THE UNIVERSITY OF LIVERPOOL

Assessing Resilience of Smart Critical Infrastructures to Deal with Emerging Risks and Threats

by

Zarif Ahmet Zaman

A thesis submitted in fulfilment of the requirements for the degree of Master of Philosophy

in the

School of Engineering

The University of Liverpool

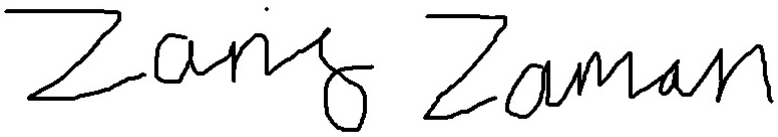
September 2022

Declaration of Authorship

I, ZARIF AHMET ZAMAN, declare that this thesis titled, “Assessing Resilience of Smart Critical Infrastructures to Deal with Emerging Risks and Threats” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:



Date: 29/09/2022

List of Publications

1. Zaman, Z.A., Patelli, E. A DC Optimal Power Flow Approach to Quantify Operational Resilience in Power Grids. *18th International Probabilistic Workshop*. Springer, 2021. Available at https://doi.org/10.1007/978-3-030-73616-3_4

THE UNIVERSITY OF LIVERPOOL

Abstract

School of Engineering
The University of Liverpool

Master of Philosophy

by Zarif Ahmet Zaman

The innovation of critical infrastructure plays a key role within sustaining a high-functioning and content society in the 21st century. Demands regarding the nature of critical infrastructure have been ever increasing in the current marketplace and new challenges are constantly being presented. Such challenges include the reliability of critical infrastructural systems in which this academic study aims to investigate, analyse and communicate a system's function with respects to a given time.

Resilience analysis simulation techniques are present in the current state of the art, however, challenges remain with regards to minimising high computational costs whilst maintaining credible accuracy of results. Such techniques range from three phase resilience modelling, consequence modelling and cost based resilience analysis. Another limitation within the current literature is the scarcity of applied uncertainty to simulation techniques. This desire for both epistemic and aleatory uncertainty is evident with high academic and industrial demand for the communication of imprecise data.

The key findings of this Master of Philosophy thesis tackle the limitations and highlight the strengths of three different resilience based techniques as applied to real world scenarios. The research found credibility within the survival signature and probability propagation for estimation of reliability within the simplified China railway network with the addition of uncertainty. The load flow method was also applied and tested to quantify operational resilience within power grids and displayed credible results for the LODF, DC-OPF, AC-OPF and a surrogate model was implemented for the Great Britain power network. A CFD based resilience model was also applied to a natural gas pipeline and displayed imprecise results in a three phase resilience curve.

Acknowledgements

I thank my parents for supporting me through a difficult journey under unprecedented circumstances. I also thank the Engineering and Physical Sciences Research Council for funding this research project and unlocking the gateway for my academic experience to research resilience alongside my academic supervisors.

Also many thanks to my friends back in Manchester for keeping in touch with me after many years following my departure from home and giving me supportive advice through the past decade. Finally, I give thanks to the candidates of my cohort who started their journeys with me and are also finishing their journeys with me, and whose company I enjoyed traveling around the country with.

Contents

List of Figures

List of Tables

Nomenclature

ABBREVIATIONS

AC-OPF	AC Optimal Power Flow
ANN	Artificial Neural Network
BOP	Blowout Preventer
BP	British Petroleum
BTU	British Thermal Unit
CFS	Computational Fluid Dynamics
CDF	Cumulative Distribution Function
CFD	Computational Fluid Dynamics
CS	Compression Stations
D	Demand Site
DAG	Direct Acyclic Graph
DC-OPF	DC Optimal Power Flow
EGIG	European Gas Pipeline Incident Data Group
GS	Gas Supplier
HPP	Homogeneous Poisson Process
IEEE	Institute of Electrical and Electronics Engineers
IMPOVER	Improved Risk Evaluation and Implementation of Resilience Concepts to Critical Infrastructure
LNG	Liquefied Nitrogen Gas
LODF	Line Outage Distribution Factors
MHMT	Momentum Heat and Mass Transfer

MILP	Mixed-Integer Linear Programming
NPI	Non-Parametric Predictive Inference
PDF	Probability Density Function
PMF	Probability Mass Function
PMS	Phased Mission System
PrPm	Analytical Probability Propagation Method
PTDF	Power Transfer Distribution Factor
RBD	Reliability Block Diagram
RCV	Remote Control Valve
RS	Regulation Station
SCADA	Supervisory Control and Data Acquisition
SMC	Sequential Monte Carlo Simulation
TSO	Transmission System Operator
UGF	Universal Generating Function
UGS	Underground Storage
UNDRR	United Nations Office for Disaster Risk Reduction

NOTATIONS

ENS	Energy-Not-Supplied
iid	Independent and Identically Distributed
LPHC	Low-Probability High-Consequence
MAOP	Maximum Allowable Overpressure
MOS	Maximum Operating Speed
mpd	Maximum packet delay
MTTF	Mean time to fail
PR_l	Power rating in a pre-contingency state
Rob	Robustness performance function
Rec	Recovery performance index
SPAD	Signal passed at danger
TTE	Time to event

TTR	Time to repair
$A - N$	Link connecting nodes A and N
$f(L_i(t))$	Gaussian model for power grid
$\min_x f(x)$	Standard objective function
$K : \mathbf{I} \rightarrow \mathbf{Y}$	ANN mathematical expression
$P(A, C, N)$	Joint probability distribution between nodes A, C and N
$P(E = e\|M)$	Bayes' theorem likelihood probability
$P(M)$	Bayes' theorem prior probability
$P(M\|E = e)$	Bayes' Rule
$P(M, E = e)$	Bayes' theorem joint probability
$P(N_e(t) = k)$	Non-homogeneous Poisson process
$P(N_f(t) = k)$	Homogeneous Poisson process
$\mathscr{G}(\mathscr{N}, \mathscr{E})$	Power grid topology
$\mathscr{P}(C_k\|\zeta)$	Probability of contingency
$\mathscr{R}(\zeta)$	Composite risk index

SYMBOLS

Chapter 2

A_i	CDF interval
$\underline{F}(x)$	Lower bound function of p-box
$\overline{F}(x)$	Upper bound function of p-box
L	Set of numeric intervals
$l(x)$	Lower bound probability distribution
m_i	Mass probability of A
$P(A)$	Probabilistic value
$\underline{P}(A)$	Lower probability bound of A
$\overline{P}(A)$	Upper probability bound of A
$P(A^c)$	Complement of $P(A)$
P_L	Credal set with interval probabilities
$P_{\underline{p}}$	Credal set

$p(x)$	Probability distribution
$u(x)$	Upper bound probability distribution
X	Gaussian input
$\mathbb{P}_X$	Set of additive capacities on X
$\mathbb{R}$	Real number
$\mathscr{F}(x)$	Credal set
μ	Gaussian mean
σ	Gaussian standard distribution
Chapter 3	
$C_k(t)$	Number of type k components functioning at time t
$c_{\underline{l}}$	Coefficient of variation
C_t	Number of components of the system
d	Node degree
$\overline{D}$	Maximum probability with the application of Bernouli data
$\underline{D}$	Minimum probability with the application of Bernouli data
E	Bayesian belief evidence variable
e	Bayesian belief evidence
$F(t)$	Failure function
f	Fraction of deleted nodes of a system
f_c	Critical fraction
K	Number of k components
k	Type of component
l	Number of components of the system
l_K	Number of components of the system of K specific failure times
m	Type of component
m_K	Type of component with K specific failure times
n_k	Number of type k with specific failure times
n_l	Number of samples
$\overline{P}$	NPI upper probability interval for Bernoulli data

$\underline{P}$	NPI lower bound probability for Bernouli data
P_1	Probability of first step of propagation sequence
P_2	Probability of second step of propagation sequence
P_3	Probability of third step of propagation sequence
P_4	Probability of fourth step of propagation sequence
p_{in}	Probability of random cluster node
$P(\mathbf{X})$	General joint probability distribution
$\overline{P}$	Upper probability interval for Bernouli data
q_j	General system signature
R	Previously propagated node reliability
R_1	Link reliability
$R(t)$	Reliability function
S	Source node
$s_k(t)$	Number of type k components still functioning at time t
S_l	State of the system
$\underline{S}_{T_S}$	NPI lower survival function
$\overline{S}_{T_S}$	NPI upper survival function
T	Terminal node
$T_{j:m}$	j-th order statistic
T_s	System random failure time
$\underline{x}$	State vector
Z	Normalisation constant
$\mathcal{O}$	Computational cost
β	Weibull shape parameter
η	Weibull scale parameter
γ	Weibull location parameter
κ	Molloy-Reed criterion
μ	Mean
Ξ	Mean distance to cluster

ξ	Set of maximal cliques
σ	Standard deviation
Φ	Survival signature
ϕ	Structural function
Ψ	Desired function for Hammersley-Clifford theorem
$\omega_{\underline{l}}$	Counter
(s)	Mean cluster size
Chapter 4	
b	ANN weighted bias
B_{bus}	Bus matrix
B_f	Branch flow
C_i	A given contingency
C_g	Output cost of generator
D_ω	Duration of the wind storm event
Dl_g	Duration of the lightning storm event
d_l^k	LODF matrix
$\mathbb{E}[ENS]$	Energy not supplied
e	Final event during simulation
F	Matrix of annual failures
F_{max}	Vector flow limits of a given branch
f	Failure index
f_k	Failure flow
f_l	Power flow
f_P^i	Real power cost function
f_Q^i	Reactive power cost function
$g(x)$	Equality Constraint
G_{is}	Single node island
G_{sh}	Shunt generator matrix
$g_i(s)$	ANN weighted function

$g_i(x)$	ANN output of node from prior layer
g_P	Nodal real power balance matrix
g_Q	Nodal reactive power balance matrix
$h(x)$	Inequality constraint
I_f	Current flow
i	Queried line i
$\mathbf{I}_j$	ANN input vector
K	ANN composition of various weighted functions
k	Number of failures of type k
L_{cut}	Load curtailed
$L_{cut,i,t}$	Load curtailed at each individual node i at a given time t
$\mathbf{L}_f$	Load profile
$L_i(t)$	Indicated load curtailed
l	Power grid transmission line
l_i	Length of queried line i
$N_f(t)$	Number of failures per km of grid line
$N(g)$	Lightning strike intensity
$N_g(t)$	Lightning strike ground density
N_s	Number of samples
n_b	Number of buses in the system
n_c	Total number of contingency failures
n_g	Number of generators in the system
n_l	Number of lines in the system
n_L	Number of loads in the system
P	Real power
$P_{Bus,shift}$	Real power of bus in a shifted state
P_d	Real power demand
P_f	Real power flow
$P_{f,shift}$	Real power flow in a shifted state

P_g	Real power output
Q	Reactive power
Q_d	Reactive power demand
Q_g	Reactive power output
R	Regression coefficient
S_f	Apparent power flow
T	Given time of interest
T_d	Time delay
T_e	Duration of severe weather model
T_{sim}	User defined simulation time
t	A given time t
W_{crt}	Critical wind speed
$W_w(t)$	Wind speed intensity at time t for event w
w_i	ANN weight of node
V_a	Voltage angle
$V_e(t)$	Time dependent probability of event
$v_e(t')$	Rate at which the disturbance occurs
V_m	Voltage magnitude
v_{norm}	Normal average repair speed
$v_e(t')$	Rate at which the disturbance occurs
v_{repair}	Recovery model
x	Optimisation vector
$\mathbf{X}_f$	State vector
Y_f	Admittance matrix
Y_j	ANN load curtailed
$\mathscr{E}$	Topological edge
$\mathscr{N}$	Node set
α_w	Regression parameter
β_{lg}	Regression coefficient

Δ_w	High winds intensity
$\Delta_w(t)$	Difference between W_{crt} and $W_w(t)$
ζ	Given operational condition
η	Recovery model parameter
$\lambda_{(lg)}$	Total line failure contribution from lightning storms
$\lambda_{lg}(N_g(t))$	Line failure rate from lightning
λ_n	Line failure rate under normal conditions
$\lambda_{n,i}$	Total failure rate
$\lambda(t)$	Total failure rate function of weather model
λ_w	Total line failure contribution from high winds
$\lambda_w(W_w(t))$	Line failure rate from high winds
$\mu_{Li}(t)$	Mean value of load
$\sigma_{Li}(t)$	Standard deviation of node i at time t
ψ	Recovery model parameter

Chapter 5

a	Angle of pipeline
A_{or}	Cross-sectional area of orifice
b	General recovery parameter
$b_{i,j}$	Semi-empirical parameter
c	Speed of sound
C_D	Charge coefficient
C_v	Specific heat capacity at constant volume
D	Pipeline diameter
D_{ij}	Diameter of the pipeline (i, j)
dx	Pipeline infinitesimal length
f	Friction factor
f_{ij}	Flow variable
g	Velocity of gravity
$g_{deg}(t)$	Degradation function

g_{rec}	Network recovery curve
G_{RecCap}	Recovery capacity
G_{RobCap}	Robustness capacity
K	Dynamic pipeline equation constant
K_{pr}	Promptness parameter
L	Load of node
L_j	Transfer of gas
l	Length
m	Pressure degradation rate
$\dot{m}$	Mass flow rate
$\dot{m}_{hole}$	Mass flow rate of hole
$\dot{m}f_{t=0}$	Mass flow rate of system prior to failure
P_{cont}	Constant pressure
$P_{ij}(t)$	Recovery function
p	Pressure
p_{init}	Initial pressure of gas pipeline system
$p_{init(i,j)}$	Initial real pressure of pipeline (i, j)
$p_{fin(i,j)}$	Final real pressure of pipeline (i, j)
p_{max}	Maximum pressure
p_{min}	Minimum pressure
$\overline{p_i}$	Mean pressure of pipe i
Q	Volumetric flow rate
Q_{cap}	Volumetric flow rate capacity
Q_{dem}	Volumetric flow rate demand
Q_{deg}	Degredation volumetric flow rate
$Q_{ij,0}$	Volumetric flow rate capacity for pipeline (i, j) at $t = 0$
$Q_{ij}(t)$	Volumetric flow rate capacity from i to j
$Q_{j,n}$	Volumetric flow rate through pipeline n to node j
Q_{max}	Maximum volumetric flow rate

R	Gas constant
Re	Reynolds Number
r	Absolute pipe roughness
S	Constant cross-sectional area
$S_{j,n}$	Cross sectional area from pipeline n to node j
S_p	Cross sectional area at constant pressure
T	Temperature
T_{cont}	Constant temperature
t	Time
t_{dest}	Time to reach destination node
t_{det}	Detection time
t_{rec}	Recovery time
$\bar{t}_{det}$	Average detection time
u_{ij}	Link capacity of (i, j)
v	Fluid velocity
v_{gas}	Gas recovery velocity
w	Capacity weighting
Z	Compressibility factor
$\mathscr{I}$	Pipeline link (i, j)
γ	Isentropic coefficient
δ	Non-dimensional conversion coefficient
∂p	Pressure partial differential
∂Q	Volumetric flow rate partial differential
∂t	Time partial differential
∂x	Partial profile differential
$\partial \rho$	Density partial differential
μ	Dynamic viscosity
ρ	Fluid density
τ	Time constant

φ	Inlet pressure constant
$\varphi(t)$	Transient resilience function
Ψ	Dimensionless factor for overpressure equation
$\dot{\Omega}$	Energy flux

Dedicated to my Family…

Chapter 1

Introduction

1.1 Background

Critical Infrastructure is an integral part of modern society with all modern civilisations reliant on various types of infrastructural systems ranging from power grids to water distribution systems and transportation networks. One method to represent the connective nature of critical infrastructure is to represent large structures as systems with nodes and links connected via the respective topological nature of such systems. Each system is allocated a specific mission time which is dictated using data from prior mission times [1] of comparably merited systems and reliability estimations can be postulated from this data.

Systems can be divided in sub-categories, such as discrete and continuous systems. Discrete systems contain nodes and links and use Boolean algebra to define an individual node's state, with either 1 indicating a functioning component or 0 indicating a failed component [2]. Continuous systems differ from discrete systems as their mission time is infinite and their outputs are not defined by Boolean algebra and are alternatively defined as continuous values representing a magnitude of a specified output parameter. The majority of systems applied in critical infrastructure analysis are classified as continuous systems and therefore a large volume of work has been researched on such systems. It is also key to analyse discrete systems prior to the analysis of continuous systems as discrete systems can provide estimations regarding system failures

in the field of reliability analysis. Such examples include fault tree analysis [3], survival signature analysis and Bayesian networks.
Systems can also be classified as either repairable or non-repairable systems with repairable systems containing components that can be restored to their optimal state during the downtime of a mission. Conversely, the components of non-repairable systems cannot be restored mid-mission and are classified as failed until the mission ends [4].

1.2 Motivation

Prior events in history have shown that disasters of critical infrastructure systems have induced detrimental impacts to communities, both socio-economically and towards quality of life. The motivation for this research project therefore, is to communicate risk in the form of resilience to experts with intentions to mitigate the likelihood of impacts and minimise the consequences of disasters in critical infrastructure systems. Before the aims and objectives are addressed, it is important to highlight the severity of a range of past events that have occurred to remind the reader of the motivation for this research project. Consequences from each event vary depending on the nature of the event and its respective severity. It is also vital to admit that a failure in a system is inevitable as a system ages and the reliability of the system decreases over time. This is when resilience analysis is applied, as resilience aims to accept the reality of possible disasters, but aims to minimise consequences. Such methods include redundancy analysis and resourceful analysis, both which assess different routes and look for other pathways to establish mission efficiency and completion despite the system taking on damage and performance loss.
This work is formulated to provide a balanced analysis to resilience in both a qualitative and quantitative angle. The scarcity of data uncertainty in the literature provides a drive to explore resilience in a potentially novel approach as researchers can work together to improve scientific literature for this particular void.
All research carried out has been trialed in an academic perspective and gaining a deep understanding of the data analysis, simulations and communicating work in conferences is a positive step into the work of resilience as all the knowledge obtained from this work is transferable into industry.

1.2.1 1999 Ladbroke Grove Rail Crash

The 1999 Ladbroke Grove rail disaster [5] currently registers as the worst British rail disaster in modern times accounting for 31 fatalities and injuring over 400 people [6]. The disaster occurred approximately two miles before the train reached its terminal destination. A three car train filled with commuters in London's early morning rush hour collided with an eight coach high speed train at a combined speed of 130 mph [7].

The direct cause of this incident can be linked to a long standing human factor problem known as a signal passed at danger (SPAD) occurring when a train driver passes through a signal with a stop sign. In this case the driver believed a "proceed" signal was delivered. The defence system of the railway was composed of multiple layers to prevent an accident occurring. The defences for this particular system are displayed on Table 1.1 [6].

Defence	**Driver response**	**Train response**
SN63 AWS horn	Cancel AWS	AWS horn cancelled
SN63 at double yellow	Select speed control notch 0 and brake level 1 applied and released	Train coasting at 45 mph then braking to 41 mph and coasting to 39 mph
SN87 AWS horn	Cancel AWS	AWS horn cancelled
SN87 at single yellow	Allow train to continue coasting for 30 s, then select speed control notch 5	Train coasting at 39 mph then cruising at 38 mph
Driver Reminder Appliance	Not applied	Train coasting but driver is still able to draw power
SN109 AWS horn	Cancel AWS Select speed control notch 7	AWS horn cancelled Train accelerating from 38 mph to 50 mph
SN109 at red	No response	Train accelerating from 38 mph to 50 mph

Table 1.1: Defences for Ladbroke Grove Rail crash

The Ladbroke Grove train disaster follows a story which violates every layer of defence designed on the system. These events are extremely rare due to the multiple layers of defence the system has to offer.

1.2.2 2012 India Blackout

The 2012 India blackout occurred on the 31st July 2012 and was the largest blackout in human history [8] with an estimated 620 million people or 9% of the world's population affected and lasted for 8 hours [9]. The event is believed to be caused by the transmission line tripping as a result of being overloaded and an estimated power loss of 48 GW was recorded.

The direct causes of the disaster are human factors and are also believed to be attributed to the engineering design of the system. The inter-regional corridors were believed to be weakened and this was thought to be caused by an overload of outages on the transmission lines with an excessively high load between the northern region to western region transmission lines in particular. The response team working on state load dispatch centres are believed to have miss-communicated information for the instructions to regional load dispatch centres regarding reducing the load emitted from the northern region to the western region. Additionally, there were issues with the protection system for safeguarding such events such as a 400 kV line being tripped due to load encroachment [10].

The other factors contributing to this event include [11];

- Frequency control
- Primary response from generators
- Operation of defence mechanisms
- Transfer capability
- Co-ordination of outage planning for transmission elements
- Reactive power compensatory tools
- Analysis tools
- Support assets for the occurrence of load contingencies
- Wide area measurement systems
- Analysis tools

1.2.3 Deepwater Horizon Oil Spill

A pipe burst of a an oil or water distribution system is a common occurrence. Most notably one of the most targeted disasters was the 2010 Deepwater Horizon oil spill which occurred in the Gulf of Mexico. During April of 2010, the semi-submersable oil rig Deepwater Horizon caught fire and exploded, taking the lives of 11 people and injuring 17 more [12]. Following this event, global media was highlighted at British Petroleum (BP), the lease operator. The consequences of the disaster were huge ranging from the fatalities caused, the long term effects such as cardiovascular health conditions caused by exposure to the leaked crude oil [13] to the adverse mental health affects for those directly or indirectly affected [14].
The system's failure was caused by multiple failures in combination. Labib [15] provides a full technical analysis of the causes of the disaster;

- Poor design of cement barriers - The primary cause of the system failure was due to an engineering design problem [16]. This cause was the leakage of hydrocarbons from the bottom of the well into the drilling tower. This event further caused the leakage of gas into the engine room. After the disaster, investigators noticed that the cement barrier was poorly designed when they tested for the robustness of pressure management and discovered from past history that the cement barrier was not even tested for pressure handling.

- Mechanical failure of the blowout prevention mechanism - The Blowout Preventer (BOP) is an important component in the system as it mimics the function of valves to regulate pressure in the system. It was soon discovered that this BOP was not tested to an adequate standard and was damaged due to the high pressure hydrocarbon entering in from the cement barrier and into the BOP gasket [17]. Following this, oil and gas was released on to the oil rig's surface, which was expected to be ventilated by mud and gas separators. In this case, ignition was initiated as the hydrocarbon gas entered the power generator room which provided the ignition source. Fire and gas mixing occurred as the planned safeguard mechanism to prevent this was not functioning, resulting in the explosion.

- Control cables damaged - The second explosion which occurred 10 seconds after the first explosion was caused by an indirect event following the first explosion.

The BOP was connected to the control room and the cables used to connect the two components were damaged from the first explosion [18], leading to no communication to the control room.

It is clear to deduce that the Deepwater Horizon oil spill was a preventable disaster given that sufficient safeguarding and audits were to be implemented. Both direct and indirect causes led to this disaster with a mixture of technical, design and financial reasons.

1.3 The role of Resilience in Critical Infrastructure

The examples listed above are just a small number of the many types of disasters for a select few critical infrastructure systems. The list is ongoing and these examples provide an insight into the possible disasters and disenfranchised performance losses resulting from a lack of awareness of resilience based studies. There are various definitions and interpretations of resilience available, both in an engineering context and in a general form [19]. The United Nations International Strategy for Disaster Reduction defines resilience as "The capacity of a system, community or society potentially exposed to hazards to adapt, by resisting or changing to reach and maintain an acceptable level of functioning and structure" [20].

Academic research carried out shows a number of different angles that have been applied to resilience analysis when tested on critical infrastructure engineering systems. These approaches can vary depending on the type of system that has been analysed and the type of resilience that has been quantified. For example, with regards to rail based infrastructure systems, an objective function can be developed based on the the delay of train times when tackling the operational resilience and with regards to the infrastructural resilience, the type of analysis could be the structural loss over time with respects to the railway track and its interaction with the train along as internal damage recorded from use of passengers [21].

An example that can be applied to Section 1.4.2 is an approach to quantify resilience as an objective function and has been carried out by Ghasemi et al. [22] who proposed a new restoration strategy to re-energise the critical loads of the system. The approach applies three objective functions to carry out a restoration plan. These three objective

functions are restored weighted energy, preparation of time plans, and switching operations cost respectively. The methodology selects the most efficient objective function at each stage of the simulation using the PROMOTHEE-II technique with a judgement matrix further applied to select the correct weighting of the critical loads. The model then displays the output curve with respects to uncertainties in the loads.

1.4 Challenges in Resilience Quantification

In the field of system simulation, there are many challenges that arise to provide accurate results in a presentable manner. The first challenge is regarding data collection and uncertainties around which data to use. Conflicting data is a very common challenge in academic research papers and the way this constraint can be overcome is through the introduction of data verification, data validation [23] and the addition of epistemic uncertainty when multiple data sources are applied for a desired system's input.
With regards to the simulation, the biggest constraint is often seen in systems with a large amount of data and through very precise algorithms incorporating many stages to achieve the output goal [24]. Realistic systems that contain thousands of nodes and links are prone to computationally expensive simulations which is further elongated depending on the methods applied for resilience quantification [25]. The challenges addressed within the simulation methods presented in this thesis tackle the issue of balancing computational expense with input data accuracy, system size and uncertainties. Each method presented provides the computational approach recorded in MATLAB 2021b and the alternative approaches to minimise computational time have been trialed out.
The challenges faced within the software that have been applied have been found to be within the applicability to provide sufficient simulation techniques in order quantify the objective purpose. In order to tackle this limitation, a combination of the various toolboxes are applied where applicable and results are transferred from one piece of software to the next. Such examples are highlighted in Chapter 3 where the GeNIe Bayesian Network software has been tested alongside the Credal Network toolbox in OpenCossan to convert precise data and deduce its respective imprecise form. A further example of this is presented in Chapter 5 which also applies conversion of data

from PipeFlow Expert into MATLAB 2021b in order to derive the simulation results as presented in the respective chapter of this thesis.

1.5 Aims and Objectives

In the field of resilience, although significant research and efforts have been dedicated to resilience analysis, there are still a number of open research questions and challenges. It is not always straightforward to develop realistic models with scarce literature and a lack of data. Therefore, the primary aim and objective of this thesis is to present an analysis of various resilience models in engineering systems with real world applications to represent an illustration of each model's respective results.
A list of these aims include;

1. Understanding the state of the art of resilience as applied to engineering systems;
2. Exploring computational reliability models with regards to discrete binary-state systems with respects to the connectivity of a system's topology;
3. Application of current resilience models to multi-state systems with regards to respective output models;
4. Linking multiple forms of resilience quantification into one output metric to enable a clear and concise method for the user to interpret;
5. Highlighting the strengths and diagnosing the limitations as discovered in the resilience methods explored.

With respects to the objectives, a series of tasks have been carried out to produce the credible data, results and information for this thesis including;

1. Investigating resilience via collecting real world data for the desired systems and analysing this data to be applied to the respective simulations;
2. Organising data for evaluation and simplifying this data for the required inputs for the respective models;

3. Implementing resilience algorithms and numerical solutions from the literature into a computational package for MATLAB 2021b;
4. Innovating current resilience based algorithms with the addition of uncertainty;
5. Writing up the results in a cohesive manner to clearly communicate and evaluate the findings of resilience as applied to critical infrastructure.

1.6 Thesis Structure

This thesis is divided into six (6) chapters, with each chapter focusing on a different aspect of resilience quantification and different applications with real world critical infrastructure systems. Chapter 1 includes the necessary background information and main motivations of this research thesis along with the examples stated before. The aims and objectives presented in the problem are clearly addressed along with the research questions needed to be asked.

Chapter 2 lists a full and updated literature review that has been carried during this research project, displaying the most important papers analysed for prior knowledge to aid with the writing of this thesis. The current resilience techniques along with their challenges are presented.

Chapter 3 for instance, combines two approaches, the survival signature and analytical probability propagation method to quantify reliability in a discrete system and expands this into resilience analysis with the addition of uncertainty.

Chapter 4 provides an application of an existing resilience quantification technique, the load flow Monte Carlo simulation model into a real world study of the UK power network. The chapter uses artificial neural networks to add a surrogate model for faster simulation techniques when applied to various optimal power flow simulations.

Chapter 5 provides a gas pipeline network application and explores resilience with regards to the mass flowrates of gas.

The final chapter concludes the various topics explored in the thesis and provides a critical analysis of the successes, failures and limitations within the study along with future work that can be carried out in more advanced scenarios to innovate the current state of the art as mentioned on this thesis.

Chapter 2

The State of the art in Reliability and Resilience Modelling

2.1 Current Reliability Estimation Techniques

Techniques to estimate reliability of a system have been developed in order to minimise computational time with respects to analysing the topology of a system. These techniques provide a different approach to reliability analysis with respects to the performance parameters being analysed and the methodology of evaluation.

The traditional method of Monte Carlo simulation has been a technique used for the estimation of components and furthermore subset simulation has been developed in order to provide an estimate of reliability. Subset simulation works by performing a limited number of samples extracted from a complete set of samples by converting small failure probabilities into a series of larger failure probabilities by introducing intermediate events in stages [26].

When quantifying failure probabilities which are considered too small for the applicability of Monte Carlo simulation (e.g.$P < 10^{-4}$), line sampling [27] has been implemented to quantify reliability in these extremely unlikely probabilistic events by computing the "Important Direction" and is estimated by the gradient of the standard normal space of the performance function.

Other techniques use fault trees to represent system reliability and furthermore include

dynamic fault trees [28] which apply graphical models that describe how failures propagate through the system and how these component failures affect the whole system leading to failure. This is represented by a Direct Acyclic Graph (DAG) with logic gates representing how failures propagate through the system. Static fault trees use "AND" and "OR" gates to represent the path set into failure for a system. Dynamic fault trees extend the idea of a static fault tree by introducing additional dynamic gates including priority sequence enforcing, functional dependency, spare gate and load sharing gates [29]. Other dynamic reliability assessment techniques include dynamic reliability block diagrams (RBDs) [30], dynamic flow graphs [31] and petri-nets [32].

Along with petri-nets, artificial neutral networks (ANNs) are used to provide a meta-model route in estimating reliability for a system and can be seen as an alternative to traditional techniques as it employs machine learning techniques to use past historical data as network architecture which involves training a genetic algorithm to mimic an original model [33][34]. The universal generating function (UGF) was introduced to evaluate reliability in multi-state systems and uses reliability optimisation algorithms to map out the system's performance distribution function [35].

2.1.1 The Survival Signature

The survival signature, a method developed by Coolen and Coolen-Maturi first published in the article [36] was first developed in 2013. The survival signature is a method which extends on the work of the general system signature by the addition of the topological analysis of systems with multiple types of components which contain independent failure times. The survival signature approach is a means invented to estimate a network's reliability given the condition that a certain number of random nodes are active for any defined number of component types. It has been developed from the general system signature, previously developed by Samaniego [37] and extends this concept to enable the signature's analysis for system's with multiple types of components that contain non-exchangeable failure times, which could include systems with two or more types of components.

The first approach to uncertainty for the survival signature was tested by Coolen et

al. [38] as an application of non-parametric predictive inference for the survival function, using bounds to denote uncertainty. Patelli et al. [39] published another approach to uncertainty for the survival signature, and takes an approach of adding bounds to the Weibull shape and scale parameters in the survival function. The authors propose three different algorithms, two on non-repairable systems and one on a repairable system. Further works on the survival signature have been carried out and include the application of marginal and joint reliability importance as applied to coherent systems consisting of multiple types of components [40] and reliability analysis of phased mission systems (PMS) with the implementation of the traditional survival signature to PMS with similar types of components in each phase [41]. Furthermore, Li et al. [42] developed a reliability-redundancy allocation optimisation and reliability sensitivity analysis for redundant components. Most recently, Behrensdorf et al. extended the work on the survival signature to enable the application of efficient computation in very large networks [43], extending the survival signature to systems that can contain thousands of nodes.

2.1.2 The Analytical Probability Propagation Method

The Analytical probability propagation method (PrPm) developed by Tien and Tong [44][45], which incorporates belief propagation in order to perform inference on a network is an innovation to current reliability techniques applied to Markov models. It computes the joint probability distribution between nodes on each stage of the propagation and carries this information on to the next propagation step until the terminal node is propagated. It can be argued that the PrPm is favourable over Monte Carlo simulation methods as Tien and Tong have proved that the deviation from the actual values are less than the Monte Carlo based methods, arguing greater accuracy with the results. The PrPm also has certain advantages over the imprecise bayesian network, known as the credal network as the computational times are significantly shorter too. The fundamental difference between the survival signature and the probability propagation method is the addition of line probabilities to the probability propagation method which is also not applied in a bayesian network. The line probabilities influence the reliability of the system's terminal probability and are included in the updating calculations for the progressive steps of the method.

Both the survival signature and the PrPm methods are applied in Chapter 3, and this thesis innovates on the current literature applied to the probability propagation method by adding the bounds for epistemic and aleatory uncertainty.

2.1.3 Limitations of the Proposed Approaches

The major limitations of these two approaches include the approximation of the uncertainties being deduced from estimation. It is not always clear that these input estimations are reliable when data is so scarce for these estimations to be applied.
The system simulation procedure for the survival signature becomes exponentially computationally expensive as more types of nodes are introduced and eventually becomes unfeasible for imprecise probabilistic quantification.
The PrPm also becomes exponentially computationally expensive when the inner approximation algorithm is applied and becomes computationally impossible in a larger size network to produce results with the exact algorithm. The outer approximation is computationally feasible for larger networks, but produces confidence bounds with a wide range of uncertainties.
These limitations do not greatly affect the credibility of the results due to the aim of Chapter 3 focusing on estimation towards reliability of systems, therefore results are deemed as credible despite potential accuracy loss throughout these assumptions and limitations.

2.2 System Resilience

The field of reliability is a well established research field in system engineering with journals such as the “Journal of Risk and Reliability” containing years of literature on this topic. However, the field of resilience is a newly established topic in recent years. Resilience plays an important part in terms of extending reliability analysis with the aims of this new analysis contributing to the creation a safer world by minimising probabilities of failures and furthermore accepting failures with the minimisation of infrastructural and operational damage.

2.2.1 The Four R's

Works on availability and reliability aim to quantify and communicate the minimisation of the probability of failures from occurring. In contrary, resilience analysis aims to accept the occurrence of these disasters, however aims to minimise the impacts and consequences of such disasters in different approaches and furthermore aims to produce recovery data in a post disaster phase, applying the process known as three-phase resilience. These models were one of the earliest academic developments in the field of resilience, and were developed by Bruneau et al. [46] and they include operational loss minimisation, rapidity modelling to analyse recovery time analysis, resource based analysis to search for alternative routes for mission time for success, and redundancy modeling to accept the loss of a system but quantifying and communicating work despite the partial loss. Each particular sub-field of resilience analysis adds a different utility in the overall picture and these individual metrics can be used to communicate the relevant information as applied to a particular system. The four R's developed by Bruneau et al. are detailed as follows;

- Robustness - The analysis of the strength and ability to withstand a given stress level without damaging the operation of a system. A robust system is able to undergo a large level of stress without reduction in performance, whereas a less robust system will fail with respects to the same levels of stress [47].
- Redundancy - Analysing the system's performance based on the desired metrics with the known state that a damaging event has been done to the system. A redundant system is able to satisfy its purpose even when post-disaster damage has occurred [48].
- Resourcefulness - The system's ability to utilise resources in the event of a disaster. This could range from finding alternative routes to a system's success, or the application of human resources to achieve certain goals [49].
- Rapidity - The speed of a recovery during the post disaster phase. This is a vital parameter to assess when analysing transient based resilience models as it is used to communicate the minimisation of future disruptions in the system [50].

This method to represent resilience can be related to technical, social, organisational and economic resilience based models as further explain by Bruneau et al. [46];

- Technical resilience - Resilience analysis based on the topology of a physical system. This could incorporate component analysis based on operational performance, connectivities and interdependencies.
- Social resilience - Focusing on analysing and therefore minimising the negative social consequences caused from disasters to critical infrastructure. This includes predictive modelling to analyse potential consequences caused to a community if a certain type of disaster occurs.
- Organisational resilience - Assessing the role and responsibilities of third party organisations to ensure resilient critical infrastructure is developed to support the four R's.
- Economic resilience - Analysing a system's resilience with respects to the four R's to quantify possible direct or indirect economic losses in the event of a disaster or downtime in the system.

2.2.2 Definitions of Resilience

Various definitions of resilience have been proposed in recent works, however the most applicable definition applied to critical infrastructure is defined by H2020 European project Improved Risk Evaluation and Implementation of Resilience Concepts to Critical Infrastructure (IMPOVER) stating resilience as "the ability of a CI system exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, for the preservation and restoration of essential societal services." [51], a definition extended from the UNDRR (United Nations Office for Disaster Risk Reduction). However when applied to specific measures of resilience, further definitions can be applied and the work of Plotnek and Slay [52] provides a table of these definitions when applied to various resilience output parameters, with the most important listed in Table 2.1.

Definition	**Application**	**Real world threat**
"The resilience of a system presented with an unexpected set of disturbances is the systems ability to reduce the magnitude and duration of the disruption. A resilient system downgrades its functionality and alters its structure in an agile way." [53]	Reduce intensity, event duration	unplanned disruptions such as earthquakes and storms
"The ability of an entity to anticipate, resist, absorb, respond to, adapt to and recover from a disturbance" [54]	absorption, response	any event
"Anticipate possible disasters, adopt effective measures to decrease system components and load losses before and during disasters, and restore power supply quickly. Additionally, valuable experience and lessons can be absorbed from disasters suffered, to prevent or mitigate the impact of similar events in future." [55]	restore, rehabilitate, adopt	natural disasters

Table 2.1: Resilience definitions

Every definition of resilience applies its significance on different parameters on the output of the system and applies a definition regarding the various stages of the simulation. This could range from initial shock in the system, to downtime and finally recovery. The table highlights which instance the particular definition of resilience is applied to and also possible examples of real world applications for which the respective study can be tested on.

2.2.3 Operational Resilience and Infrastructural Resilience

In the field of critical infrastructure, resilience analysis can be divided into two major categories, operational resilience and infrastructural resilience as defined by Panteli et al. [56]. Operational resilience is defined as the characteristics of a system that is required to provide the system's respective utility for function. In the case of water distribution systems, this could be defined as the volumetric flow rate supplied to the pumps of the system. An event which causes a decrease of volumetric flowrate

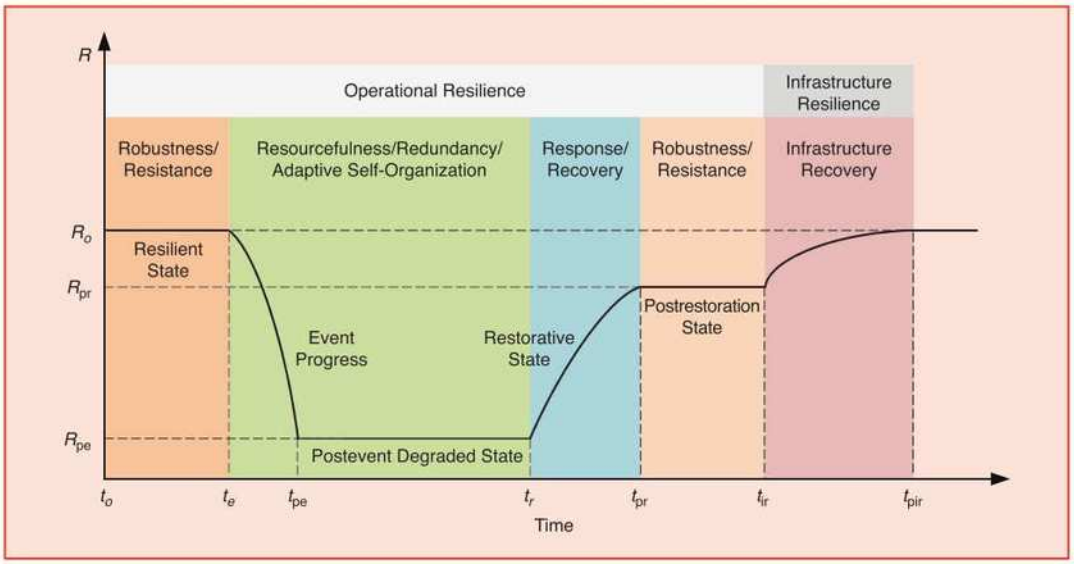


Figure 2.1: An expansion of the three phase resilience plot

quantity to a pump classifies as an operational loss and therefore a loss in functionality of the system. Infrastructural resilience on the other hand refers to the state of the physical system and specialises on the mitigation of loss in the system's structural failure. The two are interconnected and can be modelled together to quantify a combined resilience based function. Figure 2.1 [57] highlights the difference between a typical resilience simulation model between operational resilience and infrastructural resilience as displayed by Panteli et al. with the authors highlighting the stages of the four R's of resilience which have been discussed in Section 2.2.1.

The three phase resilience model is expanded into a five phase resilience model with operational resilience being the driving factor between phases 1-4 and the final phase solely containing infrastructural resilience. The primary phase is the resilient state, and represents the time of the event prior to disaster, and therefore the optimal function is retained throughout this phase. Robustness and resistance are the key parameters during this early state. This is followed by the event where the performance disintegrates and reaches the post event degraded state. In this phase, resourcefulness, redundancy and re-organisation are applied as the key parameters for resilience. Upon restoration, the performance function increases due to the repairs taken place within the interior in the system and the key parameters are the rapidity and recovery magnitude of the

system. This is followed by a post restoration phase as the final phase of operational resilience. During this phase, the performance efficiency stays static and the key parameter during this stage is the robustness parameter. Finally, the infrastructural resilience is applied to the model and is the terminal stage to rehabilitate the system.

2.3 Current Resilience Analysis Techniques

As an extension to the reliability analysis techniques mentioned in Section 2.1, the application of resilience is also implemented in system simulation. Different forms of critical infrastructure contain varying parameters to model and system safety. The analysis of system resilience requires the collection of credited data, modelling and testing of system performance and its evolution as a function of internal and external events. These techniques range from resilience analysis as applied to the connectivities of the system to time based analysis methods to quantify resilience as a transient output parameter. Infrastructural resilience has been portrayed in the field of structural engineering in the work of Chhaia et al. [58] by associating a structural resilience index to both a pre-event and a post-event state. The arbitrary structural resilience index is conformed from certain parameters deduced by the nature of the structure as stated in the article. This approach is applied to general systems, and shows a flexible approach to the quantification of multi-dimensional resilience.

Lu et al. [59] investigated the post-disaster phase of resilience with respects to urbanised areas. The work includes splitting resilience into physical resilience and socio-economic resilience. The method combines Bruneau's work [46] with the addition of Xiong et al. [60] and proposed a new framework for a city scaled transient history analysis for quantify building seismic resilience in repairable systems. The model computes post-earthquake residual functionalities using engineering demand parameters and scheduling repairs. Ma et al. [61] developed an approach to quantify resilience in general power systems by combining both infrastructural resilience and operational resilience. The infrastructural resilience is modelled by analysing the topology of the system and deducing the pertinent static and dynamic network characteristics and the operational resilience is deduced via assessing the progression of extreme events and how this affects the system. The approach applies complex network theory to analyse

the structural resilience and operational resilience of power systems and results in the evaluation of preventative, corrective and restoration strategies.

2.3.1 Resilience Applied to Train Networks

In recent works such as Bešinović et al. [62], the authors developed a technique for a passenger centred resilience assessment taking into account scenarios with multiple disruption events. Infrastructural resilience is the output measurement goal and an optimisation based restoration function has been developed combining traffic operations and passenger flows as inputs to the restoration function. The novelty of the technique combines redundancy with restoration by finding routes for the best possible infrastructural restoration path and enables decision makers to quantify the effects of multiple sources of infrastructural damage in the angle of resilience.
Fang et. al [21] applied Monte Carlo simulation to mimic multiple hazards from typhoons in a multi-phase model. The typhoons are generated by a spatially localised failure model which is simulated via Monte Carlo into both single and multi-mode failure scenarios. The authors also experimented the effect of resilience with different typhoon intensities and results displaying the different restoration scenarios of the various modes.
Chen et al. [63] applied a three phase resilience model incorporating a disturbance, response and recovery phase. The performance is measured using the author's developed indicator which is the demand-impedance indicator. This metric is defined through the passenger frequency of trips and the time of each trip. An effective path betweenness indicator is proposed which takes into account the passenger's travel path and and the stations which are represented by nodes are given various importance factors. The performance function during these phases are defined through the demand-impedance indicator and a three phase resilience triangle is formulated. The correlation coefficient matching the betweenness factor and the betweenness centrality is finally deduced. The application of this work was applied to Chengdu's subway network and can possibly be expanded using the wider China network for further investigation.
Chapter 3 applies train networks with an estimation of reliability to provide a strong foundation for resilience studies which could be expanded with applications of the work listed above.

2.3.2 Resilience Analysis applied to the Power Grid

An example of efforts applied to analyse resilience in power grids have been studied by Jufri et al. [64] who included various techniques such as transient performance modelling for the case study of typhoon Bolaven in South Korea. However the authors mentioned that the limitations in their study included only computing resilience in the form of restorative and absorptive capacity without considering anticipated and adaptive capacities and also did not include a cost-benefit analysis to analyse resilience in an economic angle.

Panteli et al. [65] developed a resilience based method for power grids with extreme weather events by developing the three phase resilience trapezoid [56]. The authors define resilience as "the network's ability to withstand high-impact low-probability events, rapidly recovering and improving operations and structures to mitigate the impact of similar events in the future". This is an extension to the traditional resilience triangle developed [66] which involves three stages to the disintegration, stagnation and recovery of the structure. Kim et al. [67] developed a novel function to analyse the South Korean power grid network using cascading failure analysis by applying three different node centrality metrics; degree, clustering coefficient and betweenness. A high clustering coefficient of a network indicates a more resilient network as it contains a higher redundancy potential to utilise alternative paths within the network.

The performance of systems and networks are usually analysed adopting complex models able to provide a range of different metrics that can be analysed individually or simultaneously. For instance, George-Williams and Patelli [68] proposed a simulation driven approach coupled with load-flow methods for estimating the reliability of complex and multi-state systems and later extended to the analysis of reconfigurable systems with interdependencies [69]. One of the main limitations of the use of models for analysing large complex systems is due to the computational time required to produce accurate results. To overcome such limitation, surrogates or emulator models are often used. Rocchetta et al. proposed the used of a Power-Flow Emulator for the resilience analyse regarding the effect of weather induced failures [70] by applying an artificial neural network model to replace the computational expensive original AC optimal power flow model [71]. The proposed approach has been further extended to deal with data deficiency and imprecision adopting p-boxes for the robust quantification of uncertainty [72, 73].

Chapter 4 presents a holistic analysis of resilience applied to power grids in the context of a load flow approach.

2.3.3 Resilience Analysis applied to Gas Pipelines

Works carried out on resilience applied to gas pipelines are relatively recent in contemporary literature with examples of works such as Sang et al. [74] proposing a novel framework for restoration strategy optimisation of gas networks. The method applies system functionality metrics to define the operational characteristics and mimics real-time performances followed by developing resilience metrics by recording past recovery features of the system performance. The authors developed a restoration sequence based optimisation model to determine the restoration phase and optimise resilience with respects to recovery nodes, repair times and recovery costs. The proposed method minimises computational expense by the application of a skeleton network based reconfiguration model to identify any critical components in the system. Linearisation methods are also used for the optimisation to transform models into mixed-integer linear programming (MILP) problems. The results concluded explaining that the method is an effective way to guide system operators and perform restoration decisions of failed components in downtime situations.

Su et al. [74] also developed an integrated dynamic model including input parameters such as pipelines, compressor stations, junctions and liquefied nitrogen gas (LNG) terminals. The properties of these parameters are integrated into graph theory to be modelled. The quantifying method for resilience applies this model by simulating system response given the conditions of different operational methods and evaluates resilience consequences from this, which is specifically applied to gas pipelines.

This work is expanded by Marino and Zio [75] who developed a robustness model to quantify resilience with the considerations of cybernetic interdependence of gas pipeline networks with the application of a supervisory control and data acquisition (SCADA) system. The quantification of resilience is achieved through specific performance metrics and the maximum flow algorithm calculates the gas network's supply which is subject to change through pressure changes in the system. This method integrates thermal-hydraulic simulation, graph theory and wireless network simulation collectively and applies a sensitivity analysis to analyse the uncertainty of the system.

These two approaches are tested in Chapter 5 which carries out work to combine the efforts placed in these two papers.

2.4 Uncertainty

In the field of resilience applied to engineering systems, the literature almost exclusively communicates data as precise figures, with lack of efforts placed into imprecise data to communicate to experts.
Uncertainty quantification for reliability analysis has been shown to cause simulation times to increase by a large fraction of original simulation time and yet for this reason remains a less studied concept in the analysis of systems. However it still remains a vital topic to explore despite this limitation. Every system in life is bound to contain some type of uncertainty within the input and output parameters.

2.4.1 Epistemic and Aleatory Uncertainty

Modern uncertainty models include both epistemic and aleatory uncertainty. Epistemic uncertainty arises due to the lack of sufficient input data into the model to obtain sufficient accuracy for the definition of the output distribution [76], whereas aleatory uncertainty arises due to the inherent variability in output data with regards to the studied phenomenon [77].
Uncertainty with regards to epistemic data is communicated in intervals, written as lower bounds and upper bounds respectively. An example of another method to communicate epistemic uncertainty includes fuzzy sets, and applying the fuzzy set theory to model uncertainties via the natural language. Zheng et al. [78] developed a fuzzy optimisation model of control design to minimise fuzzy performance with respects to uncertainty.
Gray et al. [79] proposed a new Bayesian calibration model to display uncertainty capturing both epistemic and aleatory uncertainty uncertainty via the application of a multi-dimensional second-order probability distribution. The authors proposed that all engineering design problems share common challenges with regards to uncertainty;

- Variability - It is inevitable that engineering systems will experience operation under different environments. The conditions of these environments affect the performance of the systems operation and therefore the system should be designed to withstand changes in the variability of such conditions.

- Inference - Epistemic uncertainty arising as a result of novelty in engineering design. In some scenarios, data is so sparse that it is lacking entirely, and therefore expert judgements may be applicable. A suitable design agenda incorporates this data as it progressively becomes available and innovates on a system's design to optimise performance.

- Ignorance - The design phase of engineering systems are conformed of several parameters with many unknowns, or even in some cases dependencies on parameters that are unknown. A conservative agenda for engineering design will take into account all uncertainties that could play out as a result of the lack of knowledge of the parameters in the system.

- Decision making - Uncertainties in the engineering system from decision making such as data collection contribute to uncertainty within the design phase. Such decisions are postulated with limited information and therefore must take into account the uncertainty present in the available data and adequate flexibility should be provided in the design stage.

All of these components are essential and must be considered when modelling a system with respects to uncertainty by developing a framework to quantify all these forms of uncertainty simultaneously whilst keeping output parameters optimised.

2.4.2 Characterisation of Uncertainty

Uncertainty in engineering systems is characterised in various forms depending on the proposed display outcomes. Interval probabilities are used to represent bounds in a probabilistic variable and can be defined to be the upper and lower bounds of a probability distribution. Lower bound probabilities denoted as $\underline{P}$ and upper bound probabilities denoted as $\overline{P}$ can be applied to a value X as a super additive capacity [80]. $\underline{P}$ and $\overline{P}$ are direct complements of each other such that $\overline{P}(A) = 1 - \underline{P}(A^c)$.

The credal set is given as;

$$P_{\underline{p}} = \{P \in \mathbb{P}_X | \forall A \subset X, P(A) \geq \underline{P}(A)\} \tag{2.1}$$

Probability intervals are defined as the lower and upper bounds of a probability distribution and can be defined as a set of numeric intervals, $L = \{[l(x), u(x)] \,|\, x \in X\}$ such that $l(x) \leq p(x) \leq u(x), \forall_x \in X$ where $p(x) = P(\{x\})$. A probability interval denotes the credal set as applied;

$$P_L = \{P \in \mathbb{P}_X | \forall_x \in X, l(x) \leq p(x) \leq u(x)\} \tag{2.2}$$

2.4.2.1 PDFs and CDFs

The probability density function (PDF) is used as a means of modelling the generative process for observed data [81] and is used to predict the likely outcome of a discrete random variable. Recovery and modelling the generative distribution is the objective in statistical inference, which enables the analysis of uncertainty in various parameters and can be used to produce predictions in future data. PDFs are often displayed as bell curves, with the peak of the curve being the most likely output, or specifically, the expectation of the discrete random variable. The PDF $f(x)$ has contains two important properties;

1. $f(x) \geq 0 \quad for\ all \quad x$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

A simple way of displaying a PDF is through the normal (Gaussian) distribution, in which the equation is displayed below;

$$PDF(X) = \frac{1}{\sqrt{2\pi} \cdot \sigma} exp\Big[-\frac{1}{2}\Big(\frac{X-\mu}{\sigma}\Big)^2 \Big] \tag{2.3}$$

where μ represents the Gaussian's mean and σ is the standard deviation.

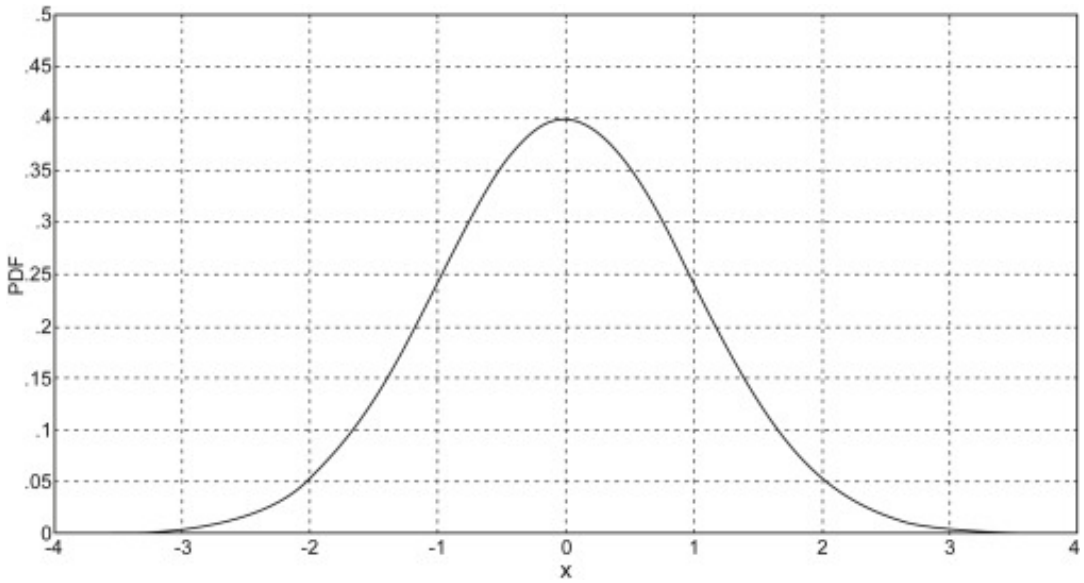


Figure 2.2: A PDF of a Gaussian distribution

Figure 2.2 displays a PDF of a Gaussian distribution for $\mu = 1$, $\sigma = 1$ [82];

Cumulative distribution functions (CDFs) can be applied to model the distribution of random variables in cases where the inputs are either discrete, continuous or mixed, and can be defined as the integral of a function's PDF;

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \tag{2.4}$$

An example of a CDF is presented as a p-box in Figure 2.3

2.4.2.2 Monte Carlo Sampling

An alternative method of deducing uncertainty is through Monte Carlo sampling of a system. This procedure consists of sampling the system simulation procedure multiple times and computing a probabilistic estimation of the reliability. The maximum a posteriori estimation is deduced as the most probable value and the uncertainty bounds provide the degree of confidence of the estimation. Zio et al. [83] developed a novel systematic procedure to evaluate the availability of multi-state, multi-output plant systems. The authors approach applied a Monte Carlo simulation model to generate a random walk which guides the system from one state to the next state and uses the

history of these samples providing a mean availability level of the production system at a given time. Lye et al. [84] applied double loop Monte Carlo sampling to estimate the reliability of a dynamical black box system under uncertainty as part of the NASA-Langley Uncertainty Quantification Challenge 2019 [85].

2.4.2.3 Probability Boxes

Ferson et al. [86] developed Probability boxes (p-boxes) as a method to represent uncertainty. P-boxes are imprecise probabilistic distributions which reflect both epistemic and aleatory uncertainty of a given quantity of interest. Beer et al. [87] states that the representation of p-boxes combines the dual ideas from evidence theory and probability intervals. Representing imprecise probability distributions as p-boxes are advantageous as the approach allows analysts to deduce the output without any precise figures for the system's input distributions. When data is in mass and is strong in confidence, output confidence bounds of the p-box is slim, and provides and approximate of the precise distribution applied in Monte Carlo simulation. However, when data is scarce and weak in confidence, output confidence bounds of the p-box is wide, representing a greater quantity of uncertainty.

P-boxes can further be divided between parametric and non-parametric models [88], with parametric p-boxes modelling the set of all the possible combinations of distributions obtained from a stated distribution function with the imprecise parameters set as known values. A non-parametric p-box models all the non-decreasing distributions that are present between upper and lower probability bounds.

A simple p-box can be mathematically represented as [89];

$$\mathscr{F}(x) = \left[\underline{F}(x), \overline{F}(x)\right], \quad \underline{F}(x) \geq \overline{F}(x) \quad \forall_x \in \mathbb{R} \tag{2.5}$$

where $\underline{F}(x)$ represents the lower bound function of the p-box and $\overline{F}(x)$ represents the upper bound function of the p-box.

Ferson also developed a method to convolute continuous p-boxes [86]. This method involves discretising the p-box into a finite list of pairs, $(A_1, m_1)...(A_n, m_n)$, where A_i represents an interval, m_i represents the mass probability of A with respects to

$\sum_{i=1}^{n} m_i = 1$.

A discretised p-box is represented in Figure 2.3 [90].

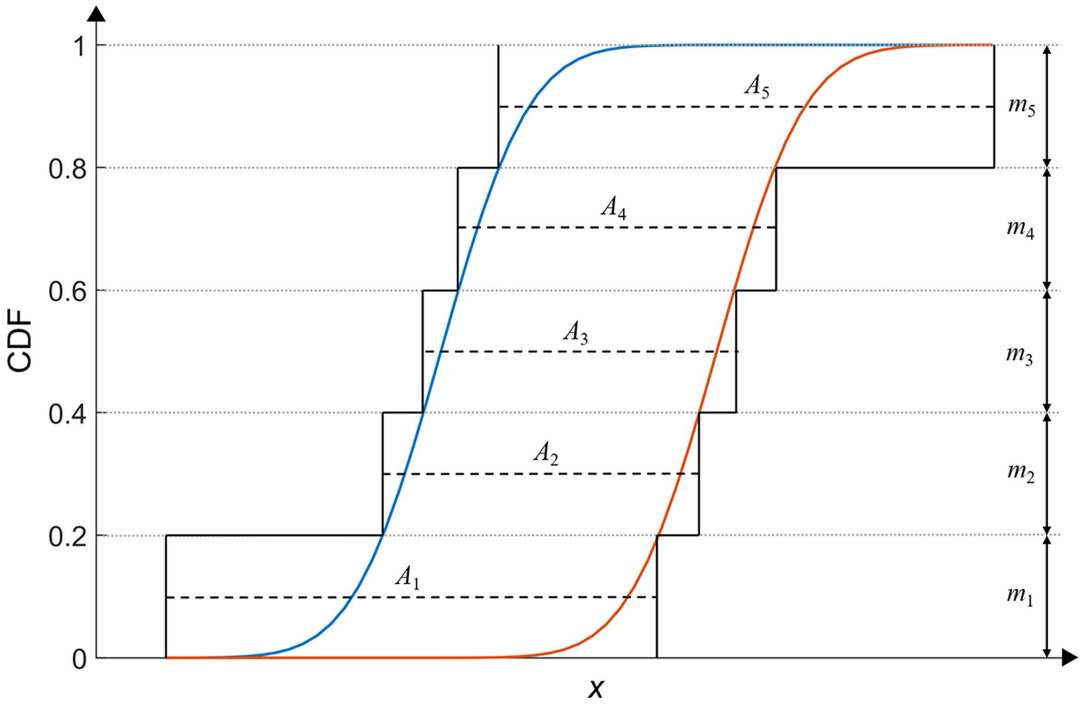


Figure 2.3: Discretisation of a p-box

2.4.3 Uncertainty in Resilience Quantification

The current state of the art is limited for uncertainty analysis applied to resilience quantification techniques. However, recent work has been carried out on imprecise resilience analysis.

Al Khaleen et al. [91] contributed to uncertainty by adding two-stage risk and optimisation models accounting for uncertainty with respects to travel times and repair times based on a given scenario based optimisation technique. Wu and Wang [92] developed a post-disruption management framework to improve resilience under uncertainty. The method uses MILP and co-ordinates recovery agents applying stochastic optimisation techniques to challenge the uncertainties in the restoration phase. The research evaluated that risk adverse optimisation produces more reliable random outcomes and applied the technique to a power system study. Cicilio et al. [93] presented an electrical grid framework for resilience and applied variance in the system through the input values of loads and generation. The uncertainty applied used Monte Carlo simulation to

sample the load uncertainties. Filippi et al. [94] developed a novel approach to quantify resilience in aerospace systems with the addition of epistemic uncertainties. The authors applied Dempster-Shafer theory to estimate belief and plausibility curves in polynomial time. A novelty in this research was the development of a min-max algorithm for worst case scenario optimisation and its development was applied to the global system reliability model with respects to epistemic uncertainty. Li et al. [95] contributed to resilience with network transportation recovery by developing a network recovery strategy during an emergency recovery phase under uncertainty for both deterministic and stochastic scenarios. The proposed resilience metrics used are rapidity of the network and the cumulative loss of the network performance. A resilience based bi-level programming model developed for both scenarios was postulated. The upper level determines the road segments required for restoration and repair times to maximise the resilience of the system, and the lower level formulates the user response to the upper decision as a user equilibrium. This is used for the input of the algorithm to integrate a genetic algorithm for the parallel machine scheduling program.

2.5 Chapter Summary

This chapter provides a summary of the current work carried out in academic literature for the current resilience analysis techniques. This reading provides a strong foundation to exploring and innovating on this thesis to obtain the optimal results and provides the reader with an insight to resilience quantification for understanding this thesis. Resilience has been defined in a multi-dimensional approach and has been separated into its respective operational and infrastructural states. This chapter has also presented the current state of the art of the literature with regards to uncertainty, highlighting both epistemic and aleatory uncertainty which has been applied to the various chapters in this thesis. It is important to note that the current research gaps in the literature are within the limitations of uncertainty as applied to resilience and the purpose of this thesis is to provide an innovation to the current state of the art of resilience as applied with respects to uncertainty and is presented in Chapters 3, 4 and 5.

Chapter 3

New Methods to Quantify Uncertainty in Binary-state Systems

3.1 Introduction

A system can be thought of as a mechanism consisting of a combination of nodes and links which contain properties that are interconnected. The goal of a system is to reach the terminal node in a mission which starts from the source node. The likelihood of this goal being attained is defined by the reliability of the system. A more general definition of reliability is the probability of success of a system has at given period of time and the knowledge of this when applied to systems enables maintenance planning to be carried out with risk-based optimal intentions [96].

A binary-state system is defined as a type of system which contains nodes characterised by Boolean algebra, with 1 representing the node's success state and 0 representing the node's failed state respectively. Binary-state systems are often used to represent systems with discrete characteristics. Such systems have been implemented into fault trees and Boolean logic gates. An example of a binary-state system includes a failure in a gate system such as the status of a valve in a water distribution network, with "1" representing a valve's open state for flow, and 0 representing the closed state of the valve. Binary-state systems are often applied to larger systems, which can contain many nodes of different types. Such systems include subsea production networks [97]

which contain a large number of nodes ranging from wells, flow lines, manifolds, x-trees, risers and pipelines.

A binary-state system can be described as coherent if removing a component from the system does not make the system worse, in other words that every component is relevant to the system's structural function [98].

Contrary to binary-state systems exist multi-state systems which can contain several different outputs defined by their failure times, operating conditions, age or another desired parameter as defined by the user [99].

This chapter focuses on binary-state systems solely, as the objective is to produce an estimate of reliability in the most computationally efficient manner.

3.1.1 Overview of the Proposed Approach

In this work, two different methods to quantify an estimate of transient reliability are tested. These two methods are the survival signature and the analytical probability propagation method (PrPm).

These two approaches are applied with the assumption of applied Boolean algebra to binary-state networks, where the individual nodes of the network can either be in a functioning or a failed state.

This chapter focuses on a case study from China's high speed train network [100]. The network is composed of 26 nodes which contain all the major cities of China and 48 links which represent the respective train lines connecting the various cities.

The novel theme of this chapter is the addition of imprecise probability to both the survival signature and PrPm. The two types of uncertainty being combined in both methods are epistemic and aleatory uncertainty, both which have been applied simultaneously. The survival signature has been applied with uncertainty based on non-parametric predictive inference and estimation is applied as applied to the survival function. The approach to epistemic uncertainty regarding the PrPm involves applying differing input probabilities into the desired nodes and links whilst for aleatory uncertainty the addition of approximation algorithms have been implemented.

3.2 Background and Theory

The idea behind reliability estimation is to minimise computational time as opposed to traditional quantification techniques which are known to be computationally expensive. The proposed approaches for this work are the survival signature technique and the PrPm. The general theory behind the two methods are outlined in this section.

3.2.1 The Survival Signature

Coolen and Coolen-Maturi developed the survival signature to extend the general system signature by separating components into different types with non-exchangeable failure times.

3.2.1.1 Theory

The following equations are taken from Coolen and Coolen-Maturi [36]. It can be assumed that a system is composed of m components, and its state vector is defined as $\underline{x} = (x_1, x_2, ..., x_m) \in \{0, 1\}^m$ with $x_i = 1$ denoting a functioning component and $x_i = 0$ denoting a failed component. The structural function $\phi\colon \{0, 1\}^m \rightarrow \{0, 1\}$ defines all possible $\underline{x}$ values as 1 if the system functions and 0 if the system fails, therefore $\phi(\underline{0}) = 0$ and $\phi(\underline{1}) = 1$. If T_s, the random failure time of the system is greater than 0 and $T_{j:\,m}$, the j-th order statistic of randomness of m independent and identically distributed (iid) component failure times for $j = 1, ..., m$, with $T_{1:\,m} \leq T_{2:\,m} \leq ...T_{m:\,m}$, the general system signature can be denoted as;

$$q_j = P(T_s = T_{j:\,m}) \tag{3.1}$$

with q_j representing the probability of system failure at the time of j-th component failure. The general system signature describes the system in a quantitative manner which can be applied for binary-state reliability calculations. Furthermore, the survival function of the system's failure can be defined as;

$$P(T_s > t) = \sum_{j=1}^{m} q_j P(T_{j:\,m} > t) \tag{3.2}$$

when Equation 3.2 is applied to a CDF, the survival function is defined as;

$$P(T_{j:\,m} > t) = \sum_{l=m-j+1}^{m} \binom{m}{l} [1 - F(t)]^l [F(t)]^{m-l} \tag{3.3}$$

The general system signature is sufficient for reliability analysis when applied to systems with m components that have iid failure times with a single type of component. In extension to the general signature, the proposed approach for the survival signature further expands the framework implemented from the general system signature and innovates this to be applied to systems with multiple types of components [36].
We denote $\phi(l)$ for $0 = 1, ..., m$ as the probability of the system functioning given that l components are functioning. The assumption for the general system signature of $\phi(0) = 0$ and $\phi(m) = 1$ still exists and we assume $\binom{m}{l}$ state vectors $\underline{x}$ with l components $x_i = 1$ and therefore, $\sum_{x=1}^{m} x_i = l$. All states of S_l are equally likely to occur due to iid for all m components, and therefore;

$$\phi(l) = \binom{m}{l}^{-1} \sum_{\underline{x} \in S_l} \phi(\underline{x}) \tag{3.4}$$

If the number of components in the system is set as a function of time $t > 0$, $C_t \in \{0, 1, ..., m\}$, the CDF of the failure time can be defined as;

$$P(C_t = l) = \binom{m}{l} [F(t)]^{m-l} [1 - F(t)]^l \tag{3.5}$$

and therefore;

$$P(T_S > t) = \sum_{l=0}^{m} \phi(l) P(C_t = l) \tag{3.6}$$

We can extend these definitions to systems containing multiple types of components with non-exchangable failure times. Consider a system with $K \geq 2$ types of m_k components, and therefore $\phi(l_1, 1_2, ..., l_K)$ with $l_k = 0, 1, ..., m_k$ for $k = 1, 2, ..., K$, the probability of the system functioning given l_k of m_k components of type k work for each $k \in \{1, 2, ..., K\}$. This means there are $\binom{m_k}{l_k}$ state vectors $\underline{x}^k$ with l_k components $x_i^k = 1$. Therefore $\sum_{i=1}^{m_k} x_i^k = l_k (k = 1, 2, ..., K)$ and $S_{l_1, l_2, ..., l_K}$ are the set of state vectors for the system.. With the assumption that the component failure times for different component types are independent, and for the same component types are exchangeable, the survival signature from Equation 3.4 can be extended to;

$$\Phi(l_1, ..., l_K) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1} \right] \sum_{\underline{x} \in S_{l_1, l_2, ..., l_K}} \phi(\underline{x}) \tag{3.7}$$

In this case, the survival function from Equation 3.2 can be extended when K types of components are applied;

$$P(T_s > t) = \sum_{l_1=0}^{m_1} ... \sum_{l_K=0}^{m_K} \Phi(l_1, ..., l_K)\, P\Big(\cap_{k=1}^{K} \{C_t(t) = l_k\} \Big) \tag{3.8}$$

where $C_t(t) \in \{0, 1, ..., m_k\}$ represents the number of components, k working at time t.

When this is applied to a CDF, the survival function from Equation 3.3 translates to;

$$P\Big(\cap_{k=1}^{K} \{C_t(t) = l_k\} \Big) = \prod_{k=1}^{K} P(C_t(t) = l_k) = \prod_{k=1}^{K} \binom{m_k}{l_k} [F_k(t)]^{m_k - l_k} [1 - F_k(t)]^{l_k} \tag{3.9}$$

The survival signature has been derived from the general system signature and its respective survival function assuming that component failures can be applied in the form of a CDF. In comparison, the survival signature is advantageous over the general system signature as it contains all the properties of the system signature applied to multiple types of components and has the flexibility to be applied to either coherent or non-coherent systems.

3.2.1.2 The Bridge System

A six node system containing two different types of components is presented by Coolen and Coolen-Maturi [36]. Figure 3.1 reconstructs this system in DAG format conformed from the connnectivities in an adjacency matrix. Type 1 components are represented in red and type 2 components are represented in blue respectively. The system contains a source node, 1 and a terminal node, 6 with the specific mission of the system to travel from the source to terminal nodes in the quickest and most efficient manner.

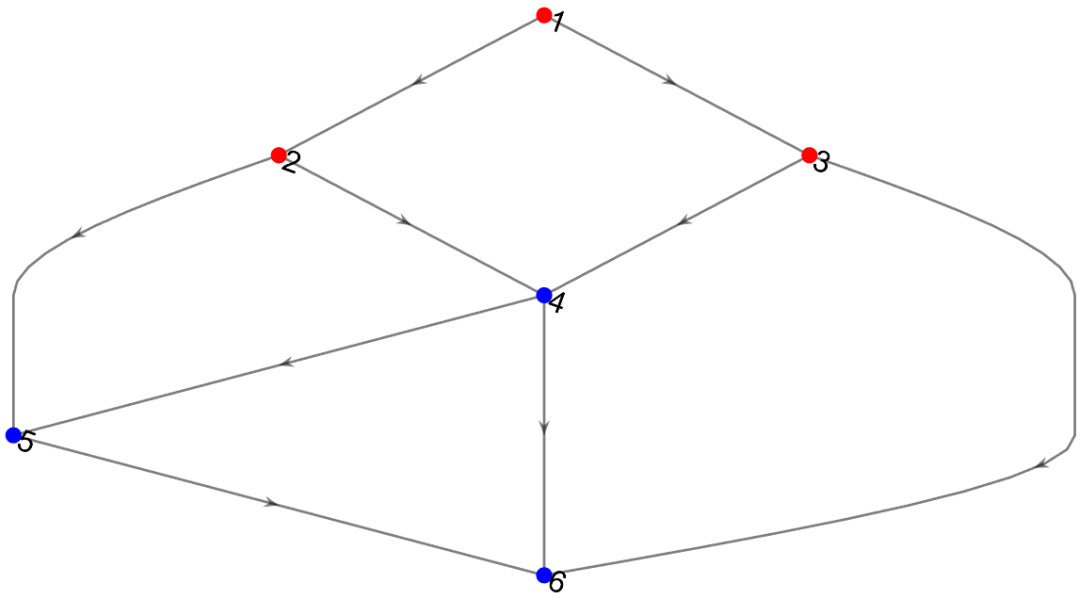


Figure 3.1: The bridge system DAG

In a system with A set of components C, Samaniego [101] defines a cut set as a system that fails when the components selected in set C fail, and conversely a path set is defined when a set of components P is applied and all components in the set P function. A minimal path set contains no subsets of Pthat are also path sets, and conversely a minimal cut set is defined as a cut set which contains no subsets of C that are also cut sets. It is also important to note the critical components in a system which is a single component that is required to function for the whole system to function. Specifically, if this component fails, the system fails regardless of functioning status of any other components the system contains. In this bridge system shown on Figure 3.1, the critical components are component 1 and component 6 as either of of these components failing will prevent any opportunity for a path set to be established.

Table 3.1 displays the survival signature values for the bridge system where l_1 represents components in red and l_2 res presents components in blue.

l_1	l_2	$\phi(l_1 l_2)$		l_1	l_2	$\phi(l_1 l_2)$
0	0	0		2	0	0
0	1	0		2	1	0
0	2	0		2	2	$\frac{4}{9}$
0	3	0		2	3	$\frac{6}{9}$
1	0	0		3	0	1
1	1	0		3	1	1
1	2	$\frac{1}{9}$		3	2	1
1	3	$\frac{3}{9}$		3	3	1

Table 3.1: Survival signature values for the bridge system

3.2.2 Analytical Probability Propagation Method

Tong and Tien developed a new method for estimating reliability for networks represented by DAGs. The approach comes in the form of Bayesian statistics and provides a more computationally efficient alternative to the traditional Bayesian Network. This approach is used for the calculation of reliability in networks when both nodes and links contain attributes which affect the reliability of the whole system, a fundamental difference to the survival signature which does not associate reliability values to the individual components and links. The method uses belief propagation to apply inference in a DAG, a message passing algorithm which produces an exact solution for propagation steps. The propagation message is sent to other neighbouring nodes in sequence. The message is then calculated using the marginal distribution of each unobserved node conditioned from the highlighted node. This message is then received and carried on to its respective neighbours. The joint distribution can be derived from the Hammersley-Clifford theorem, $p(X) = \frac{1}{Z}\prod_{c\in\xi}\Psi_{x_c}$, where Z represents the normalisation constant, ξ represents the set of maximal cliques and Ψ represents the desired function. Computational time increases exponentially as the number nodes in the given network increases. The two node joint probability distribution has been calculated and the message is passed from the source node to its direct neighbours. The message passed to the terminal node provides an estimate for the reliability of the network. The examples in this paper use a propagation sequence based on message updating from

one direct neighbour, and not multiple neighbours. All equations for this method have been obtained from Tong and Tien [44].

3.2.2.1 Bayesian Belief Updating

In order to calculate probabilistic inference for the network, a queried node, M is chosen and the calculation of the posterior probability of this queried node is computed. During the posterior probability calculation, the known variable is defined as the evidence variable, E, and the evidence probability denoted as e.The equation applied to compute the posterior probability conditioned to evidence, $P(M|E = e)$ is Bayes' rule derived from Bayes' theorem.
The joint probability $P(M, E = e)$ is calculated as follows [102];

$$P(M|E = e) = \frac{P(E = e|M)P(M)}{P(E = e)} \tag{3.10}$$

$P(M)$ represents the prior probability and $P(E = e|M)$ represents the likelihood probability.

During the computation of the queried node, M, the summation for all the combinatorics of the local joint probability distribution is known as the marginalisation process. This process becomes exponentially computationally expensive as the number of nodes required for the process increases, and therefore, the PrPm only obtains inference from one direct neighbour during each propagation step to minimise computational expense.

3.2.2.2 A 25 Node System

A 5x5 system consisting of 25 nodes with all nodes having equidistant links is proposed. The system contains one chosen source node and one chosen terminal node. The various nodes and links have been implemented in a DAG are displayed in Figure 3.2. The source node is highlighted as S and the terminal node is highlighted as T. These nodes can be ascribed to any node in the network as desired. The remaining nodes are all numbered from 1 to 25. It is also important to note that this method can

be used for both DAGs and non-DAGs as the rules are not defined by the directivity of the links.

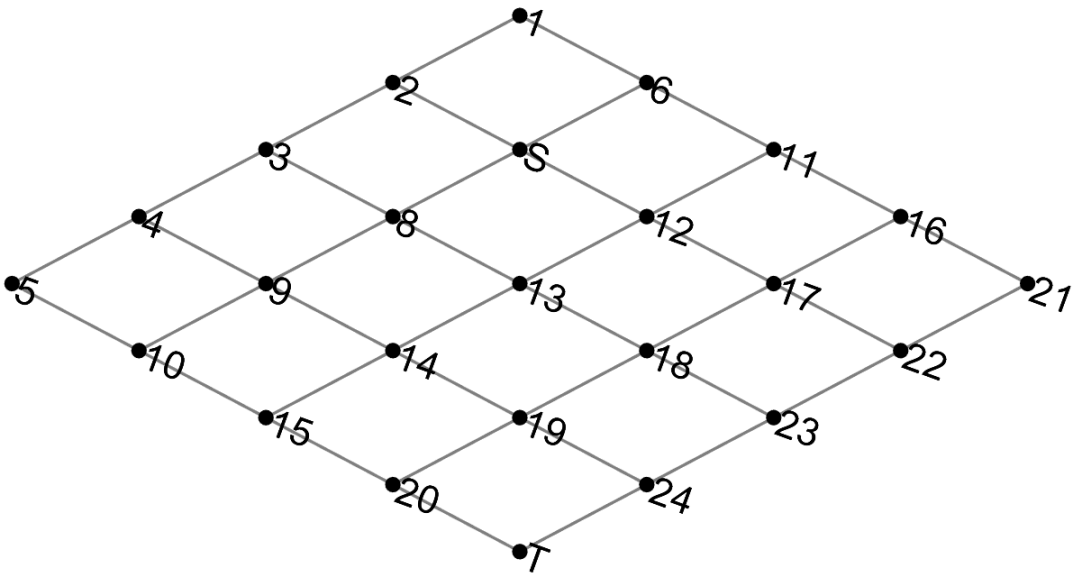


Figure 3.2: A 25 node network

A non-propagated node is defined as a node that has not yet received a propagation message which is represented by a red node in Figure 3.3. A propagated node which carries on a further propagation sequence is represented as a yellow node and a propagated node which carries no more propagation steps is represented as a green node. The sequences of nodes to be propagated for this method are based on three rules;

1. A node can only be propagated if it is a direct neighbour of a propagated node.
2. During the propagation of a selected node, this node must not separate any two non-propagated nodes to ensure that every node in the network is considered.
3. A newly defined propagated node should not connect with another propagated node to ensure that every link is considered.

It is crucial that all three rules are obeyed when propagating a sequence to ensure validity and credibility with this method.

In order to numerically represent these three updating rules as applied to a general system, Table 3.2 has been constructed to represent the rules as applied to the 5x5 example system;

A	C	Pr	A	C	N	Updates
0	0	P_1	0	0	0	$P_1(1-R_1R)$
			0	0	1	P_1R_1R
0	1	P_2	0	1	0	$P_2(1-R_1R)$
			0	1	1	P_2R_1R
1	0	P_3	1	0	0	$P_3(1-R_1R)$
			1	0	1	P_3R_1R
1	1	P_4	1	1	0	$P_4(1-R_1R)$
			1	1	1	P_4R_1R

Table 3.2: PrPm Updating rules for one direct neighbour

Using these updating rules, the steps for the network in Figure 3.2 are displayed in Figure 3.3. The first step of propagation excludes node 6 because of rule 2, as two non-propagated nodes cannot be separated during this step. Therefore only nodes 2, 8 and 12 are propagated. Node 6 is not propagated in the following step as a newly propagated node cannot be connected to another newly propagated node as explained in rule 3.This sequence and the respective rules of the sequence are applied to each step of the network's propagation until the terminal node is propagated and the terminal reliability value is deduced.

In order to provide a clear demonstration of the initial configuration, the first two propagation step and the final propagation step, Figure 3.3 has been composed;

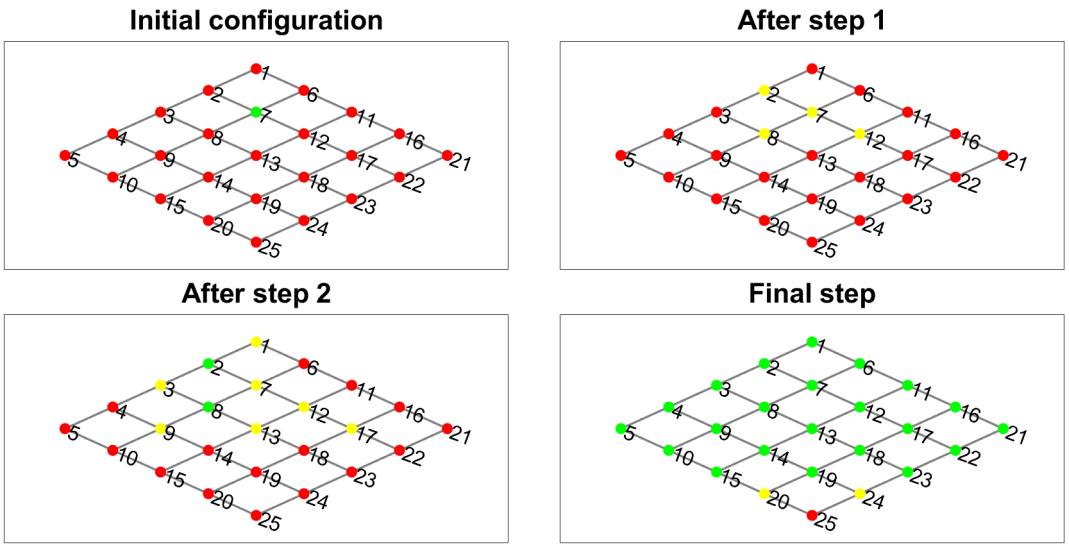


Figure 3.3: PrPm Updating steps

The rules of updating and the message passing on Figure 3.3 can be applied to any proposed DAG or non-DAG as the user desires. The node receiving the message is defined as N and the node passing the message is defined as A. A boundary node which is not a direct neighbour is defined as C and R_1 represents the reliability of the link, $A - N$.

N receives a message from its direct neighbour A, which is currently the joint distribution of A and C from the previous propagation step. During the first step, the prior probability distribution is the probability of the source node. The probability distribution during the next step is composed of the joint distribution between A, C and N which is derived from the updating rules in Table 3.2. In each step, a new joint distribution is derived from all the propagation data obtained from the prior steps of the calculation.

Each new step only uses the joint probability distribution between two nodes rather than the whole network which has been designed to reduce computational cost from an exponential cost of $\mathcal{O}(2^n)$ to a quartic cost $\mathcal{O}(n^4)$.

3.2.2.3 Message Passing

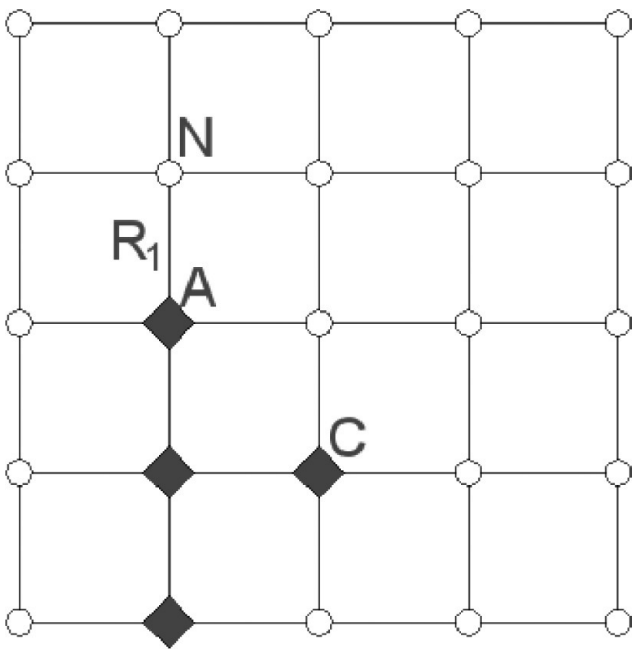


Figure 3.4: PrPm message passing

The PrPm applies joint probability distributions to pass probabilistic information from one node in the sequence to the next node. The assumption is that a single node receives a message from one direct neighbour, and that the network is binary so that "1" represents a node's success and "0" represents a node's failure. Figure 3.4 displays the process of message passing for the PrPm with nodes N and A as described in Section 3.2.2.2 and an additional node C, representing the node which is not propagated during this step. The message directed at node N is composed of the joint distribution between nodes A and C as acquired from the previous step of the propagation sequence. During the initial step, the prior probability for the joint distribution is equal to the source node. R_1 represents the reliability of the link $A-N$, therefore the reliability of node N is determined by the reliability of node A and the reliability of link R_1. This means that the new joint reliability distribution is defined as $P(A,C,N)$ and is obtained using the updating rules displayed on Table 3.1. The probability obtained from this distribution is the probability that the node is in a successful state and is listed as an exact solution without any approximations. R denotes the reliability of the previously propagated node to be applied in the next propagation step. After the joint probability distribution of $P(A,C,N)$ is obtained, the new two intermediate joint distributions $P(A,N)$ and $P(C,N)$ can also be extracted. This process is repeated constantly until the terminal node's probability is propagated.

3.3 A Framework for Imprecise Probability

Imprecise probability is key when deducing the possible confidence bounds and uncertainty targets which is a prone phenomena in real world models. Uncertainty is classified in two categories as mentioned in Section 2.4.1; epistemic and aleatory uncertainty [103][104]. Epistemic uncertainty is caused by the lack of sufficient information for input data in models, with techniques ranging from expert judgement data and estimation of uncertainties from prior models to tackle this constraint [94]. Aleatory uncertainty is also known as stochastic randomness and is caused by inherent randomness of behaviour in a model's physical system or environment [105]. Both of these types of uncertainty have been implemented for both the survival signature and PrPm, innovating on these to quantify uncertainty in both an epistemic and aleatory angle.

3.3.1 Survival Signature Uncertainty

The survival signature method has been tested in academic literature with respects to various forms and various methods to apply uncertainty. All the methods applied are listed in this subsection which aims to produce uncertainty in the form of non-parametric predictive inference, epistemic uncertainty applied to the survival function and aleatory uncertainty within the final output function. The components are assumed to be non-repairable in order to simplify the computational cost.

3.3.1.1 Non-Parametric Predictive Inference

An approach to quantify uncertainty for the survival function was first implemented by applying the non-parametric predictive inference (NPI) method [38], a statistical technique developed by Coolen which applies conditional probabilities to obtain multiple possibilities of future random observable quantities as conditioned on current data. This form of epistemic uncertainty is applied in the case with very little knowledge of uncertainties that can arise on the system.

Upper and lower probability bounds are applied in the form of imprecise probabilities as explained in Section 2.4.2 and applies consistency properties with Bayesian attributes. The outcome is to provide a solution to the goals set by Bayesian inference which cannot be obtained through the precise probabilistic approach. This approach has been used in the field of reliability for systems with restrictions on their structures [106].

A binary state system containing nodes of multiple types is proposed and its failure time is represented as T_s. The system applies the same rules of the general survival signature and combines this with NPI for Bernoulli data. This application enables the NPI method to be applied to all systems and carries out the same assumptions of iid for the general survival signature. NPI is applied to a specified type of component in the system using data and this data is derived from components that have exchangeable failure times to the selected component. For $k \in \{1, ..., K\}$, n_k represents the number of components of type k given that data specifying failure times are available. $C_k(t)$ represents the number of type k components functioning at time t and $s_k(t)$ represents the number of type k components still functioning at time t.

The various steps to apply the NPI method are displayed below;

The survival function from Equation 3.8 is applied to obtain the lower bound of the survival function;

$$P(T_s > t) \geq \sum_{l_1=0}^{m_1} ... \sum_{l_K=0}^{m_K} \phi(l_1, ..., l_K) \cdot \Big(\prod_{k=1}^{K} \overline{D}(C_k(t) = l_k) \Big) \tag{3.11}$$

$$\begin{aligned} &\overline{D}(C_k(t) = l_k) = \overline{P}(C_k(t) \leq l_k) - \overline{P}(C_k(t) \leq l_k - 1) \\ &= \binom{n_k + m_k}{n_k}^{-1} \binom{s_k(t) - 1 + l_k}{s_k(t) - 1} \times \binom{n_k - s_k(t) + m_k - l_k}{n_k - s_k(t)} \end{aligned} \tag{3.12}$$

$\overline{P}$ represents the upper probability interval for Bernoulli data, $\overline{D}$ represents the maximum possible probabilistic value with the application of this Bernoulli data and is applied at the time $C_k(t) = 0$ and therefore $\overline{D}(C_k(t) = 0) = \overline{P}(C_k(t) = 0)$. After this step, $\overline{D}(C_k(t) = 1)$ is deduced by applying the maximum possible probabilistic quantity from the total probabilistic quantity in surplus from the event $C_k(t) = 1$ and is performed by applying $\overline{D}(C_k(t) = 1) = \overline{P}(C_k(t) \leq 1) - \overline{P}(C_k(t) = 0)$. The maximum possible probabilistic quantify is now applied to an increasing l_k with $\overline{D}(C_k(t) = l_k)$. The right side of the inequality equation is therefore the lower probability interval and represents the maximum possible probabilistic quantity for the lower bound and is therefore the NPI lower bound for $T_s \geq t$ providing the lower bound survival function is at $t > 0$;

$$\underline{S}_{T_S} = \underline{P}(T_s > t) = \sum_{l_1=0}^{m_1} ... \sum_{l_K=0}^{m_K} \phi(l_1, ..., l_K) \cdot \Big(\prod_{k=1}^{K} \overline{D}(C_k(t) = l_k) \Big) \tag{3.13}$$

The NPI upper bound is then applied to define the upper bound survival function;

$$P(T_s > t) \leq \sum_{l_1=0}^{m_1} ... \sum_{l_K=0}^{m_K} \phi(l_1, ..., l_K) \cdot \Big(\prod_{k=1}^{K} \underline{D}(C_k(t) = l_k) \Big) \tag{3.14}$$

with;

$$\begin{aligned}\underline{D}(C_k(t) = l_k) &= \underline{P}(C_k(t) \leq l_k) - \underline{P}(C_k(t) \leq l_k - 1) \\ &= \binom{n_k + m_k}{n_k}^{-1} \binom{s_k(t) + l_k}{s_k(t)} \times \binom{n_k - s_k(t) + m_k - l_k - 1}{n_k - s_k(t)}\end{aligned} \tag{3.15}$$

where $\underline{P}$ represents the NPI lower bound probability for Bernouli data. This implies that a minimal probabilistic quantity is applied to $C_k(t)$, yielding the NPI upper survival function for $t > 0$;

$$\overline{S}_{T_S} = \overline{P}(T_s > t) = \sum_{l_1=0}^{m_1} ... \sum_{l_K=0}^{m_K} \phi(l_1, ..., l_K) \cdot \Big(\prod_{k=1}^{K} \underline{D}(C_k(t) = l_k) \Big) \tag{3.16}$$

3.3.1.2 Estimation of Uncertainty

The survival function obtained from Equation 3.9 is implemented and the means of imprecision are the epistemic uncertainty as applied to the Weibull parameters.

The epistemic uncertainty based on Weibull parameters in the general Weibull equation is denoted below;

$$f(t) = \frac{\beta}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{\beta - 1} e^{-\left(\frac{t-\gamma}{\eta} \right)^{\beta}} \tag{3.17}$$

The shape parameter, β and scale parameter, η are subject to epistemic uncertainty via estimation, and the location parameter, γ is set to 0. A lower bound and upper bound has been proposed for both the shape parameter $\overline{\underline{\beta}}$ and scale parameter $\overline{\underline{\eta}}$. Applying these parameters of uncertainty to Equation 3.17 yields;

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\overline{\underline{\beta}}}{\overline{\underline{\eta}}} \left(\frac{t - \gamma}{\overline{\underline{\eta}}} \right)^{\overline{\underline{\beta}} - 1} \tag{3.18}$$

3.3.1.3 Aleatory Uncertainty

This paper takes an additional approach to uncertainty by adding aleatory uncertainty in the form of a Gaussian;

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}} \left(\frac{x-\mu}{\sigma} \right) \tag{3.19}$$

The parameters of uncertainty for the Gaussian are the mean survival limit, μ and the standard deviation, σ is computed to a 95% confidence interval as this is deemed to be a suitable confidence bound with respects to the number of samples being applied for the Gaussian.

Both forms of transient uncertainty have been applied simultaneously to the survival function plotted using the survival signature attributes, with the combined equation for uncertainty yielding;

$$h(x) = \left[\frac{\overline{\beta}}{\overline{\underline{\eta}}} \left(\frac{t-\gamma}{\overline{\underline{\eta}}} \right)^{\overline{\underline{\beta}}-1} \right] + \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}} \left(\frac{x-\mu}{\sigma} \right) \right] \tag{3.20}$$

3.3.2 PrPm Uncertainty

Uncertainty applied to the the PrPm in the forms of both epistemic and aleatory uncertainty are used simultaneously. Epistemic uncertainty is applied through estimation of the input probabilities, source node reliability, link reliability and node R reliability. The input interval probabilities have been defined by the user in upper and lower bounds, and these bounds define the output probability of the terminal node. The aleatory uncertainty for the PrPm is implemented in the form of approximation algorithms, which include the inner approximation algorithm and outer approximation algorithm as applied for the credal network by Estrada-Lugo et al. [107]. These two forms of uncertainty have been combined to attain the upper and lower confidence bounds deduced from each step of the propagation sequence. The purpose of adding uncertainty for the PrPm is to attain the imprecise probabilistic value for the terminal node of the propagation sequence.

This subsection aims to present the various forms of uncertainty as applied to the PrPm and all the relevant equations required for deriving these are presented.

Consider the three node binary DAG as presented in Figure 3.5 where C represents the child node of parent nodes A and B;

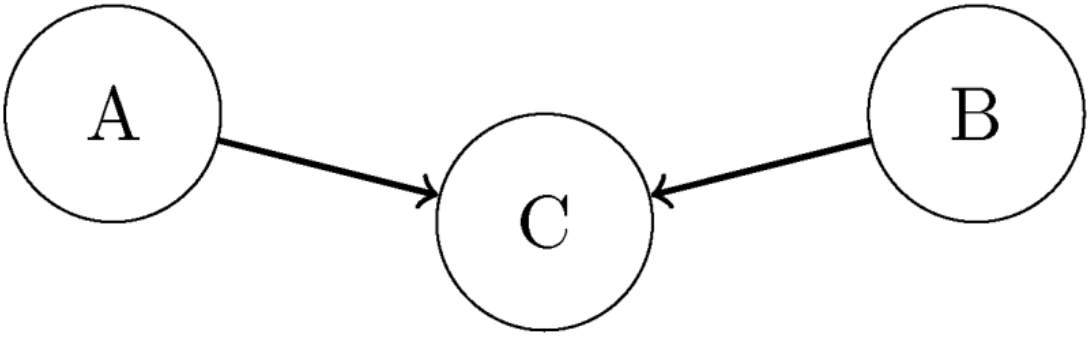


Figure 3.5: Three node DAG

The interval probabilities can be represented as $P(\overline{a_i}) = [P(\underline{a_i}), P(\overline{a_i})]$ with $P(\underline{a_i})$ representing the lower bound and $P(\overline{a_i})$ representing the upper bound respectively.

A general joint probability distribution is defined as;

$$P(\mathbf{X}) = \left\{ \bigcup_{n} \left[\bigcup_{m} \mathbf{P}(\mathbf{x_n^m}) \right] \right\} \tag{3.21}$$

In the case for the network in Figure 3.5, the lower and upper bounds can be inserted into Equation 3.21 to yield;

$$P(\underline{A}, \underline{B}, \underline{C}) = \begin{Bmatrix} [P(\underline{a_1}), (\underline{b_1}), (\underline{c_1})] & [P(\underline{a_1}), (\underline{b_2}), (\underline{c_1})] \\ [P(\underline{a_2}), (\underline{b_1}), (\underline{c_1})] & [P(\underline{a_2}), (\underline{b_2}), (\underline{c_1})] \\ [P(\underline{a_1}), (\underline{b_1}), (\underline{c_2})] & [P(\underline{a_1}), (\underline{b_2}), (\underline{c_2})] \\ [P(\underline{a_2}), (\underline{b_1}), (\underline{c_2})] & [P(\underline{a_2}), (\underline{b_2}), (\underline{c_2})] \end{Bmatrix} \tag{3.22}$$

$$P(\overline{A}, \overline{B}, \overline{C}) = \begin{Bmatrix} [P(\overline{a_1}), (\overline{b_1}), (\overline{c_1})] & [P(\overline{a_1}), (\overline{b_2}), (\overline{c_1})] \\ [P(\overline{a_2}), (\overline{b_1}), (\overline{c_1})] & [P(\overline{a_2}), (\overline{b_2}), (\overline{c_1})] \\ [P(\overline{a_1}), (\overline{b_1}), (\overline{c_2})] & [P(\overline{a_1}), (\overline{b_2}), (\overline{c_2})] \\ [P(\overline{a_2}), (\overline{b_1}), (\overline{c_2})] & [P(\overline{a_2}), (\overline{b_2}), (\overline{c_2})] \end{Bmatrix} \tag{3.23}$$

The authors outline the two approximation computational algorithms and the theory behind these as listed below.

3.3.2.1 Outer Approximation Algorithm

The outer approximation algorithm aims to provide an overestimation of the interval probabilities for the PrPm and marginalises the joint probability distribution. The interval probabilities as mentioned in Equation 3.22 and Equation 3.23 are applied to obtain the outer approximation algorithm as follows;

1. Identifying the lower and upper bounds of the initial network as applied in Equation 3.22 and Equation 3.23.
2. Computing inference by marginalising the joint probability distributions from this initial network configuration.
3. Identify the upper and lower bounds of the marginalised values.
4. Including the precise value of the network as queried into the computation of the joint probability value.

The lower bound probability in state 1, $P(\underline{a_1})$ can be applied to marginalise the joint probability distribution by marginalising B and C;

$$\begin{aligned} &P(\underline{a_1}) = \Sigma_{i=1,j=1}^{n}[P(\underline{a_1}), P(\underline{b_i}), P(\underline{c_j})] \\ &= P(\underline{a_1}), P(\underline{b_1}), P(\underline{c_1}) + P(\underline{a_1}), P(\underline{b_2}), P(\underline{c_1}) \\ &+P(\underline{a_1}), P(\underline{b_1}), P(\underline{c_2}) + P(\underline{a_1}), P(\underline{b_2}), P(\underline{c_2}) \end{aligned} \tag{3.24}$$

After normalising the probability values from the joint probability distribution, the lower bound value is computed from Equation 3.24. The same method is applied to compute the upper bound probability and the complement for the outer approximation regarding the queried variable.

3.3.2.2 Inner Approximation Algorithm

The inner approximation algorithm is used for approximating the minimum possible variance in the network, computing all the minimum possible permutations of uncertainty for every possible node in the network. This method is applied after the joint probability bounds have been obtained. If a node's marginal probability is required, one of the bounds of a single state must be fixed. For example, A can be fixed to $\overline{a}_1$ and the combinations of the bounds are fixed to this particular interval. The marginalisation combinations of $P(\overline{a}_1)$ from the joint probability of $P(\overline{a}_1, \underline{\overline{B}}, \underline{\overline{C}})$ are represented;

$$P(\overline{a}_1)_{comb1} = P(\overline{a}_1, \overline{b}_1, \overline{c}_1) + P(\overline{a}_1, \underline{b}_2, \overline{c}_1) + P(\overline{a}_1, \overline{b}_1, \underline{c}_2) + P(\overline{a}_1, \underline{b}_2, \underline{c}_2) \quad (3.25)$$

$$P(\overline{a}_1)_{comb2} = P(\overline{a}_1, \underline{b}_1, \overline{c}_1) + P(\overline{a}_1, \overline{b}_2, \overline{c}_1) + P(\overline{a}_1, \underline{b}_1, \underline{c}_2) + P(\overline{a}_1, \overline{b}_2, \underline{c}_2) \quad (3.26)$$

$$P(\overline{a}_1)_{comb3} = P(\overline{a}_1, \overline{b}_1, \underline{c}_1) + P(\overline{a}_1, \underline{b}_2, \underline{c}_1) + P(\overline{a}_1, \overline{b}_1, \overline{c}_2) + P(\overline{a}_1, \underline{b}_2, \overline{c}_2) \quad (3.27)$$

$$P(\overline{a}_1)_{comb4} = P(\overline{a}_1, \underline{b}_1, \underline{c}_1) + P(\overline{a}_1, \overline{b}_2, \underline{c}_1) + P(\overline{a}_1, \underline{b}_1, \overline{c}_2) + P(\overline{a}_1, \overline{b}_2, \overline{c}_2) \quad (3.28)$$

The $max(P(\overline{a}_1|b_2))$ and $max(P(\overline{a}_1|c_2))$ can be quantified with respects to bounds $P(\overline{A}, \overline{B}, \overline{C})$ and $P(\underline{A}, \underline{B}, \overline{C})$ from the outer approximation algorithm;

$$max(P(\overline{a}_1|B)) = max\left[\frac{\sum_C P(\overline{a}_1\overline{B}, \overline{C})}{\sum_{B,C} P(\overline{A}, \overline{B}, \overline{C})}\right] = max\left[\frac{P(\overline{a}_1, \overline{B})_{pseudoUP}}{P(\overline{B})_{PseudoUP}}\right] \quad (3.29)$$

$$max(P(\overline{a}_1|C)) = max\left[\frac{\sum_B P(\overline{a}_1\overline{B}, \overline{C})}{\sum_{A,B} P(\overline{A}, \overline{B}, \overline{C})}\right] = max\left[\frac{P(\overline{a}_1, \overline{C})_{pseudoUP}}{P(\overline{C})_{PseudoUP}}\right] \quad (3.30)$$

The full proof of the inner approximation algorithm is stated in Estrada-Lugo et al. [107].

3.4 Implementation of Techniques

The program used for both of the techniques is the OpenCossan software interface on MATLAB 2021b. This section highlights the steps taken in each method.

3.4.1 Survival Signature Implementation

The survival signature uses the application of adjacency matrix dimensions to model the topology of the system and uses mathematical probability calculations to compute the exact survival signature. The survival function is then computed by either the analytical solution or through Monte Carlo sampling.
The first step of simulation provides an approximation of the survival signature and applies percolation theory. This is divided into two stages;

1. Computing areas of the network that are negligible based on the critical percolation threshold with inputs set to 0.
2. Applying Monte Carlo simulation to provide an approximation of the surplus inputs to sample the entries.

The percolation process is carried out by deleting nodes from the DAG without repairing the edges of the graph. f denotes the fraction of deleted nodes of the system. In the event of a negligible value of f, this indicates a very high probability of the system containing a path set, and therefore a higher probability of survival. This is referred to as giant connected cluster [108]. As an increasing number of nodes are deleted from the system, the value of f increases until it reaches the point where the network contains a cut set, known as the critical fraction, f_c. The three exponents defining the interactions a network are detailed below with respects to f_c [109];

$$\begin{aligned} (s) &\sim |f - f_c|^{-\gamma} && mean\ cluster\ size \\ p_{in} &\sim (f - f_c)^{\beta} && probability\ of\ random\ node \\ \Xi &\sim |f - f_c|^{-v} && mean\ distance\ to\ cluster \end{aligned} \tag{3.31}$$

Th equations depict the phase transitions showing the topological interdependencies of the network. A path set is established until the condition, $f_c \cdot \sum_{k=1}^{K} m_k$ is revoked, and the network breaks down into various isolated parts without a functioning path set. The Molloy-Reed Criterion, which states that for a graph to contain a giant connected component, the majority of the nodes must be connected by two adjacent nodes [110]. Cohen developed a series of equations to express this [111];

$$\kappa = \frac{(d^2)}{(d)} > 2 \tag{3.32}$$

where d represents the node degree, (d) represents the first moment of degree distribution over the network and (d^2) represents the second moment of degree distribution over the network.

κ is calculated through graph representation of the adjacency matrix double loop. The ratio of the κ critical threshold obtained with the following equation;

$$f_c = 1 - \frac{1}{\kappa - 1} \tag{3.33}$$

If more than the number of components to achieve f_c has reached a failed state, the probability that the system has failed is close to zero as depicted in;

$$\sum_{k=1}^{K} = l_k < (1 - f_c) \cdot \sum_{k=1}^{K} m_k \Rightarrow \phi(\underline{l}) \approx 0 \tag{3.34}$$

The algorithm to approximate the survival signature, ϕ as available on Github [112];

Algorithm 1 Survival signature approximation algorithm

```
function APPROX(l, φ, N, C)
    c,n, ω, ϕ ← 0                                   ▷ Initialise variables
    while c > C  &  n ≤ N do
        n ← n + 1
        s ← rand state l̲
        if φ(s) = 1 then
            ω ← ω +1
        ϕ ← ω/n
        c ← √((ϕ−ϕ²)n⁻¹) / ϕ
    return ϕ, c                   ▷ Signature entry and coefficient of variation
```

Provided that the coefficient of variation target C has been achieved, generating a random network state for the vector l as dictated by s. Following this, the sample number n_l increases by one value.

If $\varphi(s) = 1$, increase the counter $\omega_{\underline{l}}$ by one value and the approximation is updated by;

$$\phi(\underline{l}) \approx \frac{\omega_{\underline{l}}}{n_{\underline{l}}} \tag{3.35}$$

and the current coefficient of variation is estimated with;

$$c_{\underline{l}} = \frac{\sqrt{\phi(\underline{l}) - \phi(\underline{l})^2\, n_{\underline{l}}^{-1}}}{\phi(\underline{l})} \tag{3.36}$$

The survival function is then simulated and the procedure is adopted through a Monte Carlo implemented simulation to estimate the reliability of the system. Random events are sampled and the outcome of the reliability is determined through the sequential mean of the Monte Carlo sampling results.

The pseudo-code for obtaining the survival function of a non-repairable system is listed below;

Algorithm 2 Survival function algorithm

Require: N, No. of simulations; dt, Discretisation time; F_k, CDF Failure times; $V_c = [m_1, m_2, ..., m_k]$, No. components per type; N_t, N0. discretisation steps

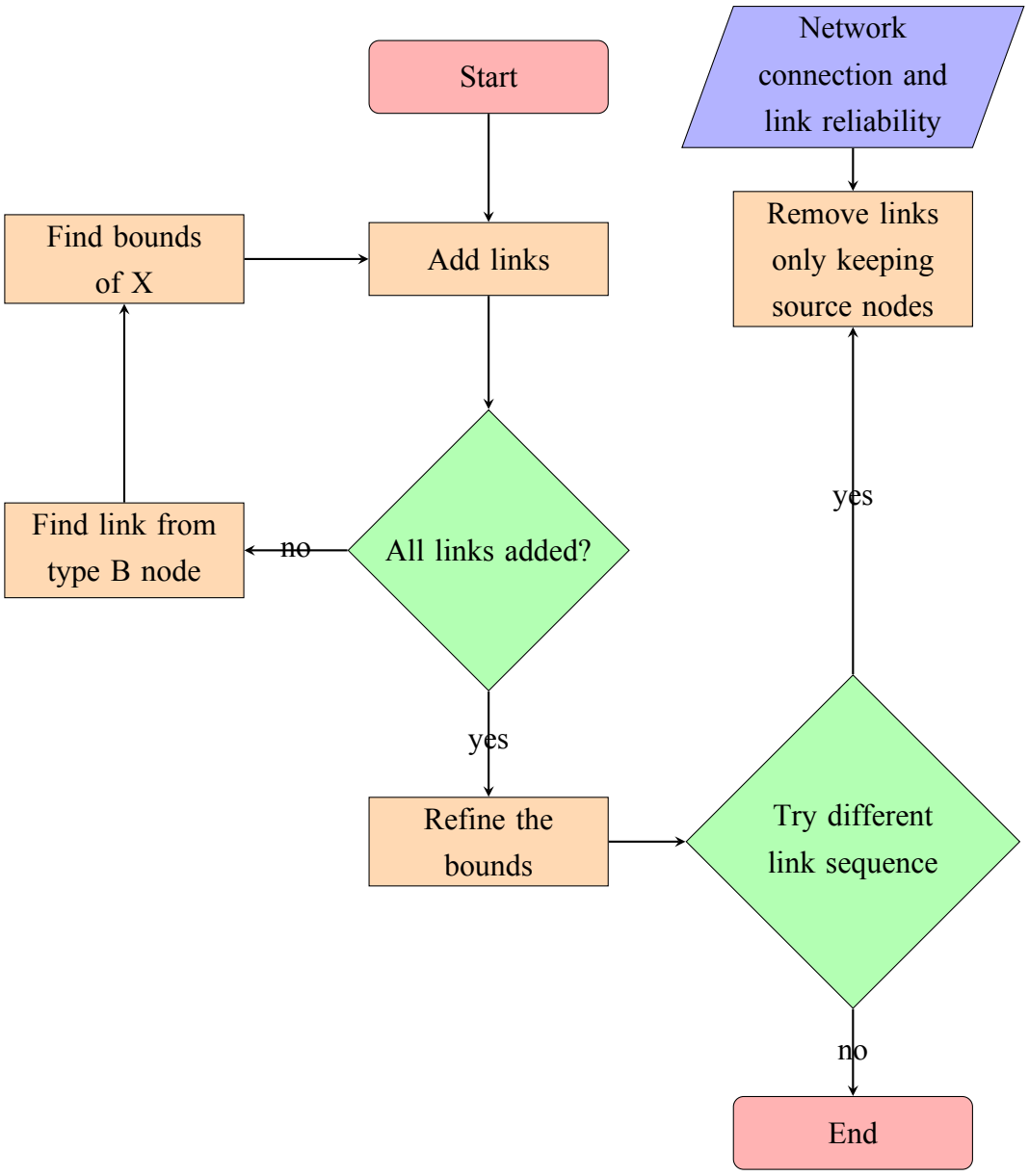
```
Set V_r(1 : N_t) = 0                                  ▷ Initialise counter
Set C=Sum(V_C)                                        ▷ Compute No. components
Set Φ = Survival Signature                            ▷ Compute survival signature
for n = 1 : N do                                      ▷ Loop over No. samples
    for k = 1 : K do                                  ▷ Loop over component type
        for j = 1 : m_k do                            ▷ Loop over No. components
            M f(j, k) ~ F_k                           ▷ Sample failure time component j of type k
    [V_t,V_i]= sort(M f)                              ▷ Reorder transition times (V_t)
                                                      ▷ Return component index vector (V_i)
    z = 1                                             ▷ Initialise index
    for m = 1 : C do                                  ▷ Loop over number of components
        V_c(V_i(m)) = V_c(V_i(m)) − 1                 ▷ Update No. working components
        while z · dt ≤ V_t(m) do
            V_r(z) = V_r(z) + Φ(V_c)                  ▷ Update counter
            z = z + 1                                 ▷ Update index
    V_r = V_r N^{−1}
```

The survival function algorithm is implemented with the desired Monte Carlo samples as chosen by the user. Increasing the number of Monte Carlo samples increases computational expense and using too few Monte Carlo samples produces a survival

function with a wide range of uncertainty. The user can experiment with the number of samples to deduce the optimum number of samples required.

3.4.2 PrPm Implementation

The PrPm explained in Section 3.2.2 applies various steps to obtain the desired output values. The reliability of the links are mimicked as nodes and all the nodes are listed on the topology with both their lower bound and upper bounds values.



The flowchart above displays steps and conditions required to obtain successful simulation of the PrPm.

3.4.2.1 Summary of the Proposed Method

The method applied is displayed on the above flowchart. The propagation sequence is first obtained based on the network's topology and the rules explained on Section 3.2.2.2 are applied. The sequence of the steps for the PrPm is obtained for the nature of message passing carried out for a given example of a network. The nodes are propagated in the network to ensure that each node received its required joint probability distribution. The messages are passed between the respective nodes and the updated joint probability distributions are derived for further message passing for the sequential nodes. The approximated PrPm terminal node probability is obtained during the final propagation step.
Tong and Tien provides a flowchart of the summary of the proposed method in Figure 3.6.

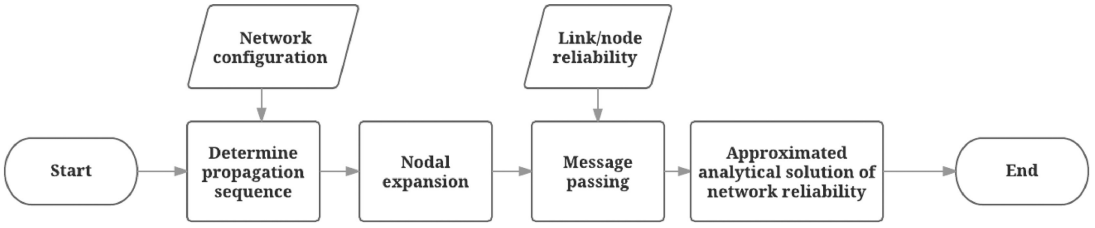


Figure 3.6: Basic PrPm work flowchart

3.5 Case Study: China's High-Speed Rail Network

Applications of both methods can be applied to both theoretical and real world systems. The case study chosen for this chapter is the China's high-speed rail network as presented by Hu et al. [100] in which the author reconstructed the current modern day topology of China's high-speed rail network and included the most important and high traffic stations as nodes updated as the 2016 version. The network is composed of 26 nodes and 48 links and is displayed in Figure 3.7. Links which do not connect a node to a neighbouring node with respects to the cities represented on the figure are

negated. The authors also innovated on the current literature by adding three separate categories for the links which are categorised by three different velocity bounds of the trains traveling within this track. This is displayed as the maximum velocity capacity which is displayed as

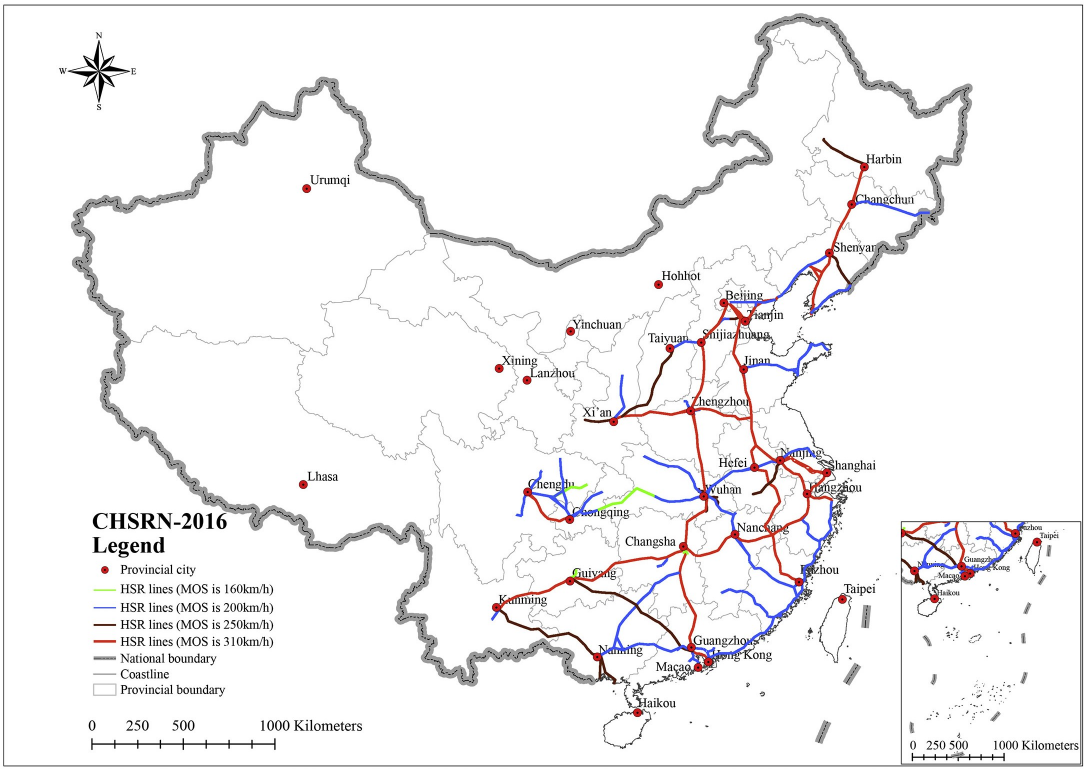


Figure 3.7: China's high-speed rail network

Table 3.3 displays the cities from Figure 3.7 that have been applied to this case study and presents these cities by their respective nodes 1-26. The table also discriminates the nodes between the three node types applied to the survival signature method. These three node types are discriminated based on the input and output train line maximum operating speed (MOS). Type 1 represents nodes with the slowest input MOS of 200 $km\,h^{-1}$, type 2 represents nodes with input an MOS of 250 $km\,h^{-1}$ and type 3 represents nodes with an input MOS exclusively of 310 $km\,h^{-1}$. The MOS lines at 160 $km\,h^{-1}$ are negated due to not contributing to any major city input and are scaled up during the rail input into a city at 200 $km\,h^{-1}$.

City	Node No.	Node type		City	Node No.	Node type
Harbin	1	3		Hefei	14	1
Changchun	2	3		Fuzhou	15	1
Shenyang	3	2		Nanjing	16	2
Beijing	4	1		Shanghai	17	3
Tianjin	5	3		Hangzhou	18	3
Shijiazhuang	6	1		Nanchang	19	1
Taiyuan	7	2		Changsha	20	3
Jinan	8	3		Guiyang	21	2
Xi'an	9	2		Kunming	22	2
Zhengzhou	10	3		Nanning	23	2
Wuhan	11	2		Guangzhou	24	1
Chongqing	12	1		Macao	25	1
Chengdu	13	3		Hong Kong	26	1

Table 3.3: Node numbers and types for China's high-speed rail network

3.5.1 Survival Signature Results

The survival signature algorithm, Algorithm 1 from section 3.4.1 has been tested for China's high-speed rail network three times and the computational simulation expense is noted in the figure below;

```
Eliminating trivial survival signature entries...
Approximating remaining survival signature entries...
Running: [====================]Done.  [72 seconds]
```

Figure 3.8: Survival signature computational time

3.5.1.1 DAG Construction

A DAG has been constructed from the information on Figure 3.7 and Table 3.3 and is shown in Figure 3.9. The types of nodes have been discriminated based on the node conditions as displayed on the table. Type 1 nodes are filled in blue, type 2 nodes are filled in black and type 3 nodes are filled in red. The arrows of the DAG display the direction of the train's flow in the network from the source node to the terminal node.

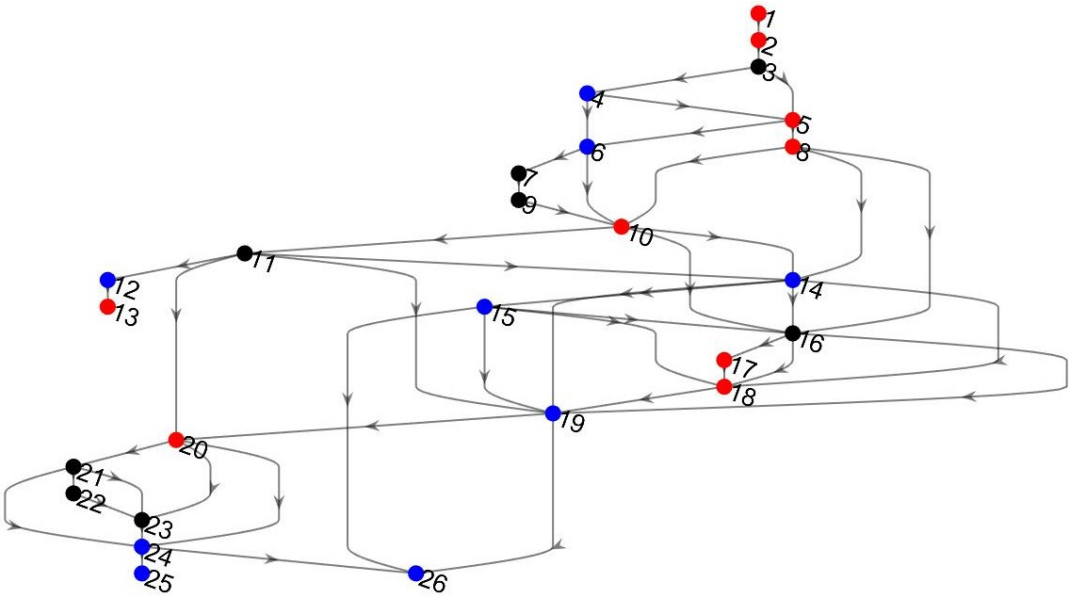


Figure 3.9: DAG of China's high-speed rail network displaying the node types

3.5.1.2 Four-Dimensional Survival Signature Results

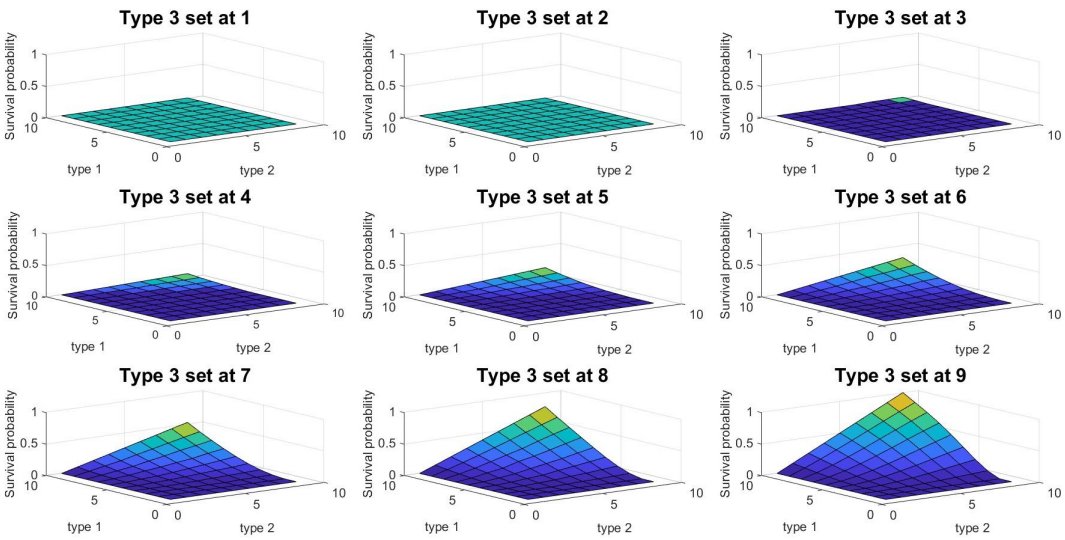


Figure 3.10: Survival signature 4D results

Since the chosen network contains three different node types, the simulation contains three input variables and one output variable. The individual surface plots in Figure 3.10 presents each combination of the states for type 3 nodes exclusively and presents the various survival probabilities (z-axis) subject to the number of nodes functioning for both type 1 nodes (x-axis) and type two nodes (y-axis).

The surface plots display the nature of the increasing likelihood of a network's path set forming as the combined number of nodes increase. In the event of only two type 3 nodes functioning, there is no probability of any path set regardless of the number of type 1 and type 2 nodes functioning. This changes when three components of type 3 are functioning and contains a small probability of a path set given that all the components of type 1 and type 2 are functioning. As the number of type 3 nodes are increasing on the individual surface plots, the observable significance of interest shows the likelihood of survival rapidly increasing.

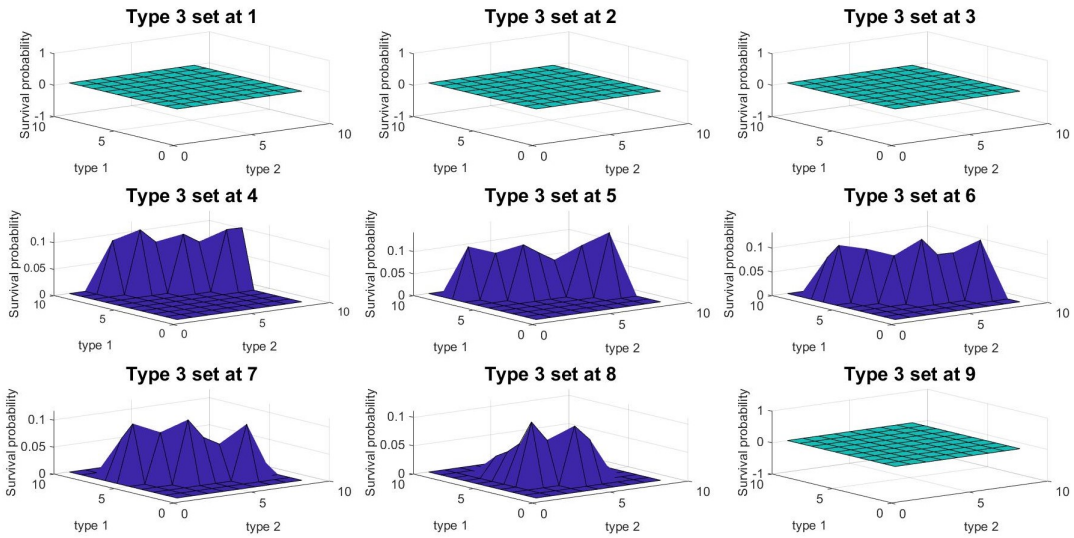


Figure 3.11: China's high-speed network

Figure 3.11 displays the surface plots for the coefficient of variation, $c_{\underline{l}}$ for every distinguishable combination of nodes trialed out from Figure 3.10. There are no $c_{\underline{l}}$ values for the simulations regarding the events concerning less than three type 3 components functioning. This is due to the negligible survival probability of the system when conditioned on this state. However, when the number of type 3 nodes functioning reaches 4, the variance is observable when the majority of type 1 and 2 nodes are also functioning. As an increasing number of type 3 nodes function, this peak of the $c_{\underline{l}}$ slowly withdraws into conditions where a decreasing value of type 1 nodes function. The $c_{\underline{l}}$ does not exceed 0.1 for any of the combinations of nodes in the survival signature, however peaks at the conditions with path sets of low probability. This is due to the uncertainty build-up for conditions of possible path sets with low likelihoods of survival. When all the nodes of type 3 are in a functioning state, the $c_{\underline{l}}$ is completely eliminated due to the very high survival probability of the system.

3.5.1.3 Survival Function Simulation

The survival function is estimated with the three types of components being assigned imprecise attributes regarding their failure characteristics. The table below shows the types of distributions selected and their respective epistemic uncertainties;

Component Type	Distribution	Scale parameter (η)	Shape parameter (β)
1	Weibull	1.9 ± 0.2	2.8 ± 0.4
2	Weibull	2.3 ± 0.3	2.6 ± 0.2
3	Weibull	1.5 ± 0.5	1.8 ± 0.4

Table 3.4: Survival function component failure types with distribution parameters

All three components for the survival function are attributed with Weibull distribution failure characteristics. The interval bounds contain varying values and both the upper and lower bounds have been simulated in the analytical solution. However, it is first important to deduce the precise survival function which has been tested with both the analytical solution and the Monte Carlo simulation method as presented in the Algorithm 2 in Section 3.4.1. The chosen number of samples for this Monte Carlo simulation is 10000 samples, costing a computational time of 158 seconds.
Figure 3.12 presents the results of both methods in one graph with respects to non-dimensional time.

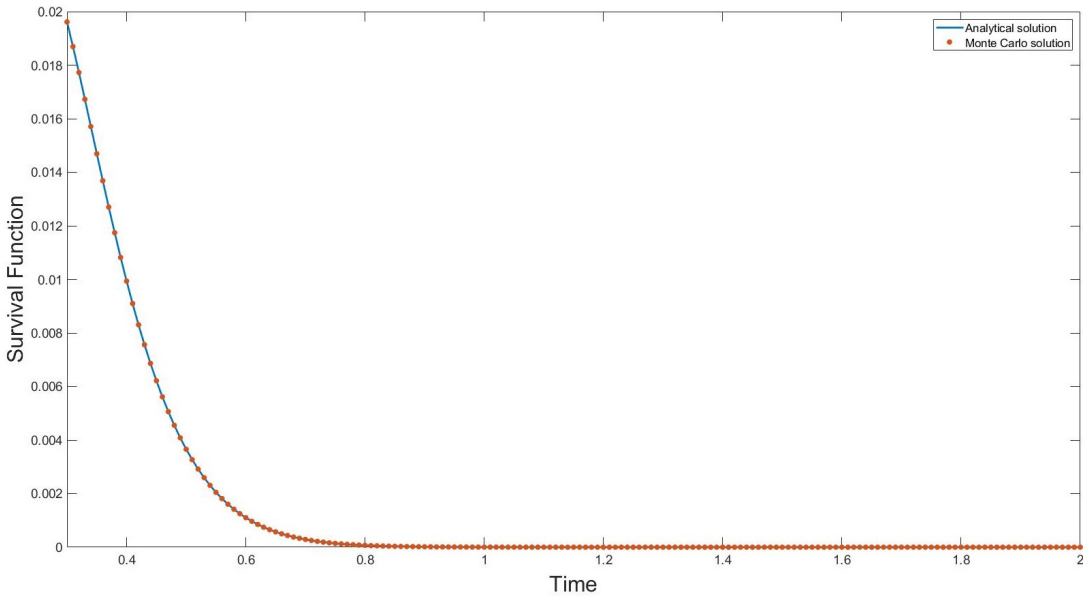


Figure 3.12: Survival function for both analytical solution and Monte Carlo

The accuracy of the Monte Carlo simulation with respects to the analytical solution is evident from Figure 3.12 via the display of congruence between both methods. The survival function for both methods declines sharply within moments of simulation and decelerates until it reaches a complete failure at 0.8 non-dimensional time. The system is assumed to be non-repairable and therefore must be reset before the system can be re-simulated.

Uncertainty from the NPI method as mentioned in Section 3.3.1.1, epistemic uncertainty from estimation as explained in Section 3.3.1.2 and is presented in Table 3.4 , and aleatory uncertainty as described as a Gaussian in Section 3.3.1.3 is also applied to the survival function as displayed in Figure 3.12.

Figure 3.13 displays the imprecise survival function with three different inputs of uncertainty applied;

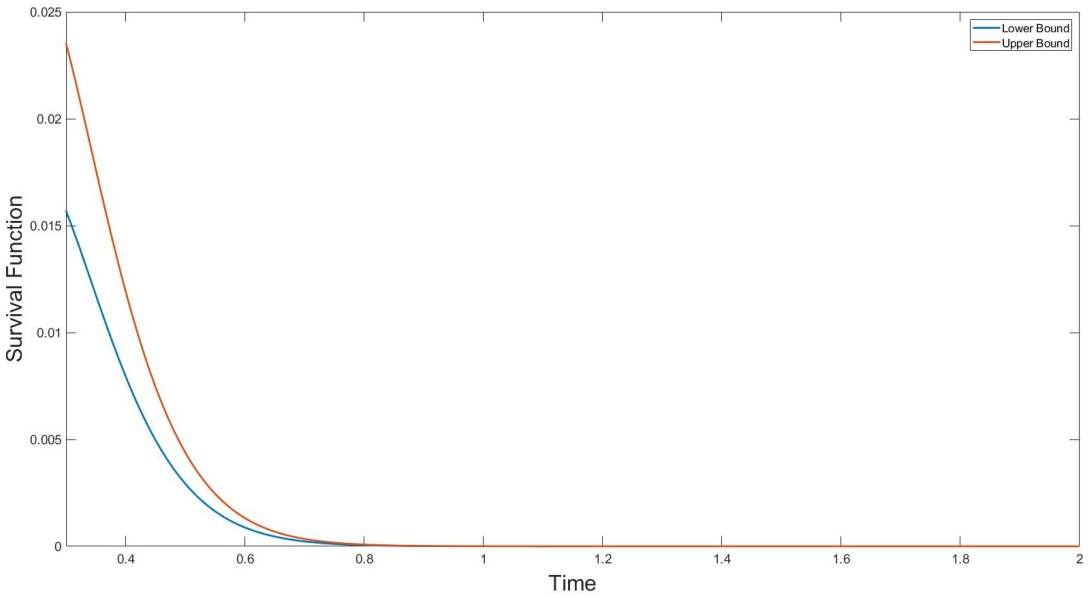


Figure 3.13: Imprecise survival function

The confidence bounds as displayed in 3.13 show the lower and upper survival functions respectively. Both bounds start at a high amount of divergence within the survival function and slowly converge until the failure time of 0.8 non-dimensional time is achieved as matched in Figure 3.12. It is also important to note that this is only the analytical solution for the survival function as it is not necessary to carry out a Monte Carlo simulation as congruence between the two methods has been proved in Figure 3.12.

3.5.2 PrPm Results

The PrPm has also been tested for China's high-speed rail network. The flowchart presented in Section 3.4.2 has been implemented to this case study and the input and output targets have been kept consistent with the survival signature implementation in Section 3.5.1. The exact solution for PrPm has been confirmed in GeNie Academic 2.5, a software to construct, implement and test out Bayesian networks.

3.5.2.1 Non-Directed Acyclic Graph

The PrPm method has been applied to the same topology of the DAG in Figure 3.9, however does not use the nodal directions as displayed on this figure. The graph is therefore a non-directed acyclic graph and displays the respective nodes and links used for the PrPm as applied to this case study. The types of nodes listed in Table 3.3 are negated as node type discrimination is only applied in the survival signature and is not applied in the PrPm. The source nodes applied to this case study are nodes 1 and 4 respectively, and the terminal node is node 26, and matches the terminal node for the survival signature implementation.

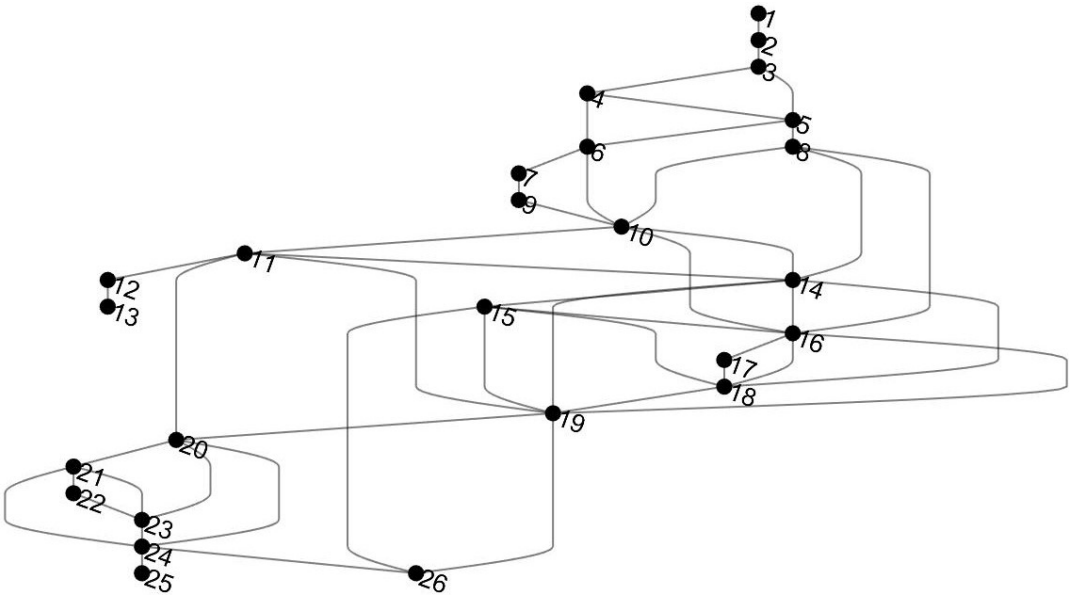


Figure 3.14: Non-DAG for China's high-speed rail network

3.5.2.2 Propagation Step Results

Each node has been propagated sequentially in numerical order and the posterior probabilities have been recorded and plotted for the both the outer approximation algorithm and the inner approximation algorithm. For the outer approximation algorithm, the data for every step in the propagation sequence is recorded and is plotted in Figure 3.15, along with the precise PrPm solution for comparison.

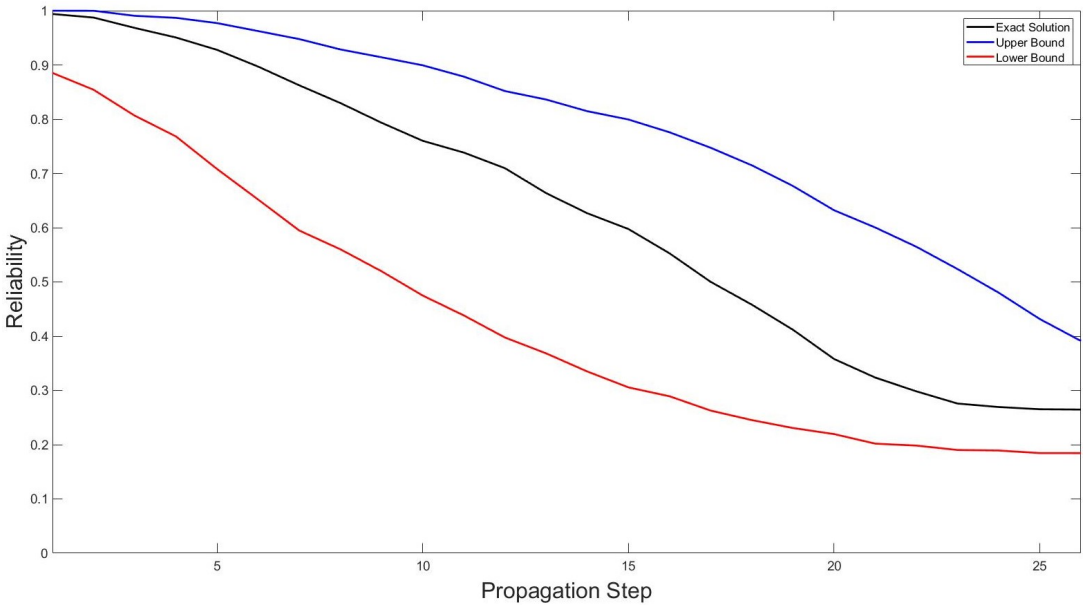


Figure 3.15: PrPm outer approximation algorithm

Figure 3.15 highlights the divergence between the upper and lower bounds from the exact solution of the propagation sequence. The outer approximation algorithm shows a relatively conservative estimate for the reliability values as shown in the sparse nature of the interval bounds. These bounds diverge as the propagation sequence progresses and peaks in divergence at propagation step 15 where the upper and lower bounds are bounded by approximately 0.5. This increasing divergence is due to the increased overload of joint probability data being accumulated for the nodes during the propagation sequence. This divergence is then reduced towards the end of the propagation sequence and the terminal node probabilities are deduced as [0.1844, 0.2647, 0.3912].

The same method has been applied for deducing the propagation sequence with the application of the inner approximation algorithm and the results for each propagation step are displayed in Figure 3.16.

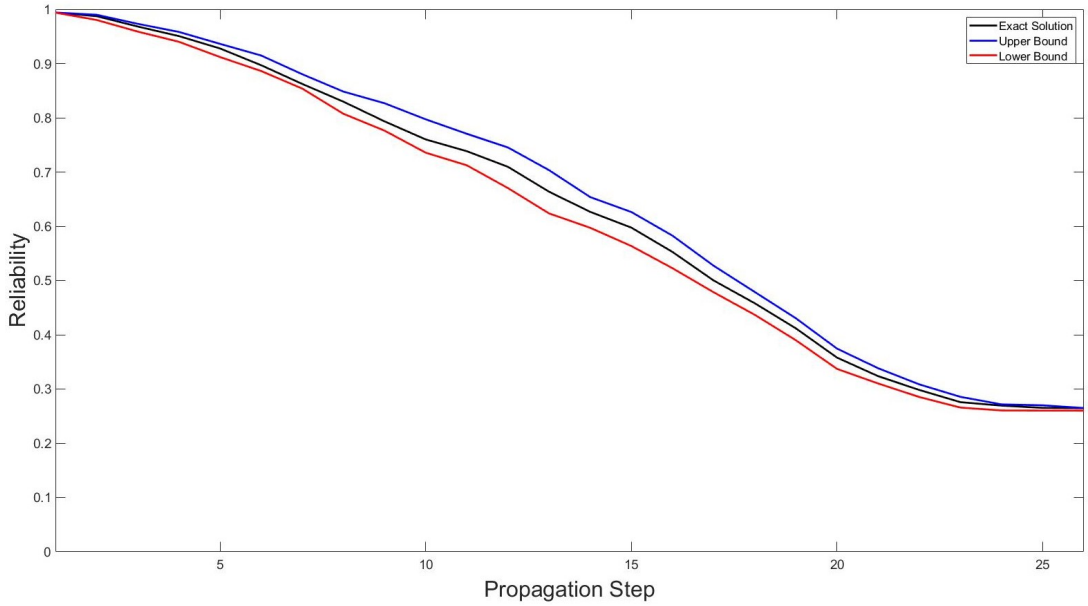


Figure 3.16: PrPm inner approximation algorithm

The approximation bounds for the inner approximation algorithm display a very low level of variance as expected when compared to the outer approximation algorithm in Figure 3.15. This is especially prevalent in the first and latter stages of the propagation sequence where the bounds barely differ from the exact solution. The terminal node's probability is finally yielded as [0.2602, 0.2647, 0.2650].

3.6 Chapter Summary

This chapter presents the theory, system simulation and a case study into both the survival signature and the analytical probability propagation method. The chapter also compares the two methods and states the advantages and limitations in each method. Both epistemic and aleatory uncertainties have been applied in different angles to the binary-state system of China's high-speed rail network. The results conclude by showing that these methods are suitable for quick and efficient simulation and can be applied as a useful tool to compute the reliability of a binary state system with minimal computational expense. The limitations of both techniques are also discussed and therefore it is recommended to use both of these methods for preliminary estimation of system reliability, rather than with in-depth analysis motives.

Chapter 4

A Weather-based Contingency Model to Quantify Resilience in Power Grids

4.1 Introduction

The power grid is an essential tool for modern society and its function is crucial for society. A single failure of the system can lead to major consequences in a socio-economic context. Events such as the 2003 British National Power Grid Corporation outage which was responsible for the load loss of 724 MW, or approximately 20 % of London's power supply have costed the UK a significant economic burden [113].
The assessment of reliability in power grid systems and the parameters incorporating reliability such a availability, consequence modelling and load outages have been of great significance for research in the IEEE (Institute of Electrical and Electronics Engineers) community with examples of works mentioned in section 2.3.3.
The constraint with power grid systems as with other realistic engineering systems is the complexity of their parameters and conditions such as the topology, interdependencies and behavioural interaction from external factors. This makes analytical techniques such as those trialed out in Chapter 3 unfeasible and inaccurate. Performance parameters in power grid analytics are varied, with the role of load supply and demand as the main analysis metrics being used for most works.
This chapter analyses how Monte Carlo simulation can be applied into resilience quantification for operational resilience in general power grid systesms containing both

buses and generators. More specifically the aim of this chapter is to provide a study into different types of simulation methods for quantification of resilience under uncertainty highlighting all the theory, simulation algorithms and limitations with these approaches. This chapter also provides the user with an insight to the basics of MATPOWER, an open-source toolbox applied on MATLAB 2021b which is used to analyse electric power simulation and optimisation.

4.1.1 Overview and Proposed Approach

In this chapter, three approaches to quantify resilience in power grids have been applied. Firstly a method using Line Outage Distribution Factors (LODFs) to measure contingencies to localised areas acting as a surrogate for DC power has been simulated. The other two methods use power flow equations which implement DC optimal power flow (DC-OPF) and AC optimal power flow (AC-OPF) calculations respectively to quantify load outage and operational recovery in proportion to the network's expected load demand.

The technique that has been trialed out is the load-flow simulation technique and utilises Monte Carlo simulation to approximate the resilience profile of the load reconstruction following the event of a disaster caused by possible weather-induced outages triggered by both high winds and lightning strikes. The chapter also proposes uncertainty in the form of both epistemic and aleatory uncertainty, and have been applied for all three of the approaches stated above.

Due to the high computational cost of the AC-OPF model, a surrogate model of the power grid performance is trained using the initial AC-OPF and has been trialed out on the same case study for comparative means. This approach is based on the use of artificial neural networks (ANNs) and is applied using the load outage data to assess the expected energy-not-supplied to the network.

Finally, the ANN surrogate model is applied to estimate the resilience of the case study and the results displayed by the this model have been been compared to the original model AC-OPF model. The validity of the ANN performance parameters with respects to its the congruence to the original AC-OPF model has been fully tested and the credibility of this surrogate is discussed.

4.2 Background and Theory

Before the model is applied, it is vital to understand the theory behind the three techniques as applied to MATPOWER, as the equations used in each method provides a different angle to quantifying the load values.
Figure 4.1 displays a simple power grid network as obtained from the MATPOWER case files.

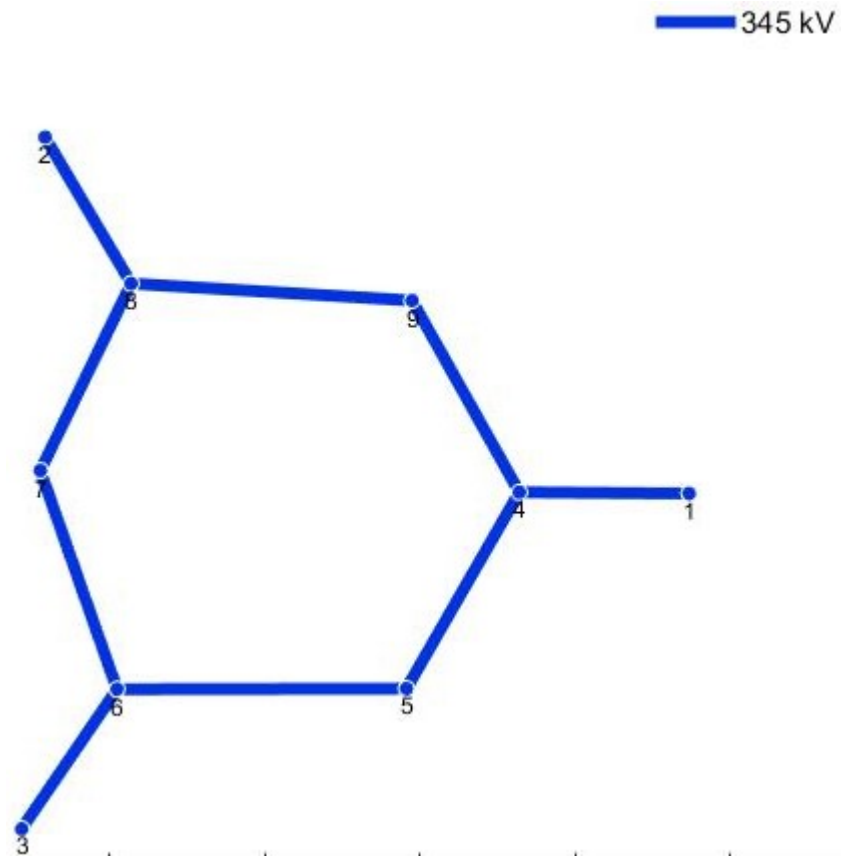


Figure 4.1: Case 9 network topology

The network displays the electrical topology of a 9 node network containing 9 links. All the nodes have been labelled and the base voltage capacity of the links are equal at 345kV. The MATPOWER values displaying the specific attributes of the network which have been applied for power flow calculations are provided on the MATPOWER file. These parameters include the real power demand represented as P_d, the reactive power demand represented as Q_d, the voltage magnitude represented as V_m and the voltage angle represented as V_a.

4.2.1 Line Outage Distribution Factors

The first optimisation problem uses line outage distribution factors (LODFs) and can be described as a sensitivity measure applied to compute a change in a line's status as influenced by changes to the adjacent lines of the system. The concept of the LODF has been developed and expanded from the power transfer distribution factor (PTDF) which is a linear based sensitivity analysis based on DC-OPF assumptions. PTDFs calculate differences in real power, P being transferred on the transmission lines of the power grid system for the user's selected source and terminal observation ranges of the system.
LODF's are obtained by expanding the PTDFs to compute the change in active power after the transmission lines are being removed sequentially. The LODF d_l^k calculates the difference in active power flow from a pre-contingency state to post-contingency state on a specific line l as a result of the complete removal of the line [114].

Assume a network line, l with power flow f_l and failure flow, f_k.
The post-contingency flow is calculated as;

$$f_l^k = f_l + d_l^k f_k \tag{4.1}$$

where d_l^k represents the value of the lth row and k^{th} column of the LODF matrix.

Equation 4.1 provides an approximation of the post-contingency power flow rate, as this applies a linear approximation to provide an estimation of the power flow equation. The degree of accuracy of the LODF calculation has a general positive consensus and is commonly used by operators in contingency analysis. This is why it is commonly used as an estimation of DC optimal power flow models and when applied sequentially, is regarded as a computationally inexpensive means to estimate the severity of the overall contingencies of the system.

Given that $N-1$ line failures occur, the composite risk index is computed as;

$$\mathscr{R}(\zeta) = \sum_{k=1}^{n_C} \mathscr{P}(C_k|\zeta) \sum_{l=1}^{n_l} S_{O,L,l}(C_i, \zeta) \tag{4.2}$$

where n_l represents the total number of transmission lines within the system, n_c represents the total number of contingency failures, ζ represents a given operational condition and C_i represents a given contingency.
Equation 4.2 is applied to compute the risk index via a power flow analysis computing f_l for the pre-contingency driven network. This provides much faster computational efficiency as opposed to the general DC-OPF algorithm as it does not require the solution of n_c to deduce power flow.

4.2.2 Power Flow Equations

With regards to power flow optimisation, the two models are the DC optimal power flow (DC-OPF) approach and the AC optimal power flow (AC-OPF) approach. In real world power grid systems, the electricity is generated in power plants using methods such as fossil fuels, converted fuels or geothermal steam and transfers this energy through the transmission network at high voltage using either DC or AC flow [115]. This high voltage steps down into a medium voltage range.
The primary difference between the DC and AC optimal power flow models is within the nature of the convexity. DC stands for direct current and the power flow is constantly in a steady state, therefore the constraints are presented as both a linear and convex optimisation problem. AC stands for alternating current and the optimal power flow calculations are non-linear and non-convex in nature leading to a significantly higher computational expense. It should also be noted that in high-fidelity models, DC-OPF simulation techniques are limited in terms of details for these networks [116]. This is due to DC-OPF models containing an estimation of AC-OPF models only accounting for real power, P, and negating the effects of reactive power, Q in the model [117].
The general optimal power flow approach is formulated as follows [118]:

The standard objective function is denoted as;

$$\min_{x} f(x) \tag{4.3}$$

subject to the equality constraint;

$$g(x) = 0 \tag{4.4}$$

subject to the inequality constraint;

$$h(x) \leq 0 \tag{4.5}$$

and subject to parameter constraint;

$$x_{min} \leq x \leq x_{max} \tag{4.6}$$

4.2.2.1 AC Optimal Power Flow

In the case of AC-OPF, the model is the AC cascade failure model (ACCF), which assumes that certain nodes in the system have failed. The failed nodes have been removed from the system and the updated system status is computed through the use of optimisation equations with necessary load outages. If steady state has still not been achieved, the model will break the overload branches and an updated calculation of the power flow distribution is performed for the whole network. This process is repeated until the system achieves a new balance satisfying all the constraints, and finally the load deficiency is computed through the final state of the system's post-cascading failure.

The assumptions of the model are as follows;

1. The line status is assumed to have binary-state characteristics which indicate either "1" for a functioning or "0" for a failed line.
2. The initial grid failures are not considered mutually dependent.
3. Artificial repair does not play a role in the process of the cascading failures.
4. Node failure will cause the branches connected with the associated node to also fail

The power grid is denoted as a network consisting of N nodes and K branches with the nodes representing the buses as displayed on Figure 4.1. The buses are either classified as load buses, generator buses or slack buses and are discriminated using the data obtained from the MATPOWER file.
In the case that the selected bus, i is a generator bus, the inequalities $0 \leq P_i^{min} \leq P_i^{max}$ and $0 \leq Q_i^{min} \leq Q_i^{max}$ apply during function.
The ACCF model identifies the isolated parts of the grid that are conformed through the initial failure. Cascading failures are then simulated until all these failures terminate and the optimisation simulation also terminates.
The optimisation equations are subject to the following assumptions;

The optimisation vector, x for AC- OPF is as follows:

$$x = \begin{bmatrix} V_a \\ V_m \\ P_g \\ Q_g \end{bmatrix} \tag{4.7}$$

where V_a represents the voltage angle, V_m represents the voltage magnitude, P_g represents the real power output and Q_g represents the reactive power output. The voltage phase angle vector for the respective buses is set at $n_b \times 1$ with n_b representing the number of buses in the power grid system. The reactive power vector is set at $n_g \times 1$ where n_g represents the number of generators in the power grid system.

For each individual generator, the cost function of real power, f_P^i and reactive power, f_Q^i are applied to Equation 4.4 to be satisfied as;

$$\min_{\Theta, V_m, P_g, Q_g} \sum_{i=1}^{n_g} f_P^i(p_g^i) + f_Q^i(q_g^i) \tag{4.8}$$

This can be expanded into the real power and reactive power balance constraints;

$$g_P(V_a, V_m, P_g) = P_{bus}(V_a, V_m) + P_d - C_g P_g = 0 \tag{4.9}$$

$$g_Q(V_a, V_m, Q_g) = Q_{bus}(V_a, V_m) + Q_d - C_g Q_g = 0 \tag{4.10}$$

The inequality constraint displayed in Equation 4.5 contains two sets of n_l branch flow equations in non-linear form with the application of V_m and V_a;

$$h_f(V_a, V_m) = |F_f(V_a, V_m)| - F_{max} \leq 0 \tag{4.11}$$

$$h_t(V_a, V_m) = |F_t(V_a, V_m)| - F_{max} \leq 0 \tag{4.12}$$

F_{max} represents the vector flow limits of the branch. The flows are active or current flows, which result in the equations;

$$F_f(V_a, V_m) = \begin{cases} S_f(V_a, V_m) \\ P_f(V_a, V_m) \\ I_f(V_a, V_m) \end{cases} \tag{4.13}$$

S_f represents the apparent power flow, P_f represents the real power flow, and I_f represents the current flow.

S_f is computed through the equation;

$$S_f(V) = [C_f V] I_f^* = [C_f V] Y_f^* V^* \tag{4.14}$$

and I_f is calculated by applying the admittance matrix Y_f;

$I_f = Y_f V$

where Y_f is composed of $n_l \times n_b$.

Equation 4.6 is extracted into multiple constraints as displayed in the equations below;

$$V_a^{i,ref} \leq V_a^i \leq V_a^{i,ref}, \qquad i \in \Gamma^{ref} \tag{4.15}$$

$$v_m^{i,min} \leq v_m^i \leq v_m^{i,max}, \qquad i = 1, 2, ..., n_b \tag{4.16}$$

$$p_g^{i,min} \leq p_g^i \leq p_g^{i,max}, \qquad i = 1, 2, ..., n_g \tag{4.17}$$

$$q_g^{i,min} \leq q_g^i \leq q_g^{i,max,} \qquad i = 1, 2, ..., n_q \tag{4.18}$$

4.2.2.2 DC Optimal Power Flow

The DC-OPF equation follows similar suit, however neglects the parameters of reactive power and voltage magnitude;

$$x = \begin{bmatrix} V_a \\ P_g \end{bmatrix} \tag{4.19}$$

which reduces Equation 4.8 to;

$$\min_{V_a, P_g} \sum_{i=1}^{n_g} f_P^i(p_g^i) \tag{4.20}$$

where f_P^i represents the real power cost function and p_g^i represents the real power output.

with respects to;

$$g_P(V_a, P_g) = B_{bus}V_a + P_{Bus,shift} + P_d + G_{sh} - C_g P_g = 0 \tag{4.21}$$

$$h_f(V_a) = B_f V_a + P_{f,shift} - F_{max} \leq 0 \tag{4.22}$$

$$h_t(V_a) = -B_f V_a - P_{f,shift} - F_{max} \leq 0 \tag{4.23}$$

$$V_a^{i,ref} \leq V_a^i \leq V_a^{i,ref}, \qquad i \in \Gamma^{ref} \tag{4.24}$$

$$p_g^{i,min} \leq p_g^i \leq p_g^{i,max}, \qquad i = 1, 2, ..., n_g \tag{4.25}$$

The following equations prove that DC-OPF provides a linear approximation solution to the AC-OPF equations as stated in Section 4.2.2.1. This is due to the various assumptions for deriving Equation 4.20 from Equation 4.8, which include the assumption of a flat voltage profile, negligible differences between V_a values and negligible voltage resistance within the transmission lines [73]. All three assumptions are not present in AC-OPF.

4.3 The Weather Based Contingency Model

The general topology of the power grid is represented in a graph as shown in Figure 4.1 and can be mathematically expressed as $\mathscr{G}(\mathscr{N}, \mathscr{E})$ [119], where $\mathscr{N}$ represents a set of nodes and $\mathscr{E}$ represents edges denoting a set of transmission lines between a given set of nodes i and j. The other parameters in this model include the number of loads denoted as n_L, the number of lines denoted as n_l and the number of generators denoted as n_g.

The proposed weather based contingency model can be applied with respects to any of the three approaches listed in Section 4.2. These power flow models are applied to obtain the solution for the power dispatch constraint [120] and failure of the nodes can cause load curtailment. The cost of this is expensive for the network and occurs on the condition that the minimisation constraint from Equation 4.3 is unsolvable. This constraint is expanded for this model as;

$$\min_{P_g, L_{cut}} f(P_g, L_{cut}) \tag{4.26}$$

where L_{cut} represents the value of the load curtailed.

4.3.1 Load Contingencies

A contingency is an event occurring that is not considered predictable at a given time. When applied to the power grid network, contingencies imply the network's architecture is is experiencing a disruption of the load transfer from one bus to the next. This is commonly caused by a failure by extremely hot weather, system failures such as outages and human errors [121].

The chosen parameter to measure the resilience quantification for the power grid system is the Expectation of Energy-Not-Supplied (ENS) which is deemed to be the most appropriate performance indicator and has historically been used as an indicator of reliability. The ENS acts as the ratio between the actual load received as compared to the maximum load demand and is averaged over a defined period of time;

$$ENS = \sum_{t=1}^{T_{sim}} \sum_{i \in \mathscr{N}} L_{cut,i,t} \cdot t \tag{4.27}$$

where T represents the time of interest, T_{sim} represents a user defined simulation period and $L_{cut,i,t}$ represents the load curtailed at each individual node i at a given time t.

The general Expectation of the ENS, $\mathbb{E}[ENS]$ is obtained by averaging the samples of N_s independent simulations and is denoted as;

$$\mathbb{E}[ENS] = \frac{\sum_{i=1}^{N_s}}{N_s} \tag{4.28}$$

4.3.2 Severe Weather Model

As an extension to the contingencies faced in the system, a weather model has been proposed in the simulation algorithm to mimic the real life application of an event. These events include lightning strikes, extremely high winds and natural disasters. The occurrence of normal weather conditions can be modelled as a homogeneous Poisson process [122];

$$P(N_f(t) = k) = \frac{[\lambda_n \cdot t]^k}{k!} e^{-\lambda_n \cdot t} \quad k = 0, 1, ..., N \tag{4.29}$$

where $P(N_f(t) = k)$ represents the probability that k failures happen within the network given the time $(0, t)$, λ_n represents the line failure rate under normal conditions and $N_f(t)$ represents the number of failures per km of grid line.

However, in a more realistic perspective, the weather model is more likely to be affected by uncertainty. This is why the occurrence of severe weather events is more suited to be modelled by a non-homogeneous Poisson process;

$$P(N_e(t) = k) = \frac{[V_e(t)]^k}{k!} e^{-V_e(t)} \quad k = 0, 1, ..., N \tag{4.30}$$

In this case, $V_e(t)$ represents the time dependent probability of the event occurring and can be obtained by applying the following equation;

$$V_e(t) = \int_0^t v_e(t')\, dt' \tag{4.31}$$

where $v_e(t')$ represents the rate at which the disturbance occurs.

Given a severe weather occurrence, the time of the event is obtained from data extracted from previous events and has been modelled using probability distribution functions.

4.3.2.1 High Winds

In the case of high winds, the wind storm intensity is obtained via the following equation;

$$W_w(t) = W_{crt} + \Delta_w(t) \tag{4.32}$$

where $W_w(t)$ represents the wind speed intensity at time t for the event w and W_{crt} represents a datum wind speed known as the critical wind speed set at 10 $m\, s^{-1}$. $\Delta_w(t)$ represents the difference between the critical wind speed and the actual wind speed during the event.

When considering individual lines, the contribution to the line failure due to high winds can be denoted in the equation;

$$\lambda_w(W_w(t)) = \lambda_n \left(\frac{W_w(t)^2}{W_{crt}^2} - 1 \right) \alpha_w \tag{4.33}$$

where α_w represents the regression parameter for failure data obtained.

4.3.2.2 Lightning Strikes

The other weather parameter involved in this model is the effect of lightning strikes. The intensity of the chosen parameter is represented by the lightning strike ground

density, $N_g(t)$, which is characterised by the units of ground flashes per unit time and area $occ\, h^{-1} km^{-2}$. The parameter is modelled with log-normal variability.

The line failure rate as a result of lightning can be denoted as:

$$\lambda_{lg}(N_g(t)) = \lambda_n \beta_{lg} N_g(t) \tag{4.34}$$

β_{lg} is the regression coefficient obtained from prior data [122].

4.3.2.3 Combined Weather Model with Uncertainty

The combined weather model contains parameters for the various weather models proposed with respects to epistemic uncertainty. These parameters include the duration of the wind storm event, D_ω and the duration of the lightning storm event, Dl_g.
The various parameters of the weather model have been assigned intervals as denoted in Table 4.1;

	Distribution	**Scale parameter (η)**	**Shape parameter (β)**
D_ω	Weibull	9.89 ± 1.42	1.17 ± 0.57
Dl_g	Weibull	0.96 ± 0.16	0.85 ± 0.09
$\Delta_\omega(t)$	Weibull	1.23 ± 0.36	1.05 ± 0.17
		Mean (μ)	**SD (σ)**
$N_g(t)$	Log-normal	-5.34	1.07

Table 4.1: Weather model parameters

Both high winds and lightning strikes are a cause of contingency and therefore it is crucial to define an equation which takes into account both forms contingencies to calculate the total failure rate;

$$\lambda(t) = \lambda_n + \lambda_w(W_w(t)) + \lambda_{lg}(N_g(t)) \tag{4.35}$$

where λ_w represents the total line failure contribution during time t due to high wind measured per km and $\lambda_{(lg)}$ represents the lightning storms contribution.

4.3.3 Weather Model Repair Speed

The recovery model [123] takes into consideration the efficiency of the repair crew as they are also affected by the adverse weather conditions. The assumptions in this model are:

1. The repair is instantly initiated following failure.
2. A line is considered instantaneously fully functioning in a post-repair state.
3. The time of for the transition of the failure is negligible with respects to the repair time.

$$v_{repair} = \begin{cases} \frac{v_{norm}}{1+\eta\cdot(W_w(t)-W_{crt})}, \\ if\ W_w(t) \geq W_{crt}, N_g = 0 \\ \\ \frac{v_{norm}}{1+\psi\cdot N_g}, \\ W_w(t) < W_{crt}, N_g > 0 \\ \\ \frac{v_{norm}}{[1+\eta\cdot(W_w(t)-W_{crt})]+[1+\psi\cdot N_g]}, \\ \text{if } W_w(t) \geq W_{crt}, N_g > 0 \end{cases} \tag{4.36}$$

In this model, ψ and η are positive parameters and the normal average repair speed, v_{norm} is set at 20 % h^{-1}. The values for ψ and η are set to 40 and 0.4 respectively.

4.3.3.1 Aleatory Uncertainty

The aleatory uncertainty applied to the repair speed is taken from a Gaussian stochastic model and is denoted as:

$$f(L_i(t)) = \frac{1}{\sqrt{2\pi\sigma_{Li}(t)}} e^{-\frac{(L_i(t)-\mu_{Li}(t)}{2\sigma(L_i)(t)^2}} \tag{4.37}$$

4.4 Implementation of Technique

The proposed framework for this study has been tested for all three power flow techniques. The AC-OPF model has also been modelled alongside its surrogate model. This proposed methodology has been implemented in OpenCossan in a MATLAB 2021b environment. The proposed steps for the implementation of the technique is explained by Rocchetta et al. [72];

4.4.1 Power Flow Analysis

The first part of implementation requires the user to compute the basic power flow attributes as applied to a given network. A pre-contingency power flow model is applied as selected by the user from Section 4.2 obtaining the line flow values, f_l and the power rating, PR_l in a pre-contingency state. This provides the required data for the demand loads which is applied within the resilience analysis framework.

The individual line failures are deduced sequentially using the power flow analysis as follows;

1. A chosen line, l is removed from the intact network.
2. Compute post-contingency f_l and PR_l using the power flow equations.
3. Identify the connected components within the post-contingency network.
4. If all lines are intact, the power equations are already solved, therefore f_l and PR_l are obtained.
5. If the network is not intact, single node islands, G_{is} are removed and negated.
6. For G_{is} clusters remaining, post-contingency load values are obtained.
7. The overload severity is calculated and the cascading failure probability is obtained for the surviving lines j.
8. Repeat steps (1)-(7) until all islands are isolated.

The general algorithm for the post-contingency power flow equation is summarised follows;

Algorithm 3 Power Flow

```
procedure (Post-Contingency Power Flow)
    Find connected components
    if cc > 1 then
        Remove isolated nodes
        ∀ G_is
        Select slack bus amongst P − V nodes
        Run selected flow equations
```

4.4.2 Power Grid Monte Carlo Simulation

The weather model described in Section 4.3 has been implemented into the system. Both homogeneous Poisson processes (HPPs) and non-homogeneous Poisson processes (non-HPPs) are applied to this model. A vector representing the "Time-To-events" (TTEs) is formulated after deducing the type of event and the time of which this event occurs. The sequential Monte Carlo simulation (SMC) is applied until all the samples of the system have been analysed.
Rocchetta et al. [70] provides a summary of the SMC technique;

1. The time, t is set as the time of the first event occuring and the failure index, $f = 0$. If a normal failure is applied, go to (2). If a non-normal failure is applied, go to (3).

2. $f = f+1$ is set and a chosen line, l is sampled from the probability mass function (PMF) distribution and values $\frac{\lambda(t)\cdot l_i \cdot X_{f,i}}{\sum_{l=1}^{N_l} \lambda(t)\cdot l_i \cdot X_{f,i}}$ with $l = 1, ..., N_l$. $X_{f,i}$ are set at 0 and a load profile, $\mathbf{L}_f$ is sampled at time t. The failed line replacement is restored and is therefore deemed to be functioning. The $\mathbf{X}_f$ and $\mathbf{L}_f$ values are saved for f. The load recovery is computed from the replacement with $TTE(e+1) - TTE(e)$ and after this is completed, go to (5).

3. The duration of the severe weather model, T_e is sampled with respects to its intensity, N_g, Δ_ω to calculate the increasing total failure rate, $\lambda(t)$ by implementing Equation 4.35. The time samples are simulated using the HPP method

with the input of $\lambda(t)$ and interval $[t\, t + T_e]$. In the event that single or multiple failure events have been sampled, t is set to the next failure event. After this is completed, go to (4). If no failure events have occurred, go to (5).

4. $f = f + 1$ is set and one failed line, l is sampled using the PMF distribution and values $\frac{\lambda(t) \cdot l_i \cdot X_{f,i}}{\sum_{l=1}^{N_l} \lambda(t) \cdot l_i \cdot X_{f,i}}$ with $l = 1, ..., N_l$. A load profile, $\mathbf{L}_f$ is sampled at time t and the replacement line, l is restored with the state vector, $\mathbf{X}_f$ and new load profile $\mathbf{L}_f$ is saved. The load recovery is computed by applying Equation 4.36 and if severe weather failures are found in the system then go to (5). If severe weather failures are not found, then repeat this step.

5. Stop the simulation if $t > T_{sim}$ or if the final event, e in the simulation has occurred. Otherwise e is set to $e + 1$ and $t = TTE(e)$. Restart the simulation from (2).

Upon completion of the simulation, the output data is saved for the use of the surrogate model as applied in Section 4.4.3. The values of $\mathbf{L}_f$ and $\mathbf{X}_f$ are stored to be applied as a vector input for the $(n_l + n_L) \cdot F$ matrix, where F represents the number of annual failures experienced by the network.

4.4.3 ANN Surrogate Model for the Power Flow models

The purpose of an emulator is to minimise computational time for the simulation and few attempts have been carried out in the context of finding a surrogate model incorporating contingencies. The artificial neural network (ANN) is a proposed method which has increased in popularity for various applications as its flexibility to be applied in various projects such as in finance [124], data validation [125] and weather modelling [126]. The applied network is modelled using the ANN toolbox on MATLAB 2021b. This toolbox enables the user to input data from the original model and mimic the simulation method using a certain set of epochs as desired. The output, which is the ENS value produces a final result based on the conditions inputted into the network's architecture.

Rocchetta et al. [70] innovated the original model from Section 4.4.2 and developed a surrogate model based on the load data obtained from the original model.

4.4.3.1 Network Architecture

An ANN can be mathematically expressed to define the function $K : \mathbf{I} \rightarrow \mathbf{Y}$, where K represents a composition of various weighted functions, $g_i(s)$. The type of neural network applied to this model is a feed-forward neural network with the Levenberg-Marquadt algorithm applied. The architecture of a feed-forward ANN consists of an input layer, at least one hidden layer and an output layer [127]. The hidden layers represent artificial neurons and the number of nodes can be selected by the user on the MATLAB 2021b toolbox. Each of the artificial neurons, represented as nodes are connected to their respective adjacent nodes and the number of hidden layers applied increases the depth of the ANN. However, increasing the depth of the ANN by a significant fraction requires more training time for the network, which contradicts the purpose of implementing a surrogate when training time is too high.
In each node, the inputs are given a weight and therefore the outputs are computed as:

$$g(x) = \sum_{i=1}^{n} w_i \cdot g_i(x) + b \tag{4.38}$$

where w_i represents the the weight of the node, $g_i(x)$ is the output of the node in the previous layer and b represents the weighted bias. The bias plays a role in the input and output layers and sets out an argument for the activation function K.
The activation function is used to produce the network's output and the equation for this function can be defined as:

$$K(g) = \frac{1}{(1 + e^{-g})} \tag{4.39}$$

The input vector is defined by the load demanded in each node represented by the jth input vector and the state vector of the lines connecting the two nodes $\mathbf{I}_j = [\mathbf{L}, \mathbf{X}]_j$. In a failed state, the load profile is defined as $\mathbf{L}_f = [L_1, ..., L_{NL}]_f$ and the state vector is denoted as $\mathbf{X}_f = [X_1, ..., X_{|\xi|}]_f$ where $L_1 \in \mathbb{R}^+$ and $X_i \in \{0, 1\}$. The minimisation problem from Equation 4.26 can be solved in the equation below to obtain the load curtailed $Y_j = \sum_{i \in \mathcal{N}} L_{cut,i,j,}$.

4.4.3.2 Pseudo-code of the SMC model

The pseudo-code for this whole process is displayed in Algorithm 4:

Algorithm 4 Risk Assessment Model

1: **procedure** ENS (*Risk Assessment based on* SMC)
2: Input = $(\lambda_n, \beta_{lg}, \alpha_w, W_{crt}, b_{Dw}, a_{Dlg}, b_{Dlg}, a_{\Delta w}, b_{\Delta_w}, v_{norm}, \mu_{Ng}, \sigma_{ng})$
3: $t = 0,\ e = 1,\ f = 0$
4: Normal failure events for $[0, T_{sim}]$ with HPP
5: Extreme weather events for $[0, T_{sim}]$ with $NHPP$
6: $t = TTE(e)$
7: **if** event i is a failure **then**
8: $f = f + 1$
9: Sample X_f and L_f for t
10: Update TTR
11: **if** $t > T_{sim}$ or e is last event **then**
12: Compute $ANNL_{cut}(f)$ ▷ Surrogate model applied
13: Compute $ENS(f)$ from $ANNL_{cut}(f)$
14: Go to (25)
15: **else** $e = e + 1$
16: Go to (6)
17: **else** Sample extreme weather T_d
18: Compute failure rates and sample TTF
19: **if** $t + TTF(f) > t + T_d$ **then**
20: Go to (11)
21: **else** $t = t + T_d$ and $f = f + 1$
22: Sample X_f and L_f for t
23: Update TTR
24: Go to (19)
25: **OUTPUT ENS**

4.4.4 Imprecision

The model can be represented with the addition of imprecise probabilities. The parameters of uncertainty include both epistemic and aleatory uncertainty. The need for imprecision is advantageous in the analysis of the power grid model in events such as varied load inputs for epistemic uncertainty and weather models containing imprecise parameters leading to fluctuating results. Table 4.1 displays the confidence bounds for the imprecise input for the implemented weather model.

4.5 Case Study: GB Power Network

The GB Power Grid is a large and complex real-world system consisting of thousands of nodes. For practical reasons, a simplified version of the GB power grid, originally developed at the University of Edinburgh [128] has been chosen as the case study for this chapter. The primary data for the model was developed and obtained through a spreadsheet from the University of Strathclyde in 2010 and was converted into a MATPower file by the University of Edinburgh. The original network contains 2224 nodes and 3204 links, and this network has been reduced to 29 nodes and 50 links with the authors only including the most significant nodes being applied from the original file. These significant nodes represent electricity generators from major cities in the UK. The network has been implemented and analysed in MATPOWER and the DAG of the system is displayed in Figure 4.2 along with the line power rating values;

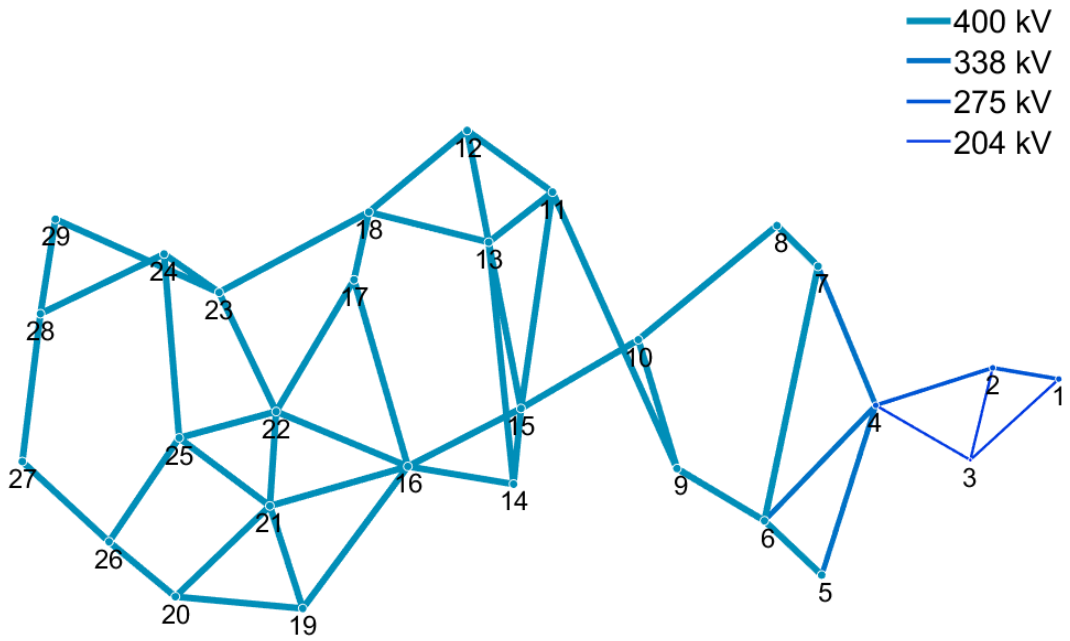


Figure 4.2: Simplified GB Power Network

The system is composed of 5 load buses, 23 generator buses and 1 slack bus. The distance of each line has been estimated using distance measurements from maps and the failure rates have been obtained from the same file. It is assumed that the transformers in each branch are working to full efficiency. The source node is set at node 1 and the terminal node is set at node 29. Table 4.2 provides details of all 50 branches on the

system with their respective distance, l_i and the total failure rate, $\lambda_{n,i}$ ($\frac{occ}{kmy}$) in occurrences per kilometer per year. The data is derived from the original paper as presented by Rocchetta and Patelli [72] and represents the expected failure occurrence per year for each kilometer of the transmission line.

Line i	l_i **(km)**	$\lambda_{n,i}$ **(**$\frac{occ}{kmy}$**)**		**Line** i	l_i **(km)**	$\lambda_{n,i}$ **(**$\frac{occ}{kmy}$**)**
1-2	135	2.760×10^{-3}		14-16	15	3.500×10^{-3}
1-3	45	6.797×10^{-3}		15-16	15	1.966×10^{-3}
2-3	100	6.551×10^{-3}		16-17	100	2.511×10^{-3}
2-4	200	1.626×10^{-3}		16-19	15	6.160×10^{-3}
3-4	20	1.190×10^{-3}		16-21	90	4.733×10^{-3}
4-5	25	4.984×10^{-3}		16-22	25	3.517×10^{-3}
4-6	25	9.597×10^{-3}		17-18	130	8.308×10^{-3}
4-7	35	3.404×10^{-3}		17-22	25	5.853×10^{-3}
5-6	20	5.853×10^{-3}		18-23	120	5.497×10^{-3}
6-7	70	2.238×10^{-3}		19-20	95	9.172×10^{-3}
6-9	100	7.513×10^{-3}		19-21	20	2.858×10^{-3}
7-8	15	2.551×10^{-3}		20-21	20	7.572×10^{-3}
8-10	80	5.060×10^{-3}		20-26	100	7.537×10^{-3}
9-10	80	6.991×10^{-3}		21-22	15	3.804×10^{-3}
9-11	135	8.909×10^{-3}		21-25	20	5.678×10^{-3}
10-15	135	9.593×10^{-3}		22-23	50	7.590×10^{-3}
11-12	30	5.472×10^{-3}		22-25	20	5.401×10^{-3}
11-13	20	1.386×10^{-3}		23-24	15	5.308×10^{-3}
11-15	110	1.493×10^{-3}		23-29	20	7.792×10^{-3}
12-13	25	2.575×10^{-3}		24-25	15	9.340×10^{-3}
12-18	130	8.407×10^{-3}		24-28	15	1.299×10^{-3}
13-14	20	2.543×10^{-3}		25-26	15	5.688×10^{-3}
13-15	100	8.143×10^{-3}		26-27	20	4.694×10^{-3}
13-18	120	2.435×10^{-3}		27-28	60	1.190×10^{-3}
14-15	15	9.293×10^{-3}		28-29	40	3.371×10^{-3}

Table 4.2: Branch properties for the GB power network

4.5.1 Original Models

The original model proposed for this simulation has been tested first without the use of an emulator. All three approaches as stated in Section 4.2 have been applied to the GB Power Network and the results obtained have been saved. The AC-OPF results

have been stored for further use of the surrogate model applied from Section 4.4.3. The simulation times for the three models have been recorded and are displayed in Figure 4.3. "CPUtime1" represents the LODF approach, "CPUtime2" represents the DC-OPF approach and "CPUtime3" represents the AC-OPF approach;

Figure 4.3: Simulation times for the power flow models

4.5.1.1 DC-OPF Results

The DC-OPF risk model has been tested with 10000 Monte Carlo samples applied. Figure 4.4 displays the results of the resilience index with respects to dimensionless time. The confidence bounds with respects to both epistemic and aleatory uncertainty is also displayed alongside the results for the original simulation.

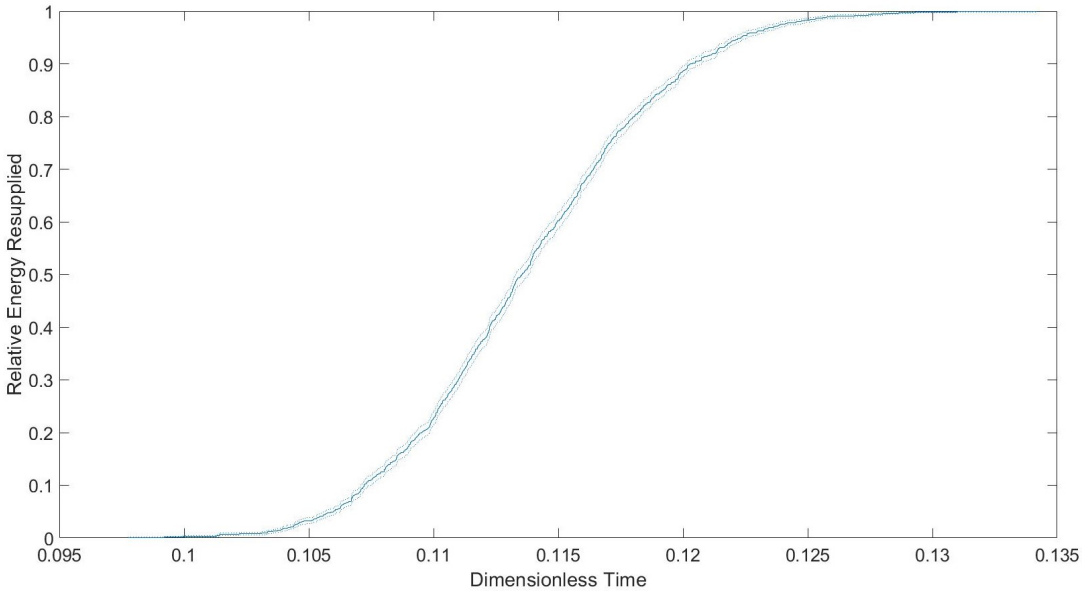


Figure 4.4: DC-OPF simulation results

The restoration function does not start immediately and only initiates after 0.1 dimensionless time. After initiation, the model slowly recovers and the uncertainty bounds are barely noticeable. The uncertainty bounds start to widen as the simulation runs as a result of the increasing load values to be analysed. This stage of restoration shows a rapidly recovering model which then decelerates due to the overload of recovery delivered to the nodes. The uncertainty bounds also diminish during the end of the simulation and the system is fully restored.

4.5.1.2 AC-OPF Results

The same model has been implemented with regards to AC-OPF. However, due to the the greater computational cost by nearly five-fold as compared to the DC-OPF model, only 1000 Monte Carlo samples have been applied.
Figure 4.5 shows the resilience profile of the case study with respects to AC-OPF;

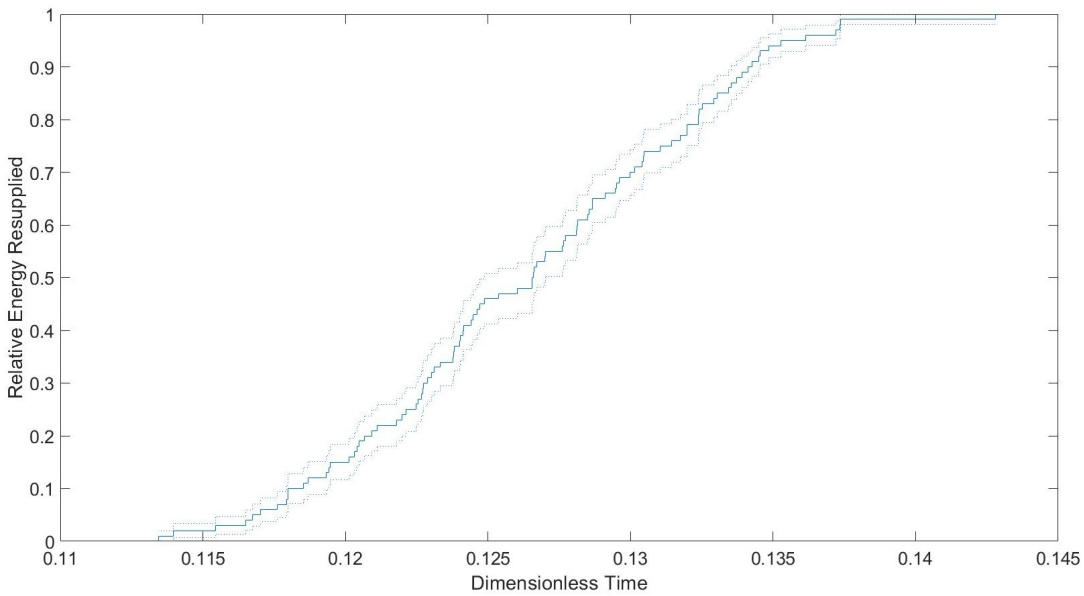


Figure 4.5: AC-OPF simulation results

As compared to the DC-OPF results, the AC-OPF starts recovery slightly at approximately 0.113 dimensionless time. This is due to the higher computational cost as the AC-OPF implements a non-linear and non-convex approximation algorithm. The model follows similar suit to the DC-OPF as recovery initiates with smaller initial confidence bounds. However, these uncertainty bounds increase more rapidly as compared to the DC-OPF technique due to the fewer Monte Carlo samples applied and

produces results with more uncertainty. The acceleration of the recovery is also slower than the DC-OPF model which shows that it requires significantly more volume to simulate the model successfully. The uncertainty bounds close near the terminal stages of the simulation and the full recovery is achieved at 0.14 dimensionless time.

4.5.2 Surrogate Models

The two surrogate models for DC-OPF and AC-OPF have also been applied for the GB power network. The LODF approach has been applied to act as a surrogate for DC-OPF and the ANN model acts as a surrogate for the highly computationally expensive AC-OPF.

4.5.2.1 LODF Results

The LODF model has been applied to the same model as the DC-OPF by applying 10000 Monte Carlo simulations and the simulation time is over 20 times faster than the regular DC-OPF which indicates that it is a useful tool for larger power grid networks. The recovery function for the LODF simulation is displayed in Figure 4.6;

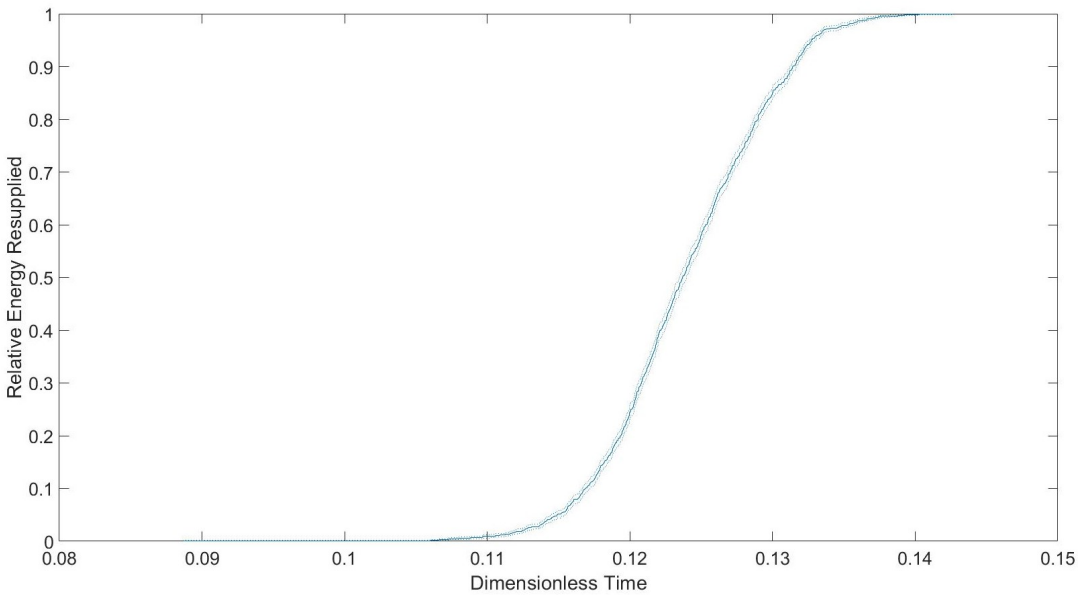


Figure 4.6: LODF simulation results

The recovery function follows the same initial pattern as the original DC-OPF results displayed in Figure 4.4. However this model shows initial stagnation stage at a slightly longer time than the DC-OPF model. The recovery function starts to slowly accelerate with minimal uncertainty bounds visible and then starts to rapidly accelerate at a much faster pace than the original DC-OPF model. The uncertainty bounds are kept tight throughout this whole simulation, suggesting a significantly lower internal volume regarding the inner parameters within the LODF algorithm as compared to the traditional DC-OPF technique. This suggests that the LODF surrogate model is a very quick and efficient computational shortcut for the estimation of DC-OPF and is appropriate for use within larger power grid networks.

4.5.2.2 ANN Surrogate Model Results

The ANN model has been developed on MATLAB 2021b and the network's architecture has been trialed with 70% training, 15% testing and 15% validation. The chosen attributes of the architecture is composed of a single input layer containing 100 nodes, 2 hidden layers both consisting of 50 nodes and a single output layer consisting of 100 nodes. Approximately 100 years of failures have been simulated and 35000 random events are selected. The line state vector, $\mathbf{X}_f$ and load samples, $L_{i,f}(t)$ are applied to train the ANN.

The surrogate model is validated using the data validation tool on MATLAB. This tool tests the surrogate model output data and selects samples from the original model. The data from the original model is compared and a regression plot is plotted displaying the regression coefficient.

The surrogate model is also tested alongside the original model to compare consistencies of the performance with the original model. This is carried out by simulating the original model and testing the performance output of the code with the behaviour of the new ANN developed. The ANN is also trained and this is an essential part of initiating the ANN simulation.

Figure 4.7 displays the regression plots of the ANN for all three simulation stages and also displays the overall regression plot for the whole simulation;

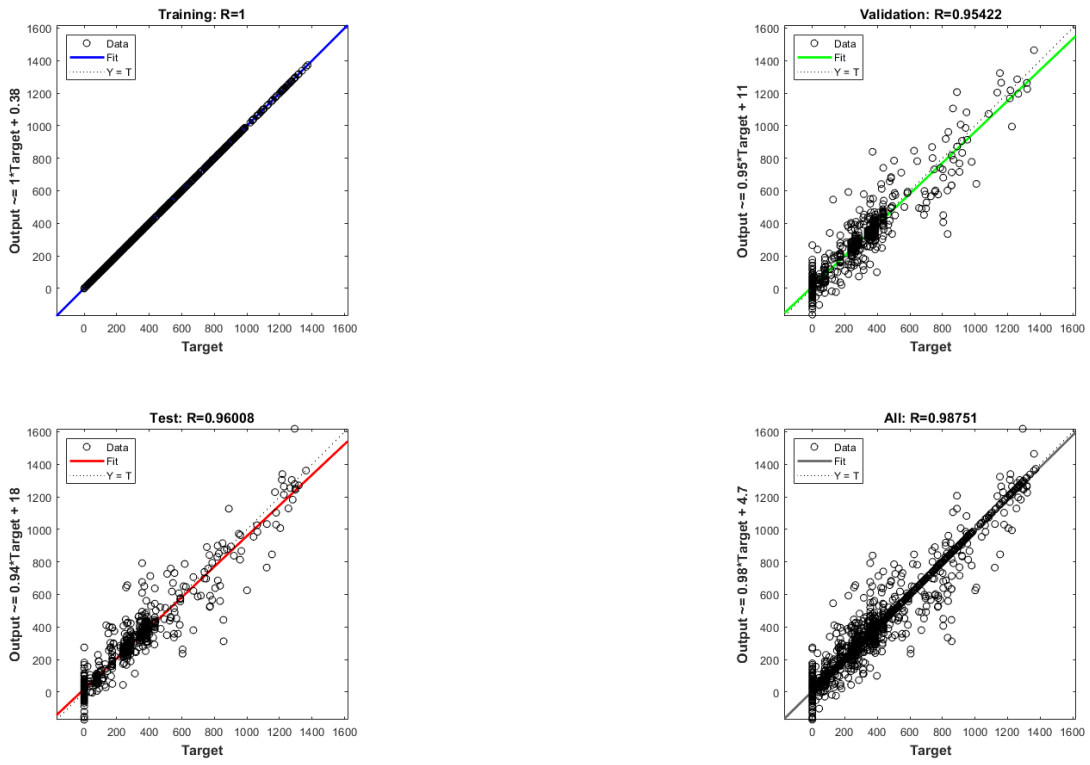


Figure 4.7: ANN regression plots

All four regression plots show correlations of $R > 0.95$ showing adequate consistent results. However, the validation and testing components have produced minor inconsistencies with obtaining the target values, producing regression values of 0.95422 and 0.96008 respectively. Due to the complexity of the original AC-OPF model with the application of a medium to large scale system, it is very difficult to obtain highly credible results without extreme computational expense. Therefore, the values shown in Figure 4.7 are considered adequate and credible with opportunities for improvement.

A more in depth approach to comparing the two simulation results is displayed in Figure 4.8 showing the two respective model targets in both a histogram and a CDF. A histogram is useful as it presents the most densely populated output frequencies of the data as compares all the outputs within bounds of the input data, providing a comparison of both the original model and the surrogate. The CDF is also useful as it provides a more precise way to communicate the performance of the surrogate model compared to the original model.

The figures below present the ANN performance on both a histogram and a CDF;

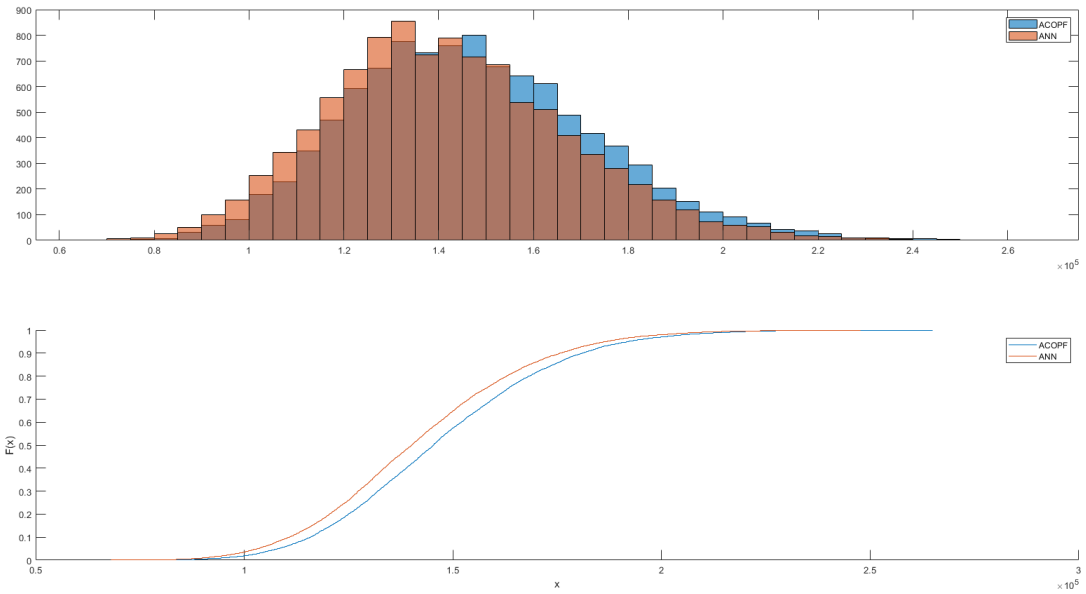


Figure 4.8: Monte Carlo simulation comparison of AC-OPF and surrogate

The histogram and CDF show congruence within the general pattern of their respective simulation. However, inconsistencies are visible with the ANN diverging from the AC-OPF results as the simulation progresses. This divergence continues throughout the heart of the simulation, and slowly converges towards the end of the simulation. This reflects on the regression results as displayed on Figure 4.7 and this divergence can be minimised by applying more nodes and hidden layers to the ANN architecture for a more in-depth simulation. Further work can be carried out in order to improve the consistencies of these results.

4.6 Chapter Summary

This chapter presents the fundamental theory of optimal power flow, the applications and implementation of various optimal power flow techniques, and a real-world case study to test these ideas. This chapter also implements a weather model to add to the contingency analysis of the power flow model incorporating for the loss of performance function and adds both high winds and lightning strikes as input parameters for this chosen model.

Both the DC-OPF and AC-OPF power flow techniques have been successfully tested on the 29 node 50 link GB network. The resilience performance index has been applied

to all three techniques in the form of relative energy resupplied. The results show that the computational expense is significantly higher in the AC-OPF model as compared to the DC-OPF model and therefore fewer Monte Carlo simulations can be applied for feasible computation. This results in a much greater amount of uncertainty within the restoration function for the AC-OPF model. It is however important to note that DC-OPF negates reactive power and therefore provides less accurate results in the precise case when applied to a network of this nature.

The two respective surrogate models, the LODF and ANN approaches have been trialed out and have proven to be efficient and computationally less expensive techniques for both the DC-OPF and AC-OPF. The LODF implementation shows consistency with the DC-OPF and has proved to be significantly less computationally expense. The ANN model has also been successful to act as a meta-model for the AC-OPF and shows adequate but consistent results for a network tested at this magnitude. There is however possibilities for improvement regarding this model as the network's architecture can be deepened with the burden of computational expense.

Chapter 5

A Failure-based CFD Model to Quantify Resilience in Gas Pipelines

5.1 Introduction

The gas network is a system designed to efficiently transport gas though a series of pipelines to achieve an industrial goal. The demand for natural gas is ever increasing with global energy inflation and the demand for natural gas alone is estimated to reach 200 quadrillion BTU by 2035 [129] making natural gas the fastest increasing source of energy. The Deepwater Horizon oil spill mentioned in Section 1.2.3 is an example of a disaster with severe consequences that has occurred as a result of a failing gas pipeline, leading to a leak and eventually an explosion.

The structure of a gas pipeline system is composed of various sub-components and when compiled together compiles the whole gas pipeline system into a single entity. This network is unique in behavior due to the physical and chemical properties of natural gas which provide additional constraints when assessing the nature of safety, risks and uncertainties of gas pipeline failure. The flow rate of natural gas is dictated through the gas pressure which is prone to dropping due to the friction which is present in the pipe's inner surface with the gas. In order to tackle these drops in pressure, compression stations (CSs) are present in areas of the network suffering from low pressure and transmission system operators (TSOs) act to control pressure as required for the

demand nodes. Underground gas storage systems are also applied to increase the flexibility of the system in events such as disruptions, high demand times and increased gas congestion.
The motivation of this chapter is to provide an insight into resilience modelling of a natural gas pipeline. This chapter aims to present a model for a natural gas pipeline as represented by nodes and links and perform computational fluid dynamic (CFD) equations to assess the resilience profile. The parameters of resilience assessed in this chapter are the network's robustness and the ability for the network to recover in post-disaster state.

5.1.1 The Proposed Approach

The proposed approach applies an integrated model to a selected gas pipeline network accounting for performance factors with respect to mass flow rates to quantify the supply, demand and uncertainty profiles within the gas pipeline system. The primary model is taken from Marino and Zio [75], and has been innovated with the addition of works to deduce the gas flow rate profile from momentum heat and mass transfer (MHMT) equations as expressed in Su et al. [74].
This approach aims to combine mathematical modelling and the complex reality of physical critical infrastructures when applied to gas pipelines. The approach taken follows a set of steps to obtain the various resilience profiles, listed as follows;

1. Modelling the network - The topology of the gas pipeline network is deduced through the construction of a DAG and the network's attributes are applied for the simulation of system.

2. Modelling the failures - The physical failures to the system once a disaster has occurred are modelled to understand the consequences of performance loss.

3. Modelling cybernetic failures - The complexity of pressure integrity of the system is modelled to understand the consequences of worst case scenarios.

4. Resilience computation - The robustness and recovery is measured as an imprecise function to quantify the resilience profile under the system's uncertainty.

5.2 Theoretical Background

When dealing with the nature of a general gas pipeline, it is crucial to underline the fundamental theory of CFD equations to compute a performance profile for mass flow. Works include full MHMT equations used from the literature such as the conservation of mass, conservation energy, conservation of momentum and the ideal gas law.
Figure 5.1 displays a general gas pipeline with the respective equations highlighting the nature of mass flow with all the theory as established by Su. et al. [74];

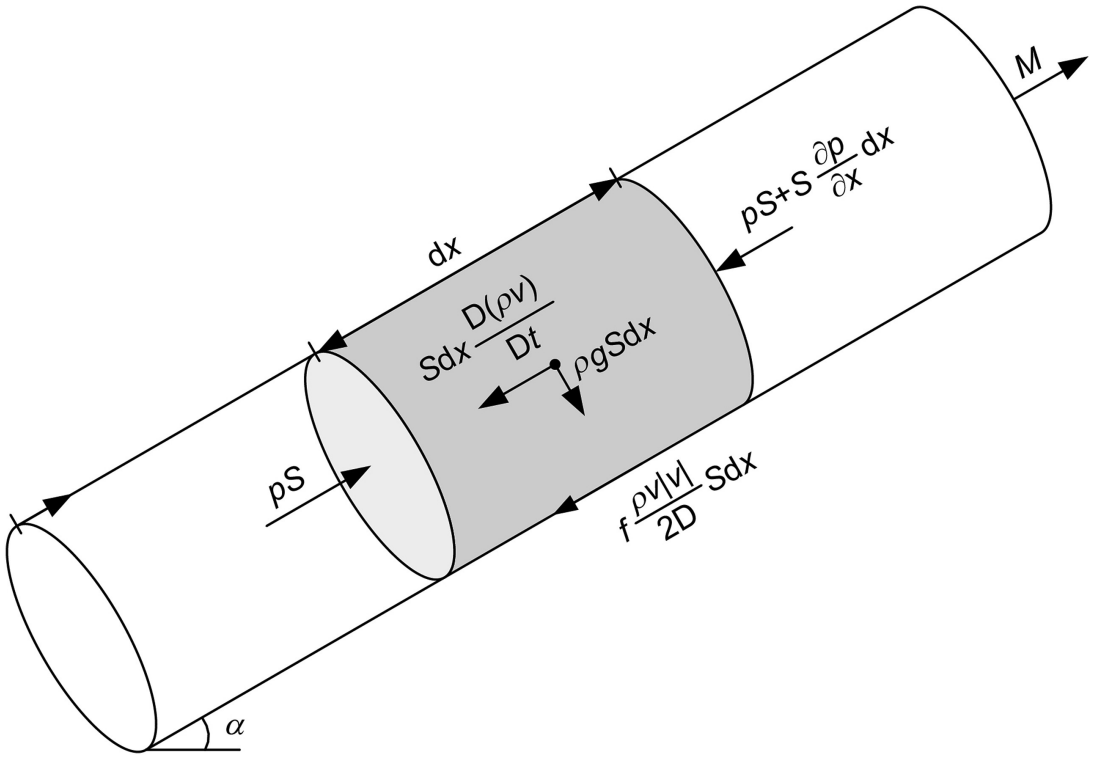


Figure 5.1: A general gas pipeline

In this model, the pipeline is represented by dx and is considered infinitesimal in length, cross-sectional area is represented as S and the pipeline diameter is represented as D. ρ represents the fluid density, p represents pipeline pressure, v represents the pipeline velocity and g represents the velocity of gravity. The assumption included in this model is that the gas flow properties are averaged over the cross-sectional area, S.
The laws that have been applied to this model produce various partial differential equations as mentioned on succeeding subsections. These equations derived are used to be

implemented to a real world gas pipeline system in order to understand the attributes of fluid flow in a topological manner.

5.2.1 General Equations for the Gas Pipeline

The continuity equation is denoted as;

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = 0 \tag{5.1}$$

where $\partial\rho$ represents the density partial differential, ∂t represents the time partial differential and ∂x represents the partial profile differential.

The momentum equation is denoted as;

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho v^2)}{\partial x} + \frac{\partial p}{\partial x} + \frac{f\rho v|v|}{2D} + \rho g \sin a = 0 \tag{5.2}$$

where f represents the friction factor and a represents the angle of the pipeline.

This equation is built up of four terms in respective order of the inertia term, the convective term, the pressure force term and the gravity term.
In order to calculate f, the friction factor, the Coolebrook-White correlation is applied;

$$\frac{1}{\sqrt{f}} = -2\log_{10}\left(\frac{2.51}{Re\sqrt{f}} + \frac{r}{3.71D}\right) \tag{5.3}$$

where r represents the absolute pipe roughness and Re represents the Reynolds number which is computed as;

$$Re = \frac{\rho v D}{\mu} \tag{5.4}$$

where μ represents the dynamic viscosity.

However, an approximation for the friction factor, f is applied in the case of turbulent flow;

$$f = \left[2\log_{10}\left(\frac{4.518}{Re}\,log_{10}\left(\frac{Re}{7}\right) + \frac{r}{3.71D}\right)\right]^{-2} \tag{5.5}$$

The energy equation is denoted as;

$$\frac{\partial}{\partial t}\left[\left(C_vT + \frac{1}{2}v^2\right)\rho S\right] + \frac{\partial}{\partial x}\left[\left(C_vT + \frac{p}{\rho} + \frac{1}{2}v^2\right)\rho S\right] + \rho v S g \sin\alpha = \dot{\Omega} \tag{5.6}$$

where C_v represents the specific heat capacity at constant volume, T represents the temperate in K and $\dot{\Omega}$ represents the energy flux.

The state equation is denoted as;

$$\frac{p}{\rho} = ZRT \tag{5.7}$$

where Z represents the comprehensibility factor, R represents the gas constant.

As a consequence of natural gas pipelines containing many nodes and links within the topology of the system, the proposed models in this chapter take assumptions to simplify the above equations which neglect certain terms to maximise computational expense. These assumptions are listed below [130];

1. The flow of gas is not affected by the temperature of the gas pipeline and is considered a negligible factor and is therefore assumed as constant. The temperature is equal to the ambient temperature and remains constant throughout the simulation.

2. For Equation 5.2, the convective term is considered negligible. All other terms in this equation are not considered to be negligible and play a role on the system's MHMT characteristics.

5.2.2 Nodal Modelling

The mass conservation law is applied to develop a dynamic model. The momentum equation and the state equation are simplified in the model;

$$\frac{\partial p}{\partial t} = -\frac{\rho c^2}{S}\frac{\partial Q}{\partial x} \tag{5.8}$$

where ∂p represents the pressure partial differential, c represents the speed of sound and ∂Q represents the volumetric flow rate partial differential.

This equation provides the correlation between pressure change and mass flow. Δx is assumed to be 0 at any j-th node and therefore the mass flow equation transforms into;

$$dQ_j = \frac{\rho c^2}{\Sigma_{n=1}^{k} S_{j,n}\Delta x_{j,n}}\Sigma_{n=1}^{k} Q_{j,n} - L_j \tag{5.9}$$

where $Q_{j,n}$ represents gas flow from pipeline n to node j and is an absolute value when gas is flowing in this direction, $S_{j,n}$ represents the cross sectional area from pipeline n to node j and L_j represents the transfer of gas when the node represents the junction $L_j = 0$.

These two equations are combined to provide the dynamic model of node j;

$$\frac{dp_j}{dt} = \frac{\rho c^2}{\Sigma_{n=1}^{k} S_{j,n}\Delta x_{j,n}}\Sigma_{n=1}^{k} Q_{j,n} - L_j \tag{5.10}$$

Equation 5.10 is further implemented in the pipeline dynamic model in Section 5.2.3;

$$\frac{d(p_j - p_{j0})}{dt} = \frac{\rho c^2}{\Sigma_{n=1}^{k} S_{j,n}\Delta x_{j,n}}\Sigma_{n=1}^{k}(Q_{j,n} - Q_{j,n0}) - (L_j - L_{j0}) \tag{5.11}$$

where p_{j0}, $Q_{j,n0}$ and L_{j0} represent the respective variables at steady state.

This finally yields the equation of model as;

$$\frac{d\Delta p_j}{dt} = \frac{\rho c^2}{\Sigma_{n=1}^{k} S_{j,n}\Delta x_{j,n}}\Sigma_{n=1}^{k}\Delta Q_{j,n} - \Delta L_j \tag{5.12}$$

The nodes represent the relevant junctions, suppliers or demand sites as characterised for the given system. The suppliers are also divided into three sub-types consisting of gas plants, LNG terminals and underground storage systems.

5.2.3 Pipeline Modelling

Only mass and momentum equations are considered for the pipeline modelling. Kirchhoff's first law at pipe nodes is negated due to the increase in computational cost of implementing transient pipeline models. Despite this assumption, the model is able to maintain a suitable level of accuracy and reduces the scale of the model enabling larger networks to be analysed. Such networks are applied when working with real-world systems.
Equations 5.2, 5.3 and 5.4 are simplified as;

$$\frac{\partial Q}{\partial t} = -S\frac{\partial p}{\rho \partial x} - \frac{f_c \rho c^2}{2\mu_c^2 D S_p} Q|Q| - \frac{gSsin(\alpha)}{\rho c^2} p \tag{5.13}$$

Where S_p represents the cross sectional area at constant pressure.

This equation is transformed into a discrete state using the finite differential approach;

$$\frac{dQ_i}{dt} = -S\frac{p_k - p_r}{2\rho \Delta x} - \frac{f_c \rho c^2}{2\mu_c^2 D S_{\overline{p_i}}} Q_i|Q_i| - \frac{gSsin(\alpha)}{\rho c^2} \overline{p_i} \tag{5.14}$$

where $\overline{p_i}$ is the mean pressure with respects to pipe i and is computed as as;

$$\overline{p_i} = \frac{2}{3}\frac{p_k^2 + p_k p_r + p_r^2}{p_k + p_r} \tag{5.15}$$

This mean pressure, $\overline{p_i}$ is further transformed into a linear equation using Taylor's formula;

$$\frac{d\Delta Q_i}{dt} = \left.\frac{\partial F}{\partial Q_i}\right|_{(Q_{i0}, p_{k0}, p_{r0})} \Delta Q_i + \left.\frac{\partial F}{\partial p_k}\right|_{(Q_{i0}, p_{k0}, p_{r0})} \Delta p_k + \left.\frac{\partial F}{\partial p_r}\right|_{(Q_{i0}, p_{k0}, p_{r0})} \Delta p_r \tag{5.16}$$

with;

$$F(Q_i, p_k, p_r) = -S\frac{p_k - p_r}{\rho\Delta x} - \frac{f_e\rho^2c^2}{2\rho\mu_e^2 DS\left(\frac{2}{3}\frac{p_k^2+p_kp_r+p_r^2}{p_k+p_r}\right)}Q_i|Q_i| - \frac{gSsin(\alpha)}{\rho c^2}\left(\frac{2}{3}\frac{p_k^2+p_kp_r+p_r^2}{p_k+p_r}\right) \tag{5.17}$$

since $\left.\frac{\partial F}{\partial Q_i}\right|_{(Q_{i0},p_{k0},p_{r0})}$, $\left.\frac{\partial F}{\partial p_k}\right|_{(Q_{i0},p_{k0},p_{r0})}$ and $\left.\frac{\partial F}{\partial p_r}\right|_{(Q_{i0},p_{k0},p_{r0})}$ are constant values, the final dynamic pipeline equations result in;

$$\frac{d\Delta Q_i}{dt} = K_{qi}\Delta Q_i + K_{pk}\Delta_{pk} + K_{pr}\Delta_{pr} \tag{5.18}$$

$$K_{qi} = \left.\frac{\partial F}{\partial Q_i}\right|_{(Q_{i0},p_{k0},p_{r0})} \tag{5.19}$$

$$K_{pk} = \left.\frac{\partial F}{\partial p_k}\right|_{(Q_{i0},p_{k0},p_{r0})} \tag{5.20}$$

$$K_{pr} = \left.\frac{\partial F}{\partial p_r}\right|_{(Q_{i0},p_{k0},p_{r0})} \tag{5.21}$$

5.2.4 Control System Modelling

The demands of the natural gas pipeline are dictated through the adjustments within the various components of the system. These components include the gas suppliers (GS), demand sites (D), regulation stations (RS), compressor stations (CS) and underground storage (UGS). For example, the RS and CS are defined by modelling the flow rate, inlet and outlet pressures with respects to the required parameters. These additional independent linear equations are crucial for modelling a system of this nature as the original nodal and pipeline modelling equations stated in Sections 5.2.2 and 5.2.3 alone are insufficient. This is due to the number of unknowns containing a higher quantity of degrees of freedom than the number of set equations.

The application of the linear equations are displayed on Table 5.1 and satisfies this constraint to negate the excess degrees of freedom;

Type	Control mode	Control equation	Constraints
GS	Flow control	$\frac{dp_{GS}}{dt} = \sum_{n=1}^{k} Q_{GS,n} - L_{GS}$	$\lvert L_{GS}\rvert \leq Max\ supply$
	Pressure control	$\sum_{n=1}^{k} Q_{GS,n} - L_{GS} = 0$	$p \leq p_{max}$
	Inactive	$L_S = 0$	
D	Flow control	$\frac{dp_D}{dt} = \sum_{n=1}^{k} Q_{D,n} - L_D$	$\lvert L_D\rvert \geq Min\ demand$
	Pressure control	$\sum_{n=1}^{k} Q_{D,n} - L_D = 0$	$p \geq p_{min}$
	Inactive	$L_D = 0$	
CS/RS	Pressure ratio	$P_{out} - \varphi P_{in} = 0$	Max outlet pressure
	Outlet pressure	$P_{out} - P_{out\_set} = 0$	Min inlet pressure
	Inlet pressure	$P_{in} - P_{in\_set} = 0$	Max flow rate
	Flow control	$Q - Q_{set} = 0$	Max pressure ratio
	Bypass	$P_{in} - P_{out} = 0$	
	Inactive	$Q = 0$	
UGS	Flow control	$\frac{dp_{UGS}}{dt} = \sum_{n=1}^{k} Q_{UGS,n} - L_{UGS} - L_{UGS\_withset}$	$\lvert L_{UGS}\rvert \geq Min\ demand$
	Pressure control	$\sum_{n=1}^{k} Q_{UGS,n} - L_{UGS} = 0$	$p \geq p_{min}$
	Inactive	$L_{UGS} = 0$	

Table 5.1: Control modes with their respective equations

The control equations are applied to the respective types of nodes and are maintained by adjusting the parameters of the constraints. For instance, the CS and RS valves are maintained through adjusting the parameters of outlet pressure, inlet pressure, flow rate and pressure ratio.

5.3 An Imprecise Failure Based Resilience Model

The resilience based model is divided between the failure stage, the stagnation stage and the recovery stage, in which all three phases have been modelled using the required equations as stated in relevant literature. The innovation of this literature is the addition of respective uncertainties that are trialed out during the input and output stages of the simulation. The approach is taken from Marino and Zio [75].

5.3.1 Probabilistic Modelling of Failure Scenarios

The gas pipeline in this model is subject to varying events contributing to the failure of the system. During the degradation stage of the resilience model, the mass flow into the respective nodes declines and therefore the overall performance disintegrates.

5.3.1.1 Pipeline Failure

The causes of pipeline failure includes leakage within the system, pipeline rupture or inadequate operation of the pipeline. European Gas Pipeline Incident Data Group (EGIG) [131] states that the mean failure rate of a European Gas pipeline is listed at $3.5 \times 10^5 (km \cdot y)^{-1}$.

5.3.1.2 Compressor Station Failure

The cause of compressor station failures are a result of the internal factors within the compressor. A decreased performance of pipeline capacity occurs during the event of compressor failure and this is believed to reduce the pressure within the pipelines by approximately 20% [132].

5.3.1.3 Gas Storage Failure

The primary trigger for gas storage failures is constituted in the build-up of facility failure and withdrawal. In the event of gas failure, it is assumed that the pipeline function has also failed [133]. Ouyang estimates that the annual failure rate of the gas storage system is at 10% [134].

5.3.1.4 Liquefied Natural Gas Terminal Failure

During the event of a Liquefied Natural Gas (LNG) terminal failure, it is also assumed that the pipeline function has failed as the supply capacity of the required gas has declined. It is estimated that the annual failure rate for this event is 15% [135].

5.3.2 Cybernetic Failure Modelling

Natural gas pipelines are vulnerable to the threat of cyber-attacks [136], and most notably pressure related cyber-attacks. This is most commonly carried out by increasing the pressure throughout the pipeline system. Consequences of overpressure can lead to economic loss and environmental disasters as demonstrated in Section 1.2.3.

5.3.2.1 Compressor Station Pressure Model

The model used for the pressure model as applied to the compressor station is presented in Figure 5.2 [75]. The types of curves represent various linear or logarithmic based responses.

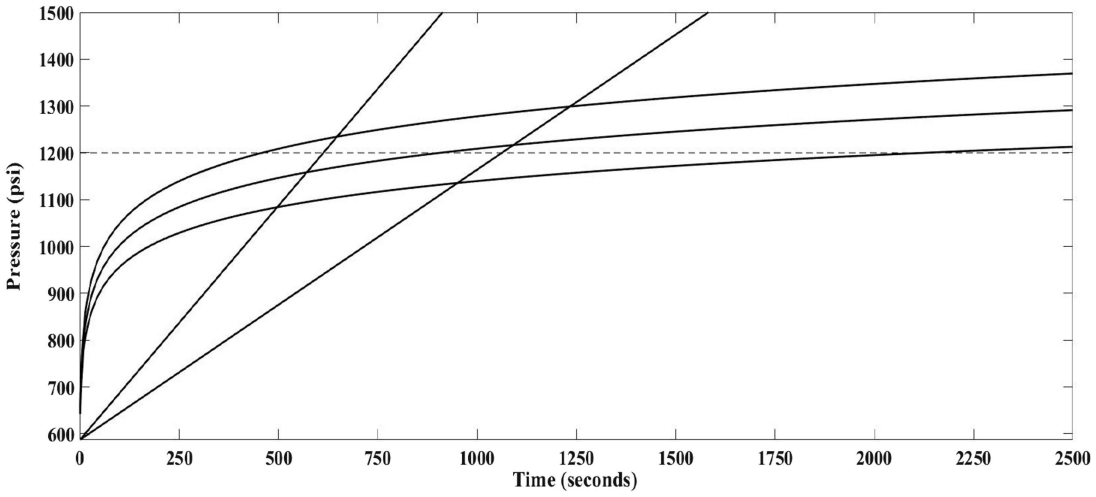


Figure 5.2: Various pressure increasing functions for the internal pipeline

5.3.2.2 Overpressure

There are various triggers which induce the overpressure of a natural gas pipeline system. The pipeline cannot deduce the irregularity of the system's dysfunction until the delivery node is met with an overpressure. The delivery time of the gas from one node to the next node across the pipeline is computed in Pipe Flow Expert after the velocity of flow in the pipe is computed. It is assumed that the pipeline fails once the pressure overrides a maximum threshold of pressure known as the maximum allowable overpressure (MAOP).

The three main common causes of this overpressure are spilling, jet fire and vapour cloud explosion.

In the case of spilling, Casal [137] states that the mass flow rate is obtained from the following equation;

$$\dot{m}_{hole} = A_{or} C_D P_{cont} \Psi \left(\gamma \left(\frac{2}{\gamma + 1} \right)^{\left(\frac{\gamma+1}{\gamma-1} \right)} \right)^{-2} \tag{5.22}$$

where $\dot{m}_{hole}$ represents the mass flow rate in $kg\, s^{-1}$, γ represents the isentropic coefficient at 1.4, C_D represents charge coefficient as a dimensionless value, A_{or} represents the cross-sectional area of the orifice in m^2, Z represents the gas compressibility factor given the conditions of constant pressure P_{cont} in Pa and constant temperature T_{cont} in K of the pipeline and Ψ represents the dimensionless factor which is influenced by the gas velocity. Natural gas as applied in this chapter is assumed to have parameters of $Z = 1$ and $\Psi = 1$.

In the event of a jet fire, irradiation values are obtained as 2.5 $kW\, m^{-2}$ for equipment and 12.5 $kW\, m^{-2}$ for people [137]. It is also assumed that the spilling diameter is approximately 20% of the pipeline diameter.
Vapour cloud explosions result in a release of energy at a rapid pace and form this cloud due to the loss of containment of flammable gas. Overpressure is caused by the mechanical energy of the explosion event which is released into the atmosphere. In the event of a vapour cloud explosion, there is a delay in ignition, and therefore a mixture of air and fuel is prone to develop into a cloud. The assumption used here is that the time taken to produce this vapour cloud is the same time of detection of the MAOP on the pipeline.

5.3.3 Parameters for Resilience

The chosen parameter for system's performance is the actual gas flow rate in relation to its respective demand. The transient resilience function for this parameter is denoted as $\varphi(t)$. Figure 2.1 from Section 2.2.3 describes a typical three phase resilience curve

which is applied as a transient function. A similar model is implemented in this study to quantify the characteristics of a three phase resilience performance profile.

5.3.3.1 Robustness Model

Robustness has been defined in Section 2.2.1 as "The analysis of the strength and ability to withstand a given stress level without damaging the operation of the system". In this case, robustness refers to the capability of the gas pipeline system to transfer its respective mass flow rates given the event of of a specific failure. In certain cases, the failure affects multiple pipelines of the system which are blocked by remote control valves (RCV's) in order aid the process of maintaining adequate pressure to the given system [138]. The respective pipelines are non-functioning when the threshold limit of depressurisation occurs, in which the failure detection time is recorded. In this case, when multiple pipelines fail, the average failure detection time is calculated.

The pipeline link(i, j) where $(i, j) \in \mathscr{I}$ is set as $\mathscr{I}$. The equation to compute the failure detection time for the respective pipelines from node i to node j is denoted as;

$$t_{det}(i, j) = \frac{K_{pr} \cdot P_{initial(i,j)}}{m} + mpd \tag{5.23}$$

where $P_{initial(i,j)}$ represents the initial pipeline pressure in psi for the link between nodes i and j, m represents the pressure degradation rate in the interior of the pipeline measured in $psi\, s^{-1}$, K_{pr} represents the promptness parameter which is set at 0.1 and mpd represents the maximum packet delay.

The pressure of the pipeline (i, j) is assumed to be the pressure at node j as an assumption to simplify the simulation. This assumption is automatically applied for the simulation software applied PipeFlow Expert to reduce excess computational expense and enabling the analysis of large pipeline networks.

The degradation function $g_{deg}(t)$ is denoted as;

$$g_{deg}(t) = \dot{m}f_{t=0} \cdot (-\frac{t}{\bar{t}_{det}}) + 1 \tag{5.24}$$

where t represents the time in s, $\dot{m}f_{t=0}$ is the maximum flow rate of the system prior to failure and $\bar{t}_{det}$ is the average detection time in s.

The robustness capacity, G_{RobCap} is computed by integrating $g_{deg}(t)$;

$$G_{RobCap}(\bar{t}_{det}) = \int_0^{\bar{t}_{det}} g_{deg}(t)\, dt \tag{5.25}$$

A linear based function to simplify of this equation is proposed [139];

$$\frac{dP}{dt} = m \tag{5.26}$$

An underpressure situation is detected when the pipeline pressure falls below a certain threshold as recorded by the remote terminal unit (RTU).

The general robustness performance function is derived [140];

$$Rob = 1 - \frac{\varphi(t_0) \cdot (t_d - t_e) - \int_{t_e}^{t_d} \varphi(t)(dt)}{\varphi(t_0) \cdot (t_d - t_e)} \tag{5.27}$$

This equation deduces the percentage of gas that is delivered in a post-failure event, with the maximum percentage representing the network's full performance at 100% .

5.3.3.2 Recovery Model

Whilst robustness deals with resilience during disaster as a quantification of potential damage to the system, rapidity and resourcefulness are indicators of recovery to the system in a post-disaster phase. Rapidity, as defined in Section 2.2.1 is referred to as "the time that a system takes to recover". The pipeline system recovers with the aid of repair until the performance has been fully restored.

The recovery function is denoted as [136];

$$P_{ij}(t) = p_{init(i,j)} + (p_{init(i,j)} - p_{fin(i,j)}) \cdot (1 - exp(-b_{i,j} \cdot t)) \tag{5.28}$$

where $P_{i,j}(t)$ represents the recovery function in s, $p_{init(i,j)}$ represents the initial real pressure of the pipeline, $p_{fin(i,j)}$ represents the final real pressure of the pipeline post-disaster in psi and $b_{i,j}$ represents the semi-empirical parameter which defines the recovery speed of the pipeline.

The lack of data with this equation falls within obtaining values for $b_{i,j}$ as without this, the estimation of the pressure response for pipeline (i, j) is unobtainable. These values are obtained from the relation between the time constant τ being inversely proportional to $b_{i,j}$.

$$b_{i,j} = \frac{1}{\tau} \tag{5.29}$$

The value of $b_{i,j}$ is related to the set point pressure $P_{init(i,j)}$. The higher the pressure of the pipeline is, the greater the value of τ.
This phenomena is graphically represented in Figure 5.3 [75];

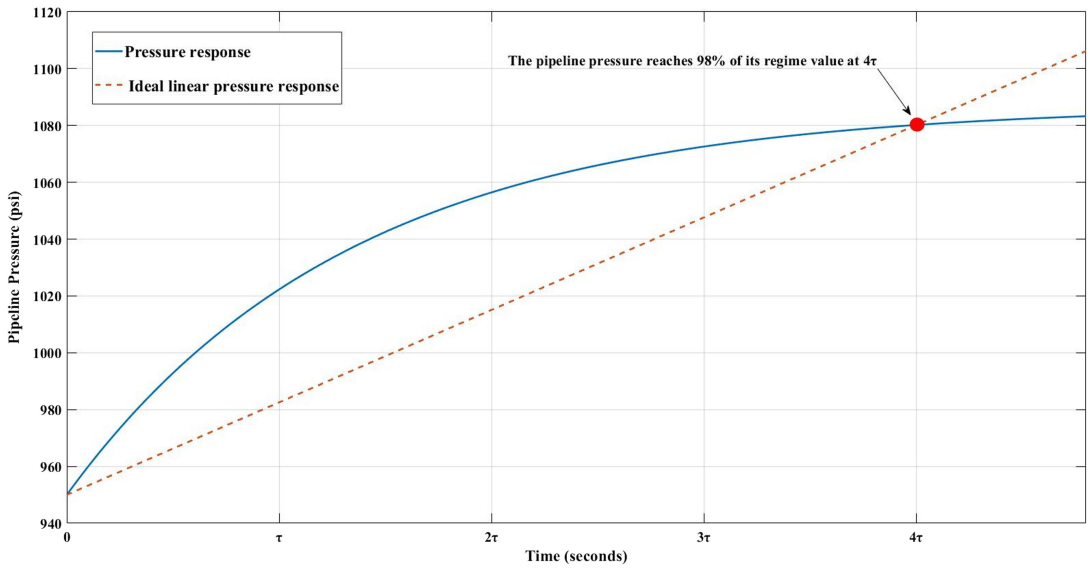


Figure 5.3: Gas pipeline pressure responses

Figure 5.3 displays the properties of the pressure response from Equation 5.28 alongside the ideal linear pressure response. The graphs intersect at 98% of the set-point pressure mark at a time constant τ at 4. This is the minimum time it is assumed that the gas has been transferred from node i to node j and therefore $t_{dest} \geq 4\tau$. This means that pipelines which exhibit a lower set point pressure and contain pipes which

are shorter in length achieve steady state at a quicker rate and therefore τ will be a lower value whilst $\frac{dP}{dt}$ will be greater. It can be assumed that the velocity of the gas flowing in the pipeline is less than 30 $m\,s^{-1}$ during the post-disaster phase [141].
t_{dest} is computed as;

$$t_{dest} = \frac{L}{v_{gas}} \tag{5.30}$$

From this information, $b_{i,j}$ is approximated;

$$b_{i,j} \sim \frac{p_{init(i,j)} - p_{fin(i,j)}}{4\tau} \tag{5.31}$$

A pipeline in a failed state recovers its transport capacity following the equation;

$$Q_{ij}(t) = Q_{ij,0} \cdot [1 - exp(-b_{ij} \cdot t)] \tag{5.32}$$

$Q_{ij}(t)$ represents the volumetric flow rate capacity for the pipeline from node i to node j during time t and $Q_{ij,0}$ represents the volumetric flow rate capacity at $t = 0$.

This equation is then rearranged to;

$$Q_{ij}(t) = D_{ij}^{\delta} \cdot \left\{ 1 - exp\left[-\left(\frac{v_{gas} \cdot (K_{pr} \cdot P_{init(i,j)} + mpd \cdot m}{L_{ij}} \right) \cdot t \right] \right\} \tag{5.33}$$

where D_{ij} represents the diameter of the pipeline (i, j) in m, δ represents the non-dimensional conversion coefficient and v_{gas} represents the velocity of the recovery gas. This is the quantity of gas required for the network to achieve sufficient nominal performance.

As with the robustness model, the area below the recovery curve is also integrated to obtain the network recovery capacity, G_{RecCap};

$$G_{RecCap}(t_{rec}) = \int_0^{t_{rec}} g_{rec}(Q_{ij}(t), t)dt \tag{5.34}$$

where t represents the time in s, g_{rec} represents the network recovery curve and t_{rec} represents the recovery time for the observation in s.

Finally, the network recovery performance index is denoted as;

$$Rec = 1 - \frac{\varphi(t_f) \cdot (t_f - t_e) - \int_{t_e}^{t_f} \varphi(t)dt}{\varphi(t_0) \cdot (t_d - t_e)} \tag{5.35}$$

Rec represents the performance index with respects to the proportion of gas that has been recovered during t_{rec}.

5.4 Implementation of the Technique

The implementation of this technique has been applied in MATLAB 2021b and the data is collected, trialed out and computed from Pipe Flow Expert.

5.4.1 Implementing the Physical Model

The physical model is designed in the form of a direct weighted graph with nodes representing demand stations such as LNG terminals, natural gas storage, and compressor stations, and links representing pipelines.
A steady state thermal hydraulic analysis has been implemented in PipeFlow Expert to obtain the data regarding the pressure of the various nodes and the vector based direction of the pipe flow. The input data that is required from literature includes the customer demand values for gas, dimensions of the pipelines and the pressure of the gas in the source nodes.

5.4.1.1 The Capacity Model

The capacity weighting of the links are computed with the equation that relates Q, the pipeline volumetric flow rate capacity and D, the pipeline diameter.

$$w = QD^{-\delta} \tag{5.36}$$

where Q represents the estimated volumetric flow rate of the pipeline in $MCM\,h^{-1}$, δ represents the constant conversion coefficient at 2.59 and D represents the diameter of the pipeline in m.

5.4.1.2 Capacity Calculation

During transmission of gas from one node to the next node, the capacity of the gas pipeline is computed by applying the Ford-Fulkerson maximum flow algorithm. This method is used to calculate the supply capacity and the performance of the network by applying the maximum possible flow as the metric.
The DAG, $G(N, L)$ is presented with N denoting the nodes and L denoting the links of the network. The links $L \in (i, j)$ represent positive values of the flow variable f_{ij} with the capacity of each link represented as u_{ij}.
The expression s represents the source node and $t(s, t \in V)$ represents the terminal node of the network's links. The aim of the maximum flow algorithm is to transfer the maximum capable flow rate from s to t satisfying the constraint $f_{ij} \leq u_{ij}$.

The maximum flow algorithm as applied to this case is expressed as;

$$\begin{aligned} &max \qquad f_{st} \\ &with \qquad \textstyle\sum_{(i,j)\in L} f_{ij} - \sum_{(i,j)\in L} f_{ji} = 0 \quad j \in N\{s,t\} \\ &0 \leq f_{ij} \leq u_{ij}, \qquad \forall (i,j) \in L \end{aligned} \tag{5.37}$$

For simulations with multiple source and terminal nodes, a supersource and superterminal node is applied. However, this chapter focuses on a single source node and these two nodes are selected by the user.

5.4.2 Uncertainty in Gas Pipeline Modelling

The approach taken to apply epistemic and aleatory uncertainty within this technique applies both imprecise inputs to the original model and confidence bounds to the final

outputs. Both of these techniques are applied to deduce the maximum performance possibility and also the minimum performance possibility for all three stages of resilience within the simulation procedure. These include the robustness and recovery function alongside the stagnation phase in between.

5.4.2.1 Epistemic Uncertainty

The form of epistemic uncertainty applied in this model is applied via estimation regarding the volumetric output from the source nodes and the volumetric input to the demand nodes. The imprecise values regarding the supply and demand capacity measured in $MCM\,h^{-1}$ are presented in Tables 5.2 and 5.3. Both the upper bound inputs and the lower bound inputs have been applied to the respective model and the two scenarios of outputs are recorded with respects to epistemic uncertainty.

5.4.2.2 Aleatory Uncertainty

The confidence bounds for the output distribution is modelled through Gaussian uncertainty distributions with a confidence bounds of 10%. The equation for the Gaussian applied to the output of this model is denoted as follows;

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}} \left(\frac{x-\mu}{\sigma} \right) \tag{5.38}$$

where μ represents the mean of the output result and σ represents the standard deviation with respects to 10% confidence. This output uncertainty is applied after the epistemic uncertainty is obtained and adds additional uncertainty bounds for the resilience function. The final results display the robustness and recovery functions with a combination of both epistemic and aleatory uncertainty.

5.4.3 The Dynamic Model

The model is presented as an algebraic differential equation dynamic model. An implicit differential model has been applied with variable-step and variable-order.
The simulation method of the dynamic model is displayed in Algorithm 5;

Algorithm 5 Dynamic flow model algorithm

1: **procedure** v (*Dynamic flow model based on CFD*)
2: Compute initial condition
3: Generate dt and dx using discretisation
4: Initiate time integration for $t_{n+1} = t_n + dt$
5: Generate differential equations for each pipeline from Section 5.2.1
6: Verify control modes from Section 5.2.4
7: Compute equations from 5.2.2 and 5.2.3 with respects to the control modes
8: Iterate t to obtain an approximation for t_{n+1}
9: **if** constraints violated **then**
10: Perform re-iteration using limitation bound as initial boundary conditions
11: Go to (8)
12: **else** Compute results at $t = t_{n+1}$
13: **if** $t_n < t_{max}$ **then**
14: Set new boundary conditions for t_{n+1}
15: Go to (4)
16: **else** End simulation

5.4.4 Limitations to the Proposed Approach

The model contains sufficient detail regarding the input parameters and included assumptions. However, even with these assumptions, computational cost is still high in networks with large numbers of nodes. Additionally, the dynamic model uses iteration to quantify the flow properties of the system and provides an estimate. Despite the iteration being constantly re-iterated until the boundary flow conditions are satisfied, these are likely to provide results with some degree of uncertainty to output flow. The other limitation within this technique is within the failure scenarios as it is assumed that the failure scenarios can only occur independently without the consideration of mutual sources of failure that happen simultaneously. An expectation of the average reliability of all three failure scenarios are estimated, however this assumes that all three scenarios are equally likely to happen.

5.5 Case Study: Natural Gas Pipeline

In another paper, Su et al. [142] presents data for a natural gas pipeline consisting of 53 nodes and 68 links. The reliability analysis of this pipeline has already been tested for this system. The approach presented in this chapter expands this and applies resilience as described in Section 5.3.3. All the equations regarding the characteristics of the system have been applied from Section 5.2 to deduce the general behaviour of flow for the system.

The DAG for the natural gas pipeline has been constructed as an adjacency matrix on MATLAB 2021b and this has been represented in graphical format;

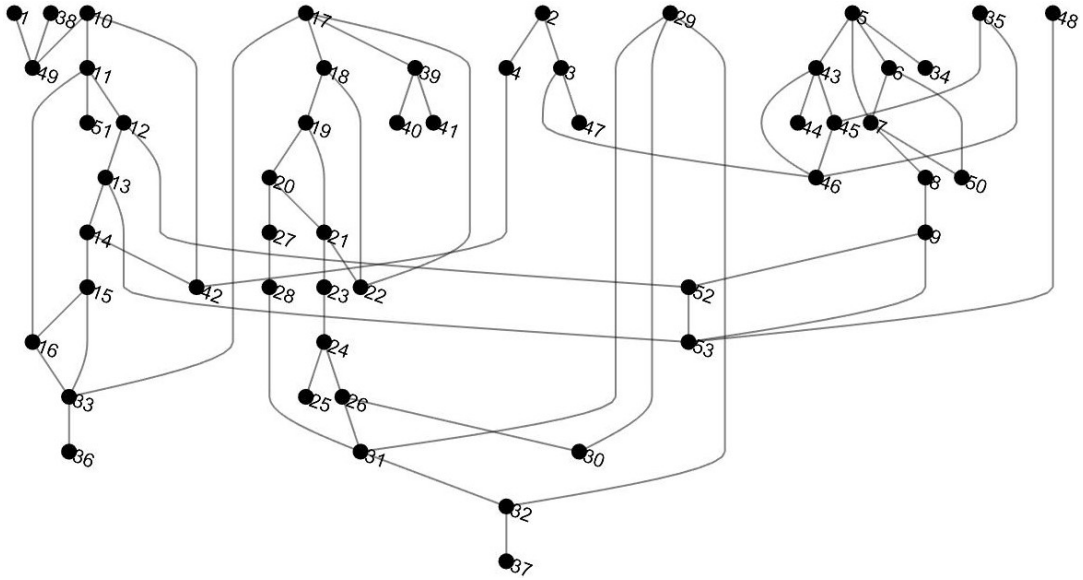


Figure 5.4: Topology of the natural gas pipeline

Nodes 1 and 53 are considered as fictitious nodes and are both supersource and supersink nodes respectively. The internal source nodes are listed on Table 5.2 and the internal demand nodes are listed on Table 5.3. It is assumed that all the pipelines are operating under negligible frictional loss factors and that the fluid loss within the nodes is also considered negligible.

The capacity model calculation from Equation 5.36 is applied and and the maximum flow algorithm from Equation 5.37 which utilises the Ford Fulkerson algorithm is also applied along with the constraint properties of the physical gas limits, node demand values and the capacities of the connected elements.

5.5.1 Input Data

The required data to successfully simulate the system is provided in different categories. Firstly, the data for the source nodes is required to estimate the maximum performance that these source nodes can provide to demand nodes. The expectation of demand nodes and their respective volumetric flow rates are also required. Both of these data sets are applied simultaneously to the model to ensure congruence within the supply and demand flow.
The other data-sets collected include collecting information regarding the failure scenarios and have been applied to the their desired equations as listed on Section 5.3.1 and the pressure response values for implementation on Figure 5.3 into the recovery model.

5.5.1.1 Source Node Volumetric Data

The first step is to obtain data regarding the capabilities of performance for the source nodes. The initial data required to simulate this system is the maximum output volumetric limit , measured in million cubic metres per hour ($MCM\,h^{-1}$). This is obtained from the basis of Su et al. [142] and is estimated with lower and upper bounds.
Table 5.2 displays the source node types with their respective imprecise volumetric properties;

Node	Type	$Q_{max}(MCM\,h^{-1})$
9	Storage	(85, 104)
10	Pipeline	(673, 789)
15	LNG terminal	(186, 303)
18	Pipeline	(500, 700)
50	LNG terminal	(5.4, 9.1)

Table 5.2: Volumetric properties for source nodes

5.5.1.2 Demand Node Volumetric Data

Along with source node data, the various demand nodes are also estimated and inputted into the simulation with their respective volumetric flow rate demands. The units for

this are maintained as the same units for the source nodes measured in $MCM\,h^{-1}$. Table 5.3 displays a list of the demand nodes with the respective volumetric flow rates as imprecise values;

Node	$Q_{dem}(MCM\,h^{-1})$		**Node**	$Q_{dem}(MCM\,h^{-1})$
4	(32.64, 37.48)		31	(17.53, 22.08)
5	(34.73, 39.96)		34	(17.52, 22.10)
7	(38.62, 41.87)		35	(23.78, 27.16)
12	(32.49, 37.75)		37	(19.42, 23.73)
16	(100.04, 109.75)		38	(37.33, 45.68)
17	(31.73, 40.97)		40	(28.78, 34.45)
20	(10.23, 13.95)		41	(44.25, 52.37)
24	(31.10, 39.71)		42	(35.27, 41.26)
25	(35.14, 42.36)		43	(30.05, 36.79)
26	(39.54, 47.83)		46	(10.58, 13.54)
28	(55.53, 66.48)		49	(26.73, 31.26)
29	(44.75, 52.09)		51	(21.04, 25.59)

Table 5.3: Volumetric properties for demand nodes

5.5.1.3 Node Pressure Data

The data for the pressure of the nodes is extracted from the simulation in PipeFlow Expert with the application of the following assumptions;

1. The source node (node 1) pressure is 1000 psi.
2. The type of gas within the pipelines is methane which at ambient temperature has as density of 0.7168 $kg\,m^{-3}$ and a dynamic viscosity of $10.9 \times 10^{-6} Pa\,s$.
3. The minimum volumetric flow rate within the pipelines is 11 $MCM\,h^{-1}$ and the maximum volumetric flow rate was found to be 1180 $MCM\,h^{-1}$.
4. The internal roughness of the carbon steel pipelines are 0.07 mm.

All of these assumptions have been accounted for and have been implemented into the Pipeflow Expert simulation. The pressures are given as precise values as multiple simulations need to be trialed out to contribute to epistemic uncertainty within the pipeline pressures which is computationally unfeasible for a network of this size.

The values of all the nodal pressures in the natural gas pipeline are displayed in Table 5.4;

Node	**Pressure** (psi)		**Node**	**Pressure** (psi)
1	1000		28	124
2	875		29	121
3	875		30	104
4	924		31	98
5	924		32	103
6	924		33	101
7	1000		34	127
8	1000		35	850
9	1000		36	850
10	945		37	265
11	944		38	126
12	944		39	99
13	1000		40	68
14	949		41	924
15	945		42	925
16	944		43	925
17	945		44	924
18	1000		45	886
19	1000		46	885
20	805		47	912
21	807		48	946
22	806		49	995
23	850		50	855
24	1000		51	920
25	839		52	954
26	102		53	955
27	100			

Table 5.4: Pressure properties for nodes

This data has been obtained from PipeFlow Expert and the bounds have been estimated.

5.5.1.4 Gas Pipeline Data

The volumetric flow rate data for the gas pipelines has been obtained from the original model used in reliability analysis from Su et. al [142] and has been converted into the units applied for the supply and demand nodes, $MCM\,h^{-1}$. This estimation is

deemed valid due to the similarities of the nature of the two topologies. The length of each pipeline is estimated and has been assigned a length from its starting node to its respective connecting node.

Table 5.5 shows the gas pipeline data with all the lengths of each pipeline in km and the respective volumetric flow rates as simulated for each pipeline;

Pipeline	$l(km)$	$Q_{cap}(MCM\,h^{-1})$		**Pipeline**	$l(km)$	$Q_{cap}(MCM\,h^{-1})$
1-49	10	744		17-33	25	1180
2-3	10	1180		17-39	15	68
2-4	5	291		18-19	15	121
3-46	15	291		18-22	5	291
3-47	10	411		19-20	10	291
4-42	20	291		19-21	10	1180
5-6	10	48		20-21	10	291
5-7	15	291		20-27	15	291
5-34	15	121		21-22	15	291
5-43	15	291		21-23	15	168
6-7	10	121		23-24	15	291
6-50	5	121		24-25	10	20
7-8	20	121		24-26	10	291
7-50	10	1180		26-30	10	1180
8-9	20	291		26-31	10	121
9-52	10	68		27-28	15	121
9-53	10	291		28-31	10	121
10-11	30	68		29-30	10	1180
10-42	10	32		29-31	10	1180
10-49	15	121		29-32	10	121
11-12	20	48		31-32	10	11
11-16	20	291		32-37	10	11
11-51	10	1180		33-36	10	24
12-13	5	1180		35-45	20	121
12-52	15	1180		35-46	10	68
13-14	15	1180		38-49	10	121
13-53	15	291		39-40	10	121
14-15	20	734		39-41	10	32
14-42	20	48		43-44	10	121
15-16	5	121		43-45	15	291
15-33	15	121		43-46	10	121
16-33	15	291		45-46	10	411
17-18	15	291		48-53	10	24
17-22	15	600		52-53	10	1180

Table 5.5: Pipeline properties

5.5.2 Failure Scenario Data

The data to characterise the robustness and recovery functions have been collected using the historical failure data from the reliability based study carried out for this original model.
The resilience of the network is trialed out under three different failure scenarios as listed on Table 5.6;

1. Gas Storage Failure at node 9.
2. Pipeline failures at nodes 15 and 18.
3. LNG terminal failures at nodes 15 and 50.

5.5.2.1 Robustness Data

The reliability data for the source nodes have been obtained from the reliability analysis carried out in the original study of this network. The type of source node failing is mapped with the maximum degredation volumetric flow rate Q_{deg}, the annual reliability, the annual mean time to fail (MTTF) and the maximum volumetric flow rate performance drop.
Table 5.6 displays the data for these proposed parameters;

Node	**Type**	Q_{deg} $(MCM\,h^{-1})$	**Reliability**	**MTTF (Y)**	$Max_{drop}(\%)$
9	Gas storage	67	0.9	9.5	58.6
10	Pipeline	238	0.999	198.2	11.8
15	LNG terminal	795	0.85	6.2	21.6
18	Pipeline	418	0.999	662.5	1.8
50	LNG terminal	23	0.85	6.2	21.6

Table 5.6: Robustness data for failure types

Failure scenarios 1 and 2 have been applied using Equation 5.28 to deduce the inner pipeline real pressure values after the RCVs have been blocked. This real pressure has been deduced using the simulation data in the case of scenario 3 as it is not viable to apply the pressure degredation within the LNG terminal nodes.

Equation 5.32 has been applied to each pipeline to compute the recovery speed parameter, and the recovery pressure for the delivery node has been computed using Equation 5.34 for each failure scenario.

5.5.2.2 Recovery Data

As mentioned in Section 5.3.1, the different types of node failure scenarios are defined, applied to the natural gas pipeline and are trialed out. These failures include pipeline failure, compressor station failure, gas storage failure and LNG terminal failure. The assumption is that the pressure degredation rate, m is 1 $psi\, s^{-1}$ is applied to obtain the simulation data for p_{init} and $\bar{t}_{det}$. The recovery speed parameter, b is also obtained and has been applied to the recovery function for the respective failure scenario.
Table 5.7 displays recovery data for the three types of failure applied to the simulation;

Failure type	**Nodes**	$-10\% p_{init}(psi)$	$\bar{t}_{det}(s)$	b
Gas storage	9	888	112	12.25
Pipeline	10, 18	940	N/A	0.01
LNG terminal	18, 50	842	131	0.02

Table 5.7: Recovery data for failure types

The recovery parameter, b is deduced from the simulation and is applied to the recovery model. It can be concluded that in the event of a gas storage failure, the recovery parameter is significantly higher than for the other two failure scenarios as the pipeline reaches the recovery threshold value at a much slower rate. The value for $\bar{t}_{det}$ for pipeline failure is considered negligible within this model and the recovery parameter for this scenario is 0.01.

5.5.3 Resilience Results

The results for the implementation of the resilience based model are split between the general resilience performance index, the graphical robustness function and the graphical recovery function. All three parts of the model are subject to both the epistemic and aleatory uncertainties as listed in Section 5.4.2. The overpressure data has also been recorded for the compressor stations by applying the Weikma method. This rate

of pressure increase has been tested out with various linear based and logarithmic-increase based recovery models. The time taken for spilling to occur and the quantity of overpressure is recorded as outputs to the model.

5.5.3.1 Resilience Performance Indexes

The three failure modes have been tested. The robustness (R1) performance index and the recovery (R2) performance index during the three different recovery time phases are obtained;

Failure type	**R1**	**R2 (60 s)**	**R2 (120 s)**	**R2 (180 s)**
Gas Storage	0.6548	0.9217	0.9684	0.9981
Pipeline	0	0.8975	0.9389	0.9738
LNG terminal	0.8473	0.8492	0.8981	0.9472

Table 5.8: Simulation reliability values

These resilience based performance indexes have been applied to the three phase resilience function. The assumption of this function is that the stagnation period lasts for a total of 60 s simulation time for all three of the failure scenarios. The simulation carried out regarding the robustness and recovery functions have been measured in intervals of 5 s. The robustness function values, $R1$ represent the terminal reliability values of the system upon completion of the failure event. The simulation time to reach this stage is found to be 120 s for all three failure scenarios and the recovery model's simulation time is 180 s.

5.5.3.2 Overpressure Results

The Wiekama method is applied to the compressor station and the values of time taken for spilling to occur have been recorded in each of the failure scenarios under the condition that the mass flow rate, $\dot{m}$ is 100 $kg\,s^{-1}$. Both linear and logarithmic distributions are applied to the pressure increase function yielding the respective spilling times and overpressure values to the system.
Table 5.9 displays the various pressure increase models with the respective time taken for spilling and overpressure values;

Pressure increase	ΔT **spilling** (s)	**Overpressure** (Pa)
Linear (θ=0.5)	261	134
Linear (θ=1)	215	142
Logarithmic (λ=50)	1024	213
Logarithmic (λ=75)	268	137
Logarithmic (λ=100)	591	186

Table 5.9: Overpressure functions with effects

5.5.3.3 Robustness Function

The imprecise robustness function has been implemented for all three failure scenarios. The precise robustness function is also implemented on the condition that all three failure scenarios have been applied and this is compared with the imprecise robustness function for the individual scenarios.

Figure 5.5 displays the various robustness functions trialed out for this case study;

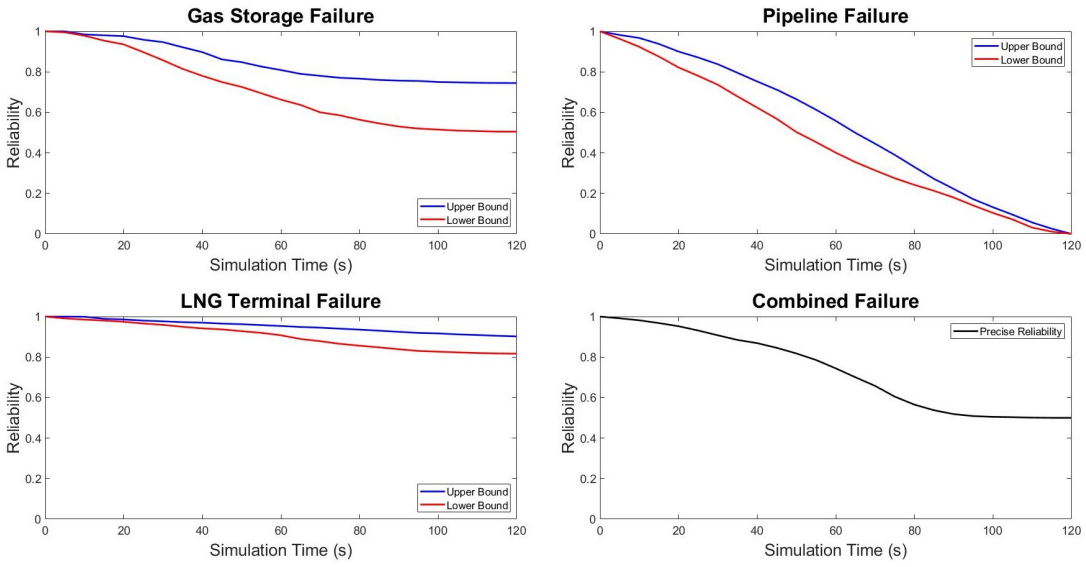


Figure 5.5: Imprecise robustness functions for various failure types

The three different failure scenarios for the robustness function are displayed on Figure 5.5 with their upper and lower bound reliability values. The gas storage failure results display increasingly diverging values from upper bound to lower bound reliability data as the simulation progresses. This contrasts with the pipeline failure robustness function which rapidly declines and converges until the reliability values reach 0. The LNG terminal failure slowly decreases and has a smaller divergence within the upper and

lower bound reliability values when compared with the gas storage failure. This type of failure generally produces the most robust loss function with a lower amount of uncertainty. This is hugely contrasted with the pipeline failure loss function which is guaranteed to drop to a minimal performance level. Finally, the precise loss function displays the expectation of failure on the condition that all three failure scenarios are equally likely to occur. The final reliability for this scenario equates to approximately half in terms of performance of the pipeline system.

5.5.3.4 Recovery Function

The imprecise recovery function is also deduced for the three failure scenarios with upper and lower bounds as displayed for the robustness function. This recovery function is implemented at a simulation time of 60 s after the robustness function ends and it is assumed that reliability values are constant until the recovery function initiates. Figure 5.6 displays the four respective recovery functions which are applied 60 s after the terminating stage of the robustness based simulation;

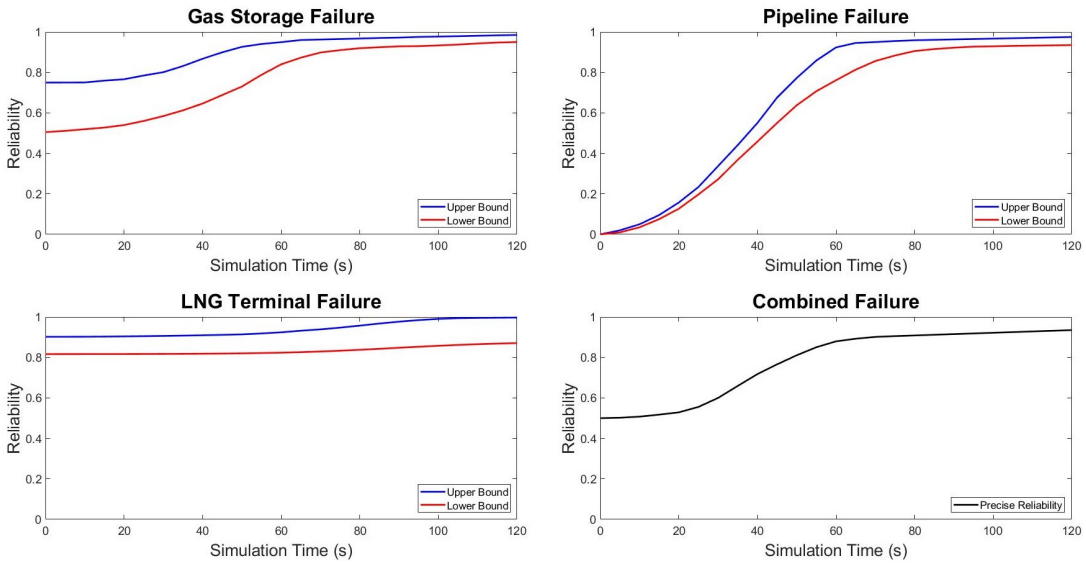


Figure 5.6: Imprecise recovery functions for various failure types

The reliability values for the recovery based simulation displays the upper and lower bounds for the recovery function when applied to the three respective failure scenarios. All four graphs start with the same reliability values as displayed in the end of the simulation from Figure 5.5. It can be deduced that the gas storage failure starts with

the widest range of uncertainty and this discrepancy slowly converges as the model recovers from its partially failed state. This is contrasted with the results for the pipeline failure recovery model which rapidly recovers from a completely failed state to a functioning state with its upper bound peaking at a faster rate than its lower bound. The LNG terminal failure recovers at a much slower rate as the initial reliability for this scenario is significantly higher than the previous two failure scenarios. The bounds of uncertainty slowly diverge until the simulation derives the terminal reliability values. A precise model to display a combination of all three recovery models is displayed with the assumption that each failure scenario is equally likely to occur.

5.6 Chapter Summary

This chapter presents the fundamental theory of momentum, heat and mass transfer (MHMT) in gas pipeline networks and applies these fundamental equations to determine the properties of dynamic flow of a basic pipeline . A CFD-based failure model has been proposed incorporating various triggers of performance loss within the gas pipeline system. A case study of a natural gas pipeline has been applied to these proposed techniques with the addition of epistemic uncertainty applied to the supply and demand nodes and aleatory uncertainty applied to the respective output performance functions.

Three different failure scenarios have been tested to show a wide range of possible outcomes within the natural gas pipeline. These three situations are defined as the gas storage failure, pipeline failure and LNG storage failure. Data has been obtained from the literature to deduce the network's properties including the volumetric flow rates for the supply and demand nodes and the pipeline properties for the pressure and flow rates which are internally present within the pipelines.

The simulations for the three models have been trialed out on PipeFlow Expert and the results were imported into MATLAB for data analysis. The final results to communicate resilience have been presented in the form of robustness for the loss function and a recovery function for post-disaster simulation. Both models have been presented with their respective uncertainty bounds and have been communicated in the form of their separate components of the three phase resilience plot.

Chapter 6

Conclusion

This chapter is written to provide a closure towards the research carried out, applications and lessons learnt from this study. This chapter is divided into two (2) sections with the concluding remarks discussed from this study and also suggestions for future academic work to be carried out.

6.1 Concluding Remarks

The field of resilience in Engineering systems is a relatively new, but it is an important component to reliability analysis. This thesis has presented the foundations of resilience, starting from the origins and motivations of carrying out this study. Communities such as the IChemE, IEEE and other Engineering entities have realised the importance of resilience in order to mitigate consequences of disaster with respects to their various applications from a social, technical and financial standpoint. As a result, many nations and states have become aware of the increasingly important study in the field of resilience, ranging from disaster prevention, consequence modelling and risk management. National Rail, for example currently has plans to innovate the UK railway network due to the old infrastructure used for British railway transport. Projects such as Great North Rail Project (GNRP) have been initiated to improve connecting network links in the north of England with intentions of providing a more reliable and better customer experience.

The introduction and literature review sections of this thesis clearly define the various angles of resilience ranging from defining ”Four R’s” and discriminating resilience between its infrastructural and operational characteristics to the motivations and applicability in real life system examples as a result of past disasters.

Probabilistic resilience analysis incorporates a diverse range of techniques, models, outputs and studies for the overall resilience framework, which is born from the base of probabilistic reliability assessment and is not exclusive solely to the three examples tested in this study. Applications of the studies carried out within this thesis are real world systems of a diverse range and have been applied to demonstrate a fraction of the complete literature and techniques to resilience quantification. The constraints of balancing computational cost with the quality of results has been highlighted in this thesis, with approaches such as estimation techniques and surrogate models being applied to combat this limitation. It has been shown that more complex, high parameter systems require more computational expense and therefore models must be formulated to reduce the computational expense whilst retaining the necessary parameter for accurate simulation outputs.

This thesis also highlights uncertainty and also stresses the importance of quantifying uncertainty in resilience based simulation techniques. Epistemic and aleatory uncertainty have been trialed out into existing precise quantification techniques and are clearly communicated in the results in the form of confidence bounds, p-boxes and discrete probability bounds. This novel application provides a relatively flexible approach to developing current resilience based simulations and communicates important information for risk based engineers with deducing possible consequences of their respective systems.

The aims and objectives from Section 1.5 have been fulfilled and the desired understanding, computational techniques and applications of current resilience models have been tested, linked and presented into a thesis. The strengths and limitations of each technique have also been highlighted and the reasons on which the limitations do not affect the credibility of the results are also discussed.

To summarise, this thesis has provided a successful and all-rounded approach to display and tackle the resilience based framework in current academic literature and provides the reader an early insight to the current state of the art and the possibilities of further resilience based academic output. This current framework is being developed on OpenCossan, an open source MATLAB toolbox for uncertainty quantification.

6.2 Recommendations for Future Work

The investment of resilience in academic literature is expanding constantly, with new challenges to face as interest grows in this field. There are still many scarcities in this literature and these voids are a talking point for potential growth within resilience based approaches. This section produces an array of current voids within this field and presents possible works that can be carried out by future scholars;

1. The limitations of the techniques applied in Chapter 3 have been mentioned in Section 3.4.3. The survival signature method becomes increasingly computationally expensive on an exponential level as more types of nodes are introduced into the system. Further work needs to be carried out to tackle this, possibly in the form of developing a survival signature meta-model for the application of networks with large numbers of components. The PrPm does not contain this limitation, however further work can be carried out in order to innovate this method and its technicalities. A possible example of this is to model the failure data using a Weibull distribution function as applied to the survival function.

2. With regards to resilience analysis as applied to the power grid in Chapter 4, the various limitations are within the lack of uncertainty for the ANN surrogate model which is a computationally arduous problem to solve. The possibility of an interval predictor to be applied to ANNs is a potential route to estimate uncertainty within ANNs and obtain credible imprecise results for the surrogate model applied in this chapter.

3. The CFD resilience-based model applied in Chapter 5 contains certain limitations within the framework as stated in Section 5.4.4. Issues regarding computational cost on PipeFlow Expert were encountered for networks with large numbers of nodes due to high volume of multi-dimensional input data applied in the case study. As a result, output data was obtained in discrete intervals of 5 s with the missing data for the values between each interval being interpolated. A more computationally efficient model can be developed using more advanced software in order to obtain more accurate results between intervals. Additionally, a model to combine the three failure scenarios without the assumption of equal probability can be developed in the future.

Bibliography

[1] Herbert S. Winokur and Lawrence J. Goldstein. Analysis of mission-oriented systems. *IEEE Transactions on Reliability*, R-18(4):144–148, 1969.

[2] Ali Grami. Chapter 8 - boolean algebra. In Ali Grami, editor, *Discrete Mathematics*, pages 131–153. Academic Press, 2023.

[3] Siwei Zhou, Luyao Ye, Shengwu Xiong, and Jianwen Xiang. Reliability analysis of dynamic fault trees with priority-and gates based on irrelevance coverage model. *Reliability Engineering & System Safety*, 224:108553, 2022.

[4] Yuxuan Cui, Kunsong Lin, Yunxia Chen, and Jiaxiao Zhu. A quantum-inspired model for statistical analysis of repairable systems. *Computers and Industrial Engineering*, 161:107613, 2021.

[5] Rebecca Lawton and Nicholas J. Ward. A systems analysis of the ladbroke grove rail crash. *Accident Analysis & Prevention*, 37(2):235–244, 2005.

[6] Neville A. Stanton and Guy H. Walker. Exploring the psychological factors involved in the ladbroke grove rail accident. *Accident Analysis & Prevention*, 43(3):1117–1127, 2011.

[7] Andrew Weyman, Rachel O'Hara, and Alan Jackson. Investigation into issues of passenger egress in ladbroke grove rail disaster. *Applied Ergonomics*, 36(6): 739–748, 2005.

[8] Kasimala Venkatanagaraju, Monalisa Biswal, and Almoataz Y. Abdelaziz. *Uncertainties in Modern Power Systems*. Academic Press, 2021.

[9] Olga P. Veloza and Francisco Santamaria. Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes. *The Electricity Journal*, 29(7):42–49, 2016.

[10] A Gailwad. Indian blackouts—july 30 & 31 2012 recommendations and further actions. In *IEEE PES General Meeting*, 2013.

[11] V. Rampurkar, P. Pentayya, H.A. Mangalvedekar, and F. Kazi. Cascading failure analysis for indian power grid. *IEEE Transactions on Smart Grid*, 7(4): 1951–1960, 2016.

[12] Robert R. Ulmer and Andrew S. Pyle. Finding renewal in the midst of disaster: The case of the deepwater horizon oil spill. *Public Relations Review*, 47(1): 102001, 2021.

[13] Hristina Denic-Roberts, Nicole Rowley, Mark C. Haigney, Kate Christenbury, John Barrett, Dana L. Thomas, Lawrence S. Engel, and Jennifer A. Rusiecki. Acute and longer-term cardiovascular conditions in the deepwater horizon oil spill coast guard cohort. *Environment International*, 158:106937, 2022.

[14] Samantha K Brooks, Rebecca Dunn, Richard Amlôt, Neil Greenberg, and G James Rubin. Social and occupational factors associated with psychological distress and disorder among disaster responders: a systematic review. *BMC psychology*, 4(1):1–13, 2016.

[15] Ashraf Labib. Chapter 5 - bp deepwater horizon. In *Learning from Failures*, pages 63–81. Butterworth-Heinemann, Oxford, 2014.

[16] Jerry Paul Sheppard and Jesse Young. An essential stupidity-based review of the deepwater horizon disaster. *Business Horizons*, 2022. URL https://www.sciencedirect.com/science/article/pii/S0007681322000209.

[17] Deborah P. French-McCay, Hilary Robinson, Michael Bock, Deborah Crowley, Paul Schuler, and Jill J. Rowe. Counter-historical study of alternative dispersant use in the deepwater horizon oil spill response. *Marine Pollution Bulletin*, 180: 113778, 2022.

[18] Xiangkun Meng, Xinhong Li, Weigang Wang, Guozheng Song, Guoming Chen, and Jingyu Zhu. A novel methodology to analyze accident path in deep-water drilling operation considering uncertain information. *Reliability Engineering & System Safety*, 205:107255, 2021.

[19] Zhonglin Wang, Marian Sorin Nistor, and Stefan Wolfgang Pickl. Analysis of the definitions of resilience. *IFAC-PapersOnLine*, 50(1):10649–10657, 2017.

[20] Intergovernmental Panel on Climate Change. *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2014.

[21] Chao Fang, Yanze Chu, Haoran Fu, and Yiping Fang. On the resilience assessment of complementary transportation networks under natural hazards. *Transportation Research Part D: Transport and Environment*, 109:103331, 2022.

[22] Sasan Ghasemi, Mokhtar Mohammadi, and Jamal Moshtagh. A new look-ahead restoration of critical loads in the distribution networks during blackout with considering load curve of critical loads. *Electric Power Systems Research*, 191: 106873, 2021.

[23] Xiao-Peng Niu, Shun-Peng Zhu, Jin-Chao He, Yang Ai, Kaikai Shi, and Liping Zhang. Fatigue reliability design and assessment of reactor pressure vessel structures: Concepts and validation. *International Journal of Fatigue*, 153: 106524, 2021.

[24] Gopalendu Pal, Ankur Gupta, Michael F. Modest, and Daniel C. Haworth. Comparison of accuracy and computational expense of radiation models in simulation of non-premixed turbulent jet flames. *Combustion and Flame*, 162(6): 2487–2495, 2015.

[25] Eklas Hossain, Shidhartho Roy, Naeem Mohammad, Nafiu Nawar, and Debopriya Roy Dipta. Metrics and enhancement strategies for grid resilience and reliability during natural disasters. *Applied Energy*, 290:116709, 2021.

[26] Aihua Liu, Ke Chen, Xiaofei Huang, Jieyun Chen, Jianfeng Zhou, and Wenbin Xu. Corrosion failure probability analysis of buried gas pipelines based on subset simulation. *Journal of Loss Prevention in the Process Industries*, 57:25 – 33, 2019.

[27] Marco de Angelis, Edoardo Patelli, and Michael Beer. Advanced line sampling for efficient robust reliability analysis. *Structural Safety*, 52:170–182, 2015. Engineering Analyses with Vague and Imprecise Information.

[28] Enno Ruijters, Daniël Reijsbergen, Pieter-Tjerk de Boer, and Mariëlle Stoelinga. Rare event simulation for dynamic fault trees. *Reliability Engineering & System Safety*, 186:220–231, 2019.

[29] Xue Lei and Cameron A. MacKenzie. Assessing risk in different types of supply chains with a dynamic fault tree. *Computers & Industrial Engineering*, 137: 106061, 2019.

[30] Lulu Jia, Yi Ren, Dezhen Yang, Qiang Feng, Bo Sun, and Cheng Qian. Reliability analysis of dynamic reliability block diagram based on dynamic uncertain causality graph. *Journal of Loss Prevention in the Process Industries*, 62: 103947, 2019.

[31] Phillip McNelles, Zhao Chang Zeng, Guna Renganathan, Greg Lamarre, Yolande Akl, and Lixuan Lu. A comparison of fault trees and the dynamic flowgraph methodology for the analysis of fpga-based safety systems part 1: Reactor trip logic loop reliability analysis. *Reliability Engineering & System Safety*, 153:135 – 150, 2016.

[32] Sohag Kabir and Yiannis Papadopoulos. Applications of bayesian networks and petri nets in safety, reliability, and risk assessments: A review. *Safety Science*, 115:154 – 175, 2019.

[33] A.A. Chojaczyk, A.P. Teixeira, L.C. Neves, J.B. Cardoso, and C. Guedes Soares. Review and application of artificial neural networks models in reliability analysis of steel structures. *Structural Safety*, 52:78–89, 2015.

[34] P.S. Rajpal, K.S. Shishodia, and G.S. Sekhon. An artificial neural network for modeling reliability, availability and maintainability of a repairable system. *Reliability Engineering & System Safety*, 91(7):809 – 819, 2006.

[35] Emil Mahler Larsen, Yi Ding, Yan-Fu Li, and Enrico Zio. Definitions of generalized multi-performance weighted multi-state k-out-of-n system and its reliability evaluations. *Reliability Engineering & System Safety*, page 105876, 2017.

[36] Frank P. A. Coolen and Tahani Coolen-Maturi. Generalizing the signature to systems with multiple types of components. In *Complex Systems and Dependability*, pages 115–130. Springer Berlin Heidelberg, 2012.

[37] Francisco J. Samaniego. *System Signatures and their Applications in Engineering Reliability*, pages 21–36. Springer US, 2007.

[38] Frank PA Coolen, Tahani Coolen-Maturi, and Abdullah H Al-nefaiee. Nonparametric predictive inference for system reliability using the survival signature. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 228(5):437–448, 2014.

[39] Edoardo Patelli, Geng Feng, Frank P.A. Coolen, and Tahani Coolen-Maturi. Simulation methods for system reliability using the survival signature. *Reliability Engineering & System Safety*, 167:327 – 337, 2017.

[40] Serkan Eryilmaz, Frank P.A. Coolen, and Tahani Coolen-Maturi. Marginal and joint reliability importance based on survival signature. *Reliability Engineering & System Safety*, 172:118 – 128, 2018.

[41] Xianzhen Huang, Louis J.M. Aslett, and Frank P.A. Coolen. Reliability analysis of general phased mission systems with a new survival signature. *Reliability Engineering & System Safety*, 189:416 – 422, 2019.

[42] Yao Li, Frank P.A. Coolen, Caichao Zhu, and Jianjun Tan. Reliability assessment of the hydraulic system of wind turbines based on load-sharing using survival signature. *Renewable Energy*, 153:766 – 776, 2020.

[43] Jasper Behrensdorf, Tobias-Emanuel Regenhardt, Matteo Broggi, and Michael Beer. Numerically efficient computation of the survival signature for the reliability analysis of large networks. *Reliability Engineering & System Safety*, 216: 107935, 2021.

[44] Yanjie Tong and Iris Tien. Analytical probability propagation method for reliability analysis of general complex networks. *Reliability Engineering & System Safety*, 189:21 – 30, 2019.

[45] Yanjie Tong and Iris Tien. Probability propagation method for reliability assessment of acyclic directed networks. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 5(3):04019011, 2019.

[46] Michel Bruneau, Stephanie Chang, Ronald Eguchi, George Lee, Thomas O'Rourke, Andrei Reinhorn, Masanobu Shinozuka, Kathleen Tierney, William Wallace, and Detlof Winterfeldt. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra - EARTHQ SPECTRA*, 19, 2003.

[47] Jesus Beyza and Jose M. Yusta. Characterising the security of power system topologies through a combined assessment of reliability, robustness, and resilience. *Energy Strategy Reviews*, 43:100944, 2022.

[48] Ali Kharrazi, Yadong Yu, Arun Jacob, Nemi Vora, and Brian D. Fath. Redundancy, diversity, and modularity in network resilience: Applications for international trade and implications for public policy. *Current Research in Environmental Sustainability*, 2:100006, 2020.

[49] Alessandro Zona, Omar Kammouh, and Gian Paolo Cimellaro. Resourcefulness quantification approach for resilient communities and countries. *International Journal of Disaster Risk Reduction*, 46:101509, 2020.

[50] Ahmed Mebarki. Safety of atmospheric industrial tanks: Fragility, resilience and recovery functions. *Journal of Loss Prevention in the Process Industries*, 49:590–602, 2017.

[51] Laura Petersen, Emma Lundin, Laure Fallou, Johan Sjöström, David Lange, Rui Teixeira, and Alexandre Bonavita. Resilience for whom? the general public's tolerance levels as ci resilience criteria. *International Journal of Critical Infrastructure Protection*, 28:100340, 2020.

[52] Jordan J. Plotnek and Jill Slay. Power systems resilience: Definition and taxonomy with a view towards metrics. *International Journal of Critical Infrastructure Protection*, 33:100411, 2021.

[53] Reza Arghandeh, Alexandra von Meier, Laura Mehrmanesh, and Lamine Mili. On the definition of cyber-physical resilience in power systems. *Renewable and Sustainable Energy Reviews*, 58:1060–1069, 2016.

[54] Zhaohong Bie, Yanling Lin, Gengfeng Li, and Furong Li. Battling the extreme: A study on the power system resilience. *Proceedings of the IEEE*, 105 (7):1253–1266, 2017.

[55] Han Zhang, Hanjie Yuan, Gengfeng Li, and Yanling Lin. Quantitative resilience assessment under a tri-stage framework for power systems. *Energies*, 11(6), 2018.

[56] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou. Metrics and quantification of operational and infrastructure resilience in power systems. *IEEE Transactions on Power Systems*, 32(6):4732–4742, 2017.

[57] Mathaios Panteli and Pierluigi Mancarella. The grid: Stronger, bigger, smarter?: Presenting a conceptual framework of power system resilience. *IEEE Power and Energy Magazine*, 13(3):58–66, 2015.

[58] Bernardino Chiaia, Emilio Barchiesi, Valerio De Biagi, and Luca Placidi. A novel structural resilience index: Definition and applications to frame structures. *Mechanics Research Communications*, 99:52 – 57, 2019.

[59] Xinzheng Lu, Wenjie Liao, Dongping Fang, Kaiqi Lin, Yuan Tian, Chi Zhang, Zhe Zheng, and Pengju Zhao. Quantification of disaster resilience in civil engineering: A review. *Journal of Safety Science and Resilience*, 1(1):19–30, 2020.

[60] Chen Xiong, Jin Huang, and Xinzheng Lu. Framework for city-scale building seismic resilience simulation and repair scheduling with labor constraints driven by time–history analysis. *Computer-Aided Civil and Infrastructure Engineering*, 35(4):322–341, 2020.

[61] Xiangyu Ma, Huijie Zhou, and Zhiyi Li. On the resilience of modern power systems: A complex network perspective. *Renewable and Sustainable Energy Reviews*, 152:111646, 2021.

[62] Nikola Bešinović, Raphael Ferrari Nassar, and Christopher Szymula. Resilience assessment of railway networks: Combining infrastructure restoration and transport management. *Reliability Engineering & System Safety*, 224: 108538, 2022.

[63] Jinqu Chen, Jie Liu, Qiyuan Peng, and Yong Yin. Resilience assessment of an urban rail transit network: A case study of chengdu subway. *Physica A: Statistical Mechanics and its Applications*, 586:126517, 2022.

[64] Fauzan Hanif Jufri, Victor Widiputra, and Jaesung Jung. State-of-the-art review on power grid resilience to extreme weather events: Definitions, frameworks, quantitative assessment methodologies, and enhancement strategies. *Applied Energy*, 239:1049 – 1065, 2019.

[65] Mathaios Panteli and Pierluigi Mancarella. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electric Power Systems Research*, 127:259 – 270, 2015.

[66] Michel Bruneau, Stephanie E. Chang, Ronald T. Eguchi, George C. Lee, Thomas D. O'Rourke, Andrei M. Reinhorn, Masanobu Shinozuka, Kathleen Tierney, William A. Wallace, and Detlof von Winterfeldt. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra*, 19(4):733–752, 2003.

[67] Dong Hwan Kim, Daniel A. Eisenberg, Yeong Han Chun, and Jeryang Park. Network topology and resilience analysis of south korean power grid. *Physica A: Statistical Mechanics and its Applications*, 465:13 – 24, 2017.

[68] Hindolo George-Williams and Edoardo Patelli. A hybrid load flow and event driven simulation approach to multi-state system reliability evaluation. *Reliability Engineering & System Safety*, 152:351–367, 2016.

[69] Hindolo George-Williams and Edoardo Patelli. Efficient availability assessment of reconfigurable multi-state systems with interdependencies. *Reliability Engineering & System Safety*, 165:431–444, 2017.

[70] Roberto Rocchetta, Enrico Zio, and Edoardo Patelli. A power-flow emulator approach for resilience assessment of repairable power grids subject to weather-induced failures and data deficiency. *Applied Energy*, 210:339–350, 2018.

[71] Roberto Rocchetta, Luca Bellani, Michele Compare, Enrico Zio, and Edoardo Patelli. A reinforcement learning framework for optimal operation and maintenance of power grids. *Applied Energy*, 241:291–301, 2019.

[72] Roberto Rocchetta and Edoardo Patelli. Assessment of power grid vulnerabilities accounting for stochastic loads and model imprecision. *International Journal of Electrical Power & Energy Systems*, 98:219 – 232, 2018.

[73] Roberto Rocchetta and Edoardo Patelli. A post-contingency power flow emulator for generalized probabilistic risks assessment of power grids. *Reliability Engineering & System Safety*, 197:106817, 2020.

[74] Huai Su, Enrico Zio, Zong-Jie Zhang, Chang-Zheng Xiong, Qian-Sheng Dai, Qing-Wei Wu, and Jin-Jun Zhang. Development of an integrated dynamic model for supply security and resilience analysis of natural gas pipeline network systems. *Petroleum Science*, 2021.

[75] Antonio Marino and Enrico Zio. A framework for the resilience analysis of complex natural gas pipeline networks from a cyber-physical system perspective. *Computers & Industrial Engineering*, 162:107727, 2021.

[76] Austin M. McKeand, Recep M. Gorguluarslan, and Seung-Kyum Choi. Stochastic analysis and validation under aleatory and epistemic uncertainties. *Reliability Engineering & System Safety*, 205:107258, 2021.

[77] Seyed Ashkan Zarghami and Jantanee Dumrak. Aleatory uncertainty quantification of project resources and its application to project scheduling. *Reliability Engineering & System Safety*, 211:107637, 2021.

[78] Yunjun Zheng, Han Zhao, Shengchao Zhen, and Hao Sun. Fuzzy-set theory based optimal robust constraint-following control for permanent magnet synchronous motor with uncertainties. *Control Engineering Practice*, 115:104911, 2021.

[79] A. Gray, A. Wimbush, M. de Angelis, P.O. Hristov, D. Calleja, E. Miralles-Dolz, and R. Rocchetta. From inference to design: A comprehensive framework for uncertainty quantification in engineering with limited information. *Mechanical Systems and Signal Processing*, 165:108210, 2022.

[80] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations – i: Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.

[81] Alexander Petersen, Chao Zhang, and Piotr Kokoszka. Modeling probability density functions as data objects. *Econometrics and Statistics*, 2021.

[82] Giora Maymon. Chapter 2 - some important statistical distributions. In *Stochastic Crack Propagation*, pages 9–18. Academic Press, 2018.

[83] Enrico Zio, Piero Baraldi, and Edoardo Patelli. Assessment of the availability of an offshore installation by monte carlo simulation. *International Journal of Pressure Vessels and Piping*, 83(4):312–320, 2006.

[84] Adolphus Lye, Masaru Kitahara, Matteo Broggi, and Edoardo Patelli. Robust optimization of a dynamic black-box system under severe uncertainty: A distribution-free framework. *Mechanical Systems and Signal Processing*, 167: 108522, 2022.

[85] Luis G. Crespo and Sean P. Kenny. The nasa langley challenge on optimization under uncertainty. *Mechanical Systems and Signal Processing*, 152:107405, 2021.

[86] Scott Ferson, Vladik Kreinovich, Lev Ginzburg, Davis Myers, and Kari Sentz. Constructing probability boxes and dempster-shafer structures. *Office of Scientific and Technical Information*, 2003.

[87] Michael Beer, Scott Ferson, and Vladik Kreinovich. Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing*, 37(1):4–29, 2013.

[88] Haibo Liu, Ming Chen, Chong Du, Jiachang Tang, Chunming Fu, and Guilin She. A copula-based uncertainty propagation method for structures with correlated parametric p-boxes. *International Journal of Approximate Reasoning*, 138:89–104, 2021.

[89] Nicholas Gray, Scott Ferson, Marco De Angelis, Ander Gray, and Francis Baumont de Oliveira. Probability bounds analysis for python. *Software Impacts*, 12:100246, 2022.

[90] Wenqing Zhu, Ning Chen, Jian Liu, and Michael Beer. A probability-box-based method for propagation of multiple types of epistemic uncertainties and its application on composite structural-acoustic system. *Mechanical Systems and Signal Processing*, 149:107184, 2021.

[91] Basem A. Alkhaleel, Haitao Liao, and Kelly M. Sullivan. Risk and resilience-based optimal post-disruption restoration for critical infrastructures under uncertainty. *European Journal of Operational Research*, 296(1):174–202, 2022.

[92] Jiaxin Wu and Pingfeng Wang. Risk-averse optimization for resilience enhancement of complex engineering systems under uncertainties. *Reliability Engineering & System Safety*, 215:107836, 2021.

[93] P. Cicilio, L. Swartz, B. Vaagensmith, C. Rieger, J. Gentle, T. McJunkin, and E. Cotilla-Sanchez. Electrical grid resilience framework with uncertainty. *Electric Power Systems Research*, 189:106801, 2020.

[94] Gianluca Filippi, Massimiliano Vasile, Daniel Krpelik, Peter Zeno Korondi, Mariapia Marchi, and Carlo Poloni. Space systems resilience optimisation under epistemic uncertainty. *Acta Astronautica*, 165:195 – 210, 2019.

[95] Zhaolong Li, Chun Jin, Pan Hu, and Cong Wang. Resilience-based transportation network recovery strategy during emergency recovery phase under uncertainty. *Reliability Engineering & System Safety*, 188:503–514, 2019.

[96] N. De Koker, C. Viljoen, R. Lenner, and S.W. Jacobsz. Updating structural reliability efficiently using load measurement. *Structural Safety*, 84:101939, 2020.

[97] A.S. Cheliyan and S.K. Bhattacharyya. Fuzzy fault tree analysis of oil and gas leakage in subsea production systems. *Journal of Ocean Engineering and Science*, 3(1):38 – 48, 2018.

[98] Francisco J. Samaniego. *System Signatures and their Applications in Engineering Reliability*, pages 7–19. Springer US, 2007.

[99] Yuchang mo, Liudong Xing, Suprasad Amari, and Joanne Dugan. Efficient analysis of multi-state k-out-of-n systems. *Reliability Engineering & System Safety*, 133:95–105, 2015.

[100] Xinlei Hu, Jie Huang, and Feng Shi. Circuity in china's high-speed-rail network. *Journal of Transport Geography*, 80:102504, 2019.

[101] J. Samaniego, F. *International Series in Operations Research & Management Science*. Springer Science+Business Media, 2007.

[102] Qingwen Lu, Ping an Zhong, Bin Xu, Feilin Zhu, Yufei Ma, Han Wang, and Sunyu Xu. Risk analysis for reservoir flood control operation considering two-dimensional uncertainties based on bayesian network. *Journal of Hydrology*, 589:125353, 2020.

[103] Fabio Antonelli, Vittorio Cortellessa, Marco Gribaudo, Riccardo Pinciroli, Kishor S. Trivedi, and Catia Trubiani. Analytical modeling of performance indices under epistemic uncertainty applied to cloud computing systems. *Future Generation Computer Systems*, 102:746 – 761, 2020.

[104] Jon C. Helton, Jay D. Johnson, William L. Oberkampf, and Cédric J. Sallaberry. Representation of analysis results involving aleatory and epistemic uncertainty. *International Journal of General Systems*, 39(6):605–646, 2010.

[105] Guijie Li, Zhenzhou Lu, Luyi Li, and Bo Ren. Aleatory and epistemic uncertainties analysis based on non-probabilistic reliability and its kriging solution. *Applied Mathematical Modelling*, 40(9):5703 – 5716, 2016.

[106] Ahmad M. Aboalkhair, Frank P. A. Coolen, and Iain M. MacPhee. Nonparametric predictive reliability of series of voting systems. *Eur. J. Oper. Res.*, 226: 77–84, 2013.

[107] Hector Estrada-Lugo, Silvia Tolo, Marco De Angelis, and Edoardo Patelli. Pseudo credal networks for inference with probability intervals. *ASCE-ASME J. Risk and Uncert. in Engrg. Sys., Part B: Mech. Engrg.*, 07 2019.

[108] Armin Bunde and Shlomo Havlin. *Fractals and Disordered Systems*, pages 59–114. Springer Berlin Heidelberg, 1996.

[109] Shlomo Havlin and Armin Bunde. *Fractals and Disordered Systems*, pages 115–176. Springer Berlin Heidelberg, 1996.

[110] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2-3):161–180, 1995.

[111] Reuven Cohen, Keren Erez, Daniel ben Avraham, and Shlomo Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85:4626–4628, 2000.

[112] Jasper Behrensdorf. Friesischscott/survivalsignature.jl: v0.2.0, 2021.

[113] Mao Shuai, Wang Chengzhi, Yu Shiwen, Gen Hao, Yu Jufang, and Hou Hui. Review on economic loss assessment of power outages. *Procedia Computer Science*, 130:1158 – 1163, 2018.

[114] Guo Chen, YuanYu Dai, Zhao Xu, ZhaoYang Dong, and YuSheng Xue. A flexible framework of line power flow estimation for high-order contingency analysis. *International Journal of Electrical Power & Energy Systems*, 70:1–8, 2015.

[115] M. Hadi Amini, Shahab Bahrami, Farhad Kamyab, Sakshi Mishra, Rupamathi Jaddivada, Kianoosh Boroojeni, Paul Weng, and Yinliang Xu. Chapter 6 - decomposition methods for distributed optimal power flow: Panorama and case studies of the dc model. In *Classical and Recent Aspects of Power System Optimization*, pages 137 – 155. Academic Press, 2018.

[116] Jinyu Li, Guoxin Tan, Bingchao Cheng, Dandan Liu, and Weisan Pan. Transport mechanism of chitosan-n-acetylcysteine, chitosan oligosaccharides or carboxymethyl chitosan decorated coumarin-6 loaded nanostructured lipid carriers across the rabbit ocular. *European Journal of Pharmaceutics and Biopharmaceutics*, 120:89 – 97, 2017.

[117] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W.L. Kling. Usefulness of dc power flow for active power flow analysis with flow controlling devices. *IET Conference Proceedings*, pages 58–62(4), 2006.

[118] Jian Li, Congling Shi, Changkun Chen, and Leonardo Dueñas-Osorio. A cascading failure model based on ac optimal power flow: Case study. *Physica A: Statistical Mechanics and its Applications*, 508:313 – 323, 2018.

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

[120] Rodrigo Mena, Martin Hennebel, Yan-Fu Li, and Enrico Zio. A multi-objective optimization framework for risk-controlled integration of renewable generation into electric power systems. *Energy*, 106:712–727, 2016.

[121] Mehdi Golari, Neng Fan, and Jianhui Wang. Two-stage stochastic optimal islanding operations under severe multiple contingencies in power grids. *Electric Power Systems Research*, 114:68 – 77, 2014.

[122] K. Alvehag and L. Soder. A reliability model for distribution systems incorporating seasonal variations in severe weather. *IEEE Transactions on Power Delivery*, 26(2):910–919, 2011.

[123] Francesco Cadini, Gian Luca Agliardi, and Enrico Zio. A modeling and simulation framework for the reliability/availability assessment of a power transmission grid subject to cascading failures under extreme weather conditions. *Applied Energy*, 185:267 – 279, 2017.

[124] Jiboning Zhang. Investment risk model based on intelligent fuzzy neural network and var. *Journal of Computational and Applied Mathematics*, page 112707, 2020.

[125] P.C. Verpoort, P. MacDonald, and G.J. Conduit. Materials data validation and imputation with an artificial neural network. *Computational Materials Science*, 147:176 – 185, 2018.

[126] Teresita Canchala-Nastar, Yesid Carvajal-Escobar, Wilfredo Alfonso-Morales, Wilmar Loaiza Cerón, and Eduardo Caicedo. Estimation of missing data of monthly rainfall in southwestern colombia using artificial neural networks. *Data in Brief*, 26:104517, 2019.

[127] E. Zio. A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. *IEEE Transactions on Nuclear Science*, 53(3):1460–1478, 2006.

[128] M. Belivanis and K. Bell. Representative gb network model, 2013. URL https://www.maths.ed.ac.uk/optenergy/NetworkData/reducedGB/.

[129] Reza Hafezi, AmirNaser Akhavan, Saeed Pakseresht, and David A. Wood. Global natural gas demand to 2025: A learning scenario development model. *Energy*, 224:120167, 2021.

[130] Kwabena Addo Pambour, Ricardo Bolado-Lavin, and Gerard P.J. Dijkema. An integrated transient model for simulating the operation of natural gas transport systems. *Journal of Natural Gas Science and Engineering*, 28:672–690, 2016.

[131] EGIG EGPIDG. Gas pipeline incidents: 8th report of the european gas pipeline incident data group, 2011.

[132] Pavel Praks, Vytis Kopustinskas, and Marcelo Masera. Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure. *Reliability Engineering & System Safety*, 144:254–264, 2015.

[133] V Kopustinskas and P Praks. Time dependent gas transmission network probabilistic simulator: focus on storage discharge modeling. In *Proceedings of the european safety and reliability conference ESREL*, 2014.

[134] Min Ouyang. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering & System Safety*, 121:43–60, 2014.

[135] Man-Jin Jung, Joseph H Cho, and Wungsang Ryu. Lng terminal design feedback from operator's practical improvements. In *The 22nd World Gas Congress*. Citeseer, 2003.

[136] Xiaoge Zhang, Sankaran Mahadevan, Shankar Sankararaman, and Kai Goebel. Resilience-based network design under uncertainty. *Reliability Engineering & System Safety*, 169:364–379, 2018.

[137] Joaquim Casal. Chapter 2 - source term. In *Evaluation of the Effects and Consequences of Major Accidents in Industrial Plants*, pages 25–74. Elsevier, 2018.

[138] Mikael I. Johansson, Martin Östling, and Günter Jagschies. Chapter 25 - downstream processing equipment. In *Biopharmaceutical Processing*, pages 477–492. Elsevier, 2018.

[139] José María Gonzalez de Durana, Oscar Barambones, Enrique Kremers, and Liz Varga. Agent based modeling of energy networks. *Energy Conversion and Management*, 82:308–319, 2014.

[140] Kathleen Tierney and M. Bruneau. Conceptualizing and measuring resilience: A key to disaster loss reduction. *TR News*, 250:14–17, 2007.

[141] Sorin Neacsu, Silvian Suditu, and Doru Stoica. Considerations on transport capacity of natural gas pipelines and its limits. *ANALELE STIINTIFICE ALE UNIVERSITATII OVIDIUS CONSTANTA-SERIA MATEMATICA*, 24(2): 67–72, 2013.

[142] Huai Su, Jinjun Zhang, Enrico Zio, Nan Yang, Xueyi Li, and Zongjie Zhang. An integrated systemic method for supply reliability assessment of natural gas pipeline networks. *Applied Energy*, 209:489–501, 2018.