ANALYSIS AND PLANNING OF POWER GRIDS

A NETWORK PERSPECTIVE

HALE ÇETİNAY İYİCİL

Analysis and Planning of Power Grids

A Network Perspective

Analysis and Planning of Power Grids

A Network Perspective

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T. H. J. J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 25 Oktober 2018 om 15:00 uur

door

Hale ÇETİNAY İYİCİL

Master of Science in Electrical and Electronics Engineering,
Middle East Technical University, Ankara, Turkey,
geboren te Keçiören, Ankara, Turkey.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie bestaat uit:

Rector Magnificus	voorzitter
Prof. dr. ir. P. F. A. Van Mieghem	Technische Universiteit Delft, promotor
Dr. ir. F. A. Kuipers	Technische Universiteit Delft, copromotor
Onafhankelijke leden:	
Prof. dr. F. M. T. Brazier	Technische Universiteit Delft
Prof. dr. ir. P. M. Herder	Technische Universiteit Delft
Prof. dr. ir. R.E. Kooij	Technische Universiteit Delft
Prof. dr. J. L. Marzo	Universitat de Girona
Prof. dr. ir. B. De Schutter	Technische Universiteit Delft

This research was funded by Alliander N.V.

Keywords: network science, power grids, cascading failures, wind power, sensitivity analyses, targeted attacks, centrality metrics

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Cover design by: Duygu Güroğlu

Copyright © 2018 by H. Çetinay İyicil

ISBN 978-94-6186-969-2

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

To my family.

CONTENTS

Summary xi

Samenvatting xiii

Özet xv

1 Introduction 1
- 1.1 Near-Future Challenges to Power Grids . . . 2
- 1.2 A Network Perspective . . . 3
- 1.3 Research Questions . . . 4
- 1.4 Thesis Outline . . . 4

2 A Topological Investigation of Power Flow 7
- 2.1 Introduction . . . 8
- 2.2 Power Flow Equations . . . 9
- 2.3 Spectral Decomposition of DC Power Flow Equations . . . 10
  - 2.3.1 Solution of DC Power Flow Equations . . . 10
  - 2.3.2 Calculation of the Effective Resistance Matrix . . . 12
- 2.4 Impact of Topology on Power Flow . . . 13
  - 2.4.1 Link Removal . . . 13
  - 2.4.2 Link Addition . . . 15
- 2.5 Numerical Analysis . . . 18
  - 2.5.1 Synthetic Example . . . 18
  - 2.5.2 IEEE 118-Bus Power Grids . . . 19
- 2.6 Conclusion . . . 23

3 Nodal Vulnerability to Targeted Attacks in Power Grids 25
- 3.1 Introduction . . . 26
- 3.2 Power Grids and Network Science . . . 27
  - 3.2.1 Power Grids Preliminaries . . . 27
  - 3.2.2 DC Power Flow Equations . . . 28
  - 3.2.3 Graph Representations of Power Grids . . . 28
- 3.3 Targeted Attacks on Power Grids . . . 30
  - 3.3.1 Ranking Nodes in the Simple Graph Representation of a Power Grid . . . 30
  - 3.3.2 Ranking Nodes in the Weighted Graph Representation of a Power Grid . . . 31
- 3.4 Identifying the Effect of Node Removals in Power Grids . . . 34
  - 3.4.1 Performance Metrics . . . 34
  - 3.4.2 Properties of the Networks Used in Simulations . . . 35
  - 3.4.3 The Effects of Targeted Node Removals in Power Grids . . . 36
  - 3.4.4 Main Lessons Learned from the Analyses . . . 39

3.5 Conclusion . . . 41
3.6 Appendix . . . 41
3.6.1 Targeted Attacks Based on Initial Centrality Metrics . . . 41
3.6.2 Calculation of Flow Betweenness Centrality in Power Grids . . . 41
3.6.3 Multiple Lines Connecting the Same Pair of Nodes . . . 43

4 Topology-Driven Performance Analysis of Power Grids 45
4.1 Introduction . . . 46
4.2 DC Power Flow Analysis in Path and Complete Graphs . . . 46
4.2.1 Electric Power Transmission in a Path Graph . . . 47
4.2.2 Electric Power Transmission in a Complete Graph . . . 48
4.3 DC Power Flow Analysis in Path and Complete Graphs After a Random Link Failure . . . 48
4.3.1 Random Link Failure in a Path Graph . . . 48
4.3.2 Random Link Failure in a Complete Graph . . . 49
4.4 The Impact of Topology on the Key Performance Indicators of Power Grids . . . 51
4.4.1 Key Performance Indicators Under Normal Operation . . . 51
4.4.2 Key Performance Indicators Under a Single Link Failure Contingency . . . 55
4.5 Conclusion . . . 57
4.6 Appendix . . . 58
4.6.1 Operating Conditions in a Path Graph . . . 58
4.6.2 Operating Conditions in a Complete Graph . . . 59
4.6.3 Single Link Failure in a Path Graph . . . 59
4.6.4 Single Link Failure in a Complete Graph . . . 60

5 Effects of Failures in Power Grids under the AC and DC Power Flow Models 61
5.1 Introduction . . . 62
5.2 Power Flow Equations . . . 63
5.2.1 AC Power Flow Equations . . . 63
5.2.2 DC Power Flow Equations . . . 64
5.3 Modeling Cascading Failures . . . 64
5.3.1 AC Cascading Failures Model . . . 66
5.3.2 DC Cascading Failures Model . . . 67
5.4 Numerical Comparison of the AC and DC Flow Models . . . 67
5.4.1 Simulations Setup . . . 67
5.4.2 No Failures Case . . . 69
5.4.3 Comparison of the Single Line Failure Effects . . . 69
5.4.4 Comparison of the Cascade Process I Evolution under the AC and DC Models . . . 72
5.4.5 Comparison Between the Three Cascade Processes under the AC and DC Models . . . 82
5.4.6 Main Lessons Learned from the Simulations . . . 86
5.5 Conclusion . . . 86

6 Planning for Wind Power Integration into Power Grids **89**
6.1 Introduction . . . 90
6.2 Probabilistic Model for the Power Output of a Wind Turbine . . . 91
6.2.1 Wind Speed Characteristics . . . 91
6.2.2 The Power Output of a Wind Turbine . . . 92
6.3 The Assessment of Potential Sites for Wind Farms from an Investor's Point of View . . . 93
6.3.1 Indicators of Suitable Locations for Wind Farms . . . 93
6.3.2 Quantifying the Criteria for Wind Farms . . . 95
6.3.3 Multiple-Criteria Decision Analysis of Wind Farms . . . 95
6.4 Placing Wind Farms under Grid Constraints . . . 96
6.5 Case Study: Wind Farm Planning in Turkey . . . 98
6.5.1 Wind Characteristics . . . 98
6.5.2 Potential Sites for Wind Farms . . . 99
6.5.3 Optimal Sites for Wind Farms . . . 100
6.6 Conclusion . . . 103

7 Conclusion **105**
7.1 Main Contributions . . . 106
7.2 Directions for Future Work . . . 107

References **109**

Acknowledgements **119**

Teşekkür **121**

Curriculum Vitæ **123**

List of Publications **125**

SUMMARY

Electric power has became an essential part of daily life: we plug our electronic devices in, switch our lights on, and expect to have power. As the availability of power is usually taken for granted in modern societies, we mostly feel annoyed at its absence and perceive the importance of power during outages which have severe effects on the public order.

Blackouts have had disastrous consequences for many countries (such as the U.S. and Canada [1], Turkey [2], India [3]) and they continue to occur frequently. In fact, the data from the North American Electrical Reliability Council show that blackouts happen on average every 15 days which leads an economic cost of in the order of tens of billion dollars per year [4]. Such examples demonstrate the necessity for careful analysis and planning of power grids, to ultimately increase the reliability of power grids.

In current practice, flow-based simulations play an essential role in both the security analyses and medium- and long-term planning of power grids. Given the generation and demand profiles, the steady-state analyses estimate the operation of power grids. Additionally, many countries require that the power grids should withstand the scheduled and unscheduled outages of its most critical lines or other components. In these contingency analyses, the component outages are also simulated to determine whether the power grids can still function properly under the failure and consequent loss of an element.

The power grids have evolved due to economic, environmental and human-caused factors. In addition to the contingency analysis, nowadays, the operation and planning of power grids are facing many other challenges (such as demand growth, targeted attacks, cascading failures, and renewable energy integration). Thus, many questions arise, including: which buses (nodes) to connect with a new line (link)? What are the impacts of malicious attacks on power grids? How may an initial failure result in a cascade of failures? How to prepare for the integration of renewable energy? Answering such questions requires developing new concepts and tools for analysing and planning of power grids.

Power grids are one of the largest and the most complex man-made systems on earth. The complex nature of power grids and its underlying structure make it possible to analyse power grids relying on network science [5, 6]. The applications of network science on power grids have shown the promising potential to capture the interdependencies between components and to understand the collective emergent behaviour of complex power grids [7, 8].

This thesis is motivated by the increasing need of reliable power grids and the merits of network science on the investigation of power grids. In this context, relying on network science, we model and analyse the power grid and its near-future challenges in terms of line removals/additions, malicious attacks, cascading failures, and renewable integration. We express the flow behaviour in power grids in terms of graph-related

matrices (Chapter 2), so that we can model power grids as simple and weighted graphs, calculate the centrality of each node in power grids (Chapter 3) and investigate the operation in various graphs (Chapter 4). Furthermore, we provide tools to investigate the current and the near-future challenges of power grids such as link failure/addition (Chapter 2), critical asset identification and targeted attacks (Chapter 3), network expansion and performance analysis (Chapter 4), cascading failures (Chapter 5), and wind power integration (Chapter 6).

The developed concepts in this thesis provide for a better understanding of the operation of the power grid, with the ultimate goal of increasing its reliability. We demonstrate the applicability of our methodologies in the synthetic power grids, in the IEEE-test power grids, and in the real-world power grids. The developed concepts extend the state of the art in the applications of network science on power grids and (i) can be the interest of researchers in the field, (ii) can support grid operators in analysing the vulnerability of their network to the current and the near-future challenges, and (iii) can assist decision makers and investors with the planning for the future trends in power grids.

SAMENVATTING

Elektriciteit is een zeer essentieel onderdeel geworden van ons dagelijks leven. We zijn er aan gewend om elektriciteit te krijgen wanneer we onze elektronische apparaten met het stopcontact verbinden, of wanneer we onze lichten aan doen. De beschikbaarheid van elektriciteit wordt als iets vanzelfsprekends gezien in onze moderne maatschappij, waardoor we voornamelijk geïrriteerd raken bij de afwezigheid ervan. Het belang van elektriciteit wordt des temeer duidelijk tijdens een stroomstoring, wat soms zelfs ernstige gevolgen kan hebben op de maatschappelijke orde.

Vele landen, zoals de V.S. en Canada [1], Turkije [2] en India [3], hebben desastreuze gevolgen ondervonden door stroomstoringen, en soortgelijke gebeurtenissen zetten zich ook tegenwoordig met regelmaat voort. Gegevens van de 'North American Electrical Reliability Council' laten zien dat stroomstoringen gemiddeld eens per 15 dagen gebeuren, en dat de economische gevolgen hiervan in de orde van tientallen miljarden dollars per jaar kunnen oplopen [4]. Dit soort voorbeelden laten de noodzaak zien voor nauwkeurige analyses en planningen van elektriciteitsnetwerken, om zo uiteindelijk de betrouwbaarheid van deze infrastructuur te verbeteren.

In het huidige toepassingsgebied spelen 'flow-based' simulaties een belangrijke rol voor zowel veiligheidsanalyses als medium- tot lange- termijn planningen van het elektriciteitsnetwerk. Bij een gegeven 'generation and demand' ('vraag en aanbod/opwekking') profiel kan via een 'steady-state' analyse een voorspelling gedaan worden van de werking van het net. Bovendien vereisen vele landen dat het elektriciteitsnet bestand moet zijn tegen geplande en ongeplande uitvallen van de meest kritieke elektriciteitslijnen, of van andere cruciale componenten. Ook het uitvallen van deze onderdelen wordt gesimuleerd, om zodoende te bepalen of het elektriciteitsnetwerk nog steeds juist kan functioneren.

Het elektriciteitsnetwerk heeft een grote ontwikkeling ondergaan door economische, milieuvriendelijke, en menselijke invloeden. Bijkomend aan eventualiteit analyses zijn er tegenwoordig vele uitdagingen betreffende de werking en het plannen van het elektrische netwerk (zoals bijvoorbeeld de groeiende vraag naar energie). Dit brengt vele vragen met zich mee. Op welke knooppunten kan een nieuwe lijn of verbinding worden aangesloten? Wat is de impact van kwaadwillige aanvallen op het netwerk? Hoe kan een initiële storing resulteren in het oplopende falen van andere componenten en een uiteindelijk uitval van het netwerk? Hoe moeten we ons voorbereiden op de integratie van hernieuwbare energie in het huidige netwerk? Het beantwoorden van deze vragen verantwoorden de ontwikkeling van nieuwe concepten en hulpmiddelen voor de analysis en planning voor elektrische netwerken.

Elektrische infrastructuren zijn één van de meest complexe door de mens gemaakte systemen op de aarde. De complexe aard van het elektriciteitsnet, en diens onderliggende structuren, maakt het mogelijk om bij de analyse gebruik te maken van 'network science' [5, 6]. Het toepassen van 'network science' op het elektriciteitsnet heeft zeer

veelbelovende resultaten laten zien bij het begrijpen van de interne afhankelijkheden tussen de componenten en bij het begrijpen van het samenhangende gedrag van complexe elektriciteitsnetwerken [7, 8].

De motivatie voor deze dissertatie komt voort uit de toenemende behoefte naar betrouwbare elektriciteitsnetten en de verdiensten van 'network science' bij het onderzoek naar elektrische infrastructuren. Met behulp van 'network science' zullen we electriciteitsnetten en de bijhorende toekomstige uitdagingen analyseren en modeleren met betrekking tot het toevoegen of verwijderen van lijnen, kwaadwillige aanvallen, opeenhopende uitvallen, en de integratie van hernieuwbare energie. Ook brengen we het stromingsgedrag in elektrische netwerken tot uiting via 'graph-related matrices' (hoofdstruk 2), zodat (i) we het elektriciteitsnet kunnen modeleren als simpele en gewogen grafieken, (ii) we de centraliteit van elk knooppunt kunnen berekenen (hoofdstuk 3), en (iii) we de werking ervan kunnen onderzoeken in verschillende grafieken (hoofdstuk 4). Tevens presenteren we hulpmiddelen voor het onderzoek van de uitdagingen betreffende het elektriciteitsnet in het heden en in de nabije toekomst, zoals het toevoegen/verwijderen van verbindingen (hoofdstuk 2), 'critical asset identification' (kritische eigendom identificatie) en doelgerichte aanvallen (hoofdstuk 3), netwerk uitbreidingen (hoofdstuk 4), opeenhopende uitvallen (hoofdstuk 5), en de integratie van wind energie (hoofdstuk 6).

De ontwikkelde concepten in deze dissertatie zorgen voor een beter inzicht in de operatie van elektriciteitsnetwerken, met het verbeteren van de betrouwbaarheid hiervan als uiteindelijk ultiem doel. We laten de toepasbaarheid van onze methodes zien voor synthetische en echte fysieke elektriciteitsnetten, alsmede voor de IEEE power grid test. De ontwikkelde concepten geven een uitbreiding aan de hedendaagse technische en conceptuele mogelijkheden, welke (i) van interesse kunnen zijn voor onderzoekers in dit vakgebied, (ii) een bijdrage kunnen leveren aan netwerk opperanten/exploitanten bij de analyse van netwerk kwetsbaarheden bij toekomstige uitdagingen, en (iii) beleidsmakers en investeerders kunnen helpen bij het plannen van toekomstige trends betreffende elektriciteitsnetwerken.

ÖZET

Elektrik günlük yaşantımızın vazgeçilmez bir parçası haline gelmiştir: elektronik eşyalarımızı prize sokup, aydınlatma anahtarlarına basıp, devrenin tamamlanmasını beklemekteyiz. Modern toplumun gelişmesinde elektriğin varlığının büyük bir etken olduğu kabul edildiği gibi, genellikle elektriğin yokluğunda veya bir arıza olduğunda rahatsız oluruz ve elektrik kesintisinin günlük yaşantımızın üzerinde ne kadar olumsuz etki bıraktığını hissederiz.

Elektrik kesintilerinin birçok ülke üzerinde feci sonuçları olmuştur (A.B.D. ve Kanada [1], Türkiye [2], Hindistan [3]'da görüldüğü gibi) ve olmaya da devam edecektir. Kuzey Amerika Elektrik Güvenilirlik Konseyi'nden elde edilmiş verilere göre her 15 günde bir oluşan elektrik kesintisi yıllık 10 milyar dolarlık bir ekonomik maliyete sebep olmaktadır [4]. Bu tip örnekler, elektrik şebekesinin güvenilirliğini arttırmak için detaylı analiz ve planlamanın yapılmasının önemini göstermektedir.

Günümüzde, şebekenin güvenlik analizlerinde ve kısa veya uzun dönem yatırım planlamasında yük akışı temelli simülasyonlar önemli bir role sahiptir. Durağan durum (steady-state) analizleri, verilmiş olan üretim ve tüketim profillerine göre şebeke işleyişini analiz eder. Buna ek olarak, birçok ülke şebekelerinin planlanmış veya planlanmamış kesintilere dayanıklı olmasını bekler. Bu tip beklenmedik olay analizlerinde, yani bir elemanın kaybında veya hata durumnda, şebekenin hala düzgün olarak çalışıp çalışmadığını bulmak için, şebeke elemanlarının devre dışı kaldığı durumlar analiz edilir.

Elektrik şebekeleri, ekonomik, çevresel, ve sosyal faktörler gibi nedenlerden ötürü evrim geçirmiştir ve geçirmektedir. Günümüzde beklenmedik olay analizlerine ek olarak şebekenin planlanması ve işleyişi çeşitli meydan okumalar ile karşı karşıya kalmaktadır (örneğin elektrik tüketiminin artması, hedeflenmiş kötücül saldırılar, kaskad arızalar ve yenilenebilir enerji entegrasyonu gibi). Böylece: "Hangi hat şebekede hangi noktaya konulmalıdır? Şebekeye hasar verici planlı saldırıların sonuçları nelerdir? Şebekede kaskad arızalar nasıl oluşur? Yenilenebilir enerji entegrasyonuna nasıl hazırlanılmalıdır?" gibi sorular ortaya çıkmaktadır. Bu tip soruları cevaplamak için şebekenin detaylı analizi ve planlaması üzerine yeni konseptler ve araçlar geliştirilmelidir.

Elektrik şebekesi dünya üzerinde bulunan ve insanlar tarafından yapılmış olan en büyük ve en karmaşık yapılardan biridir. Elektrik şebekesinin karmaşık doğası ve altyapısı, onu Ağ Bilimi (Network Science) yolu ile güvenilir şekilde analiz etmeyi mümkün kılmıştır [5, 6]. Ağ biliminin elektrik şebekeleri üzerinde kullanılması, şebekenin karmaşık doğasının ve davranışlarının anlaşılmasında yardımcı olmuştur [7, 8].

Bu tezin motivasyonunu güvenilir elektrik şebekesine olan ihtiyacın artışı ve Ağ biliminin elektrik şebekesi üzerine olan araştırmalarının katkısı oluşturmaktadır. Bu tezde Ağ bilimini kullanarak, şebekeden hat çıkarılması/şebekeye hat eklenmesi, hasar verici kötücül saldırılar, peşi sıra gelen kaskad arızalar, ve şebekeye yenilenebilir enerji entegrasyonu gibi yakın gelecekteki ortaya çıkabilecek problemleri modelliyor ve analiz ediyoruz. Şebekeleri Çizge Kuramı (Graph Theory) ile modelleyerek (Bölüm 2), şebeke

üzerindeki her baranın önemini hesaplayıp (Bölüm 3), çeşitli şebeke dizaynlarının işleyişini araştırıyoruz (Bölüm 4). Ayrıca, şebekenin şuan ve gelecekte yüzleşebileceği hat arızası veya eklemesi gibi meydan okumaları (Bölüm 2), şebekede kritik eleman belirleme ve hedeflenmiş saldırıların etkileri (Bölüm 3), şebekenin genişlemesi ve performans analizi (Bölüm 4), kaskad arızalar (Bölüm 5) ve şebekeye rüzgar gücü entegrasyonunu (Bölüm 6) incelemek için bu tezde araçlar sağlıyoruz.

Bu tezde geliştirilmiş olan konseptler daha güvenilir elektrik şebekesi operasyonuna ulaşmayı amaçlamaktadır. Bu tezdeki metodolojilerimizin uygulamasını sentetik şebekeler, IEEE test şebekeleri ve gerçek şebekeler üzerinde göstermekteyiz. Bu tezde geliştirilmiş konseptler, Ağ Bilimi'nin şebeke üzerine olan uygulamalarının içeriğini arttırmış olup (i) bu bölümdeki araştırmacıların ilgi alanı olabilir, (ii) şebeke operatörlerinin gelecekteki güvenlik açığı taramasına yardımcı olabilir (iii) oparatörlere ve yatırımcılara elektrik şebekesinin geleceğini tahmin etmede yardımcı olabilir.

1

INTRODUCTION

Every time you switch on your lights, do you realise that you are completing an electrical circuit? In fact, this circuit connects your bulb to the lines that serve your house, next, to the thicker lines that serve your neighbourhood and, finally, to a network of high capacity lines that transfer power across the country. This network –the power grid– is complex, costly and crucial to modern societies [9].

It all began with the 19th century inventors who put their small generators next to the machines that needed electricity. Then, in 1882, Thomas Edison presented a system of commercial electric lighting and power with the opening of the Pearl Street station in Manhattan. This system, like most of the systems constructed during the next few years, distributed power within a few kilometres over copper lines using direct current [10]. Thus, it might not be easy to see at that time that the "power grid" of few small power plants serving nearby demands would in the near future develop into a truly interconnected and sophisticated network with more than hundred thousand kilometres of high-voltage transmission lines, different types of power plants and various voltage levels.

Today, the availability of electric power is receiving more attention than ever. The power grid has became an essential part of a modern society. Electric power is vital not only for daily life, but also many other critical infrastructures such as public transportation, telecommunications depend on the power supply [11]. Disruptions to power supply have severe effects on public order and could lead to substantial economic cost for the society.

An example of such severe disruption happened on 14 August 2003. The United States of America experienced the largest blackout event in the North American history [1]. The outage affected nearly 50 million people and power supply was not restored for 4 days in some parts of the United States. This catastrophic blackout, with an estimated cost of $ 6 billion, also contributed to the death of at least 11 people. Other examples of blackouts in different countries [2, 3] illustrate the key importance of a reliable power grid.

1

1.1. NEAR-FUTURE CHALLENGES TO POWER GRIDS

The initial "Edison system" tied its customers to one generator. However, now, the power grid is a highly interconnected and sophisticated network, and it continues evolving. Many different power plants using diverse resources operate across the countries to serve the increasing demand of the customers who could be far from the power plants. In fact, it was reported [12] that the United States nowadays have 285 % more often disruptions to power supply than in 1984. The Government Accountability Office further stated that the reliability and security of power grids are threatened by a number of challenges and as a result, they have become increasingly fragile and vulnerable to extended disruptions [13].

INCREASED USER DEMAND

According to the International Energy Agency data [14, 15], the global power demand has continuously increased and is expected to increase in the next 20 years by nearly 30 %. However, power grids have not been invested accordingly, for mainly because of economic reasons. This continuous imbalance between demand growth and network investments has resulted in an overloaded system, increasing the number of disruptions.

AGEING INFRASTRUCTURE

The power grid is one of the oldest man-made technological systems on earth. In many developed countries, not much has changed after they were built at the beginning of 20^{th} century. In fact, it was stated [16] that 80 % of the grid in some areas of the United States has not been upgraded since 1960s. This over-ageing assets dramatically increase the failure rates in power grids.

MALICIOUS ATTACKS

Most people discuss the drastic effects of potential malicious attacks such as the 2013 attack in California, in which gunmen fired on 17 electrical transformers [17]. Despite these concerns, however, the preventive measures and progress receive relatively less attention. The power grid relies on thousands of points that are often remote and or poorly secured. Deliberate attacks in power grids could have disastrous consequences for the society.

DISTRIBUTED GENERATION

Due to the increasing environmental concerns and the technology-driven trends, nowadays, consumers are rapidly becoming producers of electricity by installing solar panels and wind turbines etc. In fact, Edison Electric Institute expects the capacity of such renewable energy sources to triple before 2040 [12]. This energy revolution and dramatic growth of renewables could negatively affect the power grids in terms of increasing need for reserve capacity, the regional overloading of assets, poor frequency performance and increasing reactive power compensation etc., making power grids more vulnerable to disturbances.

Motivated by the increasing need for reliable power grids and the above-mentioned environmental, economic and human-caused near-future challenges, our objective in

this thesis is to analyse and plan for the challenges of power grids with the ultimate purpose of improving the reliability of power grids.

1.2. A NETWORK PERSPECTIVE

A reliable power grid should be able to supply the electrical demand and comply with the requirements of its customers. This means, as well as the normal operation, it should also withstand the scheduled and unscheduled outages of its most critical lines or other components. Power system analysts often refer to this failure and the consequent loss of an element as the contingency analysis, or $N-1$ analysis. The $N-1$ criterion, i.e., affording the outage of any single component is compulsory for many transmission grid operators [18].

In current practice, power system analysts carry out the security assessments of power grids mainly via flow-based simulations. Under certain demand and generation profiles, analysts use the nonlinear AC and/or linearised DC power flow analyses [19] to estimate the steady-state operation of the power grid. Subsequently, by disabling the particular elements of the network, the impact of the outages on the system are evaluated. These analyses help to understand whether the power grid can properly function for the given generation and demand profile and to investigate what to do under contingencies.

Although $N-1$ and $N-2$ contingency analyses may be possible from the computational point of view, evaluating scenarios where more than three components fail at the same time requires substantial computational time due to the complexity of the simulation models. However, various outages do occur and could result in very large blackouts [1–3]. Thus, additional complementary measures to traditional flow-based assessments are needed to analyse and understand the subtle behaviour of power grids.

Most real-world infrastructures including power grids display non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor random [20]. Researchers define a complex network as a large group of relatively simple components with no central control and where organization and emergent non-trivial behaviour are exhibited. In other words, (i) a complex network has a large number of components (ii) these components interact with each other at different levels, and (iii) as a result, non-trivial system behaviour emerges [21]. As an example, cascading failures are an emergent phenomenon of a power grid, rather than independent and coincidental failures of its individual components [22].

The power grid is now one of the most complex technological networks. The interactions between a large number of components govern the global flow behaviour and the spread of failures. This complex nature of power grids and its underlying structure make it possible to analyse power grids relying on network science [5, 6]. The applications of network science on power grids have shown the promising potential to capture the interdependencies between components and to understand the collective emergent behaviour of complex power grids [7, 8].

Motivated by the above-mentioned observations and the merits of network science in studying power grids, our objective in this thesis is to further model and analyse the operation of power by a network perspective with the final goal of improving the reliability of power grids.

1

1.3. RESEARCH QUESTIONS

This thesis is motivated by the increasing need of reliable power grids and the merits of network science on the investigation of power grids. We regard the term *reliability* as a beneficial property for a power grid that refers to its ability to supply electric loads with a high level of probability, both during normal operations and under random or targeted failures. Similarly, we use the term *robustness* to measure to what extent a power grid has this reliability in terms of different metrics. In this context, relying on network science, the aims of this dissertation are to model and analyse the power grid and its near-future challenges in terms of line removals/additions, malicious attacks, network expansion, cascading failures, and renewable integration. Thus, the main questions this thesis aims to answer are:

Chapter 2: Can we express the flow behaviour in power grids in terms of graph-related matrices? What are the sensitivities of link flows to the topological changes in power grids? How can we identify the critical line removals and additions in power grids?

Chapter 3: How can we model power grids as a graph? How can we extend the existing centrality metrics so that they can also include the flow behaviour in power grids? What is the impact of targeted node removals in power grids?

Chapter 4: Can we derive the analytic expressions of power transmission in path and complete graphs? Is a complete graph the ultimate topology for power grids? How can we compare the impacts of power transmission in different graph types?

Chapter 5: How can one predict the evolution of cascading failures in power grids? How may an initial failure result in a cascade of failures? What are the effects of cascading failures in power grids?

Chapter 6: How can we model the long-term wind speed and wind power characteristics? What are the factors determining the location of a probable wind farm? How can we plan for the integration of wind farms into power grids?

1.4. THESIS OUTLINE

This dissertation consists of 7 chapters.

Chapter 2 combines the fundamentals of power grids with graph theory. This section expresses the linearised DC power flow equations via graph-related matrices and further derives the effective resistance matrix and the sensitivities of link flows to the changes in network topology.

Chapter 3 presents two different graph models for power grids as simple and weighted graphs. This chapter further calculates the centrality metrics of each node and analyses the critical node removals in power grids.

Chapter 4 uses a weighted graph model for power grids and in various graph types, this chapter further investigates the electric power transmission under the normal operation and under a link failure contingency using both the linearised DC and nonlinear AC power flow equations.

Chapter 5 focuses on the link failures in power grids. This chapter develops models to simulate cascading failures in power grids and investigates the effects of link failures in power grids under the linearised DC and the nonlinear AC power flow models.

Chapter 6 focuses on the renewable integration into the power grids. This chapter

presents models for the probabilistic wind speed and wind power characteristics, and further investigates the potential locations of wind farms and their integration into the power grids.

Chapter 7 concludes this dissertation by reflecting on the assessments throughout the previous chapters with the focuses of the merits of the results and possible future directions of the field.

2

A Topological Investigation of Power Flow

This chapter combines the fundamentals of an electrical network, such as the flow allocation according to Kirchhoff's laws and the effect of electrical impedance, with spectral graph theory. We express the linearised DC power flow equations using weighted graph matrices and investigate the relation between the topology and the flow behaviour of power grids. Based on the pseudo-inverse of the weighted Laplacian matrix, we further derive the effective resistance matrix in power grids and the sensitivities of active power flows to the changes in the network topology by means of link removal and link addition.

This chapter is based on a published paper [23].

2.1. INTRODUCTION

THE unavailability of electric power can severely disrupt daily life and result in substantial economic and social costs [24]. This key importance of electric power supply encourages a robust design and a careful operation of power grids [18]. Grid operators assess power system security and analyse the system's critical components during both under regular operations, but also under the events of component failures or planning to add new components.

The use of network science has opened the door to a new direction in analysing power grids, namely, a complex network perspective [5–7, 25]. A significant number of studies that deploy complex networks investigate the relationship between the topology and the specific performance metrics of power grids [7, 26, 27]. Various metrics [28] are proposed to assess the vulnerability of power grids [7, 8, 29], and to identify its critical elements [30, 31]. Most of these studies are based on classical topology metrics (such as nodal degree and clustering coefficient [8, 29]), which ignore the electrical properties, such as the flow allocation according to Kirchhoff's laws or the impedance values of transmissions elements in power grids.

Two different aspects are important in the distribution of power flows, and the consequent system's vulnerability, in an electrical grid: the operating state, including the supply and demand dispatches of the system, and the topology of the network formed by electric busbars (or busses) and their interconnection. Accordingly, some studies propose extended topological metrics (such as effective graph resistance and netability [27, 30, 32]) that reflect some of the electrical properties of grids, and some studies introduce combined topological and operational algorithms to identify critical lines [33]. Through empirical studies, those metrics, based on effective resistance, have been shown to perform better in assessing the vulnerability of power grids than purely topological approaches. Motivated by this fact and results from empirical studies with extended graph metrics, this chapter presents an analytical approach to the distribution of flows in power grids that directly analyses the impact of the topology on those flows. First, a slack-bus independent representation of power flow behaviour is introduced. Next, a closed-form expression for the effective resistance (Thevenin) matrix, which represents the topology as well as the power flow allocation behaviour, is derived. Those formulae allow the computation of the redistribution of power flows under the changes of network topology, and they provide fine-grained analysis of critical elements in power grids.

The work presented in this chapter only makes one approximation: the linearisation of the power flow equations resulting in the so-called DC power flow equations [34], which facilitates the use of enhanced linear algebra and graph theory leading to expressions that may simplify the design of robust power grids. In particular, the contributions of this chapter are: (i) A slack-bus independent expression for the linearised power flow. (ii) An analytical derivation of the effective resistance (Thevenin) matrix of a power grid. (iii) Expressions for the pseudo-inverse of the Laplacian and the redistribution of the power flow under link removal/addition.

The remainder of this chapter is organized as follows: Section 2.2 provides details about power grids and the steady-state power flow equations. Section 2.3 introduces a spectral graph perspective on the linearised power flow equations and calculates the ef-

fective resistance matrix in power grids. Section 2.4 develops expressions for the pseudo-inverse of the weighted Laplacian and the sensitivities of active power flows to link removal/addition. Section 2.5 illustrates the proposed formulations in IEEE 118-bus power grids and Section 2.6 concludes this chapter.

2.2. POWER FLOW EQUATIONS

Power grids consist of electrical buses and interconnecting elements (transmission lines and transformers). The status of each bus i can be represented by its voltage $v_i = |v_i|e^{\mathbf{i}\theta_i}$ in which $|v_i|$ is the voltage magnitude, θ_i is the phase angle, and $\mathbf{i}$ denotes the imaginary unit. In the steady-state of a power grid with N buses, the injected apparent power $s_i = p_i + \mathbf{i}q_i$ at bus i, where p_i is the active power and q_i is the reactive power, is calculated using the *AC power flow equations* [19]:

$$p_i = \sum_{k=1}^{N} |v_i||v_k|(y_{ik}^{(\mathrm{R})}\cos\theta_{ik} + y_{ik}^{(\mathrm{I})}\sin\theta_{ik}) \tag{2.1}$$

$$q_i = \sum_{k=1}^{N} |v_i||v_k|(y_{ik}^{(\mathrm{R})}\sin\theta_{ik} - y_{ik}^{(\mathrm{I})}\cos\theta_{ik}) \tag{2.2}$$

where $\theta_{ik} = \theta_i - \theta_k$ and $y_{ik}^{(\mathrm{R})}$ and $y_{ik}^{(\mathrm{I})}$ are the real and the imaginary parts of the element in the bus admittance matrix $\mathbf{Y}$ corresponding to the i^{th} row and k^{th} column, respectively. The AC power flow equations are solved to obtain voltage magnitude $|v_i|$ and voltage angle θ_i information for each bus i in power grids.

The AC power flow equations (2.1) and (2.2) are non-linear and the solution process is generally iterative. A linear set of equations is more desirable whenever fast and repetitive solutions are needed. Linearisation can be reasonably accurate when the following conditions are met [34]:

1. The difference between the voltage phase angles of two neighbouring buses is small so that $\sin\theta_{ik} \simeq \theta_{ik}$ and $\cos\theta_{ik} \simeq 1$.

2. Line resistances compared to the line reactances are negligible which causes the entries of the bus admittance matrix $\mathbf{Y}$ to be equal to the reciprocal of line reactance values, b_{ik}.

3. The variations in the bus voltage magnitudes are so small that they are assumed to be all equal to the selected system base.

4. Reactive power flows are ignored.

If these conditions are approximately met, (2.1) can be simplified to the *DC power flow equations*:

$$p_i = \sum_{k=1}^{N} b_{ik}(\theta_i - \theta_k) = b_{ik}\theta_{ik} \tag{2.3}$$

where b_{ik} is the reciprocal of the reactance between bus i and bus k.

Since **Y** is not invertible, (2.3) cannot be directly solved by inversion. The common procedure is to select a bus i as a reference bus or *slack-bus*, and drop the equation corresponding to its power injection. Then, the remaining equations of phase angles can be solved uniquely with respect to the slack-bus.

2.3. SPECTRAL DECOMPOSITION OF DC POWER FLOW EQUATIONS

This section introduces a spectral graph perspective [35] on the linearised DC power flow equations and applies the concept of the effective resistance [36] to power grids.

2.3.1. SOLUTION OF DC POWER FLOW EQUATIONS

A power grid with N buses, and L transmission lines and transformers is a complex network, whose underlying topology can be represented by a graph $G(\mathcal{N},\mathcal{L})$, where $\mathcal{N}$ denotes the set of N nodes and $\mathcal{L}$ denotes the set of L links. The $N\times N$ adjacency matrix **A** specifies the interconnection pattern of the graph $G(\mathcal{N},\mathcal{L})$: $a_{ik}=1$ only if the pair of nodes i and k are connected by a link; otherwise $a_{ik}=0$. The DC power flow equations (2.3) can be written in terms of the adjacency matrix of $G(\mathcal{N},\mathcal{L})$ as

$$p_i=\sum_{k=1}^{N}a_{ik}b_{ik}(\theta_i-\theta_k)=\theta_i\sum_{k=1}^{N}a_{ik}b_{ik}-\sum_{k=1}^{N}a_{ik}b_{ik}\theta_k$$

where b_{ik} is the reciprocal of the line reactance between the nodes i and k.

The effects of transmission line reactances are represented by the weighted adjacency matrix **W**, where each element $w_{ik}=a_{ik}b_{ik}$ is the weight of the link between nodes[1] i and k:

$$p_i=\theta_i\sum_{k=1}^{N}w_{ik}-\sum_{k=1}^{N}w_{ik}\theta_k. \tag{2.4}$$

Since (2.4) holds for every node i in the graph, the corresponding matrix representation is

$$\begin{aligned}\mathbf{P}&=\left\{\mathrm{diag}\left(\sum_{k=1}^{N}w_{ik}\right)-\mathbf{W}\right\}\Theta\\&=(\mathbf{D}-\mathbf{W})\Theta\end{aligned} \tag{2.5}$$

where $\mathbf{P}=[p_1,\ldots,p_N]^{\mathrm{T}}$ is the vector of net active power injection at the nodes under a balanced power flow i.e., $\mathbf{u}^{\mathrm{T}}\mathbf{P}=0$ where **u** is an all-one vector, **D** is the weighted degree diagonal matrix, and $\Theta=[\theta_1,\ldots,\theta_N]^{\mathrm{T}}$ is the vector of voltage phase angles. Finally, introducing the weighted Laplacian $\tilde{\mathbf{Q}}=\mathbf{D}-\mathbf{W}$ into (2.5) yields

$$\mathbf{P}=\tilde{\mathbf{Q}}\Theta \tag{2.6}$$

where the weighted Laplacian is a symmetric, positive semi-definite matrix that possesses nonnegative eigenvalues apart from the smallest eigenvalue, which is zero [35].

[1] Parallel links connecting the same pair of nodes are replaced by a single link with equivalent reactance calculated from Ohm's law.

The solution to the DC power flow equations requires finding unknown voltage phase angles at each node for the given supply and demand values, $\mathbf{P}$. Due to the zero eigenvalue of $\tilde{\mathbf{Q}}$, the matrix equation in (2.6) cannot be inverted. However, using spectral decomposition [35], the real and symmetric matrix $\tilde{\mathbf{Q}}$ can be written as $\tilde{\mathbf{Q}} = \mathbf{X}\Lambda\mathbf{X}^{\mathbf{T}}$, where $\Lambda = \text{diag}(\mu_j)_{1\leq j\leq N}$ and $\mathbf{X} = [\mathbf{x_1}, \ldots, \mathbf{x_N}]$ is an orthogonal matrix formed by the eigenvectors $\mathbf{x_1}, \ldots, \mathbf{x_N}$ of $\tilde{\mathbf{Q}}$ corresponding to the eigenvalues $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_N = 0$. The eigenvector $\mathbf{x_j}$ is normalised as $\mathbf{x_j}^{\mathbf{T}}\mathbf{x_j} = 1$. Then, expanding $\tilde{\mathbf{Q}}$

$$\tilde{\mathbf{Q}} = \sum_{j=1}^{N} \mu_j \mathbf{x_j}\mathbf{x_j^T} = \sum_{j=1}^{N-1} \mu_j \mathbf{x_j}\mathbf{x_j^T} + \frac{\mu_N}{N}\mathbf{u}\mathbf{u}^{\mathbf{T}} = \sum_{j=1}^{N-1} \mu_j \mathbf{x_j}\mathbf{x_j^T}$$

where $\mathbf{u}$ is the all-one vector, shows that the last equation corresponding to $\mu_N = 0$ can be omitted. Proceeding with the symmetric $N \times N$ matrix $\hat{\mathbf{Q}} = \hat{\mathbf{X}}\text{diag}(\mu_k)\hat{\mathbf{X}}^{\mathbf{T}}$, where the $N \times (N-1)$ matrix $\hat{\mathbf{X}}$ consists of all the eigenvectors of $\tilde{\mathbf{Q}}$ except the eigenvector $\mathbf{u}$ belonging to $\mu_N = 0$, and where the $(N-1) \times (N-1)$ diagonal matrix $\text{diag}(\mu_k)$ contains the positive eigenvalues of $\tilde{\mathbf{Q}}$, the inverse of $\hat{\mathbf{Q}}$ can be found as

$$\hat{\mathbf{Q}}^{-\mathbf{1}} = (\sum_{k=1}^{N-1} \mu_k \mathbf{x_k}\mathbf{x_k^T})^{-1} = \sum_{k=1}^{N-1} \frac{1}{\mu_k}\mathbf{x_k}\mathbf{x_k^T} = \mathbf{Q}^{\dagger}$$

where the $N \times N$ matrix $\mathbf{Q}^{\dagger} = \hat{\mathbf{X}}\text{diag}(\mu_k^{-1})\hat{\mathbf{X}}^{\mathbf{T}}$ is the pseudo-inverse of the Laplacian obeying

$$\begin{aligned}\mathbf{Q}^{\dagger}\tilde{\mathbf{Q}} &= \sum_{k=1}^{N-1} \frac{1}{\mu_k}\mathbf{x_k}\mathbf{x_k^T} \sum_{j=1}^{N-1} \mu_j \mathbf{x_j}\mathbf{x_j^T} \\ &= \sum_{k=1}^{N-1}\sum_{j=1}^{N-1} \mu_j \frac{1}{\mu_k}\mathbf{x_k}(\mathbf{x_k^T}\mathbf{x_j})\mathbf{x_j^T} = \mathbf{I} - \frac{1}{N}\mathbf{J}\end{aligned}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ the all-one matrix.

Using $\mathbf{Q}^{\dagger}$, the pseudo-inversion of (2.6) gives

$$\Theta = \mathbf{Q}^{\dagger}\mathbf{P}. \tag{2.7}$$

Equation (2.7) physically means that only the differences of voltage phase angles between the nodes matter for the power flow. Additionally, an average value of 0 has been chosen as reference for the node voltage phase angles and, consequently, the concept of slack-bus [19] becomes redundant, as a reference is already included in the graph matrix representation.

For the link flows, the active power flow f_{ik} through the link between nodes i and k can be calculated using (2.3)

$$f_{ik} = b_{ik}(\theta_i - \theta_k). \tag{2.8}$$

As (2.8) holds for every link, the corresponding matrix equation is

$$\mathbf{F} = \tilde{\mathbf{B}}^{\mathbf{T}}\Theta \tag{2.9}$$

where the $L \times 1$ vector $\mathbf{F} = [f_1, \ldots, f_L]^{\mathrm{T}}$ is the active power flow through the network links and $\tilde{\mathbf{B}}$ is the $N \times L$ weighted incidence matrix of the graph with the elements

$$\tilde{b}_{il} = \begin{cases} w_{ik} & \text{if link } e_l = i \rightarrow k, \\ -w_{ik} & \text{if link } e_l = i \leftarrow k, \\ 0 & \text{otherwise.} \end{cases}$$

Combining (2.7) and (2.9) results in the final equation for the active power flows through the graph links:

$$\mathbf{F} = \tilde{\mathbf{B}}^{\mathbf{T}} \mathbf{Q}^{\dagger} \mathbf{P}. \tag{2.10}$$

The above equation represents, assuming that the DC power flow approximation is sufficiently accurate, the relation between the active power flows through the network links under the given supply and demand values $\mathbf{P}$, and the graph-related weighted matrices $\tilde{\mathbf{B}}$ and $\mathbf{Q}^{\dagger}$.

2.3.2. CALCULATION OF THE EFFECTIVE RESISTANCE MATRIX

In graph theory, the resistance distance between a pair of nodes is the potential difference between those two nodes in an electrical network, when a unit current is injected at one node and leaves the network at the other node [36, 37]. In power grids, there are supply and demand nodes and, under the DC power flow assumptions, active power flows through the network lines resulting in phase angle differences. This analogy enables the introduction of the concept of the effective resistance matrix Ω with the elements Ω_{ab} to capture the relation between the voltage phase angle and injected active power:

$$\theta_a - \theta_b = \Omega_{ab}\, p_{ab} \tag{2.11}$$

where p_{ab} is the active power injected into the network at node a and leaving from node b, and θ_a and θ_b are the phase angles at nodes a and b, respectively.

Introducing equation (2.7) into (2.11) gives

$$(\mathbf{e_a} - \mathbf{e_b})^{\mathrm{T}} \Theta = (\mathbf{e_a} - \mathbf{e_b})^{\mathrm{T}} \mathbf{Q}^{\dagger} p_{ab} (\mathbf{e_a} - \mathbf{e_b}) \tag{2.12}$$

where $\mathbf{e_k}$ is the basic vector with the m^{th} component equal to 1 if $m = k$, else 0, and the effective resistance (or Thevenin resistance) Ω_{ab} between nodes a and b can be expressed as

$$\Omega_{ab} = (\mathbf{e_a} - \mathbf{e_b})^{\mathrm{T}} \mathbf{Q}^{\dagger} (\mathbf{e_a} - \mathbf{e_b}). \tag{2.13}$$

Multiplying out the right hand side of (2.13) yields

$$\Omega_{ab} = (\mathbf{Q}^{\dagger})_{aa} + (\mathbf{Q}^{\dagger})_{bb} - 2(\mathbf{Q}^{\dagger})_{ab} \tag{2.14}$$

from which the symmetric effective resistance matrix Ω of the electrical grid can be calculated as

$$\Omega = \mathbf{z}\mathbf{u}^{\mathbf{T}} + \mathbf{u}\mathbf{z}^{\mathbf{T}} - 2\mathbf{Q}^{\dagger} \tag{2.15}$$

where the vector $\mathbf{z} = [(\mathbf{Q}^{\dagger})_{11}, (\mathbf{Q}^{\dagger})_{22}, \ldots, (\mathbf{Q}^{\dagger})_{NN})]^{\mathrm{T}}$.

The effective resistance matrix allows to introduce the concept of electrical flow distance rather than physical distances or link weights in a graph. A strong electrical connection between a pair of nodes results in a low effective resistance [30].

2.4. IMPACT OF TOPOLOGY ON POWER FLOW

As shown so far, the electric power flow depends on the network topology as well as on the power input. In this section, effective resistances will be used to capture the flow distribution under the changes in network topology.

2.4.1. LINK REMOVAL

An electrical grid is expected to tolerate the loss of any single component at any time (which is called the $N-1$ criterion [18]). Due to the loss of a network component, the power in the electrical grid will be redistributed, and the resulting situation can lead to an increase or a decrease in the flow through a particular network link. The link removal that causes increases in remaining link flows needs to be carefully studied and necessary measures should be taken to avoid cascading failures [30].

Existing flow-based studies in power grids require the solution of system equations for each contingency under each loading scenario. Thus, computationally effective alternatives are needed, and power transfer (PTDF) and line outage (LODF) distribution factors are often used [38, 39]. These metrics capture the relative change in the flow through a particular link, after a change in injection and corresponding withdrawal at a pair of nodes (PTDF) or after a line outage (LODF). These direct calculations decrease the computation time, yet it is not possible to reflect the drivers of flow behaviour as the formulations are generally result-oriented and based on reduced matrices in the absence of the slack-bus(es). In this section, we analyse link removals in power grids from a topological point of view using the graph-related matrices in previous section.

When an arbitrary link l_{ij} in an electrical grid is removed, the network topology is changed. Following the definition of weighted adjacency matrix in Section 2.3.1, the removal of the link between the nodes i and j zeroes the entries w_{ij} and w_{ji} in the new weighted adjacency matrix, whereas the other elements remain unchanged. As a result, the weighted Laplacian will be affected in the i^{th} and j^{th} rows by the weight of the removed link on the diagonal entry and j^{th} and i^{th} columns, respectively. The relation between the old and new weighted Laplacians is essentially a rank-one update:

$$\tilde{\mathbf{Q}}' = \tilde{\mathbf{Q}} - w_{ij}\,(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}} \tag{2.16}$$

where $\tilde{\mathbf{Q}}$ is the initial Laplacian, $\tilde{\mathbf{Q}}'$ is the Laplacian after the removal of link l_{ij}, and w_{ij} is the weight of the removed link. Introducing Meyer's relation [40] between the pseudo-inverses denoted by †,

$$(\mathbf{A}+\mathbf{cd}^{\mathbf{T}})^{\dagger} = \mathbf{A}^{\dagger} - (1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{\dagger}\mathbf{c})^{-1}\mathbf{A}^{\dagger}\mathbf{cd}^{\mathbf{T}}\mathbf{A}^{\dagger} \tag{2.17}$$

allows to express the pseudo-inverse $\mathbf{Q}'^{\dagger}$ of the new Laplacian in (2.16) in terms of the

initial pseudo-inverse $\mathbf{Q}^{\dagger}$ and effective resistances in (2.13) as

$$\begin{aligned}
\mathbf{Q}'^{\dagger} &= \left(\mathbf{Q}^{\dagger} + (-w_{ij})(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\right)^{\dagger} \\
&= \mathbf{Q}^{\dagger} - \left(1+(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}(-w_{ij})(\mathbf{e_i}-\mathbf{e_j})\right)^{-1}\mathbf{Q}^{\dagger}(-w_{ij})(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger} \\
&= \mathbf{Q}^{\dagger} + \frac{w_{ij}}{1-w_{ij}(\mathbf{e_i}-\mathbf{e_j})^{T}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger} \\
&= \mathbf{Q}^{\dagger} + \frac{w_{ij}}{1-w_{ij}\Omega_{ij}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}
\end{aligned} \tag{2.18}$$

where Ω_{ij} is the effective resistance between nodes i and j.

When link l_{ij} is removed, the flow f_{ij} through the link before removal is redistributed over alternative paths between nodes i and j. Under the DC power flow approximation, which results in (2.10) being linear, the redistribution can be perceived as an additional injection of active power f_{ij} at node i and leaving node j in the new network, provided that the supply and demand values of the electrical grid, $\mathbf{P}$, remain unchanged. Hence, the final flow through an arbitrary link l_{ab} can be written as the sum of the previous state of the system, i.e., the previous flow through the link between nodes a and b when link l_{ij} is present, and the flow resulting from the change of the state due to link removal. Consequently, the change of the flow through the observed link l_{ab} can be calculated using (2.10) as

$$\Delta f_{ab} = w_{ab}(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}'^{\dagger}(\mathbf{e_i}-\mathbf{e_j})f_{ij} \tag{2.19}$$

where Δf_{ab} is the change in the flow through link l_{ab} due to removal of link l_{ij}, and w_{ab} is the weight of link l_{ab}. Inserting (2.18) into (2.19) results in

$$\begin{aligned}
\Delta f_{ab} &= f_{ij}w_{ab}(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}'^{\dagger}(\mathbf{e_i}-\mathbf{e_j}) \\
&= f_{ij}w_{ab}(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}(\mathbf{Q}^{\dagger} + \frac{w_{ij}}{1-w_{ij}\Omega_{ij}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger})(\mathbf{e_i}-\mathbf{e_j}) \\
&= f_{ij}w_{ab}\big((\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j}) + \frac{w_{ij}}{1-w_{ij}\Omega_{ij}}(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})\big) \\
&= f_{ij}w_{ab}(1+\frac{\Omega_{ij}w_{ij}}{1-w_{ij}\Omega_{ij}})(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j}).
\end{aligned}$$

Since $(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j}) = \frac{1}{2}(\Omega_{aj}-\Omega_{ai}+\Omega_{bi}-\Omega_{bj})$ according to (2.14), we have

$$\Delta f_{ab} = f_{ij}w_{ab}\frac{\Omega_{aj}-\Omega_{ai}+\Omega_{bi}-\Omega_{bj}}{2(1-w_{ij}\Omega_{ij})}$$

or

$$\frac{\Delta f_{ab}}{f_{ij}} = w_{ab}\frac{\Omega_{aj}-\Omega_{ai}+\Omega_{bi}-\Omega_{bj}}{2(1-w_{ij}\Omega_{ij})}. \tag{2.20}$$

Equation (2.20) shows that, due to the removal of link l_{ij}, the resultant change in the flow through a remaining link l_{ab} is determined by the network topology via the effective resistances between the node pairs, and the previous flow f_{ij} through the removed link. Several observations follow from equation (2.20):

- The resulting flow change Δf_{ab} through a link l_{ab} depends on and is limited by the magnitude of the previous flow f_{ij} through the removed link l_{ij}. Since the flow f_{ij} is redistributed through the network, it holds that $\left|\frac{\Delta f_{ab}}{f_{ij}}\right| \leq 1$, which forces the right-hand side of equation (2.20) to be between -1 and 1.
- If the directions of the links are defined to be the same as the direction of the initial flow through the links, a positive (negative) number in the right hand side of (2.20) indicates an increase (decrease) in the flow through the remaining link in that direction.
- From a robustness point of view, the network links whose removal increases the flows through the remaining links are critical. In addition, the network links that are consistently affected by different link removal scenarios are also critical.
- For the network links whose flows are not affected by the removal, the right-hand side of equation (2.20) must be 0, meaning the equality $\Omega_{aj} + \Omega_{bi} = \Omega_{ai} + \Omega_{bj}$ between the effective resistances of node pairs is satisfied. This equality is satisfied for the links that are in different branches[2] of the graph and for Wheatstone bridges [41] if they are present in the network.
- The denominator $(1 - w_{ij}\Omega_{ij})$ of (2.20) is zero when the effective resistance between the nodes of the link is equal to the inverse of the link weight, i.e., line reactance. It shows that there is no alternative parallel (back-up) path in the graph for the removed link. Therefore, when this link is removed, some nodes in the graph will be isolated and the underlying graph of the network will be partitioned, which can disturb the balance between supply and demand of the network. In this case, the flow cannot be redistributed without the change of supply and demand values, thus, the change in the flows through the links cannot be calculated solely from the topological values. Thus, in such a case, we rewrite (2.20) as

$$\frac{\Delta f_{ab}}{f_{ij}} = \begin{cases} \text{Network islanded (N.I.)} & \text{if } w_{ij}\Omega_{ij} = 1, \\ w_{ab}\frac{\Omega_{aj}-\Omega_{ai}+\Omega_{bi}-\Omega_{bj}}{2(1-w_{ij}\Omega_{ij})} & \text{otherwise.} \end{cases} \tag{2.21}$$

Equation (2.21) captures the final network status, i.e., islanded or not, as well as the effect of link removal on the distribution of flows through the remaining network links, when the network is not partitioned. The calculation is based on the initial graph-related matrices, and the computation of new topological matrices is avoided. Consequently, by spectral decomposition, once the effective resistance matrix is calculated, the effect of any link removal can be calculated from (2.21) for any balanced supply and demand values.

2.4.2. LINK ADDITION

The overloads in the transmission lines of power grids can be solved by supply or demand shifting in the short term. However, a long-term investment (such as addition of

[2] Here, the branch of a graph refers to its subgraph that is radially connected to the other parts of the graph, such as Barbell-like graphs.

new transmission lines) needs to be planned in the case of persistent overloads or to satisfy the $N-1$ criterion [18].

Determining the right location of a new link is challenging. It is desirable that the added link increases the robustness of the electrical grid by decreasing the critical flows through the network links. In flow-based studies, the computational complexity is high, thus alternatives which decrease the calculation time and determine the right investment for the system are sought [27], as provided in this section.

It is assumed that a new link can be added between any arbitrary two nodes i and j in the graph. Similar to Section 2.4.1, the redistribution of flows due to the link addition can be perceived in the initial network as an additional injection of the active power f_{ij} over the new link at node j and leaving from node i, i.e., in the opposite direction of the new flow. The change in the flow Δf_{ab} on an arbitrary network link l_{ab} under the DC power flow approximation is calculated as

$$\Delta f_{ab} = w_{ab}(\mathbf{e_a}-\mathbf{e_b})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(-f_{ij})$$

and, using (2.14), as

$$\frac{\Delta f_{ab}}{f_{ij}} = \frac{w_{ab}(\Omega_{ai}-\Omega_{aj}+\Omega_{bj}-\Omega_{bi})}{2} \tag{2.22}$$

where the flow f_{ij} on the new link is calculated by using the new pseudo-inverse $\mathbf{Q}'^{\dagger}$ of the Laplacian and the power input $\mathbf{P}$ of the network:

$$f_{ij} = w_{ij}(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}'^{\dagger}\mathbf{P}. \tag{2.23}$$

The addition of the link changes the Laplacian of the network and the relation between the new $\tilde{\mathbf{Q}}'$ and the old Laplacian $\tilde{\mathbf{Q}}$ becomes

$$\tilde{\mathbf{Q}}' = \tilde{\mathbf{Q}} + (w_{ij})\ (\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}.$$

Relation (2.17) shows that the new pseudo-inverse can be represented as

$$\begin{aligned}\mathbf{Q}'^{\dagger} &= (\tilde{\mathbf{Q}} + (w_{ij})(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}})^{\dagger}\\ &= \mathbf{Q}^{\dagger} - (1+(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}(w_{ij})(\mathbf{e_i}-\mathbf{e_j}))^{-1}\mathbf{Q}^{\dagger}(w_{ij})(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}\\ &= \mathbf{Q}^{\dagger} - \frac{w_{ij}}{1+w_{ij}\Omega_{ij}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}.\end{aligned} \tag{2.24}$$

Using the above derivation (2.24) of the new pseudo-inverse $\mathbf{Q}'^{\dagger}$, equation (2.23) can be rewritten as

$$\begin{aligned}f_{ij} &= w_{ij}\ (\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\big(\mathbf{Q}^{\dagger} - \frac{w_{ij}}{1+w_{ij}\Omega_{ij}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})(\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}\big)\mathbf{P}\\ &= w_{ij}\ \big((\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}\mathbf{P}\big)\big(1-\frac{w_{ij}}{1+w_{ij}\Omega_{ij}}\ (\mathbf{e_i}-\mathbf{e_j})^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e_i}-\mathbf{e_j})\big)\\ &= w_{ij}\ \theta_{ij}\big(1-\frac{w_{ij}\Omega_{ij}}{1+w_{ij}\Omega_{ij}}\big)\\ &= \frac{w_{ij}}{1+w_{ij}\Omega_{ij}}\theta_{ij}.\end{aligned} \tag{2.25}$$

Equation (2.25) shows that the new flow f_{ij} through the added link l_{ij} is related to the previous network conditions, i.e., the difference between the voltage phase angles at nodes i and j, and inversely related to the effective resistance between these nodes. As θ_{ij} and Ω_{ij} are fixed by the initial network topology, the maximum flow through the added link, $\frac{|\theta_{ij}|}{\Omega_{ij}}$, is achieved when the link weight w_{ij} tends to infinity, meaning that the reactance of the transmission line is close to zero, a short circuit of the nodes. Conversely, the flow through the new link is minimum, 0, when w_{ij} approaches zero, meaning connecting an infinite reactance between the nodes (an open circuit). Then, by adjusting the link weight w_{ij} through reconducting or replacing the conductors, it is theoretically possible to adjust the magnitude of the flow through the added link.

The term $\frac{w_{ij}}{1+w_{ij}\Omega_{ij}}$ in the right-hand side of equation (2.25) is strictly positive for passive network elements. Thus, the direction of the flow through the new link is determined only by the difference between the voltage phase angles θ_{ij} in the initial network. A positive difference in voltage phase angles θ_{ij} results in a flow from node i to node j, when the nodes are connected by a link, whereas the opposite results in a flow from node j to node i. If the voltage phase angle difference θ_{ij} is zero, there will be no flow through the link when these nodes are connected by a link (Wheatstone bridge [41]).

Inserting the result (2.25) of the flow through the new link into (2.22), the change in the flow through the observed link l_{ab} due to link addition can be calculated as

$$\Delta f_{ab} = \frac{w_{ab}w_{ij}(\Omega_{ai}-\Omega_{aj}+\Omega_{bj}-\Omega_{bi})}{2(1+w_{ij}\Omega_{ij})}\theta_{ij}. \tag{2.26}$$

Equation (2.26) shows that the change Δf_{ab} in the flow through the network links is determined by the network topology via the effective resistances and initial network conditions, whereas the relative change to the flow f_{ij} through the new link in (2.22) depends only on the network topology. Observations from (2.22) and (2.26) are as follows:

- The change in the flow through network links depends linearly on the flow through the added link and the changes in flows through the initial network links are bounded by this value. When the right-hand side of (2.22) is 1 or -1, it means that the flow through the observed link is directly affected by the link addition.
- The numerator of equation (2.22) is zero when the equality between the effective resistances $\Omega_{ai}-\Omega_{aj}=\Omega_{bi}-\Omega_{bj}$ is satisfied, meaning the added link has no effect on the flow through the observed network link. This is possible for the observed and added links that are in different branches of the network.
- If the direction of the link is defined as the direction of the initial flow through that link, then a positive (negative) number in the right-hand side of (2.26) corresponds to an increase (decrease) in the flow through the observed link in that direction. Clearly, a decrease in the flow through all network lines is desired.

Finally, equations (2.20) and (2.22) show that the effective resistances between the node pairs of the observed and the removed/added links determine the effect on the flow through the observed link. This aligns with the empirical studies that capture the relation between the effective graph resistance value [36] and the robustness of the power

grid against cascading failures [27], [30]. Additionally, the weight of the observed link w_{ab} is found to be influential in both link removal and addition calculations, whereas the weight of the added link w_{ij} does not affect the flow through the observed link relative to the flow through the added link.

From the graph-related matrices, the changes (2.20) and (2.22) in the flow through the network links relative to flow through the removed/added link can be represented. However, for the magnitude of the change, initial conditions, the supply and demand values of the network, must be known. The direction of the change in the flow through the observed link, i.e., decrease or increase in magnitude, is also determined by both the network topology and the power input of the electrical network as it depends on the existing flow and its direction. However, in electrical grids with limited generation and load variations, such as directed networks, it is possible to know the flow directions in advance. Therefore, from the effective resistances, the relative effect on the magnitude can be found. For the meshed networks with various supply and demand units, the flow directions may be unknown. Therefore, initial network conditions, the voltage phase angles, or the power input of the network must be used in the calculations regarding the direction.

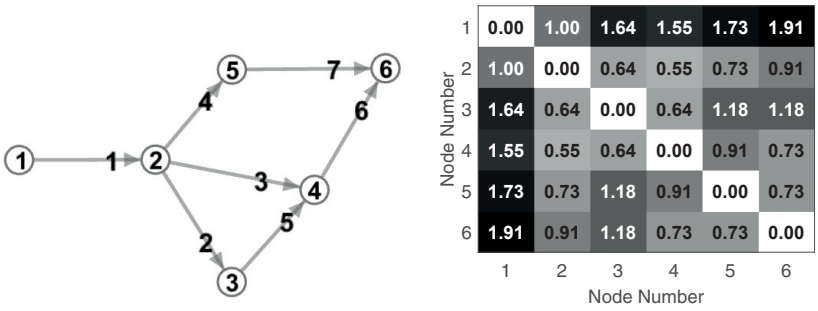
2.5. NUMERICAL ANALYSIS

This section demonstrates the results derived in previous Sections 2.3 and 2.4. For ease of inspection, first a quantitative analysis is performed for a small test network. Later, the analysis is demonstrated for the IEEE 118-bus power grids [42].

2.5.1. SYNTHETIC EXAMPLE

The network in Figure 2.1a contains 6 nodes and 7 links. For simplicity, the link weights, i.e., the reciprocal of line reactances, are set to unity. The direction of the existing flows through the links is defined to be always from lower to higher node index. The effective resistance matrix Ω is calculated according to (2.15) and the effective resistances are shown in Figure 2.1b. The minimum effective resistance is between nodes 2 and 4, whereas the largest is between nodes 1 and 6. The definition of electrical distance in (2.11) shows that the highest difference in the voltage phase angles of the network nodes occurs when the electric power is transferred between those nodes, leading to larger flows through the network links from equation (2.8). Conversely, the minimum difference in the voltage phase angles of the network nodes occurs for the same amount of power when it is transferred between nodes 2 and 4, leading to smaller flows through the network links.

Next, the effect of a link removal on the flows through the remaining network links is calculated using (2.20). Figure 2.2a illustrates how the flows through the network links are affected by a particular link removal, as compared to the previous flow through the removed link. As an example, when link 6 is removed from the network, due to the redistribution of flows, the flows through links 4 and 7 increase by the amount of the previous flow through the removed link 6. Indeed, this makes the removal of link 6 critical. In order to avoid cascading failures, it must be checked whether the excess capacity of links 4 and 7 can handle the redistributed flow. For the network links 2, 3 and 5, the removal of



(a) Graph (b) Effective resistances

Figure 2.1: Test network and its effective resistance values.

link 6 decreases the flow through them, thus there is no possibility of cascading failure due to these links.

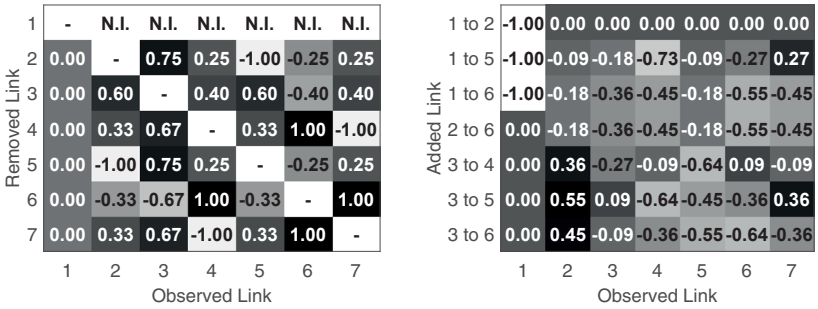
Finally, from (2.22), the effect of a link addition is calculated. Figure 2.2b displays some examples of the changes in the flows through the network links in case of a link addition, as compared to the flow through the added link. For instance, when a new link is added between nodes 2 and 6, the flows through all network links decrease except for link 1, which is connected to a pendant node. In addition, depending on the purpose of the new investment (link addition), Figure 2.2b can be used to identify the place of the added link. For example, if the aim is to decrease the flow through link 5 between nodes 3 and 4, three choices are effective: A new link parallel to link 5, a new link between nodes 3 and 5, or a new link between nodes 3 and 6 significantly decrease the flow, whereas the addition of a new link between nodes 1 and 5 has a relatively small effect on the observed link for the same amount of new flow. In some cases, the addition of new links can lead to an increase in the flow through a particular link. For instance, when a new link is added between nodes 3 and 6, the flow through link 2 increases considerably, which is the so-called Braess' paradox in power systems [43]. Therefore, such cases should be avoided or be carefully investigated before realization.

2.5.2. IEEE 118-BUS POWER GRIDS

In this section, a more realistic IEEE 118-bus power grid is considered. Figure 2.3 shows the graph representation of the network, containing in total 118 nodes and 179 links. The direction of existing flows through the links is defined according to initial conditions.

The histogram of effective resistances between all node pairs is shown in Figure 2.4a and between the observed set in Figure 2.4b, respectively. The effective resistances in the observed set are relatively small, which suggests a strong electrical connection, whereas the larger values of effective resistances in Figure 2.4a suggest the opposite, indicating the points with less back-up paths in the network.

The effect of each link removal on remaining network links is calculated using (2.21) resulting in 179 link removal cases, each with 178 observed links. The histogram of the effects of link removals relative to the flow through the removed link is shown in Fig-



(a) Link removal (b) Link addition

Figure 2.2: Effects of link removal and addition, $\frac{\Delta f_{ab}}{f_{ij}}$, in the synthetic test network.

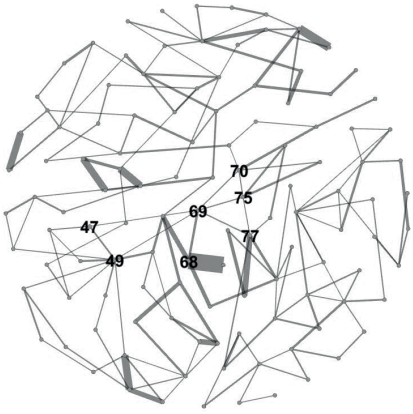
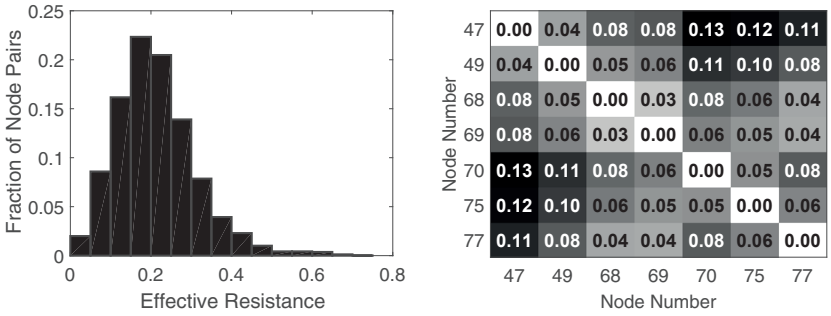


Figure 2.3: Graph representation of the IEEE 118-bus power grids. The thicknesses of the links represent the link weights, i.e., the inverse of line reactances. The average degree in the graph is 3.034, whereas the average weighted degree is 59.759. The network diameter is 14 and the average path length is 6.309. The links connected to node 69, an important generator bus serving 12% of the total demand, and their node pairs are chosen to be the observed set (a, b).



(a) All node pairs (b) Observed set

Figure 2.4: Effective resistances in the IEEE 118-bus power grids.

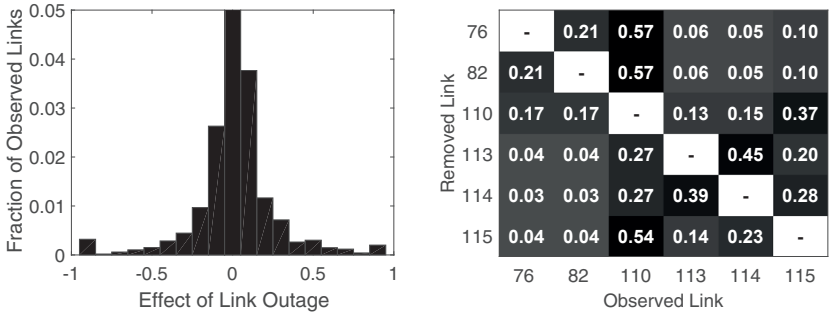
ure 2.5a. Approximately 95% of the effects have magnitude smaller than 0.2, which is a sign of a meshed network with alternative paths. However, in 3.8% of the effects, equation (2.20) results in 0, which refers to network links which are in branches of the network.

In Figure 2.5a, 0.17% of the effects of the link removals have the value 1, meaning that the previous flow through the removed link is transferred to a single alternative path. From a robustness point of view, the less frequent this is, the more robust is the network against overloads due to link removals. Therefore, these cases should be analysed in reliability assessments. Additionally, the removal of 9 links leads to isolation of one or more nodes in the network, which is again undesirable in a robust network.

In Figure 2.5b, the effect of link removals in the observed set is shown. When link 76, 82 or 115 is removed from the network, more than half of the redistributed flow goes through link 110 between nodes 68 and 69, which makes link 110 critical. As a remark, when a link is removed from the observed set, the magnitudes of changes in the flows through remaining links must sum up to the previous flow through the removed link according to Kirchhoff's law, therefore the row sums in Figure 2.5b are all 1.

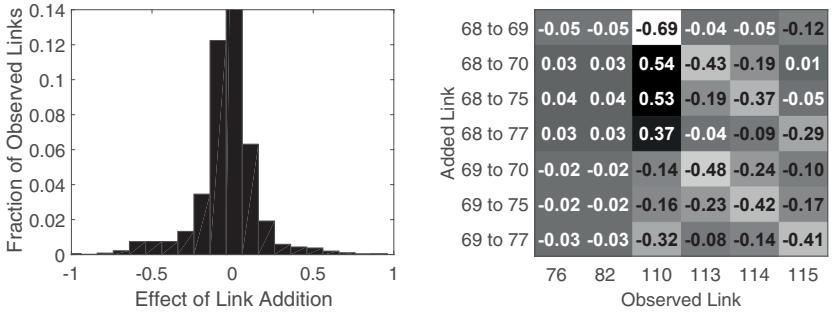
From (2.22), the effect of a link addition between each node pair in the network is calculated, resulting in $\frac{118\times117}{2} = 6903$ link addition cases, each with 179 observed links. The histogram of the effects of all possible link additions relative to the flow through the added link is presented in Figure 2.6a. 92% of the effects have magnitude smaller than 0.2, which again follows from the meshed topology and the existence of alternative paths for the redistributed flow. Due to the meshed topology, a link addition to the network can increase the flows through the network links. However, the probability of an increase in magnitude is less than compared to the probability of a decrease, which can be observed from the asymmetrical distribution in Figure 2.6a.

In Figure 2.6b, the effect of link addition in the observed set is shown. Similar to the link removal case, when a link connected to node 69 is added, the observed relative changes in the magnitudes must sum up to 1. The magnitude of the flow through link 110 may increase in 3 out of the 7 illustrated link additions, which urges detailed assessments before realization of these link additions in order to avoid Braess' paradox.



(a) All cases (b) Observed set

Figure 2.5: The effect $\frac{\Delta f_{ab}}{f_{ij}}$ of link removals in the IEEE 118-bus power grids. In Figure 2.5a, the peak corresponding to $-0.05 \leq \frac{\Delta f_{ab}}{f_{ij}} \leq 0.05$ is 0.83.



(a) All cases (b) Observed set

Figure 2.6: The effect $\frac{\Delta f_{ab}}{f_{ij}}$ of link addition in the IEEE 118-bus power grids. In Figure 2.6a, the peak corresponding to $-0.05 \leq \frac{\Delta f_{ab}}{f_{ij}} \leq 0.05$ is 0.70.

2.6. CONCLUSION

This chapter provided an extended graph approach to analyse the physical operation of a power grid from a topological point of view. Contrary to the traditional representation, the linearised DC power flow equations were expressed in terms of slack-bus-independent graph matrices. A closed formula for the effective resistance matrix, which combines the fundamentals of a power grid with the topological structure, was illustrated. We further derived the expressions for the sensitivities of link flows to link removal/addition cases to assess the topological vulnerability of power grids. Consequently, link removals that may result in cascading failures or node isolation and link additions that can decrease the critical flows or result in Braess' paradox in the power grid could be identified.

3

Nodal Vulnerability to Targeted Attacks in Power Grids

The previous chapter combines the fundamentals of power grids with topological matrices and uses the pseudo-inverse of weighted Laplacian to introduce a slack-bus independent solution to the DC power flow equations. In this chapter, we model the power grid as a simple graph, and as a weighted graph based on these flow equations. For both graph models, we present the centrality metrics of each bus (node) in a power grid. Subsequently, we formulate different node-attack strategies based on these centrality metrics and empirically analyse the impact of targeted-node attacks on the structural and the operational performance of power grids. We perform case studies in the high-voltage transmission networks of 5 European countries and in commonly used IEEE test power grids.

This chapter is based on published papers [44] and [45].

3.1. INTRODUCTION

THE unavailability of electrical power can severely disrupt daily life and result in substantial economic and social costs [24]. This vital importance encourages a robust design and operation of power grids [18]. Robust power grids are able to anticipate, adapt to and/or rapidly recover from a disruptive event or a failure.

Disruptions in networks can be caused by unintentional failures or intentional attacks. Unintentional failures can include manufacturing defects, malfunction in network elements or human error. These kinds of failures occur randomly throughout the grid and are characterized as random failures [46]. Intentional attacks or *targeted attacks*, on the other hand, are not random and are aimed at maximizing damage [47]. A major challenge in power grids is to evaluate the vulnerability of a power system to these intentional hazards, starting by quantifying the importance of electrical buses and the impact of the attacks on the network performance.

Topological investigations of power grids have demonstrated that power grids have several components with significant importance compared to the rest of the network [48]. These components are crucial for the grid as their removal can significantly disrupt the operation of the power grids. Identifying these critical components in advance can enable power grid operators to improve system robustness by monitoring and protecting these components continuously. [48, 49]

Currently, many studies use a complex networks perspective in analysing power system vulnerabilities [7, 25, 50]. A significant part of these studies investigates the relationship between the topology and specific performance metrics in the underlying graph of power grids [7, 26, 29]. Such studies focus on the basic structural properties of a graph (such as nodal degree, clustering coefficient [28]), which typically ignore the electrical properties, such as flow allocation according to Kirchhoff's laws or the impedance values of transmissions lines in the grid. Mainly, two different aspects are important in the operation and consequent robustness of power grids: the topology of the network formed by electrical buses and their interconnections, and the operating conditions such as supply and demand distributions [23, 32]. Consequently, these purely topological metrics could result in misleading research results, which may be far from real physical behaviours of power grids [48, 51, 52].

To include the electrical properties of the grid in the analyses, several studies propose extended metrics (such as effective graph resistance [27], the electrical centrality [53] and the net-ability [32]) by introducing a set of link weights (such as distance or resistance [54]) and node properties (such as the electrical demand and supply [32]). Additionally, other studies have used topological and electrical metrics to rank the electrical buses and lines in power grids as a selective contingency analysis [55, 56].

Motivated by the increasing need of alternative studies to the flow-based analyses and the merits of network science on the investigation of power grids, in this chapter, we combine both of the aforementioned approaches: First, we present two different graph representations for a power grid in Section 3.2: a simple graph and an extended graph representation that takes the electrical properties of power grids into account. Next, we develop a methodology to identify the critical electrical buses (nodes) in power grids in Section 3.3, and compare the impact of targeted node-attacks in detail for European high-voltage transmission networks [57] and for the publicly available IEEE test power

grids [42] in Section 3.4. Our contributions can be summarized as follows: (i) we consider two different graph models for power grids based on either purely topological information or by including the link weight information and the linearised DC power flow equations; (ii) we employ these two graph models to formulate the standard and the extended centrality metrics of nodes in power grids; (iii) we formulate 8 different attack scenarios according to these centrality metrics and empirically investigate the impact of targeted node-attacks on the structural and operational performance of power grids.

3.2. POWER GRIDS AND NETWORK SCIENCE

In this section, we provide details about power grids, the steady-state power flow equations and our models for power grids as simple and weighted graphs.

3.2.1. POWER GRIDS PRELIMINARIES

Power grids consist of nodes (electrical buses) and interconnecting links (transmission lines and transformers). The status of each node i is represented by its voltage $v_i = |v_i|e^{\mathbf{i}\theta_i}$ in which $|v_i|$ is the voltage magnitude, θ_i is the phase angle, and $\mathbf{i}$ denotes the imaginary unit. Each line l has a predetermined capacity $\mathscr{C}_l$ that bounds its flow f_l under a normal operation of the system. In the steady-state of a power grid with N nodes and L links, the injected apparent power $p_i + \mathbf{i}q_i$ at node i, where p_i is the active power and q_i is the reactive power, is calculated using the AC power flow equations [19]:

$$p_i = \sum_{k=1}^{n} |v_i||v_k|(y_{ik}^{(\mathrm{R})}\cos\theta_{ik} + y_{ik}^{(\mathrm{I})}\sin\theta_{ik}) \tag{3.1}$$

$$q_i = \sum_{k=1}^{n} |v_i||v_k|(y_{ik}^{(\mathrm{R})}\sin\theta_{ik} - y_{ik}^{(\mathrm{I})}\cos\theta_{ik}) \tag{3.2}$$

where $\theta_{ik} = \theta_i - \theta_k$ and $y_{ik}^{(\mathrm{R})} = \mathrm{Re}(y_{ik})$ and $y_{ik}^{(\mathrm{I})} = \mathrm{Im}(y_{ik})$ are the real and the imaginary parts of the element y_{ik} in the bus admittance matrix $\mathbf{Y}$ corresponding to the i^{th} row and k^{th} column, respectively.

Each node in a power grid contains a number of electrical devices and according to those, two basic types of nodes can be defined [58]:

- Supply node: A supply node generates the active power p_i and controls the voltage magnitude $|v_i|$ at its node i.
- Demand node: At a demand node, it is possible to specify the extracted active p_i and the reactive powers q_i from the type of the electrical loads that are connected to that node. There are also nodes without a supply or a demand connected, which can be modelled as a demand node with no injected power, i.e., $p_i = 0$ and $q_i = 0$.

Due to the impedance of transmission elements, there are power losses during the operation in power grids. As the losses are dependent on the system state –the supply and demand dispatches –they cannot be calculated in advance. Therefore, a slack node among the supply nodes is assigned in power grids to compensate for the difference between the total supply and the total demand plus the losses.

3.2.2. DC POWER FLOW EQUATIONS

The AC power flow equations (3.1) and (3.2) are non-linear and the solution process is generally iterative. A linear set of equations is more desirable whenever fast and repetitive solutions are needed. Linearisation can be reasonably accurate when the following conditions are met [34, 59]:

1. The difference between the phase angles of neighbouring nodes is small such that $\sin\theta_{ik} \approx \theta_{ik}$ and $\cos\theta_{ik} \approx 1$.
2. The active power losses are negligible, and therefore, the bus admittance matrix can be approximated as $\mathbf{Y} \approx \mathbf{i}\mathbf{Y}^{(\mathrm{I})}$ where $\mathbf{Y}^{(\mathrm{I})}$ is the imaginary part of the admittance matrix $\mathbf{Y}$, calculated neglecting the line resistances.
3. The variations in the voltage magnitudes $|v_i|$ are small and, can be assumed as $|v_i| = 1$ for all nodes.

If these conditions are approximately met, the AC power flow equations can be simplified to the so-called the DC power flow equations:

$$p_i = \sum_{k=1}^{N} y_{ik}^{(\mathrm{I})}(\theta_i - \theta_k). \tag{3.3}$$

Although the DC power flow solution is less accurate than the AC power flow solution, in practice, the differences in high-voltage transmission networks between the phase angles of neighbouring buses and the variations in voltage magnitudes are relatively small, thus the error is assumed to be negligible [34].

3.2.3. GRAPH REPRESENTATIONS OF POWER GRIDS

This section presents our models for power grids as simple and weighted graphs.

POWER GRID AS A SIMPLE GRAPH

A simple graph is an unweighted, undirected graph containing no self-loops or multiple links. A power grid can be modelled as a graph $G(\mathcal{N}, \mathcal{L})$ where $\mathcal{N}$ denotes the set of N nodes and $\mathcal{L}$ denotes the set of L links in which multiple lines connecting the same pair of nodes are modelled as one link. The $N \times N$ adjacency matrix $\mathbf{A}$ specifies the interconnection pattern of the graph $G(\mathcal{N}, \mathcal{L})$: $a_{ik} = 1$ only if the pair of nodes i and k are connected by a direct link; otherwise $a_{ik} = 0$. The $N \times N$ Laplacian matrix $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \Delta - \mathbf{A}$$

where $\Delta = \mathrm{diag}(d_1, \ldots, d_N)$ is the diagonal degree matrix with the diagonal elements $d_i = \sum_{k=1}^{N} a_{ik}$.

POWER GRID AS A WEIGHTED GRAPH

Alternatively, a power grid can be modelled as a weighted graph where each link is assigned a weight that is related to the admittance of the transmission line it represents. We model a power grid as a weighted graph $G(\mathcal{N}, \mathcal{L})$ where $\mathcal{N}$ denotes the set of N nodes and $\mathcal{L}$ denotes the set of L links[1]. By writing the DC power flow equations in (3.3)

[1] Multiple lines connecting the same pair of nodes are represented as a single equivalent link in the graph, see Appendix 3.6.3.

in terms of the adjacency matrix $\mathbf{A}$ of $G(\mathcal{N},\mathcal{L})$

$$p_i = \sum_{j=1}^{N} a_{ij} y_{ik}^{(\mathrm{I})}(\theta_i - \theta_j) = \theta_i \sum_{j=1}^{N} a_{ij} y_{ik}^{(\mathrm{I})} - \sum_{j=1}^{N} a_{ij} y_{ik}^{(\mathrm{I})} \theta_j \tag{3.4}$$

we introduce the weighted adjacency matrix $\tilde{\mathbf{A}}$, where each nonzero element $\tilde{a}_{ij} = a_{ij} y_{ik}^{(\mathrm{I})}$ represents both the connectivity and the admittance between nodes i and j. Equation (3.4) can then be written as:

$$p_i = \theta_i \sum_{j=1}^{N} \tilde{a}_{ij} - \sum_{j=1}^{N} \tilde{a}_{ij} \theta_j. \tag{3.5}$$

Since (3.5) holds for every node i, the corresponding matrix representation is

$$\begin{aligned} \mathbf{P} &= \left\{ \mathbf{diag}\left(\sum_{k=1}^{N} \tilde{a}_{ij} \right) - \tilde{\mathbf{A}} \right\} \Theta \\ &= (\tilde{\Delta} - \tilde{\mathbf{A}})\Theta \end{aligned} \tag{3.6}$$

where $\mathbf{P} = [p_1 \dots p_N]^{\mathrm{T}}$ is the vector of net active power injection at the nodes with a balanced supply and demand, i.e., $\mathbf{u}^{\mathrm{T}}\mathbf{P} = 0$ where $\mathbf{u}$ is all-one vector, $\tilde{\Delta}$ is the weighted diagonal degree matrix, and $\Theta = [\theta_1 \dots \theta_N]^{\mathrm{T}}$ is the vector of phase angles at the nodes. Finally, introducing the weighted Laplacian $\tilde{\mathbf{Q}} = \tilde{\Delta} - \tilde{\mathbf{A}}$ into (3.6) yields

$$\mathbf{P} = \tilde{\mathbf{Q}}\Theta \tag{3.7}$$

where the weighted Laplacian $\tilde{\mathbf{Q}}$ is a symmetric, positive semi-definite matrix that possesses non-negative eigenvalues apart from the smallest eigenvalue, which is zero [35].

The inversion of the *active power - phase angle* relation $\mathbf{P} = \tilde{\mathbf{Q}}\Theta$ in (3.7) is not possible due to the fact that $\det \tilde{\mathbf{Q}} = 0$, which follows from the characteristic property $\tilde{\mathbf{Q}}\mathbf{u} = 0$ of the weighted Laplacian. Although the inverse of the weighted Laplacian matrix does not exist, the *active power - phase angle* relation inversion can be shown to be $\Theta = \mathbf{Q}^{\dagger}\mathbf{P} + \frac{\mathbf{u}^{\mathrm{T}}\Theta}{N}\mathbf{u}$, where $\mathbf{Q}^{\dagger}$ is the pseudo-inverse of the weighted Laplacian $\tilde{\mathbf{Q}}$, obeying $\tilde{\mathbf{Q}}\mathbf{Q}^{\dagger} = \mathbf{Q}^{\dagger}\tilde{\mathbf{Q}} = \mathbf{I} - \frac{1}{N}\mathbf{J}$ with the identity $\mathbf{I}$ and all-one matrix $\mathbf{J} = \mathbf{u}\mathbf{u}^{\mathrm{T}}$. By choosing the average phase angle in the graph $\theta_{\mathrm{av}} = \frac{\mathbf{u}^{\mathrm{T}}\Theta}{N} = 0$ as the reference [23], the *phase angle - active power* relation takes the elegant form of

$$\Theta = \mathbf{Q}^{\dagger}\mathbf{P}. \tag{3.8}$$

While the weighted Laplacian $\tilde{\mathbf{Q}}$ and its pseudo-inverse $\mathbf{Q}^{\dagger}$ are derived here based on the linearised DC power flow equations in power grids, their applicability is far wider [45]. A weighted Laplacian $\tilde{\mathbf{Q}}$ can describe many processes, that are linear in or proportional to the network topology such as electrical circuits, water flow networks, mechanical or thermal systems. The process equivalence between those systems are given in Table 3.1.

Power grids	Phase angle	Power
Electrical circuit	Voltage	Current
Hydraulic circuit	Pressure (height of liquid)	Volume flow
Mechanical system	Force	Displacement velocity
Thermal system	Temperature	Heat flow
...	...	...

Table 3.1: Equivalence between linear systems, adopted from [45].

3.3. TARGETED ATTACKS ON POWER GRIDS

The threats for power grids can be classified by using multiple criteria considering the causes of the threat, their consequences or the preventive actions to manage the hazards [60]. One example of such threats are targeted attacks on power grids, which involve intentional, criminal actions to destroy the network. In modelling these threats, we assume that the attacks are performed with the knowledge of power grid layout and with the intention to maximally disrupt the network performance while attacking as few nodes (electrical buses) as possible. Throughout this section, we describe how network science can be employed to formulate such attack strategies, where target nodes correspond to most critical or most vulnerable nodes whose removal significantly disrupts the network functioning. We first describe the standard centrality metrics, which are purely based on the underlying topology of power grids, and then we extend these metrics to include the information on the link weights, i.e., the admittances of the transmission lines, and the DC power flow equations in power grids.

3.3.1. RANKING NODES IN THE SIMPLE GRAPH REPRESENTATION OF A POWER GRID

The topological investigation of electrical power grids demonstrates that they have scale free characteristics, suggesting that power grids can have *hub* components with significant criticality compared to the rest of the network [6]. These components are crucial for the grid. Identifying these critical components in a power grid is a major concern for power system security. In graph theory and network science, centrality metrics are used to identify the most important nodes within a graph. Different applications can include identifying the most influential person(s) in a social network, main spreaders of a disease, etc.

In this section, we review some of the existing topological centrality metrics in order to rank the importance and the centrality of nodes in the underlying simple graph of power grids.

DEGREE CENTRALITY

The degree d_i of a node i in the graph $G(\mathcal{N},\mathcal{L})$ is equal to the number of its neighbouring nodes [61]. The degree d_i can be calculated using the adjacency matrix $\mathbf{A}$:

$$d_i = \sum_{j=1}^{N} a_{ij}. \tag{3.9}$$

EIGENVECTOR CENTRALITY

The eigenvector centrality of a node is a global centrality metric that depends not only on the number of its neighbouring nodes, but also on the number of 2-hop neighbouring nodes, 3-hop neighbouring nodes, and so on [62, 63]. The eigenvector centrality x_i of node i is equal to the i^{th} component of the eigenvector corresponding to the largest eigenvalue λ_1 of the adjacency matrix **A**. The principal eigenvector centralities thus follow from the linear equations:

$$x_i = \frac{1}{\lambda_1} \sum_{k=1}^{N} a_{ik} x_k. \tag{3.10}$$

BETWEENNESS CENTRALITY

Another metric to assess node importance or centrality is the betweenness centrality [64]. In calculating the betweenness centrality, it is assumed that information or services are transmitted over shortest paths between node pairs. Hence, if many shortest paths pass through a certain node, this node takes a central role in the network. If $|\mathscr{P}_{s\to t}|$ is the number of all possible shortest paths from node s to node t, and $|\mathscr{P}_{s\to t}(i)|$ is the number of those paths that pass through node i, then the betweenness b_i of node i is equal to

$$b_i = \sum_{s,t\in\mathscr{N}\setminus\{i\}} \frac{|\mathscr{P}_{s\to t}(i)|}{|\mathscr{P}_{s\to t}|}. \tag{3.11}$$

In other words, the betweenness centrality of a node i shows the fraction of all shortest paths between any pair (s,t) of nodes, that pass through node i.

CLOSENESS CENTRALITY

In calculating the closeness centrality, the hopcount $H(\mathscr{P}_{i\to j})$ that is the number of links in the shortest path $\mathscr{P}_{i\to j}$ between a pair of nodes i and j, is used. The closeness centrality c_i of a node i is defined as [64]:

$$c_i = \frac{1}{\sum_{j\neq i} H(\mathscr{P}_{i\to j})}, \tag{3.12}$$

which is the reciprocal of the sum of the hopcounts of node i to all other nodes. A large closeness centrality value thus corresponds to a "central" node that is well-connected by a few hops to other nodes.

3.3.2. RANKING NODES IN THE WEIGHTED GRAPH REPRESENTATION OF A POWER GRID

While the standard centrality metrics are based on purely topological information, it is possible to extend the definition of these metrics by including the link weight information and the power flow equations in power grids. Different definitions of extended centrality metrics (extended betweenness [65], modified betweenness and closeness centrality [66], electical degree [53]) exist[2] and are evaluated by simulations via power flow solvers or by calculating power transfer distribution factors (PTDF) in power grids. Such

[2] A broad review of robustness studies in power grids using network science can be found in [50].

simulation-based definitions are generally computationally expensive and formulations with the absence of slack node(s) may not fully explain the analogy between the extended centrality definitions and the weighted graph model for power grids.

Extended metrics were also defined before [36, 67] based on the *voltage - current* relation in electrical circuits. Since the *phase angle - active power* relations in (3.7) and (3.8) in power grids obey the same linear relation as those in electrical circuits (as described before in Table 3.1), these metrics can identify central nodes in power grids.

We take here a graph theoretical approach using the slack-node-independent weighted graph representation for power grids described in the previous section. This weighted graph model facilitates both the analogy between the standard and the extended centrality metrics, and the enhanced linear algebra to formulate the closed-form expression of centrality metrics via graph-related matrices.

WEIGHTED DEGREE CENTRALITY

Similar to the topological definition in (3.9), the weighted degree centrality is related to the number of neighbours of a node. However, rather than only considering the number of neighbours, the weighted degree $\tilde{d}_i$ also includes the information of the admittances $\tilde{a}_{ij}$ of the transmission lines that link the nodes, which leads to the definition:

$$\tilde{d}_i = \sum_{j=1}^{N} \tilde{a}_{ij}. \tag{3.13}$$

A large value of the weighted degree $\tilde{d}_i$ corresponds to larger values of the admittance directly connected to that node, which indicate that node i is well connected to its neighbours.

WEIGHTED EIGENVECTOR CENTRALITY

In analogy with the eigenvector centrality in (3.10), the weighted eigenvector centrality $\tilde{x}_i$ not only captures the total admittance of all lines connected to node i, but is also influenced by the admittance of all lines connected to its neighbours, their neighbours and so on. The weighted eigenvector centralities correspond to the eigenvector of the highest eigenvalue $\tilde{\lambda}_1$ of the weighted adjacency matrix $\tilde{\mathbf{A}}$. Thus, the principal weighted eigenvector centrality $\tilde{x}_i$ is given by the equation:

$$\tilde{x}_i = \frac{1}{\tilde{\lambda}_1} \sum_{j=1}^{N} \tilde{a}_{ij} \tilde{x}_j. \tag{3.14}$$

FLOW BETWEENNESS CENTRALITY

While in the standard definition of the betweenness and the closeness centrality in (3.11) and (3.12), information exchange and other processes are assumed to travel over shortest paths, in the case of the DC power flow equations (or in the equivalent linear systems in Table 3.1), the flow distribution obeys Kirchhoff's and Ohm's laws. Therefore, the standard betweenness and closeness centrality based on shortest paths may not fully capture

the operation of power grids. Instead, the flow betweenness centrality $\tilde{b}_i$ of node i depends on the total flow running through that node, as proposed by [67]:

$$\tilde{b}_i = \sum_{s,t \in \mathcal{N} \setminus \{i\}} \sum_{j \in \mathcal{B}(i)} |f_{s \to t}(i,j)|, \tag{3.15}$$

where $\mathcal{B}(i)$ denotes the direct neighbours of node i, and $|f_{s\to t}(i,j)|$ is the magnitude of the power flow through the link between i and j when a unit active power is injected at node s and extracted from node t. In Appendix 3.6.2, we show how these flows can be calculated from the weighted graph representation of a power grid. Higher values of the flow betweenness centrality $\tilde{b}_i$ indicate the importance of a node with respect to the electrical power transmission in power grids.

ELECTRICAL CLOSENESS CENTRALITY

Similar to the definition of the closeness centrality in (3.12), the electrical closeness centrality of a node is an indicator of the average distance of that node to all other nodes. However, since the flow in a power grids obeys Kirchhoff's laws, the effective resistance [23, 36] is a more appropriate distance metric between nodes than the hopcount of shortest-path. The effective resistance Ω_{ij} between a pair of nodes can be calculated from the pseudo-inverse Laplacian matrix as [35]:

$$\Omega_{ij} = (\mathbf{Q}^\dagger)_{ii} + (\mathbf{Q}^\dagger)_{jj} - 2(\mathbf{Q}^\dagger)_{ij},$$

and captures the effect of the active power transfer p_{ij} and the phase angle difference $\theta_i - \theta_j$ between a pair of nodes, when active power is only injected at and extracted from nodes i or j:

$$\Omega_{ij} = \frac{\theta_i - \theta_j}{p_{ij}}.$$

Since the effective resistance satisfies the properties[3] of a distance function [68] and obeys the flow equations in power grids, it can be used to define a distance-based centrality metric. The electrical closeness centrality $\tilde{c}_i$ of a node equals the reciprocal of the total effective resistance of that node to all other nodes[4]:

$$\tilde{c}_i = \frac{1}{\sum_{j=1}^{N} \Omega_{ij}}. \tag{3.16}$$

Compared to the shortest-path hopcount $H(\mathcal{P}_{i\to j})$, the effective resistance Ω_{ij} does not depend only on the shortest path, but also incorporates the information of all possible paths between node i and j, where the contribution of each possible path follows from the linear flow equations. In the case of the unweighted tree networks, the effective

[3] The properties of a distance function $D(i,j)$ between a pair of nodes i and j are: (a) non-negativity: $D(i,j) \geq 0$, (b) zero distance for identical nodes: $D(i,j) = 0$ if and only if $i = j$, (c) symmetry: $D(i,j) = D(j,i)$ and (d) the triangle inequality: $D(i,k) + D(k,j) \geq D(i,j)$. The effective resistance $D(i,j) = \Omega_{ij}$ satisfies all four properties [68].

[4] [45] defines the node $i^\star$ which is electrically the best spreader to all other nodes based on the flow equations in electrical circuits. This best spreader node $i^\star$ corresponds here to the node with the highest electrical closeness centrality, i.e., $i^\star = \underset{i \in \mathcal{N}}{\text{argmax}}\ \tilde{c}_i$.

resistance Ω_{ij} equals the hopcount $H(\mathscr{P}_{i\to j})$ for all nodes. Thus, for tree-like power grids with equal admittances, the electrical closeness centrality closely resembles the topological closeness centrality, while for power grids with many loops (i.e., non-tree-like) both metrics could differ significantly.

Each of the centrality metrics we present captures a certain aspect of the structural and the operational centrality in the network, such as the strength of a direct connectivity (degree and eigenvector centrality), being a part of many important paths (betweenness centrality) or being close to other nodes (closeness centrality). In recent years, another conceptual definition of centrality has emerged. Based on optimal percolation theory [69], which considers the problem of "finding the smallest set of nodes whose removal fragments the network in small disconnected pieces", a number of new metrics have been proposed (such as the collective influence (CI) [69, 70], belief propagation decimation (BPD) [71] and CoreHD [72]). Such metrics reflect the importance of a node for the global structural coherence as well as their influence in spreading behaviour. However, to the best of our knowledge, *extended* metrics based on percolation theory have not been studied yet. Therefore, in this work, we focus on the generally accepted and adopted centrality metrics to the power grids.

3.4. IDENTIFYING THE EFFECT OF NODE REMOVALS IN POWER GRIDS

In this section, we empirically compare the effects of the targeted node removals based on the centrality metrics presented in the previous section. To evaluate the change in the network functioning, we use two performance metrics that can quantify both the topological and the operational characteristics of the grid after targeted attacks. We consider the networks from 5 real-world power grids of European countries [57] and 5 synthetic power grids from the IEEE test case database [42].

3.4.1. PERFORMANCE METRICS

In an ideal power grid which is robust to targeted attacks, the removal of nodes should not significantly alter the network functioning. In some cases, removing a node from the power grid can partition the network into several components, which are disconnected from each other. This is undesirable as this partition both adversely affects (i) the structure: as the size (i.e., the number of nodes) of the connected component of the network is decreased, and (ii) the operation: since the disconnected structure disrupts the service and the capacity of the network. In this work, we present two performance metrics in our case studies, the size and the capacity of the giant component, to assess the topological and the operational performance aspects in the network.

THE SIZE OF THE GIANT COMPONENT

The giant component [73] is the connected component of a graph that contains the largest fraction of the entire graph's nodes. The change in the size of the giant component reflects the disruptive effect of node removals on the structure of the network.

We assume that the underlying graph of the original network is connected, thus the original size of the giant component is N. Then, we calculate the normalized size σ of

the giant component after each node removal as the ratio between the size of the giant component after node removal and the original network size N, in other words

$$\sigma = \frac{\sum_{i=1}^{N} 1_{\{i \in G'\}}}{N} \tag{3.17}$$

where $1_{\{x\}}$ is the indicator function: $1_{\{x\}} = 1$ if the condition $\{x\}$ is true, else $1_{\{x\}} = 0$, and G' is the giant component after node removals.

THE CAPACITY OF THE GIANT COMPONENT

Each transmission line in a power grid is associated with a maximum flow carrying capability. For the safe operation of a network, the flows through the network links should be below these capability. If the flow limits are exceeded, the situation is detected by protection relays, the circuit breakers are tripped, and the corresponding element is taken out of service. The possibility and the negative impact of cascading failures in power grids increases when the operating point of a power grid is close to the flow carrying capabilities of its links [30, 59]. Consequently, a network with a high flow carrying capability is desired.

We calculate the total capacity of the network as the sum of the maximum flow carrying capabilities of links in the largest connected component of the graph. When multiple lines are connecting the same pair of nodes, we consider an equivalent capacity between those nodes. This equivalent capacity represents the maximum power that can be transferred between these nodes such that the resulting power flow through each single line is at most at its capacity. In Appendix 3.6.3, we describe how this equivalent capacity is calculated.

The capacity of the giant component depends on the number of links in the giant component as well as the flow carrying capability of links, which are closely related to the electrical demands and supplies at the neighbouring nodes[5]. We calculate the normalized capacity of the giant component γ after each node removal as the ratio between the total capacity of the giant component after the removal and the total capacity of the original network, in other words

$$\gamma = \frac{\sum_{l=1}^{L} (\mathscr{C}_l \cdot 1_{\{l \in G'\}})}{\sum_{l=1}^{L} \mathscr{C}_l}. \tag{3.18}$$

3.4.2. PROPERTIES OF THE NETWORKS USED IN SIMULATIONS

We considered the high-voltage transmission networks of 5 European countries in our case studies: Austrian, Belgian, Dutch, French and German power grids. In addition, we included 5 widely used test power grids from IEEE database [42]. In all networks, multiple lines connecting the same pair of nodes are represented as an equivalent single link using the equivalent admittance (3.23) and the equivalent maximum flow carrying

[5] Another metric that can be used to capture the operational performance of power grids is *Yield* which is the ratio of the total demand supplied at the end of an attack with respect to the initial demand of the network [59]. This electrical demand information, which is needed to calculate Yield, is not available in our data sets of real-world power grids.

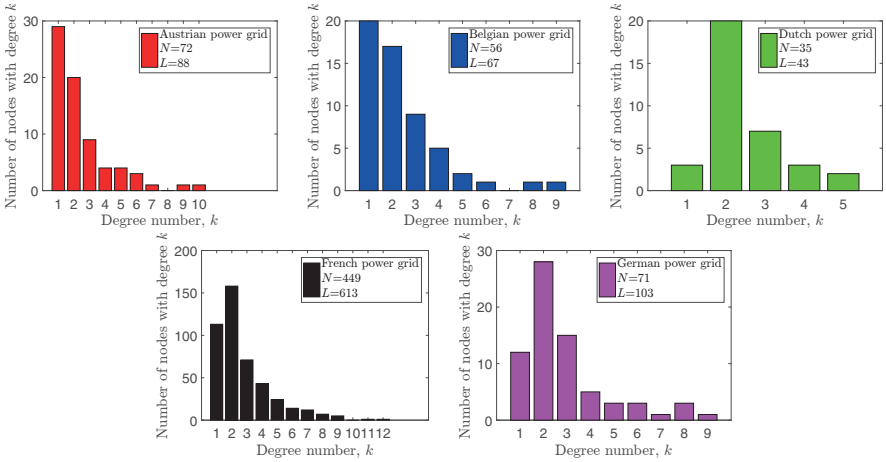


Figure 3.1: The degree distributions of the simple graphs of 5 European power grids.

capability (3.25). The degree distributions of the underlying graphs are shown Figures 3.1 and 3.2. Additionally, more details of the power grids in our case study are available on our GitHub page [57].

3.4.3. THE EFFECTS OF TARGETED NODE REMOVALS IN POWER GRIDS

We apply both the standard and the extended centrality metrics as node-attack strategies in power grids. For each centrality metric, we start the attacks by removing the node (and all its links) with the highest value of the chosen centrality metric. After each node removal, we recalculate the values of the centrality metric, and continue by removing the node with the highest value of the centrality metric in the current giant component of the graph. Note that, during the successive node removals, we do not take the cascading dynamics (such as overloading of links or demand/supply redistribution due to cascading failures [59]) into account. In other words, we focus on the instant just after the removal of nodes to identify the effects on the structure and the operational performance indicators of the power grid.

Figures 3.3 and 3.4 show the changes in the normalized size and the capacity of the giant component when we sequentially remove the nodes according to 8 different centrality metrics. We observe that the betweenness and the flow betweenness centrality are the best attack strategies as they can maximally disrupt the network functioning with fewer attacked nodes. On the other hand, the degree centrality may not always successfully assign an important node. Compared to the degree centrality, the betweenness and the flow betweenness centrality give more fine-grained centrality values for each node, whereas, multiple nodes with the same degree exist, as illustrated in Figures 3.1 and 3.2, making them indistinguishable. In addition, we observe that the eigenvector and the weighted eigenvector centrality are the least effective attack strategies: Targeted attacks according to these centrality metrics destroy the network slower than other attack strategies.

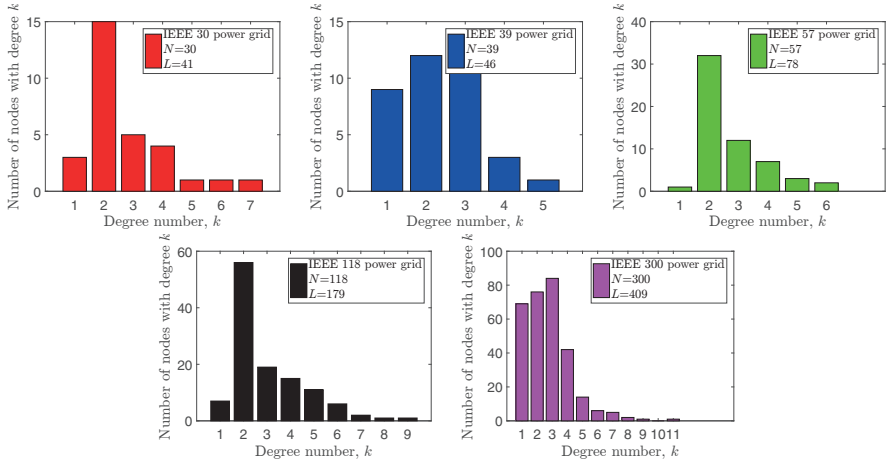


Figure 3.2: The degree distributions of the simple graphs of 5 IEEE power grids.

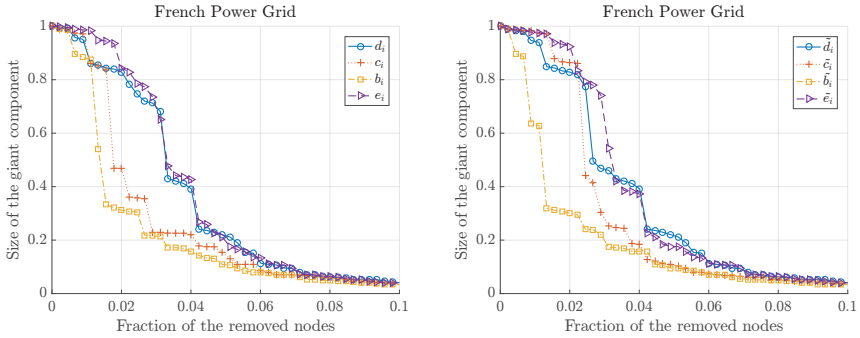


Figure 3.3: The normalized size of the giant component in the French power grid versus the removal of nodes according to the standard centrality metrics (left) and the extended centrality metrics (right). The node with optimal graph metric, computed in the resulting giant component, is removed.

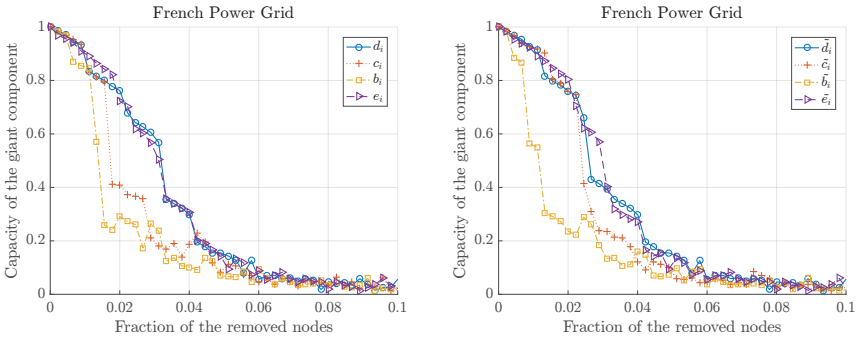


Figure 3.4: The normalized capacity of the giant components in the French power grid versus the removal of nodes according to the standard centrality metrics (left) and the extended centrality metrics (right). The node with optimal graph metric, computed in the resulting giant component, is removed.

Next, in order to compare the attack strategies and to further quantify the topological and operational changes in power grids, we calculate the average of the structural and the operational performance indicators (or the energy ε values of a graph [46]) that are the normalized sums of the size and the capacity of the giant component over successive targeted attacks, respectively. Thus, the average value $\bar{\sigma}$ of the structural performance indicator of the power grid over K successive node-attacks can be calculated as

$$\bar{\sigma} = \frac{\sum_{k=1}^{K} \sigma(k)}{K} \tag{3.19}$$

where $\sigma(k)$ is the normalized size of the giant component after k successive attacks. Similarly, the average value $\bar{\gamma}$ of the structural performance indicator of the power grid over K successive node-attacks is

$$\bar{\gamma} = \frac{\sum_{k=1}^{K} \gamma(k)}{K} \tag{3.20}$$

where $\gamma(k)$ is the normalized capacity of the giant component after k successive attacks. The structural $\bar{\sigma}$ and the operational $\bar{\gamma}$ performance indicators in (3.19) and (3.20) are evaluated on a score between 0 and 1: In an ideal power grid which is robust to targeted attacks, the node removals should have slight effects on the network performance. Thus, a performance indicator close to 1 is desirable by the network operators. On the other hand, a lower performance indicator over successive node-attacks indicates a powerful (destructive) attack-strategy in which few important nodes of the network are identified and removed, with negative operational and structural consequences.

In Figures 3.5 and 3.6, we present the average values of the performance indicators in European and IEEE test power grids (PG) after different attack strategies that remove 10% of the initial network nodes, respectively. Higher values in Figure 3.5 and 3.6 represent higher robustness to the targeted attacks, whereas lower values indicate vulnerability or a severe disrupt in network functioning. We observe that targeted attacks based on the flow betweenness and the betweenness centrality followed by the closeness and

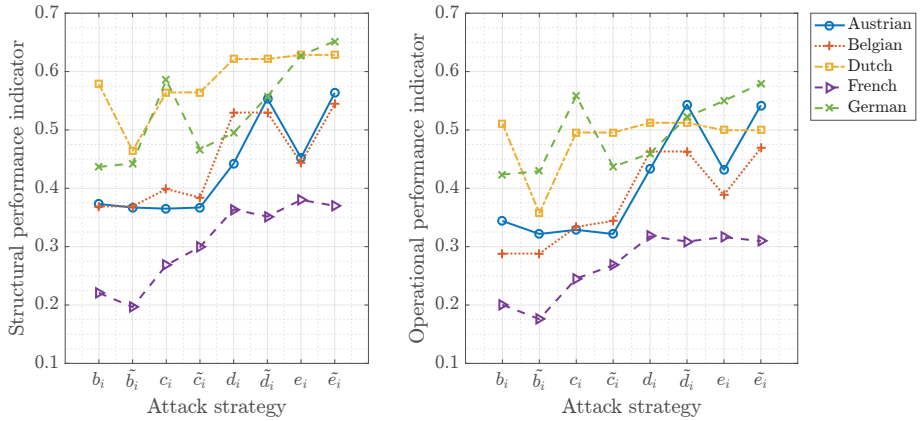


Figure 3.5: The structural performance indicator $\bar{\sigma}$ (left) and the operational performance indicator $\bar{\gamma}$ (right) in European power grids after the targeted attacks.

the electrical closeness centrality are the best attack strategies to decrease the structural and the operational performances of the power grids. As an example, the targeted node attacks according to the flow betweenness centrality of nodes destroy the Dutch power grids faster than any other attack strategy.

3.4.4. MAIN LESSONS LEARNED FROM THE ANALYSES

In this section, we summarize the insights obtained in the previous sections. The main lessons learned from the analyses of the targeted node-attacks based on the different centrality metrics from the simple and the weighted graph representations of power grids in the tested networks are as follows:

- The degree centrality (3.9) only provides information on the local structure around a node. Similarly, the weighted degree centrality (3.13) reflects the local connectivity information. Thus, a node that is connected to many other nodes (with high admittance) is not necessarily a central node for the whole network. Therefore, as illustrated by the targeted attack simulations, the degree and the weighted degree centralities cannot always indicate the important nodes.

- The betweenness centrality (3.11) incorporates information about the global network structure, and in the analyses of the test networks, high betweenness centrality values were found to efficiently indicate the nodes whose removal would significantly disrupt the network performance. While successfully indicating vulnerable nodes, the betweenness centrality (3.11) is based on the shortest paths only. This means that the betweenness centrality does not discriminate nodes that are positioned “close” to many shortest paths (and would be considered central), and peripheral nodes. This limitation is partly addressed by the flow betweenness centrality (3.15), in which the flows through the network links are distributed

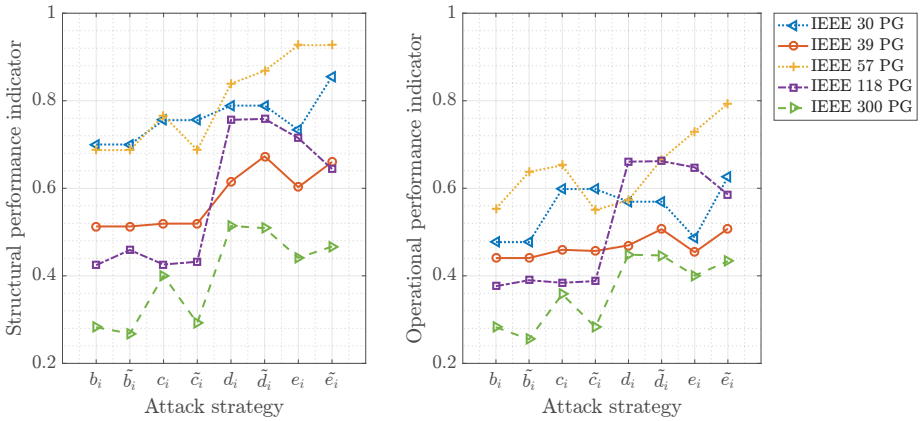


Figure 3.6: The structural performance indicator $\bar{\sigma}$ (left) and the operational performance indicator $\bar{\gamma}$ (right) in IEEE power grids after the targeted attacks.

throughout the network according to the Kirchhoff's laws. In the analyses of the test networks, removing nodes with a high flow betweenness usually resulted in the most destructive effects on the network.

- The closeness centrality (3.12) reflects the average shortest path distance from a node to all other nodes in the network. Higher closeness centrality values thus indicate nodes which can easily reach the other nodes in the network. Similarly, higher values of electrical closeness centrality (3.16) show a node that is on average close to the other nodes in the network, based on the operationally inspired effective resistance distance instead of the shortest-path distance. In the analyses of targeted attacks, the performance of the closeness and the electrical closeness centrality in identifying the important nodes in the tested power grids are found to be similar.

- The eigenvector centrality (3.10) can rarely identify the critical nodes, and thus, the targeted attacks based on the eigenvector centrality are generally the slowest destructive strategy among the traditional centrality metrics in the tested networks. Similarly, the weighted eigenvector centrality (3.14) seems not to successfully indicate important nodes.

The analyses of the targeted node-attacks show that centrality metrics, in particular the (flow) betweenness and (electrical) closeness, are very successful in indicating the critical nodes whose removals sharply decrease the selected performance indicators (the size and the capacity of the giant component) of power grids. Identifying these critical components in advance can enable power grid operators to improve system robustness by monitoring and protecting these components continuously. Additionally, although the effect of targeted attacks are more significant when the centrality information is updated after each node removal, the information based on the initial calculation of the

centrality metrics is also fairly successful in finding the important nodes. In that case, the degree centrality is a good indicator to fragment the network to decrease the structural and operational performance indicators of power grids (See Appendix 3.6.1).

3.5. CONCLUSION

In this chapter, we took a network science approach to investigate the vulnerability of power grids to malicious targeted attacks. First, we presented two different graph models for power grids: simple and weighted graphs. Subsequently, using these graph models, we ranked the importance of each node according to the standard and the extended centrality metrics that take into account the electrical properties of the grids such as the admittance of the transmission lines and the flow allocation according to the DC power flow equations. Via case studies in both real-world and test power grids, we show that the power grids are highly vulnerable to targeted attacks: sequentially removing the nodes with the highest centrality is a good strategy to fragment the power grids, and to maximally decrease its operational performance. In almost all power grids in our case study, removing approximately 15% of the nodes according to the flow betweenness centrality destroys the network almost completely. Grid operators can use the proposed methodology to analyse the current vulnerability of their network to targeted attacks and to take necessary measures by protecting the important nodes in their networks.

3.6. APPENDIX

3.6.1. TARGETED ATTACKS BASED ON INITIAL CENTRALITY METRICS

Instead of recalculating the centrality metrics after each node removal, we consider here a more simplified attack strategy based on calculating the centrality metrics only once, at the beginning of the attacks. The targeted attacks are then performed sequentially according to these initial values.

Figures 3.7 and 3.8 show the changes in the normalized size and the capacity of the giant component in French power grids after the targeted attacks, respectively. Figures 3.7 and 3.8 illustrates that even these simplified attack strategies could inflict a significant damage on the network functioning: For instance, removal of 15% of the nodes according to the initial rankings nearly destroys French power grids. Compared to Figures 3.3 and 3.4, in Figures 3.7 and 3.8, we observe that the degree centrality is the most destructive attack strategy when the centrality metrics are based on only the initial calculation of the centrality metrics, i.e., when the node-rankings are not updated after the targeted attacks.

3.6.2. CALCULATION OF FLOW BETWEENNESS CENTRALITY IN POWER GRIDS

Following the linearised DC power flow equations in power grids, the active power $\mathbf{P}$ and phase angle Θ in all nodes are related by equations (3.7) and inversely by (3.8). When a unit active power is injected at node s and extracted at node t, this corresponds to the active power input to the grid:

$$\mathbf{P}_{s\to t} = \mathbf{e}_s - \mathbf{e}_t$$

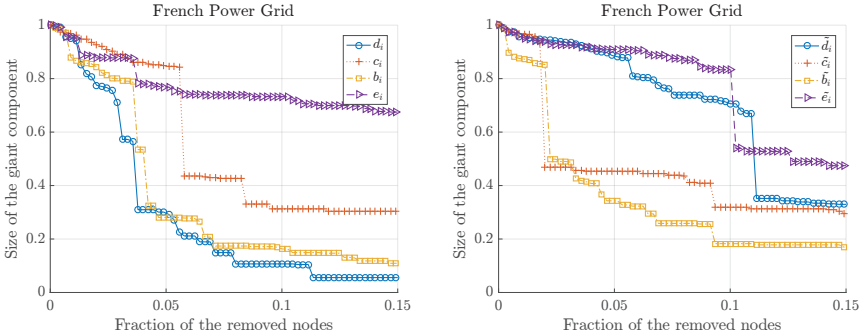


Figure 3.7: The normalized size of the giant component in the French power grid versus the removal of nodes according to the standard centrality metrics (left) and the extended centrality metrics (right). Nodes are removed sequentially according to the initial values of the centrality metrics.

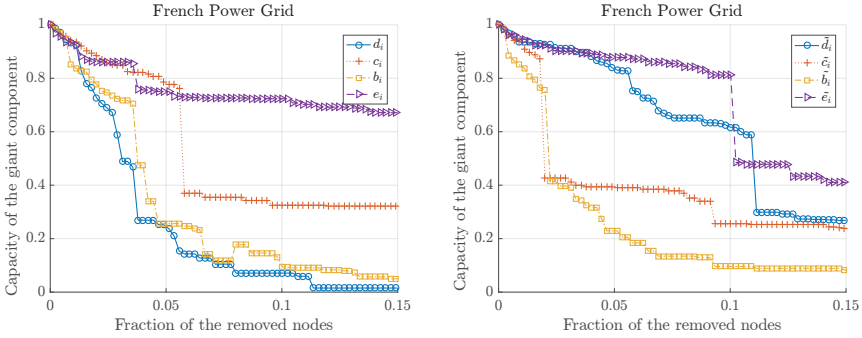


Figure 3.8: The normalized capacity of the giant component in the French power grids versus the removal of nodes according to the standard centrality metrics (left) and the extended centrality metrics (right). Nodes are removed sequentially according to the initial values of the centrality metrics.

where $\mathbf{e}_k$ is the basis vector with the k^{th} component equal to one and all other components zero. Based on equation (3.8), the resulting phase angle vector for this active power input can be calculated as

$$\Theta_{s\to t} = \mathbf{Q}^{\dagger}(\mathbf{e}_s - \mathbf{e}_t).$$

Knowing the phase angle at each node, it is then possible to calculate the flow $f_{s\to t}(i,j)$ through the link between nodes i and j as

$$f_{s\to t}(i,j) = \tilde{a}_{ij}(\mathbf{e}_i - \mathbf{e}_j)^{\mathrm{T}}\Theta_{s\to t}. \tag{3.21}$$

The flow betweenness centrality $\tilde{b}_i$ of a node i is the sum of the absolute flows that pass through that node i, over all possible pairs of source and target nodes[6]:

$$\tilde{b}_i = \sum_{s,t\in\mathcal{N}\setminus\{i\}} \sum_{j\in\mathcal{B}(i)} \left|\tilde{a}_{ij}(\mathbf{e}_i - \mathbf{e}_j)^{\mathrm{T}}\mathbf{Q}^{\dagger}(\mathbf{e}_s - \mathbf{e}_t)\right|. \tag{3.22}$$

3.6.3. MULTIPLE LINES CONNECTING THE SAME PAIR OF NODES

We consider multiple lines $\mathscr{L}'$ connecting the same pair of nodes i and j: each line l has admittance y_l and flow capacity $\mathscr{C}_l$. In the weighted graph model for power grids, those multiple lines $\mathscr{L}'$ are represented as a single equivalent link between node i and j, with admittance

$$y_{ij}^{(\mathrm{I})} = \sum_{l\in\mathscr{L}'} y_l^{(\mathrm{I})}. \tag{3.23}$$

The maximum possible flow between those nodes i and j is constrained by the capacity of each single line connecting them. If power f_{ij} flows from node i to node j, then according to the DC power flow equations in (3.3), this results in the phase angle difference

$$(\theta_i - \theta_j) = \frac{f_{ij}}{y_{ij}^{(\mathrm{I})}}, \tag{3.24}$$

where $y_{ij}^{(\mathrm{I})} = \sum_{l\in\mathscr{L}'} y_l^{(\mathrm{I})}$ is the equivalent admittance between node i and j of $\mathscr{L}'$ lines in parallel. For each single line, Ohm's law states that the flow f_l through that line is related to the phase angle difference by

$$f_l = y_l^{(\mathrm{I})}(\theta_i - \theta_j).$$

Introducing the phase angle difference from equation (3.24) then leads to

$$f_l = f_{ij}\frac{y_l^{(\mathrm{I})}}{y_{ij}^{(\mathrm{I})}}$$

[6] In power grids, the flow through network links are directed. Different from the load transmissions in other types of networks, opposite directed flows through a link can cancel out the total flow through that link in power grids. Therefore, to calculate the maximum possible flow through a link, the absolute sum is necessary. This definition under the maximum possible loading condition is also used in the definition of *full betweenness centrality* by [74].

for the flow f_l through line l. Since the maximum flow through each line is constrained by its flow capacity: $f_l \leq \mathscr{C}_l$, we find that the total flow f_{ij} between node i and j is constrained by an equivalent capacity $\mathscr{C}_{ij}$ equal to:

$$f_{ij} \leq \mathscr{C}_{ij}$$

where,

$$\mathscr{C}_{ij} = \min_{l \in \mathscr{L}'} \left(\frac{\mathscr{C}_l}{y_l^{(\mathrm{I})}} \right) y_{ij}^{(\mathrm{I})}. \tag{3.25}$$

4

Topology-Driven Performance Analysis of Power Grids

The previous chapter presents a weighted graph model for power grids to analyse the power grids from a graph theoretical point of view. By utilizing this weighted graph model, in this chapter, we investigate the relation between the topology and the electric power transmission in power grids. Initially, we focus on synthetic power grids whose underlying topology can be structured as either a path graph or a complete graph, and we analytically compute the impact of electric power transmission on link flows under the normal operation and under a link failure contingency using the linearised DC power flow equations. Subsequently, in various other graph types, we provide empirical results on the effects of electric power transmission on the link flow, the voltage magnitude and the total active power loss in power grids using the nonlinear AC power flow equations.

This chapter is based on a published work [75].

4.1. INTRODUCTION

POWER grids have been analysed from a graph topological point of view [76, 77]. Various topology metrics (such as nodal degree, clustering coefficient [28]) have been proposed to assess the vulnerability and/or to locate the critical components of power grids [29, 46, 78]. However, electric power transmission is governed by physical laws, and an assessment based on direct connections between nodes and shortest paths may not hold in power grids.

In this chapter, we take an extended graph theoretical approach [23, 32, 45, 79] by modelling the electrical properties, such as the flow allocation according to Kirchhoff's laws and the impedances of transmission lines, and investigate the impact of electric power transmission on the key performance indicators of power grids, which we take as the node voltage, the link flow, the total power loss and the served power demand.

Initially, we focus on the operation by considering two extreme graphs. In synthetic power grids whose underlying topology is either a path or a complete (full-mesh) graph, we analytically derive the steady-state operating conditions under normal operation and under a random link failure (removal) contingency using the linearised DC power flow equations. Subsequently, in various other graphs, we empirically investigate the relation between the topology and the key performance indicators using the nonlinear AC power flow equations [80].

The remainder of this chapter is organized as follows. Section 4.2 investigates the electric power transmission in path and complete graphs under normal operation. In Section 4.3, we focus on single link failure contingencies in those graphs, and derive the impact of a random link failure on the steady-state link flows and served power demand. Section 4.4 presents our empirical results on the key performance indicators in various graphs both under normal operation and single link failure contingencies. Section 4.5 concludes the chapter.

4.2. DC POWER FLOW ANALYSIS IN PATH AND COMPLETE GRAPHS

A connected simple graph (i.e., a graph with no parallel duplicate links or self-loops) lies between a tree graph and a complete graph. In a complete graph, every pair of distinct nodes is connected by a link. On the other hand, a tree has no cycles; consequently, any two nodes are connected by exactly one path (See Figure 4.1).

The direct connections between nodes usually result in an efficient transmission in a network. The distances between the nodes in a complete graph are shortest compared to the other graphs, in which multiple hops are needed to reach the destination. In power grids, different than the typical transmissions based on the shortest paths, the electric power transmission is governed by physical laws. Therefore, an assessment based on purely the direct connections between nodes may not be enough to draw conclusions. In this section, we investigate the electric power transmission in those extreme graph types.

We model a power grid with N buses (nodes), and L lines (links) by a weighted graph $G(N, L)$. We use $\mathcal{N}$ to denote the set of N nodes and $\mathcal{L}$ to denote the set of L links with equal weights, b. Every link $l_{ik} \in \mathcal{L}$ between the nodes $i, k \in \mathcal{N}$ is associated with a max-

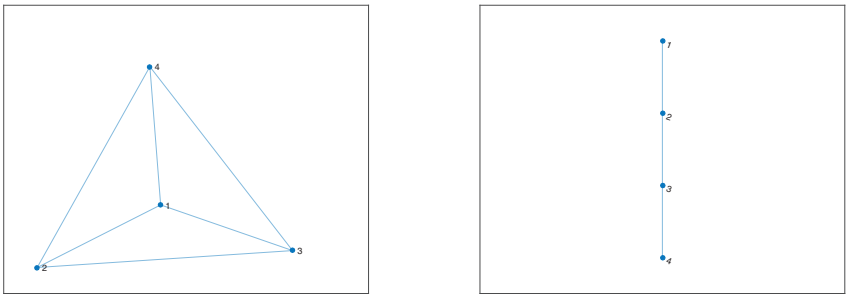


Figure 4.1: An example of a complete graph with 4 nodes (left) and a path graph (which is a type of a tree graph) with 4 nodes (right).

imum flow capacity C_{ik} that represents the maximum power flow that can be afforded by the corresponding line, and a rest flow capacity $\alpha_{ik} = C_{ik} - |f_{ik}|$ where $|f_{ik}|$ is the flow through the link l_{ik} under the normal operation. We assume a single upstream *supply* node, and treat the remaining $N-1$ downstream nodes as *demand* nodes. Without loss of generality, we label the supply node as node 1, and take the total electric power demand of the network as $(N-1)p$ where $p \geq 0$ is a constant. Throughout Sections 4.2 and 4.3, we adopt the slack-bus independent solution to the DC power flow equations [23], which could approximate the steady-state operation under the DC power flow assumptions [34].

4.2.1. ELECTRIC POWER TRANSMISSION IN A PATH GRAPH

We investigate the electric power transmission from the supply node 1 to the single demand node N in a path graph (whose nodes are labeled consecutively). The magnitude $|f_{ik}^{1\to N}|$ of the flow through a link l_{ik} between node i and node $k = i+1$ is found as (See 4.6.1)

$$\begin{aligned} |f_{ik}^{1\to N}| &= b|(\theta_i - \theta_k)| \\ &= p(N-1) \quad \forall l_{ik} \in \mathscr{L}, \end{aligned} \tag{4.1}$$

where b is the reciprocal of the line reactance and θ_i is the voltage phase angle at node i.

Equation (4.1) shows that the resulting link flows due to the electric power transmission are all the same, and their values increase with the increasing graph size N and unit power demand p. This linear correlation between the size and the magnitudes of link flows could lead to substantial flows and result in congestion problems, especially in large graphs.

On the other hand, in a path graph, the electric power is transferred through a single path between the supply and the demand node. Consequently, an assessment based on shortest paths holds, which can ease the supervision of the network operator.

4.2.2. ELECTRIC POWER TRANSMISSION IN A COMPLETE GRAPH

Similar to Section 4.2.1, we investigate the electric power transmission from the supply node 1 to a single (randomly chosen) demand node m in a complete graph. The magnitude $|f_{ik}^{1\to m}|$ of the resulting flow through a link l_{ik} is found as (See 4.6.2)

$$|f_{ik}^{1\to m}| = \begin{cases} \frac{2(N-1)p}{N} & \text{if } l_{ik} = l_{1m}, \\ \frac{(N-1)p}{N} & \text{if } l_{ik} \in \{(\mathscr{B}(1) \cup \mathscr{B}(m))\} \setminus l_{1m}, \\ 0 & \text{otherwise,} \end{cases} \tag{4.2}$$

where $\mathscr{B}(i)$ denotes the set of links that are direct neighbors of node i.

Equation (4.2) indicates that three different magnitudes of link flow exist during the electric power transmission: (a) The flow through the link between the supply and the demand node is maximum, whereas (b) the flows through the links to the other neighbors of those nodes are half of that maximum flow, and (c) the remaining links that are not direct neighbors of either the supply or the demand node have zero flows.

Comparing the magnitudes of link flows in a path graph in Equation (4.1) and in a complete graph in Equation (4.2) shows that the maximum link flow due to the electric power transmission from the supply node to a demand node is dramatically lower in a complete graph. However, the distribution of the flows through links in a complete graph is not homogeneous, thus the relation between the total decrease in the magnitudes of link flow and the total number of links added to a path graph is not linear.

4.3. DC POWER FLOW ANALYSIS IN PATH AND COMPLETE GRAPHS AFTER A RANDOM LINK FAILURE

Single line failures are common in power grids. Therefore, as well as under the normal operation, the operation after a link failure (removal) is important to assess the reliability of power grids [81]. In this section, we theoretically investigate the effects of a random link failure on the link flows and the served demand in path and complete graphs using the linearised DC power flow equations. In this context, for random link removals, we assume an equal likelihood $\frac{1}{L}$ of each link $l_{ik} \in \mathscr{L}$ to be removed from the graph $G(N, L)$. In addition, to quantify the effect of link failure contingencies in a graph, we calculate the *theoretical robustness function* of those graphs, which we define as the expected fraction of served demands after a random link failure.

4.3.1. RANDOM LINK FAILURE IN A PATH GRAPH

First, we focus on the effect of a single link failure in path graphs. Just before the link failure takes place, we assume that all demand nodes have a unit electric power demand of p, which we refer to as the *symmetrical distribution* of demands. In other words, the supply node transfers a unit electric power of p to every other demand node, resulting in the magnitude $|f_{ik}|$ of the flow through link l_{ik} between node i and node $k = i+1$ before the failure

$$|f_{ik}| = p(N-i). \tag{4.3}$$

As the graph under investigation is a path graph with no cycles, the removal of any link l_{ik} between node i and node $k = i+1$ partitions the graph (See the derivation in Section 4.6.3), and this partition removes in total $p(N-i)$ demand from the graph according to Equation (4.3). Therefore, the closer the link failure is to the supply node 1, the worse is the effect on the served demands.

The continuity of the operation of the network depends on the location of the failed link. If the failed link is adjacent to the supply node, then the supply node is isolated from the demand nodes and the network faces a complete blackout. If the failure probabilities of the links are the same in a path graph, the probability p_b that a random link failure leads to a complete blackout is $p_b = \frac{1}{L} = \frac{1}{N-1}$.

The failure and removal of any other link partitions the network and the remaining network can continue functioning, though with decreased demands. As the total demand of the network decreases after the link removal, the flows through the remaining links decrease, and thus, there is no possibility for further cascading failures [59] due to the insufficient rest flow capacity of the remaining links. Consequently, the expected fraction $\mathrm{E}[F_s]$ of served demands after a random link failure is

$$\mathrm{E}[F_s] = \frac{1}{N-1}\frac{(0+1+\cdots+N-2)}{N-1} = \frac{N-2}{2(N-1)} = \frac{1}{2} - \frac{1}{2(N-1)}.$$

4.3.2. RANDOM LINK FAILURE IN A COMPLETE GRAPH

Next, we investigate a random link failure in a complete graph. After the removal of a link, the flows are redistributed following Kirchhoff's laws and the flows through the remaining links may change. Due to the meshed topology of the graph, this redistribution can lead to an increase or a decrease in flow through a particular link [23].

Before a link failure happens, under the symmetrical distribution of demands, i.e., when the supply node transfers a unit demand of p to every other demand node, the magnitude $|f_{ik}|$ of the flow through a link l_{ik} in a complete graph is

$$|f_{ik}| = \begin{cases} p & \text{if } l_{ik} \in \mathscr{B}(1), \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

Following Equation (4.4), two different magnitudes of link flows exist in a complete graph under the symmetrical distribution of demands. As a result, a single link failure and removal can result in two cases:

Failure of a link with zero flow

When a link l_{ik} is removed from the graph, the flow $|f_{ik}|$ through the link before failure needs to be redistributed over the alternative paths between nodes i and k. Since a *redundant* link l_{ik} does not transport any flow, its removal does not cause a power redistribution.

Failure of a link with maximum flow

When a *used* link l_{ik} is removed from the graph, the flow $|f_{ik}| = p$ through the link before failure is redistributed over alternative paths between nodes i and k. As a result, the initial link l_{ik} failure may trigger further failures in the network if the increase $|\Delta f_{ab}| = \frac{p}{N-2}$ in the flow through a remaining link l_{ab} exceeds its rest flow capacity α_{ab} (See the derivation in Section 4.6.4),

$$|\Delta f_{ab}| = \frac{p}{N-2} \geq \alpha_{ab}. \tag{4.5}$$

When the rest flow capacity α_{ab} is smaller than the required value in Equation (4.5), consecutive failures occur. After the initial failure of the used link l_{ik}, the flows through all remaining used links exceed their maximum flow capacity, and fail in the next stage of the failure. This isolates the supply node. Consequently, the remaining network cannot match any demands, and it faces a complete blackout.

When the size N of the graph is 2, i.e., when there is only one link, the failure of that link destroys the graph by separating the supply node from the demand node regardless of the rest flow capacity α_{ab} of the link. For larger graphs, Equation (4.5) shows the required rest flow capacity α_{ab} of links is maximum when the size of the graph is $N = 3$, whereas it decreases as N increases. This means the effect of a link removal on the flows through the remaining links reduces with the size N of the graph.

Finally, we calculate the *theoretical robustness function* of a complete graph, which is the expected fraction $\mathrm{E}[F_s]$ of served demands after a random link removal from the underlying graph. Figure 4.2 presents the theoretical robustness function of a complete graph under the symmetrical distribution of demands. If the rest flow capacity of the links is larger than the required value in Equation (4.5), the remaining links can tolerate the redistributed flows after a random link removal. The network can continue to serve the same amount of total demand after any single link failure. Therefore, in region II in Figure 4.2, the fraction of served demands stays the same. On the other hand, when the rest flow capacity of the links is smaller than the required value in Equation (4.5) in region I, the network continues its operation only if the failed link is a *redundant* link with zero flow. Otherwise, when a *used* link with flow p fails, the network faces complete blackout and cannot serve any demand. Therefore, the expected $\mathrm{E}[F_s]$ fraction of served demands after a random link failure in region I is calculated as

$$\mathrm{E}[F_s] = 1 \cdot p_r + 0 \cdot p_u = \frac{N-2}{N} = 1 - \frac{2}{N},$$

where $p_r = \frac{N-2}{N}$ represents the probability that the failed link is a redundant link with zero flow and $p_u = \frac{2}{N}$ represents the probability that the failed link is a used link with flow p (See Equations (4.16) and (4.17)).

Comparing the effect of single link failures in a path and a complete graph shows that a path (or a tree) cannot provide a back-up path after a link removal, and the total served demand in the network definitely decreases. In addition, as the demands in the network decrease, there is no possibility of cascading failures in a tree graph with no loops. On the other hand, a meshed topology can provide back-up paths during random link removals. Yet, the correct setting of design parameters, i.e., the rest flow capacities of the links, is extremely important. If the rest flow capacity α_{ab} of the links is smaller than in Equation (4.5) and a used link fails, then the remaining links in a complete graph cannot tolerate the redistributed flows. Consequently, a random link removal in a complete graph could lead to more link and demand losses than a link failure in a simple path graph.

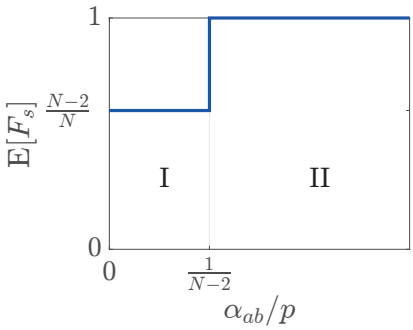


Figure 4.2: The expected $\mathrm{E}[F_s]$ fraction of served demands versus the rest flow capacity of the links after a random link failure in a complete graph under symmetrical distribution of demands. The figure is computed for $N = 5$.

4.4. THE IMPACT OF TOPOLOGY ON THE KEY PERFORMANCE INDICATORS OF POWER GRIDS

In Sections 4.2 and 4.3, we focus on the theoretical analyses of electric power transmission in path and complete graphs. In this section, we focus on many other graphs and empirically investigate the key performance indicators of power grids under the normal operation as well as under the single link failure contingencies. In our analyses, we use the AC power flow solver in Matlab [80] to calculate the steady-state operating conditions under the symmetrical distribution of demands. In addition to the line reactance x and the active power p values, we take into account the line resistance r and the reactive power q values for a more practical model of power grids.

4.4.1. KEY PERFORMANCE INDICATORS UNDER NORMAL OPERATION

For the safe and efficient operation of power grids, lower magnitudes of link flow and total power loss, and higher values of node voltage (close to 1 per unit) are desired[1]. In a power grid whose topology is modelled by the specific graph G, we define the satisfaction degree of performance indicators of the link flow $\zeta(G)$, the node voltage $\nu(G)$ and the

[1] As we focus on the impact of electric power transmission from a supply node to demand nodes, only voltage drops in the network are considered.

power loss $\eta(G)$ as

$$\zeta(G) = \begin{cases} 1 & \text{if } \max\limits_{l_{ik}\in\mathscr{L}(G)} \left(|f_{ik}|\right) < p, \\ \dfrac{\tau_f - \max\limits_{l_{ik}\in\mathscr{L}(G)} \left(|f_{ik}|\right)}{\tau_f - p} & \text{if } p \le \max\limits_{l_{ik}\in\mathscr{L}(G)} \left(|f_{ik}|\right) < \tau_f, \\ 0 & \text{if } \max\limits_{l_{ik}\in\mathscr{L}(G)} \left(|f_{ik}|\right) \ge \tau_f, \end{cases} \tag{4.6}$$

$$\nu(G) = \begin{cases} 0 & \text{if } \min\limits_{i\in\mathscr{N}(G)} \left(v_i\right) < \tau_v, \\ \dfrac{\min\limits_{i\in\mathscr{N}(G)} \left(v_i\right) - \tau_v}{1-\tau_v} & \text{if } \tau_v \le \min\limits_{i\in\mathscr{N}(G)} \left(v_i\right) < 1, \\ 1 & \text{if } \min\limits_{i\in\mathscr{N}(G)} \left(v_i\right) \ge 1, \end{cases} \tag{4.7}$$

$$\eta(G) = \begin{cases} 1 & \text{if } \tau_\sigma \le 0, \\ \dfrac{\tau_\sigma - \sigma(G)}{\tau_\sigma} & \text{if } 0 < \sigma(G) < \tau_\sigma, \\ 0 & \text{if } \sigma(G) \ge \tau_\sigma, \end{cases} \tag{4.8}$$

where $|f_{ik}|$ is the magnitude of the flow through link l_{ik}, v_i is the magnitude of voltage at node i, $\sigma(G)$ is the total active power loss, and $\mathscr{L}(G)$ and $\mathscr{N}(G)$ denote the set of links and nodes of graph G, respectively. The performances in Equations (4.6), (4.7) and (4.8) are evaluated on a scale from 0 to 1 (See Figure 4.3): The highest performance of 1 corresponds to an *ideal power grid* in which the maximum link flow is equal to the unit power demand p, the minimum voltage is equal to 1 per unit, i.e., no voltage drop, and the total power loss is 0, i.e., a lossless power grid. Conversely, the lowest performance of 0 corresponds to the maximum link flow τ_f, the minimum node voltage τ_v and the total power loss τ_σ. The requirements of τ_f, τ_v and τ_σ are usually determined by the specific grid codes of the operators.

Figure 4.4 shows the variations of the key performance indicators under the symmetrical demand distribution throughout the topological transformation of the path graph with 5 nodes and 4 links[2]. Figure 4.4 depicts that the performance indicator of the link flow is lowest in the path, and highest in the complete graph. We observe that the cycle graph *dramatically increases the performance indicator of link flow* by decreasing the maximum link flow compared to the path graph. Similar to the link flows, the minimum voltage is lowest and the total power loss is highest in the path graph. The complete graph, on the other hand, represents the operation at the minimum voltage drop and total power losses, thus corresponding to the highest values of key performance indicators.

Figure 4.4 indicates that a meshed topology can improve the key performance indicators compared to the path graph. We showed in Equation (4.2) that the flow distribution in a complete graph due to an electric power transmission is not homogeneous,

[2] We compute all possible ways to evaluate the transformation from the path to the complete graph, which is only possible for small graphs.

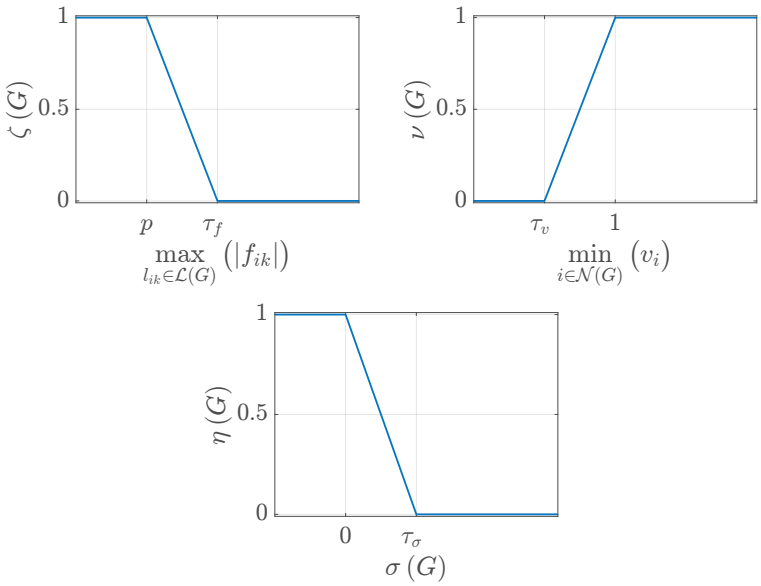


Figure 4.3: The functions of the performance indicators of the link flow $\zeta(G)$, the node voltage $\nu(G)$ and the power loss $\eta(G)$.

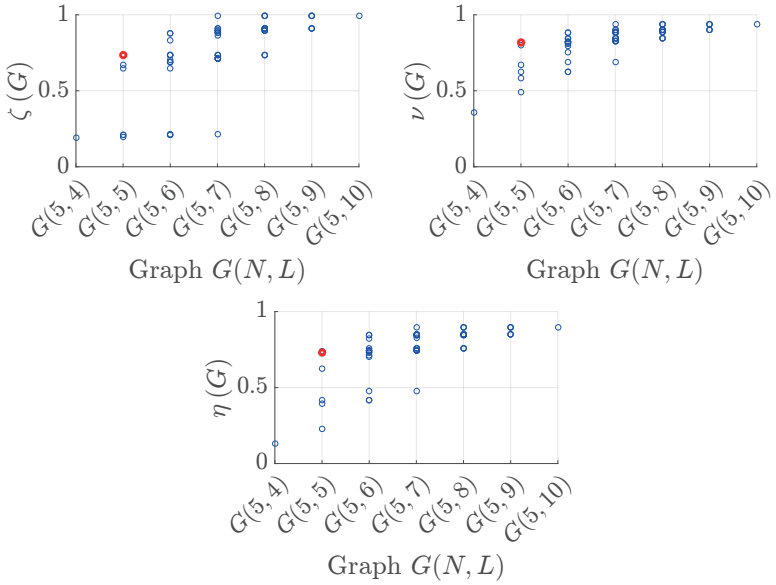


Figure 4.4: The variations of the key performance indicators throughout the topological transformation of the path graph $G(5,4)$ with 5 nodes and 4 links. The transformation towards the complete graph $G(5,10)$ requires the addition of $\frac{N(N-1)}{2}-(N-1)=6$ links. The bold red data point corresponds to the cycle graph. The performance indicators are evaluated for $r=0.1$, $x=0.1$, $p=0.05$, $q=0.01$, $\tau_f=5p$, $\tau_v=0.9$ and $\tau_\sigma=0.2p$.

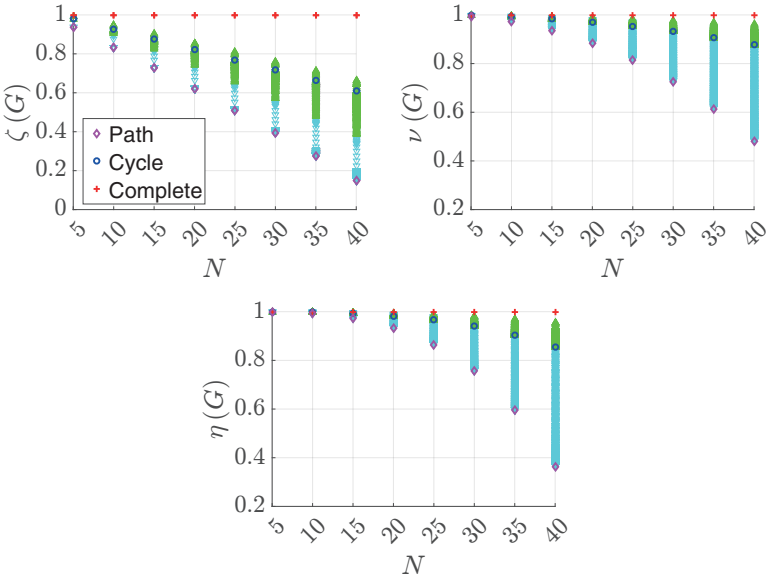


Figure 4.5: The variations of the key performance indicators in graphs with different sizes N. The blue lower triangle ($\triangledown$), and green upper triangle ($\triangle$) data points correspond to the graphs that are constructed by adding one link to the path graph, and adding one link to the cycle graph, respectively. The performance indicators are evaluated for $r = 0.02$, $x = 0.02$, $p = 0.005$, $q = 0.001$, $\tau_f = 50p$, $\tau_v = 0.8$ and $\tau_\sigma = 4p$.

which could explain the nonlinear relation between the total number of added links to the path graph and the total increase in the key performance indicators in Figure 4.4. In particular, the cycle graph and the *augmented cycles*, i.e., the graphs constructed from the cycle graph by adding links, are observed to affect the key performance indicators dramatically.

In Figure 4.5, we present the variations of key performance indicators in the graphs with different sizes N. In the complete graphs, the maximum link flows are nearly the same for all sizes N, which is in agreement with the theoretical calculations in Equation (4.4). In the other graphs, the maximum link flow increases with increasing size N, decreasing the performance indicator of the link flow.

From Figure 4.5, we observe that the minimum values of node voltages are nearly the same in the complete graphs, whereas they decrease dramatically in the path graphs with increasing size N. On the other hand, the cycle topology increases the node voltages, thus also the performance indicator of the node voltage, rapidly compared to the path graph.

Similar observations hold for the performance indicator of the power loss. In large path graphs, the total active power loss of the network is high, which decreases the related performance indicator. On the other hand, a complete graph nearly zeroes the power loss and the cycle topology significantly decreases the total loss compared to a path graph.

Similar to Figure 4.4, Figure 4.5 illustrates that a meshed topology can improve the

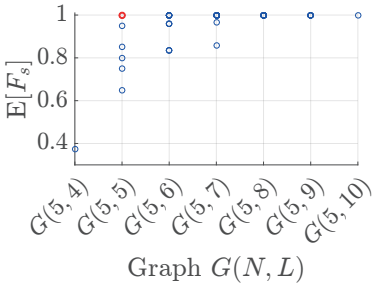


Figure 4.6: The expected E[F_s] fraction of served demands after a random link failure throughout the topological transformation of a path graph with 5 nodes. The bold red data point corresponds to the cycle graph. The remaining links are assumed to have enough rest flow capacity to handle the redistributed flows due to a random link failure.

key performance indicators compared to a path graph. We conclude that the core contributions to the key performance indicators arise from the first few links added to a path graph. In particular, for larger graphs, a cycle topology can dramatically increase the voltage magnitude and decrease the total active power loss of the network compared to the path graph. Consequently, adding a limited number of links to the tree topology can still achieve higher levels of performance during the electric power transmission between a supply and demand nodes.

4.4.2. KEY PERFORMANCE INDICATORS UNDER A SINGLE LINK FAILURE CONTINGENCY

In this section, we investigate the effect of a single link failure in different graphs. Initially, we focus on the effect of a link failure on the served demands of the network. Figure 4.6 illustrates the expected E[F_s] fraction of served demands after a random link failure throughout the topological transformation of a path graph $G(5,4)$ with 5 nodes and 4 links. In Section 4.3, we show that any link removal from a path graph partitions the graph. Figure 4.6 also depicts that only a cycle or augmented cycles can provide a back-up after any random link failure. The other graphs may partition after a random link failure and can continue their operation only with a decreased total demand, which usually improves the key performance indicators. Therefore, in this subsection, we only focus on the graphs that can provide a back-up after any random link failure.

To investigate and compare the effects of single link failures in each graph, we remove one link at a time from the graph, and calculate the changes in the performance indicators. We repeat this link failure contingency simulation for each link, and compare all changes in the link flow, the node voltage and the active power loss in the network. Figure 4.7 illustrates the maximum resulting changes in these key performance indicators after a link removal throughout the topological transformation of a path graph $G(5,4)$ with 5 nodes and 4 links. The performance indicator of link flow can significantly decrease after a link failure in a cycle graph. For the complete graph, on the other hand, we observe that the effect of a link failure on the indicator of link flow is very small. Similar to the changes in the indicator of link flow, the decreases in the performance indicators

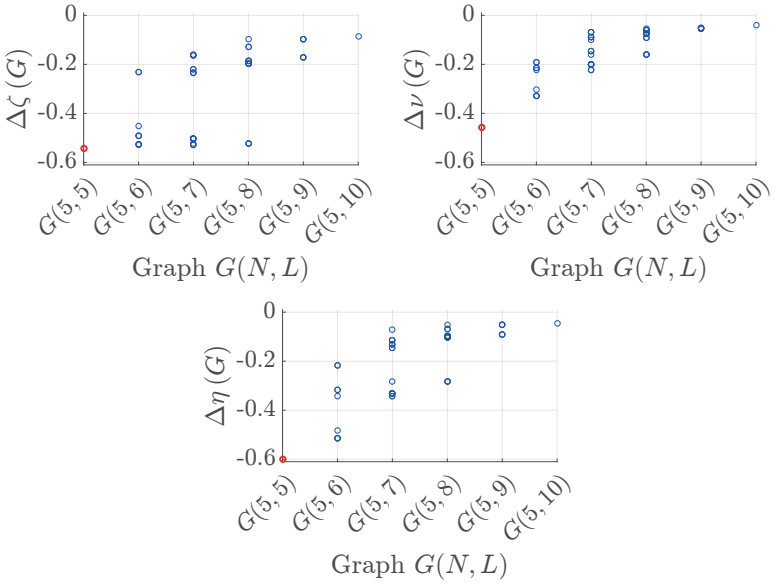


Figure 4.7: The variations of the changes Δ in the key performance indicators after a link failure throughout the topological transformation of a path graph with 5 nodes. The bold red data point corresponds to the cycle graph. The performance indicators are evaluated for $r = 0.1$, $x = 0.1$, $p = 0.05$, $q = 0.01$, $\tau_f = 5p$, $\tau_v = 0.9$ and $\tau_\sigma = 0.2p$.

of node voltage and power loss are highest in the cycle graph after a single link failure.

Finally, in Figure 4.8, we present the variations of key performance indicators after a link failure in graphs with different sizes N. Similar to the theoretical calculation in Equation (4.5), the effect of a link failure on the remaining link flows slightly decreases with the increasing size N in the complete graphs. On the other hand, in cycle graphs, the change in the magnitude of the flow through a remaining link can significantly increase with the increasing size N, which decreases the related performance indicator. In the *worst* case, when one of the links adjacent to the supply node fails in a cycle graph, it operates as a path graph with the same size N after the link failure. Therefore, the performance indicator of link flow in a cycle graph under a single link failure contingency becomes the performance indicator of link flow in a path graph under normal operation.

Similar to the changes in the performance indicator of link flow after a link failure, the indicators of node voltage and power loss slightly decrease in a complete graph with increasing size N. In the other graphs, however, the decrease in the key performance indicators could be drastic. Although under the normal operation, the cycle and the augmented cycles can provide higher values of the performance indicators; after a link failure, large drops on the key performance indicators in those graphs are expected.

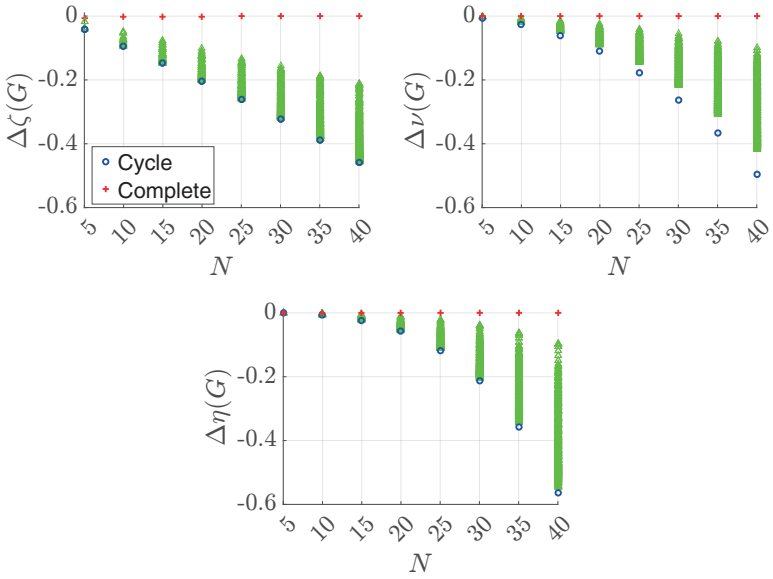


Figure 4.8: The variations of the changes Δ in the key performance indicators after a link failure in graphs with different sizes N. The green upper triangle ($\triangle$) data points correspond to the graphs that are constructed by adding one link to the cycle graph. The performance indicators are evaluated for $r = 0.02$, $x = 0.02$, $p = 0.005$, $q = 0.001$, $\tau_f = 50p$, $\tau_\nu = 0.8$ and $\tau_\sigma = 4p$.

4.5. CONCLUSION

In this chapter, we investigated the impact of topology on the electric power transmission and the performance of power grids. By utilizing a graph theoretical approach, we first focused on two extreme graphs, complete and path graphs, and analysed the electric power transmission under normal operation and under single link failure contingencies. We showed that in complete graphs, due to the redistributed flows, the survival of power grids from a random link failure depends on the rest flow capacity of the remaining links. Consequently, when the rest flow capacities are insufficient to handle the redistributed flows, a single link failure could result in more link and demand loss in a complete graph than in a path graph.

Subsequently, we empirically investigated the effect of the electric power transmission on the link flow, the node voltage and the active power loss in power grids in various other graphs. Our results show that adding few links to a path graph can significantly improve these key performance indicators of power grids compared to a path graph. However, at the same time, the performance indicators could also remarkably decrease after a link failure. Consequently, throughout a topological transformation towards a meshed topology with loops, redundancies in the design parameters of power grids are needed to ensure safety under normal operation and as well as under single link failure contingencies.

4.6. APPENDIX

We model power grids with N buses (nodes), and L lines (links) by a weighted graph $G(N,L)$. The $N \times N$ weighted adjacency matrix $\mathbf{W}$ specifies the interconnection pattern of the graph $G(N,L)$: w_{ik} is non-zero only if the nodes i and k are connected by a link; otherwise $w_{ik} = 0$. In the slack-bus independent solution of the DC power flow equations [23], the relation between the phase angles $\Theta = [\theta_1 \dots \theta_N]^{\mathrm{T}}$ of node voltages, and the power input $\mathbf{P} = [p_1 \dots p_N]^{\mathrm{T}}$ is given as [23]

$$\Theta = \mathbf{Q}^{+}\mathbf{P}, \tag{4.9}$$

where $\mathbf{Q}^{+}$ is the pseudo-inverse of the Laplacian $\mathbf{Q}$ of the weighted graph $G(N,L)$.

4

4.6.1. OPERATING CONDITIONS IN A PATH GRAPH

For a path graph, whose nodes are numbered consecutively and links have equal link weights $b > 0$, the structure of weighted Laplacian $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \begin{bmatrix} b & -b & 0 & \dots & 0 & 0 \\ -b & 2b & -b & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -b & 2b & -b \\ 0 & 0 & \dots & 0 & -b & b \end{bmatrix}.$$

In order to find $\mathbf{Q}^{+}$ in Equation (4.9), we use the definition [35] of the pseudo-inverse of Laplacian $\mathbf{Q}^{+} = \hat{\mathbf{X}}\mathbf{diag}(\frac{1}{\mu_k})\hat{\mathbf{X}}^{\mathbf{T}}$, where the $N \times (N-1)$ matrix $\hat{\mathbf{X}}$ consists of all the normalized eigenvectors of $\mathbf{Q}$, except for the eigenvector $\mathbf{u}$ belonging to eigenvalue $\mu = 0$, and where the $(N-1) \times (N-1)$ diagonal matrix $\mathbf{diag}(\frac{1}{\mu_k})$ contains the positive eigenvalues of Laplacian $\mathbf{Q}$.

The positive eigenvalues μ_k and the corresponding normalized eigenvector elements $\hat{\mathbf{X}}(v,k)$ of the weighted Laplacian of a path graph are [45]

$$\mu_k = 2b\big(1 - \cos(\frac{\pi k}{N})\big),$$
$$\hat{\mathbf{X}}(v,k) = \frac{\sqrt{2}}{\sqrt{N}} \cos(\frac{\pi k v}{N} - \frac{\pi k}{2N}),$$

where $1 \le k \le N-1$, and $1 \le v \le N$. Then, the elements q^{+}_{ik} of the pseudo-inverse of the Laplacian are

$$q^{+}_{ik} = \sum_{v=1}^{N-1} \frac{\hat{\mathbf{X}}(i,v)\hat{\mathbf{X}}(k,v)}{\mu_v}. \tag{4.10}$$

Inserting the elements of pseudo-inverse in Equation (4.10) and the power input $\mathbf{P} = [(N-1)p, 0, \dots, 0, -(N-1)p]^{\mathrm{T}}$ into the DC power flow equations in Equation (4.9) results in the operating conditions, i.e., the phase angles Θ of node voltages, when the electric power is transferred from the supply node 1 to the demand node N:

$$\begin{aligned}\theta_i &= p(N-1)(q_{i1}^{+} - q_{iN}^{+}) \\ &= \frac{p(N-1)(N-2i+1)}{2b}.\end{aligned} \tag{4.11}$$

4.6.2. OPERATING CONDITIONS IN A COMPLETE GRAPH

For a complete graph with equal link weights $b > 0$, the structure of the weighted Laplacian $\mathbf{Q}$ can be written as

$$\mathbf{Q} = b(N\mathbf{I} - \mathbf{J}), \tag{4.12}$$

where $\mathbf{J}$ is the all-one matrix, and $\mathbf{I}$ is the identity matrix. Using the definition of pseudo-inverse of the Laplacian [35]

$$\mathbf{Q}^{+} = (\mathbf{Q} + \alpha\mathbf{J})^{-1}(\mathbf{I} - \frac{1}{N}\mathbf{J}), \tag{4.13}$$

where $\alpha > 0$ is a scalar, and choosing the scalar $\alpha = b$, the pseudo-inverse of the weighted Laplacian of a complete graph can be found as

$$\begin{aligned}\mathbf{Q}^{+} &= (\mathbf{Q} + b\mathbf{J})^{-1}(\mathbf{I} - \frac{1}{N}\mathbf{J}) \\ &= \frac{1}{Nb}(\mathbf{I} - \frac{1}{N}\mathbf{J}) = \frac{1}{N^2 b}(N\mathbf{I} - \mathbf{J}).\end{aligned} \tag{4.14}$$

From Equation (4.9), the phase angles Θ of node voltages when the electric power transferred from the supply node 1 to the demand node N can be found as

$$\begin{aligned}\Theta &= \mathbf{Q}^{+}\mathbf{P} \\ &= \frac{1}{bN^2}\begin{bmatrix} (N-1) & -1 & \dots & -1 \\ -1 & \ddots & & \vdots \\ \vdots & & \ddots & -1 \\ -1 & \dots & -1 & (N-1) \end{bmatrix}\begin{bmatrix} (N-1)p \\ 0 \\ \vdots \\ -(N-1)p \end{bmatrix} = \frac{p}{bN}\begin{bmatrix} (N-1) \\ 0 \\ \vdots \\ -(N-1) \end{bmatrix}.\end{aligned}$$

4.6.3. SINGLE LINK FAILURE IN A PATH GRAPH

The failure and removal of a link l_{ik} from a network partitions its underlying graph if the equality between the reactance x_{ik} of the link and the effective resistance r_{ik} between its node pairs satisfies [23]

$$x_{ik} = r_{ik}.$$

When the underlying topology is a path graph, the effective resistance r_{ik} between nodes i and $k = i + 1$ can be written as

$$r_{ik} = |i - k| x_{ik},$$

meaning that the removal of any link l_{ik} partitions the path graph.

4.6.4. SINGLE LINK FAILURE IN A COMPLETE GRAPH

When a link l_{ik} with flow $|f_{ik}| = p$ is removed from the graph, the flow $|f_{ik}| = p$ through the link before its removal is redistributed over alternative paths between nodes i and k. Hence, the final flow through an arbitrary remaining link l_{ab} can be written as the sum of the previous state of the network, i.e., the previous flow through the link between nodes a and b when link l_{ik} is present, and the flow resulting from the change of the state due to the removal of link l_{ik}. The change in the flow Δf_{ab} through a remaining link l_{ab} can be calculated as [23]

$$\Delta f_{ab} = w_{ab} \frac{(r_{ak} - r_{ai} + r_{bi} - r_{bk})}{2(1 - w_{ik} r_{ik})} p, \tag{4.15}$$

where w_{ab} is the weight of the link l_{ab} and r_{ak} is the effective resistance between the nodes a and k.

The effective resistance between any two distinct nodes in the complete graph with equal link weights b is $\frac{2}{bN}$. Therefore, the numerator $(r_{ak} - r_{ai} + r_{bi} - r_{bk})$ of Equation (4.15) is nonzero and its magnitude is equal to $|r_{ak} - r_{ai} + r_{bi} - r_{bk}| = \frac{2}{bN}$ only when the removed link l_{ik} and the observed link l_{ab} share a node. Then,

$$|\Delta f_{ab}| = \begin{cases} 0 & \text{if } l_{ik} \cap l_{ab} = \varnothing, \\ \frac{p}{N-2} & \text{otherwise.} \end{cases}$$

If the failure probabilities of the links in a complete graph are the same, $\frac{1}{L}$, we can calculate the probability p_r that a failed link is a *redundant* link with zero flow, and the probability p_u that a failed link is a *used* link with non-zero flow as

$$p_r = \frac{(N-1)(N-2)}{2} \frac{2}{N(N-1)} = \frac{N-2}{N} = 1 - \frac{2}{N}, \tag{4.16}$$

$$p_u = (N-1) \frac{2}{N(N-1)} = \frac{2}{N}, \tag{4.17}$$

where $p_u + p_r = 1$.

5

Effects of Failures in Power Grids under the AC and DC Power Flow Models

In the previous chapter, we use the linearised DC power flow equations to capture the changes in link flows, whereas we use the nonlinear AC power flow equations for a more detailed analysis of power grids. In this chapter, we compare the effects of failures in power grids under both the nonlinear AC and the linearised DC power flow models. We numerically compare the effects of single line failures and the evolution of cascades under the AC and the DC flow models using different metrics, such as yield (the ratio of the demand supplied at the end of the cascade to the initial demand). Our numerical results in the tested networks demonstrate that the effects of a single line failure on the distribution of the flows through other lines are similar under the AC and DC power flow models. However, the cascade simulations in the tested networks suggest that the assumptions underlying the DC power flow model (e.g., ignoring power losses, reactive power flows, and voltage magnitude variations) can lead to inaccurate and overly optimistic cascade predictions.

This chapter is based on published papers [59, 82].

5.1. INTRODUCTION

POWER grids are vulnerable to external events, such as natural disasters and cyber-attacks, as well as to internal events, such as unexpected variability in load or generation, aging, and control device malfunction. The operation of a power grid is governed by the laws of physics [19], and the outage of an element may result in a cascade of failures and a blackout [24]. The recent blackouts in Turkey [2], India [3], the U.S. and Canada [1] had devastating effects and thus motivated the study of power grid vulnerabilities to cascading failures (e.g., [18, 24, 83–88]).

Some of the recent work on cascading failures considers a topological perspective where, once a network element fails, the neighboring elements also fail [78]. However, such topological models do not consider the flow behaviour of power grids. More realistic cascading failures models use the linearized DC power flow equations [30, 89]. The DC power flow equations are based on a linearization of the nonlinear AC power flow dynamics. The induced linearization error can be small in large transmission grids and high for some particular networks [34, 90]. Motivated by these observations, we study the effects of line failures and cascades under both the linearized DC model and a nonlinear AC model by performing simulations on four test networks.

Due to their complexity, the AC power flow equations are not as commonly used as the DC power flow equations in studying cascading failures in power grids. An AC power flow model is utilized in [87, 91], as well as in some (mostly commercial) software tools for modeling the evolution of the cascade [92]. Unfortunately, none of these tools is publicly available. Hence, for the evaluation in this chapter, we developed an AC cascading failures simulator [93], using the MATPOWER AC power flow solver [80].

Previous work on determining the accuracy of the DC power flow approximation includes [34, 90, 94–98]. However, these works did not consider accuracy of the DC flows in predicting the evolution of a cascade. In [87], the DC and the AC cascading failures are compared when all the buses (nodes) in the AC model are voltage controlled (PV) buses. *To the best of our knowledge, this work is the first to compare the evolution of cascades in power grids under the AC and DC power flow models in detail and for many of the publicly available power grid networks [42, 80].*

First, we numerically evaluate the accuracy of the DC power flow model when there are no failures. Then, we compare the effects of single line failures under the AC and the DC power flow models. We numerically demonstrate that the DC power flow model can capture the effects of a single line failure on the flow changes on other lines relatively close to the AC power flow model.

Subsequently, we present an AC cascading failures model that is based on the nonlinear power flow equations, and is therefore, is more realistic than the corresponding DC model. We empirically compare the AC and DC cascade models based on robustness metrics that quantify the operational and topological characteristics of the grid during a cascade for all cascading failures initiated by single and two line failures. Our simulations demonstrate that the assumptions underlying the DC model (assuming a lossless network and ignoring reactive power flows and voltage variations) can lead to inaccurate and overly optimistic cascade predictions.

Moreover, we empirically compare the AC and DC cascades under different supply and demand balancing and line outage rules. Our simulation results show that the dif-

ference between the cascade evolution under the AC and DC power flows depends on the balancing and line outage rules in power grids. In particular, the supply and demand balancing rule that separates the excess supply or demand from the grid increases the difference between the AC and DC models the most.

The remainder of this chapter is organized as follows. Section 5.2 presents the power flow equations. Section 5.3 presents the cascading failures models. Section 5.4 presents the numerical comparison of the AC and DC flow models in four different test networks and Section 5.5 concludes the chapter.

5.2. POWER FLOW EQUATIONS

In this section, we provide details on the AC and DC power flow equations.

5.2.1. AC POWER FLOW EQUATIONS

A power grid with n nodes (buses) and m transmission lines constitutes a complex network whose underlying topology can be represented by an undirected graph $\mathcal{G}(\mathcal{N},\mathcal{L})$, where $\mathcal{N}$ denotes the set of nodes and $\mathcal{L}$ denotes the set of lines. Each line l has a predetermined capacity c_l that bounds its flow $|f_l|$ under a normal operation of the system. The status of each node i is represented by its voltage $V_i = |V_i|e^{\mathbf{i}\theta_i}$ in which $|V_i|$ is the voltage magnitude, θ_i is the phase angle at node i, and $\mathbf{i}$ denotes the imaginary unit.

The goal of an AC power flow analysis is the computation of the voltage magnitudes and phase angles at each bus in steady-state conditions [99]. In the steady-state, the injected apparent power S_i at node i equals to

$$S_i = V_i(\mathbf{YV})_i^* \tag{5.1}$$

where * denotes the complex conjugation, $\mathbf{V} = [V_1,\ldots,V_n]^{\mathrm{T}}$ is the vector of node voltages, and $\mathbf{Y}$ is the $n \times n$ bus-admittance matrix [19]. The elements of the admittance matrix $\mathbf{Y}$, which depend on the topology of the grid as well as the admittance values of the lines.

Rewriting the admittance matrix as $\mathbf{Y} = \mathbf{G} + \mathbf{iB}$ where $\mathbf{G}$ and $\mathbf{B}$ are real matrices, and using the definition of the apparent power $S_i = P_i + \mathbf{i}Q_i$ in (5.1) leads to the equations for the active power P_i and the reactive power Q_i at each node i:

$$P_i = \sum_{k=1}^{n} |V_i||V_k|(G_{ik}\cos\theta_{ik} + B_{ik}\sin\theta_{ik}) \tag{5.2}$$

$$Q_i = \sum_{k=1}^{n} |V_i||V_k|(G_{ik}\sin\theta_{ik} - B_{ik}\cos\theta_{ik}) \tag{5.3}$$

where $\theta_{ik} = \theta_i - \theta_k$.

In the AC power flow analysis, each node i is categorized into one of the following three types:

1. *Slack node*: The node for which the voltage is typically 1.0. For convenience, it is indexed as node 1. The slack node compensates for network losses by emitting or absorbing power. The active power P_1 and the reactive power Q_1 need to be computed.

2. *Demand node*: The active power P_i and the reactive power Q_i at these nodes are known and the voltage V_i needs to be computed.

3. *Supply node*: The active power P_i and the voltage magnitude $|V_i|$ at these nodes are known and the reactive power Q_i and the phase angle θ_i need to be computed.

5.2.2. DC POWER FLOW EQUATIONS

The AC power flow equations are nonlinear in the voltages. The DC power flow equation provides a linearized approximation of the active power flows in the AC model. Linearization is possible under the following conditions [99]:

1. The difference between the voltage phase angles of every couple of neighboring nodes is small such that $\sin\theta_{ik} \approx \theta_{ik}$ and $\cos\theta_{ik} \approx 1$.

2. The active power losses are negligible, and therefore, $\mathbf{Y} \approx \mathbf{iB}$ where $\mathbf{B}$ is the imaginary part of the admittance matrix $\mathbf{Y}$, calculated neglecting the line resistances.

3. The variations in the voltage magnitudes $|V_i|$ are small and, therefore, it is assumed that $|V_i| = 1\ \forall i$.

Under these assumptions, given the active power P_i at each node i, the phase angle of the nodes can be estimated by $\tilde{\theta}_i$ using the DC power flow equations as follows:

$$P_i = \sum_{k=1}^{n} B_{ik}(\tilde{\theta}_i - \tilde{\theta}_k) \tag{5.4}$$

or in matrix form,

$$\tilde{\mathbf{P}} = -\mathbf{B}\tilde{\Theta} \tag{5.5}$$

where $\tilde{\mathbf{P}} = [\tilde{P}_1, P_2, \ldots, P_n]^{\mathrm{T}}$, $\tilde{\Theta} = [\tilde{\theta}_1, \ldots, \tilde{\theta}_n]^{\mathrm{T}}$. Notice that the vectors $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are equal except in the *slack node* (first entry) since in the DC power flows, the lines are lossless and therefore $\tilde{P}_1 + \sum_{i=2}^{n} P_i = 0$.

By assuming that the phase angle at the slack node is 0, the phase angle of the nodes can be estimated uniquely by solving (5.5) for the DC power flow.

In Section 5.4, we numerically compare the AC and DC power flow models.

5.3. MODELING CASCADING FAILURES

An initial failure in power grids may result in subsequent failures in other parts of the grid. These consecutive failures following an initial failure constitute a *cascading failure*. In this section, we follow [30, 84, 87, 100] and develop models for cascading failures due to line failures in power grids.

Before a cascading failure, we assume that $\mathcal{G}(\mathcal{N}, \mathcal{L})$ is connected, the power flows satisfy (5.2) and (5.3) or (5.5), and the flow magnitude $|f_l|$ of each line is at most its capacity c_l.

Next, we describe the cascading failures models. When an initial set of lines fail, they are removed from the network. As a result of this removal, the network topology is changed, and the power grid can be divided into one or more connected components.

Following [30], we assume that each connected component can operate autonomously. If there is no supply or no demand within a connected component $\mathcal{G}_k$, the component becomes a *dead component*, and all the demand or supply nodes within the component are put out of service. If there are both supply and demand nodes within a connected component $\mathcal{G}_k$, the connected component remains an *alive component*, but the supply and demand within the component should be balanced. We use two different supply and demand balancing rules [30, 84, 87]:

1. *Shedding and curtailing:* The amount of the power supply or demand are reduced at all nodes by a common factor. If the total active power supply is more than the total active power demand in a connected component $\mathcal{G}_k$, the active power outputs of supplies are curtailed. On the other hand, if the total active power supply is not sufficient to serve the total active power demand, demand shedding is performed to balance the supply and demand within $\mathcal{G}_k$.

2. *Separating and adjusting:* Excess supplies or demands are separated from the grid. In this case, we assume that the dynamic responses of the supplies (or vice versa, demands) are related to their sizes [19]. Namely, the supplies (demands) with lower amounts of power output are assumed to be faster to respond to the imbalances between supply and demand. Thus, within each component $\mathcal{G}_k$ with excess supply (demand), the supplies (demands) are separated from the grid according their sizes from the smallest to largest until the removal of one more supply (demand) results in the shortage of supply (demand). Then, the active power output (demand) of the largest supply (demand) node is reduced in order to balance supply and demand.

After supply and demand are balanced within each alive component using the selected balancing rule, the power flow equations are solved to compute new flows through the lines. Note that the line capacities are not taken into account in determining the flows. The new set of line failures are then found in all alive components. In a practical power grid, transmission lines are usually protected by relays and circuit breakers. A relay of a transmission line measures certain variables, and compares them with threshold values. When the threshold is violated, and/or this violation lasts long enough, the relay notifies the circuit breaker to trip the transmission line. To approximate this tripping mechanism, we use two different line outage rules [84, 87, 100]:

1. *Deterministic:* A line l fails when the power flow magnitude through that line, denoted by $|f_l|$, exceeds its capacity c_l.

2. *Probabilistic:* A line l fails with probability p_l at each stage of the cascade. We assume that each line l with a flow capacity c_l has also a nominal power flow level $\xi_l \in [0, c_l]$, after which the line may fail with a certain probability (due to increase in line temperature etc.). Under this model, the probability p_l is approximated as:

$$p_l = \begin{cases} 0, & \text{if } |f_l| < \xi_l \\ \frac{|f_l| - \xi_l}{c_l - \xi_l}, & \text{if } \xi_l \le |f_l| < c_l \\ 1, & \text{if } |f_l| \ge c_l. \end{cases} \tag{5.6}$$

After finding the new set of line failures using the selected line outage rule, the cascade continues with the removal of those lines. If there are no new line failures in any of the alive components, the cascade ends.

In this chapter, we study three cascade processes:

I) Cascade with shedding and curtailing balancing rule and deterministic line outage rule,

II) Cascade with separating and adjusting balancing rule and deterministic line outage rule,

III) Cascade with shedding and curtailing balancing rule and probabilistic line outage rule.

In order to study the differences between the AC and DC models, we mostly focus on the cascade process I with shedding and curtailing balancing rule and deterministic line outage rule. In order to further capture the effects of these processes on the differences obtained under the AC and DC cascade models, in Subsection 5.4.5, we briefly compare the three cascade processes.

In the following two subsections, we provide the details of the cascade models under the AC and DC power flows.

5.3.1. AC CASCADING FAILURES MODEL

In the cascade under the AC power flow model, the flows are composed of active parts P_i in (5.2) and reactive parts Q_i in (5.3). Hence, the apparent power S_i in (5.1) is used to calculate the flows. In general, due to transmission line impedances, the voltage at the sending node of a line is different than the one at the receiving node, resulting in different values of the apparent power flows at each side of the line. Hence, in the cascade under the AC model, we define the magnitude $|f_l|$ of flow through a line $l = \{i, k\}$ as follows:

$$|f_l| = \frac{|S_{ik}| + |S_{ki}|}{2}. \tag{5.7}$$

The difference, $S_{ik} - S_{ki}$, between the sent and received apparent flows through a line l represents the power loss over that line. The sum of the losses over all the lines is the total loss in the network. The total loss cannot be calculated in advance and is only known after the power flow equations in (5.1) are solved. Therefore, in the cascade under the AC flow, a part of the total supply in the network is reserved to supply the network losses and denoted by the *reserved loss factor* η.

The case of zero reserved loss factor, $\eta = 0$, means that no reserve supply is allocated for network losses, whereas a large reserved loss factor η corresponds to a large reserve supply for the network losses. Once the power flow equations are solved and the network losses are calculated, the difference between the allocated supply and the total demand with losses is compensated by the slack-node. Therefore, in the AC cascading failures model, the simulation is slack-node dependent, and for every alive component without such a node, a slack-node must be assigned. The developed model chooses the slack-node as the supply node with the maximum power output in that alive component.

The iterative process of solving the AC power flow equations (5.2) and (5.3) may result in the absence of a solution or a divergence in iterations. In such cases, it is perceived that the connected component cannot function at those operational conditions, and supply and demand shedding is applied. The amount of active and reactive power demands, and active power supply within that component are decreased until either convergence is reached in the flow equations or the component becomes a dead component with no demand.

We numerically study the three cascade processes under the AC power flow model in Section 5.4.

5.3.2. DC CASCADING FAILURES MODEL

In the cascade under the DC power flow model, the magnitude $|f_l|$ of the flow through a line $l = \{i, k\}$ is equal to the magnitude of active power flow in (5.4) on that line:

$$|f_l| = |P_{ik}| = |P_{ki}|. \tag{5.8}$$

Since the network is assumed to be lossless, the magnitude of the active power at the sending side of a line is equal to the magnitude of active power at the receiving side, $|P_{ik}| = |P_{ki}|$, and the total supply is equal to the total demand. Therefore, the supply and demand balancing is performed without a reserved loss factor η. Moreover, the no-loss assumption means that the flows in the network are slack-node independent.

Contrary to the AC power flow equations (5.2) and (5.3), which are nonlinear, the DC power flow equations (5.5) are linear, and a solution always exists for a connected network with balanced supply and demand [23]. Hence, no supply or demand shedding due to convergence issues is needed in the DC model.

We numerically study the three cascade processes under the DC power flow model in Section 5.4.

5.4. NUMERICAL COMPARISON OF THE AC AND DC FLOW MODELS

This section presents the numerical comparison of the AC and DC power flow models. After providing the simulations setup, we numerically evaluate the accuracy of the DC power flow model when there are no failures. Then, we compare the effects of single line failures, and the evolution of the cascade process I initiated by single and two line failures under the AC and DC flow models. Next, we compare the three cascade processes under the AC and DC flow models. Finally, we discuss the main lessons learned from the simulations.

5.4.1. SIMULATIONS SETUP

METRICS

We define metrics for evaluating the grid vulnerability (some of which were originally used in [23, 89, 100, 101]). To study the effects of a single line e failure on the flows through other lines we define:

▶ **Line flow change ratio** ($s_{l,e}$): the ratio $\Delta f_l / f_l$ of the change Δf_l in the flow through a line l due to the failure at line e to its original flow value f_l.

► **Line outage distribution factor** ($m_{l,e}$): the ratio $\Delta f_l / f_e$ of the change Δf_l in the flow through a line l due to the failure at line e to the flow value f_e of the failed line e.

Additionally, we define the following metrics to measure other dynamics of the system after a single line failure, which can only be captured under the AC power flow model due to the DC power flow assumptions 2 and 3 in Section 5.2.2:

► **Node voltage change** ($\Delta v_{i,e}$): the change in the voltage magnitude at node i after the failure at line e.

► **Power loss change ratio** ($\Delta p_{\mu,e}$): the ratio of the change in the active power output of the slack generator due to the failure at line e to the initial loss.

We also define metrics to evaluate the *cascade severity*:

► **Node-loss ratio** ($N_{\mathcal{G}}$): the ratio of the total number of failed nodes (i.e., nodes in dead components) at the end of the cascade to the total number of nodes.

► **Line-loss ratio** ($L_{\mathcal{G}}$): the ratio of the total number of failed lines at the end of the cascade to the total number of lines.

► **Yield** ($Y_{\mathcal{G}}$): the ratio of the demand supplied at the end of the cascade to the initial demand.

5

In addition to the previous metrics which capture the overall effect of a cascading failure on a power grid, we identify the frequently overloaded lines that may cause cascading failures to persist. Hence, we define

► **Line-vulnerability ratio** (R_l): the total number of cascading failures in which line l is overloaded over the total number of cascading failures simulations. Higher values of R_l indicate the vulnerability of the line l as a possible bottleneck in the network.

PROPERTIES OF THE NETWORKS USED IN SIMULATIONS

We considered four realistic networks: the IEEE 30-bus, the IEEE 118-bus, and the IEEE 300-bus test systems [42], as well as the Polish transmission grid [80]. The details of these networks are as follows.

► **The IEEE 30-bus test system** contains 30 nodes and 41 lines with a total power demand of 189.2 MW.

► **The IEEE 118-bus test system** contains 118 nodes and 186 lines with a total power demand of 4242 MW.

► **The IEEE 300-bus test system** contains 300 nodes and 411 lines with a total power demand of 23,525.85 MW.

► **The Polish transmission grid**, at summer 2008 morning peak, contains 3120 nodes and 3693 lines with a total power demand of 21,181.5 MW.

In the IEEE test networks, maximum line flow capacities are not present. Following [30], the line flow capacities are estimated as $c_l = (1+\alpha)\max\{|f_l|, \overline{f}\}$, where α is the line tolerance parameter, and $\overline{f}$ is the mean of the initial magnitude of line flows. For the remainder of this chapter, we assume $\alpha = 1$.

In the Polish transmission grid data, emergency ratings are used for the flow capacities of the network. In order to eliminate existing overloaded transmission lines at the base case operation, the line flow capacities of such overloaded lines are changed to $c_l = (1+\alpha)|f_l|$ where $\alpha = 1$.

POWER FLOW SOLVER

In the simulations, we used MATPOWER [80] package in MATLAB for solving the AC and DC power flows.

5.4.2. NO FAILURES CASE

In this section, we numerically evaluate the accuracy of the DC power flow model when there are no failures in four test networks. First, we check the validity of the assumptions underlying the DC power flow approximation (as mentioned in Section 5.2.2). Then, we compute the absolute difference between the AC and DC power flow models.

Figure 5.1 shows the cumulative distribution functions (CDFs) of the absolute difference between the voltage phase angle of neighboring nodes, the ratio of the real to the imaginary part of the admittance values, the deviation of the voltage magnitudes from 1 p.u., and the absolute difference of the AC and DC active power flow through links.

In particular, Figure 5.1a demonstrates that the difference between the voltage phase angles of neighboring nodes (condition 1) is less than 0.1 for 80% of the pairs in all test networks. Figure 5.1b shows that the imaginary part of the admittance values are dominant (condition 2) in the test networks. Figure 5.1c shows that the voltage magnitudes are close to 1.0 (condition 3) for all the nodes. In Figure 5.1d, the differences between the AC and DC power flows is less than 0.2 (p.u.) for nearly 80% of the lines.

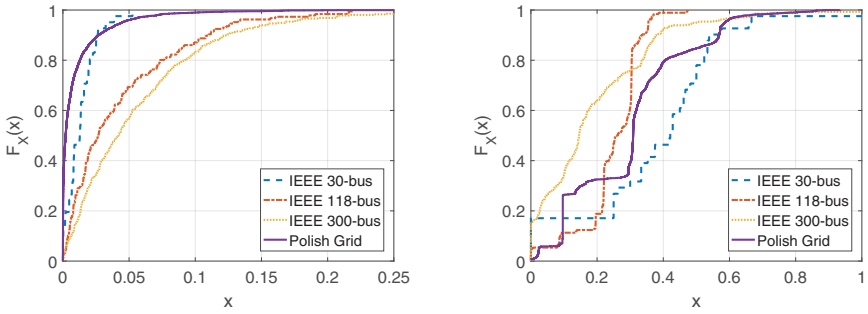
Figure 5.1 demonstrates that when the assumptions underlying the DC power flow approximation are valid, the DC power flows could approximate the AC power flows of most of the network lines relatively well when there are no failures. In the following subsections, we show that upon cascading failures, however, the DC approximation may become inaccurate. Moreover, the small differences between the AC and DC power flows in different cascade stages may lead to drastic differences at the end of the cascade.

5.4.3. COMPARISON OF THE SINGLE LINE FAILURE EFFECTS

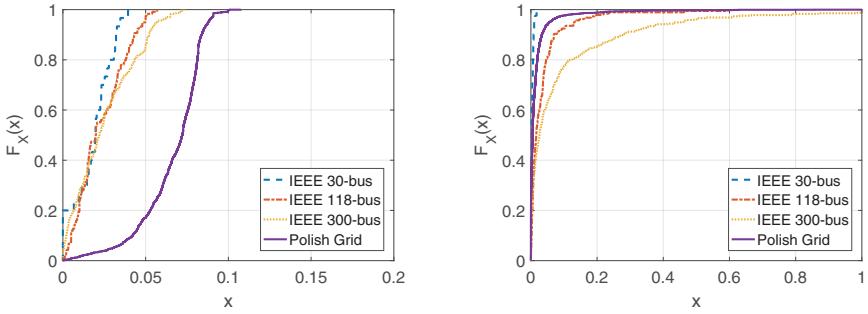
Single line failure and its consequent removal is the first stage and the triggering event of possible cascading failures. In this section, we perform empirical studies on single line failures in four realistic networks. Since the line flow change ratios $s_{l,e}$ for the lines with a low initial flow can be unreasonably high [89], these values are calculated only for the lines whose initial flow is larger than the mean flow. Additionally, to capture the variations of the line outage distribution factors $m_{l,e}$ and power loss change ratios $\Delta p_{\mu,e}$, line failures that partition the network are not considered in the set of failed lines.

Figure 5.2 presents the CDFs of the differences in the line flow change ratios and line outage distribution factors calculated based on the AC and DC flows. These results show that the differences decrease with the size of the network. In nearly 80% of the observed values in the IEEE 30-bus network, the magnitudes of the differences in the line flow change ratios $s_{l,e}$ and the line outage distribution factors $m_{l,e}$ are smaller than 0.05, whereas, in the Polish transmission grid this percentage is nearly 98%.

Since the DC power flow model cannot capture the node voltage changes $\Delta v_{i,e}$ (the node voltages are always equal to 1 under the DC model) and the power loss change ratios $\Delta p_{\mu,e}$ (the network is assumed to be lossless under the DC model) after a line failure, the CDFs of these two metrics are shown in Figure 5.3 only for the AC flow model. Figure 5.3a shows the absolute changes in the magnitude of the node voltages due to a line

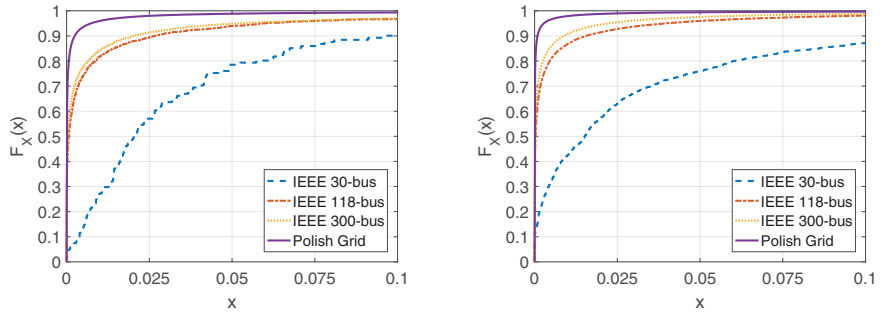


(a) $X := |\theta_i - \theta_j|$ (b) $X := |g_{ij}/b_{ij}|$



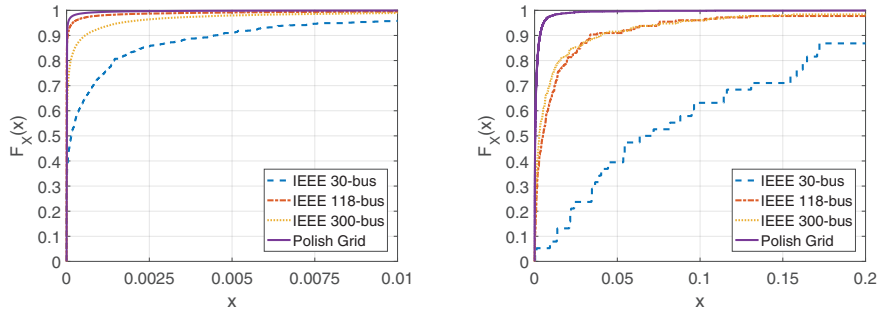
(c) $X := ||V_i| - 1|$ (d) $X := |P_{ij}^{(AC)} - P_{ij}^{(DC)}|$

Figure 5.1: The validity of the assumptions underlying the DC power flow approximation and the resulting difference between the AC and DC power flow models: the CDFs of (a) the absolute difference between the voltage phase angle of neighboring nodes, (b) the ratio of the real to the imaginary part of the admittance values, (c) the deviation of the voltage magnitudes from 1, and (d) the absolute difference of the AC and DC active power flows.



(a) $X := |s^{AC}_{l,e} - s^{DC}_{l,e}|$ (b) $X := |m^{AC}_{l,e} - m^{DC}_{l,e}|$

Figure 5.2: The CDFs of the differences in the line flow change ratios and the line outage distribution factors based on the AC and DC flow models.



(a) $X := |\Delta v^{AC}_{i,e}|$ (b) $X := |\Delta p^{AC}_{\mu,e}|$

Figure 5.3: The CDFs of the magnitudes of node voltage changes and the power loss change ratios after a single line failure for all the test networks under the AC flow model.

failure using the AC model. Both increase and decrease in the values of the node voltages are observed. However, the probability of a decrease is higher as the system continues to operate with fewer lines.

Figure 5.3b illustrates the power loss change ratios after a line failure using the AC model. A line failure can lead to an increase or a decrease in the slack node power output. However, the probability of a decrease is quite low since the system's loss generally increases when lines are removed from the grid.

Similar to our observations in Figure 5.2, the node voltage changes and power loss ratios generally become smaller with the size n of the network. For the Polish transmission grid, the obtained values of nearly all the node voltage changes and power loss change ratios are smaller than 0.005 and 0.05, respectively.

5.4.4. COMPARISON OF THE CASCADE PROCESS I EVOLUTION UNDER THE AC AND DC MODELS

The models introduced in Sections 5.3.1 and 5.3.2 are used to simulate cascading failures under the AC and DC flow models, respectively. For a fair comparison between the AC and DC models, the loss factor in the AC cascading failures model (in Section 5.3.1) is taken to be zero. Moreover, the cascade process I is used in this subsection in order to focus on the differences between the AC and DC models.

CASCADING FAILURES INITIATED BY A SINGLE LINE FAILURE

An example of a cascade initiated by a single line failure in the IEEE 118-bus network under the two cascade models is shown in Figure 5.4. The basic observation from this figure is that the evolution of the cascade under the two models can be quite different. For instance, in Figure 5.4a, there are two overloaded lines at the first stage of the cascade under the AC model which are not overloaded under the DC model. This initial difference results in a considerable difference in the evolution of the cascade: An important flow path in the AC model is failed at the first stage, resulting in more severe consecutive stages. Therefore, the differences between the AC and DC models accumulate at each cascade stage and may lead to *a drastic difference at the end of the cascade.*

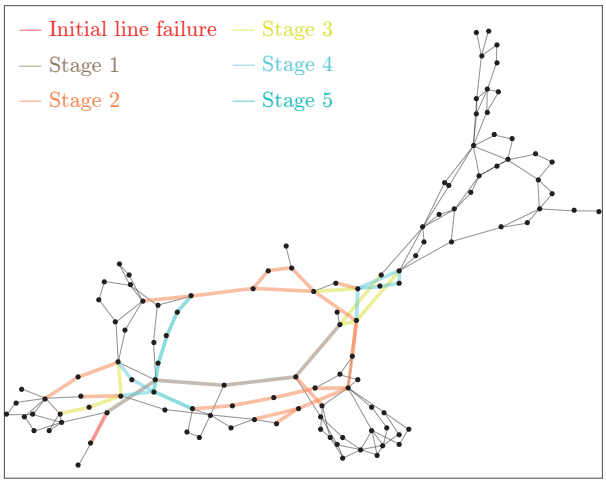
To further investigate the differences, we simulate cascading failures due to all single line failures whose initial flows were larger than the mean of initial flows in the four test networks. Figures 5.5, 5.6, 5.7, and 5.8 provide the detailed results obtained under the two cascade models.

Figure 5.5 shows the scatter plot of the yield values under the two models for the four test networks. It suggests that the yield values obtained by the DC cascade model are usually higher, specially for large networks. Moreover, Figure 5.8a, which presents the CDFs of the differences in yield values for all the test networks, also shows that the differences in the obtained yield values can grow quite high in large networks.

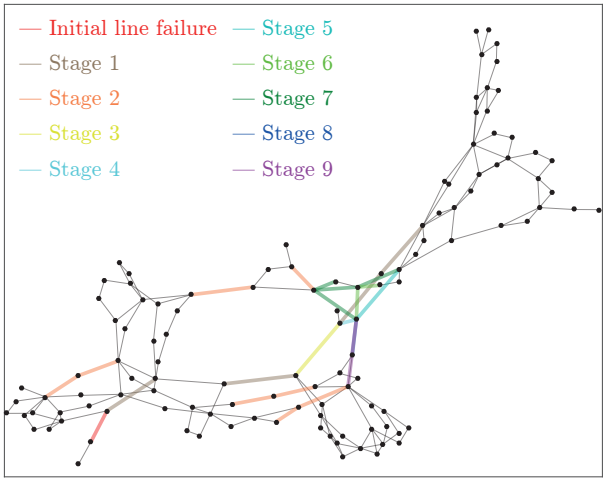
In Figure 5.6 and Figure 5.8b, however, the line-loss ratios are observed to be close under the two cascade models in all the four networks. The same is true for the node-loss ratios (see Figure 5.8c). Despite the similarity of the line-loss and node-loss ratios under the two cascade models, Figure 5.7, which presents the line-vulnerability ratios, suggests that as networks become larger, the individual lines that fail frequently under the AC model are very different from their counterparts under the DC model (see Figures 5.7c and 5.7d). Figure 5.8d also shows that the differences in the line-vulnerability ratios are close for most of the lines, but the differences may be quite large for roughly 10% of the lines in large networks.

CASCADING FAILURES INITIATED BY TWO-LINE FAILURES

We study cascades that are triggered by two-line failures. Two-line combinations of all lines whose initial flows are larger than the mean initial flows are investigated in the IEEE 30- and 118-bus networks, whereas, in the IEEE 300-bus network and the Polish transmission grid, 1000 random two-line removals are selected out of those combinations. The same set of results as in the previous section are presented in Figures 5.9, 5.10, 5.11, and 5.12. Similar observations as in the previous section can be made from these figures for the differences in the cascades initiated by two line failures under the AC and DC cascade models.

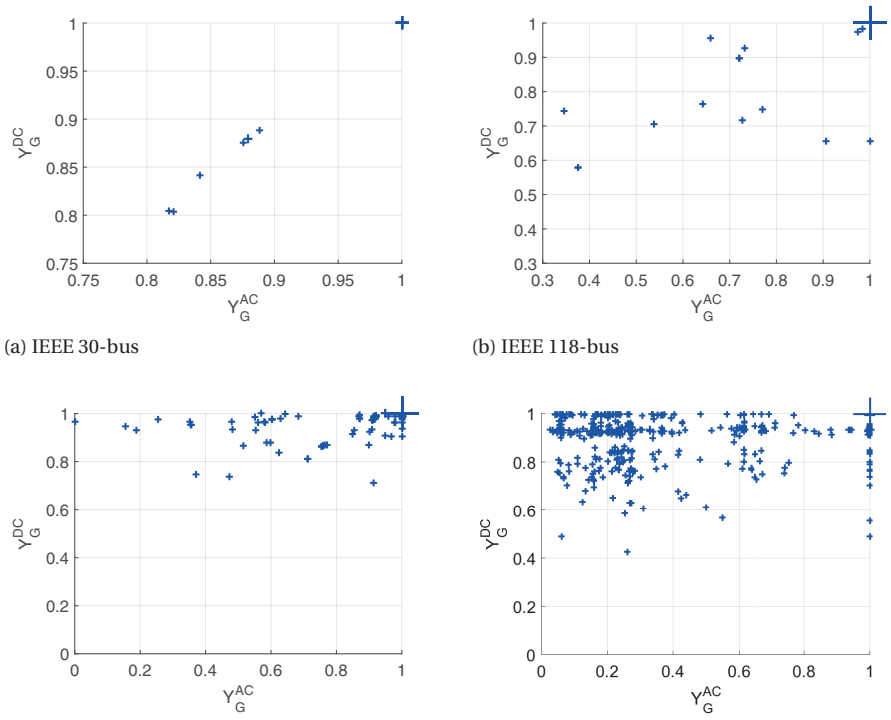


(a) AC cascading failures model



(b) DC cascading failures model

Figure 5.4: Evolution of a cascade initiated by a single line failure in the IEEE 118-bus network under the AC and DC cascade models. The remaining load at the end of the simulation is 1594.5 MW under AC cascading failures model, and 2446.3 MW under DC cascading failures model.



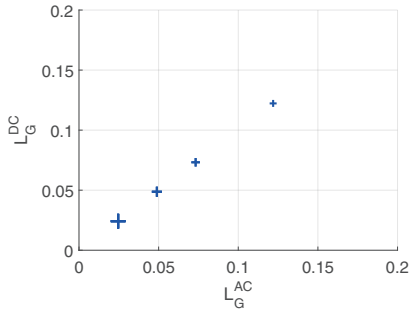
(a) IEEE 30-bus

(b) IEEE 118-bus

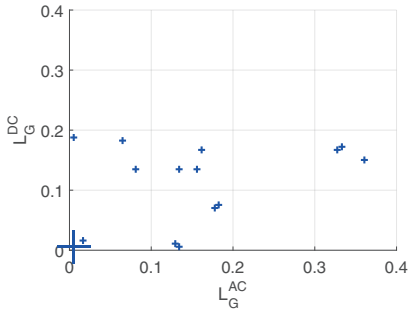
(c) IEEE 300-bus

(d) Polish Grid

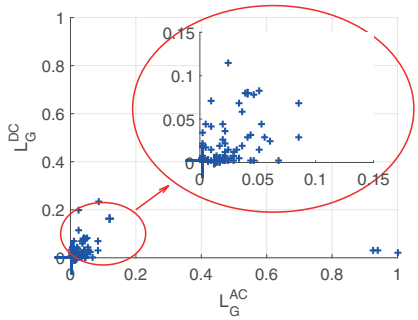
Figure 5.5: The scatter plots of the yield values under the AC versus DC cascade models initiated by single line failures. Markers are scaled according to the frequencies of corresponding data points.



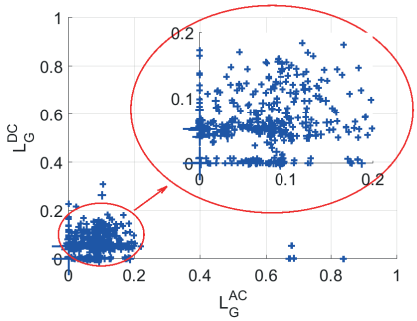
(a) IEEE 30-bus



(b) IEEE 118-bus

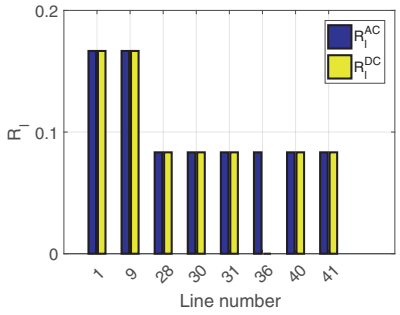


(c) IEEE 300-bus

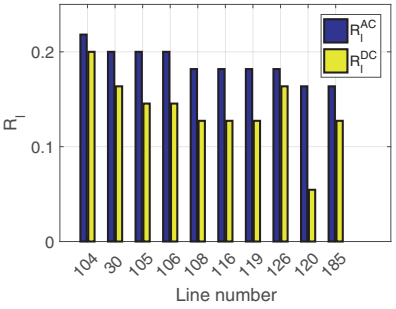


(d) Polish Grid

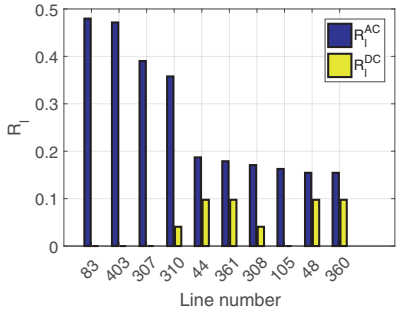
Figure 5.6: The scatter plots of the line-loss ratios under the AC versus DC cascade models initiated by single line failures. Markers are scaled according to the frequencies of corresponding data points.



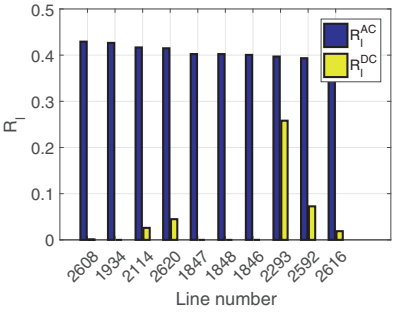
(a) IEEE 30-bus



(b) IEEE 118-bus

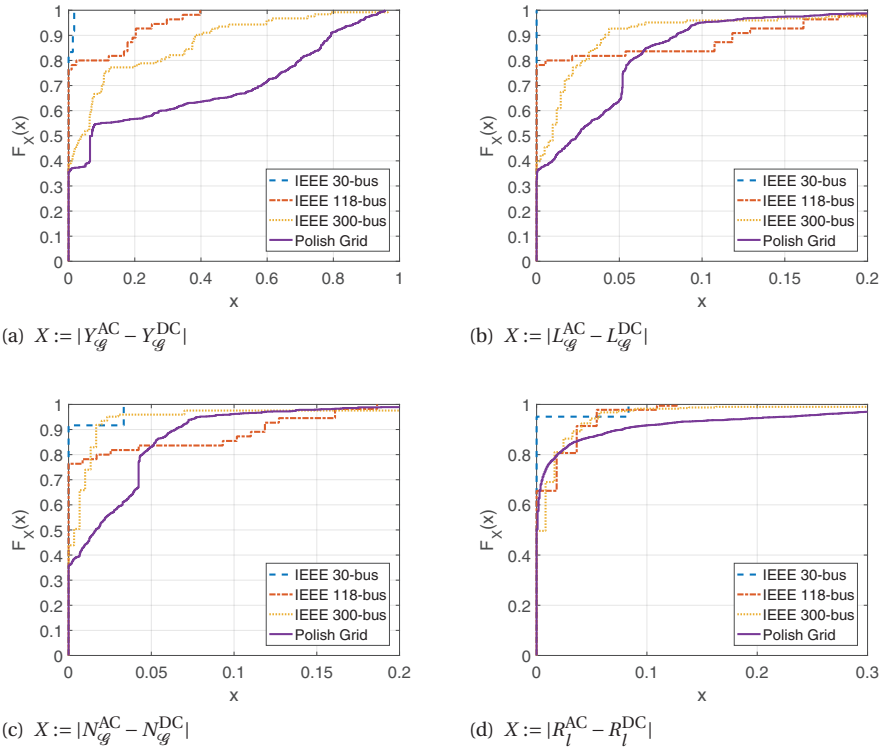


(c) IEEE 300-bus



(d) Polish Grid

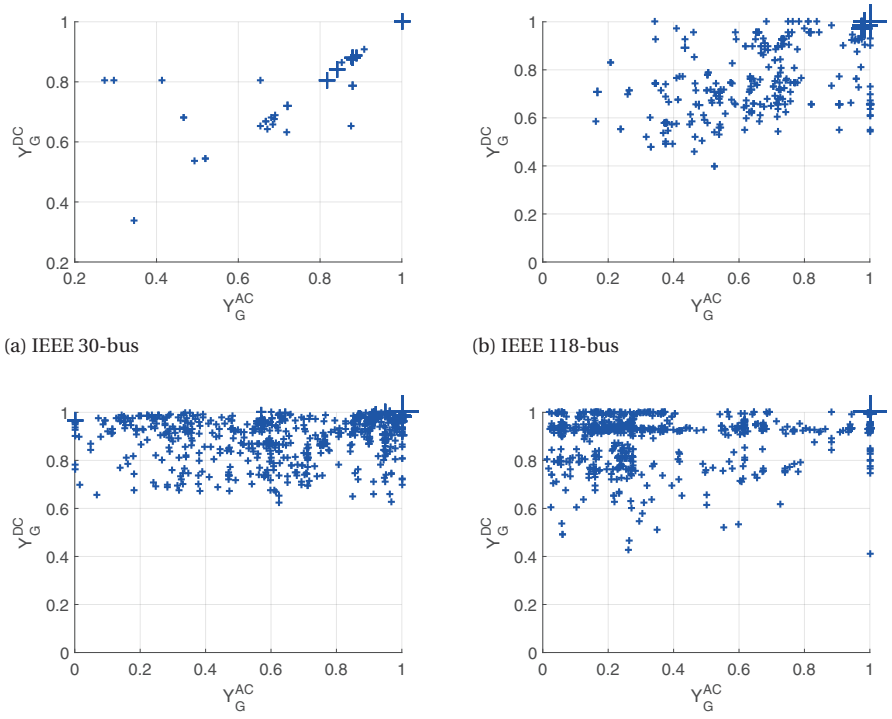
Figure 5.7: Comparison between the line-vulnerability ratios under the AC and DC cascade models initiated by single line failures. The lines with the highest line-vulnerability ratios under the AC cascade model are selected for comparison.



(a) $X := |Y_{\mathcal{G}}^{\mathrm{AC}} - Y_{\mathcal{G}}^{\mathrm{DC}}|$ (b) $X := |L_{\mathcal{G}}^{\mathrm{AC}} - L_{\mathcal{G}}^{\mathrm{DC}}|$

(c) $X := |N_{\mathcal{G}}^{\mathrm{AC}} - N_{\mathcal{G}}^{\mathrm{DC}}|$ (d) $X := |R_{l}^{\mathrm{AC}} - R_{l}^{\mathrm{DC}}|$

Figure 5.8: The CDFs of the differences between the metrics after cascading failures initiated by single line failures under the AC and DC flow models for all the test networks.



(a) IEEE 30-bus

(b) IEEE 118-bus

(c) IEEE 300-bus

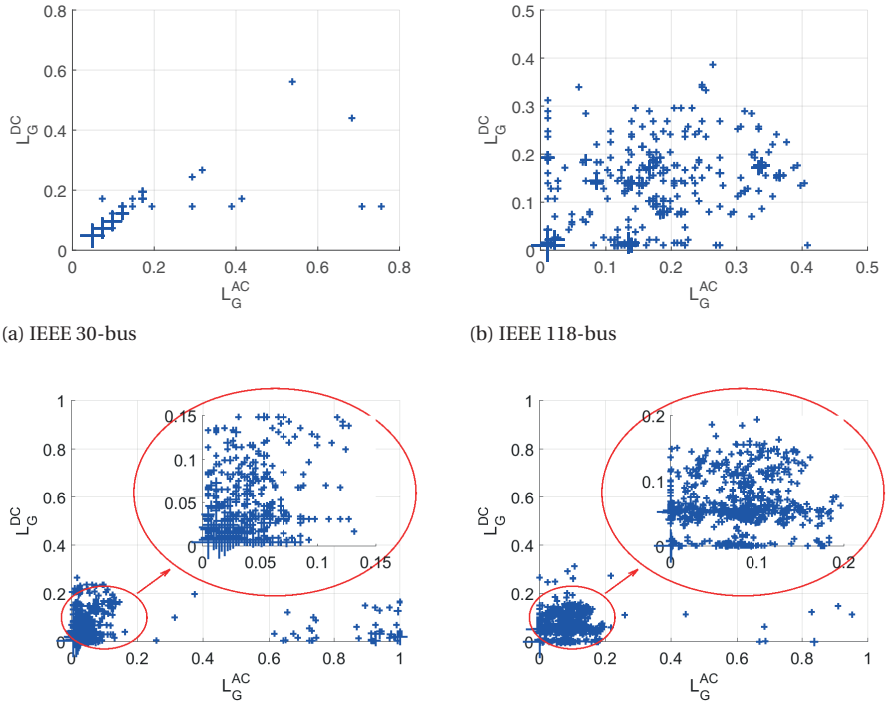
(d) Polish Grid

Figure 5.9: The scatter plots of the yield values under the AC versus DC cascade models initiated by two-line failures. Markers are scaled according to the frequencies of corresponding data points.

Figure 5.9 shows the scatter plot of the yield values under the AC and DC cascading failures models for the four test networks. Yield values obtained by the DC cascade model are usually higher, specially for large networks. Figure 5.12a presents the CDFs of the differences in yield values for all the test networks. Removal of two lines usually puts the system in a more critical condition with more cascade stages: The magnitudes of the differences in the obtained yield values are slightly higher for the cascades initiated by two line failures than by one line failure.

Figure 5.10 and Figure 5.12b show the line-loss ratios are still close under the two cascade models in all the four networks. The same is true for the node-loss ratios (see Figure 5.12c). However, similar to the yield, the differences in the line-loss and node-loss ratios are slightly higher for the cascades initiated by two line failures than by a single line failure.

Similar to the cascades initiated by single line failures, the lines that fail frequently under the AC model are also different here from their counterparts under the DC model (see Figures 5.11c and 5.11d) when the networks become larger. Figure 5.12d also suggests that the differences in the line-vulnerability ratios are also slightly higher here than in the cascades initiated by single line failures.



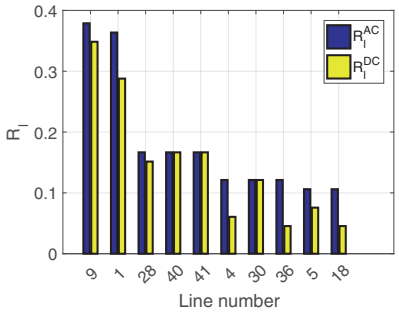
(a) IEEE 30-bus

(b) IEEE 118-bus

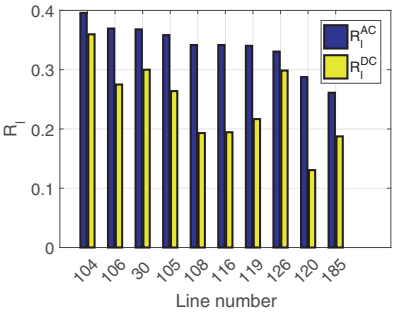
(c) IEEE 300-bus

(d) Polish Grid

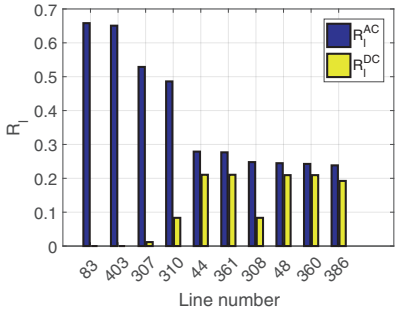
Figure 5.10: The scatter plots of the line-loss ratios under the AC versus DC cascade models initiated by two-line failures. Markers are scaled according to the frequencies of corresponding data points.



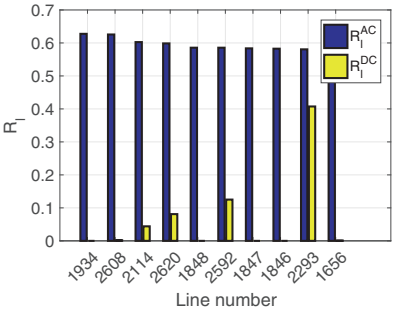
(a) IEEE 30-bus



(b) IEEE 118-bus

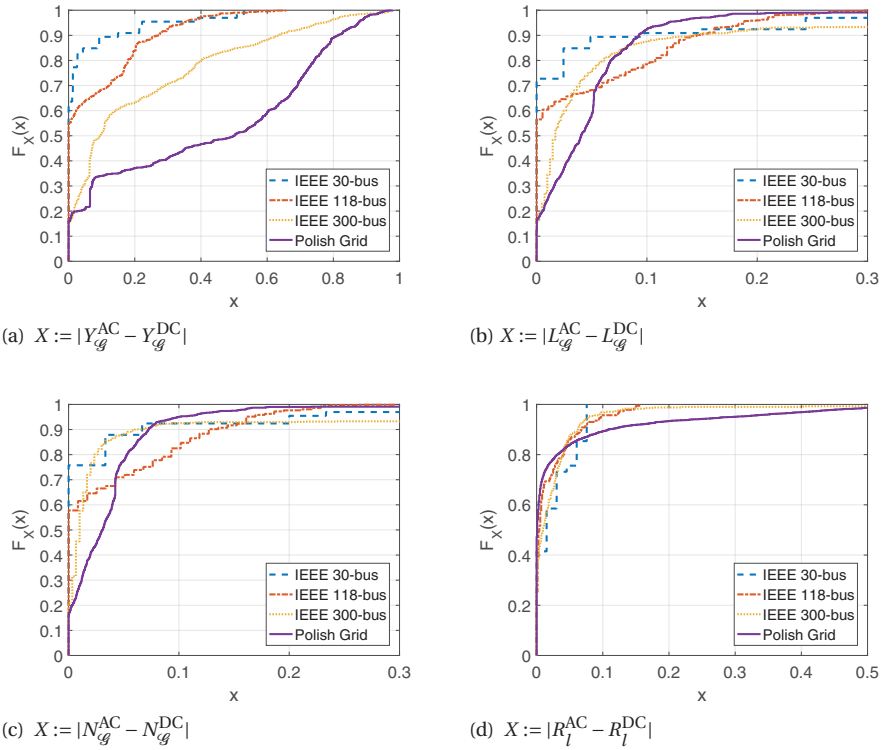


(c) IEEE 300-bus



(d) Polish Grid

Figure 5.11: Comparison between the line-vulnerability ratios under AC and DC cascade models initiated by two-line failures. The lines with the highest line-vulnerability ratios under the AC cascade model are selected for comparison.



(a) $X := |Y_{\mathcal{G}}^{\mathrm{AC}} - Y_{\mathcal{G}}^{\mathrm{DC}}|$

(b) $X := |L_{\mathcal{G}}^{\mathrm{AC}} - L_{\mathcal{G}}^{\mathrm{DC}}|$

(c) $X := |N_{\mathcal{G}}^{\mathrm{AC}} - N_{\mathcal{G}}^{\mathrm{DC}}|$

(d) $X := |R_{l}^{\mathrm{AC}} - R_{l}^{\mathrm{DC}}|$

Figure 5.12: The CDFs of the differences between the metrics after cascading failures initiated by two-line failures under the AC and DC flow models for all the test networks.

5.4.5. Comparison Between the Three Cascade Processes under the AC and DC Models

In this subsection, we compare the three cascade processes defined in Section 5.3 initiated by single line failures under the AC and DC models. For the cascade process III, we set the threshold ξ_l of a line l in (5.6) as $\xi_l = 0.8 \cdot c_l$.

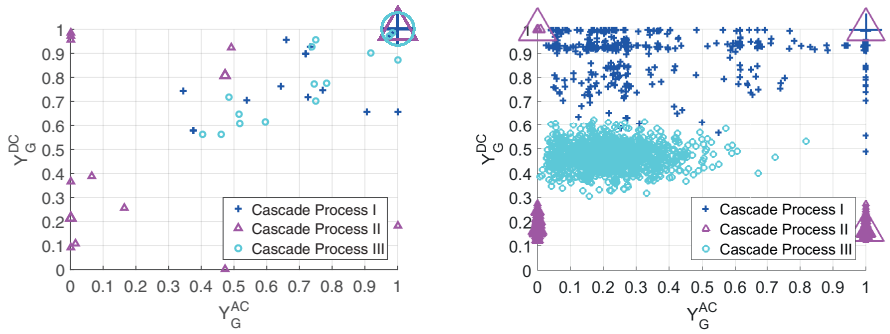
Figures 5.13-5.17 provide detailed comparisons between the results obtained under the AC and DC cascade models for the three cascade processes. Figure 5.13a and Figure 5.13b show the scatter plots of the yield values for cascades in the IEEE 118-bus network and Polish grid. They suggest that the yield values obtained by the cascade process II are generally lower than the other two cascade processes under the AC model. Figure 5.17a and Figure 5.17b, which present the CDFs of the differences in yield values under the AC and DC cascade models for the three cascade processes in the IEEE 118-bus network and Polish grid, also show that the differences in the obtained yield values under the AC and DC models can grow high for the cascade process II.

Figure 5.14a and Figure 5.14b show the scatter plots of the line-loss ratios under the AC and DC cascade models for the three cascade processes in the IEEE 118-bus network and Polish grid. Line-loss ratios obtained by the cascade process II are usually higher, leading to higher differences between the line-loss ratios obtained by the AC and DC flow models. Figure 5.17c and Figure 5.17d present the CDFs of the differences in line-loss ratios in the IEEE 118-bus network and Polish grid. Similar to Figure 5.17a and Figure 5.17b, the magnitudes of the differences in the obtained line-loss ratios under the AC and DC models are highest for the cascade process II.

Figure 5.15a and Figure 5.15b present the comparison between the highest line-vulnerability ratios under the AC and DC cascade models for the cascade process II in the IEEE 118-bus network and Polish grid. Figure 5.16a and Figure 5.16b present the comparison between the highest line-vulnerability ratios under the the AC and DC cascade models for the cascade process III in the IEEE 118-bus network and Polish grid. The difference between the individual line-vulnerability ratios in Figure 5.16b is particularly high for the cascade process III. Figure 5.17e and Figure 5.17f show that the differences in the line-vulnerability ratios may be larger for the cascade process III.

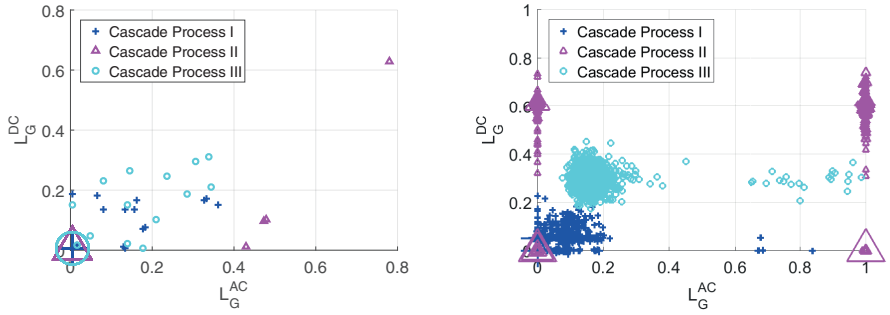
Figures 5.13-5.17 suggest that different rules for the supply and demand balancing and line outages could have different effect on the evaluation of the cascades under the AC and DC flow models. In particular, the cascade process II increases the differences between the AC and DC models the most. In this model, by disconnecting many small-sized generators distributed in the network, the demands are supplied by few large-sized generators during the cascade stages. Consequently, the remaining network suffers from low voltage magnitudes and overloaded lines, which can lead to divergence in iterations of AC power flow equations. Moreover, the reactive power flows and voltage magnitudes are not modeled by the DC flow model which can lead to higher differences between the cascades under AC and DC flow models.

Although the cascade process III does not affect the yield values and line-loss ratios very much, its effect is more significant in identifying the most vulnerable set of lines. Due to the probabilistic line tripping model in (5.6), different lines may trip at each cascade stage, which can result in detecting different sets of vulnerable lines under AC and DC flow models.



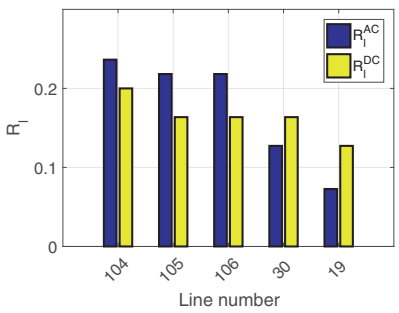
(a) IEEE 118-bus (b) Polish Grid

Figure 5.13: The scatter plots of the yield values under the AC vs DC cascade models for the three cascade processes initiated by single line failures. Markers are scaled according to the frequencies of corresponding data points.

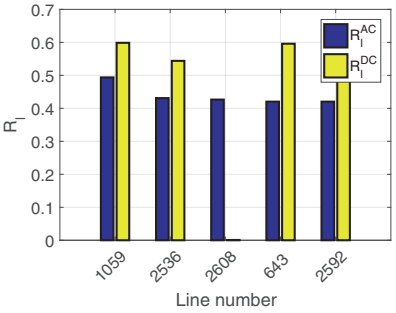


(a) IEEE 118-bus (b) Polish Grid

Figure 5.14: The scatter plots of the line-loss ratios under the AC vs DC cascade models for the three cascade processes initiated by single line failures. Markers are scaled according to the frequencies of corresponding data points.

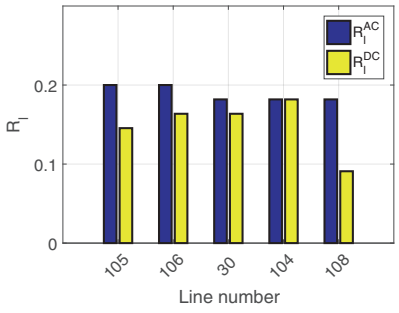


(a) IEEE 118-bus

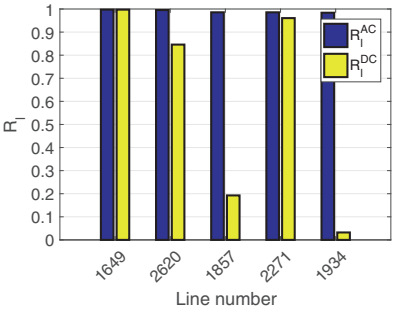


(b) Polish Grid

Figure 5.15: Comparison between the line-vulnerability ratios under the AC and DC cascade models for the cascade process II initiated by single line failures. The lines with the highest line-vulnerability ratios under the AC cascade model are selected for comparison.

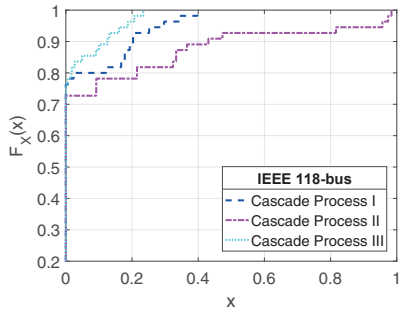


(a) IEEE 118-bus

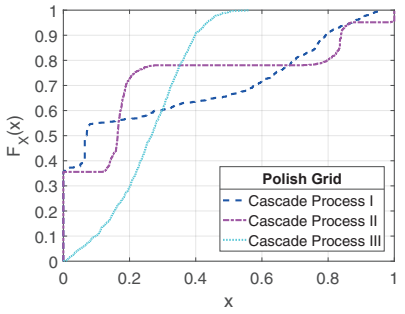


(b) Polish Grid

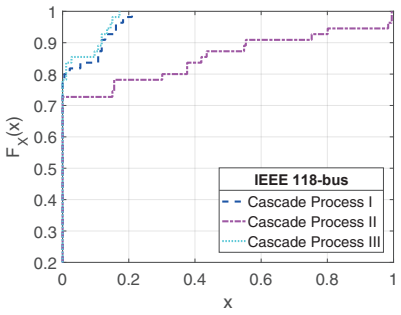
Figure 5.16: Comparison between the line-vulnerability ratios under the AC and DC cascade models for the cascade process III initiated by single line failures. The lines with the highest line-vulnerability ratios under the AC cascade model are selected for comparison.



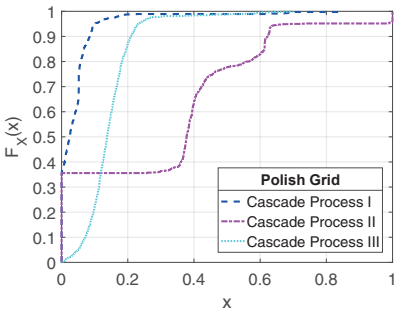
(a) $X := |Y_{\mathcal{G}}^{\mathrm{AC}} - Y_{\mathcal{G}}^{\mathrm{DC}}|$



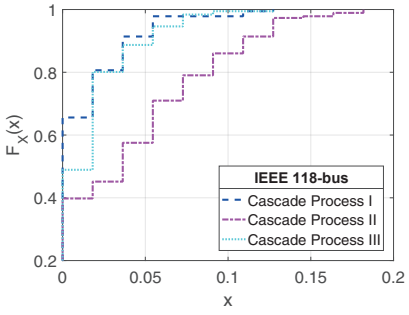
(b) $X := |Y_{\mathcal{G}}^{\mathrm{AC}} - Y_{\mathcal{G}}^{\mathrm{DC}}|$



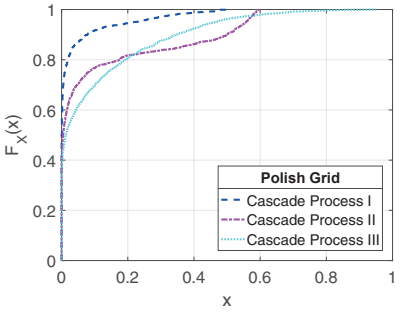
(c) $X := |L_{\mathcal{G}}^{\mathrm{AC}} - L_{\mathcal{G}}^{\mathrm{DC}}|$



(d) $X := |L_{\mathcal{G}}^{\mathrm{AC}} - L_{\mathcal{G}}^{\mathrm{DC}}|$



(e) $X := |R_l^{\mathrm{AC}} - R_l^{\mathrm{DC}}|$



(f) $X := |R_l^{\mathrm{AC}} - R_l^{\mathrm{DC}}|$

Figure 5.17: The CDFs of the differences between the metrics after cascades under the AC and DC models for the three cascade processes initiated by single line failures in IEEE 118-bus network and Polish grid.

5.4.6. MAIN LESSONS LEARNED FROM THE SIMULATIONS

In this section, we summarize the results obtained in the previous subsections. The main lessons learned from the analysis of the DC cascading failures model compared to the AC cascading failures model from the simulations are as follows:

1. When there are no failures and the assumptions underlying the DC power flow approximation are valid, the DC power flow model can approximate the AC power flow model in the network relatively well.

2. The DC power flow model can capture the effects of a single line failure on the flow changes on other lines (i.e., line flow change ratios and line outage distribution factors) relatively accurately. However, because of their limitations, they fail to capture other dynamics such as node voltage changes and power loss change ratios.

3. The AC and DC cascade models with the cascade process I provide similar line- and node-loss ratios (i.e., total number of line and node failures) most of the time.

4. The AC and DC cascade models with the cascade process I provide similar yield for small networks. However, for large networks (e.g., the Polish grid) the DC cascade model tends to overestimate the yield.

5. The AC and DC cascade models with the cascade process I agree on the most vulnerable lines under the line-vulnerability ratios in small networks, most of the time. However, for larger networks (i.e., the Polish grid) they tend to detect different sets of lines.

6. The DC cascade model with the cascade process II could underestimate the severity of the cascade compared to AC model with the same cascade process, as the effects of node voltage changes and reactive power flows are neglected under the DC flow model.

7. The AC and DC cascade models with the cascade process III provide similar yield, line-loss, and vulnerability ratios for small networks. However, for larger networks (e.g., the Polish grid) they result in different sets of most vulnerable lines.

Overall, the obtained results suggest that due to the voltage constraints, the divergence problems, and the reactive power flows, the cascades under the AC flow models are more significant compared to the ones under the DC flow model. Hence, the DC model may underestimate the severity of the cascade, especially for larger networks.

5.5. CONCLUSION

In this chapter, we thoroughly compared the AC and DC power flow models in describing the state of the grid when there are no failures as well as in predicting the effect of single line failures and the evolution of cascades. We numerically compared the AC and DC power flow models and numerically demonstrated in the tested networks that when there are no failures, the DC power flow model provides relatively accurate approximation of the AC power flow model. Upon failures, numerical results for the single line

failure analysis show that the DC power flow model provides a similar flow redistribution after single line failures as the AC flow model. On the other hand, the cascading failures simulation demonstrates that even slight errors in individual line flows can turn out to be important at cascade stages, and the metrics that capture the operational and topological aspects of the cascade can differ significantly under the two models. These results suggest that special care should be taken when drawing conclusions based on the DC cascade model in power grids. Overall, the DC cascade model can provide an overly optimistic estimation compared to the AC cascade model.

6

PLANNING FOR WIND POWER INTEGRATION INTO POWER GRIDS

In previous chapter, we show that the generation dispatch of the network is important as well as the underlying topology in the analyses of power grids. Many countries aim to integrate a substantial amount of wind power in the near future. This requires meticulous planning, as the integration of wind farms and their power outputs into the power grid affect the operation of network. In this chapter, we investigate the optimal placement of wind farms, thereby taking into account wind characteristics and power grid constraints. We model the long-term variability of wind speed using a Weibull distribution according to wind direction intervals, and formulate the metrics that capture wind characteristics at a specific location, namely the arithmetic mean of wind speed, the theoretical wind power density and the capacity factor of a prospective wind power plant. Subsequently, we construct a linear optimization to determine the geographical locations and the installed capacities of wind farms in order to maximize the expected annual wind power generation while obeying the constraints from the power grid and the transmission system operator.

This chapter is based on published papers [102] and [103].

6.1. INTRODUCTION

DRIVEN by the long-term goals to achieve a sustainable energy system, the utilization of renewable energy, especially wind power, is rising. Wind farm investments are usually large in size (installed capacity), and their output can be integrated into the power grid from the high-voltage level. The intermittent nature of wind power can challenge the stability and reliability of the power grid. The wind speed characteristics at a wind farm determine the power generation from wind turbines. Therefore, for prospective investors and for power system analysts who carry out the reliability analyses, modelling the variation of wind speed is essential [104].

Many researchers have conducted case studies to investigate the distribution of wind speed for various purposes [105]: In risk analyses concerning extreme or maximum wind speeds, extreme value distributions [106] are typically used [107], whereas in grid integration studies, the Weibull distribution is widely used because of its flexibility and satisfactory results in fit tests [108, 109]. This chapter investigates wind power integration with a long-term focus. Consequently, the Weibull distribution according to wind direction intervals [110] is utilised to model the annual wind speed distributions. We further formulate the metrics [111], average wind speed, wind power density and the capacity factor of a wind power plant, to assess the wind characteristics at a geographical location.

The main motivation for investing in a wind farm is the expected profit. On sites with strong wind, investors prefer to establish wind farms of substantial sizes. In order to avoid network congestion due to the power outputs of such wind farms, the grid operators are conservative about the integration of new power plants into the power grid and may impose limits on the maximum installed capacities in certain regions [112]. Various methodologies have been proposed to facilitate the determination of wind farm locations [113]. A number of these studies are based on the maximization of the profit of investors [114], which ignore the integration effects of wind farms into the power grid. Therefore, some studies propose methods for wind power integration according to the needs of the power grid (such as loss reduction, voltage regulation [115, 116]). Nevertheless, those theoretical integration plans may fail to be realised as the proposed wind farm locations do not necessarily attract investors. To solve this wind farm placement problem, this chapter presents a combined methodology for countrywide optimal wind power integration: Initially, the metrics that capture the quality of wind are utilised to assess the feasible locations for establishing wind farms from an investor's point of view. Subsequently, those feasible geographical sites are mapped on the power grid, and the optimal siting and sizing of wind farms to maximize annual wind power generation, while obeying the constraints from the power grid and the transmission system operator are determined. Therefore, the proposed placement of wind farms is of interest to investors, and the prospective integration of the power outputs of wind farms does not violate the transmission grid constraints.

The remainder of this chapter is organised as follows: Section 6.2 explains the models for wind speed and the power output of a wind turbine. The criteria for the evaluation of potential sites for wind farms are discussed in Section 6.3. Section 6.4 formulates the linear optimization problem to investigate the optimal integration of wind power plants, and the results of the optimization are presented in Section 6.5. Finally, Section 6.6 con-

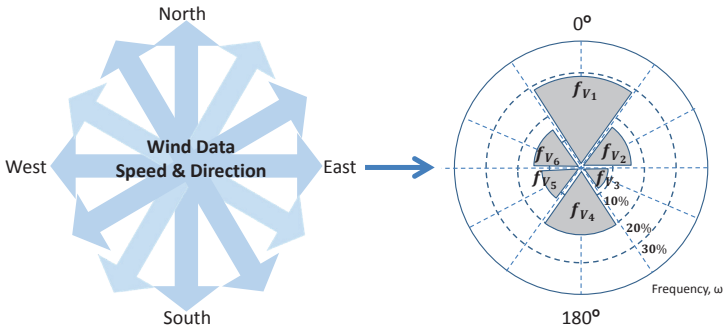


Figure 6.1: Model for annual variability of wind speed (N_d=6).

cludes the chapter.

6.2. PROBABILISTIC MODEL FOR THE POWER OUTPUT OF A WIND TURBINE

This section presents probabilistic models for wind speed and formulates the long-term variability of the power output of a wind turbine.

6.2.1. WIND SPEED CHARACTERISTICS

The annual variability of wind speed is used in assessing the integration of wind farms into the power grid [104], and typically, the Weibull distribution is used to represent the annual variation of wind speed [108, 109]. The probability density function (pdf) $f_V(v)$ and the cumulative distribution function (cdf) $F_V(v)$ of the Weibull distribution are defined as

$$f_V(v) = b\, a^{-b}\, v^{b-1}\, e^{-(\frac{v}{a})^b}$$

$$F_V(v) = 1 - e^{-(\frac{v}{a})^b}$$

where v denotes the Weibull random variable (wind speed), a is a scale parameter and b is a shape parameter [109].

Measurements show that wind speed characteristics depend on wind direction [110, 117, 118]. In this chapter, the dependence of wind speed on wind direction is incorporated into the probabilistic model of wind speed as follows: Annual wind measurement data (usually on an hourly basis) at a specific site are divided into N_d intervals according to wind direction. Subsequently, the wind speed values clustered for each interval are represented by a fitted Weibull distribution and a frequency value that captures how often wind blows from this direction interval as compared to all intervals [110]. Figure 6.1 illustrates a long-term wind speed model at an arbitrary site.

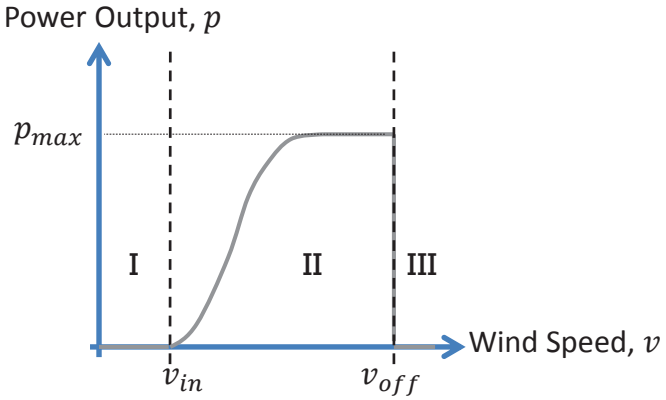


Figure 6.2: A typical p-v characteristic of a wind turbine.

As a result, the probability density function of wind speed at a site is defined as

$$f_V(v) = \sum_{i=1}^{N_d} f_{V_i}(v)\,\omega_i \tag{6.1}$$

where N_d is the total number of direction intervals, $f_{V_i}(v)$ is the Weibull probability density function of wind speed for the i^{th} interval, and ω_i is the frequency of the i^{th} interval.

6.2.2. THE POWER OUTPUT OF A WIND TURBINE

The power available in wind is converted to a useful form of energy by wind turbines. The power output of a wind turbine depends on wind speed and the characteristics of the wind turbine, such as efficiency, size and power curve. The power curve or the p-v characteristic of a wind turbine defines how the power output of the wind turbine varies with wind speed [119]. In Figure 6.2, a typical power curve of a wind turbine is illustrated.

The power curve of a wind turbine can be analysed in three regions: In order for the wind turbine to start generating power, wind speed must be greater than the cut-in speed v_{in}. Consequently, below the cut-in speed, in region I, the power output of a wind turbine is zero. Similarly, in region III, the wind turbine stops operating to prevent damage at higher speeds than the cut-off speed v_{off} and does not generate power. Therefore, the wind turbine generates power when wind speed is between the cut-in and the cut-off speeds, in region II. In this region, the power output of a wind turbine increases with increasing wind speed till the rated speed at which the maximum power output of the wind turbine is reached and generated till the cut-off speed. Hence, the power output p of a wind turbine can be expressed as

$$p = \begin{cases} 0 & \text{if } v < v_{\text{in}} \text{ or } v > v_{\text{off}}, \\ PC(v) & \text{if } v_{\text{in}} \le v \le v_{\text{off}} \end{cases} \tag{6.2}$$

where $PC(.)$ represents the power curve of the wind turbine.

From (6.2), the probability of zero power output can be calculated as the sum of the probabilities that wind speed is smaller than the cut-in speed or larger than the cut-off speed

$$\begin{aligned} F_P(0) =& \Pr(v < v_{\text{in}}) + \Pr(v > v_{\text{off}}) \\ =& F_V(v_{\text{in}} - \varepsilon) + 1 - F_V(v_{\text{off}} + \varepsilon) \end{aligned}$$

where $F_P(p)$ is the cdf of the power output of the wind turbine, and ε is a small positive number.

The power curve of a wind turbine is a non-decreasing function in regions I and II. Therefore, from the change of variables technique in probability theory [120], the cdf of the power output can be expressed as the sum of the cdf of wind speed and the probability that power output is zero due to speeds higher than the cut-off speed:

$$F_P(p) = F_V(v) + 1 - F_V(v_{\text{off}} + \varepsilon) \tag{6.3}$$

where $p = PC(v)$, i.e., the power output of the wind turbine at wind speed v.

From the cdf of the power output, the expected annual power generation from the wind turbine G_W can be found by integration:

$$G_W = \int_0^{p_{\max}} \big(1 - F_P(p)\big)\, dp \cdot 8760 \text{ [MWh]} \tag{6.4}$$

where $p_{\max}$ is the size of the wind turbine in MW and 8760 represents the hours in a year.

6.3. THE ASSESSMENT OF POTENTIAL SITES FOR WIND FARMS FROM AN INVESTOR'S POINT OF VIEW

The location of a wind farm influences the power generated from wind turbines and the impact on the power grid. Therefore, the assessment of potential sites for wind farms is crucial to the wind power integration analyses. Based on the models for wind speed and the power output of a wind turbine presented in Section 6.2, this section develops the criteria for evaluating a potential site for wind farm construction from an investor's point of view.

6.3.1. INDICATORS OF SUITABLE LOCATIONS FOR WIND FARMS

The decision to invest in a wind farm at a certain location depends on two main criteria: wind power potential and investment disincentives. The wind power generative potential of an area can be captured by indicators such as average wind speed, wind power density, and the capacity factor of a prospective wind turbine. On the other hand, disincentive indicators can include economic and environmental criteria such as high values of land cost or altitude levels, and the proximity to urban areas.

THE ARITHMETIC MEAN OF WIND SPEED

The arithmetic mean of wind speed $\bar{v}$ at a site is calculated using the expected value theorem [120] in the model for wind speed (6.1) as

$$\begin{aligned}\bar{v} &= \sum_{i=1}^{N_d} \int_0^{\infty} b_i \, a_i^{-b_i} \, v^{b_i} \, e^{-(\frac{v}{a_i})^{b_i}} \, dv \, \omega_i \\ &= \sum_{i=1}^{N_d} a_i \, \Gamma(1+\frac{1}{b_i}) \, \omega_i \end{aligned} \tag{6.5}$$

where $\Gamma(.)$ is the Gamma function, and a_i and b_i are the Weibull parameters of $f_{V_i}(v)$.

THE THEORETICAL WIND POWER DENSITY

The theoretical power available in wind p_w at an instant of time [111] is calculated as

$$p_w = \frac{1}{2} \, \rho_a \, v^3 A_{\perp} \tag{6.6}$$

where ρ_a is the air density and $A_{\perp}$ is the area perpendicular to wind, i.e., the blade sweep area of a wind turbine. As the calculation in (6.6) depends on the size of the wind turbine due to the cross-sectional area $A_{\perp}$, wind power density ρ is defined [111] so that wind power potential can be captured regardless of the turbine size:

$$\rho = \frac{1}{2} \, \rho_a \, v^3. \tag{6.7}$$

The expected theoretical wind power density $\bar{\rho}$ at a site can be calculated by introducing the proposed model for wind speed in (6.1) and the expected value theorem into (6.7) as

$$\begin{aligned}\bar{\rho} &= \sum_{i=1}^{N_d} \frac{1}{2} \, \rho_a \int_0^{\infty} b_i \, a_i^{-b_i} \, v^{b_i+2} \, e^{-(\frac{v}{a_i})^{b_i}} \, dv \, \omega_i \\ &= \frac{1}{2} \, \rho_a \sum_{i=1}^{N_d} {a_i}^3 \, \Gamma(1+\frac{3}{b_i}) \, \omega_i. \end{aligned} \tag{6.8}$$

THE CAPACITY FACTOR OF A WIND TURBINE

The capacity factor [111] of a wind turbine is defined as the ratio of actual power generation over a period of time, to the potential power generation if it were possible to operate at full capacity indefinitely:

$$\text{Capacity Factor} = \frac{\text{Total Generation}}{\text{Turbine Size} \cdot \text{Operating Hours}}. \tag{6.9}$$

For the proposed model, the annual capacity factor η of a wind turbine at a site can be calculated using the expected annual power generation (6.4) in the definition of capacity factor (6.9) as

$$\eta = \frac{\int_0^{p_{\max}} \big(1 - F_P(p)\big) dp}{P_{\max}}. \tag{6.10}$$

ECONOMIC AND ENVIRONMENTAL CRITERIA

Strong wind characteristics are good indicators of potential sites for wind farms. However, it is not possible to construct a wind farm at every promising site. Additional criteria for site selection may be imposed due to economic or environmental concerns. Economic criteria could comprise the lack or difficulty of transportation to the site, the land cost or the distance to the power grid, whereas environmental factors could be the site being close to city centres, airports, or forested areas, having high altitude and so on. Indeed, in some countries, for wind power plants to have a generation license, investors have to submit an assessment showing that the wind farm to be established does not result in any harm to the nature and environment. Therefore, these criteria should be included in the studies related to the determination of wind farm locations and the integration of wind farms into the power grid.

6.3.2. QUANTIFYING THE CRITERIA FOR WIND FARMS

In previous section, we show that two main criteria are important while deciding on suitable locations for wind farms: wind power potential and investment disincentives due to economic and environmental criteria. Ideally, an investor should review all M related indicators $\{r_1(k),\dots,r_M(k)\}$ before investing in a wind farm at a site k. In this section, we model those indicators of a wind farm investment using fuzzy sets [121], which enables us to quantify the satisfaction degree of each indicator during the decision process of a wind farm at a site k.

We use increasing fuzzy function $\overline{F}\big(r_i(k)\big)$ in (6.11) and decreasing fuzzy function $\underline{F}\big(r_i(k)\big)$ in (6.12) to evaluate the satisfaction degree of each indicator $r_i(k)$ for a wind farm in site k. The increasing fuzzy function represents the incentive indicators, whereas the decreasing fuzzy function represents the disincentive indicators. The resulting fuzzy membership degrees take values between 0 and 1 corresponding to the unsatisfactory and full-satisfactory evaluations of a site k, respectively.

$$\overline{F}\big(r_i(k)\big) = \begin{cases} 0 & \text{if } r_i(k) < q_i, \\ \frac{r_i(k)-q_i}{p_i-q_i} & \text{if } q_i \le r_i(k) \le p_i, \\ 1 & \text{if } r_i(k) > p_i, \end{cases} \tag{6.11}$$

$$\underline{F}\big(r_i(k)\big) = \begin{cases} 1 & \text{if } r_i(k) < p_i, \\ \frac{r_i(k)-q_i}{p_i-q_i} & \text{if } p_i \le r_i(k) \le q_i, \\ 0 & \text{if } r_i(k) > q_i, \end{cases} \tag{6.12}$$

where for each indicator r_i, q_i and p_i correspond to the thresholds of unsatisfactory and full-satisfactory evaluations, respectively.

6.3.3. MULTIPLE-CRITERIA DECISION ANALYSIS OF WIND FARMS

Since we have to deal with and optimize for multiple fuzzy parameters, we focus on multiple-criteria decision analysis in this section.

The perspective of an investor is important when assessing the criteria for a wind farm at a site k. For instance, an investor could consider a worst-case scenario of the related indicators or could, as the other extreme, consider a best-case scenario. Following

[122], [123], we employ fuzzy logic aggregation operators to allow for variability in perspective. We use the *and* $\wedge$ and the *or* $\vee$ aggregation operators to map two extreme cases of an investor's stance on multiple-criteria decisions. The *and* operator $\wedge$ of the fuzzy membership degrees requires the satisfaction of all desired criteria, in other words, a conservative perspective when evaluating the satisfaction degrees of related indicators:

$$\min_{1\le i\le M} F\big(r_i(k)\big). \tag{6.13}$$

The *or* operator $\vee$ is appropriate to model a more optimistic or lenient perspective. The implementation of the *or* operator in (6.14) passes over the less satisfactory indicators of site k:

$$\max_{1\le i\le M} F\big(r_i(k)\big). \tag{6.14}$$

Lastly, to model the perspective of an investor in between those two extreme cases, we can use a *weighted mean* operator μ in (6.15):

$$\sum_{i=1}^{M} w_i F\big(r_i(k)\big) \tag{6.15}$$

where the ultimate decision is the convex combination of the satisfaction degrees of the decision indicators, such that $\sum_i w_i = 1$.

By applying these aggregation operators to each site k, we can obtain an *suitability value* $\in [0,1]$ of that site for building wind farms. The higher suitability values of a site indicate a potential location of wind farms from an investors point of view. In the next section, an optimization problem is formulated to find the optimal siting and sizing of wind farms among those potential sites which maximizes the expected total annual power generation while adhering to the rules by transmission system operator.

6.4. PLACING WIND FARMS UNDER GRID CONSTRAINTS

The integration of wind farms and their power outputs into the power grid affects its operation. Therefore, grid operators demand that newly integrated wind power plants into the power grid do not violate the transmission system constraints. Additionally, grid operators could declare regional upper bounds on wind farms due to economic and geographical concerns. On the other hand, the expected profit from a wind farm is the main concern for power plant investors, and the revenue from a wind farm is closely related to the expected total generation. Therefore, on sites with strong wind, investors prefer to establish wind farms of substantial sizes whose power output integration into the power grid can cause problems. In Section 6.3, the criteria governing the potential sites for wind power plants from an investor's point of view are determined. In this section, an optimization problem is formulated to find the best solution to the wind farm placement problem: The expected total annual power generation from the selected sites is maximized, while adhering to the rules established by transmission system operator. Consequently, the objective function $f(S_k)$ to be maximized is defined as

$$f(S_k) = \sum_{k=1}^{N} \int_0^{S_k} \big(1 - F_{P_k}(p)\big)\, dp \cdot 8760 \tag{6.16}$$

where $f(S_k)$ represents the total annual generation from the wind farms, N is the total number of potential sites for wind farms satisfying the criteria according to Section 6.3, $F_{P_k}(p)$ is the cdf of the wind farm power output at the k^{th} site and S_k represents the size of the wind farm at the k^{th} site.

When wind power outputs are integrated into the power grid, power flow in the network changes, and the final state of the power grid must satisfy the power flow laws. We assume that a prospective wind farm at the k^{th} site will be electrically connected to the power grid at a pre-defined network node (substation) i, and we use the linearised DC power flow [34] to represent the power flow behaviour in the grid. Consequently, the maximization of (6.16) is subjected to the linearised power flow equations (6.17) for each network link l, and the power balance equations (6.18) for each network node i:

$$P_l - \frac{\Theta_s - \Theta_r}{X_l} = 0 \tag{6.17}$$

$$P_{C_i} + P_{W_i} + P_{l_i} - L_i = 0 \tag{6.18}$$

where P_l is the active power flow over the network link l, Θ_s is the voltage phase angle at the sending node of link l, Θ_r is the voltage phase angle at the receiving node of link l, X_l is the reactance of link l, P_{C_i} is the total power output of the existing generators at node i, P_{W_i} is the total wind power output at node i, i.e., the sum of the power outputs of the wind farms that are electrically connected to node i, P_{l_i} is the net link flow received at node i, and L_i is the electrical load consumed at node i.

The power output of a wind turbine is shown to be a random variable in Section 6.2. Therefore, total wind power output P_{W_i} integrated into the power grid from node i at an instant is not known in advance. We deploy a diversity factor[1] f_D to define the relation between the power output at an instant and the total installed capacity of wind farms connected to a node i

$$f_{D_i} = \frac{P_{W_i}}{\sum_{k\to i}^{N} S_k} \tag{6.19}$$

where $\sum_{k\to i}^{N} S_k$ represents the sum of the installed capacities of wind farms electrically connected to node i, and f_{D_i} is the diversity factor of node i. The diversity factor needs to be input to the optimization problem and reflects the level of conservatism wanted by the transmission system operator. The maximum diversity factor of a node is 1, demonstrating the instant at which all wind farms connected to node i are producing their maximum power output. Consequently, assuming a diversity factor of 1 in the optimization formulation is the most conservative integration planning, as it is expected to be most confined by the power grid constraints.

As the power outputs of the wind farms are not known in advance, the balance between the generated power and the consumed electrical load in (6.18) is achieved by giving a margin of safety to the selected existing reserve power plants, whereas for the other existing plants the generation output is fixed

$$P_{C_{i\min}} \le P_{C_i} \le P_{C_{i\max}} \tag{6.20}$$

[1] Demand diversity factor in power system analyses is defined as the ratio of the peak demand of the loads as a whole to the sum of the non-coincident peak demands of all individual units [124].

where $P_{C_{i\min}}$ represents the minimum generation from existing generators at node i, and $P_{C_{i\max}}$ represents the maximum generation from existing generators at node i.

Finally, the objective function in (6.16) is subjected to the constraints which are imposed by the transmission system operator: The final flows over each network link l must be smaller than the maximum flow limit, and the total size of wind farms at each site k and in each region r should not exceed the limits set by the transmission system operator

$$|P_l| \leq P_{l,\max} \tag{6.21}$$

$$S_k \leq S_{k,\max} \tag{6.22}$$

$$\sum_{k \in r}^{N} S_k \leq U_{r,\max} \tag{6.23}$$

where $P_{l,\max}$ is the maximum flow limit of the link l, $S_{k,\max}$ is the maximum total size of the wind farm at the k^{th} site, $\sum_{k \in r}^{N} S_k$ represents the sum of the installed capacities of wind farms in region r and $U_{r,\max}$ is the regional upper bound for the wind farms in region r.

The resulting linear optimization, i.e., the maximization of the objective function (6.16), subject to the linear equality constraints (6.17), (6.18), and the linear inequality constraints (6.20), (6.21), (6.22), (6.23) for different values of diversity factors in (6.19), can be solved by the linear programming.

6.5. CASE STUDY: WIND FARM PLANNING IN TURKEY

This section presents a case study to demonstrate how the proposed methodology can be applied to assess the potential sites for wind farms in Turkey and to plan for the wind power integration into the power grid.

6.5.1. WIND CHARACTERISTICS

The countrywide wind data of Turkey that include historical annual hourly wind speed and direction values for every 6×6 km^2 geographical area have been utilised to define the probability density function of wind speed (6.1) for each site[2].

For the power curve model in (6.2), the power curve of a practical wind turbine[3] is used and the cdf of the power output of a prospective wind turbine at each geographical area in Turkey is calculated according to (6.3). As an example, Figure 6.3 illustrates the pdf of the wind speed according to wind direction intervals at a specific area, and Figure 6.4 shows the cdf of the power output of a wind turbine at the specific area for the wind speed characteristics in Figure 6.3: The unpromising wind speed characteristics in the 1$^{\text{st}}$, 2$^{\text{nd}}$, 5$^{\text{th}}$, 6$^{\text{th}}$ and 9$^{\text{th}}$ direction intervals in Figure 6.3 result in higher probabilities of zero power output of the wind turbine.

[2] The data were retrieved from TÜBITAK Marmara Research Center. The total number of geographical areas is 21,983. The total number of direction intervals N_d in (6.1) is selected as 12, since the minimum error in the Weibull fitting is obtained for $N_d = 12$ for randomly chosen test areas, which is in-line with [110].

[3] The power curve of the Vestas V112-3.0 MW wind turbine was retrieved from: http://www.vestas.com.

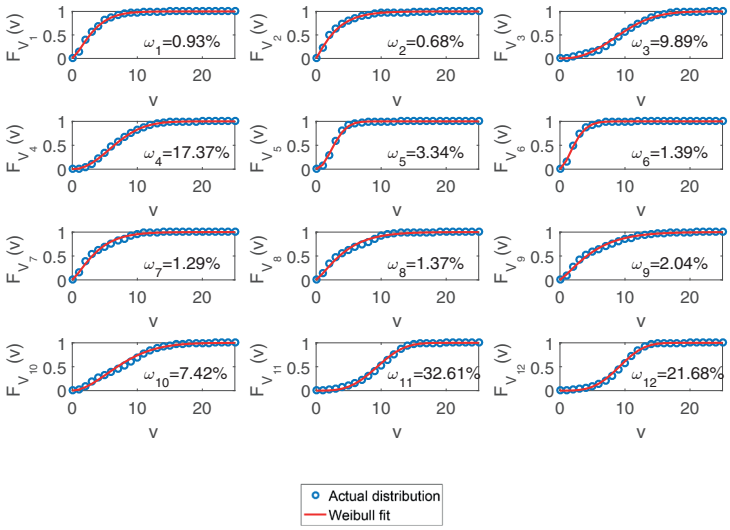


Figure 6.3: The wind speed variation according to wind direction intervals at a specific location.

6

6.5.2. POTENTIAL SITES FOR WIND FARMS

Following from Section 6.3, the arithmetic mean of wind speed (6.5), the theoretical wind power density (6.8), and the capacity factor of a prospective wind turbine (6.10) are calculated for each geographical area. Due to the positive correlation between the promising wind energy potential and the investment criteria for wind farms, the increasing fuzzy function in (6.11) is used to calculate corresponding satisfaction degrees of those indicators. The landscape of Turkey contains heterogeneously distributed mountainous regions with varying altitudes. High altitude regions and high slope lands are undesirable for establishing wind farms. Thus, we use the altitude of a site k as a disincentive indicator for wind farms. Due to the negative correlation between the altitude and the investment criteria for wind farms, the decreasing fuzzy function in (6.12) is used. The resulting membership functions of selected indicators are shown in Figure 6.5. The full-satisfactory and unsatisfactory thresholds of indicators are determined based on related works [125] and the wind power characteristic of Turkey [126].

The potential sites of wind farms are evaluated based on their suitability values. We used *and* operator (6.13) to aggregate the indicators of wind farms and set the minimum suitability value of a potential wind farm site to 0.5. In other words, the sites that have higher values of 6 m/s for the arithmetic mean of wind speed, 200 W/m^2 for the theoretical wind power density, and 15% for the capacity factor of a prospective wind turbine and lower values of 2000 m for the altitude are selected as potential sites. Under these constraints, more than three quarters of the geographical areas in Turkey are not feasible for the establishment of wind farms. We also eliminated ineligible geographical areas which can contain urban areas, natural parks, airports, etc.[4] Finally, we show the

[4]The related lists of national parks, natural monuments, protected areas in Turkey were retrieved from the

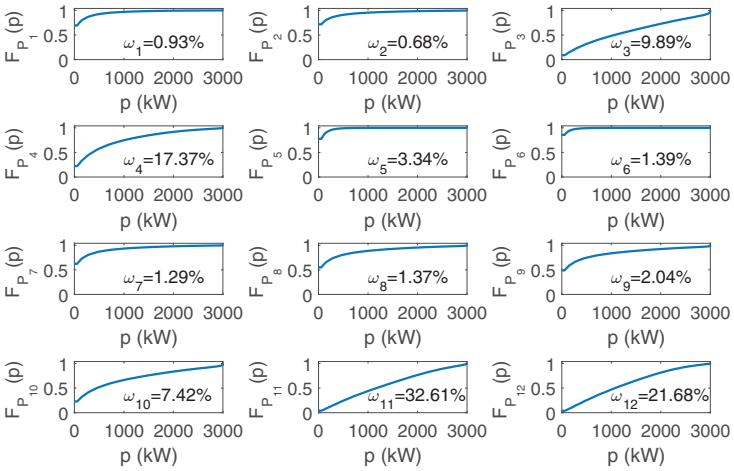


Figure 6.4: The cdf of power output of a wind turbine at the specific location in Figure 6.3.

6

potential sites for wind farms Figure 6.6. The white coloured areas in Figure 6.6 show the geographical areas in Turkey that satisfy the proposed criteria in Section 6.3 and each of them is treated as a potential site for the establishment of wind farms in the optimization formulation.

6.5.3. OPTIMAL SITES FOR WIND FARMS

The optimal integration of wind farms into the Turkish electricity grid is investigated according to the proposed methodology in Section 6.4 for the declared 3 GW of the wind power integration. The weighted graph representation of the high-voltage (400 kV and 154 kV) transmission grid is used. The substations are modelled as nodes, whereas the transmission lines and the transformers are modelled as links. The constructed model for the power grid has in total 1499 nodes and 2479 links. Year 2017 forecasts for the supply and demand are used in the model for the power grid in (6.17) and (6.18) at the instant of peak load[5] conditions. Flexible generation is allowed for the largest 2 hydropower plants of Turkey. The thermal ratings of transmission elements are used as the flow limits in (6.21). The geographic coordinates of substations are utilised to couple each feasible site k with the closest node i in the power grid, and with the region r. The maximum size of a wind farm in (6.22) is restricted to 30 MW and maximum wind farm connection capacities announced by the Turkish electricity transmission operator (TEİAŞ) are used as the regional constraints in (6.23). Lastly, the optimization problem in Section 6.4 is solved by the linear programming solver (*linprog*) in MATLAB for the maxi-

websites of Republic of Turkey, Energy Market Regulatory Authority, and Republic of Turkey, Ministry of Energy and Natural Resources:
http://www.epdk.gov.tr, and http://www.enerji.gov.tr, respectively.

[5] Peak load describes the period in which the power requirement of a power system is expected to be maximum.

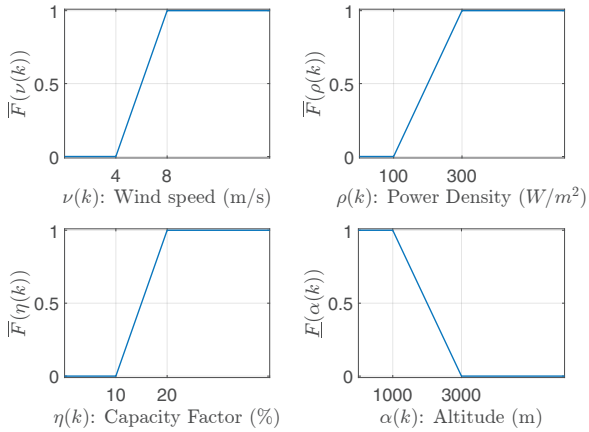


Figure 6.5: The membership functions of the selected indicators of wind farms.

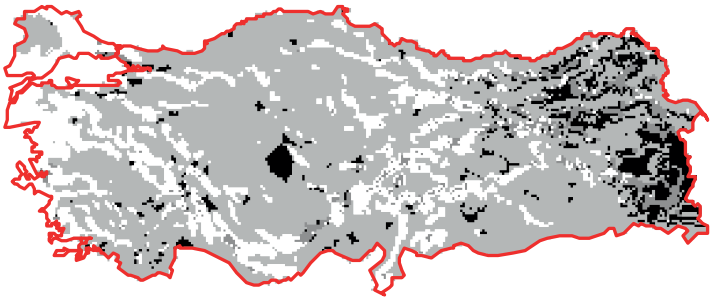


Figure 6.6: The determination of potential sites for wind farms in Turkey. White areas satisfy the criteria in Section 6.3, whereas light grey, dark grey and black areas are ineligible for the establishment of wind power plants due to the quality of wind criterion, economic and environmental criteria, and both of the criteria, respectively.

mization of the objective function (6.16), subject to the linear equality constraints (6.17), (6.18), and the linear inequality constraints (6.20), (6.21), (6.22), (6.23) for different values of diversity factors in (6.19).

Figure 6.7 illustrates the optimal siting and sizing of wind farms for the diversity factor $f_D = 1$. The selected placement maximizes the annual wind power generation from the wind farms, while complying with the power grid constraints. The thermal ratings of 16 transmission elements are found to be binding[6] in the optimal solution, and the whereabouts of those transmission elements indicate the bottlenecks of the power grid. The expected total annual wind power generation is calculated as 8.2796 TWh with the average capacity factor of 31.9%, which support the feasibility of the proposed place-

[6]A constraint is binding if changes in its value change the optimal solution. In other words, at the optimal solution, a binding inequality constraint is satisfied at its limit. Less severe constraints that do not affect the optimal solution are called non-binding.

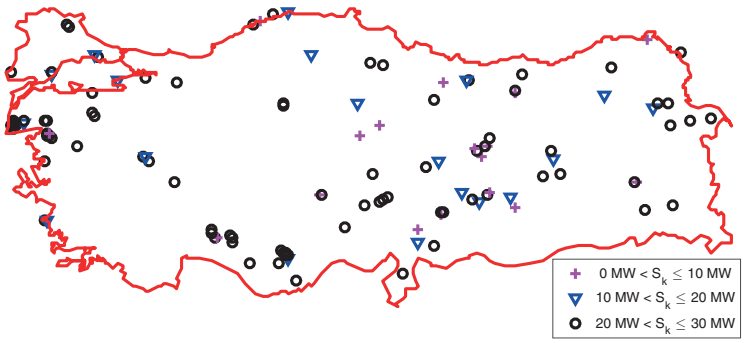


Figure 6.7: The optimal wind farm placement when $f_D = 1$.

Table 6.1: The value of the objective function in different power output scenarios.

f_D	$f(S_k)$ [TWh]	f_D	$f(S_k)$ [TWh]
0.1	8.4136	0.6	8.4136
0.2	8.4136	0.7	8.4107
0.3	8.4136	0.8	8.3953
0.4	8.4136	0.9	8.3503
0.5	8.4136	1.0	8.2796

ment plan from an investor's point of view.

Finally, the impact of the diversity factor on the optimal locations is analysed. Table 6.1 presents the results of the objective function for 10 different values of diversity factor (6.19) in the optimization problem. The same optimal sizes and locations, which produce 8.4136 TWh of annual wind generation, are selected in 6 out of 10 cases. The power grid inequality constraints (6.21) are observed to be non-binding for that solution, which means that the most contributing sites to the objective function (6.16) in each region are selected for the integration, and their simultaneous power output up to 60% of the total installed capacity does not violate the transmission grid constraints. However, in the remaining cases with diversity factor equal to or larger than 0.7, the power grid inequality constraints also become binding. Consequently, new sets of sizes and locations with less-favourable wind characteristics that comply with the power grid constraints are selected, which decreases the annual wind power generation. The diversity factor could be determined by transmission system operators depending on the planned degree of flexibility in the operation of transmission system. If minimum risk of overloads is desired by the transmission system operators, the solution of the optimization problem for maximum diversity factor, $f_D = 1$, could be used for the long-term integration plan of wind farms.

6.6. CONCLUSION

This chapter presented a comprehensive methodology to investigate the wind power integration into the power grid. We calculated the probabilistic functions for the long-term variability of wind speed and the power output of a wind turbine. We presented the expressions for three metrics, arithmetic mean of wind speed, theoretical wind power density, and the capacity factor, which capture the quality of wind at a specific area and used them while determining the feasible locations for wind farm establishment. We further utilised a map-based approach to couple each wind farm site with the nodes in the power grid. Subsequently, we formulated an optimization problem to find the optimal sizing and siting of wind farms, such that the selected sites can generate maximum expected annual wind power generation while satisfying the regional and the power grid constraints. The constructed optimization formulation facilitates the assessment of the placement of wind power plants, and could be used by transmission system operators as a long-term transmission network planning tool for the grid integration of wind power plants.

7

Conclusion

Network science provides a complementary approach to the traditional flow-based methods for investigating the subtle behaviour of power grids. Motivated by the increasing need of reliable power grids and the merits of network science on the analyses of power grids, this thesis develops concepts and tools for modelling, analysing and planning of power grids relying on network science. The developed concepts for investigating the power grids and its current and/or near-future challenges can be used to exploit the relationship between the topology, the operation and the performance of power grids, and help to address the near-future challenges of power grids by simulating their impacts and by assisting in network planning processes.

7.1. MAIN CONTRIBUTIONS

THE main contribution of this thesis is the additional insight, concepts and measures for a better understanding of (i) the network science approach for power grids, and (ii) how to utilize it to solve the current and/or the near-future challenges of power grids in terms of line removals/additions, malicious attacks, network expansion, cascading failures, and renewable integration. The operation of a power grid, thus also its reliability, is mainly governed by the operative state and the topology. In this context, we diverge from the purely topological approaches, and take an extended-topological approach in the analysis and planning of power grids: We model the underlying topology of power grids as simple or weighted graphs (according to the characteristics of the transmission lines), and we take the electric power as the physical quantity that flows through the network according to the power flow equations. Such an approach makes it possible both to assess the interdependencies between the components of power grids, and to investigate the emergent behaviours of a power grid (such as cascading failures) while obeying the main assumptions of power grids.

This thesis addresses the challenges of power grids with an interdisciplinary approach and extends the state of the art in the applications of network science on power grids. The main contributions of each chapter are as follows:

Chapter 2 expresses the linearised DC power flow equations in power grids in terms of graph-related matrices. This result opens the door to the direction in analysing power grids by an extended-topological approach, which includes the fundamentals of a power grid, such as the flow allocation according to Kirchhoff's laws and the effect of transmission line reactances. The chapter further derives the expressions for the sensitivities of link flows to link removals and additions. These results can be used to support decision makers in deciding which lines or flow capacities to upgrade, how to prepare for contingencies and how to choose the location of a new line.

Chapter 3 presents two different graph models for power grids as simple and weighted graphs. Based on the weighted graph model for power grids, this chapter extends the traditional centrality metrics, i.e., degree, closeness, betweenness and eigenvector centrality, by including the flow behaviour in power grids. Via case studies, we show that targeted node attacks based on those traditional and extended centralities are extremely effective to destruct the power grids. Thus, the methodology and results of this chapter can assist the grid operators with analysing the current vulnerability of their network to targeted attacks and with taking necessary measures by protecting the important nodes in their networks.

Chapter 4 investigates the power transmission in various graphs and aims to understand the impact of power transmission on the magnitude of link flows, node voltages and active power losses in power grids. Many other criteria related to the performance of power grids, such as the effect of flow capacities of links, the operation under single line failure contingencies and the ease of the control over the network, as well as the underlying topology, are discussed in this chapter, which could assist in planning for a topological transformation and network expansion in power grids.

Chapter 5 develops models to predict how cascading failures in power grids evolve using both the nonlinear AC and linearised DC power flow equations. This section contributes the related field of cascading failures that considers a topological perspective

where, once a network element fails, the neighbouring elements also fail. Moreover, the models presented in this chapter could help to identify the vulnerable parts of power grid not only under $N-1$ contingencies but also the consequent $N-k$ contingencies.

Chapter 6 presents a methodology to plan for the wind power integration into power grids. The two-fold approach in this chapter, i.e., first identifying the potential wind farm locations from an investors point of view and then selecting the locations under the constraints from power grids, provides a more realistic perspective on the wind farm placement problem. The metrics that capture the promising wind power potential and the constructed optimization formulation facilitate the planning for renewable energy integration and could be used by transmission system operators as a medium- or long-term planning tool.

7.2. DIRECTIONS FOR FUTURE WORK

The developed concepts in the previous chapters and their results open doors to possible future research directions:

Chapter 2 uses the weighted Laplacian and its pseudo-inverse to express the linearised DC power flow equations in power grids. However, a weighted Laplacian can describe many other processes, that are linear –or could be linearised– to the network topology such as water flow networks, mechanical or thermal systems. It would be an interesting research direction to apply the similar concepts such as effective graph resistance to those other types of networks.

Chapter 3 shows that the power grids are vulnerable to targeted attacks. What to do after the targeted attacks is an attractive direction for future work. Given that the effects of the attacks are temporary (i.e., network components are not permanently damaged), how to optimally (in terms of cost, time etc.) restore the network back to the initial conditions?

Chapter 4 investigates the steady-state operation of power grids under given deterministic generation and demand profiles. It would be a challenge to analyse the performance indicators of power grids and decide on the 'optimal' topology under probabilistic generation and demand profiles.

Chapter 5 proposes models to simulate the evolution of cascading failures in power grids initiated by single- or double-line failures. This work can be extended to analyse the effects of different sets and patterns of line and/or substation outages or generation and/or demand changes, for instance during the case of natural disasters.

Chapter 6 proposes a long-term wind power integration plan under the constraints from power grids. It would be possible to turn the integration plan into a multi-objective optimization problem, and decide on feasible network investments which can welcome more wind farm integration at a certain locations. In addition, for other near-future challenges, such as the placing the electric vehicle charging stations, can one apply a similar methodology?

REFERENCES

[1] *Final Report on the August 14th Blackout in the United States and Canada: Causes and Recommendations*, Tech. Rep. (U.S.- Canada Power System Outage Task Force, April 2004).

[2] *Report on Blackout in Turkey on 31st March 2015*, Tech. Rep. (ENTSO-E, 2015).

[3] A. Bakshi, A. Velayutham, S. Srivastava, K. Agrawal, R. Nayak, S. Soonee, and B. Singh, *Report of the enquiry committee on grid disturbance in Northern Region on 30th July 2012 and in Northern, Eastern & North-Eastern Region on 31st July 2012,* New Delhi, India (2012).

[4] P. Hines, J. Apt, H. Liao, and S. Talukdar, *The frequency of large blackouts in the united states electrical transmission system: an empirical study,* in *2nd Carnegie Mellon Conference on Electric Power, January* (Citeseer, 2006).

[5] S. H. Strogatz, *Exploring complex networks,* Nature **410**, 268 (2001).

[6] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks,* Science **286**, 509 (1999).

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

[8] D. P. Chassin and C. Posse, *Evaluating North American electric grid reliability using the Barabasi–Albert network model,* Physica A **355**, 667 (2005).

[9] M. H. Brown and R. P. Sedano, *Electricity transmission: A primer,* (National Conference of State, 2004).

[10] C. Sulzberger, *Thomas Edison's 1882 Pearl Street generating station,* IEEE Global History Network (2010).

[11] M. Van Eeten, A. Nieuwenhuijs, E. Luiijf, M. Klaver, and E. Cruz, *The state and the threat of cascading failure across critical infrastructures: The implications of empirical evidence from media incident reports,* Public Administration **89**, 381 (2011).

[12] J. M. Barrett, *Challenges and Requirements for Tomorrow's Electrical Power Grid* (Lexington Institute, June 2016).

[13] U. S. G. A. Office, *Actions Needed to Improve the Identification and Management of Electrical Power Risks and Vulnerabilities to DOD Critical Assets,* Tech. Rep. (GAO-10-147, October 2009).

[14] *Monthly energy review: Electricity sales and power sector generating capacity in the reference case, 1949-2040*, Tech. Rep. (U.S. Energy Information Administration, 2013).

[15] *World energy outlook 2017*, Tech. Rep. (International Energy Agency, 2017).

[16] B. Warshay, *Upgrading the grid*, Foreign Affairs (2015).

[17] R. Smith, *Assault on California power station raises alarm on potential for terrorism*, Wall Street Journal **5** (2014).

[18] F. Alvarado and S. Oren, *Transmission system operation and interconnection*, National transmission grid study–Issue papers (2002).

[19] J. J. Grainger and W. D. Stevenson, *Power system analysis* (McGraw-Hill, 1994).

[20] M. E. Newman, *The structure and function of complex networks*, SIAM review **45**, 167 (2003).

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

[22] Y. Koç, *On Robustness of Power Grids*, Ph.D. thesis, Delft University of Technology (2015).

[23] H. Cetinay, F. A. Kuipers, and P. Van Mieghem, *A topological investigation of power flow*, IEEE Systems Journal **12** (2018).

[24] R. Baldick, B. Chowdhury, I. Dobson, Z. Dong, B. Gou, D. Hawkins, H. Huang, M. Joung, D. Kirschen, F. Li, *et al.*, *Initial review of methods for cascading failure analysis in electric power transmission systems*, in *Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE* (IEEE, 2008) pp. 1–8.

[25] E. Bompard, M. Masera, R. Napoli, and F. Xue, *Assessment of structural vulnerability for power grids by network performance based on complex networks*, in *International Workshop on Critical Information Infrastructures Security* (Springer, 2008) pp. 144–154.

[26] E. Negeri, F. Kuipers, and N. Baken, *Assessing the topological structure of a smart low-voltage grid*, Int. J. of Critical Infrastructure Protection **9**, 24 (2015).

[27] Y. Koç, M. Warnier, P. Van Mieghem, R. E. Kooij, and F. M. Brazier, *The impact of the topology on cascading failures in a power grid model*, Physica A: Statistical Mechanics and its Applications **402**, 169 (2014).

[28] J. M. Hernández and P. Van Mieghem, *Classification of graph metrics*, Delft University of Technology, Tech. Rep (2011).

[29] M. Rosas-Casals, S. Valverde, and R. V. Solé, *Topological vulnerability of the European power grid under errors and attacks,* International Journal of Bifurcation and Chaos **17**, 2465 (2007).

[30] Y. Koç, T. Verma, N. A. Araujo, and M. Warnier, *Matcasc: A tool to analyse cascading line outages in power grids,* in *Intelligent Energy Systems (IWIES), 2013 IEEE International Workshop on* (IEEE, 2013) pp. 143–148.

[31] X. Chen, K. Sun, Y. Cao, and S. Wang, *Identification of vulnerable lines in power grid based on complex network theory,* in *Power Engineering Society General Meeting, 2007. IEEE* (IEEE, 2007) pp. 1–6.

[32] E. Bompard, E. Pons, and D. Wu, *Extended topological metrics for the analysis of power grid vulnerability,* IEEE Systems Journal **6**, 481 (2012).

[33] S. Soltan, D. Mazauric, and G. Zussman, *Cascading failures in power grids: Analysis and algorithms,* in *Proceedings of the 5th international conference on Future energy systems* (ACM, 2014) pp. 195–206.

[34] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, *Usefulness of DC power flow for active power flow analysis with flow controlling devices,* in *Proc. IET ACDC'06* (2006) pp. 58–62.

[35] P. Van Mieghem, *Graph spectra for complex networks* (Cambridge University Press, 2010).

[36] W. Ellens, F. Spieksma, P. Van Mieghem, A. Jamakovic, and R. Kooij, *Effective graph resistance,* Linear algebra and its applications **435**, 2491 (2011).

[37] A. Ghosh, S. Boyd, and A. Saberi, *Minimizing effective resistance of a graph,* SIAM review **50**, 37 (2008).

[38] T. Guler, G. Gross, and M. Liu, *Generalized line outage distribution factors,* IEEE Transactions on Power Systems **22**, 879 (2007).

[39] P. W. Sauer, K. E. Reinhard, and T. J. Overbye, *Extended factors for linear contingency analysis,* in *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on* (IEEE, 2001) pp. 697–703.

[40] C. D. Meyer, Jr, *Generalized inversion of modified matrices,* SIAM Journal on Applied Mathematics **24**, 315 (1973).

[41] L. O. Chua, C. A. Desoer, and E. S. Kuh, *Linear and nonlinear circuits* (McGraw-Hill College, 1987).

[42] IEEE, *Power systems test case archive,* (2018), available at: http://www.ee.washington.edu/research/pstca/.

[43] X. Wang, Y. Koç, R. E. Kooij, and P. Van Mieghem, *A network approach for power grid robustness against cascading failures,* in *Reliable Networks Design and Modeling (RNDM), 2015 7th International Workshop on* (IEEE, 2015) pp. 208–214.

[44] H. Cetinay, K. Devriendt, and P. Van Mieghem, *Nodal vulnerability to targeted attacks in power grids,* Applied Network Science, Springer Open **3** (2018).

[45] P. Van Mieghem, K. Devriendt, and H. Cetinay, *Pseudoinverse of the Laplacian and best spreader node in a network,* Physical Review E **96**, 032311 (2017).

[46] S. Trajanovski, J. Martín-Hernández, W. Winterbach, and P. Van Mieghem, *Robustness envelopes of networks,* Journal of Complex Networks **1**, 44 (2013).

[47] D. F. Rueda, E. Calle, and J. L. Marzo, *Robustness comparison of 15 real telecommunication networks: Structural and centrality measurements,* Journal of Network and Systems Management **25**, 269 (2017).

[48] Y. Koç, M. Warnier, R. Kooij, and F. Brazier, *Structural vulnerability assessment of electric power grids,* in *Networking, Sensing and Control (ICNSC), 2014 IEEE 11th International Conference on* (IEEE, 2014) pp. 386–391.

[49] E. Bompard, R. Napoli, and F. Xue, *Analysis of structural vulnerabilities in power transmission grids,* International Journal of Critical Infrastructure Protection **2**, 5 (2009).

[50] L. Cuadra, S. Salcedo-Sanz, J. Del Ser, S. Jiménez-Fernández, and Z. W. Geem, *A critical review of robustness in power grids using complex networks concepts,* Energies **8**, 9211 (2015).

[51] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, *Do topological models provide good information about electricity infrastructure vulnerability?* Chaos **20**, 033122 (2010).

[52] S. Arianos, E. Bompard, A. Carbone, and F. Xue, *Power grid vulnerability: A complex network approach,* Chaos: An Interdisciplinary Journal of Nonlinear Science **19**, 013119 (2009).

[53] P. Hines and S. Blumsack, *A centrality measure for electrical networks,* in *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual* (IEEE, 2008) pp. 185–185.

[54] X. Qi, E. Fuller, R. Luo, and C.-q. Zhang, *A novel centrality method for weighted networks based on the Kirchhoff polynomial,* Pattern Recognition Letters **58**, 51 (2015).

[55] I. Gorton, Z. Huang, Y. Chen, B. Kalahar, S. Jin, D. Chavarría-Miranda, D. Baxter, and J. Feo, *A high-performance hybrid computing approach to massive contingency analysis in the power grid,* in *Fifth IEEE International Conference on E-Science* (IEEE, 2009) pp. 277–283.

[56] A. Nasiruzzaman, H. Pota, and M. Mahmud, *Application of centrality measures of complex network framework in power grid,* in *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society* (IEEE, 2011) pp. 4660–4665.

[57] H. Cetinay, et al., *Targeted attacks simulator in power grids,* (2018), available at: https://github.com/hcetinay/Targeted-node-attacks-in-power-grids.

[58] A. R. Bergen and V. Vittal, *Power Systems Analysis* (Prentice-Hall, 1999).

[59] H. Cetinay, S. Soltan, F. A. Kuipers, G. Zussman, and P. Van Mieghem, *Comparing the effects of failures in power grids under the AC and DC power flow models,* IEEE Transactions on Network Science and Engineering (2017).

[60] E. Ciapessoni, D. Cirio, A. Pitto, P. Marcacci, M. Lacavalla, S. Massucco, F. Silvestro, and M. Sforna, *A risk-based methodology and tool combining threat analysis and power system security assessment,* Energies **11**, 83 (2017).

[61] L. C. Freeman, *Centrality in social networks conceptual clarification,* Social networks **1**, 215 (1978).

[62] P. Bonacich, *Power and centrality: A family of measures,* American journal of sociology **92**, 1170 (1987).

[63] P. Bonacich, *Simultaneous group and individual centralities,* Social networks **13**, 155 (1991).

[64] L. C. Freeman, *A set of measures of centrality based on betweenness,* Sociometry, 35 (1977).

[65] E. Bompard, D. Wu, and F. Xue, *The concept of betweenness in the analysis of power grid vulnerability,* in *Complexity in Engineering, 2010. COMPENG'10.* (IEEE, 2010) pp. 52–54.

[66] F. Gutierrez, E. Barocio, F. Uribe, and P. Zuniga, *Vulnerability analysis of power grids using modified centrality measures,* Discrete Dynamics in Nature and Society **2013** (2013).

[67] M. E. Newman, *A measure of betweenness centrality based on random walks,* Social networks **27**, 39 (2005).

[68] D. J. Klein and M. Randić, *Resistance distance,* Journal of mathematical chemistry **12**, 81 (1993).

[69] F. Morone and H. A. Makse, *Influence maximization in complex networks through optimal percolation,* Nature **524**, 65 (2015).

[70] F. Morone, B. Min, L. Bo, R. Mari, and H. A. Makse, *Collective influence algorithm to find influencers via optimal percolation in massively large social media,* Scientific reports **6**, 30062 (2016).

[71] S. Mugisha and H.-J. Zhou, *Identifying optimal targets of network attack by belief propagation,* Physical Review E **94**, 012305 (2016).

[72] L. Zdeborová, P. Zhang, and H.-J. Zhou, *Fast and simple decycling and dismantling of networks,* Scientific reports **6**, 37954 (2016).

[73] M. Molloy and B. Reed, *The size of the giant component of a random graph with a given degree sequence,* Combinatorics, probability and computing **7**, 295 (1998).

[74] J. Martín Hernández, Z. Li, and P. Van Mieghem, *Weighted betweenness and algebraic connectivity,* Journal of Complex Networks **2**, 272 (2014).

[75] H. Cetinay, Y. Koç, F. A. Kuipers, and P. Van Mieghem, *Powerweb: Intelligent energy systems,* (Springer, 2018) Chap. Topology-Driven Performance Analysis of Power Grids.

[76] G. A. Pagani and M. Aiello, *The power grid as a complex network: A survey,* Physica A: Statistical Mechanics and its Applications **392**, 2688 (2013).

[77] Y. Koç, M. Warnier, P. Van Mieghem, R. E. Kooij, and F. M. Brazier, *A topological investigation of phase transitions of cascading failures in power grids,* Physica A: Statistical Mechanics and its Applications **415**, 273 (2014).

[78] P. Crucitti, V. Latora, and M. Marchiori, *A topological analysis of the Italian electric power grid,* Physica A: Statistical mechanics and its applications **338**, 92 (2004).

[79] X. Wang, E. Pournaras, R. E. Kooij, and P. Van Mieghem, *Improving robustness of complex networks via the effective graph resistance,* The European Physical Journal B **87**, 221 (2014).

[80] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, *Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,* IEEE Trans. Power Syst. **26**, 12 (2011).

[81] Y. Koç, M. Warnier, R. E. Kooij, and F. M. Brazier, *An entropy-based metric to quantify the robustness of power grids against cascading failures,* Safety science **59**, 126 (2013).

[82] H. Cetinay, S. Soltan, F. A. Kuipers, G. Zussman, and P. Van Mieghem, *Analyzing cascading failures in power grids under the AC and DC power flow models,* in *The 35th International Symposium on Computer Performance, Modeling, Measurements and Evaluation (IFIP Performance'17)* (2017).

[83] S. Soltan, A. Loh, and G. Zussman, *Analyzing and quantifying the effect of k-line failures in power grids,* to appear in IEEE Trans. Control Netw. Syst. (2017).

[84] I. Dobson, *Encyclopedia of Systems and Control*, Cascading network failure in power grid blackouts (Springer, 2015) pp. 105–108.

[85] B. Schäfer, D. Witthaut, M. Timme, and V. Latora, *Dynamically induced cascading failures in supply networks,* arXiv preprint arXiv:1707.08018 (2017).

[86] P. Hines, K. Balasubramaniam, and E. C. Sanchez, *Cascading failures in power grids,* IEEE Potentials **28**, 24 (2009).

[87] D. Bienstock, *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*, Vol. 22 (SIAM, 2016).

[88] P. D. Hines and P. Rezaei, *Cascading failures in power systems,* Smart Grid Handbook (2016).

[89] S. Soltan, D. Mazauric, and G. Zussman, *Analysis of failures in power grids,* in IEEE Trans. Control Netw. Syst. **4**, 288 (2017).

[90] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, *Usefulness of DC power flow for active power flow analysis,* in *IEEE PES-GM'05* (2005).

[91] D. P. Nedic, I. Dobson, D. S. Kirschen, B. A. Carreras, and V. E. Lynch, *Criticality in a cascading failure blackout model,* Int. J. Elec. Power **28**, 627 (2006).

[92] M. Papic, K. Bell, Y. Chen, I. Dobson, L. Fonte, E. Haq, P. Hines, D. Kirschen, X. Luo, S. Miller, *et al.*, *Survey of tools for risk assessment of cascading outages,* in *IEEE PES-GM'11* (2011).

[93] H. Cetinay, et al., *Cascading failures simulator in power grids,* (2017), available at: https://github.com/TUDelftNAS/AC-Cascade-Sim.

[94] R. Kaye and F. Wu, *Analysis of linearized decoupled power flow approximations for steady-state security assessment,* IEEE Trans. Circuits and Sys. **31**, 623 (1984).

[95] T. J. Overbye, X. Cheng, and Y. Sun, *A comparison of the AC and DC power flow models for LMP calculations,* in *Proc. IEEE HICSS'04* (2004).

[96] P. Yan and A. Sekar, *Study of linear models in steady state load flow analysis of power systems,* in *IEEE PES-WM'02* (2002).

[97] B. Stott, J. Jardim, and O. Alsaç, *DC power flow revisited,* IEEE Trans. Power Syst. **24**, 1290 (2009).

[98] S. Deckmann, A. Pizzolante, A. Monticelli, B. Stott, and O. Alsac, *Numerical testing of power system load flow equivalents,* IEEE Trans. Power App. and Sys., 2292 (1980).

[99] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version* (Cengage Learning, 2012).

[100] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, *Power grid vulnerability to geographically correlated failures - analysis and control implications,* in *Proc. IEEE INFOCOM'14* (2014).

[101] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*, 3rd ed. (John Wiley & Sons, 2012).

[102] H. Cetinay, F. A. Kuipers, and A. N. Guven, *Optimal siting and sizing of wind farms,* Renewable Energy, 51 (2017).

[103] H. Cetinay, T. Kekeç, F. A. Kuipers, and D. Tax, *Markov random field for wind farm planning,* in *5th IEEE International Conference on Smart Energy Grid Engineering (SEGE)* (IEEE, 2017).

[104] D. R. Chandra, M. S. Kumari, and M. Sydulu, *A detailed literature review on wind forecasting,* in *2013 International Conference on Power, Energy and Control (ICPEC)* (IEEE, 2013) pp. 630–634.

[105] J. A. Carta, P. Ramirez, and S. Velazquez, *A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands,* Renewable and Sustainable Energy Reviews **13**, 933 (2009).

[106] J. Caers and M. A. Maes, *Identifying tails, bounds and end-points of random variables,* Structural Safety **20**, 1 (1998).

[107] Y. Xiao, Q. Li, Z. Li, Y. Chow, and G. Li, *Probability distributions of extreme wind speed and its occurrence interval,* Engineering Structures **28**, 1173 (2006).

[108] A. Garcia, J. Torres, E. Prieto, and A. De Francisco, *Fitting wind speed distributions: A case study,* Solar energy **62**, 139 (1998).

[109] J. Seguro and T. Lambert, *Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis,* Journal of Wind Engineering and Industrial Aerodynamics **85**, 75 (2000).

[110] N. G. Mortensen, L. Landberg, I. Troen, and E. Lundtang Petersen, *Wind Atlas Analysis and Application program (WAsP): Vol. 1: Getting started,* Tech. Rep. (Risø National Laboratory, 1998).

[111] T. Ackermann, *Wind power in power systems* (John Wiley & Sons, 2005).

[112] M. Albadi and E. El-Saadany, *Overview of wind power intermittency impacts on power systems,* Electric Power Systems Research **80**, 627 (2010).

[113] H. Ibrahim, M. Ghandour, M. Dimitrova, A. Ilinca, and J. Perron, *Integration of wind energy into electricity systems: Technical challenges and actual solutions,* Energy Procedia **6**, 815 (2011).

[114] H. Falaghi and M.-R. Haghifam, *ACO based algorithm for distributed generation sources allocation and sizing in distribution systems,* in *Power Tech, 2007 IEEE Lausanne* (IEEE, 2007) pp. 555–560.

[115] C. L. Borges and D. M. Falcao, *Optimal distributed generation allocation for reliability, losses, and voltage improvement,* International Journal of Electrical Power & Energy Systems **28**, 413 (2006).

[116] S. H. Gilani, H. Afrakhte, and M. J. Ghadi, *Probabilistic method for optimal placement of wind-based distributed generation with considering reliability improvement and power loss reduction,* in *Thermal Power Plants (CTPP), 2012 4th Conference on* (IEEE, 2012) pp. 1–6.

[117] F. Türksoy, *Investigation of wind power potential at Bozcaada, Turkey,* Renewable Energy **6**, 917 (1995).

[118] R. Barthelmie, M. Courtney, J. Højstrup, and S. E. Larsen, *Meteorological aspects of offshore wind energy: Observations from the Vindeby wind farm,* Journal of Wind Engineering and Industrial Aerodynamics **62**, 191 (1996).

[119] S. V. Dhople and A. D. Domínguez-García, *A framework to determine the probability density function for the output power of wind farms,* in *North American Power Symposium (NAPS), 2012* (IEEE, 2012) pp. 1–6.

[120] P. Van Mieghem, *Performance analysis of complex networks and systems* (Cambridge University Press, 2014).

[121] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, Vol. 4 (Prentice hall New Jersey, 1995).

[122] D. Latinopoulos and K. Kechagia, *A GIS-based multi-criteria evaluation for wind farm site selection. A regional scale application in Greece,* Renewable Energy **78**, 550 (2015).

[123] Y. Noorollahi, H. Yousefi, and M. Mohammadi, *Multi-criteria decision support system for wind farm site selection using GIS,* Sustainable Energy Technologies and Assessments **13**, 38 (2016).

[124] T. Jayaweera and H. Haeri, *The uniform methods project: Methods for determining energy efficiency savings for specific measures* (National Renewable Energy Laboratory, 2013).

[125] N. Y. Aydin, E. Kentel, and S. Duzgun, *GIS-based environmental assessment of wind energy systems for spatial planning: A case study from western Turkey,* Renewable and Sustainable Energy Reviews **14**, 364 (2010).

[126] *Wind power atlas of Turkey* (Republic of Turkey, Ministry of Energy and Natural Resources).

ACKNOWLEDGEMENTS

First of all, I owe my deepest gratitude to my two supervisors and (co)-promotors, Prof. Piet Van Mieghem and Dr. Fernando A. Kuipers. It has been my honour to have been working with you both. Piet has always impressed me with his passion and his rigorous approach to science and scientific problems. I am grateful to have my Ph.D. studies in his group. My sincere gratitude goes to Fernando as he is the one who originally hired me to the NAS group and has been supervising me from the beginning. He gave me the freedom and encouragement to pursue my ideas and his kind advises helped me to formulate my interesting questions in a scientific manner. His keen eye for details has also helped me a lot to improve my academic writing skills.

I would like to express my gratitude to the committee members of my thesis defence for their time and efforts on my thesis. I want to thank my brother-in-law Robert de Korte for helping me with the translation of the summary into Dutch. I would like to also acknowledge Alliander N.V., the Netherlands which has funded the Ph.D. project I have worked on and my contact person Frans Campfens who acted as a mentor and helped me a lot to start my research.

During my Ph.D. studies, I have been fortunate to work together and collaborate with many distinguished researchers, Prof. Nezih Güven, Prof. Jose L. Marzo, Prof. Robert E. Kooij, Dr. Gil Zussman, Dr. David Tax, Dr. Yakup Koç, Dr. Saleh Soltan, Karel Devriendt, Sergio Gomez Cosgaya and Taygun Kekeç. I am very grateful for the opportunity.

Next, I would like to thank the members of NAS group and my office mates. Time has flown fast with them. Dr. Marcus Märtens always made my day with his humour and Zhidong He was frequently eager to share his Chinese tea with me. I am glad to have Misa Taguchi as an office mate at least in my final year of my studies. I enjoyed spending time with her and our walks during the lunch break. I also need to specifically mention Dr. Jaron Sanders and Dr. Jil Meier, who were ready to help and listen, and my office mates Dr. Xiangrong Wang, Peng Sun, and Dr. Ruud van de Bovenkamp – although only for a few months. I would like to thank Prof. Robert E. Kooij who has inspired me with his research attitude on the fruitful applications of science to solve real-life problems. Many thanks also go to Dr. Edgar van Boven, Walter Knoop, Dr. Huijuan Wang, Dr. Remco Litjens, Dr. Eric Smeitink, Prof. Nico Baken, Karel Devriendt, Bastian Prasse, Long Ma, Qiang Liu, Rogier Noldus, Dr. Niels van Adrichem, Dr. Norbert Blenn, Dr. Farabi Iqbal, Dr. Cong Li, Dr. Song Yang, Dr. Stojan Trajanovski, and our former and current secretaries Wendy Murtinu-van Schagen, Marloes van der Krogt-van Lier, Rani Ramnares, Lidwina Tromp and Joyce van Velzen to make our lives easier.

My sincere gratitude also goes to my friends in Delft who have made my time here more enjoyable. Duygu Güroğlu has been a good companion for me in Delft. I have always enjoyed our discussions over both scientific and non-scientific subjects with Taygun Kekeç. During my Dutch courses, I met Marina Pogos and since then we enjoyed our lunch walks during the campus. Many thanks also go to Dr. Hamdi Dibeklioğlu, Dr.

Görkem Saygılı, Dr. Kasım Sinan Yıldırım, and Günay Aslan. I want to also mention my friends from Turkey, Dr. Itır Önal Ertuğrul, Buket Büyüksaraç, Serdar Büyüksaraç, Necati Çağan, Fulya Ağırnas, Ercan Çandır and of course Kübra Aşık Akdemir.

Finally, I would like to thank my mother Nurhan Çetinay and my father İbrahim Çetinay. I am grateful to my family for their unconditional love, care and support. Behind my every accomplishment stand their help and pure love. They were always with me despite the 3000 km between us. My warmest feelings also go to my grandmother Mahmure Köse, my aunts Reyhan Arkan and Şükran Hun, and my cousin Gökhan Arkan. They were one of the reason why I was looking for next my travel to Turkey. I had the luxury to have my sister, Dr. Hande de Korte, here in the Netherlands. Her fruitful Ph.D. studies in the Eindhoven University of Technology had inspired to come to the Netherlands to pursue my own Ph.D. She and her husband Robert de Korte have provided support whenever I needed. Last but not least, I need to especially thank to my dear husband İsmail Mert İyicil who has been the best supporter and motivator throughout my journey. His presence makes me wonder how I could make it till the end as a single Ph.D.! I owe my ongoing happiness even during the difficult times to him.

Time flies fast. It has been nearly four years since I started my Ph.D. journey, but feels much shorter. The famous adage of Piet "Ph.D. is *just* four years" makes much more sense now.

Hale Çetinay İyicil
Delft, 2018

TEŞEKKÜR

Öncelikle her iki danışmanım Prof. Piet Van Mieghem ve Dr. Fernando A. Kuipers'a en içten gönül borcumu belirtmek isterim. Her ikisiyle de çalışmış olmak benim için bir onurdu. Piet'in bilime ve bilimsel problemlere dikkatli ve tutkulu yaklaşımı beni hep olumlu olarak etkiledi. Doktora çalışmamı onun grubunda yapmış olmaktan müteşekkirim. Fernando'ya teşekkür ediyorum, beni bu grupta işe alan ve başından beri günlük danışmanlığımı yapan o oldu. Bana bu doktora çalışmasında özgürlük verip kendi fikirlerimi onun tavsiyeleri ve bilimsel yöntemler çerçevesinde ortaya koymamı hep teşvik etti. Ayrıca onun dikkati akademik yazımımı geliştirmemde çok etkili oldu.

Komite üyelerine göstermiş oldukları ilgi ve ayırdıkları zamandan dolayı şükranlarımı sunmak istiyorum. Enişte Robert de Korte'ye özet kısmını Hollandaca'ya çevirmesinden dolayı teşekkür ediyorum. Doktoram boyunca projeme kaynak sunmuş olan Alliander N.V.'yi belirtmek ve şirketteki danışmanım olan Frans Campfens'a bana vermiş olduğu danışmanlık ve yardımdan dolayı teşekkür etmek istiyorum.

Doktoram süresince birlikte çalışma fırsatı bulmuş olduğum Prof. Nezih Güven, Prof. Jose L. Marzo, Prof. Robert E. Kooij, Dr. Gil Zussman, Dr. David Tax, Dr. Yakup Koç, Dr. Saleh Soltan, Karel Devriendt, Sergio Gomez Cosgaya ve Taygun Kekeç'e teşekkürlerimi sunar ve vermiş oldukları bu fırsat için müteşekkir olduğumu belirtmek isterim.

NAS grubu ve ofis arkadaşlarıma teşekkür etmek istiyorum. Zaman onlarla çok hızlı ilerledi. Dr. Marcus Märtens'in şaka anlayışı ve Zhidong He'nin Çin çayı paylaşma isteği beni her zaman eğlendirdi ve mutlu etti. Son senemde Misa Taguchi'yi ofis arkadaşı edinmekten çok mutluyum. Onunla zaman geçirmekten ve öğle yürüyüşlerimizden çok zevk aldım. Ayrıca özellikle Dr. Jaron Sanders ve Dr. Jil Meier ve kisa sureli ofis arkadaşlarım Dr. Xiangrong Wang, Peng Sun, ve Dr. Ruud van de Bovenkamp'a beni her zaman dinlemeye ve yardım etmeye hazır oldukları için teşekkürlerimi iletmek isterim. Prof. Robert E. Kooij'un günümüz sorunlarını çözmeye yönelik bilimsel tutumu ve uygulamaları bana ilham kaynağı oldu. Ayrıca Dr. Edgar van Boven, Walter Knoop, Dr. Huijuan Wang, Dr. Remco Litjens, Dr. Eric Smeitink, Prof. Nico Baken, Karel Devriendt, Bastian Prasse, Long Ma, Qiang Liu, Rogier Noldus, Dr. Niels van Adrichem, Dr. Norbert Blenn, Dr. Farabi Iqbal, Dr. Cong Li, Dr. Song Yang, Dr. Stojan Trajanovski'a ve önceki ile şu anki sekreterlerimiz Wendy Murtinu-van Schagen, Marloes van der Krogt-van Lier, Rani Ramnares, Lidwina Tromp ve Joyce van Velzen'e hayatımızı kolaylaştırdıkları için teşekkür ederim.

Ayrıca Delft'teki arkadaşlarıma benim hayatımı daha eğlenceli yaptıkları için teşekkür ederim. Duygu Güroğlu Delft'teki çok iyi bir yoldaşımdı. Taygun Kekeç ile olan bilimsel ve bilimsel olmayan tartışmalarımızdan çok keyif aldım. Hollandaca dersleri sayesinde tanışmış olduğum Marina Pogos'a kampüsteki öğle yürüyüşlerinde vermiş olduğu keyifli sohbetten dolayı teşekkür ediyorum. Ayrıca Dr. Hamdi Dibeklioğlu, Dr. Görkem Saygılı, Dr. Kasım Sinan Yıldırım, ve Günay Aslan'a birçok kez teşekkür etmek istiyorum. Türkiye'deki arkadaşlarım Dr. Itır Önal Ertuğrul, Buket Büyüksaraç, Serdar

Büyüksaraç, Necati Çağan, Fulya Ağırnas, Ercan Çandır ve Kübra Aşık Akdemir'den de söz edip teşekkür etmek istiyorum.

Aramızdaki 3000 kilometre mesafeye rağmen her zaman bana vermiş oldukları kıymet, destek ve saf sevgiden dolayı canım annem Nurhan Çetinay ve canım babam İbrahim Çetinay'a teşekkürlerimi sunuyorum. Hayattaki her başarım onların bana verdiği sevgi ve emekler sayesinde mümkün oldu. Türkiye'ye yapacağım her geziyi dört gözle beklememe sebep oldukları için, en sıcak duygularımı anneannem Mahmure Köse, teyzelerim Reyhan Arkan ve Şükran Hun, ve kuzenim Gökhan Arkan'a iletiyorum. Ablam Dr. Hande de Korte'nin Hollanda'da olması benim için bir avantajdı. Onun Eindhoven University of Technology'de yapmış olduğu doktora çalışmaları, benim kendi doktoramı yapmaya karar vermemde büyük bir etken oldu. Ablam ve eşi Robert de Korte bizden desteğini hiç bir zaman esirgemedi. Son olarak değerli eşim İsmail Mert İyicil'e özellikle teşekkür ediyorum. Kendisi bizzat bu yoldaki en büyük destekçim ve motivasyon kaynağım oldu. Mert olmadan doktora çalışmalarımı burada tek başıma nasıl sürdürebilirdim bilemiyorum. Zorlu zamanlarımda bile sürekli mutluluğumu ona borçluyum.

Zaman çok hızlı uçup gidiyor. Doktora yapma yoluna çıkalı neredeyse 4 yıl oldu ama zaman çok hızlı gecti. Piet'in her yeni öğrencisine söylediği ünlü "Doktora çalışmalarınız için *sadece* 4 yılınız var." deyişi yolun soluna gelince daha çok anlam kazandı.

Hale Çetinay İyicil
Delft, 2018

CURRICULUM VITÆ

Hale Çetinay was born in Ankara, Turkey on 14 March 1988. She completed her Bachelor's degree at Middle East Technical University, Ankara in June 2011 in the Electrical and Electronics Engineering. She continued her studies at Middle East Technical University and received her Master's degree in the Electrical and Electronics Engineering in May 2014.

During the last semester of her Bachelor studies, she worked as a part-time researcher in the Power Systems Analysing and Planning group of the Scientific and Technological Research Council of Turkey (TÜBİTAK), Ankara, where she continued working as a full-time power system engineer during her Master studies. Since September 2014, she has been pursuing her Ph.D. in the Network Architectures and Services (NAS) group in the Faculty of Electrical Engineering, Mathematics and Computer Science at Delft University of Technology, the Netherlands, where she has performed researches on the modelling, analyses and robustness of power grids.

During her Ph.D. studies, she assisted Prof. Piet Van Mieghem with teaching the final-exercise classes of the Master's course *Performance Analysis* and with supervising the students projects of the Master's course *Networking*. She has reviewed several articles for peer-reviewed international journals and conferences. She also performed a Short Term Scientific Mission (STSM) over the robustness and critical elements in real-world structures in Girona University, Spain, granted by European COST action.

LIST OF PUBLICATIONS

JOURNAL PUBLICATIONS

9. **H. Cetinay**, S. Soltan, F. A. Kuipers, G. Zussman, and P. Van Mieghem, *Comparing the Effects of Failures in Power Grids under the AC and DC Power Flow Models*, to appear in IEEE Transactions on Network Science and Engineering, 2017.

8. P. Van Mieghem, K. Devriendt and **H. Cetinay**, *Pseudo-inverse of the Laplacian and Best Spreader Node in a Network*, Physical Review E, vol. 96, no. 3, pp. 032311, 2017.

7. **H. Cetinay**, F. A. Kuipers, and A. N. Guven, *Optimal Siting and Sizing of Wind Farms*, Renewable Energy, Elsevier, vol. 101, pp. 51-58, 2017.

6. **H. Cetinay**, F. A. Kuipers and P. Van Mieghem, *A Topological Investigation of Power Flow*, IEEE Systems Journal, vol. 12, no. 3, pp. 2524-2532, 2018.

5. **H. Cetinay**, K. Devriendt and P. Van Mieghem, *Nodal Vulnerability to Targeted Attacks in Power Grids*, Applied Network Science, Springer Open, vol. 3, no. 34, 2018.

CONFERENCE PUBLICATIONS

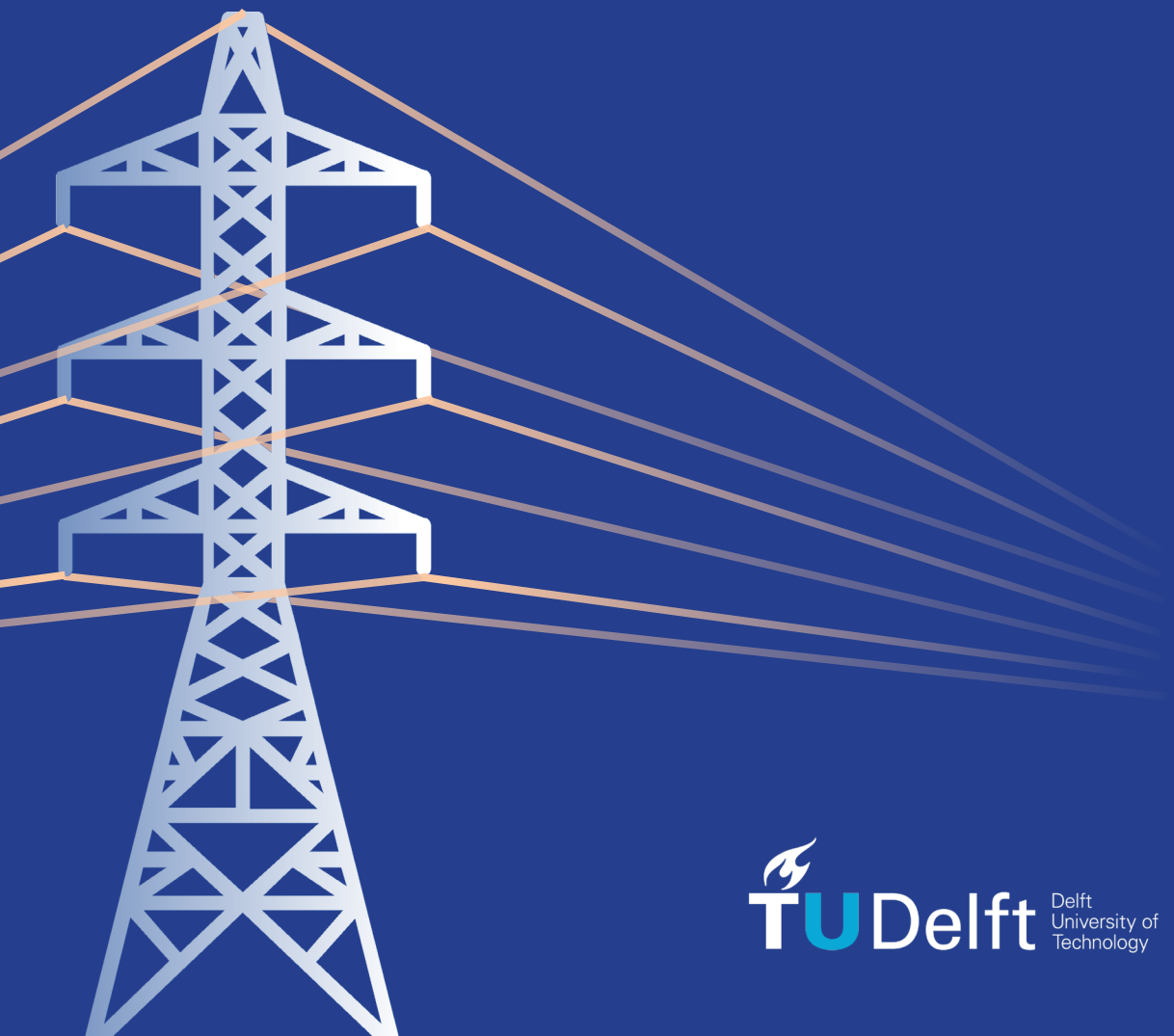
4. **H. Cetinay**, T. Kekec, F. A. Kuipers, and D. M. J. Tax, *Markov Random Field for Wind Farm Planning*, The 5th IEEE International Conference on Smart Energy Grid Engineering, 2017.

3. **H. Cetinay**, S. Soltan, F. A. Kuipers, G. Zussman, and P. Van Mieghem, *Analyzing Cascading Failures in Power Grids under the AC and DC Power Flow Models*, The 35th International Symposium on Computer Performance, Modeling, Measurements and Evaluation, 2017.

BOOK CHAPTERS

2. **H. Cetinay**, Yakup Koç, F. A. Kuipers and P. Van Mieghem, *Topology-Driven Performance Analysis of Power Grids*, to appear in Intelligent Integrated Energy Systems, The PowerWeb Program at TU Delft, 2018.

1. **H. Cetinay**, S. Gomez, J. L. Marzo, R. Kooij and P. Van Mieghem *The Most Destructive Strategies for Attacking Networks*, in preparation, to appear in European Cooperation in Science and Technology RECODIS Action Book, 2018.



TU Delft Delft University of Technology