete data availability on the system's operation, its component mechanical failure distributions, consumer consumptions, and its operational strategies in failure modes.

2.2.1 Indirect/surrogate vulnerability assessment methods

Here, the vulnerability of the system is quantified through heuristic surrogate measures aimed at indirectly assessing the system redundancy and through that, its vulnerability.

Two major approaches were suggested in the distribution systems research literature for this category: the Todini index (Todini, 2000), and Entropy (Awumah et al., 1990).

Todini (2000) proposed the resilience index (Eq. 1) as a measure for redundancy of water distribution systems, with the following rationale: assume water is supplied to each consumer while exactly meeting design demands at required pressures. In such circumstances, whenever the demand at one node will increase or a device (such as a pipe or pump) will fail, the direction of flow in the system will change, the original network will shift into a new network containing higher internal energy loses. In such a case, it will be impossible to deliver the desired flow rates at the required minimum pressures.

Given this observation each node must have a higher energy level than required in order to have sufficient surplus to be dissipated in case of failure. The Todini index quantifies these surpluses to defining a heuristic intrinsic capacity measure for coping with failures. It should be noted that the Todini index does not involve statistical considerations on failures. However, its increase will raise the system cost and reduce the network vulnerability. This was shown by several studies (Todini, 2000, Ostfeld et al., 2014).

The Todini index (TI) is defined in Eq. (1):

$$TI = \frac{\sum_{j=1}^{n_n} d_j(h_j - h_{aj})}{\sum_{i=1}^{n_0} q_i h_i + \left(\frac{1}{\gamma_w}\right)\sum_{k=1}^{n_p} P_k - \sum_{j=1}^{n_n} d_j h_{aj}} \quad (1)$$

where: n_n = number of nodes in the network, d_j = demand at node j, h_j = hydraulic head at node j, h_{aj} = required minimum hydraulic head at node j, n_0 = number of reservoirs in the systems, q_i = outflow from reservoir i, h_i = hydraulic head at reservoir I, n_p = number of pumps in the network, P_k = power of pump k, and γ_w = water specific weight.

Several studies (Prasad and Park, 2004; Farmani et al., 2005; Reca et al., 2008; Jayaram and Srinivasan, 2008; Raad et al., 2010; Baños et al., 2011; Tanyimboh et al., 2011; Greco et al., 2012; Pandit and Crittenden, 2012) suggested modifications to the resilience index of Todini and compared its performance (Atkinson, 2014) against other heuristic vulnerability surrogates for water distribution systems such as Entropy (Awumah et al., 1990; Tanyimboh et al., 2011):

Entropy is a thermodynamic property representing the number of possible configurations of a system. Awumah et al. (1990) demonstrated how maximizing entropy reduces the vulnerability of water distribution systems.

Considering this approach and that the value of maximum achievable entropy has no standard range and depends on the number of nodes and attached pipes within a network, Tanyimboh et al. (2011) formulated the Entropy (S) of the system (Eq. 2) to be maximized for reducing the system vulnerability:

$$S = -\sum_{i \in IN}^{nn} \left(\frac{Q_i}{T}\right) \ln\left(\frac{Q_i}{T}\right) - \frac{1}{T}\sum_{i \in IN}^{nn} T_i \left[\left(\frac{q_i}{T_i}\right) \ln\left(\frac{q_i}{T_i}\right) + \sum_{j \in Ni}^{nn} \left(\frac{q_{ij}}{T_i}\right) \ln\left(\frac{q_{ij}}{T_i}\right)\right] \quad (2)$$

where: nn = number of nodes in the system; IN = set of links entering node i; Q_i = total flow into node i; T = total network inflow from reservoir/tanks; T_i = the total flow reaching node i; Ni = set of direct upstream nodes j connected to node i; q_i = demand at node i; and q_{ij} = flow in link from node i to node j.

Gheisi and Naser (2015) (Table 1) summarized the major surrogate vulnerability measures for water distribution systems.

Table 1: Surrogate vulnerability measures (Gheisi and Naser, 2015)

Surrogate measure	**Description**
Entropy statistical flow	Degree of flow uniformity and redundancy in a WDS
Resilience index	Surplus power available at demand nodes as a percentage of net input power
Modified resilience index	Surplus power available at demand nodes as a percentage of required power
Network resilience index	Surplus power available at demand nodes as a percentage of net input power considering reliable loops and redundancy
Mixed reliability surrogate	Mixture of a statistical flow entropy approach and resilience index

Note that all the measures in Table 1 take a system perspective without considering which components contribute to the vulnerability measure. In Section 2.3.2 an approach for treating component vulnerability contribution is presented.

2.2.2 Topological vulnerability assessment methods

Topological vulnerability refers to the probability that a given network is physically connected, given its components' mechanical vulnerabilities (i.e., the components' probabilities to stay operational over a given time interval and given environmental conditions).

Wagner et al. (1988a) used Reachability and Connectivity to assess the vulnerability of a water distribution system, where Reachability is defined as the probability that a given demand node is connected to at least one source, and Connectivity is defined as the

probability that all demand nodes are connected to at least one source. Shamsi (1990), and Quimpo and Shamsi (1991) used node pair reliability (NPR), where NPR is defined as the probability that a given source node is connected to a given demand node. Ormsbee and Kessler (1990) used graph theory for designing invulnerable water distribution systems.

Measures used within this category consider only the connectivity between nodes (as in transportation or telecommunication network vulnerability models), and therefore do not consider the level of service provided to consumers during a failure. It should be emphasized that the existence of a path between a source and a consumer node is only a necessary condition for providing consumers required demands.

Yazdani and Jeffrey (2012) presented some topological vulnerability measures for water distribution systems which were further implemented in Jung and Kim (2018) for trading off cost versus topological vulnerability. Torres et al. (2017) nicely summarized the existing topological system-based measures.

Table 2: Topological vulnerability measures (Torres et al., 2017)

Topological measure	**Description**
Algebraic connectivity	The second smallest eigenvalue of the normalized graph Laplacian which ranges between 0 and 2. Greater values indicate higher invulnerability.
Average degree	Average number of edges or pipes connecting a node at a given network.
Average shortest path length	Average number of links traversed between two nodes. Shorter average path length is an indicator of network efficiency.
Betweenness centrality	The number of shortest paths crossing a node. Higher values indicate high importance as a bottleneck node.
Edges	The total number of edges in a network.
Network density	The maximum number of possible edges versus how many edges are actually present, Higher network densities imply pipe networks of higher connectivity.
Network diameter	The length of the longest geodesic path between any pair of vertices. Higher values may indicate higher system-level head loss.
Network efficiency	The harmonic-mean physical distance between nodes. Ranges between 0% for least-efficient and 100% for most-efficient networks and may be used as proxy for average water travel time.
Network radius	The length of the smallest geodesic path between any pair of vertices. Lower values may indicate lower system-level head loss
Meshedness	Density of general loops in planar graph. Ranges between zero for tree-like and 0.5 for grid-like networks. May be used as a local redundancy measure for pipe networks
Single degree nodes	The total number of nodes in a network with a node degree equal to one. This is equivalent to the number of dead-end connections within a pipe network.

Note also here that all the measures in Table 2 take a system perspective without considering which components contributing to the vulnerability measure. In Section 2.3.2 an approach for treating component vulnerability contribution is presented.

2.2.3 Stochastic simulations vulnerability assessment methods

Stochastic simulations vulnerability assessment methods (Wagner et al., 1988b; Ostfeld et al., 2002; Yang et al., 1996; Gheisi et al., 2016) refer to quantifying the hydraulic

vulnerability of a water distribution system which is the probability of the system to provide its consumers a required level of service in terms of water quantities, pressures, and water qualities, over a given time period under specified environmental conditions. As such, assessing the hydraulic vulnerability of a WDS refers directly to its major objective: conveying to its consumers required water quantities at minimum pressures and adequate water qualities, at desired probabilities. The defined probabilities outline the system level of service, similarly to the Return Period/Recurrence Interval in surface hydrology.

To compute the hydraulic vulnerability of a water distribution system, stochastic simulations need to be performed. These involve generation of random events out of the mechanical component vulnerabilities through random number generators, evaluation of the resulted events' impact on the system performance, and as a result computation of statistic vulnerability measures, such as the frequency of pressure reduction at consumer nodes.

Essentially, any system vulnerability measure can be computed through stochastic (Monte Carlo) simulations, as long as the necessary data is available. While stochastic simulation is the most accurate approach to assess the true vulnerability of a system, it is the most difficult to extrapolate (i.e., interpret its physical outcomes), and practically unfeasible. The reason for it being practically unfeasible is because performing stochastic simulations requires the availability of probability density functions for all of its components, and a calibrated extended period hydraulic model of the distribution system. Both the probabilities and the hydraulic model data are rarely available. In addition, the problem of how to operate the distribution system at a reduced/failure mode is another substantial issue that is often neglected in assessing the system vulnerability. Uncertainty inclusion poses additional challenges to this problem (Shafiqul et al., 2014; Goharian et al., 2018).

Fig. 1 in Ostfeld et al. (2002) provides a general framework for stochastic simulations of water distribution systems for vulnerability assessment.

2.3 Interpretation of vulnerability and related terms

This section discusses various interpretations of vulnerability as basis for proposing various component-based vulnerability indices in Section 3.4. A methodology for calculation of the indices is proposed in Section 3.5. Appendix A gives the required formulas for standard and advanced reliability calculation in network systems, applied in connection with the indices.

The definition of vulnerability by Rausand (2011) emphasizes a *weakness* that can be *exploited* by *threat agents*. This definition emphases three aspects:

1. A weakness, i.e., some property
2. A threat agent that can exploit this weakness. The agent is not necessarily malicious acts. It could for example also refer to "lack of competence" and "bad weather"
3. The term may be used at a system level, or at a component level

2.3.1 System level – system vulnerability

At the system level we may say a system, for example the water distribution system, is vulnerable due to its inability to withstand a hostile environment. This "hostile environment" could then be an aging infrastructure with limited redundancy. The original use of Todini's index was proposed in such a context, i.e. the vulnerability index is measuring lack of reserve capacity in the network by an aggregation over consumers and pipes, valves,

pumps etc. The AVAT tool presented in this report has as its main focus a system level understanding of vulnerability.

2.3.2 Component vulnerability

We may also say that a component or subsystem is vulnerable, for example a pumping station. From the vulnerability definition this would be to say that the pumping station is vulnerable due to its weakness to withstand attack from a threat agent. For example, we do not have many physical barriers (fences, locked doors etc.) Therefore, the pumping station is vulnerable. In such a context we do not consider criticality of the pumping station, it is vulnerable because it is easy to “attack”. In order to assign a vulnerability index for components following this definition, we therefore need to investigate explicitly the asset with respect to “inherent barriers” implemented to withstand a hostile environment.

2.3.3 Component importance

In risk and reliability analyses the term component reliability importance is introduced. A reliability importance measure attributes system performance on a component level. When defining an importance measure, we can ask two different questions:

1. To what extend will the system be affected if component i fails or has a reduced capacity?
2. Will component i fail or experience reduced capacity, and what is then the system impact?

If systems are simplified such that we can treat both system and component performance as binary (as is the case for fault tree analysis), two different importance measures are defined to reflect these two situations:

1. Birnbaum’s measure, IB(i) = The probability that component i is critical = The probability that the other components are in such a state that it is decisive whether component i is functioning or not.
2. Criticality importance, ICR(i) = The probability that component i is critical and is failing given that the system is failing.

Birnbaum’s measure is stressing the importance of component i independently of the performance of the component itself, whereas the criticality importance measure also takes the performance of component i into account.

Note that neither of the two importance measures address vulnerability aspects on component level, i.e., the measures do not express explicitly inability of a component to withstand something. If the system perspective is taken, and we ask the question regarding the system inability to withstand a component failure we can say that the Birnbaum’s measure is a very relevant measure, it states basically the probability that the system is unable to withstand a component failure.

This means that a Birnbaum like measure could be relevant for screening of components contributing to system vulnerability. We should then have in mind that the measure is not addressing the vulnerability of the components. That would then be the topic at the next stage, i.e., further analysis of the vulnerability contributing components.

2.3.4 What to include in vulnerability indexes at component level?

A vulnerability index at component level should include “inability" for that particular component or asset to withstand a hostile environment” and at first, leave out the impact it

will cause on the system as whole if an attack, or hostile environment is able to “kick down” the component. Some others might include aspect of criticality or importance of component, whereas, others again would consider “if”, in the meaning of “will there be a hostile environment or different kinds of an attack?”.

For a component (or subsystem) the following element could therefore be considered:

1. Likelihood of attack (hostile environment)
2. Inability / ability to withstand an attack
3. Consequence / severity if such an attack succeeds.

The combination of these three aspects is what we often include in a risk measure. All three aspects are important, but whether they should be denoted “risk”, “vulnerability”, “criticality” is not obvious.

In STOP-IT a two-step procedure is proposed with respect to a vulnerability analysis:

1. From the system perspective, identify components (assets or subsystems) that contribute to system vulnerability, i.e., a (water) system’s inability to withstand deficiencies in those components. This is implemented within the AVAT tool.
2. For components contributing to system vulnerability, investigate the vulnerability of those components. This is developed as a method, but not fully implemented in the current version of the AVAT tool.

Both procedures are described in the following chapter.

3 The Asset Vulnerability Assessment Tool (AVAT): measures and methodologies implementation

In AVAT, two system vulnerability, one node vulnerability, and one link/element vulnerability indices are computed. The system vulnerability indices are the Todini Index (TI), and the Connectivity Index (CI), where the node vulnerability index is the Reachability Index (RI), and the link/element vulnerability index is the Link Critical Index (LCI).

A description of each of the above measures and the underlying methodology for their computations are described further below.

3.1 System vulnerability measures

3.1.1 The Todini Index (TI)

The Todini index is a system relative aggregated measure defining how close a water distribution network operates compared to its minimum required level (see Eq. 1 and the description above).

The Todini index is very easy to calculate and to intuitively interpret, and as such has become one of the most common deterministic measures for water distribution systems redundancy/vulnerability (cited 348 times according to SCOPUS).

To compute the Todini index a steady state water distribution systems solution should be available. Such a solution should be provided to AVAT which then calculates the Todini index according to Eq. 1.

It should be noted that the Todini index can mainly serve as a tool for comparing the redundancy/vulnerability of different design/operational scenarios of a given system (e.g., connection to an additional source, pumps addition, or altering minimum pressure requirements), less for comparing the redundancy/vulnerability of different systems.

The Todini index can also assist in ranking elements vulnerability (e.g., pumps, major pipes). This can be accomplished by running the system with and without elements, followed by ranking the Todini indices outcome. Such computations are especially important for identifying critical system assets, thus helping in prioritizing their attractiveness to be attacked.

Such calculations can be easily performed in AVAT.

3.1.2 The Connectivity Index (CI)

The Connectivity Index (CI) is the probability that all nodes in the system are connected to at least one source.

$$\mathrm{CI} = \mathrm{P}(\{\text{All nodes connected to at least one source}\}) \quad (3)$$

The implementation of equation (3) in AVAT is described in **Algorithm 1** below.

3.1.2.1 Algorithm 1. Connectivity Index

Input G [N, E (P)], E (P), Itermax
For i = 1 to itermax

Check if any of the elements in E (P) failed through failure randomization of each element (i.e., Monte Carlo Simulations for each element).

- If none of the elements failed: i = i +1
- If at least one element failed, check if all nodes are connected to at least one source. If all nodes are connected to at least one source: i = i +1, else: NCF = NCF + 1

Until itermax
CI = NCF/Itermax

---

where: N = the set of system nodes; E (P) = the set of system elements; P = vector of probabilities for all links; G [N, E (P)] = the graph of the system; itermax = maximum number of iterations; NCF = number of connectivity failures.

3.2 Node and link vulnerability measures

3.2.1 The Reachability Index (RI)

The Reachability Index (RI) is the probability that a given node in the system is connected to at least one source.

$$\mathrm{RI} = \mathrm{RI}(i) = \mathrm{P}(\{\text{Node i is connected to at least one source}\}) \quad (4)$$

The implementation of equation (4) in AVAT is described in **Algorithm 2** below.

3.2.1.1 Algorithm 2. Reachability Index

Input G [N, E (P)], E (P), Itermax
For i = 1 to itermax
Check if any of the elements in E (P) failed through failure randomization of each element (i.e., Monte Carlo Simulations for each element).

- If none of the elements failed: i = i +1
- If at least one element failed, check if all nodes are connected to at least one source. If all nodes are connected to at least one source: i = i +1, else: check which nodes are not connected to at least one source. All j nodes which are not connected to at least one source: $NRF_j = NRF_j + 1$

Until itermax
$RI_j = 1 - (NRF_j / Itermax)$, for each node j

---

where: NRF_j = number of reachability failures of node j, RI_j = the reachability index of node j

3.2.2 The Link Critical Index (LCI)

The Link Critical Index (LCI) is a link/element index identifying the number of disconnected nodes resulted from an element outage in an undirected graph representation of the distribution system. As the element importance increases (such as a solitary pipe connecting the system to a single source), so are the number of disconnected nodes occasioned from its failure, and its corresponding LCI.

$$\mathrm{LCI} = \mathrm{LCI}(i) = \mathrm{P}(\{\#\text{ of disconnected nodes upon an outage of link } i\}) \quad (5)$$

Several observations can be made here regarding the LCI:

- In a tree liked system, any element outage, disconnects all downstream nodes. A tree liked system thus holds the most vulnerable water distribution system layout.
- In a looped water distribution system, multiple paths can be available to each of the system nodes. A failure of an element in a looped water distribution system may thus cause no disconnection of nodes. However, as the LCI does not incorporate any hydraulics of the system, the existence of a path between nodes does not guarantee the supply of any level of service (i.e., flow at minimum pressure head).
- Through the LCI computation, the damage incurred to the system as a result of a link/element failure can be easily calculated in terms of the non-supplied water, by summing up all flows associated with all the disconnected nodes. Multiplying this figure by the probability of that link failure results the risk related to that element. This, in conjunction with the Todini index as described above, can serve for identifying critical system assets, thus helping in prioritizing their attractiveness to be attacked.

Currently in AVAT the LCI holds the number of disconnected nodes for each link/element outage. Further extensions as described above will be incorporated in the next AVAT updates.

3.3 Methodological clarifications

A few observations to note on AVAT:

1. All threats are considered via the failure probabilities. For example: if a pump's PLC is particularly open to a cyberattack, then this should be reflected in its probability to fail.
2. The user is responsible for entering the probability of failure associated with each link. Common data of pipe failure are related to failure per unit length (e.g., Mays 1989). The user is in-charge of performing the appropriate multiplications, thus entering the correct probability of failure figure for each link.
3. Since Monte Carlo simulations are involved, sensitivity analysis should be performed for a sufficient number of Monte Carlo simulations to receive stable results.
4. Only "links" are assigned failure probabilities. To consider "node" assets (e.g., Treatment Plants, Tanks, Sources) vulnerabilities, an additional single link should be added to connect the asset to the network. That link probability will then hold the probability of that asset to fail.
5. In a much broader sense the probability of failure can be related to the attractiveness of an asset to be attacked. For example, if an asset has a probability of a physical/mechanical failure of 1%, but its likelihood to be attacked due to its attractiveness is higher than that, that the initial probability of 1% can be increased to reflect this phenomenon (e.g., from 1% to 5%). Guarding and surveillance facilities, as well as maintenance actions, are means of reducing failure probabilities, as well as reducing this aspect of "attractiveness" to be attacked.

3.4 Component Vulnerability Contribution- and Inherent Vulnerability Indices

The component vulnerability contributing index measures how vulnerable a system, i.e., a WDN is in relation to deficiency in component i. In STOP-IT a vulnerability contributing index can be calculated based on either:

- A deterministic hydraulic model for the WDN. EPANET is one such model.
- A (simplified) skeleton model of the network. A reliability model based on such a model could be based on reliability block diagrams and/or fault trees. A reliability model is a probabilistic model.
- A combination of 1 and 2.

3.4.1 Deterministic Importance Measures

Deterministic importance measures (index) are related to the performance of the WDN in a deterministic way, i.e., without treating failure probabilities and other random events. A main principle pursued here is to investigate the performance of the WDN when the component under investigation is set to a fault state. For the various component types this means:

- For pipes we assume that the pipe is disconnected from the network
- For valves we assume that the valve is left in the position it normally has (i.e., *open* if it normally is open, and *closed* if it is normally closed)
- For pumping stations, we assume that the pumping station is in a fault state, i.e., cannot pump water
- For tanks we assume that the tank is empty.

It is also required to define performance of the WDN. Several performance measures could be defined. In STOP-IT is recommended to use the Todini's resilience index as a basis for the performance of the WDS:

$$I^{\mathrm{R}} = \frac{n_n \sum_{j=1}^{n_n} w_j d_j \left(h_j - h_j^*\right)}{\sum_{k=1}^{n_o} q_k h_k - n_n \sum_{j=1}^{n_n} w_j d_j h_j^*} \tag{6}$$

where n_n are the number of nodes in the network, h_j is the nodal hydraulic head, h_j^* is the minimum allowable hydraulic head, d_j is the nodal demand, n_o is the number of reservoirs in the network, q_k is the outflow from reservoir k, and h_k is the hydraulic head in reservoir k. Compared to the original definition of the index, Equation (6) also allows to give a normalised weight, w_j to each node ($\Sigma_j w_j = 1$).

The resilience index in Equation (6) is measuring the amount of excess pressure in the network and could be calculated by for example EPANET for the steady state situation, or for a specific point of time.

The resilience index in equation (6) does not give any reference to the various components since it is a system resilience index. To define a (deterministic) component vulnerability contributing index we introduce:

$$I^{\mathrm{DVC}}(i,t) = 1/I^{\mathrm{R}} \text{ calculated } t \text{ time units after a failure of component } i \tag{7}$$

The index in Equation (7) may be calculated for various point of times. In the steady state situation this corresponds to a situation where all water tanks are disconnected (empty).

Since water tanks are "dynamic" they might need special attention in the WDN vulnerability assessments, beyond considering only steady state conditions.

Although, the time dependent solution in Equation (7) brings insight into how much water tanks can compensate for a component failure, it will in most cases be required to have only one importance measure for a component. In those cases, it is natural to integrate $I^{\mathrm{DVC}}(i,t)$ over the downtime distribution $f_D(d)$ of component *i*:

$$I^{\mathrm{DVC}}(i) = \int f_D(d) I^{\mathrm{DVC}}(i,d) dt \tag{8}$$

If computational effort is a challenge, we could simplify:

$$I^{\mathrm{DVC}}(i) = I^{\mathrm{DVC}}(i, \mathrm{MDT}_i) \tag{9}$$

where, MDT_i is the mean downtime and represents a typical down-time. The indexes in Equations (8) and (9) indicate the expected consequences given a component failure. Hence, if the probability of a component failure is given, the measure can also be used to assess the associated risk, but this is not pursued here.

3.4.2 Probabilistic Importance Measures

Probabilistic importance measures (index) take the probabilistic nature of the WDN into account. Above it is argued that a Birnbaum like measure for component *i* is appropriate, i.e., a measure of the probability that component *i* is critical to the system.

To find a Birnbaum like measure for each component we need an unavailability measure of the WDN. In the following we use Q_0 and F_0 to denote WDN unavailability and WDN failure frequency respectively. Appendix A outlines the required calculation formulas. Depending on the situation we focus on either WDN unavailability or WDN failure frequency. If unavailability is the main focus, we use the following (probabilistic) component vulnerability contributing index:

$$I^{\mathrm{PVC}}(i) = Q_0(0_i) - Q_0(1_i) \tag{10}$$

where the notation 0_i means that component *i* is in a fault state, and 1_i means that component *i* is in a functioning state. It should be noted that in order to calculate $I^{\mathrm{B}}(i)$ in equation (10), Q_0 is defined for one critical end user, i.e., one of the nodes in the WDN. In some cases, it would be required to calculate Q_0 for several critical end users, and then take a *weighted average over the various end users*. If WDN system frequency is our main concern, we rather use:

$$I^{\mathrm{PVC}}(i) = F_0(0_i) - F_0(1_i) \tag{11}$$

3.4.3 Combining the Deterministic and Probabilistic Vulnerability Measures

The deterministic and probabilistic vulnerability contributing indexes in Equations (8) and (10) or (11) are not on the same scale. In case we are combining the deterministic and probabilistic measures into a (weighted) average a normalization is required. Let $I^{\mathrm{DVC}}_{\mathrm{Max}}$ be the maximum value of indexes calculated by Equation (8) and similarly $I^{\mathrm{PVC}}_{\mathrm{Max}}$, be the maximum value of indexes calculated by Equations (10) or (11). The recommended

normalization is to divide the indexes by $I_{\mathrm{Max}}^{\mathrm{DVC}}$ and $I_{\mathrm{Max}}^{\mathrm{PVC}}$ for the deterministic and probabilistic vulnerability contributions respectively.

3.4.4 The Inherent Vulnerability Index

The vulnerability contribution index $I^{\mathrm{VC}}(i)$ is a system index indicating components that have a major impact on the WDN as such. Taking a component perspective, it is also important to establish an inherent vulnerability index $I^{\mathrm{IV}}(i)$ indicating the likelihood of attack or other types of hostile environments that threaten a specific component, and the inability to withstand such an attack. The proposed inherent vulnerability index comprises:

$$I^{\mathrm{IV}}(i) = f_i^{V}\left(1 - p_i^{\mathrm{V}}\right) + \lambda_i = f_i^{\mathrm{V}}\, q_i^{\mathrm{V}} + \lambda_i \tag{12}$$

where $f_i^{\mathrm{V}}$is the frequency of an attack or other types of hostile environment, and $p_i^{\mathrm{V}} = 1 - q_i^{\mathrm{V}}$is the probability to withstand such an attack. Further, λ_i is the total failure rate of "non-intended" failures, both physical and cyber events. In the following, vulnerability factors affecting $f_i^{\mathrm{V}}$and $q_i^{\mathrm{V}}$are discussed. Some factors are rather general and affect all components in the WDN, whereas some factors might have stronger influence on a specific component. Therefore, a system frequency vulnerability index $f^{V}$is introduced for common factors, and we let $f_i^{\mathrm{V}} = f^{\mathrm{V}} f_i^{\mathrm{C}}$ where f_i^{C} is a component specific factor, with respect to attack frequency.

In the following we distinguish between *a factor* like "objectives" or "capabilities" and the assessed *score* for the factor. A score is a number between 0 and 1, the higher score the more severe the condition is for the actual WDN. Some factors are linked to more than one of the quantities f^{V}, f_i^{C}, and q_i^{V}, but most factors are linked to one of them.

Vulnerability factors influencing f^{V} (general, independent of components):

- Objectives
- Facility attractiveness/Symbolism
- Historical evidence, e.g., number of incidents/attempts
- Recognisability

Vulnerability factors influencing f_i^{C}(component specific to frequency)

- How vulnerable the component seems from the attackers point of view
- Required access (easy to access the particular component)
- Required skills (Attacker's skill vs required skill to make an attempt)
- Required resources (Attacker's available resources vs required resources to make an attempt)
- Proximity to the component
- Software vulnerability

Vulnerability factors influencing q_i^{V}(component specific to probability of not withstanding)

- Required skills (Attacker's skill vs required skill to succeed in an attempt)
- Required resources (Attacker's available resources vs required resources to succeed in an attempt)
- Lack of preventive measures and monitoring

3.5 Calculation of Vulnerability Indices at Component level

3.5.1 Component Vulnerability Contribution Indices

To calculate the component vulnerability contribution index we assume in the following that we are able to both find the deterministic indexes based on e.g., EPANET, and the probabilistic indexes based on fault tree analysis or reliability block diagram analysis. If only one of the measures are available, the weighted average is replaced by the index we have available.

1. Establish the EPANET model for the WDN under consideration. Simplification could be acceptable in this context.
2. Calculate the deterministic vulnerability indexes by Equation (8) or Equation (9) for each component
3. Normalize the calculated deterministic indexes by dividing all indexes by the largest index
4. Establish a simplified (skeleton) reliability model of the WDN under consideration. Identify one or more critical end users. If more than one end user is considered, give weights to each end user considered.
5. Calculate the probabilistic vulnerability indexes by Equation (10) or Equation (11) for each component. If more than one end user is treated, repeat for each end user and then calculate weighted averages.
6. Normalize the calculated probabilistic indexes by dividing all indexes by the largest index
7. Calculate the final component vulnerability contributing indexes, $I^{\mathrm{VC}}(i)$ by taking the average of the normalized deterministic and normalized probabilistic indexes for each component.
8. Sort in descending order the vulnerability indexes in order to screen components.

3.5.2 Inherent vulnerability indexes of components

Table 3 below is the basis for calculation of an inherent vulnerability index for each component. By default, all scores are set to $S_i = 1$. The form in *Table 3* is to be filled out for all components with a high vulnerability contributing index (i.e., after the initial component screening). Note that the scores for the general conditions ($S_0 \dots S_4$), are the same for all components.

The score S_0 has a different interpretation compared to the other scores. It is a "worst case" *frequency* of an attack given that all related vulnerability factors were in their worst state. S_0 can be both lower and higher than 1. A natural dimension here is number of attacks per year. A value of $S_0 = 1$ is reasonable even if no explicit analysis is conducted. Calculation formulas are provided both in the form, and for the final calculation if the inherent vulnerability index. No explicit procedure is used for assessment of the "non-intended" failure rate λ_i.

Table 3 Specification of scores for each vulnerability factor

Vulnerability factor	**#**	S_i
Baseline frequency (worst case frequency)	0	
Objectives	1	
Facility attractiveness/Symbolism	2	
Historical evidence, e.g., number of incidents/attempts	3	

Recognisability	4	
General conditions, frequency of attack:	$f^V =$	$\prod_{i=0}^{4} S_i$
How vulnerable the component seems from the attacker's point of view	5	
Required access (easy to access the particular component)	6	
Required skills (Attacker's skill vs required skill to make an attempt)	7	
Required resources (Attacker's available resources vs required resources to make an attempt)	8	
Proximity to the component	9	
Software vulnerability	10	
Component specific, frequency of attack:	$f_i^C =$	$\prod_{i=5}^{10} S_i$
Required skills (Attacker's skill vs required skill to succeed in an attempt)	11	
Required resources (Attacker's available resources vs required resources to succeed in an attempt)	12	
Lack of preventive measures and monitoring	13	
Component specific, probability of notwithstanding:	$q_i^V =$	$\prod_{i=11}^{13} S_i$

The final inherent vulnerability factor is now calculated by:

$$I^{\mathrm{IV}}(i) = f^{\mathrm{V}} f_i^{\mathrm{C}} \, q_i^{\mathrm{V}} + \lambda_i \tag{13}$$

3.5.3 Total vulnerability indices

The component vulnerability contributing index $I^{\mathrm{VC}}(i)$ and the inherent component vulnerability index $I^{\mathrm{IV}}(i)$ point to vulnerability aspects of a component. To establish a total vulnerability index, it is recommended to multiply the two indexes:

$$I^{\mathrm{V}}(i) = I^{\mathrm{IV}}(i) I^{\mathrm{VC}}(i) \tag{14}$$

$I^{\mathrm{VC}}(i)$ can be interpreted as a risk measure, where $I^{\mathrm{IV}}(i)$ represents the "probability" of an event, and then $I^{\mathrm{VC}}(i)$ represents the "consequence" of the event on the WDS. Generally, the terms risk and vulnerability should not be mixed up, but the way of arguing in terms of first a vulnerability contributing index and then an inherent vulnerability index component by component is a normal approach for risk calculation, hence the similarity.

These indices are means to link components such as pipes, pumping stations and water tanks to system vulnerability. The inherent vulnerability assessment included identification of vulnerability factors, a scoring regime, and a result compilation framework. For a more comprehensive vulnerability assessment these approaches should be part of the AVAT methodology.

AVAT (further described in Section 4) implements the methodologies described in Section 3.1 and 3.2. The methodology described in Section 3.4 was developed for completeness and could be implemented at a later stage if interest from the water utilities is manifested.

3.6 Summary of indexes

Table 4 summarises the presented indices. In addition to definitions, comments are given to whether the indices relate to a system- vs. or component perspective, and if the model is simulation-based (EPANET) or an analytical reliability-based approach (e.g. made possible by a "Skeleton"-model).

Table 4 Summary of indexes

Measure/Equation	Definition	Comment
TI = Todini index/ Eq (1)	System related measure calculating the redundancy of a water distribution system	This index is an overall measure of the system invulnerability. It is calculated by the AVAT tool, and it requires an EPANET model run for a steady state situation. This is a ***deterministic system index***.
CI = Connectivity Index Eq (3)	Probability that all nodes (users) are connected to at least one source	This index is calculated by the AVAT Tool. It requires an EPANET input file and an Excel file with link failure probabilities. This is a ***probabilistic system index***.
RI(*j*) = Reachability Index / Eq (4)	Probability that node *j* is connected to at least one source	This index is calculated by the AVAT Tool. It requires an EPANET input file and an Excel file with link failure probabilities. This is a ***probabilistic*** index for each ***node*** (end users) in the network.
LCI(i) = Link Critical Index / Eq (5)	Number of disconnected nodes resulted from an element outage	This index is calculated by the AVAT Tool. It requires an EPANET input file. This is a ***deterministic*** index for each ***link*** in the network, and will therefore not treat link probabilities.
I^{R} = Weighted Todini index / Eq (6)	Similar to the Todini Index in equation (1) but allowing giving weights to each node	This index is an alternative to the Todini Index if we would like to give different weights to each node, corresponding to for example giving higher weights to critical infrastructures such as hospitals, industry and so on. This measure is not calculated by the AVAT Tool. It will require an EPANET model run in the steady state situation. This is a ***deterministic system index***.
$I^{DVC}(i)$ = Deterministic Vulnerability	Reduced redundancy of a water distribution system upon a fallout or failure of	This index is not calculated by the AVAT Tool. It will require an EPANET model. Further for each

Importance Index (I^{DVC})/ Eqs (8) or (9)	component *i*. The index is measuring how much a component is contributing to the vulnerability of the system.	component the model need to be rerun for a situation where this component is "taken out" of the model. This is a ***deterministic*** index for each ***component*** in the system.
$I^{PVC}(i)$ = Probabilistic Vulnerability Importance Index / Eqs (10) or (11).	Increase in system performance if component performance of *i* is increased. The index is measuring how much a component is contributing to the vulnerability of the system.	This index is not calculated by the AVAT Tool. It requires one or more skeleton of the system corresponding to a fault tree / reliability block diagram. Further it requires failure data and repair times for each component. This is a ***probabilistic*** index for each ***component*** in the system.
$I^{IV}(i)$ = Inherent Vulnerability Index / Eq (13)	An index aggregating conditions that can cause a failure or an outage of component *i*, i.e., measuring inherent vulnerability.	This index is not calculated by the AVAT Tool. It requires a systematic evaluation of vulnerability attributes as shown in Table 1. This index is presenting the ***failure probability*** for each ***component*** based on inherent conditions.
$I^{V}(i)$ = Total Vulnerability Index / Eq (14)	A combination of the vulnerability contributing measure and the inherent vulnerability measure	This index is not calculated by the AVAT Tool. It is based on the contributing and inherent vulnerability indexes. This index is a ***component index***.

4 The Asset Vulnerability Assessment Tool (AVAT): technical description and demonstration

In this section AVAT's technical implementation is described and demonstrated through a water distribution system case study application.

The required data for AVAT consists of two parts:

- A steady state hydraulic simulation EPANET ((https://www.epa.gov/water-research/epanet) file which runs without any errors, and
- A data MS-Excel file with a structure as described below.

According to the user's settings, a vulnerability assessment will be performed including the calculation of Todini's Resilience Index for the network, the network's Connectivity Index and the Reachability Index for each node in the network, thus highlighting and ranking the most vulnerable areas of the system. In addition, the Criticality of each link asset in the network will be evaluated, which forms for each link/element its Link Critical Index (LCI).

The output of AVAT consists of tabular data exported to MS-Excel and color-bar figures which may be exported as images. In addition, some results are exported to an EPANET INP file for presentation purposes.

4.1 The AVAT tool: technical description

4.1.1 System requirements

AVAT was developed in MATLAB® and compiled as a standalone application and as a web application (see details below). As such, it mainly relies on MATLAB's Runtime libraries[2].

MATLAB Runtime Prerequisites are:

- The MATLAB Runtime installer requires administrator privileges to run.
- The version of the MATLAB Runtime that runs your application on the target computer must be compatible with the version of MATLAB Compiler or MATLAB Compiler SDK that built the deployed code.
- Do not install the MATLAB Runtime in MATLAB installation directories.
- The MATLAB Runtime installer requires approximately 2 GB of disk space.
- The MATLAB version used to develop AVAT is 2018b, thus the MATLAB runtime version needed is 9.5 which may be downloaded from MathWorks web site: https://www.mathworks.com/products/compiler/matlab-runtime.html. However, there is no need to install the Runtime module separately. AVAT's installation package will download and install the required libraries if they are not already installed on the client computer.
- In addition, MS-Excel® must be installed on the machine AVAT is installed on.

[2] The MATLAB® Runtime is a standalone set of shared libraries that enables the execution of compiled MATLAB applications or components on computers that do not have MATLAB installed. MATLAB Production Server™ requires a MATLAB Runtime instance to execute the deployed MATLAB applications it hosts.

4.1.2 Installing AVAT

AVAT may be used as a standalone program or as a web application. Generally, it is preferable to run AVAT in the standalone version due to possible security issues with MATLAB's application web server. However, it is possible to install and use the web version in cases were many users will need to use the program given that installing and updating the program on many machines is problematic.

Please note that there is an issue with the web application saving one of the result files. This issue will be resolved in future updates.

The next sections detail the installation processes for both the standalone and the web application.

4.1.2.1 Standalone version

The installation of the standalone version of AVAT is done by running the setup program named "AVAT_Setup.exe" (Figure 2). The program should be run by a user with administrator permissions.

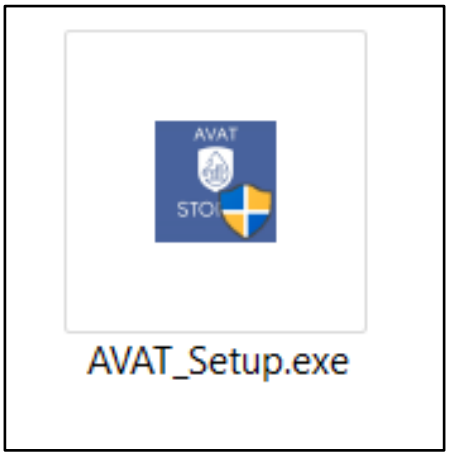

Figure 2: Setup program

After running the setup program, the AVAT splash screen will briefly appear as shown in Figure 3.

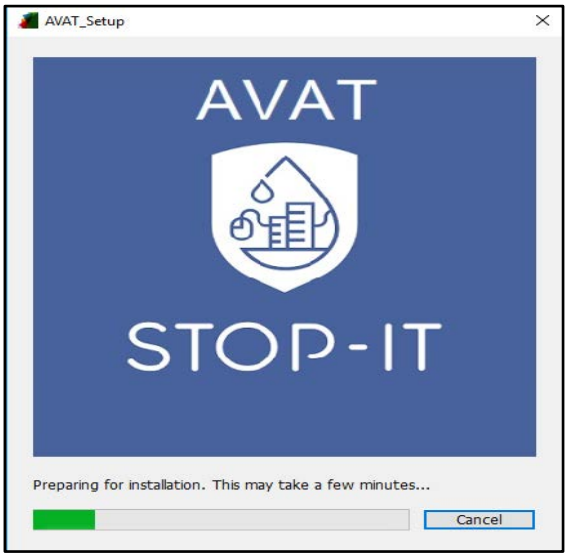

Figure 3: Installation splash screen

After initiation steps the AVAT installer screen will appear (Figure 4) with some general details about the program, to start the installation wizard click "Next".

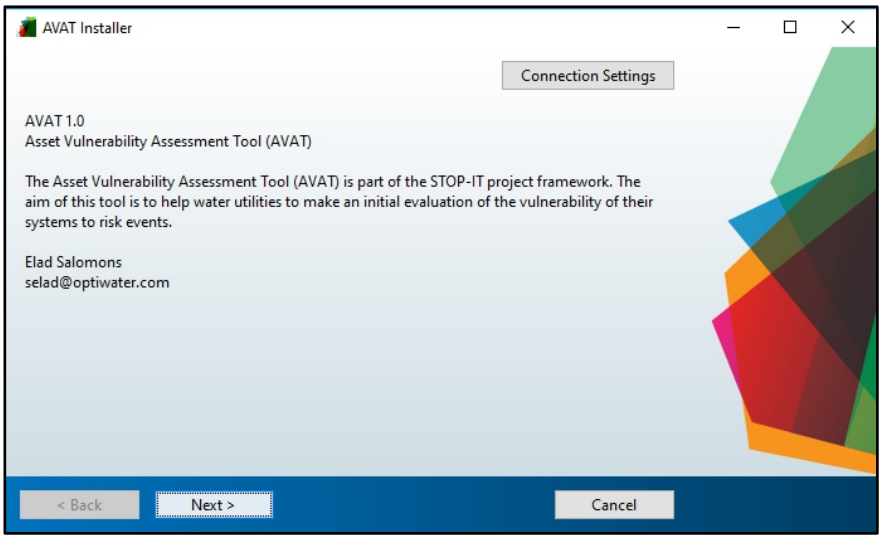

Figure 4: AVAT initial installation screen

In the installation option screen (Figure 5) the user can select the installation directory for the AVAT program and the option to create a shortcut for the program on the desktop screen.

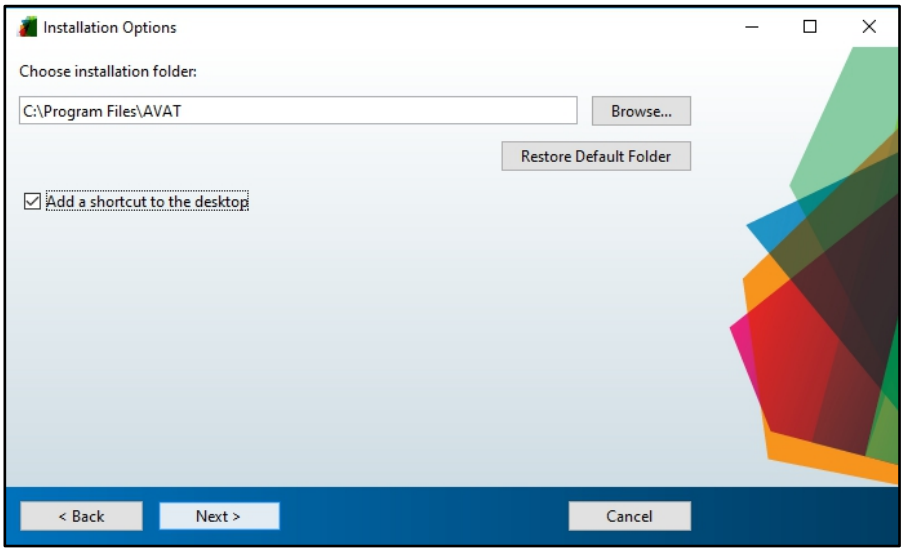

Figure 5: AVAT installation options screen

As detailed in the program requirements section, MATLAB runtime is required. If the runtime libraries are not already installed, the user should select the installation path for the MATLAB runtime (Figure 6). It is recommended to keep the installation path suggested by the installation utility.

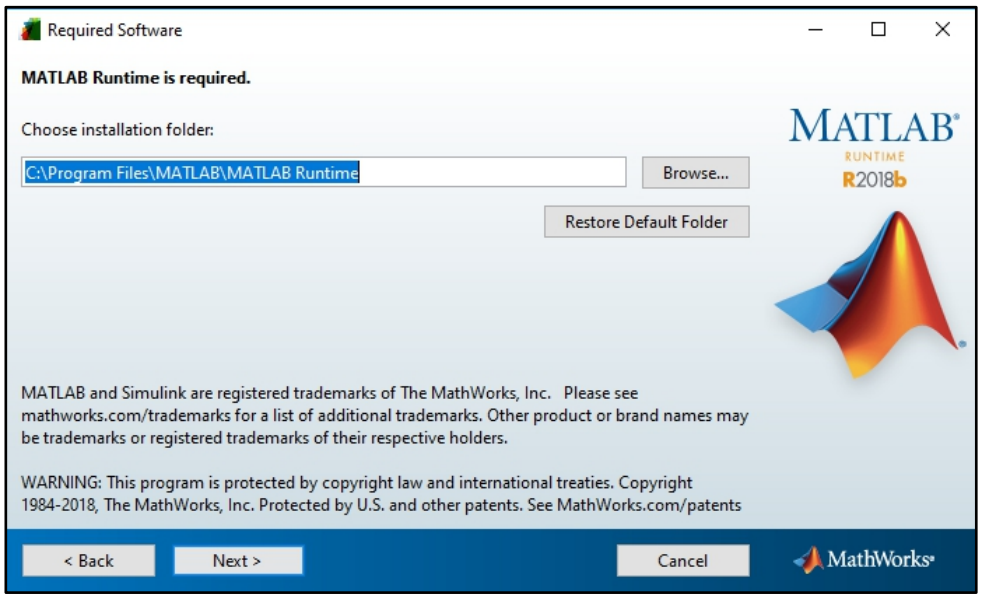

Figure 6: MATLAB runtime installation path

If the required runtime is already installed on the computer the user will be notified, as shown in Figure 7.

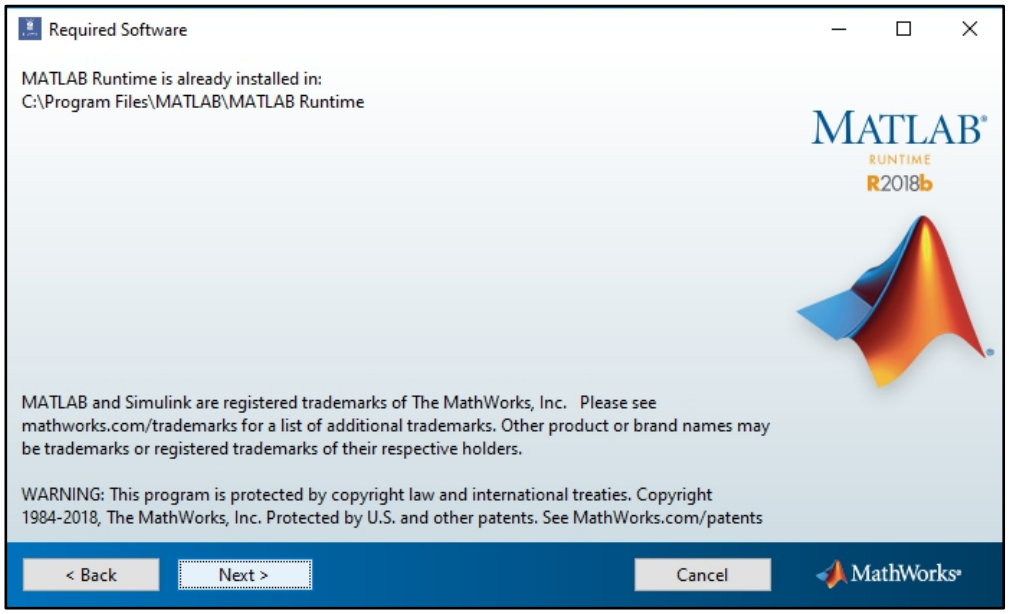

Figure 7: MATALB runtime is already installed

The general MATLAB license agreement must be approved (Figure 8).

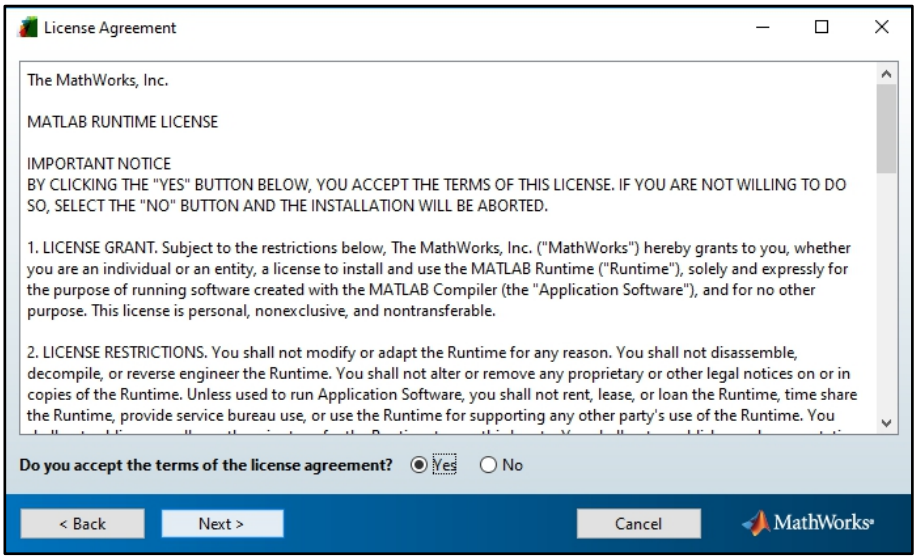

Figure 8: MATLAB license agreement

As a last step before installation, a summary of the installation option will be displayed. To begin the installation, click "Install" (Figure 9).

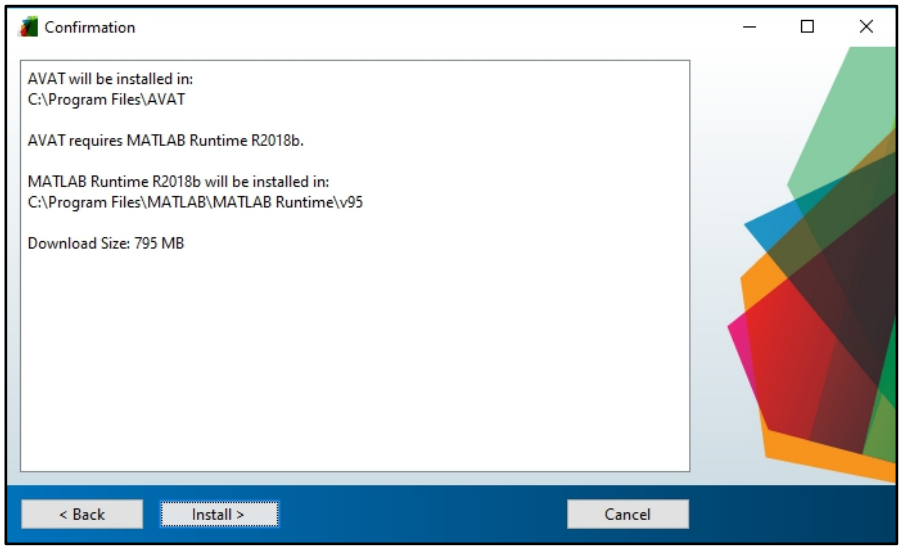

Figure 9: Installation confirmation

At this time the required software will be downloaded and installed (Figure 10).

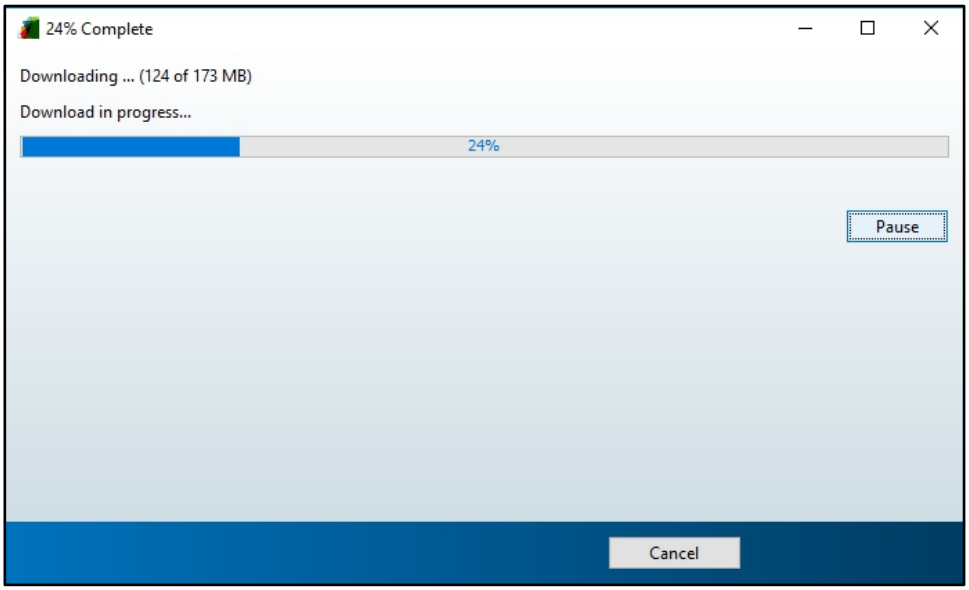

Figure 10: Installation progress

When the installation is complete a notice will be shown (Figure 11).

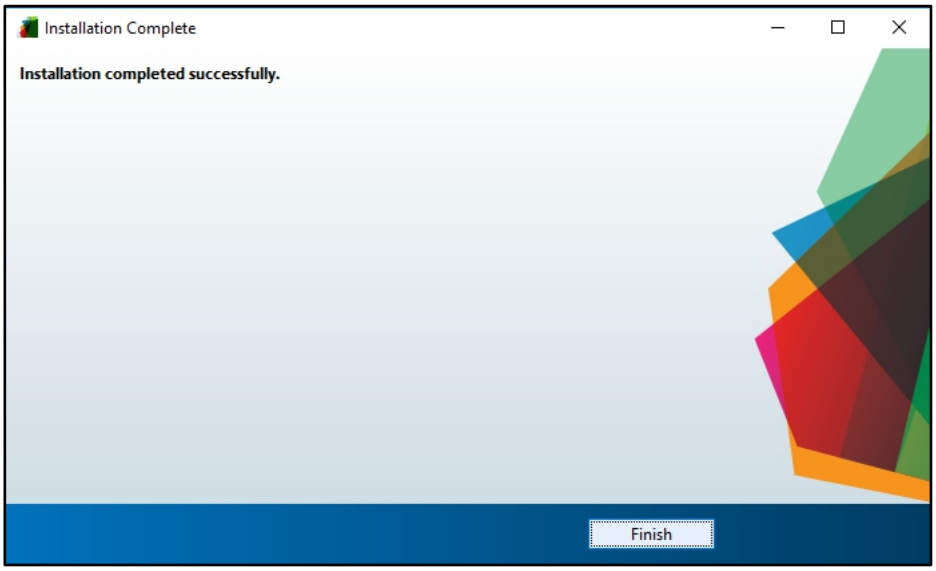

Figure 11: Installation complete screen

If requested during the installation steps, a shortcut for AVAT will be placed on the user's desktop (Figure 12).

Figure 12: AVAT shortcut

It should be noted that in some cases certain MS-Excel Add-ins may cause issues for AVAT to run properly. If this happens try the following: Open Excel > Office button > Excel options > Add ins > Manage > COM Add-inns > Go > Uncheck the add-ins there. This issue will be resolved in future updates.

4.1.2.2 AVAT Web version

The AVAT web application comes in the form of one file "AVAT.ctf". Installing the application on a MATLAB's web application server is as simple as dropping the ".ctf" file in

the application directory of the application server. However, the MATLAB's web application server must first be installed.

From Matworks website: MATLAB® Web App Server must be installed in a trusted intranet environment on dedicated hardware. The only purpose of the physical or virtual machine where the server is installed must be to host web apps that connect to the server. The server must never be exposed to the open Internet.

The MATLAB Web App Server hosts web apps, packaged using the Web App Compiler app. For web apps to work, the server must be installed and configured. The server mediates the HTTP/HTTPS communication between the client web browser and the packaged MATLAB web app. It has a home page listing all the available hosted web apps. The home page can be accessed from a browser using a URL.

Further installation information can be found on Mathworks web site:
https://www.mathworks.com/help/compiler/webapps/install-matlab-web-app-server.html

Following a successful installation of the web application server, the server should be configured by the steps detailed here:
https://www.mathworks.com/help/compiler/webapps/configure-matlab-web-app-server.html
Mainly, the server must be registered as a service in the server machine (Figure 13) and a few settings such as the server port number must be set (Figure 14).

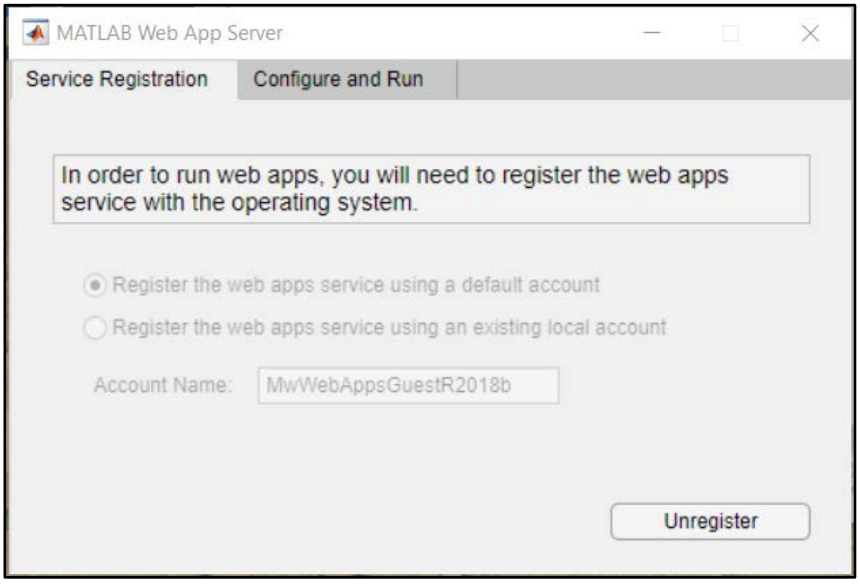

Figure 13: MATLAB web application server registration

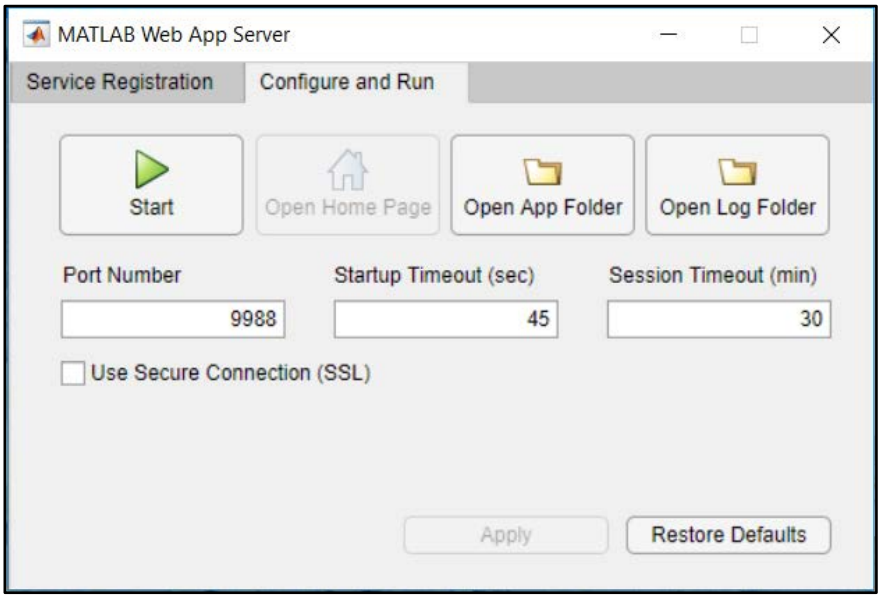

Figure 14: MATLAB web application server settings

To start the MATLAB web application server, click the "Start" button. As described above, to install AVAT on the web server, the "AVAT.ctf" should be placed in the "App Folder" (see Figure 15).

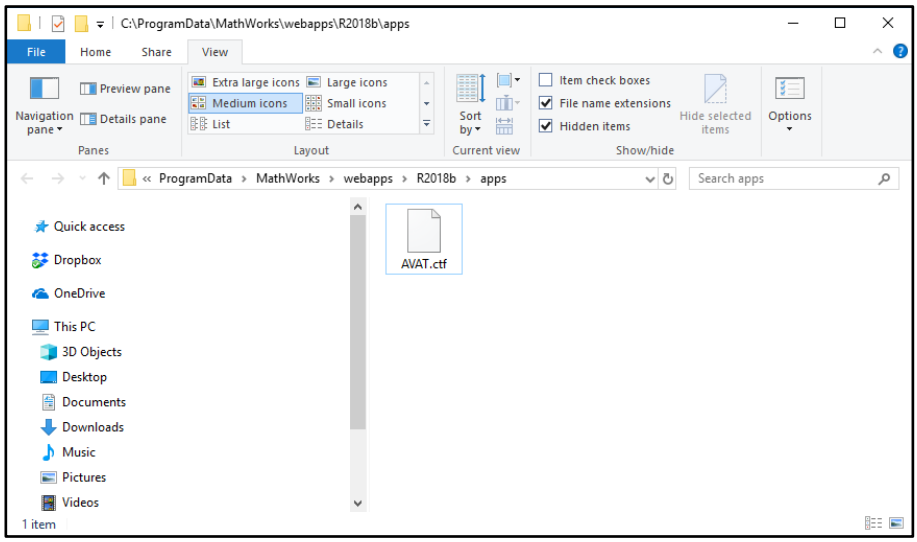

Figure 15: MATLAB web applications folder

To see the available applications on the web server, click the "Open home page" button (Figure 16).

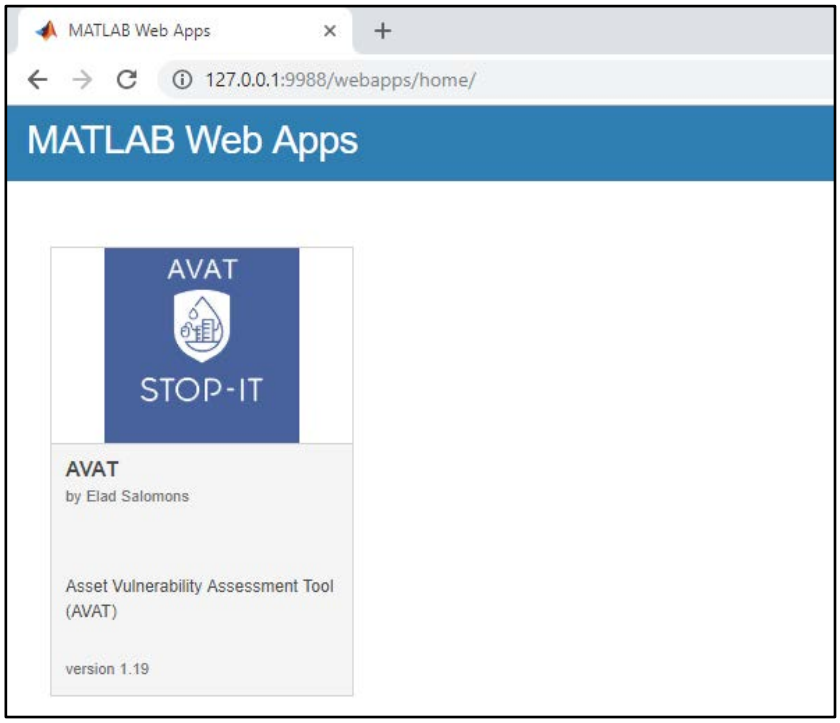

Figure 16: MATLAB web applications server home page

4.1.3 MATLAB Web App Server Security

It should be noted that using the MATLAB web applications server may cause security issues. As Mathworks suggests[3]: Installing and running the server on your network exposes your network and file system to risks. The machine running the server is most at risk from accidental or deliberate misuse of deployed web applications. Therefore, you must install the server software only on dedicated hardware. This can be a physical or virtual machine whose only purpose is to host web applications that connect to the server software. Using a physical or virtual machine limits the risk in the event the machine is compromised.

The MATLAB Web App Server alters the security profile of the machine on which it is running. The installation process creates a server user account with low privileges. This new account has read-only permission to the app folder created during the installation of the server. However, through a process known as privilege escalation, attackers may be able to exploit bugs in the operating system or network to obtain the privileges of ordinary or even administrative users. They can then attempt to access files or other intellectual property without permission.

The server relies on the authentication and authorization scheme of its host machine and network. Other than supporting HTTPS, it does not contain any additional mechanisms for authenticating or authorizing web application users.

[3] https://www.mathworks.com/help/compiler/webapps/matlab-web-app-server-security.html

You may be able to mitigate some of these risks by:

- Restricting network access: Only trusted users can access the server and its associated applications.
- Executing only trusted applications: Trust applications developed by only well-known, trusted, and authenticated sources.
- Limiting application functionality: Include in the application only those features of MATLAB required for the application to perform its function

4.1.4 AVAT input data requirements

As noted in the above sections, AVAT requires limited data to run. The main requirement is a steady-state EPANET model of the network analyzed. Figure 17 shows the EPANET model of the C-Town network (Ostfeld et al., 2012) which will be used throughout this guide as a demo network. As can be seen in Figure 17, the total time duration of the network's simulation is set to 0:00 hours which indicates a steady-state simulation. The hydraulic model must be solved within EPANET successfully without errors (some warnings are allowed). The demo CTOWN.INP file is attached to the AVAT installation package.

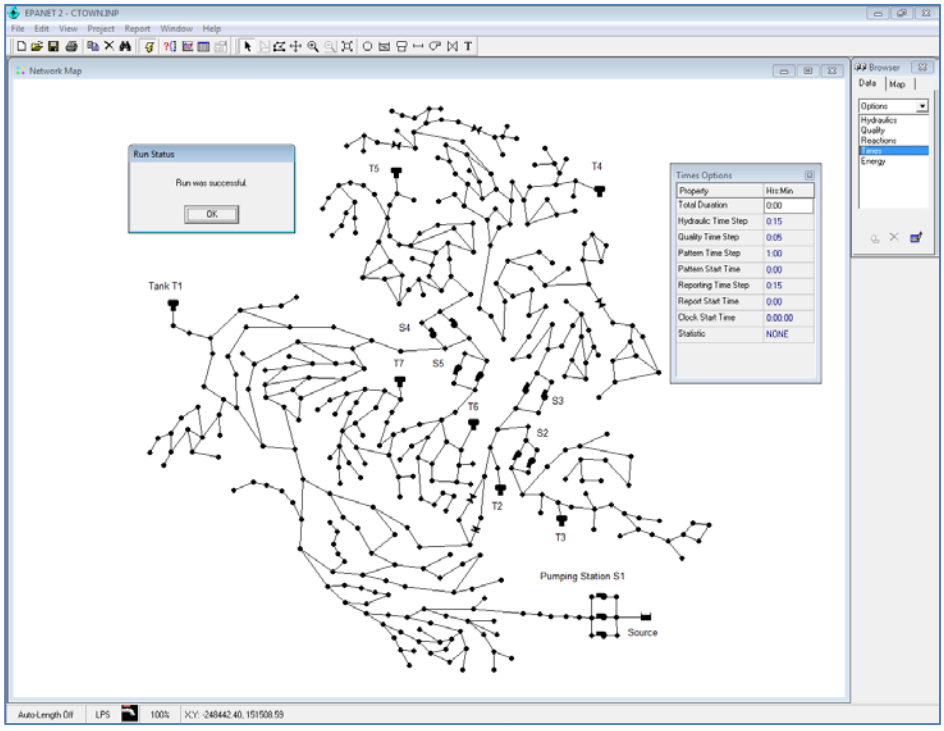

Figure 17: EPANET model of C-Town

In addition to the EPANET model of the network, an Excel file with additional information is required. The Excel file must include at least the following three sheets:

- "Defaults" – default values for pipes, pumps and valves failure probabilities and the minimum pressure required for calculating the Todini's Index (Figure 18).

- "Elements probabilities" – specific elements failure probabilities which overrule the default settings including the element "Link ID" from the EPANET model and the specific failure probability (Figure 19).
- "Sources" – a list of the EPANET model Node IDs which are the networks sources (Figure 20).

	A	B
1	**Parameter**	**Value**
2	Pipe failure probability	1.00E-04
3	Pump failure probability	0.001
4	Valve failure probability	0.001
5	Minimum pressure for Todini	30
6		

Figure 18: Default settings for AVAT

	A	B	C
1	**Link ID**	**Probability**	
2	P53	1.00E-04	
3	P399	1.00E-04	
4			
5			

Figure 19: List of specific line failure probabilities

	A	B	C
1	**Sources ID**		
2	R1		
3			
4			
5			

Figure 20: List of the Networks sources

The demo CTOWN.XLSX file is attached to the AVAT installation package. The data above are for demonstration purposes. Real data on water distribution system component failure probabilities can be found in various references (e.g., Mays, 1989).

4.1.5 Running AVAT

As described above, AVAT can be run as a standalone application or as a web application. Running the program as a standalone application is accomplished by clicking the desktop shortcut or by navigating to the installation directory and running the AVAT.exe file.

To run the program as a web application, the URL of the MATLAB web application server must be used. In both cases the program's user interface is the same. The rest of this guide assumes that the standalone version is used (recommended).

Upon running the program, a splash screen will be shortly displayed and then the first screen of the program will be presented as shown in Figure 21.

The program is built as a step-by-step wizard which guides the user throughout the process.

Each screen includes a "Next" and a "Previous" button to guide the user. These buttons are enabled only if the necessary tasks have been fulfilled successfully in the current screen.

In general, AVAT includes 3 stages:

- Input file selection and validation
- Simulation options
- Results

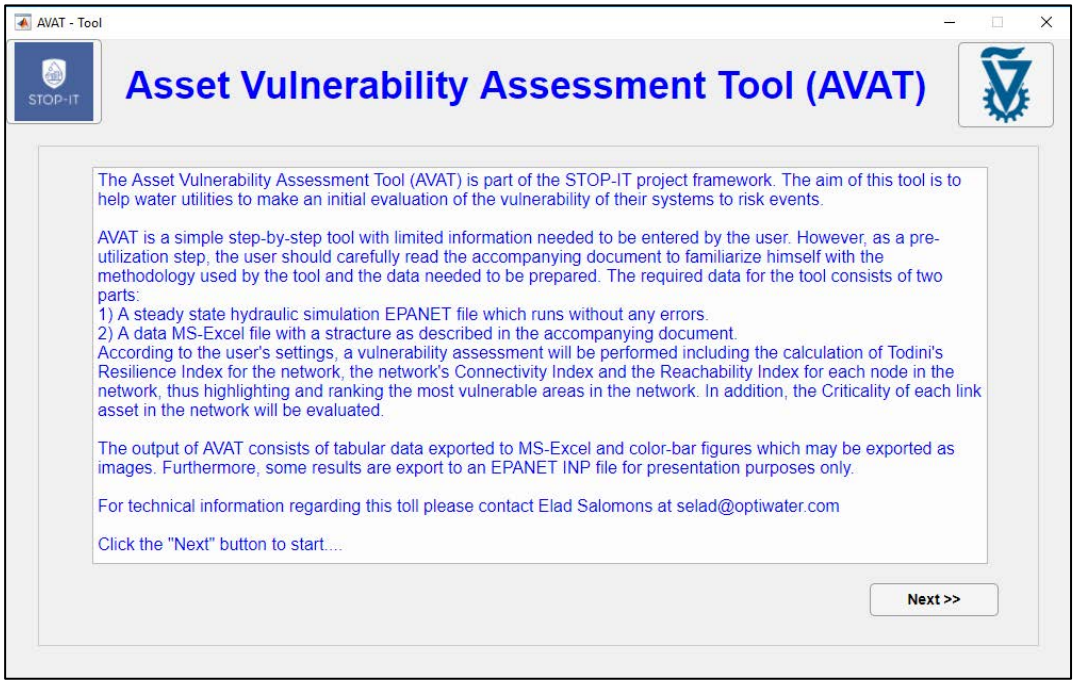

Figure 21: AVAT first screen

4.1.5.1 Input file selection and validation

In the first step in the process, the EPANET INP file and the Excel data file must be selected and loaded. First click the "Select INP file" button (Figure 22).

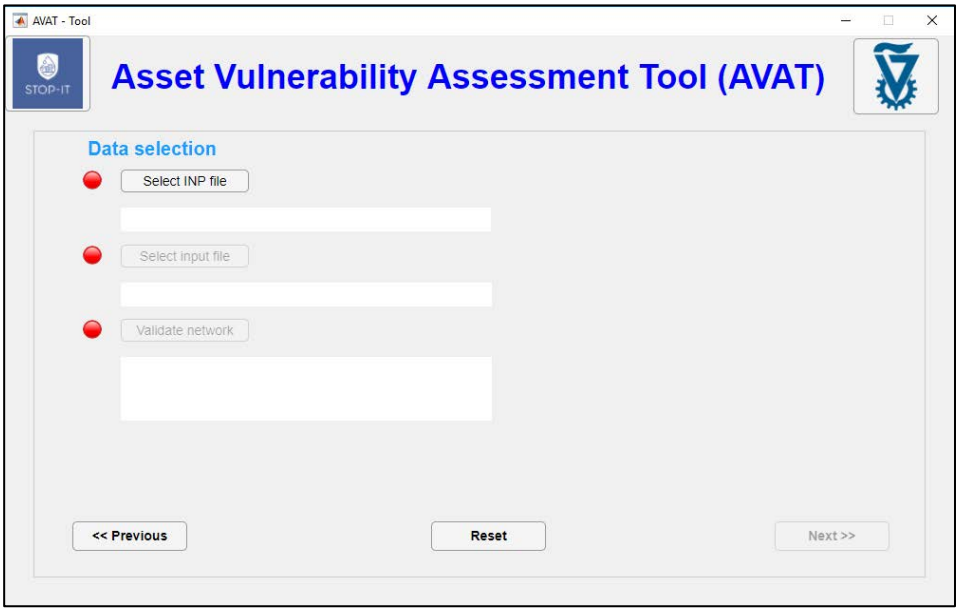

Figure 22: Input data selection and validation screen

From the file selection form, select the networks INP file (Figure 23). For this demo the CTOWN.INP file is selected.

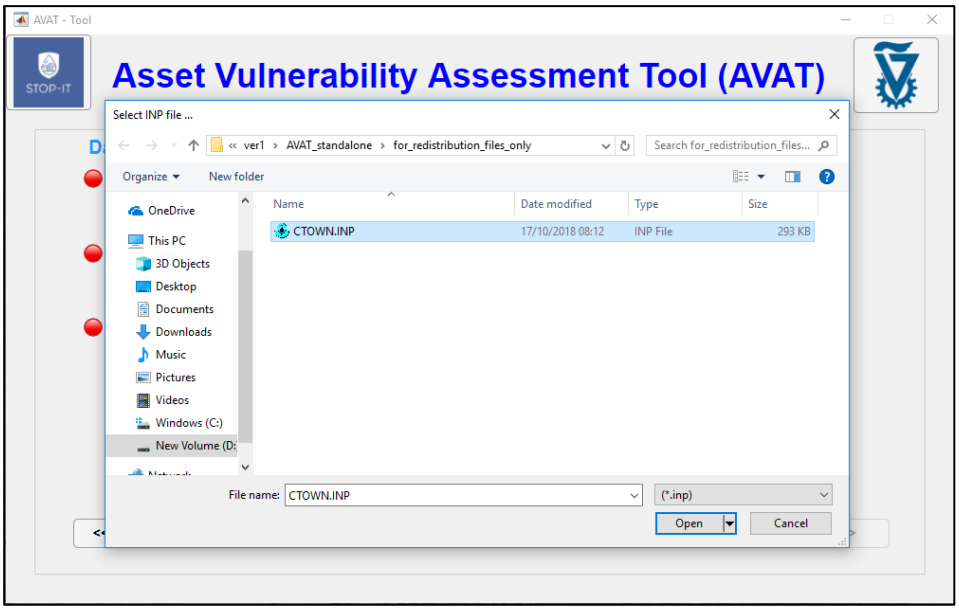

Figure 23: INP file selection form

Once an INP file is chosen, a green symbol will appear to the left of the selection button and the name of the file will be shown under it. The Excel data file should be selected next. The program automatically searches for an Excel file with the same name as the selected INP file (excluding the file extension). If such a file is found, its name will be shown below the "Select input file" button and the second green symbol will be shown (Figure 24).

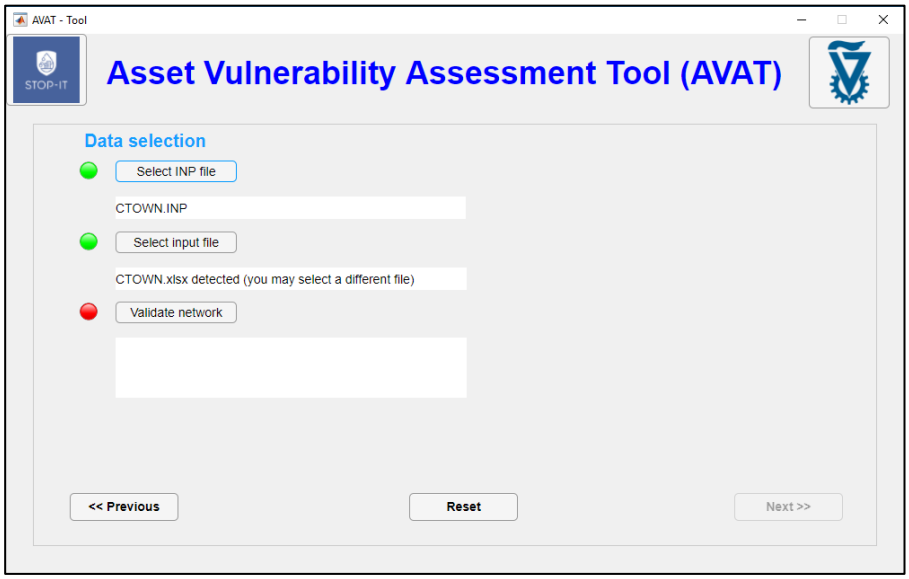

Figure 24: INP and data files selection

If an Excel file with a matching name is not found, or the user wishes to select a different Excel data file, the "Select input file" button should be clicked and the requested file should be selected (Figure 25).

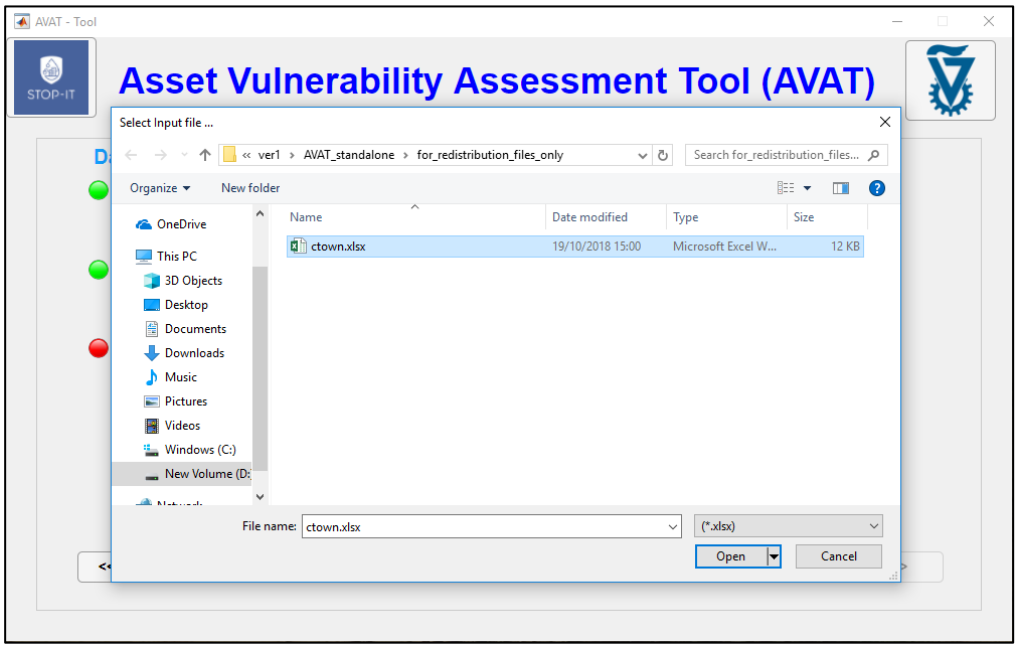

Figure 25: Excel data file selection

For this demo the CTOWN.XLSX data file is used. The next step is to validate the two input files. This is done by clicking the "Validate network" button (Figure 26).

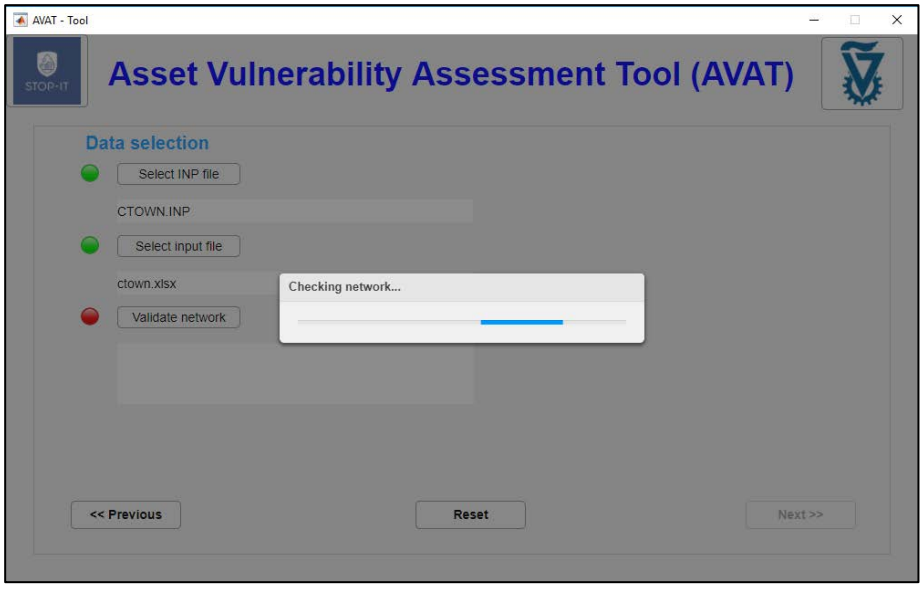

Figure 26: Network validation process

The validation process takes a few seconds during which the INP file is loaded and the structure of the Excel file is checked. In addition, the list of link IDs in the "Elements probabilities" sheet is validated as well as the list of source IDs in the "Sources" sheet. If an error is detected it will be shown below the validation button. If all validation procedures end

successfully, the "Network data is OK" message will appear, the network layout will be plotted, a green symbol will be shown, and the "Next" button will be enabled (Figure 27). At any stage, the "Reset" button may be clicked and all the input selections will be deleted and reset.

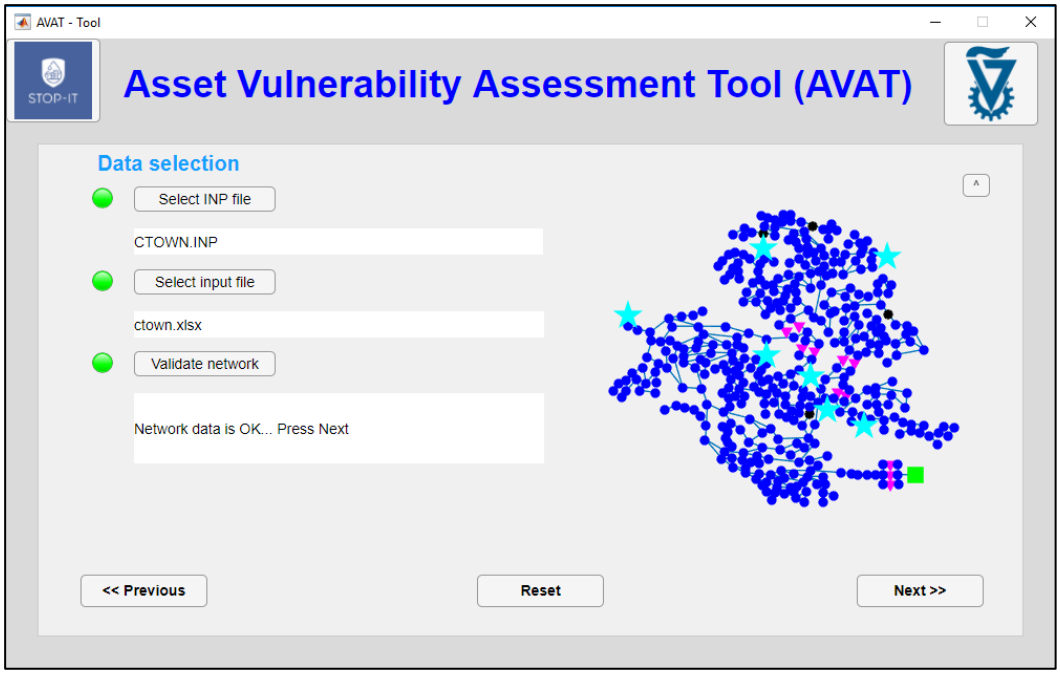

Figure 27: INP file loaded

4.1.5.2 Simulation options

In the next screen of the wizard, the simulation options can be set. First, the requested network wide and element specific indexes can be selected for calculation. Since calculating most of the indexes requires mainly the same procedures, it does not make a lot of difference if some indexes are not selected.

It is recommended to keep all options selected as shown in Figure 28. However, selecting the number of simulations utilized in the Monte Carlo simulations, effects the calculation time and the AVAT output. For the C-Town network, running 100,000 simulations takes less than one minute and provides stable results.

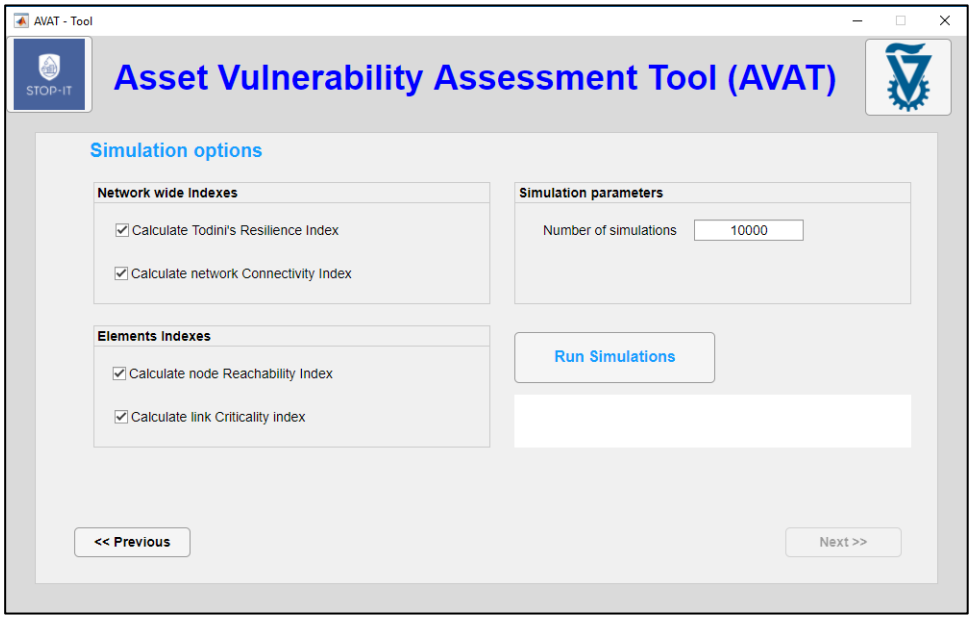

Figure 28: AVAT simulation options screen

Once the simulation options have been set, the "Run Simulations" button can be pressed to perform the calculations (Figure 29).

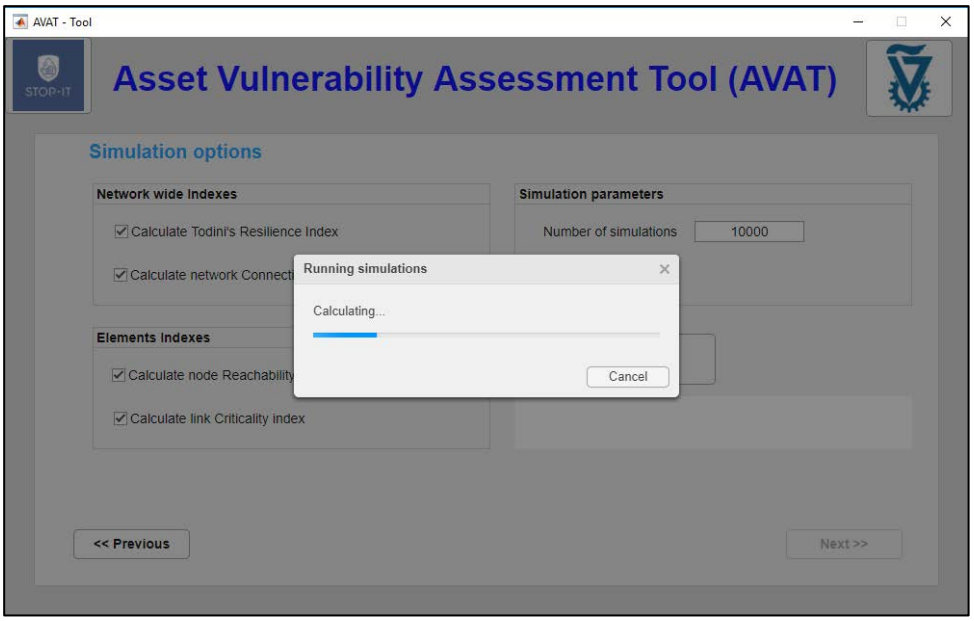

Figure 29: AVAT running simulation screen

If the simulations end successfully, a message "All is fine…" is shown (Figure 30) and the "Next" button can be pressed to access the results screen.

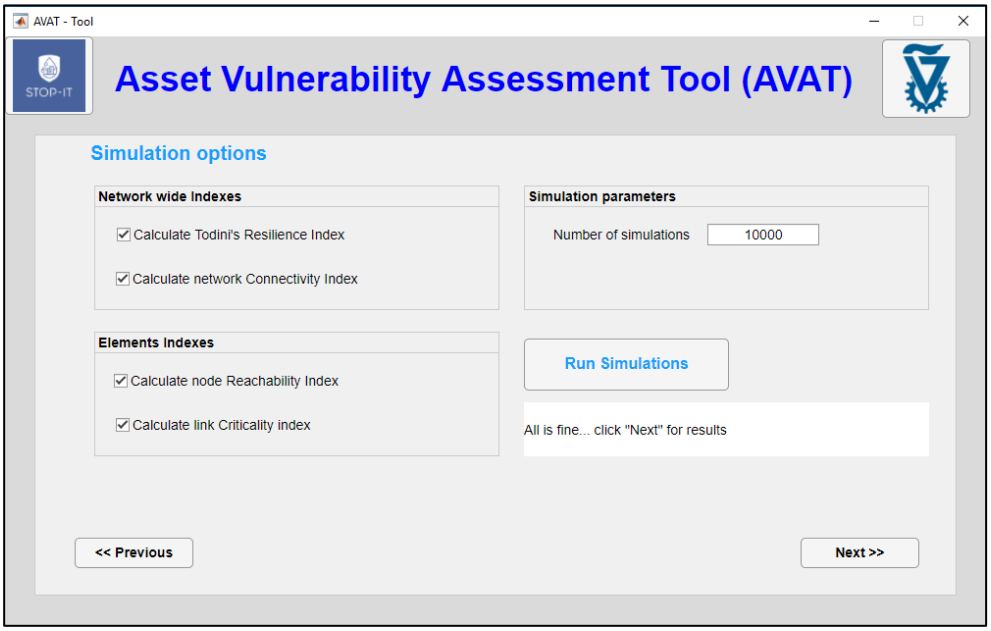

Figure 30: AVAT simulations ended

4.1.5.3 Simulation results

Following a successful simulation run, the AVAT results screen is shown (Figure 31). In this screen the network wide indexes, Todini's resilience index and the Connectivity index, are presented.

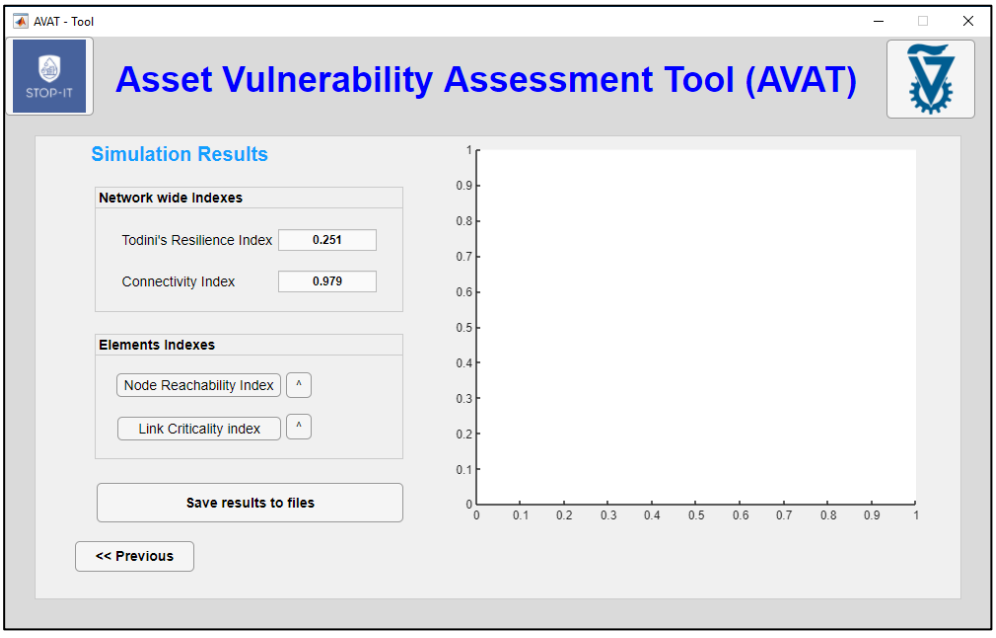

Figure 31: AVAT simulations results screen

The elements indexes, Nodes Reachability and Links Criticality, can be presented as figures by pressing the appropriate button (Figure 32 and Figure 33).

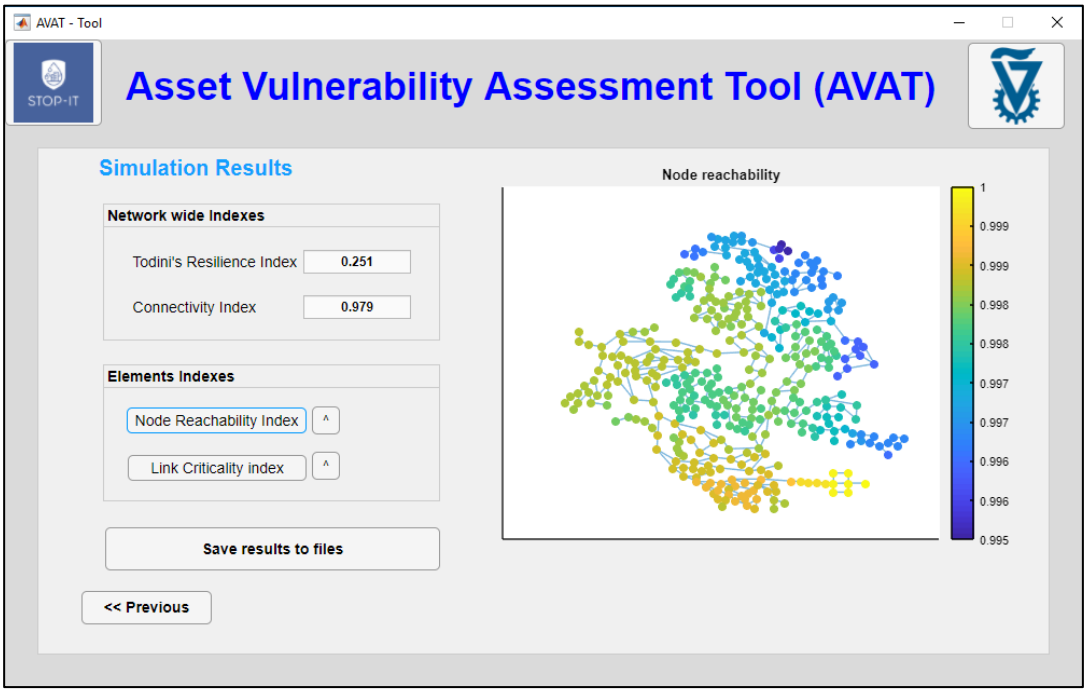

Figure 32: AVAT results – nodes reachability index

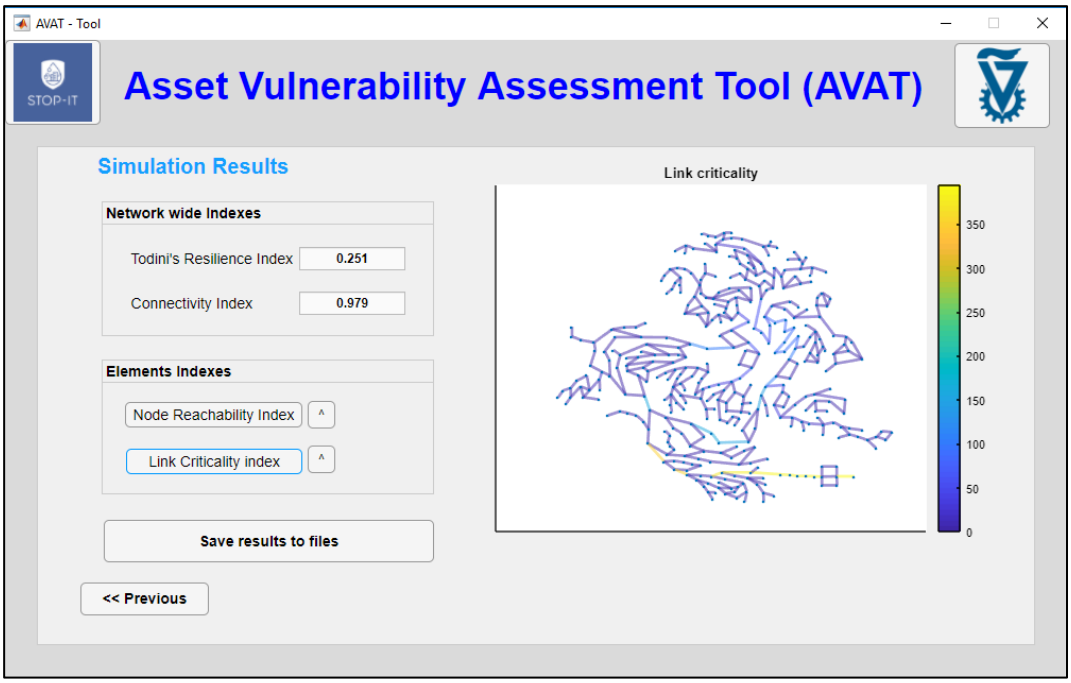

Figure 33: AVAT results – links criticality index

These two result figures can be enlarged by pressing the buttons marked as "^". As a result, a new window will be opened with the requested figure (Figure 34). From this new window, the figure can be edited, saved or printed.

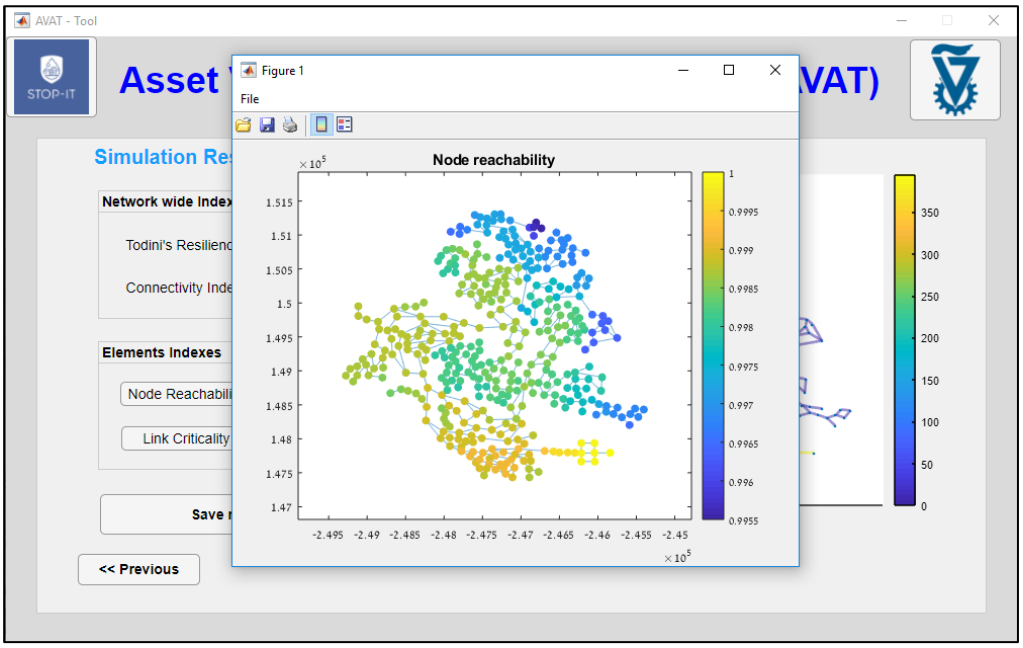

Figure 34: AVAT results - enlarged figure

AVAT results may be exported to files by clicking the "Save results to file" button. First, an INP file will be saved (Figure 35) with the Nodes Reachability index entered as the "Initial Quality" of each node. This option is intended to be used as a "real" simulation file but as a way to use the EPANET graphical user interface (GUI) to present AVAT results in combination with other options available within the EPANET GUI. For example, a contour map can be plotted for the Node Reachability index (Figure 36).

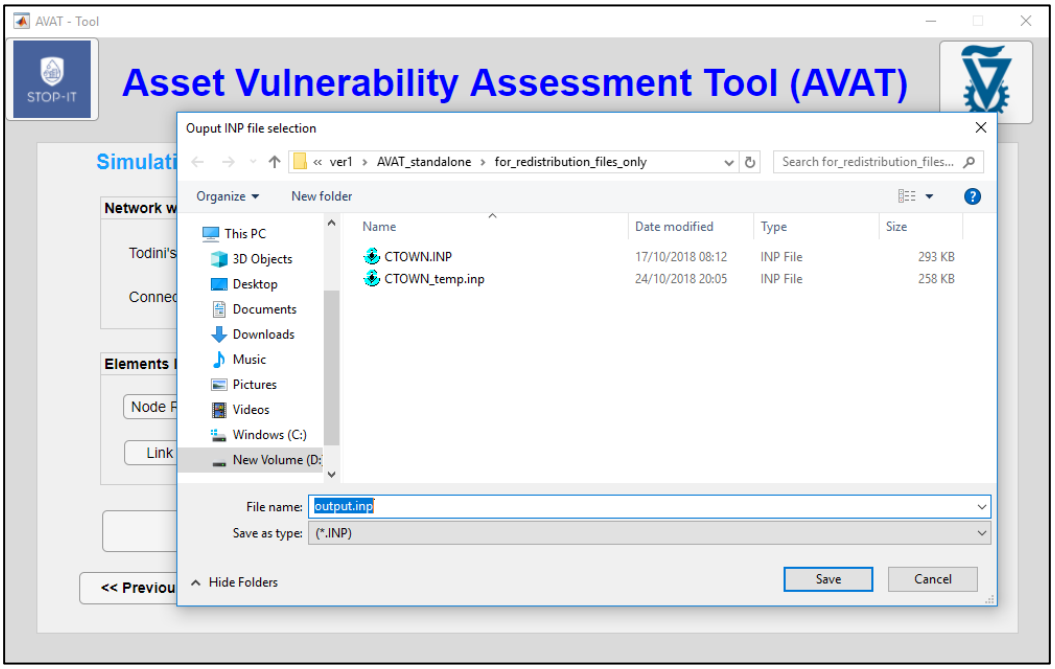

Figure 35: AVAT results – Reachability index output as INP file

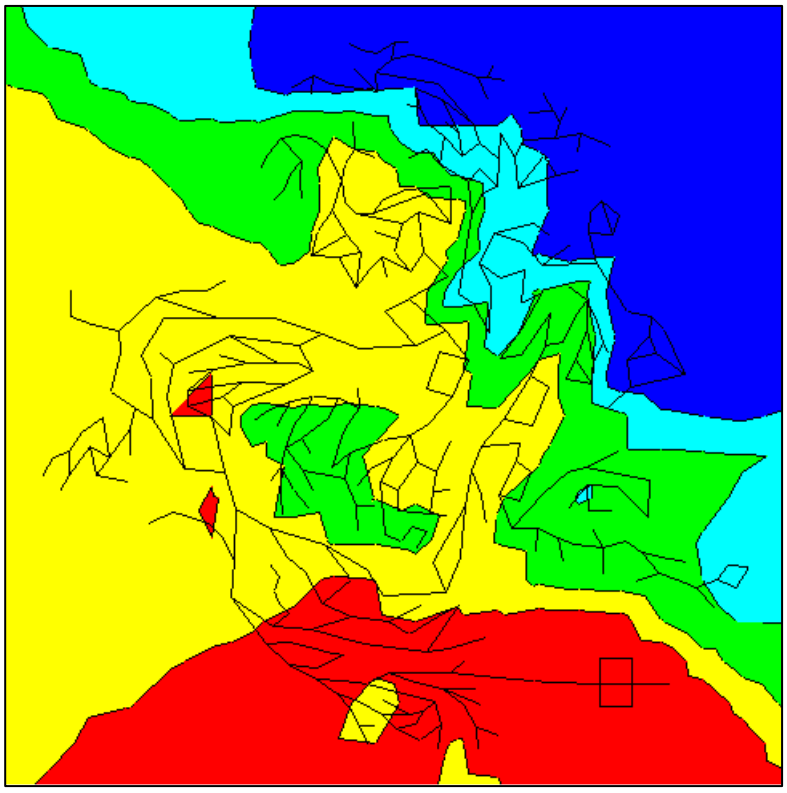

Figure 36: Nodes Reachability index as a contour map in EPANET GUI

In addition, an Excel file with the full numeric results will be saved (Figure 37). The exported Excel file will include a sheet with the network wide indexes, a sheet with each node's reachability index, and a sheet with each link's criticality index. These values could then be used for external calculations and reports.

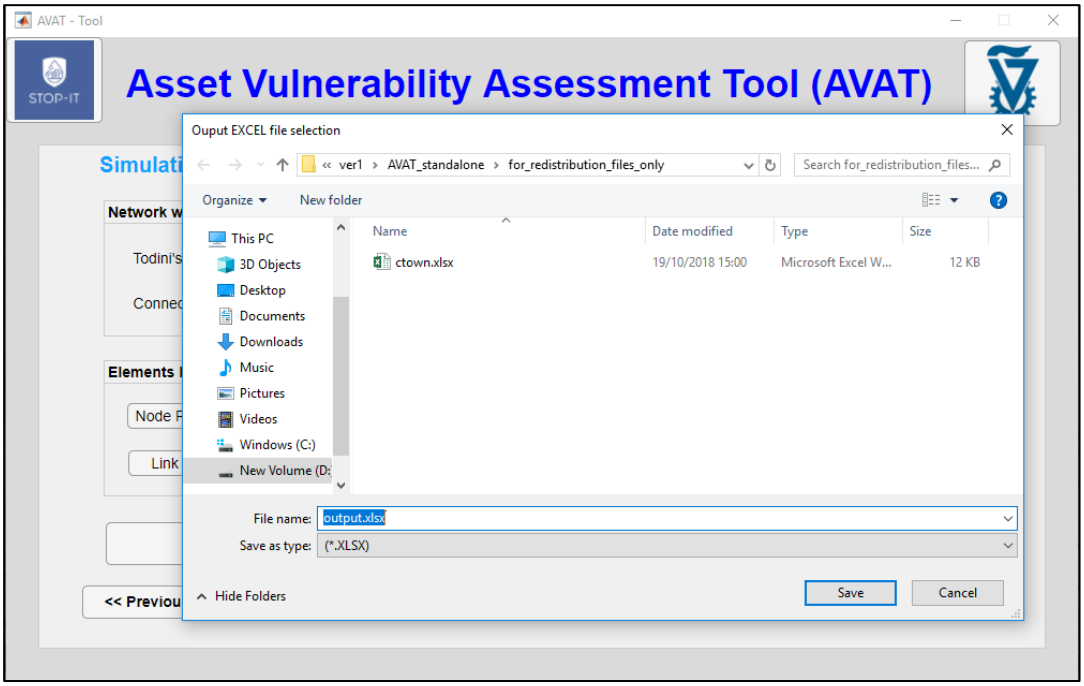

Figure 37: AVAT results – export to Excel file

4.2 Case study demonstration

In this section AVAT is further demonstrated on C-Town (Ostfeld et al., 2012) through a base run and sensitivity analyses.

4.2.1 Base run

A base run was executed using the sample water distribution network C-Town for 10,000 failure situations. The AVAT software calculated a Todini's resilience index of 0.251 and a connectivity index of 0.979.

Maps of the RI and the LCI are presented in Figure 38 and Figure 39 below.

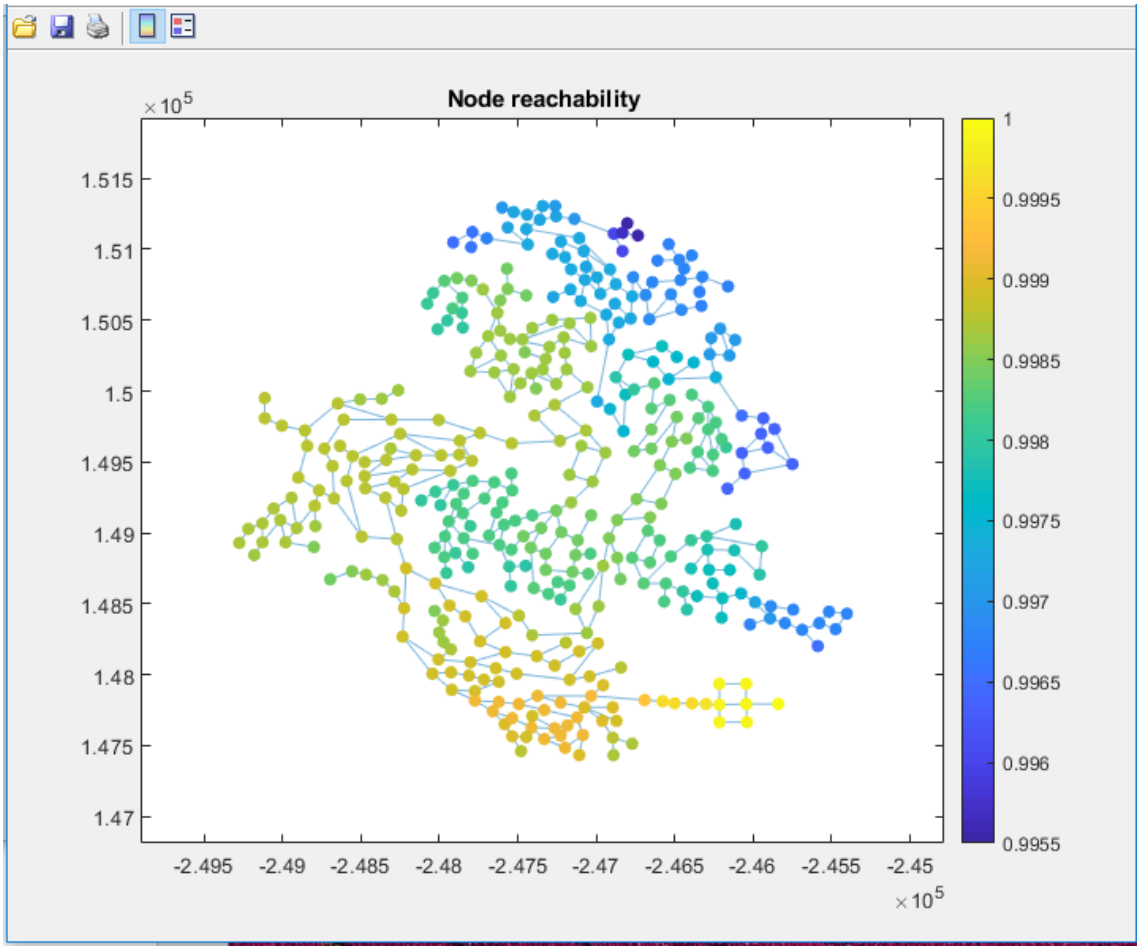

Figure 38: Basic run - node reachability index

As outlined in section 2, reachability is defined as the probability that a given demand node is connected to at least one source. In the RI map it can be noted that nodes with greater proximity to the source have a higher node reachability whereas nodes further from the source have a lower node reachability. This result is logical given that in C-Town all pipes have the same probability of failure and given that the further a node is from the source the more opportunities there are for failure of the nodes along the way.

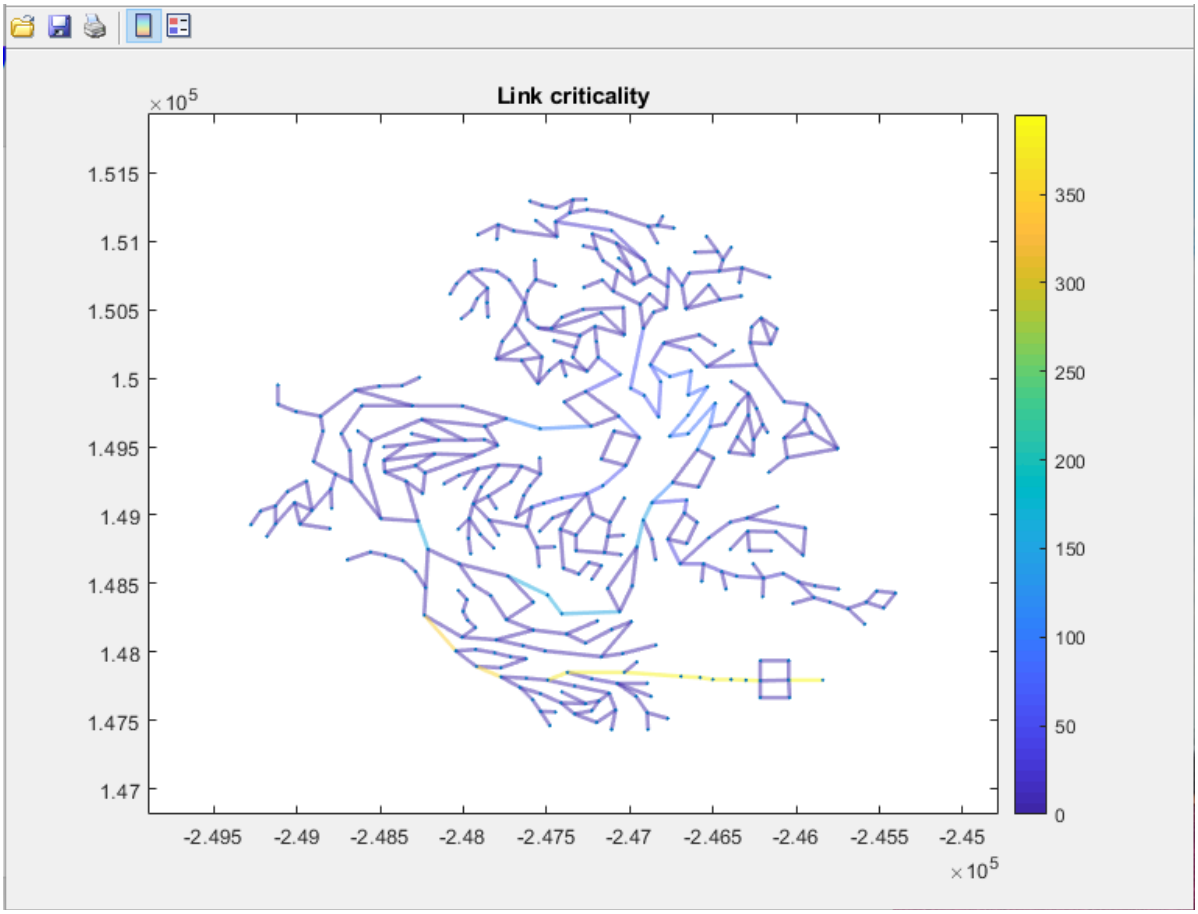

Figure 39: Basic run - link criticality index

The LCI identifies the number of disconnected nodes caused by an element failure. As can be expected lines which connect the source to the network which appear in yellow have the highest link criticality. The light blue lines have lower but still elevated link criticality index given that they connect between concentrated water distribution areas.

4.2.2 Sensitivity analyses

In order to perform a sensitivity analyses minimum pressure demand requirements and pipe failure probability were altered.

4.2.2.1 Sensitivity analyses of minimum pressure demand

As can be seen in Figure 40, the minimum pressure was gradually elevated from 30 meters in the base run to 40 meters. As the minimum pressure rises, the TI drops from 0.25 to 0.17. The TI measures how close a water distribution network operates relative to its minimum required level therefore, the higher the TI the more excess energy can be found in the system. Thus, as can be expected the higher the minimum pressure demand, the less excess energy found in the system.

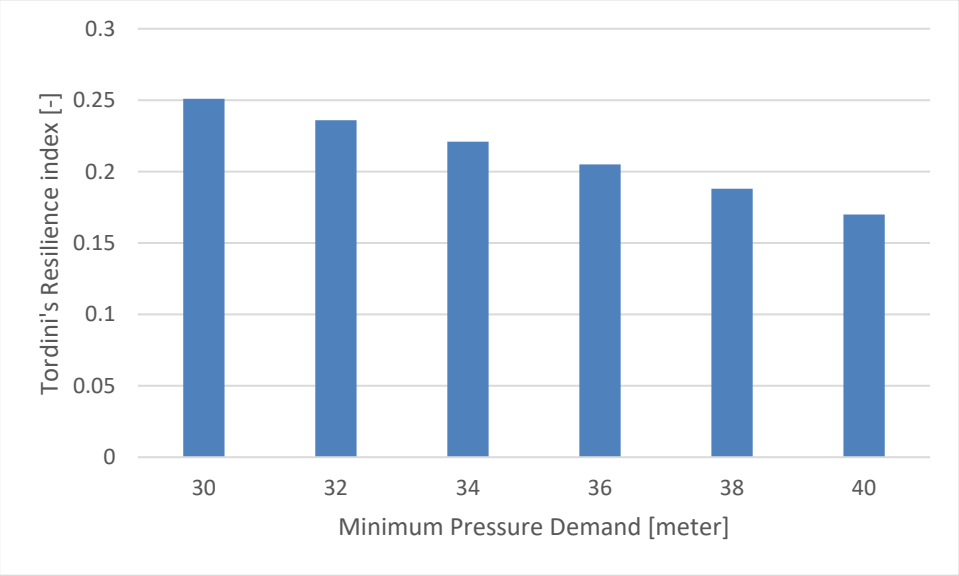

Figure 40: Effect of changing minimum pressure demand on the TI

4.2.2.2 Sensitivity analysis of probability of failure

In this sensitivity analysis the probability of failure of a cluster of lines was raised from 0.0001 in Figure 41 to 0.01 in Figure 42 with the expectation that more failures would occur in the cluster. As predicted, the nodes in the cluster (encircled below) are lighter and therefore have a lower reachability in Figure 41 of approximately 0.9987 compared to a reachability of approximately 0.9975 in Figure 42.

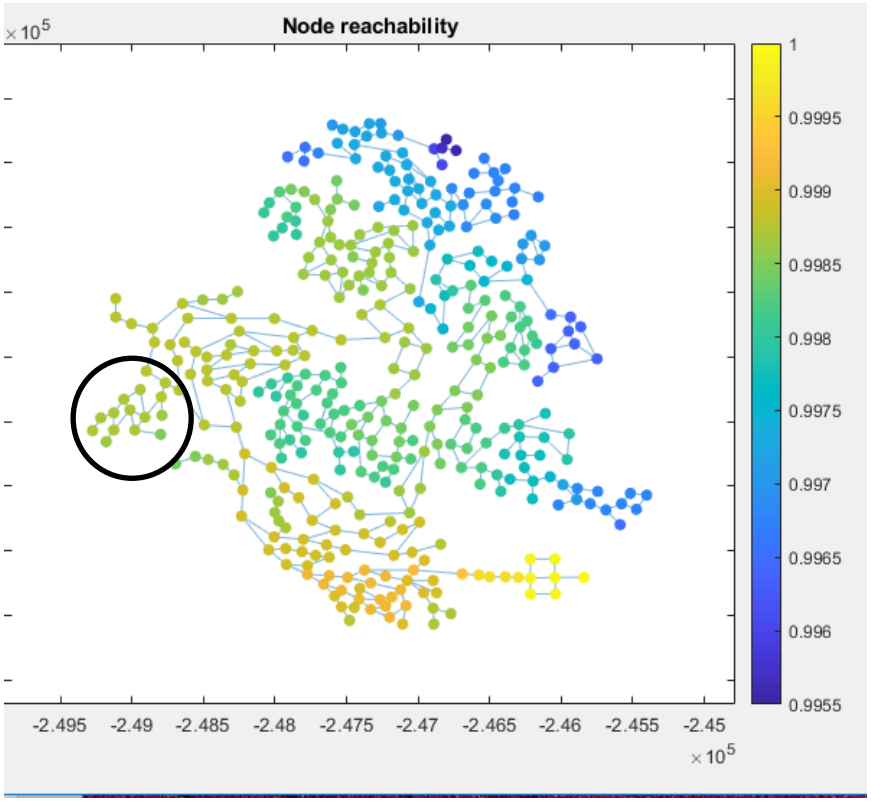

Figure 41: Base node reachability

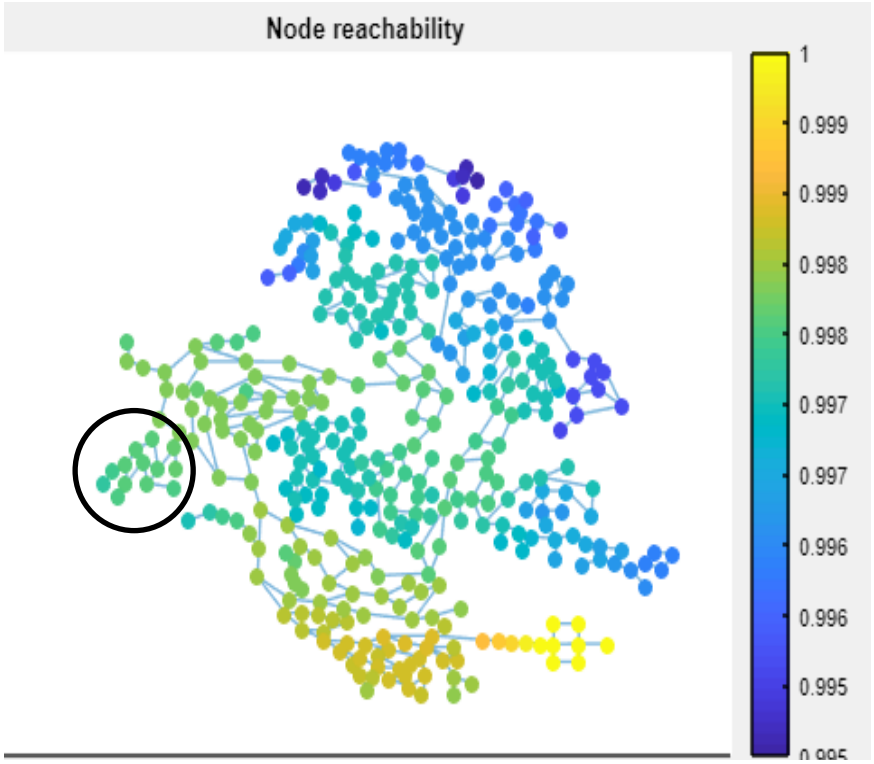

Figure 42: Node reachability after failure probability adjustment

References

Atkinson S., Farmani R., Memon F. A.; and Butler D. (2014). "Reliability indicators for water distribution system design: comparison." Journal of Water Resources Planning and Management, Vol. 140, No. 2., pp. 160-168.

Awumah K., Goulter I. C., and Bhatt, S. K. (1990). "Assessment of reliability in water distribution networks using entropy based measures." Stochastic Hydrology and Hydraulics, Vol. 4, Issue 4, pp. 309-320.

Baños R., Reca J., Martínez J., Gil C., and Márquez A. L. (2011). "Resilience indexes for water distribution network design: a performance analysis under demand uncertainty." Water Resources Management, Vol. 25, pp. 2351-2366, doi: 10.1007/s11269-011-9812-3.

Farmani R., Walters G. A., Savic D. A. (2005). "Trade-off between total cost and reliability for anytown water distribution network." Journal of Water Resources Planning and Management, Vol. 131, No. 3, pp. 161-171.

Gheisi A. and Naser G. (2015). "Multistate reliability of water distribution systems: comparison of surrogate measures." Journal of Water Resources Planning and Management, Vol. 141, No. 10, pp. 04015018-1 - 04015018-9.

Gheisi A., Forsyth M., and Naser G. (2016). "Water distribution systems reliability: a review of research literature." Journal of Water Resources Planning and Management, Vol. 142, No. 11, pp. 04016047-1 – 04016047-13.

Goharian E., Burian S. J., and Karamouz M. (2018). "Using joint probability distribution of reliability and vulnerability to develop a water system performance index." Journal of Water Resources Planning and Management, Vol. 144, No. 2, pp. 04017081-1 - 04017081-12.

Greco R., Nardo Di A. and Santonastaso G. (2012). "Resilience and entropy as indices of robustness of water distribution networks." Journal of Hydroinformatics, Vol. 14, No. 3, pp. 761-771.

Jayaram N. and Srinivasan K. (2008). "Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing." Water Resources Research, Vol. 44, W01417, doi: 10.1029/2006WR005316

Jung D. and Kim J. H. (2018). "Water distribution system design to minimize costs and maximize topological and hydraulic reliability." Journal of Water Resources Planning and Management, Vol. 144, No. 9, pp. 06018005-1 – 06018005-9.

Mays L. (Ed.) (1989). "Reliability analysis of water distribution systems." Published by the American Society of Civil Engineers, ISBN 0-87262-712-8

Ormsbee L. and Kessler A. (1990). "Optimal upgrading of hydraulic network reliability." Journal of Water Resources Planning and Management, Vol. 116, No. 6, pp. 784-802.

Ostfeld A., Kogan D., and Shamir U. (2002). "Reliability simulation of water distribution systems-single and multiquality." Urban Water, Vol. 4, No. 1, pp. 53-61.

Ostfeld A., Salomons E., Ormsbee L. et al. (+ 41 co-authors) (2012). "Battle of the water calibration networks", Journal of Water Resources Planning and Management Division, ASCE, Vol. 138, No. 5, pp. 523 – 532

Ostfeld A., Oliker N. and Salomons E. (2014). "Multi-objective optimization for least cost design and resiliency of water distribution systems", Journal of Water Resources Planning and Management Division, ASCE, Vol. 140, No. 12, 04014037, http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000407

Prasad D. T. and Park N. (2004). "Multiobjective genetic algorithms for design of water distribution networks." Journal of Water Resources Planning and Management, Vol. 130, No. 1, pp. 73-84.

Quimpo R. G. and Shamsi U. M. (1991). "Reliability based distribution system maintenance." Journal of Water Resources Planning and Management, Vol. 117, No. 3, pp. 321-339.

Raad D. N., Sinske A. N., Vuuren J. H. (2010). "Comparison of four reliability surrogate measures for water distribution systems design." Water Resources Research, Vol. 46, W05524, doi: 10.1029/2009WR007785

Reca J., Martinez J., Banos R., and Gil C. (2008). "Optimal design of gravity-fed looped water distribution networks considering the resilience index." Journal of Water Resources Planning and Management Division, ASCE, Vol. 134, No. 3, pp. 234-238.

Rausand, M. (2011). “Risk Assessment: Theory, Methods, and Applications”. WILEY. ISBN: 9780470637647.

Shafiqul I. M., Sadiq R., Rodriguez M. I., Najjaran H., and Hoorfar M. (2014). "Reliability assessment for water supply systems under uncertainties." Journal of Water Resources Planning and Management, Vol. 140. No. 4. pp. 468-479.

Shamsi U. (1990). "Computerized evaluation of water supply reliability." IEEE Transaction on Reliability, Vol. 39, No. 1, pp. 35-41.

Tanyimboh T. T., Tietavainen M. T., and Saleh S. (2011). "Reliability assessment of water distribution systems with statistical entropy and other surrogate measures." Water Science and Technology: Water Supply, Vol. 11, No. 4, pp. 437-443.

Todini E. (2000). "Looped water distribution networks design using a resilience index based heuristic approach." Urban Water, Vol. 2, No. 2, pp. 115-122.

Torres J. M., Duenas-Osorio L., Li Q., and Yazdani A. (2017). "Exploring topological effects on water distribution system performance using graph theory and statistical models." Vol. 143, No. 1, 04016068-1 - 04016068-18.

Wagner J. M., Shamir U., and Marks, D. H. (1988a)."Water distribution reliability: analytical methods." Journal of Water Resources Planning and Management, Vol. 144, No. 3, pp. 253-275.

Wagner J. M., Shamir U., and Marks, D. H. (1988b). "Water distribution reliability: simulation methods." Journal of Water Resources Planning and Management Division, Vol. 114, No. 3, pp. 276-294.

Yang S.-L., Hsu N.-S. Louie P. W. F., and Yeh W. W.-G. (1996). "Water distribution network reliability: stochastic simulation." Journal of Infrastructure Systems, Vol. 2, No. 2, pp. 65-72.

Yazdani A. and Jeffrey P. (2012). "Applying network theory to quantify the redundancy and structural robustness of water distribution systems." Vol. 138, No. 2, pp. 153-161.

Appendix A: Calculation formulas for reliability analysis of WDN

If we consider a water distribution network from “source” to “tap” we can see this as a reliability block diagram (RBD) where the network is “functioning” if it is possible to transmit sufficient water from the source to tap through functioning pipes. By introducing pipe reliabilities, it is possible to calculate the reliability of the network. A dual approach to an RBD analysis is a fault tree analysis (FTA).
Several challenges are encountered when making such a model realistic:

- There could be more than one source, and more than one end users
- Transmitting sufficient water is not easy to describe by Boolean structures such as “series”, “parallels” and “*K*-out-of-*N*s”. The need for hydraulic models is evident, but not easy to combine with the RBD approach.
- Pumping stations and water treatment need to be included in the model
- Water tanks (buffers) need to be included in the model
- Control systems which are a main source of “cyber physical threats” need to be included in the model
- A WDN comprises typically thousands of components, and a complete RBD analysis is usually not practicable.

In the STOP-IT project is recommended to establish simplified models of the network, a so-called skeleton and then use of Boolean structures supported by more qualitative hydraulic understanding of the network. In the following basic aspects of the modelling is presented.

Reliability block diagram

A reliability block diagram (RBD) is a graphical representation of a system where there is one “source” and one “sink”. The RBD comprises arcs and nodes. The nodes can be seen as components in the system, and these are either in a functioning state or in a fault state. The arcs visualize the connection between the nodes. The system is in a functioning state if it is a connection between the source and the sink through functioning nodes. In very many cases we can construct an RBD by means of the following “structures”:

- Series structure – A series structure is functioning only if all the components in the structure are functioning
- Parallel structure – A parallel structure is functioning if one or more of the components in the structure are functioning
- *K*-out-of-*N* structure – A *K*-out-of-*N* structure is functioning if at least *K* out of the *N* components in the structure is functioning. The notation *K*oo*N* is often used (oo = Out Of).

To analyse an RBD it is possible to construct the so-called structure function. The structure function is a mathematical function linking the system state to component state. However, for large systems it is not possible to use the structure function for calculations due to the exploding number of terms when the structure function is “resolved”. Therefore, an alternative approach is to convert the reliability block diagram into a fault tree, and use standard fault tree algorithms to find the so-called “minimal cut sets”.

A cut set is a set of events whose (simultaneous) occurrence ensures that the system is in a fault state. A cut set is said to be minimal if the set cannot be reduced without losing its status as a cut set. An event here represents that a particular component is in a fault state. Algorithms exist for obtaining the minimal cut set. One such algorithm is implemented in the CAFTAN code (computerized fault tree analysis).

Basic quantitative measures of an RBD

For simple RBD analyses where there are no "buffers" or other dynamic components the typical reliability measures of interest are:

- $Q0$ = Probability that the system is in a fault state. $Q0$ represents the unavailability of the system
- $F0$ = The expected number of times the system enters a system fault state per time unit. $F0$ is often referred to as the system failure frequency
- $I^{B}(i)$ = Birnbaums measure of component reliability. $I^{B}(i) = \frac{dQ_0}{dq_i}$, where q_i is the component failure probability. Birnbaums measure is therefore a sensitivity measure.

To calculate these measures, we need component reliability parameters. These are:

- λi = 1/MTTFi = component failure rate, where MTTF = Mean Time To Failure
- μi = 1/MDTi = component repair rate, where MDT= Mean Down Time after a failure

A reasonable approximation for the component unavailability is given by:

$$q_i = \lambda_i \mathrm{MDT}_i \tag{15}$$

Now assume that the minimal cut sets are obtained, and denote these by $K_1, K_2, \ldots, K_J$. If the components are stochastically independent, the probability that cut set j is in a fault state is given by:

$$\check{Q}_j = \prod_{i \in K_j} q_i \tag{16}$$

The upper bound approximation is then used to find the system failure probability:

$$Q_0 \approx 1 - \prod_{j=1:J} \left(1 - \check{Q}_j\right) \tag{17}$$

To find Birnbaums measure for each component it can be shown that:

$$I^{B}(i) = Q_0(0_i) - Q_0(1_i) \tag{18}$$

where the notation 0_i means that component i is in a fault state, and 1_i means that component i is in a functioning state. An alternative interpretation of Birnbaums measure is that $I^{B}(i)$ is the probability that the system is in such a state that component i is critical. A component is critical means that the system is in such a state that the system is functioning if component i is functioning, and in a fault state if component i is in a fault state. Then it follows that the system failure frequency is found by:

$$F_0 = \sum_i I^{B}(i)(1 - q_i)\lambda_i \tag{19}$$

Modelling of water tanks (buffers)

Up to now we have assumed that there are no buffers in the system. A water tank is such a buffer where a failure upstream of the buffer is not critical if the component could be repaired before the water tank is empty. However, a buffer will not help if the failure is downstream of the buffer. For an appropriate modelling we need two RBDs in this situation. One RBD for the system where buffers are not included, and then one RBD downstream of each buffer. In this presentation we only consider the situation of one buffer. Let K be the cut sets for the entire RBD without the buffer, and let K_D be the cut sets for the RBD downstream of the buffer. Further let $\mathbf{0}_D$ represent the event that one or more of the cut sets in K_D is in a fault state, and $\mathbf{1}_D$ represent the situation that none of the cut sets are in a fault state.

From equation (19) we may calculate the system failure frequency, say $F_0(K)$ when buffers in the system are ignored. Further let $F_0(K, \mathbf{0}_D)$ be the system failure frequency given that one or more of the cut sets downstream of the buffer is in a fault state. $F_0(K, \mathbf{0}_D)$ thus represents the system failure frequency where the buffer will have no impact. An approximation for obtaining $F_0(K, \mathbf{0}_D)$ is given by

$$F_0(\mathrm{K}, \mathbf{0}_\mathrm{D}) = \Sigma_j F_0(\mathrm{K}, \mathbf{0}_j)\ \check{Q}_j \tag{20}$$

Where the sum is taken over the cut sets j in K_D, and the notation $\mathbf{0}_j$ is used to express that all components in cut set j are in a fault state.

Finally let:

$$F_0(\mathrm{K}, \mathbf{1}_\mathrm{D}) = F_0(\mathrm{K}) - F_0(\mathrm{K}, \mathbf{0}_\mathrm{D}) \tag{21}$$

be the failure frequency where the buffer will prevent loss of water at the sink in the network given that the buffer is not empty. Let $q_E = q_E(b)$ be the probability that repair time is higher than buffer capacity, say b.

Given a buffer capacity b the total system failure frequency of a system including one buffer is given by:

$$F_{0,\mathrm{B}}(b) = F_0(\mathrm{K}, \mathbf{0}_\mathrm{D}) + F_0(\mathrm{K}, \mathbf{1}_\mathrm{D})\ q_\mathrm{E}(b) \tag{22}$$

To obtain $q_E(b)$ we need to find the “repair time”. However, the various components upstream of the buffer have different repair times. Given that it is component i that is repaired first and thus “saves the day” the probability that the buffer is empty is given by $\mathrm{e}^{-b\mu_i}$. Birnbaums measure is then used to obtain a weighted probability for the empty buffer situation:

$$q_\mathrm{E}(b) = \frac{\sum_i I^\mathrm{B}(i) q_i \mathrm{e}^{-b\mu_i}}{\sum_i I^\mathrm{B}(i) q_i} \tag{23}$$

where the sum is taken over components not included in K_D.

By inserting Equation (23) in Equation (22) we obtain the system failure rate taking the buffer capacity into account. If more buffers exist, the method becomes much more

complex. An approximation would be to allocate all buffer capacity to one virtual buffer and let K_D be a representative set of minimal cuts downstream of this virtual buffer.

With respect to cyber physical threats it is natural to take these into account when considering what is the actual buffer size. Physically each water tank has a limited capacity in terms of cubic metres. However, operational procedures and efficiency in the operations of valves etc. may influence the number of hours these cubic metres could be available for a critical end user. If efficient procedures are in place we can prioritize important "customers", hence the number of hours is higher compared to if we are not able to prioritize. In the calculations we could therefore model *b* in equation (22) as a random variable, say *B*. Cyber physical threats and other aspects such as situational awareness, coordination efficiency etc could then be taken into account in when establishing the probability density function, $f_B(b)$. Equation (22) should then be integrated over the probability density function $f_B(b)$ to include also the cyber physical threats and other factors identified:

$$F_{0,\mathrm{B}} = \int_0^\infty F_{0,\mathrm{B}}(b) f_B(b)\mathrm{d}b = \int_0^\infty [F_0(\mathbf{0}_\mathrm{D}) + F_0(\mathbf{1}_{\boldsymbol{D}}) q_\mathrm{E}(b)] f_B(b)\mathrm{d}b \qquad (24)$$

Appendix B: Worked example – Probabilistic approach

In this appendix we demonstrate the methodology for calculation of probabilistic vulnerability contributing indexes. A simple skeleton of a given network is first presented (Figure 43). In the presented case study only one end user is considered. In a real study it would be natural to treat 5-10 end users, where each end user could be given a weight representing the importance of that user. As examples, higher weights are given to hospitals, schools, police station and so on. These calculations could be performed by a Network Reliability module. Possible inclusion of such a module is either implemented as part of InfraRisk-CP (in T4.2), or later development of the STOP-IT platform (e.g. in WP6).

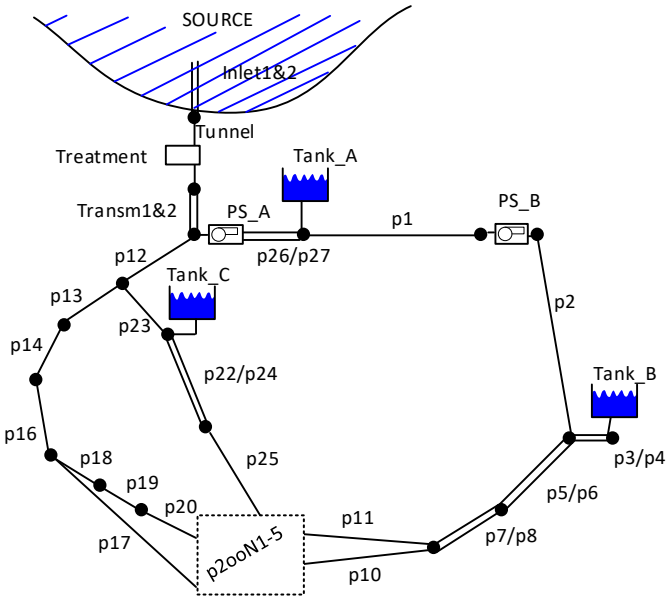

Figure 43 Skeleton of a water distribution network

Figure 43 shows the skeleton of the network used to demonstrate the full probabilistic approach. Only one end user is considered. This end user is considered as an important part of the critical infrastructure for the area considered. The dotted rectangle in Figure 43 represents the end user. There are 5 pipes from the surrounding network leading water to the end user. It is assumed that it if at least two out of five (2oo5) pipes are available, and each of these are connected to one of the “functioning” main lines, there will be sufficient water to the user considered.

As can be seen from Figure 43 it is three “main lines” from the source to the critical end user. What also is seen is that some of these main lines have partly redundancy, for example pipes p22 and p24.

Application of the NetworkReliability module will allow reliability assessments of water networks as shown in Figure 43. The specification of the module is presented in Table 5, and given as a standard text file. Each line represents a “block”. The assignment operator is “:=”, and the following structures are allowed:

- Series(<comma separated list of sub-structures>)
- Parallel(<comma separated list of sub-structures>)
- KooN(<comma separated list of sub-structures>)

For the KooN structure the letter K is replaced by the actual *K* for the structure as shown in Table 5. The sub-structures are either lower level structures, or “components” like pipes, valves or pumps. The first line is always the top level.

Table 5 Specification of the main network shown in Figure 43

```
TOP := Series(Parallel_1,2ooN_15,Series_16)
Parallel_1 := Parallel(Series_2,Series_5,Series_8)
Series_2 := Series(p12,p13,p14,p16,Parallel_3)
Parallel_3 := Parallel(Series_4,p17)
Series_4 := Series(p18,p19,p20)
Series_5 := Series(p12,Parallel_6,p25)
Parallel_6 := Parallel(p22,Series_7)
Series_7 := Series(p24,p23)
Series_8 := Series(Series_9,p2,p1)
Series_9 := Series(Parallel_10,Parallel_11,Parallel_12,Parallel_13,Series_14)
Parallel_10 := Parallel(p27,p26)
Parallel_11 := Parallel(p6,p5)
Parallel_12 := Parallel(p7,p8)
Parallel_13 := Parallel(p11,p10)
Series_14 := Series(PS_A,PS_B)
2ooN_15 := 2ooN(p2ooN1,p2ooN2,p2ooN3,p2ooN4,p2ooN5)
Series_16 := Series(Parallel_17,Parallel_18,Tunnel,Treatment)
Parallel_17 := Parallel(Inlet1,inlet2)
Parallel_18 := Parallel(Transm2,Transm1)
```

As can be seen from Figure 43 there are three “main routes” from the source to the end user. In the modelling we assume that it is only required one of these to be open in order to provide sufficient water. It could be cases where the hydraulic modelling shows that for example at least two out of these tree (2oo3) would be required. The second line in Table 5 would then read:

Parallel_1 := 2ooN(Series_2,Series_5,Series_8)

This is not implemented in this example, but could easily be done. This is a way how the hydraulic model could play together with the reliability model.

The model described in Table 5 does not include the water thanks. Failures “upstream” of the water tanks will have no immediate impact. However, if repair times are long, the water tanks will drain out. In the reliability modelling we need to specify a reliability block diagram

for the network “downstream” of the water thanks. Here we simplify and only consider one of the water thanks, but add the total capacity for all three water thanks for the one included in the model. The reliability block diagram “downstream” of Tank_B is shown in Table 6.

Table 6 Structure downstream of Tank_B

```
TOP := Series(2ooN_1,Parallel_2,Parallel_3,Parallel_4,Series_5)
2ooN_1 := 2ooN(p2ooN1,p2ooN2,p2ooN3,p2ooN4,p2ooN5)
Parallel_2 := Parallel(p6,p5)
Parallel_3 := Parallel(p7,p8)
Parallel_4 := Parallel(p11,p10)
Series_5 := Series(Parallel_4,Parallel_2,Parallel_3)
```

Reliability data

Table 7 shows reliability data for each pipe, for the water treatment and pumping stations. For pipes and the tunnel the failure rate is given per km per year, whereas for objects (tanks and treatment) the failure rate is given per year.

Table 7 Reliability data for components

Component name	Length (m)	Type	λ (/km/yr)	MDT (hr)
Inlet1	100	Rep. pipe	0.1	169
inlet2	100	Rep. Pipe	0.1	168
Tunnel	5000	Rep. Pipe	0.001	168
Treatment		rep. Object	0.0001	2
Transm1	1250	Rep. Pipe	0.01	24
Transm2	1250	Rep. Pipe	0.01	24
PS_A		rep. Object	0.00001	24
p26	3500	Rep. Pipe	0.01	24
p27	2900	Rep. Pipe	0.01	25
Tank_A		Tank		
p1	7200	Rep. Pipe	0.01	24
PS_B		rep. Object	0.00001	24
p2	3000	Rep. Pipe	0.1	8
p3	800	Rep. Pipe	0.1	24
p4	800	Rep. Pipe	0.1	24
Tank_B		Tank		
p5	3500	Rep. Pipe	0.1	24
p6	3500	Rep. Pipe	0.1	24
p7	1450	Rep. Pipe	0.1	24
p8	14570	Rep. Pipe	0.1	24
p9	300	Rep. Pipe	0.1	24
p10	4800	Rep. Pipe	0.1	24
p11	1200	Rep. Pipe	0.1	24
p12	2300	Rep. Pipe	0.1	24
p13	800	Rep. Pipe	0.1	24
p14	2000	Rep. Pipe	0.1	24
p15n	2000	Rep. Pipe	0.1	24
p16	1200	Rep. Pipe	0.1	24
p17	2800	Rep. Pipe	0.2	8
p18	900	Rep. Pipe	0.2	8
p19	550	Rep. Pipe	0.2	8
p20	300	Rep. Pipe	0.2	8
p21	1600	Rep. Pipe	0.1	24

Component name	Length (m)	Type	λ (/km/yr)	MDT (hr)
Tank_C		Tank		
p22	1100	Rep. Pipe	0.1	24
p23	400	Rep. Pipe	0.1	24
p24	800	Rep. Pipe	0.2	16
p25	1700	Rep. Pipe	0.2	16
p2ooN1	300	Rep. Pipe	0.2	8
p2ooN2	70	Rep. Pipe	0.2	8

Minimal cut sets

There are some eighty minimal cuts for this system when we ignore the water thanks. Table 8 shows the minimal cut sets up to order 3. Table 9 shows the minimal cut sets for the downstream structure, i.e., downstream of Tank_B.

Table 8 Minimal cut sets up to order 3

{Tunnel}
{Treatment}
{p12,PS_A}
{p12,PS_B}
{p12,p2}
{p12,p1}
{Inlet1,inlet2}
{Transm2,Transm1}
{p12,p27,p26}
{p12,p6,p5}
{p12,p7,p8}
{p12,p11,p10}
{p13,p25,PS_A}
{p13,p25,PS_B}
{p13,p25,p2}
{p13,p25,p1}
{p14,p25,PS_A}
{p14,p25,PS_B}
{p14,p25,p2}
{p14,p25,p1}
{p16,p25,PS_A}
{p16,p25,PS_B}
{p16,p25,p2}
{p16,p25,p1}

Table 9 Minimal cut sets for structure downstream Tank_B

{p6,p5}
{p7,p8}
{p11,p10}
{p2ooN1,p2ooN2,p2ooN3,p2ooN4}
{p2ooN1,p2ooN2,p2ooN3,p2ooN5}
{p2ooN1,p2ooN2,p2ooN4,p2ooN5}
{p2ooN1,p2ooN3,p2ooN4,p2ooN5}
{p2ooN2,p2ooN3,p2ooN4,p2ooN5}

Reliability measures without taking water tanks into account

From equation (17) we obtain $Q_0 \approx$ 9.6E-05 which corresponds to an average unavailability of water in 50 minutes per year. The frequency of such an event is found from equation (19) to be $F_0 \approx$ 5.5E-03 per year, or in average every 182 years.

Reliability measures taking water tanks into account

We simplify and consider all the tree tanks to be one thank with capacity equal to the sum of the capacities.

The frequency of the lack of water situation when one of the minimal cut sets downstream the water tank(s) is in a fault state is given by equation (20) and found to be $F_0(\mathcal{K}, \mathbf{0}_D)$ ≈2.3E-07. The frequency of the lack of water situation when none of the minimal cut sets downstream the critical tank is in a fault state is given by equation (21) and found to be $F_0(\mathcal{K}, \mathbf{1}_D) = F_0 - F_0(\mathcal{K}, \mathbf{0}_D) \approx$ 5.5E-03 which is the same as F_0. The reason for this is the fact that the likely cause of a system failure is events upstream the buffer(s). Note that $F_0(\mathcal{K}, \mathbf{1}_D)$ represents the frequency of events where we may utilize the buffer capacities in the water tanks.

From equation (23) we obtain the probability that the buffer is empty as a function of the buffer capacity. Table 10 shows the result.

Table 10 Probability of empty buffer as a function of the buffer capacity

b (hours)	$q_E(b)$
0	1
30	0.83
60	0.69
90	0.58
120	0.48
150	0.40
180	0.34
210	0.28
240	0.23

We observe that we need a buffer capacity of more than four to five days in order to have some significant reduction in $q_E(b)$. The reason for this is that system failure is likely to be caused by a failure or a combination of failures where the repair times are long.

The implication of this is that depending on the real buffer capacity in an actual failure situation, the frequency of loss of water situation to the end user can vary from one per two hundred years to one per thousand year.

The physical buffer capacity in the three water tanks in the case study is 25 hours. In the worst-case situation, we can imagine a complete failure in utilizing the buffer capacity. For example, a failure in opening the required valves. In this case $b = 0$, and $q_E(b=0) = 1$. In the case of opening the required valves but without making any restrictions on the water usage, $b = 25$, and $q_E(b=25) \approx 0.8$. In the best case where both the valves to the water tanks are operated successfully and where the buffer capacity is "reserved" for the critical end user one can imagine $b = 200$, and $q_E(b=200) \approx 0.25$.

Probabilistic vulnerability contribution index

Table 11 shows the probabilistic vulnerability contribution index calculated by equation (10).

Table 11 Probabilistic vulnerability contribution index

Component	I^{PVC}
Tunnel	0.9999996
Treatment	0.9999037
p12	4.74E-04
PS_A	6.30E-04
PS_B	6.30E-04
p2	6.30E-04
p1	6.30E-04
Inlet1	1.92E-04
inlet2	1.93E-04
Transm2	3.42E-05
Transm1	3.42E-05

As expected, the Tunnel and the Treatment contribute most to the vulnerability since a failure of these will give the system failure.

STOP-IT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 740610.