REAL-TIME WEATHER-DEPENDENT PROBABILISTIC RELIABILITY ASSESSMENT OF THE ICELANDIC POWER SYSTEM

Samuel Perkin

January 2018
School of Science and Engineering
Reykjavík University

Ph.D. Dissertation

Abstract

Power system operation is concerned with supplying electricity to society in a reliable manner. Reliability has traditionally been managed through use of the deterministic N-1 criterion. The EU FP7 GARPUR research project proposes that Transmission System Operators (TSOs) transition to the use of a probabilistic reliability management approach and criteria. This proposal is based on the N-1 criterion's insensitivity to: the variable probability of exogenous threats to power systems; varying spatio-temporal and socio-economic impacts of service outages; and the trade-off between preventive and corrective controls, among other aspects.

This work aimed to implement the theoretical framework presented by GARPUR on the Icelandic transmission system in the context of real-time operations. This was achieved by first interpreting the GARPUR framework into an ideal implementation. The ideal implementation was then simplified, by considering present day data and tool availability, into a first-step implementation of the GARPUR framework, which was prototyped on academic test systems. This formed the basis for developing a pilot test of the GARPUR framework on the Icelandic transmission system. The pilot test was extended beyond present day tools and data, through the creation of weather-dependent failure rate models to determine the importance of weather in real-time risk assessment, and a custom-built system response model to assess the impact of various assumptions and system protection devices.

The core of this work is a validation that the approach proposed by the GARPUR project can be implemented as a tractable and practically useful reliability assessment process for real-world system operation. Additionally, this work shows the influence of weather on real-time reliability through the use of weather-dependent failure rate models. As a result, recommendations are made on the practical and gradual transition from present TSO practices towards probabilistic approaches.

Útdráttur

Rekstur raforkukerfa snýst um að tryggja áreiðanlega afhendingu á raforku til þjóðfélagsins. Hið hefðbundna N-1 viðmið hefur til þessa verið notað við að stýra áreiðanleika. Í EU FP7 rannsóknaverkefninu GARPUR, er lagt til að kerfisstjórnendur raforkukerfa byggi frekar á líkindafræðilegum aðferðum og viðmiðum til að ná fram auknum áreiðanleika. Tilmæli um notkun líkindafræðilegra aðferða eru þar sett fram vegna þess að hið hefðbundna N-1 viðmið er ónæmt fyrir breytilegum líkum á utanaðkomandi ógn við raforkukerfið; breytilegum þjóðhagslegum afleiðingum raforkuskorts; og jafnvægi fyrirbyggjandi og leiðréttandi aðgerða, meðal annars.

Sú rannsókn sem lýst er í þessari ritgerð hefur haft það að markmiði að yfirfæra hina fræðilegu hugmyndafræði sem GARPUR byggir á yfir á íslenska raforkukerfið og er þá átt við rekstur kerfisins í rauntíma. Þetta var gert með því að útfæra fyrst fræðilegar hugmyndir GARPUR verkefnisins á fræðilega fullkominn hátt. Þessi útfærsla var síðan einfölduð með því að taka tillit til þess hvaða gögn og verkfæri væru aðgengileg í reynd. Þetta myndaði fyrsta skrefið í hagnýtingu GARPUR hugmyndafræðinnar þar sem þróuð var frumgerð fræðilegs prófunarkerfis.

Með frumgerð fræðilegs prófunarkerfis var myndaður grundvöllur að því að sannprófa GARPUR hugmyndirnar í litlum mælikvarða á íslenska raforkukerfinu. Prófunarkerfið var svo víkkað út fyrir þau gögn og verkfæri sem til eru í dag. Til að fá betri mynd af mikilvægi veðurs í rauntíma áhættumati var veðurháð bilanalíkan þróað. Einnig var útbúið kerfislíkan til að meta áhrif ýmissa forsendna og verndaraðgerða raforkukerfisins.

Kjarni þessarar rannsóknavinnu er sannprófun þess að hægt er að útfæra í reynd aðferðafræði GARPUR verkefnisins sem nothæfa og gagnlega leið til að meta og stjórna áreiðanleika raforkukerfis í rauntíma. Auk þess hefur þessi rannsókn dregið fram áhrif veðurs á áreiðanleika raforkukerfisins í rauntíma með því að nota veðurháð bilanalíkön. Þetta leiðir til þess að settar hafa verið fram ráðleggingar um hagnýta yfirfærslu frá núverandi aðferðum við mat á áreiðanleika og rekstur raforkukerfa yfir í líkindafræðilegar aðferðir.

Samuel Perkin Thesis Committee Final Report

Samuel Perkin has submitted a thesis and presented his defense Friday January 12 2018 in respect of "Real-Time Weather-Dependent Probabilistic Reliability Assessment of the Icelandic Power System".

The candidate's main premises were that:

1. The weather-dependence of faults can and should be taken into account when quantifying power system risk;
2. It is feasible to operate power systems with a quantitative risk-based approach.

Based on the final submitted thesis, the committee believes that there is sufficient original work and evidence of the candidate's ability to independently conduct useful research. The context is well-described in respect of both the motivation and prior work. The main arguments are well presented and supported by suitable evidence.

The impact of real-time, weather-dependent overhead line failure rates on quantified power system risk is shown. Examples of application to the Icelandic power system are given. The work submitted is of a level that is consistent with what the external examiner would consider to be a suitable standard at his own institution, the University of Strathclyde in the UK.

The committee was particularly impressed with the presentation of his work given by the candidate at his defense. A number of good analogies were described that succeeded in communicating the main concepts to non-specialists in the audience.

The candidate's engagement in discussion was good showing a deep understanding with good suggestions for further work. The committee is also encouraged that, given the candidate's success in gaining a position at Landsnet, there is every prospect of much of the suggested further work being taken forward.

Reykjavík January 17 2018

The thesis committee:

Páll Jensson

Dr. Páll Jensson, Professor of Engineering Management, Reykjavik University

Hlynur Stefánsson

Dr. Hlynur Stefánsson, Associate Professor of Engineering Management, Reykjavik University

Ragnar Kristjánsson

Dr. Ragnar Kristjánsson, Assistant Professor of Electrical Engineering, Reykjavik University

1

L Wehenkel

Dr. Louis Wehenkel, Professor of Electrical Engineering and Computer Science, University of Liege, Belgium

The external examiner:

Keith Bell

Dr. Keith Bell, Professor of Electrical Engineering, University of Strathclyde.

Acknowledgements

I would like to thank my three thesis supervisors from Háskólinn í Reykjavík: Páll Jensson, Hlynur Stefánsson and Ragnar Kristjánsson. Their approach and attitude to supervision made the entire process enjoyable, and kept me from going off of the rails on multiple occasions, for which I greatly appreciate their advice and continual support. Similarly I would like to thank Íris Baldursdóttir for trusting me with the opportunity to work with Landsnet on the GARPUR project, and for supporting and encouraging me throughout the process, especially for providing many opportunities to learn, grow and challenge myself with. I also acknowledge the continual support and collaboration from Guðjón Hugberg Björnsson, specifically in suffering all my technical questions and bad humour. Magni Þór Pálsson and Guðlaugur Sigurgeirsson also provided greatly appreciated support in learning about aspects of TSOs beyond System Operation. Many staff at Landsnet (too many to name individually) deserve acknowledgement, in particular everyone in and around the control room, and those others who accepted my presence and entertained my bad Icelandic.

I have many thanks for Louis Wehenkel and Efthymios Karangelos who graciously allowed me to spend time at the University of Liege and to learn from them, without which I would never have wrapped my head around the GARPUR project. I also must thank Kjetil Uhlen, Vijay Venu Vadlamudi and Camille Hamon for inviting me to NTNU and providing great hospitality and collaborations. I also thank Nicolaos Cutululis at DTU who provided a significant amount of historical weather data, which greatly enhanced the pilot test and allowed me to expand the scope of my research. Again, numerous others from the GARPUR project deserve appreciation and acknowledgement for engaging with me in many stimulating conversations, and for being so effortless and enjoyable to work with. I would also like to thank Arne, Trond and Ingrid from Promaps for collaborating with me on a publication.

The emotional support of friends here in Iceland, Australia, and those elsewhere in the world, has been monumental in keeping me sane throughout the PhD process. In particular Stefán, Morgan and Mike whose friendship has been a constant source of stability, as has the long-distance support of my family for encouraging me, believing in me and mostly for understanding me. Lastly, I would like to thank dear Glódís for reminding me that there's a person somewhere inside every engineer, and her extended family for being the kind people that they naturally are.

The research leading to this PhD dissertation has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 608540, and further support from Landsnet.

"INSUFFICIENT DATA FOR A MEANINGFUL ANSWER"
-Isaac Asimov, The Last Question

Publications

This PhD dissertation is based upon the journal publications and conference papers listed below, and referred to throughout the text by their Roman numerals (I-V). Note that these papers do not appear in the online version of this thesis.

I. **Framework for threat based failure rates in transmission system operation**
Samuel Perkin, Gudjon Hugberg Bjornsson, Magni Palsson, Iris Baldursdottir, Hlynur Stefansson, Ragnar Kristjansson, Pall Jensson, Efthmyios Karangelos and Louis Wehenkel
Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO), pp. 150-158, IEEE (2016)

II. **Modelling weather dependence in online reliability assessment of power systems**
Samuel Perkin, Arne Brufladt Svendsen, Trond Tollefsen, Ingrid Honve, Iris Baldursdottir, Hlynur Stefansson, Ragnar Kristjansson, and Pall Jensson
Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, Vol 231, Issue 4, pp. 364-372 (2017)

III. **Framework for trajectory-based security assessment of power systems**
Samuel Perkin, Camille Hamon, Ragnar Kristjansson, Hlynur Stefansson, and Pall Jensson
Submitted to IET Generation, Transmission & Distribution

IV. **Near real-life pilot testing of real-time probabilistic reliability assessments**
Samuel Perkin, Ragnar Kristjansson, Hlynur Stefansson, Pall Jensson, Gudjon Hugberg Bjornsson, Iris Baldursdottir, Magni Palsson, Efthmyios Karangelos and Louis Wehenkel
International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), IEEE (2016)

V. **Case-study of real-time probabilistic reliability assessment on the Icelandic power system**
Samuel Perkin, Gudjon Hugberg Bjornsson, Camille Hamon, Efthymios Karangelos, Ragnar Kristjansson, Hlynur Stefansson, and Pall Jensson
Submitted to International Journal of Electrical Power & Energy Systems

Other dissemination activities

Other Conference Papers:

- **On improving data and models on corrective control failures for use in probabilistic reliability management**
  Vijay Venu Vadlamudi, Camille Hamon, Gerd Kjolle and Samuel Perkin
  International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), IEEE (2016)
  (minor paper contributions, presentation of results at conference)

Other Presentations:

- *The Garpur Project: moving beyond N-1 for real-time security assessment*
  Samorka conference
  in Akureyri, Iceland on 5 May 2017

- *Case-study of probabilistic risk assessment on the Icelandic Power System*
  14th Electric Power Control Centers (EPCC) Workshop
  in Weisloch, Germany on 14 May 2017

- *Application of RMAC to real time operation: case study at Landsnet*
  Section of GARPUR Training Webinar 1
  recorded live in Reykjavík, Iceland on 23 May 2017

- *Comparing reliability assessment in Transmission System Operation with the STAMP model*
  European STAMP (Systems-Theoretic Accident Model and Processes) Workshop
  in Reykjavík, Iceland on 15 September 2017

GARPUR Technical Report Contributions:

- D2.2 [1]: minor contributions
- D5.1 [2]: minor contributions
- D6.1 [3]: major contributions to Chapter 4
- D6.2 : major contributions to Chapters 3.3, 4.3 and 5
- D8.1: major contributions to Chapter 3.3
- D8.3: major contributions to Chapter 3

Declaration of contribution

Paper I - Samuel Perkin made the simulations/calculations, wrote the manuscript, created the figures, and presented the paper.

Paper II - Samuel Perkin designed the study, performed the failure rate calculations, wrote the manuscript, created the figures, and corresponded with the journal.

Paper III - Samuel Perkin designed the study, made the simulations/calculations, wrote the majority of the manuscript, created the figures, and corresponds with the journal.

Paper IV - Samuel Perkin wrote the manuscript, created the figures, and presented the paper.

Paper V - Samuel Perkin designed and programmed the majority of the pilot test, performed the simulations/calculations, created the figures, wrote the manuscript and corresponds with the journal.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Modern society depends on reliable access to electricity to function, and therefore depends upon a reliable power system. This dependence ranges from the electrification of critical transport and communication infrastructure and industries, to the growing connectedness of individuals and businesses via cloud-based services. This is most apparent from goal 7 of the United Nations Sustainable Development Goals [4], aiming for universal access to energy services by 2030.

Similarly, the real consequences of blackouts [5, 6] and subsequent public concern imply a growing reliance on power systems and vulnerability to their failure, and an expectation of high reliability. Transmission System Operators (TSOs), and power system regulators, aim to balance the expectation of reliability with the real cost of developing, maintaining and operating a power system. Historically this reliability has been managed through the application of the N-1 principle, which is defined by [7] as:

> *The power system is designed and operated to withstand a set of contingencies referred to as "normal contingencies" selected on the basis that they have a significant likelihood of occurrence. In practice, they are usually defined as the loss of any single element in a power system either spontaneously or preceded by a single-, double-, or three-phase fault. This is usually referred to as the criterion because it examines the behaviour of an N-component grid following the loss of any one of its major components.*

A similar definition is given in the ENTSOe handbook on operational security [8], but also extending to the requirement that events do not cascade across cross-border connections. This N-1 principle is presently applied to various aspects of TSO decision-making, ranging from the decision to invest in new infrastructure, to the planning of maintenance activities, and to how control room operators manage the grid in real-time. Through this, the power system is designed and operated to ensure that no single unplanned outage will interrupt supply or degrade power quality.

The European Union Framework 7 research project "GARPUR", which this dissertation is a part of, was initiated to determine if better approaches to reliability management are possible in a European context[9]. Specifically the project aimed to develop new generally acceptable approaches to reliability management that handle uncertainty modelling and are based upon probabilistic risk assessments.

The term risk is frequently used in this dissertation, and therefore it is useful to clearly define it here. In simple terms, risk is defined for a single event as its probability multiplied by its consequence. It can therefore be thought of as a probability-weighted consequence. Adapting the definition of risk provided by [10], the equation below provides a mathematical definition:

$$R = \sum_{\boldsymbol{\xi} \in \Xi} \mathbb{P}(\boldsymbol{\xi})\, C(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\xi}) \tag{1.1}$$

where:

R	: risk [$\mathbb{E}(\$)$]
$\boldsymbol{x}$	: system state variables [-]
$\boldsymbol{u}$	: controlled inputs to the system [-]
$\boldsymbol{\xi}$	: uncontrolled inputs to the system [-]
Ξ	: uncertainty space of uncontrolled inputs [-]
$\mathbb{P}(\boldsymbol{\xi})$	: probability of a realisation of uncontrolled inputs [-]
$C(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\xi})$	: consequence of $\mathbf{x}$ with inputs $\mathbf{u}$ and $\boldsymbol{\xi}$ [\$]

where the units of consequence and risk are arbitrarily shown in units of dollars. Uncontrolled inputs include changes in load, generation or component operational states (due to faults, for example), as well as variables describing the exogenous environment. In reality, the controlled inputs to the systems are broken down into preventive and corrective controls. Preventive controls are those which are made in anticipation of some uncontrolled inputs to the system, whilst corrective controls are apply after some realisation of uncontrolled inputs. Consequence functions may measure monetary impact (loss), lost energy (MWh), lost power (MW), a binary measure of whether the outcome is acceptable, or any other function that is meaningful for the given context.

1.1 Research focus

This research, as part of the GARPUR project, focuses upon the application of the GARPUR framework to real-time operation of power systems. Further, this research focuses on reliability assessment, rather than the extended reliability control problem. This choice of focus is based upon the need to first propose tractable and practical approaches to reliability assessment, that TSOs may possibly use and accept, prior to using the approach for optimising their control actions. For a TSO to accept any such assessment method, it must also adequately model the main risks to the transmission system. Therefore, a major focus of this research is on implementing and testing weather-dependent failure rate models within a probabilistic reliability assessment. In order to ensure the practicality of the proposed reliability assessment implementation, the method is then tested in a pilot test on the Icelandic transmission system using operational data.

Generally this work is based upon the GARPUR framework proposed within Work Package 2 of the GARPUR project [1]. It also builds heavily on the author's collaborative work in Work Package 6 of the GARPUR project on describing general TSO processes in a short-term and real-time context [3] and on formulating algorithms to apply the GARPUR framework in real-time operation [11]. Some parts of the work in this dissertation are derived from the author's prior work in internal and restricted deliverables, reports and memos within the GARPUR project.

1.2 Motivation

The primary motive for undertaking this research is that power systems, a socially and economically important piece of infrastructure, are undergoing significant change. Power systems are affected by changes in the generation mix (more intermittent generation), in how electricity is used (increased transport electrification), how they are operated (increasing automation and adoption of phasor measurement units), and how they are viewed by society (environmental impacts and reliability). These changes require a shift in how the industry considers reliability, how decisions are justified for the sake of reliability, and how this is communicated with stakeholders.

Society expects perfect reliability and more renewable energy, yet at the same time opposes the infrastructure expansion and associated cost increases required to facilitate near-perfect reliability with increasing proportions of non-dispatchable power generators. It is not possible to determine the optimal balance of reliability and the socio-cost of infrastructure when armed only with the N-1 criterion, particularly under uncertainty. In the Icelandic case, although the reliability of the system has been above 99.99% for 7 of the last 8 years [12], the N-1 criterion is not achieved at all times. This is mainly due to the presence of radial lines, but also due to maintenance activities that cannot be performed whilst still ensuring the N-1 criterion is satisfied. Decision making often considers the vulnerability of the system to weather, which constitutes a great source of uncertainty in Iceland. Therefore, the motivation is to find an approach to assessing reliability that can assess the changes that the Icelandic power system is experiencing, can quantify the benefit of recent improvements to the operation of the Icelandic transmission system, and can better inform future debates on reliability.

1.3 Aims and objectives

The research focus and motivation lead to the following two hypotheses:

1. The weather-dependence of faults can and should be taken into account when quantifying power system risk.

2. It is feasible to operate power systems with a quantitative risk-based approach;

To address these hypotheses, the following research objectives were defined:

i. Review the state of the art in power system operation and reliability management

ii. Study the importance of including weather-dependence in failure rate models

iii. Determine a modelling architecture to achieve the GARPUR framework in real-time reliability assessment

iv. Build, perform and validate a pilot test of the method on the Icelandic System

v. Derive insights from the pilot test for the future implementation of probabilistic reliability assessment methods

This research makes an original contribution to the existing work on the topic of TSO reliability management, as it converts a novel probabilistic reliability framework into a real-world implementation, which is then tested in an operational setting. This work develops an innovative approach to probabilistic reliability assessment, then through its application provides practical and meaningful feedback to the academic community, highlighting practical barriers, challenges, and topics requiring further research. It also provides insight into the present day tractability of the GARPUR approach on real-world power systems, which is valuable to those who intend to continue developing probabilistic reliability assessment methods for operational use.

1.4 Outline of thesis

This thesis is based upon a number of published journal and conference papers, which constitute the main content of the research. This document provides context and support to these papers. The Background chapter provides a review of literature and concepts that are central to the research. This is followed by a chapter on methods and results. In this chapter, each section summarises one or two papers on a particular topic of study, describing the general method and the key results in the broader context of this dissertation. Extended results, beyond what was included in the papers, are also included where relevant. The three sections of this chapter relate to the modelling of threat-based failure rates, the formulation of a probabilistic reliability assessment implementation, and finally to the pilot testing of this assessment implementation on the Icelandic power system. The final chapter concludes the research, and provides recommendations on further study and future work.

Chapter 2

Background

This chapter provides a literature review of papers, reports and books on topics relevant to this research. Additional, more detailed backgrounds can be found within each of the papers associated with this thesis.

2.1 Challenges for modern power systems

Modern power systems are complex interconnected networks of generators and loads, with a wide variety of components in between to manage the reliability, power quality, power balance, and stability of the network. The primary goal of power system operation is to maintain service to all end-users. Transmission system operators must also ensure that the power provided is of a particular quality (often defined by regulations), that the system is stable, and that disturbances will not result in blackouts.

The power system must also be operated to ensure that all connected components are operated within their specified limits. Such limits are often expressed in terms of real power loading limits, over/under voltage, or over/under frequency. These limits are often proxies for real physical limits, for example that the insulation of an underground cable will begin to deteriorate above some temperature, or that an overhead transmission line will sag too close to grounded structures as its temperature increases. Given that there is not a perfect linear relationship between loading and component temperature (e.g. due to wind cooling, solar irradiation, etc.), there is often a factor of safety built into these loading limits. Operating components outside of their limits may either cause component failure, short circuits, or activation of system protection systems, which can lead to a cascading outage.

Often the terms *disturbance*, *outage*, *failure*, and *fault* are used when discussing power system reliability, and are used extensively throughout this thesis. These terms are defined by [13] as:

> **Disturbance:** An unplanned event that may cause the transmission system to divert from the normal state.
>
> **(Forced) Outage:** The unplanned removal from service of a relevant asset for any urgent reason that is not under the operational control of the operator of the concerned relevant asset.
>
> **Faults:** All types of short-circuits (single-, double-, and triple-phase, with and without earth contact), a broken conductor, interrupted circuit, or an intermittent connection, resulting in the permanent non-availability of the affected transmission element.

Notably the term *failure* is not defined in [13]. The usage of the term *failure* throughout this thesis is equivalent to the unplanned event that results in a disturbance, and is inclusive of faults, as defined above.

The challenges of managing power system stability are defined and discussed in detail by [14]. The real-time interaction of power system components, loads and generators can lead to instabilities of voltage (magnitude and angle), frequency, transient power flows, and of oscillatory instabilities. This complex interaction, especially when considering the timing of system protection and automated control devices, can lead to unpredictable system responses to disturbances.

This complexity is most evident by investigating why and how blackouts occur [6]. Interestingly, blackouts tend to occur due to a single disturbance, which seems to be at odds with the N-1 criterion. As noted in [15], a single disturbance is often followed by a short period in which the system remains intact. In such a period, transient, small signal or voltage instabilities may exist, and if left undiagnosed and untreated, may cause triggering of system protection or loss of system stability, causing a cascading outage. Cascading may occur due to hidden failures on components or system protection devices, or when the grid is already weakened by either ongoing maintenance or due to a prior disturbance (multiple outages may be credible during storms, for example). In some cases, a cascading outage may be triggered by the complex interaction of various control devices and protection schemes. For example, the droop response of generators following a disturbance may cause the power flow on some underground cable to exceed its limits, causing the correct tripping of protection devices. It is likely that as power system protection becomes more sophisticated (e.g. Wide Area Control Systems) that the potential for undesirable interactions between protection devices will increase.

This challenge is exacerbated by the presence of electricity markets. Electricity markets typically organise the commitment of generator units to produce real power during specified operational periods, typically single hour blocks. Markets however are settled ahead of operation, on expectations of demand and of the output of variable generators (e.g. wind and solar power plants). Given that TSOs must have the ability to balance load and generation in all instances, they require reserve generation capacity.

As detailed in [16], reserves are defined as either Frequency Containment Reserves (FCR), Frequency Restoration Reserves (FRR) or Replacement Reserves (RR). Note that FRR is split into reserves which are activated automatically and those which are activated manually. FCR is bought as ancillary services, and rapidly respond to changes in system frequency to reduce the frequency peak or nadir. FRR may also be bought as ancillary services, or through activation of the balancing market, and is used to return the frequency to nominal levels. Finally, RR uses the balancing market to replace any automatically activated reserves with generation on the balancing market, such that FCR and FRR can continue to act as reserves.

In normal operation reserve generators and controllable power system components (e.g. transformers, circuit breakers, capacitor banks, etc.) represent the main controls that transmission system operators have in ensuring that the power balance is maintained. Following a disturbance, operators may also operate other non-reserve generators, and even loads, in order to maintain system stability following a disturbance. A number of automation devices are also used by TSOs in order to maintain stability, particularly in managing transient instabilities which may propagate faster than humans can react. This automation primarily consists of Automatic Generator Controllers (AGCs), Power System Stabilizers (PSSs), distance relays, System Integrity Protection Schemes (SIPS), and more recently demand-side response and Wide Area Control Systems (WACS). These protection devices range in scope from monitoring and acting upon single components to automated actions on numerous devices due to regional instabilities.

Beyond the scope of operation, the reliability of the power system is managed through system development and asset management. The grid infrastructure must be constantly assessed and new infrastructure must be planned in the long-term and constructed ahead of time, in anticipation of changing power system demand and generator availability. Existing infrastructure and spare assets must also be constantly maintained and/or replaced, to ensure that components maintain operability and do not degrade into failed states, as well as ensuring that any failed component can be quickly repaired or replaced. The decisions made in these contexts can greatly define the operational limits in which reliability is managed in real-time - that is, ensuring the system is adequate for future operation.

2.1.1 The Icelandic power system

As detailed in [12], the average power production on the Icelandic transmission system is 2073 MW, with a maximum of 2291 MW in 2016 (the maximum load was 2241 MW). In the same year, 85 unexpected component outages occurred, of which 37 resulted in lost load. The total lost load during the year was 170 MWh of primary load, and 724 MWh of curtailable load, together amounting to approximately 0.005% of the total annual energy demand. Curtailable loads are a subset of end-users that allow Landsnet to shed their demand during disturbances, accepting a lower reliability level than primary loads.

The reliability of the Icelandic system is commonly reported in terms of System Minutes [17], calculated as:

$$SMS = \frac{\sum E_i}{E_{all}}(8760)(60), [min/yr] \tag{2.1}$$

where:

E_i : curtailed load during the ith disturbance [MWh]
E_{all} : total annual energy delivered to consumers [MWh]

The lost load statistics above can be represented in SMS as 5 minutes for primary load and 21 minutes for curtailable load. For reference, Landsnet aims to ensure SMS is below 50 minutes each year, as stated in [12]. This target was met in each year from 2013 to 2016, but not in 2012 due to a contingency that resulted in a significant loss of load. Additionally, the power system in Iceland is governed by Icelandic Energy Regulation 1048/2004 [17], which contains operational constraints related to frequency and voltage, among others. Notably, the regulation does not mention the N-1 criterion, but refers to article 28 of the Electricity Act 65/2003 [18] which states that Landsnet (among other stakeholders) shall establish internal controls on power system security and reliability, subject to supervision by the National Energy Authority and to any additional requirements set forth in regulation.

There are a number of challenges related to operating the Icelandic power system, shown in Figure 2.1. Notably, the grid consists of two 220 kV regions (one in the South-West, one in the East) which are connected by a long distance 132 kV ring, within a single synchronous area. The majority of power generation and consumption on the power system is contained in these two areas, with a few power plants and demand centres located on radial connections along the ring, or on lower voltage connections to the 220 kV regions.

The generation capacity on the power system is mostly hydro power and geothermal power, with a small number of wind turbines and diesel generators. Despite this large proportion of spinning machines, the system has quite low inertia. Low system inertia results in sudden disturbances (loss of connected loads or generators) having a relatively large Rate Of Change Of Frequency (ROCOF). That is, the time between the initial disturbance and an excursion of frequency outside of operational limits is relatively short. The calculation of ROCOF is defined in [19] as:

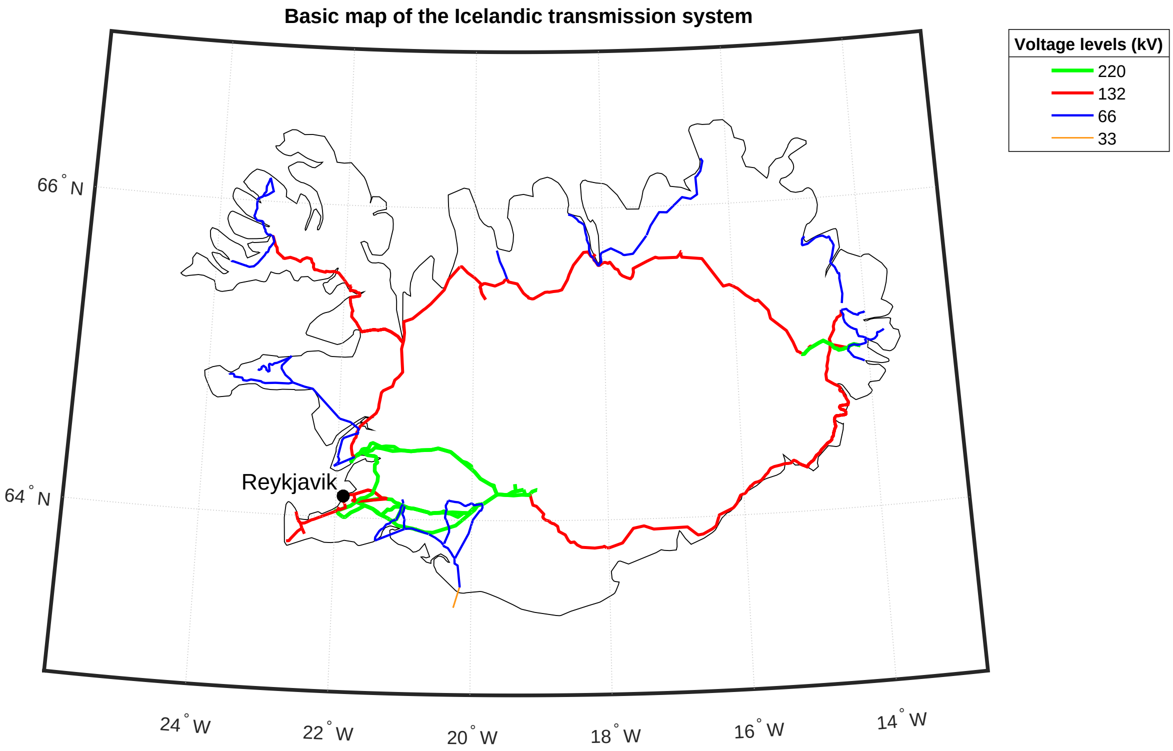


Figure 2.1: High voltage transmission lines on the Icelandic power system, as of 2016, showing the two separate 220 kV areas joined by a 132 kV ring

$$\left.\frac{d\Delta f}{dt}\right|_{t=0^+} = \frac{f^0 P_k}{2\sum_{i=1,i\neq k}^{N} H_i S_i} \tag{2.2}$$

where:

f^0	: nominal system frequency [Hz]
Δf	: change in frequency from nominal value [Hz]
t	: time since initial disturbance at t=0 [s]
$i \in [1, \ldots, N]$	: index of synchronous machine [-]
$k \in [1, \ldots, N]$	: index of lost synchronous machine [-]
P_k	: lost load or generation [W]
H_i	: Inertia constant of machine i [W/VA]
S_i	: Apparent power rating of machine i [VA]

The connection between ROCOF, system protection, and inertia is best shown by Figure 2.2. System protection devices and their associated measurement devices have some delay before any automatic control action can occur (note not all devices have the same reaction time). Within this period, shown on the figure as reaction time, the ROCOF will drive the frequency towards some operational and/or regulatory limits. As system inertia decreases, the change in frequency prior to any reaction of the system will be greater. This then requires faster reacting ancillary services and system integrity protection schemes, and hence also requires that protection triggers on shorter measurement periods (e.g. calculating ROCOF over 100ms rather than 200ms). Notably the ROCOF is also dependent upon the size of the initial disturbance, and so protection devices are normally designed to react to dimensioning disturbances (e.g.

large N-1 contingencies). As such, the reliability of the Icelandic power system is heavily dependent upon automation.

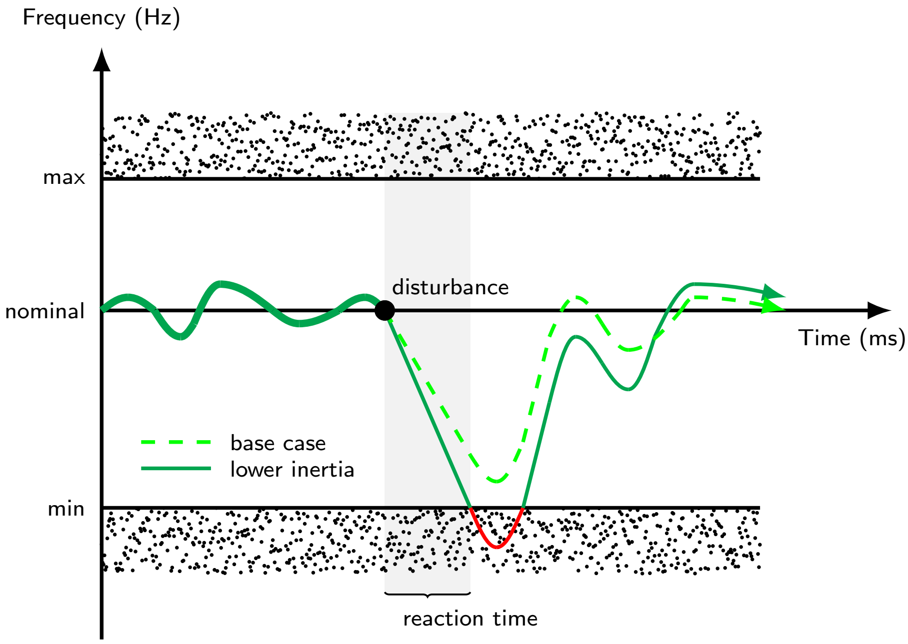


Figure 2.2: Impact of inertia on frequency following a disturbance, for a system with a fixed reaction time

The reliability of the system is also governed by the presence of radial connections. Isolated demand centres, typically fishing villages and small municipalities, are serviced by long distance transmission lines (132 kV or 66 kV). Such connections are not N-1 secure in transmission, and as such are islanded whenever there is a fault on the radial connection. These radial lines are often located in areas that are subject to extreme weather, and therefore fail relatively frequently due to storms. The reliability of these locations is commonly managed through SIPS installations and the presence of back-up diesel generation.

The main threats to the Icelandic power system are wind, snow and icing, as discussed in [20]. Such threats often cause multiple concurrent component outages, and also cause outages with a long restoration time due to the extent of physical damage and difficult travelling and working conditions. Changes in the system state due to maintenance activities and sudden load disconnections also greatly affect the system reliability. The gradual deterioration of grid components, due to use in operation, and due to weather exposure, also presents a threat to the system.

2.1.2 Comparison to other European systems

Given the lack of solar power plants, the small number of wind turbines, and the presence of substantial hydropower resources in Iceland, there is not much uncertainty in generation in the short-term. Further, the significant proportion of stable heavy industrial load on the power system (approximately 80%) means that there is also relatively little uncertainty in demand. The variable residential, commercial and industrial demands only represent 20% of the demand. This lack of variability in demand is not necessarily common in Europe, given that there is no uncertainty related to cross-border connections or to intermittent generators in Iceland. The presence of large loads and large generator plants however represents a significant threat to the power system, where an individual disturbance may result in the loss of 25% of the power system's load or generation. Disturbances of such proportions are uncommon in larger transmission systems.

A year long study of frequency disturbances in the pan-European study by [21] assessed the impact of similar sized disturbances (300 to 1500 MW). The study calculated that the ROCOF of the pan-European system ranged between 0.0051 to 0.0132 Hz/s. By contrast, the ROCOF of the Icelandic system during large disturbances has exceeded values of 1 Hz/s. Similar large ROCOF values may be found in small transmission systems with minimal cross-border connections, or island systems with a large proportion of wind and solar power production, such as on the Irish transmission system [22].

ENTSO-E releases an annual report which compares the disturbance statistics of the Nordic and Baltic power systems, including Iceland [23]. Compared to other countries, Iceland experiences far fewer disturbances due to lightning (<5% in Iceland, compared to 10-15% in Sweden and Norway). Yet Iceland experiences a greater proportion of disturbances due to other environmental causes (70% in Iceland, compared to 45% in Norway), which is primarily due to wind, ice and snow.

2.2 Transmission system operation

TSOs perform a number of parallel and inter-dependent processes to ensure continual supply of electricity. These processes are summarised in Figure 2.3. In order to guarantee the reliable operation of the transmission system in real-time, decisions must be made days, months and years ahead of time to ensure that the system has adequate supply and transmission capacity, and that it can be operated under maintenance.

The three main contexts of TSO operations are highlighted on the figure. Broadly, system development is concerned with system adequacy, asset management is concerned with the replacement of components and timing of maintenance/construction activities, and system operation is concerned with ensuring security of the power system and managing disturbances. An over-arching approach to reliability management is applied to all of these contexts through the development of policies, which are constrained and guided by regulations and laws.

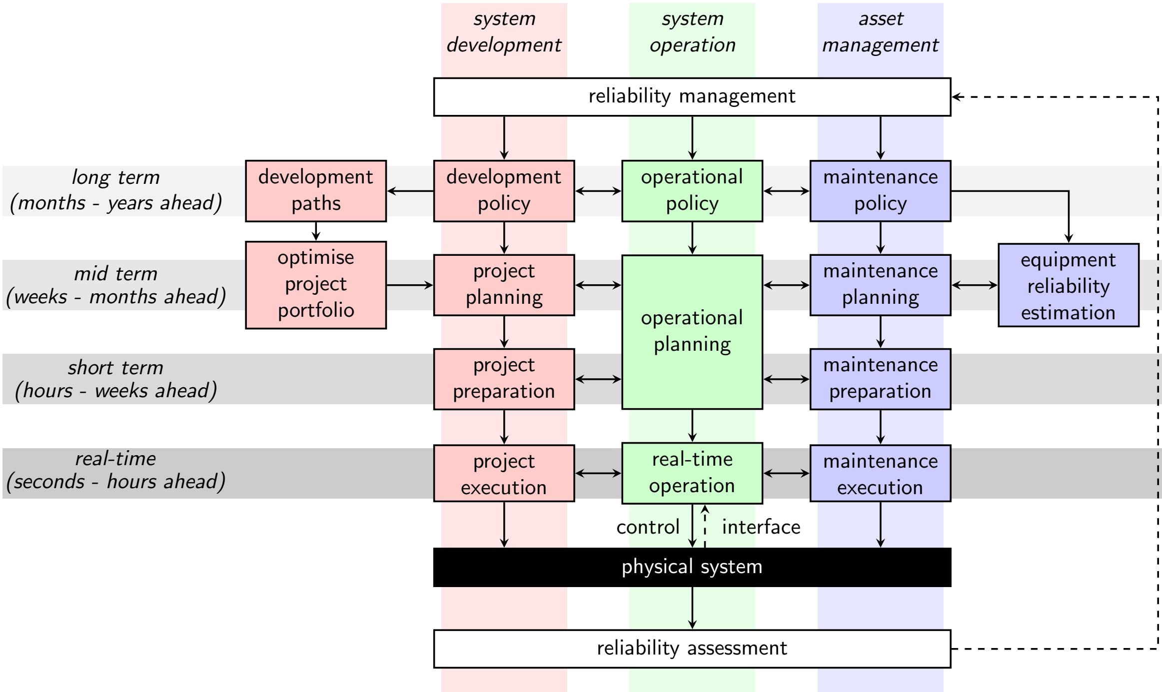


Figure 2.3: Simple organisational chart of TSO tasks, adapted from [3]

The actual organisational structure of TSOs varies from this general diagram, but functions similar to those shown are likely to exist at all TSOs. For example, Figure 2.4 below, adapted from [24], provides a hierarchical control structure for the Landsnet. This alternate high-level interpretation shows TSO processes and stakeholder interactions specific to the Icelandic case, where the blue shaded region is equivalent to the tasks shown in Figure 2.3.

Although reliability management is embedded in the foundation of all TSO processes, this dissertation focuses on the context of system operation. Of particular interest is real-time system control, but operational policy and operational planning will be briefly discussed given that they provide the context within which system control occurs.

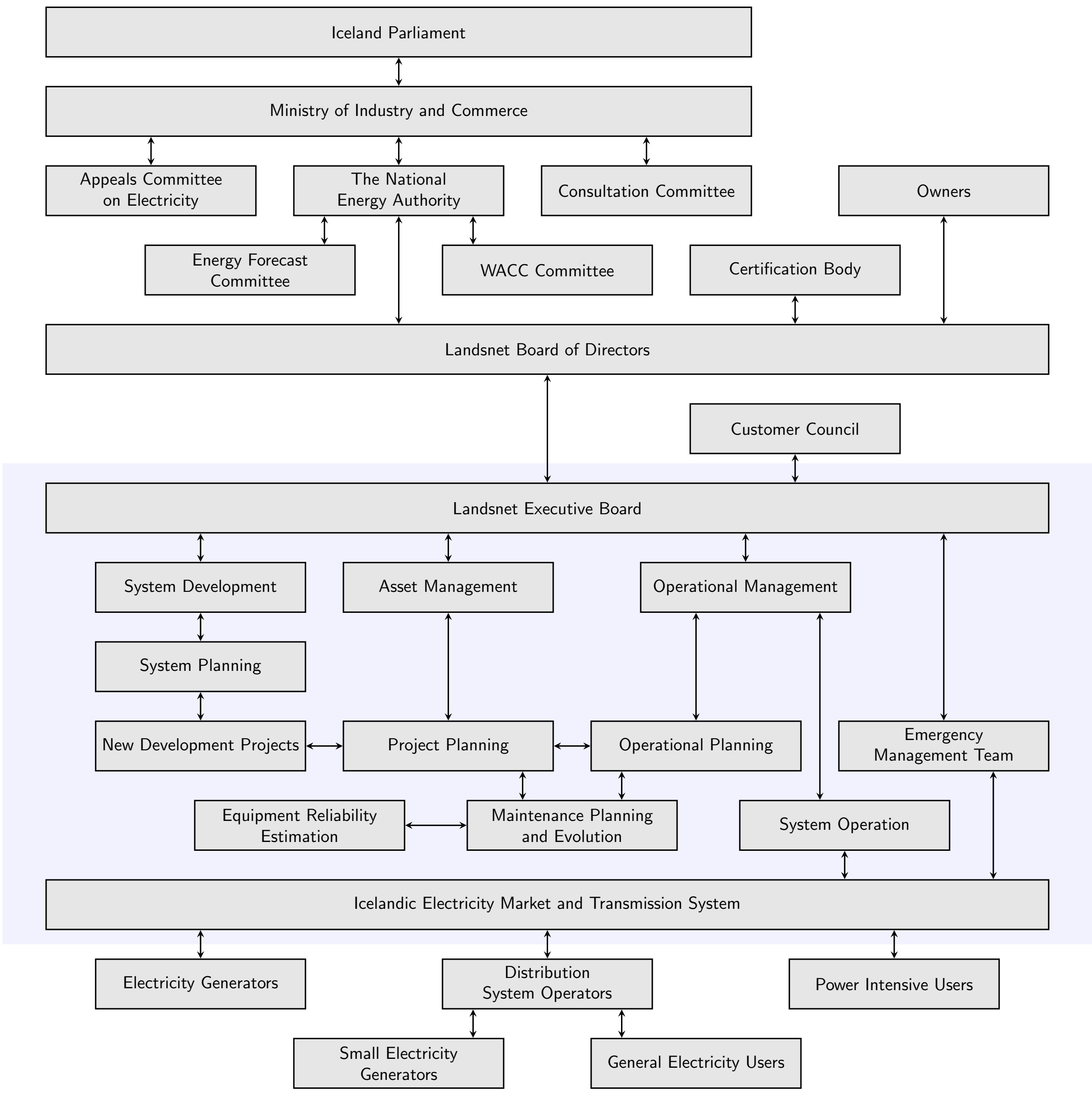


Figure 2.4: High-level organisational chart of Landsnet, adapted from [24]

2.2.1 Operational policy and planning

Operational policies are defined in [3] as, *"... the doctrine developed for interconnected system operation, forming the main part of the ENTSOe operational handbook [8]. Each doctrine consists of criteria, standards, requirements, guides, and instructions, and applies to all control areas"*. These doctrines dictate the expectations and constraints on actions in the context of operational planning and system control. As noted in Section 2.1.1, the Icelandic TSO is responsible for setting internal targets and controls to ensure a reliable power system. Such targets are set within the context of operational policy, and policies are used to ensure subsequent system operation tasks result in compliance with these targets.

Similarly, operational planning is defined in [3] as, *"... the group of reliability management activities linked to system optimisation ahead of real-time operation, within the short-term and mid-term horizons"*. In the mid-term horizon, this mainly relates to checking the operational security given expected component and generator outage plans. In the short-term this relates to further refinement of outage plans, congestion forecasting, ensuring adequate reserve generation availability, preparation of preventive control measures, and system protection management, among others. Broadly, the objective of these activities is to anticipate and manage the reliability of the system ahead of time, and to ensure that sufficient control opportunities are available for real-time control.

2.2.2 Real-time operation

Real-time system operation is concerned with ensuring that the power system remains within operational security limits through corrective and preventive control measures, anticipating and managing threats to the system, and managing planned outages. The specific tasks and control actions available are reviewed and elaborated in detail within [3]. At a high-level, real-time operation can be considered as a control loop between human operators and the physical system, where the operator views the system state via an Energy Management System (EMS).

An EMS is a suite of tools that are connected to SCADA and other sensing and communication devices, that display near real-time information on the power system. The primary tool within an EMS is the state estimator, which displays voltage, power flows, topology, breaker positions, loads, generation and other steady-state power system variables. Additional tools commonly relate to anticipation of future system states (e.g. load forecasting and market tools), identifying dynamic instability, operational alarms, and contingency assessment. In some cases, the EMS may also include real-time measurements from phasor measurement units (PMUs) which allows real-time system frequency to be observed, as well as higher resolution displays of the previously mentioned system variables.

2.3 System modelling

Given the complexity of power systems, it is not trivial to predict how disturbances or state changes will affect the power system by human inspection. Any tool that assesses how the transmission system will react to changes, either in component states or to a change in load or generation, must model the physical transmission system. Power systems are complex, and therefore a number of models of varying accuracy and simulation speed exist for a variety of purposes.

The most fundamental power system simulation model is the Power Flow (PF) calculation, described in [25] for example. A PF calculation requires data related to the state and connectivity of power system components, generators and loads, as well as their electrical parameters (e.g. impedance of transmission lines, power output of generators, etc.). The calculation itself solves for the voltage magnitude and angle and real and reactive power injections at all nodes of the transmission system. From this solution it is possible to determine power flows between nodes. PF calculations are often solved using Newton-Rhapson methods when considering AC power flow, although other techniques exist. Further, PF computation can be sped up by only considering real power flows, as if the power system were a DC network, but such a solution comes at the cost of accuracy. This loss of accuracy is due to the assumptions that: the system has zero conductance; there is a small difference in voltage angles; voltage magnitudes are negligibly different and can be set to values of 1.0 per unit; and that reactive power flow is of negligible importance when considering component overloads. These assumptions result in a linear power flow problem, which is significantly faster and useful for coarse simulations of large power systems.

An algorithm which aims to optimise the output of generators given some load, generator cost curves, and congestion constraints, is called an Optimal Power Flow (OPF) algorithm. OPFs can also be extended to allow additional controlled variables, such as the switching of circuit breakers, the operational state of generators, tap changer positions, among other control actions. Such extensions will be discussed further in sub-section 2.4.3 on reliability control.

In order to assess the dynamic stability of the power system following disturbances, as detailed in [14], time-domain models are required that extend the steady-state algebraic equations to a system of differential algebraic equations that describe the temporal evolution of the system. The formulation and computation of such models is elaborated in [26]. Such an approach requires additional data related to the dynamic behaviour of generators, their governors and excitation systems, as well as dynamic power system components (e.g. dynamic-var compensators), and system protection devices. By considering power system dynamics, such approaches provide greater accuracy and more detailed descriptions of how the system may respond to disturbances or changes. However, as noted in [6] using an example of the 1996 blackout in the North America Western-Interconnected system, modelling errors may incorrectly portray an operational state as secure.

Power system models and data are generally categorised as either bus-branch or node-breaker models. Both model types represent the power system as a graph, consisting of edges and vertices. In a bus-branch model, the vertices represent aggregated nodes on the power system (e.g. a substation, or a power plant) where power injections may be present, and edges that represent connections between nodes. The nodes and edges in a bus-branch model are aggregate models of multiple components. For example, the branch may be an abstraction of a transmission line, switching equipment, a transformer, and some cable segments. In a node-breaker model, the vertices represent discrete components (primarily, both buses and breakers), whilst the edges represent branches that consist only of transmission lines or cables. As argued in [27], there is a real need to migrate to node-breaker models in practice in order to achieve realistic modelling results. They argue that traditional bus-branch models, "ignore the substation breaker configuration and thus limit the assessment of the substation equipment's impact on system reliability".

In spite of this, a bus-branch model of the Icelandic power system was created and used for this research, based on existing bus-branch models at Landsnet. To provide some context, this model consists of 85 buses and 107 branches, whilst the largest public model of the French power system [28] contains 6515 buses and 9037 branches. This context must be kept in mind when considering the remainder of this dissertation, as the specific approaches chosen here are tailored to a small transmission system, and may not be tractable in a suitable time-frame for larger transmission systems.

The following subsections elaborate on the modelling of system response, system restoration, and interruption costs.

2.3.1 System response

Cascading failure tools, also known as models of system response, aim to predict how the physical transmission system will react to some disturbance or state change. Such models aim to simulate the physical response of the system with the purpose of estimating loss of load, capturing all additional failures and processes resulting from the initial disturbance.

A recent review of system response models [29] provides a comprehensive discussion on the different approaches, their validation and needs for future research. It states the importance of validation through testing on real-world systems, and notes that there is no clear consensus in literature on approaches, or on actionable outputs from such models.

A number of system response models were reviewed throughout the course of the research described in this dissertation [30, 31, 32, 33, 34, 35], but the so-called Manchester Model [36, 37] was chosen as an appropriate basis for developing a system response model specific to the Icelandic power system.

2.3.2 System restoration

Models of power system restoration broadly cover the process of restoring a degraded system state (following a disturbance and system response) back to a normal operational state [38]. The purpose of such models can vary from estimating the duration of service outages (such that predictions of lost load can be converted to energy not served) to analysing and optimising restorative control actions. As discussed by [39], overhead line restoration times often have a heavy-tailed distribution for TSOs, which matches Icelandic data covering 2004 to 2015, shown in Figure 2.5. Heavy-tailed distributions are those which have tails that are not exponentially bounded. In this particular case, this refers to the right-side tail of the distribution extending to the right rather than tapering off exponentially. A study by [40] highlights the importance of considering system load, as well as additional disturbances, during the restoration process.

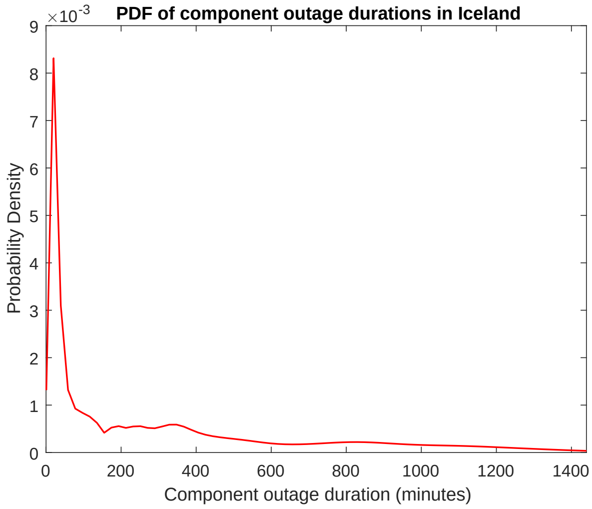


Figure 2.5: Probability Density Function for component restoration times for overhead lines in Iceland from 2004 to 2015. Note that restoration times greater than 5 days are not shown for visibility.

A number of papers on system restoration models for TSOs have been produced. For example, [41] studies the importance of the order in which restorative actions occur. Such approaches aim to search for optimal ordering of events, and can therefore be computationally expensive. Computation of energy not served (ENS), for the purpose of reliability studies, requires fast algorithms that are capable of predicting restoration times for a large number of degraded system states. Heuristic models that relate restoration time to disturbance size, such as in [42], provide a fast alternative to approaches that model specific restorative actions.

2.3.3 Socio-economic models

When considering the impact of service interruptions, it is useful to consider stakeholders as part of the power system. Models of socio-economic impact aim to convert ENS into real costs to stakeholders, as well as modelling the impact of state changes on the environment, generator owners, distribution system operators, and on TSO operational costs. A general framework for assessing the economic impact was developed within the GARPUR project [43], that considers all of these factors. In the case of assessing very large disturbances, given the lack of data difficulty in predicting consequences for such events, methods such as the 'disutility criterion', discussed in [44], may provide ways of including the TSOs aversion to disasters.

The formulation and validation of such methods was not studied in detail for this dissertation, in particular due to a shortage of economic data relevant to the Icelandic power system. The START group in Iceland produces annual reports on the cost of service outages [45]. Such data is sufficient for building value of lost load (VOLL) curves, that are dependent upon the outage duration. Given some assumptions on the demographic composition of load at various nodes on the transmission system, it is then possible to use these VOLL curves to convert ENS into interruption cost estimates.

2.4 Reliability assessment

The concept of using probabilistic approaches to evaluate power system reliability is not new [46, 47]. Numerous books have also been written on the topic of system reliability theory [48, 49] as well as on transmission system reliability [50, 51]. Despite this history of research into probabilistic methods, they are still not widely adopted in practice. The following subsections review the current state-of-the-art for reliability assessment, discussing both deterministic methods and probabilistic methods. Extensions to reliability control are also discussed for completeness, however control was not within the scope of this dissertation.

2.4.1 Deterministic methods

As noted in Chapter 1, the N-1 principle is embedded in in the present reliability assessment methods of TSOs, and applied in all contexts and processes shown in Figure 2.3. In Europe, network security is governed by ENTSOe network codes [52]. A report comparing international standards [53] found that generally all electricity markets require a standard of N-1 reliability or greater, and apply deterministic rules. Similarly, a review of current practices within the GARPUR project [54] also found that the N-1 criterion is applied by TSOs with some variation in how it is interpreted. The main variation being whether the N-1 criterion is applied immediately after a disturbance, or after some period of time (e.g. 15 minutes) for corrective control actions.

The studies of [55, 56] discuss the long-term impact of following the N-1 approach. The work suggests that approaches that aim to limit the impact of a small set of events (e.g. preventing the N-1 contingencies from cascading into larger events), may increase the risk of blackouts due to the implicit assumption that higher order contingencies are independent.

2.4.2 Probabilistic methods

The system security handbook of the French TSO, RTE [57], discusses risk at a high-level, shown by the adapted diagram below in Figure 2.6. This figure already shows a consideration of risk in industry, with sensitivity to how likely particular contingencies may be, and the extent of their contingencies. Importantly, this figure suggests that there exists some acceptable risk threshold. Notably, probabilistic methods have been used by NERC (North American Electric Reliability Corporation) to compute probabilistic metrics such as loss of load hours (similar to system minutes, SMS) and expected unserved energy [58].

Within the context of power system operation, it has been shown [10] through comparison of deterministic and probabilistic methods that deterministic reliability constraints are inconsistent when compared against measures of risk. That is, strict constraints may either result in over mitigation of risks, or lead to unintended high risk situations, by failing to account for contingency probabilities or the severity of any constraint violation. Such a conclusion is supported by [59] which also compares deterministic and probabilistic methods, arguing for risk as a better basis for decision making and for the need to consider weather. The studies of [60, 61], further support the need to migrate to probabilistic methods, by stressing the trade-off between preventive and (potentially unreliable) corrective control.

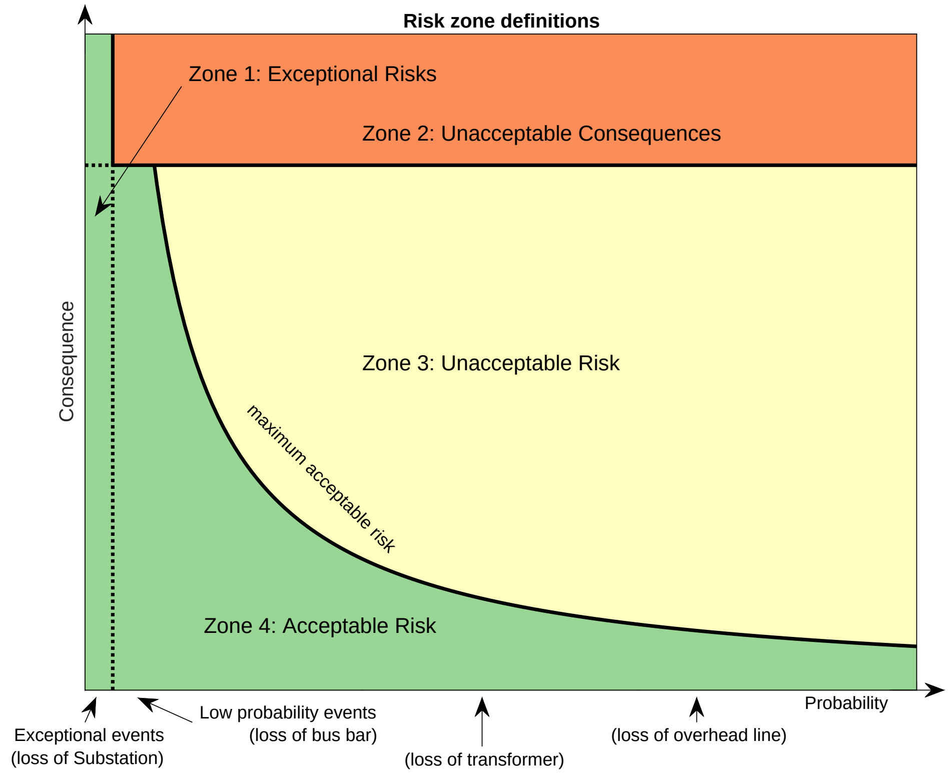


Figure 2.6: Visualisation of risk zones based on probability and consequence, adapted from the RTE reliability handbook [57]

A paper on the practical implementation of probabilistic reliability assessments [62] raises a number of useful critiques. Firstly, that unlike the N-1 principle, probabilistic methods require some definition of what a 'low' risk is. Attempts to define numerical risk thresholds may still necessitate exceptions for special cases, which may burden the operator. As stated by the author, in such cases operators may be more concerned with compliance than with taking the necessary actions to save the power system. Further, care must be taken that the output of a probabilistic reliability system is clear in what actions should be taken, or what events should be secured against, given that operators work under significant time pressure.

2.4.3 Reliability control

Beyond assessing the risk of the power system, a body of work exists on optimising risk (i.e. lowering the expected interruption cost) by suggesting control actions. Within the scope of this problem are an extension of PF and OPF algorithms commonly referred to as Security-Constrained Optimal Power Flow (SCOPF) algorithms. Traditionally SCOPF algorithms are solved over the set of deterministic N-1 events.. More recently SCOPF algorithms have been extended into stochastic and robust optimisation problems, solving over a set of probabilistic scenarios, as discussed in [63].

A review of the state-of-the-art of such models [64] shows the wide range of interpretations of such models, including risk-based and stochastic approaches. The author points out the gap that exists between the MINLP (mixed integer non-linear problem) formulation and the real-world power system, and the need for more precise industry data. As evident by applications at the Swiss TSO [65], SCOPF algorithms are also being studied within the industry.

2.5 Uncertainty modelling

Probabilistic reliability assessment and control approaches depend upon the estimated probability of future uncertain events. These uncertain events can be divided into two groups: continuous uncertainty, and discrete uncertainty. The following subsections describe present approaches to estimating and managing exogenous and discrete uncertainty on power systems.

2.5.1 Continuous uncertainty

The main sources of continuous uncertainty on power systems are those of uncertainty in load and generation, particularly related to wind and solar power injections. Continuous uncertainty also extends to weather, which affects both load and generation, but also may also influence component failures, however this will be discussed in more detail in following subsections. In all cases, the most important aspect of quantifying uncertainty is in the production of forecasting models, in particular models which provide a probability distribution of forecast errors for all cases [66, 67, 68, 69].

A significant body of research has been produced in the last decade on the topic of quantifying and managing the impact of wind power uncertainty on power systems. Studies on wind power uncertainty cover topics such as its impact on reserve management [70, 71], congestion management [72, 73], and also on risk assessment and control [74, 75]. An important aspect of predicting the potential impact of wind forecast errors is on the potential spatial correlation between wind farms, particularly with the use of copulas [76, 77]. Similar studies also exist for quantifying the impact of uncertainty for solar power and load, for example [78].

Given some forecasts of continuous uncertainty over time, methods are required to aggregate these into a set of multi-modal scenarios of future system states, considering all spatio-temporal correlations. Approaches to producing these scenarios are reviewed in [79], and it is suggested that the most feasible approach is one in which correlations are applied to the innovations of all forecast models as Gaussian white noise.

Such models can be used to generate a large number of scenarios of the continuous uncertainty of a power system. The computation of risk for a power system can be quite computationally expensive for a single scenario, yet any sampling based approach may require thousands of scenarios in order to cover the uncertainty space. As such, ongoing research has investigated the possibility to cluster scenarios into a smaller representative set [80, 81]. Improving and validating such approaches are necessary to improve the tractability of long-term reliability assessments.

Continuous uncertainty is however not within the scope of this dissertation. Given that real-time reliability assessment is applied at the present time, the uncertainty in load and generation has been considered negligible. As suggested in [62], significant forecast errors may be reinterpreted as discrete events that occur with some probability.

2.5.2 Discrete uncertainty

Discrete uncertainty relates to specific events that occur with some probability, and can be modelled using a boolean or integer variable. Examples of such events are the sudden failure of a component, sudden load disconnection, or even a specific forecast error (as mentioned at the end of the prior subsection).

Common approaches to predicting the probability of component outages, particularly in industry, are through the use of static failure rate (or hazard rate) models [82, 83]. Such models aggregate years of historical failures for components, or groups of components, into a static failure rate. Such a failure rate can then be converted into a probability of occurrence within some time-frame using a cumulative density function for the exponential distribution, as in [49]:

$$Q(t) = 1 - e^{\lambda t} \tag{2.3}$$

where:

$Q(t)$: probability that failure occurs within the time interval $[0, t]$
t : length of time period [yrs]
λ : failure rate of the component [1/yrs]

Approaches that assume a constant failure rate provide for straightforward models of failure probabilities, and are particularly useful for long-term studies given their simplicity. These models however fail to capture the influence of exogenous factors on the probability of failures. As discussed earlier, the range of threats to the Icelandic power system are well understood [20], as well as for other power systems [84], and evident through the classification of faults based on their root cause [85]. Therefore any models of failure probabilities, particularly applied in the short-term, should consider the spatio-temporal variability of these threats.

Research on the weather-dependence of failure rates and failure probabilities considered a state-based approach, as shown in [86, 87]. This approach defines two or more weather states, and then considers the probability that the system will move from one state (good weather) to another (adverse weather). Each component then has a failure rate defined for each weather state. Such approaches are useful in long-term studies, particular in showing the potential for correlated and dependent contingency probabilities during extreme weather events.

As stated by [27], such models may be too simplistic, since adverse weather can refer to a wide range of situations (lightning, ice storms, hurricanes, etc.) that all have different affects on power system components. Further, characterising failure and repair probabilities in the form of a Markov model may be misleading in the context of real-time and short term operation. An alternative approach to state-based models, is to model the failure rates and probabilities resulting from specific threats.

Studies such as [88, 89, 90] investigate the modelling of particular weather and component health variables on the failure rate of components. Another study applies threat-based failure rates to the Portuguese power system [91]. Such models are normally produced by identifying the causes of historical failures, segmenting failure statistics based on their causes/threats, and then performing regression on these data segments against variables that describe the threat (e.g. the threat of failures due to wind load are modelled using wind speed data). It is also common to include some Bayesian prior failure rates, particularly for new components or those missing data sets, to inform the regression process [92, 93, 94].

2.5.3 Contingency filtering models

Contingencies, which include but are not limited to the weather-dependent failures discussed in the prior section, are central to power system reliability assessment. Traditionally, reliability assessment only considers the set of single component contingencies (the N-1 set). Studies without a fixed contingency list, such as most probabilistic approaches, must use some method to select and filter possible contingencies. In fact, selecting only the N-1 contingencies can be thought of as a simple filtering approach. Contingency filtering is necessary due to limited computational power, limited time (particularly in a real-time context), and an intractably large number of possible contingency combinations. For example, if all possible component failure combinations were considered for a system with 100 fallible components, the contingency list would have 10^{30} contingencies.

A review of some contingency filtering approaches is provided by [64]. Some research [95, 96] has focused on assessing the applicability of heuristic approaches to contingency filtering, such as filtering by spatial separation (e.g. assessing only N-2 contingencies that include adjacent components) or other topological methods. One study on cascading [35] suggests that approaches that consider proximity in terms of electrical influence, rather than spatial proximity, may better represent real world power systems. Alternative approaches to topological filtering are those that filter based on probability or risk. Note that the latter requires some initial estimate of risk prior to the actual reliability assessment (in which assessing the consequence of contingencies is computationally expensive).

It is important to note that variable contingency probability models (both state-based and threat-based) depend upon the realisation of exogenous uncertainty. As such, contingency probabilities, and therefore contingency lists, will vary with exogenous scenarios. Therefore there exists a need to consider scenario clustering and contingency filtering in combination, particularly for short-term applications. To the best of the author's knowledge, no studies exist on such approaches.

2.6 GARPUR framework

The GARPUR research project aimed to produce generally acceptable approaches for probabilistic reliability assessment and control. The term Reliability Management Approach and Criteria (RMAC) resulted from the GARPUR project to describe the framework. The fundamental formulations of GARPUR can be found in the public deliverables of [1, 43].

In short, reinterpreting the description provided in [61] an RMAC consists of three ingredients:

1. **reliability target:** consisting of a set of constraints defining an acceptable operational state, with a TSO target for probability of compliance;

2. **socio-economic objective:** an aim to minimize a cost function that quantifies the cost to both society and the TSO given some system state (as discussed in section 2.3.3);

3. **discarding principle**: a method of selecting a subset of possible future system trajectories to assess the above two items over, given that the residual risk associated with discarded trajectories is below some pre-defined threshold;

To assist in understanding the above ingredients, it is useful to consider how the N-1 principle would be interpreted as an RMAC, shown in Table 2.1 below. Note that this specific representation does not reflect the actual N-1 implementation of any particular TSO, but instead is used for illustrative purposes. This interpretation shows that the N-1 principle can be interpreted within the context of a probabilistic reliability assessment, and highlights some of the implicit assumptions. The main assumptions are that all service outages are treated equally, and that higher-order contingencies are considered to be negligible, and all N-1 contingencies are considered to be non-negligible.

Extension of the GARPUR RMAC to a probabilistic implementation in the context of real-time and short-term operation is described in [11], considering the general organisational structure of TSOs [3], and the present data and tool availability. This was achieved by first describing the ideal algorithmic implementation of the GARPUR RMAC in an industrial setting, and then simplifying individual lines based on real-world constraints. For example, modelling of unexpected corrective control and system response were simplified to heuristic system-wide estimates given the lack of data pertaining to their occurrence. The practical implementation of the GARPUR framework in real-time operation represents a core part of this thesis, discussed in more detail in Section 3.2 and the relevant publications.

Table 2.1: Basic interpretation of the N-1 principle in terms of the RMAC ingredients from the GARPUR project

RMAC Ingredient	N-1 Interpretation
Reliability target	**Acceptability constraints:** Operational security constraints **Probability target:** 100% *Reliability target is to achieve the operational security constraints 100% of the time*
Socio-economic objective	**Function:** Boolean identifier of service outages **Aim:** Minimize *Minimize number of events that are expected to result in service outages*
Discarding principle	**Residual risk target:** Residual risk not quantified **Underlying assumption:** Higher order contingencies (N-2 and above) have negligible risk *Discard all events except for the N-1 contingencies*

Chapter 3

Methods and results

This chapter provides a summary of the publications by introducing the purpose of each paper in the broad context of the PhD project, then summarising and discussing the main results. Additional results that were omitted from the papers (due to length, scope and/or time constraints) are included in these sections.

The first section describes the work related to threat-based failure rates and methods of constructing these models in practice are discussed. The second section covers the conversion of the GARPUR framework into algorithms for real-world implementation, and the theoretical basis for pilot testing. The third and final section focuses on the pilot test design and its results.

3.1 Importance of weather-dependent failure rates

This section relates **Papers I** and **II**, investigating the importance of including weather-dependent failure rates into probabilistic reliability assessment. Only the key results are discussed and interpreted here, with the inclusion of some additional results.

3.1.1 Link between threats and failures

As noted in subsections 2.1.1 and 2.5.2, there exist a few main time-varying threats that affect the Icelandic power system. As shown by Figure 3.1, component outage data showed that the occurrence of faults varies over time, particularly when considering only specific threats (recorded for all outages). Often the threat leading to a particular component outage is estimated by control room operators, and not validated by maintenance staff or with any measured data or visual inspection. As such, the connection of threats to failures is not flawless, and may be prone to errors. Further, in the case of overhead transmission line (OHL) outages, only the line name is recorded and not a specific location (e.g. GPS location or span/tower number). Therefore attempts to connect historical data with outages is significantly limited by data resolution. Notably, even if the labels of threats assigned to outages are incorrect, there still exists a general tendency of more outages to occur during Winter, and fewer during Summer.

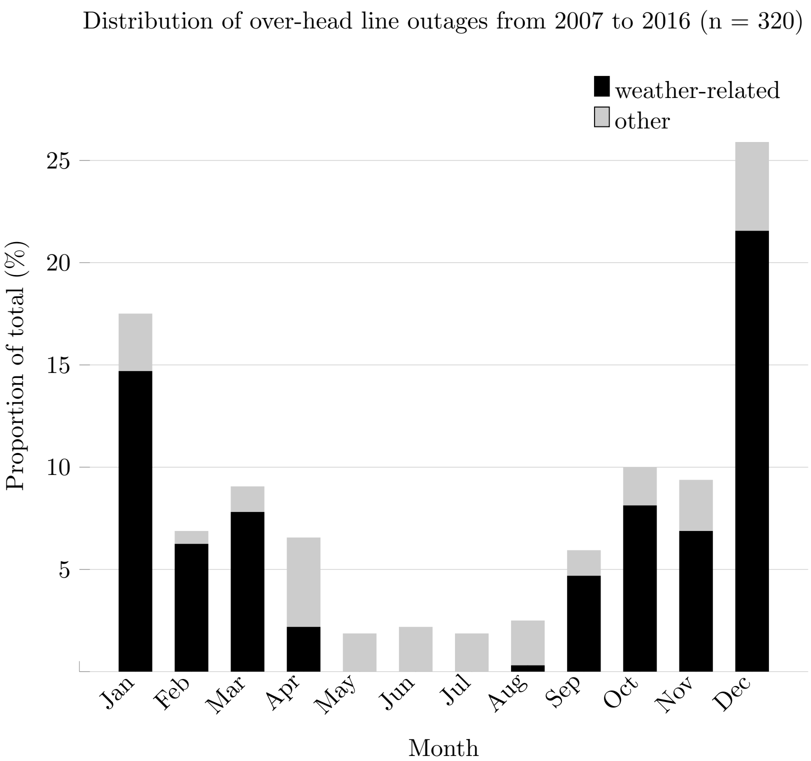


Figure 3.1: Monthly variation of the occurrence of unplanned outages on overhead lines caused by weather on the Icelandic transmission system

The number of faults in any given month can vary significantly from year to year. To highlight this, Figure 3.2 shows the same data as in Figure 3.1, but in the form of boxplots. Notably, the number of faults due to non weather-related causes is relatively constant. Faults due to caused by weather phenomena however occur primarily in Winter, and rarely occurring in Summer. Of the disturbances that have affected OHLs

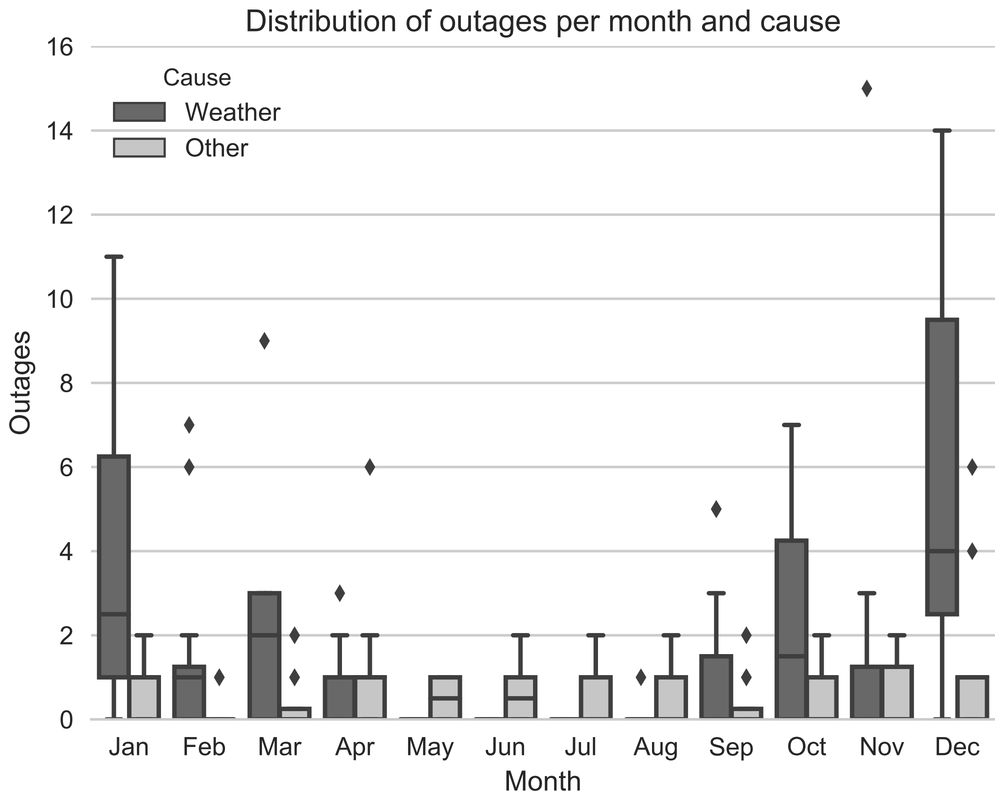


Figure 3.2: Boxplots of unplanned outages per month and cause (where diamonds represent outliers)

from 2007 to 2016, the most frequent underlying threats stated in the database are wind (32%) and snow and ice (18%), with a large proportion that are considered to be due to both (21%). Other noted spatio-temporally varying threats include salt, lightning and landslides/avalanches (naturally varying due to proximity to coasts or slopes), with a couple of outages listed as being due to weather but without any specific threat identified.

A number of modelling approaches were reviewed in **Paper I**, and it was determined that threat-based models are more applicable to the Icelandic case than state-based models. The main reason for this determination was that storms in Iceland vary in the threats they pose to the power system, as does the exposure of power system components to various threats. Aggregation of multiple threats into a single 'adverse weather' state would therefore make weather-dependent failure rate models difficult to understand. Similarly, it is unclear how adverse weather could be defined in an Icelandic case, given the varying exposure of lines to particular threats.

In order to determine the feasibility of building weather-dependent failure rate models, and to determine their potential impact on real-time reliability assessment, **Paper II** applied an existing wind-dependent failure rate model [97] to the Icelandic power system. Failure rate models were constructed using the outage data discussed above, and using 10 years of hindcast weather data of Iceland provided by DTU. Individual wind-dependent failure rate models for each OHL were trained on 8 years of data (2005-2012), and tested on the subsequent 2 years of data (2013-2014).

The errors (in estimating the occurrence of faults in the test set) between the wind-dependent and the static failure rate models were found to be negligibly different. In order to understand these errors better, particularly that the studied model did not significantly reduce errors, individual OHL models were studied, as in Figure 3.3.

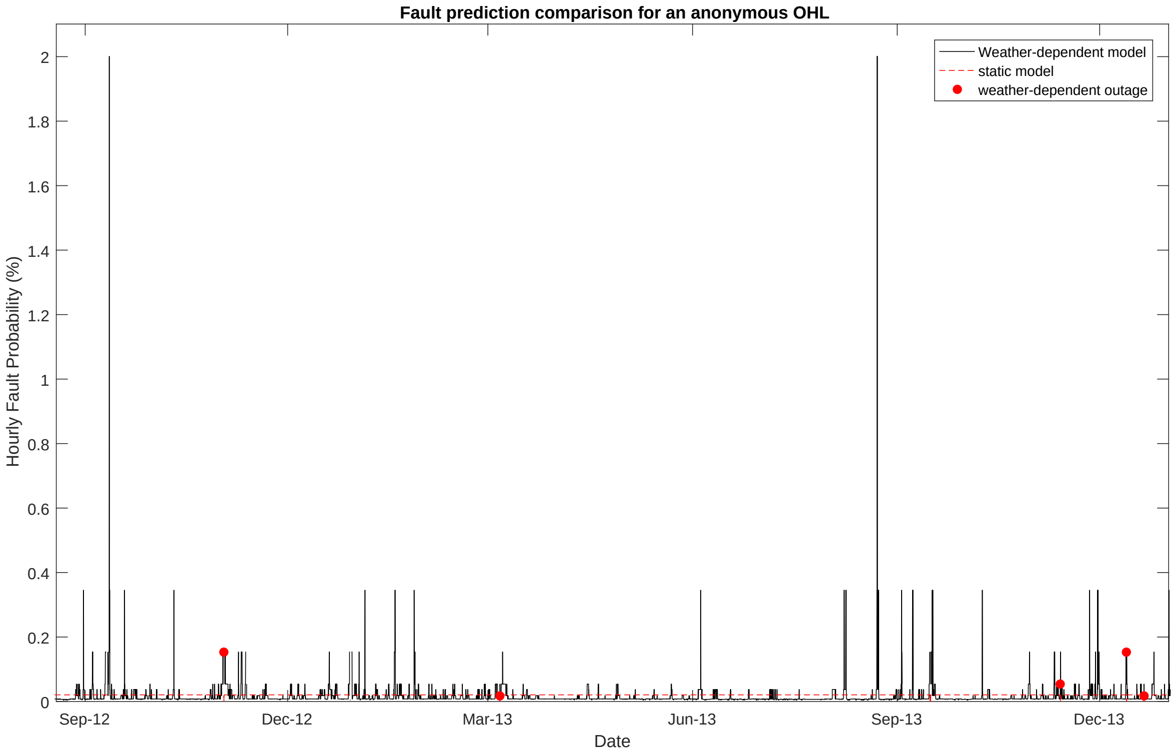


Figure 3.3: Example of errors resulting from failure rate models used in **Paper II**, comparing weather-dependent and static models with actual fault occurrences

One measure of accuracy, or of errors, in a probabilistic forecast is the Brier Score [98], which can be expressed as:

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i) \tag{3.1}$$

where:

BS : The Brier Score
N : Number of forecasted instances
f_i : forecasted probability of the i^{th} occurrence $\in [0, 1]$
o_i : i^{th} observation $\in 0, 1$

When predicting the occurrence of component faults, the ideal probabilistic model would output a probability of 100% when faults occur, and 0% when no faults occur. In such a case, the Brier Score would produce a result of 0, implying that the model has no error.

When considering the actual occurrence of failures, the weather-dependent failure rate models tend to slightly out-perform static models (i.e. they predict slightly larger probability of occurrence). Yet the data set primarily consists of hours in which no failures occur. As such, a model's Brier Score (i.e. errors) are dominated by its ability to predict that no failures occur. Figure 3.3 contains two significant spikes in probability (up to 2%) and a large number of smaller spikes, during which no outage occurred. The errors generated by these false positives negate the slight benefit gained when outages do occur. It is expected that with improved data resolution (particularly with precise locations of faults) the accuracy of failure rate models could be improved considerably.

Further unpublished work was performed on extending the model from **Paper II** into a continuous failure rate model, and testing the model's sensitivity to various forms of data pre-processing. This pre-processing consisted of grouping data for similar lines together (e.g. grouping outages of all 220 kV lines together to make a single 220 kV OHL failure rate model), considering icing, and on threat correction (e.g. assuming wind labels are erroneous on faults occurring with low wind speeds, or missing on faults occurring in high winds). The structure of these failure rate models is defined by:

$$\lambda(v) = \lambda_o + c(v)\lambda_w \tag{3.2}$$

where:

v : maximum mean hourly wind speed along length of OHL [m/s]
$\lambda(v)$: failure rate at a point in time, given v [1/yrs]
λ_o : failure rate due to other faults [1/yrs]
λ_w : failure rate due to weather-dependent faults [1/yrs]
$c(v)$: multiplication factor for weather-dependent faults, given v [-]

The values of λ_o and λ_w were calculated by employing a Bayesian inference approach as shown in [93]. The correction factors were calculated using a method similar to that in **Paper II**, except the discrete histograms were replaced with kernel density functions. That is, the correction factors are calculated from probability density functions as:

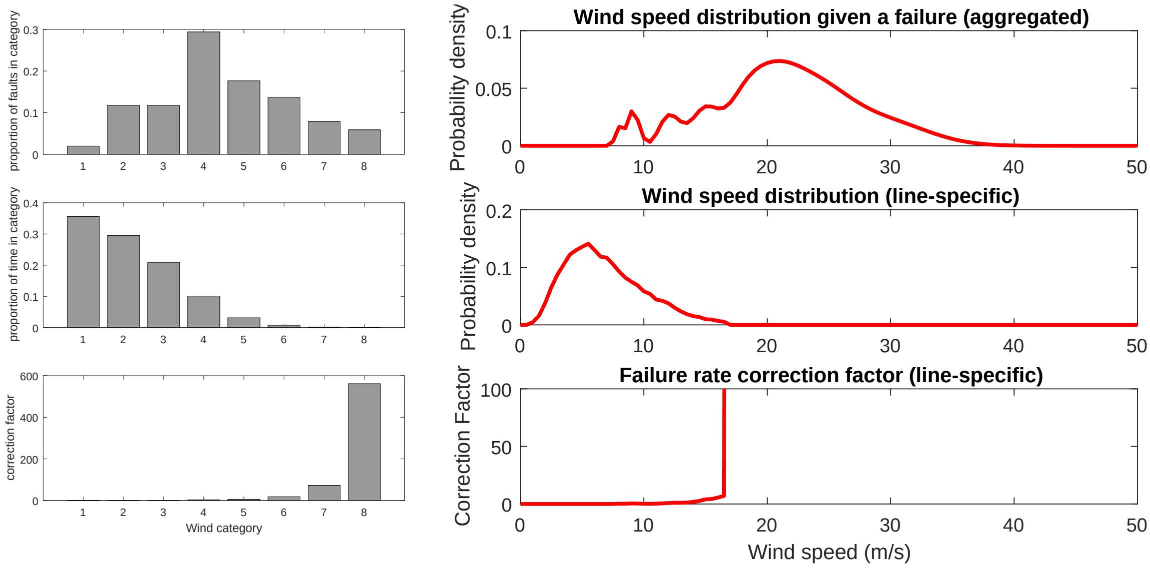
$$c(v) = \frac{f(v|x)}{f(v)} \tag{3.3}$$

where:

$f(v)$: probability density function of wind speeds
$f(v|x)$: probability density function of wind speed given outages

The difference between the discrete approach used in Paper II and the continuous approach based upon kernel density functions is visualised in Figure 3.4. Primarily, there is no longer any need to abstract wind speeds by converting them into wind categories. Notably, the correction factor curve in the continuous approach, when aggregating failure data, is dominated by the asymptote when approaching wind speeds that have not been observed for the particular line. Note that the top and middle probability density functions in Figure 3.4 represent $f(v|x)$ and $f(v)$ in (3.3), respectively.

Icing (and hence icing credibility) was not considered in these models. This was due to there being insufficient weather observations to compute the icing credibility in real-time, and therefore it would not be usable in pilot testing (discussed later in Section 3.3). As shown by 3.5, the resulting correction factors for OHLs vary significantly. Note the significant range of curves between individual lines, and that the grouped curves (e.g. all 66 kV lines aggregated into a single model) clearly don't represent all lines. For example, the most robust line (grey line furthest to the right) would have its failure probability over-estimated when using aggregated statistics.



(a) discrete (b) continuous

Figure 3.4: Comparison of discrete correction factor calculation from **Paper II** with the continuous correction factor calculation used for pilot testing (fault data aggregated by voltage level)

Similarly, the weakest lines will have their failure probabilities under-estimated, and will skew the reliability of stronger OHLs.

Similar results are stated in [83], that pooling of components into a single failure rate model can provide misleading results. Therefore future work needs to concentrate on adding granularity to failure rate models.

Given these results, the failure rate models employed in **Paper V** and discussed in 3.3 do not employ any aggregation.

3.1.2 Relevance to real-time reliability

The discrete failure rate models in **Paper II** were used along with an annual time series of wind speeds to assess the reliability and risk of the Icelandic transmission system. This study used the commercial software Promaps Realtime [99], which was already in use at Landsnet. At the time of this study, the Promaps Realtime software computed aggregated risk metrics for the power system (such as System Minutes, defined in (2.1))based on subcomponent failure rate models, combined with a transport flow algorithm. This study showed that in periods of calm weather, such as in Summer periods, the use of variable failure rate models will result in a lower estimate of system risk compared to static failure rate models. This is due to the failure rate correction factors providing multipliers with values less than 1 in periods of calm wind.

Counter-intuitively, most winter periods also have a lower risk when using variable failure rate models. Instead, the risk is concentrated in short periods of high wind, particularly when high winds are experienced by a large number of OHLs. Notably there is always a non-zero risk on the power system due to the presence of threats that are not entirely dependent on wind.

Histograms of risk distributions over Summer and Winter are shown in Figure 3.6, reproduced from **Paper II**. In both Summer and Winter the risk distribution is shifted

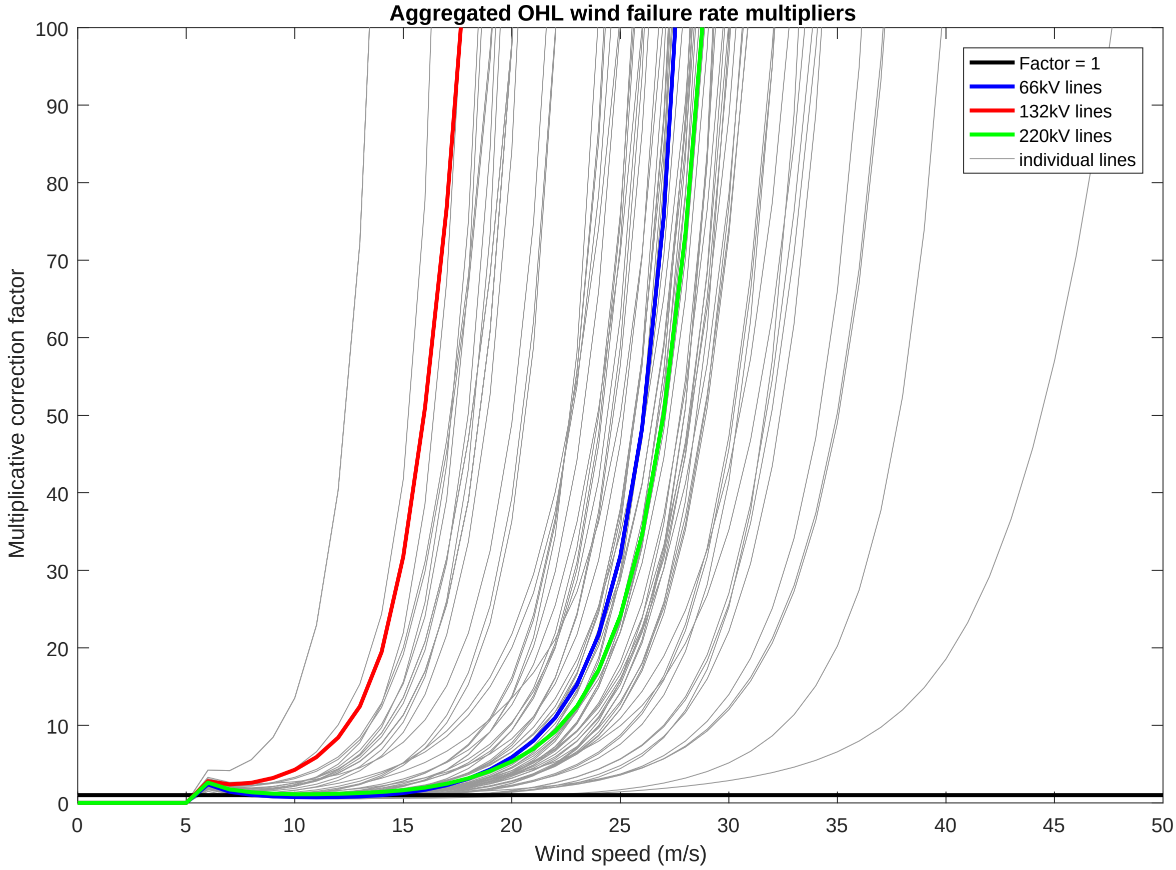


Figure 3.5: Example of errors resulting from failure rate models used in **Paper II**, comparing weather-dependent and static models with actual fault occurrences

to the left (i.e. median risk is 33% smaller) when using weather-dependent failure rate models. The exceptions being outlier events in winter where risk is far greater than the maximum risk in the static failure rate model, due to high wind speeds experienced by multiple OHLs.

The impact of maintenance is also studied for both the static and weather-dependent reliability assessments. For a maintenance outage of a particular OHL, the risk increases by a factor of 2.4 in both assessments. Given that the weather-dependent assessment results in a lower risk during Summer, then the absolute increase in risk during maintenance is also decreased. In the example shown in **Paper II** the risk in the static case increases from 70 to 170 minutes, whilst in the weather-dependent case the risk increases from 40 to 95 minutes. As such, by ignoring the relationship between wind and component outages in this summer case, the risk increase due to maintenance would be over-estimated by a factor of 2.4.

These results implied that the pilot testing of the GARPUR framework in Iceland could not provide realistic risk assessments unless it considered weather-dependence in some degree. By modelling additional threats, it is anticipated that the estimated risk during certain periods where few threats are credible will significantly decrease, and the height of risk peaks will significantly increase. Similarly, the improvement of failure rate models by collecting higher resolution data (temporally, but more-so spatially) will improve risk estimates.

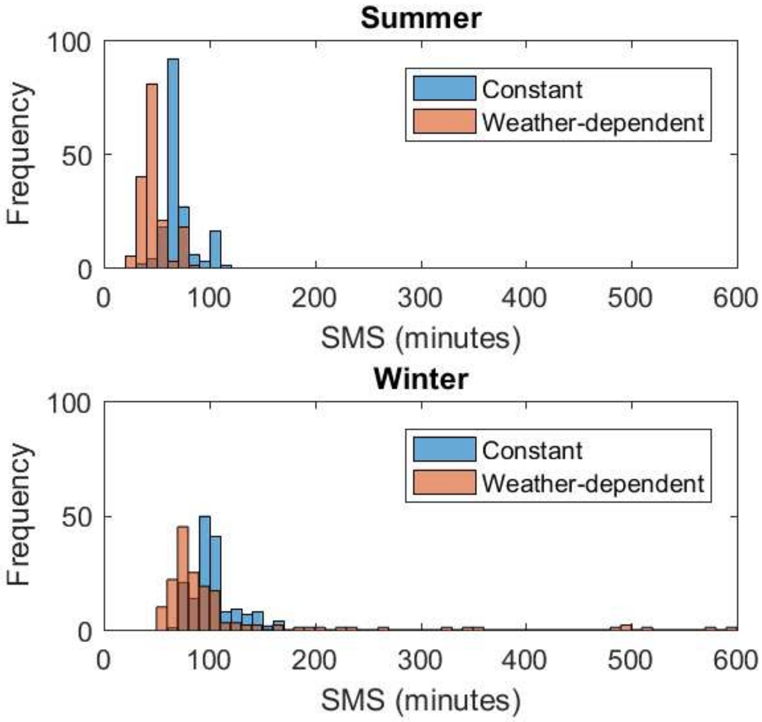


Figure 3.6: Distribution of Expected SMS over Winter and Summer, for constant and weather-dependent failure rate models. Overlapping portions of the distributions are shown in grey, from **Paper II**.

3.1.3 General framework for failure rate models

At present, there is no general framework for the modelling of variable failure rates for real-time and short-term applications that is generally applicable to TSOs. Significant literature exists however in developing these models in specific circumstances. The development of such a framework is required in order to then develop a clear framework for data sharing between TSOs.

A framework was provided in **Paper I** for a single TSO to determine how to model a particular threat. In general, an industry-wide approach to managing failure rate models could be in the form shown below in Figure 3.7. In short, it is possible to define a set of threats that may affect a particular power system, such as the list below for the Icelandic power system, reproduced from **Paper I**, and based upon [20]:

- Wind (galloping)
- Wind (structural failure)
- Ice loading
- Lightning strikes
- Earthquakes
- Landslides/Avalanches
- Volcanic eruptions

- Glacial floods (jokulhlaup)
- Snow accumulation
- Salt pollution
- Generator failure
- Load failure
- Solar flares
- Human error (random)
- Human error (control errors)
- Human error (due to proximity of planned work)
- Sabotage
- Technical
- Other (unknown/invisible cause)

If each of these threats could be defined in a way that they are mutually exclusive (that is, 'ice' and 'wind' may need to be defined as 'ice only', 'wind only' and 'ice and wind'), then each threat can be defined by a threat-specific model. Such models may be generalised across all TSOs. That is, a general model for estimating the probability (or credibility) of failures due to lightning should be possible. Further, such general models could be coupled with some prior parameters that are derived from data that is shared across the industry.

General models can then be adjusted using any available TSO-specific TSO data. Then, by taking appropriate observational measurements of any model parameters, it should be possible to estimate the real-time failure rate for a particular component due to a particular threat. Each threat-specific failure rate can then be aggregated into a single real-time component failure rate. Importantly, it is suggested that all threat models should be applicable to all TSOs. For example, if such a system were in place, Iceland may have models that account for threats that haven't historically threatened the system. The recent growing threat of lightning would not come as a surprise if predictive models were already in place based on data from countries that already experience frequent outages due to lightning. Longer-term, models for forest fires may one day be required in parts of Iceland, which could be based on prior data from countries such as Portugal where the threat of fires already exists [100].

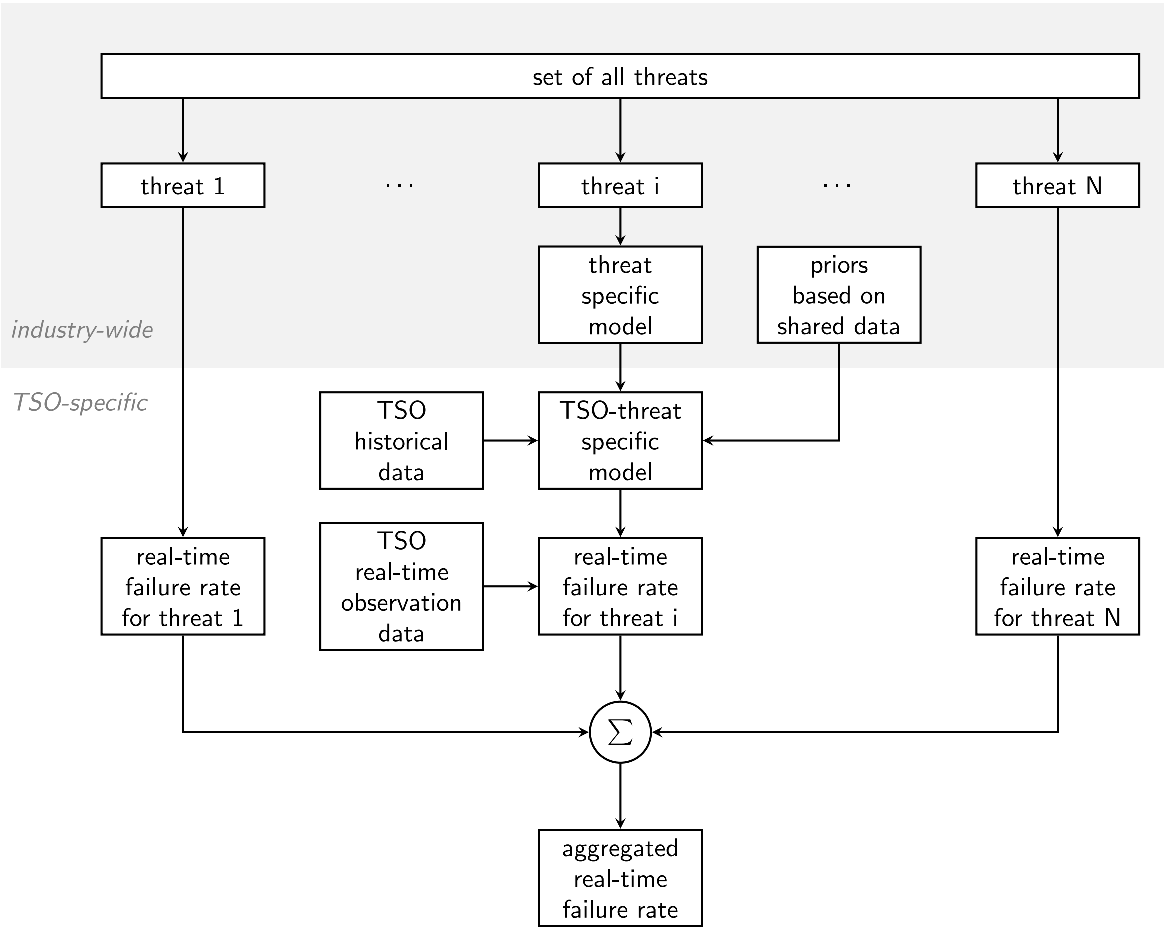


Figure 3.7: General framework for failure rate models, considering possible industry-wide sharing of data and models

3.2 Translating the GARPUR framework into an industrial prototype

The GARPUR project provided a framework (see Section 2.6) for probabilistic reliability assessment and control. In order to perform the pilot testing, described in Section 3.3, the framework had to be converted into a practical algorithm. This was achieved by first producing an ideal formulation, which was then used to determine the possible limitations and roadblocks. By knowing these limitations and roadblocks, and by considering existing TSO data and tools, it was then possible to determine a first-step algorithm for implementation in the time frame of the GARPUR project. The following subsections describe the ideal prototype, the resulting algorithm for implementation, the theoretical implications of the results, and finally a discussion on the need for further validation.

3.2.1 Formulating the ideal prototype

The work related to the formulation of the ideal prototype is detailed within restricted internal deliverables and memos of the GARPUR project. The core contributions to this work are represented in **Paper III**, which elaborates on system security based on the concept of system trajectories. Specifically, the term system trajectory is used to denote a sequence of transitions, where each transition is a quartet of realisations of the following four processes:

1. realisation of exogenous uncertainty;
2. realisation of one or more contingencies;
3. a resulting (potentially uncertain) system response;
4. a resulting (potentially uncertain) corrective control;

In the context of real-time security assessment, only transitions in the first time step are considered. The remainder of the system trajectory is characterised by restoration actions, and additional contingencies or other unexpected transitions are not explicitly considered. Instead, these trajectories are grouped together as 'unassessed overlapping trajectories' where the term 'overlap' refers to one contingency overlapping the system restoration of a prior contingency.

Notably, the number of potential trajectories that can be experienced is intractably large, hence why the GARPUR framework incorporates a discarding principle. Figure 3.8 below from **Paper III** visualises the scope of real-time security assessment. With each of the four realisations in the first transition, there is some probability that is discarded in order to make the assessment tractable. System restoration is also formulated in a way that implicitly assumes no further contingencies, and as such there are a set of trajectories that are implicitly discarded, referred to as 'unassessed overlapping transitions'.

A system trajectory should be considered until the system state has converged with the planned trajectory. This however is not entirely clear, given that an outage may affect spare part inventories, or may be due to an irreversible event (e.g. following a volcanic eruption, it may be inconceivable to rebuild the system such that it is equivalent to the initial system). In **Paper III** it was suggested that a reasonable

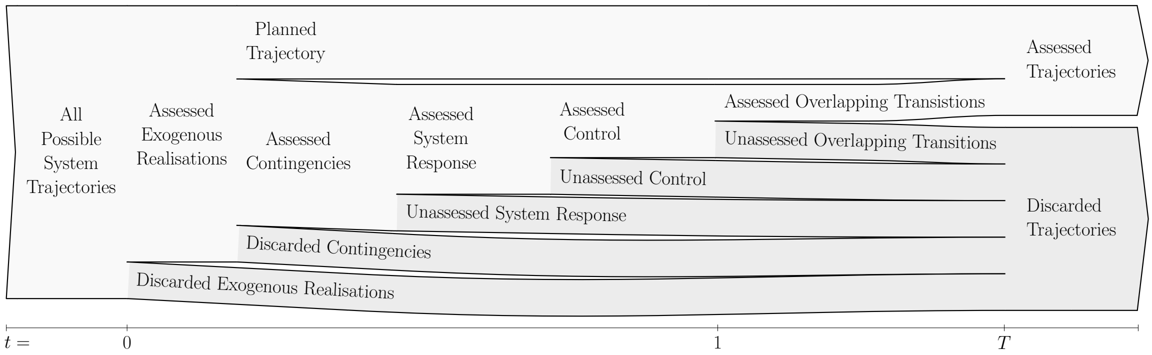


Figure 3.8: Probability space partitioning, based on trajectories

simplification is that the assessment period should be equivalent to the maximum expected service outage duration. In reality expected service outage durations are an uncertain quantity. As such, the determination of an assessment period is undefined. Traditional reliability assessments for real-time applications only assess the initial disturbance or the first operational period (e.g. 15 minutes) following the disturbance. This implicitly assumes that the increased vulnerability of the system during restoration (particularly for disturbances that take some time to restore) is negligible. For example, if an OHL and a cable fault don't result in any operational security limit violations, then they are considered to be secured against by the N-1 principle, despite the cable fault likely having a significantly longer outage duration and hence a long period of increased system vulnerability.

The definition of trajectories in **Paper III** resulted in the following formulation of the probability of an N-k trajectory for real-time probabilistic assessment:

$$\pi_s(\boldsymbol{x}_0) = \mathbb{P}(\boldsymbol{x}_0)\pi_\Delta(\boldsymbol{\xi}_0, c_0, r_0, b_0|\boldsymbol{x}_0) \prod_{t=1}^{d_c-1} \pi_\emptyset(\boldsymbol{x}_t, \boldsymbol{\xi}_t), \tag{3.4}$$

where:

π_s : Probability of the s^{th} trajectory
π_Δ : Unplanned transition probability function
$\pi_\emptyset$: Planned transition probability function
$\boldsymbol{x}_t$: System configuration at time t
$\boldsymbol{\xi}_t$: Realisation of exogenous variables at time t
c_t : Occurrence of a contingency at time t
r_t : System response at time t
b_t : Corrective control behaviour at time t
d_c : maximum service outage duration of all assessed contingencies

This equation provides the probability of the s^{th} trajectory from a potentially uncertain initial system configuration (denoted by $(\boldsymbol{x}_0)$), given that it represents some initial transition (denoted by $(\boldsymbol{\xi}_1, c_1, r_1, b_1|\boldsymbol{x}_0)$) and a probability of no unplanned transitions during restoration (denoted by $\pi_\emptyset(\boldsymbol{x}_t, \boldsymbol{\xi}_t)$). Ideally, the function of the probability of a transition is formulated as:

$$\begin{aligned}\pi_\Delta(\boldsymbol{\xi}_t, c_t, r_t, b_t|\mathbf{x}_t) = &\mathbb{P}\left(\boldsymbol{\xi}_t \in \Xi^t(\xi_{[t-1]})\right)\\ &\times \mathbb{P}\left(c_t \in \mathcal{C}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t\right)\\ &\times \mathbb{P}\left(r_t \in \mathcal{R}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t, c_t\right)\\ &\times \mathbb{P}\left(b_t \in \mathcal{B}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t, c_t, r_t\right),\end{aligned} \tag{3.5}$$

where:

$\mathcal{C}^t$: Set of all possible contingencies at time t
$\mathcal{R}^t$: Set of all possible system responses at time t
$\mathcal{B}^t$: Set of all possible corrective control behaviours at time t
$\Xi^t(\xi_{[t-1]})$: Set of all possible exogenous realisations at time t
$\xi_{[t-1]}$: Sequence of realisations up until time $t-1$

This formulation provides a description of all possible transitions at a single point in time, but essentially describes an intractably large set of unique transitions, given that each process is conditioned on the prior process. Further, the set of possible system responses and of corrective control behaviours is not well understood, and there is insufficient data to provide meaningful probability estimates. To resolve these issues, the formulation was simplified to:

$$\begin{aligned}\pi_\Delta(\boldsymbol{\xi}_t, c_t, r_t, b_t|\mathbf{x}_t) = &\mathbb{P}\left(\boldsymbol{\xi}_t \in \Xi^t(\xi_{[t-1]})\right)\\ &\times \mathbb{P}\left(c_t \in \mathcal{C}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t\right)\\ &\times \mathbb{P}\left(r_\emptyset\right)\\ &\times \mathbb{P}\left(b_\emptyset\right),\end{aligned} \tag{3.6}$$

where:

$r_\emptyset$: The planned/expected system response
$b_\emptyset$: The planned/expected corrective control behaviour

Rather than enumerating the sets of system responses and corrective control behaviours, it is suggested in **Paper III** to instead set an overall estimate of their respective probabilities. For example, assuming that the expected (i.e. modelled) system response occurs 99.9% of the time. The accuracy of this estimate is studied with a theoretical simulation. The simulation uses a cascading system response model, where each component has some probability of sympathetic tripping due to adjacent tripping events, referred to as θ. The value of θ is varied and the resulting risk of a test system is estimated. The sympathetic tripping module of the simulation is then disabled, and instead a value for $\mathbb{P}(r_\emptyset)$ is calculated, such that the simulation results in the same risk estimate. In other words, the complexity of sympathetic component tripping is replaced by a simple system-wide estimate of the probability of an 'expected system response'.

This process was performed for five values of θ from 0% to 2%, as shown below in Table 3.1, reproduced from **Paper III**. The relationship between θ and $\mathbb{P}(r_\emptyset)$ was then projected for larger values of θ using a simple linear regression. It is then found that such a simple regression is valid up until $\theta = 5\%$, shown in Figure 3.9. This suggests that, if the probability of sympathetic tripping is small, that it may be a reasonable assumption to model the uncertainty related to it through an estimate of $\mathbb{P}(r_\emptyset)$ rather than through extensive Monte Carlo simulations. Importantly, in this study in **Paper III**, the use of $\mathbb{P}(r_\emptyset)$ instead of θ reduced computation times by a factor of 130 to 190.

Table 3.1: Assumptions of probabilistic coverage given modelling decisions (from **Paper III**)

Sympathetic tripping probability: θ	Expected ENS (MWh)	Computation time (seconds)	Equivalent Assumed Value of $\mathbb{P}(r_\emptyset)$
0%	0.2541	14.24	100%
0.5%	0.2594	1780	99.973%
1.0%	0.2646	2641	99.948%
1.5%	0.2697	2735	99.923%
2.0%	0.2741	2541	99.901%

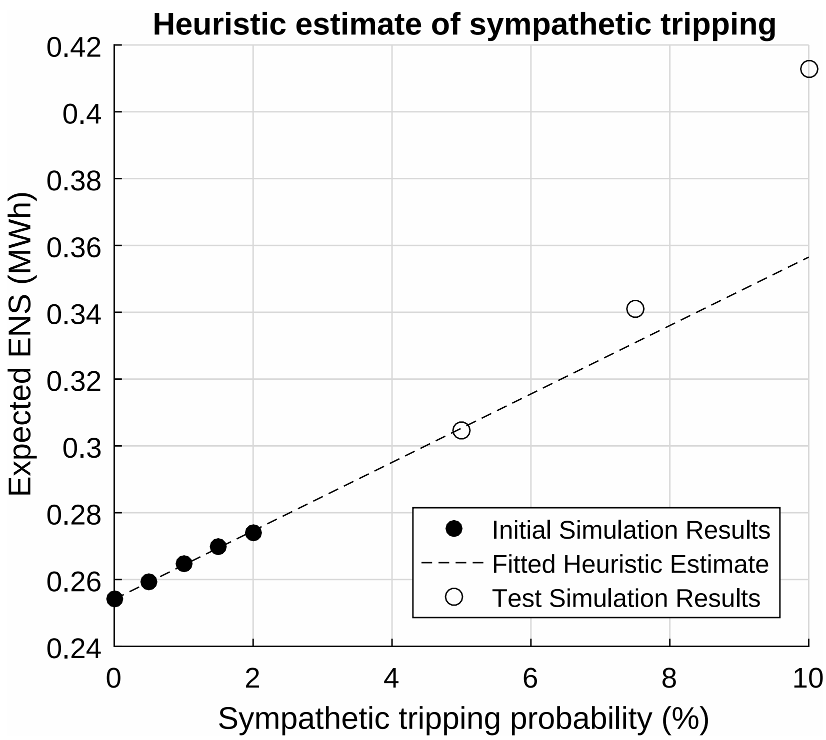


Figure 3.9: Comparison of heuristic estimates with simulated results for sympathetic tripping (from **Paper III**)

3.2.2 Algorithm for real-time reliability assessment

As noted above, the algorithm developed for real-time reliability assessment is elaborated in GARPUR deliverable D6.2 [11]. The algorithm describing the ideal implementation that formed a basis for the algorithm in D6.2 can be found in Appendix A. In developing this algorithm, the main consideration was tractability, particularly in reducing the problem's complexity, but also considering limited data availability, and in proposing something that could fit into present day TSO processes. The reliability assessment algorithm was also supported by the high-level algorithmic descriptions of:

- Contingency discarding;
- Uncertainty modelling;
- System response;
- System restoration;
- Socio-economic impact assessment.

As part of the author's contribution to an internal GARPUR report (deliverable ID6.3.1), and in collaboration with colleagues from other European TSOs on the GARPUR project, a list of feasibility issues was generated. This list was generated by considering each line of the ideal algorithm (in Appendix A). Each line was considered in terms of tractability, data availability and whether it was generally applicable. The collaborative discussions on these aspects were then distilled to specific feasibility issues.

This list described issues that would render the ideal approach impractical for industrial use. Table 3.2 below lists these issues, as well as the solutions that were implemented in D6.2 [11] and the subsequent GARPUR pilot test at Landsnet.

The synthesis of the design constraints above is that any new reliability assessment approaches should be designed around existing system response tools, and additional tools should be relatively simple. Although in some cases, particularly in the case of corrective control, the new models will be overly-simplified, they should allow for TSOs to perform sensitivity testing to determine how important it may be to improve such models and to collect the requisite data. The design constraints also imply that uncertainty modelling (related to continuous uncertainty) should be discarded in a first-step implementation of real-time reliability assessment, and that socio-economic impact assessment should be greatly simplified.

Tractable contingency discarding could simply consist of selecting only the N-1 contingency list. Risk-based contingency filtering approaches are ideal, but may be too time consuming for real-time applications. That is, the discarding process exists to reduce excessive calculation of contingency consequences, and would therefore defeat its own purpose if it also calculated consequences. Such approaches should only be used if they can be performed off-line, in anticipation of RT situations, or if suitable proxies of contingency risk assessment can be created.

As outlined in [101], the first-step approach described in [11] can be partially described as a migration from existing N-1 contingency analysis approaches towards a probabilistic N-1 reliability assessment. Assuming a TSO presently computes loss of load with a system response model as part of their N-1 contingency assessment, this gradual migration can be summarised as:

1. Develop a heuristic model to convert lost load into energy not served for each contingency;
2. Develop economic models (e.g. nodal VOLL curves) to convert energy not served into interruption costs;
3. Compute N-1 contingency probabilities using failure rate models to convert interruption costs into risk;
4. Assess the worst-case consequences and the discarded probability in order to compute residual risk;

Table 3.2: List of main feasibility issues for implementing the GARPUR framework in the context of real-time reliability assessment and resulting design constraints

Feasibility issue	Design constraints
1. Algorithm must be computable within a useful time-frame (i.e. a few minutes)	Algorithm should use existing TSO models, else additional models should be simple or computable off-line
2. Definition of acceptability constraints is not clear	Existing operational security limits should be used as acceptability constraints until term is clearly defined
3. Reliability target may be variable over time	Algorithm should be defined such that TSO can vary target if required (setting value requires pilot testing)
4. Acceptable risk threshold (for the discarding principle) is undefined	Same as above
5. Generating a list of all possible trajectories is not possible	Contingency discarding algorithm should only generate a list of trajectories to be assessed
6. Need an approach for estimating the worst-case consequences	System response model should be able to quickly estimate consequences of total system loss
7. Not feasible to model the system response for discarded contingencies	Aggregate into a single worst-case estimate (see above)
8. Consequence of HILP events are difficult to assess	Ability to assess consequence depends on TSO system response models
9. Considering continuous uncertainty in RT is not presently feasible	Assume no continuous uncertainty in RT formulation, consider significant forecast errors as contingencies
10. TSOs lack data on uncertainty and correlations thereof	As above, continuous uncertainty not considered in first-step approach
11. Data on system response and corrective action behaviour reliability is difficult to obtain	Formulate approach to include general assumptions on system and control reliability
12. A large number of trajectories will make the approach intractable	Discarding approach should allow TSOs to set hard limits on the number of contingencies/trajectories
13. Not tractable to model each trajectory as a stochastic dynamic simulation with control	Build model around existing TSO system response tools used for contingency analysis
14. A model of corrective control must essentially be a model of the operator, or an SC-OPF model, which is not tractable in RT	First-implementation should limit modelling of corrective control to existing models used by TSOs (if any)
15. Challenging to model human error, particularly delayed action or failure to observe faults	Simplify by using general assumptions related to probability of unexpected control behaviours for first-implementation
16. No good models for estimating disconnection durations	Initial model should use simple heuristic estimates based on historical data
17. The duration of some disconnections extend beyond the RT timeframe	Allow TSO to set assessment period
18. Insufficient data to perform Socio-Economic Impact Assessment analysis described in [43]	Implementation should allow for different consequence functions, dependent upon TSO data availability

5. Aggregate the N-1 contingency probabilities for contingencies that do not result in operational security limit violations in order to represent reliability as the probability of an acceptable system state.

Note that, in cases where sufficient economic data is not available, it may be sufficient to simply express consequences in terms of ENS and therefore risk in terms of expected ENS (also referred to as Expected Unserved Energy (EUE) in [58]). Additional steps required to achieve the GARPUR approach, by improving the probabilistic N-1 reliability assessment, outlined in [101] are:

6. Upgrade the failure rate models used in item 3 to be variable and sensitive to the changing threats of the power system;
7. Incorporate a contingency discarding/filtering algorithm, dependent upon the variable failure rates, to have a dynamic contingency list;
8. Use real-time system state data;
9. Assess the reliability across multiple exogenous scenarios (of load or generation) or include forecast errors as contingencies;
10. Extend interruption cost estimate into a full socio-economic impact assessment that is in line with [43];
11. Optimise the socio-economic impact of the power system by varying controls, such that a reliability target is maintained (based on the probability of an acceptable system state).

The ability of a particular TSO to achieve any of the steps above may vary significantly, as may their existing process in already implementing such steps in practice. Similarly, the time and resources required in order to achieve each of the above steps will vary. It is likely that the size of the transmission system will be the main indicator of the relative speed with which a TSO could make the above changes. Larger systems would have increased problem complexity, requiring greater computational power to perform a probabilistic reliability assessment. Similarly, the speed at which such changes take place are likely to be limited by or facilitated by regulatory barriers or incentives. These mainly relate to the costs of data collection and storage to facilitate such data-driven methods of reliability assessment.

3.2.3 Need for real-world validation

The study in **Paper III** showed that the approach described in [11] is tractable on a small test power system, given some simplifying assumptions. However, test systems are not constrained by data availability and data quality, and therefore do not necessarily represent real world power systems. As such, the proposed approach required validation through real life pilot testing.

3.3 Pilot test to validate the GARPUR framework

Section 3.1 provided an argument for, and an analysis of, threat-based failure rate models implemented into probabilistic reliability assessment. Section 3.2 provided a discussion on **Paper III** and on [11], highlighting key features of the first-step interpretation of the GARPUR framework for real-time probabilistic reliability assessment. This section relates to **Paper IV** and **Paper V**, which synthesise and validate the two prior sections by discussing the real-world implementation of the GARPUR framework with threat-based failure rate models on the Icelandic transmission system.

The objective of the near real-life GARPUR pilot testing, as interpreted in [102] and based upon the GARPUR Description of Works, was:

> *to validate the practicalities of incorporating... new families of criteria into a TSO's actual decision making processes. [In order to achieve this] a pilot test runs a selected set of reliability management applications using the new criterion with actual system data, and evaluates decisions within the particular system and context(s).*

For the Icelandic pilot test of the GARPUR methodology in the real-time context, the following pilot-test specific objectives were defined:

1. Implement the RT reliability assessment algorithm (as described in Section 3.2 and [11]), and comment on tractability/speed.
2. Comment on the interpretation of outputs from the assessment.
3. Investigate methods of visualising outputs.
4. Test the sensitivity of the outputs and speed to modelling assumptions and decisions.

In achieving these four objectives, the pilot testing could also achieve the overall intended objective specified in [102]. The pilot test itself began development in late 2016, and studies performed on the pilot test in 2017 from March to June. As of January 2018 the pilot test is still running in real-time on the Icelandic transmission system. The main results, including a functional description of the methodology, of the pilot testing can be found in **Paper V**, and as such will not be repeated here. The remainder of this section will provide extended information regarding the data used, the main assumptions, the pilot test design, extended results related to the four objectives above, and finally a section on the implications for future work.

3.3.1 Data and tools

The main assumptions and simplifications required to implement the GARPUR approach on the Icelandic transmission system are listed in Table B.1 in Appendix B.

The main data sets required to perform the pilot test were:

- Static (base) bus-branch model of the transmission system
- Live system state data
- High resolution historical weather data (provided by DTU)
- Historical outage data (i.e. time, component name, cause, etc.)

- Operational security limits (e.g. Acceptable voltage limits, inter-area congestion limits)
- Nodal Value of Lost Load (VOLL) curves
- Nodal end-user demographics (e.g. proportion of residential load)
- Generator inertia and droop constants
- Under Frequency Load Shedding (UFLS) rules
- System Integrity Protection Scheme (SIPS) logic
- Component locations (i.e. GPS coordinates)
- Generator operational modes (to determine participation in LFC)
- Pre-planned control actions

The additional tools, and their links to the main data sources are shown below in Figure 3.10. A number of data sources are queried with each iteration of the pilot test, referred to as *live data* and shown by solid arrows. Other data is only queried manually when necessary (due to updates or changes in downstream tools), which are referred to as *static data* and shown by dashed arrows. The three main live data sources are those related to the system state, the exogenous uncertainty (in this case weather, but should be extended to load, maintenance and market data), and user inputs. Note that two separate weather data sources were used for the pilot test. Historical hindcast data was used to build base failure rates, whilst live meteorological observational data was used to inform the live failure rates. Ideally both historical and live weather data would come from the same source. The pilot test tool is described in detail in the next section.

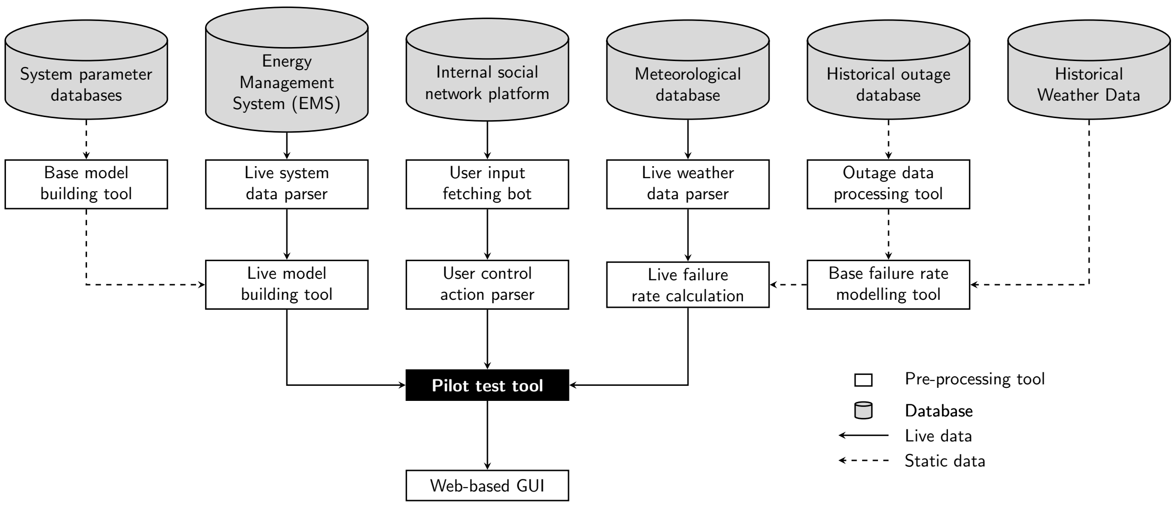


Figure 3.10: Overview of main data sources and processing tools

The live model building tool, user control action parser and live failure rate calculation are embedded in the same matlab code as the pilot test tool, and are therefore computed in series with the reliability assessment. Similarly, visualisations of outputs are also created in series. As discussed later in subsection 3.3.4, such computation should be performed in parallel to the reliability assessment, to avoid inefficient resource allocation.

3.3.2 Pilot test tool

This subsection contains a description of the pilot test tool, reproduced from **Paper V**. Figure 3.11 provides a high-level flowchart of the pilot test tool. Importantly, the loop of 'System Response', 'System Restoration', and 'Assess Contingency Consequences' modules are computed in parallel, on the CPU pool, over the set of contingencies. The tool itself is designed to be run in a continuous loop, illustrated by the dashed line. When one iteration of the reliability assessment is complete, the system state data is updated using real-time information from the local Energy Management System (EMS), and the next iteration commences. The following subsections provide detailed descriptions of each module within the pilot test tool, with emphasis placed on the system response tool.

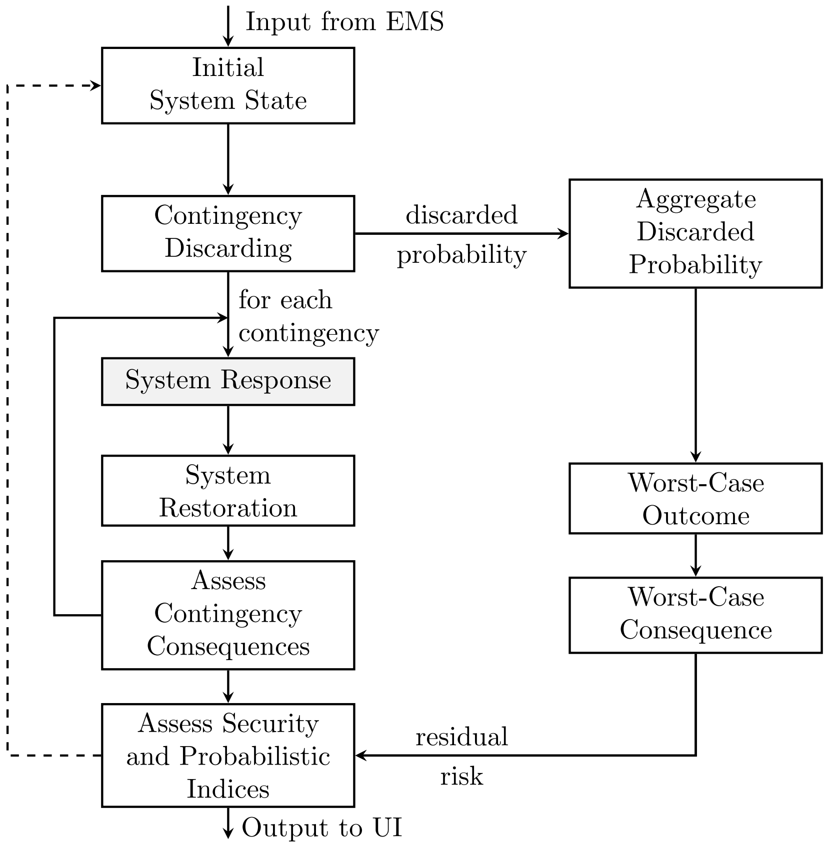


Figure 3.11: Flowchart of the pilot test tool, highlighted in black in Figure 3.10

3.3.2.1 Initial system state

The main element of the initial system state data is a bus-branch model of the transmission system. This model uses live operational data exported from the Energy Management System (EMS), which is then parsed onto a base model of the system in Matpower format [103]. Additional static system information that is loaded to the initial system state relates to under-frequency load shedding, consumer types, Value of Loss Load (VOLL), generator droop settings, failure rate models (as detailed in subsection 3.1.1), minimum component repair times, and security limits.

3.3.2.2 Contingency discarding

Contingency discarding selects a subset of all possible contingencies (mutually exclusive and collectively exhaustive combinations of component failures) based on the discarding principle presented in [61]. Given a limit on acceptable residual risk, this module discards the greatest number of contingencies that keep the residual (discarded) risk below the user-specified limit. The residual risk is calculated using a pessimistic assumption, where all discarded contingencies are assumed to result in worst-case consequences (total system blackout).

Note that risk-based discarding of contingencies is an alternative to selecting an N−1 contingency set, in which the list of contingencies is fixed and the residual risk is variable and hidden.

3.3.2.3 System response

The system response model is shown in Figure 3.12. Many approaches are used in both industry and in the literature to approximate this process, depending on the size of the system, the specific application, and the intricacies of the modelled system. Existing techniques for modelling system response are reviewed in [104]. Commonly such models are based on a mix of Power Flow (PF), Optimal Power Flow (OPF), and heuristic models. For tractability, it is also common to perform system-wide studies using DC or even transport-flow models rather than AC models, particularly in reliability assessment or optimisation. This system response model is an adaption of the model presented in [36]. The major additions to the model of [36] are the simulation of system protection schemes, heuristic dynamic response models, replacing the power flow with a Load Freqency Control algorithm, and performing load shedding based on existing Under-Frequency Load Shedding (UFLS) rules. The functions of the system response model are elaborated on below.

First, the contingency is applied to the initial system state. Immediately following the disturbance, System Protection Schemes and Dynamic Response are simulated. System Protection Schemes and Dynamic Response are treated as a set of deterministic if-then statements that effectively approximate the response of the Icelandic power system following a contingency. This includes automated islanding schemes, undervoltage tripping of industrial loads, and activation of backup generation, for example. Given that these are not time-domain simulations they do not model transient effects and therefore do not accurately reflect the system dynamics in all situations. In the context of real-time reliability assessment, given that results must be produced within a few minutes, it is not presently feasible to perform a dynamic simulation of the system in this context. Potential alternative approaches include offline dynamic simulations of system stability to learn stability constraints or event trees which can be efficiently checked against in an online manner [105, 106].

Load frequency control (LFC) is modelled as a distributed-slack AC power flow. In cases where the LFC model does not converge, Under Frequency Load Shedding (UFLS) is applied until a power flow converges. The load-shedding rules defined in the UFLS match those used in operation. Over-current protection of components is checked after LFC convergence, and tripped as necessary, where unprotected lines are assumed to trip if loaded over 130% of their thermal limits.

In the case that any lines have tripped since the System Protection Scheme model, the algorithm repeats until no further tripping occurs. Once the post-contingency system response has converged, the system is run through an AC OPF with load shedding as an estimate of operator actions immediately following the disturbance. The underlying purpose of this AC OPF is to convert any breaches of operational security limits (e.g. voltage limits) into an equivalent energy not served (if load shedding is necessary) such that the breach can be monetized.

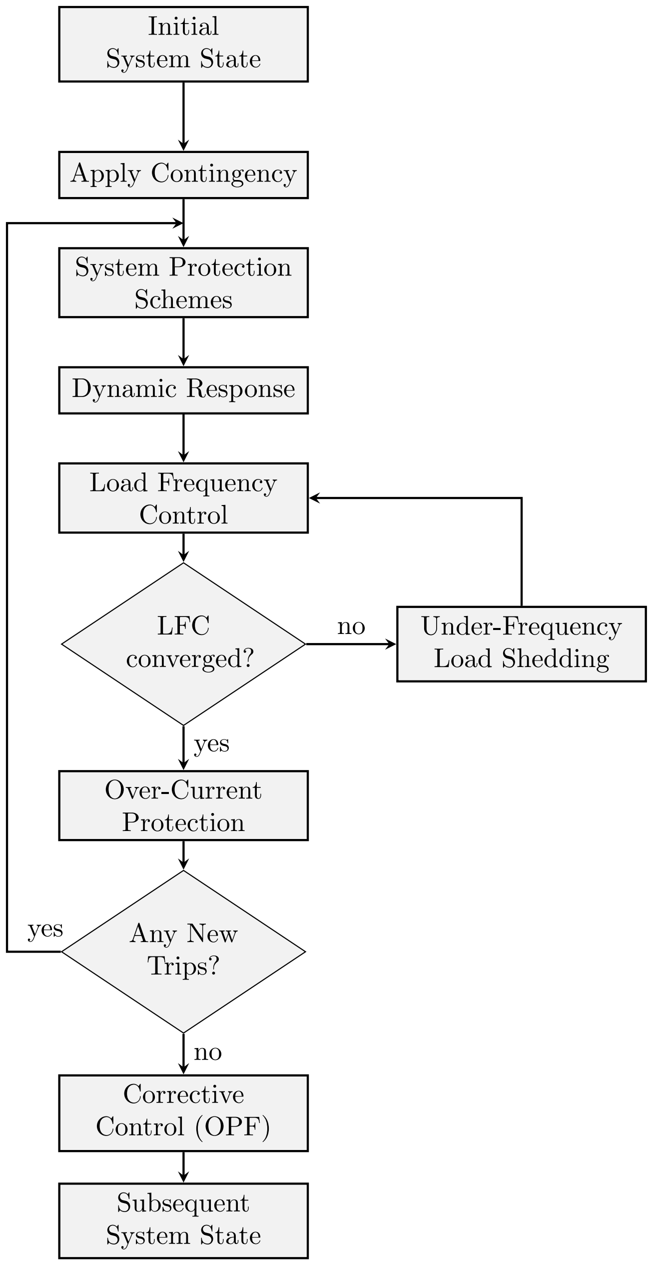


Figure 3.12: Flowchart of the system response model, highlighted in grey in Figure 3.11

3.3.2.4 System restoration

Given a post-contingency system state in which some loss of load has occurred, the system restoration model converts the nodal (or regional) lost load estimate into an estimate of energy not served (ENS). Ideally this process would be a sequential optimisation model that is constrained by human resources and the availability/location of spare parts, and possibly exogenous factors. However, a heuristic model is used in the pilot test similar to the model in [42]. Such a heuristic model takes the post-contingency lost load as a percentage of initial system load, and provides as output the expected outage duration. The lost load at any point of time between the initial disturbance and service restoration is interpolated linearly. Integrating the curve of lost load over time results in an estimate of ENS.

3.3.2.5 Consequence assessment

This module converts the ENS of a particular post-contingency system state into a cost of service interruption, using the approach outlined in [43] and discussed in subsection 2.3.3. The value of lost load curves used to estimate interruption costs are based upon the data found in [45]. The worst case consequences, used in the calculation of residual risk, are estimated by simulating a blackout of the system using the above system restoration model. The modelled consequences can then be aggregated over the set of assessed contingencies, in order to estimate the total expected system consequences in a given operational period.

3.3.2.6 Model outputs

The reliability assessment results in a few key output parameters. The estimates of assessed risk and residual risk are measured in units of thousand Icelandic krona per hour (kISK/hour). Risk is presented in a 'per hour' unit given that the probability of a contingency occurring is defined over the next hour. The assessed risk defines the expected interruption cost for the coming hour, aggregated over the set of assessed contingencies, as well as the event in which no contingency occurs. Note that the cost of control actions are not considered within the risk term.

The residual risk however is a worst-case estimate of expected interruption cost for the coming hour, aggregated over the non-assessed contingencies. The total system risk can be considered to lie somewhere between the assessed risk and the sum of the assessed and residual risk. This implicitly assumes that the potential over-estimate of assessed risk is negligible relative to the risk of non-assessed contingencies. Therefore residual risk can be thought of as an error margin in the risk assessment, which can be reduced by assessing a larger set of contingencies at the cost of increasing computation time.

Reliability is measured as the probability that the system will comply with its acceptability constraints. This is calculated by determining compliance of each individual post-contingency system state, and then aggregating the probability of all those states that comply with the acceptability constraints. As is the case with risk, unassessed contingencies are pessimistically assumed to result in an unreliable system state.

As per consultation with the control room operators, other useful outputs are the number of contingencies in the assessed set of contingencies, the probability of at least one contingency occurring within the hour, the computation time, and indicators of any significant state changes.

3.3.3 Software and hardware

The hardware used to perform the pilot test consisted of a desktop server with a 10 core 2.2+GHz CPU (Intel Xeon Processor E5-2630 v4), 32GB of RAM, a 512GB SSD, and an NVIDIA Quadro M2000 GPU.

The majority of the software was implemented in Matlab, using various power flow functions from within Matpower [103], and using IPOPT as a solver for optimal power flow calculations [107].

3.3.4 Computational performance

The first objective of pilot testing was to determine whether the probabilistic reliability assessment could be performed within a practical time-frame. To assess the runtime of the pilot test, a set of 102 real-time cases were assessed over a set of 1000 contingencies. The mean runtime was approximately 58 seconds, ranging mostly from 55 to 65 seconds, with a couple of outliers, shown in Figure 3.13. The first iteration was 15 seconds longer than the average due to initialization code and checking of network connections. The other significant outlier (at iteration 6) was due to an error in the live system data that reduced the complexity of the system.

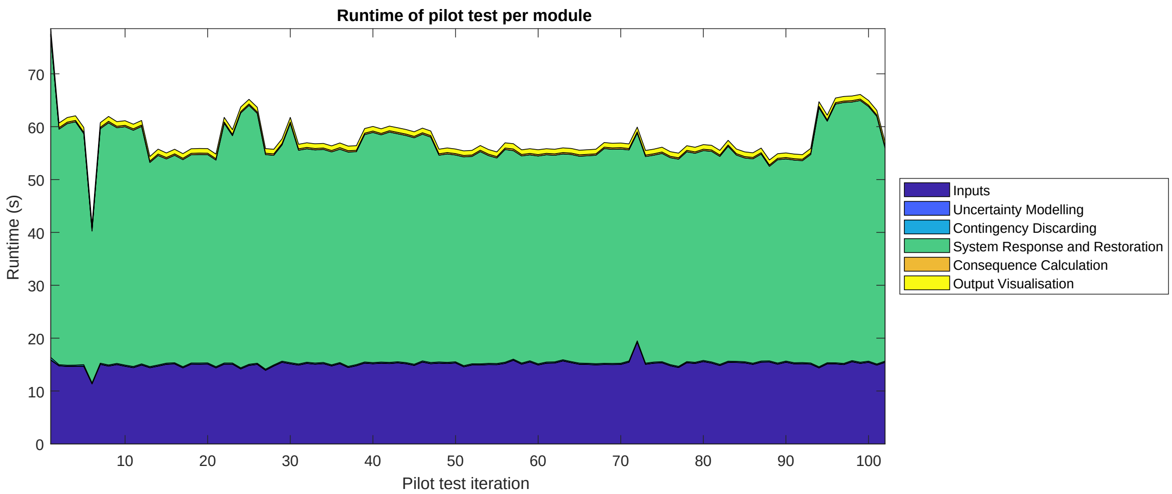


Figure 3.13: Computational performance of the pilot test calculation, with runtime of individual modules shown, for 1000 contingencies. Note that some modules are not visible due to negligibly small runtime

It should be noted that the Uncertainty Modelling (failure rate updating and probability calculations), Contingency Discarding and Consequence Calculation modules did not require significant computational resources. Further, the restoration algorithm used was a simple linear heuristic model, and also required negligible computational resources. Therefore the runtime of the probabilistic reliability assessment is most sensitive to the speed of the system response model. From the runtime assessment, it was computed that the average time required to compute the system response of a contingency was 0.8 seconds. A simple relationship for determining the trade-off between model complexity, computational resources, and contingency list size is described by:

$$\frac{R \times N}{C} + O = T \tag{3.7}$$

where:

R : Computation time for system response of a single contingency [s]
N : Number of contingencies assessed [-]
C : Number of computational threads [-]
O : Time required by other functions (overhead) [s]
T : Target iteration time [s]

This relationship over-simplifies the queueing of parallel tasks and their efficiency, but provides a relatively accurate rule of thumb for reliability assessments. Notably, the overhead of the Icelandic pilot test was quite significant, as highlighted by Figure 3.14. The time required to prepare inputs and to produce visualisations of outputs can be considered as overhead. When these processes are performed, the system response modelling stops, and only a portion of computational threads are used. This highlights the importance of considering software (and hardware) architecture in the design of a probabilistic reliability assessment. If the Inputs and Outputs were processed on a separate machine to the system response and restoration computation, the pilot test runtime could be reduced by approximately 30% (in the case of 1000 contingencies).

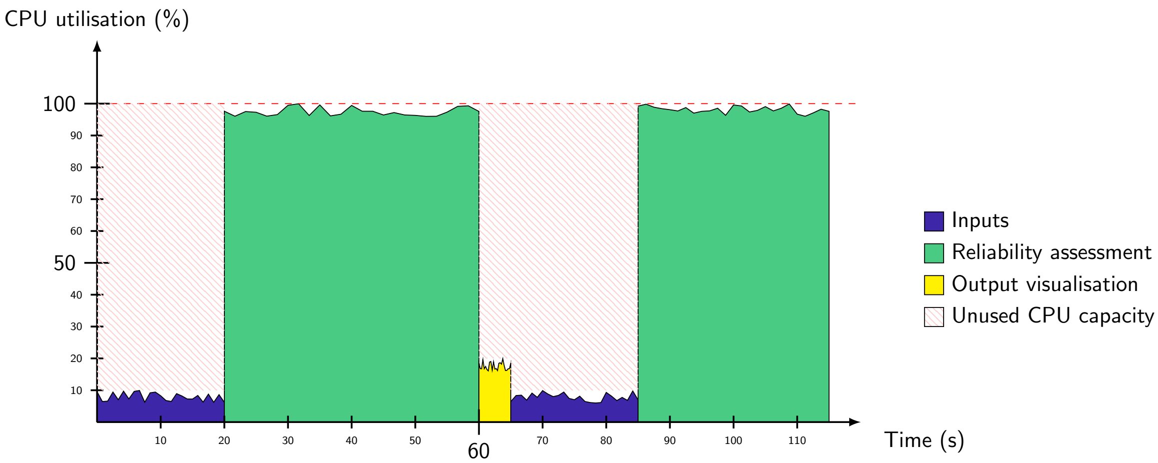


Figure 3.14: Simplified diagram of inefficiencies caused by serial software architecture

Further, both the Input and Output tools (as shown in Figure 3.10) are connected to various databases and services via a network, and are therefore subject to additional bottlenecks. Issues on IT networks, given the present software architecture (see **Paper V** for more detail on software architecture), could further reduce the efficiency with which computational resources are used. For example, a network interruption may result in the simulation failing to access the latest system snapshot, which would presently block any further system response modelling.

It should be noted that, although some effort was put into ensuring the pilot test was computationally efficient, the software was a first-step prototype, rather than an industrial grade tool produced by software engineers. It is anticipated that a professional tool would be an order of magnitude more efficient. Given that the pilot test could perform an iteration of the probabilistic reliability assessment in approximately 1 minute for 1000 contingencies (5-6 times more contingencies than in an N-1 assessment) the approach is considered tractable for the Icelandic case. Given that the system state is unlikely to change significantly within a minute (ignoring the occurrence of any disturbance) the resulting risk assessment will still be relevant to system operators.

Finally, from one iteration to the next, there is often minimal change to the system state, and any changes often only affect parts of the power system. As such, a lot of computational power is spent on re-evaluating the same (or negligibly different) trajectories again and again, and arriving at the same outputs. The assessment could be significantly sped up by identifying familiar system trajectories and recalling prior system responses, especially for small systems with fewer state variables, such as the Icelandic one.

3.3.5 Understanding reliability indicators

The main reliability indicators displayed by the pilot test GUI are shown below in Figure 3.15, over 1000 continuous real-time operational cases (snapshots of the system state, roughly one every 4 to 5 minutes) in a 72 hour period. Note that the average runtime of these results was 4.3 minutes per assessment, which is greater than the runtime reported in 3.13 due to significant inefficiencies in the system state parsing tool during this particular period of time, which were later corrected.

The top plot shows the number of contingencies assessed in each period, ranging from approximately 100 to 500 contingencies, depending on the present weather conditions and number of operational components. That is, contingencies are not considered if components are already out of operation, and during bad weather conditions, more contingencies are considered to ensure the residual risk of unassessed contingencies is below the required threshold (elaborated upon below).

The second plot shows the estimated probability of one or more faults occurring in the next hour, based on the live failure rates (and including the probability of discarded events). The third plot shows the assessed system risk over time, in units of one thousand Icelandic krona (kISK). Note that throughout this section, risk refers specifically to *expected interruption costs*, as described in sub-section 2.3.2. The residual risk is not shown as it is approximately equal to 68 000 ISK in all time steps, and displaying it would unnecessarily stretch the y-axis. Note that a residual risk of 68 000 ISK is equivalent to the residual risk of an N-1 reliability assessment (that is, the average probability of all N-k contingencies, for $k > 2$, multiplied by the worst-case consequence of a total system blackout).

The final plot shows the probability that the system state will be acceptable after one hour. The probability of an acceptable system state could be interpreted as an indicator of system reliability. Notably, this is highly dependent upon the definition of acceptability constraints.

The main critique from operators of this type of output is that it lacked transparency. That is, when risk increases, or more contingencies are added to the contingency list, the reason is not immediately clear. For example, an increase in risk could be due to changing in failure probabilities (e.g. worsening weather in part of the country), changing consequences (e.g. due to changing system loads causing congestion, or due to a component outage), or both. Whilst such high level indicators are useful for general indication of the system state, they are not particularly useful for fast decision making. Similarly, such figures are not useful for debugging the pilot test and its models.

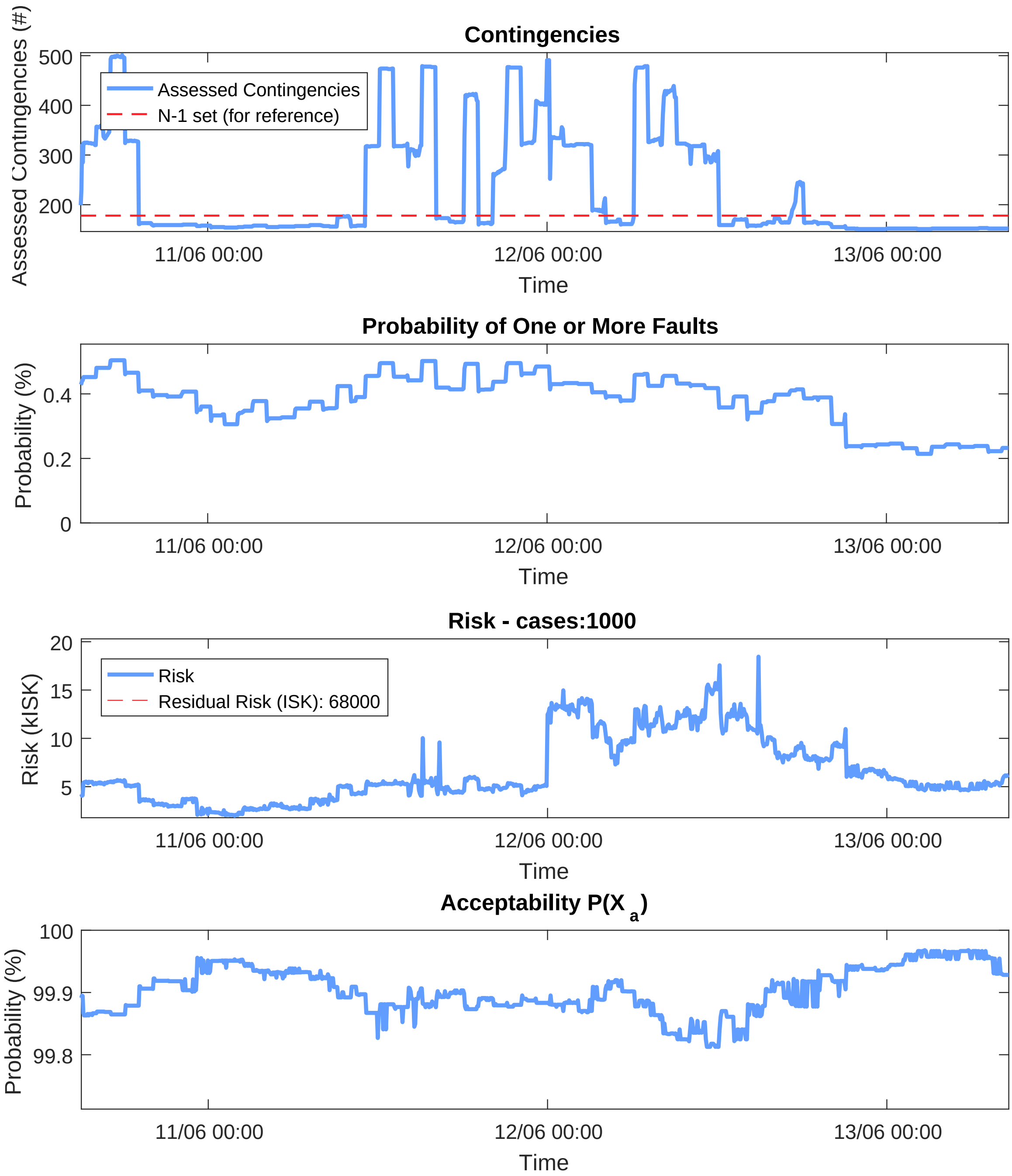


Figure 3.15: High-level outputs of the pilot test

In order to provide transparency to such reliability assessments, there is a need to show users only the important information, but also allow for users to delve into specific details where necessary. The work by [108] on dynamic decision-event trees suggests a possible approach. Just as they suggest the visualisation of real-time, dynamic events as decision-event trees, there is a need to describe model predictions in a similar way. Behind each quantitative output of a reliability assessment, there should be a narrative (qualitative) description of how the models arrived at that outcome. Such an approach was attempted in the pilot test, through the use of a detailed log for each contingency, available from within the user interface. An example of the type of information captured in these logs is shown below:

```
        Contingency: OHL1 + OHL2

Occurrence:
        Event probability in the next hour is 0.00002%

System response:
        System split into three islands (Region A, B and C).
        Region C SIPS activated.
        Region B has no swing bus.
        Region B swing bus changed to bus BUS1.
        Region C does not have enough generation capacity for LFC, running UFLS.
                UFLS shed 0.6 MW at BUS2,
                UFLS shed 0.1 MW at BUS3,
                UFLS shed 2.0 MW at BUS4,
                UFLS shed 1.1 MW at BUS5,
        UFLS successful in Region C, LFC converged.

System restoration:
        0.08 MW of primary load was shed (0.08 MW at BUS3)
        3.83 MW of secondary load was shed (2.04 at BUS4)
        Service expected to be restored after 0 hour(s) and 16 minute(s).
```

Pseudo-algorithm 3.1: Example reliability assessment log for a contingency

From such information it is possible to quickly identify if the modelled sequence of events matches operator descriptions of events. In cases where the descriptions do not match, this allows for fast identification of specific modules that may be inaccurate or missing. Such an approach was used throughout the pilot testing for debugging purposes. Notably the information captured by these logs is incomplete, and does not describe why the contingency may occur, what generator outputs change as a result of LFC, or any specific details on the logic that caused protection devices to trip.

As mentioned in Section 2.6, a key ingredient of the GARPUR framework is the discarding principle. This principle discards contingencies and exogenous scenarios (or trajectories), with the requirement that the aggregate risk associated with discarded events is below a user-defined threshold. Therefore, by varying the residual risk target, the TSO can vary the size of the contingency list.

For this pilot test, the residual risk of unassessed contingencies was estimated using the pessimistic assumption described in [61]. That is, that unassessed contingencies are assumed to result in worst-case scenario consequences, or have maximum severity. As a result, an assessed contingency will always have a consequence less than or equal to that which is assumed when it is unassessed. This results in the residual risk being over-estimated. In the following results, the term 'residual risk' actually refers to the estimate of residual risk, given that the actual residual risk is unknown (would require assessment of an intractable number of contingencies).

Figure 3.16 shows the trade-off between selecting some residual risk target on the size of the contingency list and on the assessed risk. Expectedly, as the residual risk target is decreased, the assessed risk increases, given that previously discarded events are now included in the reliability assessment. Note that, given the pessimistic assumption that discarded events result in worst-case consequences, the increase in assessed risk is far smaller than the decrease in residual risk. As noted on the figure, the left most point (the most accurate assessment) corresponds to a residual risk of 5 kISK and assessed risk of 16.7 kISK. As such, if all 10^{48} contingencies were included in the contingency set (residual risk target equal to 0), the assessed risk would converge to a value between 16.7 and 21.7 kISK. This can be compared to assessing the N-1 contingencies, which would give an assessed risk range of 16.3 to 150 kISK (where the latter term is equal to the assessed risk plus the residual risk). These results show quite clearly that the pessimistic approach to estimating residual risk can be drastically improved upon. Note that a perfect residual risk estimation algorithm would result in the top plot of Figure 3.16 being a line with a slope of -1 equivalent to: $R_{assessed} = R_{actual} - R_{residual}$.

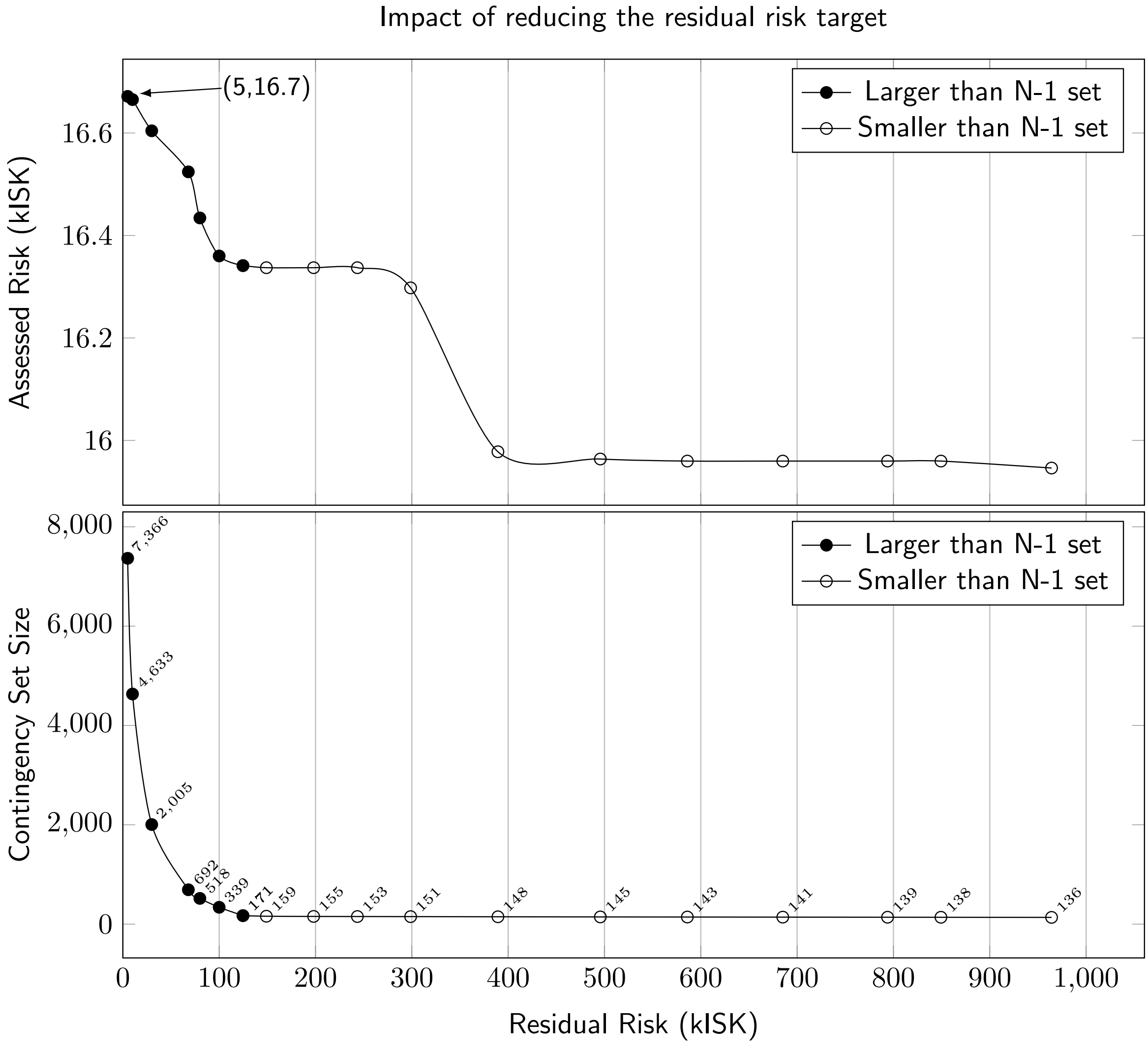


Figure 3.16: Impact of reducing the residual risk on assessed risk estimates and on contingency list size

Reducing residual risk comes at a cost of increasing the contingency set size, as shown by the lower plot in Figure 3.16. Recall that each additional contingency requires 0.8s of a single CPU-core to compute the system response. With this in mind, it is clear that reducing residual risk below some threshold will result in significant computational costs, with very little improvement in the accuracy of the reliability assessment. Choosing a residual risk target above a value of 120 kISK in this case would be inefficient, given that it reduces the accuracy of the reliability assessment, without any significant reduction in contingency set size. Similarly, choosing a residual risk target below 5 kISK would greatly inflate computation time, with minimal affect on the assessed risk. It is not possible to determine an optimal residual risk target from this single figure alone, given that the shape of these figures will change as system state and exogenous variables change. Further work is required on both the improvement of discarded risk assessment and of the optimal setting of discarded risk targets. In the meantime, it seems practical to set residual risk targets as low as reasonably possible, given the limits of available computational hardware and how long operators are willing to wait for results.

3.3.6 Sensitivity of outputs to variations in input data

In order to further study the importance of weather-dependent failure rate models on the risk assessment method, a year of hourly system states were assessed for the following three cases:

- Weather Only: A constant system state (all components in operation) with real-time weather-dependent failure rates
- State and Weather: real-time system state and weather-dependent failure rates
- State Only: real-time system state data, with static failure rates

The resulting risk distribution is shown below in Figure 3.17 (similar to those shown in Figure 3.6). The 'State and Weather' case (red distribution) represents the normal pilot test model. If only a static system state is used (ignoring maintenance, generation and load change, and other changes in system state), risk is significantly under-estimated. The resulting distribution (blue distribution) however provides a system-wide measure of how weather is affecting the system. Notably, by neglecting weather data (yellow distribution) the median risk is over-estimated. Inclusion of weather data also results in an increase in high-risk outliers, shown by the bump in the distribution between 80 and 100 kISK per hour. These results support those from the study in **Paper II**, shown in Figure 3.6. It is important to note when comparing these results that both studied the same system, but using different failure rate models, different system response models, and data from different years.

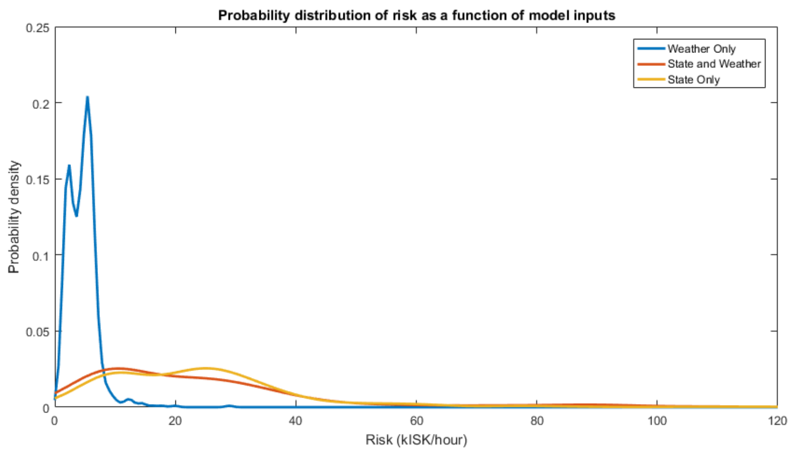


Figure 3.17: Distribution of risk given changes in input data (legend refers to which data is dynamic)

3.3.7 Visualising the dimensions of risk

To provide additional transparency to the indicators presented in subsection 3.3.5, a risk map visualisation was created as shown in Figure 3.18, similar to the diagram shown in Figure 2.6. The axes of the figure are probability and consequence (interruption costs) of specific contingencies, which are represented by labelled squares on the diagram. As in (1.1), risk is probability multiplied by consequence, and therefore it is possible to draw iso-risk lines on such a diagram. Visualising risk this way allows for different types of contingencies to be delineated. For example the High Impact Low Probability (*HILP*, also known as High Impact Low Frequency or HILF events) contingencies are evident on the upper-left of the figure, in this example Contingency A. Similarly, Low Impact High Probability (*LIHP*) events, like Contingency C are visible in the lower-right. Such events are discussed in [109].

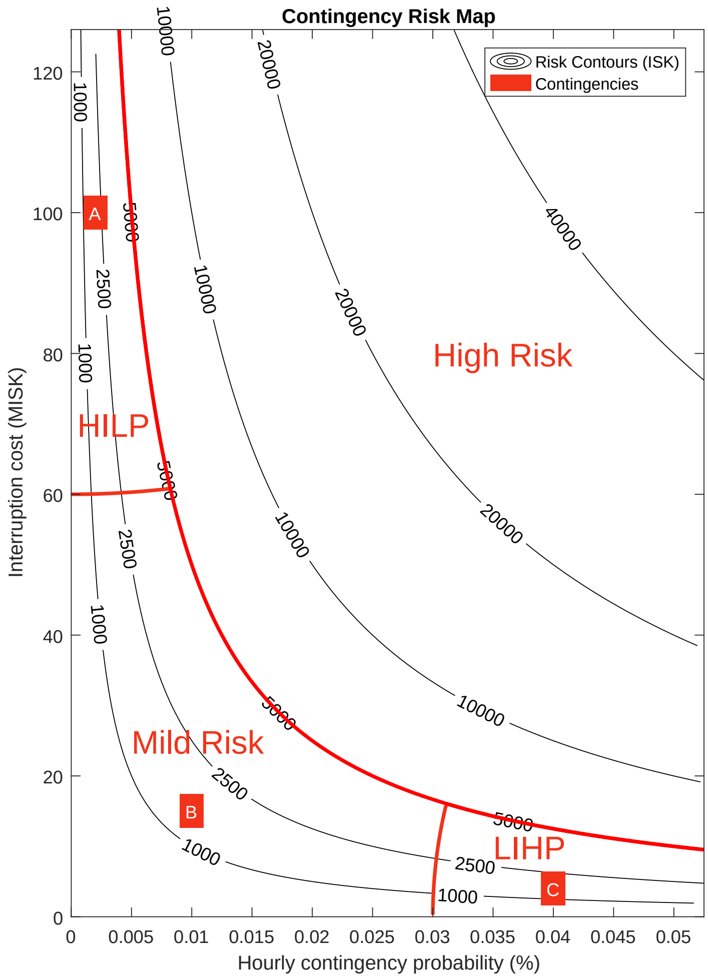


Figure 3.18: Hypothetical risk map, highlighting the differences between three hypothetical contingencies of equal risk (MISK: Million ISK)

Notably, contingencies which are have no consequence or are not possible, do not appear on the risk map as they have zero consequence. Residual risk also does not appear on the risk map, as it is an aggregation of a large number of contingencies, many of which are extreme HILP events (near zero probabilities, and large consequences).

The risk map can be further enhanced by including temporal changes, as in Figure 3.19, showing how an in increase in contingency probabilities affects contingencies A, B, and C. Such an increase in probability may occur due to worsening weather, for example. Similarly, an upward movement of a contingency over time may imply increasing load, or some other change in system state that exacerbates the consequences of the contingency. Including such temporal information into the figure allows for fast appraisal of the risk state of the system by the operator, and allows them to notice alarming trends or unusual changes rapidly.

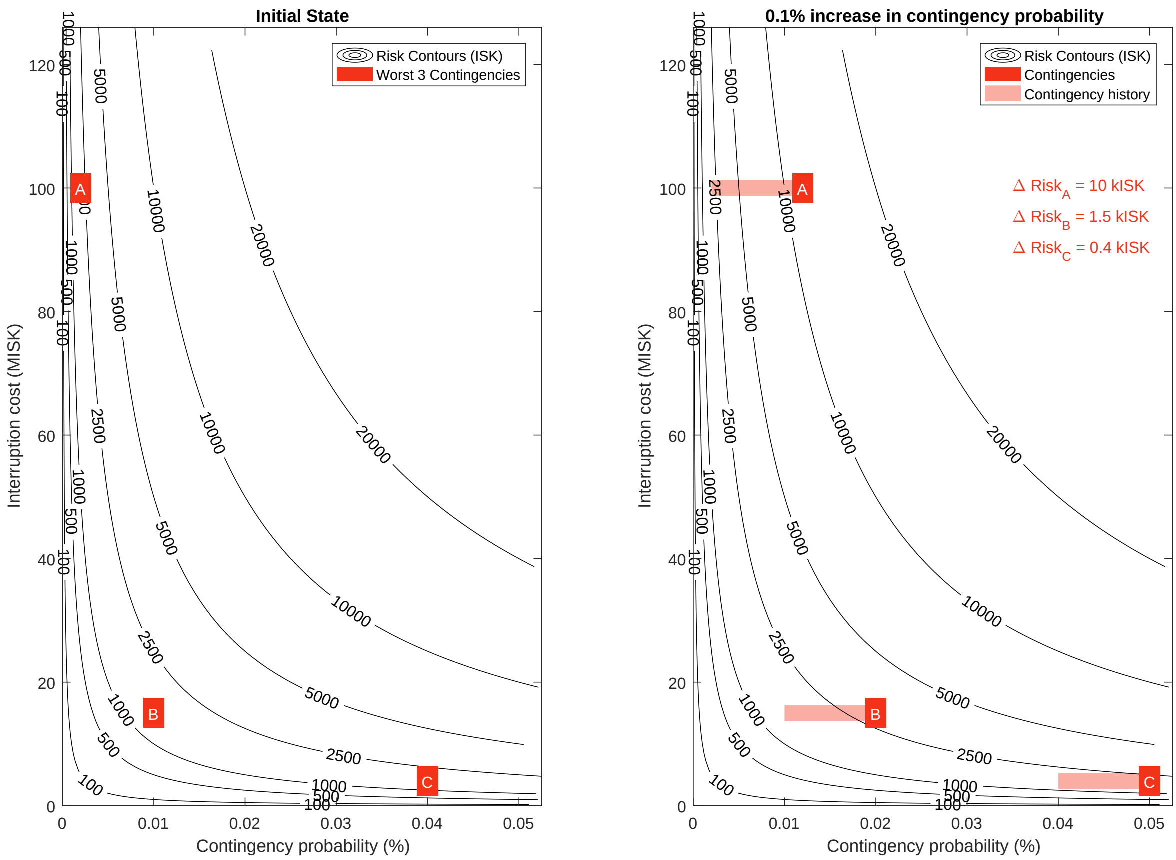


Figure 3.19: Effect of increased probability of contingencies (e.g. due to weather)

Figure 3.19 also shows that worsening weather affects contingencies differently, given their initial position. A HILP contingency has a far greater potential to increase in risk due to worsening weather than a LIHP contingency. Similarly, a LIHP contingency can quickly become a high risk contingency if exacerbated by a change in the system state (e.g. due to a nearby maintenance activity, or a spatial shift in generation/reserves).

For HILP events, the risk contours are almost parallel to the vertical axis. Therefore, the risk of a HILP event is more sensitive to changes to its probability of occurrence than to its consequence. As suggested in [104], decisions related to HILP can be based on credible probability ranges, or on whether there are factors present that may increase the probability of occurrence (e.g. presence of a storm causing an increased probability of a common mode of failure for multiple grid components). In any case, the degree to which HILP events should be mitigated depends highly upon the TSO's ability to estimate its credibility, either using probabilistic or deterministic models. Such events may be better managed through system theory approaches, such as STPA (Systems-Theoretic Process Analysis) described in [110].

Conversely, the risk of LIHP events is primarily sensitive to estimates of their consequences. For example, failure to model significant dynamic instability due to contingency C would result in it being misclassified as a LIHP event rather than a high risk contingency. The risk of such events is commonly managed through deployment of System Integrity Protection Schemes (SIPS), that aim to contain the disturbance to a small region, or to reduce the service outage times.

This visualization approach is shown using operational data in Figure 3.20, covering a period of 6 hours. The top 10 contingencies are shown as coloured boxes with encrypted names/labels, and each with a unique colour and a matching trail colour. Lower risk contingencies are shown as blue scatter points (mainly in the lower-left of the figure) and contingencies with no risk are not shown on the map. The position of each contingency is based on the pilot test estimates of its hourly probability of occurrence and estimated interruption cost. Note that these values are the results of a prototype model and are not claimed to be official estimates of system risk.

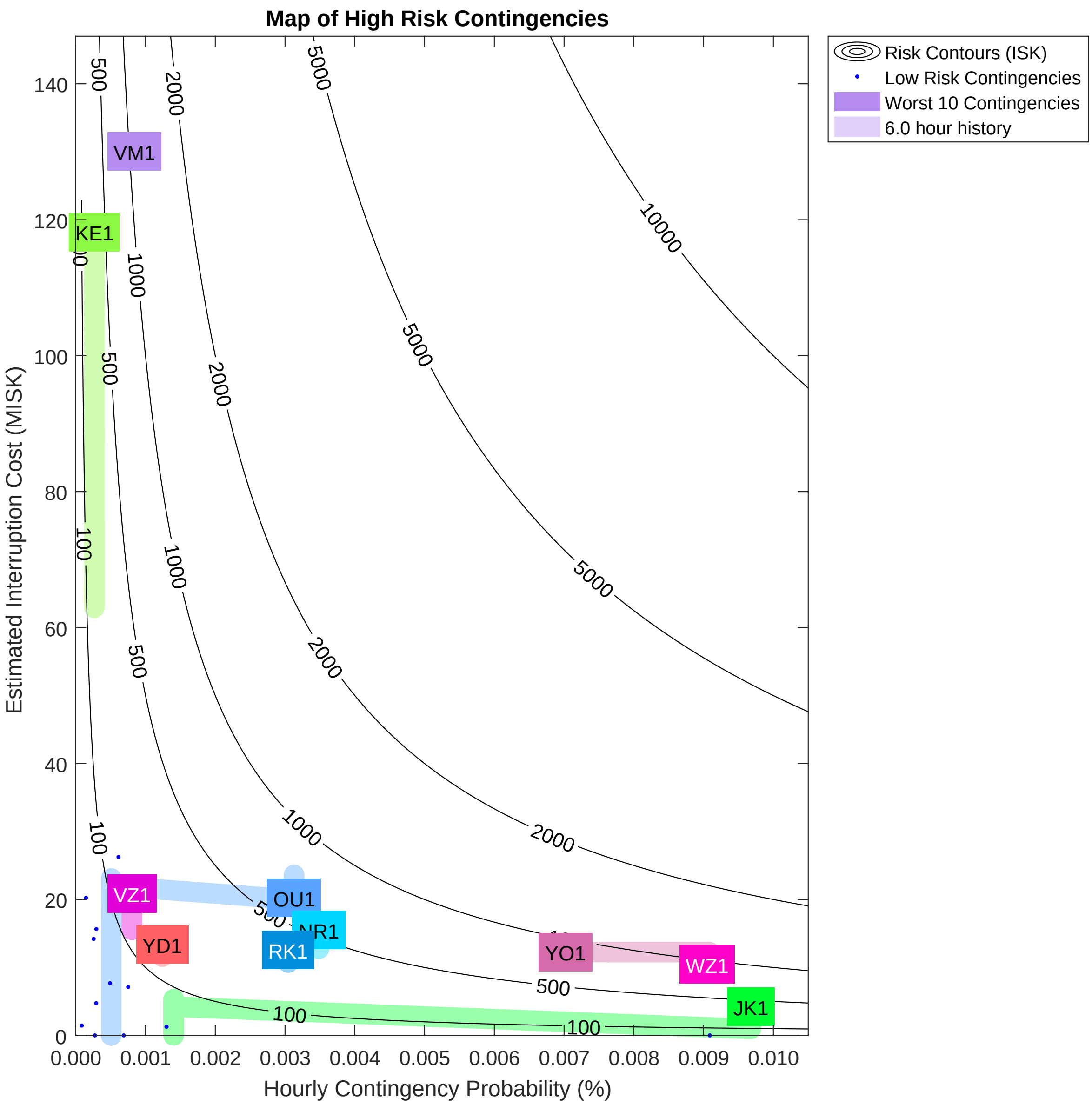


Figure 3.20: Risk map from the pilot test, with encrypted contingency labels, covering 6 hours of operation in mid-June 2017. The top 10 contingencies are shown with labels, with lower risk contingencies shown as blue points.

The contingencies are distributed roughly below the 1000 ISK risk contour over the observed period. This suggests that the existing approach to risk management methods has adequately identified and addressed major risks to the system during normal operation. This however does not apply in all cases. Extreme weather events may significantly increase the probability of some contingencies increasing by a couple of orders of magnitude. Similarly, certain planned maintenance outages may greatly increase the consequences of some contingencies. As such there exist operational states, most of which are unavoidable, which can cause contingencies to become high risk. Approaches such as the one studied in this thesis may provide useful information for weighing the risk of such rare events with the costs of mitigating them.

3.3.8 Support to decision-making

The final objective of the pilot test was to show how such a probabilistic approach to reliability assessment could assist in decision making. As discussed in **Paper V**, a sequence of operational system states over an 8 hour period were assessed for the following three cases:

- Base operational case;
- Base case + the removal of SIPS in a particular region of the power system;
- Base case + the redispatch of 30MW from West Iceland to East Iceland.

The impact of removing SIPS is shown in Figure 3.21, and the impact of redispatch is shown in Figure 3.22. In all operational states, the assessed risk of the system is reduced by the inclusion of SIPS. The decrease in risk due to SIPS varies over the operational period, given that it protects the system against a subset of events, and therefore its benefit varies as the probability of this subset of events vary.

The benefit of redispatching generation during this operational period however is not as clear. By redispatching generation from the West of Iceland to the East, this would increase reserve capacity in the West of Iceland but decrease it in the East. As such, the value of moving generation between two generators depends upon the production and commitments of all other generators, as well as the location of loads. Notably there are a few cases where such a redispatch action would result in a decrease of risk by more than 1000 ISK for a sustained period of time (e.g. 5:00 to 6:00 and 10:30 to 11:30).

Both of these comparisons show that the pilot test implementation could form a basis for comparing different types of decisions based on their overall impact on system risk. Comparisons such as these, however, face the same problem discussed above of a lack of transparency. In each case, when comparing two actions, any change in reliability indicators should be accompanied by a concise explanation. Similarly, comparing and choosing between control actions should not be based solely upon risk, but also on system reliability (compliance with acceptability constraints).

3.3.9 Implications for other TSOs and the GARPUR framework

The pilot test performed at Landsnet showed a first-step implementation of the GARPUR methodology on a real-world power system. The process of implementing this pilot test

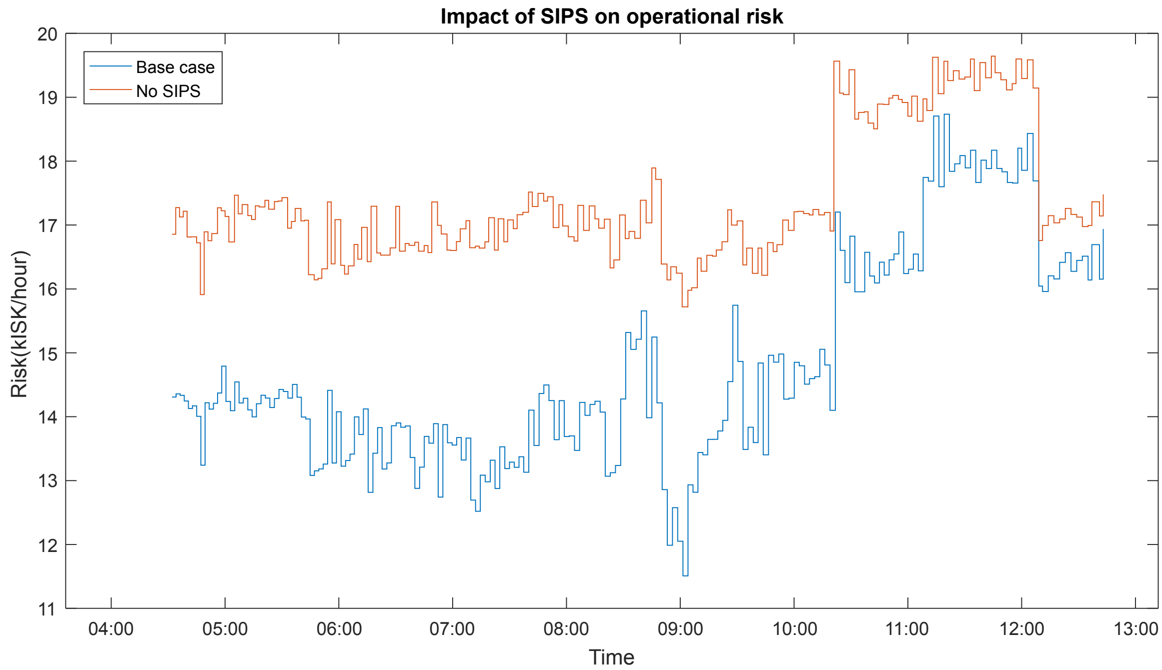


Figure 3.21: Impact of implementing SIPS in a particular region on risk during operation

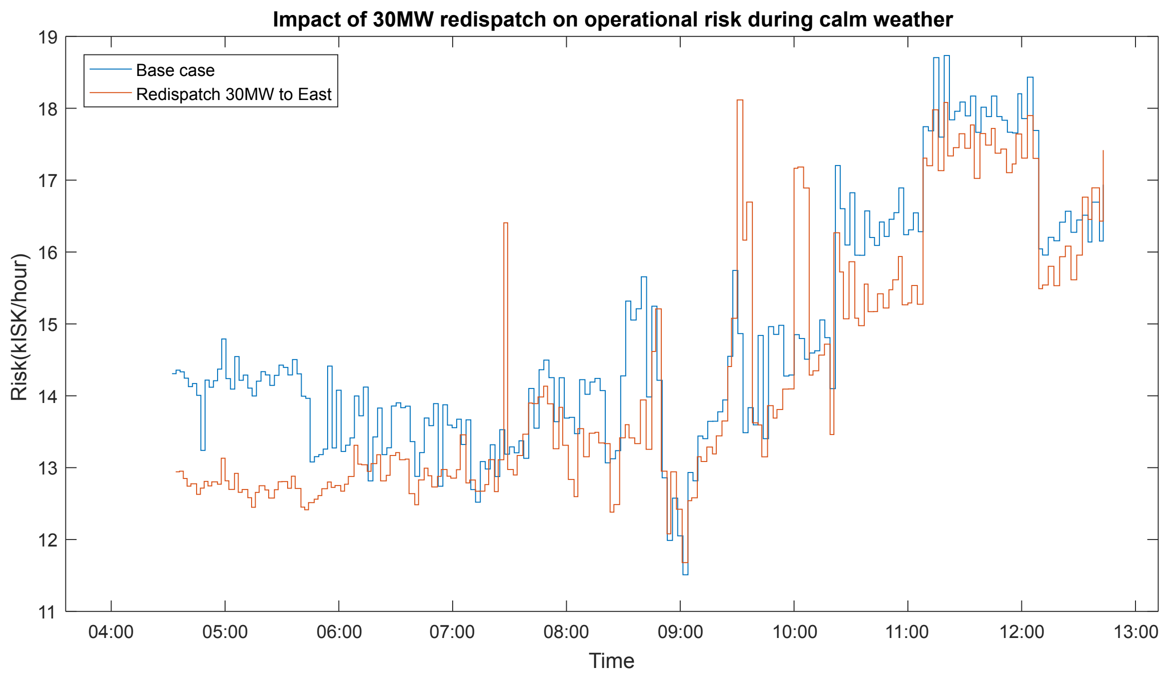


Figure 3.22: Impact of redispatching generation on risk during operation

provided a number of insights specific to the Icelandic power system. Yet the intent of the GARPUR project is to produce a generally applicable approach to reliability management, and thus this section attempts to provide insights useful to both other TSOs and to understanding/developing the GARPUR framework further.

Control room staff at Landsnet often commented upon the accuracy of the system response model, quickly noticing when the method was over/under-estimating the consequence of particular contingencies. Often such errors were due to the system response model not being able to describe the dynamics of the Icelandic power system, as transient instabilities were not modelled accurately (the pilot test was not a time-domain simulation). The Icelandic system has relatively low inertia and disturbances that affect a relatively larger proportion of the system (by load), compared to higher inertia systems for which frequency perturbations are more benign [111].

For larger systems in the short-term, the system response model required to estimate the eventual consequence of given contingencies may be simpler, and thus less computationally intensive. This may outweigh (or at least reduce) the increased computation required for larger systems with more interconnected components and hence more contingencies to assess, and the far larger cascading events that they may experience. In the longer term, the gradual loss of inertia on the European power system (due to increased penetration of dispersed wind and solar generation) may necessitate more detailed system response models.

In some cases, TSOs may only perform N-1 contingency assessment through simple load flow calculations, and determine whether any branch loadings or voltage limits have been breached. That is, an assessment that does not attempt to estimate the quantity of lost load, given that securing against events that breach security limits will result in security against events that cause lost load. Existing cascading simulations may not be presently tractable in practice [33], to the extent that they provide a useful amount of fidelity without discarding too many system trajectories. Notably, a cascading simulator is not required to calculate initial contingency probabilities, and therefore it may still be feasible to estimate reliability (the probability that the system state is acceptable at the end of the hour). Research, development and validation of more efficient and effective cascading models (as well as high-level statistical models) should continue to be supported, as their existence and adoption will assist in the adoption of risk estimation by TSOs.

In cases where the TSO is capable of efficiently estimating lost load for contingencies, then existing methods can be quite easily extended to a probabilistic N-1 reliability assessment, following the methodology presented in [101]. In such cases, the difficulty is likely to come from determination of failure rates, particularly the creation of time-varying failure rate models. The approach presented in [112] may give a basic starting point, by modelling time-varying phenomena using credibility functions.

The outputs of the GARPUR methodology are measures of risk, residual risk and reliability, among other fine-grained indicators. Feedback from control room operators stated that these notions are not necessarily clear, in that the true value of these indicators cannot be known. Whilst the actual value of the indicators was not immediately useful to operators, the relative change in the indicators provide indication of significant changes to the state of the system or to the external environment.

Similarly, the connection between residual risk, risk and contingency list size was not clear until a sustained period of pilot testing. That is, the setting of residual risk is a somewhat arbitrary value (given that the value of risk is not necessarily trusted or validated), which then acts as a sensitivity for the contingency list size. In the pilot test, a value of residual risk was chosen such that the average contingency list was equivalent to the N-1 contingency set. At any point in time the number of contingencies would vary from half the size of the N-1 contingency set, to an order of magnitude greater, depending on the weather. Notably, if residual risk is set to a reduced value, then the calm weather contingency set will increase slightly, but the bad weather sets will increase dramatically in size, as shown in 3.16. Enforcing strict limits on the size of the contingency list (as done in the pilot test, by setting a maximum contingency list size of 5000) then results in significant variation to residual risk. As such, the value of residual risk (relative to the set limit) can be used as an alarm indicator for situations in which the system is more vulnerable.

Finally, the setting of a reliability target is not suggested in early implementations of probabilistic reliability assessments. During the pilot test, the probability of the system being in an acceptable system state throughout the next hour varied slightly in most cases, but with significant reductions during storms, particularly when coinciding with maintenance. By reducing the percent value of reliability into a boolean indicator of whether it is above or below the reliability target, important information regarding the degree of system vulnerability is lost. Note however that reliability targets are useful (and necessary) when searching for optimal control actions within the GARPUR framework.

Chapter 4

Conclusions and future work

This chapter provides concluding remarks on the research, its usefulness, and retrospective insights. This section also includes a description of suggested future work to continue progress towards the development of practical approaches for probabilistic reliability assessment and control.

4.1 Conclusion

The two hypotheses of this research, stated in section 1.3, were:

1. The weather-dependence of faults can and should be taken into account when quantifying power system risk.
2. It is feasible to operate power systems with a quantitative risk-based approach;

The results of this research are described below in terms of these two hypotheses.

4.1.1 Hypothesis 1: importance of modelling weather-dependence in risk assessment

The link between threats and failures was established in subsection 3.1.1 and in **Paper I**, by analysing Icelandic data on overhead line failures and historical weather data. This link is shown in Figure 3.2, showing that unplanned outages attributed to weather-based threats in the last 10 years occur in Winter, whilst other outages seem to be uniformly distributed over the year. The discrete and continuous failure rate models that are then elaborated upon also show, through correction factors, that there is a positive correlation between increasing wind speeds and the occurrence of outages. The grey curves in Figure 3.5 show that all overhead lines have an exponential relationship between wind speed and correction factor. These models show that weather-dependent failure rate models can be created for the Icelandic transmission system using existing data.

These weather-dependent failure rate models were applied to both existing commercial reliability assessment software, and used within the pilot test of the GARPUR methodology. The study in **Paper II** showed that using weather-dependent failure rate models in risk assessments result in the median risk decreasing in both Winter and Summer (by approximately 33%), with high risk outlier events occurring in Winter. Further, it was shown that these models have significant implications for assessing risk of maintenance activities, where the risk of a studied Summer maintenance activity was over-estimated by a factor of 2.4 when using static failure rate models. These results strongly suggest that such models should be used in real-time and short-term quantitative risk assessments.

The pilot test also showed the importance of using weather-dependent failure rate models, as described in subsection 3.3.6. The results of **Paper II** discussed above were reflected in the pilot test, which used different weather-dependent failure rate formulations and risk assessment software, and used input data from different years. Figure 3.19 showed the importance of modelling weather-dependence, particularly for HILP contingencies. Figure 3.20 highlights the importance of considering weather-dependence, showing how the top 10 contingencies change in risk over a six hour operational period. By failing to consider weather-dependence, a number of changes in risk would have been neglected.

This research therefore provided strong arguments that weather-dependence can, and should, be modelled in quantitative risk assessments. More generally, as outlined in subsection 3.1.3, these models should be generalized as threat-dependent failure rate models.

4.1.2 Hypothesis 2: feasibility of the risk-based approach

Section 3.2 formulated a quantitative risk-based approach, as part of the GARPUR project. In order to show that the approach is feasible, the model was pilot tested on the Icelandic power system. Subsections 3.3.1 to 3.3.3 describe the pilot test methodology in detail. This shows that the existing data at the Icelandic TSO is sufficient for computation of quantitative risk assessments.

The runtime of the pilot test was shown to be approximately 1 minute, on average, for a set of 1000 contingencies (5-6 times more contingencies than in an N-1 assessment) in subsection 3.3.4. It was also shown that inefficiencies in the software architecture of the pilot test severely impacted the runtime, showing that it could be easily reduced by 30%. Improvements to the pilot test code, particularly system response modelling, may also provide significant speed-ups. Therefore, the method is considered to be tractable for the Icelandic power system.

As noted in subsection 3.3.9, control room staff assisted in ensuring the system response model provided realistic estimates of contingency consequence. Notably the risk outputs often agreed with staff intuition on what the highest risk contingencies are. Therefore the presented framework of models can result in quantitative risk estimates that adequately match expert opinion. Replacement of the prototype models within this framework with validated industrial tools and software should improve these outcomes further.

Given that the pilot test of the GARPUR approach computed in a reasonable time frame, and produced results that were accepted by operational experts, the risk-based approach may be considered as feasible for the Icelandic transmission system. It is likely that the approach is also feasible on other small isolated systems, or applied to regional networks (regional sections of larger TSO networks, or distribution networks). For larger TSOs or for multi-TSO networks, the approach may also be feasible, but with some concessions on the accuracy of the system response model and may require larger computational resources, and improvements both of the software architecture and of the various algorithms/tools. In other words, the barriers are technical and are not due to any theoretical issue with the GARPUR framework itself.

Real-time probabilistic reliability assessments, as shown in this study, may provide significant improvements to control room operation and management. Primarily, the assessment aggregates information related to threats, system state, and contingency assessment, into a single overall indicator of system reliability. Such an indicator allows for the vulnerability of the power system to be assessed at a glance. Visualisations developed during this research also allow for particular contingencies to be assessed at a glance, including how they have varied over the last few hours. Simple visualisations like the risk map can quickly prompt operators to begin investigating why the probability has increased (encouraging them to look at weather reports when relevant) or why consequences have increased (encouraging them to run more detailed simulations or to discuss how to control such an event).

4.2 Future work

The following subsections provide short descriptions of suggested future work on research and development. Each sub-section provides a short description and justification for a specific topic. These topics are not exhaustive, as common topics are excluded for brevity, such as the important topic of probabilistic forecasting.

4.2.1 Modelling uncertainty

There already exists a significant body of research into the topic of probabilistic forecasting (wind power, load, etc.) as well as capturing spatio-temporal and modal correlations. Firstly, prototypes similar to the one modelled in this research should be applied to small power systems that also contain considerable wind power and/or load uncertainty, such as the power system of Ireland. In doing so, further insights may be gained regarding residual risk. There is a need for the development of trajectory discarding methods that can jointly consider exogenous scenarios and contingency probabilities. For example, a discarding algorithm that captures that significant wind power forecast errors may also coincide with increased contingency probabilities, in cases of high wind.

4.2.2 Residual risk estimation

As shown in 3.16, assuming worst-case consequences for contingencies in the discarding algorithm (closely related to contingency filtering) was largely inaccurate. Overestimating the residual risk during discarding results in a larger contingency list, and thus greatly impacts the runtime of the reliability assessment. It is likely that significant efficiency and accuracy gains can be made with discarding algorithms.

Firstly, worst-case consequence estimates should be contingency specific, rather than assuming that all discarded contingencies result in a total system blackout. For example, in particular regions of the Icelandic transmission system, it is inconceivable that any number of combined faults (N-k, where $k \geq 2$) contained within that region could propagate into a total system blackout. Therefore the worst-case consequence for contingencies within that region could be reduced from ‘total system blackout’ to ‘total region blackout’.

Secondly, the approach applied in this research assumes that every operational state is unique, and therefore all contingencies must be re-evaluated at each moment in time. Figure 3.20 represents a six hour period operation, representing more than seventy risk assessments. During this period, the consequences of contingencies ‘VM1’, ‘YO1’ and ‘WZ1’ were unchanging, with a few others only changing slightly. This implies that these contingencies were not sensitive to the changes in system state over this 6 hour period of operation. By determining the sensitivity of various contingencies to changes in system state, it should be feasible to determine whether a contingency needs to be reassessed, or whether the last consequence estimate is sufficient. By building memory into the approach, these contingencies may have only needed to be simulated once during this operational period, freeing up computation time to assess other contingencies which would have otherwise been discarded.

Finally, it is evident from Figure 3.16 that most discarded contingencies have a negligible impact on residual risk. If the pessimistic assumption employed in this research was correct, the top figure would show a straight line with a slope of -1.

Clearly this is not the case. The relationship between *Assessed Risk* vs *Residual Risk*, and how sensitive this shape is to changes in system state or exogenous variables, should be studied further to determine better heuristics for residual risk estimation. It is likely that there exists a relatively simple function (i.e. polynomial or exponential) that could provide an upper-bound to the *Assessed Risk* vs *Residual Risk* curve.

4.2.3 Threat assessment tools

In order to make significant progress in threat-based failure rate modelling, there is a need to retain and record data related to threats, both during contingencies and during uneventful periods. Until such data is collected, it will be difficult to progress past coarse state-based failure rate models. To encourage collection of this data, an intermediate step may be to produce threat assessment tools. That is, GIS tools that aim to identify areas in which threats are credible, to display in the control room. Such tools may reduce pressure on operators to identify threats (or external experts to notify of threats) and may improve response times, and would result in the definition and collection of data required for improved failure rate modelling. As threat assessment tools are created, they can be integrated into a threat-based failure rate framework, as described in subsection 3.1.3.

4.2.4 Missing data

There are a few types of missing data that limit the ability of TSOs to implement the GARPUR methodology to the intended degree. Primarily, this data relates to corrective control actions, and the various types of corrective control failure [113]. Further, for the development of proxies of real-time operation, data on corrective control needs to be connected to the system state and exogenous state data.

Data related to system restoration may also be lacking. Service outage times and component restoration times are often recorded, but there is a lack of data on the time taken for various phases of restoration processes, or on the exogenous processes that may delay restoration. The lack of such data makes it difficult to estimate the delay in service outages that may result from a severe storm, and therefore may result in significant errors of estimating ENS.

4.2.5 Probabilistic test cases

In order to develop practical approaches for probabilistic reliability assessment and control, there is a need for more detailed test cases for the sake of academia. Present test cases lack data or algorithms for the modelling of exogenous threats, and as such, limit the ability to study variable failure rates. At best, test cases include static failure rates or state-based failure rates that have been pre-computed. Similarly, test cases lack information on uncertainty, and often uncertainty is injected by researchers, each using different values. Importantly, there is also a need for test cases that are based upon node-breaker system models, due to shortcomings in bus-branch models, such as their inability to reflect substation topology as noted in [27].

4.2.6 Capturing systemic failures

The GARPUR approach is based upon defining a set of events, and estimating their probability of occurrence and consequence. Such event based approaches are capable of estimating reliability in terms of component failures and of unexpected exogenous changes (e.g. forecast errors). Yet failures may also occur due to failures in ICT (Information and Communication Technology) or due to failures of the human operational system. In addition, systemic failures may occur out of interaction between the various processes that manage the power system, which are not necessarily captured by an event-based approach. It is expected that the GARPUR approach could capture the risk of, for example, the action of the AGC ramping generators in a way that triggers system protection. It is however not presently clear how to determine the risk or benefit of changing operational procedures, or by changing the flow of information between TSO departments. For example, will communicating information on real-time system risk to the system development department affect their decision making, and how will those changes affect operational security in the future? Approaches such as STPA (Systems Theoretic Process Analysis) provide a qualitative method to assessing how changing processes may lead to system failure [110]. Further research is required in linking such methods to quantitative reliability methods, in order to quantify procedural risk and the benefit of changing processes.

4.2.7 Large-scale implementation

In order to develop tractable implementations for the pan-European transmission system, there is a clear need to develop less computationally intensive methods of estimating reliability. Of most interest is the formulation of modular or hierarchical probabilistic reliability assessments. That is, if reliability can be assessed for individual regions at a high level of detail, and then system-wide reliability can be computed based on the regional results. For example, rather than computing reliability on the entire Icelandic system in one simulation, the East and West portions of the power system could be assessed separately.

Such an approach would require that events cascading from one region to the other could be described as uncertainty on the lines connecting the two regions. By splitting the assessment along regional connections, the cardinality of the problem could be significantly reduced. Such an approach would be particularly useful in the European case, given that countries could assess their own internal risk/reliability, and then communicate uncertainty to neighbours in terms of failures or changes of flows on their cross-border connections. The difficulty in implementing such a methodology is to capture potential correlations in exogenous forecast errors between regions.

Further case-studies and pilot tests are required to be performed on the European scale, to better understand the computational bottlenecks, and need for improved models. None-the-less, the Icelandic pilot test described by this research may provide a useful reference point for future implementations of probabilistic reliability assessments.

Bibliography

[1] GARPUR consortium, "D2.2 guidelines for implementing the new reliability assessment and optimization methodology," 2016. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d2.2-guidelines-for-implementing-the-new-reliability-assessment-and-optimization-methodology.pdf`.

[2] GARPUR consortium, "D5.1 functional analysis of asset management processes," 2015. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d5.1-functional-analysis-asset-management.pdf`.

[3] GARPUR consortium, "D6.1 Functional analysis of System Operation processes," 2015. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d6.1-functional-analysis-system-operation.pdf`.

[4] United Nations, "Progress towards the sustainable development goals," 2017.

[5] UCTE, "Final report: System disturbance on 4 november 2006," 2006.

[6] P. Pourbeik, P. S. Kundur, and C. W. Taylor, "The anatomy of a power grid blackout," *IEEE Power and Energy Magazine*, vol. 4, no. 5, pp. 22–29, 2006.

[7] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, *et al.*, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *IEEE transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, 2004.

[8] ENTSO-E, "Continental europe operation handbook p3 - policy 3: Operational security," 2009.

[9] GARPUR consortium, "The garpur project results: Using probability and risk to allow power systems to operate closer to their optimum." *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-brochure.pdf`.

[10] J. McCalley, S. Asgarpoor, L. Bertling, R. Billinton, H. Chao, J. Chen, J. Endrenyi, R. Fletcher, A. Ford, C. Grigg, G. Hamoud, D. Logan, A. Meliopoulos, M. Ni, N. Rau, L. Salvaderi, M. Schilling, Y. Schlumberger, A. Schneider, and C. Singh, "Probabilistic security assessment for power system operations," in *IEEE Power Engineering Society General Meeting*, vol. 2, (Denver, Colorado, USA), pp. 212–220, IEEE, 2004.

[11] GARPUR consortium, "D6.2 how to upgrade reliability management for short-term decision making," 2016. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d6.2-how-to-upgrade-reliability-management-for-st-decision-making.pdf`.

[12] Landsnet, "Reliability of supply and quality of delivered electricity - performance report 2016," 2016.

[13] ENTSOE, "ENTSOE glossary," 2017.

[14] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*, vol. 7. McGraw-hill New York, 1994.

[15] CIGRE working group C1.17, "Planning to manage power interruption events," 2010.

[16] ENTSO-E, "Network Code on Load-Frequency Control and Reserves," 2013.

[17] Icelandic Ministry of Industry, "Regulation on the quality of electricity and security of supply," 2004.

[18] Althingi, "Electricity act nr.65/2003," 2003.

[19] ENTSOE, "Rate of change of frequency (ROCOF) withstand capability," 2017.

[20] A. J. Eliasson, "Natural Hazards and The Icelandic Power Transmission Grid," in *7. konferenca slovenskih elektroenergetikov*, (Velenje, Slovenia), CIGRE, 2005.

[21] S. Canevese, E. Ciapessoni, A. Gatti, and M. Rossi, "Monitoring of frequency disturbances in the european continental power system," in *AEIT International Annual Conference (AEIT), 2016*, pp. 1–6, IEEE, 2016.

[22] R. Doherty, A. Mullane, G. Nolan, D. J. Burke, A. Bryson, and M. O'Malley, "An assessment of the impact of wind generation on system frequency control," *IEEE transactions on power systems*, vol. 25, no. 1, pp. 452–460, 2010.

[23] ENTSOe, "Nordic and baltic grid disturbance statistics 2015," 2017.

[24] S. H. Bjornsdottir, "Comparison of risk analysis methodologies in an electrical grid." European STAMP Workshop, 2015.

[25] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version.* Cengage Learning, 2012.

[26] F. Milano, *Power system modelling and scripting.* Springer Science & Business Media, 2010.

[27] M. Papic, S. Agarwal, R. N. Allan, R. Billinton, C. J. Dent, S. Ekisheva, D. Gent, K. Jiang, W. Li, *et al.*, "Research on common-mode and dependent (cmd) outage events in power systems: A review," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1528–1536, 2017.

[28] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, "Ac power flow data in matpower and qcqp format: itesla, rte snapshots, and pegase," *arXiv preprint arXiv:1603.01533*, 2016.

[29] J. Bialek, E. Ciapessoni, D. Cirio, E. Cotilla-Sanchez, C. Dent, I. Dobson, P. Henneaux, P. Hines, J. Jardim, S. Miller, *et al.*, "Benchmarking and validation of cascading failure analysis tools," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4887–4900, 2016.

[30] G. McClure and M. Lapointe, "Modeling the structural dynamic response of overhead transmission lines," *Computers & Structures*, vol. 81, pp. 825–834, May 2003.

[31] P. Henneaux, P.-E. Labeau, J.-C. Maun, and L. Haarla, "A two-level probabilistic risk assessment of cascading outages," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2393–2403, 2016.

[32] N. Samaan, J. Miller, LE an Dagle, B. Vyakaranam, Y. Makarov, S. Wang, R. Diao, F. Tuffner, M. Vallem, and M. Pai, "Dynamic contingency analysis tool – phase 1," Tech. Rep. PNNL-24843, Pacific Northwest National Laboratory, 2015.

[33] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk Assessment of Cascading Outages: Methodologies and Challenges," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 631–641, 2012.

[34] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with dc power flow model and transient stability analysis," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 285–297, 2015.

[35] P. D. Hines, I. Dobson, and P. Rezaei, "Cascading power outages propagate locally in an influence graph that is not the actual grid topology," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 958–967, 2017.

[36] M. A. Rios, D. S. Kirschen, D. Jayaweera, D. P. Nedic, and R. N. Allan, "Value of security: modeling time-dependent phenomena and weather conditions," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 543–548, 2002.

[37] M. Rios, K. Bell, D. Kirschen, and R. Allan, "Computation of the value of security," 1999.

[38] M. Adibi, J. Borkoski, and R. Kafka, "Power system restoration-the second task force report," *IEEE Transactions on Power systems*, vol. 2, no. 4, pp. 927–932, 1987.

[39] S. Kancherla and I. Dobson, "Heavy-tailed transmission line restoration times observed in utility data," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1145–1147, 2018.

[40] C. Liu, D. Li, E. Zio, and R. Kang, "A modeling framework for system restoration from cascading failures," *PloS one*, vol. 9, no. 12, p. e112363, 2014.

[41] C. Coffrin, P. Van Hentenryck, and R. Bent, "Last-mile restoration for multiple interdependent infrastructures.," in *AAAI*, vol. 12, pp. 455–463, 2012.

[42] P. Henneaux and D. S. Kirschen, "Probabilistic security analysis of optimal transmission switching," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 508–517, 2016.

[43] GARPUR consortium, "D3.1 quantification method in the absence of market response and with market response takin into account," 2016. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d3.1-quantification-method-in-the-absence-of-market-response-and-with-market-response-taken-into-account.pdf`.

[44] W. Bukhsh, K. Bell, and T. Bedford, "Risk and reliability assessment of future power systems," in *European Safety and Reliability Conference 2016*, 2016.

[45] EFLA, "Kostnadar Vegna Raforkuskorts: Tolur til notkunar arin 2016 og 2017," tech. rep., EFLA Consulting, 2017.

[46] J. Endrenyi, *Reliability modeling in Electric Power Systems.* Wiley, 1978.

[47] J. Dodu and A. Merlin, "New probabilistic approach taking into account reliability and operation security in ehv power system planning at edf," *IEEE transactions on power systems*, vol. 1, no. 3, pp. 175–181, 1986.

[48] M. Raisand and A. Hoyland, *System reliability theory: models, statistcal methods, and applications.* John Wiley & Sons, 2004.

[49] R. Billinton and R. N. Allan, *Reliability Evaluation of Engineering Systems: Concepts and Techniques.* New York: Springer, 2nd edition ed., June 1992.

[50] R. Billinton and R. Allan, *Reliability Evaluation of Power Systems.* Plenum Press, 1996.

[51] Billinton and W. Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods.* Springer Science & Business Media, June 2013.

[52] ENTSO-E, "Network Code on Operational Security," 2013.

[53] P. Palermo, M. Wilks, and D. Korinek, "International review of transmission reliability standards," *Summary Report prepared for the Australian Energy Market Commission Reliability Panel*, pp. 53–4617, 2008.

[54] "D1.2 Current practices, drivers and barriers for new reliability standards," tech. rep.

[55] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization," *CHAOS*, vol. 17, no. 2, p. 26103, 2007.

[56] H. Ren, I. Dobson, and B. A. Carreras, "Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid," *IEEE transactions on power systems*, vol. 23, no. 3, pp. 1217–1225, 2008.

[57] RTE, "Memento de la surete du system electrique," 2004.

[58] NERC, "Probabilistic assessment: technical guideline document," 2016.

[59] D. Kirschen and D. Jayaweera, “Comparison of risk-based and deterministic security assessments,” *IET Generation, Transmission & Distribution*, vol. 1, no. 4, pp. 527–533, 2007.

[60] E. Karangelos, P. Panciatici, and L. Wehenkel, “Whither probabilistic security management for real-time operation of power systems?,” in *Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*, pp. 1–17, Aug. 2013.

[61] E. Karangelos and L. Wehenkel, “Probabilistic reliability management approach and criteria for power system real-time operation,” in *19th Power Systems Computation Conference (PSCC)*, pp. 1–9, 2016.

[62] K. Bell, “Issues in integration of risk of cascading outages into utility reliability standards,” in *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–7, IEEE, 2011.

[63] P. Panciatici, M. C. Campi, S. Garatti, S. Low, D. K. Molzahn, A. Sun, and L. Wehenkel, “Advanced optimization methods for power systems,” in *Power Systems Computation Conference (PSCC), 2014*, pp. 1–18, IEEE, 2014.

[64] F. Capitanescu, “Critical review of recent advances and further developments needed in ac optimal power flow,” *Electric Power Systems Research*, vol. 136, pp. 57–68, 2016.

[65] P. C. López, R. Sadikovic, H. Pinto, and F. Magnago, “Swiss tso experience with an ac security-constrained optimal power flow application for real-time security management,” in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–6, IEEE, 2015.

[66] M. Lange *et al.*, “On the uncertainty of wind power predictions-analysis of the forecast accuracy and statistical distribution of errors,” *Transactions of the ASME-N-Journal of Solar Energy Engineering*, vol. 127, no. 2, pp. 177–184, 2005.

[67] P. Bacher, H. Madsen, and H. A. Nielsen, “Online short-term solar power forecasting,” *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, 2009.

[68] G. Gross and F. D. Galiana, “Short-term load forecasting,” *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.

[69] S. Rahman and O. Hazim, “Load forecasting for multiple sites: development of an expert system-based technique,” *Electric Power Systems Research*, vol. 39, no. 3, pp. 161–169, 1996.

[70] H. Holttinen, “Impact of hourly wind power variations on the system operation in the nordic countries,” *Wind Energy*, vol. 8, no. 2, pp. 197–218, 2005.

[71] S. Jaehnert, T. Aigner, G. Doorman, and T. Gjengedal, “Impact of large scale wind integration on power system balancing,” in *PowerTech, 2011 IEEE Trondheim*, pp. 1–6, IEEE, 2011.

[72] A. Vergnol, J. Sprooten, B. Robyns, V. Rious, and J. Deuse, “Real time grid congestion management in presence of high penetration of wind energy,” in *Power Electronics and Applications, 2009. EPE’09. 13th European Conference on*, pp. 1–10, IEEE, 2009.

[73] J. C. Villumsen, G. Bronmo, and A. B. Philpott, “Line capacity expansion and transmission switching in power systems with large-scale wind power,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 731–739, 2013.

[74] C. Hamon, M. Perninge, and L. Soder, “A computational framework for risk-based power systems operations under uncertainty. Part I: Theory,” *Electric Power Systems Research*, vol. 119, pp. 45–53, Feb. 2015.

[75] M. T. Do, J. Sprooten, S. Clenet, and B. Robyns, “Influence of wind turbines on power system reliability through probabilistic studies,” in *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, pp. 1–6, IEEE, 2010.

[76] H. Louie, “Evaluation of bivariate archimedean and elliptical copulas to model wind power dependency structures,” *Wind Energy*, vol. 17, no. 2, pp. 225–240, 2014.

[77] G. Papaefthymiou and D. Kurowicka, “Using copulas for modeling stochastic dependence in power system uncertainty analysis,” *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 40–49, 2009.

[78] A. P. Douglas, A. M. Breipohl, F. N. Lee, and R. Adapa, “Risk due to load forecast uncertainty in short term power system planning,” *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1493–1499, 1998.

[79] eHighway-2050 Consortium, “D8.2.a Enhanced methodology for the computation of Demand/Generation scenarios,” tech. rep., EU Commission grant agreement 308908, 2015.

[80] H. Kile, K. Uhlen, and G. Kjølle, “Scenario selection in composite reliability assessment of deregulated power systems,” *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 124–131, 2014.

[81] N. Growe-Kuska, H. Heitsch, and W. Romisch, “Scenario reduction and scenario tree construction for power management problems,” in *Power tech conference proceedings, 2003 IEEE Bologna*, vol. 3, pp. 7–pp, IEEE, 2003.

[82] M. H. Bollen, *Literature search for reliability data of components in electric distribution networks.* Eindhoven University of Technology, Faculty of Electrical Engineering, 1993.

[83] R. Stillman, “Modeling failure data of overhead distribution systems,” *IEEE Transactions on power delivery*, vol. 15, no. 4, pp. 1238–1242, 2000.

[84] S. de Almeida, *Portuguese Transmission Grid Incidents Risk Assessment.* PhD thesis, FUEP, Portugal, 2010.

[85] ENTSO-E, “Guidelines for the classification of grid disturbances above 100 kV,” 2017.

[86] R. Billinton and L. Wenyuan, "A novel method for incorporating weather effects in composite system adequacy evaluation," *IEEE Transactions on Power Systems*, vol. 6, pp. 1154–1160, Aug. 1991.

[87] R. Billinton and G. Singh, "Application of adverse and extreme adverse weather: modelling in transmission and distribution system reliability evaluation," *Generation, Transmission and Distribution, IEE Proceedings*, vol. 153, pp. 115–120, Jan. 2006.

[88] J. McCalley, Fei Xiao, Yong Jiang, and Qiming Chen, "Computation of Contingency Probabilities for Electric Transmission Decision Problems," in *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems*, pp. 540–545, IEEE, 2005.

[89] K. Alvehag and L. Soder, "A reliability model for distribution systems incorporating seasonal variations in severe weather," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 910–919, 2011.

[90] F. Xiao, J. McCalley, Y. Ou, J. Adams, and S. Myers, "Contingency Probability Estimation Using Weather and Geographical Data for On-Line Security Assessment," in *International Conference on Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006*, pp. 1–7, June 2006.

[91] N. Machado, S. A. de Graaff, and R. Pestana, "Risk assessment methodology – Running tests at the Portuguese TSO," in *CIGRE*, (Paris, France), CIGRÉ, 2014.

[92] A. Shafaghi, "Equipment failure rate updating—bayesian estimation," *Journal of hazardous materials*, vol. 159, no. 1, pp. 87–91, 2008.

[93] Ø. R. Solheim, G. Kj, *et al.*, "Wind dependent failure rates for overhead transmission lines using reanalysis data and a bayesian updating scheme," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 International Conference on*, pp. 1–7, IEEE, 2016.

[94] A. Moradkhani, M. R. Haghifam, and M. Mohammadzadeh, "Bayesian estimation of overhead lines failure rate in electrical distribution systems," *International Journal of Electrical Power & Energy Systems*, vol. 56, pp. 220–227, 2014.

[95] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about electricity infrastructure vulnerability?," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 3, p. 033122, 2010.

[96] P. Kaplunovich and K. S. Turitsyn, "Statistical properties and classification of n-2 contingencies in large scale power grids," in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pp. 2517–2526, IEEE, 2014.

[97] T. Tollefsen, A. Svendsen, R. Pedersen, P. Skeie, T. Lunde, and J. Maelan, "Online reliability calculations of power systems with forecasted and real time weather influence," in *Safety and Reliability of Complex Engineered Systems*, ch. 215, pp. 1739–1744, CRC Press, 2015.

[98] G. W. Brier, "Verification of forecasts expressed in terms of probability," *Monthey Weather Review*, vol. 78, no. 1, pp. 1–3, 1950.

[99] Promaps Technology, "Promaps real time," 2017. `http://www.http://promaps.no/`.

[100] S. de Almeida, R. Pesta, and F. Barbosa, "The main causes of incidents in the Portuguese Transmission System - Their characterization and how they can be used for risk assessment," in *6th International Conference on the European Energy Market, 2009. EEM 2009.*, pp. 1–6, IEEE, May 2009.

[101] GARPUR Consortium, "TSO reliability management: a probabilistic approach for better balance between reliability and costs," 2017. *Available at* `https://www.youtube.com/watch?v=ZkBTjyJlcw8`.

[102] GARPUR consortium, "D8.1 Pilot testing methodologies, models, scenarios and validation approach," 2016. *Available at* `http://www.sintef.no/globalassets/project/garpur/deliverables/garpur-d8.1-pilot-testing-methodologies-models-scenarios-and-validation-approach.pdf`.

[103] R. D. Zimmerman, C. E. Murllo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[104] CIGRE working group C4.601, "Review of the current status of tools and techniques for risk-based and probabilistic planning in power systems," 2010. Technical Brochure 434.

[105] E. Ciapessoni, D. Cirio, S. Massucco, A. Morini, A. Pitto, and F. Silvestro, "Risk-based dynamic security assessment for power system operation and operational planning," *Energies*, vol. 10, no. 4, p. 475, 2017.

[106] M. Vasconcelos, L. Carvalho, J. Meirinhos, N. Omont, P. Gambier-Morel, G. Jamgotchian, D. Cirio, E. Ciapessoni, A. Pitto, I. Konstantelos, *et al.*, "Online security assessment with load and renewable generation uncertainty: The itesla project approach," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 International Conference on*, (Beijing, China), pp. 1–8, IEEE, 2016.

[107] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[108] Q. Chen, K. Zhu, and J. D. McCalley, "Dynamic decision-event trees for rapid response to unfolding events in bulk transmission systems," in *Power Tech Proceedings, 2001 IEEE Porto*, vol. 2, pp. 5–pp, IEEE, 2001.

[109] G. Doorman, G. Kjolle, K. Uhlen, E. S. Huse, and N. Flatabo, "Vulnerability of the Nordic Power System," Tech. Rep. TR F5962, SINTEF, Trondheim, Norway, 2004.

[110] T. Ishimatsu, N. G. Leveson, J. Thomas, M. Katahira, Y. Miyamoto, and H. Nakao, "Modeling and hazard analysis using STPA," in *Proceedings of the 4th International Association for the Advancement of Space Safety (IAASS) Conference*, no. SP-680, 2010.

[111] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7290–7297, 2014.

[112] S. Perkin, G. Bjornsson, I. Baldursdottir, M. Palsson, R. Kristjansson, H. Stefansson, P. Jensson, E. Karangelos, and L. Wehenkel, "Framework for threat based failure rates in transmission system operation," in *2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO)*, (Beer Sheva, Israel), pp. 150–158, IEEE, 2016.

[113] V. V. Vadlamudi, C. Hamon, G. Kj, S. Perkin, *et al.*, "On improving data and models on corrective control failures for use in probabilistic reliability management," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 International Conference on*, pp. 1–6, IEEE, 2016.

Appendix A

Algorithm for RT trajectory-based security assessment

The algorithm below provides a mathematical formulation of the ideal trajectory-based security assessment in a real-time context. This algorithm formed the basis for the work in **Paper III** (listed on page xi) and also led to the formulation of the implementation described in GARPUR deliverable D6.2 [11]. This appendix is referenced in Sub-section 3.2.2. Variable definitions can be found on the next page.

Algorithm 1: Ideal formulation of RT trajectory-based security assessment

Data: $s_\emptyset, T, d_c, s_r, \mathcal{X}_a^0, \Xi_a^T, \mathcal{C}_a^0, \mathcal{R}_a^0, \mathcal{B}_a^0, C(\cdot), f(\cdot)$
$\mathbb{P}\left(\boldsymbol{\xi}_t \in \Xi^t(\xi_{[t-1]})\right), \mathbb{P}\left(c_t \in \mathcal{C}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t\right), \mathbb{P}\left(r_t \in \mathcal{R}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t, c_t\right),$
$\mathbb{P}\left(b_t \in \mathcal{B}^t \;\middle|\; \mathbf{x}_t, \boldsymbol{\xi}_t, c_t, r_t\right)$
Result: $\mathcal{P}, R, R_r$

$s = 1,\ R = \mathcal{P} = 0$ (initialise variables)
foreach $\boldsymbol{x}_0 \in \mathcal{X}_a^0$ **do**
    foreach $\xi \in \Xi_a^T$ **do**
        $R = R + \mathbb{P}(\mathbf{x}_0) \cdot \mathbb{P}\left(\boldsymbol{\xi}_0 \in \Xi_a^0\right) \cdot C(\mathbf{x}_0, \boldsymbol{\xi}_0)$
        foreach $(c_0, r_0, b_0) \in \{\mathcal{C}_a^0 \times \mathcal{R}_a^0 \times \mathcal{B}_a^0\}$ **do**
            $\pi_\Delta(\boldsymbol{\xi}_0, c_0, r_0, b_0|\mathbf{x}_0) \leftarrow$ Eq. (3.5)
            $\pi_\emptyset^s = 1$
            foreach $t \in [0, d_c]$ **do**
                if $t = 0$ **then**
                    $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \boldsymbol{\xi}_t, c_t, r_t, b_t)$
                else
                    $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \boldsymbol{\xi}_t, c_\emptyset, r_\emptyset, b_\emptyset)$
                end
                $\pi_\emptyset^s = \pi_\emptyset^s . \pi_\emptyset(\mathbf{x}_t, \boldsymbol{\xi}_t)$
            end
            $\pi_s = \mathbb{P}(\mathbf{x}_0) \cdot \pi_\Delta(\boldsymbol{\xi}_0, c_0, r_0, b_0|\mathbf{x}_)\pi_\emptyset^s$
            $\mathcal{P} = \mathcal{P} + \pi_s$
            $R_s = \pi_s \cdot \sum_{t=1}^{T} C(\mathbf{x}_t, \boldsymbol{\xi}_t)$
            $R = R + R_s$
            $s = s + 1$
        end
    end
end
$R_r = \big(1 - \mathcal{P}\big).C(s_r)$
$R = R + R_r$

Definition of variables used in the algorithm:

$s_\emptyset$	: The planned system trajectory
$t \in [0, T]$	: Operational period time index
T	: Assessment period $(= \max_{c \in \mathcal{C}_a^0} d_c)$
ξ	: a sequence of scenario realisations
d_c	: maximum service outage duration of contingency c
s_r	: estimated worst-case system trajectory (r: residual)
$\mathcal{X}_a^0$	: Set of assessed initial system configurations
Ξ_a^T	: Set of exogenous scenarios over the assessment period
$\mathcal{C}_a^0$	: Set of assessed initial contingencies
$\mathcal{R}_a^0$	: Set of assessed initial system responses
$\mathcal{B}_a^0$	: Set of assessed initial corrective control behaviours
$C(\cdot)$	: Function used to compute consequences (e.g. SEIA)
$f(\cdot)$	: System response and restoration model (see **Paper V**)
$\mathbb{P}\left(\boldsymbol{\xi}_t \in \Xi^t(\xi_{[t-1]})\right)$	: Prob. of an exogenous realisation at time t given prior realisations
$\mathbb{P}\left(c_t \in \mathcal{C}^t \mid \mathbf{x}_t, \boldsymbol{\xi}_t\right)$	: Cond. prob. of a particular contingency at time t
$\mathbb{P}\left(r_t \in \mathcal{R}^t \mid \mathbf{x}_t, \boldsymbol{\xi}_t, c_t\right)$	: Cond. prob. of a particular system response at time t
$\mathbb{P}\left(b_t \in \mathcal{B}^t \mid \mathbf{x}_t, \boldsymbol{\xi}_t, c_t, r_t\right)$	: Cond. prob. of a particular corr. control behaviour at time t
$\mathcal{P}$	: Probabilistic coverage $= \mathbb{P}(s \in \mathcal{S}_a^T)$
R	: System risk (subscript r denotes residual risk)

Appendix B

Main pilot test simplifications and assumptions

The table below provides a list of the main simplifications and assumptions that were required in order to perform the pilot test within the available time frame. Each assumption is also accompanied by a justification to explain why such assumptions and simplifications were necessary. Although the table is not an exhaustive list of simplifications and assumptions, it contains those which are considered to be major.

Table B.1: Main assumptions for pilot testing, adapted from prior internal reporting on pilot testing

Context	Simplification/assumption	Reasoning
System model	Substations modelled as nodes (bus bars aggregated and breakers not modelled)	Required simple model to facilitate fast debugging
Failure rate models	Only considered wind speed as an input variable	Limited real-time weather observation data was available, and insufficient historical data/models for other threats
Pilot test method	Acceptability constraints are only checked on the post-control system state	System response model over-estimated voltage excursions, and insufficient time to resolve issue
Pilot test method	Transient trips not modelled	Only sustained trips were included due to time constraints, which were believed to be sufficient to show the general approach
Pilot test method	Only single component failure modes considered	Same as above
Pilot test method	Probability of unexpected system response and corrective control behaviours set to 0%	Although functionality was included and tested in **Paper III**, it was excluded to simplify the results
Pilot test method	Only a single exogenous scenario considered	Functionality was implemented but not used to simplify outputs, reduce runtime during debugging, and due to a lack of requisite data for the Icelandic system
Pilot test method	No market settling modelled	Model runs on real-time data, therefore market-defined unit commitments are implicitly included
Pilot test method	Ignore overlapping faults	Although this functionality was included (inline with Equation (3.4)), the probability of overlapping faults was set to 0% to reduce output complexity
Contingency discarding	Maximum contingency list size set to 5000	Initial tests suggested that inclusion of more than 5000 contingencies resulted in significant increase in computational time with negligible impact on reliability indicators
System response	Straight-forward control actions and system protection was modelled heuristically	Most protection devices and corrective control actions could be simplified to if-then-else statements. More complex control actions were not modelled, but instead an OPF with load-shedding was used to quantify the cost of inaction (unless corrective controls were pre-specified)
System response	Sympathetic tripping probability set to 0%	Although implemented, the functionality was not used for the sake of transparency and to avoid the need for Monte Carlo simulations (reduce computational speed)
System response	Cables and transformers assumed to trip deterministically if loaded above 130% of short-term MVA limits	Same reasoning as above
System response	Non-convergence of the Load Frequency control (LFC) algorithm is assumed to imply the need for Under Frequency Load Shedding (UFLS)	In reality, LFC of generators and UFLS can occur in parallel due to frequency deviations, but for modelling simplicity they are considered separately
System restoration	Assume maximum service outage duration is 8 hours	Expert estimates suggested that nearly all conceivable outages can be resolved in this time frame, and those extending beyond 8 hours cannot be resolved in the context of system operation
System restoration	No load forecasts were included	For simplicity and to reduce the number of interfaced systems, power demand was assumed to be constant over the restoration process
Economic Assessment	Time-step for assessment set to 1 hour	impact was assessed at hourly resolution to simplify outputs
Economic Assessment	Loads are split into primary and secondary loads, where secondary loads can be curtailed and don't contribute to interruption costs	Reflects some loads on the Icelandic transmission system

School of Science and Engineering
Reykjavík University
Menntavegur 1
101 Reykjavík, Iceland
Tel. +354 599 6200
Fax +354 599 6201
www.ru.is