HAL
open science

Cadres pour l'analyse multi-perspective des infrastructures critiques

Fangyuan Han

To cite this version:

Fangyuan Han. Cadres pour l'analyse multi-perspective des infrastructures critiques. Autre. Université Paris Saclay (COmUE), 2018. Français. NNT : 2018SACLC011 . tel-01721557

HAL Id: tel-01721557

https://theses.hal.science/tel-01721557

Submitted on 2 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2018SACLC011

Frameworks for the multi-perspective analysis of critical infrastructures

Thèse de doctorat de l'Université Paris-Saclay
préparée à Nom de l'établissement

École doctorale n°573 approches interdisciplinaires, fondements, applications et innovation (Interfaces)
Spécialité de doctorat : sciences et technologies industrielles

Thèse présentée et soutenue à Gif-sur-Yvette, le 23 janvier 2018, par

Mme Fangyuan HAN

Composition du Jury :

Robert Kooij Professor, Delft University of Technology	Rapporteur
Daqing Li Professor, Beihang University	Rapporteur
Juan Chen Professor associé, Beihang University	Examinateur
Vytis Kopustinskas European Commission, Joint Research Centre	Examinateur
Franck Marle Professeur, CentreSupélec, Université Paris-Saclay	Président
Antonio Scala Professeur associé, CNR-ISC Uos "La Sapienza"	Examinateur
Gaoxi Xiao Professeur associé, NTU Singapore	Examinateur
Enrico Zio Professeur, CentreSupélec, Université Paris-Saclay	Directeur de thèse
José Sanchez Torres Chercheur, EDF	Invité

Thèse de doctorat

Acknowledgements

I would like to express my sincere appreciation to my thesis supervisor, Professor Enrico Zio, for picking me as his Ph.D. student at the first place and for the continuous support for my PhD research work. Without his guidance, this PhD dissertation would not have materialized. His strong enthusiasm and passion for scientific research, his immense knowledge in different domains, his prudent attitude and rigorous organization of the work had been inspiring and had set a great model for me. His caring and lovely personality had made my Phd an enjoyable experience. I feel so lucky and honored to be supervised by such a researcher and professor as him.

My deep appreciation goes to all the jury members, Professor Robert Kooij, Professor Daqing Li, Dr. Juan Chen, Dr. Vytis Kopustinskas, Professor Franck Marle, Dr. Antonio Scala and Dr. Gaoxi Xiao. I am very pleased to have presented my Ph.D. work in front of such a high qualified jury and have had such an interesting Question and Answer section. Special thanks to the reviewers, Professors Robert Kooij and Professor Daqing Li, for the evaluation of the manuscript and all the constructive and helpful remarks.

I want to thank Dr. Vytis Kopustinskas for offering me the opportunity to spend three months in the European Commission Joint Research Centre. Thank him for his advice, patience and great support. It is my pleasure to work in such a great place with such a great research team.

I would like to acknowledge Professor Bernard Yannou and Professor Jean-Claude Bocquet, director and ex-Director of Laboratoire Génie Industriel (LGI), for their hearty welcome and extremely pleasant working environment: it has been my pleasure working in this laboratory.

My warmest thanks go to Corinne Ollivier, Delphine Martin and Sylvie Guillemain, the secretaries of LGI. They have been lovely and helpful and their work has smoothed mine. I would like to thank all of my colleagues from LGI. It's my great pleasure to work here, to be part of the family. And especially, I want to thank every chair member: Yanfu, Ionela, Elisaveta, Ronay, Nicola, Yiping, Tairan, Jie, Lo, Elisa, Rodrigo, Yanhui, Xing, Pietro, Muxia, Shanshan, Zhiyi, Tasneem, Jinduo, Islam, Phuong, Zhiguo, Mengfei, Xiangyu, Hongping and Daogui. I have been so happy to work with these wonderful people! We are more than colleagues. Besides, I would like to give my special thanks to my former officemates, dear Tairan, Yiping, Pietro and Islam. It is a funny, fruitful and unforgettable memory to work with them together in the same office.

Finally, I would like to dedicate this thesis to my parents who have been there to support me and guide me under all circumstances. Their love and inspiration will always be the motivation for me to carry on.

Abstract

Critical infrastructures (CIs), like electricity and gas transmission and distribution systems, rail and road transportation, communication networks, etc., provide essential goods and service for modern society. Their safety and reliability are primary concerns. The complexity of CIs calls for approaches of system analysis capable of viewing the problem from multiple perspectives. The focus of the present thesis is on the integration of the control perspective into the safety and reliability analysis of CIs.

The integration is first approached by investigating the control properties of a small network system, i.e., an electric power microgrid. A simulation-based scheme is developed for the analysis from different perspectives: supply service, controllability and topology. An optimization-based model predictive control framework is proposed to analyze the microgrid under various failure scenarios.

Then, a multi-perspective framework is developed to analyze CIs with respect to supply service, controllability and topology. This framework enables identifying the role of the CI elements and quantifying the consequences of scenarios of multiple failures, with respect to the different perspectives considered. To demonstrate the analysis framework, a benchmark network representative of a real gas transmission network across several countries of the European Union (EU) is considered as case study.

At last, a multi-objective optimization framework is proposed for complex CIs design: design of network topology and allocation of link capacities are performed in an optimal way to minimize the non-supplied demand and the structural complexity of the system, while at the same time to maximize the system controllability. Investigation on the multiple objectives considered is performed to retrieve useful insights for system design.

The findings of this thesis demonstrate the importance of developing frameworks of analysis of CIs that allow considering different perspectives relevant for CIs design, operation and protection.

Keywords: Critical infrastructures, Complex networks, Supply, Controllability, Multi-perspective analysis, Multi-objective optimization, Gas transmission network.

Résumé

Les infrastructures critiques (CIs), telles que les réseaux électriques, les réseaux de gaz, les réseaux de transport ou encore les rseaux de communication, sont essentielles au fonctionnement de la société moderne. Leur sécurité et leur fiabilité sont les principales proccupations. La complexité des CIs exige des approches d'analyse de systèm capables de voir le problème de plusieurs points de vue. La présente thèse porte sur l'intégration de la perspective de contrôle dans l'analyse de sécurité et de fiabilité des éléments de configuration.

L'intégration est d'abord abordée par examiner les propriétés de contrôle d'un microgrid d'alimentation électrique. Un schéma basé sur la simulation est développé pour l'analyse sous différentes perspectives: le service d'approvisionnement, la contrôlabilité et la topologie. Un cadre de la commande prédictive (model predictive control) est proposé pour analyser le microrgrid dans divers scénarios de défaillance.

Ensuite, un cadre multi-perspectif est développé pour analyser les CIs considérant le service d'approvisionnement, la contrôlabilité et la topologie. Ce cadre permet d'identifier le rôle des éléments de CI et de quantifier les conséquences de scénarios de défaillances multiples, par rapport aux différentes perspectives considérées. Afin de présenter le cadre d'analyse, un réseau de référence représentatif d'un réseau de transport de gaz réel à travers plusieurs pays de l'Union européenne est considéré comme une étude de cas.

Enfin, un cadre d'optimisation à trois objectifs est proposé pour la conception de CI complexes: la conception de la topologie du réseau et l'allocation des capacités de liaison sont optimisées pour minimiser la demande non fournie et la complexité structurelle du systéme, et en même temps maximiser la contrôlabilité du système. Une investigation

approfondie sur les multiples objectifs considérés est effectuée pour tirer des informations utiles pour la conception du système.

Les résultats de cette thèse démontrent l'importance de développer du cadre d'analyse des CIs qui permettent de prendre en considération de plusieurs perspectives pertinentes pour la conception, l'opération et la protection des CIs.

Contents

1 Introduction **1**

1.1 Challenges in the modeling and analysis of CIs 2

1.1.1 Definition of CIs 2

1.1.2 Challenges in the modeling and analysis of CIs 3

1.2 Safety and reliability analysis of CIs 3

1.2.1 Topological analysis 4

1.2.2 Reliability analysis 5

1.2.3 System safety from control perspective 7

1.3 Research objectives 8

1.4 Structure of the thesis 9

2 Model predictive control framework for CI analysis **11**

2.1 Microgrids 11

2.1.1 Microgrid control 12

2.1.2 Microgrid safety 13

2.2 Microgrid modeling and Model predictive control 15

2.2.1 Dynamic modeling of the microgrid 15

2.2.2 Model predictive control generalities 16

2.3 System-level indexes 18

2.3.1 Supply index (Non-supplied demand) 18

2.3.2 Controllability Index 18

2.3.3 System capacity efficiency 19

2.4 Case study 20
2.4.1 System modeling 20
2.4.2 Optimization-based control for system safety analysis 22
2.4.3 Analysis of the results 26
2.4.4 Summary 31

3 A multi-perspective framework for the analysis of CIs 33

3.1 State of the art 33
3.1.1 CIs as complex networks 33
3.1.2 Controllability of complex networks 35
3.2 System-level indexes 36
3.3 Application 38
3.3.1 Network description 38
3.3.2 Quantification of link importance 40
3.3.3 Simulation and analysis 43
3.3.4 Summary 51

4 Multi-objective optimization of CIs 53

4.1 State of the art 54
4.1.1 Multi-objective optimization 54
4.1.2 Evolutionary algorithms for network optimization 55
4.2 Optimization for CI design 57
4.2.1 Three system-level indexes considered as optimization objective . . 57
4.2.2 Multi-objective optimization problem formulation 58
4.3 Case study 59
4.3.1 Correlations among the three objectives 60
4.3.2 Optimization results 64
4.3.3 Analysis of node importance 66
4.3.4 Summary 68

5 Conclusions and future research **71**

5.1 Conclusions 71

5.2 Prospective works 72

References **90**

List of Figures

1.1 Research objectives of the presented thesis 9

2.1 Microgrid 20

2.2 Consumer load profile 25

2.3 Wind power profile 26

2.4 Supply sources of Scenario 1.0. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator. 28

2.5 Supply sources of Scenario 1.1. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator. 29

2.6 Supply sources of Scenario 2.0. The black area represents the power from the battery and the shaded area represents the power from the renewable generator. 30

3.1 Gas transmission network(Praks and Kopustinskas, 2016) 39

3.2 Link importance 43

3.3 Histogram: number of cases and frequency 44

3.4 Non-supplied demand for the 335 cases 46

3.5 Non-supplied demand for different categories 46

3.6 Controllability index for the 335 cases 48

3.7 Controllability index for different categories 48

3.8 Network topological efficiency for the 335 cases 49

3.9 Network topological efficiency for different categories 50

3.10 CDF of NSD for all configurations and for the configurations without failure of node 19 51

4.1 Mean of average node degree versus number of driver nodes N_D 61

4.2 Mean of the three objectives 62

4.3 Mean of structural complexity C versus N_D 63

4.4 Mean of NSD versus N_D 64

4.5 Pareto front in 3-D space (a) and 2-D projections (b)-(d) 66

List of Tables

2.1 Numerical data for the microgrid components 24

2.2 Index values for the grid connected mode 27

2.3 Index values for the stand-alone mode 29

3.1 Sources properties 40

3.2 Indexes values for the nominal configuration 41

3.3 Indexes values associated to the removal of the most critical links 41

3.4 30 most frequent cases 45

4.1 Parameters of the NSGA-II algorithm 60

4.2 Index values for the original network 60

4.3 Objective functions values for the Pareto-optimal solutions 65

4.4 Average node importance values for the optimal solutions 67

List of Symbols

$< k >$ Average node degree

α_i Complexity of the ith component

β_{ij} Interface complexity between the ith and j^{th} components

γ Normalization factor

Adj Adjacency matrix

G Graph representing the network

$\mathbf{G}(base)$ Graph representing the original network

$\mathbf{G}'(x_i = 1)$ Graph obtained by removing the node i from $\mathbf{G}(base)$

$\mathbf{G}'(x_{ij} = 1)$ Graph obtained by removing the link ij from $\mathbf{G}(base)$

K Capacity matrix of the network

Len Length matrix of the network

$\mu(\lambda_i)$ Geometric multiplicity of the eigenvalue λ_i of matrix $\mathbf{A}$

ω_i Weight of the ith user

C Structural complexity

C_1 Sum of individual components complexities

C_2 Sum of interface complexities

C_3 Topological complexity

C_{ind} Controllability index

D_i Demand of the ith user

$E(\mathbf{Adj})$ Matrix energy of **Adj** (i.e. sum of its singular values)

$IM_{ij}^{C_{ind}}$ Link importance with respect to C_{ind}

IM_{ij}^{E} Link importance with respect to E

IM_{ij}^{NSD} Link importance with respect to NSD

L Number of links in the network

N Number of nodes in the network

N_D Minimum number of driver nodes

N_G Maximum number of generations

N_P Number of candidate solutions

N_s Number of sources in the network

N_y Number of user nodes within the network

$NI_i^{C_{ind}}$ Node importance with respect to C_{ind}

NI_i^{C} Node importance with respect to structural complexity

NI_i^{NSD} Node importance with respect to NSD

NSD Normalized non-supplied demand

P_t Parent population of the tth generation

Q_t Offspring population of the tth generation

R_t Union population of the tth generation

y_i Supply to the ith user

List of Acronyms

ACO Ant Colony Optimization

BDD Binary Decision Diagram

CDF Cumulative distribution function

CI Critical infrastructure

CIP Critical infrastructure protection

DER Distributed energy resources

DoD Depth of Discharge

DPT Disjoint Product Technique

EMS Energy Management System

EU European Union

FTA Fault tree analysis

GA Genetic Algorithms

MAS Multi-agent system

MBDE Modified binary differential evolution

MOEA Multi-objective evolutionary algorithm

MPC Model predictive control

MSR Matrix-based System Reliability

NSBDE Non-dominated sorting binary differential evolution

NSGA-II Non-dominated sorting genetic algorithm II

PSO Particle Swarm Optimization

RA Risk Achievement

RDA Recursive Decomposition Algorithm

Appended Papers

Paper I: F. Han, E. Zio. Modeling an electric power microgrid by model predictive control for analyzing its characteristics from reliability, controllability and topological perspectives. Journal of Risk and Reliability, 2017. (Accepted).

Paper II: F. Han, E. Zio. Optimization of critical infrastructures with respect to supply service, structural complexity and controllability. Reliability Engineering and System Safety, 2017. (Under review).

Paper III: F. Han, E. Zio. A multi-perspective framework of analysis of critical infrastructures with respect to supply service, controllability and topology. International Journal of Critical Infrastructure Protection, 2017. (Under review).

Chapter 1

Introduction

Critical infrastructures (CIs), like electricity and gas transmission and distribution systems, rail and road transportation, communication networks, etc., are essential to the operation of modern society (Kröger and Zio, 2011). They need to be designed, maintained and protected to provide optimal performance, reliable operation and functional safety for long periods of time (Ottino, 2004; Rouse, 2003). Hence, the great attention and priority are given to the "care" of these systems by the EU, US and other national and transnational administrations (Clinton, 1998; EU, 2014; Lewis, 2014; Lindström and Olsson, 2009), which calls for risk assessment and resilience evaluation of CIs (EU, 2010; Rigaud and Guarnieri, 2006).

However, many questions and challenges rise from the increasing complexity of CIs: How to analyze the control, reliability and safety properties of CIs? How to identify the critical elements whose failure and loss of control can lead to large consequences? How to design CIs seeking the optimal balance between different goals? The objective of the present thesis is to address the above questions and to develop a multi-perspective modeling, analysis and optimization framework for the safe and reliable operation of CIs.

This chapter aims to provide a general overview of the problems addressed in this thesis, and is organized as follows. Section 1.1 discusses the challenges in the analysis of CIs. Section 1.2 introduces the key issues related to the safety of CIs. Section 1.3 explains the objectives of this study. Finally, Section 1.4 presents the structure of the thesis.

1.1 Challenges in the modeling and analysis of CIs

1.1.1 Definition of CIs

Critical infrastructures (CIs) are generally understood as comprising the facilities and services that are vital to the basic operations of society and whose disruption or destruction could greatly impair the functioning of the society (Zhang et al., 2015). From a European Union perspective, CIs are defined as network systems that provide life-essential services and their disruption or destruction would have a significant impact on the health, safety, security, economics, and social well-being, including the effective functioning of governments (Directive, 2008).

CIs are divided into a number of sectors, from traditional areas such as defense, transportation and energy, to new areas such as banking and finance, healthcare and information technology. The focus of this thesis is on networked CIs for supply, such as networks providing energy (electricity, oil, and gas), transportation (by rail, road, air, and sea), information and telecommunication (e.g., the Internet) and state and local services (e.g., water and emergency services)(Kröger and Zio, 2011).

Large-scale CIs have several characteristics in common (Kröger, 2008):

- They are consisted of networked human-made systems that function synergistically to produce a continuous flow of goods or services to customers.
- They are designed to satisfy specific social needs but also shape social change at a much broader and more complex level.
- They are subject to multiple threats (technical-human, physical, natural, cyber, contextual; unintended or malicious) and have inherent vulnerabilities.
- They are inter-dependent, both physically and through a host of ICT.
- Disruptions may cascade, i.e. local interruptions may lead to wide spread cascading failures.

- They have no single owner, operator or regulator and their operating environment is based on different goals and logics.

1.1.2 Challenges in the modeling and analysis of CIs

Due to the important role played by CIs and their characteristics which make them difficult to control or to operate safely and reliably, concerns have been arising on their modeling, analysis and protection.

The problem is that the conventional mathematical methodologies behind today's modeling, simulation, and control paradigms are unable to handle their complexity and interconnectedness and the classical methods of system vulnerability and risk analysis cannot capture the heterogeneity and (structural, dynamic, and operational) complexities of CIs: the analysis of these systems cannot be carried out with classical methods of system decomposition and logic modeling. As Zio (Zio, 2007) and Kröger (Kröger, 2008) point out, in order to address the complexities of CIs, new methods for their analysis are needed, since the current quantitative methods of risk analysis seem not to be fully equipped to deal with the level of complexity inherent in such systems" (Zio, 2007). A framework is needed for the integration of methods capable of viewing the problem from different perspectives (e.g., topological and functional, static and dynamic), suitable for coping with the high complexity of the system and the related uncertainties, and capable of offering a holistic viewpoint (Kröger and Zio, 2011; Zio, 2016).

1.2 Safety and reliability analysis of CIs

CIs are witnessing more and more system-level breakdowns, which emerge from small perturbations followed by cascades of failures, which lead to large-scale consequences. CIs are exposed to many types of hazards, such as natural hazards, component aging and failure, sharp load demand increase, climatic changes and intentional attacks. For this reason, Critical Infrastructure Protection (CIP) has gained great importance in all

nations, with particular focus being placed traditionally on physical protection and asset hardening.

Then, it is not surprising that CI protection and resilience have become a national and international priority, which calls for vulnerability analysis and CI properties evaluation, for ensuring their protection and resilience (Rigaud and Guarnieri, 2006). To ensure their safe and reliable operation, system analysis, reliability engineering, graph theory and others have been propounded to study the behavior and performance of CIs, also with respect to failure events, protection practice and system resilience (Fang and Zio, 2013a; Limiao et al., 2016; Zio, 2009).

1.2.1 Topological analysis

The fact that CIs are complex networks of interacting components raises the interest in studying their topological characteristics(Fang and Zio, 2013a,b; Lewis et al., 2013; Lu et al., 2016; Ouyang, 2014). In addition, as the CI networks grow in size and complexity, it is extremely difficult, if not infeasible, to perform an analytic description or simulation of the behavior or physical processes of the whole network. The less computationally demanding topological perspective is, therefore, proposed to investigate the network properties.

Topological analysis based on complex network theory can unveil relevant properties of the structure of a networked system (Albert et al., 2000; Strogatz, 2001) and can be used to the analysis of CIs for identifying the role of its components and evaluating the network properties in the presence of failures mainly represented by the removal of nodes and edges (Kröger and Zio, 2011).

A number of recent studies have proposed various measures to evaluate the structural properties of networks and addressed topological investigations to identify critical elements. Among these measures, topological centrality (including degree, closeness, betweenness and information centrality) (Freeman, 1978; Nieminen, 1974) and network efficiency (Latora and Marchiori, 2001) are two important and classical measures, which

quantify the importance of individual network elements and evaluate the connectivity of the whole network, respectively.

These topological properties have been studied in relation to the safety and reliability issues of CIs. The authors of (Albert et al., 2004) have investigated the ability of the North American power grid to transfer power between generators and consumers when certain transmission substations are disrupted, and the results show that the system is vulnerable to disturbances affecting the key transmission substations, while robust to most perturbations. Large interest of considering the topological perspective has ever since been seen in the study of the vulnerability of power grid (Crucitti et al., 2004; Eusgeld et al., 2009; Holmgren, 2006), safety features of urban transport networks (Crucitti et al., 2006; Zio et al., 2008), and finding critical component of Internet (Latora and Marchiori, 2005). These studies show that the structure properties of CIs provide important information which helps understand their global behavior.

However, the insights gained from the topological studies can be limited from the point of view of the description of physical processes and phenomena in the network (Boccaletti et al., 2006). The functioning of some networks like communication networks is less dependent on the physical process for the transfer of information, but the physical process is particularly important in the networks with physical flow, in which the pipeline capacity directly affects the maximum possible flow passing through it, and component failures (e.g. compressor stations in a gas network) can be critical. Thus, it is important to narrow the gap between the highly conceptualized analysis based on network topology and the highly detailed system physical modeling.

1.2.2 Reliability analysis

Reliability is a fundamental attribute for the safe operation of any modern technological system and the concept of reliability has been used in the context of engineering system for more than 60 years (Zio, 2009). The reliability of an infrastructure system can be defined as the probability (or the ability, more generally) of the system to provide its

services to its customers (Johansson et al., 2013).

Reliability analysis is commonly used in the context of CIs. The goal in reliability analysis is to obtain a picture of a system's likely behavior(Johansson et al., 2013), by quantifying the probability of failure, calculating different reliability indices, for example. Given the relationship between the topology of the complex networks and their vulnerability and safety features, the association between network structure and system reliability is also of relevance. A common measure of network reliability is the so called k-terminal reliability, which calculates the probability that every two nodes in a specific subset of K nodes are connected by a path of operational edges (Kelleher, 1991). Of particular interest are the 2-terminal reliability and the all terminal reliability. The former is also known as the st-reliability, which is the probability of successful communication between a specified source node and a terminal node in a network, given the probability of success of each link and node in the network. The latter is the probability that the network is fully connected. In the worst case, computing the exact reliability of a network is NP-hard (Shier, 1991).

Given the complexity and scaling issue of CIs, effective and rapid network reliability analysis methods are required to appropriately address the complexities and to timely calculate system reliability. Network reliability analysis is often performed by simulation-based approaches, typically relying on Monte Carlo simulation strategies. These methods are based on random samples of hazard intensities and the corresponding network component responses. They are suitable for large networks because their computational efficiency depends more on the convergence of probability than the number of network components. The non-simulation-based methods are originally developed for the node pair connectivity analysis of generic networks, including Disjoint Product Technique (DPT) (Yeh, 2007), Binary Decision Diagram (BDD) (Singh et al., 1996), Recursive Decomposition Algorithm (RDA) (Li and He, 2002), Matrix-based System Reliability (MSR) (Song and Kang, 2009), etc. In these methods, analytical insight is sought and guaranteed approximations or bounds are also unraveled despite their high computational complexity.

Their applications can be found in various studies (e.g. (Fang and Zio, 2013a; Helseth and Holen, 2006; Kim and Kang, 2013; Pino et al., 2016; Yeh et al., 2010; Zio and Golea, 2012)).

1.2.3 System safety from control perspective

Control theory provides another angle of viewing the issues of system safety and reliability of CIs. Under a general control perspective, system safety can be framed as a control "problem" (Bakolas and Saleh, 2011; Leveson, 2004; Rasmussen, 1997). Rasmussen's work has been influential (Rasmussen, 1997), in which he has mentioned that:"*many levels [...] are involved in the control of safety by means of laws, rules, and instructions that are formalized means for the ultimate control of some hazardous physical process.*" and that "*safety depends on the control of work process*". Then, Leveson (Leveson, 2004) expanded and built the system-theoretic accident model and process (or STAMP) model for accident causation and system safety and highlighted that accidents result from inadequate control actions or insufficient enforcement of safety-related constraints on the development, design, and operation of the system, leading to their violation and subsequently to accidents.

Notions from Control Theory, such as controllability and observability, have been introduced in relation to the problem of accident causation and system safety (Saleh et al., 2010). According to Control Theory, a dynamical system is controllable if, by a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time (Kalman, 1959; Liu et al., 2011). From system safety perspective, controllability is the ability to guide the system's behavior towards a safe state through the appropriate choice of a few input variables (Bakolas and Saleh, 2011). Accidents are, therefore, seen as the result of inadequate control or insufficient enforcement of safety-related constraints on the development, design, and operation of the system, leading to their violation and subsequently to accidents(Lussier et al., 2004). In the case of an accident, if the system is controllable, there exists at least one decision/action that could steer the system back to safe operation mode; otherwise, there is no such guarantee.

The efforts poured in developing analysis frameworks of CIs help us to retrieve insights of their behavior and structural and dynamic characteristics. The ultimate proof of our understanding of CIs is reflected in our ability to control them (Liu et al., 2011), which leads us to investigate the controllability of CIs.

1.3 Research objectives

The focus of this thesis is to propose a framework integrating the control perspective to complement the analysis of CIs for safety and reliability consideration of CIs. Then, the mission is to retrieve useful information from such multi-perspective analysis in order to guide the design, improvement and protection of CIs.

The research objectives, which represent also the main contributions of this thesis, include:

- Development and application of a simulation-based framework for CI analysis from different perspectives: topology, reliability, controllability.
- Consideration of the controllability property of CIs.
- Investigation of the relations among different system-level indexes.
- Identification and classification of important elements for failure dynamics, with respect to different perspectives.
- Quantification of the consequences of multiple failure scenarios with respect to different perspectives.
- Three-objective optimization framework for complex CIs design.

Figure 1.1 presents a schematic representation summarizing the main objectives.

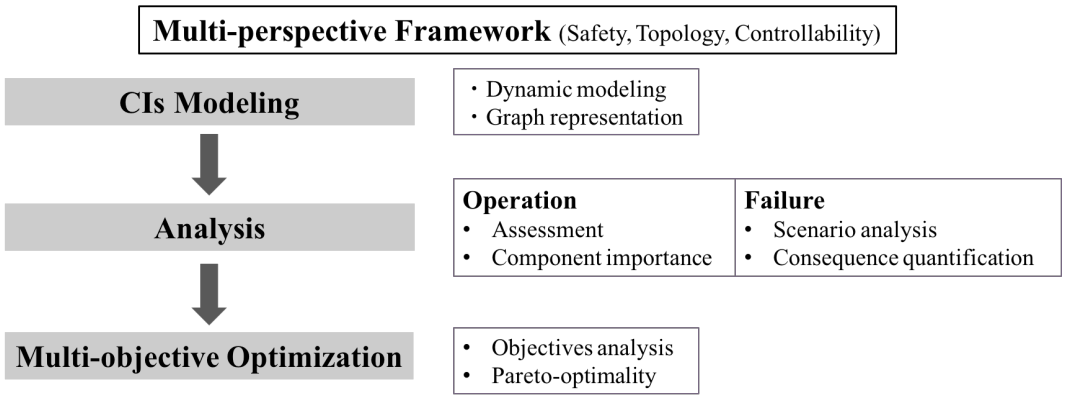


Figure 1.1: Research objectives of the presented thesis

1.4 Structure of the thesis

This thesis is divided into two parts. The first part, composed of five chapters, introduces the research context, describes the problems addressed, presents the approaches developed and applied in this work, discusses some of the results obtained in the case studies and provides general conclusions and some future work perspectives.

Chapter 2 begins with a brief overview of model predictive control. Then, the relevant issues of microgrids, such as their control and safety, are introduced. Case study on a test microgrid is provided, including elaboration of the dynamic modeling and discussion on the simulation results for different faulty scenarios.

Chapter 3 proposes a multi-perspective framework of analysis of CIs. It starts with an introduction to the field of complex network theory and controllability of complex networks. Then, three system-level indexes considered for the analysis of CIs are introduced. The case study of a real gas transmission network is presented to demonstrate the analysis of link importance and consequences of failure scenarios by applying the proposed framework.

Chapter 4 focuses on the optimization of CIs design. A three-objective optimization problem is formulated and solved by the non-dominated sorting genetic algorithm II (NSGA-II). A gas transmission network is taken as case study to demonstrate the proposed approach. Relevant results analysis is also presented.

Chapter 5 draws conclusions of this Ph.D work and presents perspectives for future research.

The second part of the thesis includes the collection of three papers submitted or prepared for submission, reporting on the outcomes of the work, which the readers can refer to for further details. Papers [1] and [2] presents the multi-perspective analysis framework of CIs under various failure scenarios. Specially, Paper [1] concerns the dynamic modeling and analysis of microgrids characteristics by model predictive control (See Chapter 2 of Part I). Paper [2] proposes a framework of analysis of critical infrastructures (CIs) of larger scale, with the objective of identifying the most critical elements and failure scenarios and evaluating their consequences (See Chapter 3). Paper [3] considers the optimization of complex supply networks design, with respect to the objectives of minimizing the non-supplied demand and their structural complexity, while at the same time maximizing their controllability (See Chapter 4).

Chapter 2

Model predictive control framework for CI analysis

The aim of this chapter is to develop a simulation-based scheme for CI analysis from different perspectives: supply service, topology, and controllability. This chapter focuses on power distribution systems and distributed generation, and in particular microgrids which offer an interesting solution for integrating renewable and distributed energy resources. A model predictive control (MPC) framework is proposed to analyze the microgrid under various faulty scenarios.

Section 2.1 briefly introduces the field of microgrids and the relevant issues, such as the control and safety of microgrids. Section 2.2 provides the modeling of microgrid by model predictive control. Section 2.3 introduces the three system-level indexes considered for CIs analysis in this thesis. At last, Section 2.4 presents the modeling of the microgrid considered and discusses the simulation results for different faulty scenarios.

2.1 Microgrids

Green energy (solar and wind in particular) production is supposed to increase significantly in the coming years, since the traditional energy supplies of Earth are finite and suffer from a diminishing returns curse. This requires a smartgrid system capable of deal-

ing with distributed production/intermittent variations of output and optimal scheduling of demand. Distributed renewable energy sources are increasingly connected to power distribution networks as a remedy to environmental and economic concerns.

Microgrids can be a key solution for integrating renewable and distributed energy resources (DER) (Lasseter et al., 2002). A microgrid is a cluster of micro-sources, energy storage systems and loads, which can be connected to the power grid as a single entity that can respond to central control signals (Lasseter and Paigi, 2004). It can improve reliability and security of power distribution system, especially for sensitive loads, because micro-sources will ensure that the sensitive loads will receive enough power in any operating condition (Olivares et al., 2014; Wang et al., 2013). The increasing interest on microgrids is triggered by the potential benefits of the microgrid that may provide reliable, secure, efficient, environmentally friendly, and sustainable electricity from renewable energy resources.

2.1.1 Microgrid control

Microgrid control is one of the key issues related to microgrid techniques and must be able to ensure the reliable and economical operation of the system by overcoming the difficulties. The responsibilities of microgrid control include the following points ((Lasseter et al., 2002) and (Zamora and Srivastava, 2010)):

- micro-sources work properly at predefined operating point or slightly different from the predefined operating point but still satisfy the operating limits;
- active and reactive powers are transferred according to necessity of the microgrids and/or the distribution system;
- voltage sag and system imbalances can be corrected;
- isolation and reconnection to the main grid are conducted in a rapid and seamless fashion;
- the use of resources by both the microgrid and grid is optimized.

A hierarchical control, which represents a compromise between fully centralized control and fully decentralized control, is an interesting solution to realize control on microgrids while answering to the above requirements(Guerrero et al., 2011; Mohamed and Radwan, 2011). The three hierarchical control levels are as follows (Vasquez et al., 2010):

- Primary control, or local or internal control, includes output control and power sharing. The former is responsible for the adequate share of active and reactive power mismatches in the microgrid, while the latter regulates the output voltages and currents (Blaabjerg et al., 2006; Gao and Iravani, 2008; Karimi et al., 2008; Lopes et al., 2006). Primary controls are designed to operate independently and react in predefined ways instantaneously to local events.

- Secondary control, also referred to as the Energy Management System (EMS), is responsible for the reliable, secure and economical operation of microgrids in both grid-connected and stand-alone operation modes. Its objectives include restoring any steady-state error introduced by the action of the primary control and finding the optimal (or near optimal)dispatch of the available DER. Secondary controls, on the other hand, coordinate the internal primary controls within the microgrids and subsystems in the span of a few minutes.

- Tertiary control concerning global responsibilities decides the import or export of energy for the microgrid and typically operates in the order of several minutes, providing signals to secondary level controls at microgrids and subsystems and other subsystems that form the full grid.

2.1.2 Microgrid safety

Microgids have been proposed to improve reliability and stability of electrical system and to ensure power quality of modern grid. Risks exist both inside and outside the microgrids mainly due to the high uncertainty and high variability associated with the system components and environmental factors (Zhang et al., 2013). Microgrid protection

against failures, threats and hazards is an important issue. Various studies have addressed the safety issue of microgrid systems, such as the electrical safety (Jayawarna et al., 2005; Jiayi et al., 2008), the risk-based performance evaluation (Gabbar et al., 2012), availability (Kwasinski, 2011), resilience (Hamilton et al., 2016), etc.

The characteristics of microgrids lead to the following safety requirements to ensure the safe and reliable operation (Islam and Gabbar, 2012):

- Sensitivity: appropriate threshold value should be set to identify any abnormal condition;
- Selectivity: in the presence of fault detected, the protection/control system should disconnect the smallest possible part of the microgrid;
- Speed: protective relay should respond in the least possible time, in order to minimize the damage to the system
- Security: the protection/control system should reject abnormal events and transients which are not fault and avoid misoperation while experiencing failures;
- Redundancy: backup protection is needed to ensure the level of safety;
- Reliability: high reliability level is required for the power system.

The reliability evaluation and risk analysis of microgrids is of great importance for system design and operation. Various studies have studied the impact of microgrids on the distribution networks. Reference (Costa and Matos, 2005) proposes a reliability analysis method for distribution networks with microgrids and investigates the impacts of microgrids on the distribution network. Reference (Bae and Kim, 2007) analyzes the effect of different operation modes of DGs. The authors of reference(Yokoyama et al., 2008) propose a method for the evaluation of supply reliability of microgrids including wind power and photovoltaics by simulation. The reliability of interior microgrid has also been discussed. Reference (Li et al., 2010) proposes a fault tree analysis (FTA)-based approach to evaluate the reliability of islanded microgrid operating in an emergency mode.

Reference (Wang et al., 2013) proposes a series of new metrics for the reliability assessment of microgrids which takes into consideration the effect of outages in a distribution network and the island switching process.

Proper microgrid protection and safety are achieved through proper control, and the need and interest of considering the control perspective for the analysis of microgrids emerge.

2.2 Microgrid modeling and Model predictive control

2.2.1 Dynamic modeling of the microgrid

We adopte a state-space model based on differential equations to describe the response of the system states to operational and environmental changes. The dynamic of the system is described by the following linear time-invariant equations:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{2.1}$$

where $\mathbf{x}$ is the state vector, representing the states of the links and storages devices, $\mathbf{u}$ is the vector of control input, $\mathbf{y}$ is the output vector. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are state transition matrices.

For a link l_i, its dynamic can be described as following:

$$x_{l_i}(t+1) = (1-\alpha_i)x_{l_i}(t) + \alpha_i \sum_{l_{in} \in I_{l_i}} u_{l_{in}}(t) \tag{2.2}$$

where α_i characterizes the inertia of flow transmission that depends on the physical characteristics of the link l_i, $u_{l_{in}}(t)$ is an input flow of the link l_i, and I_{l_i} is the set of input flows of the the link l_i.

For a storage device s_i, the dynamic is described as:

$$x_{s_i}(t+1) = (1-\tau_i)x_{s_i}(t) + \sum_{s_{in} \in I_{s_i}} u_{s_{in}}(t) - \sum_{s_{out} \in O_{s_i}} u_{s_{out}}(t) \tag{2.3}$$

with the mixed-integer conditions (Prodan and Zio, 2014a):

$$\begin{cases} 0 \leq \sum u(t) \leq Ma(t), \\ 0 \leq \sum x(t) \leq M(1-a(t)), \end{cases} \tag{2.4}$$

where τ denotes the hourly self-discharge decay. $u_{s_{in}}(t)$ and $u_{s_{out}}(t)$ are input and output flow of the storage device s_i respectively, and I_{s_i} and O_{s_i} the sets of inout and output flows of the storage device s_i, respectively. M represents a constraint and $a(t) \in \{0, 1\}$ is an auxiliary binary variable, characterizing the battery state of charge: $a(t) = 1$ when the battery is in discharge mode, $a(t) = 0$ when the battery is in charge mode.

2.2.2 Model predictive control generalities

The energy management of microgrids is particularly difficult as it is necessary to consider both exogenous factors (e.g. variations of wind speed, consumer demand) and the structural properties and internal dynamics of individual components (e.g. links, storage devices), which may change due to degradation, failure and other factors. Various approaches for control and energy management of microgrids are reported in the literature (Zamora and Srivastava, 2010). Optimal dispatching of distributed generators and storages (Hernandez-Aramburo et al., 2005) has been proposed to deal with the problem of energy management, and solved by various heuristic optimization techniques, such as Genetic Algorithms (GA) (Conti et al., 2012), Particle Swarm Optimization (PSO)(Colson et al., 2010b), Ant Colony Optimization (ACO) (Colson et al., 2009). Multi-agent system (MAS) framework (Dimeas and Hatziargyriou, 2005) has been used to model microgrids and to analyze by simulation the interactions between individual intelligent decision-makers (Colson et al., 2010a; Jimeno et al., 2011; Krause et al., 2006; Kuznetsova et al., 2011; Logenthiran et al., 2008; Oyarzabal et al., 2005; Weidlich and Veit, 2008). The authors of (Prodan and Zio, 2014a) and (Prodan and Zio, 2014b) apply Model Predictive Control (MPC) to develop an optimization-based control approach for the reliable energy management of a microgrid system.

Model predictive control (MPC), a widely used optimization-based technique in the

control community (Rawlings and Mayne, 2011; Richalet and O'Donovan, 2009).

MPC considers a cost function over a finite prediction horizon N_p and provides a sequence of control inputs. The control action $u(k)$ for a given state $x(k)$ is obtained from the control sequence $\mathbf{u} \triangleq \{u(k|k), u(k+1|k), \ldots, u(k+N_P-1|k)\}$ as the result of the optimization problem (Prodan and Zio, 2014a):

$$\arg \quad \min_{\mathbf{u}} \quad V_f(x(k+N_P|k) + \sum_{s=0}^{N_p-1} V_n(x(k+s|k), u(k+s|k)),$$

subject to:

$$\begin{cases} x(k+s+1|k) = f(x(k+s|k), u(k+s|k)), & s = 0, \cdots, N_p - 1, \\ h(x(k+s|k), u(k+s|k)) \leq 0, & s = 0, \cdots, N_p - 1, \\ h_f(x(k+N_P|k) \leq 0, & s = 0, \cdots, N_p - 1, \end{cases}$$

$V_f(\cdot)$ is the terminal cost function; $V_n(\cdot)$ is the cost per step within the horizon. $f(\cdot,\cdot)$ describes the dynamics of the system states. The objective (or cost) function is constructed to penalize deviations of the states and inputs from their reference values, while the constraints are enforced explicitly (Goodwin et al., 2005). The constraints include states and control inputs $h(\cdot)$ and terminal constraints $h_f(\cdot)$.

Various applications of MPC can be found in literature, for example, refrigeration systems (Hovgaard et al., 2011), heating systems (Halvgaard et al., 2012), power production plants (Edlund et al., 2011) and transportation networks (Negenborn et al., 2006). Recently, MPC has been attracting interest in the energy management of microgrid systems (Hooshmand et al., 2012; Negenborn et al., 2009; Parisio and Glielmo, 2011; Perez et al., 2013), because it is able to handle control and state constraints and deal with the uncertainties in the behavior of their components, such as variations in power outputs from non-dispatchable DERs, energy prices and instantaneous demand (Olivares et al., 2014; Prodan and Zio, 2014a), while offering good control performance.

2.3 System-level indexes

2.3.1 Supply index (Non-supplied demand)

Microgids have been proposed to improve reliability and stability of electrical systems and to ensure power quality of modern grids, and have the responsibility to ensure the supply to the essential loads(Zamora and Srivastava, 2010). Supply performance is a fundamental functional requirement for the microgrid. In this paper, we call the system "safe" if it ensures the satisfaction of the consumers essential demands. We introduce the non-supplied demand (NSD) as a measure of the system's capacity to satisfy its users' demands. The normalized NSD is introduced as a system-level index:

$$NSD = 1 - \frac{\sum_{i=1}^{i=N_m} \omega_i y_{m_i}}{\sum_{i=1}^{i=N_m} \omega_i D_{m_i}} \tag{2.5}$$

where, ω_i is the weight of the ith of the N_m users, y_{m_i} is the supply to user i and D_{m_i} is its demand, which is considered as the target supply to user i. Then, the second term in Equation (3.2) represents the satisfied proportion of users' demands. Since $y_{m_i} \leq D_{m_i}$, the index NSD is normalized to take values in $[0, 1]$. NSD equals 0 when the users' demands are fully supplied.

2.3.2 Controllability Index

A dynamic system is controllable if, with a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time (Liu et al., 2011). Taking a system safety perspective, controllability is the ability to guide the system's behavior towards a safe state through the appropriate manipulation of a few input variables (Bakolas and Saleh, 2011; Han et al., 2015).

It is proved in Control Theory that the system (as described by equation 2.1) is controllable if and only if its controllability matrix has full rank(Kalman, 1963):

$$rank[B \quad AB \quad ... \quad A^{n-1}B] = n \tag{2.6}$$

where n is the number of state variables of the system. This criteria is called Kalman's

controllability rank condition. The rank of the controllability matrix provides the dimension of the controllable subspace of the system.

In this work, the controllability index CI measures the controllable proportion of a dynamic system. It is defined as the ratio of the rank of the controllable subsystem to the rank of the system:

$$CI = \frac{R_C}{n} \tag{2.7}$$

where $R_C = rank[B \quad AB \quad ... \quad A^{n-1}B]$.

2.3.3 System capacity efficiency

We introduce the system capacity efficiency to measure how much flow a system topology allows to exchange. The capacity of flow exchange from nodes i to j is determined by the capacity of the widest-capacity path between them, k_{ij}, which is further determined by the minimum edge capacity in the path between the two nodes that maximizes the capacity of the minimum-capacity edge. Then, the capacity efficiency of the whole system E_c is given by:

$$E_C[\mathbf{G}] = \frac{1}{N(N-1)} \sum_{i \neq j \in \mathbf{G}} k_{ij} \tag{2.8}$$

The source-terminal capacity efficiency E_c^{st}, which only takes into account the transmission capacity between a source node and a terminal (demand) node, is given by:

$$E_C^{st}[\mathbf{G}] = \frac{1}{N_{st}} \sum_{s \in \mathbf{S}, t \in \mathbf{T}} k_{st} \tag{2.9}$$

Then, we define the source-terminal capacity efficiency index (EI^{st}) as the normalized E_C^{st}:

$$EI^{st}[\mathbf{G}'] = \frac{E_C^{st}[\mathbf{G}']}{E_C^{st}[\mathbf{G}]} \tag{2.10}$$

where $\mathbf{G}'$ is the graph obtained by the removal of certain components from $\mathbf{G}$.

2.4 Case study

We consider the microgrid system in Figure 2.1, taken from (Prodan and Zio, 2014b). This microgrid system contains one renewable generator (wind turbine), one storage device (battery) and one local consumer. The microgrid system is connected to the external power grid through a transformer. All the components are characterized by the dynamic models, constraints and reference profiles presented in the following.

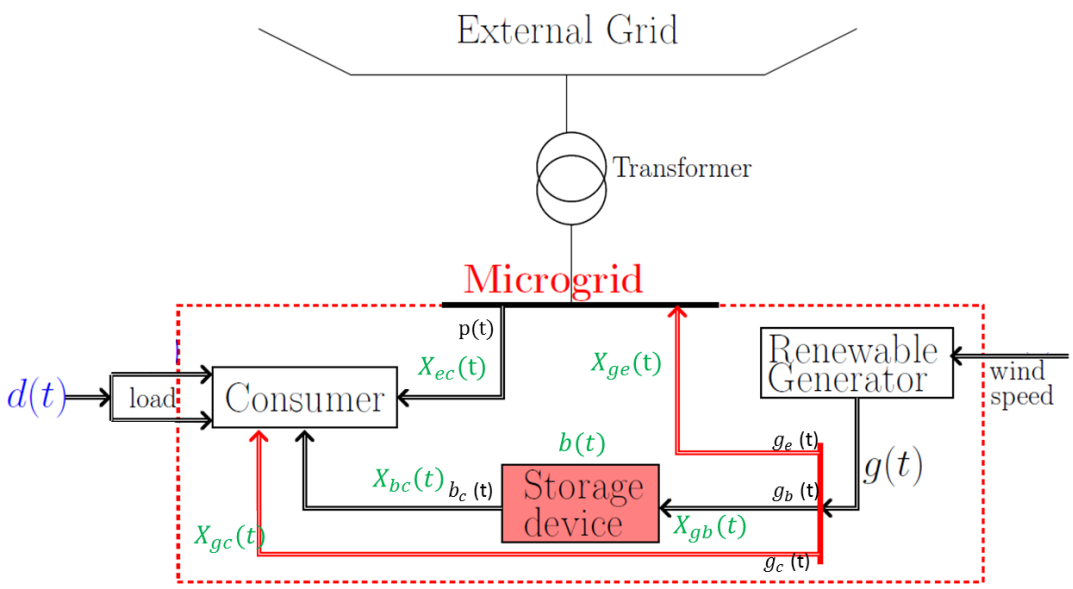


Figure 2.1: Microgrid

2.4.1 System modeling

The description of the system dynamics leads to a 6 elements state vector: 5 states contain the values of energy in the links that can propagate to the next node and the sixth state represents the battery energy level (Han et al., 2015):

$$\mathbf{x}(t) = [x_{ec}(t) \quad x_{ge}(t) \quad x_{gc}(t) \quad x_{gb}(t) \quad x_{bc}(t) \quad b(t)]^T$$

The corresponding dynamic models are:

$$\text{External grid to consumer:} \quad x_{ec}(t+1) = (1-\alpha)x_{ec}(t) + \alpha p(t) \tag{2.11}$$

$$\text{Generator to external grid:} \quad x_{ge}(t+1) = (1-\alpha)x_{ge}(t) + \alpha g_e(t) \tag{2.12}$$

$$\text{Generator to consumer:} \quad x_{gc}(t+1) = (1-\alpha)x_{gc}(t) + \alpha g_c(t) \tag{2.13}$$

$$\text{Generator to battery:} \quad x_{gb}(t+1) = (1-\alpha)x_{gb}(t) + \alpha g_b(t) \tag{2.14}$$

$$\text{Battery to consumer:} \quad x_{bc}(t+1) = (1-\alpha)x_{bc}(t) + \alpha b_c(t) \tag{2.15}$$

where $\alpha \in [0, 1]$ is a fixed constant, mainly dependent upon the size of the discretization step, and,

$$\text{Battery:} \quad b(t+1) = (1-\tau)b(t) + x_{gb}(t) - b_c(t) + w(t) \tag{2.16}$$

with the mixed-integer conditions (Prodan and Zio, 2014a):

$$\begin{cases} 0 \leq b_c(t) \leq Ma(t), \\ 0 \leq x_{gb}(t) \leq M(1-a(t)), \end{cases} \tag{2.17}$$

The above six state variables can be inferred by the vector of system control inputs $\mathbf{u}(t)$: $\mathbf{u}(t) = [p(t) \quad g_e(t) \quad g_c(t) \quad g_b(t) \quad b_c(t)]^T$ where (Han et al., 2015):

- $p(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the external grid to the consumer, at time step t.
- $g_e(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the external grid, at time step t.
- $g_b(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the battery, at time step t.
- $g_c(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the consumer, at time step t.
- $b_c(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the battery to the consumer, at time step t.

The consumer has the possibility to take electrical power from the external grid, the renewable generator and the battery. Thus, the sum of powers received by the consumer

is $x_{ec}(t) + x_{gc}(t) + x_{bc}(t)$. Finally, the system output $y(t)$ is the total power received by the consumer:

$$y(t) = x_{ec}(t) + x_{gc}(t) + x_{bc}(t)$$

In the end, the microgrid can be described by the following global dynamic model:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \tag{2.18}$$

where

$$\mathbf{A} = \begin{bmatrix} 1-\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-\alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\alpha & 0 \\ 0 & 0 & 0 & 1 & 0 & 1-\tau \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

2.4.2 Optimization-based control for system safety analysis

The safety performance of the microgrid is measured in terms of the satisfaction of consumer power demand and the optimization problem is to find the appropriate control inputs that minimize the cost function $Cost(t)$ defined as the difference between the power demanded by the consumer and that actually received. Thus, the objective function for

MPC is:

$$\min_{[\mathbf{u}(t)]_{t=k:k+N_p}} \sum_{t=k}^{k+N_p} d(t) - y(t)$$

with the set of constraints defined in the following equations (2.19) - (2.23).

- Satisfaction of consumer power demands

$$0 \leq x_{ec}(t) + x_{gc}(t) + x_{bc}(t) \leq d(t) \tag{2.19}$$

where $d(t)$ is the consumer's demand.

- Battery storage

Batteries have their physical characteristics: the minimum capacity B_{min} determined by the Depth of Discharge (DoD) (Prodan et al., 2015) and the capacity B_{max}. The rate of the battery charge is also limited by some bounds:

$$B_{min} \leq b(t) \leq B_{max}, \tag{2.20}$$

$$Br_{min} \leq \Delta b(t) \leq Br_{max}, \tag{2.21}$$

where $B_{min} \in \mathbb{R}$, $B_{max} \in \mathbb{R}$, $Br_{min} \in \mathbb{R}$, $Br_{max} \in \mathbb{R}$.

- Generator

The power taken from the generator by the battery, $g_b(t)$, the consumer, $g_c(t)$, and the external grid, $g_e(t)$, is bounded by the total power generated by the renewable generators:

$$0 \leq g_b(t) + g_c(t) + g_e(t) \leq g(t), \tag{2.22}$$

with $g_b(t) \geq 0$, $g_c(t) \geq 0$, $g_e(t) \geq 0$.

- Link capacities

Table 2.1: Numerical data for the microgrid components

Battery parameters	
τ	$1.3 \cdot 10^{-4}$
M	$9 \cdot 10^3$
B_{min} $[Wh]$	$1.2 \cdot 10^3$
B_{max} $[Wh]$	$9 \cdot 10^3$
Br_{min} $[W]$	$-1.5 \cdot 10^3$
Br_{max} $[W]$	$1.5 \cdot 10^3$
Control input constraints	
U_{min}	$[0 \quad 0 \quad 0 \quad 0 \quad 0]^T$
U_{max}	$[2 \quad 1.5 \quad 2 \quad 1.5 \quad 1.5]^T \cdot 10^3$
Prediction horizon	
N_p	7
Simulation steps	
N	300

$$\mathbf{u}_{min} \leq \mathbf{u}(t) \leq \mathbf{u}_{max}, \tag{2.23}$$

where $u(t) \in \mathbb{R}$.

The numerical values of the parameters used for the simulations are taken from (Grigg et al., 1999) and are presented in Table 2.1.

We consider two reference profiles characterizing the microgrid components (i.e. the consumer and the renewable generator) based on real numerical data taken from (Grigg et al., 1999) and the details can be found in (Prodan and Zio, 2014a). The consumer load takes into account seasonal numerical data, and is predicted by using weekly, daily and hourly peaks. Figure 2.2 shows the consumer load profile $d(t) \in \mathbb{R}$. The wind speed used to calculate the wind power profile is estimated based on meteorological data. Figure 2.3 shows the power generator profile: the electrical power generated $g(t) \in \mathbb{R}$ is obtained from the wind speed profile (Justus et al., 1976).

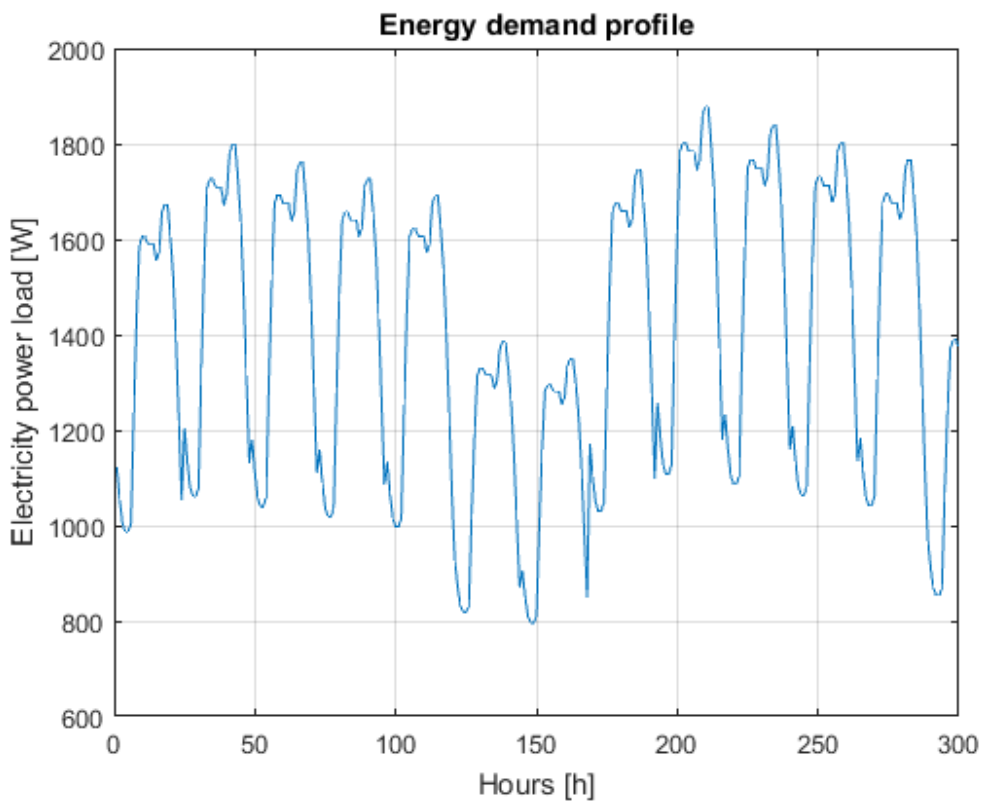


Figure 2.2: Consumer load profile

Figure 2.3: Wind power profile

2.4.3 Analysis of the results

We consider two operation modes for the microgrid: grid-connected and stand-alone. Under these two modes, the microgrid is designed to satisfy consumer's demand. We assume that the external grid and the renewable generator are fault-free. Then, threats to the microgrid service may come from failure of the links from the three sources (i.e. the external grid, the renewable generator and the battery) to the consumer. The two other links from the renewable generator are also considered, since they impact on the cost of the microgrid. The links failures are described as removal of the links and no recovery action is taken into account.

We analyze the system property indexes introduced in Section 2.3 for each scenario. For the grid-connected mode, we also analyze the difference between cost and profit of the microgrid for each scenario. The simulation of each scenario is considered for the period of 300 hours, during which the microgrid experiences almost all extreme conditions of

consumer demands and winder power.

Grid-connected mode

During the grid connected mode, the consumer takes electrical power from two sources: the external grid and the renewable generator, and the demands can be fully satisfied. Note that the failure of the battery is not considered for the grid-connected mode, since in that mode it is assumed that the battery is not used.

The scenarios considered are as follows:

- Scenario 1.0: the nominal functioning case, i.e. fault-free. The consumer takes electrical power from the external grid and the renewable generator.

- Scenario 1.1: the link from the generator to the external grid is disconnected (i.e. g_e is removed), and it's impossible to sell electricity to recompense the expenses on electricity bought from the external grid, therefore, instead of making a profit, the microgrid exclusively spends money to buy electricity from the external grid.

- Scenario 1.2: the link from the generator to the consumer is disconnected (i.e. g_c is removed). The consumer is supplied completely by the external grid.

Table 2.2: Index values for the grid connected mode

Scenario	NSD	CI	EI^{st}	Profit
1.0	0	1	1	+203.2
1.1	0	0.83	1	-77.0
1.2	0	0.83	1	-79.2

From Table 2.2, we can see that in the grid-connected mode, the demands of the consumers can always be satisfied. The two links g_e and g_c have identical influence on the system controllability: with the removal of each link, CI drops from 1 to 0.83, which means that the microgrid is no longer controllable, and the rank of controllability matrix decreases by 1, which means that one component is out of control. The capacity efficiency index, EI^{st}, remains the same for all the scenarios. However, the cost differs a lot: when

the generator is able to provide power to the consumer and sell power to the external grid, the microgrid is profitable (Scenarios 1.0); otherwise, the microgrid spends money on buying electricity from the external grid (Scenarios 1.1 and 1.2).

Figures 2.4 and 2.5 show the sources of the power actually received by the consumer for Scenario 1.0 and Scenario 1.1, respectively.

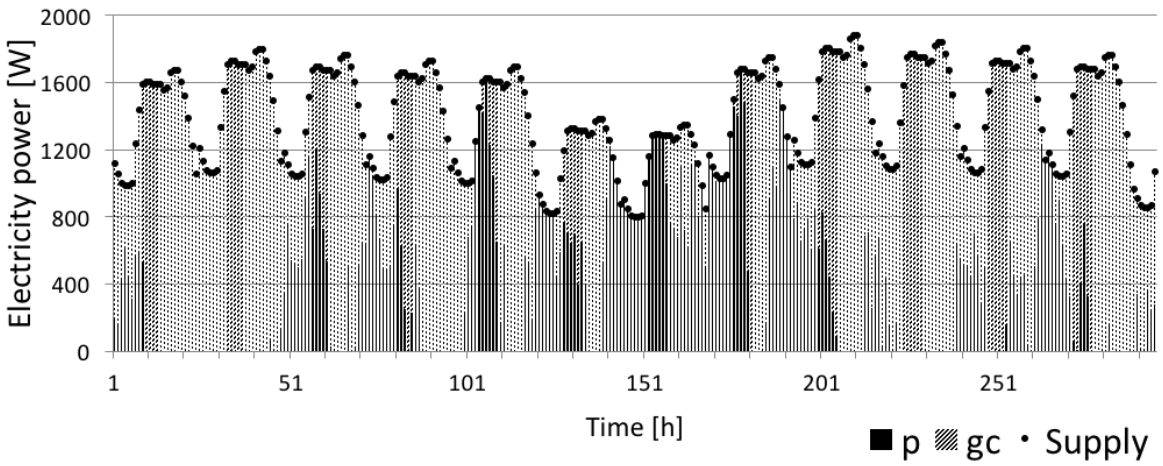


Figure 2.4: Supply sources of Scenario 1.0. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator.

In the grid-connected mode, the consumer's demand is always satisfied, which is reasonable. Furthermore, the introduction of the microgrid (renewable generator) decreases the cost on electricity and even makes a profit and it is, thus, interesting for economic considerations.

Stand-alone mode

In the the stand-alone mode, the microgrid is disconnected from the external power grid and the power generation should by itself satisfy the consumer's demand.

The scenarios considered are as follows:

- Scenario 2.0: the stand-alone functioning case (i.e. only p is disconnected).

- Scenario 2.1: the link from the generator to the consumer is disconnected (i.e. g_c is

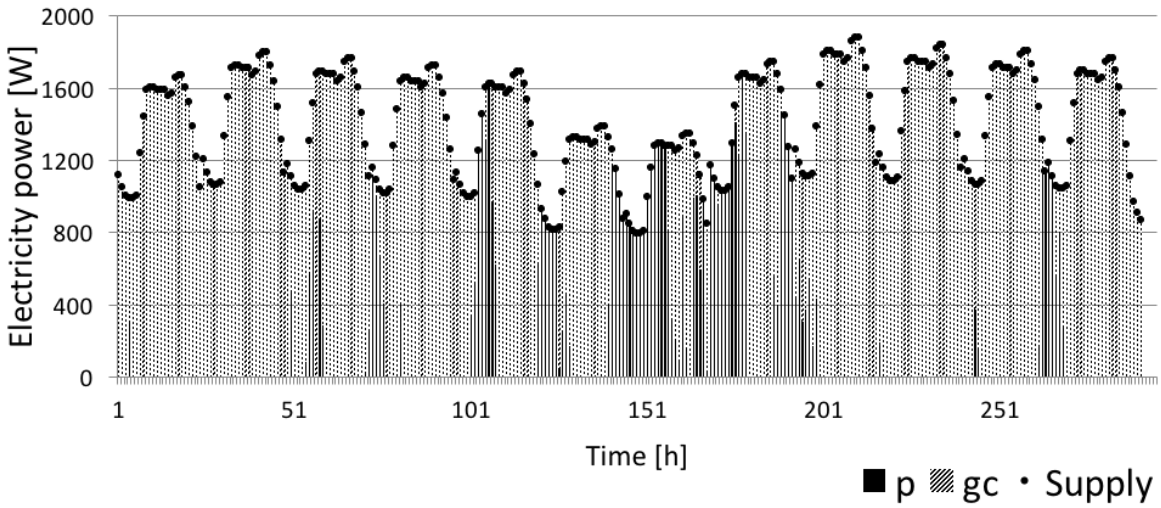


Figure 2.5: Supply sources of Scenario 1.1. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator.

removed).

- Scenario 2.2: the link from the generator to the battery is disconnected (i.e. g_b is removed).

- Scenario 2.3: the link from the battery to the consumer is removed (i.e. b_c is removed)

Table 2.3: Index values for the stand-alone mode

Scenario	NSD	CI	EI^{st}
2.0	0.1065	0.83	0.667
2.1	0.6562	0.83	0.583
2.2	0.4190	0.83	0.667
2.3	0.4364	0.83	0.333

From Table 2.3, we can see that the non-supplied demand NSD increases in this mode and the demands are never fully satisfied. More detailed analysis is given below. The capacity efficiency index EI^{st} decreases compared to the stand-alone because the disconnection of the link p reduces the transmission capacity of the microgrid; this also contributes to the inadequate supply of the consumer's demands.

Figure 2.6 shows the sources of the power actually received by the consumer for Scenario 2.0. From the initial time to $t = 145$ hours under this scenario, the supply is similar to that of Scenarios 1.0 and 1.1: the renewable generator provides most of the demanded power; but, instead of the the external power grid, the battery fills the unsatisfied portion, except for certain periods of low wind power generation. Then, the microgrid arrives at a relatively long period when there is no wind power at all, the battery reaches its lower limit, and the supply is totally cut off. At around t=200, the microgrid continues to function when the wind comes back to its usual level.

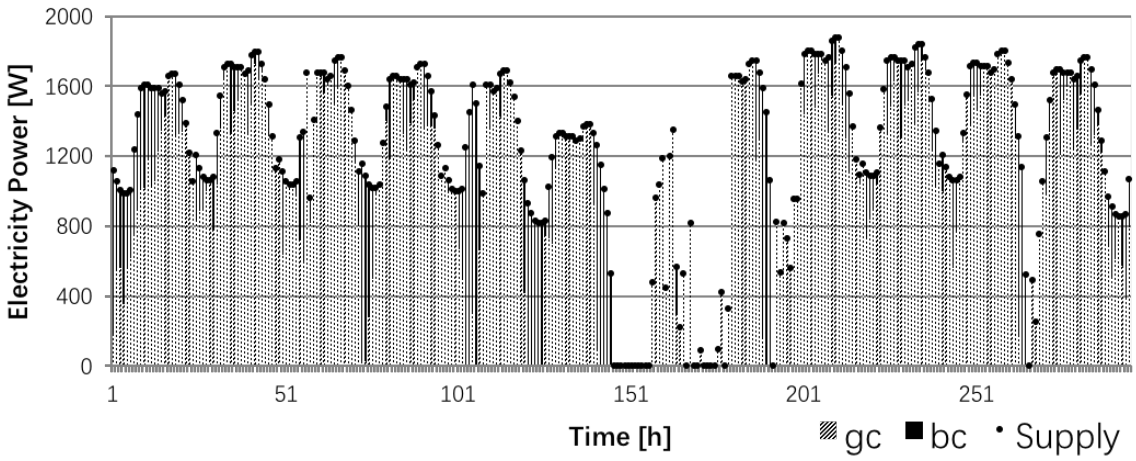


Figure 2.6: Supply sources of Scenario 2.0. The black area represents the power from the battery and the shaded area represents the power from the renewable generator.

In Scenario 2.1, the control input of the link from the generator to the consumer g_c is lost, while the battery provides power to the consumer and can be charged by g_b. In this scenario, the value NSD is the highest of all the stand-alone modes. This is due to the fact that the supply to the consumer is dominated by the capacity of the link from the battery (g_b has smaller capacity than g_c), which is also reflected by the fact that EI^{st} is lower than that of Scenario 2.0. In addition, when the battery reaches its lower limits and switches to charge mode, there is no supply to the consumer.

In Scenario 2.2, the link from the generator to the battery g_b is removed, i.e. the battery can no longer be charged. in this scenario, the battery can provide part of the demanded power in the beginning and when the battery reaches its lower limit at $t = 20$ hour, the renewable generator becomes the only source of supply to the consumer. NSD is much higher than that of Scenario 2.0. The consumer's demand can be satisfied only when there is enough wind power. During the period of high wind speed (the power generated can reach 6KW, which is much higher than the largest demand), the redundant power generated can not be stored. The capacity efficiency index EI^{st} is the same as Scenario 2.0, since the failed link does not affect the supply to the consumer directly; however, without the link, the battery does not have income any more, which decreases the supply.

In Scenario 2.3, the link from the battery to the consumer b_c is failed, and the renewable generator is the only source to supply the consumer for the whole period. Under this scenario, the non-supplied demand NSD is similar to Scenario 2.2; but, without the contribution of the initial energy stored in the battery, NSD is slightly higher. In addition, the capacity efficiency is the lowest for Scenario 2.3, since there is only one source-terminal path left, i.e. the link from the generator to the consumer.

In the stand-alone mode, we have $NSD > 0$ for all the scenarios considered. But, because of the integration of the renewable generator into the microgrid, a large part of the demand can still be supplied. Furthermore, the importance of the storage device is highlighted through the comparison of the Scenarios 2.0, 2.2 and 2.3.

2.4.4 Summary

In this chapter, we have adopted model predictive control for describing microgrid dynamics and analyzed system performance under grid-connected and stand-alone modes, for different failure scenarios. This analysis enables quantitative evaluation of microgrid performance with respect to different perspectives of reliability, controllability and topology.

We have considered a specific case study, which confirms the fact that the microgrid being connected to the external power grid is important to insure supply to the consumer under different failure scenarios and that the introduction of microgrids composed of renewable generators and storage devices improves reliable performance of the power grid. The instability of the wind power and the limited capacity of the battery or links can be a barrier to the reliable service of the microgrid, especially when in stand-alone mode.

The findings of analyses of this kind provide information for the design and operation of the microgrid, seeking the right balance of multiple characteristics and least cost for the safe and reliable functioning of the microgrid system.

Chapter 3

A multi-perspective framework for the analysis of CIs

In this chapter, we propose a framework for analyzing large-scale CIs, with the objective of identifying the most critical elements and failure scenarios and evaluating their consequences. Three perspectives of analysis are considered: supply service, topology and controllability.

Section 3.1 briefly reviews complex network theory and controllability of complex networks. Section 3.2 introduces the three system-level indexes considered for CIs analysis in this thesis. Section 3.3 considers a case study of a real gas transmission network and analyzes of link importance and consequences of failure scenarios from the three perspectives.

3.1 State of the art

3.1.1 CIs as complex networks

CIs are complex networks of interacting components. Complex network theory has been used to study a wide range of systems, such as: social networks, technical networks, cellular networks, written human language, etc. (Albert and Barabási, 2002; Newman,

2003). Complex network theory builds a model of real-world networks and describes the form and, in various degree, the functionality of the network by different measures. Complex network theory can be applied to the analyze CIs and identify preliminary vulnerabilities by topology-driven and dynamical analysis, and critical areas that need further detailed analysis (Kröger and Zio, 2011).

Network representation

Graph theory provides a natural mathematical framework to represent complex networks (Fang and Zio, 2013a; Lombardi and Hörnquist, 2007). A graph is an ordered pair $\mathbf{G}(V, E)$ comprising a set of nodes (vertices) $V = v_1, v_2, \ldots, v_n$, together with a set of links (also called edges or arcs) $E = e_1, e_2, \ldots, e_m$, which are two-element subsets of V. Network structures are usually defined by an $n \times n$ adjacency matrix $\mathbf{Adj}$, which is constructed as follows: if there is an edge from node i to node j, then we have $Adj_{ij} = 1$; otherwise, $Adj_{ij} = 0$.

Topological characteristics

Within the complex network framework, failures of CIs are typically modeled topologically by removing nodes and links. Component importance is usually quantified by centrality measures, which describes the relative importance of an individual component in terms of supporting the interaction and communication that occurr in the network. Important and classical topological centrality measures include (Freeman, 1978; Nieminen, 1974):

- Degree centrality: the degree of a node v, $k(v)$ normalized over the maximum number of its possible neighbors: $C_D(v) = \frac{k(v)}{N-1}$.
- Betweenness centrality: the number of shortest paths from all nodes in the network to all others that passes through the given node. $C_B(v) = \sum_{i \neq j \neq v \in V} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$.
- Closeness centrality: $C_{Cl}(v) = \frac{N-1}{\sum_{j \in V} d_{ij}}$ the average length of the shortest path d_{ij} between the node and other nodes in the graph. The more central a node is, the closer it is to all other nodes.

- Eigenvector centrality: $C_{Eig}(v)$ uses the eigenvector corresponding to the largest eigenvalue of the graph adjacency matrix to assign relative scores to all nodes.

3.1.2 Controllability of complex networks

As CIs evolve and become more and more dependent on information technologies, it is essential to understand their controllability. Therefore, it is desirable to develop a control framework to steer the network dynamics towards optimal performance, while avoiding undesired or unfavorable states. Control is a fundamental property for the safe and reliable operation of CIs. However, control of complex network systems that make up a CI is challenging, due to their large scale and complexity.

For complex network systems, the controllability problem can be formulated as follows: find a suitable control matrix $\mathbf{B}$ consisting of a minimum number of driver nodes (nodes controlled by an outside signal), so that the Kalman's rank condition (2.6) is satisfied. However, this requires evaluating of the rank of the controllability matrix for all 2^N possible combinations of the driver nodes (Lombardi and Hörnquist, 2007). For real CIs, such a brute-force search is computationally prohibitive.

In (Liu et al., 2011), analytical tools have been developed to characterize the controllability of directed networks. Maximum matching is used to determine the minimum number of inputs (or driver nodes) that can guide the system's entire dynamics. Full control can be achieved if and only if each unmatched node is directly controlled and there are directed paths from the input signals to all the matched nodes. The unmatched nodes are the driver nodes. The number of driver nodes, N_D, is a measure of the controllability of the network and it influences the resources needed for controlling the network. If $N_D = N$, i.e., the total number of nodes in the network, (this means that the external control signal is applied to every node in the network), the likelihood of gaining full system control is high, while the associated cost is also high (Wang et al., 2012). A small value of N_D, instead, indicates a more controllable network system, in the sense that it requires less effort to obtain full control of the network.

Several related topics can be considered under this framework, such as control centrality (Liu et al., 2012), achieving full control by using only one controller (Cowan et al., 2012), optimization (Wang et al., 2012), control energy(Yan et al., 2012), control capacity (Jia and Barabási, 2013), control mode (Jia et al., 2013), control of edge dynamics(Nepusz and Vicsek, 2012) etc. Structural controllability of temporal networks are also studied (Pan and Li, 2014; Pósfai et al., 2012).

In (Yuan et al., 2013), the exact controllability for arbitrary network structures and link weights (say arbitrary matrix $\mathbf{A}$) is introduced to calculate N_D:

$$N_D = \max_i \{\mu(\lambda_i)\} \tag{3.1}$$

and the minimum number of driver nodes N_D is determined by the maximum geometric multiplicity $\mu(\lambda_i)$ of the eigenvalue λ_i of $\mathbf{A}$. In fact, these are the nodes corresponding to the linearly-dependent rows: the controllers should be imposed on the linearly-dependent rows to eliminate all linear correlations and ensure the controllability condition.

3.2 System-level indexes

In this chapter, we consider three perspectives of CI assessment: control, supply and topology. For each perspective, we propose an index to evaluate the network performance.

Supply index (Non-supplied demand)

We propose to use the non-supplied demand (NSD), similar to that introduced in Equation 2.5 in Chapter 2.3.1, as a measure of the network's capability to satisfy users' demands. Consider a CI network of N nodes and L links, which supplies service or products from N_s supply nodes (sources) to its N_y user nodes (users) through a number of transmission nodes. The supply to each user can be computed by maximum flow algorithms (Deo, 2016). Then, the normalized NSD is:

$$NSD = 1 - \frac{\sum_{i=1}^{i=N_y} \omega_i y_i}{\sum_{i=1}^{i=N_y} \omega_i D_i} \tag{3.2}$$

where, ω_i is the weight of the ith user, y_i is the supply to user i and D_i is its demand, N_y is the number of users. The second term in Equation 3.2 is the fraction that the user's demand is satisfied. Since $y_i \leq D_i$, the index NSD is normalized to $[0, 1]$, where NSD equals 0 when the users' demands are fully supplied.

Topological index (Network topological efficiency)

To measure a network's performance from topological perspective, network topological efficiency is proposed. Network topological efficiency is a measure of the connectivity of the network, i.e. of how well the nodes exchange flow (Latora and Marchiori, 2001). This measure is based on the assumption that the flow travels along the shortest routes, and that the efficiency of the communication between two nodes i and j, denoted by ε_{ij}, is inversely proportional to their shortest path length d_{ij}, which is defined as the smallest sum of *physical distances* throughout all the possible paths in the weighted network. When there is no path between i and j, $d_{ij} = +\infty$, i.e $\varepsilon_{ij} = 0$. Then, the topological efficiency of the whole network is given by:

$$E[\mathbf{G}] = \frac{\sum_{i \neq j \in \mathbf{G}} \varepsilon_{ij}}{N(N-1)} = \frac{1}{N(N-1)} \sum_{i \neq j \in \mathbf{G}} \frac{1}{d_{ij}} \tag{3.3}$$

Controllability index

To measure the structural controllability of the network system, we adopt the controllability index (C_{ind}) first introduced in (Li et al., 2016):

$$C_{ind} = \frac{N - N_D}{N} \tag{3.4}$$

where N is the number of nodes in the network and N_D is the minimum number of driver nodes that are needed to fully control the network. Also, the index C_{ind} is normalized to $[0, 1]$. The occurrence of failures (represented as the removal of links) is likely to increase the number of the linearly-dependent rows in matrix $\mathbf{A}$ and, thus, N_D would increase and C_{ind} decreases; when the current number of control nodes is insufficient to obtain full control over the whole system, there is no guarantee that the system can be brought

back to the designed operation condition. Thus, a larger value of C_{ind} indicates a more controllable network system.

3.3 Application

3.3.1 Network description

We consider the case study from (Praks and Kopustinskas, 2016) to demonstrate the use of the proposed indexes. The system is visualized in Figure 3.1 and represents a real gas transmission network for supply across several countries in the EU. The gas transmission network is modeled as an undirected graph, including 56 nodes and 74 links, where nodes represent sources or substations and links represent the gas transmission pipelines connecting the nodes. Among the 74 links, 10 links are virtual links representing the virtual connection of parallel pipelines, and their failure is not considered. Its connectivity structure can be defined by its $N \times N$ adjacency matrix **Adj**, whose entries $[Adj_{ij}]$ are equal to 1 if there exists a link from node i to node j and 0 otherwise.

Each link in the network is characterized by its capacity, i.e., the maximum amount of flow that it is able to transmit, and its length. The capacity matrix **K** contains information about the capacity constraints on the network elements including source nodes, demand nodes and pipelines. The length matrix **Len** contains the lengths of the edges between nodes: entry Len_{ij} is the length of the pipeline connecting the i-th and j-th nodes; an entry of 0 indicates that the i-th and j-th nodes are not connected.

We distinguish $N_y = 35$ demand nodes, with deterministic daily demands 45.9 millions of cubic meters (mcm) for the total system, one LNG terminal (node 10), two compressor stations (nodes 11 and 12), two storage devices (nodes 10 and 19) and two pipeline source nodes (nodes 2 and 29).

The properties of the four supply nodes (sources), i.e. nodes 2, 10, 19 and 29, are presented in Table 3.1. The capacities and demands are expressed in millions of cubic meters per day (mcm/d). The data of supply and demand are realistic and they are

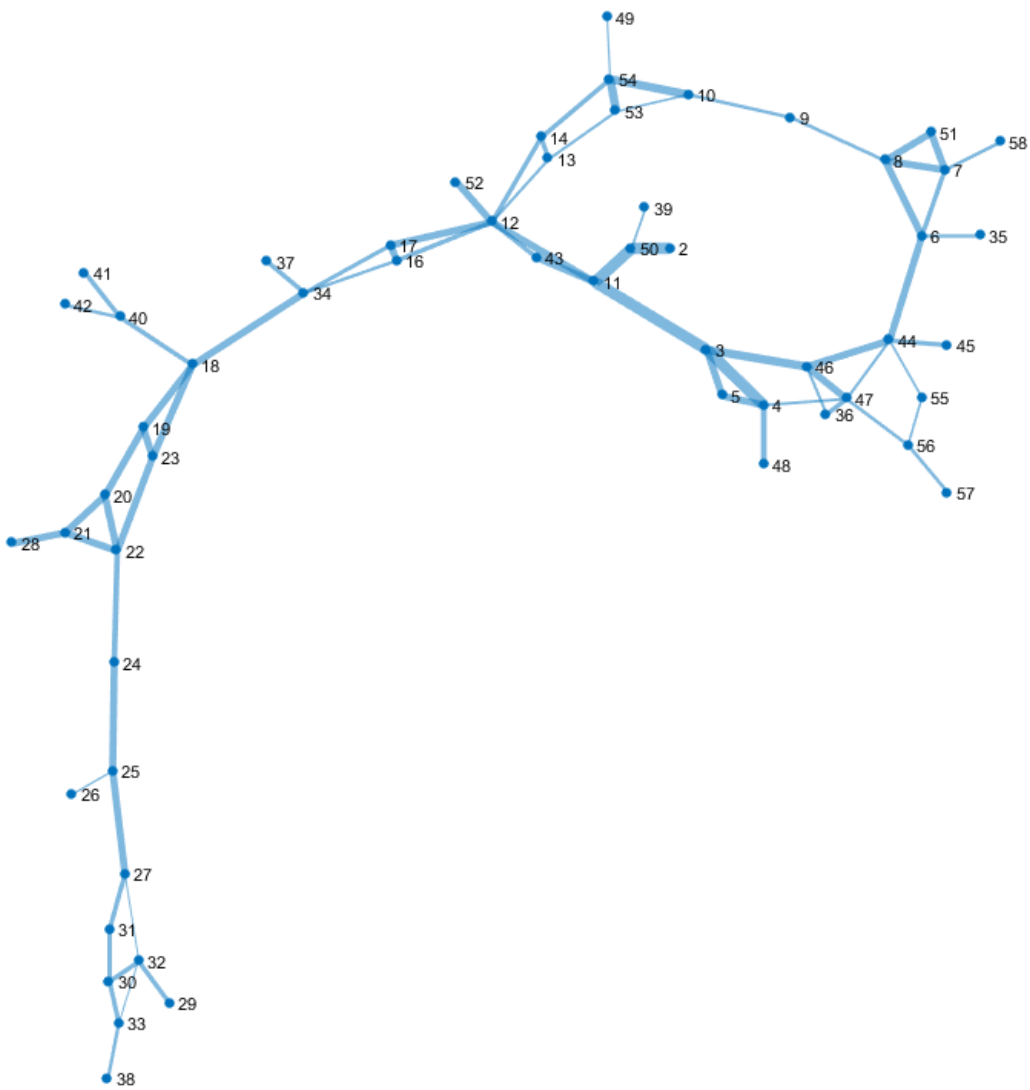


Figure 3.1: Gas transmission network(Praks and Kopustinskas, 2016)

expressed on a daily scale, in order to capture the fluctuation of peak gas demand during one year (high gas demand in winter) (Praks et al., 2015). These data are intended to represent the most stressed conditions for the gas transmission network. Depending on the purpose of the analysis, variable values of demand or supply can be considered, and the user demands satisfaction can be evaluated as the average over a simulation horizon, with associated uncertainty.

Given the capacities of the links connecting the nodes and the constraints on the sources and users, the supply to each user is used to calculate the non-supplied demand NSD defined in Equation 3.2 by maximum flow algorithms (Praks et al., 2015).

Node	Capacity	Type
2	31	Pipeline source
10	10.5	LNG terminal
19	25	Underground storage
29	4	Pipeline source

Table 3.1: Sources properties

We consider the failure of the LNG station, compressor stations, storages and 64 pipelines.

The failure of an LNG terminal and storage devices is modeled as the total capacity loss of each pipeline connected to it. According to (Praks and Kopustinskas, 2016), the monthly failure frequency of the LNG (node 10) is set to $f_{LNG} = 1.25E - 2$, and the monthly failure frequency of the storage (node 19) is $f_S = 8.33E - 3$.

When a compressor station fails, the capacity of each pipeline connected to it will be reduced by 20%. The monthly failure frequency of the two compressor stations (nodes 11 and 12) are $f_{CS} = 2.08E - 2$.

According to the EGIG report (EGIG, 2011), the average failure frequency of a European gas transmission pipeline is $3.5E - 4$ per kilometer-year. We consider the total rupture of a pipeline and we assume that 10% of the failures reported in a year cause such a rupture. Thus, the monthly failure frequency of a pipeline is $f_P = 2.92E - 6$ per kilometer (Praks and Kopustinskas, 2016).

3.3.2 Quantification of link importance

Table 3.3.2 presents the values of the three indexes calculated for the nominal network configuration. To analyze link importance, we systematically disconnect one link at a time from the original network and compute the three indexes of the new network configuration $\mathbf{G}'$.

Table 3.3.2 presents the ten most critical links in terms of NSD, the three most critical links in terms of C_{ind}, and the single most critical link in terms of E.

With the removal of single links, the NSD value ranges from 0 to 0.363. Pipelines

NSD	C_{ind}	E
0	0.9107	0.6327

Table 3.2: Indexes values for the nominal configuration

Link	NSD	C_{ind}	E
3-11	0.363	0.9107	0.6319
3-46	0.209	0.9107	0.6327
21-28	0.131	0.8929	0.6327
2-50	0.126	0.9107	0.6325
11-50	0.120	0.9107	0.6323
6-44	0.106	0.9107	0.6321
44-46	0.081	0.9107	0.6327
36-47	0.048	0.8929	0.6327
4-48	0.039	0.8929	0.6318
34-37	0.028	0.8929	0.6327
44-45	0.028	0.9107	0.6318
18-40	0.026	0.9107	0.6325
6-35	0.002	0.8750	0.6325
11-43	0	0.8750	0.6326
18-23	0	0.8750	0.6327
18-34	0	0.9107	0.6318

Table 3.3: Indexes values associated to the removal of the most critical links

3-11 and 3-46 are of large capacity, so they are essential to supplying the demand nodes in their neighborhood, and thus their importances are significant in terms of supply. A similar explanation applies for the removal of links 6-44 and 44-46. Node 28 is a large demand node, and therefore, the removal of link 21-28, which is its only connection to the network, will affect the overall network NSD. The same explanation also applies for the impact of links 4-48, 34-37 and 44-45. Links 2-50 and 11-50 connect the main source (node 2) to the rest of the network, and their removal leads to a deficit in supply capacity, since the remaining sources 10, 19 and 29 are not capable of fully supplying the total demands. Link 4-48 and link 44-45 are critical in terms of topological efficiency E, because their removal disconnects the end nodes 48 and 45, respectively and, thus, decreases the network efficiency; moreover, considering that they are relatively short pipelines, the value of E

drops much more than for the removal of links 34-37 and 21-28, which are long pipelines. Link 18-34 is also a critical link in terms of topological efficiency: when it fails, the network will break into two separate parts and no gas flow can be transferred between them, so that, the topological efficiency E drops.

To rigorously quantify the importance of a link, we consider Risk Achievement (RA), an importance metric that measures the contribution of the failure of the generic i^{th} component ($x_i = 1$) to the system risk index (Van der Borst and Schoonakker, 2001): $RA = R(x_i = 1) - R(base)$, where $R(x_i = 1)$ is the increased risk index when the ith component fails and $R(base)$ is the risk of the nominal conditions. Considering the fact that a smaller value of NSD indicated lower risks, the RA of NSD is defined by:

$$IM_{ij}^{NSD} = NSD[\mathbf{G}(base)] - NSD[\mathbf{G}'(x_{ij} = 1)] \tag{3.5}$$

Similarly, we define the RA for C_{ind} and E as:

$$\begin{aligned} IM_{ij}^{C_{ind}} &= C_{ind}[\mathbf{G}(base)] - C_{ind}[\mathbf{G}'(x_{ij} = 1)] \\ &= \frac{N_D[\mathbf{G}'(x_{ij} = 1)] - N_D[\mathbf{G}(base)]}{N} \end{aligned} \tag{3.6}$$

$$IM_{ij}^{E} = E[\mathbf{G}(base)] - E[\mathbf{G}'(x_{ij} = 1)] \tag{3.7}$$

where $\mathbf{G}'(x_{ij} = 1)$ is the graph of the network obtained by removing the link $i - j$ from the original network $\mathbf{G}(base)$.

Figure 3.2 shows the links importance metrics in terms of the three indexes: the left vertical axis is the values for NSD (triangles) and C_{ind} (squares), while the right vertical axis is the topological efficiency E (circles). It is seen that NSD has the widest range, while a single link disconnection has little impact on controllability and topological efficiency.

Among the 64 links, only 23 links have impact on NSD and most of them also have an impact on E, but only five links influence the controllability index.

The influence of a link is not the same from the three perspectives, which confirms the need of a multi-perspective framework of analysis.

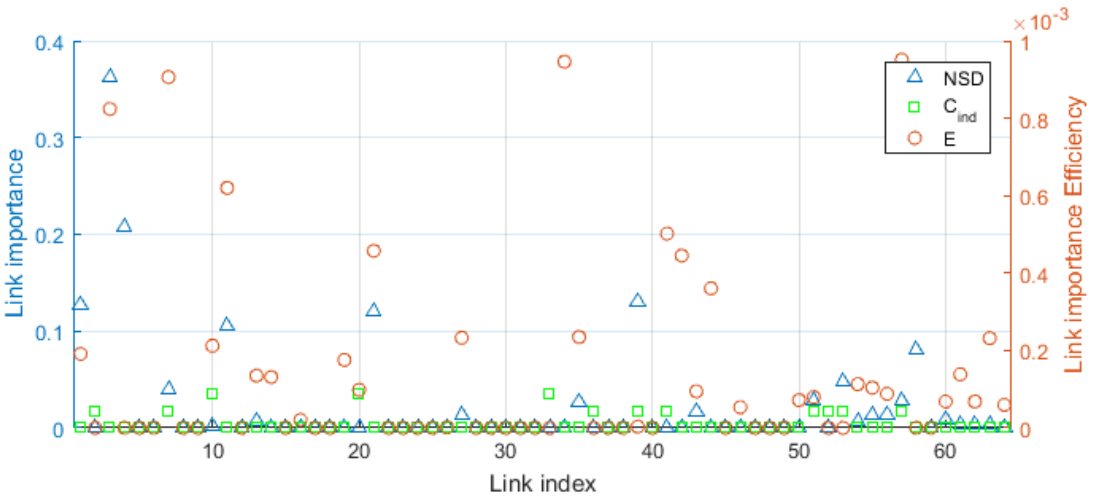


Figure 3.2: Link importance

3.3.3 Simulation and analysis

We have run 10^6 dynamic simulations by ProGasNet (Praks and Kopustinskas, 2016), sampling nodes and links according to their occurrence probabilities, as introduced in Section 3.3.1. A total of 335 different gas transmission states (cases) emerge from the sampled configurations. The most frequent sample state is the one with no link failures.

We classify the 334 failure cases into different categories by their causes. We consider the thirty most frequent states and investigate how these affect the three system-level indexes considered. For each of the indexes, we quantify their consequences and analyze the impacts of different types of failures.

Both links and nodes of the gas transmission network may fail and multiple failures may occur. In order to understand the influence of different types of failures, we classify the 334 failure cases into seven classes as:

- Single link failure (SL)
- Single node failure (SN)
- Single link failure and single node failure (SL-SN)
- Single link failure and multiple node failures (SL-MN)
- Multiple link failures (ML)

- Multiple link failures and single node failure (ML-SN)
- Multiple node failures (MN)

Single node failure (SN) includes 4 cases (only four nodes may fail according to our definition), but they cover 83.23% of the failure contributions. Single link failure (SL) includes 61 cases and is the second most frequent class. The cases of SL-MN, ML and ML-SN only occur once in all simulations performed. Figure 3.3 shows the number of cases for each class and their counts.

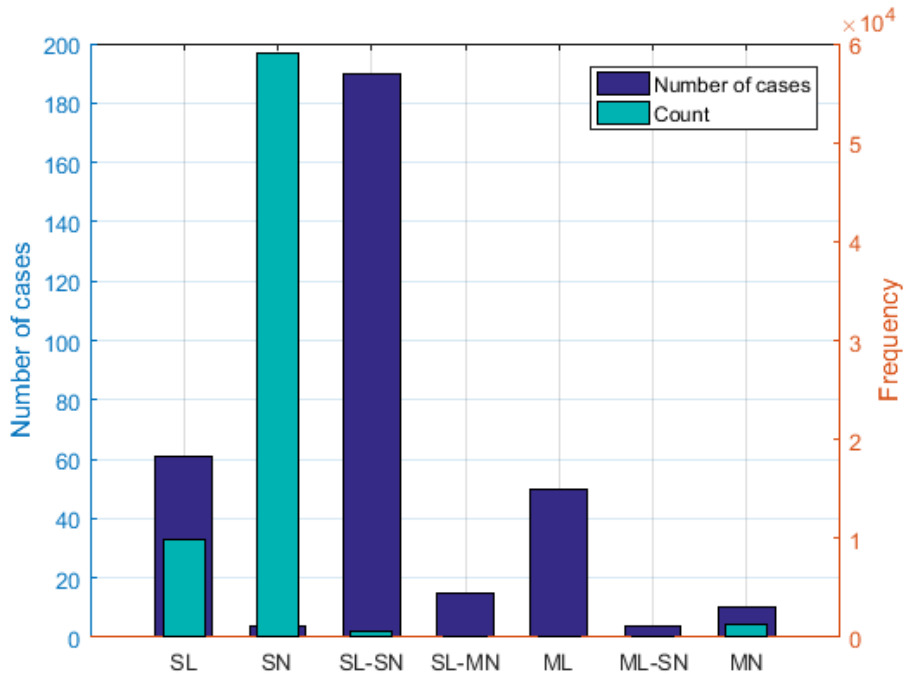


Figure 3.3: Histogram: number of cases and frequency

As we are analyzing an existing gas transmission network, it is reasonable to have low probabilities for multiple failures scenarios; however, the failures of low frequency of occurrence may have an important impact on the properties of the system, which is one of our interests in this study. Therefore, although the probability of their occurrence is small, it is interesting to consider such multiple-element failures and understand the corresponding consequences, which provides additional information for CIs design.

We also consider the 30 most frequent cases and apply the analysis framework. Table

3.4 summarizes the failure types and frequencies of the 30 most frequent cases. Node failure is the most common failure type, the four most frequent failure cases being the four single node failure (SN) cases.

Type	Cases	Frequency (over 10^6 simulations)
Failure free	1	929013
SN	4	59040
SL	21	6238
MN	4	1098

Table 3.4: 30 most frequent cases

We analyze the three indexes separately, with the objectives of identifying the failures affecting each index, quantifying their consequences in terms of loss in the properties considered and calculating their frequency.

Non-supplied demand

Figure 3.4 shows the NSD index value for the 335 cases, where the abscissa axis is the frequency rank of the 335 cases. The non-supplied demand NSD ranges from 0 to 0.64. In the presence of multiple failures, the network may reach a much higher level of non-supplied demand NSD. The highest value 0.64 corresponds to the SN-SL case where both node 19 and link 2-50 are failed. Node 19 is the second largest source and its failure alone results in $NSD = 0.2261$, since without it, some demand nodes far from the main source (e.g. node 2) are not fully supplied due to the limited capacity of pipelines connecting different areas (e.g. link 18-34), even though the total supply capacity of the sources is able to cover all demands. Combining with the failure of link 2-50, which cuts the supply from the main source (node 2) to other nodes, the supply of the whole network drops even more.

$NSD = 0$ is the most frequent value. Generally speaking, high values of NSD ($NSD > 0.3$) tend to have low frequency. For 146 out of 335 cases (43.6%), the demand can not be fully supplied, i.e. $NSD > 0$.

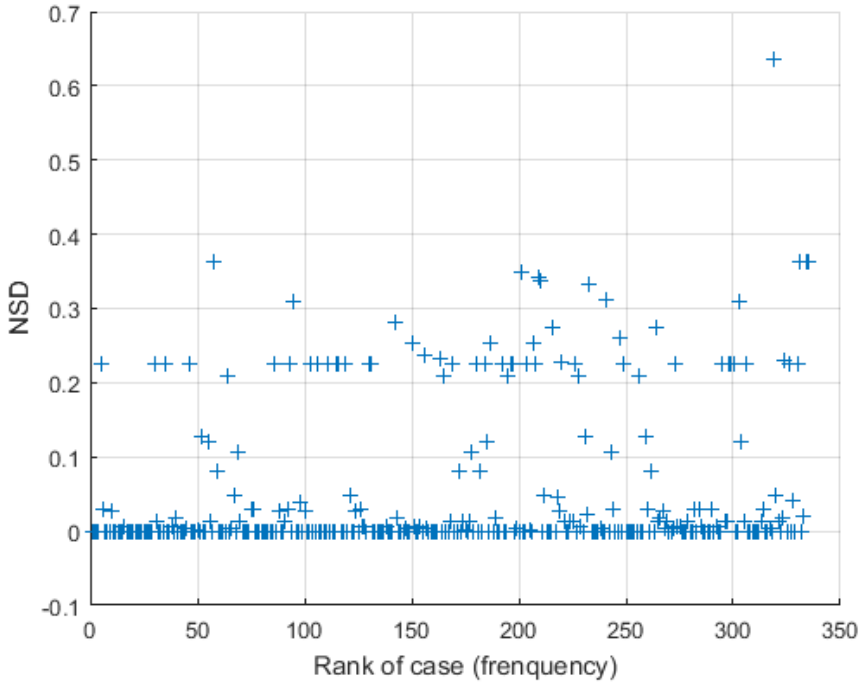


Figure 3.4: Non-supplied demand for the 335 cases

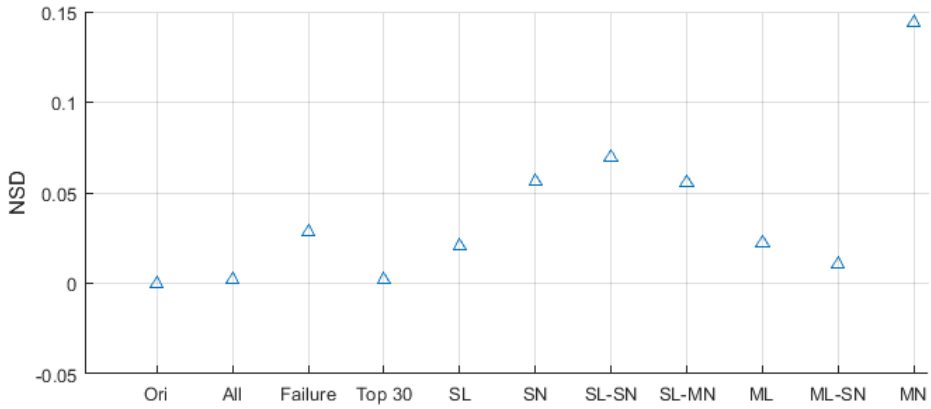


Figure 3.5: Non-supplied demand for different categories

Figure 3.5 shows the NSD of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN, ML, ML-SN, MN).

The NSD for the top 30 cases is higher than that for the nominal configuration and comparable to that of all configurations. This indicates that, the most frequent cases have

non-negligible impact on the demand supply.

If we compare the seven failure classes, we see that, as a whole, node failures have a more important impact in terms of non-supplied demand. In presence of node failure (SN, SL-SN and SL-MN), the non-supplied demand NSD is higher than for the cases with single link failure (SN). The combination of link failures only increases slightly the NSD: the mean NSD of ML is slightly higher than that of SL. As for ML-SN, among all the possible combinations of failures, only four cases occur once each, and they happen to have relatively small influence in terms of NSD: the mean being 0.010. The class of multiple node failures (MN) has a significantly high value of non-supplied demand. All MN cases without failure of node 19 have $NSD = 0$. The failure of node 19 alone would lead to $NSD = 0.2261$, and for the case where node 10, 11 and 19 all fail at the same time, $NSD = 0.3111$. In fact, the link 3-11 represents a large capacity pipeline and in absence of gas supply from the storages nodes 10 and 19, the reduction of its capacity due to the failure of node 11 would result in the non-supply to demand nodes in the vicinity, depending mainly on the main source (node 2). Considering the relatively high failure probability of node 19, the MN class has a high value of NSD.

Controllability index

Figure 3.6 shows the index value of C_{ind} for the 335 cases and the abscissa axis is the frequency rank of the 335 cases. The controllability index takes three values 0.9107, 0.8929 and 0.875 for the 335 cases, as for the removal of single links. $C_{ind} = 0.9107$ is the most frequent value. For 58 out of 335 cases (17.3%), that is, 1559 out of the 70987 failure configurations (2%), the controllability index C_{ind} is lower than that of the failure-free network configuration. The lowest value 0.875 is more present for the less frequent cases (i.e. Rank>150).

Figure 3.7 shows the mean value of C_{ind} of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN,

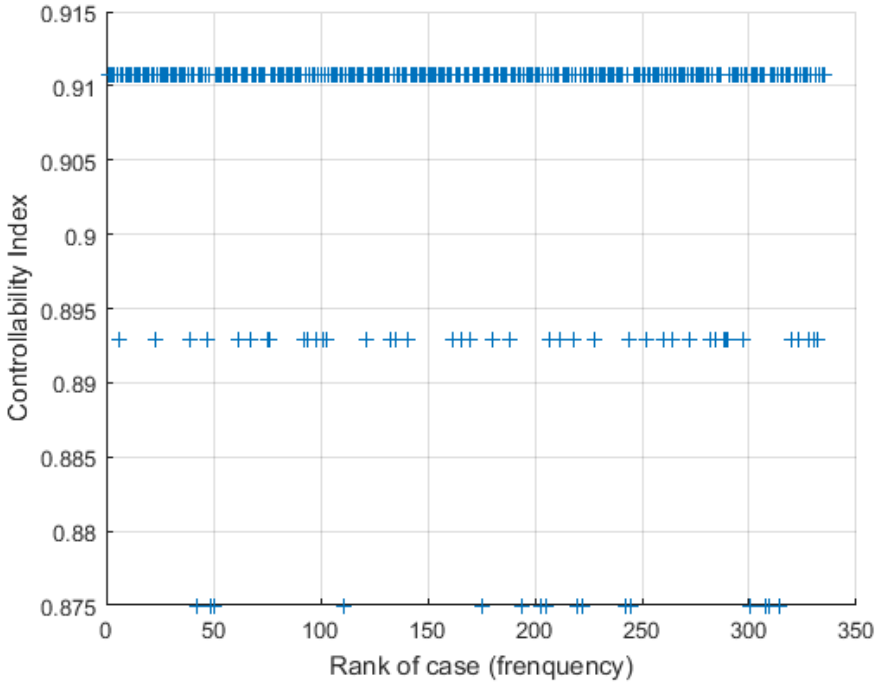


Figure 3.6: Controllability index for the 335 cases

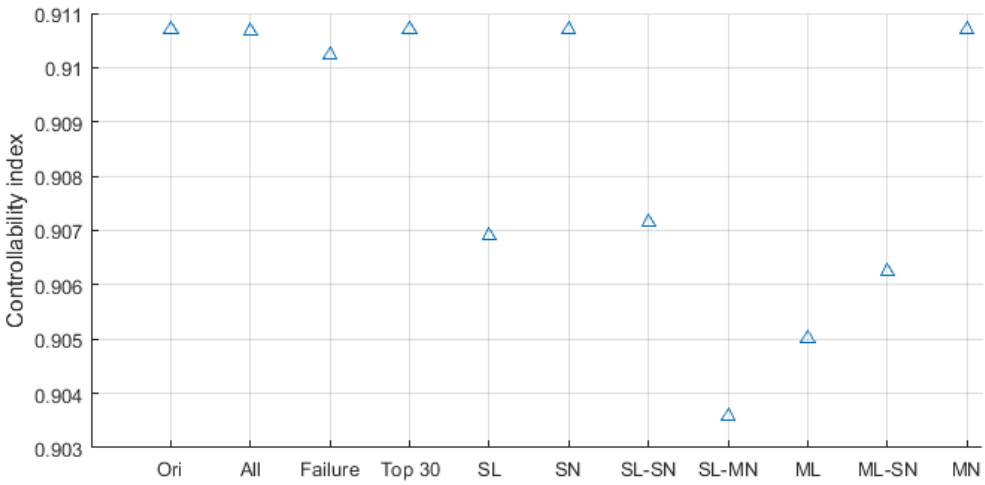


Figure 3.7: Controllability index for different categories

ML, ML-SN, MN). The mean of C_{ind} of the top 30 cases is slightly lower than but still close to that of the nominal network configuration.

Node failures have no impact on the controllability index, since they only concern the reduction of pipelines capacity but not their disconnection. Thus, link failures are the only contribution to the loss in controllability. The mean of all the cases containing

single link failures (i.e. SL, SL-SN and SL-MN) is 0.9069, slightly higher than that of the cases containing multiple link failures (i.e. ML and ML-SN) cases, which equals to 0.9051. This indicates that multiple failures have a more important impact on the controllability index, with C_{ind} reaching values no lower than 0.8750. For the SL, SL-SN and SL-MN cases, this lowest value results from the failures of links 6-35, 11-43 or 18-23, the removal of each of which decreases C_{ind} to 0.8750. As for ML and ML-SN cases, the lowest value comes from the combination of two link failures with no separate impact on C_{ind} (links 9-10 and 10-53), or the combination of two links whose removal decreases C_{ind} to 0.9829 (links 22-24 and 34-37).

Network topological efficiency

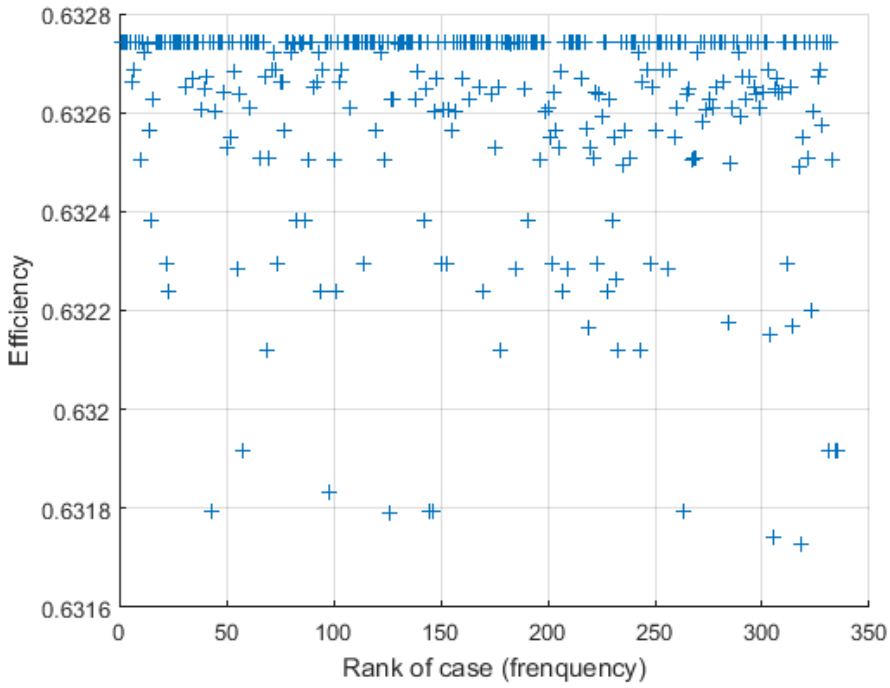


Figure 3.8: Network topological efficiency for the 335 cases

Figure 3.8 shows the values of E for the 335 cases, and the abscissa axis is the frequency rank of the 335 cases. The network topological efficiency ranges from 0.6317 to 0.6327. Failure of any link will influence the efficiency. All the seven values below 0.6318 are related to the failure of link 44-45 or link 18-34: the five first cases include single link

failures alone or together with single node failure, while the last two cases with low values are multiple link failures.

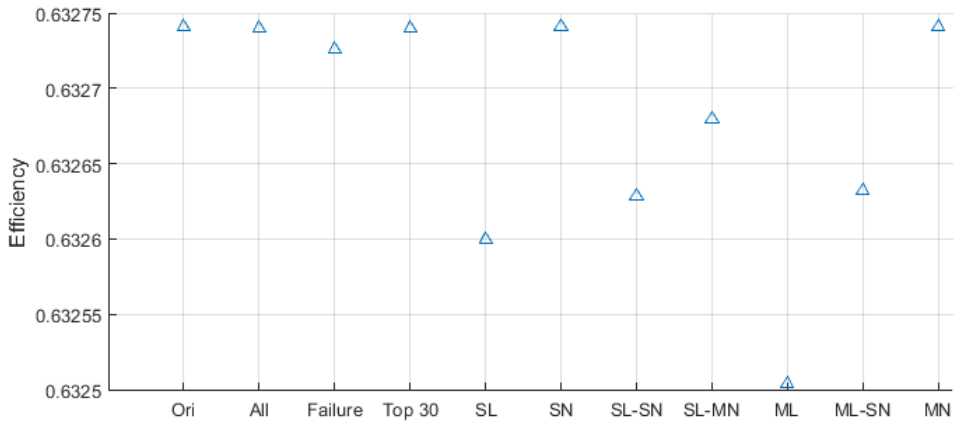


Figure 3.9: Network topological efficiency for different categories

Figure 3.9 shows the mean value of E of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN, ML, ML-SN, MN). Similar to the controllability index, node failures alone have no influence on topological efficiency and multiple link failures have a more important influence than single link failures. The mean of E over all cases of single link failure (i.e. SL, SL-SN and SL-MN) is 0.6326 and for the cases containing multiple link failures (i.e. ML and ML-SN) it is 0.6325.

Generally speaking, the variation of E is not significant, much less than the other two indexes. This is reasonable, because the network is not a sequential one and multiple paths exist between any two nodes: when one link fails, the gas flow can still be transferred via an alternative path, although of longer distance.

Protective actions

From the above analysis, we understand that node failures have significant importance in terms of supply, but do not affect other properties, that the link failures influence on NSD is less important than that of node failures and that the consequences of failures on

controllability and topological efficiency are limited compared to NSD. This means that, supply is the primary concern with respect to protection from failures, whereas network connections and control are not so sensitive and more fault tolerant.

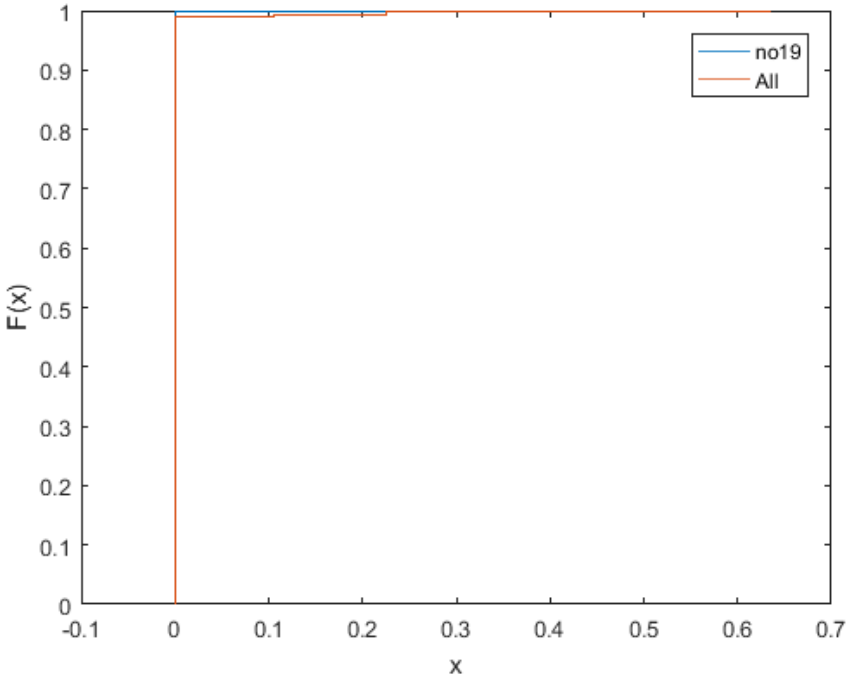


Figure 3.10: CDF of NSD for all configurations and for the configurations without failure of node 19

Node 19 is found to be an important node which may require protection priority, since it has a relatively high failure probability and at the same time has a significant impact on supply.

Figure 3.10 compares the cumulative distribution function of NSD over all simulated configurations and the 991570 configurations in which node 19 is not failed: the mean value of NSD of all configurations drops from 0.0020 to 0.0001, which is much lower.

3.3.4 Summary

In this chapter, we have presented an analysis framework considering three perspectives: supply service, controllability and topology. By performing such an analysis, we are able to identify the most critical elements within the network and quantify the consequences of

failure scenarios. The analysis framework has been applied to a gas transmission network. The results of the analysis show that the influence of a single link is not the same for the three perspectives and neither are the various failure scenarios. Supply turns out to be the most affected by failures, and can, thus, be used as the objective for the prioritization of investments for CI protection. The findings of this work show the importance of considering several perspectives of analysis for CIs.

Chapter 4

Multi-objective optimization of CIs

The design and operation of CIs are multi-objective problems, in which the multiple objectives account for different perspectives. In this chapter, we consider CIs for supply and investigate their optimization with respect to minimizing non-supplied demand and structural complexity, while at the same time maximizing their controllability. In order to understand the nature of the tradeoffs among the three objectives and extract useful information from the optimal solutions, a thorough analysis is performed to investigate the correlation among the three objectives and the impact of topological properties of the supply network (e.g. the average node degree). A benchmark case study of a real gas transmission network across several countries of the European Union (EU) is considered to illustrate the optimization framework.

Section 4.1 provides a brief literature review on the multi-objective optimization of CIs. Section 4.2 presents the multi-objective optimization framework. Section 4.3 presents the optimization results and analysis of the considered gas transmission network .

4.1 State of the art

4.1.1 Multi-objective optimization

Optimal design of a real world CI network typically involves simultaneously optimizing multiple objective functions and the considering several equality and/or inequality constraints.

Previous works on analysis and optimization of CIs typically focused on reliability and cost. In (Ramírez-Rosado and Bernal-Agustín, 2001), a multi-objective optimization methodology, using an evolutionary algorithm, is proposed to find out the best distribution network reliability while simultaneously minimizing the system expansion costs. The paper (Farmani et al., 2006) investigates the application of multi-objective evolutionary algorithms identifying the pay-off characteristic between total cost, reliability and water quality of Anytown's water distribution system. Reference (Üster and Dilaveroğlu, 2014) considers the problem of minimizing total investment and operating costs for designing a new natural gas transmission network or expanding an existing network. Reference (Baghaee et al., 2016) presents the multi-objective optimization for designing hybrid wind-solar generating microgrid systems under varying weather conditions, with the minimum cost and maximum reliability.

Due to the requirement of reliable operation and the complex nature of CIs, all-terminal reliability is often considered as a constraint or a second objective for the optimization problem. The reliability and cost of these systems are important and are largely determined by network topology. Reference (Deeter and Smith, 1997) presents a heuristic approach to design networks when considering all-terminal reliability, where the optimization problem is formulated as minimizing cost given a reliability constraint. In (Dengiz et al., 1997), a heuristic search algorithm inspired by evolutionary methods is presented to solve the all-terminal network design problem when considering cost and reliability. Reference (Ramirez-Marquez and Rocco, 2008) considers the minimization of the network design cost subject to a known constraint on all-terminal reliability by as-

suming that the network contains a known number of functionally equivalent components to provide redundancy.

Selecting of optimal network topology is an NP-hard combinatorial problem where classical enumeration-based methods grow exponentially with network size (Dengiz et al., 1997). Various approaches have been proposed to solve the multi-objective optimization problem considering cost and supply, such as heuristics (Dengiz et al., 1997; Ramírez-Rosado and Bernal-Agustín, 2001), neural networks (NN) (Papadrakakis and Lagaros, 2002) and particle swarm optimization algorithm (Baghaee et al., 2016).

4.1.2 Evolutionary algorithms for network optimization

Non-dominated sorting genetic algorithm II (NSGA-II) (Deb et al., 2002) is a computationally fast and elitist multi-objective evolutionary algorithm (MOEA) based on a non-dominated sorting approach. NSGA-II has been shown to outperform other optimization algorithms in terms of the spread of solutions and efficiency of convergence near the true Pareto-optimal front (Deb et al., 2002). Various applications to CIs can be found in literature. In (Farmani et al., 2006), NSGA-II is applied to search a set of optimal solutions for the rehabilitation of water distribution systems with objectives of total cost, reliability and water quality. In (Wu et al., 2009), this algorithm is adopted for water distribution systems to solve a multi-objective optimization problem of their design considering greenhouse gas emission. In (Li et al., 2013), it has been used to improve the modified binary differential evolution (MBDE) algorithm and the resulting non-dominated sorting binary differential evolution algorithm (NSBDE) has, then, been used to solve the multi-objective optimization of cascading failures protection in complex networks, cascade-resilient electrical infrastructures design (Fang et al., 2015), etc. Reference (Hu et al., 2016) investigates the combined gas and electricity network planning, with the objectives to minimize both investment cost and production cost of the combined system while taking into account network security criterion. Reference (Kamjoo et al., 2016) considers the multi-objective design of a hybrid renewable energy system under

uncertainty using NSGA-II.

The NSGA-II proceeds as follows (Deb et al., 2002):

1. *Generation of initial population*

   Initialize the population P_0 of N_P candidate solutions.

2. *Objective function evaluation*

   Evaluate each of the N_P solutions in the population P_0, by calculating the three objective functions presented above.

3. *Generation of an offspring population*

   Apply the binary tournament selection operator to the population P_0 to create an offspring population Q_0 of size N_P, which undergoes the evolution operations of mutation and crossover.

4. *Evaluation*

   Evaluate each of the N_P solutions in the population Q_0.

5. *Union and ranking: non-dominated sorting*

   (For the tth generation) Combine P_t and Q_t to form a union population $R_t = P_t \cup Q_t$ of size $2N_P$. Then, the population R_t is sorted by the fast non-dominated sorting and the ranked non-dominated fronts F_1, F_2,...,F_k are identified (F_1 is the best, F_2 is the second best, and so on).

6. *Comparison and selection*

   To select the first N_P members of R_t for the new population P_{t+1}, the crowded-comparison operator (Deb et al., 2002) is used: between two solutions, the one with better rank is preferred; otherwise, if they belong to the same front, the solution with larger crowding distance has higher priority, where crowding means the density of solutions surrounding a solution of specified radius.

7. *Iteration*

   Increase the generation number t by 1, and the algorithm stops when it reaches the

maximum number of generations N_G.

4.2 Optimization for CI design

In the present work, we propose an optimization framework for the design (i.e. allocation of links and capacities) of the complex network that make up a CI, with the objective of minimizing the non-supplied demand (NSD) and structural complexity, and maximizing controllability (i.e minimizing the number of driver nodes). In previous works, the optimization typically only considered NSD and cost. We include the control perspective by introducing of an objective function that minimizes the number of driver nodes, which implies the minimization of the cost related to achieving full control for the network.

4.2.1 Three system-level indexes considered as optimization objective

Non-supplied demand

We use the non-supplied demand (NSD) as a measure of the network's capacity to satisfy its users' demand, as defined in equation 3.2 in Section 3.2. The objective function is to minimize NSD.

Structural complexity

System complexity is another important property of CIs. In this study, we adopt the structural complexity metric introduced in (Sinha and de Weck, 2013b). This metric, hereafter denoted C, accounts for the complexities of the individual components C_1, the complexities linked to the connections among the components C_2 and the topological complexity of the system structure C_3:

$$\begin{aligned} C &= C_1 + C_2 C_3 \\ &= \sum_{i=1}^{N} \alpha_i + \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} \mathbf{Adj}_{ij} \right] \gamma E(\mathbf{Adj}) \end{aligned} \tag{4.1}$$

where α_i is the complexity of the ith component, β_{ij} is the interface complexity between the ith and j^{th} components, **Adj** is the adjacency matrix defining the connections among the components, γ is a normalization factor and matrix energy of the network$E(\mathbf{Adj})$, used to measure the topological complexity, is defined as the sum of the singular values of its adjacency matrix **Adj**, thus accounting for the topological complexity of the system structure.

Empirical and experimental evidence reveals that the system development cost grows with the structural complexity, suggesting that a low structural complexity for low development cost is preferred, if the design satisfies all of the other constraints (Sinha and de Weck, 2013a).

Controllability

We propose to use the minimum number of driver nodes N_D to indicate the controllability of the network and aim at minimizing N_D for a less costly network in terms of control.

4.2.2 Multi-objective optimization problem formulation

In this study, we seek to: (1) minimize the non-supplied demand (NSD) so as to satisfy the users' demands as much as possible; (2) minimize the structure complexity (C), which relates to the cost of the system design and the development efforts; (3) minimize the number of driver nodes (N_D) to make the network as controllable as possible. The three objectives are calculated by equations (3.2), (4.1) and (3.1), respectively. The formulation of the optimization problem with regards to the three objectives is, then, simply given as follows:

$$\begin{cases} \min\ NSD(\mathbf{X}) \\ \min\ C(\mathbf{X}) \\ \min\ N_D(\mathbf{X}) \end{cases} \tag{4.2}$$

The decision variable vector $\mathbf{X}$ contains the capacity of the links connecting node i and node j, X_{ij}, $i, j = 1, ...N$, $i \neq j$. The element X_{ij} is equal to 0 if the nodes i and j are

not connected; otherwise, it indicates the capacity of the link between the two nodes.

The optimization is done by considering the connection patterns among the nodes and allocating capacities to the connecting links. The variables to be optimized are, then, the link capacities $X_{ij} \in [0, x_{max}]$, which can take any value inferior to the limit capacity x_{max}. The adjacency matrix **Adj** derived from X_{ij} is used to calculate C and N_D as defined in Equations 4.1 and 3.1.

The constraints are: (1) the nodes remain unchanged, including their quantity, location and functionalities; (2) the users' demands and supply capacities of the sources stay the same.

The final goal of the optimization is to identify a set of solutions, in which no solution can be regarded better than another with respect to all objective functions. This can be achieved by adopting the concepts of Pareto optimality and dominance (Sawaragi et al., 1985). We propose to use the non-dominated sorting genetic algorithm II (NSGA-II) to search the solutions of the Pareto-optimal set, given its proven performance. Details can be found in Paper [III].

4.3 Case study

We consider the same gas transmission network as described in Chapter 3.3.

Note that, since the number of nodes and and their functionality remain the same (and so does the component complexities), the term C_1 in equation (4.1) can be neglected. As there exists only one type of connection between any two nodes, i.e. gas flow, the interface complexity β_{ij} is assumed to be the same for all pipelines and it is set to 0.5. However, this is a simplification of the reality, and it can be estimated differently considering the effect from distance, pressure, etc. γ is arbitrarily set to $1/N$.

NSGA-II is applied to solve the previously defined multi-objective optimization problem. The parameters of the algorithm are set as in Table 4.1, based on the results of a number of test runs.

Table 4.2 presents the three index values of the original gas transmission network.

Parameters	Values
Population size N_P	40
Crossover rate CR	0.9
Scaling factor F	0.2
Maximum generation N_G	300

Table 4.1: Parameters of the NSGA-II algorithm

NSD	N_D	C
0	4	196.2

Table 4.2: Index values for the original network

We run the NSGA-II algorithm 10 times and consider all the different Pareto-optimal solutions found during the ten runs, in order to investigate the relationship among the objectives and extract useful information for the selection of the CI network configuration. Then, we select the overall six non-dominated solutions to construct the Pareto front.

4.3.1 Correlations among the three objectives

We choose the networks whose N_D is between 1 and 6 and divide them into six classes accordingly. We, then, calculate the mean of NSD and C for each class.

Studies in (Li et al., 2016; Liu et al., 2011; Pósfai et al., 2012; Wang et al., 2012) show that there exists a correlation between the average node degree and the controllability of networks. We also calculate the mean average node degree for each class. In our study, we find that the average node degree $< k >$ of the generated networks ranges from 2.4 to 3.1. This can represent most cases of the real-world complex networks, which are typically sparse (Li et al., 2016). Indeed, the degree of a node is the number of edges connected to the node and the average node degree $< k >$ represents the link density of the network. When adding links to the network (which means that $< k >$ increases), the structural complexity increases, as the complexity metric C takes into account both the quantity

of links and the topological complexity (matrix energy) does not reduce. Therefore, it is intuitive to understand that the higher $< k >$, the higher is its structural complexity. The Pearson correlation coefficient between the mean of $< k >$ and C is 0.9987.

Figure 4.1 shows the mean of the average node degree for each class of networks obtained: as the number of driver nodes increases, the average node degree decreases. For sparse networks, N_D is determined by the rank of $\mathbf{A}$ and drivers nodes correspond to linearly dependent rows (Yuan et al., 2013). Adding links to the network is possible to eliminate the linearly dependent rows in $\mathbf{A}$, which explains the negative correlation between N_D and average node degree $< k >$. The correlation coefficient between them is -0.9929.

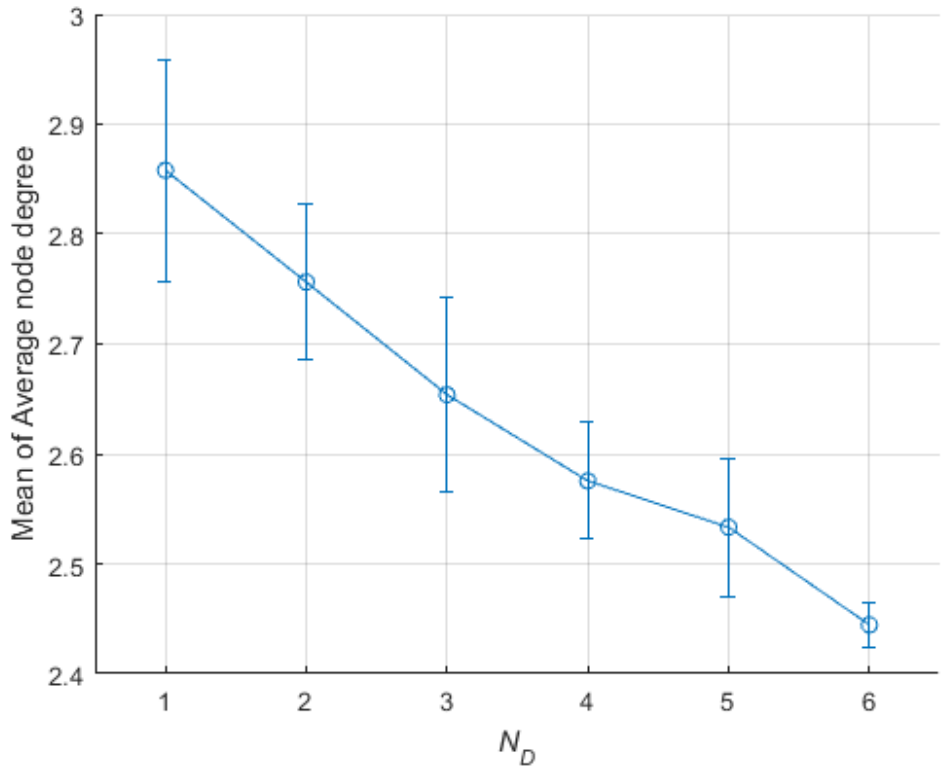


Figure 4.1: Mean of average node degree versus number of driver nodes N_D

As for the non-supplied demand NSD, the correlation coefficient is -0.9139. Adding links (which increases the average node degree) may increase the supply capacity of a network and, thus, NSD decreases. However, this is not guaranteed since the supply capacity between two nodes can be limited by a small capacity link on the path. The

supply depends more on the capacity of transmission rather than on the number of links within the network. This explains why NSD correlates with $< k >$ less than the other two properties.

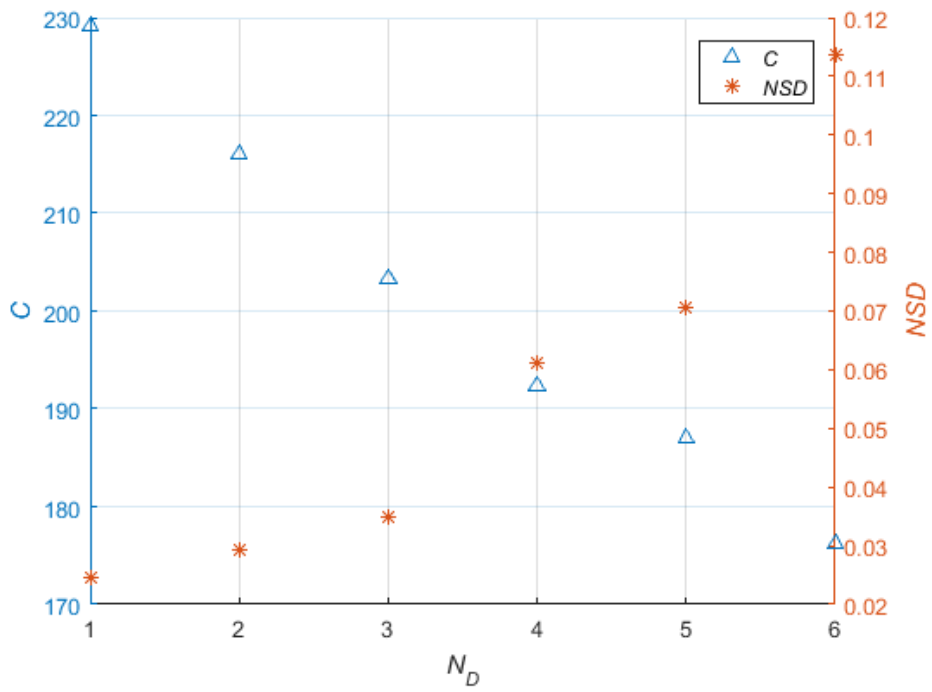


Figure 4.2: Mean of the three objectives

The mean values of the three objectives for the six classes are shown in Figure 4.2. We observe that, as the number of driver nodes N_D increases, the structural complexity C decreases and the NSD increases.

The number of driver nodes N_D and the structural complexity C are highly correlated. We can see from Figure 4.3 that as N_D increases, C increases, while the standard deviation decreases. This can be explained by the fact that these two objectives are affected by the average node degree in opposite directions.

Figure 4.4 shows the correlation between NSD and N_D. We can observe that NSD and N_D have relatively weak correlation, and large standard deviation. The supply depends more on the capacity of transmission rather than on the number of links within the network and these two objectives are more independent.

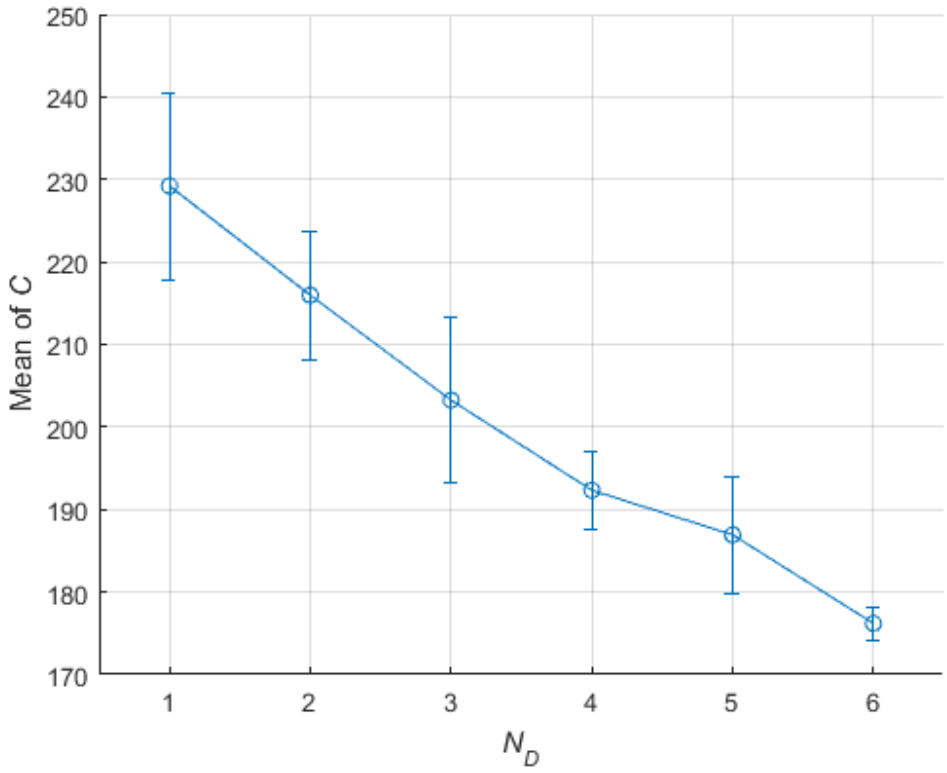


Figure 4.3: Mean of structural complexity C versus N_D

Remarks

To sum up the previous analysis, we find that structural complexity and controllability are influenced by the topological property of the network system (average node degree) in opposite directions: for a sparse complex network, such as is the gas transmission network considered, a relatively dense one is preferable for the consideration of controllability, but it is a more structurally complex network, and therefore it comes at a cost; if we seek to choose a less complex network configuration, we tend to have a less controllable network, i.e. that requires more efforts to keep the system under control; yet, the topology has less impact on the supply to the users. And for the purpose of demand satisfaction, link capacity is a more important factor to consider. For selection among the Pareto-optimal solutions, it would be interesting to choose those with $NSD = 0$ and, then, to seek the balance between complexity and controllability.

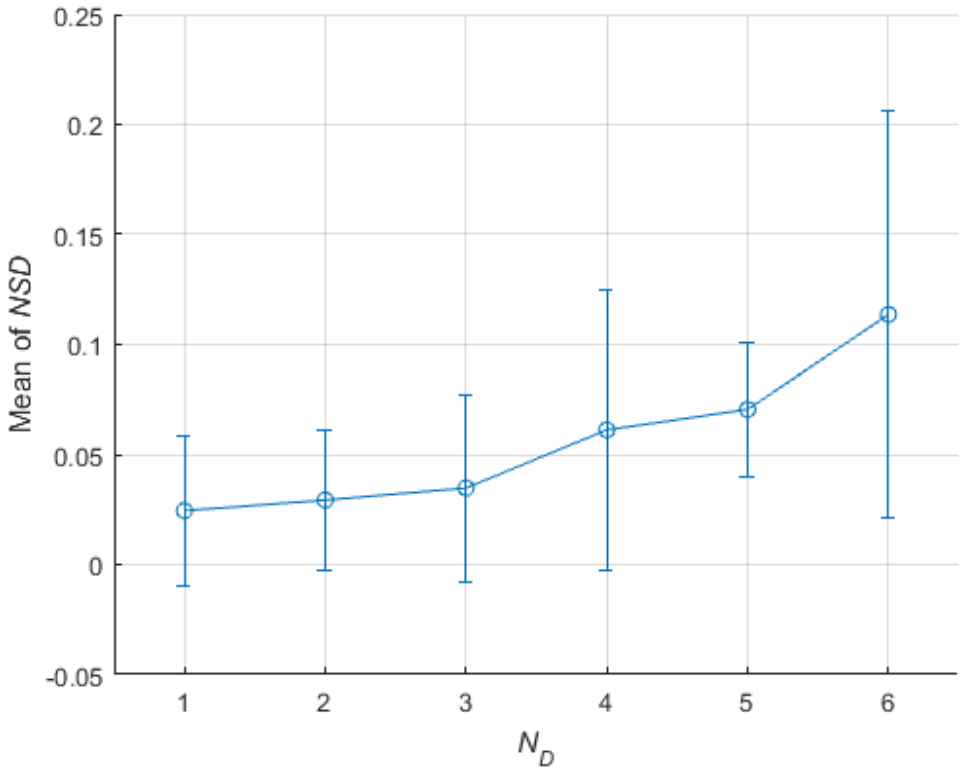


Figure 4.4: Mean of NSD versus N_D

4.3.2 Optimization results

Table 4.3 presents the values of the three objective functions for each solution identified as Pareto-optimal by the algorithm. The Pareto front obtained by the NSGA-II is illustrated in Figure 4.5.

The original real network (indicated by the red triangle) is close to the solution 1 with $N_D = 4$, $NSD = 0$. In fact, the two networks coincide in the plane N_D, NSD of Figure 4.5(d). The difference is that the optimal solution 1 has lower structural complexity than the original network, obtained by removing the link connecting nodes 44 and 55. Removing certain links could be an improvement to the original network, for the purpose of minimizing the complexity, for instance. The optimal solutions found may provide other possibilities of network configurations and capacity allocations.

Among the six solutions of Table 4.3, N_D varies from 1 to 6 and the complexity metric C varies from 178.3 to 232.3. Solution 2 has the lowest C and largest N_D. Solution 5 has

No.	NSD	N_D	C
1	0	4	192.3
2	0.024	6	178.3
3	0	2	222.4
4	0	3	216.0
5	0	1	232.3
6	0.102	2	219.4

Table 4.3: Objective functions values for the Pareto-optimal solutions

the smallest number of driver nodes and largest C values. As for NSD, it takes value between 0 and 0.102. The Solutions 2 and 6 include separate nodes, and the fact that they are not able to be supplied leads to $NSD > 0$. Four out of six solutions are able to fully satisfy all users' demands.

Let us focus on these latter four solutions, capable of fully satisfying the demands (i.e. solutions 1, 3, 4 and 5, with $NSD = 0$). We can see that, as the number of driver nodes decreases, the structural complexity increases. We can define a rate of change to choose the most efficient optimal solution network in terms of corresponding objectives. If controllability is the primary concern, for example, we define the ratio of the changes in the number of driver nodes and in the structural complexity $\Delta N_D/\Delta C$: the larger this ratio, the more preferable in terms of gain the network is. Then, solution 1 is the best, and the original network also has a rather satisfying configuration.

We, then, compare the two solutions with $N_D = 2$ (solutions 3 and 6). Solution 3 has $NSD = 0$ and is, thus, better in terms of supply performance than solution 6, which has the highest non-supplied demand $NSD = 0.102$. On the other hand, the difference in terms of structural complexity is not significant. This indicates that, in this case, we do not necessarily need a particularly complex structure to satisfy the demand.

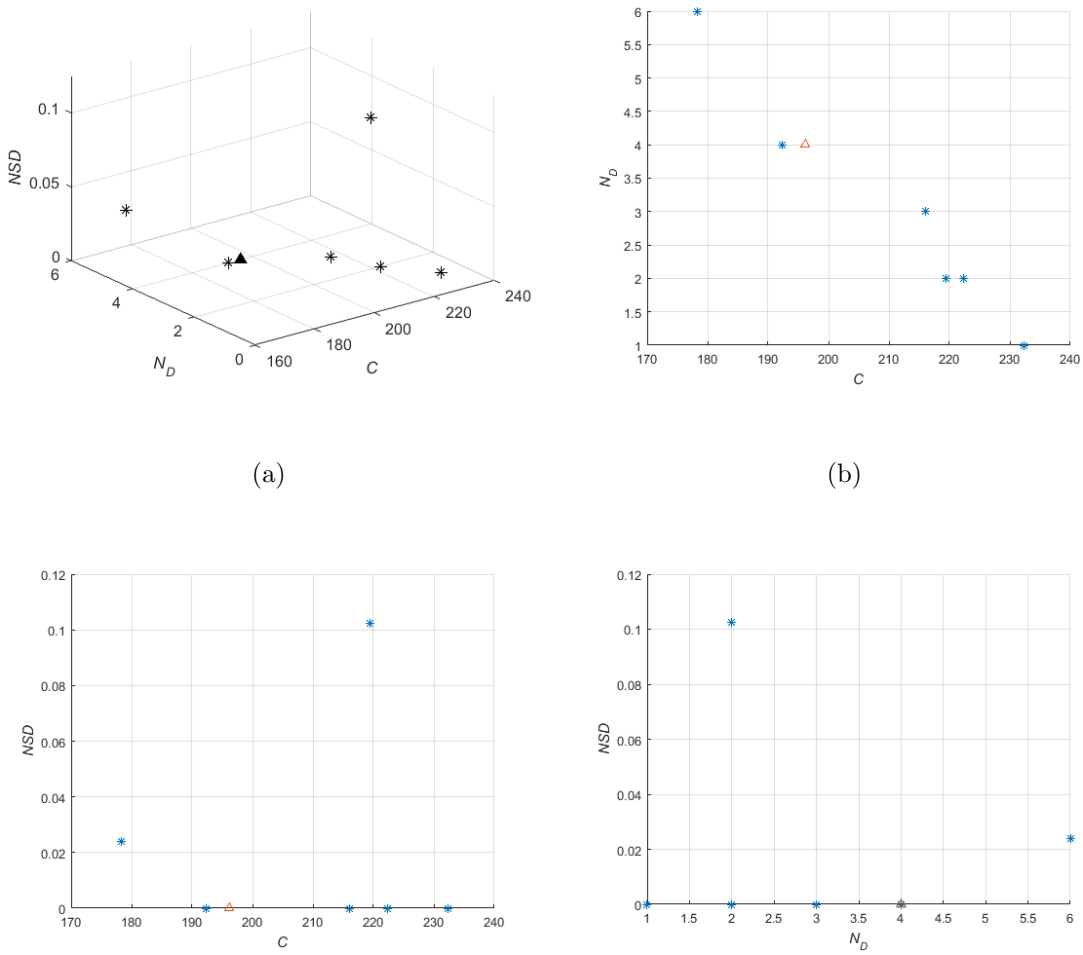


Figure 4.5: Pareto front in 3-D space (a) and 2-D projections (b)-(d)

4.3.3 Analysis of node importance

Different components may have different contributions to the normalized non-supplied demand (NSD), the structural complexity (C) and controllability (N_D). In this section, we investigate whether some nodes may be important independently of the network configuration. We also analyze and compare the average node importance of each solution network, in order to understand if certain network configurations or certain properties are more sensitive to the removal of single nodes.

To quantify the node importance of each solution network with respect to NSD and

N_D, We compute RA, as introduced in Chapter 3.3.2:

$$NI_i^{NSD} = NSD[\mathbf{G}] - NSD[\mathbf{G}'(x_i = 1)] \tag{4.3}$$

$$NI_i^{N_D} = \frac{N_D[\mathbf{G}'(x_i = 1)] - N_D[\mathbf{G}]}{N} \tag{4.4}$$

where $\mathbf{G}'(x_i = 1)$ is the graph obtained by removing from the solution network $\mathbf{G}$ the node i (i.e. all edges incident in node i).

As for the node importance with respect to the structural complexity, NI_i^C, it is defined as the relative drop in C caused by the deactivation of node i from the solution network $\mathbf{G}$:

$$NI_i^C = \frac{\Delta C[\mathbf{G}]}{C[\mathbf{G}]} = \frac{C[\mathbf{G}] - C[\mathbf{G}'(\mathbf{x_i} = \mathbf{1})]}{C[\mathbf{G}]} \tag{4.5}$$

The average node importances across all nodes are shown in Table 4.4.

No.	NI^{NSD}	$NI^{C_{ind}}$	NI^C
1	0.0550	0.0249	0.0619
2	0.0288	0.0203	0.0614
3	0.0279	0.0219	0.0615
4	0.0253	0.0237	0.0618
5	0.0408	0.0252	0.0616
6	0.0228	0.0225	0.0612

Table 4.4: Average node importance values for the optimal solutions

Solution 1 and the original network give the same result for the node importance analysis: they both have the highest average node importance with respect to NSD, NI^{NSD}, and are most sensitive to node failure in terms of supply, which indicates that every link within this network is relatively critical. This means cost-effective design but good protection is required during operation to provide reliable service. Solution 5 also has a relatively high NI^{NSD}, in fact: with the removal of source node 19, NSD reaches a value of 0.48, which is the highest value among all optimal network configurations. Solution 6 has a relatively low average node importance: however, it is not an interesting

solution, since it can not satisfy all the users' demand. Solutions 3 and 4 are less sensitive to the removal of single nodes in terms of supply. We note that nodes 2 and 19 always have an important impact on NSD for any network configuration, because of their role as (large) sources. Node 18 is also a very important node for supply in all six solution networks, because it has a large demand and for most of the solutions, it connects the source node 19 to other demand nodes: thus, its removal usually has a more important impact on supply than source node 19.

The average node importance values with respect to structural complexity NI^C turn out to be close for all Pareto-optimal networks. The solutions are all sparse networks and the nodes have relatively low degree; thus, the removal of single nodes does not influence significantly the structural complexity. We also observe that node 18 is important for all the six solutions in terms of structural complexity, as it connects a relatively large number of nodes and has a high node degree.

The removal of single nodes can at most increase the number of driver nodes N_D of 2 units. Solution 5 with $N_D = 1$ has the largest average node importance with respect to N_D and it is the most sensitive to node failure in terms of N_D. In contrast, Solution 2 with $N_D = 6$, is the least sensitive one.

Solutions 3 and 4 could be considered as reasonable alternatives to the original network configuration, since they are most resistant to node failures.

4.3.4 Summary

In this chapter, we have considered three objectives for optimal complex supply networks design: minimizing the non-supplied demand and the structural complexity, and maximizing the controllability (i.e. minimizing the number of driver nodes). We have proposed to use the non-dominated sorting genetic algorithm II (NSGA-II) to tackle the multi-objective optimization problem. A gas transmission network has been taken as a reference case study. A comparative evaluation has been performed to analyze the optimal solutions, with respect to how the allocation of link capacities can improve the desired

properties. At last, an investigation of the impact of topological properties (i.e. the average node degree) on the three objectives and the correlation among them has been performed to draw insights on the tradeoffs among the three objectives.

In conclusion, this work develops a feasible optimization framework for designing critical supply infrastructures, taking into account different desired properties of the system.

Chapter 5

Conclusions and future research

5.1 Conclusions

The present thesis focuses on the integration of the control perspective into the safety and reliability analysis of critical infrastructures (CIs). We have developed frameworks for multiple-perspective modeling, analysis, and optimization of CIs. Three perspectives are considered, including supply, topology and controllability.

Firstly, a simulation-based scheme was developed for the modeling and analysis of small-scale CIs. The developed simulation-based scheme integrates Control Theory with CI modeling and allows analysis from different perspectives, including supply service, controllability and topology. An optimization-based model predictive control framework is proposed to analyze the CI performance under various failure scenarios. An application on an electrical power microgrid shows that the developed methods are able to integrate control perspectives in CI modeling and analysis and allow multi-perspective analysis.

Then, a multi-perspective framework was developed to analyze CIs of larger scale with respect to supply service, controllability and topology. A risk achievement worth-based method is developed to analyze the importance of different elements in the CI, from different perspectives. The analysis framework has been applied to a benchmark network of a real gas transmission network across several countries of the European Union (EU). The results of the application show that the developed methods are able to identify the

most critical elements within the network and quantify the consequences of scenarios of multiple failures, with respect to the different perspectives considered.

Finally, the multi-perspective analysis framework enables the optimization of CIs with respect to different properties of the system. A three-objective optimization framework for complex CIs design was developed: design of network topology and allocation of link capacities are performed with the objectives of minimizing the non-supplied demand and the structural complexity of the system, while at the same time maximizing system controllability. NSGA-II was used to identify the optimal solutions of the developed model. A comparative evaluation has been performed to analyze the optimal solutions, with respect to how the allocation of link capacities can improve the desired properties. We have also performed an investigation of the correlation among the multiple objectives considered to draw some useful insights for system design.

In conclusion, the findings of this work demonstrate the feasibility and the importance of the developed frameworks for the analysis and optimization of CIs, taking into account different perspectives relevant for CIs design, operation and protection.

5.2 Prospective works

Different research directions can be considered to extend the work.

Firstly, from the component-level, in the developed framework, we only consider binary-state failures. Only with such simplifications, the presented framework can be able to identify the critical elements in terms of different perspectives considered. In future work, a multi-state model will be considered to capture the multi-state nature of failure and repair behaviors.

Secondly, from the system-level, the developed frame only considered independent component failures. Cascading effects and targeted attacks will be considered in future works, since they are quite common in practice and can provide more information for understanding the correlation among the different properties.

Lastly, this work focuses on single-CI networks; however, as CIs are more and more

interconnected and automated, interdependencies and multi-CIs modeling become of great interest. A relatively small failure could lead to a catastrophic breakdown of the interconnected system. Over the past decade, there have been advances in the field of interdependent networks; however, understanding and quantifying the effects of interdependencies among various types or real-life engineered infrastructure systems in their response to systemic risks still remain challenge for CI protection. The framework will be expanded for interconnected networks in the future works.

References

Albert, R., Albert, I., and Nakarado, G. L. (2004). Structural vulnerability of the north american power grid. *Physical review E*, 69(2):025103.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *nature*, 406(6794):378–382.

Bae, I.-S. and Kim, J.-O. (2007). Reliability evaluation of distributed generation based on operation mode. *IEEE Transactions on Power Systems*, 22(2):785–790.

Baghaee, H., Mirsalim, M., Gharehpetian, G., and Talebi, H. (2016). Reliability/cost-based multi-objective pareto optimal design of stand-alone wind/pv/fc generation microgrid system. *Energy*, 115:1022–1041.

Bakolas, E. and Saleh, J. H. (2011). Augmenting defense-in-depth with the concepts of observability and diagnosability from control theory and discrete event systems. *Reliability Engineering & System Safety*, 96(1):184–193.

Blaabjerg, F., Teodorescu, R., Liserre, M., and Timbus, A. V. (2006). Overview of control and grid synchronization for distributed power generation systems. *IEEE Transactions on industrial electronics*, 53(5):1398–1409.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308.

Clinton, W. (1998). Presidential Decision Directive PDD-63, Protecting America's Critical Infrastructure. *The White House, Washington, DC.*

Colson, C., Nehrir, M., and Gunderson, R. (2010a). Multi-agent microgrid power management. In *Proceedings of the 18th IFAC World Congress*, pages 77–82.

Colson, C., Nehrir, M., and Pourmousavi, S. (2010b). Towards real-time microgrid power management using computational intelligence methods. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–8. IEEE.

Colson, C., Nehrir, M., and Wang, C. (2009). Ant colony optimization for microgrid multi-objective power management. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–7. IEEE.

Conti, S., Nicolosi, R., Rizzo, S., and Zeineldin, H. (2012). Optimal dispatching of distributed generators and storage systems for mv islanded microgrids. *IEEE Transactions on Power Delivery*, 27(3):1243–1251.

Costa, P. M. and Matos, M. A. (2005). Reliability of distribution networks with microgrids. In *Power Tech, 2005 IEEE Russia*, pages 1–7. IEEE.

Cowan, N. J., Chastain, E. J., Vilhena, D. A., Freudenberg, J. S., and Bergstrom, C. T. (2012). Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PloS one*, 7(6):e38398.

Crucitti, P., Latora, V., and Marchiori, M. (2004). A topological analysis of the Italian electric power grid. *Physica A: Statistical Mechanics and its Applications*, 338(1):92–97.

Crucitti, P., Latora, V., and Porta, S. (2006). Centrality measures in spatial networks of urban streets. *Physical Review E*, 73(3):036125.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197.

Deeter, D. L. and Smith, A. E. (1997). Heuristic optimization of network design considering all-terminal reliability. In *Reliability and Maintainability Symposium. 1997 Proceedings, Annual*, pages 194–199. IEEE.

Dengiz, B., Altiparmak, F., and Smith, A. E. (1997). Efficient optimization of all-terminal reliable networks, using an evolutionary approach. *IEEE transactions on Reliability*, 46(1):18–26.

Deo, N. (2016). *Graph theory with applications to engineering and computer science.* Courier Dover Publications.

Dimeas, A. L. and Hatziargyriou, N. D. (2005). Operation of a multiagent system for microgrid control. *IEEE Transactions on Power systems*, 20(3):1447–1455.

Directive, C. (2008). 114/ec of 08 december 2008 on the identification and designation of european critical infrastructures and the assessment of the need to improve their protection. *Official Journal of the European Union, L*, 345.

Edlund, K., Bendtsen, J., and Jørgensen, J. (2011). Hierarchical model-based predictive control of a power plant portfolio. *Control Engineering Practice*, 19(10):1126–1136.

EGIG (2011). 8th report of the european gas pipeline incident data group.

EU (2010). Regulation no. 994/2010 of the european parliament and of the council of 20 october 2010 concerning measures to safeguard security of gas supply and repealing council directive 2004/67/ec. *Official Journal of the European Union.*

EU (2014). European Energy Security Strategy. *Communication from the Commission to the European Parliament and the Council. COM (2014) 330 final. Brussels, 2014.*

Eusgeld, I., Kröger, W., Sansavini, G., Schläpfer, M., and Zio, E. (2009). The role of network theory and object-oriented modeling within a framework for the vulnerability analysis of critical infrastructures. *Reliability Engineering & System Safety*, 94(5):954–963.

Fang, Y., Pedroni, N., and Zio, E. (2015). Optimization of cascade-resilient electrical infrastructures and its validation by power flow modeling. *Risk Analysis*, 35(4):594–607.

Fang, Y.-P. and Zio, E. (2013a). Hierarchical modeling by recursive unsupervised spectral clustering and network extended importance measures to analyze the reliability characteristics of complex network systems. *American Journal of Operations Research*, 3(1A):101–112.

Fang, Y.-P. and Zio, E. (2013b). Unsupervised spectral clustering for hierarchical modelling and criticality analysis of complex networks. *Reliability Engineering & System Safety*, 116:64–74.

Farmani, R., Walters, G., and Savic, D. (2006). Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality. *Journal of Hydroinformatics*, 8(3):165–179.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.

Gabbar, H. A., Islam, R., Isham, M. U., and Trivedi, V. (2012). Risk-based performance analysis of microgrid topology with distributed energy generation. *International Journal of Electrical Power & Energy Systems*, 43(1):1363–1375.

Gao, F. and Iravani, M. R. (2008). A control strategy for a distributed generation unit in grid-connected and autonomous modes of operation. *IEEE Transactions on Power Delivery*, 23(2):850–859.

Goodwin, G., Seron, M., and De Dona, J. (2005). *Constrained control and estimation: an optimisation approach*. Springer Verlag.

Grigg, C., Wong, P., Albrecht, P., Allan, R., Bhavaraju, M., Billinton, R., Chen, Q., Fong, C., Haddad, S., and Kuruganty, S. (1999). The ieee reliability test system-1996. a

report prepared by the reliability test system task force of the application of probability methods subcommittee. *Power Systems, IEEE Transactions on*, 14(3):1010–1020.

Guerrero, J. M., Vasquez, J. C., Matas, J., De Vicuña, L. G., and Castilla, M. (2011). Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. *IEEE Transactions on Industrial Electronics*, 58(1):158–172.

Halvgaard, R., Poulsen, N. K., Madsen, H., and Jorgensen, J. (2012). Economic model predictive control for building climate control in a smart grid. In *IEEE PES Innovative Smart Grid Technologies (ISGT)*, pages 1–6. IEEE.

Hamilton, M. C., Lambert, J. H., Connelly, E. B., and Barker, K. (2016). Resilience analytics with disruption of preferences and lifecycle cost analysis for energy microgrids. *Reliability Engineering & System Safety*, 150:11–21.

Han, F., Prodan, I., and Zio, E. (2015). A framework of model predictive control for the safety analysis of an electric power microgrid. In *ESREL 2015, 25th European Safety and Reliability Conference.*

Helseth, A. and Holen, A. T. (2006). Reliability modeling of gas and electric power distribution systems; similarities and differences. In *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on*, pages 1–5. IEEE.

Hernandez-Aramburo, C. A., Green, T. C., and Mugniot, N. (2005). Fuel consumption minimization of a microgrid. *IEEE transactions on industry applications*, 41(3):673–681.

Holmgren, Å. J. (2006). Using graph models to analyze the vulnerability of electric power networks. *Risk analysis*, 26(4):955–969.

Hooshmand, A., Poursaeidi, M. H., Mohammadpour, J., Malki, H. A., and Grigoriads, K. (2012). Stochastic model predictive control method for microgrid management. In *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, pages 1–7. IEEE.

Hovgaard, T., Larsen, L., and Jorgensen, J. (2011). Robust economic mpc for a power management scenario with uncertainties. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 1515–1520. IEEE.

Hu, Y., Bie, Z., Ding, T., and Lin, Y. (2016). An nsga-ii based multi-objective optimization for combined gas and electricity network expansion planning. *Applied Energy*, 167:280–293.

Islam, M. R. and Gabbar, H. A. (2012). Analysis of microgrid protection strategies. In *Smart Grid Engineering (SGE), 2012 IEEE International Conference on*, pages 1–6. IEEE.

Jayawarna, N., Jenkins, N., Barnes, M., Lorentzou, M., Papthanassiou, S., and Hatziagyriou, N. (2005). Safety analysis of a microgrid. In *Future Power Systems, 2005 International Conference on*, pages 7–pp. IEEE.

Jia, T. and Barabási, A.-L. (2013). Control capacity and a random sampling method in exploring controllability of complex networks. *Scientific reports*, 3.

Jia, T., Liu, Y.-Y., Csóka, E., Pósfai, M., Slotine, J.-J., and Barabási, A.-L. (2013). Emergence of bimodality in controlling complex networks. *Nature communications*, 4.

Jiayi, H., Chuanwen, J., and Rong, X. (2008). A review on distributed energy resources and microgrid. *Renewable and Sustainable Energy Reviews*, 12(9):2472–2483.

Jimeno, J., Anduaga, J., Oyarzabal, J., and de Muro, A. (2011). Architecture of a microgrid energy management system. *European Transactions on Electrical Power*, 21(2):1142–1158.

Johansson, J., Hassel, H., and Zio, E. (2013). Reliability and vulnerability analyses of critical infrastructures: comparing two approaches in the context of power systems. *Reliability Engineering & System Safety*, 120:27–38.

Justus, C., Hargraves, W., and Yalcin, A. (1976). Nationwide assessment of potential output from wind-powered generators. *Journal of Applied Meteorology*, 15(7):673–678.

Kalman, R. (1959). On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):110–110.

Kalman, R. E. (1963). Mathematical description of linear dynamical systems. *Journal of the Society for Industrial & Applied Mathematics, Series A: Control*, 1(2):152–192.

Kamjoo, A., Maheri, A., Dizqah, A. M., and Putrus, G. A. (2016). Multi-objective design under uncertainties of hybrid renewable energy system using nsga-ii and chance constrained programming. *International Journal of Electrical Power & Energy Systems*, 74:187–194.

Karimi, H., Nikkhajoei, H., and Iravani, R. (2008). Control of an electronically-coupled distributed resource unit subsequent to an islanding event. *IEEE Transactions on Power Delivery*, 23(1):493–501.

Kelleher, J. J. (1991). Tactical communications network modelling and reliability analysis: An overview. Technical report, DTIC Document.

Kim, Y. and Kang, W.-H. (2013). Network reliability analysis of complex systems using a non-simulation-based method. *Reliability Engineering & System Safety*, 110:80–88.

Krause, T., Beck, E., Cherkaoui, R., Germond, A., Andersson, G., and Ernst, D. (2006). A comparison of nash equilibria analysis and agent-based modelling for power markets. *International Journal of Electrical Power & Energy Systems*, 28(9):599–607.

Kröger, W. (2008). Critical infrastructures at risk: A need for a new conceptual approach and extended analytical tools. *Reliability Engineering & System Safety*, 93(12):1781–1787.

Kröger, W. and Zio, E. (2011). *Vulnerable systems*. Springer Science & Business Media.

Kuznetsova, E., Culver, K., and Zio, E. (2011). Complexity and vulnerability of smartgrid systems. In *Proceedings of the European Safety and Reliability Conference*, pages 2474–2482.

Kwasinski, A. (2011). Quantitative evaluation of dc microgrids availability: Effects of system architecture and converter topology design choices. *IEEE Transactions on Power Electronics*, 26(3):835–851.

Lasseter, R., Akhil, A., Marnay, C., Stephens, J., Dagle, J., Guttromson, R., Meliopoulous, A., Yinger, R., and Eto, J. (2002). The certs microgrid concept. *White paper for Transmission Reliability Program, Office of Power Technologies, US Department of Energy*, 2(3):30.

Lasseter, R. H. and Paigi, P. (2004). Microgrid: a conceptual solution. In *Power Electronics Specialists Conference, 2004. PESC 04. 2004 IEEE 35th Annual*, volume 6, pages 4285–4290. IEEE.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701.

Latora, V. and Marchiori, M. (2005). Vulnerability and protection of infrastructure networks. *Physical Review E*, 71(1):015103.

Leveson, N. (2004). A new accident model for engineering safer systems. *Safety science*, 42(4):237–270.

Lewis, A. M., Ward, D., Cyra, L., and Kourti, N. (2013). European reference network for critical infrastructure protection. *International journal of critical infrastructure protection*, 6(1):51–60.

Lewis, T. G. (2014). *Critical infrastructure protection in homeland security: defending a networked nation*. John Wiley & Sons.

Li, J., Dueñas-Osorio, L., Chen, C., and Shi, C. (2016). Connectivity reliability and topological controllability of infrastructure networks: A comparative assessment. *Reliability Engineering & System Safety*, 156:24–33.

Li, J. and He, J. (2002). A recursive decomposition algorithm for network seismic reliability evaluation. *Earthquake engineering & structural dynamics*, 31(8):1525–1539.

Li, Y.-F., Sansavini, G., and Zio, E. (2013). Non-dominated sorting binary differential evolution for the multi-objective optimization of cascading failures protection in complex networks. *Reliability Engineering & System Safety*, 111:195–205.

Li, Z., Yuan, Y., and Li, F. (2010). Evaluating the reliability of islanded microgrid in an emergency mode. In *Universities Power Engineering Conference (UPEC), 2010 45th International*, pages 1–5. IEEE.

Limiao, Z., Daqing, L., Pengju, Q., Bowen, F., Yinan, J., Zio, E., and Rui, K. (2016). Reliability analysis of interdependent lattices. *Physica A: Statistical Mechanics and its Applications*, 452:120–125.

Lindström, M. and Olsson, S. (2009). The European programme for critical infrastructure protection. In *Crisis Management in the European Union*, pages 37–59. Springer.

Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, 473(7346):167–173.

Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2012). Control centrality and hierarchical structure in complex networks. *Plos one*, 7(9):e44459.

Logenthiran, T., Srinivasan, D., and Wong, D. (2008). Multi-agent coordination for der in microgrid. In *Sustainable Energy Technologies, 2008. ICSET 2008. IEEE International Conference on*, pages 77–82. IEEE.

Lombardi, A. and Hörnquist, M. (2007). Controllability analysis of networks. *Physical Review E*, 75(5):056110.

Lopes, J. P., Moreira, C., and Madureira, A. (2006). Defining control strategies for microgrids islanded operation. *IEEE Transactions on power systems*, 21(2):916–924.

Lu, W., Su, M., Fath, B. D., Zhang, M., and Hao, Y. (2016). A systematic method of evaluation of the Chinese natural gas supply security. *Applied Energy*, 165:858–867.

Lussier, B., Chatila, R., Ingrand, F., Killijian, M.-O., and Powell, D. (2004). On fault tolerance and robustness in autonomous systems. In *Proceedings of the 3rd IARP-IEEE/RAS-EURON Joint Workshop on Technical Challenges for Dependable Robots in Human Environments*.

Mohamed, Y. A.-R. I. and Radwan, A. A. (2011). Hierarchical control system for robust microgrid operation and seamless mode transfer in active distribution systems. *IEEE Transactions on Smart Grid*, 2(2):352–362.

Negenborn, R., De Schutter, B., and Hellendoorn, H. (2006). Multi-agent model predictive control of transportation networks. In *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pages 296–301. IEEE.

Negenborn, R., Houwing, M., De Schutter, B., and Hellendoorn, J. (2009). Model predictive control for residential energy resources using a mixed-logical dynamic model. In *Networking, Sensing and Control, 2009. ICNSC'09. International Conference on*, pages 702–707. IEEE.

Nepusz, T. and Vicsek, T. (2012). Controlling edge dynamics in complex networks. *Nature Physics*, 8(7):568–573.

Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

Nieminen, J. (1974). On the centrality in a graph. *Scandinavian journal of psychology*, 15(1):332–336.

Olivares, D. E., Mehrizi-Sani, A., Etemadi, A. H., Cañizares, C. A., Iravani, R., Kazerani, M., Hajimiragha, A. H., Gomis-Bellmunt, O., Saeedifard, M., Palma-Behnke, R., et al. (2014). Trends in microgrid control. *IEEE Transactions on smart grid*, 5(4):1905–1919.

Ottino, J. M. (2004). Engineering complex systems. *Nature*, 427(6973):399–399.

Ouyang, M. (2014). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability engineering & System safety*, 121:43–60.

Oyarzabal, J., Jimeno, J., Ruela, J., Engler, A., and Hardt, C. (2005). Agent based micro grid management system. In *Future Power Systems, 2005 International Conference on*, pages 6–pp. IEEE.

Pan, Y. and Li, X. (2014). Structural controllability and controlling centrality of temporal networks. *PloS one*, 9(4):e94998.

Papadrakakis, M. and Lagaros, N. D. (2002). Reliability-based structural optimization using neural networks and monte carlo simulation. *Computer methods in applied mechanics and engineering*, 191(32):3491–3507.

Parisio, A. and Glielmo, L. (2011). Energy efficient microgrid management using model predictive control. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 5449–5454. IEEE.

Perez, E., Beltran, H., Aparicio, N., and Rodriguez, P. (2013). Predictive power control for pv plants with energy storage. *IEEE Transactions on Sustainable Energy*, 4(2):482–490.

Pino, W., Worm, D., van der Linden, R., and Kooij, R. (2016). The reliability of a gas distribution network: A case study. In *System Reliability and Science (ICSRS), International Conference on*, pages 122–129. IEEE.

Pósfai, M., Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2012). Effect of correlations on network controllability. *arXiv preprint arXiv:1203.5161*.

Praks, P. and Kopustinskas, V. (2016). Identification and ranking of important elements in a gas transmission network by using progasnet. In *Risk, Reliability and Safety: Innovating Theory and Practice*, pages 1573–1579. CRC Press.

Praks, P., Kopustinskas, V., and Masera, M. (2015). Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure. *Reliability Engineering & System Safety*, 144:254–264.

Prodan, I. and Zio, E. (2014a). A model predictive control framework for reliable microgrid energy management. *International Journal of Electrical Power & Energy Systems*, 61:399–409.

Prodan, I. and Zio, E. (2014b). On the microgrid energy management under a predictive control framework. In *Control Applications (CCA), 2014 IEEE Conference on*, pages 861–866. IEEE.

Prodan, I., Zio, E., and Stoican, F. (2015). Fault tolerant predictive control design for reliable microgrid energy management under uncertainties. *Energy*, 91:20–34.

Ramirez-Marquez, J. E. and Rocco, C. M. (2008). All-terminal network reliability optimization via probabilistic solution discovery. *Reliability Engineering & System Safety*, 93(11):1689–1697.

Ramírez-Rosado, I. J. and Bernal-Agustín, J. L. (2001). Reliability and costs optimization for distribution networks expansion using an evolutionary algorithm. *IEEE Transactions on Power Systems*, 16(1):111–118.

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety science*, 27(2):183–213.

Rawlings, J. and Mayne, D. (2011). Postface to model predictive control: Theory and design.

Richalet, J. and O'Donovan, D. (2009). *Predictive Functional Control: Principles and Industrial Applications*. Springer.

Rigaud, E. and Guarnieri, F. (2006). Proposition of a conceptual and a methodological modelling framework for resilience engineering. In *2nd Symposium on Resilience Engineering*, pages 8–pages.

Rouse, W. B. (2003). Engineering complex systems: Implications for research in systems engineering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2(33):154–156.

Saleh, J. H., Marais, K. B., Bakolas, E., and Cowlagi, R. V. (2010). Highlights from the literature on accident causation and system safety: Review of major ideas, recent contributions, and challenges. *Reliability Engineering & System Safety*, 95(11):1105–1116.

Sawaragi, Y., NAKAYAMA, H., and TANINO, T. (1985). *Theory of multiobjective optimization*, volume 176. Elsevier.

Shier, D. R. (1991). *Network reliability and algebraic structures*. Clarendon Press.

Singh, H., Vaithilingam, S., Anne, R., and Anneberg, L. (1996). Terminal reliability using binary decision diagrams. *Microelectronics Reliability*, 36(3):363–365.

Sinha, K. and de Weck, O. L. (2013a). A network-based structural complexity metric for engineered complex systems. In *Systems Conference (SysCon), 2013 IEEE International*, pages 426–430. IEEE.

Sinha, K. and de Weck, O. L. (2013b). Structural complexity quantification for engineered complex systems and implications on system architecture and design. In *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages V03AT03A044–V03AT03A044. American Society of Mechanical Engineers.

Song, J. and Kang, W.-H. (2009). System reliability and sensitivity under statistical dependence by matrix-based system reliability method. *Structural Safety*, 31(2):148–156.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268–276.

Üster, H. and Dilaveroğlu, Ş. (2014). Optimization for design and operation of natural gas transmission networks. *Applied Energy*, 133:56–69.

Van der Borst, M. and Schoonakker, H. (2001). An overview of psa importance measures. *Reliability Engineering & System Safety*, 72(3):241–245.

Vasquez, J. C., Guerrero, J. M., Miret, J., Castilla, M., and De Vicuna, L. G. (2010). Hierarchical control of intelligent microgrids. *IEEE Industrial Electronics Magazine*, 4(4):23–29.

Wang, S., Li, Z., Wu, L., Shahidehpour, M., and Li, Z. (2013). New metrics for assessing the reliability and economics of microgrids in distribution system. *IEEE Transactions on Power Systems*, 28(3):2852–2861.

Wang, W.-X., Ni, X., Lai, Y.-C., and Grebogi, C. (2012). Optimizing controllability of complex networks by minimum structural perturbations. *Physical Review E*, 85(2):026115.

Weidlich, A. and Veit, D. (2008). A critical survey of agent-based wholesale electricity market models. *Energy Economics*, 30(4):1728–1759.

Wu, W., Simpson, A. R., and Maier, H. R. (2009). Accounting for greenhouse gas emissions in multiobjective genetic algorithm optimization of water distribution systems. *Journal of water resources planning and management*, 136(2):146–155.

Yan, G., Ren, J., Lai, Y.-C., Lai, C.-H., and Li, B. (2012). Controlling complex networks: How much energy is needed? *Physical review letters*, 108(21):218703.

Yeh, W.-C. (2007). An improved sum-of-disjoint-products technique for the symbolic network reliability analysis with known minimal paths. *Reliability Engineering & System Safety*, 92(2):260–268.

Yeh, W.-C., Lin, Y.-C., Chung, Y. Y., and Chih, M. (2010). A particle swarm optimization approach based on monte carlo simulation for solving the complex network reliability problem. *IEEE Transactions on Reliability*, 59(1):212–221.

Yokoyama, R., Niimura, T., and Saito, N. (2008). Modeling and evaluation of supply reliability of microgrids including pv and wind power. In *Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–5. IEEE.

Yuan, Z., Zhao, C., Di, Z., Wang, W.-X., and Lai, Y.-C. (2013). Exact controllability of complex networks. *Nature communications*, 4.

Zamora, R. and Srivastava, A. K. (2010). Controls for microgrids with storage: Review, challenges, and research needs. *Renewable and Sustainable Energy Reviews*, 14(7):2009–2018.

Zhang, P., Li, W., Li, S., Wang, Y., and Xiao, W. (2013). Reliability assessment of photovoltaic power systems: review of current status and future perspectives. *Applied energy*, 104:822–833.

Zhang, Z., Li, X., and Li, H. (2015). A quantitative approach for assessing the critical nodal and linear elements of a railway infrastructure. *International Journal of Critical Infrastructure Protection*, 8:3–15.

Zio, E. (2007). From complexity science to reliability efficiency: a new way of looking at complex network systems and critical infrastructures. *International Journal of Critical Infrastructures*, 3(3-4):488–508.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, 94(2):125–141.

Zio, E. (2016). Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering & System Safety*, 152:137–150.

Zio, E. and Golea, L. R. (2012). Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. *Reliability Engineering & System Safety*, 101:67–74.

Zio, E., Sansavini, G., Maja, R., and Marchionni, G. (2008). An analytical approach to the safety of road networks. *International Journal of Reliability, Quality and Safety Engineering*, 15(01):67–76.

Paper I

F. Han, E. Zio. Modeling an electric power microgrid by model predictive control for analyzing its characteristics from reliability, controllability and topological perspectives. Journal of Risk and Reliability, 2017. (Accepted).

Modeling an electric power microgrid by model predictive control for analysing its characteristics from reliability, controllability and topological perspectives

Fangyuan Han[1] and Enrico Zio[1,2]

[1]Chair on Systems Science and the Energetic Challenge, Fondation EDF, CentraleSupelec, France

[2]Department of Energy, Politecnico di Milano, Italy.

Abstract

Microgrids can be a key solution for integrating renewable and distributed energy resources. This paper analyzes microgrids characteristics adopting model predictive control. We study the microgrid performance under two operation modes: grid-connected and stand-alone. For each mode, we consider different faulty scenarios and by dynamic simulations we investigate the importance of the microgrid components from different perspectives: topological, reliability and controllability. This analysis enables evaluation of the microgrid performance and quantification of the importance of each component with respect to the different perspectives considered. The findings provide information for the design and operation of a microgrid, seeking the right balance of multiple characteristics.

Keywords: Microgrid modeling, Model predictive control, Reliability, Controllability, Network topology.

1 Introduction

Critical infrastructures, like electricity or gas transmission and distribution systems, rail and road transport or communication networks, are essential to the functioning of modern

society [1]. They are designed to perform reliably and safely for long periods of times [2,3]. The complexity of these systems calls for new approaches of analysis and a framework is needed to integrate a number of methods capable of viewing the problem from different perspectives [4–6]. Integration of different perspectives has been sought. For example, in [7] an electrical transmission system is analyzed with the objective of identifying the most critical elements in terms of four different perspectives; in [8], the correlation between connectivity reliability and controllability of network systems has been studied; in [9], the authors perform network reliability analysis considering spatial constraints.

In this paper, we focus on power distribution systems and distributed generation, and in particular microgrids which offer an interesting solution for integrating renewable and distributed energy resources. A microgrid is a cluster of micro-sources, storage systems and loads, which can be connected to the power grid as a single entity that can respond to central control signals [10]. The control problem for this kind of systems is particularly difficult as it is necessary to consider both exogenous factors (e.g. variations of wind speed, consumer demand) and the structural properties and internal dynamics of individual components (e.g. links, storage devices), which may change due to degradation, failure and other factors. [11]. Various approaches for control and energy management of microgrids are reported in the literature [12]. In [13–16] an agent-based modeling approach is proposed to model microgrids and to analyze by simulation the interactions between individual intelligent decision-makers. [11,17] and [18] develop an optimization-based control approach.

It is considered that microgrids can improve the reliability of servicing local loads [19,20], besides that of the power grid to which they are connected.

In this work, we consider a microgrid system and adopt a graph representation and dynamic modeling for capturing its structural properties and internal dynamics. Different perspectives are considered for the analysis of the system topology, reliability and controllability.

Model predictive control (MPC) [21, 22], a widely used technique in the control community, can be used to manage the dynamics of systems affected by uncertainties in the behavior of their components [11]. It is able to handle control and state constraints, while offering good control performance. For this, in MPC the objective (or cost) function is constructed to penalize deviations of the states and inputs from their reference values, while the constraints are enforced explicitly [23]. Recently, MPC has been considered for refrigeration systems [24], heating systems [25], power production plants [26] and transportation networks [27].

In this paper, the MPC framework is proposed to analyze the microgrid under various faulty scenarios.

The original contributions are the following:

- The formulation of an optimization-based model predictive control problem for safety and reliability considerations of microgrid systems.
- The development and application of a simulation-based scheme for the analysis of a microgrid from different perspectives: reliability, controllability and topology

The remainder of the paper is organized as follows. Section 2 describes in details the representation of the microgrid system, including its network representation, the dynamic model and the formulation of the optimization-based control problem. In Section 3, three system indicators are presented from different perspectives. Section 4 presents simulation results for different faulty scenarios. Finally, conclusions are drawn in Section 5.

2 System modeling and description

2.1 Graph representation of the microgrid

Graph theory provides a natural framework for the mathematical representation of complex networks [28, 29]. A graph is an ordered pair $\mathbf{G}(V, E)$ comprising a set of nodes (vertices) $V = v_1, v_2, \ldots, v_n$, together with a set of links (also called edges or arcs) $E = e_1, e_2, \ldots, e_m$, which are two-element subsets of V. The topological structure of a microgrid can be represented by a directed graph: the nodes represent the components or subsystems of the microgrid and the directional edges represent the functional links between the microgrid components.

The network structure is usually defined by the $n \times n$ adjacency matrix $\mathbf{A}$, which can be constructed as follows: if there is an edge from node i to node j, then we put a value of 1 in the entry on row i, column j of the matrix; otherwise, we put a value of 0. The $n \times n$ capacity matrix $\mathbf{K}$ contains information about the capacity constraints of the links.

2.2 Dynamic modeling of the microgrid

We introduce the dynamic model for representing the characteristics of the nodes and the flow in the links of the microgrid. We adopt a state-space model based on differential equations to describe the response of the system states to operational and environmental

changes. The dynamic of the system is described by the following linear time-invariant equations:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{1}$$

where $\mathbf{x}$ is the state vector, $\mathbf{u}$ is the vector of control input, $\mathbf{y}$ is the output vector. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are state transition matrices.

The components of the microgrid can be divided into different groups according to their functionalities: links, storage devices, suppliers (renewable generators), transporters and consumers. The state vector $\mathbf{x}$ represents mainly the states of the links and storages devices, which are treated as dynamic, whereas the states of other nodes are considered static.

2.2.1 Link dynamics

For a link l_i, the state $x_{l_i}(t)$ indicate the capacity of the link to deliver power from one component to the other connected one at time t. It is assumed to be determined by both the input flow and the state of the previous time step. Its dynamic can be described as following:

$$x_{l_i}(t+1) = (1-\alpha_i)x_{l_i}(t) + \alpha_i \sum_{l_{in} \in I_{l_i}} u_{l_{in}}(t) \tag{2}$$

where α_i is a coefficient of small value characterizing the inertia of flow transmission that depends on the physical characteristics of the link l_i, $u_{l_{in}}(t)$ is an input flow of the link l_i, and I_{l_i} is the set of input flows of the the link l_i.

2.2.2 Storage device dynamics

For a storage device s_i, its state $x_{s_i}(t)$ represents the energy storage level at time step t and depends on the energy level of the previous time step and the charge or discharge rates. The dynamic is described as:

$$x_{s_i}(t+1) = (1-\tau_i)x_{s_i}(t) + \sum_{s_{in} \in I_{s_i}} u_{s_{in}}(t) - \sum_{s_{out} \in O_{s_i}} u_{s_{out}}(t) \tag{3}$$

with the mixed-integer conditions [17]:

$$\begin{cases} 0 \leq \sum u(t) \leq Ma(t), \\ 0 \leq \sum x(t) \leq M(1-a(t)), \end{cases} \tag{4}$$

where τ denotes the hourly self-discharge decay. $u_{s_{in}}(t)$ and $u_{s_{out}}(t)$ are input and output flow of the storage device s_i respectively, and I_{s_i} and O_{s_i} the sets of inout and output flows of the storage device s_i, respectively. M represents a constraint and $a(t) \in \{0, 1\}$ is an auxiliary binary variable, characterizing the battery state of charge: when $a(t) = 1$ the battery is in discharge mode, when $a(t) = 0$ the battery is in charge mode [11].

2.3 Optimization-based control for system safety analysis

The microgrid safety performance measured in terms of satisfaction of consumer power demand and solve an optimization problem in order to find the control input that minimize the cost function $Cost(t)$ (e.g. the difference between the power demanded by the consumer and that actually received), subject to a set of system constraints and considering predicted profiles.

$$\min_{[\mathbf{u}(t)]_{t=k:k+N_p}} \sum_{t=k}^{k+N_p} Cost(t)$$

2.3.1 Constraints on sources (renewable generators)

Renewable generators are considered as the sources of flows in a microgrid. Therefore, the total amount of output flows of a supplier p_i should be lower than its supply capacity:

$$\sum_{p_{out} \in O_{p_i}} u_{p_{out}}(t) \leq K_{p_i} \tag{5}$$

where $u_{p_{out}}(t)$ is an output flow of the source p_i, O_{p_i} is the set of output flows of the source p_i, and K_{p_i} is the capacity of of the source p_i.

2.3.2 Constraints on transporters

A transporter t_i is a static node related to transmission or distribution, where the dynamic flows follow basic conservation laws, i.e., the total amount of the output flows is equal to that of the input flows:

$$\sum_{t_{in} \in I_{t_i}} u_{t_{in}}(t) = \sum_{t_{out} \in O_{t_i}} u_{t_{out}}(t) \tag{6}$$

where $u_{t_{in}}(t)$ and $u_{t_{out}}(t)$ are input and output flows of the transporter t_i respectively, and $I_{t_i}$and O_{t_i} are the sets of input and output flows of the transporter t_i.

2.3.3 Constraints on consumers

The objective of the microgrid system is to supply power to satisfy consumers demand. Then, the amount of flows received by the consumer never exceeds its demand:

$$y_{m_i} \leq D_{m_i}(t) \tag{7}$$

where $y_{m_i}(t)$ is the output of the system corresponding to the state of the consumer m_i and D_{m_i} is the demand of the consumer m_i.

2.3.4 Constraints on links

The flow through the link l_i is constrained by its maximum capacity K_{l_i}:

$$0 \leq x_{l_i} \leq K_{l_i} \tag{8}$$

where K_{l_i} is the maximum capacity of the link l_i.

2.3.5 Constraints on storage devices

The amount of flow charged in a storage device s_i is limited by its storage capacity S_{max_i} and by its base storage S_{min_i}:

$$S_{min_i} \leq x_{s_i}(t) \leq S_{max_i} \tag{9}$$

Similar constraints apply to the rates of the battery charge:

$$Sr_{min_i} \leq \Delta x_{s_i}(t) \leq Sr_{max_i}, \tag{10}$$

3 System property indexes

3.1 Non-supplied demand

Microgids have been proposed to improve reliability and stability of electrical systems and to ensure power quality of modern grids, and have the responsibility to ensure the supply to the essential loads [12]. Supply performance is a fundamental functional requirement for the microgrid. In this paper, we call the system "safe" if it ensures the satisfaction of the consumers essential demands. We introduce the non-supplied demand (NSD) as a measure of the network's capacity to satisfy its users' demands. The normalized NSD is introduced as a system-level index:

$$NSD = 1 - \frac{\sum_i^{i=N_m} \omega_i y_{m_i}}{\sum_i^{i=N_m} \omega_i D_{m_i}} \tag{11}$$

where, ω_i is the weight of the ith of the N_m users, y_{m_i} is the supply to user i and D_{m_i} is its demand, which is considered as the target supply to user i. Then, the second term in Equation (11) represents the satisfied proportion of users' demands. Since $y_{m_i} \leq D_{m_i}$, the index NSD is normalized to take values in $[0, 1]$. NSD equals 0 when the users' demands are fully supplied.

3.2 Controllability Index

A dynamic system is controllable if, with a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time [30]. Taking a system safety perspective, controllability is the ability to guide the systems behavior towards a safe state through the appropriate manipulation of a few input variables [11, 31].

From control theory, the system (as described by equation 1) is controllable if and only if its controllability matrix has full rank [32]:

$$rank[B \quad AB \quad ... \quad A^{n-1}B] = n$$

where n is the number of state variables of the system. This criteria is called Kalmans controllability rank condition. The rank of the controllability matrix provides the dimension of the controllable subspace of the system.

In this work, the controllability index CI measures the controllable proportion of a dynamic system. It is defined as the ratio of the rank of the controllable subsystem to the rank of the system:

$$CI = \frac{R_C}{n} \tag{12}$$

where $R_C = rank[B \quad AB \quad ... \quad A^{n-1}B]$.

3.3 System capacity efficiency

We introduce the system capacity efficiency to measure how much flow the system topology allows to exchange. The capacity of flow exchange from nodes i to j through a path is determined by the capacity of the widest-capacity path between them, k_{ij}, which is the minimum edge capacity in the path between the two nodes maximizing the capacity of the minimum-capacity edge. Then, the capacity efficiency of the whole system E_c is given by:

$$E_C[\mathbf{G}] = \frac{1}{N(N-1)} \sum_{i \neq j \in \mathbf{G}} k_{ij} \tag{13}$$

The source-terminal capacity efficiency E_c^{st}, which only takes into account the transmission capacity between a source node and a terminal (demand) node, is given by:

$$E_C^{st}[\mathbf{G}] = \frac{1}{N_{st}} \sum_{s \in \mathbf{S}, t \in \mathbf{T}} c_{st} \tag{14}$$

Then, we define the source-terminal capacity efficiency index (EI^{st}) as the normalized E_C^{st}:

$$EI^{st}[\mathbf{G}'] = \frac{E_C^{st}[\mathbf{G}']}{E_C^{st}[\mathbf{G}]} \tag{15}$$

where $\mathbf{G}'$ is the graph obtained by the removal of certain components from $\mathbf{G}$.

4 Case study and simulation results

4.1 Case study: microgrid

We consider the microgrid system in Figure 1, taken from [11]. This microgrid system contains one renewable generator (wind turbine), one storage device (battery) and one local consumer. The microgrid system is connected to the external power grid through a transformer.All the components are characterized by the dynamic models, constraints and reference profiles presented in the following.

The nature of the paper is methodological and the considered system is a simplified example to illustrate the methods proposed. It is a system including a variety of components that make up a microgrid and define its characteristics. Other components, links, profiles and constraints can be added and modeled. Then, the optimization problem for the model predictive control can be regarded as a mixed-integer linear programming program, for which various efficient solvers exist.

The microgrid can be modeled as a graph of four nodes (the external grid, the renewable generator, the battery and the consumer) and five links (from external grid to consumer, generator to external grid, generator to consumer, generator to battery and battery to consumer). In this work, the node representing the battery and the five links are considered dynamic.

4.1.1 Dynamic model of the microgrid

We consider the dynamic models for six components including the five links and the storage device (battery). The description of the system dynamics leads to a 6 elements state vector:

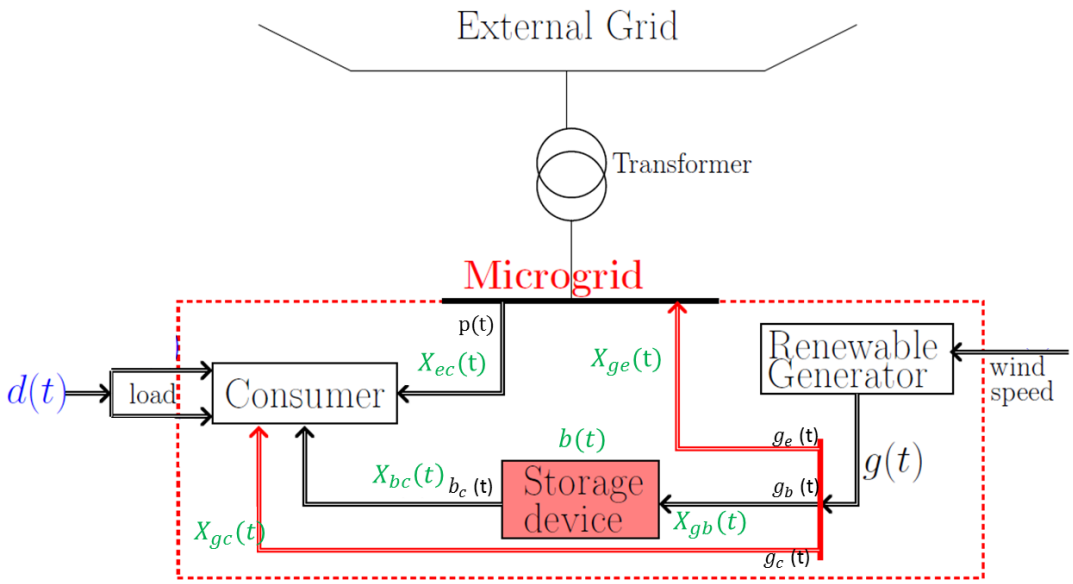


Figure 1: Microgrid

5 states contain the values of energy in the links that can propagate to the next node and the sixth represents the battery energy level [11]:

$$\mathbf{x}(t) = [x_{ec}(t) \quad x_{ge}(t) \quad x_{gc}(t) \quad x_{gb}(t) \quad x_{bc}(t) \quad b(t)]^T$$

The corresponding dynamic models are:

$$\text{External grid to consumer:} \quad x_{ec}(t+1) = (1-\alpha)x_{ec}(t) + \alpha p(t) \tag{16}$$

$$\text{Generator to external grid:} \quad x_{ge}(t+1) = (1-\alpha)x_{ge}(t) + \alpha g_e(t) \tag{17}$$

$$\text{Generator to consumer:} \quad x_{gc}(t+1) = (1-\alpha)x_{gc}(t) + \alpha g_c(t) \tag{18}$$

$$\text{Generator to battery:} \quad x_{gb}(t+1) = (1-\alpha)x_{gb}(t) + \alpha g_b(t) \tag{19}$$

$$\text{Battery to consumer:} \quad x_{bc}(t+1) = (1-\alpha)x_{bc}(t) + \alpha b_c(t) \tag{20}$$

where $\alpha \in [0,1]$ is a fixed constant, mainly dependent upon the size of the discretization step, and,

$$\text{Battery:} \quad b(t+1) = (1-\tau)b(t) + x_{gb}(t) - b_c(t) + w(t) \tag{21}$$

with the mixed-integer conditions [17]:

$$\begin{cases} 0 \leq b_c(t) \leq Ma(t), \\ 0 \leq x_{gb}(t) \leq M(1-a(t)), \end{cases} \tag{22}$$

The above six state variables can be inferred by the vector of system control inputs $\mathbf{u}(t)$: $\mathbf{u}(t) = [p(t) \quad g_e(t) \quad g_c(t) \quad g_b(t) \quad b_c(t)]^T$ where [11]:

- $p(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the external grid to the consumer, at time step t.
- $g_e(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the external grid, at time step t.
- $g_b(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the battery, at time step t.
- $g_c(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the renewable generator to the consumer, at time step t.
- $b_c(t) \in \mathbb{R}$ [W] represents the electrical power transmitted by the battery to the consumer, at time step t.

The consumer has the possibility to take electrical power from the external grid, the renewable generator and the battery. Thus, the sum of powers received by the consumer is $x_{ec}(t) + x_{gc}(t) + x_{bc}(t)$. Finally, the system output $y(t)$ is the total power received by the consumer:

$$y(t) = x_{ec}(t) + x_{gc}(t) + x_{bc}(t)$$

In the end, the microgrid can be described by the following global dynamic model:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \tag{23}$$

where

$$\mathbf{A} = \begin{bmatrix} 1-\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-\alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\alpha & 0 \\ 0 & 0 & 0 & 1 & 0 & 1-\tau \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

4.1.2 Reference profiles

We consider two reference profiles characterizing the microgrid components (i.e. the consumer and the renewable generator) based on real numerical data taken from [33] and the details can be found in [17]. The consumer load takes into account seasonal numerical data, and is predicted by using weekly, daily and hourly peaks. Figure 2 shows the consumer load profile $d(t) \in \mathbb{R}$. The wind speed used to calculate the wind power profile is estimated based on meteorological data. Figure 3 shows the power generator profile: the electrical power generated $g(t) \in \mathbb{R}$ is obtained from the wind speed profile [34].

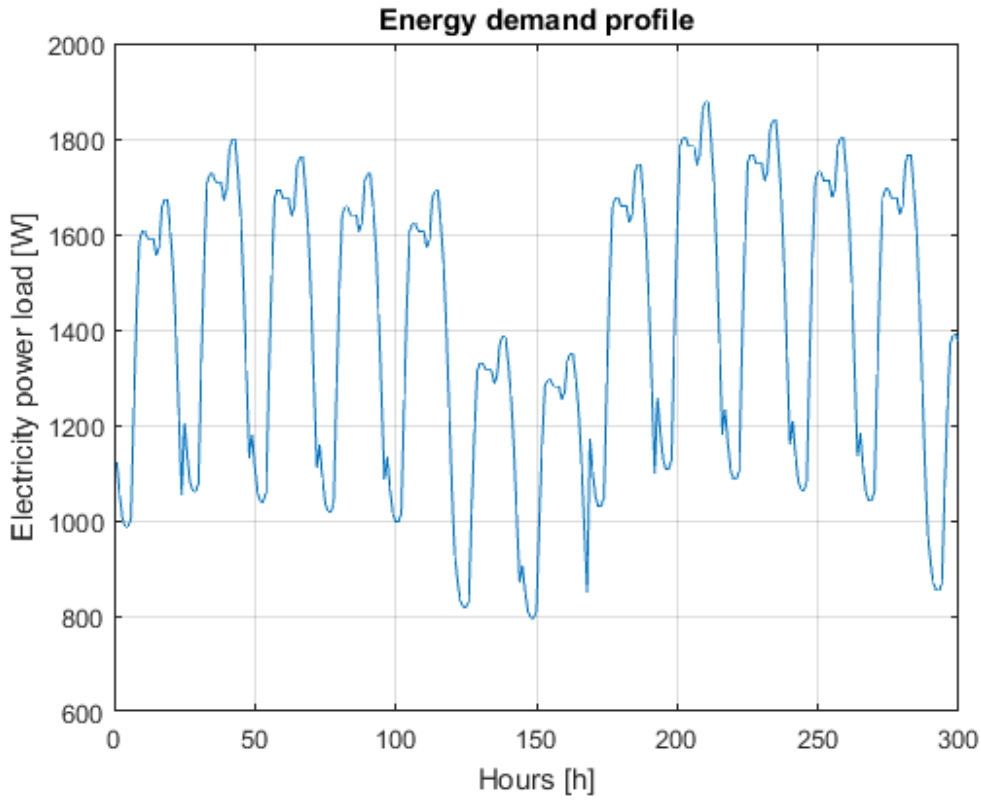


Figure 2: Consumer load profile

4.1.3 Optimization problem of the microgrid

The optimization problem is to find the appropriate control inputs that minimize the difference between the power demanded by the consumer and that actually received. Thus,

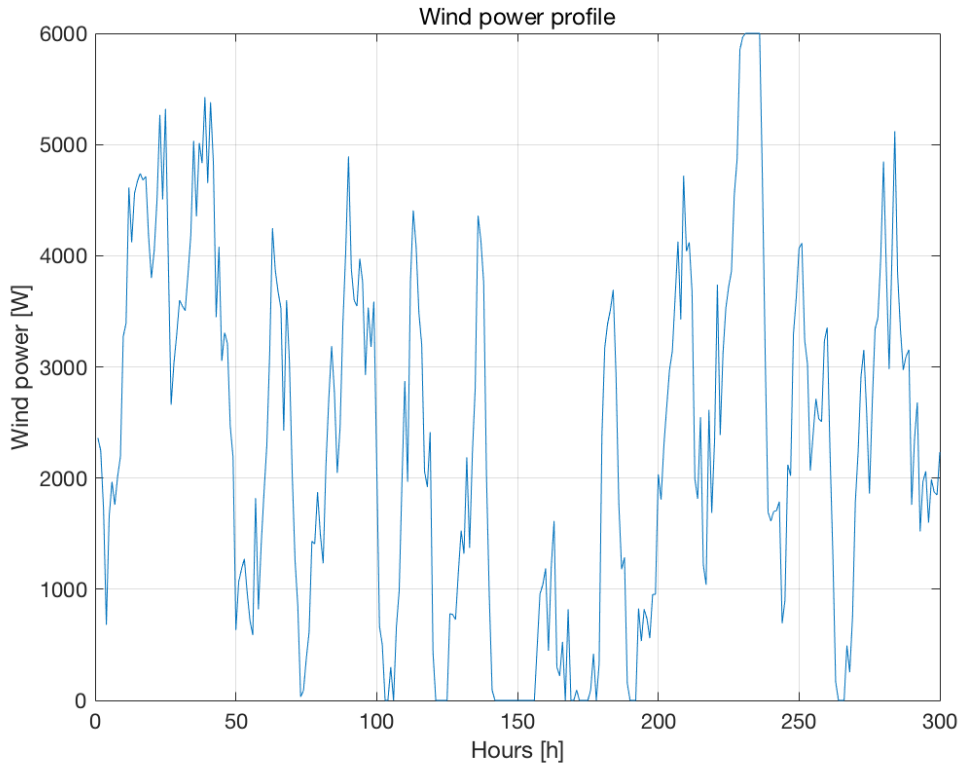


Figure 3: Wind power profile

the objective function for MPC is:

$$\min_{[\mathbf{u}(t)]_{t=k:k+N_p}} \sum_{t=k}^{k+N_p} d(t) - y(t)$$

with the set of constraints defined in the following equations (24) - (28).

- Satisfaction of consumer power demands

$$0 \leq x_{ec}(t) + x_{gc}(t) + x_{bc}(t) \leq d(t) \tag{24}$$

where $d(t)$ is the consumer's demand.

- Battery storage

Batteries have their physical characteristics: the minimum capacity B_{min} determined by the Depth of Discharge (DoD) [35] and the capacity B_{max}. The rate of the battery charge is also limited by some bounds:

$$B_{min} \leq b(t) \leq B_{max}, \tag{25}$$

$$Br_{min} \leq \Delta b(t) \leq Br_{max}, \tag{26}$$

where $B_{min} \in \mathbb{R}$, $B_{max} \in \mathbb{R}$, $Br_{min} \in \mathbb{R}$, $Br_{max} \in \mathbb{R}$.

- Generator

The power taken from the generator by the battery ($g_b(t)$), the consumer ($g_c(t)$) and the external grid ($g_e(t)$) is bounded by the power generated:

$$0 \leq g_b(t) + g_c(t) + g_e(t) \leq g(t), \tag{27}$$

with $g_b(t) \geq 0$, $g_c(t) \geq 0$, $g_e(t) \geq 0$.

- Link capacities

$$\mathbf{u}_{min} \leq \mathbf{u}(t) \leq \mathbf{u}_{max}, \tag{28}$$

where $u(t) \in \mathbb{R}$.

The numerical values of the parameters used for the simulations are taken from [33]. See Table 1.

Table 1: Numerical data for the microgrid components

Battery parameters	
τ	$1.3 \cdot 10^{-4}$
M	$9 \cdot 10^3$
B_{min} $[Wh]$	$1.2 \cdot 10^3$
B_{max} $[Wh]$	$9 \cdot 10^3$
Br_{min} $[W]$	$-1.5 \cdot 10^3$
Br_{max} $[W]$	$1.5 \cdot 10^3$
Control input constraints	
U_{min}	$[0 \quad 0 \quad 0 \quad 0 \quad 0]^T$
U_{max}	$[2 \quad 1.5 \quad 2 \quad 1.5 \quad 1.5]^T \cdot 10^3$
Prediction horizon	
N_p	7
Simulation steps	
N	300

4.2 Scenarios

We consider two operation modes for the microgrid: grid-connected and stand-alone. Under these two modes, the microgrid is designed to satisfy consumers demand. We assume that the external grid and the renewable generator are fault-free. Then, threats to the microgrid service may come from failure of the links from the three sources (i.e. the external grid e_c, the renewable generator g_c and the battery b_c) to the consumer. The two other links from the renewable generator (i.e. g_e and g_b) are also considered, since they impact on the cost of the microgrid. The links failures are represented as the removal of the links and no recovery action is taken.

4.2.1 Grid-connected mode

During the grid connected mode, the consumer takes electrical power from two sources: the external grid and the renewable generator.

The scenarios considered are the following:

- Scenario 1.0: the nominal functioning case, i.e. fault-free.
- Scenario 1.1: the link from the generator to the external grid is disconnected (i.e. g_e is removed).
- Scenario 1.2: the link from the generator to the consumer is disconnected (i.e. g_c is removed).

Note that the failure of the battery is not considered for the grid-connected mode, since in that mode it is assumed that the battery is not used.

4.2.2 Stand-alone mode

In stand-alone mode, the microgrid is disconnected from the external power grid and generation should by itself satisfy consumers demand.

The scenarios considered are as following:

- Scenario 2.0: the stand-alone functioning case (i.e. only p is disconnected).
- Scenario 2.1: the link from the generator to the consumer is disconnected (i.e. g_c is removed).
- Scenario 2.2: the link from the generator to the battery is disconnected (i.e. g_b is removed).
- Scenario 2.3: the link from the battery to the consumer is removed (i.e. b_c is removed)

4.3 Analysis of the results

We analyze the system property indexes introduced in Section 3 for the scenarios described above. For the grid-connected mode, we also analyze the difference between cost and profit of the microgrid for each scenario. The simulation of each scenario is considered for the period of 300 hours, during which the microgrid experiences almost all extreme conditions of consumer demands and winder power. The simulation is taken for the period of 300 hours.

4.3.1 Grid-connected mode

Table 2: Index values for the grid connected mode

Scenario	NSD	CI	EI^{st}	Profit
1.0	0	1	1	+203.2
1.1	0	0.83	1	-77.0
1.2	0	0.83	1	-79.2

From Table 2, we can see that, in grid-connected mode, the demands of the consumers can always be satisfied. The two links have identical influence on the system controllability: with the removal of each link, the microgrid is no longer controllable, and the rank of controllability matrix decreases by 1, which means that one component is out of control. The capacity efficiency index, EI^{st}, remains the same for all the scenarios. However, the cost differs a lot: when the generator is able to provide power to the consumer and sell power to the external grid, the microgrid is profitable (Scenarios 1.0); otherwise, the microgrid spends money on electricity (Scenarios 1.1 and 1.2).

Figures 4 and 5 show the sources of the power actually received by the consumer for Scenario 1.0 and Scenario 1.1, respectively.

In Scenario 1.0, the consumer takes electrical power from the external grid and the renewable generator. We can see from the Figure 4 that the renewable generator (g_c) provides most of the demanded power (represented by the shaded bar). If compared with the load profile and wind power profile (see Figure 2 and Figure 3), we can see that the consumer takes electricity from the external grid when the wind power is low; and when the wind power is high, the microgrid sells electricity to the external grid: this is why in the case of nominal functioning the microgrid can make a profit.

In Scenario 1.1, the power sources remains the same as Scenario 1.0. The system prop-

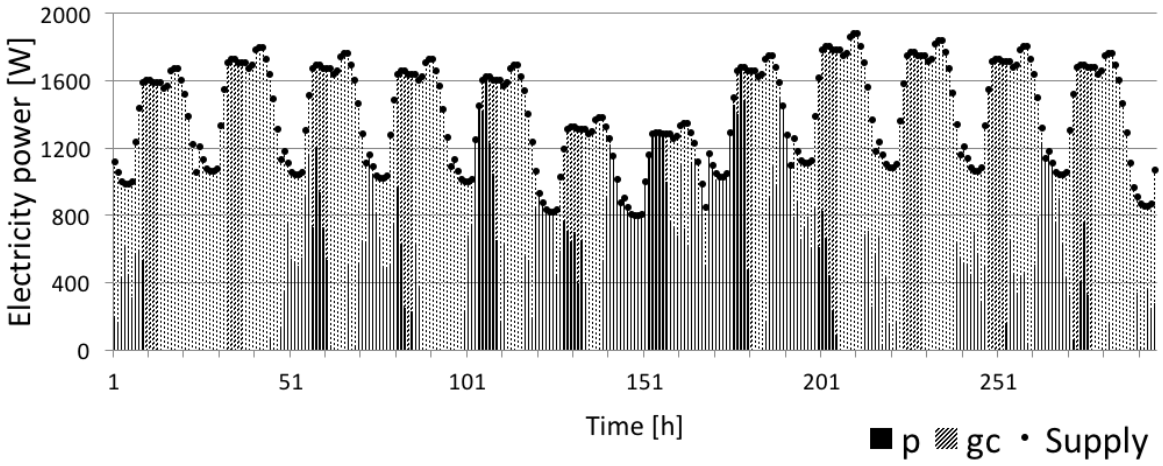


Figure 4: Supply sources of Scenario 1.0. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator.

erties remain the same as the nominal functioning case, and the sources of the supply to the consumer is almost the same as the in Scenario 1.0 (see Figure 5). However, with the disconnection of the link g_e, it's impossible to sell electricity to recompense the expenses on electricity bought from the external grid, therefore, instead of making a profit, it spends money to buy electricity.

As for Scenario 1.2, where the link from the generator to the consumer g_c is disconnected, the consumer is supplied completely by the external grid, which also explains the fact that the non-supplied demand NSD is always equal to 0. In addition, since the capacity of link g_e is limited, the wind power generated cannot be fully sold, i.e. the profit by selling electricity is not able to compensate the expenses.

In the grid-connected mode, the consumer's demand is always satisfied, which is natural reasonable. However, the introduction of the microgrid (renewable generator) decreases the cost on electricity or even makes a profit and it is, thus, interesting for economic considerations.

4.3.2 Stand-alone mode

From Table 3, we can see that the non-supplied demand NSD increases in this mode and the demands are never fully satisfied. More detailed analysis is given in the following. The capacity efficiency index EI^{st} decreases, because the disconnection of the link p reduces the transmission capacity of the microgrid; this also contributes to the inadequate supply of the

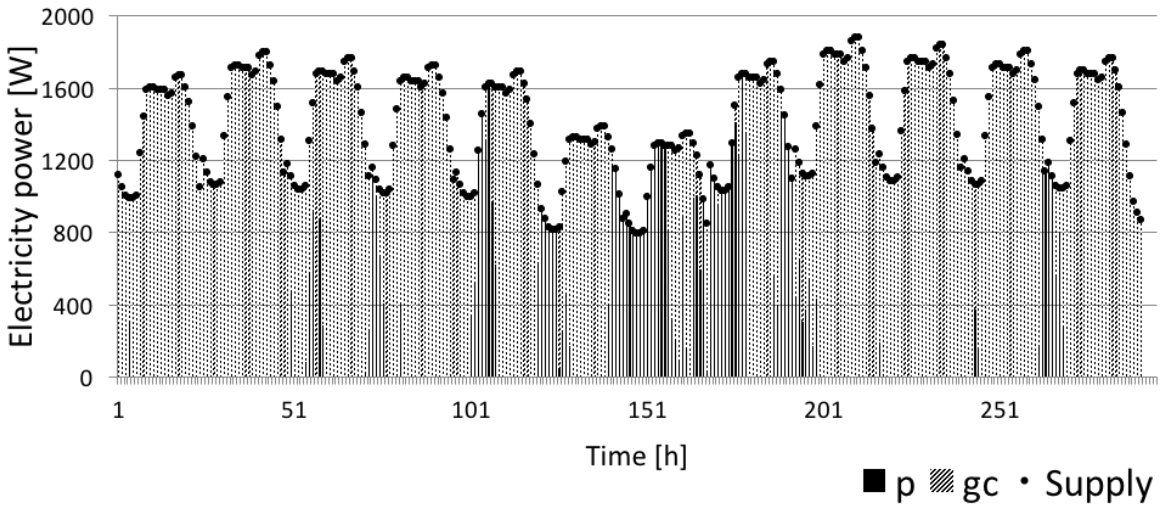


Figure 5: Supply sources of Scenario 1.1. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator.

Table 3: Index values for the stand-alone mode

Scenario	NSD	CI	EI^{st}
2.0	0.1065	0.83	0.667
2.1	0.6562	0.83	0.583
2.2	0.4190	0.83	0.667
2.3	0.4364	0.83	0.333

consumer's demands.

Figure 6 shows the sources of the power actually received by the consumer for Scenario 2.0. From the beginning to $t = 145$ hours under this scenario, the supply is similar to that of Scenarios 1.0 and 1.1: the renewable generator provides most of the demanded power, and the battery fills the unsatisfied portion, except for certain periods of low wind power generation. Then, the microgrid arrives at a relatively long period when there is no wind power at all, the battery reaches its lower limit, and the supply is totally cut. Around t=200, when the wind comes back to its usual level, the microgrid continues to function.

In Scenario 2.1, the control input of the link from the generator to the consumer g_c is lost, while the battery provides power to the consumer and can be charged by g_b. NSD is the highest of all the stand-alone modes. This is due to the fact that the supply to the consumer is dominated by the capacity of the link from the battery (g_b has smaller capacity than g_b), which is also reflected by the fact that EI^{st} is lower than that of Scenario 2.0. In

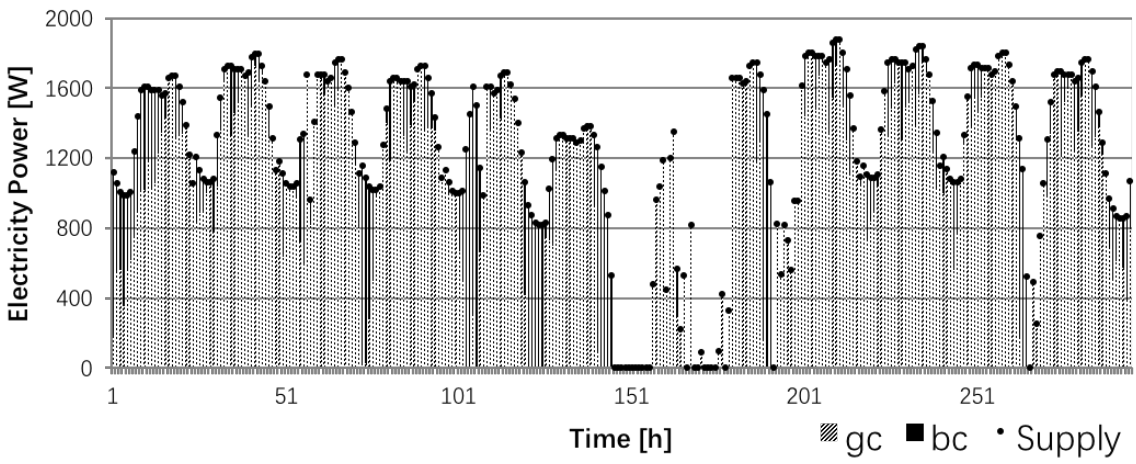


Figure 6: Supply sources of Scenario 2.0. The black area represents the power from the external grid and the shaded area represents the power from the renewable generator.

addition, when the battery reaches its lower limits and switches to charge mode, there is no supply to the consumer.

In Scenario 2.2, the link from the generator to the battery g_b is removed, i.e. the battery can no longer be charged. The battery can provide part of the demanded power in the beginning and when the battery reaches its lower limit at $t = 20$ hour, the renewable generator becomes the only source of supply to the consumer. NSD is much higher than that of Scenario 2.0. The consumer's demand can be satisfied only when there is enough wind power. During the period of high wind speed (the power generated can reach 6KW, which is much higher than the largest demand), the redundant power generated can not be stored. The capacity efficiency index EI^{st} is the same as Scenario 2.0, since the failed link does not affect the supply to the consumer; however, without it, the battery does not have income any more, which decreases the supply.

In Scenario 2.3, the link from the battery to the consumer b_c is failed, and the renewable generator is the only source to supply the consumer for the whole period. Under this scenario, the non-supplied demand NSD is similar to Scenario 2.2; but, without the contribution of the energy stored in the battery, NSD is slightly higher. In addition, the capacity efficiency is the lowest for Scenario 2.3, since there is only one source-terminal path left, i.e. the link from the generator to the consumer.

In the stand-alone mode, we have $NSD > 0$ for all the scenarios considered. But, because

of the integration of the renewable generator, a large part of the demand can be supplied. Furthermore, the importance of the storage device is shown through the comparison of the Scenarios 2.0, 2.2 and 2.3.

5 Conclusion

In this paper, we have adopted model predictive control for describing microgrid dynamics and analyzed system performance under grid-connected and stand-alone modes, for different failure scenarios. This analysis enables quantitative evaluation of microgrid performance with respect to different perspectives of reliability, controllability and topology.

We have considered a specific case study, which confirms the fact that the microgrid be connected to the external power grid is important to insure supply to the consumer under different failure scenarios and that the introduction of microgrids composed of renewable generators and storage devices improves reliable performance of the power grid. The instability of the wind power and the limited capacity of the battery or links can be a barrier to the reliable service of the microgrid, especially when in stand-alone mode.

The findings of analyses of this kind provide information for the design and operation of the microgrid, seeking the right balance of multiple characteristics and least cost for the safe and reliable functioning of the microgrid system.

References

[1] Wolfgang Kröger and Enrico Zio. *Vulnerable systems.* Springer Science & Business Media, 2011.

[2] William B. Rouse. Engineering complex systems: Implications for research in systems engineering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2(33):154–156, 2003.

[3] Julio M Ottino. Engineering complex systems. *Nature*, 427(6973):399–399, 2004.

[4] Enrico Zio. Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering & System Safety*, 152:137–150, 2016.

[5] E. Zio. Reliability analysis of systems of systems. *IEEE Reliability Magazine*, pages 1–6, February 2016.

[6] Enrico Zio. Critical infrastructures vulnerability and risk analysis. *European Journal for Security Research*, (DOI 10.1007/s41125-016-0004-2):1–18, 2016.

[7] Enrico Zio and Lucia R Golea. Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. *Reliability Engineering & System Safety*, 101:67–74, 2012.

[8] Jian Li, Leonardo Dueñas-Osorio, and Changkun Chen. Reliability and controllability of infrastructure networks: do they match? 2015.

[9] Zhang Limiao, Li Daqing, Qin Pengju, Fu Bowen, Jiang Yinan, Enrico Zio, and Kang Rui. Reliability analysis of interdependent lattices. *Physica A: Statistical Mechanics and its Applications*, 452:120–125, 2016.

[10] Robert H Lasseter and Paolo Paigi. Microgrid: a conceptual solution. In *Power Electronics Specialists Conference, 2004. PESC 04. 2004 IEEE 35th Annual*, volume 6, pages 4285–4290. IEEE, 2004.

[11] F Han, Ionela Prodan, and E Zio. A framework of model predictive control for the safety analysis of an electric power microgrid. In *ESREL 2015, 25th European Safety and Reliability Conference*, 2015.

[12] Ramon Zamora and Anurag K Srivastava. Controls for microgrids with storage: Review, challenges, and research needs. *Renewable and Sustainable Energy Reviews*, 14(7):2009–2018, 2010.

[13] J. Jimeno, J. Anduaga, J. Oyarzabal, and A.G. de Muro. Architecture of a microgrid energy management system. *European Transactions on Electrical Power*, 21(2):1142–1158, 2011.

[14] E. Kuznetsova, K. Culver, and E. Zio. Complexity and vulnerability of smartgrid systems. In *Proceedings of the European Safety and Reliability Conference*, pages 2474–2482, 2011.

[15] T. Krause, E.V. Beck, R. Cherkaoui, A. Germond, G. Andersson, and D. Ernst. A comparison of nash equilibria analysis and agent-based modelling for power markets. *International Journal of Electrical Power & Energy Systems*, 28(9):599–607, 2006.

[16] A. Weidlich and D. Veit. A critical survey of agent-based wholesale electricity market models. *Energy Economics*, 30(4):1728–1759, 2008.

[17] Ionela Prodan and Enrico Zio. A model predictive control framework for reliable microgrid energy management. *International Journal of Electrical Power & Energy Systems*, 61:399–409, 2014.

[18] Ionela Prodan and Enrico Zio. On the microgrid energy management under a predictive control framework. In *Control Applications (CCA), 2014 IEEE Conference on*, pages 861–866. IEEE, 2014.

[19] Daniel E Olivares, Ali Mehrizi-Sani, Amir H Etemadi, Claudio A Cañizares, Reza Iravani, Mehrdad Kazerani, Amir H Hajimiragha, Oriol Gomis-Bellmunt, Maryam Saeedifard, Rodrigo Palma-Behnke, et al. Trends in microgrid control. *IEEE Transactions on smart grid*, 5(4):1905–1919, 2014.

[20] Shouxiang Wang, Zhixin Li, Lei Wu, Mohammad Shahidehpour, and Zuyi Li. New metrics for assessing the reliability and economics of microgrids in distribution system. *IEEE Transactions on Power Systems*, 28(3):2852–2861, 2013.

[21] J.B. Rawlings and D.Q. Mayne. Postface to model predictive control: Theory and design, 2011.

[22] J. Richalet and D. O'Donovan. *Predictive Functional Control: Principles and Industrial Applications*. Springer, 2009.

[23] G.C. Goodwin, M. Seron, and J. De Dona. *Constrained control and estimation: an optimisation approach*. Springer Verlag, 2005.

[24] T.G. Hovgaard, L.F. Larsen, and J.B. Jorgensen. Robust economic mpc for a power management scenario with uncertainties. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 1515–1520. IEEE, 2011.

[25] R. Halvgaard, N. K. Poulsen, H. Madsen, and J.B. Jorgensen. Economic model predictive control for building climate control in a smart grid. In *IEEE PES Innovative Smart Grid Technologies (ISGT)*, pages 1–6. IEEE, 2012.

[26] K. Edlund, J.D. Bendtsen, and J.B. Jørgensen. Hierarchical model-based predictive control of a power plant portfolio. *Control Engineering Practice*, 19(10):1126–1136, 2011.

[27] R.R. Negenborn, B. De Schutter, and H. Hellendoorn. Multi-agent model predictive control of transportation networks. In *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pages 296–301. IEEE, 2006.

[28] Yi-Ping Fang and Enrico Zio. Hierarchical modeling by recursive unsupervised spectral clustering and network extended importance measures to analyze the reliability characteristics of complex network systems. *American Journal of Operations Research*, 3(1A):101–112, 2013.

[29] Anna Lombardi and Michael Hörnquist. Controllability analysis of networks. *Physical Review E*, 75(5):056110, 2007.

[30] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[31] Efstathios Bakolas and Joseph H Saleh. Augmenting defense-in-depth with the concepts of observability and diagnosability from control theory and discrete event systems. *Reliability Engineering & System Safety*, 96(1):184–193, 2011.

[32] Rudolf Emil Kalman. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial & Applied Mathematics, Series A: Control*, 1(2):152–192, 1963.

[33] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, and S. Kuruganty. The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *Power Systems, IEEE Transactions on*, 14(3):1010–1020, 1999.

[34] CG Justus, WR Hargraves, and Ali Yalcin. Nationwide assessment of potential output from wind-powered generators. *Journal of Applied Meteorology*, 15(7):673–678, 1976.

[35] Ionela Prodan, Enrico Zio, and Florin Stoican. Fault tolerant predictive control design for reliable microgrid energy management under uncertainties. *Energy*, 91:20–34, 2015.

Paper II

F. Han, E. Zio. A multi-perspective framework of analysis of critical infrastructures with respect to supply service, controllability and topology International Journal of Critical Infrastructure Protection, 2017. (Under review).

A multi-perspective framework of analysis of critical infrastructures with respect to supply service, controllability and topology

Fangyuan Han[1] and Enrico Zio[1,2]

[1]Chair on Systems Science and the Energetic Challenge, Fondation EDF, CentraleSupelec, France

[2]Department of Energy, Politecnico di Milano, Italy.

Abstract

In this work, we propose a multi-perspective framework of analysis of critical infrastructures (CIs) with respect to supply service, topology and controllability. The framework enables identifying the role of CI elements and quantifying the consequences of scenarios of multiple failures, with respect to the different perspectives considered. To present the analysis framework, a benchmark network representative of a real gas transmission network across several countries of the European Union (EU) is considered. The information extracted from such analysis can help us to identify the critical elements and how the properties of the network are affected by failures, and to propose corresponding improvements for CIs. The findings of this paper demonstrate the interest of considering several perspectives in the analysis of CIs for providing useful information for ensuring their safe and reliable operation.

Keywords: Critical infrastructures, Multi-perspective analysis, Complex networks, Supply, Controllability, Gas transmission network.

List of Symbols

$\mathbf{A}$	Coupling matrix
$\mathbf{B}$	$N \times M$ input matrix
$\mathbf{Ctrb}$	Controllability matrix
$\mathbf{G}$	Graph representing the network
$\mathbf{G}(base)$	Graph representing the original network
$\mathbf{G}'(x_{ij} = 1)$	Graph obtained by removing the node i from $\mathbf{G}(base)$
$\mathbf{K}$	Capacity matrix of the network
$\mathbf{Len}$	Length matrix of the network
$\mathbf{u}(\mathbf{t})$	Input vector of M independent control signals at time t
$\mathbf{x}(\mathbf{t})$	System state vector at time t
$\mu(\lambda_i)$	Geometric multiplicity of the eigenvalue λ_i of matrix $\mathbf{A}$
ω_i	Weight of the ith user
C_{ind}	Controllability index
D_i	Demand of the ith user
$IM_{ij}^{C_{ind}}$	Link importance with respect to C_{ind}
IM_{ij}^{E}	Link importance with respect to E
IM_{ij}^{NSD}	Link importance with respect to NSD
L	Number of links in the network
N	Number of nodes in the network
N_D	Minimum number of driver nodes
N_s	Number of sources in the network
N_y	Number of user nodes within the network
NSD	Normalized non-supplied demand

y_i Supply to the ith user

List of Acronyms

CDF Cumulative distribution function

CI Critical infrastructure

EU European Union

mcm million of cubic meters

NSD Non-supplied demand

RA Risk Achievement

1 Introduction

Critical infrastructures (CIs), like power grids or gas transmission and distribution systems, rail and road transport or communication networks, are essential to the operation of modern society [1]. They need to be designed, maintained and protected to provide optimal performance, reliable operation and functional safety for long periods of time [2, 3]. Hence, the great attention and priority given to the "care" of these systems by the EU, US and other national and transnational administrations [4–7], which calls for the risk assessment and resilience evaluation of CIs [8, 9].

As CIs evolve and rely on information technologies more intensively, it is essential to understand their controllability and it is desirable to develop a control framework able to steer the network dynamics toward states with optimal performance, while avoiding undesired or unfavorable states. Control is a fundamental property for the safe and reliable operation of CIs, under a general control perspective, system safety can be framed as a control "problem" [10, 11], whereby, accidents result from inadequate control or insufficient enforcement of safety-related constraints on the development, design, and operation of the system, leading to their violation and subsequently to accidents. According to Control Theory, a dynamical system is controllable if, by a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time [12,13]. However, the control of the complex network systems that make up a CI remains a challenging problem. Studying the controllability of complex networks requires an integration of classical control theory and network theory. In this perspective, the notion of structural controllability

has been introduced in [14]. In [13], analytical tools have been developed to characterize the controllability of directed networks. In [15], an exact controllability measure has been proposed to generalize the determination of the set of driver nodes to arbitrary network structures and link weights. Several related topics can be considered under this framework, such as control centrality [16], achieving whole control by using only one controller [17], minimization of control inputs [18], control capacity [19], control mode [20], control of edge dynamics [21], structural controllability of temporal networks [22], control energy [23], etc.

Supply performance is the fundamental functional requirement of a CI and the security of supply is being addressed by an increasing number of researchers. Paper [24] presents a probabilistic model to study the security of supply in a gas network. The model is developed into a Monte Carlo simulation and graph-based tool aiming at the evaluation of CIs for different purposes, including reliability, vulnerability, bottleneck analysis, etc.

The fact that CIs are complex networks of interacting components raises the interest in studying their topological characteristics [25–29]. A number of recent studies have proposed various measures to evaluate the structural properties of networks and addressed topological investigations to identify critical elements. Among these measures, topological centrality (including degree, closeness, betweenness and information centrality) [30, 31] and network efficiency [32] are two important and classical measures, quantifying the importance of individual network nodes and evaluating the connectivity of the whole network, respectively. Topological properties have also been studied in relation to vulnerability and risk analysis. For example, in [33] the authors analyze the structural vulnerability of the Italian GRTN power grid. In [34], electric power delivery networks are modelled as graphs and their topological characteristics are studied. In [35], centrality analysis is applied to identify the most important components of a railway infrastructure. Given the relationship between the topology of a network and its vulnerability and safety properties, the association between network topological features and system reliability is also of relevance. A common measure of network reliability is the so called K-terminal reliability, which calculates the probability that every two nodes in a specific subset of K nodes are connected by a path of operational edges [36]. Specifically, due to the requirement of reliable operation and the complex nature of CIs, the all-terminal reliability is of particular interest and often considered as a necessary condition for function-based reliability. In various applications, it is used as constraint or objective for the design optimization and operation management of CIs [37–41].

The complexity of CIs calls for approaches capable of viewing the problem from multiple perspectives [42–44]. System analysis, reliability engineering, graph theory and others have

been propounded to study the behavior and performance of complex systems, also with respect to failure events, their protection and resilience [27, 45, 46]. Integration of the different perspectives and analysis of their relations is necessary. For example, in [47] an electrical transmission system is analyzed with the objective of identifying the most critical elements in terms of four different perspectives: topological, reliability, electrical and electrical-reliability. In [48], the correlation between connectivity reliability and controllability of network systems is studied. In [46], the authors perform network reliability analysis considering spatial constraints. The authors of [49] consider a three-objective optimization of economic cost, hydraulic reliability and greenhouse gas emissions, and the nature of the tradeoffs among the objectives is also studied.

In this work, we develop a framework of analysis considering several perspectives (supply service, controllability and topology). Compared to previous works, which typically consider reliability and topology only, we include the control perspective into the safety and reliability analysis of CIs. The analysis is performed by simulation and the failure scenarios are generated by the software ProGasNetwork proposed in [50]. The complex network representative of a real EU gas transmission system supplying several countries is considered as case study to illustrate the analysis framework. The main contributions of this work are:

- Development of a multi-objective framework of analysis of CIs.
- Identification of the role of each component and quantification of the consequences of multiple failure scenarios, with respect to the different perspectives considered.
- Proposals for CIs reliable performance improvement.

The rest of the paper is organized as follows: Section 2 introduces the three system-level indexes considered in this paper. Section 3 describes the modelling of the considered gas transmission network. Section 4 presents the analysis of link importance and Section 5 presents the consequences of the failure scenarios from the three perspectives. Finally, conclusions and ideas for future work are provided in Section 6.

2 System-level indexes

In this work, we consider three perspectives of the CI assessment: supply, control and topology. For each perspective, we propose an index to evaluate the network performance.

2.1 Non-supplied demand

Supply performance is the fundamental functional requirement of a CI. Consider a CI network of N nodes and L links, which supplies service or products from N_s supply nodes (sources) to N_y user nodes (users) through a number of transmission nodes.

We introduce the non-supplied demand (NSD) as a measure of the network's capacity to satisfy its users' demands. The normalized NSD is introduced as a system-level index:

$$NSD = 1 - \frac{\sum_{i=1}^{i=N_y} \omega_i y_i}{\sum_{i=1}^{i=N_y} \omega_i D_i} \tag{1}$$

where, ω_i is the weight of the ith of the N_y users, y_i is the supply to user i and D_i is its demand, which is considered as the target supply to user i. Then, the second term in Equation (1) represents the satisfied proportion of users' demand. Since $y_i \leq D_i$, the index NSD is normalized to take values in $[0, 1]$. NSD equals 0 when the users' demands are fully supplied.

2.2 Controllability

In control theory, a system is defined controllable if, by a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time [12, 13]. From a system safety perspective, controllability is the ability to guide the system's behavior towards a safe state through the appropriate choice of a few input variables [10].

Considering the network of N nodes, we describe its state dynamics by:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{2}$$

where $\mathbf{x}(\mathbf{t})$ is the system state vector, describing the state of each node in the network at time t; $\mathbf{A}$ is an $N \times N$ coupling matrix, in which a_{ij} represents the weight of the directed link from node i to node j (i.e. the interaction strength between node i and node j, for example, the flow in the pipeline of a gas transmission network); $\mathbf{B}$ is an $N \times M$ input matrix ($M \leq N$), identifying the nodes that are controlled by the time-dependent input vector $\mathbf{u}(\mathbf{t})$, made of M independent control signals.

Based on dynamic control theory, the above system is controllable if and only if the $N \times NM$ controllability matrix $\mathbf{Ctrb} = (\mathbf{B} \quad \mathbf{AB} \ldots \mathbf{A}^{n-1}\mathbf{B})$ has full rank (the so-called Kalman's rank condition) [51]:

$$rank(\mathbf{Ctrb}) = N \tag{3}$$

For complex network systems, the controllability problem can be formulated in terms of finding a suitable control matrix $\mathbf{B}$ consisting of a minimum number of driver nodes (N_D) so as to satisfy the Kalman's rank condition (3). However, this requires the evaluation of the rank of $\mathbf{C}$ for 2^N possible combinations of the driver nodes [52]: for real CI network systems, such a brute-force search is computationally prohibitive.

To overcome this problem, in [13], the authors have developed analytical methods to determine the minimum number of driver nodes (N_D) that are needed to fully control the network, by finding the maximum matching, i.e. the maximum set of links that do not share start or end nodes. Full control can be achieved if and only if each unmatched node is directly controlled and there are directed paths from the input signals to all matched nodes. The unmatched nodes determined by maximum matching are the so called driver nodes.

In [15], the exact controllability for arbitrary network structures and link weights (say arbitrary matrix $\mathbf{A}$) is introduced to calculate N_D:

$$N_D = \max_i \{\mu(\lambda_i)\} \tag{4}$$

and the minimum number of driver nodes N_D is determined by the maximum geometric multiplicity $\mu(\lambda_i)$ of the eigenvalue λ_i of $\mathbf{A}$. In fact, these are the nodes corresponding to the linearly-dependent rows: the controllers should be imposed on the linearly-dependent rows to eliminate all linear correlations and ensure the controllability condition.

The number of driver nodes N_D can be taken as a measure of the controllability of the network, indicating how many driver nodes are needed to control the network and directly relating to the cost of the resources needed to keep or bring the system under control. If $N_D = N$, i.e. the total number of nodes in the network (this means that the external control signal is applied to each node of the network), the likelihood of gaining full system control is high, but so is the associated cost [18]. A small N_D, instead, indicates a more controllable network system, in the sense that it requires less effort to obtain full control over the network.

To measure the structural controllability of the network system, we adopt the controllability index (C_{ind}) first introduced in [48]:

$$C_{ind} = \frac{N - N_D}{N} \tag{5}$$

Also, the index C_{ind} is normalized to take values in $[0, 1]$.

The occurrence of failures (represented as the removal of links) is likely to increase the number of the linearly-dependent rows in matrix $\mathbf{A}$ and, thus, N_D would increase and C_{ind} decreases; when the current number of control nodes is insufficient to obtain full control

over the whole system, there is no guarantee that the system can be brought back to the designed operation condition.

2.3 Network topological efficiency

Network topological efficiency is a measure of the connectivity of the whole network, i.e. of how well the nodes of a network exchange flow [32]. This measure is based on the assumption that the flow in a network travels along the shortest routes and that the efficiency in the communication between two nodes i and j, ε_{ij}, is inversely proportional to their shortest path length d_{ij}; this latter is defined as the smallest sum of *physical distances* throughout all the possible paths in the weighted network: $\varepsilon_{ij} = 1/d_{ij}$. When there is no path between i and j, $d_{ij} = +\infty$, i.e $\varepsilon_{ij} = 0$. Then, the topological efficiency of the whole network is given by:

$$E[\mathbf{G}] = \frac{\sum_{i \neq j \in \mathbf{G}} \varepsilon_{ij}}{N(N-1)} = \frac{1}{N(N-1)} \sum_{i \neq j \in \mathbf{G}} \frac{1}{d_{ij}} \tag{6}$$

3 Case study

3.1 Network description and graph representation

We consider the case study from [50]. The system is visualized in Figure 1 and represents a real gas transmission network for supply across several countries in the EU. The gas transmission network includes 56 nodes and 74 links, where nodes represent sources or substations and links represent the gas transmission pipelines connecting the nodes. Among the 74 links, 10 links are virtual links representing the virtual connection of parallel pipelines, and their failure is not considered.

The gas transmission network is modeled as an undirected graph $\mathbf{G}$. Its connectivity structure can be defined by its $N \times N$ adjacency matrix $\mathbf{Adj}$, whose entries $[Adj_{ij}]$ are equal to 1 if there exists a link joining node i to node j and 0 otherwise.

Each link in the network is characterized by its capacity, i.e the maximum amount of flow that it is able to supply, and its length. The capacity matrix $\mathbf{K}$ contains information about the capacity constraints of the network elements including source nodes, demand nodes and pipeline capacities. The length matrix $\mathbf{Len}$ contains the lengths of the edges between nodes: entry Len_{ij} is the length of the pipeline connecting the i-th and j-th nodes; an entry of 0 indicates that the i-th and j-th nodes are not connected.

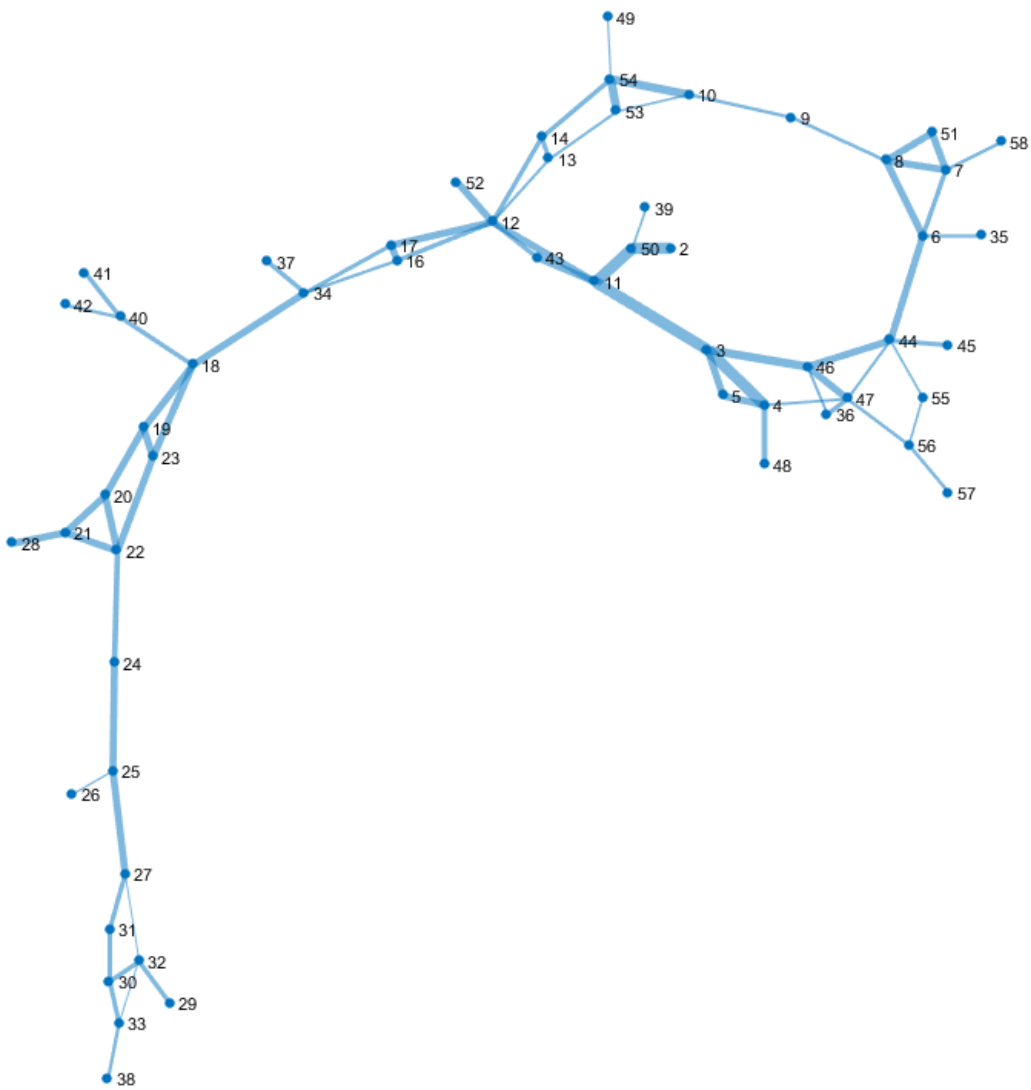


Figure 1: Gas transmission network [50]

We distinguish: $N_y = 35$ demand nodes with deterministic daily demands for a total system daily demand of 45.9 millions of cubic meters (mcm), one LNG terminal (node 10), two compressor stations (nodes 11 and 12), two storage devices (nodes 10 and 19) and two pipeline source nodes (nodes 2 and 29).

The properties of the four nodes considered as supply nodes (sources), numbered 2, 10, 19 and 29, are presented in Table 1. The properties of the 35 demand nodes are shown in Table 2. The capacities and demands are expressed in millions of cubic meters per day (mcm/d). The data of supply and demand are realistic and they are expressed at a daily scale, in order to assume peak gas demand during one peak day (in winter) with extreme high gas demand [24]. These data are intended to represent the most stressed conditions for the gas transmission network. Depending on the purpose of the analysis, variable values

of demand or supply can be considered, and the user demands satisfaction can be evaluated as the average over a simulation horizon, with associated uncertainty.

Given the capacities of the links connecting the nodes and the constraints on the sources and users, the supply to each user is used to calculate the non-supplied demand NSD defined in Equation 1 by maximum flow algorithms [24].

Node	Capacity	Type
2	31	Pipeline source
10	10.5	LNG terminal
19	25	Underground storage
29	4	Pipeline source

Table 1: Sources properties

Node	Demand	Node	Demand	Node	Demand
4	0.1	27	3.0	44	0.7
5	3.2	28	6.0	45	1.3
6	0.1	30	0.5	47	0.1
8	0.1	33	0.5	48	1.8
9	0.1	34	0.5	49	0.2
10	1.0	35	0.1	51	7.0
13	0.5	36	4.2	52	0.6
17	0.1	37	1.3	53	0.1
18	8.5	39	0.3	55	0.2
20	0.6	41	0.6	57	0.2
25	0.5	42	0.6	58	0.3
26	0.8	43	0.2		

Table 2: Demands of the 35 users

3.2 Failure modelling

We consider the failure of the LNG station, compressor stations, storages and 64 pipelines.

The failure of the LNG terminal and of the storage devices is modeled as the total capacity loss of each pipeline connected to it. According to [50], the monthly failure frequency of the LNG (node 10) is set to $f_{LNG} = 1.25E-2$, and the monthly failure frequency of the storage (node 19) is $f_S = 8.33E-3$.

In case of a compressor station failure, the capacity of each pipeline connected to it will reduce by 20%. The monthly failure frequency of the two compressor stations (nodes 11 and 12) are $f_{CS} = 2.08E - 2$.

According to the EGIG report [53], the average failure frequency of a European gas transmission pipeline is $3.5E - 4$ per kilometer-year. We consider the total rupture of a pipeline and we assume that 10% of the failures reported in a year cause such a rupture. Thus, the monthly failure frequency of a pipeline is $f_P = 2.92E - 6$ per kilometer [50].

4 Analysis of link importance

We focus on the importance of a link in terms of its influence on the three system properties considered.

NSD	C_{ind}	E
0	0.9107	0.6327

Table 3: Indexes values for the nominal configuration

Table 3 presents the values of the three indexes introduced in Section 2, calculated for the nominal network configuration $\mathbf{G}$. For the analysis of link importance, we systematically disconnect one link at a time from the original network to obtain and compute the indexes of the new network configuration reached, $\mathbf{G}'$. We identify the most important nodes in terms of NSD, C_{ind} and E, respectively. Table 4 presents the ten most critical links in terms of NSD, the three most critical links in terms of C_{ind}, and the single most critical link in terms of E.

With the removal of single links, the NSD value ranges from 0 to 0.363. Pipelines represented by links 3-11 and 3-46 are of large capacity, so they are essential to supplying the demand nodes in their neighborhood, and thus their importances are significant in terms of supply. A similar explanation applies for the removal of links 6-44 and 44-46.

Node 28 is a large demand node, and therefore, the removal of link 21-28, which is its only connection to the network, will affect the overall network NSD. The same explanation also applies for the impact of links 4-48, 34-37 and 44-45. Link 4-48 and link 44-45 are critical in terms of efficiency E, because their removal disconnects the end nodes 48 and 45, respectively and, thus, decreases the network efficiency; moreover, considering that they are relatively short pipelines, the value of E drops much more than for the removal of links

Link	NSD	C_{ind}	E
3-11	0.363	0.9107	0.6319
3-46	0.209	0.9107	0.6327
21-28	0.131	0.8929	0.6327
2-50	0.126	0.9107	0.6325
11-50	0.120	0.9107	0.6323
6-44	0.106	0.9107	0.6321
44-46	0.081	0.9107	0.6327
36-47	0.048	0.8929	0.6327
4-48	0.039	0.8929	0.6318
34-37	0.028	0.8929	0.6327
44-45	0.028	0.9107	0.6318
18-40	0.026	0.9107	0.6325
6-35	0.002	0.8750	0.6325
11-43	0	0.8750	0.6326
18-23	0	0.8750	0.6327
18-34	0	0.9107	0.6318

Table 4: Indexes values associated to the removal of the most critical links

34-37 and 21-28, which are long pipelines.

Links 2-50 and 11-50 connect the main source (node 2) to the rest of the network, and their removal leads to a deficit in supply capacity, since the remaining sources 10, 19 and 29 are not capable of fully supplying the total demands.

Link 18-34 is also a critical link in terms of topological efficiency: when it fails, the network will break into two separate parts and no gas flow can be transferred between them, so that, the topological efficiency E drops.

To rigorously quantify the importance of a link, we compute its Risk Achievement (RA) metric [54] with respect to NSD, C_{ind} and E:

$$IM_{ij}^{NSD} = NSD[\mathbf{G}(base)] - NSD[\mathbf{G}'(x_{ij} = 1)] \tag{7}$$

$$\begin{aligned} IM_{ij}^{C_{ind}} &= C_{ind}[\mathbf{G}(base)] - C_{ind}[\mathbf{G}'(x_{ij} = 1)] \\ &= \frac{N_D[\mathbf{G}'(x_i = 1)] - N_D[\mathbf{G}(base)]}{N} \end{aligned} \tag{8}$$

$$IM_{ij}^{E} = E[\mathbf{G}(base)] - E[\mathbf{G}'(x_{ij} = 1)] \tag{9}$$

where $\mathbf{G}'(x_{ij} = 1)$ is the graph of the network obtained by removing the link $i - j$ from the original network $\mathbf{G}(base)$.

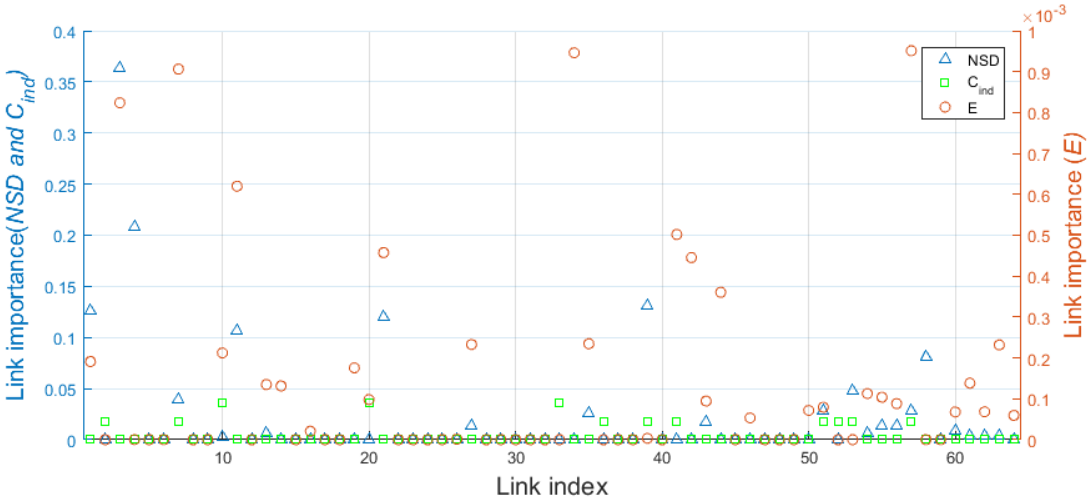


Figure 2: Link importance in terms of NSD, C_{ind} and E

Figure 2 shows the links importance values in terms of the three indexes: the left vertical axis is the values for NSD (triangles) and C_{ind} (squares), while the right vertical axis is the topological efficiency E (circles). It is seen that NSD presents the largest range, while a single link disconnection has little impact on controllability and E has the smallest range.

Among the 64 links, only 23 links have impact on NSD and most of them also have an impact on E, but only five links influence the controllability index.

The influence of a link is not the same for the three perspectives, which confirms the need of a multi-perspective framework of analysis.

5 Simulation and analysis

We have run 10^6 dynamic simulations by ProGasNet [50], sampling nodes and links according to their occurrence probabilities, as introduced in Section 3.2. A total of 335 different gas transmission states (cases) emerge from the sampled configurations. The most frequent state sample is the one with no link failures.

We classify the 334 failure cases into different categories by their combination of failures. We consider the thirty most frequent states and investigate how these affect the three system-level indexes considered. For each of the indexes, we quantify their consequences and analyze the impacts of different types of failures.

5.1 Classification by failure types

Both links and nodes of the gas transmission network may fail and multiple failures may occur. In order to understand the influence of different types of failures and of their combinations, we classify the 334 failure cases into seven classes as:

- Single link failure (SL)
- Single node failure (SN)
- Single link failure and single node failure (SL-SN)
- Single link failure and multiple node failures (SL-MN)
- Multiple link failures (ML)
- Multiple link failures and single node failure (ML-SN)
- Multiple node failures (MN)

Single node failure (SN) includes 4 cases (only four nodes may fail according to our definition), but they cover 83.23% of the failure configurations. Single link failure (SL) includes 61 cases and is the second most frequent class. The cases of SL-MN, ML and ML-SN only occur once in all simulations performed. Figure 3 shows the number of cases for each class and their counts.

As we are analyzing an existing gas transmission network, it is reasonable to have low probabilities for multiple failures scenarios; however, the failures of low frequency of occurrence may have an important impact on the properties of the system, which is one of our interests in this study. Therefore, although the probability of their occurrence is small, it is interesting to consider such multiple-element failures and understand the corresponding consequences, which provides additional information for CIs design.

5.2 30 most frequent cases

We consider the 30 most frequent cases and apply the analysis framework. Table 5 summarizes the failure types and frequencies of the 30 most frequent cases. Node failure is the most common failure type, the four most frequent failure cases being the four single node failure (SN) cases.

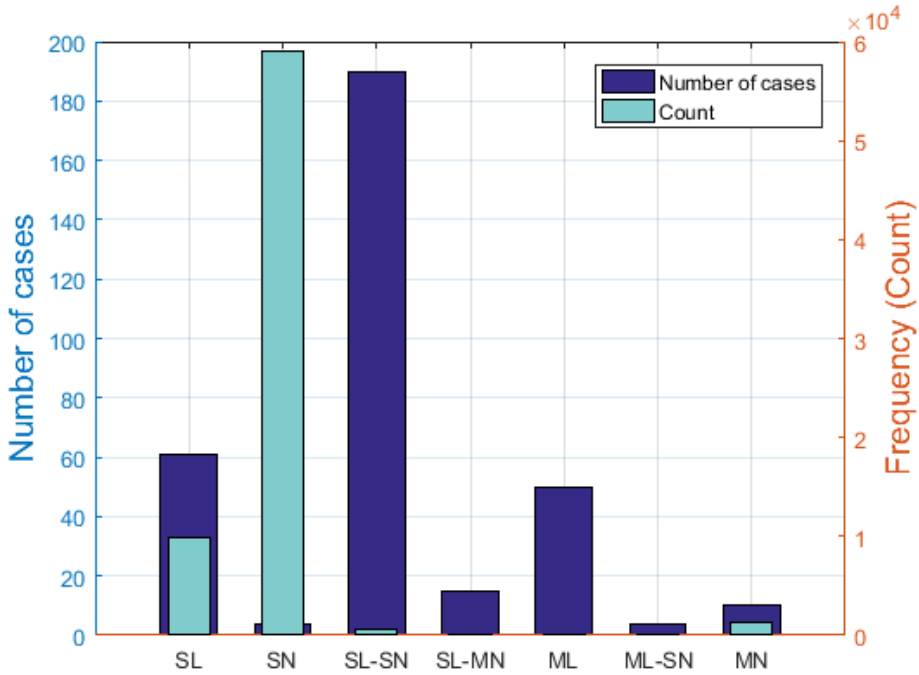


Figure 3: Histogram: number of cases and frequency

Type	Cases	Frequency (over 10^6 simulations)
Failure free	1	929 013
SN	4	59 040
SL	21	6 238
MN	4	1 098

Table 5: 30 most frequent cases

5.3 The three indexes

We analyze the three indexes separately, with the objectives of identifying the failures affecting each index, quantifying their consequences in terms of loss in the properties considered and calculating their frequency.

5.3.1 Non-supplied demand

Figure 4 shows the NSD index value for the 335 cases, where the abscissa axis is the frequency rank of the 335 cases. The non-supplied demand NSD ranges from 0 to 0.64. In the presence of multiple failures, the network may reach a much higher level of non-supplied demand NSD. The highest value 0.64 corresponds to the SN-SL case where both node 19

and link 2-50 are failed. Node 19 is the second largest source and its failure alone results in $NSD = 0.2261$, since without it, some demand nodes far from the main source (e.g. node 2) are not fully supplied due to the limited capacity of pipelines connecting different areas (e.g. link 18-34), even though the total supply capacity of the sources is able to cover all demands. Combining with the failure of link 2-50, which cuts the supply from the main source (node 2) to other nodes, the supply of the whole network drops even more.

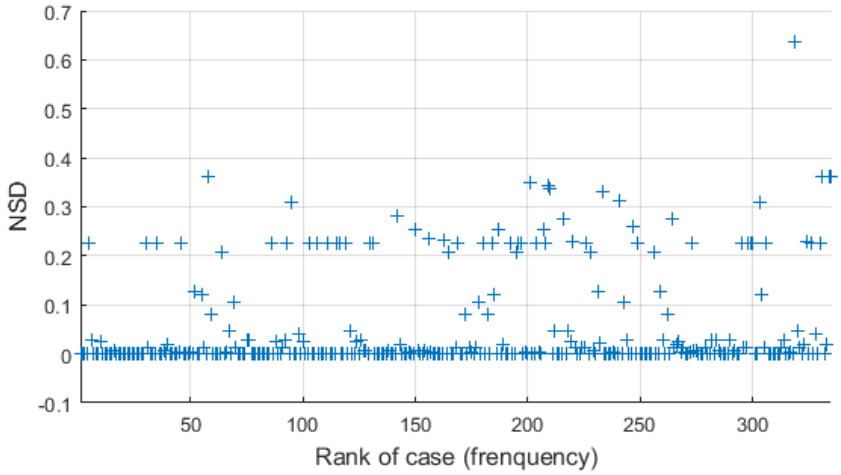


Figure 4: Non-supplied demand for the 335 cases

$NSD = 0$ is the most frequent value. The case ranked 5 is the failure of node 19 and has $NSD = 0.2261$. It occurs 7 888 times out of the one million simulations. Generally speaking, high values of NSD ($NSD > 0.3$) tend to have low frequency. For 146 out of 335 cases (43.6%), the demand can not be fully supplied, i.e. $NSD > 0$.

Figure 5 shows the cumulative distribution function (CDF) of NSD for the failure configurations. The mean value of NSD is 0.0285 over the 70 987 configurations with failures and 0.0020 over all 10^6 configurations simulated.

Figure 6 shows the NSD of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN, ML, ML-SN, MN).

The NSD for the top 30 cases is higher than that for the nominal configuration and comparable to that of all configurations. This indicates that, the most frequent cases have non-negligible impact on the demand supply.

If we compare the seven failure classes, we see that, as a whole, node failures have a more important impact in terms of non-supplied demand. In presence of node failure (SN, SL-SN and SL-MN), the non-supplied demand NSD is higher than for the cases with single

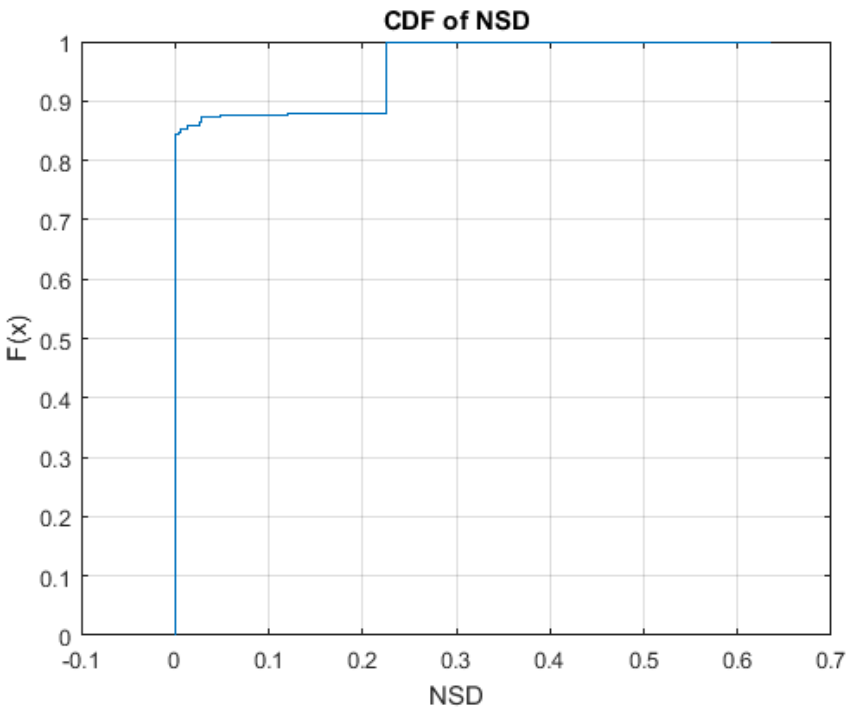


Figure 5: CDF of non-supplied demand for the failure configurations

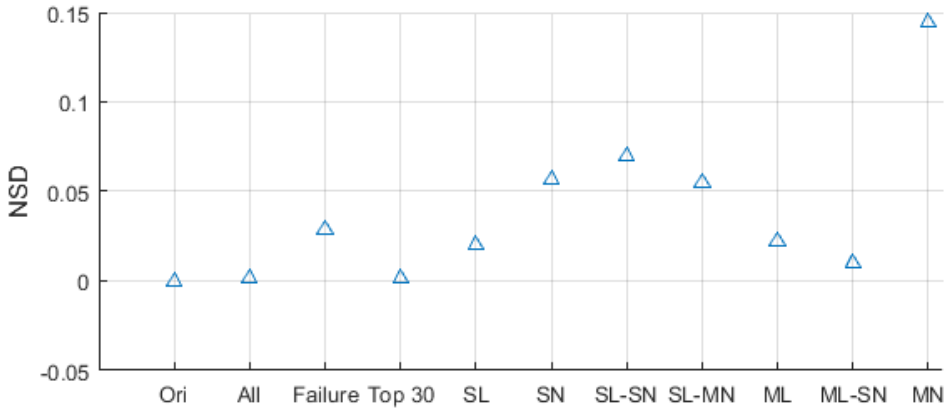


Figure 6: Non-supplied demand for different categories

link failure (SL). The combination of link failures only increases slightly the NSD: the mean NSD of ML is slightly higher than that of SL. As for ML-SN, among all the possible combinations of failures, only four cases occur once each, and they happen to have relatively small influence in terms of NSD, the mean being 0.010. The class of multiple node failures (MN) has a significantly high value of non-supplied demand. All MN cases without failure of node 19 have $NSD = 0$. The failure of node 19 alone would lead to $NSD = 0.2261$, and for the case where node 10, 11 and 19 are all failed at the same time, $NSD = 0.3111$.

In fact, the link 3-11 represents a large capacity pipeline and in the absence of gas supply from the storages represented by nodes 10 and 19, the reduction of its capacity due to the failure of node 11 would result in the non-supply to demand nodes in the vicinity, depending mainly on the main source (node 2). Considering the relatively high failure probability of node 19, the MN class has a high value of NSD.

5.3.2 Controllability index

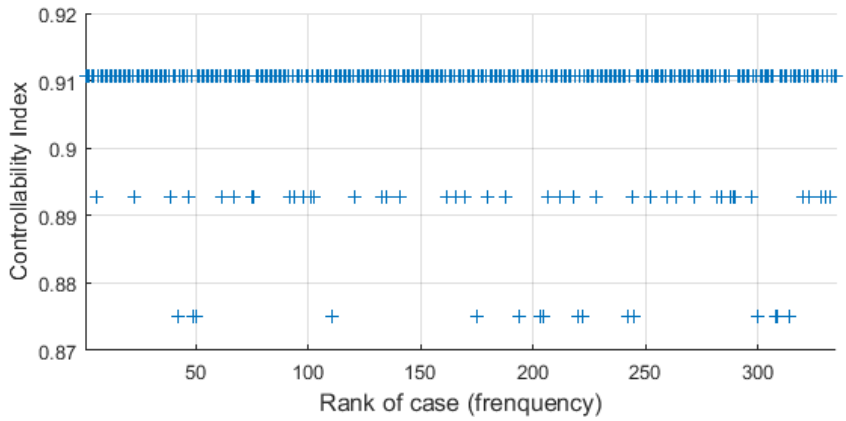


Figure 7: Controllability index for the 335 cases

Figure 7 shows the index value of C_{ind} for the 335 cases and the abscissa axis is the frequency rank of the 335 cases. The controllability index takes three values 0.9107, 0.8928 and 0.875 for the 335 cases, as for the removal of single links. $C_{ind} = 0.9107$ is the most frequent value. For 58 out of 335 cases (17.3%), the controllability index C_{ind} is lower than that of the failure-free network configuration. The lowest value 0.875 is more present for the less frequent cases (i.e. Rank>150).

Figure 8 shows the cumulative distribution function (CDF) of C_{ind}. Only 1 559 out of the 70 987 failure configurations (2%) have C_{ind} lower than 0.9107. The mean value of C_{ind} is 0.9102 over the 70 987 failure configurations and very close to 0.9107 over the 10^6 simulations.

Figure 9 shows the mean value of C_{ind} of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN, ML, ML-SN, MN). The mean of C_{ind} of the top 30 cases is slightly lower than but still close to that of the nominal network configuration.

Node failures have no impact on the controllability index, since they only concern the

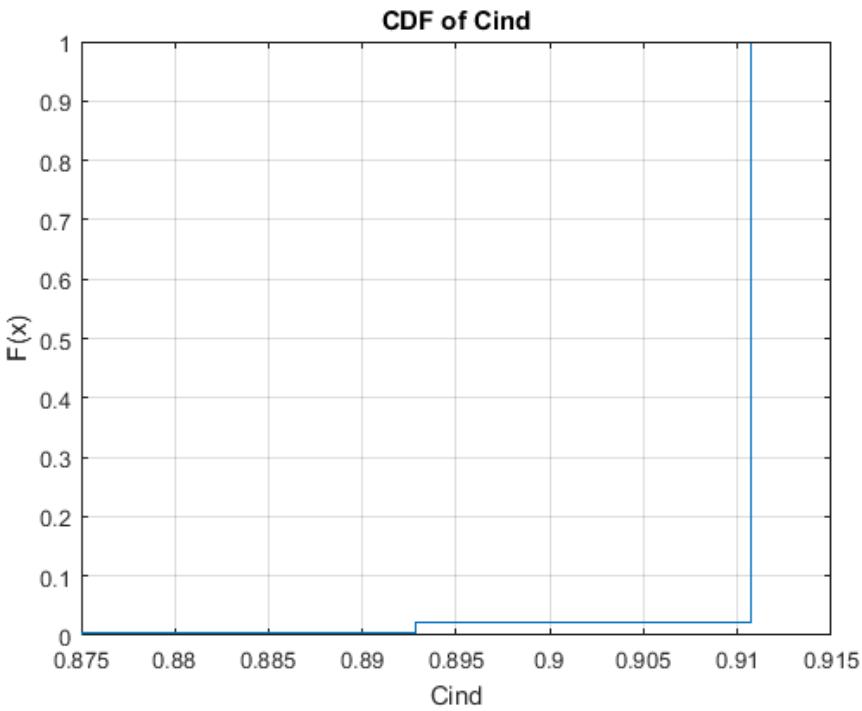


Figure 8: CDF of the controllability index for the failure configurations

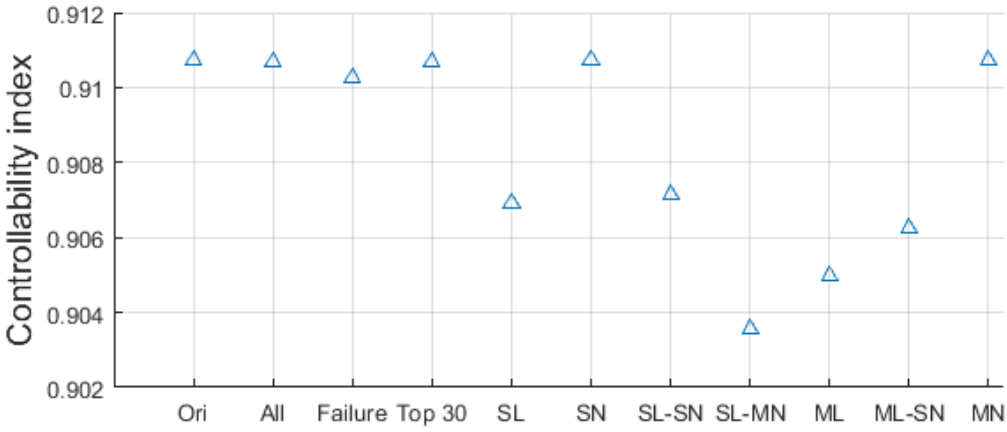


Figure 9: Controllability index for different categories

reduction of pipelines capacity but not their disconnection. Thus, link failures are the only contribution to the loss in controllability. The mean of all cases containing single link failures (i.e. SL, SL-SN and SL-MN) is 0.9069, slightly higher than that of the cases containing multiple link failures (i.e. ML and ML-SN) cases, which equals to 0.9051. This indicates that multiple failures have a more important impact on the controllability index, with C_{ind} reaching values no lower than 0.8750. For the SL, SL-SN and SL-MN cases, this lowest value results from the failures of links 6-35, 11-43 or 18-23, the removal of each of

which decreases C_{ind} to 0.8750. As for ML and ML-SN cases, the lowest value comes from the combination of two link failures with no separate impact on C_{ind} (links 9-10 and 10-53), or the combination of two links whose removal decreases C_{ind} to 0.8929 (links 22-24 and 34-37).

5.3.3 Network topological efficiency

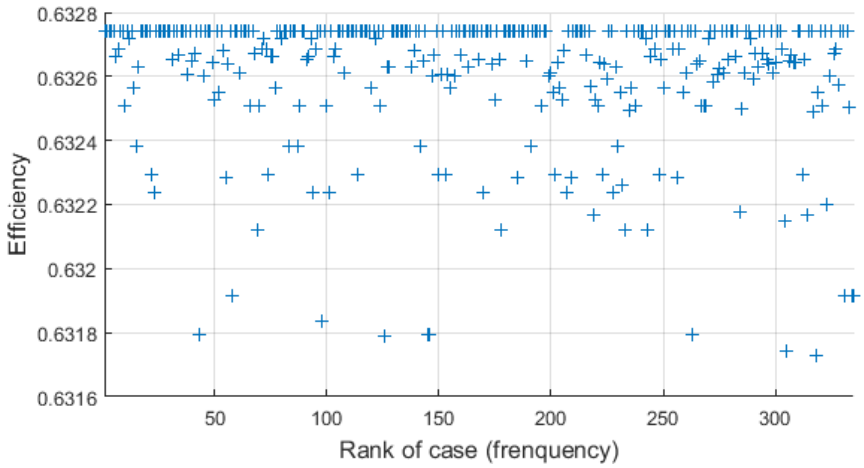


Figure 10: Network topological efficiency for the 335 cases

Figure 10 shows the values of E for the 335 cases, and the abscissa axis is the frequency rank of the 335 cases. The network topological efficiency ranges from 0.6317 to 0.6327. The most critical links are link 44-45 and link 18-34, whose removal decreases the efficiency to 0.6318. Multiple failures decrease the lowest value of E slightly. In fact, all the seven values below 0.6318 are related to the failure of link 44-45 or link 18-34: the five first cases include single link failures alone or together with single node failure, while the last two cases with low values are multiple link failures.

Figure 11 shows the cumulative distribution function of E. The topological efficiency E stays close to the value of the failure-free configuration. However, 10 581 out of the 70 987 failure configurations (14.91%) have E lower than the value of the failure-free configuration, and 10581 is the number of configurations with at least one link failure. This means that failure of any link will influence the efficiency.

Figure 12 shows the mean value of E of the original network (Ori) and the mean values for all sampled configurations (All), the failure configurations only (Failure), the 30 most frequent cases (Top 30) and the seven classes of failures (SL, SN, SL-SN, SL-MN, ML, ML-SN, MN).

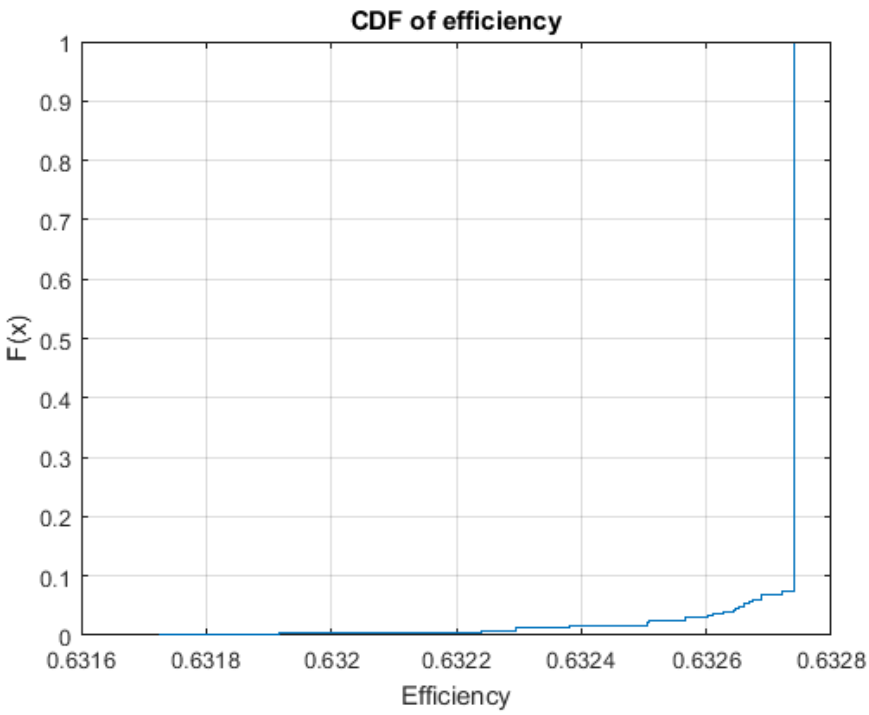


Figure 11: CDF of the network topological efficiency for the failure configurations

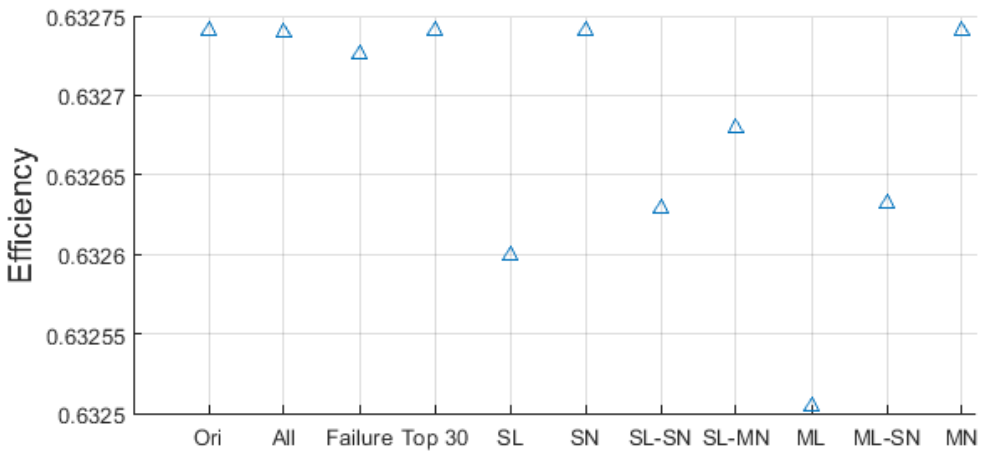


Figure 12: Network topological efficiency for different categories

Similar to the controllability index, node failures alone have no influence on topological efficiency and multiple link failures have a more important influence than single link failures. The mean of E over all cases of single link failure (i.e. SL, SL-SN and SL-MN) is 0.6326 and for the cases containing multiple link failures (i.e. ML and ML-SN) it is 0.6325.

Generally speaking, the variation of E is not significant, much less than the other two indexes. This is reasonable, because the network is not a sequential one and multiple paths exist between any two nodes: when one link fails, the gas flow can still be transferred via

an alternative path, although of longer distance.

5.4 Protective actions

From the above analysis, we understand that node failures have significant importance in terms of supply, but do not affect other properties, that the link failures influence on NSD is less important than that of node failures and that the consequences of failures on controllability and topological efficiency are limited compared to NSD. This means that, supply is the primary concern with respect to protection from failures, whereas network connections and control are not so sensitive and more fault tolerant.

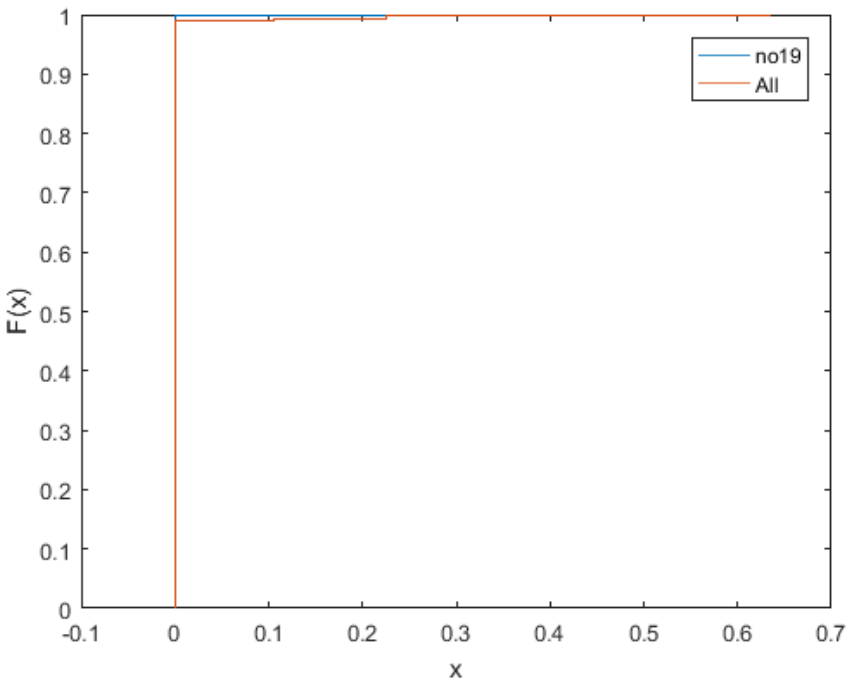


Figure 13: CDF of NSD for all configurations and for the configurations without failure of node 19

Node 19 is found to be an important node which may require protection priority, since it has a relatively high failure probability and at the same time has a significant impact on supply. Figure 13 compares the cumulative distribution function of NSD over all simulated configurations and the 991 570 configurations in which node 19 is not failed: the mean value of NSD of all configurations drops from 0.0020 to 0.0001, which is much lower.

6 Conclusions

Critical infrastructures, such as power grids, gas or water distribution networks, etc., are complex networks designed and operated to supply the service demanded. The increasing threats to the safety and security of their functions make it crucial to ensure safe and reliable performance.

In this paper, we have developed an analysis framework considering three perspectives: supply service, controllability and topology. By performing such an analysis, we are able to identify the most critical elements within the network and quantify the consequences of failure scenarios. The analysis framework has been applied to a gas transmission network. In the current work, repair is not taken into account and only one failure level is simulated. Even with such simplifications, the presented framework is able to identify the critical elements in terms of different perspectives considered. In future work, a multi-state approach and repair will be considered.

The results of the analysis show that the influence of a single link is not the same for the three perspectives and neither are the various failure scenarios. Supply turns out to be the most affected by failures, and can, thus, be used as the objective for the prioritization of investments for CI protection. However, other key performance indicators should be considered, including flexibility, recovery capacity, etc., which the other perspectives are sensitive to; then, the integration of the three perspectives should be considered within a multi-objective optimization for the multi-KPI design of CIs. This will be the object of future work. The findings of this work show the importance of considering several perspectives of analysis for CIs.

In the current paper, we emphasize the identification of the elements critical to the perspectives considered and the stochastic approach is adopted. Consideration of cascading effects and targeted attacks will complete the study in the continuation of this work, since this can provide more information for understanding the correlation among the different properties. Also, as CIs are more and more interconnected and automated, interdependencies and multi-CIs modeling become of great interest. The framework will be expanded and improved by taking into consideration the interdependencies in future works.

Acknowledgement

The authors would like to thank Dr. Vytis Kopustinskas and Dr. Pavel Praks of the Joint Research Center of the European Union for providing the data of the case study.

References

[1] Wolfgang Kröger and Enrico Zio. *Vulnerable systems.* Springer Science & Business Media, 2011.

[2] William B. Rouse. Engineering complex systems: Implications for research in systems engineering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2(33):154–156, 2003.

[3] Julio M Ottino. Engineering complex systems. *Nature*, 427(6973):399–399, 2004.

[4] EU. European Energy Security Strategy. *Communication from the Commission to the European Parliament and the Council. COM (2014) 330 final. Brussels, 2014*, 2014.

[5] Madelene Lindström and Stefan Olsson. The European programme for critical infrastructure protection. In *Crisis Management in the European Union*, pages 37–59. Springer, 2009.

[6] Ted G Lewis. *Critical infrastructure protection in homeland security: defending a networked nation.* John Wiley & Sons, 2014.

[7] William Clinton. Presidential Decision Directive PDD-63, Protecting America's Critical Infrastructure. *The White House, Washington, DC*, 1998.

[8] EU. Regulation no. 994/2010 of the european parliament and of the council of 20 october 2010 concerning measures to safeguard security of gas supply and repealing council directive 2004/67/ec. *Official Journal of the European Union*, 2010.

[9] Eric Rigaud and Franck Guarnieri. Proposition of a conceptual and a methodological modelling framework for resilience engineering. In *2nd Symposium on Resilience Engineering*, pages 8–pages, 2006.

[10] Efstathios Bakolas and Joseph H Saleh. Augmenting defense-in-depth with the concepts of observability and diagnosability from control theory and discrete event systems. *Reliability Engineering & System Safety*, 96(1):184–193, 2011.

[11] Nancy Leveson. A new accident model for engineering safer systems. *Safety science*, 42(4):237–270, 2004.

[12] Rudolf Kalman. On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):110–110, 1959.

[13] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[14] Ching-Tai Lin. Structural controllability. *Automatic Control, IEEE Transactions on*, 19(3):201–208, 1974.

[15] Zhengzhong Yuan, Chen Zhao, Zengru Di, Wen-Xu Wang, and Ying-Cheng Lai. Exact controllability of complex networks. *Nature communications*, 4, 2013.

[16] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Control centrality and hierarchical structure in complex networks. *Plos one*, 7(9):e44459, 2012.

[17] Noah J Cowan, Erick J Chastain, Daril A Vilhena, James S Freudenberg, and Carl T Bergstrom. Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PloS one*, 7(6):e38398, 2012.

[18] Wen-Xu Wang, Xuan Ni, Ying-Cheng Lai, and Celso Grebogi. Optimizing controllability of complex networks by minimum structural perturbations. *Physical Review E*, 85(2):026115, 2012.

[19] Tao Jia and Albert-László Barabási. Control capacity and a random sampling method in exploring controllability of complex networks. *Scientific reports*, 3, 2013.

[20] Tao Jia, Yang-Yu Liu, Endre Csóka, Márton Pósfai, Jean-Jacques Slotine, and Albert-László Barabási. Emergence of bimodality in controlling complex networks. *Nature communications*, 4, 2013.

[21] Tamás Nepusz and Tamás Vicsek. Controlling edge dynamics in complex networks. *Nature Physics*, 8(7):568–573, 2012.

[22] Yujian Pan and Xiang Li. Structural controllability and controlling centrality of temporal networks. *PloS one*, 9(4):e94998, 2014.

[23] Gang Yan, Jie Ren, Ying-Cheng Lai, Choy-Heng Lai, and Baowen Li. Controlling complex networks: How much energy is needed? *Physical review letters*, 108(21):218703, 2012.

[24] Pavel Praks, Vytis Kopustinskas, and Marcelo Masera. Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure. *Reliability Engineering & System Safety*, 144:254–264, 2015.

[25] Min Ouyang. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability engineering & System safety*, 121:43–60, 2014.

[26] Adam M Lewis, David Ward, Lukasz Cyra, and Naouma Kourti. European reference network for critical infrastructure protection. *International journal of critical infrastructure protection*, 6(1):51–60, 2013.

[27] Yi-Ping Fang and Enrico Zio. Hierarchical modeling by recursive unsupervised spectral clustering and network extended importance measures to analyze the reliability characteristics of complex network systems. *American Journal of Operations Research*, 3(1A):101–112, 2013.

[28] Yi-Ping Fang and Enrico Zio. Unsupervised spectral clustering for hierarchical modelling and criticality analysis of complex networks. *Reliability Engineering & System Safety*, 116:64–74, 2013.

[29] Weiwei Lu, Meirong Su, Brian D Fath, Mingqi Zhang, and Yan Hao. A systematic method of evaluation of the Chinese natural gas supply security. *Applied Energy*, 165:858–867, 2016.

[30] Juhani Nieminen. On the centrality in a graph. *Scandinavian journal of psychology*, 15(1):332–336, 1974.

[31] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

[32] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701, 2001.

[33] Paolo Crucitti, Vito Latora, and Massimo Marchiori. A topological analysis of the Italian electric power grid. *Physica A: Statistical Mechanics and its Applications*, 338(1):92–97, 2004.

[34] Åke J Holmgren. Using graph models to analyze the vulnerability of electric power networks. *Risk analysis*, 26(4):955–969, 2006.

[35] Zili Zhang, Xiangyang Li, and Hengyun Li. A quantitative approach for assessing the critical nodal and linear elements of a railway infrastructure. *International Journal of Critical Infrastructure Protection*, 8:3–15, 2015.

[36] James J Kelleher. Tactical communications network modelling and reliability analysis: An overview. Technical report, DTIC Document, 1991.

[37] José Emmanuel Ramirez-Marquez and Claudio M Rocco. All-terminal network reliability optimization via probabilistic solution discovery. *Reliability Engineering & System Safety*, 93(11):1689–1697, 2008.

[38] Darren L Deeter and Alice E Smith. Heuristic optimization of network design considering all-terminal reliability. In *Reliability and Maintainability Symposium. 1997 Proceedings, Annual*, pages 194–199. IEEE, 1997.

[39] Berna Dengiz, Fulya Altiparmak, and Alice E Smith. Efficient optimization of all-terminal reliable networks, using an evolutionary approach. *IEEE transactions on Reliability*, 46(1):18–26, 1997.

[40] Manolis Papadrakakis and Nikos D Lagaros. Reliability-based structural optimization using neural networks and monte carlo simulation. *Computer methods in applied mechanics and engineering*, 191(32):3491–3507, 2002.

[41] Dan M Frangopol and Min Liu. Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost. *Structure and infrastructure engineering*, 3(1):29–41, 2007.

[42] E. Zio. Reliability analysis of systems of systems. *IEEE Reliability Magazine*, pages 1–6, February 2016.

[43] Enrico Zio. Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering & System Safety*, 152:137–150, 2016.

[44] Enrico Zio. Critical infrastructures vulnerability and risk analysis. *European Journal for Security Research*, (DOI 10.1007/s41125-016-0004-2):1–18, 2016.

[45] Enrico Zio. Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, 94(2):125–141, 2009.

[46] Zhang Limiao, Li Daqing, Qin Pengju, Fu Bowen, Jiang Yinan, Enrico Zio, and Kang Rui. Reliability analysis of interdependent lattices. *Physica A: Statistical Mechanics and its Applications*, 452:120–125, 2016.

[47] Enrico Zio and Lucia R Golea. Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. *Reliability Engineering & System Safety*, 101:67–74, 2012.

[48] Jian Li, Leonardo Dueñas-Osorio, Changkun Chen, and Congling Shi. Connectivity reliability and topological controllability of infrastructure networks: A comparative assessment. *Reliability Engineering & System Safety*, 156:24–33, 2016.

[49] Wenyan Wu, Holger R Maier, and Angus R Simpson. Multiobjective optimization of water distribution systems accounting for economic cost, hydraulic reliability, and greenhouse gas emissions. *Water Resources Research*, 49(3):1211–1225, 2013.

[50] P Praks and V Kopustinskas. Identification and ranking of important elements in a gas transmission network by using progasnet. In *Risk, Reliability and Safety: Innovating Theory and Practice*, pages 1573–1579. CRC Press, 2016.

[51] Rudolf Emil Kalman. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial & Applied Mathematics, Series A: Control*, 1(2):152–192, 1963.

[52] Anna Lombardi and Michael Hörnquist. Controllability analysis of networks. *Physical Review E*, 75(5):056110, 2007.

[53] EGIG. 8th report of the european gas pipeline incident data group. December 2011.

[54] M Van der Borst and H Schoonakker. An overview of psa importance measures. *Reliability Engineering & System Safety*, 72(3):241–245, 2001.

Paper III

F. Han, E. Zio. Optimization of critical infrastructures with respect to supply service, structural complexity and controllability. Reliability Engineering and System Safety, 2017. (Under review).

Optimization of critical infrastructures with respect to supply service, structural complexity and controllability

Fangyuan Han[1] and Enrico Zio[1,2]

[1]Chair on Systems Science and the Energetic Challenge, Fondation EDF, CentraleSupelec, France

[2]Department of Energy, Politecnico di Milano, Italy.

Abstract

We consider critical infrastructures for supply and their optimization with respect to the objectives of minimizing the non-supplied demand and their structural complexity, while at the same time maximizing their controllability. The multi-objective optimization is performed by the efficient heuristics of the non-dominated sorting genetic algorithm II (NSGA-II). In order to grasp the nature of the tradeoffs among the three objectives and extract useful information from the optimal solutions, a thorough analysis is performed to investigate their correlations and the impact of topological properties of the supply network (e.g. the average node degree). A benchmark network representative of a real gas transmission network across several countries of the European Union (EU) is considered as case study to illustrate the optimization framework and to present the associated analysis. The findings of this paper demonstrate the usefulness of considering the three objectives for providing information for supply network design.

Keyword: Complex networks, Controllability, Structural complexity, Multi-objective optimization, Gas transmission network.

List of Symbols

$< k >$	Average node degree
α_i	Complexity of the ith component
β_{ij}	Interface complexity between the ith and j^{th} components
Γ	Number of inequality functions
γ	Normalization factor
Λ	Number of equality constraints
$\mathbf{Adj}$	Adjacency matrix
$\mathbf{A}$	Coupling matrix
$\mathbf{B}$	$N \times M$ input matrix
$\mathbf{Ctrb}$	Controllability matrix
$\mathbf{G}$	Graph representing the network
$\mathbf{G}(base)$	Graph representing the original network
$\mathbf{G}'(x_i = 1)$	Graph obtained by removing the node i from $\mathbf{G}(base)$
$\mathbf{K}$	Capacity matrix of the network
$\mathbf{u(t)}$	Input vector of M independent control signals at time t
$\mathbf{x(t)}$	System state vector at time t
$\mathbf{X}$	Decision variable vector
$\mu(\lambda_i)$	Geometric multiplicity of the eigenvalue λ_i of matrix $\mathbf{A}$
ω_i	Weight of the ith user
C	Structural complexity
C_1	Sum of individual components complexities
C_2	Sum of interface complexities
C_3	Topological complexity
D_i	Demand of the ith user
$E(\mathbf{Adj})$	Matrix energy of $\mathbf{Adj}$ (i.e. sum of its singular values)
f_o	o^{th} objective function

g_l	l^{th} of the Λ equality constraints
h_k	k^{th} of the K inequality constraints
L	Number of links in the network
N	Number of nodes in the network
N_D	Minimum number of driver nodes
N_G	Maximum number of generations
N_P	Number of candidate solutions
N_s	Number of sources in the network
N_y	Number of user nodes within the network
$NI_i^{C_{ind}}$	Node importance with respect to C_{ind}
NI_i^C	Node importance with respect to structural complexity
NI_i^{NSD}	Node importance with respect to NSD
NSD	Normalized non-supplied demand
O	Number of objective functions
P_t	Parent population of the tth generation
Q_t	Offspring population of the tth generation
R_t	Union population of the tth generation
y_i	Supply to the ith user

List of Acronyms

CI	Critical infrastructure
EU	European Union
MBDE	Modified binary differential evolution
mcm	million of cubic meters
MOEA	Multi-objective evolutionary algorithm
NSBDE	Non-dominated sorting binary differential evolution

NSD Non-supplied demand

NSGA-II Non-dominated sorting genetic algorithm II

RA Risk Achievement

1 Introduction

Critical infrastructures (CIs), like power grids, gas transmission and distribution systems, rail and road transport or communication networks, supply services that are essential to the operation of modern society [1]. They need to be designed, maintained and protected to provide optimal performance, reliable operation and functional safety for long periods of time [2, 3]. Hence, the great attention and priority given to the "care" of these systems by the EU, US and other national and transnational administrations [4, 5].

CIs are large-scale engineered systems of significant complexity for satisfying the continuously increasing demand on functionality and performance, improved lifecycle properties and safety, etc [1, 6]. Adequate understanding of the complexity of the systems is, thus, needed.

The analysis of the complexity of natural and engineered systems has long been an active field of research and application, and various measures have been introduced to describe complexity, including entropy [7], randomness [8] and predictability [9]. In [6], the authors formulated a quantitative structural complexity metric, taking into account the heterogeneity and quantity of different elements, and their connectivity structure. Compared to counting-based measures of complexity, this metric captures the 'global effect' of the system structure. In addition, this metric has been proved empirically [10, 11] and experimentally [12] to follow the same increasing trends as the development costs of complex systems, thus providing a way to look at the fundamental characteristics of system architecture in the process of system selection and design.

But the complexity of CIs calls for approaches capable of viewing the problem from multiple perspectives [13–15]. System analysis, reliability engineering, graph theory and others have been propounded to study the behavior and performance of complex systems, also with respect to failure events, their protection and resilience [16–18]. Integration of the different perspectives and analysis of their relations is necessary. For example, in [19] an electrical transmission system is analyzed with the objective of identifying the most critical elements in terms of four different perspectives: topological, reliability, electrical and electrical-reliability. In [20, 21], the correlation between connectivity reliability and

controllability of network systems is studied. In [18], the authors perform network reliability analysis considering spatial constraints. The authors of [22] consider a three-objective optimization of economic cost, resilience and greenhouse gas emissions, and the nature of the tradeoffs among the objectives is also studied.

Control is a fundamental property for safe and reliable operation of CIs, it is, then, important to develop a control framework able to steer the network dynamics toward states with optimal performance, while avoiding undesired or unfavorable states [23]. To achieve this, it is important to understand the controllability of complex networks. However, the control of the complex network systems that make up CIs remains a challenging problem. Studying the controllability of complex networks requires an integration of classical control theory and network theory. In this perspective, the notion of structural controllability has been introduced in [24]. In [25], analytical tools have been developed to characterize the controllability of directed networks. Several related topics can be considered under this framework, such as control centrality [26], optimization [23], control energy [27], control capacity [28], control mode [29], control of edge dynamics [30] etc.

The primary objective of the present work is to provide a framework for the analysis of CIs and their optimization, with respect to different objectives. In previous works on analysis and optimization of CIs, attention has typically focused on reliability and cost [31–33]. Specifically, due to the requirement of reliable operation and the complex nature of CIs, all-terminal reliability is often considered as a constraint or a second objective for the optimization problem [34–36]. Various approaches have been proposed to solve the multi-objective optimization problem considering cost and supply, such as, heuristics [32, 36], neural networks (NN) [37] and particle swarm optimization algorithm [38].

In the present work, we include the control perspective in the optimization through the introduction of the objective of minimizing the number of driver nodes, which relates to the minimization of the cost of controlling the system, which is another relevant property of a supply network.

The multi-objective optimization problem is solved by the non-dominated sorting genetic algorithm II (NSGA-II) [39]. This is a computationally efficient, elitist multi-objective evolutionary algorithm (MOEA) based on a non-dominated sorting approach. NSGA-II has been shown to outperform other optimization algorithms in terms of the spread of solutions and efficiency of convergence near the true Pareto-optimal front [39]. Various applications to CIs can be found in literature. In [40], NSGA-II is applied to search a set of optimal solutions for the rehabilitation of water distribution systems with the objectives of total cost, reliability

and water quality. In [41], NSGA-II is adopted for water distribution systems to solve the multi-objective optimization problem of their design considering cost and greenhouse gas emission. In [42], NSGA-II is used to search for the optimal capacity allocation pattern of a power transmission network to optimize its resilience against cascading failures while minimizing investment costs.

In this work, we present an optimization framework for the design (i.e. allocation of links and capacities) of the complex network making up a CI, with the objective of minimizing the non-supplied demand (NSD) and structural complexity, and maximizing structural controllability (i.e minimizing the number of driver nodes). The correlation among the three objectives considered is also investigated to understand the nature of their tradeoffs. The results of the analysis can provide useful information for the decision maker to compare and choose alternative designs of complex-networked CIs of different characteristics.

A complex network representative of a real EU gas transmission system supplying several countries is considered as case study to illustrate the optimization framework and to investigate the correlations among the different system-level properties.

The main contributions of this work are:

- Formulation and efficient solution of a three-objective optimization framework for complex CIs design.
- Consideration of the system controllability property in the optimization.
- Investigation of the multiple objectives considered and their correlations, to retrieve insights useful for system design.

The rest of the paper is organized as follows: Section 2 introduces the three system-level property indexes considered as objectives of the optimization in this paper. In section 3, the multi-objective optimization framework is presented. Section 4 describes the modelling of the considered gas transmission network. Section 5 presents the optimization results and analysis. Finally, conclusions are drawn in Section 6.

2 System-level property indexes

In this work, we consider three properties of the system from the supply, topological and control perspectives, non-supplied demand, structural complexity and controllability, and introduce three corresponding system-level indexes.

2.1 Non-supplied demand

Supply performance is the fundamental functional requirement of a CI. Consider a CI network, which supplies service or products from N_s production nodes (sources) to its N_y user nodes (users) through a number of transmission nodes.

We propose to use the non-supplied demand (NSD) as a measure of the network's capacity to satisfy its users' demand. The normalized NSD is introduced as a system level index:

$$NSD = 1 - \frac{\sum_{i=1}^{i=N_y} \omega_i y_i}{\sum_{i=1}^{i=N_y} \omega_i D_i} \tag{1}$$

where, ω_i is the weight of the ith of the N_y users, y_i is the supply to user i and D_i is its demand, which is considered as the target supply to user i. Then, the second term in Equation 1 represents the satisfied proportion of users' demand. Since $y_i \leq D_i$, the index NSD is normalized to take values in $[0, 1]$. Note that NSD equals 0 when the users' demands are fully supplied. As the mission of the systems is to satisfy users' demands, the objective of the optimization is to minimize NSD.

2.2 Structural complexity

System complexity is another important property of CIs. In this study, we adopt the structural complexity metric introduced in [6]. This metric, hereafter denoted C, accounts for the complexities of the individual components C_1, the complexities linked to the connections among the components C_2 and the topological complexity of the system structure C_3:

$$\begin{aligned} C &= C_1 + C_2 C_3 \\ &= \sum_{i=1}^{N} \alpha_i + \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} \mathbf{Adj}_{ij} \right] \gamma E(\mathbf{Adj}) \end{aligned} \tag{2}$$

where α_i is the complexity of the ith component, β_{ij} is the interface complexity between the ith and j^{th} components, depending on the complexities of the two components and the type of the interface, $\mathbf{Adj}$ is the adjacency matrix defining the connections among the components, γ is a normalization factor and $E(\mathbf{Adj})$ is the matrix energy of the network, defined as the sum of the singular values of its adjacency matrix $\mathbf{Adj}$, thus accounting for the topological complexity of the system structure.

The term C_1 is related to component engineering and it does not hold structural information; C_2 is related to interface design, including the quantity and complexity of each pair-wise connection between components; C_3, measured by matrix energy, describes the

interaction between nodes of the network and depends on the connectivity structure; topological complexity increases from centralized to more distributed networks and is related to system integration: higher topological complexity will likely lengthen system integration efforts significantly.

The structural complexity metric C introduced, captures information on the fundamental characteristics of system architecture, providing quantitative measurement, and at the same time, specifying the origin of such complexity. Empirical and experimental evidence reveals that the system development cost grows with the system structural complexity, suggesting that a low structural complexity for low development cost is preferred, if the design satisfies all of the other constraints [12].

2.3 Controllability

Safe and reliable operation of CIs stands on understanding their behavior and structural and dynamic characteristics. The ultimate assurance stands on the ability to control them [25], which leads to the need of characterizing the controllability of CIs.

A system is controllable if, by a suitable choice of inputs, it can be driven from any initial state to any desired final state within finite time [25,43]. In the specific case of interest here, the control of a gas supply network is intended to guide the supply of the demanded amount of gas to its users, while operating in safe conditions. This can be achieved by various control actions, e.g. pressure and gas flow regulation in the pipelines, etc. In this work, we do not consider the effects of different control actions and investigate controllability from a system perspective.

Considering the network of N nodes, we describe its state dynamics by:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{3}$$

where $\mathbf{x}(\mathbf{t})$ is the system state vector, describing the state of each node in the network at time t; $\mathbf{A}$ is an $N \times N$ coupling matrix, in which a_{ij} represents the weight of the directed link from node i to node j (i.e. the interaction strength between node i and node j, for example, the flow in the pipeline of a gas transmission network); $\mathbf{B}$ is an $N \times M$ input matrix ($M \leq N$), identifying the nodes that are controlled by the time-dependent input vector $\mathbf{u}(\mathbf{t})$, made of M independent control signals.

Based on dynamic control theory, the above system is controllable if and only if the $N \times NM$ controllability matrix $\mathbf{Ctrb} = (\mathbf{B} \quad \mathbf{AB} \ldots \mathbf{A}^{n-1}\mathbf{B})$ has full rank (the so-called

Kalman's rank condition) [44]:

$$rank(\mathbf{Ctrb}) = N \tag{4}$$

For complex network systems, the controllability problem can be formulated in terms of finding a suitable control matrix $\mathbf{B}$ consisting of a minimum number of driver nodes (N_D) so as to satisfy the Kalman's rank condition (4). However, this requires the evaluation of the rank of $\mathbf{C}$ for 2^N possible combinations of the driver nodes [45]: fot real CI network systems, such a brute-force search is computationally prohibitive.

To overcome this problem, in [25] analytical methods have been developed to determine the minimum number of inputs (or driver nodes) N_D that are needed to fully control the network, by finding the maximum matching, i.e. the maximum set of links that do not share start or end nodes. Full control can be achieved if and only if each unmatched node is directly controlled and there are directed paths from the input signals to all matched nodes. The unmatched nodes determined by maximum matching are the driver nodes.

In [46], the exact controllability for arbitrary network structures and link weights (say arbitrary matrix $\mathbf{A}$) is introduced to calculate N_D:

$$N_D = \max_i \{\mu(\lambda_i)\} \tag{5}$$

and the minimum number of driver nodes N_D is determined by the maximum geometric multiplicity $\mu(\lambda_i)$ of the eigenvalue λ_i of $\mathbf{A}$. In fact, this number is that of the nodes corresponding to the linearly-dependent rows. The controllers should be imposed on the linearly-dependent rows to eliminate all linear correlations to ensure the controllability condition.

The number of driver nodes N_D can be taken as a measure of the controllability of the network, indicating how many driver nodes are needed to control the network and directly related to the cost of the resources needed to keep or bring the system under control. If $N_D = N$, i.e. the total number of nodes in the network, (this means that the external control signal is applied to each node of the network), the likelihood of gaining full system control is high, but so is the associated cost [23]. A small N_D, instead, indicates a more controllable network system, in the sense that it requires less effort to obtain full control over the network.

In this work, we use the minimum number of driver nodes N_D to indicate the controllability of the network and aim at minimizing N_D for a less costly network in terms of control.

3 Multi-objective Optimization problem

The optimization of the design of a real world CI network typically involves the simultaneous optimization of multiple objective functions and the consideration of several equality and/or inequality constraints. In all generality, such a multi-objective problem can be formulated as follows (considering minimization of the objectives):

$$Minimize \quad f_o(\mathbf{X}), o = 1, \ldots, O \tag{6}$$

$$Subject \quad to \quad \begin{cases} g_l(\mathbf{X}) = 0, & l = 1, \ldots, \Lambda \\ h_k(\mathbf{X}) \leq 0, & k = 1, \ldots, \Gamma \end{cases} \tag{7}$$

where f_o is the oth objective function, $\mathbf{X}$ is the decision variable vector to be optimized, O is the number of objective functions, g_l is the lth of the Λ equality constraints and h_k is the kth of the Γ inequality constraints.

In this study, we seek to: (1) minimize the non-supplied demand (NSD) so as to satisfy the users' demands as much as possible; (2) minimize the structure complexity (C), which relates to the cost of the system design and the development efforts for integration of the system components and the system management; (3) minimize the number of driver nodes (N_D) to obtain an as controllable as possible network with the minimum effort. The three objectives are calculated by equations (1), (2) and (5), respectively. The formulation of the optimization problem (6) specific to the three objectives is, then, given as follows:

$$\begin{cases} \min\ NSD(\mathbf{X}) \\ \min\ C(\mathbf{X}) \\ \min\ N_D(\mathbf{X}) \end{cases} \tag{8}$$

The decision variable vector $\mathbf{X}$ contains the capacity of the links connecting node i and node j, X_{ij}, $i, j = 1, ...N$, $i \neq j$. The element X_{ij} is equal to 0 if the nodes i and j are not connected; otherwise, it indicates the capacity of the link between the two nodes.

The constraints (7) to be met are specified as: (i) the nodes remain unchanged, including their quantity, location and functionalities; (ii) the users' demands and supply capacities of the sources stay the same; (iii) each demand node should be connected to at least one source node through a path.

In this study, for simplicity of illustration, the supply and demand data are fixed values. The extension to the case of variable supply or demand implies defining the simulated horizon T of the analysis and partitioning it into intervals t in which the supply and demand

is assumed to be steady. For each such period, $NSD(t)$ can be evaluated by Equation 1. Then, the objective function becomes the sum of $NSD(t)$ values over the time horizon T.

The final goal of the optimization is to identify a set of solutions in which no solution can be regarded better than another with respect to all objective functions. This can be achieved by adopting the concepts of Pareto optimality and dominance [47]. Solution $\mathbf{x}_a$ is regarded to dominate solution $\mathbf{x}_b$ if both following conditions are satisfied (in the case of minimization):

$$\begin{aligned} &\forall i \in \{1, 2, \ldots, O\}, \quad f_i(\mathbf{x}_a) \leq f_o(\mathbf{x}_b) \\ &\exists j \in \{1, 2, \ldots, O\}, \quad f_i(\mathbf{x}_a) < f_o(\mathbf{x}_b) \end{aligned} \tag{9}$$

The non-dominated solutions within the entire search space are denoted Pareto-optimal solutions and the corresponding values of the objective functions form the Pareto-optimal front [47].

We propose to use the non-dominated sorting genetic algorithm II (NSGA-II) to search the solutions of the Pareto-optimal set, given its proven performance [48]. NSGA-II uses a fast non-dominated sorting procedure, an elitist strategy, a parameter-less approach and a simple yet efficient constraint-handling method, which solves complex multi-objective optimization problems with satisfactory performance [39]. The NSGA-II proceeds as follows [39]:

1. Initialize the population P_0 of N_P candidate solutions and evaluate each of the N_P solutions in the population P_0, by calculating the three objective functions presented above.

2. Apply the binary tournament selection operator to the population P_0 to create an offspring population Q_0 of size N_P, which undergoes the evolution operations of mutation and crossover. Evaluate each of the N_P solutions in the population Q_0.

3. (For the tth generation) Combine P_t and Q_t to form a union population $R_t = P_t \cup Q_t$ of size $2N_P$. Then, the population R_t is sorted by the fast non-dominated sorting and the ranked non-dominated fronts F_1, F_2,...,F_k are identified (F_1 is the best, F_2 is the second best, and so on).

4. To select the first N_P members of R_t for the new population P_{t+1}, using the crowded-comparison operator [39].

5. Increase the generation number t by 1, and the algorithm stops when it reaches the maximum number of generations N_G.

4 Case study

4.1 Network description and graph representation

We consider the case study from [49]. The system is visualized in Figure 1 and represents a real gas transmission network for supply across several countries in the EU.

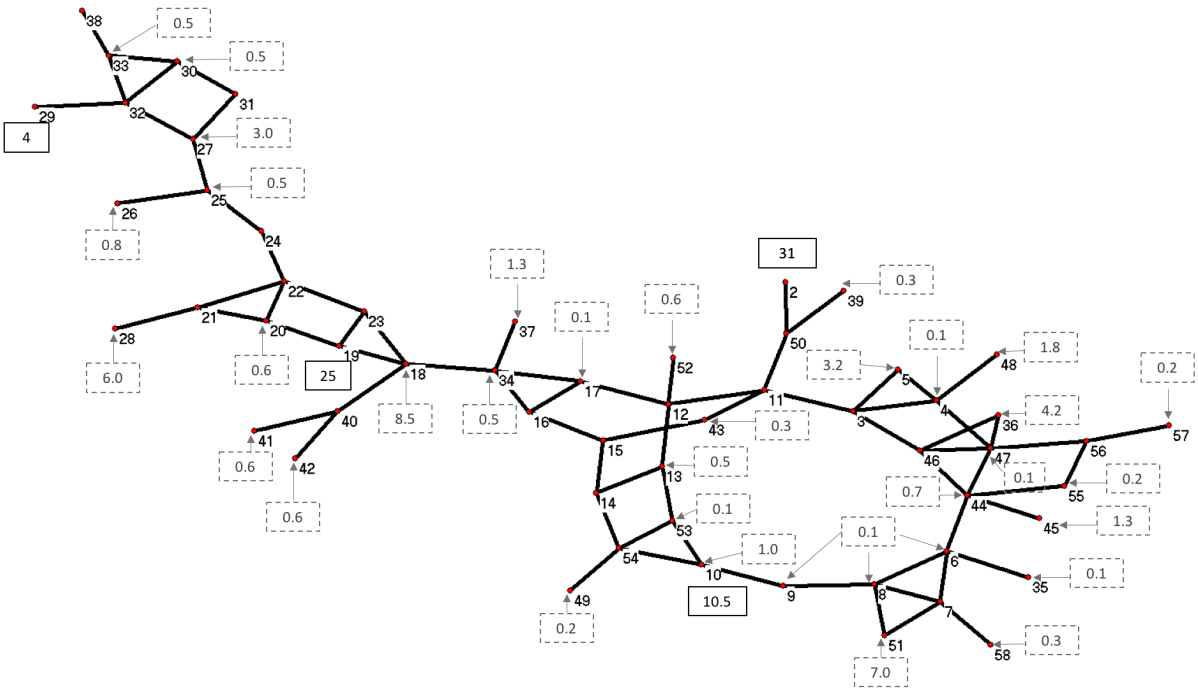


Figure 1: Gas transmission network: topology and data of supply (solid boxes) and demand (dashed boxes).

The gas transmission network is modeled as an undirected graph $\mathbf{G}$ of $N = 57$ nodes, connected by $L = 74$ links. Its connectivity structure can be defined by its $N \times N$ adjacency matrix $\mathbf{Adj}$, whose entries $[Adj_{ij}]$ are equal to 1 if there exists a link joining node i to node j and 0 otherwise.

We distinguish $N_y = 35$ demand nodes with deterministic daily demands for a total system daily demand of 45.9 millions of cubic meters (mcm), one LNG terminal (node 10), two compressor stations (nodes 12 and 13), two storage devices (nodes 10 and 19) and two pipeline source nodes (nodes 2 and 29). Note that an LNG terminal is a storage device of liquefied natural gas and at the same time, it consumes energy (gas) to operate, since the liquefied gas is turned back into gaseous state at LNG terminals. The storage devices and pipeline sources are considered supply nodes (numbered 2, 10, 19 and 29) , and their properties are presented in Table 1.

Node	Capacity	Type
2	31	Pipeline source
10	10.5	LNG terminal
19	25	Underground storage
29	4	Pipeline source

Table 1: Sources properties

The properties of the 35 demand nodes are shown in Figure 1. The capacities and demands are expressed in millions of cubic meters per day (mcm/d). The data of supply and demand are realistic and they are expressed at a daily scale, in order to assume peak gas demand during one peak day (in winter) with extreme high gas demand [49]. These data are intended to represent the most stressed conditions for the gas transmission network.

The links of the gas transmission network represent the gas transmission pipelines connecting the nodes. Each link in the network is characterized by its capacity, i.e the maximum amount of flow that it is able to supply, which is contained in the capacity matrix $\mathbf{K}$. Given the capacities of the links connecting the nodes and the constraints on the sources and users, the supply to each user can be computed by maximum flow algorithms [50].

In the case of multi-source and multi-terminal networks, such as the studied gas network, we consider a virtual source node (denoted node 1) with directed links to all of the N_s nodes representing sources, and a virtual sink node (denoted node 59) with directed links from the N_y demand nodes [49]. Nodes 1 and 59 do not have topological properties and are not shown in Figure 1. The capacity of the virtual links represents the capacity of supply and demand, which are also included in the capacity matrix $\mathbf{K}$. Thus, $\mathbf{K}$ is of size $(N+2) \times (N+2)$. The problem is converted into a single-source and single-terminal maximum flow problem. The amounts of flow of the virtual links from the demand nodes to node 59 are the supply to these demand nodes, and are used to calculate the non-supplied demand NSD defined in Equation 1. Various approaches can be applied to solve the maximum flow problem, and we use the Boykov-Kolmogorov algorithm, which is based-on augmenting path algorithms and constructs two search trees associated with the source node and the sink node [51].

4.2 Optimization problem and parameters setting

For the optimization with respect to the structural complexity and controllability, which are topological properties of the network, and the non-supplied demand, which depends on the capacities of the pipelines, we act on the connection patterns among the nodes and the

allocation of capacities to the connecting links. The variables to be optimized are, then, the link capacities $X_{ij} \in [0, x_{max}]$, which can take any value inferior to the limit capacity x_{max} assumed the same for any link joining node i and node j; if no link exists between i and j, $X_{ij} = 0$. The virtual links are not optimized and, thus, not included in the decision variables. Then, the adjacency matrix **Adj** derived from X_{ij} is used to calculate C and N_D as defined in Equations 2 and 5.

The three objectives to minimize are the normalized non-supplied demand (NSD), the structural complexity (C) and the number of driver nodes (N_D) of Equation (8) in Section 3. Note that, since the number of nodes and and their functionality remain the same (and so does the component complexities), the term C_1 in equation (2) can be neglected. As there exists only one type of connection between any two nodes, i.e. gas flow, the interface complexity β_{ij} is assumed to be the same for all pipelines and it is set to 0.5. However, this is a simplification of the reality, and it can be estimated differently considering the effect from distance, pressure, etc. γ is arbitrarily set to $1/N$.

NSGA-II is applied to solve the previously defined multi-objective optimization problem. The parameters of the algorithm are set as in Table 2, based on the results of a number of test runs.

Parameters	Values
Population size N_P	40
Crossover rate CR	0.9
Scaling factor F	0.2
Maximum generation N_G	300

Table 2: Parameters of the NSGA-II algorithm

Table 3 presents the three index values of the original gas transmission network.

NSD	N_D	C
0	4	196.2

Table 3: Index values for the original network

5 Results and discussion

In this section, we present the optimization results and analyze the Pareto-optimal solutions.

We run the NSGA-II algorithm 10 times and consider all the different Pareto-optimal solutions found during the ten runs, in order to investigate the relationship among the objectives and extract useful information for the selection of the CI network configuration. Then, we select the overall six non-dominated solutions to construct the Pareto front.

5.1 The impact of network topology

We choose the networks whose N_D is between 1 and 6 and divide them into six classes accordingly. We, then, calculate the mean average node degree, the mean of NSD and C for each class.

Studies in [21, 23, 25, 52] show that there exists a correlation between the average node degree and the controllability of networks.

In our study, we find that the average node degree $< k >$ of the generated networks ranges from 2.4 to 3.1. This can represent most cases of the real-world complex networks, which are typically sparse [21]. Indeed, the degree of a node is the number of edges connected to the node and the average node degree $< k >$ represents the link density of the network. When adding links to the network (which means that $< k >$ increases), the structural complexity increases, as the complexity metric C takes into account both the quantity of links and the topological complexity (matrix energy) does not reduce. Therefore, it is intuitive to understand that the higher $< k >$, the higher is its structural complexity. The Pearson correlation coefficient between the mean of $< k >$ and C is 0.9987.

Figure 2 shows the mean of the average node degree for each class of networks obtained: as the number of driver nodes increases, the average node degree decreases. For sparse networks, N_D is determined by the rank of $\mathbf{A}$ and drivers nodes correspond to linearly dependent rows [46]. Adding links to the network is possible to eliminate the linearly dependent rows in $\mathbf{A}$, which explains the negative correlation between N_D and average node degree $< k >$. The correlation coefficient between them is -0.9929.

As for the non-supplied demand NSD, the correlation coefficient is -0.9139. Adding links (which increases the average node degree) may increase the supply capacity of a network and, thus, NSD decreases. However, this is not guaranteed since the supply capacity between two nodes can be limited by a small capacity link on the path. The supply depends more on the capacity of transmission rather than on the number of links within the network. This

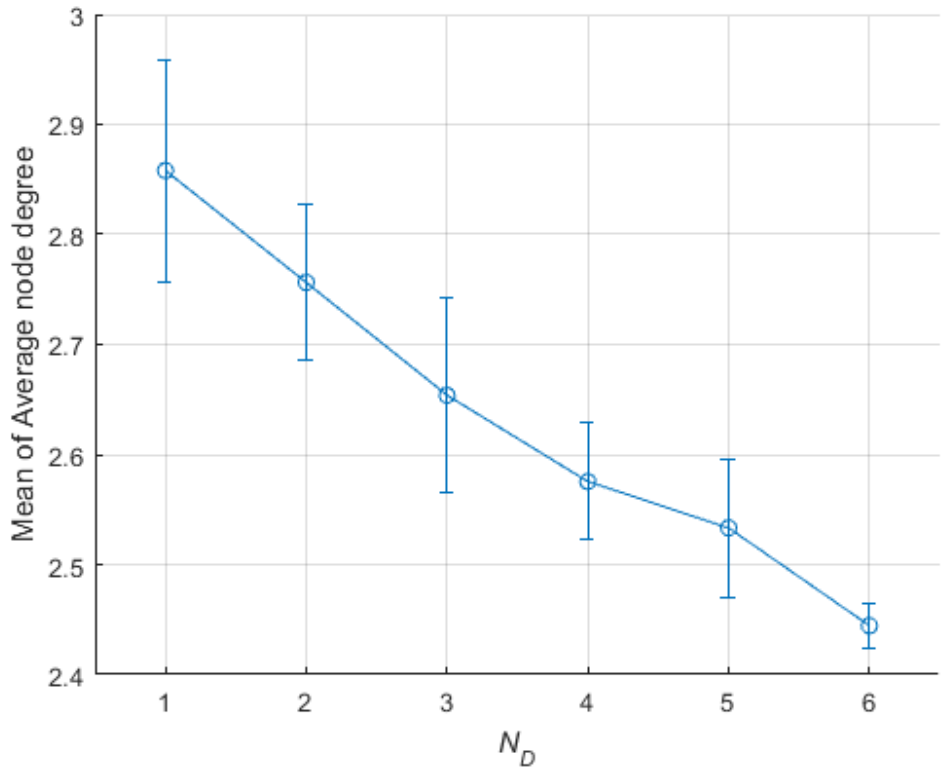


Figure 2: Mean of average node degree versus number of driver nodes N_D

explains why NSD correlates with $< k >$ less than the other two properties.

5.2 Correlations among the three objectives

The mean values of the three objectives for the six classes are shown in Figure 3. We observe that, as the number of driver nodes N_D increases, the structural complexity C decreases and NSD increases. The linear Pearson correlation coefficient r is computed to quantify the correlation among the objectives (Table 4).

	r
C and N_D	-0.9917
NSD and N_D	0.9418
C and NSD	-0.9202

Table 4: Pearson correlation coefficient among the three objectives

The number of driver nodes N_D and the structural complexity C are highly correlated.

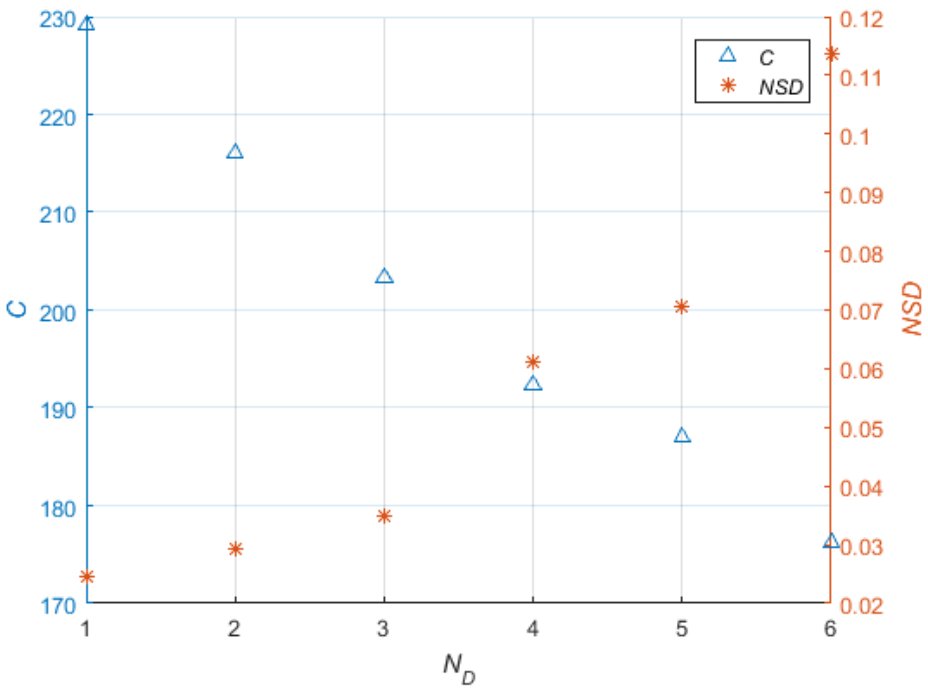


Figure 3: Mean of the three objectives

We can see from Figure 4 that as N_D increases, C increases, while the standard deviation decreases. This can be explained by the fact that these two objectives are affected by the average node degree in opposite directions, as mentioned in 5.1.

Figure 5 shows the correlation between NSD and N_D. We can observe that NSD and N_D have relatively weak correlation, and large standard deviation, which indicates that these two objectives are more independent.

5.2.1 Remarks

To sum up the previous analysis, we find that structural complexity and controllability are influenced by the topological property of the network system (average node degree) in opposite directions: for a sparse complex network, such as is the gas transmission network considered, a relatively dense one is preferable for the consideration of controllability, but it is a more structurally complex network, and therefore it comes at a cost; if we seek to choose a less complex network configuration, we tend to have a less controllable network, i.e. that requires more efforts to keep the system under control; yet, the topology has less impact on the supply to the users. And for the purpose of demand satisfaction, link capacity is a more important factor to consider. For selection among the Pareto-optimal solutions, it would be interesting to choose those with $NSD = 0$ and, then, to seek the balance between

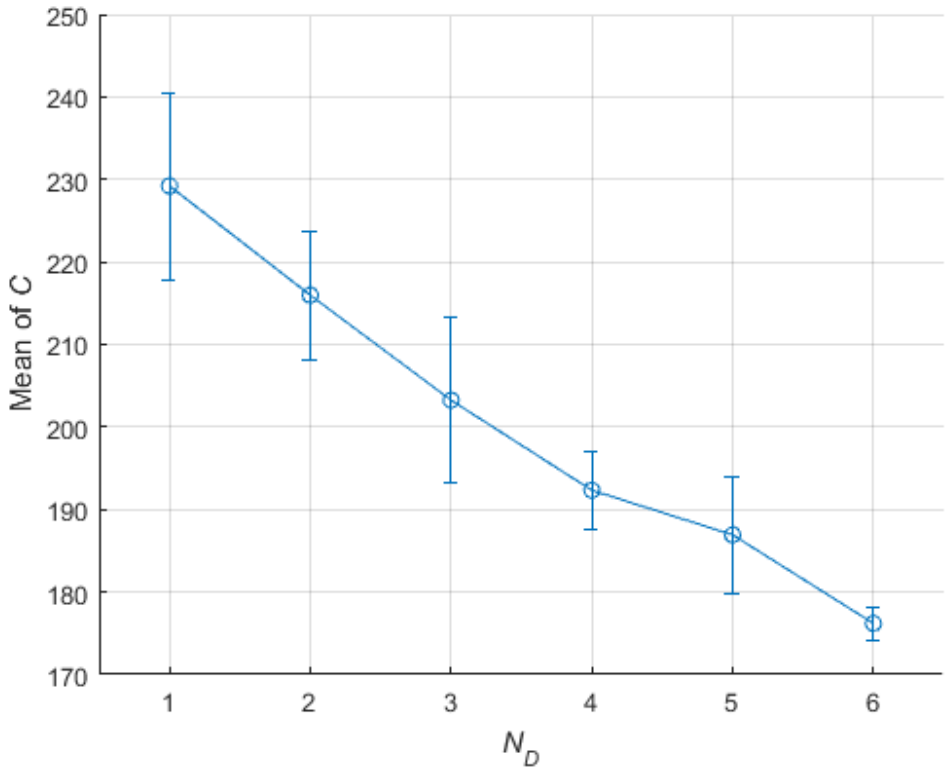


Figure 4: Mean of structural complexity C versus number of driver nodes N_D

complexity and controllability.

5.3 Optimization results

Table 5 presents the values of the three objective functions for each solution identified as Pareto-optimal. The Pareto front obtained by the NSGA-II is illustrated in Figure 6, and the topology of the six solution networks is shown in Figure 7.

No.	NSD	N_D	C
1	0	4	192.3
2	0.024	6	178.3
3	0	2	222.4
4	0	3	216.0
5	0	1	232.3
6	0.102	2	219.4

Table 5: Objective functions values for the Pareto-optimal solutions

The original real network (indicated by the triangle) is close to the solution 1 (Figure

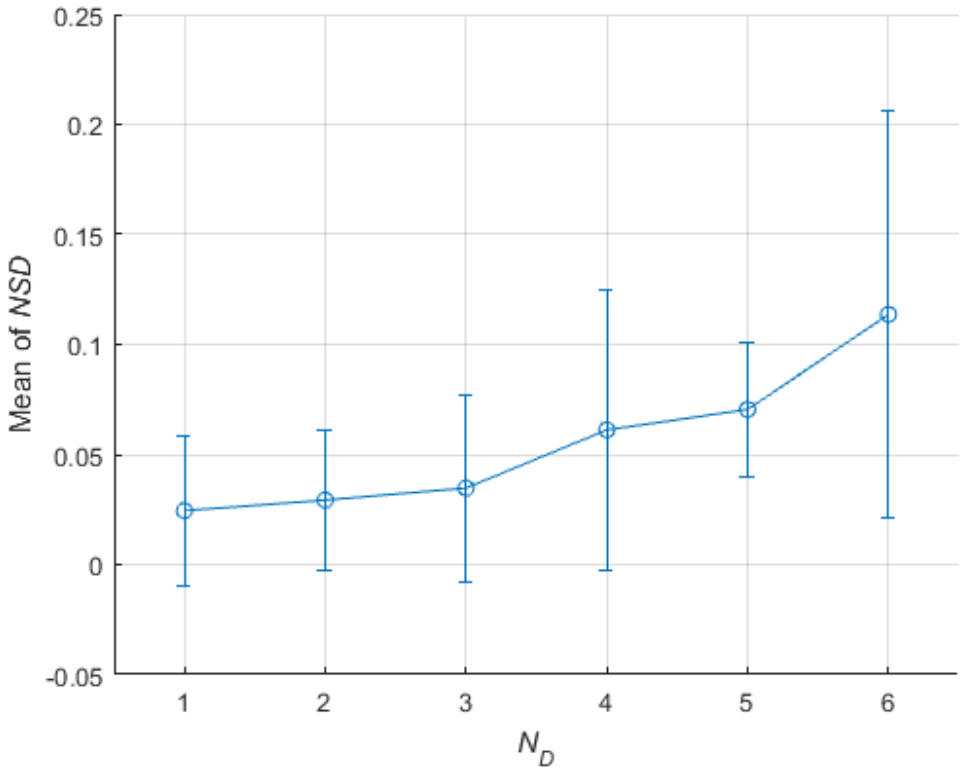


Figure 5: Mean of NSD versus number of driver nodes N_D

7(a)) with $N_D = 4$, $NSD = 0$. In fact, the two networks coincide in the plane N_D, NSD of Figure 6(d). The difference is that the optimal solution 1 has lower structural complexity than the original network, obtained by removing the link connecting nodes 44 and 55. Removing certain links could be an improvement to the original network, for the purpose of minimizing the complexity, for instance. The optimal solutions found may provide other possibilities of network configurations and capacity allocations.

Among the six solutions of Table 5, N_D varies from 1 to 6 and the complexity metric C varies from 178.3 to 232.3. Solution 2 has the lowest C and largest N_D. Solution 5 has the smallest number of driver nodes and largest C values. As for NSD, it takes value between 0 and 0.102. The Solutions 2 and 6 (Figures 7(b) and 7(f), respectively) include separate nodes, and the fact that they are not able to be supplied leads to $NSD > 0$. Four out of six solutions are able to fully satisfy all users' demands.

Let us focus on these latter four solutions, capable of fully satisfying the demands (i.e. solutions 1, 3, 4 and 5, with $NSD = 0$). Figure 6(c)). We can see that, as the number of driver nodes decreases, the structural complexity increases. We can define a rate of change to choose the most efficient optimal solution network in terms of corresponding objectives.

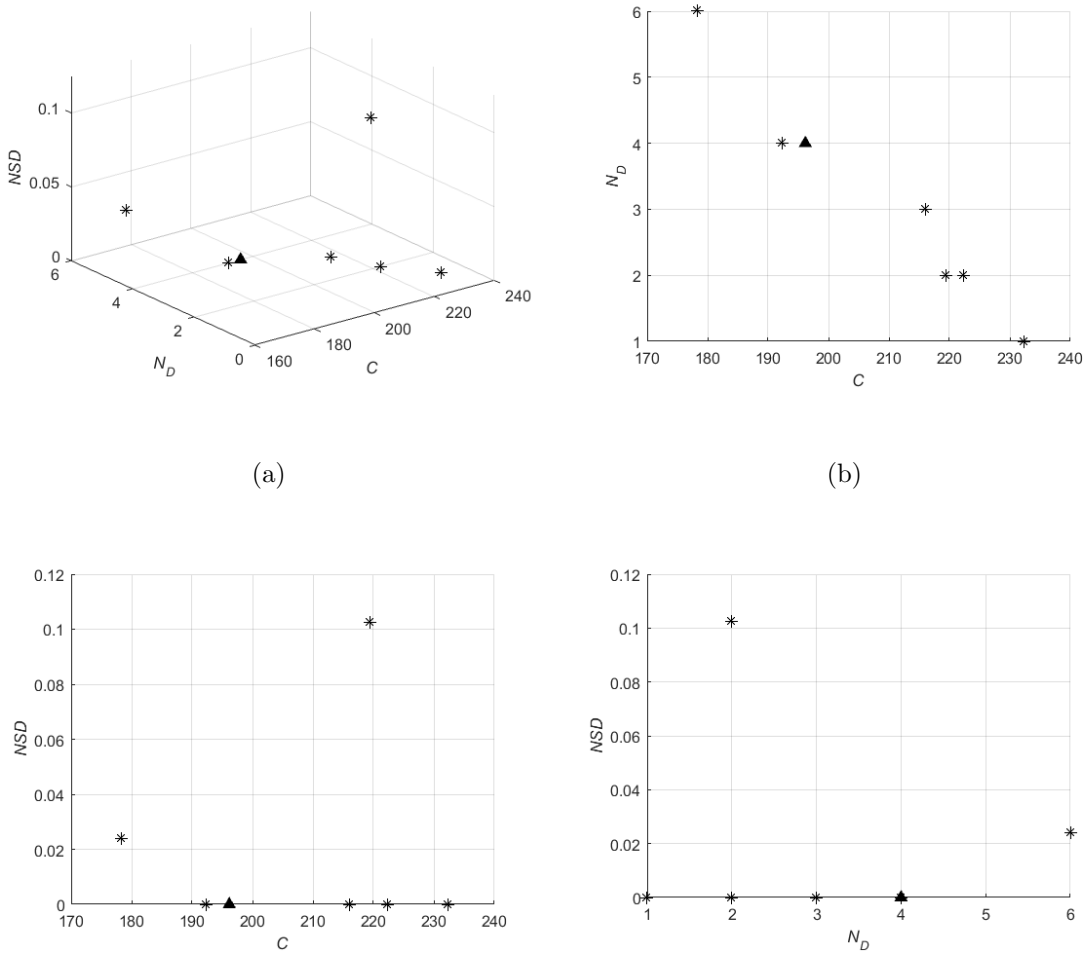


Figure 6: Pareto front in 3-D space (a) and 2-D projections (b)-(d)

If controllability is the primary concern, for example, we define the ratio of the changes in the number of driver nodes and in the structural complexity $\Delta N_D/\Delta C$: the larger this ratio, the more preferable in terms of gain the network is. Then, solution 1 is the best, and the original network also has a rather satisfying configuration.

We, then, compare the two solutions with $N_D = 2$ (solutions 3 and 6). Solution 3 has $NSD = 0$ and is, thus, better in terms of supply performance than solution 6, which has the highest non-supplied demand $NSD = 0.102$. On the other hand, the difference in terms of structural complexity is not significant. This indicates that, in this case, the structural complexity does not have a significant influence on the supply; instead of adding links, link capacity is a more important factor to consider.

5.4 Analysis of node importance

Different components may have different contributions to the normalized non-supplied demand (NSD), the structural complexity (C) and controllability (N_D).

To analyze this, we consider Risk Achievement (RA), an importance metric that measures the contribution of the failure of the generic i^{th} component ($x_i = 1$) to the system risk level [53]: $RA = R(x_i = 1) - R(base)$, where $R(x_i = 1)$ is the increased risk level brought by the fact that the ith component is taken out of the system or is failed and $R(base)$ is the risk of the original network with the ith component working. Considering that in this work the goal of the optimization is to obtain a network configuration with low NSD and N_D, we take the increase of these two indexes as the increase of risk level and compute the RA to quantify the node importance of each solution network with respect to them:

$$NI_i^{NSD} = NSD[\mathbf{G}] - NSD[\mathbf{G}'(x_i = 1)] \tag{10}$$

$$NI_i^{N_D} = \frac{N_D[\mathbf{G}'(x_i = 1)] - N_D[\mathbf{G}]}{N} \tag{11}$$

where $\mathbf{G}'(x_i = 1)$ is the graph obtained by removing from the solution network $\mathbf{G}$ the node i (i.e. all edges incident in node i).

As for the node importance with respect to the structural complexity, NI_i^C, it is defined as the relative drop in C caused by the deactivation of node i from the solution network $\mathbf{G}$:

$$NI_i^C = \frac{\Delta C[\mathbf{G}]}{C[\mathbf{G}]} = \frac{C[\mathbf{G}] - C[\mathbf{G}'(\mathbf{x_i} = \mathbf{1})]}{C[\mathbf{G}]} \tag{12}$$

In this section, we identify the most critical nodes and investigate whether some nodes may be important independently of the network configuration. We also analyze and compare the average node importance of each solution network, in order to understand if certain network configurations or certain properties are more sensitive to the removal of single nodes. The average node importances across all nodes are shown in Table 6.

Solution 1 and the original network give the same result for the node importance analysis: they both have the highest average node importance with respect to NSD, NI^{NSD}, and are most sensitive to node failure in terms of supply, which indicates that every link within this network is relatively critical. This means cost-effective design but good protection is required during operation to provide reliable service. Nodes 3 and 11 are the two most important nodes in terms of supply, representing important connections between the main source at node 2 and the rest of the network with pipelines of large capacity (11-13, 2-46). Node 19, the second largest source, is another important node: even though the total source

No.	NI^{NSD}	$NI^{C_{ind}}$	NI^{C}
1	0.0550	0.0249	0.0619
2	0.0288	0.0203	0.0614
3	0.0279	0.0219	0.0615
4	0.0253	0.0237	0.0618
5	0.0408	0.0252	0.0616
6	0.0228	0.0225	0.0612

Table 6: Average node importance values for the optimal solutions

capacity is enough to cover all demands, due to the limited capacity of pipelines connecting different areas, without it some nodes far from the main source at node 2 would not be supplied. These results match those of [54]. Solution 5 also has a relatively high NI^{NSD}, in fact: with the removal of source node 19, NSD reaches a value of 0.48, which is the highest value among all optimal network configurations. Solution 6 has a relatively low average node importance: however, it is not an interesting solution, since it can not satisfy all the users' demand. Solutions 3 and 4 are less sensitive to the removal of single nodes in terms of supply. We note that nodes 2 and 19 always have an important impact on NSD for any network configuration, because of their role as (large) sources. Node 18 is also a very important node for supply in all six solution networks, because it has a large demand and for most of the solutions, it connects the source node 19 to other demand nodes: thus, its removal usually has a more important impact on supply than source node 19.

The average node importance values with respect to structural complexity NI^{C} turn out to be close for all Pareto-optimal networks. The solutions are all sparse networks and the nodes have relatively low degree; thus, the removal of single nodes does not influence significantly the structural complexity. We also observe that node 18 is important for all the six solutions in terms of structural complexity, as we can see in Figure 7, as it connects a relatively large number of nodes and has a high node degree.

The removal of single nodes can at most increase the number of driver nodes N_D of 2 units. Solution 5 with $N_D = 1$ has the largest average node importance with respect to N_D and it is the most sensitive to node failure in terms of N_D. In contrast, Solution 2 with $N_D = 6$, is the least sensitive one.

Solutions 3 and 4 could be considered as reasonable alternatives to the original network configuration, since they are most resistant to node failures.

6 Conclusions

Critical infrastructures, such as power grids, gas or water distribution networks, are complex networks designed and operated to reliably and safely supply the service demanded. Thus, it is crucial to guarantee the control of such systems, so as to ensure safe and reliable performance under different operating conditions.

With respect to this, in this paper we have considered three objectives for the multi-objective optimization of complex supply networks for minimizing the non-supplied demand and the system structural complexity, and maximizing the system controllability (i.e. minimizing the number of driver nodes). We have proposed to use the non-dominated sorting genetic algorithm II (NSGA-II) to tackle the multi-objective optimization problem. A gas transmission network has been taken as reference case study. A comparative evaluation has been performed to analyze the optimal solutions, with respect to how the allocation of link capacities can improve the desired system properties. At last, an investigation of the impact of topological properties (i.e. the average node degree) on the three objectives and the correlation among them has been performed, to draw insights on possible tradeoffs among the objectives.

In conclusion, the findings of this work demonstrate the opportunity and possibility of developing frameworks of optimization and analysis for the design of critical infrastructures for supply, taking into account different properties desired for the system.

Acknowledgement

The authors would like to thank Dr. Vytis Kopustinskas and Dr. Pavel Praks of the Joint Research Center of the European Union for providing the data of the case study.

References

[1] Wolfgang Kröger and Enrico Zio. Vulnerable systems. Springer Science & Business Media, 2011.

[2] William B. Rouse. Engineering complex systems: Implications for research in systems engineering. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2(33):154–156, 2003.

[3] Julio M Ottino. Engineering complex systems. Nature, 427(6973):399–399, 2004.

[4] Madelene Lindström and Stefan Olsson. The European programme for critical infrastructure protection. In Crisis Management in the European Union, pages 37–59. Springer, 2009.

[5] Ted G Lewis. Critical infrastructure protection in homeland security: defending a networked nation. John Wiley & Sons, 2014.

[6] Kaushik Sinha and Olivier L de Weck. Structural complexity quantification for engineered complex systems and implications on system architecture and design. In ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pages V03AT03A044–V03AT03A044. American Society of Mechanical Engineers, 2013.

[7] Dan Braha and Oded Maimon. The measurement of a design structural and functional complexity. In A Mathematical Theory of Design: Foundations, Algorithms and Applications, pages 241–277. Springer, 1998.

[8] Peter Grassberger. Randomness, information, and complexity. arXiv preprint arXiv:1208.3459, 2012.

[9] Guido Boffetta, Massimo Cencini, Massimo Falcioni, and Angelo Vulpiani. Predictability: a way to characterize complexity. Physics reports, 356(6):367–474, 2002.

[10] Mike J Van Wie, James L Greer, Matthew I Campbell, Robert B Stone, and Kristin L Wood. Interfaces and product architecture. In Proceedings of DETC, volume 1, pages 9–12, 2001.

[11] James Richard Wertz and Wiley J Larson. Reducing space mission cost. Microcosm Press Hawthorne, 1996.

[12] Kaushik Sinha and Olivier L de Weck. A network-based structural complexity metric for engineered complex systems. In Systems Conference (SysCon), 2013 IEEE International, pages 426–430. IEEE, 2013.

[13] E. Zio. Reliability analysis of systems of systems. IEEE Reliability Magazine, pages 1–6, February 2016.

[14] Enrico Zio. Challenges in the vulnerability and risk analysis of critical infrastructures. Reliability Engineering & System Safety, 152:137–150, 2016.

[15] Enrico Zio. Critical infrastructures vulnerability and risk analysis. European Journal for Security Research, (DOI 10.1007/s41125-016-0004-2):1–18, 2016.

[16] Enrico Zio. Reliability engineering: Old problems and new challenges. Reliability Engineering & System Safety, 94(2):125–141, 2009.

[17] Yi-Ping Fang and Enrico Zio. Hierarchical modeling by recursive unsupervised spectral clustering and network extended importance measures to analyze the reliability characteristics of complex network systems. American Journal of Operations Research, 3(1A):101–112, 2013.

[18] Zhang Limiao, Li Daqing, Qin Pengju, Fu Bowen, Jiang Yinan, Enrico Zio, and Kang Rui. Reliability analysis of interdependent lattices. Physica A: Statistical Mechanics and its Applications, 452:120–125, 2016.

[19] Enrico Zio and Lucia R Golea. Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. Reliability Engineering & System Safety, 101:67–74, 2012.

[20] Jian Li, Leonardo Dueñas-Osorio, and Changkun Chen. Reliability and controllability of infrastructure networks: do they match? 2015.

[21] Jian Li, Leonardo Dueñas-Osorio, Changkun Chen, and Congling Shi. Connectivity reliability and topological controllability of infrastructure networks: A comparative assessment. Reliability Engineering & System Safety, 156:24–33, 2016.

[22] Wenyan Wu, Holger R Maier, and Angus R Simpson. Multiobjective optimization of water distribution systems accounting for economic cost, hydraulic reliability, and greenhouse gas emissions. Water Resources Research, 49(3):1211–1225, 2013.

[23] Wen-Xu Wang, Xuan Ni, Ying-Cheng Lai, and Celso Grebogi. Optimizing controllability of complex networks by minimum structural perturbations. Physical Review E, 85(2):026115, 2012.

[24] Ching-Tai Lin. Structural controllability. Automatic Control, IEEE Transactions on, 19(3):201–208, 1974.

[25] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. Nature, 473(7346):167–173, 2011.

[26] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Control centrality and hierarchical structure in complex networks. Plos one, 7(9):e44459, 2012.

[27] Gang Yan, Jie Ren, Ying-Cheng Lai, Choy-Heng Lai, and Baowen Li. Controlling complex networks: How much energy is needed? Physical review letters, 108(21):218703, 2012.

[28] Tao Jia and Albert-László Barabási. Control capacity and a random sampling method in exploring controllability of complex networks. Scientific reports, 3, 2013.

[29] Tao Jia, Yang-Yu Liu, Endre Csóka, Márton Pósfai, Jean-Jacques Slotine, and Albert-László Barabási. Emergence of bimodality in controlling complex networks. Nature communications, 4, 2013.

[30] Tamás Nepusz and Tamás Vicsek. Controlling edge dynamics in complex networks. Nature Physics, 8(7):568–573, 2012.

[31] Halit Üster and Şebnem Dilaveroğlu. Optimization for design and operation of natural gas transmission networks. Applied Energy, 133:56–69, 2014.

[32] Ignacio J Ramírez-Rosado and José L Bernal-Agustín. Reliability and costs optimization for distribution networks expansion using an evolutionary algorithm. IEEE Transactions on Power Systems, 16(1):111–118, 2001.

[33] Yiping Fang, Nicola Pedroni, and Enrico Zio. Optimization of cascade-resilient electrical infrastructures and its validation by power flow modeling. Risk Analysis, 35(4):594–607, 2015.

[34] José Emmanuel Ramirez-Marquez and Claudio M Rocco. All-terminal network reliability optimization via probabilistic solution discovery. Reliability Engineering & System Safety, 93(11):1689–1697, 2008.

[35] Darren L Deeter and Alice E Smith. Heuristic optimization of network design considering all-terminal reliability. In Reliability and Maintainability Symposium. 1997 Proceedings, Annual, pages 194–199. IEEE, 1997.

[36] Berna Dengiz, Fulya Altiparmak, and Alice E Smith. Efficient optimization of all-terminal reliable networks, using an evolutionary approach. IEEE transactions on Reliability, 46(1):18–26, 1997.

[37] Manolis Papadrakakis and Nikos D Lagaros. Reliability-based structural optimization using neural networks and monte carlo simulation. Computer methods in applied mechanics and engineering, 191(32):3491–3507, 2002.

[38] HR Baghaee, M Mirsalim, GB Gharehpetian, and HA Talebi. Reliability/cost-based multi-objective pareto optimal design of stand-alone wind/pv/fc generation microgrid system. Energy, 115:1022–1041, 2016.

[39] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE transactions on evolutionary computation, 6(2):182–197, 2002.

[40] Raziyeh Farmani, Godfrey Walters, and Dragan Savic. Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality. Journal of Hydroinformatics, 8(3):165–179, 2006.

[41] Wenyan Wu, Angus R Simpson, and Holger R Maier. Accounting for greenhouse gas emissions in multiobjective genetic algorithm optimization of water distribution systems. Journal of water resources planning and management, 136(2):146–155, 2009.

[42] Yi-Ping Fang, Nicola Pedroni, and Enrico Zio. Comparing network-centric and power flow models for the optimal allocation of link capacities in a cascade-resilient power transmission network. IEEE Systems Journal, 2014.

[43] Rudolf Kalman. On the general theory of control systems. IRE Transactions on Automatic Control, 4(3):110–110, 1959.

[44] Rudolf Emil Kalman. Mathematical description of linear dynamical systems. Journal of the Society for Industrial & Applied Mathematics, Series A: Control, 1(2):152–192, 1963.

[45] Anna Lombardi and Michael Hörnquist. Controllability analysis of networks. Physical Review E, 75(5):056110, 2007.

[46] Zhengzhong Yuan, Chen Zhao, Zengru Di, Wen-Xu Wang, and Ying-Cheng Lai. Exact controllability of complex networks. Nature communications, 4, 2013.

[47] Yoshikazu Sawaragi, HIROTAKA NAKAYAMA, and TETSUZO TANINO. Theory of multiobjective optimization, volume 176. Elsevier, 1985.

[48] Abdullah Konak, David W Coit, and Alice E Smith. Multi-objective optimization using genetic algorithms: A tutorial. Reliability Engineering & System Safety, 91(9):992–1007, 2006.

[49] Pavel Praks, Vytis Kopustinskas, and Marcelo Masera. Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure. Reliability Engineering & System Safety, 144:254–264, 2015.

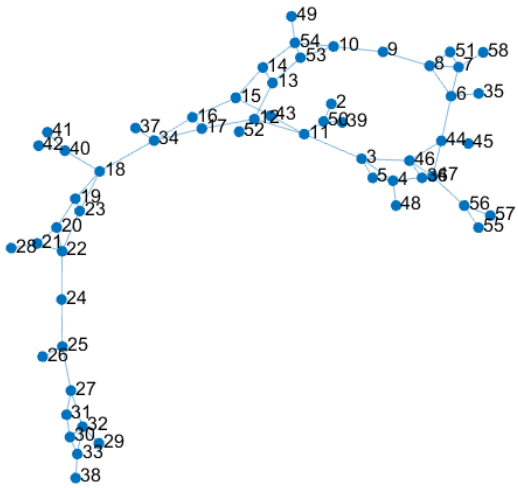
[50] Narsingh Deo. Graph theory with applications to engineering and computer science. Courier Dover Publications, 2016.

[51] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE transactions on pattern analysis and machine intelligence, 26(9):1124–1137, 2004.

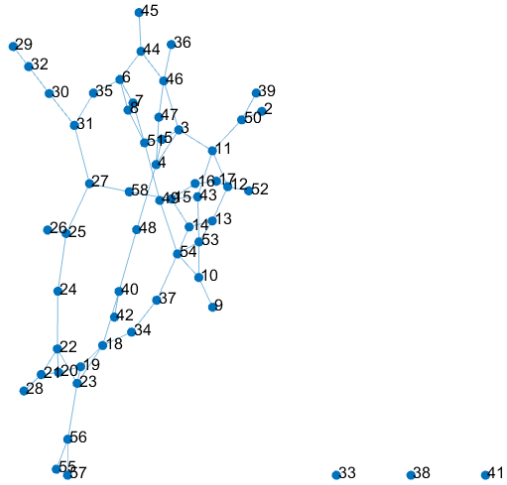
[52] Márton Pósfai, Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Effect of correlations on network controllability. arXiv preprint arXiv:1203.5161, 2012.

[53] M Van der Borst and H Schoonakker. An overview of psa importance measures. Reliability Engineering & System Safety, 72(3):241–245, 2001.

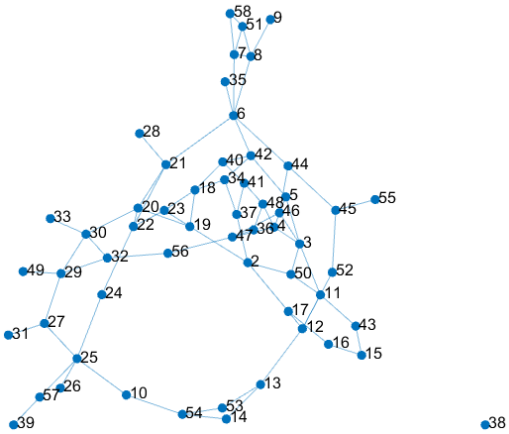
[54] P Praks and V Kopustinskas. Identification and ranking of important elements in a gas transmission network by using progasnet, 2017.



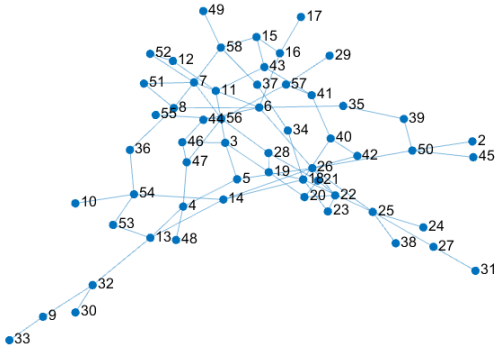
(a) Solution 1



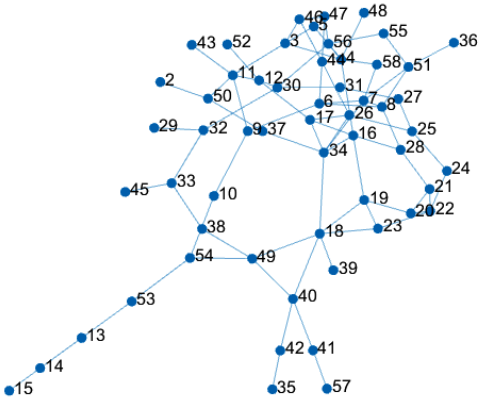
(b) Solution 2



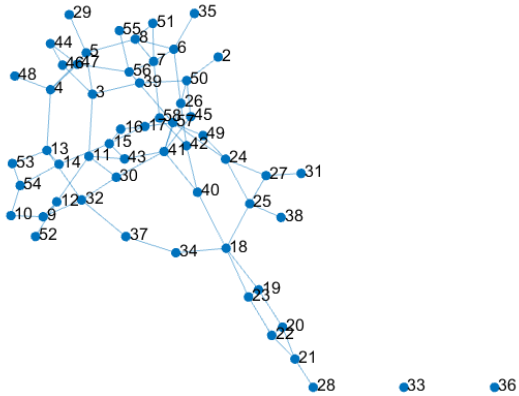
(c) Solution 3



(d) Solution 4



(e) Solution 5



(f) Solution 6

Figure 7: Topology of the 6 Solutions

ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Titre : Cadres pour l'analyse multi-perspective des infrastructures critiques

Mots clés : Infrastructures critiques, Réseau complexe, Fiabilité, Contrôlabilité, Analyse multi-perspective, Optimisation multi-objectif.

Résumé : Les infrastructures critiques (CIs) sont essentielles au fonctionnement de la société moderne. Leur sécurité et leur fiabilité sont les principales préoccupations. La complexité des CIs exige des approches d'analyse de système capables de voir le problème de plusieurs points de vue. La présente thèse porte sur l'intégration de la perspective de contrôle dans l'analyse de sécurité et de fiabilité des éléments de configuration. L'intégration est d'abord abordée par examiner les propriétés de contrôle d'un microgrid d'alimentation électrique. Un schéma basé sur la simulation est développé pour l'analyse sous différentes perspectives : le service d'approvisionnement, la contrôlabilité et la topologie. Un cadre basé sur la commande prédictive (MPC) est proposé pour analyser le microrgrid dans divers scenarios de défaillance. Ensuite, un cadre multi-perspectif est développé pour analyser les CIs considérant le service d'approvisionnement, la contrôlabilité et la topologie. Ce cadre permet d'identifier le rôle des éléments de CIs et de quantifier les conséquences de scénarios de défaillances, par rapport aux différents perspectives considérées. Afin de présenter le cadre d'analyse, un réseau de transport de gaz réel à travers plusieurs pays de l'Union européenne est considéré comme une étude de cas. En fin, un cadre d'optimisation a trois objectifs est proposé pour la conception de CI : la topologie du réseau et l'allocation des capacités de liaison sont optimisées minimisant la demande non fournie et la complexité structurelle du système, et en même temps maximisant la contrôlabilité du système. Une investigation approfondie sur les multiples objectifs considérés est effectuée pour tirer des informations utiles pour la conception du système. Les résultats de cette thèse démontrent l'importance de développer du cadre d'analyse des CIs considérant de plusieurs perspectives pertinentes pour la conception, l'opération et la protection des CIs.

Title : Frameworks for the multi-perspective analysis of critical infrastructures

Keywords : Critical infrastructures, Complex networks, Controllability, Multi-perspective analysis, Multi-objective optimization.

Abstract : Critical infrastructures (CIs) provide essential goods and service for modern society. Their safety and reliability are primary concerns. The complexity of CIs calls for approaches of system analysis capable of viewing the problem from multiple perspectives. The focus of the present thesis is on the integration of the control perspective into the safety and reliability analysis of CIs. The integration is first approached by investigating the control properties of a small network system, i.e., an electric power microgrid. A simulation-based scheme is developed for the analysis from different perspectives: supply service, controllability and topology. An optimization-based model predictive control framework is proposed to analyze the microgrid under various failure scenarios. Then, a multi-perspective framework is developed to analyze CIs with respect to supply service, controllability and topology. This framework enables identifying the role of the CI elements and quantifying the consequences of scenarios of multiple failures, with respect to the different perspectives considered. To demonstrate the analysis framework, a benchmark network representative of a real gas transmission network across several countries of the European Union (EU) is considered as case study. At last, a multi-objective optimization framework is proposed for complex CIs design: design of network topology and allocation of link capacities are performed in an optimal way to minimize the non-supplied demand and the structural complexity of the system, while at the same time to maximize the system controllability. Investigation on the multiple objectives considered is performed to retrieve useful insights for system design. The findings of this thesis demonstrate the importance of developing frameworks of analysis of CIs that allow considering different perspectives relevant for CIs design, operation and protection.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France