European Commission

JRC TECHNICAL REPORTS

Guidelines for stress-test design for non-nuclear critical infrastructures and systems: Methodology

STREST Reference Report 4

Editors: Stojadinovic B, Esposito S
Contributors:, Cotton F, Giardini D, Iqbal S, Mignan A, Selva J
Reviewers: Dolšek M, Babič A
Publishing editor: Tsionis G

2016

EUR 28342 EN

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

JRC Science Hub
https://ec.europa.eu/jrc

JRC104653

EUR 28342 EN

Print	ISBN 978-92-79-64603-4	ISSN 1018-5593	doi:10.2788/607
PDF	ISBN 978-92-79-64602-7	ISSN 1831-9424	doi:10.2788/659118

Luxembourg: Publications Office of the European Union, 2016

How to cite this report: Stojadinovic B, Esposito S, Babič A, Cotton F, Dolšek M, Giardini D, Iqbal S, Mignan A, Selva J, Tsionis G, *Guidelines for stress-test design for non-nuclear critical infrastructures and systems: Methodology*, EUR 28342 EN, doi:10.2788/659118

D 7.6.4

DELIVERABLE

PROJECT INFORMATION

Project Title:	Harmonized approach to stress tests for critical infrastructures against natural hazards
Acronym:	STREST
Project N°:	603389
Call N°:	FP7-ENV-2013-two-stage
Project start:	01 October 2013
Duration:	36 months

DELIVERABLE INFORMATION

Deliverable Title:	Guidelines for stress-test design for non-nuclear critical infrastructures and systems: Methodology
Date of issue:	31 July 2016
Work Package:	WP7 – Dissemination and stakeholder interaction
Editor/Author:	Bozidar Stojadinovic Simona Esposito (ETH Zürich)
Reviewer:	Matjaž Dolšek (ETH Zürich) Anže Babič (UL)

REVISION: Final

SEVENTH FRAMEWORK PROGRAMME

Project Coordinator:	Prof. Domenico Giardini
Institution:	ETH Zürich
e-mail:	mailto: giardini@sed.ethz.ch
fax:	+ 41 446331065
telephone:	+ 41 446332610

Table of contents

Abstract iii
Acknowledgments v
Deliverable contributors vii
1. Introduction 1
2. ST@STREST methodology for stress testing of critical non-nuclear infrastructures.. 3
   2.1 ST@STREST Multiple-expert Integration: EU@STREST.... 4
   2.2 ST@STREST Workflow 6
      2.2.1 PHASE 1: Pre-Assessment phase 7
      2.2.2 PHASE 2: Assessment phase.... 9
      2.2.3 PHASE 3: Decision phase.... 9
      2.2.4 PHASE 4: Report phase 10
   2.3 ST@STREST Test Levels 10
      2.3.1 Component Level Assessment 12
      2.3.2 System Level Assessment 16
   2.4 ST@STREST Data Structures 21
      2.4.1 Application of BNs to natural hazards and CIs 22
      2.4.2 Discussions.... 24
   2.5 ST@STREST Grading System 25
      2.5.1 Risk limits and boundaries between grades.... 26
      2.5.2 Grading system in time domain.... 27
      2.5.3 Grading of the components.... 28
      2.5.4 Grading of the system with consideration of epistemic uncertainties . 30
      2.5.5 Discussion and future developments 31
   2.6 ST@STREST Penalty System 31
      2.6.1 Proposed Penalty System 32
3. Incorporating ST@STREST into the life cycle management of non-nuclear critical CIs 39
   3.1 Introduction 39
   3.2 LCC including natural hazard risk 39
   3.3 Unified life cycle management of CI 40
      3.3.1 Life cycle analysis including stress test and post-event data 41
   3.4 Discussions 44
4. Using ST@STREST to enhance societal resilience 47
   4.1 Modelling resilience of critical infrastructures against natural hazards.... 48
   4.2 Resilience-based stress test for critical non-nuclear infrastructures 51
      4.2.1 Future research and discussions 52

5. Conclusions and recommendations 55
References 57
List of abbreviations and definitions 63
List of figures 65
List of tables 67

Abstract

Critical infrastructures (CIs) are of essential importance for modern society: these systems provide the essential functions of public safety and enable, through their services, the higher-level functions of a community, such as housing, education, healthcare and the economy. A harmonized approach for stress testing critical non-nuclear infrastructures, ST@STREST, has been developed. The aims of the ST@STREST methodology and framework are to quantify the safety and the risk of individual components as well as of whole CI system with respect to extreme events, and to compare the expected behavior of the CI to acceptable values.

This report summarizes the ST@STREST methodology and framework, and addresses the extensions of the proposed methodology towards life-cycle management of civil infrastructures and evaluation of civil infrastructure system post-disaster resilience. A detailed elaboration of these topics is presented in the accompanying Work Package 5 reports. The ST@STREST methodology has been applied to six key representative Critical Infrastructures (CIs) in Europe, exposed to variant hazards, namely: a petrochemical plant in Milazzo, Italy, large dams of the Valais region in Switzerland, hydrocarbon pipelines in Turkey, 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 outcomes of these stress tests are presented in the STREST Reference Report 5.

Acknowledgments

The work presented in this report 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 [FP7/2007-2013] under grant agreement no. 603389. The authors gratefully acknowledge this funding. The methods, results, opinions, findings and conclusions presented in this report are those of the authors and do not necessarily reflect the views of the European Commission.

Deliverable contributors

ETH Zurich	Simona Esposito
	Bozidar Stojadinovic
	Arnaud Mignan
	Domenico Giardini
UL	Matjaž Dolšek
	Anže Babič
INGV	Jacopo Selva
	Sarfraz Iqbal
GFZ	Fabrice Cotton

1. Introduction

Critical infrastructures (CIs) are of essential importance for modern society: these systems provide the essential functions of public safety and enable, through their services, the higher-level functions of a community, such as housing, education, healthcare and the economy. Extreme natural events can interrupt services, cause damage, or even destroy such CI systems, which consequently trigger disruption of vital socio-economic activities, extensive property damage, and/or human injuries or loss of lives. Recent catastrophic events showed that the CI systems rarely recover their functionality back to the pre-disaster state, significantly increasing the concerns of the public.

In the context of the STREST project, a harmonized approach for stress testing critical non-nuclear infrastructures, ST@STREST, has been developed. The aims of the ST@STREST methodology and framework are to quantify the safety and the risk of individual components as well as of whole CI system with respect to extreme events and to compare the expected behavior of the CI to acceptable values. In particular, a multi-level stress test methodology has been proposed. Each level is characterized by a different scope (component or system) and by a different level of risk analysis complexity (starting from design codes and ending with state-of-the-art probabilistic risk analyses, such as cascade modelling). This allows flexibility and application to a broad range of infrastructures. The framework is composed of four main phases and nine steps. First the goals, the method, the time frame, and the total costs of the stress test are defined. Then, the stress test is performed at component and system level; additionally, the outcomes are checked and analyzed. Finally, the results are reported and communicated to stakeholders and authorities. The ST@STREST data framework, used to store and manage the data about the CI under test, is also flexible, in that it allows the use of data structures that support frequentist (event and fault trees, bow ties) and belief-based notions of probability.

The stress test approach proposed in this project addresses the vulnerabilities of CIs to catastrophic but rare (high-consequence low-probability) natural hazard events. An extension of the proposed ST@STREST methodology and framework to integrate the results of stress tests and the data retrieved after disastrous events with the data collected during every-day operation of the system and its degradation (low-consequence persistent events) into a unified life-cycle management strategy for CIs has been proposed. In particular, the results of the risk analysis conducted in the scope of a stress test in terms of system performance and expected costs of natural events, may be incorporated in a life-cycle cost analysis of the CI system and optimization of its operations and maintenance. Further, the evaluation of risk reduction strategies resulting from a loss disaggregation may make it possible to improve the full management and maintenance plan of the CI itself. Moreover, the evaluation of the state of civil infrastructures after the occurrence of a natural event, and the collection and processing of post-event data, such as typology, location, component's features and the assessed physical damages, can be useful to update the state condition history of the inspected components of the CI and to estimate and/or update performance prediction models used in a future risk analysis.

The CIs support the vital functions of public safety, and provide energy, water, communication and transportation services. By doing so, the CIs are an essential layer of front-line systems that support the economic functions of a community, such as employment opportunities, adequate wages and affordable housing options, as well as social functions like community ownership and participation, education and training opportunities, and a sense of community and place. Therefore, the CIs play a crucial role in enabling a community to successfully function by providing the physical foundations for much of the economic and social activities that characterize a modern society.

An extension of the ST@STREST methodology and framework to evaluate not only the vulnerability but also the resilience of CIs, i.e. "the ability to prepare and plan for, absorb, recover from and more successfully adapt to adverse events" (TNA, The National Academy 2012) has also been proposed. This extension builds on the ST@STREST methodology by modelling the post-disaster recovery process of a CI system and by quantifying the lack of resilience and the attributes of a resilient system using a novel compositional supply/demand CI resilience quantification framework. This extension enables a new role of a stress test, that of examining the ability of a community and its CIs to bounce back after a natural disaster.

This report is structured in the following way:

- in Chapter 2, the main aspects of the proposed engineering risk-based methodology for stress tests of non-nuclear CIs, ST@STREST, are presented. First, the workflow and the interaction among the main actors of the process are discussed. Then the multi-level approach and the different levels of analysis are presented. Finally, a possible grading system for quantifying the outcomes of a CI stress test is introduced. This system enables uniform grading of stress test outcomes across a broad spectrum of CIs as well as indicating how much the risk of the CI should be reduced in the next periodical verification of the CI.
- in Chapter 3 a method to integrate the results of stress tests and the data collected after disastrous events into a unified life-cycle management strategy for CIs is introduced. This method enables management of both long-term degradation and instantaneous natural hazard-induced stressors during the lifetime of a CI system.
- in Chapter 4 the link between societal resilience and the CIs of the community affected by a disaster is established first. Then, a time-varying metrics of resilience of a system is adopted in order to represent the pre-event state of the community, the phases of disaster-induced loss accumulation and absorption, followed by the recovery phase and finishing with the post-event adapted state of the community. A novel compositional supply/demand resilience quantification framework is presented.

The proposed ST@STREST methodology has been applied to six key representative Critical Infrastructures (CIs) in Europe, exposed to variant hazards, namely: a petrochemical plant in Milazzo, Italy, large dams of the Valais region in Switzerland, hydrocarbon pipelines in Turkey, 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 outcomes of these stress tests are presented in the STREST ERR5 (Pitilakis et al, 2016).

2. ST@STREST methodology for stress testing of critical non-nuclear infrastructures

The aims of the proposed methodology are to assess the performance of individual components as well as of whole CI systems with respect to extreme events, and to compare this response to acceptable values (performance objectives) that are specified at the beginning of the stress test. ST@STREST is based on probabilistic and quantitative methods for best-possible characterization of extreme scenarios and consequences (Cornell and Krawinkler, 2000; Mignan et al, 2014; 2016a).

Further, it is important to note that CIs cannot be tested using only one approach: they differ in the potential consequence of failure, the types of hazards, and the available resources for conducting the stress tests. Therefore, multiple stress test levels are proposed (Section 2.3). Each Stress Test Level (ST-L) is characterized by different focus (component or system) and by different levels of risk analysis complexity (starting from design codes and ending with state-of-the-art probabilistic risk analyses, such as cascade modelling, Mignan et al, 2016a). The selection of the appropriate Stress Test Level depends on regulatory requirements, based on the different importance of the CI, and the available human/financial resources to perform the stress test.

In order to allow transparency of the ST@STREST process, a description of the assumptions made to identify the hazard and to model the risk (consequences) and the associated frequencies is required. The data, models and methods adopted for the risk assessment and the associated uncertainties are clearly documented and managed by different experts involved in the stress test process, following a pre-defined process for managing the multiple-expert integration (Selva et al, 2015, Selva et al, in prep.). This allows defining how reliable the results of the stress test are (i.e. level of “detail and sophistication”) of the stress test (Section 2.6).

Different experts are involved in the implementation of stress test process and different roles and responsibilities are assigned to different actors, as described in Section 2.1 and Section 2.2. In particular, several participants may be involved, with different background knowledge. The size of such groups depends on selected ST-Level (see Section 2.3).

The workflow of ST@STREST comprises four phases (Fig. 2.1): Pre-Assessment phase; Assessment phase; Decision phase; and Report phase. In the Pre-Assessment phase the data available on the CI (risk context) and on the phenomena of interest (hazard context) is collected. Then, the goal, the time frame, the total costs of the stress test, and the most appropriate Stress Test Level to apply to test the CI are defined. In the Assessment phase, the stress test is performed at Component and System Levels. In the Decision phase, the stress test outcomes are checked, i.e. the results of risk assessment are compared to the objectives defined in Pre-Assessment phase. Then critical events, i.e. events that most likely cause a given level of loss, are identified and risk mitigation strategies and guidelines are formulated based on the identified critical events and presented in the Report phase.

All the aspects characterizing the ST@STREST methodology are described in the following sections, in particular:

- The use of multiple experts (Section 2.1): to guarantee the robustness of stress test results, to manage subjective decisions and quantify epistemic uncertainty.
- The workflow of the process (Section 2.2): description of the sequence of phases and steps which have to be carried out in a stress test.
- The multi-level framework (Section 2.3): the different levels of the risk analysis to test the CI response to natural hazards.
- Data structures (Section 2.4): different representations of complex systems for a probabilistic risk analysis.
- The grading system (Section 2.5): to compare the results of the risk assessment with acceptance criteria and define the outcome of the test.

- The penalty system (Section 2.6): to acknowledge the limitation of the methods and models used to assess the performance of the CI and eventually penalize the output of the risk assessment.

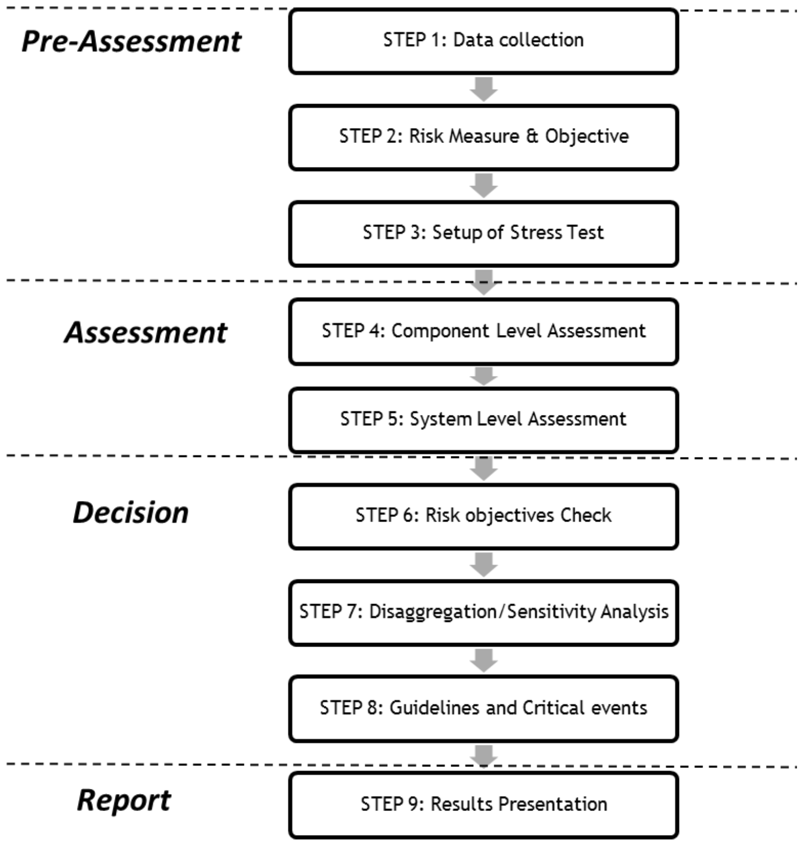

Fig. 2.1 Workflow of ST@STREST methodology

2.1 ST@STREST multiple-expert integration: EU@STREST

The involvement of multiple experts is critical in a risk assessment when potential controversies exist and the regulatory concerns are relatively high. In order to produce robust and stable results, the integration of experts plays indeed a fundamental role in managing subjective decisions and in quantifying the epistemic uncertainty capturing *'the center, the body, and the range of technical interpretations that the larger technical community would have if they were to conduct the study'* (SSHAC, 1997). To this end, the experts' diverse range of views and opinions, their active involvement, and their formal feedbacks need to be organized into a structured process ensuring transparency, accountability and independency.

EU@STREST, a formalized multiple expert integration process has been developed within STREST (Selva et al 2015, Selva et al, in prep.) and integrated into the ST@STREST Workflow (Section 2.2). This process guarantees the robustness of stress test results, considering the differences among CIs with respect to their criticality, complexity and ability to conduct hazard and risk analyses, manages subjective decision making, and enables quantification of the epistemic uncertainty. With respect to the different levels in the SSHAC process developed for nuclear critical infrastructures (SSHAC, 1997), the proposed process is located between SSHAC levels 2 and 3 in terms of expert interaction. EU@STREST also makes an extensive use of classical Expert Elicitations, and is extended to single risk and multi-risk analyses.

The core actors in the multiple expert process are the Project Manager (PM), the Technical Integrator (TI), the Evaluation Team (ET), the Pool of Experts (PoE), and the Internal Reviewers (IR). The interactions among these actors are well-defined in the process. The descriptions and the roles of these actors are given below.

- *Project Manager (PM)*: Project manager is a stakeholder who owns the problem and is responsible and accountable for the successful development of the project. It is the responsibility of the PM that his/her decisions appear rational and fair to the authorities and public. The PM specifically defines all the questions that the ST should answer.
- *Technical Integrator (TI)*: The technical integrator is an analyst responsible and accountable for the scientific management of the project. The TI is responsible for capturing the views of the informed technical community in the form of trackable opinions and community distributions, to be implemented in the hazard and risk calculations. Thus, the TI explicitly manages the integration process. The TI should have: i) expertise on managing classical Expert Elicitation (cEE), preparing questionnaires and analyzing the results in order to manage the interviews to extract the information from the larger community feedbacks regarding critical choices/issues that any test involves (e.g., the selection of appropriate scientifically acceptable models); ii) experience in hazard and risk calculations; iii) experience in expert integration techniques, in order to manage the quantification and the propagation of epistemic uncertainty out of acceptable models.
- *Evaluation Team (ET)*: The Evaluation Team is a group of analysts that actually perform the hazard, vulnerability and risk assessments required by the ST, under the guidelines provided by the TI. The team is selected by consensus between the TI and PM, and it may be formed by internal CI resources and/or external experts. In this sense, the ET represents also the interface between the project and the CI authorities, guaranteeing the successful and reciprocal acknowledgement of choices and results.
- *Pool of Experts (PoE)*: This pool is formed only if required by the ST-Level. For most ST-Ls, the role of the PoE is covered by the TI. It has the goal of representing the larger technical community within the process. Two sub-pools are foreseen, which can partially overlap: PoE-H (a pool of hazard analysts) and PoE-V (a pool of vulnerability and risk analysts). The PoE-H should have either site-specific knowledge (e.g., hazards in the area) and/or expertise on a particular methodology and/or procedure useful to the TI and the ET team in developing the community distribution regarding hazard assessments. The PoE-V should have expertise on the specific CI and/or on the typology of CI and/or on a particular methodology and/or procedure useful to the TI and the ET team regarding fragility and vulnerability assessments. Individual experts of the pool may also act as proponent and advocate a particular hypothesis or technical position, in individual communications with the TI (referring to SSHAC (1997) documents, the PoE includes both resource and proponent experts). They participate to the interviewing processes (either in remote or through specific meetings) lead by the TI as pool of experts, providing the TI for their opinions on critical choices/issues. If requested by the CI authorities or if irreconcilable disagreements among the experts of the pool emerge during the interviewing processes (in both PHASE 1 and PHASE 2 of the Workflow, see Section 2.2), the TI and PM may decide to organize meetings with the PoE (or parts of it), in order to openly discuss about controversial issues. In this case, the pool acts as a panel, and the TI is responsible for moderating the discussion.
- *Internal Reviewers (IR)*: One expert or a group of experts on subject matter under review that independently peer reviews and evaluates the work done by the TI and the ET. This group provides constructive comments and recommendations during the implementation of the project. In particular, IR reviews the coherence between TI choices and PM requests, the TI selection of the PoE in terms of expertise

coverage and scientific independence, the fairness of TI integration of PoE feedbacks, and the coherence between TI requests and ET implementations. In particular, IR reviews the project both in terms of technical and procedural aspects of the project (actor's independency, transparency, consistency with the project plan). The IR makes sure that the TI has captured the center, body and range of technically defensible interpretations when epistemic uncertainty is accounted for in the ST level. Note that the IR actively plays an important role during the project and thus is part of the project. If regulators or external authorities foresee an external review of the project results, this further review is performed independently and after the end of the project. Here, the internal review by the IR is considered essential also in this case, in order to increase the likelihood of a successful external ex post review.

The CI authorities select the PM. The PM selects the TI and IR and, jointly with the TI, the components of the ET and of the PoE. PM and TI are, in principle, individuals. The ET and IR may involve several participants, with different background knowledge, but in specific cases may be reduced to individuals. The PoE is, by definition, a group of experts. In all cases, the size of groups depends on the purpose and the given resources of the project.

The PM interacts only with the TI and specifically defines all the questions that the project should answer to, taking care of the technical and societal aspects (e.g., selection of the ST level, definition of acceptable risks, etc.). The TI coordinates the scientific process leading to answer to these questions, coordinating the ET in the implementation of the analysis, organizing the interaction with the PoE (through elicitations and individual interactions), and integrating PoE and IR feedbacks into the analysis. The ET implements the analysis, following the TI choices. The IR reviews the whole process, in order to maximize the reliability of the results and to increase their robustness. The basic interactions among the core actors are shown in Fig. 2.2.

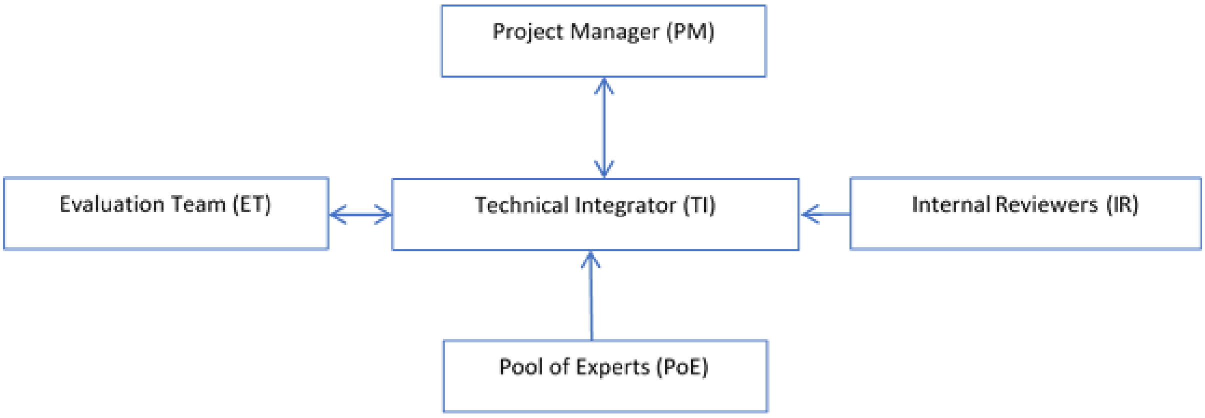

Fig. 2.2 The basic interactions among the core actors in the process of EU@STREST

2.2 ST@STREST workflow

The workflow represents a systematic sequence of steps (processes) which have to be carried out in a stress test. As mentioned before, the ST@STREST workflow comprises four phases: Pre-Assessment phase; Assessment phase; Decision phase; and Report phase. Each phase is subdivided into a number of specific steps, with a total of 9 steps.

In the Pre-Assessment phase all the data available on the CI and on the phenomena of interest (hazard context) are collected. Then, the goal (i.e. the risk measures and objectives), the time frame, the total costs of the stress test and the most appropriate Stress Test Level to apply are defined. In the Assessment phase, the stress test is performed at Component and System Levels. The performance of each component of the CI and of the whole system is checked according to the Stress Test Level selected in Phase

1. In the Decision Phase, the stress test outcomes are checked i.e. the results of risk assessment are compared to the risk objectives defined in Phase 1. Then critical events, i.e. events that most likely cause a given level of loss value are identified through a disaggregation analysis. Finally, risk mitigation strategies and guidelines are formulated based on the identified critical events. In the Reporting Phase the results are presented to CI authorities and regulators.

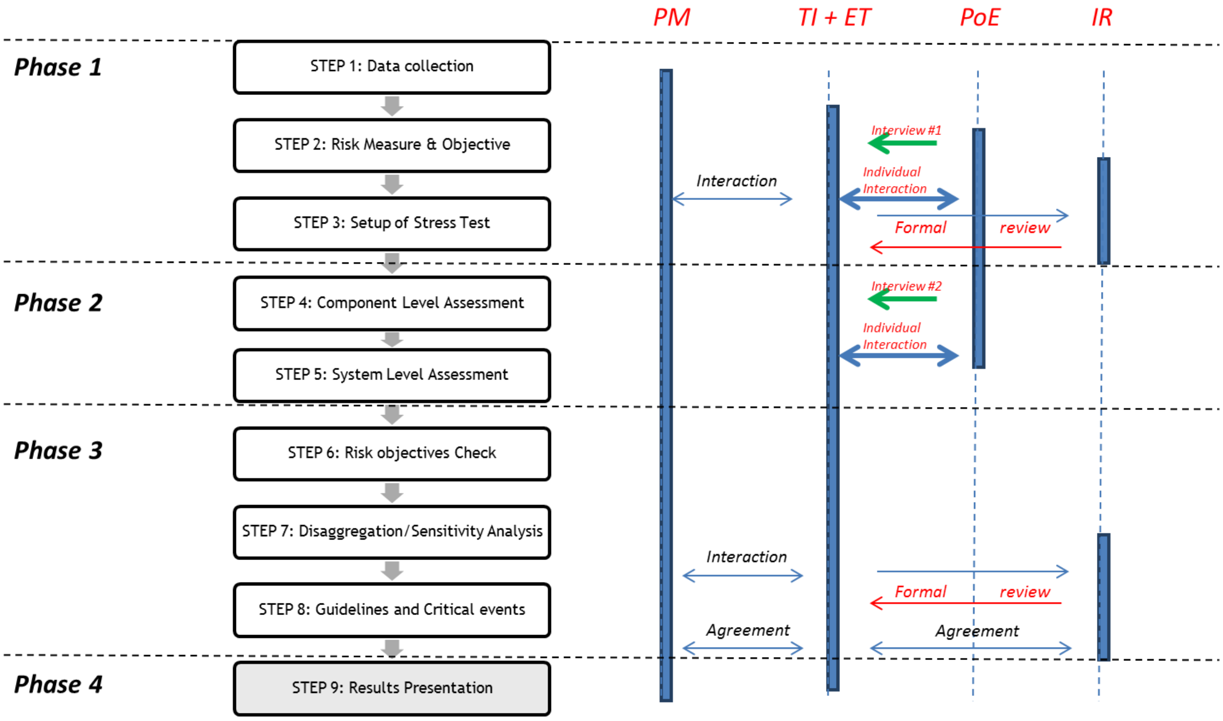

Fig. 2.3 Interaction among the main actors during the multiple-expert process EU@STREST. The PoE is present only in ST sub-levels c and d. For sub-levels a and b, the role of the PoE is assumed directly by the TI

The participation of the different actors significantly changes along the different phases of the Stress Test (Fig. 2.3). The PM and TI are the most active participants in the ST workflow. The PM participates in all the steps of the Stress Test until the end (reporting of the results), while the role of TI ends at the end of the Decision phase. The TI is constantly assisted by the ET and supported by the PM, while the level of assistance depends on the ST level. The PoE (if present, see Section 2.3) participates in the Assessment and Decision phases. The IR performs a participatory review at the end of Phase 1 and 3. The final agreement, at the end of the Decision phase, is made among the PM, TI and IR.

The workflow and the involvement of main actors and their phase-wise interactions are shown in Fig. 2.3. In the following, a detailed description of the four phases is provided together with a specification of the involvement of the different experts in process.

2.2.1 PHASE 1: Pre-Assessment phase

The Pre-Assessment phase comprises the following three steps:

- *STEP 1 Data collection*: collection of all the data available on the CI (risk context) and on the phenomena of interest (hazard context). Also data coming from Stress Tests performed on other similar CI and/or in the same area are collected. In this step, the test participants are selected: the PM selects the TI and the IR; the TI and the PM jointly select the ET. Then, the TI, with the technical assistance of the ET, collects data and relevant information about hazards and CI, and about previous Stress Tests. The TI pre-selects the potential target hazards and the relevant CI components.

- <u>*STEP 2 Risk Measures and Objectives*</u>: definition of one or more risk measures (e.g. fatalities, economic loss, etc.) and objectives (e.g. expected loss, annual probability etc.). This definition is performed by the PM, based on the regulatory requirements, the technical and societal considerations, and previous Stress Tests.
- <u>*STEP 3 Set-up of the Stress Test*</u>: selection of the Stress Test Level, and Timing and Costs of the project and definition of the "level of detail and sophistication" used for the computation of the assessment phase, as presented in Section 2.6. The selection of the ST-Level is made by the PM with the assistance of the TI, based on the regulatory requirements.

STEP 3 may be a long process and may differ substantially depending if the PoE is in place or not, according to the ST-L selected. The presence of the PoE allows for a robust set-up of the ST, based on the quantitative feedbacks of multiple experts. In this case, the PM and TI set an initial costs and timeframe for the assessments to be performed in STEP 3. The TI selects the PoE and organizes a one-day kick-off meeting with PoE, ET, and PM. With the assistance of PoE, through classical Expert Elicitation, the TI selects the target single and multiple hazards and the relevant CI components and their interactions. If significant disagreements emerge from the elicitation result, the TI may promote specific topical discussions among the members of the PoE, enabling a final decision. Based on this selection, the TI and PM integrate the ET and the PoE to have a complete coverage of the required expertise. The TI collects applicable scientific models and data needed for hazard, vulnerability and risk assessment, with the technical assistance of the ET (and through potential individual interaction with the PoE, if required). At this stage, also potential lacks in modelling procedures are identified by the TI. If technically possible, such lacks should be filled by the TI based on quantification through classical Expert Elicitation of the PoE, which is at this point planned for PHASE 2. Otherwise, a complementary scenario-based assessment should be planned (see Section 2.3). The specification of this scenario-based assessment (e.g., the definition of scenarios to be considered) is made through a specific classical Expert Elicitation planned for PHASE 2.

To complete the planning of actions in PHASE 2, the TI also plans the classical Expert Elicitation of the PoE for ranking alternative models to be used in the stress test, in order to enable the quantification of epistemic uncertainty.

If the selected ST-Level does not foresee the presence of the PoE, this process becomes simplified since all critical decisions are taken directly by one single expert, the TI. The TI selects the target hazards and the relevant CI components. Based on this selection, the TI and PM integrate the ET, to have a complete coverage of the required expertise.

In either case, at the end of these basic choices, the TI collects applicable scientific models and data needed for hazard, vulnerability and risk assessment, with the technical assistance of the ET. Based on this collection, the TI and PM jointly identify the "level of detail and sophistication" used for the computation of the assessment phase (see Section 2.6) based on target costs and model availability. As mentioned above, one of the main goal of this assessment phase is to capture the center and range of technical interpretations that the larger technical community would have if they were to conduct the study. A preliminary sensitivity analysis may help to identify the key parameters which controls the results in order to focus the uncertainty analysis and experts discussions on these key inputs.

All decisions/definitions are specifically documented by the TI. The IR reviews such documents and provides his/her feedbacks regarding the decisions/definitions made thus far. The PM and TI finalize all documents, based on this review. At this point, the final costs and the exact timing for PHASE 2 and PHASE 3 are established. Further, based on the IR review, the PM and TI may evaluate potential changes to the analysis implementation along the assessment phase, in order to avoid potential penalties suggested by the reviewers. In fact, in the case the "level of detail and sophistication" reached in the final implementation is lower than the *level* required, a Penalty System is applied to the output of the risk assessment (STEP 6 *Risk objectives Check).*

2.2.2 PHASE 2: Assessment phase

The Assessment phase is characterized by two steps in which the stress test is performed at Component and System levels according to the Stress Test Level selected in Phase 1. In particular:

- *STEP 4 Component Level Assessment*: the performance of each component of the CI is checked by the hazard-based assessment, design-based assessment or risk-based assessment approach (see Section 2.3). This check is performed by the TI or by one expert of the ET selected by the TI.
- *STEP 5 System Level Assessment*: the stress test at the system level is performed. At first, the TI finalizes all the required models. In particular, if the PoE is in place (sub-levels c), the TI organizes the classical Expert Elicitations in order to: i) fill potential methodological gaps, ii) quantify the potential scenario for the scenario-based risk assessment (SBRA), and iii) rank the alternative models to enable the quantification of the epistemic uncertainty. The PoE performs the elicitation remotely. Open discussions among the PoE members (moderated by the TI) are foreseen only if significant disagreements emerge in the elicitation results. If the PoE is not in place but EU assessment is required (sub-level b), the TI directly assigns scores on the selected models for ranking. Then, the ET (coordinated by the TI) actually implements all the required models and performs the assessment. If specific technical problems emerge during the implementation and application, TI may solve them through individual interactions with members of the PoE (if foreseen at the ST-Level).

2.2.3 PHASE 3: Decision phase

The Decision Phase is characterized by three steps:

- *STEP 6 Risk objectives Check*: comparison of results of the Assessment phase to the risk objectives. This task is performed by the TI, with the technical assistance of the ET. Depending on the type of risk measures and objectives defined by the PM (F-N curve, expected value, etc.) and on the level of "detail and sophistication" adopted to capture the center and range of technical interpretations, the comparison between results from probabilistic risk assessment with these goals may differ (see Section 2.6). One possibility to assess the difference between the obtained risk measures and the adopted risk objectives is presented in Section 2.5 where the outcome of the stress test is presented by grades (e.g. AA – negligible risk, A – as low as reasonably practicable (ALARP) risk, B – possibly unjustifiable risk, C – intolerable risk).
- *STEP 7 Disaggregation/Sensitivity Analysis*: identification of critical events. This task is performed by the ET coordinated by the TI. Critical events that most likely cause the exceedance of the considered loss value are identified through a disaggregation analysis1 (Esposito et al 2016) and based on them, risk mitigation strategies and guidelines are then formulated. If specific technical problems emerge during the application, the TI may solve them through individual interactions with the PoE (if present). This step is not mandatory. It depends on the results of STEP

[1] See Appendix B, Deliverable 5.1 (Esposito et al, 2016).

6 (Risk objectives Check). For example, if the outcome of STEP 6 is that the critical infrastructure passes the stress tests, performing STEP 7 may be informative, but is not required.

- *STEP 8 Guidelines and Critical events*: risk mitigation strategies and guidelines are formulated based on the identified critical events. This task is performed by the TI, with the technical assistance of the ET.

All the results in all the steps of PHASE 2 and PHASE 3 are specifically documented by the TI. The IR reviews the activities performed in assessments from STEP 4 to STEP 8. The TI, with the technical assistance of the ET, update to the final assessments for such steps accounting for the review. Final assessments and decisions are documented by the TI. Based on such documents, the PM, TI and IR reach the final agreement.

2.2.4 PHASE 4: Report phase

The Report phase comprises one step:

- *STEP 9 Results Presentation*: presentation of the outcome of stress test to CI authorities, regulators and community representatives. This presentation is organized and performed by PM and TI. The presentation includes the outcome of stress test in terms of the grade, the critical events, the guidelines for risk mitigation, and the level of "detail and sophistication" of the methods adopted in the stress test.

Note that the time for this presentation is set in PHASE 1, and it cannot be changed during PHASE 2 and 3.

2.3 ST@STREST test levels

Due to the diversity of types of CIs and the potential consequence of failure of the CIs, the types of hazards and the available resources for conducting the stress tests, it is not optimal to require the most general form of the stress test for all possible situations. Therefore, three stress test variants, termed Stress Test Levels (ST-Ls) are proposed:

- Level 1 (ST-L1): single-hazard component check;
- Level 2 (ST-L2): single-hazard system-wide risk assessment;
- Level 3 (ST-L3): multi-hazard system-wide risk assessment.

Each ST-L is characterized by a different scope (component or system) and by a different complexity of the risk analysis (e.g. the consideration of multi-hazard and multi-risk events) as shown in Fig. 2.4.

The aim of the ST-L1 (Component Level Assessment) is to check each component of a CI independently in order to show whether the component passes or fails the minimum requirements for its performance, which are defined in current design codes. The performance of each component of the CI is checked for the hazards selected as the most important (e.g. earthquake or flood). At component-level there are three methods to perform the single-hazard component check. These methods differ for the complexity and the data needed for the computation. The possible approaches are: the hazard-based assessment, design-based assessment, and the risk-based assessment approach.

Since a CI is a system of interacting components, ST-L1 is inherently not adequate. Nevertheless, ST-L1 is obligatory because design of (most) CI components is regulated by design codes, and the data and the expertise are available. Further, for some CIs, the computation of system-level analysis (single- and multi-risk) could be overly demanding in terms of available knowledge and resources.

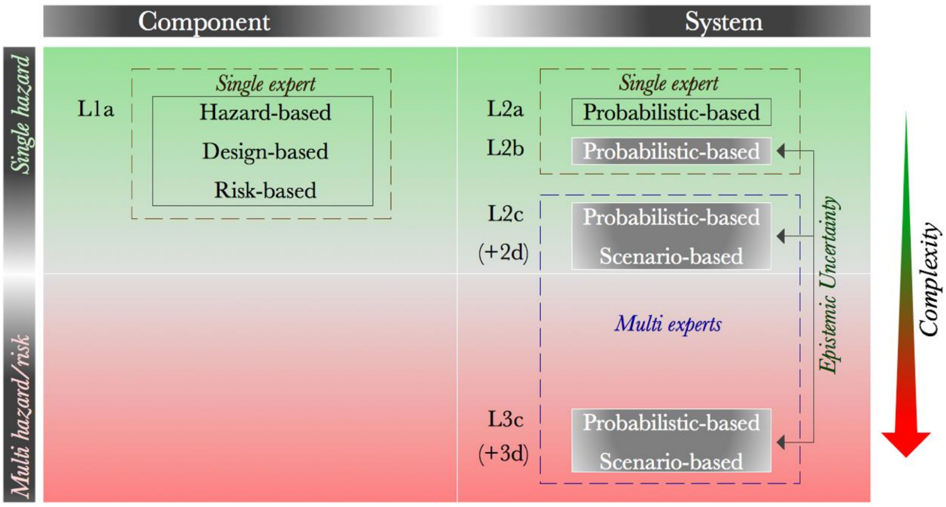

Fig. 2.4 ST-Levels in the ST@STREST methodology

The stress test assessment at the system level of the CI is foreseen at ST-L2 or ST-L3 where the probabilistic risk analysis of the entire CI (system) is performed. The system level assessment is highly recommended, since it is the only way of revealing the majority of the mechanisms leading to potential unwanted consequences. However, note that it requires more knowledge and resources (e.g. financial, staff) for conducting the stress test, thus it is not made obligatory (if not required by regulations).

At these 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). Indeed, a more accurate quantification of the technical community knowledge distribution (describing the epistemic uncertainty) can be reached if more experts are involved in the analysis and, in particular, when dealing with all critical choices.

Further, in case specific needs have been identified in the Pre-assessment phase (e.g. important methodological/modelling gaps) and such requirements cannot be included into the risk assessment for whatever reason, scenario-based analysis should be also performed as complementary to the system ST-L selected (sub-level d). Levels 2d and 3d are complementary to L2c and L3c, respectively. In this case, multiple experts define and evaluate possible scenarios that, for whatever reason, cannot be included into probabilistic risk analysis. In this case, the choice of performing a scenario-based assessment should be justified and documented by the TI, and reviewed by the IR. If scenario-based assessment is finally selected, the choice of the scenarios should be based on ad hoc expert elicitation experiments of the PoE (SSHAC 1997). These additional scenarios are meant to further investigate the epistemic uncertainty by including events otherwise neglected only for technical reasons. Indeed, L2d and L3d are performed to evaluate the potential impact of epistemic gaps identified by experts, eventually increasing the capability of exploring the effective epistemic uncertainty. Thus, it is foreseen only as complementary to a full quantification of epistemic uncertainty in a multiple-expert framework.

The system level analysis is thus performed according to: 1) the degree of complexity of the analysis (single vs. multi hazards), and 2) the degree of involvement of the technical community in taking critical decisions and in the quantification of the epistemic uncertainty for the computation of risk. According to these two aspects a subdivision for ST levels has been introduced (Fig. 2.4). The selection of the actual procedure to be implemented is performed in the Pre-Assessment (Phase 1). These two choices essentially depend on regulatory requirements, on the different importance of the CI, and on the available human/financial resources to perform the stress test. A practical tool to support the choice of the appropriate ST level may be a criticality assessment aimed at identifying and ranking

CIs (for example at a national scale). In Esposito et al 2016 (Appendix A) some key factors that may be considered to define the criticality of the CIs and a possible methodology to rank CIs are presented and discussed.

In the following, a specific description of all ST-Ls and sub-levels is reported.

2.3.1 Component level assessment

At Component Level Assessment only one implementation is foreseen, i.e. the ST-L1a. This level requires less knowledge and resources (financial, staff, experts) for conducting the stress test in comparison to the system level assessment, but it is obligatory because design of (most) CI components is regulated by design codes, and usually, both the data and the experts are available. Further, for some CIs, the computation of system-level analysis (single- and multi-risk) could be overly demanding in terms of available knowledge and resources.

Only the TI is required as expert contributing to critical scientific decision, while the whole process may require up to five experts to assist the TI in technical decisions. The TI selects the most important hazard to consider in the component-level analysis but, if more than one hazard is considered critical for the CI under study, more than one Level 1 check should be performed, one for each hazard.

Three methods to perform the single-hazard component check are proposed in ST@STREST, and they differ for the complexity and the data needed for the computation. The possible approaches are: the hazard-based assessment, design-based assessment and the risk-based assessment approach. A detailed description is provided in the following.

The main aspects characterizing the ST-L1a are summarized in Table 2.1.

Hazard-based assessment: The performance of the component is checked by comparing the design value of intensity of the hazard which was actually used in the design of the component (building, pipeline, storage tank, etc.), $I_{Design\ phase}$, to the design value of intensity of the hazard prescribed in current regulatory documents or to the value of intensity according to the best possible knowledge, $I_{Assessment\ phase}$. The complexity of such an assessment phase is not high. As a consequence, the level of "detail and sophistication" of this type of assessment is considered moderate, since all other design factors (e.g. minimum requirements for detailing, material safety factors, design procedures, type of analysis, safety margin) and their impact on the performance of the components, which can also change from different versions of regulatory documents, are neglected in the assessment. The outcome of this type of assessment phase is qualitative:

- In compliance with the design level of hazard ($I_{Assessment\ phase} \leq I_{Design\ phase}$);
- Not in compliance with the design level of hazard ($I_{Assessment\ phase} \geq I_{Design\ phase}$);
- The design level of hazard is unknown. This outcome is assigned when there is no regulatory document which would require design of the component for considered type of hazard at the time of performing the stress test.

Hazard-based assessment may be used in cases when the component has not been designed using modern design codes and when the component is not significant for the system response. In such cases, the target level of detail is expected to be set to Moderate (see Section 2.6), which would allow the method to be used. However, if the target level is set to High or Advanced, a more accurate method should be used (i.e. design- or risk-based assessment, respectively) to evaluate the components and avoid imposition of penalty factors. Moreover, due to the trend of increasing design levels of hazard over time, the outcome of the hazard-based assessment is expected, in a vast majority of cases, to be "Not in compliance with the design level of hazard", which would, again, require a more accurate method to be utilized.

Design-based assessment: The level of "detail and sophistication" of this type of assessment is higher than the previous method since it is based on the design state-of-practice. The expert compares the demand, *D*, with the capacity, *C*, (expressed in terms of forces, stresses, deformations or displacements). The assessment can be based on factoring the results from the existing design documentation or by performing design (assessment) of the component according to the current state-of-practice. The decision-making regarding the sufficiency of the investigated component is sometimes difficult, since the demand in the design is most often based on linear-elastic analysis while the performance objectives of the component are often associated with its nonlinear behavior. Alternatively, the performance assessment can be based on nonlinear methods of analysis. The complexity of this type of assessment may differ, depending on the type of analysis (linear, nonlinear) used. The outcome of the design-based assessment is qualitative:

- In compliance with the code ($D \leq C$);
- Not in compliance with the code ($D \geq C$);
- The design objectives for this type of hazard are not defined. This outcome is assigned when there is no regulatory document which would require design of thecomponent for considered type of hazard at the time of performing the stress test.

Risk-based assessment: The hazard function at the location of the component and the fragility function of the component are required for this type of performance assessment. The level of "detail and sophistication" of this type of assessment varies from Moderate to Advanced, which depends on the level used for evaluation of the hazard function and the fragility function. These two functions can then be convolved in the risk integral in order to obtain the probability of exceedance of a designated limit state in a period of time (P_{LS}). In general, the risk integral can be solved numerically. Under some conditions, simple closed-form solutions of risk integral also exist. The target probability of exceedance of a designated limit state for a period of time ($P_{LS,t}$) also has to be defined for each component and different limit states (e.g. loss of function, low/medium/high damage, collapse) if they are considered in this assessment (e.g. probability of exceedance implied by the code). The complexity of risk-based assessment is in general high, but it can be reduced to low when the hazard and fragility functions are already available. Such situation occurs if the ST-L2 or ST-L3 assessments are also foreseen in the stress test. In this case the ST-L1 assessment and system level assessment should be partly performed in parallel. The outcome of the risk-based assessment is quantitative, since the performance of the component is measured by the estimated P_{LS}, which is then used as a basis for the grading (see *Section 2.5*).

Table 2.1 Main aspects characterizing the Component Level Assessment (STL-1a)

Level	**ST-L1a**
Events considered	Single hazard check. Hazard selected as the most important (e.g. earthquake or flood, etc.). If more than one hazard is important, more than one Level 1 check should be performed.
Number of experts contributing to critical scientific decisions	1 (the TI)
Total number of experts involved in the process	< 5 (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI, and an IR with 1 expert)
Method:	The performance of each component of the CI is checked using the hazard-based assessment, design-based assessment or risk-based assessment approach. Design-based assessment is recommended when only ST-L1 is performed. In the case when ST-L1 is followed by ST-L2, in which component-specific fragility functions are used, it makes sense to perform risk-based assessment of the components since fragility function are anyway required in ST-L2.
Core actors	PM, TI + ET, IR

The three methods described above for a single hazard check at component level assessment is demonstrated by means of an example of a precast reinforced concrete industrial building (single-storey precast reinforced concrete building with masonry infills on the perimeter, Fig. 2.5) located in Ljubljana (Slovenia).

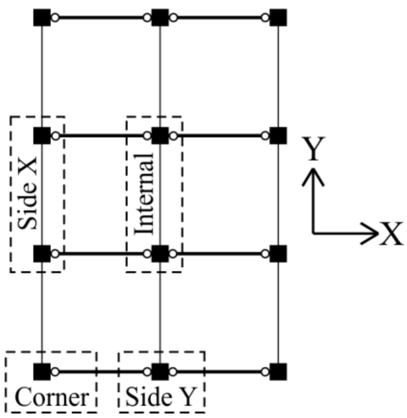

Fig. 2.5 Plan view of the case study building from ST-L1 assessment

The structure was designed before the introduction of the Eurocode standards and it consists of cantilever columns, which are connected by an assembly of roof elements. It has two bays in the X direction and three bays in the Y direction. The distance between the columns in the X and Y directions are 17.4 m and 8.7 m, respectively, whereas the height of the columns amounts to 10.3 m. The critical components of the building are columns and beam-to-column connections. The ratios of longitudinal and transverse reinforcement in all columns amount to 1.29 % and 0.10 %, respectively. No connections between beams and columns are provided. The total mass of the structure amounts to 237 t.

The design peak ground acceleration for the 475 and 2475 year earthquakes amount to 0.25 g and 0.35 g, respectively. The ground is classified as B (CEN, 2005a).

Hazard-based assessment

No information on the hazard used in the design of the component is available. It is therefore concluded that the component was not designed to withstand seismic loading. The outcome of the hazard-based assessment is: "The design level of hazard is unknown".

Design-based assessment for Limit State of Near Collapse

Eurocode 8 Part 3 (CEN, 2005b) is used to conduct the design-based assessment of the component. The knowledge level, as required by the code, is identified as »limited«. Consequently, confidence factor amounts to 1.35. The limit state of Near Collapse, which corresponds to the return period of 2475 years, is checked. Lateral force analysis is selected.

Table 2.2 summarizes the results of the assessment in terms of verification of beam-to-column connections. As shown in the table, connections above the corner columns do not meet requirements, which means that the building does not comply with the code. Further assessment of the columns is not performed since it can be concluded that the outcome of the design-based assessment is: "Not in compliance with the code (D ≥ C)".

The detailed explanation of the calculations are reported in Deliverable 5.1 (Esposito et al, 2016).

Table 2.2 Verification of beam-to-column connections

Column	Capacity of columns in terms of M_{Rd} [kNm] (material characteristics are multiplied by the confidence factor)	D_{conn} [kN]	C_{conn} [kN] (material characteristics are divided by the confidence factor)
Internal	527	26	143
Side X	487	47	72
Side Y	487	24	36
Corner	466	45	36

Risk-based assessment

Fragility function for the collapse limit state and hazard function are required in order to estimate the risk. In this case fragility function is determined by conducting non-linear dynamic analyses using a set of hazard consisting ground motions. The numerical model of the building is defined using the principles described in Babič and Dolšek (2016) and Crowley et al (2015). Based on the results of the numerical simulations a regression analysis is carried out by assuming a lognormal distribution and by using the maximum likelihood method as proposed in previous studies (e.g. Baker, 2015). The geometric mean of the spectral accelerations in both horizontal components at 1.9 s is chosen as an intensity measure. The parameters of the resulting fragility function (Fig. 2.6a), i.e. the median $\overline{IM_C}$ and the standard deviation in the log domain β_C, are 0.22 g and 0.40, respectively.

The seismic hazard curve (Fig. 2.6b) is determined based on the probabilistic seismic hazard analysis (PSHA) used for the development of seismic hazard maps in Slovenia (Lapajne et al 2003). It is idealized by a linear function on a log-log plot, expressed as:

$$\mathrm{H}(IM) = k_0 \cdot IM^{-k} \tag{2.1}$$

Interval from $0.25\cdot\overline{IM}$ and $1.25\cdot\overline{IM}$ was chosen for the idealization of the hazard curve, as proposed by Dolšek and Fajfar (2008). The parameters of the idealized hazard curve $k_{0,C}$ and k_C amount to $4.8\cdot10^{-5}$ and 1.75, respectively. The resulting probability of exceedance of the collapse limit state is determined as follows:

$$P_C = H(\overline{IM_C})\ \exp(0.5\ k_C^{\ 2}\beta_C^{\ 2}) = 8.5 \cdot 10^{-4} \tag{2.2}$$

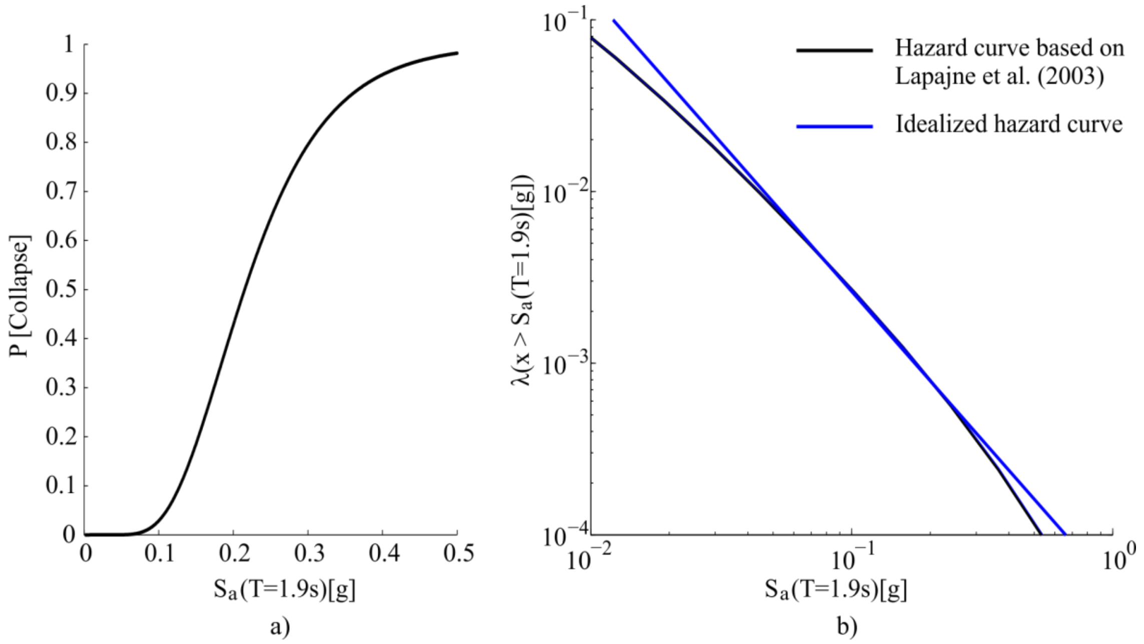

Fig. 2.6 a) Fragility function of the building and b) seismic hazard on the location of the building

2.3.2 System level assessment

The system level assessment requires more knowledge and resources for conducting the stress test compared to the Component Level Assessment. Thus, it is not made obligatory. However, the system level assessment represents the only way of revealing the paths that lead to potential unwanted consequences. Therefore, it is highly recommended.

Different implementations are possible, according to:

- The consideration of a single hazard (STL-2) or of multiple-hazard/risks (STL-3).
- The quantification of epistemic uncertainty may not be performed (sub-level a).
- The use of a single expert (sub-level b) or of multiple-experts (sub-level c) to quantify the epistemic uncertainty.

Single hazard (ST-L2)

For the single hazard system level check, three sublevels are foreseen according to the degree of involvement of the technical community in taking critical decisions and in the quantification of the Epistemic Uncertainty (EU) for the computation of risk. The quantification of EU may not be performed (ST-L2a). If performed, it may be either based on the evaluations of a single expert (ST-L2b) or of multiple-experts (ST-L2c).

As for ST-L1a for the ST-L2a, only the TI is required as expert contributing to the critical scientific decision, while the whole process may require up to five experts to assist the TI in technical decisions. ST-L2b, instead, requires the use of up to nine experts (the ET formed by few individuals internal to the CI and a few external experts, and an IR with more than one expert) to assist the TI. The ST-L2c requires even more knowledge and resources. In this case more than six experts are required to contribute to scientific decisions (the TI and a PoE formed by at least six experts), while the whole process may require more than ten experts.

Regarding the methods to apply for the risk analysis, for all the sublevels the aim is to evaluate the performance of the whole CI. In a generic format, the process is independent of the field of application. The process can be divided into the following steps (AS/NZS 4360): definition of context, definition of system, hazard identification, analysis of consequences and analysis of probability (or frequency), risk assessment and risk treatment. Several methods/techniques exist for each of the main steps. They can be classified as qualitative or quantitative (Faber and Stewart, 2003). A list of some of the methods usually applied in risk assessment of engineering facilities is provided in Esposito et al, 2016.

In ST@STREST, for all sub-levels of the system-level assessment, probabilistic (i.e. probabilistic risk analysis, PRA) methods are foreseen. PRA is a systematic and comprehensive methodology to evaluate risks associated with every life-cycle aspect of a complex engineered entity, where the severity of consequence(s) and their likelihood of occurrence are both expressed qualitatively (Bedford and Cooke, 2001). It can be also found in the literature under the names of quantitative risk assessment (QRA) or probabilistic safety assessment (PSA).

The final result of a PRA is a risk curve and the associated uncertainties (aleatory and epistemic). The risk curve generally represents the frequency of exceeding a consequence value as a function of that consequence values. PRA can be performed for internal initiating events (e.g. system or operator errors) as well as for external initiating events (e.g. natural hazards).

Main applications of PRA have been performed in different fields such as civil, aeronautic, nuclear, and chemical engineering. The specific quantitative method to use depends upon the context in which the risk is placed (the hazard context), and upon the system under consideration. In civil engineering, PRA methods were developed for the analysis of structural reliability, using analytical or numerical integration, simulation, moment-based methods, or first- and second-order methods (FORM/SORM). In earthquake engineering, the state of the art of probabilistic and quantitative approaches for the estimation of seismic risk relies on performance-based earthquake engineering (PBEE). PBEE is the framework that enables engineers to assess if a new or an existing structure is adequate in the sense that it performs as desired at various levels of seismic excitation. Different analytical approaches to PBEE have been developed in the last years: the approach pursued by the Pacific Earthquake Engineering Research (PEER) Center is the most representative (Cornell and Krawinkler, 2000). This approach was originally developed for buildings (i.e., point-like structures). However, in the years, a significant body of research was developed focusing on risk assessment of infrastructure systems. PBEE was extended to spatially distributed systems such as gas or electric networks (Esposito et al, 2015; Cavalieri et al, 2014), transportation networks (Argyroudis et al, 2015), and telecommunication networks (Esposito et al, in prep.).

For the three CI classes identified in STREST and for the specific hazard considered, the detail list and explanation of possible methods that may be applied to assess the performance and the risk of the CI, are provided in STREST Deliverable 4.1 (Salzano et al, 2016), 4.2 (Kakderi et al, 2015) and 4.3 (Crowley et al, 2015).

Epistemic uncertainties are treated only at ST-L2b and ST-L2c. The goal is the assessment of the "community distribution", that is, a distribution describing "the center, the body, and the range of technical interpretations that the larger technical community would have if they were to conduct the study" (SSHAC). Here, "community distribution" means "the probability distribution representing the epistemic uncertainty within the community". This assessment goal is achieved by: selecting a number of appropriate alternative scientifically acceptable models, and weighting them according to their subjective credibility. The selection of models may be based on the development of Alternative Trees (more details can be found in Deliverable 3.1, Selva et al 2015), where the analysis is divided into a number of consecutive steps, and alternative models are defined at each step. The procedure to be followed in these tasks is different for ST-L2b and ST-L2c. In ST-L2b, the

TI (supported by the ET) selects the models based on a literature review, and assigns the weights to each one of them. At ST-L2c, a more robust procedure is foreseen (see STREST Deliverable 3.1, Selva et al 2015). In ST@STREST PHASE 1 (pre-assessment), a preliminary list of models is prepared by the TI (supported by the ET), which is formally screened by the PoE and reviewed by the IR. Then, at the beginning of ST@STREST PHASE 2 (assessment), an expert elicitation of the PoE is organized by the TI to assign the weights of the models (for example, following an AHP procedure, see STREST Deliverable 3.1, Selva et al 2015). Then, the ET implements models and weights in order to produce the "community distribution", implementing methodologies like the Logic Tree (e.g., Bommer and Scherbaum, 2008) or the Ensemble Modelling (Marzocchi et al 2015). Note that the selection of the models depends on the adopted strategy for their integration (see STREST Deliverable 3.1, Selva et al, 2015). For example, Logic Trees require that models form a MECE (Mutually Exclusive and Collectively Exhaustive) set, while Ensemble Modelling simply requires that models form an unbiased set of alternatives representing the epistemic uncertainty into the community.

The main aspects characterizing each sub level of the ST-L2 are summarized in Tables 2.3-2.5.

Table 2.3 Main aspects characterizing the System Level Assessment, STL-2a

Level	**ST-L2a**
Events considered	Single hazard, selected as the most important (e.g. earthquake, flood)
Number of experts contributing to critical scientific decisions	1 (the TI)
Total number of experts involved in the process	Up to 5 (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI, and an IR with 1 expert)
Method:	Probabilistic Risk Analysis (PRA, e.g. PBEE framework for seismic hazard)
Core actors	PM, TI + ET, IR

Table 2.4 Main aspects characterizing the System Level Assessment, ST-L2b.

Level	**ST-L2b**
Events considered	Single hazard, selected as the most important (e.g. earthquake, flood)
Number of experts contributing to critical scientific decisions	1 (the TI)
Total number of experts involved in the process	Up to 10 (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI and a few external experts, and an IR with > 1 experts)
Method:	Probabilistic Risk Analysis (PRA, e.g. PBEE framework for seismic hazard) + epistemic uncertainty
Core actors	PM, TI + ET, IR

Table 2.5 Main aspects characterizing the System Level Assessment, ST-L2c.

Level	**ST-L2c**
Events considered	Single hazard, selected as the most important (e.g earthquake, flood)
Number of experts contributing to critical scientific decisions	> 6 (the TI and a PoE formed by > 5 experts)
Total number of experts involved in the process	> 10 (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI and a few external experts, the PoE formed by > 5 experts, and an IR with > 1 experts)
Method:	Probabilistic Risk Analysis (PRA, e.g. PBEE framework for seismic hazard) + epistemic uncertainty
Core actors	PM, TI + ET, PoE , IR

Multiple hazards/risks

As for the ST-L2c, the assessment process requires more than six experts to contribute to scientific decisions (the TI and a PoE formed by at least six experts), and a total of more than ten experts (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI and a few external experts, the PoE formed by more than five experts, and an IR with more than one expert) to complete the whole process.

There is no standard approach for multi-risk assessment. Different methods could be used, taken from the scientific literature. For example, Liu et al (2015) could be used to identify the multi-risk assessment level required (semi-quantitative vs. quantitative); Marzocchi et al (2012) combined with Selva (2013) could be used when the number of interactions at the hazard and/or risk levels remains limited, and Mignan et al (2014; 2016a) could be used when the number of interactions becomes significant (roughly more than 3-4 domino effects). The Bayesian approach of Marzocchi et al (2012) has the advantage of coupling to the PEER PBEE method (Cornell and Krawinkler 2000; Der Kiureghian 2005), which is already well known to the seismic engineers and relates to the lower stress test levels (L1 and L2). Selva (2013) proposed a method to test potential individual interactions at the risk level (in vulnerability and/or exposure), and to eventually include them into the assessment using the PEER PBEE formula. Moreover, other PEER PBEE-based methods, such as damage-dependent vulnerability methods (Iervolino et al, 2016) and loss disaggregation, can easily be added to such a general multi-risk framework. The Generic Multi-Risk (GenMR) framework developed by Mignan et al (2014), on the other hand, is purely stochastic (a variant of a Markov Chain Monte Carlo method) and not derived from existing single-risk assessment approaches. It is therefore more flexible when including a multitude of perils (i.e., it is not earthquake-focused) but at the same time, requires some adaptation from the modeler to develop a multi-risk model on GenMR (i.e., all events defined in a stochastic event set, all interactions defined in a hazard correlation matrix, process memory defined from time-dependent or event-dependent variables). While GenMR could be used for a seismic multi-risk analysis (see Mignan et al, 2015), advantages become more obvious in more complex cases, such as interactions between different hazards (e.g., earthquake, flooding, erosion) and different infrastructure elements (e.g., hydropower, spillway and bottom outlet failures) at a hydropower dam (Matos et al, 2015; Mignan et al, 2015). Whatever the method used, the final output should be a probabilistic risk result in the form of probabilities of exceeding different loss levels, a risk or a loss curve. The multi-risk loss curves shall then be compared to the ones generated in stress test levels L1 and L2, and differences identified. The main cause of risk should be investigated, by disaggregation (e.g., Iervolino et al, 2016) or by GenMR time series ranking and metadata analysis (Mignan et al, 2014; Matos et al, 2015; Mignan et al, 2015; 2016a).

The treatment of EUs in ST-L3c is similar to the one described for ST-L2c. In addition, in ST@STREST PHASE 1, it is foreseen that the selection of the hazards and hazard interactions to be included is based on the results of an expert elicitation procedure of the PoE (for example, based on a qualitative risk analysis made through verbal scale, see the case study of the Harbor facilities of Thessaloniki in STREST ERR5, Pitilakis et al, 2016).

The main aspects characterizing the ST-L3 are summarized in Table 2.6.

Table 2.6 Main aspects characterizing the System Level Assessment, ST-L3c.

Level	**ST-L3c**
Events considered	Multi-hazards (multi-hazard, i.e. coinciding events and multi-risk)
Number of experts contributing to critical scientific decisions	> 6 (the TI and a PoE formed by > 5 experts)
Total number of experts involved in the process	> 10 (the TI, along with the technical assistance of the ET formed by few individuals internal to the CI and a few external experts, the PoE formed by > 5 experts, and an IR with > 1 experts)
Method:	Multi-risk analysis (extension of PRA methodology for multi-risk) + epistemic uncertainty
Core actors	PM, TI + ET, PoE , IR

Scenario-based assessment

Scenario-based analysis may be performed as complementary to ST-L2c and ST-L3c due to methodological gaps identified for specific events/hazards that cannot be formally included into the PRA. This means that it should be considered only if, for technical reasons, one important phenomenon cannot be included into a formal probabilistic framework (e.g., PRA for ST-L2c). In this case, the choice of performing a scenario-based assessment should be justified and documented by the TI, and reviewed by the IR. If scenario-based assessment is finally selected, the choice of the scenarios should be based on ad-hoc expert elicitation experiments of the PoE (see Selva et al, 2015).

Different strategies can be adopted in organizing the elicitation experiment and in preparing the documentation for the PoE. For example, the hazard correlation matrix (HCM), one of the main inputs to the GenMR framework (see above), can also be used qualitatively to build more or less complex scenarios of cascading hazardous events. The HCM is a square matrix with trigger events defined in rows and target events (the same list of events) in columns. In ST-L3c, each cell of the HCM is defined as a conditional probability of occurrence. In a deterministic view, cells can be filled by plus "+" signs for positive interactions (triggering), minus "-" signs for negative interactions (inhibiting) and empty "Ø" signs for no known interactions (supposedly independent events). The HCM has recently been shown to be a cognitive tool that promotes transformative learning on extreme event cascading. In other words, it allows defining more or less complex scenarios from the association of simple one-to-one interacting couples. Once the modus operandi is understood, more knowledge on multi-risk can be generated (Mignan et al, 2016b). The core actors could use the HCM tool to define the list of relevant events as well as to discuss the space of possible interactions in an intuitive interactive way. ST-L3d scenarios would then emerge from the HCM tool.

The main aspects characterizing the scenario-based assessment are summarized in Table 2.7.

Table 2.7 Main aspects characterizing the complementary scenario-based assessment

Level	ST-L2d ST-L3d
Events considered	"Black swan", i.e. events not previously considered (e.g. multi-hazard, correlated events) and for which a PRA is not feasible due to lack of procedures and basic knowledge. This possibility should be confirmed by the IR (Internal Reviewers). The PoE (Pool of Experts) is asked to define such scenarios.
Number of experts contributing to critical scientific decisions	Same of ST-L2c or ST-L3c.
Total number of experts involved in the process	Same of ST-L2c or ST-L3c.
Method:	Scenario-based risk assessment (SBRA)
Core actors	PM, TI + ET, PoE , IR

2.4 ST@STREST data structures

A CI is a complex assembly of components, structures and systems designed to provide a service, in terms of generation and flow of water, electric power, natural gas, oil, or goods in the scope of the built environment of a community. The data on the components, structures and systems of the CI needs to be assembled and held in a framework to facilitate the application of the proposed stress test methodology and the execution of a stress test. The data on the CI includes not only the information about the hazard and the vulnerability of the components and structures, but also the information about the functioning of the system that includes the topology of the system, the links that describe the interactions between the components and structures, and the causal relations between the events in the system.

Representation of complex systems for a probabilistic risk analysis in general, and accident sequence investigation in particular, has been done since the early 1970's in the nuclear industry. There, the event and fault trees are used to represent the system information necessary to conduct a probabilistic risk analysis.

An event tree is a graphical representation of the various accident sequences that can occur as a result of an initiating event (USNRC 2012). It is an essential tool in analyzing whether a complex system satisfies its system-level design targets. It provides a rational framework for enumerating and, subsequently, evaluating the myriad of events and sequences that can affect the operation of the CI system.

A fault tree is an analytical model that graphically depicts the logical combinations of faults (i.e., hardware failures and/or human errors) that can lead to an undesired state (i.e., failure mode) for a particular subsystem or component (Vesely et al 1981). This undesired state serves as the topmost event in the fault tree, and usually corresponds to a top event in an event tree. Thus, a fault tree provides a rational framework for identifying the combinations of hardware failures and/or human errors that can result in a particular failure mode of a subsystem or component. Once fully developed, a fault tree can be used to quantitatively evaluate the role of a CI subsystem or component in the operation and failure of the CI system.

A particular graphical combination of a fault tree and an event tree, called the bow-tie model (De Dianous and Fiévez 2006), has been used in risk management since the 1980's to visually represent the possible causes and consequences of an accident. Typically, the causes of an accident are shown on the left side using a fault tree, while the consequences of an accident are shown on the right side using an event tree.

Bayesian networks (BNs) are probabilistic models that provide an efficient framework for probabilistic assessment of component/system performance and can be used to model multiple hazards and their interdependencies[2]. They may also facilitate information updating for near-real time and post-event applications. Evidence on one or more variables can be entered in a BN model to provide an up-to-date probabilistic characterization of the performance of the system. BN is nowadays used for infrastructure risk assessment and decision support, particularly in the aftermath of a natural event (Bensi, 2010).

Similar to event and fault trees, thus also bowties, the topology of a BN is derived from an analysis of the system and remains static. This means that the component, structures and subsystems and the causal links and conditional dependencies among them are pre-determined and do not change during the probabilistic risk analysis process. There are, however, so-called adaptive (Pascale and Nicoli 2011) or reconfigurable (Mirmoeini and Krishnamurthy 2005) BNs whose topology changes (among several pre-determined topologies) to best match an estimate of the varying state of the modeled system. Finally, there are modular BNs (Niel et al 2000), built out of many BN modules, with each module representing a functionally independent component or subsystem of a system-level BN (Park and Cho, 2012).

More important, the probabilistic nature of the two frameworks is different: the event/fault tree framework is based on the notion of probability as a frequency, while the BN framework represents the state of knowledge or belief. Fundamentally, the BN framework naturally allows for introduction of new knowledge, for example, from observations of the CI system behavior during its normal operation, from inspections, or from previous stress tests. This enables a fundamental aspect of the proposed stress test methodology, that of repeating a stress test in certain intervals depending on the outcome of the previous stress test in order to reduce the risk exposure of the CI through the practice of continuous improvement.

2.4.1 Application of BNs to natural hazards and CIs

The use of BNs for natural hazard assessment has increased in recent years. Straub (2005) presented a generic framework for the assessment of the risks associated with natural hazards using BNs and applied it to the rockfall hazard. BNs have also been applied to the modeling of risks due to typhoon (Nishijima and Faber 2007), geotechnical and hydrological risks posed to a single embankment dam (Smith, 2006), avalanches (Grêt-Regamey and Straub 2006), liquefaction modeling (Bayraktarli et al 2005, 2006, Tasfamariam and Liu, 2014), tsunami early warning (Blaser et al 2009) and seismic risk (Bayraktarkli et al 2005, 2006, 2011; Bensi, 2010; Broglio, 2011).

In particular, regarding seismic risk, Bayraktarkli et al (2005, 2006) proposed a three components framework (Fig. 2.7) for earthquake risk management using BNs, composed of an exposure model that is an indicator of hazard potential, a vulnerability model which is an indicator of direct/immediate consequences, and a robustness model to quantify indirect consequences. However, the framework proposed by the authors does not include many aspects which complicate the applications of BNs to seismic hazard and risk analysis of infrastructure systems such as the modeling of ground motion random fields, directivity effects, or issues associated with the modeling of system performance.

Bensi (2010) proposed a more comprehensive BN methodology for performing infrastructure seismic risk assessment that includes also a decision model for post-event

[2] A short introduction to BNs terminology and probabilistic structure is available in Deliverable 5.2 (Esposito and Stojadinovic, 2016a).

decision making (Fig. 2.8). The methodology developed by Bensi (2010) consists of four major components: i) a seismic demand model where ground motion intensities are modelled as Gaussian random field accounting for multiple seismic sources and including finite fault rupture and directivity effects; ii) a performance model of point-like and distributed components; iii) models of system performance as a function of component states; and iv) the extension of the BN to include decision and utility nodes to aid post-earthquake decision-making.

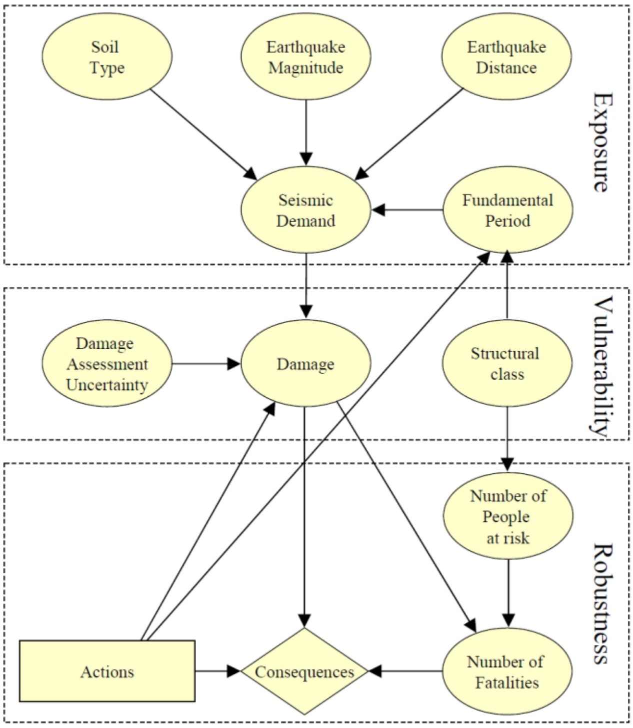

Fig. 2.7 BN framework for seismic risk management (source: Bayraktarli et al, 2005)

In addition to demonstrating the value of using Bayesian networks for seismic infrastructure risk assessment and decision support, the study proposed models necessary to construct efficient Bayesian networks with the goal of minimizing computational demands, which represent one of the weak points of BN frameworks.

More recently Grauvogl and Steentoft (2016) and Didier et al (2017) proposed a BN-based model to evaluate the seismic resilience of infrastructure systems. The model is based on the compositional supply/demand resilience quantification framework presented in the STREST Deliverable 4.5 report (Stojadinovic and Esposito, 2016).

A schematic overview of this BN-based model used to evaluate the resilience of the electric power supply system in Nepal after the 2015 Gorkha earthquake is shown in Fig. 2.9.

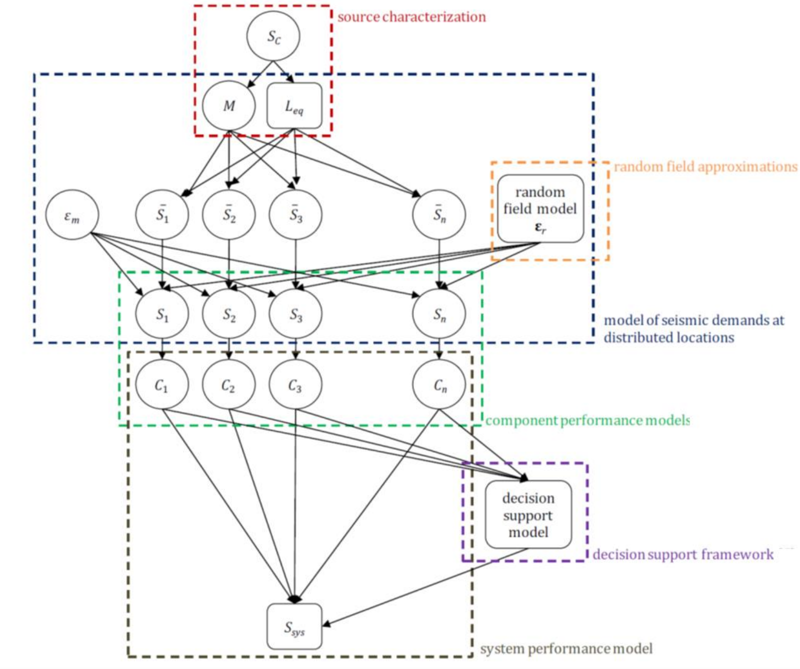

Fig. 2.8 Bayesian network methodology for seismic infrastructure risk assessment and decision support proposed by Bensi (2010)

2.4.2 Discussions

Bayesian networks are useful tools in engineering risk analysis because they facilitate the computation, the understanding and the communication of complex problems subject to uncertainty.

BNs offer several important advantages. BNs provide an efficient framework for probabilistic assessment of component/system performance and can be used to model multiple hazards and their interdependencies. They are an efficient and intuitive graphical tool that enable representation of the components and subsystems and the causal links and conditional dependencies among them and assessment of systems under uncertainty. They provide a consistent and clear treatment of the joint probability distributions of multiple random variables, and an efficient framework for probabilistic real-time updating in light of new evidence. BN can be also be extended to include utility and decision nodes, thus providing a decision tool for ranking different alternatives. Complex BNs can be constructed using verified and validated modules that represent components and subsystems of the CI system.

Fundamentally, the BN framework naturally allows for introduction of new knowledge, for example, from observations of the CI system behavior during its normal operation, from inspections, or from previous stress tests. This enables a crucial aspect of the proposed stress test methodology, that of repeating a stress test in certain intervals depending on the outcome of the previous stress test in order to reduce the risk exposure of the CI through the practice of continuous improvement.

However, Bayesian networks have limitations. Calculations in Bayesian networks can be highly demanding and the application to distributed systems characterized by a complex topology is not always feasible. An accurate modeling via BNs requires thorough understanding of the problem. The need for expert knowledge in generating the preliminary BN structure represents one of the most salient points of this tool. Modeling complex systems via BNs may require trade-offs between accuracy, transparency, computational complexity, and detail of modeling (Friis-Hansen 2004).

Further, the availability of statistical data to develop robust models to relate random variables in a BN is often scarce in civil engineering and infrastructure system analysis (Bensi, 2010). Thus, dependence relations between parents and children and the marginal distributions of root nodes should be based on theoretical models and/or expert judgement.

Although BNs represent an appropriate framework to handle uncertainty for pre- and post-event risk assessment and decision support analysis, it is important to acknowledge that challenges remain, particularly with respect to computational demands for application to large civil infrastructure systems.

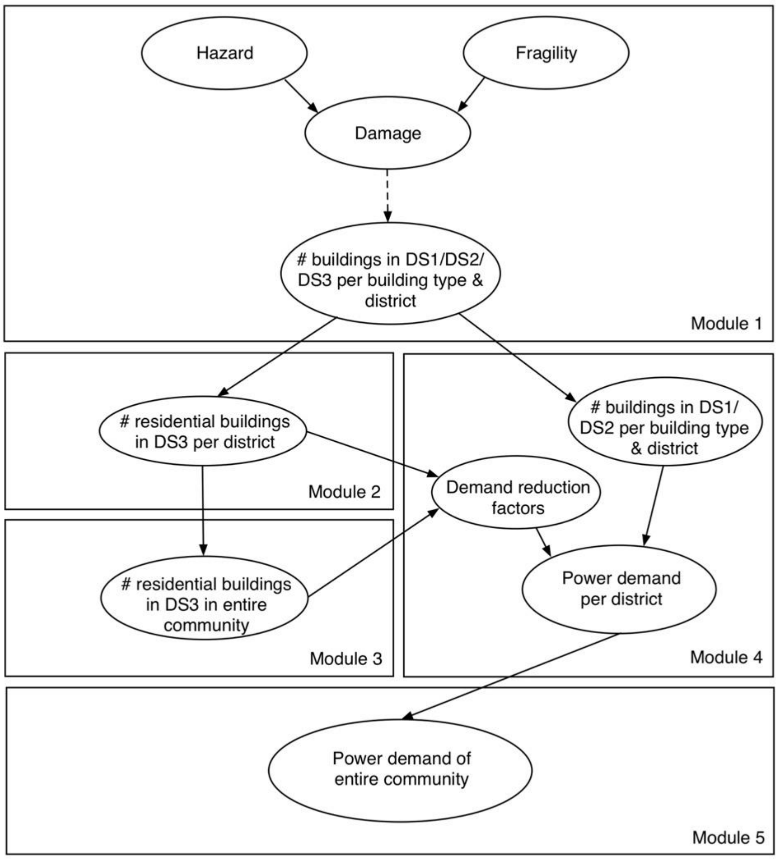

Fig. 2.9 BN model proposed by Didier et al (2017)

2.5 ST@STREST grading system

The first outcome of the stress test, obtained in the STEP 6 (Risk Objectives Check), is described using a grading system (Esposito et al 2016). This grading systems is based on the comparison of the results of risk assessment with the risk objectives (i.e. acceptance criteria) defined at the beginning of the test in STEP 2 (Risk Measures and Objectives).

The proposed grading system (Fig. 2.10) is composed of three different outcomes: Pass, Partly Pass, and Fail. The CI passes the stress test if it attains grade AA or A. The former grade corresponds to negligible risk and is expected to be the attained risk objective for new CIs, whereas the latter grade corresponds to risk being as low as reasonably practicable (ALARP, Helm, 1996; Jonkman et al, 2003) and is expected to be the attained risk objective for existing CIs. Further, the CI partly passes the stress test if it receives grade B, which corresponds to the existence of possibly unjustifiable risk. Finally, the CI fails the stress test if it is given grade C, which corresponds to the existence of intolerable risk.

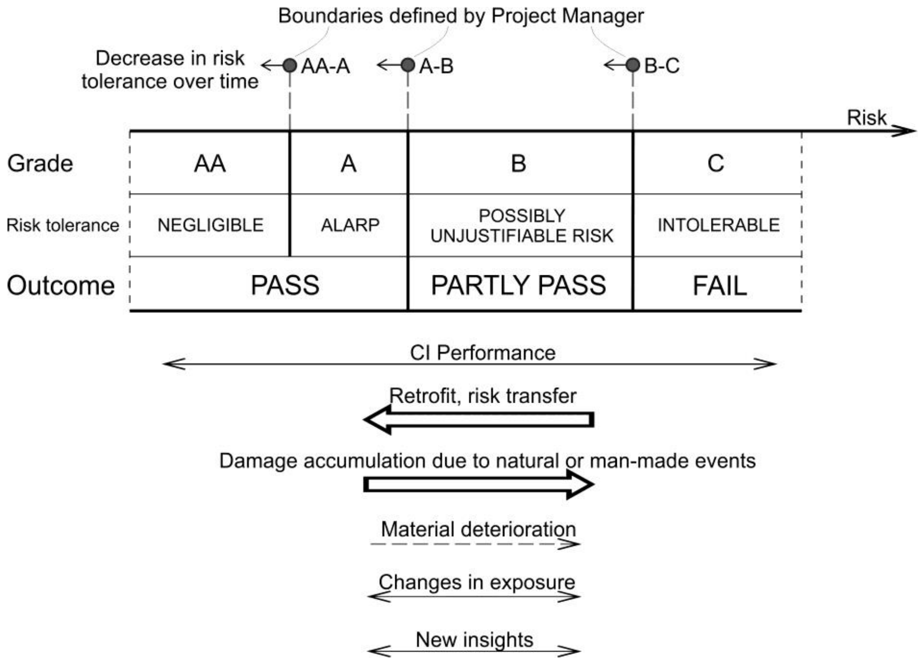

Fig. 2.10 An example of grading system for the outcome of stress test. The CI may pass, partly pass, or fail the stress test

In the following sections, the risk limits and the boundaries between grades are first discussed. This is followed by the description of how the grading system is extended considering the time dimension. The guidelines for the grading of individual components are then given. A generalization of the grading system is made in order to apply it to those ST levels which take into account epistemic uncertainties and system analysis. Finally, a brief discussion is given.

2.5.1 Risk limits and boundaries between grades

The project manager (PM) of the stress test defines the boundaries between grades (i.e. the risk objectives) by following requirements of the regulators. The boundaries (i.e. the acceptance risk levels, see STREST Deliverable 5.1*, Esposito et al 2016)* can be expressed using scalar (Fig. 2.11 top) or continuous (Fig. 2.11 bottom) risk measures. Examples of the former include the annual probability of the risk measure (e.g. loss of life) and the expected value of the risk measure (e.g. expected number of fatalities per year), whereas the latter is often represented by an F-N curve, where F represents the cumulative frequency of the risk measure (N) per given period of time. In several countries, an F-N curve is defined as a straight line on a log-log plot. However, the parameters of these curves, as well as parameters of scalar risk objectives (i.e. regulatory boundaries in general) may differ between countries and industries (STREST Deliverable 5.1, Esposito et al 2016). Harmonizing the risk objectives of risk measures across a range of interests on the European level remains to be done. This is a task for regulatory bodies and for industry association: they should reconcile the societal and industry interest and develop mutually acceptable risk limits. When acceptance criteria are defined as continuous measures, the grade is assigned based on the position of the farthest point of the CI loss curve from the F-N limits (Fig. 2.11).

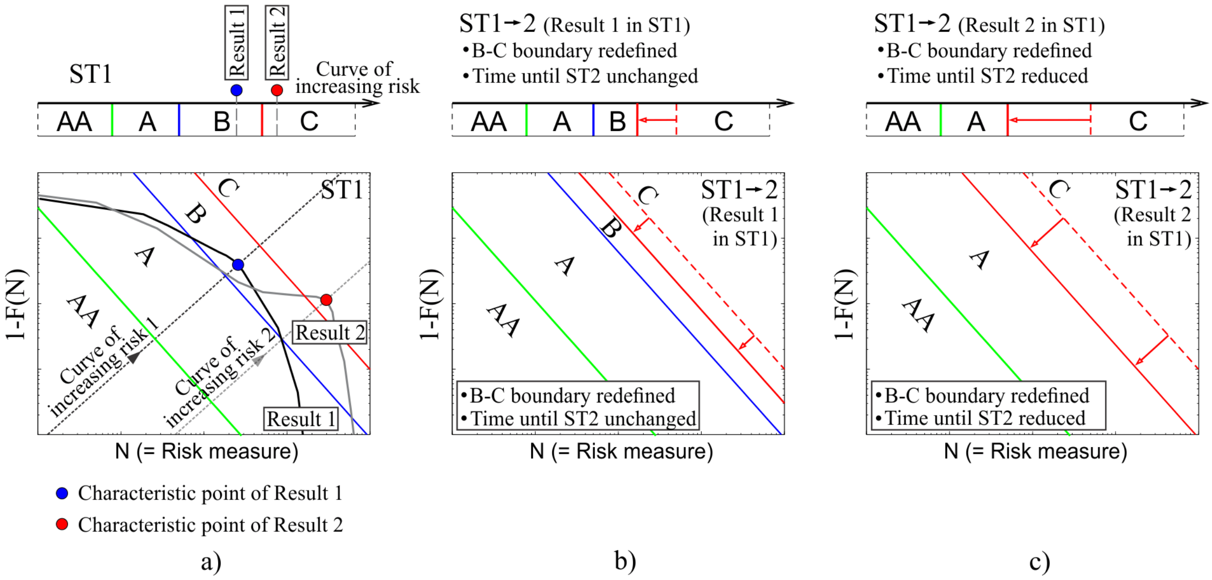

Fig. 2.11 Grading system in time domain using scalar risk objectives (top) and limit F-N curves (bottom): a) two different results of the first evaluation of stress test (ST1), b) redefinition of the parameters of the grading system due to Result 1 in ST1, and c) redefinition of the parameters of the grading system due to Result 2 in ST1

2.5.2 Grading system in time domain

In general, the CI performance can be understood as time-variant. It may change due to, for example, ageing through use, long-term degradation process such as corrosion, effects of previous hazard events, man-made events (e.g. terroristic attacks), and change in exposure (e.g. population). Such change in performance may lead to an increase of the probability of failure or loss of functionality, or exacerbate the consequences of failure during the CI system's lifetime (Fig. 2.10).

In the proposed grading system, it is foreseen that the performance of the CI and/or the performance objectives can change over time. Consequently, the outcome of the stress test is also time-variant. For this reason, the stress test is periodic, which is also accounted for by the grading system. If the CI passes a stress test (grade AA or A), the risk objectives for the next stress test do not change until the next stress test. The longest time between successive stress tests should be defined by the regulator considering the cumulative risk. However, most of existing CIs will probably obtain grade B or even C, which means that the risk is possibly unjustifiable or intolerable, respectively. In these cases, the grading system has to stimulate the stakeholders to upgrade the existing CI or to start planning tor a new CI in the following stress test cycle. It is proposed that stricter risk objectives are used or that the time between the successive stress tests is reduced in order to make it possible that stakeholders adequately mitigate the risks posed by the CI in as few repetitions of the stress test as possible, which means that the CI will eventually obtain grade A or the regulator will require that the operation of CI be terminated.

The basis for the redefinition of risk objectives in the next stress test is the so-called *characteristic point of risk*. In the case when scalar risk measures are used, the characteristic point of risk is represented directly by the results of the risk assessment (Fig. 2.11, top). In the case when result of risk assessment is expressed by a loss curve in F-N space, the characteristic point is defined by one point of the F-N curve. In general, each curve of increasing risk (see Fig. 2.11) results in one point of the F-N curve. The curve of increasing risk, associated with the characteristic point is denoted as the *characteristic curve of increasing risk*. It is recommended that the point associated with the greatest risk above the ALARP region be selected as the characteristic point (see Fig. 2.11a). In this case the characteristic point is defined as the point of the F-N curve which is the farthest from the limit F-N curve that represents the boundary between grades (for example, grades A and B, and the A-B boundary are shown by the blue line in Fig. 2.11a).

Once the characteristic point is determined, the grading system parameters for the next repetition of the stress test can be defined. If the CI obtains grade B in the first evaluation of stress test (ST1, blue dot in Fig. 2.11a), the grading system foresees the reduction of the distance between grades B and C (the B-C boundary) in the next stress test (ST2, Fig. 2.11b). This reduction should be equal to the amount of cumulative risk beyond the ALARP region assessed in ST1. This ensures risk equity over two cycles, which may be expressed by the following expression:

$$R_{ST1} - R_{(A\text{-}B)} = R_{(B\text{-}C),\,ST1} - R_{(B\text{-}C),\,ST2} \tag{2.3}$$

where $R_{(A\text{-}B)}$ is the A-B boundary, $R_{(B\text{-}C),ST1}$ and $R_{(B\text{-}C),ST2}$ are the B-C boundary in ST1 and ST2, respectively, and R_{ST1} is the value of the risk measure assessed in the ST1. Note that the left side of the Eq. 2.3 is equal to the amount of risk beyond the ALARP region assessed in ST1. Furthermore, if grade C (red dot in Fig. 2.11a) is given in ST1, both the B-C boundary and the period until the next stress test ST2 are reduced (Fig. 2.11c). In this case, the B-C boundary is set equal to the A-B boundary, since this is the maximum possible reduction of the region of possibly unjustifiable risk. Moreover, the reduced period until ST2 ($t_{cycle,redefined}$) is determined on the basis of equity of risk above the ALARP region over the two cycles and can be calculated using the following expression:

$$t_{cycle,redefined} = t_{cycle,initial} \cdot \frac{R_{(B\text{-}C),\,ST1} - R_{(A\text{-}B)}}{R_{ST1} - R_{(A\text{-}B)}} \tag{2.4}$$

where $t_{cycle,initial}$ is the initial amount of time between two stress tests.

2.5.3 Grading of the components

Each component is assessed by at least one method (hazard-based, design-based or risk-based assessment). Objectives of a hazard-based assessment and a design-based assessment are obtained directly from the design codes, whereas the risk objectives need to be defined in Step 2 (*Risk Measures and Objectives*). Similar to the case of system level assessment, three thresholds need to be defined (between grades AA and A, between grades A and B and between grades B and C) in order to consistently evaluate the components of a CI.

If a less detailed and sophisticated method assessment (see *Section 2.6*) results in the component not being in compliance with the requirements or the requirements are unknown, a more sophisticated method may be used. For the Component-Level Assessment (STEP 4), three levels of detail and sophistication are defined as Moderate, High and Moderate-Advanced for hazard-based assessment, design-based assessment and risk-based assessment, respectively. Different levels in the case of risk-based assessment exist due to various levels of complexity of hazard and fragility analysis. If the result of a hazard-based assessment or a design-based assessment is that the component is in compliance with the requirements, a grade A is assigned to the component. If these types of assessment result in the component not being in compliance with the requirements or the requirements are unknown, a grade C is assigned to the component, or a higher Level assessment is required. Note that, if the risk-based assessment is used, the grading system at the component level is same as that proposed for the system-level assessment. The proposed procedure for the progressive approach in the case of the assessment at the level of component and the corresponding grading system is illustrated in Fig. 2.12.

If a component is assigned grade C, mitigation actions need to be taken. The time in which the grade needs to be improved depends on the type of assessment. If a hazard-based or a design-based assessment is used, the mitigation has to be made immediately, as the component is not in compliance with the current regulatory requirements. If a risk-based assessment is used, the time in which the grade has to be improved is determined on the basis of the amount of risk corresponding to the component reaching the designated limit state in the time period considered (see *Section 2.5.2*).

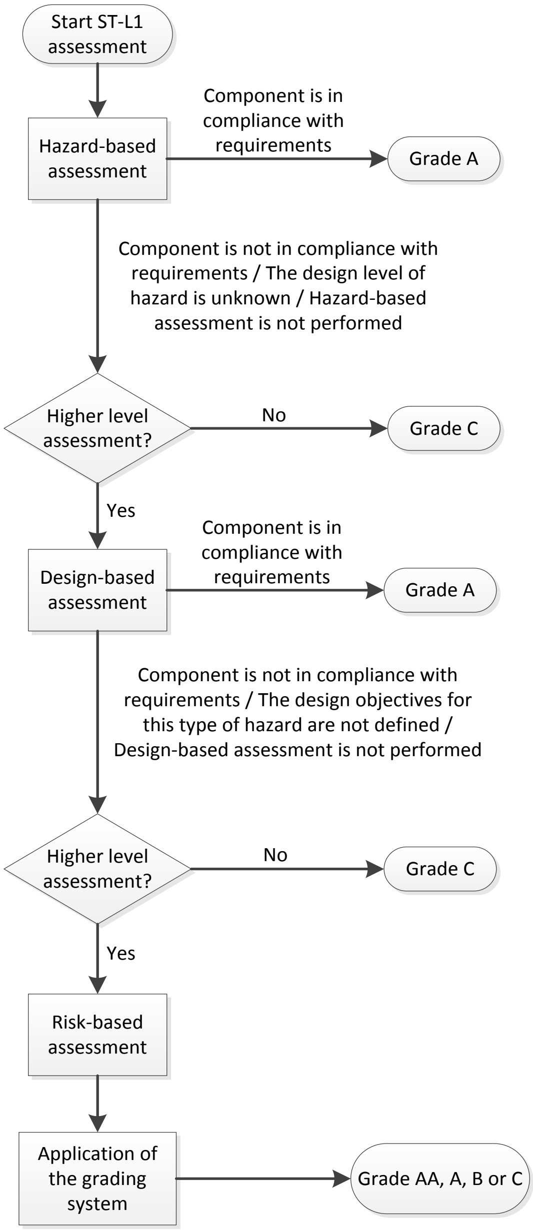

Fig. 2.12 Grading of components of the system (ST-L1)

In this section the grading of the components is applied to the example provided in *Section 2.3.1*.

Risk objectives for the component in terms of probability of collapse, which needed to be defined in Step 2, are as follows: 10^{-6} between grades AA and A, 10^{-4} between grades A and B, and 10^{-3} between grades B and C. The most stringent risk boundary is approximately equal to the target probability of collapse, which is foreseen in building codes for frequent or permanent loads, e.g. in Eurocode 0 (CEN, 2004). Those values of acceptable probability of collapse are within a magnitude of 10^{-6}. Such a low probability of collapse cannot be achieved by employing building codes for earthquake-resistant design since the nature of seismic action is completely different than the nature of frequent or permanent load. The probability of collapse for buildings designed according to Eurocode 8 is around magnitude of 10^{-5}. A significantly larger value of target collapse risk (1% in 50 years ($2\cdot10^{-4}$)) was assumed for new buildings in USA (Luco et al, 2007). As a consequence, the risk boundary between grades A and B was set to 10^{-4}, while the risk boundary between grades B and C was increased 5 fold. The probability of collapse 5% in 50 years approximately corresponds to buildings, which were designed and constructed in the third quarter of 20th century.

The procedure is initiated by performing the hazard-based assessment. Since the design level of hazard is unknown, there are two options: settle with grade C or move on to the design-based assessment. We choose the latter. The design-based assessment results in the component not being in compliance with the code, then two options are possible: settle with grade C or move on to the risk-based assessment. We choose the latter. This results in the probability of collapse equal to $8.5\cdot10^{-4}$. Thus, the component receives grade B, which means that no risk mitigation actions are required, but the threshold between grades B and C will be reduced to $2.5\cdot10^{-4}$ in the next stress test.

2.5.4 Grading of the system with consideration of epistemic uncertainties

The grading system presented in Sections 2.5.1 and 2.5.2 assumes that no epistemic uncertainties are related to the assessed risk. Since ST-L2c and ST-L3c consider the effect of epistemic uncertainties, the grading system needs to be generalized in a way that it accounts for a distribution of values of the risk measure. The grading criteria based on a distribution of risk measure values can be formulated in a variety of ways. In this project, it is recommended that the mean value of the risk measure distribution be used to assess the CIs. Other options, which should be examined in future studies, are discussed in Section 2.5.5.

Furthermore, the grading system for consecutive stress tests, described in Section 2.5.2, is based on the cumulative probability of risk measure exceedance in the selected time period between two stress tests. For this reason, we determine the left side of Eq. 2.3, i.e. the total value of risk above the ALARP region, as the sum of all possible risk values above the ALARP region (dashed area in Fig. 2.13), which are weighted by their probability:

$$R_{ST1} - R_{(A\text{-}B)} = \int_{R_{(A-B)}}^{\infty} p(R)(R - R_{(A\text{-}B)})\, dR \qquad (2.5)$$

In the case of risk measure based on an F-N curve, each curve of increasing risk corresponds to a distribution of points from different F-N curves (Fig. 2.13b). The characteristic curve of increasing risk is the curve that corresponds to the greatest amount of risk above the ALARP region, i.e. where the integral in Eq. (2.4) produces the highest value.

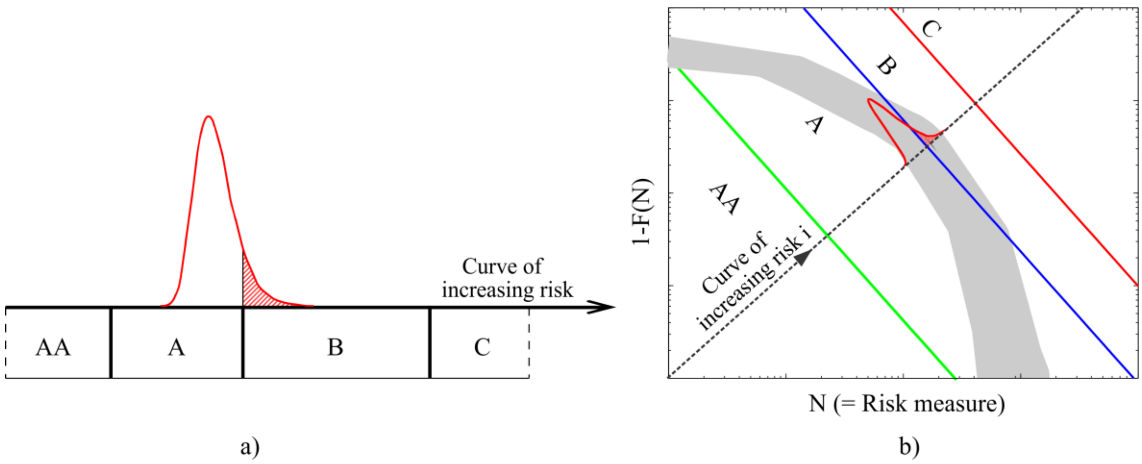

Fig. 2.13 Distribution of a risk measure with boundaries of grades in the case of a) a scalar risk measure and b) an F-N curve

2.5.5 Discussion and future developments

There are some points of the grading system that need to be discussed and further developed as a part of the future studies.

Firstly, it is yet to be determined how grades of single components should affect the global outcome of stress test. For example, if the CI is assigned grade B in the ST-L2 assessment, the outcome is a partly pass. However, one or several components may receive grade C in the component level assessment. It is unclear how this should affect the global outcome. One option would be to change the global outcome of stress test to "fail", since the stakeholders would be required to reduce risk of those components. However, such an approach may be too conservative. Another option would be to introduce a complementary outcome of stress test, which would address only single components and would be independent of the outcome obtained based on systemic level assessment. In this case, risk mitigation strategies and guidelines would be defined separately for individual components as well. A third option would be to require a system-level assessment in this or the subsequent stress test that explicitly accounts for the effects of the offending components on the behavior of the system using, for example, the bow-tie approach, to identify the causes and the effects of failure of such components.

Secondly, in case epistemic uncertainty analysis is of concern, it is currently recommended that the mean value of the designated risk measure is used. However, other options should be investigated. The grade could be based on other quantiles of the risk measure distribution, which should be determined by the PM. Guidelines for selection of the appropriate risk measure distribution quantile should be developed based on a comprehensive parametric study as a part of future developments. Grades could also be assigned based on a value of the risk measure corresponding to a specific number of standard deviations above the mean, i.e. the confidence level that a specific value of the risk measure will not be exceeded. Such approach, the high confidence of low probability of failure (HCLPF) is used by the nuclear industry. Again, comprehensive parametric studies would be required to select the appropriate number of standard deviations for non-nuclear critical CIs. Furthermore, grades could depend on the type of adjustments of the grading system parameters and the time between successive stress tests. For example, if a redefinition of the boundary between grades B and C is required (based on the amount of risk above the ALARP region, see Fig. 2.11 and 2.13), grade B would be assigned. If the reduction of the time before the next stress test is also required (again based on the amount of risk above the ALARP region), grade C would be assigned.

Thirdly, the proposed grading system requires boundaries (acceptance criteria) to be defined between the regions of negligible, ALARP, possibly unjustifiable, and intolerable risk. The PM will often need to rely on his or her own judgement when defining these boundaries, especially in situations where regulatory requirements do not yet exist. It is the matter of future developments to create recommendations for the boundaries of different types of performance measures that can be used by the PM of the stress test as the guidelines.

2.6 ST@STREST penalty system

There is a wide range of methods and models for assessing performance of critical infrastructures against natural hazards. These methods cover different levels of detail and complexity for each hazard, vulnerability, and risk computation. All models are necessarily a simplification of the reality. However, the level of simplification may vary significantly. In fact, different models and methods have to be assumed or introduced to describe how the hazard and vulnerability interact in time and in space. Furthermore, each combination corresponds a different level of detail of the analysis.

For example, regional seismic hazard assessments and site-specific hazard assessments may both represent the input for the risk assessments, however they do differ in the level of details related to the hazard analysis (e.g., the description of the natural variability of sources, the details in modelling the propagation from source to target, etc.). In a similar way, generic fragility functions and element specific fragility function may be used, but again they largely differ in the level of details considered in their quantification. Such differences are expected to significantly influence the reliability of the risk results.

In the STREST methodology, the "level of detail and sophistication" used for the risk computation reflects the level of complexity of the methods adopted for the component and system-level risk assessment. In a general sense, it may be defined as *the trueness and precision, and the repeatability and reproducibility of the results of the risk assessment*.

The selection of the "level of detail and sophistication" to be used in a particular stress test, namely, to perform the hazard and risk analysis, is important because it allows defining how reliable are the results of the Assessment phase of the stress test. At the same time, this is a challenge, since it requires experts that need to have a clear idea about all of the models and methods available in the scientific literature to perform each step of the analysis, i.e. the *"center, body and range"* of the methods and models. The state-of-the-practice methods and models are expected to have the trueness, precision, repeatability and reproducibility that can be achieved within the established state of knowledge and within a reasonable engineering and analysis effort. The experts need to characterize the trueness and precision of the state-of-practice methods using multiplicative factors (to shift the mean and adjust trueness) and dispersions (to characterize the precision). More advanced methods should be promoted and less advanced methods should be discouraged by adjusting the factors used to characterize them. Thus, a penalty system is proposed as a part of the ST@STREST methodology.

During the Pre- Assessment Phase (STEP 3: *Set-up of the Stress Test*) the TI and PM select the most appropriate ST-Level for the given CI. As each ST-Level corresponds to a different level of complexity of the hazard and risk analysis, a different level of "detail and sophistication" should be required as a minimum to perform the required analysis.

In particular, in the proposed ST@STREST, a *Target Level (TL)* of "detail and sophistication", has been associated with each ST-Level, according to the judgement about the complexity of the required hazard and risk analysis. This target value represents the state of knowledge of the community and characterizes the state-of-practice of assessing the CI at the component and the system level.

Then, data, models and methods needed to perform each step of the risk analysis are identified by the TI. These models and methods are characterized by a level of detail that reflects the grade of complexity among the wide range of available methods in the scientific literature. The level of "detail and sophistication" of the Stress Test depends on the specific models selected for the particular test. This selection is mainly based on a scientific ground, but also has practical consequences, such as the requirement of the necessary duration and resources for the stress test. Therefore, the choice of the models should be taken (and documented) jointly by the TI and the PM. Based on the choices made, the TI evaluates the *Effective Level* of detail of the analysis *(EL)*. This assessment is reviewed by the Internal Review (IR) team, and compared with the *TL*. The *EL* should be at least as high as the *TL*. Based on the IR review, PM and TI may evaluate if changes to the hazard and risk analysis complexity are needed, principally to avoid potential penalties suggested by the reviewers. In fact, if the *EL* attained in the conducted stress test is lower than the *TL* required, the ST@STREST Penalty System is applied.

In the following, the ST@STREST Penalty System, based on the difference between the *EL* and the *TL*, is proposed.

2.6.1 Proposed penalty system

The proposed ST@STREST Penalty System aims to penalize the results of the hazard and risk assessment of the conducted stress test by evaluating a *Penalty Factor* (PF). This factor penalizes simplistic approaches (with respect to the state-of-practice) that cannot guarantee a sufficiently accurate analysis.

The PF is defined by the TI in STEP 6 (*Risk Objectives Check)* of the methodology based on the difference between *EL* and *TL*. Namely, if the *EL* is greater or equal to *TL* the penalty system is not applied.

Levels of "detail and sophistication" and a *Penalty Factor* scheme are proposed in the following. This is just one of the possible schemes that the PM and TI need to determine, the IR to review and confirm, with a possibility to involve the PoE to arrive at the broadest possible consensus. However, the proposed *Levels* and *Penalty Factor* system is general and can be applied in stress test.

Proposed levels

Three categories are defined to describe the trueness, precision, repeatability and reproducibility of the hazard and risk analysis in a stress test:

- **Advanced**: making use of detailed information and advanced state-of-the-art methods and models in most of the steps of the assessment;
- **High**: making use of commonly detailed information and state-of-the-practice methods and models in most of the steps of the assessment;
- **Moderate**: making use of coarse information and simplified methods in most of the steps of the assessment.

Starting from this classification, a *Target Level* has been associated to each ST-Level (Table 2.8) according to the grade of complexity of the risk analysis required. In case a quantitative scale is adopted, a *Factor* interval ($F \in [0,1]$) is set up by the experts and associated to each Level. An example is provided in the following:

1. **Advanced** : $F \in [0.7,1]$
2. **High** : $F \in [0.4,0.7)$
3. **Moderate:** $F \in [0.2,0.4)$

These values associated to each level are indicative: in a particular stress test, they need to be determined by consensus between the PM, TI and IR. In general, these values can be studied in more detail, for example, in a study to account for different parameters that affect the results of stress tests. In this case, the resulting *Effective Level* identified for the hazard and risk assessment, i.e. the *EL,* should be at least equal to the lowest *F* (lower bound of the interval) corresponding to the *Target Level*. In this case the *TL* is characterized by lower (TL_{lb}) and upper (TL_{ub}) bounds.

Table 2.8 Target Levels for each ST-Level

ST-Level	Target Level (TL)
1a	Moderate
2a	Moderate
2b	High
2c	Advanced
2d	Advanced
3c	Advanced
3d	Advanced

Effective level (EL)

At component-level (ST-L1) there are three methods to perform the single-hazard component check. These methods differ in the complexity and the data needed for the computation. Therefore, the associated "level of detail and sophistication" is set as follows:

- Hazard-based assessment: **Moderate**
- Design-based assessment: **High**
- Risk-based assessment: **Moderate to Advanced**

The *Effective Level* for the component hazard-based and design-based assessments is moderate (lowest of all possible) and high, respectively. This means that, according to Table 2.8, the hazard-based assessment represents the minimum level of analysis required.

If a risk-based component assessment approach is required, the *Effective Level* may vary according to the level of trueness, precision, repeatability and reproducibility used for the evaluation of hazard and vulnerability. Therefore, the resulting *EL* is a function of the trueness, precision, repeatability and reproducibility of the method adopted for hazard and vulnerability analysis and it may vary from Moderate to Advanced.

For system–level (ST-L2 or ST-L3) stress tests, the evaluation of *EL* is a function of the level of detail selected for each hazard, the method adopted for the epistemic uncertainty quantification, and the method adopted for the multi-hazard/risk evaluation. Furthermore, evaluation of *EL* for each hazard is a function of the level of each step and sub-step needed for the computation of the performance and risk of the CI. In other words, if the computation of risk comprises three principal steps *i* (hazard, vulnerability and risk), and each one of the steps is characterized by *j* different layers*,* the resulting *EL* is a function of the level of "detail and sophistication" of each step *i* and layer *j.* Thus, if a qualitative scale is adopted, the *EL* corresponds to the most frequent (mode) value of the level of detail adopted in each step and layer. If a quantitative scale is adopted (i.e. a quantitative factor is associated with the analysis), the *EL* may be computed (for a single hazard analysis, ST-L2) following Eq. (2.6):

$$EL = W_1 \frac{\sum_{j=1}^{n} w_{1,j} EL_{1,j}}{n} + W_2 \frac{\sum_{j=1}^{m} w_{2,j} EL_{2,j}}{m} + W_3 \frac{\sum_{j=1}^{p} w_{3,j} EL_{3,j}}{p} \quad (2.6)$$

where *n, m* and *p* are the number of layers in each step (hazard, vulnerability, risk); W_i represent the weight of each step *i* of the risk analysis *and* $w_{i,j}$ the weight of each layer *j* (for each step *i*) set up by experts. If all layers (for each step) are considered equally important, then $w_{1,1} = w_{1,2} = ... = w_{1,n} = 1$ $w_{2,1} = w_{2,2} = ... = w_{2,m} = 1$, $w_{3,1} = w_{3,2} = ... = w_{3,p} = 1$. If all steps are considered equally important, then $W_1 = W_2 = W_3 = 1/3$.

In case of a multi-hazard analysis (ST-L3), AL_E may be obtained as in Eq. (2.7):

$$EL = \frac{H_1 EL^{H_1} + H_2 EL^{H_2} + ... + H_s EL^{H_{1s}}}{s} \quad (2.7)$$

where H_q represents a weight of each hazard *q* set up by experts. Thus, a multi-hazard *EL* corresponds to the weighted mean of the level of detail evaluated for each hazard EL^{Hq}. If all hazards are considered equally important, then the weights $H_1 = H_2 = ... = H_s = 1$. If the epistemic uncertainty analysis is also of concern, the method of accounting for epistemic uncertainties could be considered as an additional layer.

Penalty factor (PF)

The penalty factor (PF) is defined as the difference between the *EL* and the *TL* of the ST level selected. If a qualitative scale (i.e. Moderate, High, Advanced) is considered, three cases are possible:

a) *TL*=High, *EL* =Moderate
b) *TL* =Advanced, *EL* = High
c) *TL*= Advanced, *EL* =Moderate

The penalty factor may be computed using the reference values that may be associated to the three cases.

For example, in the cases above: a) $PF_{H\text{-}M}$=0.2, b) $PF_{A\text{-}H}$=0.2, c) $PF_{A\text{-}M}$=0.4. These values are indicative: the actual values need to be set by experts' consensus for each stress test.

If the "level of detail and sophistication" is expressed using a quantitative scale, *PF* is defined as the difference between the *EL* and the lower bound of the *TL* of the ST level selected $\left(TL_{lb}^{ST}\right)$,

$$\begin{cases} PF = \left(TL_{lb}^{ST} - EL\right) & if \quad EL < TL_{lb}^{ST} \\ 0 \quad otherwise \end{cases} \tag{2.8}$$

Note that the penalty system could be also applied to penalize the CIs that reach the minimum target but just barely, i.e. when $TL_{lb}^{ST} \leq EL < TL_{ub}^{ST}$. In this case, the *PF* may be evaluated considering the upper bound of the *TL*, i.e.,

$$PF = \left(TL_{ub}^{ST} - EL\right) \tag{2.9}$$

Penalized loss, LP

Consider that the outcome of the risk assessment at the system level is expressed by the annual exceedance rate of losses (*L*), $\lambda(l)$. For example, in case seismic hazard is of concern, according to the PEER performance based earthquake engineering (PBEE) framework (Cornel and Krawinkler, 2000), $\lambda(l)$ is formulated as:

$$\lambda(l) = \int_d \int_{edp} \int_{im} G(l \mid d) \left| dG(d \mid edp) \right| \left| dG(edp \mid im) \right| \left| d\lambda(im) \right| \tag{2.10}$$

where *im* is an intensity measure (e.g., peak ground acceleration, peak ground velocity, spectral acceleration, etc.), *edp* is an engineering demand parameter (e.g., interstorey drift), *d* is a damage measure (e.g., minor, medium extensive, etc.), *l* is the loss variable (e.g., monetary losses, down-town time, etc.), and $G(\mathrm{y} \mid \mathrm{x})$ is a conditional complementary cumulative distribution function (CCDF) relating the variables.

As mentioned before, the risk analysis can be performed at different levels of "detail and sophistication". In Eq. 2.9 it is possible to include an extra uncertainty, here named penalty uncertainty, to penalize simplistic analysis approaches. Therefore, a new metric is introduced, named penalized loss L_P expressed (in the logarithmic scale) as:

$$\log(L_P) = \log(L) + \varepsilon_P \tag{2.11}$$

where ε_P is the penalty uncertainty. Observe that penalty uncertainty ε_P acts exactly as model error. In fact, the objective is to amplify the uncertainties introduced by simplistic approaches that cannot guarantee an analysis with desirable level of "detail and sophistication". A convenient choice for the probability distribution of ε_P is the Normal distribution, i.e. $\varepsilon_P \sim N(0, \sigma(l))$, where $\sigma(l)$ is defined as:

$$\sigma(l) = |PF \cdot \log(l)|, l > 0 \tag{2.12}$$

where *PF* is the penalty factor defined previously. Observe that *PF* acts as a coefficient of variation (c.o.v). Further, in order to focus on the tails of the risk curve, no error is added to the penalty factor for $l = 0$.

Considering that the support of L is usually $[0, +\infty)$ or bounded as $[0, l_{max}]$, the distribution of ε_P must be truncated according to the support of *L*. It is of interest to observe that $\sigma(l)$ is proportional to the loss; consequently, the tails are penalized both by the presence of an extra-uncertainty and by a higher $\sigma(l)$.

Then, the penalized loss L_P is a new random variable, defined conditionally with respect to the loss value *l* obtained from the risk assessment. Given this, the conditional cumulative complementary distribution of L_P can be written as:

$$G(l_P \mid l) = 1 - F(l_P \mid l) = P(L_P > l_P \mid L = l) \tag{2.13}$$

and the annual exceedance rate of L_P can be written as:

$$\lambda(l_P) = \int_l G(l_P \mid l) |d\lambda(l)| \tag{2.14}$$

An example is provided in Fig. 2.14, where the annual exceedance curve of a hypothetical CI has been penalized using different *PF* values. The blue curve corresponds to *PF*=0, i.e. the annual exceedance rate of *L* (Eq. 2.9), while the other curves represent the annual exceedance rate of the penalized loss L_P expressed in Eq. 2.13.

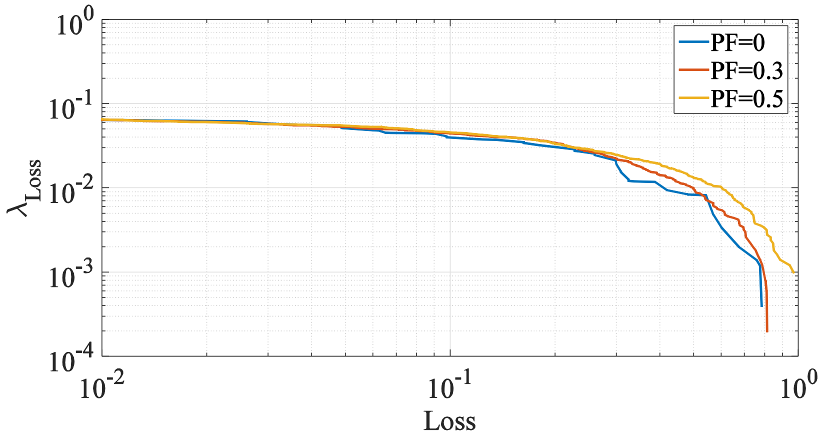

Fig. 2.14 Annual exceedance curves of penalized loss considering different penalty factor values

Discussion

There are some points of the penalty system that need further discussion and investigations as a part of future studies.

Firstly, the proposed penalty system requires levels of "detail and sophistication" (qualitative and/or quantitative) to be properly set by experts' consensus. Experts must have a clear idea about models and methods available in the scientific literature and their applicability to perform each step of the risk analysis. This may not be feasible for all perils that have to be considered for the stress test. Further, this evaluation should change in each stress test, reflecting the progress of the scientific research. The level of knowledge between two stress tests may change and the levels of "detail and sophistication" scheme should reflect this change.

Secondly, the computation of the *Effective Level* (Eq. 2.5 and 2.6) does not take into account the level of "detail and sophistication "associated to the approach adopted for the multi-risk analysis. This is because the current level of knowledge does not allow ranking these approaches, even though different multi-risk methods have been proposed recently.

Finally, the distribution of the penalized loss has been selected as a Lognormal distribution in this project. Other probability distributions, for example, a Gaussian distribution on the normal scale can be justified as well. Further studies on the determination of the appropriate distribution of the penalized loss should be done.

3. Incorporating ST@STREST into the life cycle management of non-nuclear critical CIs

3.1 Introduction

Structures and civil infrastructure systems are subjected to time-varying environmental stressors. These stressors can be low-consequence persistent stressors such as aging, fatigue or corrosion, as well as high-consequence low-probability-of-occurrence stressors such as natural or man-made disastrous events. Both types of stressors may induce huge economic losses and result in significant environmental impacts on the community these CI systems serve. In order to increase the long-term performance of such systems against rare events and long-term degradation process, it is very important to implement adequate strategies for maintaining such systems during their lifetimes.

These activities may include periodic inspections, maintenance and retrofit actions, structural health monitoring, and performance and risk analysis (Frangopol and Soliman, 2016). These actions are rationally scheduled along the life-cycle of the systems using a life-cycle management (LCM) procedure. Life cycle cost (LCC) and optimization tools are usually adopted to predict the performance of an infrastructure system subjected to long-term degradation process during its lifetime and to plan maintenance interventions. In particular, the performance profile (performance indicator graphed against time) resulting from the life cycle analysis allows planning the necessary interventions (maintenance, inspection and repair) in order to maintain the structural performance at an acceptable level. Establishing the best schedules requires a robust optimization process. The complexity of this process depends on the scale of the problem and on the type of deterioration phenomena considered (long-term processes and/or extraordinary events).

A brief overview of different aspects of LCM (i.e. life cycle analysis and cost optimization, degradation processes and modelling as well as the role of structural health monitoring and inspection techniques in supporting life cycle management decisions) may be found in STREST Deliverable 5.3 (Esposito and Stojadinovic 2016b).

3.2 LCC including natural hazard risk

The main aim of a LCC is to predict the performance of a CI system subjected to all environmental stressors during its lifetime. However, it is noted that the seismic risk analysis, and natural hazard risk analysis in general, has not devoted enough attention to the structural maintenance optimization problem (Furuta et al, 2011), although some examples exist. In regions exposed to frequent catastrophic natural events, LCC optimization analysis should account for the effects of these hazards. The model proposed by Chang and Shinozuka (1996) represents one of the first attempts to include natural hazard (in particular seismic risk) in the LCC framework. The framework is shown in Fig. 3.1. It includes two innovative aspects:

i. First, in addition to the initial costs of construction and costs attributed to maintenance action, the costs due to service interruption are considered. The latter are called “user costs” that represent the societal costs that are imposed when the functionality of a system is reduced mostly during the routine maintenance work or the retrofit action. For example, during a maintenance intervention of a bridge, the serviceability of the road network (flow of goods and people) is reduced, imposing an increment of travel time for each user. The total extra travel cost due to the maintenance action represents the user cost.
ii. Second, the expected costs associated to seismic risk of a CI system (i.e. discount cost for seismic retrofit and damage/repair costs) during the lifecycle of a structure or a system are combined with the initial capital and discounted maintenance cost.

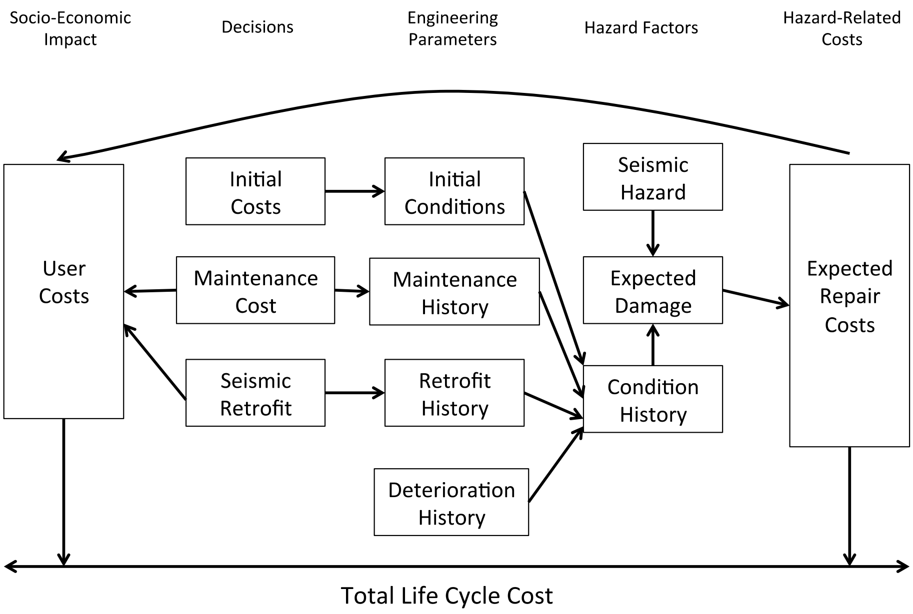

Fig. 3.1 Life-Cycle Cost framework including natural hazard risks adapted from Chang and Shinozuka (1996)

The life-cycle costs *C* is divided in four categories as expressed in the following equation:

$$C = C_1 + C_2 + C_3 + C_4 \tag{3.1}$$

where C_1 and C_2 represents the planned costs, C_3 and C_4 the unplanned costs.

Planned costs to owners (C_1) involves initial construction, subsequent expected discounted maintenance costs and discounted seismic retrofit costs, considering that a seismic retrofit is only applied once in the lifetime of a structure. In addition to planned costs paid by the owner of the structure/infrastructure system, maintenance and seismic retrofit actions may also impose user costs (C_2) due to the interruption of normal service (e.g. travel delay in a road network). This cost imposed to the society is function of the extent and the duration of the usage disruption during the maintenance activities and the retrofit action.

In addition to the costs associated to maintenance and design choices (planned costs), the framework includes unplanned life cycle costs related to the structural performance and associated repair costs due to a seismic event. Unplanned costs to owners (C_3) consist of expected discounted repair costs of earthquake damage over the life span of the structure. These costs are evaluated performing a probabilistic seismic risk/performance analysis of the system, conditional to its physical state at time t. The performance evaluation changes over time due to natural deterioration as well as mitigation actions. The unplanned hazard-related user costs (C_4) constitute the final category of this life-cycle cost framework and they are also based on a probabilistic condition/performance analysis of the system under study. These user costs are related to the service disruption due to earthquake damage and repairs and depend on the expected duration of repair/reconstruction activity over the life span of the structure.

3.3 Unified life cycle management of CI

Through the life cycle of the CI, systems operators have the objective to maintain the infrastructure systems and mitigate degradation of system components over time all the while achieving an economically justified operation of the system. To this aim, LCC and optimization tools are usually adopted to predict the performance of an infrastructure subjected to long-term degradation process during its lifetime and planning maintenance interventions. However, in regions exposed to natural events, LCC analysis should also take into account the effects of extreme natural events that may increase the probability of failure or loss of functionality during their lifetime.

The multi-level framework ST@STREST has been proposed with the aim of providing a multi-level systematic and harmonized approach for the evaluation of the performance of these systems against extreme and disastrous natural events.

In order to increase and optimize the long-term performance of CIs, the outcomes and findings of a stress test (e.g. results of risk analysis and identified risk mitigation strategies) should be included in the long-term maintenance plan of a CI. Results of the risk analysis (i.e. Assessment phase, Phase 2) in terms of system performance and expected costs of natural events may be incorporated in a LCC analysis and optimization problem. Furthermore, the evaluation of risk reduction strategies (Decision Phase) may make it possible to reconsider the full management and maintenance plan of the CI itself.

Therefore, the possibility to include the data on the current state of a CI in the aftermath of an actual disastrous event is another important aspect of the proposed framework. The state of civil infrastructures after the occurrence of a natural event is usually assessed through rapid visual inspection or automatic screening tools (e.g. close-circuit television). Through the use of standardized survey forms (e.g. EERI, 1996), data on the typology, location, component's features and the assessed physical damages are then collected to provide an estimate of the extent of the service disruption, costs and repair times and to define the repair/replacement strategy to apply. At the same time, the processing of these data can be useful to update the state condition history of the inspected components of the CI and for estimating and/or updating of the performance prediction models used in the risk analysis.

In this section, a framework to integrate stress test outcomes and findings and data gathered from post-event damage survey into a unified life-cycle management strategy is proposed and discussed. In particular, an extended version of the model proposed by Chang and Shinozuka (1996) is proposed. The proposed framework aims to include the stress test outcomes (i.e. loss curves, safety assessment and risk mitigation strategies) and information that can be retrieved from post-event damage survey into a life cycle cost evaluation and optimization procedure.

3.3.1 Life cycle analysis including stress test and post-event data

In order to optimize the life-cycle costs in a CI management strategy, the outputs of a stress test are going to be considered in the proposed framework.

As shown in Fig. 3.2, the outcomes of a Stress Test have an impact on:

- *Expected damages*: unplanned life cycle costs related to the structural performance and associated repair costs due to extreme natural events. A stress test allows to evaluate the performance of the CI against extreme natural events (according to the ST-Level adopted). In this way it is possible to quantify the expected costs caused by extreme natural events and then evaluate the associated unplanned owner and user costs (C_3 and C_4 in equation 3.1) to be included in the LCC analysis and optimization.
- *Mitigation history*: another outcome of a stress test is represented by the evaluation of risk reduction strategies based on a disaggregation analysis (Decision Phase, Phase 3). A disaggregation analysis is aimed at obtaining the probability that a specific value of a variable involved in the risk assessment is causative for the exceedance of a loss value of interest. The loss may be disaggregated with respect to system's response, which may help identifying the component the damage of which most likely causes the exceedance of the loss value of interest. Then, risk

mitigation strategies are formulated based on the results of the disaggregation analysis with the aim of increasing the long-term performance of CIs.

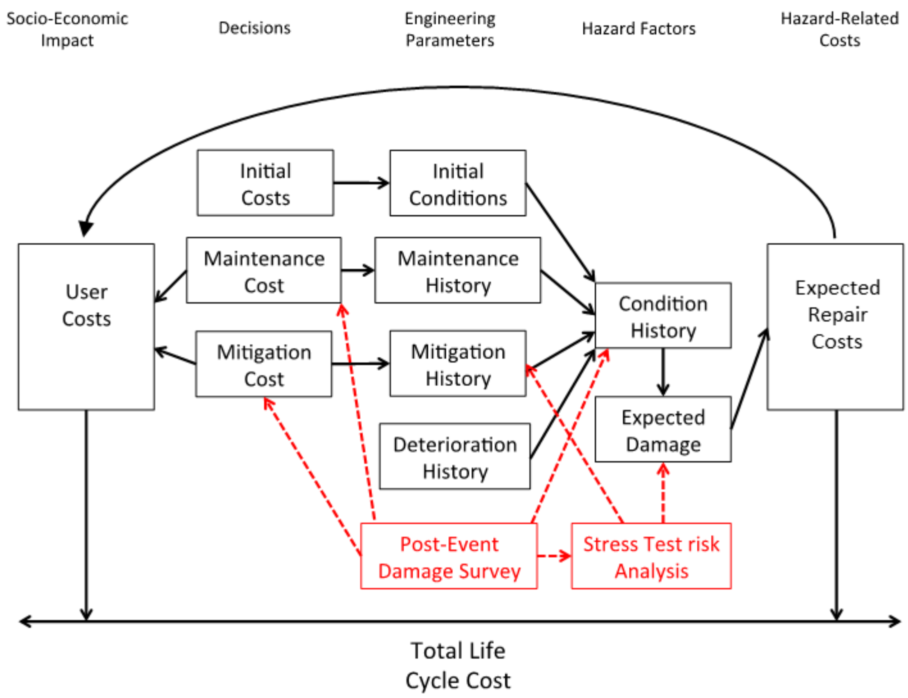

Fig. 3.2 Proposed framework for assimilating stress test and post-event data in a total life cycle cost analysis

In order to demonstrate how the outputs of a stress test may be incorporated in the life cycle management of a CI, the proposed framework was applied to the case study of L'Aquila (Italy) gas network. A Stress Test Level 2a was performed on the L'Aquila network as it was before the 2009 earthquake event to assess the performance of the network due to earthquake hazard. Risk is expressed in terms of annual probability of exceedance of service disruption levels, measured by a connectivity-based performance indicator (PI), i.e. the Connectivity Loss CL. Risk boundaries for the case study were defined in terms of F-N limits, according to the equation reported in STERST Deliverable 5.1 (Esposito et al 2016). Then, a disaggregation analysis was performed and possible risk mitigation strategies were identified. Finally, in order to evaluate the consequences of the risk reduction actions (e.g. seismic retrofit of some components of the gas network), the seismic performance of the gas network was assessed again, and results of the risk analysis were compared with the risk objectives identified at the beginning of the stress test.

Results in terms of annual exceedance curve of the assessed performance loss considering three mitigation actions are shown in Fig. 3.3.

More details on this application study are presented in STREST Deliverable 5.3 (Esposito and Stojadinovic 2016b).

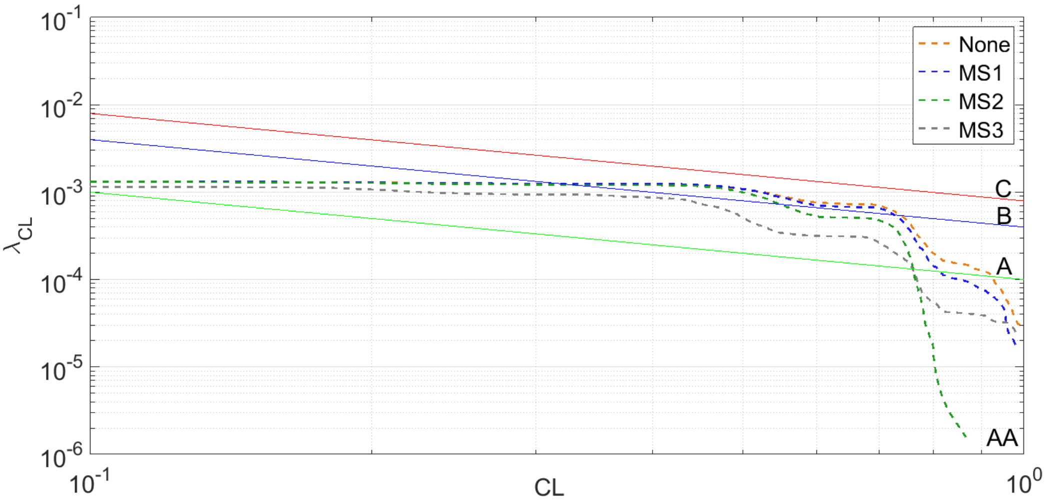

Fig. 3.3 Annual rate of exceedance of CL considering mitigation strategies applied to the stations (MS1 and MS2) and to buried pipelines (MS3)

Another important aspect of the proposed framework is represented by the use of collected data on the state of a CI in the aftermath of a disastrous event. As shown in Fig. 3.2, the information gathered for post-event inspection and survey has an impact on:

- *Condition history*: after a disastrous even, new information about the CI is available. Through the collection and the processing of on-site data the state condition of the inspected components of the CI may be updated.
- *Risk analysis*: information gathered after the occurrence of a natural disaster can be used to estimate and/or update performance model parameters adopted in the risk analysis through the use of statistical regression methods or the more advanced Bayesian approaches.
- *Mitigation history*: the main purpose of post-event damage surveys is to assess the functionality of system's components and the repair/replacement strategy to apply.
- *Maintenance costs*: the updated state condition of the CI may be used to redefine the intervention maintenance schedule, i.e. to determine whether a maintenance action is needed or not.

Information gathered after the occurrence of a natural disaster can be of extraordinary importance for the estimation and/or updating of performance prediction models adopted in the risk analysis. Through the gathering and the processing of the post event damage data, it is possible to derive empirical estimate of performance models (Basoz et al, 1999, Shinozuka et al, 2000; O'Rourke and So, 2000).

Statistical regression methods or more advanced Bayesian approaches can be used to estimate model parameters. In particular, Bayesian procedures are adopted to update model parameters estimates when new data becomes available, combining the likelihood function with the prior information on these parameters (Straub and Der Kiureghian, 2008). This approach has also the ability to handle all types of information and to include engineering expert opinion through a prior distribution.

An example of empirical estimation and Bayesian updating of a fragility model for buried pipelines is provided in STREST Deliverable 5.3 (Esposito and Stojadinovic 2016b). Pipeline damage data retrieved after the 2009 L'Aquila earthquake were used to estimate a fragility function for buried steel pipes caused by seismic ground shaking. In particular, a Bayesian estimation model along with the use of Importance sampling technique for numerical efficiency has been adopted to estimate the parameters of the fragility function considering as a-priori distribution of the model parameters a non-informative one. Fig. 3.4 shows the results of the Bayesian estimation for all sets of simulations. The results of the estimations were compared with a pipeline fragility relation considered suitable (in terms of pipe material and diameter) for the L'Aquila gas network., i.e. the ALA (2001) for steel arc welded pipes.

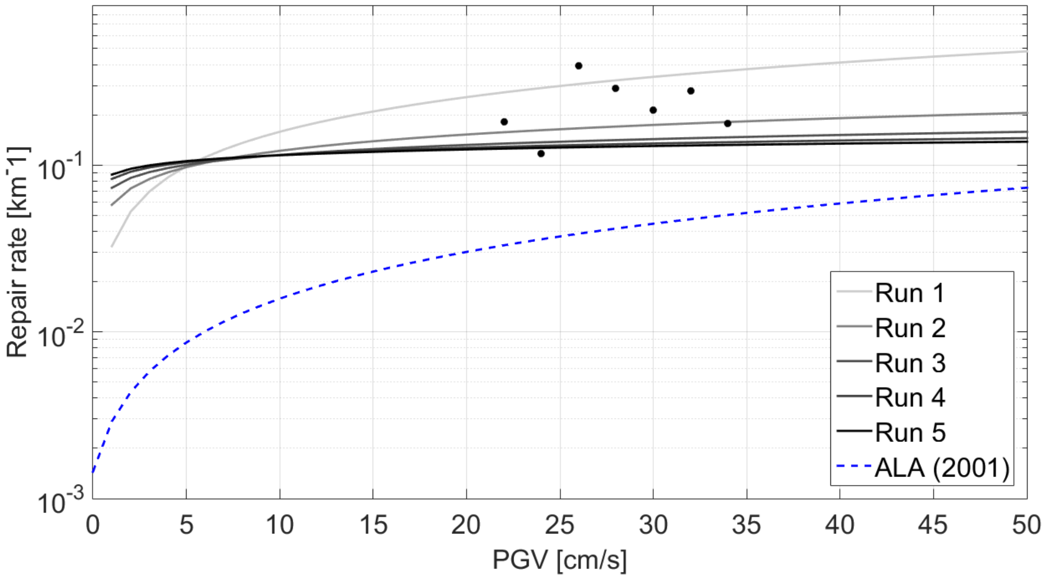

Fig. 3.4 Comparison of existing and updated fragility curves for L'Aquila gas steel pipes

This case study shows that valuable data from the model-based stress tests and actual post-event investigations can be inserted into a total life cycle cost analysis and that the effects of high-consequence low-probability events can be combined with the effects of low-consequence persistent degradation processes in a comprehensive model to better plan the life-cycle management of critical civil infrastructure systems.

3.4 Discussions

In regions exposed to natural events, LCC analysis should also take into account the effects of extreme natural events that may increase the probability of failure or loss of functionality during their lifetime. However, very few studies have been focused on the possibility to include the risk associated to extreme natural events in a LCCframework.

Stress tests for civil infrastructure systems have been proposed in with the aim of providing a multi-level systematic and harmonized approach for the evaluation of the performance of these systems against extreme and disastrous natural events. In particular, the ST@STERST multi-level framework has been proposed to verify the risk of CI systems respect to extreme natural events and to support decision makers in the evaluation of strategies to improve the performance of CIs along the life cycle. Each Stress Test Level is characterized by different objectives (component or system) and by different levels of risk analysis complexity (starting from design codes and ending with state-of-the-art risk analyses, such as modeling cascading failures). This makes the stress test adaptable to different hazard contexts and application to a broad range of civil infrastructure systems. Further, the level of complexity is tuned accordingly to types of critical infrastructures, the potential consequence of failure of the CIs, the types of hazards, and the available resources for conducting the stress tests.

A possible framework to integrate the results of stress tests and the data retrieved after disastrous events into a unified life-cycle management strategy of CIs has been introduced in order to manage both long degradation and instantaneous natural hazard-induced stresses during the lifetime of a civil infrastructure system.

In particular, results of the risk analysis conducted in the scope of a stress test in terms of system performance and expected costs of natural events, may be incorporated in a LCC analysis and optimization problem. Further, the evaluation of risk reduction strategies resulting from a loss disaggregation may make it possible to reconsider the full management and maintenance plan of the CI itself.

On the other hand, the evaluation of the state of civil infrastructures after the occurrence of a natural event, and the collection and processing of post-event data, such as typology, location, component's features and the assessed physical damages, can be useful to update the state condition history of the inspected components of the CI and to estimate and/or update performance prediction models used in the risk analysis.

4. Using ST@STREST to enhance societal resilience

When CIs are affected by extreme natural events, such as earthquakes, floods, tsunami, etc., they are more and more often unable to quickly recover their functionality, either back to the pre-disaster original state, or just to a level sufficient to satisfy the post-disaster demand. With the increasing density and interconnectedness of the communities today, the demand on and the importance of the CIs is growing; thus, the consequences of CI failure (to meet the demands) can be devastating both from the standpoint of human life endangerment and from the economic standpoint. The post-disaster performance of CIs has a high impact on the coordination and the execution of emergency actions, and at the same time it influences both the long-lasting post-disaster recovery process of the community, and the eventual post-disaster resumption of community functions.

Today, after decades of development of probabilistic risk assessment techniques, there are solid probabilistic engineering risk assessment methods and tools that provide practical estimates of instantaneous CI performance (service) loss due to direct and indirect disaster-induced damage. However, the instantaneous loss by itself does not reveal how a community served by the CIs responds to a disaster. The *time dimension* represents a key aspect: the time-evolution of community needs and the ability of the CIs to fulfill these needs (e.g. water, gas, and electricity) is best represented and modelled using the concept of resilience rather than of risk.

The term resilience has increasingly been seen in the research literature in many fields, from psychology, biology, economy, social studies, and also engineering. Definition and modelling or disaster resilience of engineered systems is the topic of an increasing amount of recent research work. Nevertheless, there is still a substantial diversity among the definitions and the modelling of resilience. In this project, we will define *resilience* in general as "the ability (of CIs and communities) to prepare and plan for, absorb, recover from and more successfully adapt to adverse events" (The National Academy 2012).

Conducting a stress test to assess only the risk of a civil infrastructure system, i.e. to relate the losses with the total probability of their occurrence due to one or more hazards, does not provide enough information on the ability of the CI system to function and recover after a disaster. The system, and systems of systems that form the built environment of our society, are non-linear.

The ST@STREST framework proposed in Chapter 2 aims at evaluating the CI system risk from natural hazards. However, this framework was designed to also serve as a basis for the development of a new stress test concept that may support decision makers in the evaluation of strategies to not only decrease the risk exposure, but also to enhance the resilience of CIs against natural hazards.

It is clear that a new *resilience-oriented stress test methodology and framework* for civil infrastructure systems must include the recovery process and, furthermore, include models of how the systems function and deliver their service to the community, and how the community recovers its needs for such services. The ST@STREST framework was developed while keeping in mind such an extension, to make it possible to test the resilience of CIs to extreme events, i.e. to verify the capacity of CIs to anticipate, absorb and adapt to events disruptive to its function, and recover either back to its original state or another state consistent with the needs of the community during, and at the end of the post-disaster recovery process.

The extension of the proposed framework requires the pursuit of two main goals:

- Identification of resilience metrics and standardized methodologies to model the resilience of CIs; and
- Understanding how stakeholders' needs depend on CIs, defining resilience-based acceptance criteria.

Regarding the identification of resilience-based acceptance criteria, understanding how the different stakeholders' needs depend on the functionality of the CIs represents the key issue. Business activities need suitable facilities and their supply chains and delivery networks; everyone needs a transportation network, electricity, water, gas, and communication networks but, in the aftermath of a disastrous event, some of these services (e.g. water) are more needed than others. There are also different, and competing priorities for services to critical facilities (e.g. hospitals). Reconciling these factors to develop CI resilience acceptance criteria, taking into account not only the instantaneous losses but also the time evolution of the CIs and the community systems during the recovery process, is not trivial.

In this chapter these aspects will be argued in more details. A detailed overview of a possible approach to integrate the evolution of the CI performance in time in the ST@STREST framework is presented in STREST Deliverable 5.4 (Esposito and Stojadinovic 2016c) and STREST Deliverable 4.5 (Stojadinovic and Esposito 2016).

4.1 Modelling resilience of critical infrastructures against natural hazards

The probabilistic resilience assessment of CIs is gaining increasing importance in a research effort toward assessing the risk and resilience of communities to natural hazards because the CIs are essential to the functioning of a community.

Several definitions of resilience have been offered in various disciplines. Many of them are similar and they overlap with existing concepts as robustness, flexibility, agility, etc. The concept of resilience has been also approached across application domains, including psychology, ecology, enterprises, and engineering, among others. In the engineering domain, in particular in the subdomain of infrastructure systems, the National Infrastructure Advisory Council (NIAC, 2009) defined the resilience of infrastructure systems as "the ability to predict, absorb, adapt and/or quickly recover from a disruptive event such as natural disasters". Infrastructure systems are also considered as subdomain of social science, in which lack of CI resilience can lead to huge consequences on communities.

In the civil infrastructure domain, in a field-defining paper, Bruneau et al (2003) conceptualized the resilience as a metric that "can be understood as the ability of the system to reduce the chances of a shock, to absorb a shock if it occurs, and to recover quickly after a shock". The authors defined four dimensions of resilience in the well-known resilience triangle model: 1) robustness, the strength of the system, 2) rapidity, i.e. the speed at which the system could return to its original state or at an acceptable level of functionality, 3) resourcefulness, the level of capability in applying material and human resources to respond to a disruptive event, and 4) redundancy, the extent to which carries by a system to minimize the likelihood and the impact of disruption.

Bruneau et al (2003) proposed a deterministic static metric of the resilience loss of a community with respect to a specific event, as the expected degradation in quality (probability of failure), over time (that is, time to recovery), R, formulated in the following equation.

$$R = \int_{t_0}^{t_1} \left[100 - Q(t)\right] dt \qquad (4.1)$$

where Q(T) represents the quality of the system, t_0 the time when the specific damaging event occurs and t_1 is the time when the restoration of the system is completed (indicated by a quality of 100%). The notion of system quality was left open to interpretation, but is often understood as the ability of the system to perform, which, in the case of CIs, may mean the quantity of delivered service.

Following the Bruneau et al (2003) pioneering study, considerable attention has been focused towards developing frameworks to assess the resilience of civil facilities or infrastructures; among these, notable works are: Chang and Shinouzuka (2004), Franchin and Cavalieri (2015), Bocchini et al (2014), Francis and Bekera, (2014), Broccardo et al (2015), Iervolino and Giorgio (2015), Sun et al (2015a,b).

Among them, the most innovative are Francis and Bekera, (2014) and Broccardo et al (2015)

Francis and Bekera (2014) proposed a resilience factor as quantitative metric of an infrastructure system's resilience. This factor depends on a speed recovery factor, the original stable system performance level (pre-disaster performance of the system), the performance level immediately post-disruption (before recovery starts), and the performance at a new stable state level (after recovery efforts have been exhausted). The speed recovery factor is defined as a function of the time that is acceptable to elapse after a disaster before recovery starts, the time to complete initial recovery actions and the time to final recovery.

Broccardo et al (2015) investigated all the statistical assumptions and limitations to integrate the quantification of seismic resilience of a given civil facility or system in a stochastic Markovian framework. In particular, the study revisited the PEER framework formula by imposing the resilience of a facility as a final decision variable, analyzing then the limitations and the range of applicability evaluating the probability of interaction between the recovery time and the inter-arrival time of seismic events.

However, despite the increasing importance of the role of system resilience in various disciplines of system engineering, and the recent efforts by many authors, there is a substantial diversity among the definitions and the modelling of resilience (Hosseini et al, 2016, Henry and Ramirez-Marquez, 2012, Ouyang and Duenas-Osorio 2012, Bruneau et al, 2003).

As also reported by Bruneau et al (2003) "there is no explicit set of procedures that suggests how to quantify resilience in the context of earthquake hazard, how to compare communities with one another in terms of resilience, or how to determine whether individual communities are moving in the direction of becoming more resilient in the face of earthquake hazards."

Communities and infrastructure systems are complex systems of systems. Modelling their resilience against natural hazards in a probabilistic way and with a single metric is not quite straightforward. Further, most of the resilience quantification frameworks proposed in literature impose the point of view of the infrastructure owner, i.e. to recover the initial functionality of the system as fast as possible. However, CIs are built to deliver a service to a community, then the resilience assessment should also take into account the ability of a CI to supply the time-varying community demand for the services provided by the assessed CI (Mieler et al, 2015). A CI resilience quantification framework needs to explicitly account for the evolution of the supply (i.e. the service supply capacity of the system) and for the evolution of the demand of the community and other CIs for its services in the aftermath of a disaster.

To this end, a compositional demand/supply resilience quantification framework to evaluate the post-disaster resilience of CISs that supply their services to satisfy the demand of a community was recently proposed (Dider et al, 2015; Sun et al 2015a, 2015b; Didier et al, *in prep*). The framework accounts explicitly for the evolution of the demand of a community and the demand of other CIs during the post-disaster recovery process. The framework consists of three main elements:

1. The evolution of the potential demand for the service of the investigated CI over time after a disaster. The potential demand is the amount of demand of all consumers of the service of the assessed CI, if there were no limitations on the supply side (i.e. assuming an unlimited supply of service). Consumers include, for example, the community (composed by its residential building stock, industries,

businesses and critical facilities, used by the population) and all other CIs (e.g. electric power demand of the water supply system in order to run water pumps). Potential demand depends on:

- the vulnerability of the components of the set of demand systems (e.g. a community and/or another CI) during the loss accumulation and absorption phase of a disaster; and,
- the recovery of the components of the set of demand systems during the recovery phase after a disaster.
- potential extraordinary or high-priority needs in the aftermath of a disaster (e.g. hospital, telecommunications networks)

2. The evolution of the potential supply for the service of the investigated CI over time after a disaster. The potential supply is the amount of service supply available to satisfy the demand of the system. Potential supply depends on:
   - the vulnerability of the components of the service supply and distribution systems during the loss accumulation and absorption phase of a disaster; and
   - the recovery of the components of the service supply and distribution systems during the recovery phase after a disaster.
3. A system operation model, regulating the allocation (or dispatch) of the service supply in order to satisfy the demand of the consumers. It accounts for the capacity limitations and interactions of the different elements of the CI: the service production system, the distribution system, the technical functioning and control of the system, and the system or network effects. These include, for example, the topology of the system, operator service allocation policies, or possible demand distribution strategies.

The compositional resilience quantification framework allows for the assessment of the resilience of a combined, interacting set of demand/supply CI systems. CI system resilience is the time-varying ability to cover the demand for its services, while subjected to disruptive events that may occur over the system's lifetime. The framework allows, thus, to account for the impact of a disaster on both the demand and the supply side and to track the post-disaster evolution of demand and supply at both component and system levels. The evolution of the demand, supply and consumption after a disaster is sketched in Fig. 4.1. The shaded area represents the integral lack of resilience, i.e. the time during which the demand of the community is larger than the service provided by the CI. Also shown are the measures of time required to satisfy the demand, and times needed to fully recover the demand and the supply to the pre-disaster level, accounting for some reserve margin.

One advantage of the proposed compositional resilience quantification framework is the component and system level evaluation strategy, which follows the ST@STREST methodology. Another advantage is the explicit account of the evolution of demand, supply and consumption of a CI service. This makes it possible to treat event sequences, such as aftershocks, community depopulation, such as the aftermath of the 72 AD Pompeii volcanic eruption, permanent demand changes, such as the effect of the 1995 Kobe earthquake on the recovery of Kobe port operations or the effect of demand surges, such as those that occur on the cellular phone network immediately after a disaster.

A detailed overview of this framework is presented in STREST deliverable 4.5 (Stojadinovic and Esposito, 2016).

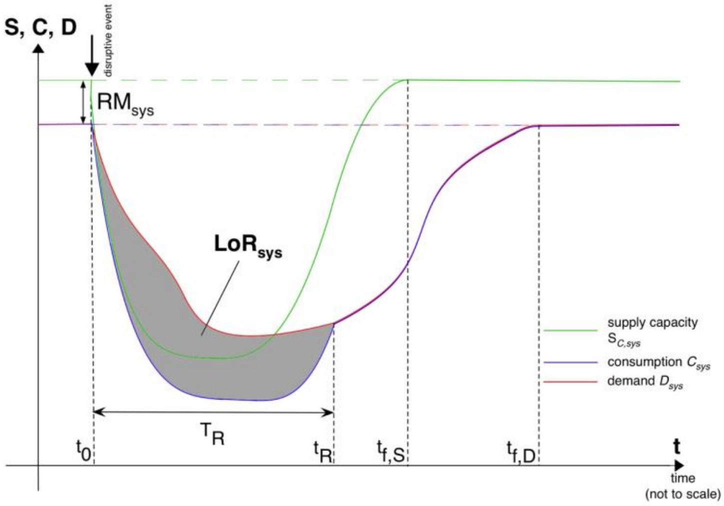

Fig. 4.1 Representation of the network supply, demand and consumption curves after a natural disaster event (STREST Deliverable 4.5 (Stojadinovic and Esposito 2016))

4.2 Resilience-based stress test for critical non-nuclear infrastructures

In this section the ST@STREST framework proposed in Chapter 2 is analyzed to define a new concept of stress test aimed at testing the resilience (and not only the risk) of CIs to extreme events and comparing the probability of loss of resilience (not just the probability of instantaneous losses) to acceptable levels.

A resilience-based stress test concept may support decision makers in the evaluation of strategies to improve the capacity of CIs to anticipate, absorb and adapt and/or quickly recover from a disruptive event.

In order to define a new concept of stress test aimed at verifying and mitigating the resilience of CIs, some aspects of the four-phase ST@STREST workflow have to be reviewed and the scope of each phase of the methodology modified, in particular:

- Pre-Assessment phase (Phase 1)

  The collection of all data available on the CI also has to include all the information required for the estimation of the resilience metrics selected for the assessment, e.g. the information on the rate of recovery, conditioned on the incurred damage state (the recovery curves, a counterpart to vulnerability curves), the funds, materials and manpower availability for the recovery and restoration process, the pre- and post-event demand patterns for the service of the investigated CI, the characterization of the community the CI serves, and the operation models of the CI in both normal and emergency conditions.

  Further, resilience-based objectives/acceptance criteria have to be defined for each resilience metric and according to the specific perspective considered, i.e. the network operator and/or the community the CI serves. Here, the competing interests of the network operator (e.g. maximizing profit) and the community (e.g. minimizing disruption to the population) need to be reconciled in an aggregate acceptance criterion.

- Assessment phase (Phase 2)

  In the Assessment phase, the resilience (and not only the performance) of each component of the CI (Component analysis) and the whole system (System Analysis) should be evaluated according to the ST-Level selected. One possible way to perform this task is to use the compositional demand/supply resilience quantification framework (Didier et al, in prep).

More efforts should be devoted to develop standardized methodologies aimed at verifying the resilience of CIs in the natural hazard context, both at component and system level.

- Decision Phase (Phase 3)

  The results of the resilience assessment are compared with the objectives defined in the pre-assessment phase and resilience mitigation strategies and guidelines are formulated. An effort to disaggregate the resilience of a CI-Community system to find which elements and systems and which events cause the largest amount of impact.

- Report phase (Phase 4)

  Results of the resilience analysis and mitigation strategies are presented to CI authorities, regulators and representatives of the community. An effort to communicate resilience (and its probabilistic nature), building on the ongoing work on communication or risk, should be undertaken.

4.2.1 Future research and discussions

The extension of the proposed framework requires the pursuit of two main goals:

- Identification of resilience metrics and development of standardized methodologies to model the resilience of CIs; and
- Definition of resilience-based acceptance criteria, understanding how community's needs depend on critical infrastructures.

Definition and modelling of disaster resilience of engineered systems is the topic of an increasing amount of recent research. Nevertheless, there is still a substantial diversity among the definitions and the modelling of resilience. In particular, there is no standardized approach that suggests how to quantify the natural disaster resilience of CIs in the context of natural hazard. As future research, we foresee the need of defining a taxonomy of resilience metrics mainly based on the following aspects:

- The identification of quantifiable time-dependent system delivery functions that specify the system functionality of the infrastructure system under study, such as the flow, the connectivity, the time delay, etc.
- The modelling of interdependencies between networks within a community.
- Including the social perspective in the definition of resilience metrics accounting for time-varying community demand for the services provided by the assessed CI.

The definition of resilience metrics requires a deep understanding of the CI's functionality and the parameters that are important for both the operator and the society the CI serves.

An example of possible resilience metrics for a gas distribution network (Bellagamba, 2015) is provided in Fig. 4.2. In this case, the system functionality is expressed in terms of daily gas flow. The resilience metrics are identified comparing the required system functionality by the community (demand, red line) to the effective capacity of the network after an earthquake (blue line), in particular:

- The non-supplied demand $S_{Nonsupplied}$, defined as the area between the capacity and the demand curves when the demand curve is above the capacity curve.
- The recovery time of the gas distribution network $T_{Recovery}$, defined as the time needed for the gas distribution network to recover its full functionality.

- The time needed for the capacity to be equal or greater than the demand, called resilience time, $T_{Resilience}$.

Further, according to the different possible metrics, standardized approaches aimed at modelling the resilience of non-nuclear CIs should be identified and/or developed. This implies, first a review of the existing approaches in the field of quantitative risk analysis and a classification based on, for example, the use of analytic or simulation-based approaches for the quantification of the aleatory uncertainties, the inclusion of inter-dependencies with other infrastructure systems and the interaction with the community the CI serves, etc.

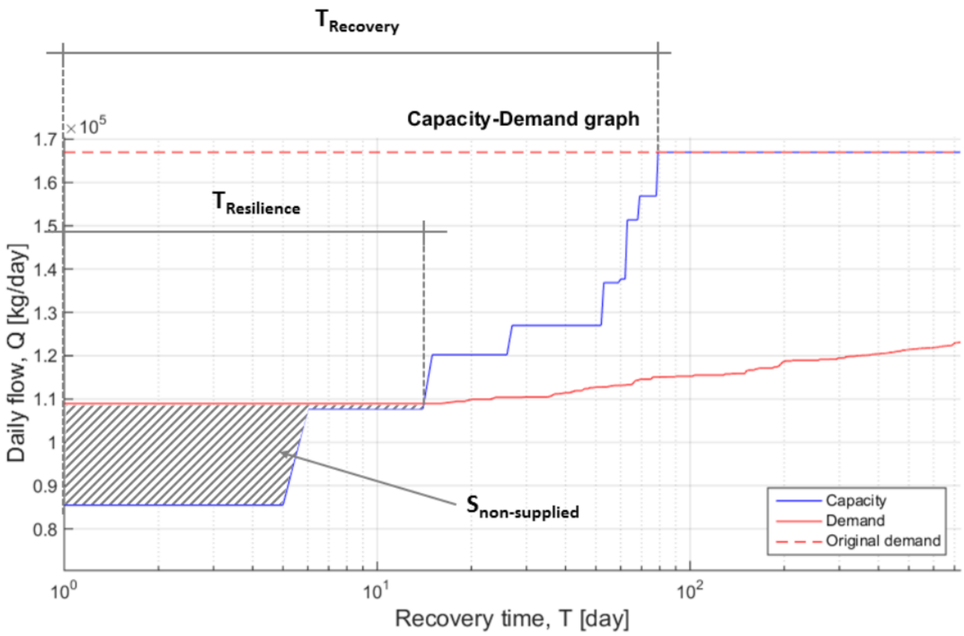

Fig. 4.2 Resilience metrics defined for a gas network, from Bellagamba (2015)

Another important aspect for the development of a resilience-based stress test concept is represented by the definition of acceptance resilience-based criteria to be identified in the Phase 1 of the workflow. The key question to be answered is:

- When and how do the CI systems need to be restored before adversely affecting the different stakeholders, (e.g. community, the infrastructure operator)?

Understanding how community's needs depend on the functionality of the CIs (now and in the future) is the key. Business activities need suitable facilities and their supply chains and delivery networks; everyone needs a transportation network, electricity, water, gas, and communication networks but, in the aftermath of a disastrous event, some of these services (e.g. water) are more needed than others.

A first attempt toward this direction is represented by the report published in April 2015 by the National Institute of Standards and Technology (NIST, 2015): "Community Resilience Planning Guide for Buildings and Infrastructure Systems". The Guide provides a methodology for a local government, as the logical convener, to bring together the relevant stakeholders and incorporate resilience into the long-term community development planning processes. In particular, it identifies the ways social organizations depend on buildings and infrastructure systems to help support community recovery by establishing recovery sequencing and the degree of functionality needed in the built environment at different points in time after a hazardous event. The guide also provides examples of resilience goals that communities might set for their social institutions.

Further examination of extending the stress test concept to societal resilience is presented in STREST Deliverable 4.5 (Stojadinovic and Esposito 2016).

However, developing a CI resilience-targeted stress test is, as of today, beyond the state of the art. Foremost, there is a need to develop a harmonized definition of societal resilience applicable to a CI-community system. Second, a set of societal resilience targets need to be established and transformed into acceptance criteria for the CI systems and their elements. Third, a transparent method for modeling and evaluating the CI resilience needs to be established. Only then could a stress tests targeting the resilience of a CI system be constructed. The ST@STREST methodology and framework that targets CI vulnerability, developed in this project, can be used as the prototype for such CI resilience-targeted stress test.

5. Conclusions and recommendations

In the context of the STREST project, a harmonized approach for stress testing critical non-nuclear infrastructures, ST@STREST, has been developed. The aims of the ST@STREST methodology and framework are to quantify the safety and the risk of individual components as well as of whole CI system with respect to extreme events and to compare the expected behavior of the CI to acceptable values.

In particular, a multi-level stress test methodology has been proposed. Each level is characterized by a different scope (component or system) and by a different level of risk analysis complexity (starting from design codes and ending with state-of-the-art probabilistic risk analyses, such as cascade modelling). This allows flexibility and application to a broad range of infrastructures The framework is composed of four main phases and nine steps. The goals, the method, the time frame, and the total costs of the stress test are defined in the Pre-Assessment Phase. In the Assessment Phase, the stress test is performed at component level and system level. The outcomes of the stress test are checked and analyzed in the Decision Phase. Finally, the results are reported and communicated to stakeholders and authorities (Report Phase). The ST@STREST data framework, used to store and manage the data about the CI under test, is also flexible, in that it allows the use of data structures that support frequentist (event and fault trees, bow ties) and belief-based notions of probability.

Further, in order to allow transparency of the stress test process, the data, models, methods adopted for the risk assessment and the associated uncertainty are clearly documented and managed by different experts involved in the stress test process. This allows to define how reliable the results of the stress test are. In particular, a penalty system has been proposed to acknowledge the limitation of the methods and models used to assess the performance of the CI and eventually penalize the output of the risk assessment. In particular, the proposed system penalizes the results of the hazard and risk assessment of the conducted stress test by evaluating an extra uncertainty, here named penalty uncertainty, to amplify the uncertainties introduced by simplistic approaches that cannot guarantee a sufficiently accurate analysis.

The outcome of the stress test is defined using a grading system based on the comparison of the results of risk assessment with the risk objectives (i.e. acceptance criteria) defined at the beginning of the test. The proposed system is composed of three different outcomes: Pass, Partly Pass, and Fail. The CI passes the stress test if it attains grade AA or A. The former grade corresponds to negligible risk whereas the latter grade corresponds to risk being as low as reasonably practicable. The CI partly passes the stress test if it receives grade B, which corresponds to the existence of possibly unjustifiable risk. Finally, the CI fails the stress test if it is given grade C, which corresponds to the existence of intolerable risk. Guidelines for the grading of individual components are also proposed together with a generalization of the grading system to take into account epistemic uncertainties and system analysis.

The stress test approach proposed in this project addresses the vulnerabilities of CIs to catastrophic by rare (high-consequence low-probability) natural hazard events. An extension of the proposed ST@STREST methodology and framework to integrate the results of stress tests and the data retrieved after disastrous events with the data collected during every-day operation of the system and its degradation (low-consequence persistent events) into a unified life-cycle management strategy for CIs has been proposed. In particular, the results of the risk analysis conducted in the scope of a stress test in terms of system performance and expected costs of natural events, may be incorporated in a life-cycle cost analysis of the CI system and optimization of its operations and maintenance. Further, the evaluation of risk reduction strategies resulting from a loss disaggregation may make it possible to improve the full management and maintenance plan of the CI itself. Moreover, the evaluation of the state of civil infrastructures after the occurrence of a natural event, and the collection and processing of post-event data, such as typology, location, component’s features and the assessed physical damages, can be useful to update the state condition history of the inspected components of the CI and to estimate and/or update performance prediction models used in a future risk analysis.

An extension of the ST@STREST methodology and framework to evaluate not only the vulnerability but also the resilience of CIs, i.e. "the ability to prepare and plan for, absorb, recover from and more successfully adapt to adverse events" (The National Academy 2012) has also been proposed. This extension builds on the ST@STREST methodology by modelling the post-disaster recovery process of a CI system and by quantifying the lack of resilience and the attributes of a resilient system using a novel compositional supply/demand CI resilience quantification framework. This extension enables a new role of a stress test, that of examining the ability of a community and its CIs to bounce back after a natural disaster.

However, there are some points of the proposed ST@STREST methodology and framework that need to be discussed and further developed as a part of the future studies.

The stress test has been classified in three (macro) conceptual frameworks for the safety of non-nuclear CIs. The selection of the appropriate ST-L and sub-levels, made by the PM during the Pre-Assessment phase, depends on regulatory requirements that should account for the importance/criticality of the type CI. CIs are complex and diverse in nature. It is important to rank them, if the number of CIs being considered is greater than one for performing the stress test. The ranking of CIs is a challenging task due to their diverse nature, the potential consequence of failure, the types of hazards posing threat to them, vulnerability state etc. A criticality assessment of the CIs, aimed at identifying and ranking CIs (for example at a national scale), may represent a practical tool to support the choice of the appropriate ST-level.

The proposed penalty system requires "level of detail and sophistication" to be properly set by experts' consensus. Experts must have a clear idea about models and methods available in the scientific literature and their applicability to perform each step of the risk analysis. This may not be feasible for all perils that have to be considered for the stress test. Secondly, the computation of the penalty uncertainty does not take into account the complexity of the approach adopted for the multi-risk analysis. This is because the current level of knowledge does not allow ranking these approaches, even though different multi-risk methods have been proposed recently.

To establish a common grading system, the risk objectives of the risk measures across a range of interests should be harmonized on the European level. This is a task for regulatory bodies and for industry association: they should reconcile the societal and industry interest and develop mutually acceptable risk limits. Further, it is yet to be determined how grades of single components should affect the global outcome of stress test. Moreover, in case epistemic uncertainty analysis is of concern, it is currently recommended that the mean value of the designated risk measure is used. However, other options such as a grade based on other quantiles of the risk measure distribution should be investigated.

Finally, developing a CI resilience-targeted stress test is, as of today, beyond the state of the art. The ST@STREST methodology and framework that targets CI vulnerability, developed in this project, can be used as the prototype for such CI resilience-targeted stress test once CI and societal resilience is defined in a harmonized way, acceptable levels of resilience agreed on, and ways to transparently and consistently evaluate CI resilience developed and accepted in practice.

References

1. TNA (2012). Disaster Resilience: A National Imperative, The National Academies, Washington D.C. (www.nap.edu) DOI 10.17226/13457
2. Pitilakis K, Argyroudis S, Fotopooulou S, Karafagka S, Esposito S, Stojadinovic B, Giardini D, Dolšek M, Babič A, Selva J, Iqbal S, Volpe M, Tonini R, Romano F, Brizuela B, Piatanesi A, Basili R, Lorito S, Salzano E, Basco A, Crowley H, Rodrigues D, Matos JP, Schleiss AJ, Courage W, Reinders J, Akkar S, Uckan E, Erdik M (2016). Reference Report 5: Guidelines for stress-test design for non-nuclear critical infrastructures and systems: Applications. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards
3. Cornell C, Krawinkler H (2000). Progress and challenges in seismic performance assessment. PEER Center News, 3 (2)
4. Mignan A, Wiemer S, Giardini D (2014). The quantification of low-probability–high-consequences events: part I. A generic multi-risk approach, Nat. Hazards, 73, 1999-2022, doi: 10.1007/s11069-014-1178-4
5. Mignan A, Komendantova N, Scolobig A, Fleming K (2016a). Multi-Risk Assessment and Governance, Handbook of Disaster Risk Reduction and Management, World Sci. Press & Imperial College Press, London, chapter 16, in press
6. 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 project: Harmonized approach to stress tests for critical infrastructures against natural hazards
7. Selva J, Iqbal S, Cotton F, Giardini D, Esposito S, Stojadinovic B, Argyroudis S, Pitilakis K, Mignan A. The Epistemic Uncertainty Process (EU-P): management of critical choices and uncertainty in single and multi hazard/risk assessments based on multiple experts, *in prep*.
8. SSHAC (1997). Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts (No. U.S. Nuclear Regulatory Commission Report, NUREG/CR-6372), U.S. Nuclear Regulatory Commission Report, NUREG/CR-6372. Washington, D.C.
9. Esposito S, Stojadinovic B, Mignan A, Dolšek M, Babič A, Selva J, Iqbal S, Cotton F, Iervolino I (2016). Deliverable D5.1: Report on the proposed engineering risk assessment methodology for stress tests of non-nuclear CIs. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards
10. CEN (2005a). European Standard EN 1998-1: 2005 Eurocode 8: Design of structures for earthquake resistance. Part 1: General rules, seismic actions and rules for buildings, Brussels, Belgium.
11. CEN (2005b). European Standard EN 1998-3: 2005 Eurocode 8: Design of structures for earthquake resistance. Part 3: Assessment and retrofitting of buildings, Brussels, Belgium.
12. Babič A, Dolšek M (2016). Seismic fragility functions of industrial precast building classes. Engineering Structures, 118: 357-70.
13. Baker JW (2015). Efficient analytical fragility function fitting using dynamic structural analysis. Earthquake Spectra, 31: 579-99.

14. Dolšek M, Fajfar P (2008). The effect of masonry infills on the seismic response of a four storey reinforced concrete frame-a probabilistic assessment. Engineering Structures. 30 (11): 3186–3192.

15. Faber MH, Stewart MG (2003). Risk assessment for civil engineering facilities: critical overview and discussion. Reliability Engineering & System Safety. 80:173–184.

16. Bedford T, Cooke R (2001). Probabilistic Risk Analysis. 1st ed. Cambridge: Cambridge University Press, 2001. Cambridge Books.

17. 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

18. Cavalieri F, Franchin P, Buritica Cortes JAM, Tesfamariam S (2014). Models for seismic vulnerability analysis of power networks: comparative assessment, Computer-Aided Civil and Infrastructure Engineering. 29(8): 590–607.

19. Argyroudis S, Selva J, Gehl P, Pitilakis K (2015). Systemic Seismic Risk Assessment of Road Networks Considering Interactions with the Built Environment. Computer-Aided Civil and Infrastructure Engineering 30 (7): 524–540.

20. Esposito S, Botta A, De Falco M, Iervolino I, Pescape A, Santo A. Seismic Risk Analysis of a Regional Telecommunication Network, *in prep*.

21. 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 project: Harmonized apporach to stress tests for critical infrastructures against natural hazards

22. Kakderi K, Fotopoulou S, Argyroudis S, Karafagka S, Pitilakis K, Anastasiadis A, Smerzini C, Selva J, Giannopoulos G, Galbusera L, Courage W, Reinders J, Cheng Y, Akkar S, Erdik M, Uckan E (2015). Deliverable D4.2: Guidelines for performance and consequences assessment of geographically distributed, non-nuclear critical infrastructures exposed to multiple natural hazards. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards

23. Crowley H, Casotto C, Dolšek M, and Babič A (2015). Deliverable D4.3: Guidelines for performance and consequences assessment of multiple-site, low-risk, high-impact, nonnuclear critical infrastructures (exposed to multiple natural hazards, etc.). STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards

24. Bommer JJ, Scherbaum F (2008). The use and misuse of logic-trees in probabilistic seismic hazard analysis. Earthquake Spectra 96: 1967–1977.

25. Marzocchi W, Taroni M, Selva J (2015). Accounting for Epistemic Uncertainty in PSHA: Logic Tree and Ensemble Modeling. Bulletin of the Seismological Society of America, 105 (4), doi: 10.1785/0120140131.

26. Liu Z, Nadim F, Garcia-Aristizabal A, Mignan A, Fleming K, Luna BQ (2015). A three-level framework for multi-risk assessment, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 9, 59-74, doi: 10.1080/17499518.2015.1041989

27. Marzocchi W, Garcia-Aristizabal A, Gasparini P, Mastellone ML, Di Ruocco A (2012). Basic principles of multi-risk assessment: a case study in Italy, Nat. Haz., 62, 551-573, doi: 10.1007/s11069-012-0092-x

28. Selva J (2013). Long-term multi-risk assessment: statistical treatment of interaction among risks, Nat. Hazards, doi: 10.1007/s11069-013-0599-9.

29. Iervolino I (2016). Soil-invariant seismic hazard and disaggregation. Bulletin of the Seismological Society of America. DOI: 10.1785/0120160072 (in press).

30. Mignan A, Danciu L, Matos JP, Schleiss A (2015). Deliverable D3.5: Report on cascading events and multi-hazard probabilistic scenarios. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards.

31. Matos JP, Mignan A, Schleiss AJ (2015). Vulnerability of large dams considering hazard interactions: conceptual application of the Generic Multi-Risk framework. In: 13th ICOLD International Benchmark Workshop on Numerical Analysis of Dams. Lausanne, Switzerland, 285-292

32. Mignan A, Scolobig A, Sauron A (2016b). Using reasoned imagination to learn about cascading hazards: a pilot study, Disaster Prevention and Management, 25 (3), 329-344, doi: 10.1108/DPM-06-2015-0137.

33. USNRC (2012). Probabilistic Risk Analysis. Retrieved from http://www.nrc.gov/about-nrc/regulatory/risk-informed/pra.html 31.07.2016.

34. Vesely WE, Goldberg FF, Roberts NH, Haasl DF (1981). Fault tree handbook (NUREG-0492). United States Nuclear Regulatory Commission, Washington, DC.

35. De Dianous V, Fiévez C (2006). ARAMIS project: A more explicit demonstration of risk control through the use of bow–tie diagrams and the evaluation of safety barrier performance, Journal of Hazardous Materials, Volume 130, Issue 3, 31 March 2006, Pages 220-233.

36. Bensi M (2010). A Bayesian Network methodology for infrastructure seismic risk assessment and decision support. PhD Dissertation, University of California, Berkeley.

37. Pascale A, Nicoli M (2011). Adaptive Bayesian network for traffic flow prediction. 2011 IEEE Statistical Signal Processing Workshop (SSP), Nice, pp. 177-180.

38. Mirmoeini F, Krishnamurthy V (2005). Reconfigurable Bayesian Networks for Hierarchical Multi-Stage Situation Assessment in Battlespace, Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 2005, pp. 104-108.

39. Esposito S, Stojadinovic B (2016a). Deliverable 5.2 Report on the proposed Bayesian network framework for conducting stress tests of non-nuclear critical infrastructures. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards.

40. Niel M, Fenton N, Nielson L (2000). Building large-scale Bayesian networks. The Knowledge Engineering Review, 15: 257-284.

41. Park HS, Cho SB (2012). A modular design of Bayesian networks using expert knowledge: Context-aware home service robot, Expert Systems with Applications, Volume 39, Issue 3, 15 February, Pages 2629-2642

42. Nishijima K, Faber M (2007). A Bayesian framework for typhoon risk management. Proceedings of 12th International Conference on Wind Engineering (12ICWE), Cairns, Australia.

43. Smith M (2006). Dam Risk Analysis Using Bayesian Networks. *Proceedings of 2006 ECI Conference of GeoHazards*, Lillehammer, Norway.

44. Grêt-Regamey A, Straub D (2006). Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS. Natural Hazards and Earth System Sciences, 6: 911-926.

45. Bayraktarli YY, Ulfkjaer J, Yazgan U, Faber M (2005). On the application of Bayesian probabilistic networks for earthquake risk management. Proceedings of the 9th International Conference on Structural Safety and Reliability, Rome, Italy.

46. Bayraktarli YY, Yazgan U, Dazio A, M Faber (2006). Capabilities of the Bayesian probabilistic networks approach for earthquake risk management. Proceedings of First European Conference on Earthquake Engineering and Seismology, Geneva, Switzerland.
47. Tasfamariam S, Liu Z (2014). Chapter 7: Seismic risk analysis using Bayesian belief networks. Handbook of Seismic Risk Analysis and Management of Civil Infrastructure Systems. Woodhead Publishing Series in Civil and Structural Engineering. Pages 175-208.
48. Blaser L, Ohrnberger M, Riggelsen C, Scherbaum F (2009). Bayesian Belief Network for Tsunami Warning Decision Support. Lecture Notes In Artificial Intelligence; Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Springer-Verlag Berlin Heidelberg, Verona, Italy, 757-768.
49. Bayraktarli YY, Baker J, Faber M (2011). Uncertainty treatment in earthquake modelling using Bayesian probabilistic networks. Georisk. 5(1):44.
50. Broglio S (2011). Bayesian Network framework for macro-scale seismic risk assessment and decision support for bridges. PhD Dissertation, ROSE Programme, UME School, IUSS Pavia.
51. Grauvogl B., and A. Steentoft. 2016. Seismic Resilience of Communities during the 2015 Nepal earthquake events. Master Thesis, IBK, ETH Zurich.
52. Didier M, Grauvogl B, Steentoft A, Broccardo M, Ghosh S, Stojadinovic B (2017). Assessment of post-disaster community infrastructure services demand using Bayesian networks. 16th World Conference on Earthquake Engineering, January 9th to 13th 2017, Chile.
53. Stojadinovic B, Esposito S (2016). Deliverable 4.5 Development of a coherent definition of societal resilience and its attributes. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards.
54. Friis-Hansen P (2004). Structuring of complex systems using Bayesian Networks. Proceedings of Two Part Workshop at DTU, O. Ditlevsen and P. Friis-Hansen, eds., Technical University of Denmark
55. Helm, P (1996). Integrated Risk Management for Natural and Technological Disasters. Tephra. 15(1): 4-13.
56. 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
57. CEN (2004). European Standard EN 1990: 2004 Eurocode: Basis of structural design, Brussels, Belgium.
58. Luco N, Ellingwood BR, Hamburger RO, Hooper JD, Kimball JK, Kircher CA (2007). Risk-targeted versus current seismic design maps for the conterminous United States. Structural Engineers Association of California 76th Annual Convention, Lake Tahoe, California.
59. Frangopol DM, Soliman M (2016). Life-cycle of structural systems: recent achievements and future directions. Structure and Infrastructure Engineering. 12:1: 1-20.
60. Esposito S, Stojadinovic B (2016b). Deliverable 5.3 Tools and strategies to incorporate stress tests into the long-term planning and life cycle management of non-nuclear CIs. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards.

61. Furuta H, Frangopol, DM, Nakatsu K (2011). Life-cycle cost of civil infrastructure with emphasis on balancing structural performance and seismic risk of road network. Structure and Infrastructure Engineering. 7: 65-74.

62. Chang SE, Shinozuka M (1996). Life-cycle cost analysis with natural hazard risk. Journal of Infrastructure Systems, 23: 118-126.

63. Basoz NI, Kiremidjian AS, King SA, Law KH (1999). Statistical analysis of bridge damage data from the 1994 Northridge, CA, earthquake. Earthquake Spectra 15(1):25–54.

64. Shinozuka M, Feng MQ, Lee J, Naganuma T (2000). Statistical analysis of fragility curves. Journal of Engineering Mechanics. 126(12):1224–31.

65. O’Rourke MJ, So P (2000). Seismic fragility curves for on-grade steel tanks. Earthquake Spectra. 16(4):801–15.

66. Straub D, Der Kiureghian A (2008). Improved seismic fragility modeling from empirical data. Structural Safety. 30 (4): 320-336.

67. American Lifelines Alliance (2001). Seismic fragility formulations for water systems. Part 2 – Appendices, ASCE-FEMA

68. Esposito S, Stojadinovic B (2016c). Deliverable 5.4. Report on strategies for stress test implementation at community level and strategies to enhance societal resilience using stress tests. STREST project: Harmonized approach to stress tests for critical infrastructures against natural hazards.

69. National Infrastructure Advisory Council (NIAC) (2009). Critical Infrastructure Resilience. Final Report and Recommendations. U.S. Department of Homeland Security.

70. Bruneau M, Chang SE, Eguchi RT, Lee GC, O’Rourke TD, Reinhorn AM, Shinozuka M, Tierney K, Wallace WA, von Winterfeldt D (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. Earthquake Spectra 19: 733–752

71. Chang SE, Shinozuka M (2004). Measuring Improvements in the Disaster Resilience of Communities. Earthquake Spectra. 20: 739-755.

72. Franchin P, Cavalieri F (20159. Probabilistic assessment of civil infrastructure resilience to earthquakes. Computer Aided Civil And Infrastructure Engineering. DOI: 10.1111/mice.12092.

73. Bocchini P, Frangopol DM, Ummenhofer T, Zinke T (2014). Resilience and sustainability of civil infrastructure: Toward a unified approach. Journal of Infrastructure System. 20 (2).

74. Francis R, Bekera B (2014). A metric and frameworks for resilience analysis of engineered and infrastructure systems. Reliability Engineering and System Safety. 121: 90-103.

75. Broccardo M, Galanis P, Esposito S, Stojadinović B (2015). Resilience-based risk assessment of civil systems using the PEER framework for seismic hazard. European Safety and Reliability Conference ESREL 2015, Zurich, Switzerland. 7-10 September.

76. Iervolino I, Giorgio M (2015). Stochastic Modeling of Recovery from Seismic Shocks. 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12 Vancouver, Canada, July 12-15.

77. Sun L, Didier M, Déle E, Stojadinovic B (2015a). Probabilistic Demand and Supply Resilience Model for Electric Power Supply System under Seismic Hazard, ICASP12, Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering, paper 267, 12-15. July, Vancouver, Canada.

78. Sun L, Didier M, Delé E, Stojadinovic B (2015b). Probabilistic Demand and Supply Resilience Model for Electric Power Supply System under Seismic Hazard. European Safety and Reliability Conference ESREL, Zurich, Switzerland. 7-10 September.

79. Hosseini S, Barker K, Ramirez-Marquez JE (2016). A review of definitions and measures of system resilience. Reliability Engineering and System Safety. 145: 47-61.

80. Henry D, Ramirez-Marquez JE (2012). Generic metrics and quantitative approaches for system resilience as a function of time. Reliability Engineering and System Safety, 99: 114-122.

81. Ouyang M, Duenas-Osario L (2012). A three stage resilience analysis framework for urban infrastructure systems. Structural Safety. 36: 23-31.

82. Mieler M, Stojadinovic B, Budnitz R, Comerio M, Mahin S (2015). A Framework for Linking Community-Resilience Goals to Specific Performance Targets for the Built Environment, Earthquake Spectra, 31 (3): 1267-1283.

83. Didier M, Sun L, Ghosh S, Stojadinovic B (2015). Post-Earthquake Recovery of a Community and its Power Supply System, CompDyn 2015, Proceedings of the 5th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, M. Papadrakakis, V. Papadopoulos and V. Plevris, editors, Paper #C869, Crete Island, Greece, 25-27. May 2015.

84. Didier M, Broccardo M, Esposito S, Stojadinović B. A compositional demand/supply framework to quantify the resilience of civil infrastructure systems. In prep.

85. Bellagamba X (2015). Seismic Resilience of a Gas Distribution Network. Master Thesis, ETH Zurich 2015. Advisor: Prof. B. Stojadinović.

86. National Institute of Standards and Technology (NIST) (2015). Community Resilience Planning Guide 3 for Buildings and Infrastructure 4 Systems. NIST Special Publication 1190.

List of abbreviations and definitions

ALARP	As Low As Reasonably Practicable
BN	Bayesian network
CIs	Critical Infrastructures
CL	Connectivity Loss
cEE	classical Expert Elicitation
EL	Effective Level
ET	Evaluation Team
EU	Epistemic Uncertainty
EU@STREST	Epistemic Uncertainty at STREST
GenMR	Generic Multi-Risk
IM	Intensity Measure
IR	Internal Reviewer
PBEE	Performance-Based Earthquake Engineering
PF	Penalty Factor
LCM	Life-Cycle Management
LCC	Life Cycle Cost
PEER	Pacific Earthquake Engineering Research
PIs	Performance Indicators
PM	Project Manager
PoE	Pool of Experts
PRA	Probabilistic Risk Analysis
PSHA	Probabilistic Seismic Hazard Analysis
QRA	Quantitative Risk Analysis
SBRA	Scenario-Based Risk Analysis
SHARE	Seismic Hazard Harmonization in Europe
ST	Stress Test
ST-L	Stress Test Level
STREST	harmonized approach to stress tests for critical infrastructures against natural hazards
ST@STREST	Stress Test at STREST
TI	Technical Integrator
TL	Target Level

List of figures

Fig. 2.1 Workflow of ST@STREST methodology 4

Fig. 2.2 The basic interactions among the core actors in the process of EU@STREST 6

Fig. 2.3 Interaction among the main actors during the multiple-expert process EU@STREST. The PoE is present only in ST sub-levels c and d. For sub-levels a and b, the role of the PoE is assumed directly by the TI 7

Fig. 2.4 ST-Levels in the ST@STREST methodology 11

Fig. 2.5 Plan view of the case study building from ST-L1 assessment 14

Fig. 2.6 a) Fragility function of the building and b) seismic hazard on the location of the building .. 16

Fig. 2.7 BN framework for seismic risk management (source: Bayraktarli et al, 2005) 23

Fig. 2.8 Bayesian network methodology for seismic infrastructure risk assessment and decision support proposed by Bensi (2010) 24

Fig. 2.9 BN model proposed by Didier et al (2017) 25

Fig. 2.10 An example of grading system for the outcome of stress test. The CI may pass, partly pass, or fail the stress test 26

Fig. 2.11 Grading system in time domain using scalar risk objectives (top) and limit F-N curves (bottom): a) two different results of the first evaluation of stress test (ST1), b) redefinition of the parameters of the grading system due to Result 1 in ST1, and c) redefinition of the parameters of the grading system due to Result 2 in ST1 27

Fig. 2.12 Grading of components of the system (ST-L1) 29

Fig. 2.13 Distribution of a risk measure with boundaries of grades in the case of a) a scalar risk measure and b) an F-N curve 30

Fig. 2.14 Annual exceedance curves of penalized loss considering different penalty factor values.. 36

Fig. 3.1 Life-Cycle Cost framework including natural hazard risks adapted from Chang and Shinozuka (1996).. 40

Fig. 3.2 Proposed framework for assimilating stress test and post-event data in a total life cycle cost analysis.. 42

Fig. 3.3 Annual rate of exceedance of CL considering mitigation strategies applied to the stations (MS1 and MS2) and to buried pipelines (MS3)... 43

Fig. 3.4 Comparison of existing and updated fragility curves for L'Aquila gas steel pipes 44

Fig. 4.1 Representation of the network supply, demand and consumption curves after a natural disaster event (STREST Deliverable 4.5 (Stojadinovic and Esposito 2016)) 51

Fig. 4.2 Resilience metrics defined for a gas network, from Bellagamba (2015).......... 53

List of tables

Table 2.1 Main aspects characterizing the Component Level Assessment (STL-1a) 14
Table 2.2 Verification of beam-to-column connections 15
Table 2.3 Main aspects characterizing the System Level Assessment, STL-2a 18
Table 2.4 Main aspects characterizing the System Level Assessment, ST-L2b. 18
Table 2.5 Main aspects characterizing the System Level Assessment, ST-L2c............. 19
Table 2.6 Main aspects characterizing the System Level Assessment, ST-L3c............. 20
Table 2.7 Main aspects characterizing the complementary scenario-based assessment 21
Table 2.8 Target Levels for each ST-Level 33

LB-NA-28342-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2788/659118

ISBN 978-92-79-64602-7