A risk-based multi-level stress test methodology: Application to six critical non-nuclear infrastructures in Europe

Sotirios A Argyroudis[1], Stavroula Fotopoulou[1], Stella Karafagka[1], Kyriazis Pitilakis[1], Jacopo Selva[2], Ernesto Salzano[3], Anna Basco[4], Helen Crowley[5], Daniela Rodrigues[5], José P. Matos[6], Anton J. Schleiss[6], Wim Courage[7], Johan Reinders[7], Yin Cheng[8], Sinan Akkar[9], Eren Uckan[9], Mustafa Erdik[9], Domenico Giardini[10], Arnaud Mignan[10]

(1) Department of Civil Engineering, Aristotle University, Thessaloniki, Greece
(2) National Institute of Geophysics and Volcanology (INGV), Italy
(3) Department of Civil, Chemical, Environmental and Materials Engineering, University of Bologna, Italy
(4) AMRA, Italy
(5) EUCENTRE, Pavia, Italy
(6) Civil Engineering Institute, EPFL, Switzerland
(7) TNO, Netherlands
(8) School of Civil Engineering, Southwest Jiaotong University, China
(9) Bogazici University, Kandilli Observatory and Earthquake Research Institute, Turkey
(10) Department of Earth Sciences, Institute of Geophysics, ETH Zürich, Switzerland

Abstract

Recent natural disasters that seriously affected Critical Infrastructure (CI) with significant socio-economic losses and impact, revealed the need for the development of reliable methodologies for vulnerability and risk assessment. In this paper, a risk-based multi-level stress test method that has been recently proposed, aimed at enhancing procedures for evaluation of the risk of critical non-nuclear infrastructure systems against natural hazards, is specified and applied to six key representative CIs in Europe, exposed to variant hazards. The following CIs are considered: an oil refinery and petrochemical plant in Milazzo, Italy, a conceptual alpine earthfill dam in Switzerland, the Baku-Tiblisi-Ceyhan pipeline in Turkey, part of the Gasunie national gas storage and distribution network in the Netherlands, the port infrastructure of Thessaloniki, Greece, and an industrial district in the region of Tuscany, Italy. The six case studies are presented following the workflow of the stress test framework comprised of four phases: Pre-Assessment phase, Assessment phase, Decision phase and Report phase. First the goals, the method, the time frame, and the appropriate stress test level to apply are defined. Then, the stress test is performed at component and system levels and the outcomes are checked and compared to risk acceptance criteria. A stress test grade is assigned and the global outcome is determined by employing a grading system. Finally, critical components and events and risk mitigation strategies are formulated and reported to stakeholders and authorities.

1. Introduction

Critical infrastructure (CI) provides essential services to society and represents the backbone of economy, security and health. Recent examples from key CIs have revealed that natural hazards can cause significant economic and social damage, severely affect the provided services and lead to disasters, whilst cascading failures of CIs can cause multi-infrastructure collapse and widespread consequences even in developed countries (Pescaroli and Alexander 2016). Representative paradigms from Japan can be highlighted, i.e. the Tohoku earthquake, tsunami and Fukushima nuclear release in 2011 (Krausmann and Cruz 2013) and the Hyogo-Ken Nanbu (Kobe) earthquake in 1995 that caused extended damage to port and other critical infrastructure with long term consequences (Chang 2000). Among past events in Europe, devastating flash floods in the spring of 2010 caused extended dam failures in Poland (Reuters 2010), while major damage to industrial facilities was reported after the 2009 L'Aquila and 2012 Emilia earthquakes in Italy (Grimaz 2014).

The increase and intensity of such natural disasters over the last two decades (EMDAT 2019), which is correlated to the ageing infrastructure and in some cases its inadequate design as well as to urban growth, climate change and environmental degradation, has increased the interest of policy makers, practitioners and researchers toward the understanding of infrastructure vulnerability and risk (Giannopoulos et al. 2012; Theocharidou and Giannopoulos 2015; Opdyke et al. 2017). There is a remaining need to address gaps in existing knowledge in order to better understand and assess the vulnerability and risk of CIs and improve their resilience against natural hazards. In this respect, advanced and standardized tools for hazard and risk assessment of CIs are required, such as the stress test tools , that include both low-probability high-consequences (LP-HC) events and so-called extreme events, as well as the systematic application of these new tools to whole classes of critical infrastructure. In particualr, stress testing is the process of assessing the ability of a CI to maintain a certain level of functionality under unfavorable conditions. Stress tests consider LP-HC events, which are not always accounted for in the risk assessment procedures and tools, commonly adopted by public authorities or industrial stakeholders. They have been initially developed for the financial and nuclear sectors, e.g. to check whether the safety and design standards applied to nuclear power plants are sufficient to cover unexpected extreme events (Kutkov and Tkachenko, 2017). In Europe, after the accident at the Fukushima nuclear power plant in Japan, a comprehensive safety and risk assessment in the form of stress tests was performed on all nuclear plants (ENSREG 2012). Stress tests contribute to the improvement of prevention and preparedness of critical infrastructure, providing the roadmap for strengthening measures of the high-risk components and the improvement of emergency response planning. Hence, stress tests contribute toward the resilience enhancement of the CIs, i.e. how they can adapt to and recover from shocks.

In this context, an engineering risk-based multi-level stress test framework has recently been developed (Esposito et al. 2016; 2019), aimed at enhancing the current procedures for evaluating the risk of critical non-nuclear infrastructure against natural hazards, considering single or multi-hazards, probabilistic or scenario based approaches, systemic analysis, interactions between components, cascading effects and an advanced grading system.

The main objective of this paper is to demonstrate the applicability of this methodology, which is summarized in Section 2, through six case studies of CIs in Europe exposed to different hazards: (1) an oil refinery and petrochemical plant in Milazzo, Italy, by taking into account the impact of earthquakes and tsunami (Section 3); (2) a conceptual alpine earthfill dam in Switzerland under multi-hazard effects (Section 4); (3) the Baku-Tiblisi-Ceyhan pipeline in Turkey, focusing on seismic threats at pipe-fault crossing locations (Section 5); (4) part of the Gasunie national gas storage and distribution network in the Netherlands, exposed to earthquake and liquefaction effects (Section 6); (5) the port infrastructures of Thessaloniki in Greece, subjected to seismic, tsunami and liquefaction hazards (Section 7); and (6) an industrial district in the region of Tuscany, Italy, exposed to seismic hazard (Section 8). These applications are representative of the following CI types: (i) single-site (case studies 1, 2 and 5), (ii) geographically extended (case studies 3 and 4), (iii) distributed multi-site (case study 6). The key elements and output of the six applications are summarised in Section 9.

2. Methodology

2.1 ST workflow and phases

A harmonized framework for stress testing critical non-nuclear infrastructure systems has been recently proposed (Esposito et al. 2019) aiming to quantify the safety and risk of individual components as well as of whole CI system with respect to natural events and to compare the behavior of the CI to acceptable values. The multi-level framework combines probabilistic and quantitative methods to characterise both extreme and common scenarios and consequences, including potential multi-hazards and systemic amplification effects (e.g., Mignan et al. 2014; 2016a; 2016b). To manage subjectivity and uncertainty, the proposed framework includes a multiple-expert integration (Selva et al. 2015), in which data, models and methods adopted for the risk assessment and the associated uncertainty quantification are clearly documented and managed by different experts. Different roles and responsibilities are assigned to different actors, namely the *project manager (PM)*, *technical integrator (TI)*, *evaluation team (ET)*, *pool of experts (PoE)* and *internal reviewers (IR)*. Their roles and interactions are illustrated in Figure 1, along with the workflow of the framework.

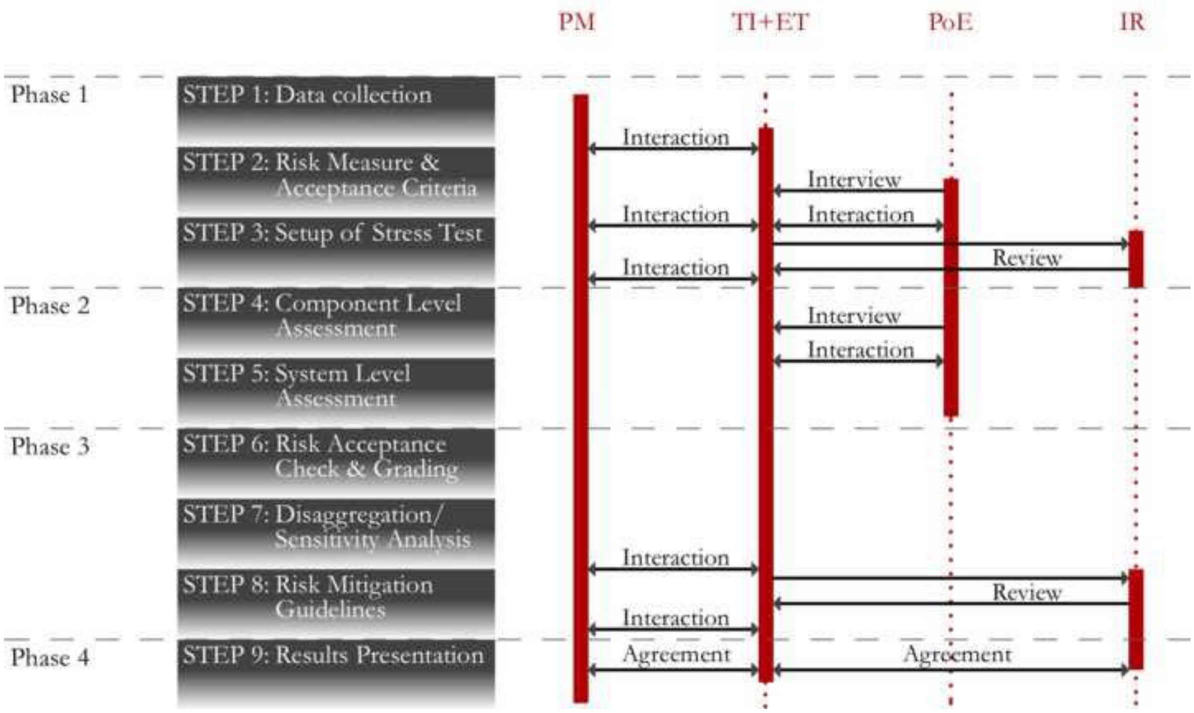

Figure 1. Workflow of the stress testing framework (Esposito et al. 2019)

The proposed framework is implemented in four main phases:

1. Pre-Assessment Phase (steps 1 to 3): the necessary data on the CI and hazards are collected. The risk measures and acceptance criteria, the time frame, the most appropriate stress test level(s) and level of detail of the analysis are defined depending on potential regulatory and stakeholder requirements as well as available resources and data (Esposito et al. 2019). .

2. Assessment Phase (steps 4 to 5): the stress test at component and system levels is performed following state-of-the-art methods for the hazard, vulnerability and risk analysis.

3. Decision Phase (steps 6 to 8): the results of the Assessment phase are compared to the acceptance criteria that have been defined in the Pre-Assessment Phase. This comparison results in a grade that informs about the degree of the risk posed by the infrastructure, and, if the risk is unjustifiable or intolerable, how much the safety of the CI should be improved until the next periodical verification. Critical events that most likely cause the exceedance of a loss value of interest are identified through a disaggregation and/or sensitivity analysis. Risk mitigation strategies and guidelines are formulated.

4. Report Phase (step 9): the experts present the stress test results to authorities and regulators of the CI. The presentation includes the outcome of stress test in terms of the grade, the critical trigger events, the guidelines for risk mitigation, and level of detail adopted in the stress test.

2.2 Stress test levels

Three *Stress Test Levels (ST-Ls)* are proposed. *Level 1 (ST-L1)*: single-hazard component check (hazard-based, design-based, risk-based); *Level 2 (ST-L2)*: single-hazard system-wide risk assessment; *Level 3 (ST-L3)*: multi-hazard system-wide risk assessment. Each level is characterized by a different scope (component or system) and by a different complexity of the risk analysis. Within these three levels, potentially different implementations are possible. The quantification of epistemic uncertainty may not be performed (sub-level a). If performed, it may be based either on the evaluations of a single expert (sub-level b) or of multiple experts (sub-level c). In Levels ST-L2 (sub-levels a, b and c) and ST-L3 (sub-levels a, b and c) probabilistic risk analysis (PRA) of the entire CI (system) is performed. Complementary scenario-based analysis (sub-level d) may be performed for specific conditions, events or hazards that cannot be included into the PRA due to methodological gaps. It is noted that ST-L1 should be the routinely check for each CI and it might be deterministically (hazard or design-based) or/and probabilistically (risk-based) defined.

2.3 Penalty and grading system

The stress test can result to three outcomes: *Pass*, *Partly Pass*, and *Fail* (Figure 2). In particular, the CI passes the stress test if it attains grade AA or A. Grade AA corresponds to negligible risk and is expected to be the risk objective for new CIs. Grade A corresponds to risk being as low as reasonably practicable (ALARP) (Helm, 1996; Jonkman et al. 2003), and is expected to be the risk objective for existing CIs. The CI partly passes the stress test if it gains grade B, which corresponds to the existence of possibly unjustifiable risk. The CI fails the stress test when grade C is assigned, corresponding to the existence of intolerable risk. The boundaries between grades, i.e. the risk acceptance criteria, are defined by the project manager of the stress test based on the requirements of the regulators and societally acceptable risk norms. The form of the boundaries can be expressed using point estimates, e.g. expected number of fatalities per year, or continuous functions, e.g. F-N curves, representing cumulative frequency of the risk measure per given period of time. These boundaries may differ between countries and industries. Further details can be found in Esposito et al. (2019).

The application of the stress test concepts to six CIs in Europe is summarized in the following sections. It is noted that these applications include different ST levels based on the available data and resources in the framework of a research study and they should not be considered as formal or complete stress tests. For a more elaborated description of the case studies reference is made to Pitilakis et al. (2016).

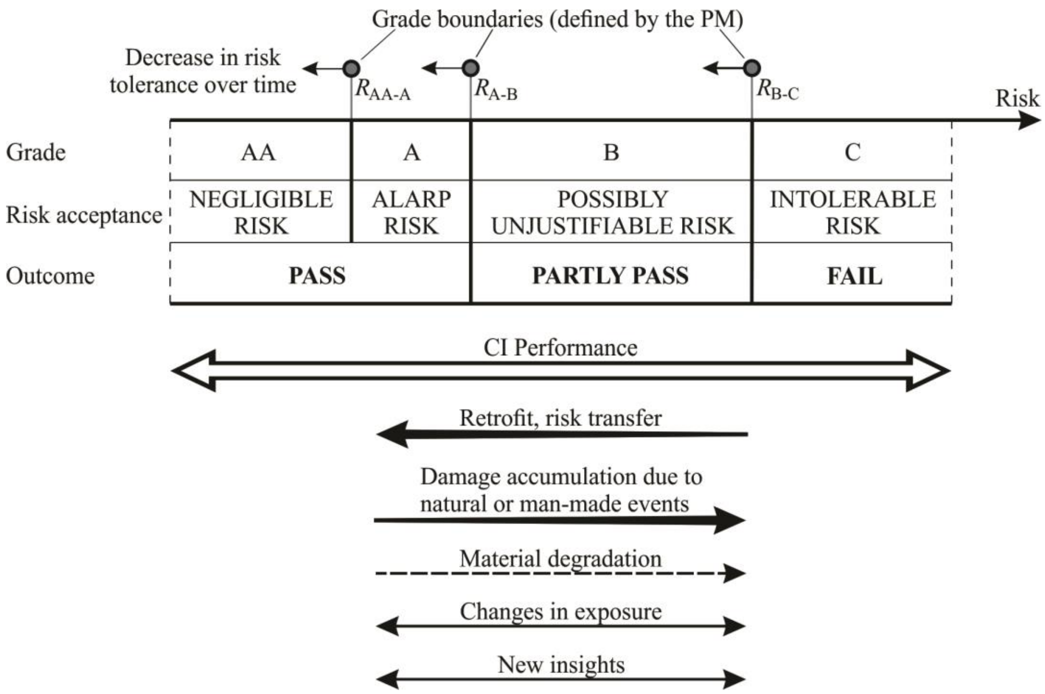

Figure 2. Grading system for the outcome of stress test (Esposito et al. 2019)

3. Application to oil refinery and petrochemical plant in Italy

3.1 Pre-Assessment phase

Natural events may dramatically interact with industrial equipment with different intensity and hazards. Structural failures may be indeed induced by seismic waves or tsunami waves, flooding and other combined scenarios. Hence, industrial accidents may derive, such as fires, explosions, toxic dispersion or environmental disasters. These scenarios are nowadays defined as Natech (Krausmann et al. 2011; Salzano et al. 2013; Renni et al. 2010; Krausmann et al. 2016). Natech risks should be included in the industrial risk assessment (Quantitative Risk Assessment, QRA), which is normally performed in early-design phase, during the licensing and land use planning procedures, and other civil protection applications. Quite typically, results are given in terms of locational risk and societal risks. The first is defined as the frequency per year that a hypothetical person will be lethally affected by the consequences of possible accidents during an activity involving hazardous materials, e.g. a chemical plant or transport activities. This risk indicator is a function of the distance between the exposed person and the activity, regardless of whether people are actually living in the area, or at the specified location. Societal risk is defined as the cumulative frequency that a minimum casualties due to possible accidents during an activity with hazardous materials.

The refinery of Milazzo (Raffineria di Milazzo) is located in the north part of the island of Sicily, in Italy. It is an industrial complex, which transforms crude oil into a series of oil products currently available on the market (LPG, gasoline, jet fuel, diesel and fuel oil) and comprises a number of auxiliary services. The refinery has many storage tanks containing a large variety of hydrocarbons, such as LPG, gasoline, gasoil, crude oil and atmospheric and vacuum residues. The capacities of the tanks vary from 100 m^3 (fuel oil, gasoil, gasoline, kerosene) to 160 000 m^3 (crude oil). All tanks are located in catch basins (bunds) with concrete surfaces. The LPG is stored in pressurised spheres, while all other substances are stored in single containment tanks.

In the following, a Natech QRA for this installation, based on public information regarding the industrial process, has been performed.

3.2 Assessment phase

Probabilistic Hazard Analysis was performed for both tsunami and earthquake (ST-L2). For the tsunamis, we have focused only on tsunami of seismic origin, which is the dominant component in most areas of the world. The impact of natural hazards on the accident or release scenarios and frequencies is given in Table 1. These frequencies have been calculated by taking into account the methodology described in several previous works (Salzano et al. 2015; Basco and Salzano 2016), where equipment vulnerability with respect to the intensity of the natural events has been assessed by taking into account the construction characteristics of equipment and, more important, the new limit states based on the release of content.

Table 1. Scenarios and frequencies for stationary vessels due to natural hazards

Scenario	**Frequency (-/yr)**		
	Atmospheric vessels	*Pressurized vessels*	*Pipelines*
Earthquake			
Instantaneous release of the complete inventory	$3.70\cdot10^{-3}$	$1.16\cdot10^{-9}$	-
Continuous release of the complete inventory in 10 min at a constant rate of release	$3.70\cdot10^{-3}$	$1.16\cdot10^{-9}$	-
Continuous release from a hole with an effective diameter of 10 mm	$7.33\cdot10^{-2}$	0	-
Full bore rupture	-	-	$5.56\cdot10^{-2}$
Tsunami	*Atmospheric vessels*	*Pressurized vessels*	*Pipelines*
Instantaneous release of the complete inventory	$1.85\cdot10^{-5}$- $3.47\cdot10^{-4}$	0	-
Continuous release of the complete inventory in 10 min at a constant rate of release	$1.85\cdot10^{-5}$- $3.47\cdot10^{-4}$	0	-
Continuous release from a hole with an effective diameter of 10 mm	0	0	-
Full bore rupture	-	-	0
Earthquake + Tsunami			
Instantaneous release of the complete inventory	$3.7\cdot10^{-3}$- $4.05\cdot10^{-3}$	$1.16\cdot10^{-9}$	-
Continuous release of the complete inventory in 10 min at a constant rate of release	$3.7\cdot10^{-3}$- $4.05\cdot10^{-3}$	$1.16\cdot10^{-9}$	-
Continuous release from a hole with an effective diameter of 10 mm	$7.33\cdot10^{-2}$	0	-
Full bore rupture	-	-	$5.56\cdot10^{-2}$

3.3 Decision phase

Results obtained for the Natech QRA for the refinery of Milazzo, in terms of locational risk and societal risk is presented in Figure 3. The isorisk curves take into account the combination of all natural and industrial hazards. The right part of the same figure allows the evaluation of the contribution of either industrial or natural events, separately, and their relative weights. The fact that the curves for Industrial and Earthquake and Industrial, Earthquake and Tsunami coincide, means that tsunami adds a negligible contribution to risk. This methodology can be then used for the decision phase, in terms of licensing, land use planning, civil protection plan (emergency plan), early design and industrial and environmental authorizations.

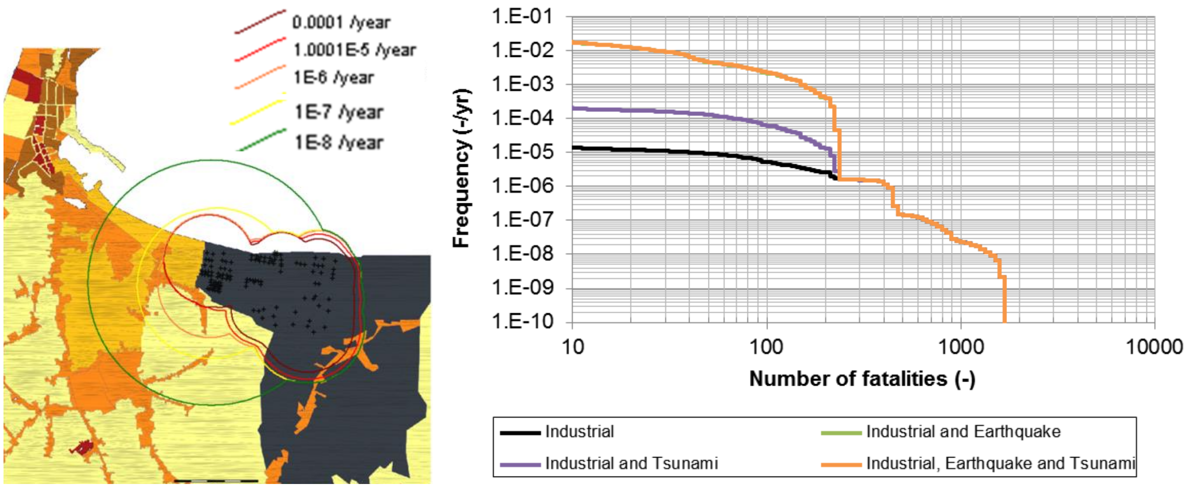

Figure 3. Locational risk (left) and societal risk (right) – Hazard combinations (industrial, seismic, tsunami)

3.4 Report phase

Naturally induced hazards can play an important role in the total risk associated with the presence of installations with dangerous goods. For the specific site analyzed, our stress testing results indicate that the predicted tsunamis can only damage a limited number of the atmospheric storage vessels along the shoreline. Hence the increase on the total risk is limited. Nonetheless, the overloading of emergency response should be considered, at least for the tanks along the coastline.

Of more importance is the effect of an earthquake, which significantly increases the failure frequency of atmospheric storage tanks. Therefore, reinforcing the emergency response for multiple fire scenarios would be beneficial, together with structural improvement of the tanks. Neither an earthquake nor a tsunami significantly increases the failure frequency of, and hence risk imposed by, pressurized vessels (like LPG spheres). As for the considered site, the risk is largely dominated by the LPG tanks when failing due to industrial-related causes, whereas the impact of the natural hazards is limited. All in all though, naturally induced hazards should be considered when determining the overall risk and the risks associated with natural disaster. The communication among key actors (emergency responders, public authorities, industrial stakeholders) is deemed mandatory, according to the Seveso directive (EC 2012). In particular, the communication should be improved by re-thinking of the information to the population related to the industrial risks, which is still mandatory by the Seveso directive (EC 2012), but completely lacking for the natural-technological interaction.

4. Application to a large dam in Switzerland

Dams operate by storing water (and its potential energy) in their reservoirs and releasing it when convenient. Often, that potential energy can produce massive damages if not controlled adequately. In the event of a failure or breach a large amount of water travels downstream in the form of a dam-break wave, affecting downstream areas more seriously than natural floods. To fully understand the risks associated with large dams one should therefore take into account the dam, the reservoir, the downstream areas, and the multiple elements and interactions that characterize what can be called the dam-reservoir system.

Dam safety is most commonly tested using deterministic frameworks where the system's response is simulated and analyzed in detail for a given number of limit cases (Zenz and Goldgruber 2013; Gunn et al. 2016). Although proven very successful, the deterministic approach's focus on limit cases leaves countless possibly disastrous combinations of events unchecked. Furthermore, the probability of occurrence of the limit cases under test is not necessarily known and, therefore, even if the risk associated with the infrastructure can qualitatively be inferred to be small when the test succeeds, it remains unknown in quantitative terms. This justifies further investments on probabilistic alternatives.

The present application aimed to develop a flexible probabilistic framework that separates the risk assessment for large dams in two sequential steps: the analysis of the dam-reservoir system, that provides information about the frequency of failures and the conditions under which water is released; and the downstream areas, where the progression of each dam-break wave is accounted for and damages are evaluated. The modeled system includes a dynamical representation of the dam-reservoir system that relies on the Generic Multi-Risk (GenMR) framework (Mignan et al. 2014; 2016) and accounts for multiple hazards, multiple elements, and a large number of non-linear influences and feedbacks between them. Also included in the system is a module capable of efficiently predicting inundation parameters for each simulated dam failure case to roughly 30 km downstream, where a sizable urban agglomeration is assumed to exist (Figure 4).

A large conceptual alpine earthfill dam was taken as a case study. The infrastructure is approximately 100 m high, with a reservoir capable of holding over 100 000 000 m^3 of water. It is equipped with a spillway in order to cope with excessive water levels, a bottom outlet that allows for the control of the volume of water stored, and a hydropower system through which the main purpose of the dam is fulfilled, i.e. producing energy.

4.1 Pre-Assessment phase

The considered hazards included earthquakes, floods, internal erosion, and electromechanical malfunctions in key systems. Regarding elements, the dam and foundation, the bottom outlet, the hydropower system, the spillway, and the reservoir were modeled. The most relevant interactions considered were the damages induced on elements, the damage states that lead to changes in operations, the probability of internal erosion events and how it is affected by reservoir levels and damage through overtopping. Focusing on the downstream area, the response of each building to the inundation was also modeled resorting to fragility curves. Hazards were defined according to statistical distributions and, for each case, epistemic uncertainty on the parameters of those distributions was assumed. The response of each element to relevant hazards was also defined probabilistically, according to fragility functions.

The objectives of the stress test were two-fold. First, regarding the frequency of failures. Second, the expected damages downstream as a direct consequence of such failures. Risk measures were, accordingly, the expected return period of dam failure events and the expected built volume downstream of the dam that would be destroyed as a result.

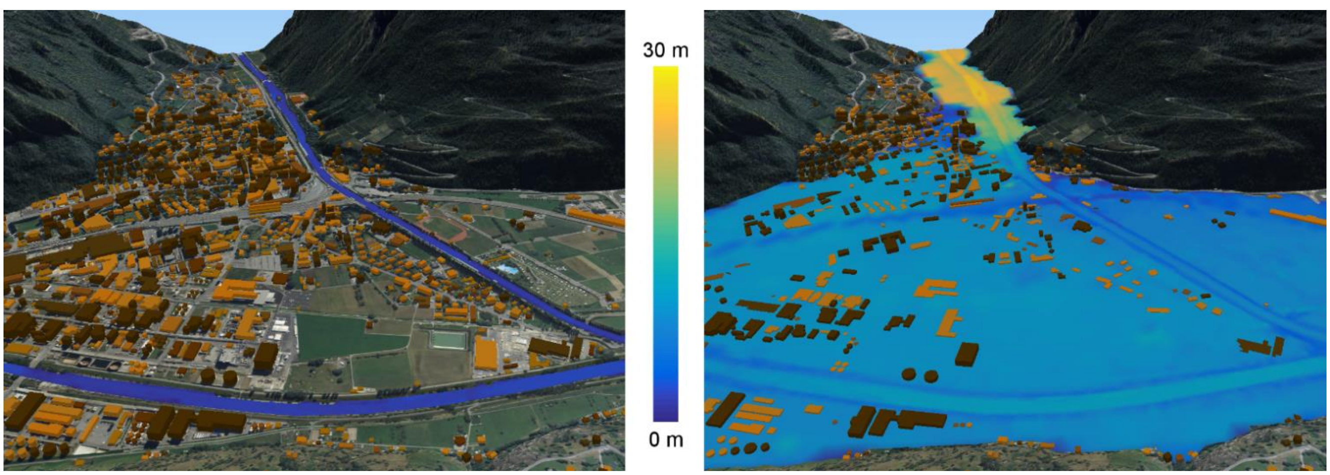

Figure 4. Illustration of the impact of a specific dam-break wave on an urban area downstream

4.2 Assessment phase

The backbone of the assessment phase is the component level assessment (ST-L1). Here, as the case study is conceptual, it was admitted that all the elements of the system comply to and slightly exceed regulatory requirements. The ST-L2 system level assessment for a single hazard (earthquake) was undertaken in both deterministic and probabilistic models. In the ST-L3 system level assessment, for multiple hazards, the full integration of the dam-reservoir and downstream analysis realms was made. Through the simulation of 20 000 000 years of dam operation a number of failures with different characteristics was sampled. It should be clear that the simulations are not extended 20 000 000 years into the future; rather, it is different possibilities for “next” year that are simulated. The number of simulations should be large enough to sample events of the order of magnitude of the return period intended for the infrastructure. For example, in 20 000 000 simulations it can be expected to find, on average, 2 000 events with a return period of 10 000 years or above. For each one of these, inundations parameters were estimated throughout the downstream valley. Computations were performed by a machine learning meta-model trained based on detailed 2D hydraulic simulations of representative dam-break events. Integrating the information from all the simulations, including aleatoric and epistemic uncertainty, it was possible to gain remarkable insight on the system. Figure 5a, for example, illustrates the return period of individual buildings collapsing or being washed away as a result of dam failures.

4.3 Decision phase

With a return period of 25 000 years for failures, the conceptual dam was shown to meet the first of three risk objectives by having, on average, less than one failure per 10 000 years. In what concerns damages in the downstream area, the goal was to limit expected damages to the loss of one household per 100 years. In concrete terms, this was assumed to be equivalent to an average built volume loss of 7.5 m^3 per year due to dam failures. After integrating expected losses in the downstream area, however, a substantially higher value of 200 m^3 of built volume lost per year was estimated. As a consequence, the second risk objective was not met. Despite this the expected losses were deemed acceptable as the undesirably high value is more a product of the number of households exposed to the dam-break wave than on the frequency of dam failures, being therefore and to some extent beyond the influence of changes that the dam may undergo. The third objective is bound to the analysis of a F-N curve (Figure 5b), in this case prepared to show the cumulative frequency of built volume collapsing or being washed away as a consequence of a dam break upstream. The threshold AA-A corresponds to a risk of 7.5 m^3/yr, the A-B threshold, corresponding to the third risk objective, to 75 m^3/yr, and finally B-C to 750 m^3/yr (roughly equivalent to a household per year).

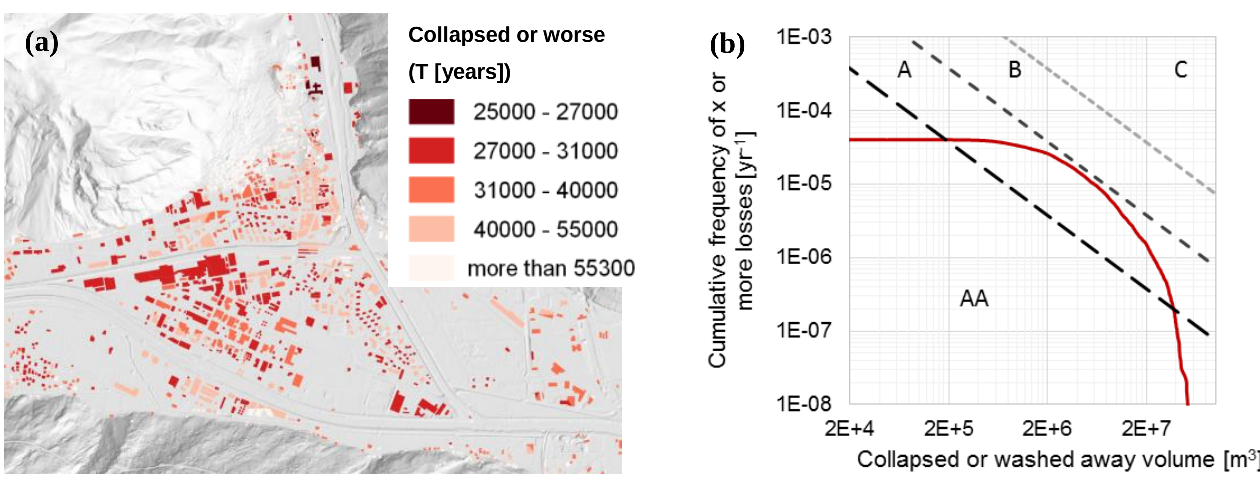

Figure 5. Illustration of results from the study on large dams. a) return periods of individual buildings being collapsed or washed away following a dam failure upstream. b) F-N curve based on collapsed or washed away built volume following a dam failure upstream (adapted from Pitilakis et al. 2016)

4.4 Report phase

The flexibility of the GenMR framework (e.g., Mignan et al. 2014), particularly when combined with machine learning methods that allow extraordinary gains in computational performance, makes this inclusive and formally correct estimate of risk attainable. Obviously, this is a highly desirable feature when performing a stress test.

From the three objectives established in this stress test, one, concerning the dam-reservoir system and the probability of failures taking place, was met with a failure return period of 25 000 years, safely above the 10 000 years mark. The second, focusing on the expected losses downstream, was not. Quantitatively the chosen risk metric was more than 25 times over the objective of 7.5 m^3/year of built infrastructure collapsed or washed away. The third objective, defined on the basis on an F-N curve, classified the risk as ALARP.

In this conceptual case, earthquakes appear to be responsible for the most part of the expected losses. They have a direct impact on the dam but can also lead to the catastrophic elevation of water in the reservoir through damages to the outlet structures. Investing in a more resilient bottom outlet would virtually prevent all overtopping events, being perhaps the most direct and cost-effective way to reduce risk. Regarding the downstream losses, a possible use of the analysis results and maps is to reinforce, provide with shelters or relocate the buildings which are assessed as high risk. However, the risk downstream is not only dependent on the probability of failure of the dam-reservoir system, but also on the amount of people and infrastructure exposed at risk. Once the CI is considered safe, it may be more cost-effective to invest on the protection of the downstream area than on the dam itself (for example, providing better warning systems and escape routes). For the conceptual CI that was studied, some potential failures could be averted by drawing down the reservoir. Therefore, beyond the notions of fragility that were explored, the resilience of the dam-reservoir system beyond the design requirements is very much defined by the capacity to perform a successful and timely drawdown operation. Cascade effects become important when the possibility of drawing down the reservoir is lost, and a substantial inflow arrives.

Concluding, to evaluate the risk associated with the failure of a large dam it is important to bring together a number of experts in different fields, relevant to the structure itself, the foundation and hydrology. For an accurate quantification of impacts downstream it is essential to collect data and knowledge on infrastructure, property and populations, including the evacuation in case of a failure. Compared to other CIs, the verification of the safety of large dams is quite developed as these infrastructures are built not to fail. An improvement of existing approaches is the consideration of uncertainty, in the quantification of hazards and fragility.

5. Application to major hydrocarbon pipeline in Turkey

The hydrocarbon pipelines usually extend over very long distances by crossing borders, hard geographic conditions and geo-hazard areas. As of seismic effects, they are prone to permanent fault displacement (PFD) hazard because fault crossings (upon their rupture) may cause large deformations on the hydrocarbon pipelines and impose a major risk for their structural integrity. When such pipelines are exposed to PFD, typical damage is in the form of local buckling due to axial compression and/or bending (in normal burial depths) and global (beam) buckling (in shallow burial depths) or in submarine pipelines. The rupture of the pipe could be due to severe compressive buckling of the pipe wall or tensile fracture.

This section implements the stress testing methodology for seismic risk assessment of pipeline failure due to PFD. The Baku-Tiblisi-Ceyhan (BTC) pipeline is used as the case study that crosses several strike-slip fault segments in the Eastern Anatolia Fault.

5.1 Pre-Assessment phase

The diameter and thickness of the pipes at the five main fault segments are 42 inches (1.0688 m) and 20.62 mm, respectively. The pipeline trench is trapezoidal-shaped and packed with loose-to-medium granular cohesionless backfill with minimum soil cover. The pipeline crosses five fault segments along Eastern and North Anatolia Fault zones with fault-pipe crossing angles varying between 30° and 90°. All other compiled data information about the mechanical features of BTC pipeline as well as fault properties important in PFD computations are given in Pitilakis et al. (2016).

The risk measure in this case study is the pipeline rupture or loss of pressure integrity due to fault offsets. Table 2 lists the probability ranges of different risk tolerances according to the grading system of the stress test methodology.

Table 2. The risk tolerance levels and the probabilities defined for the stress test grade

Grade	**AA**	**A**	**B**	**C**
Risk tolerance	Negligible	ALARP	Possibly Unjustifiable risk	Intolerable
Probability range in 2475-year return periods	0%-2%	2%-10%	10%-50%	50%-100%
CI performance	pass		partly pass	fail

Eidinger and Avila (1999) propose four performance classes (life safety, key operational, other operational and disruption) to represent severity of pipe failure at pipe-fault crossings. These four performance goals are set to four pipeline failure probabilities (Table 2) that are defined as 1% (life safety), 2% (key operational), 10% (other operational) and 50% (disruption) against PFD underground-motions represented by 2475-year return period uniform hazard spectral ordinates.

The stress tests comprise of three steps at component level (ST-L1), performing hazard-based (moderate accuracy), design-based (advanced accuracy) and risk-based assessment (high accuracy).

5.2 Assessment phase

5.2.1 Hazard-based assessment: the 2475-year PFDs (recommended by ALA 2001; 2005) at five pipe-fault crossings are computed from the Monte-Carlo based probabilistic PFD hazard (Chen and Akkar, 2017; third row in Table 3). They are compared with the prescribed ALA hazard requirements (second row in Table 3). The comparisons indicate that of the five pipe-fault crossings, the computed 2475-year PFD hazard at #2, #3 and #4 pipe-fault crossings are larger than the ALA requirements (last row in Table 3). The potential impact of mega-ruptures in the region (Mignan et al., 2015) was not included in this analysis, since the mechanism of mega ruptures is complicated and models to estimate the fault displacement are yet to be proposed.

Table 3. Hazard-based assessment – comparison of 2475-year return period PFD hazard with ALA requirements

	Pipe-fault crossings				
	#1	#2	#3	#4	#5
2475-year ALA2001 (design)	1.31m	1.18m	1.61m	3.84m	0.63m
2475-year ALA2001 (assessment)	0.73m	2.25m	3.91m	4.49m	0.44m
Compliance (design ≥ assessment)	yes	no	no	no	yes

5.2.2 Design-based assessment: The tensile pipe strain under the 2475-year PFD is compared with the allowable tensile pipeline strain provided in ALA (2001). The allowable tensile pipe strain is designated as 3% in these design provisions. The comparisons are done for all five pipe-fault crossings and the tensile strains at these pipe-fault crossings comply with the code requirements (Table 4).

Table 4. Calculated tensile strains at the designated fault offsets

Pipe-fault crossing	**Crossing angle**	**2475-year fault offset (m)**	**Tensile strain**	**Compliance (≤ 3%)**
#1	60°	0.73	0.33%	Yes
#2	70°	2.25	0.85%	Yes
#3	30°	3.91	2.18%	Yes
#4	45°	4.49	2.00%	Yes
#5	90°	0.44	0.18%	Yes

5.2.3 Risk-based assessment: The annual exceedance rate of pipeline failure is compared with the suggested allowable pipeline failure rates in the literature. The probabilistic pipeline failure is achieved by integrating the probabilistic fault displacement hazard, mechanical response of pipe due to fault displacement and empirical pipe fragility function (Cheng and Akkar 2017). The aggregated effects of tensile and compressive strains developed along the pipe are considered in the seismic pipe failure risk. The annual failure probability (P_f) for pipelines at fault crossings is computed for different pipe-fault crossing angles (α) by considering the uncertainty in α. The inaccuracy in fault-pipe crossing angle is modeled by a truncated normal probability with alternative standard deviations of 2.5° and 5°.

The acceptable annual pipe failure rate of $4.0 \cdot 10^{-5}$ (Honegger and Wijewickreme 2013) is compared with the pipe failure rates at five designated pipe-fault crossings (Table 5). The comparisons indicate that pipe-fault crossings #3 and #4 are critical as their computed failure rates are larger than the allowable annual failure rate. The listed annual failure rates are also used to compute the aggregated failure risk along the whole BTC pipeline to complete the probabilistic risk assessment. Two marginal probabilities are computed: (a) perfect correlation between pipe failures at the five pipe-fault crossings (P_{fc}) and (b) independent pipe failures at the five pipe-fault crossings (P_{fi}). The aggregated marginal failure probabilities are very high and they range between 40% and 50% (Table 6) that fall into grade B: possibly unjustifiable risk according to Table 2.

Table 5. Comparisons of annual pipe failure exceedance rates with the allowable pipe failure rate

Pipe-fault crossings	**σ (standard deviation) : Uncertainity to pipe-fault crossing angles (α)**			**Compliance (≤4.0 10^{-5})**
	0	**2.5**	**5**	
#1	$3.142 \cdot 10^{-6}$	$3.183 \cdot 10^{-6}$	$3.304 \cdot 10^{-6}$	Yes
#2	$1.833 \cdot 10^{-6}$	$2.256 \cdot 10^{-6}$	$3.293 \cdot 10^{-6}$	Yes
#3	**$1.967 \cdot 10^{-4}$**	**$1.964 \cdot 10^{-4}$**	**$1.955 \cdot 10^{-4}$**	**No**
#4	**$5.987 \cdot 10^{-5}$**	**$5.981 \cdot 10^{-5}$**	**$5.962 \cdot 10^{-5}$**	**No**
#5	$1.973 \cdot 10^{-5}$	-	-	Yes

Table 6. Aggregated failure probabilities of BTC pipeline under 2475-year PFD hazard before and after the risk mitigation strategies

	P_{fc} (perfectly correlated case)	**P_{fi} (statistically independent case)**
Before retrofit	38.56 %	51.0 %
After retrofit	0.775 %	2.206 %

5.3 Decision and Report phase

The probabilistic pipe failure risk assessment yields higher probabilities of pipe failure at #3 and #4 pipe-fault crossings. Therefore, these pipeline segments are identified as critical components and it is decided to be upgraded.

The effective retrofitting of the pipeline segments at the critical crossings is to change the pipe-fault intersection angle. When the intersection angles of these three pipe-fault intersection angles are changed to ~80°, the resulting aggregated risk probability is reduced to negligible levels

Table 6). The disaggregation and sensitivity analysis of BTC pipe failure assessment bring forward the higher PFD hazard and small pipe-fault crossing angles (resulting in higher tensile strain) as the main sources of large failure probabilities at the pipe-fault crossings #3 and #4.

6. Application to gas storage and distribution network in Netherlands

This section summarizes the application of the stress test methodology to part of the main gas distribution network of Gasunie Gas Transport Services (Gasunie-GTS). The Groningen field is a large natural gas field located in the northern Netherlands, contributing to approximately half of the natural gas production in the country. The gas distribution relies on a major gas pipeline infrastructure, with a total length of over 12 000 km of installed pipes in. Located in an area of very low tectonic seismicity, gas extraction in the region has led to an increase in seismicity since the early 1990s. A sub-network (Figure 6) is studied located in the induced earthquake prone area, directly above the main gas field covering an area of approximately 3360 km^2.

6.1 Pre-Assessment phase

Numerous seismic hazard studies dedicated to the Groningen area have been performed over the past several years and are still ongoing. In the current stress test one of the earlier model versions was adopted: the so-called Z1 model from Dost et al. (2013) for the seismic zonation (four zones), the Akkar et al. (2014a; 2014b) modified ground motion model (Bommer 2013) and the classical Gutenberg-Richter (GR) relation (Gutenberg and Richter 1956). A maximum magnitude (for the stress test only) value of 6.0 was applied and the annual event rate for events with $M \geq 1.5$ is set to 30 events per year (Dost et al. 2013).

Serviceability Ratio (SR) and Connectivity Loss (CL) are used as risk measures (Esposito et al. 2015). The Serviceability Ratio is directly related to the number of demand nodes in the network, which remain accessible from at least one source node following an earthquake. Connectivity Loss measures the average reduction in the ability of demand nodes to receive flow from source nodes due to an earthquake event.

An as low as reasonable practicable (ALARP) grade of the risk measures is targeted for the existing gas transport network to pass the stress test (Jonkman et al. 2003). In the Netherlands a standard for quantified risk assessment (QRA) exists, issued by the national "Committee for the Prevention of Disasters" (CPR 18E 1999). In the current application of the stress test methodology to the Gasunie-GTS case, no full QRA was performed for the 1000 km sub-network. However, values for the annual failure rates originally prescribed in (CPR 18E 1999) and adjusted values nowadays used for the Gasunie network are selected to define grade boundaries (Table 7):

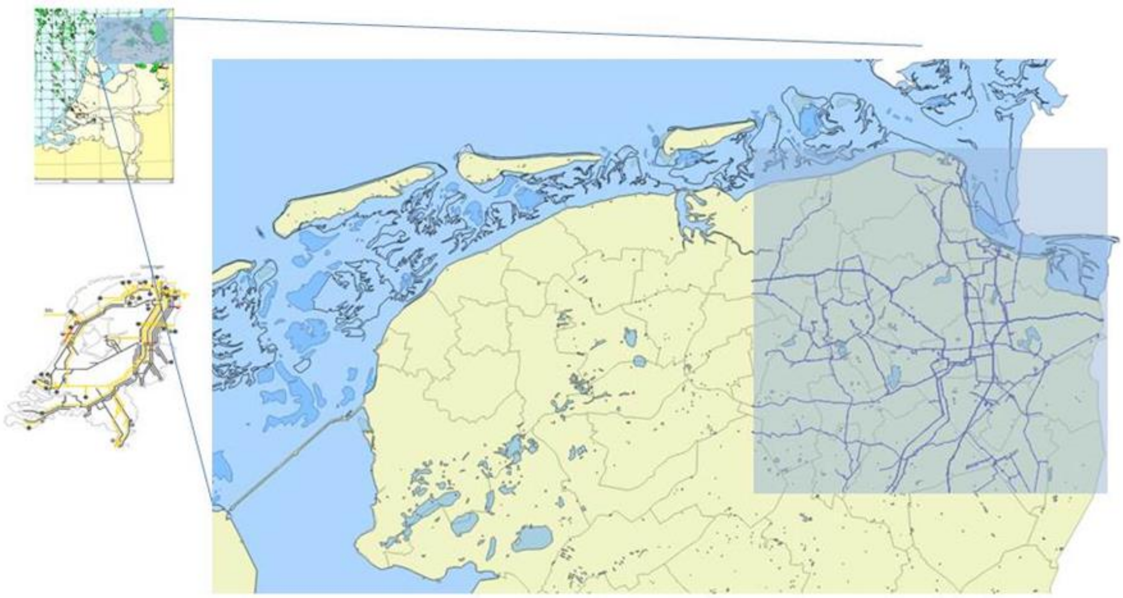

Figure 6. Selected sub system of the gas distribution network (right) located above main natural gas field (top left)

Table 7. Definition of grading boundaries for the gas network

Boundary	**Pipe [$yr^{-1}km^{-1}$]**	**Station [yr^{-1}]**
AA-A	$8 \cdot 10^{-6}$	$8 \cdot 10^{-6}$
A-B	$6 \cdot 10^{-5}$	$6 \cdot 10^{-5}$
B-C	$1.4 \cdot 10^{-4}$	$1.4 \cdot 10^{-4}$

For illustrative purposes only, indicative grading boundaries are attributed to the values of the performance parameter connectivity loss (CL). The boundaries used are taken from (Esposito et al. 2016). No actual calibrations for these bounds with respect to economic loss or fatalities exist yet for the sub-network at hand and the grading is indicative and provisional.

The stress test has been performed up to ST-L2 considering earthquake as single hazard and conducting a full Probabilistic Risk Analysis using Monte Carlo simulations for the network analysis. ST-L1 considers individual components for which also a risk based approach is applied. As the methods in this case for ST-L1 and ST-L2 are both Monte Carlo based, ST-L1 makes use of the ST-L2 results.

Accuracy levels targeted are classified as advanced: the stress test is risk based for the network as well as for the individual components with site specific hazard analyses, structure specific fragility functions and using outcomes of dedicated studies by, among others, the NAM, KNMI, TNO and Deltares as well as by an international community of experts (WINN_TA-NAM 2016).

6.2 Assessment phase

The ST-L2 for the evaluation of the seismic network performance is consisted of five major steps:

- Seismic hazard assessment of the region considering gas depletion as source of the seismic activities.
- Evaluation of ground motion hazard in terms of PGA, PGV and permanent ground displacement (due to liquefaction).
- Seismic demand evaluation at each station and pipe section to obtain the failure using fragility curves.
- Vulnerability analysis through the use of a connectivity algorithm to assess the network performance.
- Probabilistic risk assessment in terms of mean network functionality and annual exceedance curves.

The likelihood of liquefaction given the soil conditions in the Groningen area, the Netherlands, was first assessed (Miraglia et al. 2015) in which a model based on the Idriss-Boulanger model (Idriss and Boulanger 2008) is used. Two soil profiles based on CPT tests were analyzed by describing the soil properties as stochastic parameters and sampling the liquefaction response of the layers with earthquake events. Sampling results were then summarized as fragility curves as a function of PGA values for the two soil profiles. Soil liquefaction can cause permanent soil displacements as well as floating or sinking of pipe segments due to gravity. Structural reliability calculations are performed for distinct pipe configurations and probabilities of failure are calculated conditional on liquefied soil. For transient load effects, again structural reliability calculations are performed based on Newmark's formulae of seismic strain for buried pipelines (Newmark and Rosenblueth 1971). As a result transient load fragility curves were obtained as function of PGV values. For the stations, a generic fragility curve from the HAZUS methodology (NIBS 2004) was adopted.

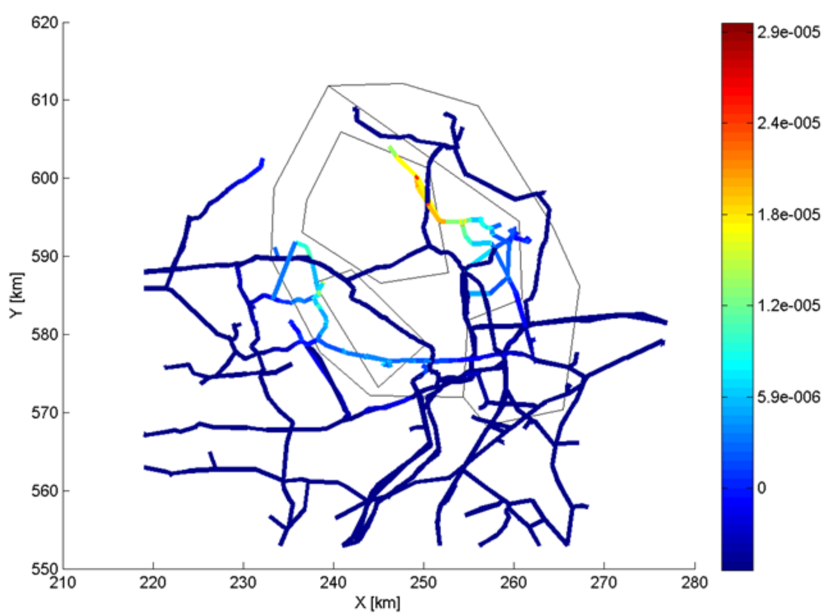

Figure 7. ST-L1: annual failure frequencies (per km) for the pipe sections (black lines on background indicate the earthquake zones)

Seismicity, network and network properties are modeled with the OOFIMS (Franchin et al. 2011) tool and Monte Carlo simulations are performed. The results show a good performance with respect to CL (Figure 8): the annual probability of having a connectively loss of e.g. 50% or more is $3.6 \cdot 10^{-5}$. For the serviceability ratio very high exceedance frequencies for all values of the serviceability ratio are found, with only a drop near SR reaching one. Hence the results show a high robustness of the network, indicating a vast redundancy in possible paths between demand and source nodes. Sampled results (failure, no failure) per component (pipes/stations) from the ST-L2 Monte Carlo analysis of the network are used to calculate ST-L1 annual failure probabilities per component (e.g. Figure 7). Pipes as well as stations showed satisfactory performance in terms of reliability.

6.3 Decision phase

With respect to the grading on component level the following results are obtained:

- Pipe sections: Most pipe sections obtain grade AA, some obtain grade A. The pipe sections pass the stress test.
- Stations: Most stations are classified with grade AA or A. Some, near or within the seismic zone, obtain grade B. The stations partly pass the stress test.

With respect to the network performance Figure 8 presents the values for the connectivity loss relative to the indicative grading boundaries. The network performance is shown to comply with grade AA and passes the stress test.

These findings are obtained despite a number of conservative assumptions made with respect to fragilities. Also the seismic demand was modeled in a conservative way with e.g. a maximum magnitude of 6 and an annual rate of occurrence for $M_L > 1.5$ equal to 30. Reducing these assumptions to a maximum magnitude of 5 and or an annual rate equal to 23 leads to all stations complying with grade AA or A.

With respect to components, both types (pipe sections and stations) are found to contribute evenly to the network performance. From these:

- Specific pipe sections can to some extend be identified as being a weakest link in the network. These sections should be checked on their current actual state assessing the need for upgrading.
- For the stations a rather strong assumption is made with respect to the fragility curve adopted. These should be quantified in more detail and depending to findings retrofitting of stations might be necessary.

In the current analysis soil liquefaction is the dominant failure mechanism. As much uncertainty still exists in the liquefaction fragilities for the Groningen area, further studies into these fragilities and their geographical distribution is recommended.

6.4 Report phase

The stress test is performed as being initiated by the asset owner, the Gasunie Transportservices and as such reported to the asset owner. No formal presentation of the outcome of the stress test to (other) CI authorities and/or regulators is foreseen. Reporting, in terms of the grade, the critical events, the guidelines for risk mitigation, and the accuracy of the methods adopted in the stress test is accomplished in Pitilakis et al. (2016).

Most pipe sections and stations conform to grade AA or A, except for few stations that reach grade B. Turning points are magnitude M_L=5 or annual rate $N_{M>1.6}$=23 at which all components comply to grade AA or A and pass the stress test, see Table 8.

At the time of performing the stress test, no governing earthquake specific design requirements existed in the Netherlands. The CI's safety and resilience will be improved by reassessing the need for retrofitting of a confined number of pipe sections identified. The stress test also revealed the need for site-specific fragility functions for the Gasunie-GTS stations as well as further research into the liquefaction mechanisms for the Groningen site conditions.

Table 8. Stress test results for Gasunie-GTS sub-network

Item	**M_{max}**	**$N_{M>1.6}$**	**Grading**	**Result**
Pipe sections	6	30	AA, A	Pass
Stations	6	30	AA, A, B	Partly pass
	5	30	AA, A	Pass
	6	23	AA, A	Pass
Network *CL*	6	30	AA, A	Pass

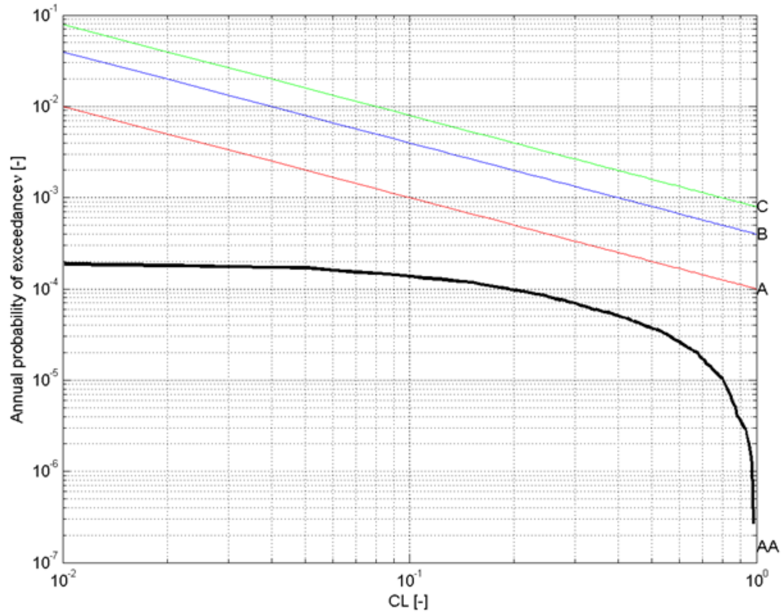

Figure 8. Exceedance frequencies for connectivity loss relative to (indicative) grading boundaries

7. Application to port infrastructures of Thessaloniki in Greece

This section outlines the application of the stress test methodology to the port of Thessaloniki, one of the most important ports in Southeast Europe and the largest transit-trade port in Greece. Ground shaking, liquefaction and tsunami hazards have been considered in the case study. Readers are referred to Pitilakis et al. (2017) for more details on this stress test.

7.1 Pre-Assessment phase

A GIS database for the examined port facilities, i.e. waterfront structures, cargo handling equipment, buildings (offices, sheds, warehouses etc.) and the electric power supply system has been developed. The Port subsoil conditions are characterized by soft alluvial deposits, sometimes susceptible to liquefaction. All necessary information to perform site specific ground response analyses were obtained by a comprehensive set of in-situ geotechnical tests (e.g. drillings, sampling, SPT and CPT tests), detailed laboratory tests and measurements, as well as geophysical surveys (cross-hole, down-hole, array microtremor measurements) at the port broader area. A topobathymetric model was also produced for the tsunami simulations, based on nautical and topographic maps and satellite images (Cotton et al. 2016; Selva et al. 2016).

The vulnerability of the port infrastructures to the given target hazards is assessed using site and case specific or generic fragility functions. New seismic fragility curves have been developed for typical quay walls and gantry cranes subjected to ground shaking based on dynamic numerical analyses. Analytical tsunami fragility curves as a function of inundation depth have been developed for representative typologies of RC buildings, warehouses and gantry cranes (Karafagka et al. 2016; Salzano et al. 2015). For simplicity reasons, the waterfront structures were considered as non-vulnerable to tsunami forces. The electric power lines were also assumed as non-vulnerable for the three hazards.

The stress test includes a component level risk based assessment of the key components (ST-L1) and a probabilistic risk analysis at the system level (ST-L2). A complementary scenario-based system wide risk assessment is also conducted associated to two earthquake return periods. Specific risk measures and acceptance criteria have been defined related to the functionality of the port at system level and the structural losses at component level. Since two terminals (container, bulk cargo) were assumed herein, the system performance is measured through the total number of containers handled (loaded and unloaded) per day (TCoH), in Twenty-foot Equivalent Units (TEU), and the total cargo handled (loaded and unloaded) per day (TCaH), in tones. Risk measures related to structural and economic losses of the buildings were also set for the tsunami case and the scenario-based assessment. Since no regulatory boundaries exist for the moment for port facilities, continuous (Figure 9) and scalar boundaries (Table 9) were defined based on general judgment criteria for the probabilistic and scenario-based system-wide risk assessment respectively.

7.2 Assessment phase

In the component level assessment, a risk-based assessment of each component is carried out for earthquake and tsunami hazards to check whether the component passes or fails the minimum requirements for its performance. The hazard function at the location of the component and the fragility function of the component are convolved in risk integral in order to obtain the probability of exceedance of a designated limit state in a period of time. To check whether or not the component is safe against collapse, the target probability was compared with the corresponding probability of exceeding the ultimate damage state. A reference target probability of collapse equal to $1 \cdot 10^{-5}$ has been pre-defined based on the existing practice.

In the system level assessment, a probabilistic risk analysis (PRA) is conducted separately for earthquake and tsunami hazards considering specific interdependencies between network and components. The objective is to evaluate the mean annual frequency (MAF) of events with the corresponding loss in the performance of the port operations. The analysis was based on an object-oriented paradigm where the system is described through a set of classes, characterized in terms of attributes and methods, interacting with each other (Franchin et al 2011; Kakderi et al. 2014). A Monte Carlo simulation is carried out sampling events and corresponding damages for the given hazards. The seismic hazard is based on the 2013 European Seismic Hazard Model - ESHM13 (Woessner et at 2015, Giardini et al 2013) and the modeling procedure described in Weatherill et al. (2014). The tsunami hazard analysis was performed considering tsunamis generated by co-seismic sea floor displacements due to earthquakes (e.g., Grezio et al. 2017; Davies et al. 2017; Lorito et al. 2015) and obtaining 253 representative scenarios based on inundation simulation of the Thessaloniki area (Selva et al. 2016). The performance indicators (PIs) of the port system for both the container and cargo terminal were evaluated for each simulation based on the damages and corresponding functionality states of each component and considering the interdependencies between components. The final computed PIs are normalized to the value referring to normal (non-seismic) conditions (P_{max}) assuming that all cranes are working at their full capacity 24 hours per day while the performance loss is defined as 1-PI/PI_{max}.

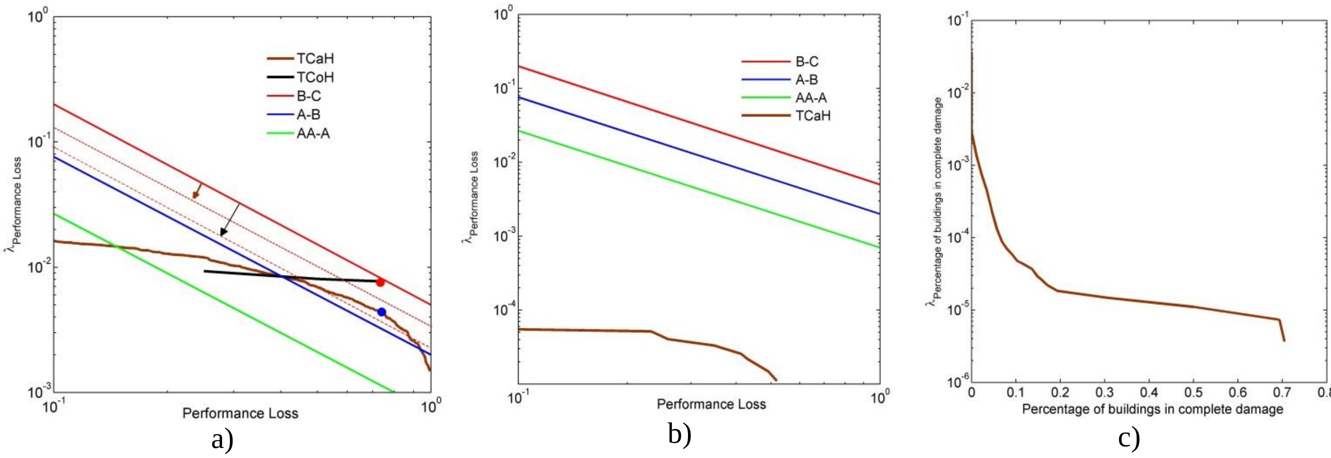

Figure 9. MAF of exceedance curves for the port system PIs (TCoH, TCaH) in terms of normalized performance loss (1-PI/PImax) for the seismic (a) and tsunami (b) hazard case and the buildings in collapse state for the tsunami case (c).

For the seismic hazard case, Figure 9a shows the MAF of exceedance curves ("performance curve") for the normalized performance loss in terms of TCoH and TCaH. The green, blue and red continuous lines correspond to the boundaries between risk grades AA (negligible), A (ALARP), B (possibly unjustifiable risk), and C (intolerable). For performance loss values below 40% TCaH yields higher values of exceedance frequency, while for performance loss over 40% TCoH yields higher values of exceedance frequency. For the tsunami hazard case, an example for one of the alternative models (i.e. the epistemic uncertainty is not considered here) is presented in Figure 9b. The container terminal is not expected to experience any loss (TCoH), while the loss in the cargo terminal (TCaH) is very low. This is due to the non-vulnerable condition of waterfront structures, the high damage thresholds for the cranes (i.e. inundation values that are not expected in the study area) and the distance of the electric power substations from the shoreline. The annual probabilities for buildings collapses are also low (Figure 9c). As an example 10% of the total buildings in the Port (~9 structures) will be completely damaged under tsunami forces with annual probability equal to $5 \cdot 10^{-5}$.

The scenario-based risk analysis (SBRA) is performed complementary to the classical PRA approach described previously, to quantify the potential impact of the local site response at the port area and to reduce the corresponding uncertainties. This type of effects may be of major importance in port areas, and by adopting specific scenarios is possible to model the site response more accurately than in standard PRA. Two different seismic scenarios were defined in collaboration with a pool of experts: the standard seismic design scenario and an extreme scenario corresponding to return periods of T_m=475 years and T_m=4975 years respectively. For the 475 years scenario a set of 15 accelerograms was selected to fit the target spectrum defined based on the disaggregation of the probabilistic seismic hazard analysis results (SRM-LIFE 2007; Papaioannou 2004) and the median plus 0.5 standard deviation of Akkar and Bommer (2010) spectrum (Pitilakis et al. 2019). For the 4975 years scenario the selection of ground motion requires special attention considering that it might be an extreme event that has not been recorded yet. Thus, two different approaches were considered: 4975 years scenario I and II (Pitilakis et al. 2019). In particular, 10 synthetic accelerograms were computed to fit the target spectrum (median plus one standard deviation Akkar and Bommer (2010) spectrum) (4975 years scenario I) and broadband ground motions were generated (Smerzini et al. 2016) using 3D physics-based "source-to-site" numerical simulations (4975 years scenario II). 1D equivalent-linear (EQL) and nonlinear (NL) site response analyses including also the potential for liquefaction are carried out. It is observed that the EQL approach is associated with higher number of non-functional components for all considered seismic scenarios whereas for the NL approach non-functional components are present only for the 4975 years scenario I (Table 9). This is due among other factors to the significantly higher PGA values calculated using the EQL approximation, which lead to higher damage probabilities and consequently higher performance loss. Thus, even though the vulnerability using the NL approach is assessed considering both ground shaking and liquefaction hazards, the estimated combined exceedance probabilities and the corresponding performance loss are still lower compared to the ones predicted by the EQL approach (Pitilakis et al. 2019).. As also evidenced by the estimated functionality state of each component, the port system is non-functional both in terms of TCaH and TCoH for the 4975 years scenario I. A 100% and 67% performance loss is estimated for the TCoH and TCaH respectively when considering the EQL approach for the 475 and 4975 years II scenarios, while the port is fully functional when considering the NL approach both in terms of TCaH and TCoH for the latter scenarios.

Table 9. Estimated normalized performance loss of the port system for TCaH and TCoH and comparison with risk acceptance criteria for the scenario-based assessment

Scenario	**Analysis type**	**Performance loss (1-PI/PImax)**		**Risk acceptance criteria**			**Stress test outcome**	
		TCaH	**TCoH**	**AA-A**	**A-B**	**B-C**	**TCaH**	**TCoH**
475 years	EQL	0.67	1.00	0.10	0.30	0.50	Fail	Fail
	NL	0.00	0.00				Pass	Pass
4975 years I	EQL	1.00	1.00	0.30	0.50	0.70	Fail	Fail
	NL	1.00	1.00				Fail	Fail
4975 years II	EQL	0.67	1.00				Partly pass	Fail
	NL	0.00	0.00				Pass	Pass

7.3 Decision phase

With reference to seismic hazard for both bulk cargo and container terminals, the port obtains grade B, meaning that the risk is possibly unjustifiable and the CI partly passes this evaluation. The basis for redefinition of risk objectives in the next stress test evaluation is the characteristic point of risk, which is defined as the point associated with the greatest risk above the ALARP region. The CI receives grade AA (negligible risk), and as expected in this example application, passes the stress test for the tsunami hazard. Based on the proposed grading system, for the case which the port obtains grade B and partly passes the stress test, the B-C boundary in the next stress test is reduced (i.e. B-C: 53% performance loss) while the other boundaries remain unchanged (Figure 9a). The scenario-based assessment showed that the CI may pass, partly pass or fail for the specific evaluation of the stress test (receiving grades AA, B and C respectively) depending on the selected seismic scenario, the analysis approach and the considered risk metric. This level of analysis is complementary to the PRA and shows that a detailed modelling of local site effects is of major importance for the outcome of the stress test. It is also worth noting that the risk objectives and the time between successive stress tests should be defined by the CI authority and regulator. Since regulatory requirements do not yet exist for the port infrastructures, the boundaries need to rely on judgments.

7.4 Report phase

For the selected target probabilities of collapse, all port components are deemed as unsafe towards seismic hazards at the component level assessment (ST-L1), while only few cranes are characterized as safe against exceedance of the collapse limit state for the tsunami hazard. These results cannot be judged unconditional to the fact that subjective boundaries relying on expert judgments are used, since regulatory requirements for port infrastructures do not yet exist.

For ST-L2, and for the seismic case, several electric power distribution substations present high failure risk and contribute to the performance loss of the port due to loss of power supply to the cranes. It is recommended to investigate further the response of the substations under seismic shaking and consider potential upgrade or/and alternative power sources. The systemic tsunami risk connected to direct damages from waves results not significant. This is primarily connected to the physical position of the port (with relatively low tsunami hazard) and the low fragility of components to tsunami waves. However, the potential effects of debris collisions have not been accounted for. Therefore, a careful check of preparedness against tsunami should be suggested, ranging from the connection to efficient tsunami warning systems as well as the definition of actions to secure ships and port equipment.

For the scenario-based assessment the estimated losses are significantly dependent on the analysis approach. In particular, the EQL approach is associated with higher losses even for the design scenario (475 years), while for the NL approach the losses to the cranes, waterfronts and electric power substations are expected solely for the 4975 scenario I. Therefore, the impact of local site effects on the stress test outcome is very important and should be considered in the PRA through advanced seismic site response analysis.

The risk mitigation and resilience planning for the port infrasrtructure include preventive, e.g. early warning systems for earthquakes and tsunami, retrofitting of high-risk facilities, improvement of foundation soil, updating of contingency plans and training exercises, and reactive measures, e.g. efficient emergency and restoration plans, back-up capabilities for such as use of mobile cranes or diesel generators for power supply. In this context, Galbusera et al. (2018) performed a resilience analysis for the port infrastructures of Thessaloniki, considering the fragility and importance of each component, the interdependencies, the recovery priorities and the buffering capabilities for given seismic scenarios. Stress testing can further benefit the resilience planning, while the effective communication between the key actors (e.g. port authority, operators, experts) is essential.8. Application to industrial district in Italy

The performance and consequences assessment of an industrial building stock in Northern Italy, and more specifically in the region of Tuscany, is presented in this case study. Only seismic hazard has been considered, as it is the predominant hazard to which the industrial building stock in Tuscany is exposed. The limited budget for a stress test of an industrial district (given that these facilities do not serve the same critical functions as other infrastructure considered herein) has conditioned the level of detail and complexity of the stress test. Nevertheless, the simplicity of the case study allows the full probabilistic risk assessment and disaggregation methodology to be fully demonstrated. Readers are referred to Rodrigues et al. (2017) for more details on this stress test.

8.1 Pre-Assessment phase

The exposure data of the industrial infrastructure in Tuscany have been provided by the Seismic Section of the Tuscany Region. The details related to 300 industrial buildings in the province of Arezzo were used for the case study, which included the geographical location (represented by a pair of coordinates), year of construction, floor area, structural type, non-structural elements, and other data useful for identifying the value of contents, type of business, and geographical extent of the facility's customer base (Figure 10).

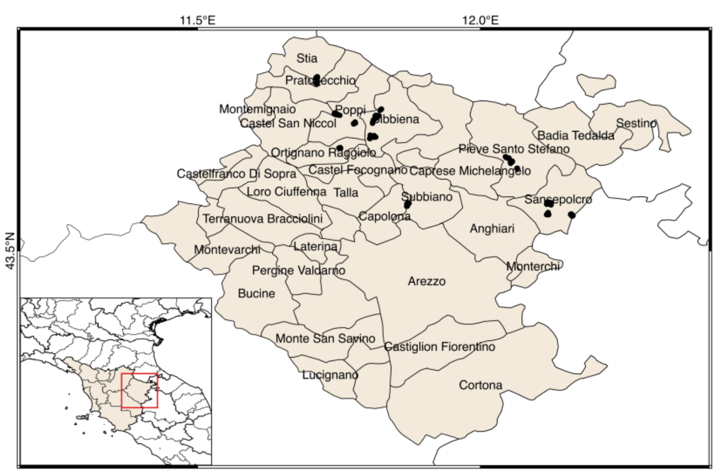

Figure 10. Location of the 300 industrial facilities in the province of Arezzo. Due to the close proximity of some of the buildings, each point that is shown on this map could represent up to 20 buildings

The majority of reinforced concrete precast industrial buildings in the Tuscany region can be categorized into three classes as a function of the design code level (pre-code or low-code, depending on whether the buildings were constructed before or after 1996), type of structure (type 1 buildings with long saddle roof beams, and type 2 with shorter rectangular beams and larger distances between the portals) and type of cladding (vertical precast panels (V), horizontal panels (H) and concrete masonry infills (M)). Once the building subclass has been assigned to each building in the exposure model, it is then necessary to add the value of the structural components, non-structural components, contents and business interruption (in terms of revenue per day). Typical construction costs for an industrial facility are used to assign the value of the structural and non-structural elements, estimated using the mean market prices of industrial/typical warehouses as a function of their location within so-called OMI zones "*Osservatorio del Mercato Immobiliare*" (Italian Revenue Agency 2016). The industrial sector in the Tuscany region is dominated by mining due to the abundance of underground resources, but also textiles industries, chemicals/pharmaceuticals, metalworking and steel, glass and ceramics, clothing and printing/publishing sectors have a strong presence in the region. Specific data on the contents of each industrial building was not available in the current database, and so the contents categories that are commonly damaged in Italian industrial buildings have been considered to be present in all buildings (until more reliable information on the contents of each building is available): i.e. fragile stock and supplies on shelves, computer equipment, industrial racks and movable manufacturing equipment. The cost of the contents has been estimated according to FEMA (2012), where it states that the value of the contents for the type of facilities considered herein can be assumed to be 44% of the total value of the construction. Finally, business interruption costs have been estimated using the HAZUS methodology (FEMA 2003).

Structural and non-structural fragility functions were derived using the analytical framework as described in Babič and Dolšek (2016). The contents fragility functions were derived using a simplification of the procedure in ATC-58 (ATC 2012), as proposed by Porter et al. (2012). Business interruption is defined herein as the time needed to repair building damage, and so median downtimes have been estimated for each damage state in the structural fragility functions. The downtime is currently only related to the structural damage as it is assumed that any non-structural damage can be addressed in parallel during the time required to recover from structural damage.

For the hazard model, the three source models (area sources, fault sources and distributed seismicity) of the 2013 European Seismic Hazard Model, ESHM13 (Woessner et al. 2015) were used together with a ground-motion prediction tree (GMPE) logic tree described in Rodrigues et al. (2017).

The stress test includes a component level risk based assessment of the key components, i.e. the industrial buildings, (ST-L1) and a probabilistic risk analysis to assess the economic losses at the system level, combining structural, non-structural, and contents damage as well as business interruption (ST-L2).

8.2 Assessment and Decision Phase

The annual probability of collapse for the component-based assessment only considers the structural components of the facilities (as these are the only components that need to be legally considered in design). This risk-based component level assessment has been undertaken for the 300 industrial facilities in Arezzo (see Figure 10) using hazard curves (i.e. PGA versus annual probability of exceedance) estimated with the OpenQuake-engine (Pagani et al. 2014) using the ESHM13 (Woessner et al. 2015), amplified considering topography-based V_{s30} estimates (USGS 2016), together with the complete damage structural fragility functions for each sub-class of structure (Babič and Dolšek 2016). According to the proposed grading system none of the structures has an annual probability of collapse below $1 \cdot 10^{-5}$ (the specified A-B boundary), which means that all facilities are classified as "partly pass" or "fail". More specifically, 260 facilities are assigned grade B (partly pass) and the others 40 facilities are assigned grade C (and thus fail), as they had an annual probability of collapse below $2.0 \cdot 10^{-4}$ (the specified B-C boundary).

For the system level assessment (where the seismic damage to a whole industrial district is estimated), economic loss-based measures and objectives have been used due to the large losses that were experienced in Italy following the Emilia-Romagna earthquakes (see Krausmann et al. 2014). The economic loss has been estimated considering the losses due to structural damage, non-structural damage, contents damage and associated direct business interruption (due to downtime). Specific objectives for these risk metrics have not yet been defined by stakeholders in the industrial facilities, and thus hypothetical values have been considered herein for illustrative purposes of the methodology. The threshold for the total AAL at the A-B boundary was defined as 0.05% of the total exposure value, and 0.10% for the B-C boundary. For the second objective, the loss due to business interruption at a mean annual rate of 10^{-4} (i.e. 1 in 10,000 years) should not be higher than 7 times the daily business interruption exposure (i.e. 10 million €) for the A-B boundary, and not greater than 30 days for the B-C boundary (42 million €).

In order to calculate probabilistic seismic risk for the spatially distributed portfolio of assets in Arezzo, the Probabilistic Event-Based Risk (PEBR) calculator from the OpenQuake-engine (Silva et al. 2014) has been employed. This calculator generates loss exceedance curves and risk maps for various return periods based on probabilistic seismic hazard, within a Monte Carlo event-based approach.

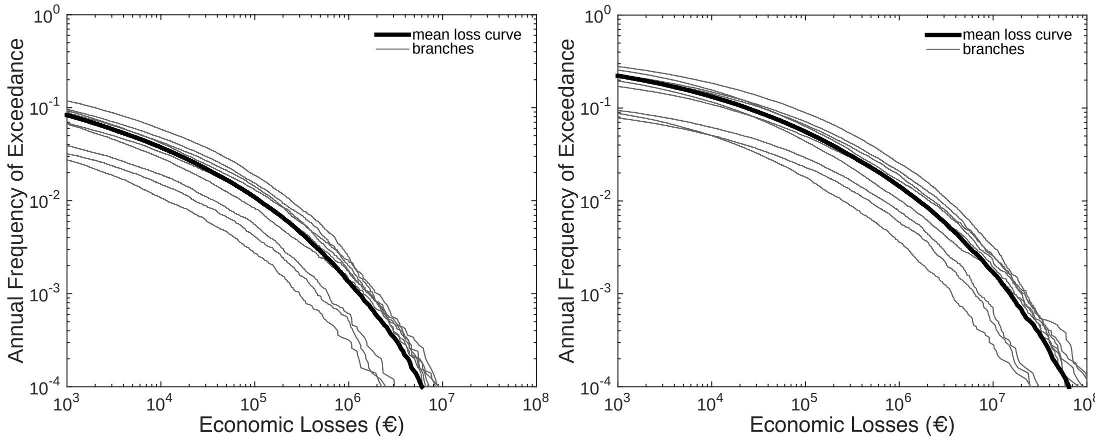

Figure 11. Loss exceedance curves for (left) structural and (right) business interruption losses in Arezzo

The average annual losses (AAL) have also been calculated from the loss exceedance curves in Figure 11 and these results show that the largest component of loss is given by business interruption. The results also indicate that the A-B system level assessment objective is not met (as the total AAL percentage is 0.052%), but the B-C level is instead met. Hence the grading would be B (partly pass) for this objective. The business interruption loss at a mean annual rate of exceedance of 10^{-4} is 64 million € (which can be translated as an average of 45 days of business interruption), and so the grading would be C (fail) for this objective.

In order to develop a potential risk reduction strategy, it is relevant to better understand which sub-classes of the industrial facilities are contributing most for the average annual losses, and to identify the type of hazard events that contribute to different loss levels. The disaggregation of the average annual loss, in terms of the critical components for each loss, are given in Figure 12. The sub-typologies that contribute most to the total average annual losses are V2 (i.e. pre-code type 2 portal frame with vertical cladding), H1 (i.e. pre-code type 1 portal frame with horizontal cladding) and V3 (i.e. low-code type 2 portal frame with vertical cladding).

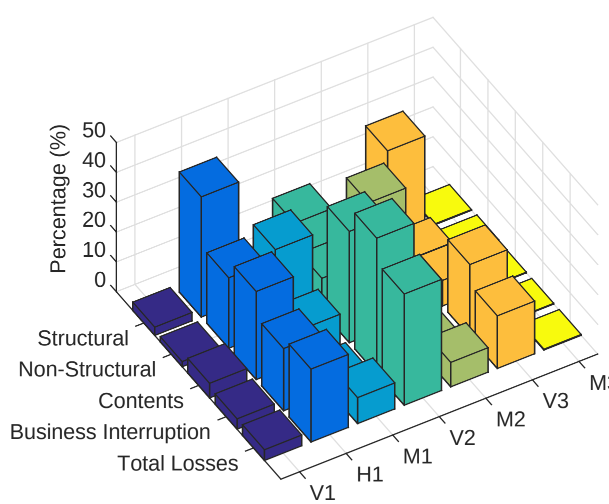

Figure 12. Disaggregation of AAL according to building sub-class for each component of loss

8.3 Report Phase

Although each industrial building has not been assessed individually in detail according to current Italian/European design requirements for single buildings, the results of the component-based assessment indicate that a significant percentage would not meet current design requirements. The final overall outcome of the stress test is driven by the system-level test and is deemed to be C (intolerable/fail), and thus this should stimulate stakeholders to upgrade the existing industrial districts such that they will improve their grading in the following stress test cycle. The performance of these pre-cast buildings would be significantly improved by strengthening the weak beam-column connections. Collaborative action from a large number of stakeholders, represented by the owners of each industrial warehouse, is required to improve the grading of the stress test, and this should be encouraged and enforced by the regulatory authorities.

However, it is noted that the outcomes of the stress test presented herein are highly influenced by the assumptions made in developing the exposure model as well as the definition of the target objectives, which have been defined herein by the authors rather than the relevant stakeholders. Hence, further comments on the outcome of the stress test are not made in these conclusions and instead it is stressed that additional efforts are needed in the future to work with the owners of the industrial facilities to collect reliable content and annual revenue data, and to identify the most appropriate target objectives.

9. Discussion - Conclusions

An engineering risk-based methodology for conducting stress tests of critical non-nuclear infrastructures has been applied to six CIs in Europe. Different stress test levels were selected according to the characteristics of the particular CIs and the available resources. The objective was to demonstrate the efficiency of the methodology and how the proposed framework can be specified and implemented with regard to different types of CIs, i.e. single site, geographically extended, distributed multi-site, each one exposed to varying hazards. These case studies can be used as a basis for similar types of CIs, while the proposed framework can be adjusted and implemented to other sectors. However, risk measures and acceptance criteria may vary depending on the peculiarities of each CI, even if of similar type. For example, in case of port facilities, a risk measure in terms of economic loss could be an alternative, instead of the loss in terms of cargo or container handled that is used in the present application. Inevitably, the heterogeneity of the different CIs justifies the reasonable assumptions and/or simplifications made in some steps of the applications. In this context, the authors disavow a quantitative interpretation of the results provided, as these applications were not, nor should they be, considered formal stress tests in each particular CI. In Table 10, the key elements of the six case studies are summarised, i.e. CI data, hazard data, risk measures, risk acceptance criteria (component, system), stress test level, risk acceptance check, and risk mitigation guidelines.

*The Table is provided at the end of the document***Table 10.** Overview of the key elements of the six stress test case studies

The stress test to the oil refinery of Milazzo showed that the earthquake impact is important for the atmospheric storage tanks. The tsunami effect on the atmospheric storage vessels along the shore line is relatively negligible in terms of cascading effects and increase of the overall risk on population. Neither an earthquake nor a tsunami significantly increases the failure frequency of, and hence the risk imposed by, pressurized vessels. Despite this, the risk remains largely dominated by the LPG tanks failures due to industrial-related causes, whereas the impact of the natural hazards is limited. Mitigation measures include the enhancement of the emergency preparedness for multiple fire scenarios and the structural upgrade of tanks.

The stress test to a large dam in Switzerland exposed to multi-hazard effects, considering earthquakes, floods, internal erosion and electromechanical malfunctions in key systems, showed that the first of three risk objectives concerning the dam-reservoir system and the probability of failures taking place was met. The second objective, related to the expected losses downstream was not met, while the third one, defined on the basis of an F-N curve, classified the risk as ALARP (as low as reasonably practicable). The most efficient mitigation measure is to upgrade the bottom outlet of the dam to prevent all overtopping events. Also, the resilience of the dam-reservoir is very much defined by the capacity to perform a successful and timely drawdown operation, therefore cascade effects become important when the possibility of drawing down the reservoir is lost, and a substantial inflow arrives. The mitigation measures for the downstream area include the reinforcement or relocation of the high-risk buildings, the installation of early warning systems and the improvement of emergency planning, e.g. shelthes, escape routes.

The application to Baku-Tiblisi-Ceyhan pipeline that crosses strike-slip fault segments in the eastern Anatolia in Turkey, indicated that two pipe-fault crossings are critical as their failure rates exceed the allowable rate. The risk assessment showed that risk is classified possibly unjustifiable. The risk mitigation guidelines are focused at the retrofitting of the pipelines at the critical crossings by changing the angle of the pipe-fault intersection.

The stress test to the Gasunie gas distribution network in Groningen, Netherlands, exposed to earthquake and liquefaction effects, showed that soil liquefaction is the dominant failure mechanism. In particular, with respect to components, the pipe sections pass the stress test, while stations pass the stress test only partially. With respect to the systemic risk the stress test is passed. The safety and resilience of this CI will be improved by reassessing the need for retrofitting of the critical pipe sections identified in this study. The stress test also revealed the need for site-specific fragility functions for the stations and the need for further research into the liquefaction mechanisms for the Groningen site conditions.

The stress test to the port infrastructures of Thessaloniki exposed to seismic, tsunami and liquefaction hazards showed a variation in the outcomes depending on the type of analysis. Most of the port components do not pass the safety test against collapse for both earthquake and tsunami hazards in the case of a component level assessment. The systemic risk is possibly unjustifiable and negligible for the PRA of earthquake and tsunami hazards respectively, meaning that the port partly passes or passes this evaluation of the stress test. The scenario-based assessment showed the importance of the modelling approach of local site effects in the outcome of the stress test. The proposed mitigation planning includes the potential upgrade of the electric power substations due to their criticality for the port operations or/and the installation of alternative power sources. Moreover, the resilience planning of the port should consider the fragility and importance of each component, interdependencies, recovery priorities and buffering capabilities.

The stress test to an industrial district in Northern Italy, exposed to seismic hazard, concluded that the facilities partly pass or fail to pass the component-based assessment. For the system level assessment, where economic loss-based measures and objectives have been used, the industrial district partly passes or fails to pass the test depending on the considered boundaries used as thresholds of loss due to business interruption. Risk mitigation can be achieved on the basis of strengthening building sub-classes that contribute most to the total losses, in particular the weak beam-column connections of pre-cast buildings.

In summary, standardized actions and results are foreseen in the proposed framework, which are defined based on the level of stress testing and the level of detail that is applied. For example, if a low level is adopted leading to lack of risk acceptance then a more advanced method should be used, while if a component fails the assessment risk mitigation actions must be applied. In all six case studies the risk objectives boundaries have been set mainly based on expert judgment. However, formulation of risk acceptance criteria is not a straightforward task. In practice, setting objectives and establishing risk measures is difficult and strongly dependent on legal, socio-economic and political contexts and they should be defined by the corresponding stakeholders. Nevertheless, when needed, the results of the stress tests have the potential to stimulate stakeholders to take action to upgrade the existing infrastructure aiming to improve their grading in the following stress test cycle toward improving the resilience and preparedness of CIs. Lessons learned through the six applications is the need for improvement of existing assessment approaches considering the uncertainties in the quantification of hazard, vulnerability and loss estimates as well as the need for site-specific fragility models. An important issue is also the collaborative action and effective communcation of the key actors, i.e. stakeholders, experts, owners and operators of the CIs and regulatory authorities.

Acknowledgements

The work presented in this paper was conducted within the project "*STREST: Harmonized approach to stress tests for civil infrastructures against natural hazards*" funded by the European Community's Seventh Framework Programme under grant agreement no. 603389. The authors gratefully acknowledge this funding. The authors acknowledge the contributions of the Work Package leaders, Mr Peter Zwicky, Prof Fabrice Cotton, Prof Iunio Iervolino, Prof Bozidar Stojadinovic, Dr Fabio Taucer as well as Dr Simona Esposito, Prof Matjaž Dolšek and Dr George Tsionis. The methods, results, opinions, findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the European Commission or the owners and stakeholders of the studied CIs.

References

Akkar S, Sandıkkaya MA, Bommer JJ (2014a) Empirical ground-motion models for point-and extended-source crustal earthquake scenarios in Europe and the Middle East. Bulletin of Earthquake Engineering 12(1):359–387

Akkar S, Sandıkkaya MA, Bommer JJ (2014b) Erratum to: Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. Bulletin of Earthquake Engineering 12(1):389–390

American Lifelines Alliance (ALA) (2001). Seismic Fragility Formulations for Water Systems, Part 1–Guideline, http://www.americanlifelinesalliance.org

ATC - Applied Technology Council (2012) ATC-58: Guidelines for Seismic Performance Assessment of Buildings, 100% Draft. Redwood City, CA

Babič A, Dolšek M (2016) Seismic fragility functions of industrial precast building classes. Engineering Structures 118: 357-370

Basco A, Salzano E (2016) The vulnerability of industrial equipment to tsunami. Journal of Loss Prevention in the Process Industries (in press), DOI: 10.1016/j.jlp.2016.11.009

Bommer JJ (2013) Proposals for New GMPEs for the Prediction of PGA and PGV in the Groningen Gas Field, NAM internal note, 37pp

Chang SE (2000) Disasters and transport systems: Loss, recovery, and competition at the Port of Kobe after the 1995 earthquake. J Transp Geogr 8(1): 53–65

Cheng Y and Akkar S (2017) Probabilistic permanent fault displacement hazard via Monte Carlo simulation and its consideration for the probabilistic risk assessment of buried continuous steel pipelines. Earthquake Engineering and Structural Dynamics 46 (4), 605-620

Cotton F et al. (2016) Deliverable 3.7: Multi-hazard assessment of low-probability hazard and LP-HC events for six application areas. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

CPR 18E (1999) Guidelines for quantitative risk assessment. Committee for the Prevention of Disasters (CPR)

Davies G, Griffin J, Lovholt F, Glimsdal S, Harbitz C, Thio HK, Lorito S, Basili R, Selva J, Geist E, Baptista MA (2017) A global probabilistic tsunami hazard assessment from earthquake sources, in Tsunamis: Geology, Hazards and Risks (Scourse EM, Chapman NA, Tappin DR & Wallis SR Eds), Geological Society, London, Special Publications, 456, https://doi.org/10.1144/SP456.5

Dost B, Caccavale M, Van Eck T, Kraaijpoel D (2013) Report on the expected PGV and PGA values for induced earthquakes in the Groningen area, KNMI report, 26pp. (http://bibliotheek.knmi.nl/knmipubDIV/Report_on_the_expected_PGV_and_PGA_values_for_induced_earthquakes.pdf)

EC (2012) Directive 2012/18/EU of the European Parliament and of the Council of 4 July 2012 on the control of major-accident hazards involving dangerous substances, amending and subsequently repealing Council Directive 96/82/EC. Off J Eur Union, pp. 1-37

EMDAT (2019) OFDA/CRED International disaster database, Université Catholique de Louvain, Brussels, Belgium, https://www.emdat.be/

ENSREG (2012) Stress tests performed on European nuclear power plants. Peer Review Report. European Nuclear Safety Regulators Group, available online: http://www.ensreg.eu/node/407

Eidinger JM, Avila EA (1999) Guidelines for the seismic evaluation and upgrade of water transmission facilities (Vol. 15). ASCE Publications

Esposito S, Iervolino I, d'Onofrio A, Santo A, Franchin P, Cavalieri F (2015) Simulation-based seismic risk assessment of a gas distribution network. Computer-Aided Civil and Infrastructure Engineering. DOI: 10.1111/mice.12105

Esposito S, Stojadinovic B, Babič A, Dolšek M, Iqbal S, Selva J, Broccardo M, Mignan A, Giardini D (2019) A risk-based multi-level methodology to stress test critical non-nuclear infrastructure systems, Journal of Infrastructure Systems (accepted).

Esposito S, Stojadinovic B, Mignan A, Dolšek M, Babič A, Selva J, Iqbal S, Cotton F, Iervolino I (2016) Reference Report RR4: Guidelines for stress-test design for non-nuclear critical infrastructures and systems: Methodology. STREST EC/FP7 project: Harmonized approach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

FEMA (2003) Multi-hazard Loss Estimation Methodology, Earthquake Model, HAZUS. Federal Emergency Management Agency and National Institute of Buildings Sciences, Washington DC

FEMA (2012) FEMA E-74: Reducing the risks of nonstructural earthquake damage – a practical guide, Report by Federal Emergency Management Agency

Franchin P, Cavalieri F, Pinto PE, Lupoi A, Vanzi I, Gehl P, Kazai B, Weatherill G, Esposito S, Kakderi K (2011) General methodology for systemic seismic vulnerability assessment. Deliverable 2.1 SYNER-G EC/FP7 project, http://www.vce.at/SYNER-G/

Galbusera L, Giannopoulos G, Argyroudis S, Kakderi K (2018). A Boolean Networks approach to modeling and resilience analysis of interdependent critical infrastructures. Computer-Aided Civil and Infrastructure Engineering 33(12): 1041-1055.

Giannopoulos G, Filippini R, Schimmer M (2012) Risk assessment methodologies for critical infrastructure protection. Part I: A state of the art. Publications Office of the European Union, Luxembourg. doi: 10.2788/22260

Giardini D. et al. (2013) Seismic Hazard Harmonization in Europe (SHARE). Online Data Resource, http://portal.share-eu.org:8080/jetspeed/portal/, doi: 10.12686/SED-00000001-SHARE

Grezio A, Babeyko A, Baptista MA, Behrens J, Costa A, Davies G, Geist EL, Glimsdal S, González FI, Griffin J, Harbitz CB, LeVeque RJ, Lorito S, Løvholt F, Omira R, Mueller C, Paris R, Parsons T, Polet J, Power W, Selva J, Sørensen MB, Thio HK (2017) Probabilistic Tsunami Hazard Analysis (PTHA): multiple sources and global applications, Reviews of Geophysics 55: 1158-1198, DOI: 10.1002/2017RG000579

Grimaz S (2014) Can earthquakes trigger serious industrial accidents in Italy? Some considerations following the experiences of 2009 L'Aquila (Italy) and 2012 Emilia (Italy) earthquakes. Bollettino di Geofisica Teorica ed Applicata 55(1): 227-237. doi: 10.4430/bgta0116

Gunn R, Balissat M, Manso P, et al. (eds) (2016) Proceedings of the 13th International Benchmark Workshop on Numerical Analysis of Dams. In: ICOLD Proceedings. Swiss Committee on Dams, Lausanne, Switzerland.

Gutenberg B, Richter CF, (1956). Magnitude and Energy of Earthquakes. Annali di Geofisica 9: 1–15. Helm P (1996) Integrated Risk Management for Natural and Technological Disasters. Tephra. 15(1): 4-13.

Honegger DG, Wijewickreme D (2013) Seismic risk assessment for oil and gas pipelines. In Handbook of Seismic Risk Analysis and Management of Civil Infrastructure Systems pp.682-715. Elsevier

http://www.nexus.globalquakemodel.org/gem-vulnerability/posts/draft-content-vulnerability-guidelines

Idriss IM, Boulanger RW, 2008. Soil Liquefaction during earthquake, EERI monograph MNO-12 on earthquake engineering. Earthquake Engineering Research Institute, Oakland (CA), USA

Jonkman SN, Van Gelder PHAJM, Vrijling JK (2003) An overview of quantitative risk measures for loss of life and economic damage. Journal of Hazardous Materials 99(1):1-30

Kakderi K, Selva J, Pitilakis K (2014) Application in the Harbor of Thessaloniki, in: K. Pitilakis et al. (eds). Systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications, Springer Science+Business Media, Dordrecht, doi: 10.1007/978-94-017-8835-9_12

Karafagka S, Fotopoulou S, Pitilakis K (2016) Tsunami fragility curves for seaport structures, In the 1st International Conference on Natural Hazards & Infrastructure, Chania, Greece, June 28-30

Krausmann E, Cozzani V, Salzano E, Renni E (2011) Industrial accidents triggered by natural hazards: an emerging risk issue. Natural Hazards and Earth System Sciences 11:921–929

Krausmann E, Cruz AM (2013) Impact of the 11 March 2011, Great East Japan earthquake and tsunami on the chemical industry. Nat Hazards 67(2): 811–828. doi: 10.1007/s11069-013-0607-0

Krausmann E, Piccinelli R, Ay BÖ, Crowley H, Uckan E, Erdik M, Lanzano G, Salzano E, Iervolino I, Esposito S, Pistolas A, Kakderi K, Pitilakis D, Pitilakis K, Steenbergen R (2014) Deliverable D2.3: Report on lessons learned from recent catastrophic events. STREST EC/FP7 project: Harmonized approach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

Krausmann E, Cruz AM, Salzano E (2016). Natech Risk Assessment and Management - Reducing the Risk of Natural-Hazard Impact on Hazardous Installations, Elsevier, 1st Ed., ISBN-10: 0128038071, 268

Kutkov V.A, Tkachenko V.V. (2017). Fukushima Daiichi accident as a stress test for the national system for the protection of the public in event of severe accident at NPP. Nuclear Energy and Technology 3(1):38-42.

Lorito S, Selva J, Basili R, Romano F, Tiberti MM. Piatanesi A (2015) Probabilistic hazard for seismically-induced tsunamis: accuracy and feasibility of inundation maps, Geophys J Int 200(1):574-588

Mignan A, Wiemer S, Giardini D (2014) The quantification of low-probability–high-consequences events: part I. A generic multi-risk approach. Nat Hazards 1–24. doi: 10.1007/s11069-014-1178-4

Mignan A, Danciu L, Giardini D (2015) Reassessment of the Maximum Fault Rupture Length of Strike-Slip Earthquakes and Inference on M_{max} in the Anatolian Peninsula, Turkey. Seismol. Soc. Am., 86, 890-900, doi: 10.1785/0220140252

Mignan A, Danciu L, Giardini D (2016a) Considering large earthquake clustering in seismic risk analysis. Nat. Hazards, doi: 10.1007/s11069-016-2549-9

Mignan A, Scolobig A, Sauron A (2016b) Using reasoned imagination to learn about cascading hazards: a pilot study. Disaster Prevention and Management, 25, 329-344, doi: 10.1108/DPM-06-2015-0137

Miraglia S, Courage W, Meijers P (2015) Fragility functions for pipeline in liquefiable sand: a case study on the Groningen gas-network. In Haukaas T (Ed.) In the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12), Vancouver, Canada, July 12-15

National Institute of Building Sciences, NIBS (2004) Direct physical damage-general building stock. HAZUS-MH Technical manual, Chapter 5. Federal Emergency Management Agency, Washington, D.C.

Newmark NM, Rosenblueth E (1971) Fundamentals of earthquake engineering. Englewood Cliffs: Prentice-Hall.Opdyke A, Javernick-Will A, Koschmann M (2017) Infrastructure hazard resilience trends: an analysis of 25 years of research. Nat Hazards 87(2): 773–789. doi:10.1007/s11069-017-2792-8

Pagani M, Monelli D, Weatherill G, Danciu L, Crowley H, Silva V, Henshaw P, Butler L, Nastasi M, Panzeri L, Simionato M and Vigano D (2014) OpenQuake Engine: An open hazard (and risk) software for the Global Earthquake Model. Seismological Research Letters, 85(3): 692-702.

Pescaroli G, Alexander D (2016) Critical infrastructure, panarchies and the vulnerability paths of cascading disasters. Nat Hazards 82(1): 175–192. doi: 10.1007/s11069-016-2186-3

Pitilakis K, Argyroudis S, Fotopoulou S, Karafagka S, Kakderi K, Selva J (2017). Application of new stress test concepts for port infrastructures against natural hazards. The case of Thessaloniki port in Greece, Reliability Engineering & System Safety 184:240-257.

Pitilakis K, Argyroudis S, Fotopoulou S, Karafagka S, Anastasiadis A, Pitilakis D, Raptakis D, Riga E, Tsinaris A, Mara K, Selva J, Iqbal S, Volpe M, Tonini R, Romano F, Brizuela B, Piatanesi A, Basili R, Salzano E, Basco A, Schleiss AJ, Matos JP, Akkar S, Cheng Y, Uckan E, Erdik M, Courage W, Reinders J, Crowley H, Rodrigues D (2016). Deliverable D6.1: Integrated report detailing analyses, results and proposed hierarchical set of stress tests for the six CIs. STREST EC/FP7 project: Harmonized approach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

Porter K, Cho I, Farokhnia K (2012) Contents seismic vulnerability estimation guidelines. Global Vulnerability Consortium

Renni E, Basco A, Businі V, Cozzani V, Krausmann E, Rota R, Salzano E (2010). Awareness and mitigation of NaTech accidents: Toward a methodology for risk assessment. Chemical Engineering Transactions 19:383-389

Reuters (2010) Flash floods inundate central Europe, http://www.reuters.com/article/2010/08/08/us-europe-floods-idUSTRE67617F20100808

Rodrigues D, Crowley H and Silva V (2017) Earthquake loss assessment of precast RC industrial structures in Tuscany (Italy), Bulletin of Earthquake Engineering 16(2), DOI: 10.1007/s10518-017-0195-6

Salzano E, Basco A, Busini V, Cozzani V, Renni E, Rota R (2013) Public Awareness Promoting New or Emerging Risk: Industrial Accidents Triggered by Natural Hazards. Journal of Risk Research 16:469-485

Salzano E, Basco A, Karafagka S, Fotopoulou S, Pitilakis K, Anastasiadis A, Matos JP, Schleiss AJ (2015). Deliverable D4.1: Guidelines for performance and consequences assessment of single-site, high-risk, non-nuclear critical infrastructures exposed to multiple natural hazards. STREST EC/FP7 project: Harmonized apporach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

Selva J, Iqbal S, Taroni M, Marzocchi W, Cotton F, Courage W, Abspoel-Bukman L, Miraglia S, Mignan A, Pitilakis K, Argyroudis S, Kakderi K, Pitilakis D, Tsinidis G, Smerzini C (2015) Deliverable D3.1: Report on the effects of epistemic uncertainties on the definition of LP-HC events. STREST EC/FP7 project: Harmonized approach to stress tests for critical infrastructures against natural hazards, www.strest-eu.org

Selva J, Tonini R, Romano F, Volpe M, Brizuela B, Piatanesi A, Basili R, Lorito S (2016) From regional to site specific SPTHA through inundation simulations: a case study for three test sites in Central Mediterranean. EGU General Assembly, Vienna, Austria, 17-22 April, Abstract #EGU2016-16988

Silva V, Crowley H, Pagani M, Pinho R (2014) Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment. Natural Hazards, 72(3):1409-1427

Smerzini C, Pitilakis K, Hasmemi K (2016) Evaluation of earthquake ground motion and site effects in the Thessaloniki urban area by 3D finite-fault numerical simulations. B Earthq Eng, 15(3):787–812

Theocharidou M, Giannopoulos G (2015) Risk assessment methodologies for critical infrastructure protection. Part II: A new approach. Publications Office of the European Union, Luxembourg. doi: 10.2788/621843

USGS (2016) Earthquale hazards program (Predefined Vs30 Mapping). Available at URL: http://earthquake.usgs.gov/hazards/apps/vs30/predefined.php

Weatherill G, Esposito S, Iervolino I, Franchin P, Cavalieri F (2014) Framework for seismic hazard analysis of spatially distributed systems, in: K. Pitilakis et al. (eds). Systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications, Springer Science+Business Media, Dordrecht, doi: 10.1007/978-94-017-8835-9_3

WINN_TA-NAM (2016). Technical Addendum to the Winningsplan 2016. NAM, April 1st (in Dutch)

Woessner J, Danciu L, Giardini D, Crowley H, Cotton F, Grunthal G, Valensise G, Arvidsson R, Basili R, Demircioglu M, Hiemar S, Meletti C, Musson R, Rovida A, Sesetyan K, Stucchi M (2015) The 2013 European Seismic Hazard Model-Key Components and Results. *Bulletin of Earthquake Engineering*, DOI 10.1007/s10518- 015-9795-1

Zenz G, Goldgruber M (eds) (2013) Proceedings of the 12th International Benchmark Workshop on Numerical Analysis of Dams. ICOLD, Graz, Austria

Table 10

	Oil refinery and petrochemical plant in Milazzo, Italy	Alpine earthfill dam, Switzerland	Baku-Tiblisi-Ceyhan pipeline,Turkey	Gasunie national gas storage and distribution network, Netherlands	Port infrastructure of Thessaloniki, Greece	Industrial district in the region of Tuscany, Italy
CI data						
Number of components	177 hydrocarbons storage tanks (gasoline, gasoil, crude oil, LPG)	conceptual dam-reservoir system: dam and foundation, spillways, bottom outlet, hydropower system and reservoir. downstream area: approx. 1000 buildings	1,800 km buried pipeline of crude oil	~1,000 km pipe network for gas transport; 11 M&R stations; 15 feeding points and 91 receiving points as end nodes	25 waterfront structures (WS); 35 cargo handling equipment (CE); 4 gantry cranes (GC); 85 building and storage facilities (BD); electric power lines (EL) and 17 distribution substations (ES)	300 buildings (structural elements)
Typology	steel storage tanks with variable capacities: 100 m^3 (fuel oil, gasoil, gasoline, kerosene) to 160 000 m^3 (crude oil) located in catch basins (bunds) with concrete surfaces. LPG: pressurised spheres	dam-reservoir system: custom typology; height: 100 m, reservoir capacity: 100 000 000 m^3 downstream area: buildings grouped according to the number of storeys	material: steel API XL Grade X65; diameter: 42 inches; thickness: 20.62mm; buried depth: 1.5m	main gas transmission pipes (4 to 8 MPa); diameters ranging from 114 mm to 1219 mm; M&R stations: predominately small masonry buildings	WS: concrete gravity block type quay walls, non-anchored components; CE: non-anchored components without backup power supply; GC: capacity 45 tons; EL: non-vulnerable; ES: low-voltage, with non-anchored components	pre-code or low-code reinforced concrete precast industrial buildings
Hazard data						
Hazard type	seismic (ground shaking and liquefaction), tsunami	seismic, flood, internal erosion, bottom outlet malfunction, and hydropower system malfunction	permanent fault displacement (PFD)	seismic (ground shaking and liquefaction)	seismic (ground shaking and liquefaction), tsunami	seismic
Model	probabilistic seismic hazard analysis (PSHA); seismic probabilistic tsunami hazard analysis (SPTHA)	seismic hazard maps of the Swiss Seismological service; extreme flood analyses; expert knowledge and historically observed malfunction frequencies	PFD at 5 fault crossings: scenario-based (2475 years); uncertainty in pipe-fault angle crossing	PSHA: Z1 model for Groningen area (Dost et al. 2013); modified GMPE (Bommer 2013)	PSHA: ESHM13 and Weatherill et al. (2014). SPTHA: 253 representative scenarios based on inundation simulation. seismic scenario-based: 475 years (EQL, NL), 4975 years (EQL, NL)	2013 European Seismic Hazard Model (ESHM13)
Risk measures						
Component	annual probability of release of content of hazardous materials (defined standard mass flow rate)	1) annual probability of failure and 2) annual probability of household loss	annual probability of loss of pressure integrity	annual failure probability	annual probability of collapse	annual probability of collapse
System	locational and individual risk (fatalities/year); societal risk (F/N curve)	obj1: uncontrolled release of the reservoir; obj2: probability of a household being collapsed or washed away (lost	N/A	connectivity loss (Esposito et al., 2016)	normalized loss of: total number of containers handled (loaded and unloaded) per day (TCoH); total cargo handled (loaded	obj1: average annual loss; obj2: mean annual rate of economic loss (due to structural damage, non-structural damage, contents

		volume/year)			and unloaded) per day (TCaH)	damage and direct business interruption)
Risk acceptance criteria						
Component	target probability of collapse of equipment with the instantaneous release of content less or equal to $1.0x10^{-8}$	all the components comply to and slightly exceed regulatory requirements (assumption of a conceptual dam)	target probability of pipeline failure in 2475 years: $4.0x10^{-5}$ scenario-based: AA-A: 2-10, A-B: 10-50, B-C: 50-100 (% loss for 2475 years)	AA-A: $8.0x10^{-6}$ A-B: $6.0x10^{-5}$ B-C: $1.4x10^{-4}$ Pipes (km/year); M&R stations (object/year)	target probability of collapse: $1.0x10^{-5}$	target probability of collapse: A-B: $1.0x10^{-5}$ B-C: $2.0x10^{-4}$
System*	fatality rate as given by acceptability parameters either for the locational risk ($1.0x10^{-4}$ f/year for workers, $1.0x10^{-6}$ f/year for population) or for the societal curve ($F<1.0x10^{-3}/N^2$)	obj1: p(failure)≤ $1.0x10^{-5}$ obj2: AA-A: 7.5 m^3/yr (~one household lost per 100 years), A-B: 75 m^3/yr, B-C: 750 m^3/yr (~one household lost per year)	N/A	AA-A: $1.0x10^{-4}$, A-B: $4.0x10^{-4}$, B-C: $8.0x10^{-4}$, for annual probability of 100% loss (Esposito et al., 2016)	PSHA & SPTHA: AA-A: $7.5x10^{-4}$, A-B: $2.0x10^{-3}$, B-C: $4.5x10^{-3}$, for annual probability of 100% loss. scenario-based: AA-A: 10, A-B: 30, B-C: 50 (% loss for 475 years) AA-A: 30, A-B: 50, B-C: 70 (% loss for 4975 years)	obj1: A-B: 0.05%, B-C: 0.1% obj2: A-B: 10^{-4} for 10 Million Euro loss (7 days), B-C: 10^{-4} for 42 million Euro loss (30 days)
Stress test level	ST-L1a; ST-L2b; ST-L2d	ST-L2b/L2d; ST-L3c/L3d	ST-L1	ST-L1; ST-L2a	ST-L1; ST-L2b; ST-L2d/L3d	ST-L1; ST-L2a
Risk acceptance check (AA-A: pass, B: partly pass, C: fail)	ST-L1a, ST-L2b, ST-L2d: AA-C (earthquake), AA-C (tsunami)	ST-L2 (seismic): AA-A ST-L3 (all five hazards): AA-A	ST-L1: (hazard based assessment), fault crossings #2, #3 and #4 do not comply with the code requirements. ST-L1: (design based assessment), all five crossings comply with the code requirements. ST-L1: B (risk based assessment), fault crossings #3 and #4 do not comply with the target risk tolerance	ST-L1 (components): AA-B ST-L2a (system): AA-A (see Table 8)	ST-L1: C (seismic), AA-C (tsunami) ST-L2b: B (seismic), AA (tsunami) ST-L2d/L3d: AA, A, B, C (depending on the scenario and analysis type, see Table 9)	ST-L1: B (260 facilities), C (40 facilities) ST-L2a: B for obj1, C for obj2
Risk mitigation guidelines	reinforcing of tanks for earthquake and tsunami-induced structural damage and defining the emergency response in case of release of hazardous materials	invest on the resilience of the reservoir drawdown mechanism (bottom outlet)	upgrading of two critical fault crossing points by changing the orientation angles of pipes (after upgrading, the risk is classified as AA)	retrofitting pipe sections identified; site-specific fragility functions for the Gasunie-GTS stations	consider potential upgrade of substations or/and alternative power sources; measures to secure ships and port equipment	upgrade of building sub-classes that contribute most to the total losses

** boundaries for system risk acceptance criteria were based on expert judgement*