Resilience study research for NIC

Systems analysis of interdependent network vulnerabilities

Final Report
April 2020

Dr. Raghav Pant
Mr. Tom Russell
Dr. Conrad Zorn
Dr. Edward Oughton
Prof. Jim W. Hall

Environmental Change Institute
University of Oxford
South Parks Road
Oxford, OX1 3QY

This report may be cited as follows:
Pant, R., Russell, T., Zorn C., Oughton, E., and Hall, J.W. (2020). Resilience study research for NIC – Systems analysis of interdependent network vulnerabilities. Environmental Change Institute, Oxford University, UK.

This report was produced to inform the National Infrastructure Commission's study on resilience. The views expressed and recommendations set out in this report are the authors' own and do not necessarily reflect the position of the National Infrastructure Commission.

The materials have been prepared by Oxford University. Whilst every care has been taken by Oxford University to ensure the accuracy and completeness of the reports and maps, the reader must recognise that errors are possible through no fault of Oxford University and as such the parties give no express or implied representations or warranty as to:

(i) the quality or fitness for any particular purpose of the report or maps supplied or of any design, workmanship, materials or parts used in connection therewith or correspondence with regard to any description or sample; or

(ii) the accuracy, sufficiency or completeness of the reports or maps provided. In particular, there are hereby expressly excluded all conditions, warranties and other terms which might otherwise be implied (whether by common law, by statute or otherwise).

Oxford University, its employees, servants and agents shall accept no liability for any damage caused directly or indirectly by the use of any information contained herein and without prejudice to the generality of the foregoing, by any inaccuracies, defects or omissions.

Contents

1. INTRODUCTION

1.1 Background and main objectives

The resilience of economic infrastructure is critical to the continued provision of services on which everyday socio-economic activities depend. Economic infrastructure, such as electricity, digital communication, water supply, railways, and roads, are large interdependent networked systems. The vulnerability and resilience assessment of each of these infrastructure networks depends on understanding how failures in one network can result in cascading impacts across others. Quantifying vulnerabilities requires a system-of-systems approach underpinned by data on real-world networks' physical structure, their operational characteristics, and failure characteristics. Such analysis allows improved decision-making from the knowledge and tools to geospatially identify vulnerable locations and assets that have the most impact on systemic performance. Understanding of vulnerabilities, possible modes of failure and consequences provides the rationale for actions required for enhancing infrastructure system resilience.

This report describes the work done, in developing a system-of-systems modelling approach, by Oxford University in the 'Resilience Study Research for NIC' project, which was commissioned by the National Infrastructure Commission (NIC). The project timeline was from September 2019 till May 2020. The system-of-systems approach is demonstrated for UK with national-scale network representations of electricity, road and rail transport, public water supply and digital communication networks, capturing their interdependencies.

As outlined by the NIC, the purpose of the project was three-fold[1]:

1. To pilot an approach to assess the key physical vulnerabilities of the current UK economic infrastructure system
2. To draw out vulnerabilities that arise from network architecture and how these are likely to change in the future.
3. To inform the development of a framework to identify actions to assess, improve and monitor the resilience of the system.

In response to the above, we satisfied the NIC's main requirements[1] for us, which were to:

1. Identify a range of vulnerabilities characteristics that arise from the architecture of the UK economic infrastructure network, in consultation with the NIC. Each characteristic should be accompanied by criteria to establish the relative importance of the characteristic in different parts of the system as well as compared with others, for example based on impacts.
2. Develop a model to assess the most relevant of these characteristics for the current UK economic infrastructure system, and likely changes in the future.
3. Use the model to produce a preliminary assessment of these characteristics and their relative importance.
4. Identify some resilience enhancing options for reducing network vulnerabilities and evaluate the effectiveness of these options.

Specifically, to assess national infrastructure network vulnerabilities, the main questions answered during the project included:

1. What are the different (inter)dependencies between networks and how do these affect failure propagation?

[1] From NIC Terms of Reference

2. Can we see a difference in the failure propagation if we increase the connections between networks?
3. What is the effect of adding backups to the different interdependent nodes? What are the failure sequences and over what timeframe do they occur?
4. Can we identify a list of possible characteristics of the UK infrastructure networks that provide indications of the vulnerabilities of the system, as well as its resilience?
5. How do we establish criteria to identify the relative importance of each characteristic in different parts of the system as well as compared to other characteristics?
6. How would the network vulnerabilities change in the future under different planning scenarios?

1.2 Key findings

1.2.1 Effects of different resilience enhancing options

To understand how network interdependencies influence failure cascades, we looked at all *single point* (node) initiating failure events in electricity and telecoms networks and their propagation into other networks. Throughout the analysis it was assumed that for utility networks of electricity, water supply and telecoms the network nodes were considered to have failed only when they lost all their service. Partial failure states, where nodes might still be operating at below 100% operational levels and providing reduced service were not considered. For transport networks of railways and roads we assumed that failures were initiated in a way similar to the utility networks with nodes completely losing their ability to provide service, and we also accounted for disruptions to nodes that lost part of their pre-disruption journeys due to network failure propagation. The assumption that failure led to total loss of service was considered appropriate because we were interested in understanding worst-case scenarios of large-scale widespread disruptions.

We looked at two types of resilience options: (1) the effects of adding more connections between networks, which would provide alternative ways of providing essential infrastructure services; and (2) the effects of incorporating backup electricity supply into telecoms, water and road assets, which would substitute for lost electricity network supply but only for limited durations.

We first considered the case where networks were connected such that each dependent node of one network derived its supply from only one node of the other network. This case, called 'single connections with no backup supply', signified the *baseline case* for representing networks connections and resilience. Subsequently we considered the following resilience enhancing options:

1. Two connections (2C) – By connecting each dependent node of one network to two nodes of the supplying network;
2. Three connections (3C) – By connecting each dependent node of one network to three nodes of the supplying network;
3. Backup supply (B) – By assuming that some assets had backup electricity supply lasting a certain duration based on random gamma distribution survival rates. The telecoms and water nodes were assumed to have backup electricity supply lasting from 2 hours to 72 hours, while some roads with tunnels were assumed to have 24 hours of backup electricity supply;
4. Two connections and with backup supply (2C+B) – which combined options 1 and 3 above;

5. Three connections and with backup supply (3C+B) - which combined options 1 and 3 above.

Our analysis showed that in the baseline case of all failure events *initiated in the electricity network*, about 40% of failure events led to further disruptions to telecoms and at least one of rail and water. A further 20% of failure events led to further electricity failures, and 5.7% to another order of telecoms failures. By enhancing resilience to the 2C option electricity initiated cascading failures were reduced significantly, with about 5.6% events leading to telecoms and at least one of rail and water disruptions, with further 0.9% events leading to electricity failures, and 0.11% to another order of telecoms failures. Further improvements were created with the 3C option, though they were only marginal relative to the 2C case.

Similar analysis of failures *initiated in the telecoms network* showed that in the baseline case about 7.8% failure events led to electricity and at least one of rail and water disruptions, with 1.8% events leading to further order of telecoms failures. With the 2C and 3C resilience enhancing options cascading failures from telecoms to other networks were almost eliminated, with about 0.3% events leading to electricity and at least one of rail and water disruptions, and a further 0.02% events leading to another order of telecoms failures.

We also compared the failure impacts for each network, and cumulatively in terms of two metrics: (1) the numbers of disrupted users (residential customers over a day); and (2) the macroeconomic input-output (IO) losses in £million/day over the UK economy comprising 129 industry sectors.

The analysis showed that, in the baseline case, single failures initiated from the electricity network had the potential to cause the largest disruption of about 8 million users/day cumulative across all networks. This was mainly due to a knock-on effect on the water network. But the highest macroeconomic output losses, across the whole UK economy, of about £6.7 million/day were mainly due to railways with disruptions affecting a significant proportion of its total capacity. With the 2C resilience enhancing option the highest cumulative failure impacts were reduced to around 2.6 million user disruptions or £4.9 million/day, which was a different event from the baseline case. Most of the high impact failures in the water network were eliminated in comparison to the baseline case, while railway disruptions were still producing largest economic losses. The 3C option further reduced the highest cumulative failure impact to 1.3 million user disruptions or £3.8 million/day due to telecoms and railway failures initiated from electricity failures. Similar analysis for failures initiated in the telecoms network showed that in the baseline case the largest cumulative disruption of about 7 million users or £7 million/day economic output losses were mainly from knock-on effects on the water and railway networks in terms of user disruptions. But these were completely eliminated with the 2C and 3C resilience options, where the highest cumulative failure impacts resulted in 280,000 user disruptions or £0.36 million/day economic output losses mainly due to failure being confined to the telecoms network with some disruptions propagating towards the electricity networks only.

The economic loss analysis also showed that direct economic demand losses from infrastructure user disruptions led to total output losses that were between 1.41 – 2.36 times of the direct losses, which signified the economic multiplier effects of infrastructure driven demand side disruptions to the macroeconomic IO system.

To understand the effectiveness of the electricity backup supply option (B) in a systemic way, we re-simulated each of the 50 worst-case failure events in the baseline case, ranked by their total user disruptions across all networks. For each event we performed 20 simulations assuming a failure lasting over a 100-hour timeframe and with different gamma distribution-based survival times for backup supply durations of telecoms, water and road assets. We then estimated the time-averaged values of the disruptions across the 50 events with 20 simulations per event. Our analysis showed that on average backup supply effects prevented worst-case disruptions from growing until around 10 hours after which the impacts grew significantly to around 24 hours and further until up to 42 hours when the electricity backup supply of telecoms exchanges was first exhausted, followed by road and water backups being exhausted. Over 100 hours backup electricity supply helped reduce systemic worst-case electricity-initiated network failure impacts by 17% and systemic worst-case telecoms-initiated network failure impacts by 7%. About 33%-75% of the total avoided disruptions occurred between the first 10-30 hours when most of the backup supply was still working. This highlighted the importance of having backup supply and crucially also showed that if the original disrupted networks were to be restored then there are significant gains that can be made if the repairs occurred within 10-30 hours after the initiating failure event. Especially, if the repairs happened closer to 10 hours then most of the cascading disruptions could be avoided.

Overall applying all resilience options to the systemic analysis of the 50 worst-case electricity-initiated disruptive events, ranked by total customer disruptions across all networks, in the baseline case showed that for the 2C and 3C options disruptions from electricity networks were reduced by about 70%, telecoms by 91%-95%, water and road disruptions by at least 90% and at most 100%, and railways 82%-93%. The backup supply (B) options were most effective for roads where on average disruptions are reduced by about 40%, from the baseline and for other networks the gains were between 10%-23%. For combined backup and increased connection options, the biggest gains are made in the electricity networks where the 2C+B option reduced disruptions on average by 78% and the 3C+B option reduces disruptions on average by 81%, a gain of 10%-13% over the options with no backup supply. This showed that adding backup electricity supply to other networks could in turn reduce and delay further cascading impacts on the electricity network and help avoid disruptions. The total cumulative disruptions were reduced on average by 89% (2C+B) and 94% (3C+B) when considering the combinations of backup supply and increased network redundancies. Since all these worst-case disruption events in the baseline scenario resulted in cumulative disruptions between 1 - 8 million users and £0.5 - £6.7 million/day such gains were quite significant.

1.2.2 Future network vulnerabilities and resilience options

We analysed the resilience of future configurations of national infrastructure systems, based on NIC recommendations in the National Infrastructure Assessment (see Section 3.8), mainly by creating future electricity networks for the year 2050 based on supply and demand projections for the UK. Two future electricity scenarios were considered, where 70% of the generation mix in the electricity supply would be made up of renewables: (1) *Hydro70* – Where domestic heating would be predominantly provided through hydrogen gas; and (2) *Elec70* – Where demand for heating by electrification would be very high. The future electricity network has about 820 more new links due to adding new interconnectors and renewable energy (solar, batteries, onshore and offshore wind) sources to the current electricity network.

We performed a systemic assessment of the future network failures in a similar manner to the current networks. The analysis showed that, for the baseline single connection case, in comparison to the current electricity network-initiated failures there are about 199 (2.7%) fewer instances of cascading failures in the future networks, which was due to some additional network redundancy created in future electricity networks by adding new renewable sources. When the degrees of connections were increased to two (2C) there were about 104 (8.3%) fewer instances of cascading failures in the future networks than the current networks, and for the three degree of connections case (3C) there are about 78 (8%) fewer instances of cascading failures in the future networks. Few differences were seen in future failure propagation initiated in the telecoms networks. For all the high impact events the user disruptions in the future increased in proportion to increased demands from projected population increases in the future. But there were significant numbers of events where the impacts were almost eliminated. These instances were the ones where adding future generation capacity seems to have provided gains in terms of reducing the impacts.

We assumed that future economic impacts would grow based on compounded GDP growth forecasts for the UK. Assuming 1.9% GDP growth rate projection till 2050, the analysis showed that the worst-case economic output losses in the future baseline case would be as high as £14 million/day and mostly economic losses would be 1.9 – 2 times current baseline loss levels. Applying the resilience enhancing options, explored in the current scenarios, to the future networks showed similar gains across sectors when reducing the averaged disruptions for the 50 worst-case future baseline events. The future baseline disruptions were reduced by 85%-92% with a combination of increased connections and backup supply (2C+B and 3C+B) being most effective. All these disruptive impacts in the future baseline case were in excess of 1 million users/day and £1 million/day added across all networks and economy and were as high as 10 million user/day and about £14 million/day.

Another possible option for enhancing resilience of the future electricity networks was to consider the possibility that Electric vehicles (EV) could be used as backup supply options for residential consumption, when the grid supply would be disrupted. We explored this option by analysing the total disrupted electricity demand load in MW versus the user disruptions and the proportion of this demand that could be satisfied by the installed EV capacities in MW that existed at the locations of disruptions. The analysis showed that the installed EV capacity had more potential of being effective as a backup in the Hydro70 future scenario, in comparison to the heat demand intensive Elec70 scenario. For the Hydro70 scenario between 20%-40% of the disrupted MW demand load could be satisfied by installed EV capacity for some of the high user disruption events, and the percentages were in excess of 60% for some instances where user disruptions were between 1,300 – 170,000 residential customers. Generally lower values of user disruptions would occur at locations of sparse populations, where the electricity grid connections and accessibility might not be very good. Hence, repairs to restore the electricity supply to such locations might take time, making in worthwhile to explore the EV's as a source of supply to households.

1.3 Quality assurance

This study explored the possible impacts of infrastructure failure events that have not been observed in the past. Because the analysis deals with rare events that have not been observed it is challenging to validate it. Nonetheless, to help ensure that the results are robust and provide a credible basis for policy decisions, we have done a series of quality assurance (QA) checks throughout the duration of this study. Some of the QA actions are described below:

1. The methodology is based on previous research that has been published in peer-review journals and widely cited in the scientific and practitioner communities. These papers are cited throughout this report. Thus, the methodology has passed the standards of independent academic peer review.
2. The infrastructure data used in this study has been created from the latest best-known open-source resources on each sector, such as Ordnance Survey, Google Maps, OpenStreetMap, UK government websites, and network operators' data portals. In several instances geospatial network assets locations and connections information were verified with satellite imagery to improve the network spatial accuracy. Because our data sources are open and publicly available, they can be verified by third parties. See Appendix D for data sources.
3. We have conducted a thorough internal peer review of this report with team members who are well-known experts in infrastructure network modelling and systems analysis.
4. There has been continued dialogues and weekly meetings with the NIC throughout this project. NIC have arranged expert review of some aspects, which has been documented and discussed with the research team.
5. The NIC arranged face-to-face and virtual stakeholder meeting with academics and sector experts to assist with data collection, model assumptions, model validation and review of the interim results.
6. All assumptions and limitations of this study have been clearly stated throughout this report and are also summarised in Appendix C.

2. METHODOLOGY

2.1 Network modelling

We define *infrastructure systems* as *the collection and interconnection of all physical facilities and human systems that operate in a coordinated way to provide infrastructure services*[2]. This definition is relevant here because the scope of our study is specific to understanding the impacts of physical vulnerabilities to physical infrastructure systems. The continuous availability of reliable infrastructure services is crucial for economic prosperity and long-term sustainability[3]. Hence, the use of the term *economic infrastructure*[4] to refer to the systems under consideration in the study.

Economic infrastructure are large-scale spatially distributed systems with complex interactions that deliver essential services to society and the economy. It is difficult to develop unifying models that can completely represent the underlying collection and interconnection of all physical facilities and human systems to a suitable level of complexity. Several modelling approaches, each with their strengths and limitations, have been used for modelled infrastructure systems in the context of risk and resilience analysis. For most recent detailed literature reviews of different models and methods see Ouyang (2014)[5], Hosseini et al. (2016)[6], Saidi et al. (2018)[7]. We have adopted a *network modelling approach* to suitably represent the infrastructure systems for the purposes of this analysis. Such an approach, embedded in network-science theories[8,9] and widely applied to real world cases[5,6,7,10], is most suitable for this study because we can leverage upon previously created data and models[11,12,13,14,15]. Some of these are discussed later in this document.

A *network* here is defined as a *collection of nodes joined together by a collection of links*. *Nodes* are point representations of key locations of physical facilities and human systems in the infrastructure systems – electricity substations, water treatment plants, rail stations, etc. *Links* are line representations of physical connections between node pairs – electricity overhead cables, road sections, railway lines, etc. Links could also represent notional connections by joining straight lines between node pairs, to represent interactions that are not physical. The term *asset* is also frequently used here in this report to refer to network nodes and links. The

2 Hall, J.W., Tran, M., Hickford, A.J., & Nicholls, R.J. eds. (2016). *The Future of National Infrastructure: A System-of-Systems Approach*. Cambridge University Press.
3 https://www.nic.org.uk/wp-content/uploads/CCS001_CCS0618917350-001_NIC-NIA_Accessible.pdf
4 https://www.nic.org.uk/wp-content/uploads/NIC_Resilience_Scoping_Report_September_2019-Final.pdf
5 Ouyang, M. (2014). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability engineering & System safety*, *121*, 43-60.
6 Hosseini, S., Barker, K., & Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience. *Reliability Engineering & System Safety*, *145*, 47-61.
7 Saidi, S., Kattan, L., Jayasinghe, P., Hettiaratchi, P., & Taron, J. (2018). Integrated infrastructure systems—A review. *Sustainable cities and society*, *36*, 1-11.
8 Lewis, T. G. (2011). *Network science: Theory and applications*. John Wiley & Sons.
9 Barabási, A. L. (2016). *Network science*. Cambridge university press.
10 Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, *94*(2), 125-141.
11 Thacker, S., Pant, R., & Hall, J. W. (2017). System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures. *Reliability Engineering & System Safety*, *167*, 30-41.
12 Pant, R. Hall, J.W. and Blainey, S.P. (2016). Vulnerability assessment framework for interdependent critical infrastructures: case study for Great Britain's rail network. *EJTIR*, 16(1): 174-194, ISSN 1567-7141.
13 Thacker, S., Barr, S., Pant, R., Hall, J. W., & Alderson, D. (2017). Geographic hotspots of critical national infrastructure. *Risk Analysis*, *37*(12), 2490-2505.
14 Pant, R., Thacker, S., Hall, J. W., Alderson, D., & Barr, S. (2018). Critical infrastructure impact assessment due to flood exposure. *Journal of Flood Risk Management*, *11*(1), 22-33.
15 Oughton, E. J., Ralph, D., Pant, R., Leverett, E., Copic, J., Thacker, S., ... & Hall, J. W. (2019). Stochastic Counterfactual Risk Analysis for the Vulnerability Assessment of Cyber-Physical Attacks on Electricity Distribution Infrastructure Networks. *Risk Analysis*.

description or quantification of the arrangement of nodes and links is called the network *topology*.

In addition to topology the *functional attributes* of network nodes are also needed to be able to assign the direction of *flow* of resources[11]. There are three types of node functions that are included in the network model: (1) *source/origin nodes* – from where network services are generated or originate; (2) *sink/destination nodes* – from where network services are delivered to users or other networks or where the end of the service happens; and (3) *intermediary nodes* – that transmit network services from the source nodes towards the sink nodes. Between a chosen source and sink the flow of services is traced along a *directed flow path*, which includes all the assets traversed in the direction from the source to the sink. Overall all possible directed flow paths that can be traced between sources and sinks provide us with a complete understanding of how the network topology facilities the flow of services.

With growing recognition that infrastructure systems do not exist in isolation, the main interest in research[5] and policy (especially for the NIC)[16,17] is in understanding their *interdependencies*[18], which represent the mutual interactions between different types of infrastructure systems. For this study as well, the key consideration is to understand and model how interdependencies between networks influence vulnerabilities. While there have been several ways in which infrastructure interdependencies have been conceptualized[5], the interpretations of Rinaldi et al. (2001)[18] apply the most to the context of this study because they are described in the context of disruptions. Utilising Rinaldi's characterizations, network interdependencies of interest include: (1) *Physical* – where two nodes are physically connected by a link to exchange material outputs, so the failed state of one influences the other; (2) *Cyber* – where the state of a network asset depends on information transmitted through information infrastructure, so it fails due to cyber failures; (3) *Geographic* – when multiple network assets are in close geographical proximity, making them susceptibility to fail from the same external shock events; and (4) *Logical* – which explain how network asset failures link to users (customers) and economic systems (industry sector) that go beyond physical, cyber or geographic interdependencies. The flow path mapping also creates *functional interdependencies*[11,12,13,14] which include the functional understanding of flow of resources across physical systems using the wider network topology.

In the network models built for this study the interdependencies (or dependencies) are translated into directed network links to infer the flow of services between networks. *In most cases the network representations capture functional (inter)dependencies, which result from physical (inter)dependencies*. Having considered telecoms as one of the infrastructures, we also account for cyber-physical dependencies on telecom assets. By mapping customers and the economic impacts of infrastructure disruptions we also account for logical (inter)dependencies. One of the key challenges of modelling networks connectivity to represent their real-world connections is the lack of data to inform such connectivity. This is especially and most critically true to mapping interdependent connections. For example, if we knew that a particular railway station derived its electricity from a known electricity substation, then we can create a *notional link* between the two in the network model if the actual overhead/underground cable information is not known. This level of accurate data might be available for some locations in the country, but it is currently next to impossible to procure for the whole national-scale

16 https://www.apm.org.uk/media/18859/national-infrastructure-briefing-lr-v2.pdf
17 https://www.nic.org.uk/wp-content/uploads/NIC_Resilience_Scoping_Report_September_2019-Final.pdf
18 Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE control systems magazine*, *21*(6), 11-25.

analysis. Hence, where data is not available, but it is known that two types of sector assets should be connected, we assume that they connect by creating straight line links between the right kind of assets nearest to each other. In most cases this assumption is quite valid because the nearest connection represents the path of least resistance of service flows and is also most cost effective in terms of materials and design of systems.

Figure 2-1 shows a schematic representation of the network topology and directed connections between sources and sinks within a network and the dependent links across networks.

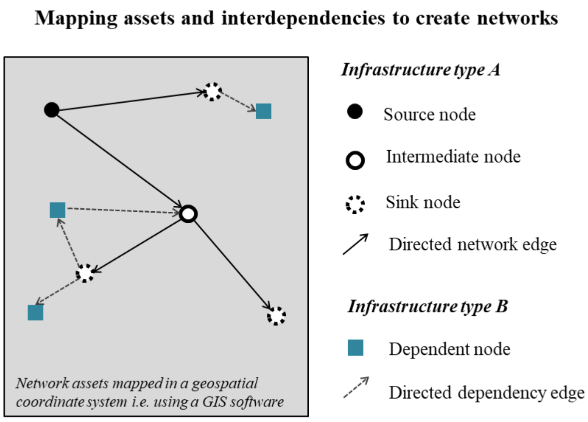


Figure 2-1: Schematic representation of network topology and directed dependencies across sectors.

While conceptualising infrastructure networks it is also assumed that they are organised in a layered hierarchical structure, where larger nodes with wider national-scale network influence are at the top of the hierarchy and smaller nodes with localised network influence are at the bottom[11,14]. A typical example of this is the electricity network (see Figure 3-1) in which the big power generation sites form the top layer, followed by the transmission network (400kV) substations layer below, going all the way to the lowest substations (6.6kV) that supply power to customers/households.

Based on the definitions outlined above, Figure 2-2 (adapted from Thacker et al. 2017[11]) shows a final generalised system-of-systems (network-of-networks) representation of all networks built for this study. As can be seen from the figure each network can be conceptualised in a layered network structure where goods and services are delivered to the *customers* who are the common metric across sectors. While mapping interdependencies between different infrastructure networks the appropriate layer of connections is selected to represent the flow of services across systems.

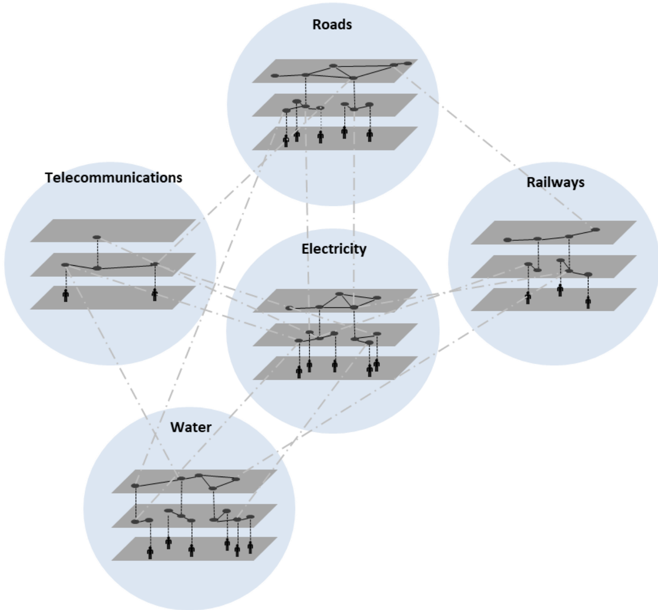
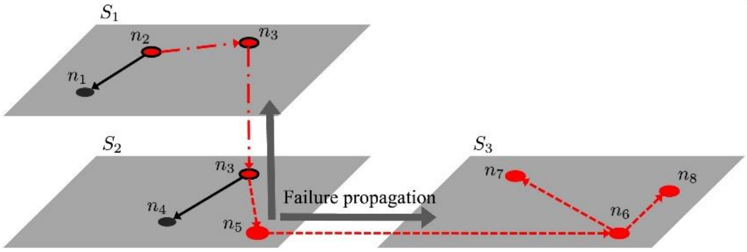


Figure 2-2: System-of-systems conceptualisation of infrastructure networks and their interdependencies (adapted from Thacker et al. 2017[11]).

2.2 Failure and impact analysis

Following the creation of network models, failure analysis involves removing nodes or links, individually or several, to trigger an *initiating event* that might lead to further failure cascades. Throughout the analysis it is assumed that failure meant that a node completely lost its service. Partial failure states, where nodes might still be operating at below 100% operational levels and providing reduced service were not considered. The assumption of total loss of service is considered appropriate because we are interested in understanding worst-case scenarios of large-scale widespread disruptions. There are two ways in which the cascading effects proceed: (1) to the nodes and links in the closest neighborhood of the initiating asset; and (2) assets farther away that stop receiving service because their flow paths included the initiating asset, which is now discontinued. An illustration of failure initiation and propagation conceptualized across multiple networks (layers) is shown in Figure 2-3 (from Thacker et al. 2017[11]), where edges is another term used for links. Here the failure is initiated in node n_5 in system S_2, following which all nodes in system S_3 fail because they either lose their dependency (node n_6) or all flow paths directed towards them (n_7,n_8). The failure propagation also affects nodes (n_2,n_3) directed towards to n_5 because the services delivered by them cannot reach further, due to which there might be some loss of service.



Operational nodes		$((n_1, S_1), (n_4, S_2))$
Failed nodes (complete)		$((n_5, S_2), (n_6, S_3), (n_7, S_3), (n_8, S_3))$
Failed nodes (incomplete)		$((n_3, S_2), (n_3, S_1), (n_2, S_1))$
Operational edges		$(((n_2, S_1), (n_1, S_1)), ((n_3, S_2), (n_4, S_2)))$
Failed edges (complete)		$(((n_3, S_2), (n_5, S_2)), ((n_5, S_2), (n_6, S_3)), ((n_6, S_3), (n_7, S_3)), ((n_6, S_3), (n_8, S_3)))$
Failed edges (incomplete)		$(((N_3, S_1), (n_3, S_2)), ((n_2, S_1), (n_3, S_1)))$

Figure 2-3: Schematic representation of failure propagation across networks (from Thacker et al. 2017[11])

The failure impact or *network vulnerability* is the measure of the service provision affected due to failures of network nodes and links from external shock events[12]. In this study affected service provision is measured in terms of the aggregated numbers of customers disrupted or value of service lost over a *service demand area* associated with each disrupted sink node. The aggregated numbers of customers disrupted or value of service, called *service demands*, are first grouped at detailed spatial disaggregations, which differs for each sector. For example electricity service demands assocaited with sink substations are all first grouped at the Local Super Output Area (LSOA) which are roughly 41,000 area polygons across Great Britain, while water service demand areas are grouped to sink nodes at coarser resolutions of 128 Water Resource Zones (WRZs). The electrtcity and water servcie demand areas are then grouped at their sink nodes. The service demands in terms of customer numbers depend of census data on only residential customers that can be mapped and grouped to the service demand areas for the specific sector’s sink nodes. For transport networks the service demands are estimated only as total passenger (customer) flows along nodes and links, since one of the main services provided by transport is the mobility of people. Unlike utility networks the service demand areas of transport assets are not limited to fixed areas. Hence, we model transport origin-destination (OD) flows in this study and assign them statically along the flow paths to infer the volumes of passenger (customer) trips assigned and subsequently disrupted. This also creates a distinction in the way the impacts are estimated in utility networks and transport networks. In the former impacts are measured for only those nodes that lose all service when they no longer have acess to any flow path, while in the latter impacts are measured for nodes that also lose part of their pre-disruption journeys as there might be reduced numbers of flow paths through them. Details of each sector’s demand mapping are provided in Section 3.1 – 3.5.

In order to capture the cascading effect of interdependent network failures, a distinction is made between the network of the initiating event and every subsequent failure propagation to other networks. Figure 2-4 shows the schematic representation of a *direct service demand disruptions* in the network where the initiating event (marked X) takes place, while the *indirect service demand disruptions* happen in the dependent network due to loss of service from the initiating failure network. In this study we are interested in tracking the number of failure sequences that trigger indirect service demand disruptions. Hence, we use the term *Order 0* to represent a direct (initating) service disruption network effect and subsequently *Order n* (>0) to track futher sequences of indirect service demand disruptions. In the exaple demonstration of Figure 2-4 there is an Order 1 indirect service demand disruption, showing the failure propagated once across networks.

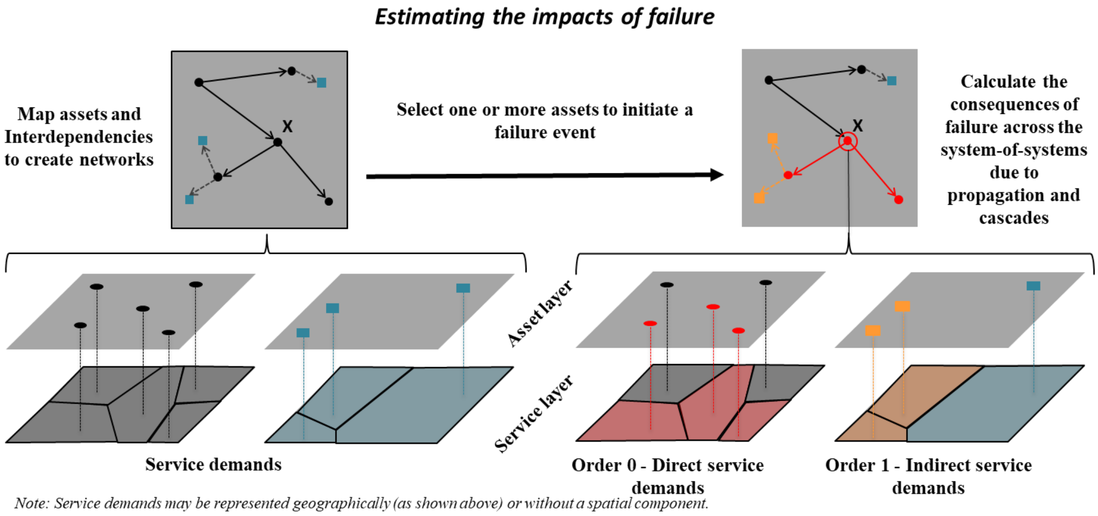


Figure 2-4: Representation of direct and indirect service disruptions across interdependent networks.

Another vulnerability metric estimated in this study is the macroeconomic loss occurring in the whole economy comprised of infrastructure and non-infrastructure sectors. We use a demand-side Leontief Input-Output (IO) model[19] for estimating the macroeconomic losses across 129 sectors that make up the UK national accounts[20]. The macroeconomic model is not spatially disaggregated below the UK-scale. The model translates the customer disruptions due to infrastructure failures into household demand losses, which signify *direct economic losses*. Subsequently *indirect economic losses* are estimated by balancing the economic output supply to meet reduced demands. The final outcome of the IO analysis is to produce loss estimates in £/day. Details of the IO model are given in Section 3.10.

2.3 Incorporating resilience

2.3.1 Adding backup supply

The term *resilience*, which has gained a lot of prominence in literature[6], involves assessing the ability of the system to provide infrastructure services including the ability to adsorb, adapt and recover from shocks or gradual changes[21]. Infrastructure network resilience is quantified by measuring the vulnerability along with the duration of recovery of assets and networks. In this study the recovery dimension of resilience is not considered, mainly due to lack of data and understanding of how long disruptions last and what measures of recovery planning are put in place by infrastructure operators, regulators, and users (households and businesses). Nonetheless another approach to quantify some resilient behavior in systems is considered by assuming the disruptions last over a certain time frame and are delayed in some assets due to the provision of backup supply to maintain service if the supplying network fails. These backup supply options characterize two elements of resilience here: (1) *Robustness* – The ability of a network to absorb the initial shock and continue operating at a certain level of functionality after disruption; and (2) *Redundancy* – The ability of the network to absorb the initial shock

[19] Leontief, W. (Ed.). (1986). *Input-output economics*. Oxford University Press.
[20] https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/datasets/ukinputoutputanalyticaltablesdetailed
[21] From NIC Terms of Reference

impact by providing alternative connection options when disrupted. For all asset backup durations are assumed to be probabilistic, which reflects the uncertainty in the durations of backup supply in providing resilience against the spread of disruptions.

2.3.2 Changing degrees of interdependencies

Network redundancy is also captured in the network's flow path characteristics by tracing allowable flow paths from one source to several sinks and vice versa, thereby guaranteeing connectivity of the flows if a single source-sink flow path is affected. In the case of interconnectivity, the redundancy is very low because mostly it is assumed that between two networks assets connect in single pairs. What this means is that a single railway station is assumed to derive its electricity for a single electricity substation. In reality this might not be the case, especially for large nodes (major power plants, stations, telecoms) in the system. To overcome the data gap in the model we have tested the failure outcomes under three varying degrees of connections described as following:

1. *One connection mapping* – where each selected asset of one infrastructure is connected to one asset of the infrastructure it is dependent upon. For example, linking each railway station to its nearest electricity substation for electricity supply.
2. *Two connection mapping* – where each selected asset of one infrastructure is connected to two assets of the infrastructure it is dependent upon. For example, linking each railway station to its nearest two electricity substation for electricity supply.
3. *Three connection mapping* – where each selected asset of one infrastructure is connected to three assets of the infrastructure it is dependent upon. For example, linking each railway station to its nearest three electricity substations for electricity supply.

The aim of adding more connections is mainly to test if there are any gains in reducing the disruptive impacts across sectors if there were more redundancy between networks. In some cases, it might not represent the actual cross-sectoral connections, especially if the second or third nearest dependency node might be much farther than the nearest one. In such cases we have assumed a distance threshold of 10km to truncate the creation of dependency links, assuming that links longer than this will be unrealistic.

2.4 Changing networks in the future

A key interest for the NIC was to know how network vulnerabilities might evolve under specific future planning scenarios. The information for future scenarios mainly comes from the National Infrastructure Assessment (NIA) published by the NIC[22]. For understanding changing vulnerabilities due to these scenarios, some high-level recommendations from the scenarios were taken and translated into changes in the network models and their interdependencies.

Though different NIA scenarios have different timelines, we assumed they would be achieved fully by the time at which we analysed the changes to the networks and their resulting vulnerabilities and resilience. Hence, the general principle was to represent 'one state' of the networks each in the present and the future where:

1. *State* is the static representation of: (A) Network topology; (B) Network flows; (C) Customer demands; (D) Economic losses.
2. *Current state* – Whatever latest data we can get show the state as a representative of the year 2017, which was chosen based most of the current data was latest to this year.
3. *Future states* – Inferred data based on the scenarios in the year 2050.

22 https://www.nic.org.uk/wp-content/uploads/CCS001_CCS0618917350-001_NIC-NIA_Accessible.pdf

Based on the data received from the NIC three future planning options were considered in the analysis, as shown in Table 2-1. The Hydro70 and Elec70 are electricity specific scenarios where 70% of the generation mix in the electricity supply would be made up of renewables but heating is predominantly provided by hydrogen gas and electrification respectively. The choice of 70% was based on the NIC's assessment that these would be the most realistic futures given the current renewable energy trajectory and future nuclear phasing decisions being made in the UK. Their implications on the network analysis are discussed below, while the details of the underlying data are described later. In our analysis we study the effects of Hydro70 + 100% EV sales as one case, and Elec70 + 100% EV sales as another.

Table 2-1: Future scenarios from the NIA and their translation in network topology, flow, and failure models.

<table>
<tr><th>Future scenarios</th><th>Network topology modifications</th><th>Flow/demand modifications</th><th>Implications on failure analysis</th></tr>
<tr><td>1. Hydro70 – Electricity generation is mainly driven by increased renewable uptake with lower gas, oil and coal uptake and domestic heating is predominantly provided through hydrogen gas</td><td rowspan="2">• Electricity network topology changes due to adding and removing new source nodes
• All other networks topologies remain the same
• New interdependent connections added due to new electricity nodes</td><td rowspan="2">• A 2050 electricity demand profile from aggregated estimates[23] is merged with a spatial electricity demand model
• All sector customer demands change based on future population projections</td><td rowspan="2">• Topologically changes in the electricity network will change the flow paths and hence disruption outcomes
• Increased customer disruptions due to population in increases will be seen for other networks</td></tr>
<tr><td>2. Elec70 – Electricity generation is mainly driven by increased renewables supported by gas and demand for heating by electrification is very high</td></tr>
<tr><td>3. Preparing for 100 per cent electric vehicle sales</td><td>• No changes to road or electricity topology
• All other networks remain the same</td><td>• Added transport EV demand will add more load onto the electricity network</td><td>• Will increase electricity service demand losses
• EV demands will be tested as alternative backup supply options</td></tr>
</table>

[23] https://www.ofgem.gov.uk/ofgem-publications/55666/157018blensappendices.pdf

2.5 Methodology implementation

To build the networks models and estimate the network vulnerability outcomes, under different assumptions described above, the methodology and implementation steps for spatial vulnerability assessment are explained in Table 2-2. These steps are based on system-of-systems methodological approaches built previously to inform assessments at the national-scale (Great Britain)[24].

Table 2-2: Methodology and implementation steps in estimating the relative importance of vulnerability characteristics

Step 1. Topology creation	Assemble disjointed spatial nodes (points) and edges (line) assets
	Connect nodes pairs by physical or notional edges
	Identify connections between networks
Step 2. Flow assignment	Assemble data to assign attributes to nodes and edges – source-sink characteristics
	Get data on flow performance metric of network – source supply volumes – sink demand values
	Map all source-sink paths and assign static flows on paths
Step 3. Customer assignments	Assemble data on customer demands at sink nodes
	Infer customer demands by combining asset service areas with census/building stock data
Step 4. Economic losses	Build economic Input-Output (IO) model
	Link infrastructures to economic sectors
	Translate flow and customer disruptions to direct economic flow losses
	Estimate indirect economic flow losses from IO model
Step 5. Estimate vulnerability characteristics/metrics	Quantify characteristic/metric in 3 stages – Only based on topology – Topology + static flows
Step 6. Failure analysis	Rank nodes and edges based on failure outcomes
Step 7. Results	**Direct and indirect estimates of** – **Number of nodes/edges affected; proportion of the network affected** – **Number of people affected** – **Macroeconomic impacts** – **Spatial location of the impacts** – **Spatial clustering of the impacts** – **Spatial extension of the impacts**
Step 8: Incorporating backups	– Perform the analysis by assuming the failures last over a certain time period and some disruptions are delayed due to backup supply – Incorporate uncertainty in the durations of backup supply for each asset
Step 9: Future network changes	Incorporate all future scenario changes in Step 1-8

24 Pant, R., Thacker, S., Hall, J., Barr, S., Alderson, D., & Kelly, S. (2016). Analysing the risks of failure of interdependent infrastructure networks. *The Future of National Infrastructure: A System-of-Systems Approach*, p241.

3. UNDERLYING DATA AND ASSUMPTIONS

This section describes the infrastructure network data on electricity, telecoms, water, railways, and roads assembled for this study. For each infrastructure the network topology structure is explained, with the flow metrics, failure modelling assumptions taken in this study, the spatial aggregations of customers, and the spatial scale of the models. The assumptions taken in creating interdependencies across networks are also explained, along with the assumptions about backup supply. The future network changes and data are also discussed in detail. The data accessibility issues associated with harnessing data to build the models are described throughout. It is noted here that although most of the raw data available for such models is available online, such data were in various formats and contained data gaps that had to be corrected in order to translate them into the network models. Hence, while raw data was obtained from existing open-source resources, the final network created is an original ITRC product that cannot be found anywhere else.

3.1 Electricity network

The electricity network representation in this study consisted of identifying the power generation sites and substations and joining them with overhead and underground cables. The main aim of this model was to capture the possible ways in which electricity is delivered from power generation sites to the transmission grid, and then from the distribution networks towards the final users. The model represented the locations of key power generation sites, smaller embedded generation sites, 400kV and 275 kV substations in the transmission network, and 132kV, 66kV, 33kV, 11kV and 6.6kV substations in the distribution networks. The 11kV and 6.6kV substations represented the lowest voltages that connect to customers.

3.1.1 Network topology

The network topology, represented as a hierarchical network, is shown in Figure 3-1. Here each hierarchy is connected to the one below it, but there might be several connections that skip one or two hierarchies and connect to the lower levels directly. The overall network consisted of 18,061 geolocated nodes out of which 2,565 represented the generation sites. There were 13,245 links representing overhead lines and underground cables. The locations of the nodes were collected and verified from several sources and meticulously checked with satellite imagery as best as possible. Several of the substation locations data at the distribution level were simply obtained from Google Maps and OpenStreetMap. Similar data sources were used for geolocating the link information, which has lesser accuracy in terms of the geometries but more accuracy in terms of connecting the right types of nodes to each other.

The links within the same layer in the hierarchy were bidirectional to represent the possibility that electricity would flow in both directions. But the links between with the transmission (275kV – 400kV), High Voltage (HV) (66 kV – 132 kV) and Low Voltage (LV) (< 66kV) distribution layers were directed to show the step-up and step-down transformers that convert electricity voltages before they are distributed. This meant that in the creation of source-sink flow paths the direction of flow was always from transmission to high voltage to low voltage network nodes.

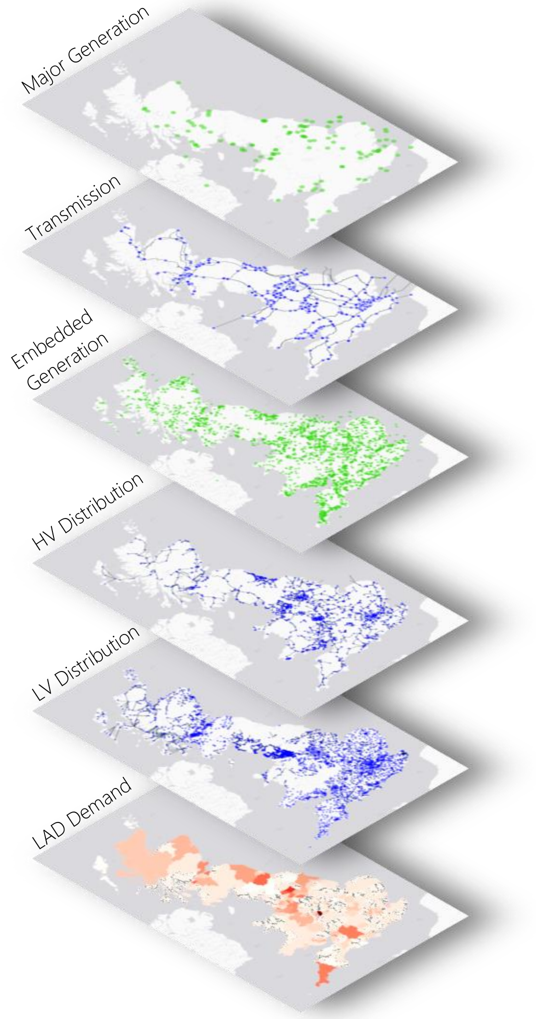


Figure 3-1: Topological representation of a hierarchical electricity network for Great Britain.

3.1.2 Demand allocation

There were two types of demand allocations for electricity nodes: (1) in terms of the loads in MW; and (2) the numbers of customers of electricity. Both these demands were estimated at 4,897 sink nodes corresponding to mostly the 11kV and 6.6kV substations. Also, data on the supply capacities of the generation sites was collected to identify the source nodes and also to check that supply was greater than the demand. The allocations of demands in MW was first done at the 380 Local Authority District (LAD)[25] administrative area levels for Great Britain, using an energy demand model[26] that accounted for household and industry usage of electricity at every hour throughout the year. We extracted the peak hourly demand over the whole year from this model, because we were only concerned with assessing one state of the system and the peak load would be the state when the network is under most stress.

[25] https://geoportal.statistics.gov.uk/datasets/local-authority-districts-december-2017-full-clipped-boundaries-in-great-britain

[26] Eggimann S, Hall JW, & Eyre N (2019). A high-resolution spatio-temporal energy demand simulation to explore the potential of heating demand side management with large-scale heat pump diffusion. Applied Energy, 236, 997-1010.

The LAD level data was further disaggregated and grouped to the Local Super Output Area (LSOA)[27] level of which there were 41,667 polygons in Great Britain. The disaggregation at this finer scale was done by assuming the energy usage within each LSOA was in proportion to its building areas, where the data from building footprints was obtained from the Ordnance Survey (OS) MasterMap[28]. From the LSOA levels the demands were aggregated or grouped at the sink nodes based on identifying the nearest nodes for each LSOA. Both MW and population demands were allocated with this method, which in the end resulted in allocating demands at the sink node levels of the network.

3.1.3 Failure analysis

Electricity network failures were estimated in terms of the numbers of customers at the demand nodes disrupted when some nodes were removed from the network. Since each demand node had customers on it, it was straightforward to assume that all those customers would be disrupted if their demand node failed. For every other node failure, the possible disruptions in all flow paths through the node was checked to infer if there would be any resulting disruption.

First, we mapped all the possible *directed flow paths* between every source node and sink node in the network. This was done because it was assumed that if there were a failure anywhere in the network then electricity service flow would still be maintained as long as there was a source to supply electricity and a functioning path to the sink nodes. Given the large numbers of sources (2,565) and sinks (4,897) the path mapping resulted in creating 1,002,837 unique source-sink paths. By mapping so many flow paths we are accounting for the redundancies in the network, in terms of maintaining electricity supply when some source-sink flows would not work. Given that the links are directed from the transmission to HV and LV distribution levels, the flow paths are directed accordingly, with no sources connected at the lower levels supplying to sinks at the upper levels in accordance to the expected flow of electricity. Previous studies[11,13] have shown that this approach gives a reasonable estimate for realistic failure outcomes of network failures.

When a failure was initiated in the electricity network all the paths containing the failed nodes were considered disrupted and removed from the set of flow paths. If there were further nodes that lost all their flow connectivity due to the removal of the disrupted flow paths, then these were also considered to have failed due to complete loss of any flows through them. If any of the final set of disrupted nodes were demand nodes, then their allocated demands were summed up to estimate the disrupted customers.

3.2 Digital communications network

Digital communications consist of three main types of technologies including fixed networks (fibre/coaxial/copper etc.), wireless terrestrial networks (cellular, WiFi, Tetra, etc.), and satellite networks (geosynchronous, low or medium earth orbit)[29]. In this analysis we focussed on the main fixed and wireless terrestrial networks. The coverage of these technologies in this study was over Great Britain.

[27] https://data.gov.uk/dataset/fa883558-22fb-4a1a-8529-cffdee47d500/lower-layer-super-output-area-lsoa-boundaries
[28] https://www.ordnancesurvey.co.uk/business-government/tools-support/open-mastermap-programme
[29] Oughton, E.J., Tran, M., Jones, C.B., Ebrahimy, R., 2016. Digital communications and information systems, in: The Future of National Infrastructure: A System-of-Systems Approach. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781107588745.010

Figure 3-2 illustrates the system modelled in this analysis consisting of:

- *A core network* – a high-capacity long-distance transportation network consisting of fibre optic cables.
- *An internet exchange network* – local access consisting of either fixed fibre, coaxial cable or copper.
- *A cellular network* – consisting of wide-area macro cells, as well as a smaller number of local high-capacity small cells.

Figure 3-2 also shows the connections between the exchanges and macro cells to other network assets, which is discussed in detail later.

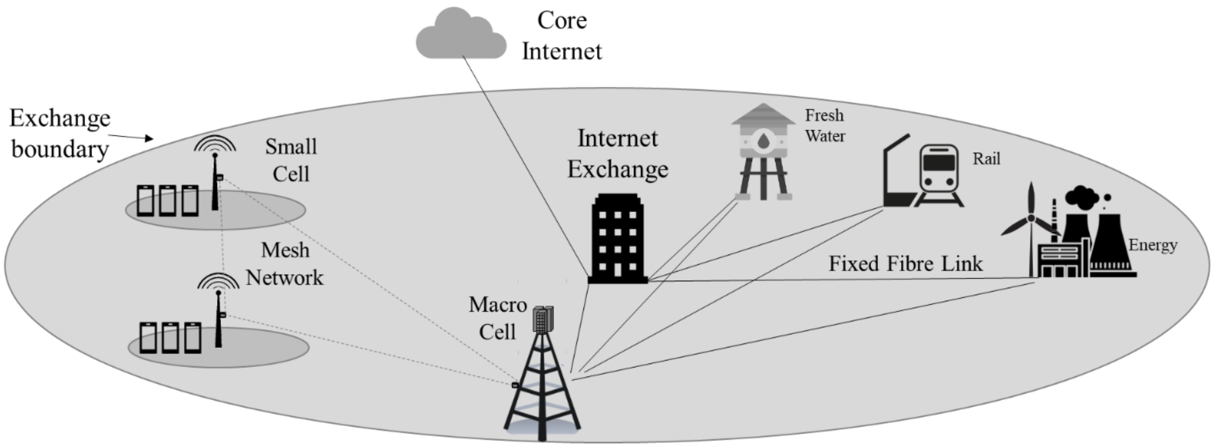


Figure 3-2: Schematic model of the digital communications system structure and its connections to other sectors.

Digital communications assets are cheaper and easier to deploy than other infrastructure sectors. For example, it can take numerous decades to plan, design and build a high-speed railway or nuclear power plant. In contrast, the deployment of a new generation of cellular technology, such as 5G, is estimated to take ~7 years to reach most of the population (90%)[30]. Hence, the digital communications sector experiences generational changes on a decadal basis. Data availability is a serious problem which constrains the type of analysis that can be undertaken for digital communications networks[31]. Most digital assets are deployed by private companies and therefore data on precise location, or capacity of coverage information, can be limited as this is treated as commercially sensitive. Although governments do have the power to obtain this data from private operators, as there can be hundreds of operators this is usually only undertaken for the largest asset owners.

Considering this context, this analysis focused on the main operators. One of the largest owners is BT, formally known as British Telecom, which owns the previously nationalised networks of telephone *exchange assets*. We had limited information on the network topology of the BT network, except for some information reported on several open websites. Figure 3-3 reports the information we had, which included approximately 20 core node locations, 86 metro nodes, 1000 Tier 1 Multi-Service Access Nodes (MSANs) and 4,400 small and medium exchanges.

[30] Oughton, E.J., Frias, Z., 2018. The cost, coverage and rollout implications of 5G infrastructure in Britain. Telecommunications Policy, The implications of 5G networks: Paving the way for mobile innovation? 42, 636–652. https://doi.org/10.1016/j.telpol.2017.07.009
[31] Oughton, E.J., Frias, Z., Dohler, M., Whalley, J., Sicker, D., Hall, J.W., Crowcroft, J., Cleevely, D.D., 2018. The strategic national infrastructure assessment of digital communications. Digital Policy, Regulation and Governance 20, 197–210. https://doi.org/10.1108/DPRG-02-2018-0004

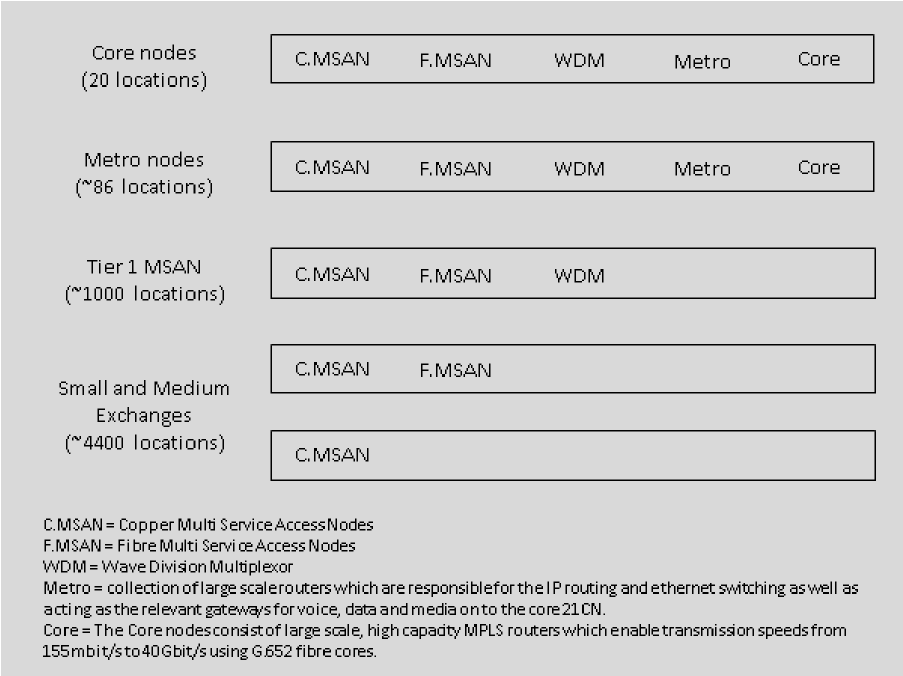


Figure 3-3: Hierarchical network architecture of the BT telecoms exchanges.

3.2.1 Network topology

To translate the network concept into spatially network topology, several datasets were used in the analysis. Firstly, we obtained information on the approximate service areas of over 5,000 exchange (the fixed network) by mapping them to ~1.5 million postcodes served across all exchanges in Great Britain. Postcode data was also required to map this information into exchange boundary areas (as emphasised already in Figure 3-2). After the service areas of the exchanges were created their node locations were approximated as the centroids of each area polygon.

For estimating core locations and other layers of the fixed network, information on the BT's 21st Century Network (21CN) was obtained. A total of 85 exchanges were identified as metro nodes, with 12 of these being outer code nodes, and 8 being inner core nodes. Inner core nodes were fully meshed (connected) to all other inner core nodes and outer core nodes were triple parented (connected) to the inner core. Metro nodes were dual parented (connected) to the nearest core nodes, and then all lower level exchanges were dual connected to the nearest two exchanges. Remote areas and islands were treated separately, and exchanges on such areas were connected to each other via a minimum spanning tree[32] (by connecting all exchanges with the least number of links, such that each exchange pair connects only to its closest exchange) and then connected to the mainland via the nearest Tier 1 MSAN exchange.

Cellular asset data was taken from online sources and pre-processed to identify single site *macro cell* locations by buffering all points by 50 meters[33], dissolving overlapping shapes and estimating the site location by using the polygon centroid. This resulted in creating 33,062 macro cell nodes. Cellular site traffic was routed ('backhauled') into the internet exchange network using the straight-line path to the nearest serving exchange. This created a radial

[32] Graham, R. L., & Hell, P. (1985). On the history of the minimum spanning tree problem. *Annals of the History of Computing*, *7*(1), 43-57.
[33] Oughton, E.J., Frias, Z., Russell, T., Sicker, D., Cleevely, D.D., 2018. Towards 5G: Scenario-based assessment of the future supply and demand for mobile telecommunications infrastructure. Technological Forecasting and Social Change 133, 141–155. https://doi.org/10.1016/j.techfore.2018.03.016

network structure between the exchanges and clusters of macro cells dependent links to each exchange. Since the flow of information in the network could take place in both directions between any two connected node pairs, overall the network had 97,992 links to represent the connections and information flow between nodes in the network. Figure 3-4 shows the result of creating the different layers of the telecoms networks.

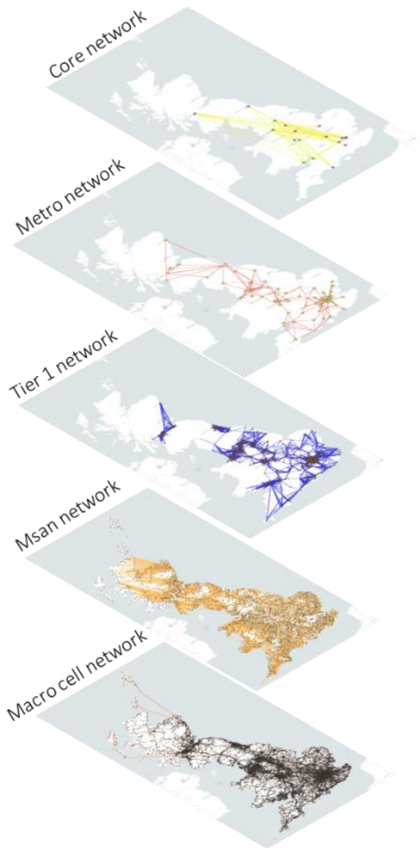


Figure 3-4: Topological representation of the different layers of the digital telecoms network for Great Britain.

3.2.2 Demand allocation

The demands allocated to the exchanges and macro cells were in terms of the numbers of customers assigned over their service areas. While the service areas of the exchanges were created from the data described in the previous sections, for the macro cells service areas were created by assuming each macro cell served locations nearest to it. This resulted in creating Voronoi polygons[11] as service area for the macro cells.

The population layer used to allocate customers to the telecoms nodes was at the Local Authority Distract (LAD) level, of which there were 380 polygons covering Great Britain with population estimates for 2017. First the LAD populations were disaggregated at the postcode level based on weighing by each postcode's coverage density intersecting a particular LAD polygon. This coverage density was estimated in terms of number of address point connections for telecoms at the post code levels. 4G information on coverage by local authority was also taken from Ofcom's Connected Nation report (2018)[34]. Postcode sector coverage was estimated by disaggregating local authority coverage, based on the thesis that Mobile Network Operators (MNOs) rationally upgrade sites in the highest population density areas first in order

[34] Ofcom, 2018. Connected nations 2018: UK report. Ofcom, London.

to serve as many potential customers as possible, but also to serve areas of high traffic with the most efficient technology. The postcode assigned populations were then intersected with the respective exchange and macro cell polygons to estimate the total populations assigned to these assets.

The above allocation of customer demands only accounted for fixed residential populations within LAD's. Given that mobile connectivity is variable throughout the day and in fact might be highest during the working hours of the day due to commercial usage, for the macro cells a further day time working population allocation was done to compare with the residential population estimates. Data for the working population was obtained from official labour market statistics[35] and Scottish Census data[36], but unfortunately was only updated to 2011 as the latest figures. Hence, for each LAD we compared the ratio of the working and residential populations in 2011 and multiplied by the 2017 residential census numbers to get the working population estimates. The final population chosen for a LAD was the maximum of the two estimates, which was then disaggregated to the postcodes and macro cell services areas as described before.

3.2.3 Failure estimation model

Failures in the telecoms network were estimated in terms of the numbers of customers of macro cells or exchanges disrupted when nodes are removed from the network. Since clusters of macro cells were assumed to be radially dependent upon one exchange, if the exchange failed then all the macro cells also lost service and hence customers. The failures of the exchanges not failed directly depended upon their connectivity to the core network, which had a lot of redundancies. Hence, in the network model we assumed that as long as there was a flow path (route) connecting an exchange to at least one of the core nodes, the exchange would not fail indirectly from failures at other locations of the network.

The numbers of telecoms customers disrupted due to failures of both macro cells and exchanges were estimated to be the minimum from the two types of nodes, based on the assumption that the least spatial coverage would be affected due to such failures. If an exchange failed and there were macro cells service areas within its boundary that were linked to another working exchange then the customers within those areas would be still able to get mobile coverage and assumed not disrupted. Although we acknowledge that this might not fully represent the impact on the customer, it is the best assumption possible without further data.

3.3 Water network

There was no detailed national-scale geospatial water pipe network data available in this study, which showed connections from water supply networks to water distribution networks to customer demand location points at disaggregated spatial scales (either household/building locations or even some postcode level). The best available model was a water supply resource system model of England and Wales developed for a previous study[37]. The data from this model was modified and adopted for this study.

The data included all major *public water supply nodes* (reservoirs, boreholes, transfers, water treatment works, pumped storage, desalination plants and river abstraction points) that were

[35] https://www.nomisweb.co.uk/census/2011/workplace_population
[36] https://www.scotlandscensus.gov.uk/news/workplace-population-and-daytime-population-council-areas
[37] http://www.mariusdroughtproject.org/

connected into England and Wales's wider water network via any river or transfer of significance (i.e. > 2Ml/d). This included more than 90% of England and Wales's population and water demand, and more than 80% of the combined land area. Some population and land areas were not accounted for because their either were not covered by the public water supply network or the water transfers in such areas were below 2Ml/d and were not considered significant for modelling. The nodes were connected with links representing rivers and pipes. The model included: pipe capacities, treatment works capacities, reservoir capacities, abstraction and operational licence conditions, operational preferences, control curves, system connectivity, and asset locations where necessary (e.g. for river abstractions or boreholes).

3.3.1 Network topology

For the purposes of this study we needed the network topology information from the water supply network model, with the assigned sources and sinks. Figure 3-5 shows the network topology, with the identified source nodes (inflow points, abstraction, reservoir) and the sink nodes (demand). The inflow points show the locations on the rivers from which surface water is being extracted for water supply. Several of these points were not linked to the network in the original data and model but were created by us. In the end the water supply network topology consisted of 931 nodes and 700 links. The links in the network were all directed links representing the direction of flow to water. For example, we might have water flowing from an inflow point towards the reservoir and then towards a demand node. Hence in the network we had links directed from the inflow point towards the reservoir, and then from the reservoir towards a demand node.

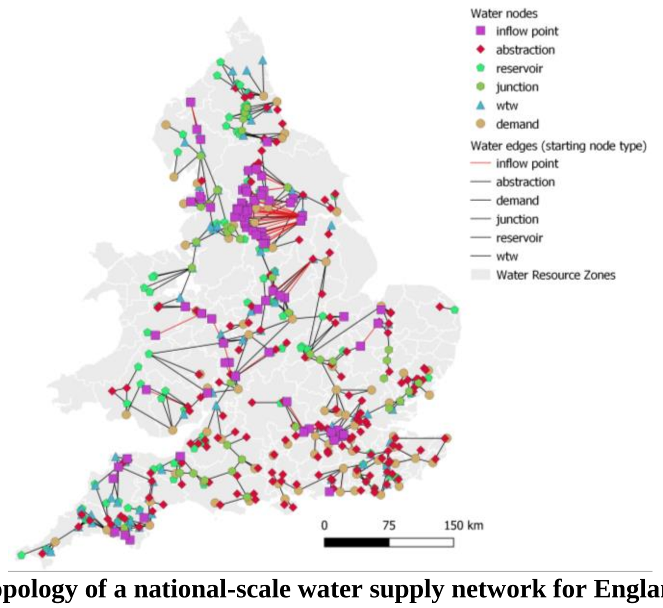


Figure 3-5: Topology of a national-scale water supply network for England and Wales.

3.3.2 Demand allocation

Demands in the water supply network model were allocated at 128 Water Resource Zones (WRZs) levels over England and Wales. All water companies do their planning at the WRZ levels, and estimate demands in terms of total residential populations within WRZs. We did the same by intersecting the LAD level residential census polygon with WRZ polygons and then aggregating the resulting customer demands to nodes within these WRZs. While most WRZ's had only one demand node to which its population was allocated, some demand nodes extracted water from surrounding WRZs. These were identified and the population of their allocated WRZs were also assigned to the nodes. Some big WRZs had more than one demand node and the Water Companies had indicated how the water was proportionally divided to demand nodes within such WRZs. The population within the WRZs were assumed to be divided into similar proportions to the demand nodes. Figure 3-6 shows the result of the customer demand allocation. As can be noted from the figure, the demand nodes are highly aggregated . For example, the whole demand around the London region is represented by one node to which about 8 million customers are assigned.

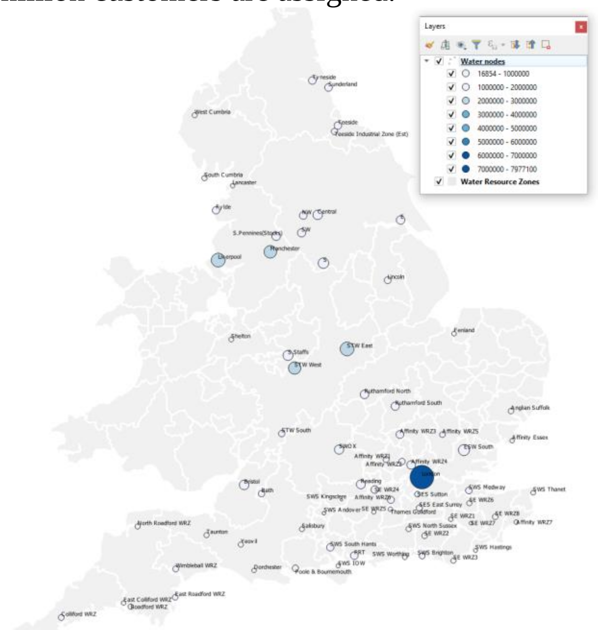


Figure 3-6: Customer demands from WRZ's allocated to demand nodes in the water supply network model.

3.3.3 Failure estimation model

Failures in the water network were estimated similar to the approach followed from the electricity network. These failures were estimated in terms of the numbers of customers at the demand nodes disrupted when some nodes were removed from the network.

We mapped all the possible directed flow paths between all source (40) node and sink (80) node in the network, which in creating 520 unique source-sink paths. Since, the water network was a completely directed network, there were very few feasible source-sink paths. This might also imply low redundancy in the water network, but that is expected for such a high-level sparse network representation of the water system in the country.

When a failure was initiated in the water network all the paths containing the failed nodes were considered disrupted and removed from the set of flow paths. If there were further nodes that lost all their flow connectivity due to the removal of the disrupted flow paths, then these were also considered to have failed due to complete loss of any flows through them. If any of the final set of disrupted nodes were demand nodes, then their allocated demands were summed up to estimate the disrupted customers.

3.4 Railway network

The railways model created for this study relied on a previous vulnerability assessment of Great Britain's railways[38]. This model has been used in several other peer-reviewed studies[39,40] on infrastructure risk analysis. The model shows the railway network for Great Britain owned and operated by Network Rail.

3.4.1 Network topology

Data on the locations of all existing 2,564 railways station was first collected along with the geospatial information on the line geometries of different railway routes in Great Britain. The line geometries showed the single-track routes, which were sufficient for this analysis. The underlying data gave very accurate geospatial information on the node and route locations, as verified by matching with satellite imagery. But this data set has not been updated since 2016, so new railway stations and routes were identified through open data sources, to plug the gaps in the existing data.

The raw data had to be post-processed to be able to join the station nodes onto the line routes and add junctions where two lines intersected, which was done using a novel Python library, for network data cleaning and processing, we have developed and used in several previous projects[41]. The post-processed version resulted in a topologically connected network of 4,024 nodes and 4,524 links.

3.4.2 Demand allocation

The demands on the railway network were estimated in terms of the numbers of passenger journeys over a typical 24-hour period on a weekday, which was similar to an average annual daily count. No freight flow or commercial travel allocation was considered, as there was no data available on such types of travel. While data on station-station journey counts does exist[42], it is a proprietary dataset that was not available to us for this study. Instead we created a trip assignment model using openly available train timetable data and annual station-usage statistics. The train timetable data gave the codes for all station stops made by trains running in the country, which we translated into a spatial routing map based on the location of station and routes in our network. This results in creating 15,038 train flow paths across the whole rail network. From the timetable data we also estimated the numbers of trains on each day of a week over the whole year. The station-usage statistics gave the annual number of entries, exits and interchanges at all station in the country, which we mapped spatially onto our network. The annual station-usage numbers were converted into daily numbers by dividing by 52 weeks and then within the week by the numbers of trains on the day. The daily station entries and interchanges were then proportionally distributed along routes, weighted by the frequency of trains on each route and the numbers of exits and interchanges to all subsequent stops on the routes. For details of the model see Pant et al. 2016[38]. Figure 3-7 shows the result of the railway

[38] Pant, R. Hall, J.W. and Blainey, S.P. (2016). Vulnerability assessment framework for interdependent critical infrastructures: case study for Great Britain's rail network. *EJTIR*, 16(1): 174-194, ISSN 1567-7141.
[39] Lamb, R., Garside, P., Pant, R., & Hall, J. W. (2019). A Probabilistic Model of the Economic Risk to Britain's Railway Network from Bridge Scour During Floods. *Risk Analysis*, *39*(11), 2457-2478.
[40] Oughton, E. J., Ralph, D., Pant, R., Leverett, E., Copic, J., Thacker, S., ... & Hall, J. W. (2019). Stochastic Counterfactual Risk Analysis for the Vulnerability Assessment of Cyber-Physical Attacks on Electricity Distribution Infrastructure Networks. *Risk Analysis*, *39*(9), 2012-2031.
[41] https://github.com/tomalrussell/snkit
[42] https://orr.gov.uk/__data/assets/pdf_file/0014/26600/regional-rail-usage-odm-methodological-report-2017.pdf

flow allocation model, where the highest flows are mostly concentrated around links coming and going out of London.

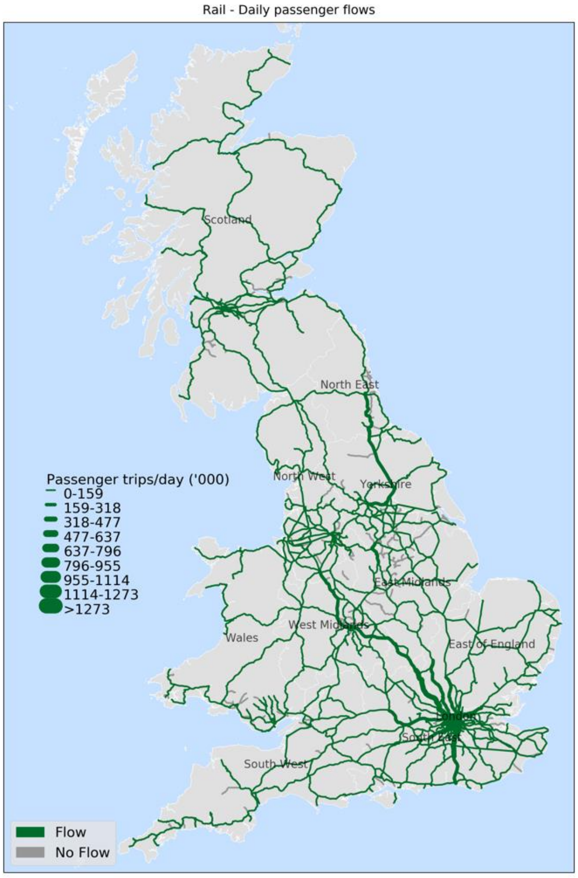


Figure 3-7: Network representation of the Great Britain's railways with the model estimates of numbers of daily passenger flows along nodes and links.

3.4.3 Failure analysis

The flow paths on the railway network were the routes indicated by the timetable data, as that is what the trains would be adhering to. From the trip assignment model, we knew the numbers of passenger journeys on each flow path. We assumed that when a node or link failed it would knock out the whole train journey, thereby disrupting the entire flow paths passengers. This is a worst-case assumption but is not quite unrealistic because in major real big failure events entire train journeys have been cancelled[43,44]. Hence, when one or more nodes or links were removed from the network, we estimated all the disrupted train journey paths and added up the numbers of passengers on these journey paths to get the total disruptions.

[43] https://www.theguardian.com/uk-news/2020/feb/08/uk-rail-firms-reduce-services-as-storm-ciara-approaches
[44] https://www.lner.co.uk/travel-information/travelling-now/travel-alerts/storm-dennis/

3.5 Road network

The road network model for this study was derived from a long-term planning model developed in the ITRC project[45,46]. The network coverage was over Great Britain.

3.5.1 Network topology

The road network topology was derived from road traffic statistics data, using only the geospatial data provided for the major road network for Great Britain. This included all motorways, trunk roads, A roads, and some B roads. The network links do not show the actual geometry of the roads but gives straight line connections between junctions and roundabouts. The original data was post-processed to fill all gaps in connections between road links, and in some instances, this was done by also adding ferry links over waterways. The data also contained traffic statistics of vehicle counts by direction of travel on roads, which was merged with the spatial network topology. Hence, a distinction was made in the network topology as well to represent the direction of travel on roads, which resulted in creating two links between most node pairs. The final network topology consisted of 13,685 nodes representing junctions 36,382 directed links with traffic counts. Another attribute added to the network was the identification of road links which had tunnels in them, because we were interested in mapping the electricity substations supplying power to these tunnels (discussed later). We used other open data sources to identify all major roads with tunnels and matched them to our road network for this study.

3.5.2 Demand allocation

While the traffic counts on roads already gave an indication of their usage, they did not give any information on the where the traffic was coming from and going. For the failure analysis we needed such information to create flow paths. Hence, the demands on the road network were estimated in terms of the numbers of passenger journeys over an average annual daily of traffic patterns in Great Britain. For this we used an Origin-Destination (OD) matrix derived from the National Trip End Model (NTEM) of the Trip End Model Presentation Program (TEMPRO). The NTEM provided an OD matrix of vehicle trips between 7,000 geographical area zones in Great Britain.

The OD matrix was disaggregated to the network level by first finding the network nodes within each OD geographical area. Next the trips created in the origin zone were disaggregated to the road nodes in proportion to the traffic counts on the nodes. Similarly, the destination zones nodes were also given weights in proportion to traffic counts through them. This resulted in dividing each origin zones nodes trip flow to all destination nodes in proportion to their weights, resulting in a final node-node OD matrix. As an example, if an origin zone generating 100 trips, had two origin nodes (O_1, O_2) which attracted 60% and 40% of the traffic respectively, then 60 trips were assigned to one node and 40 to the other. Similarly, if the 100 trips from origin was delivered to a destination zone with two nodes (D_1, D_2) that attracted 70% and 30% traffic counts respectively, then 70 trips were delivered to one node and 30 to another. Overall in this example there are four OD pairs with assigned trips estimated as $\{O_1D_1 = 42\,, O_1D_2 = 18\,, O_2D_1 = 28\,, O_2D_2 = 12\}$.

[45] https://www.itrc.org.uk/highlights/nismod-v2-transport-model/

[46] Lovrić, M., Blainey, S., & Preston, J. (2017). A conceptual design for a national transport model with cross-sectoral interdependencies. *Transportation Research Procedia*, *27*, 720-727.

Following this, the route comprising all links travelled between the O_1D_1 pair was estimated, based on finding the least time route based on speeds on the roads. The whole 42 trips were assigned to the least time route. Similar calculations were done for the O_1D_2 , O_2D_1, O_2D_2 pairs and their trips were assigning to each OD pair's route. For the whole network more than 1 million trip were created, several of which had very small numbers of trips on them. To reduce the set of possible OD trip routes only those routes were chosen that had in excess of 5 trips per day resulting in 182,528 unique trip routes being created, with each having an estimated count of trips. This was converted to passenger numbers by assuming an average occupancy factor of 1.6 across all types of vehicles[47,48]. Figure 3-8 shows the results of the flow allocation on the major road network of Great Britain, where flows are mostly concentrated around big urban conurbations.

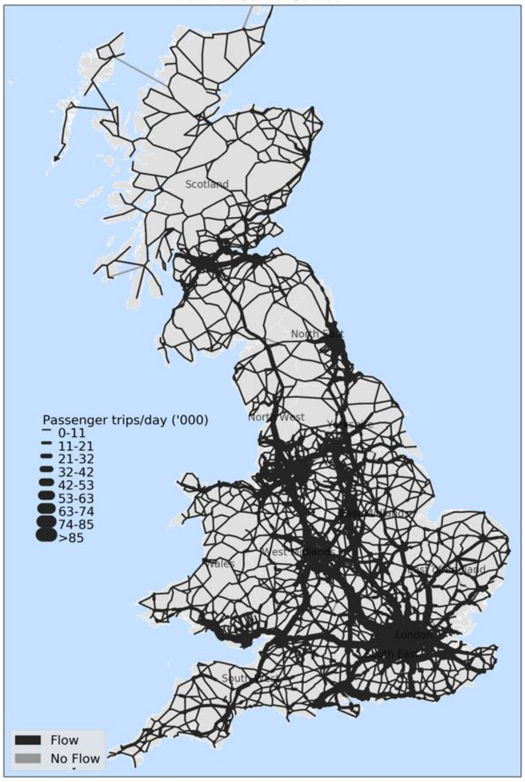


Figure 3-8: Network representation of the Great Britain's major roads.

3.5.3 Failure analysis

The failure analysis algorithm for the roads was similar to the railways, where the flow paths were used to indicate the routes traffic will be adhering to. We assumed that when a node or link failed it would knock out the whole trip, thereby disrupting the entire flow paths passengers. Hence, when one or more nodes or links were removed from the network, we estimated all the disrupted OD trip paths and added up the numbers of passengers on these journey paths to get the total disruptions. We note that again this is a worst-case disruption analysis because the road network has several rerouting options which are used in realistic disruptions. But since the purpose of this analysis was to highlight large scale cascading failures, we chose to not account for trip rerouting in the model, though we could have done it.

[47] https://www.statista.com/statistics/314719/average-car-and-van-occupancy-in-england/
[48] https://www.gov.uk/government/statistical-data-sets/nts09-vehicle-mileage-and-occupancy

3.6 Mapping interdependencies

Data on network interdependencies was extremely difficult to collect, because: (1) there is no existing practice or regulation that makes network operators share their data on links to other networks, in a similar way they would make some of their network data open-access; and (2) most network operators might not have information which goes beyond their own networks. Hence, most of these interdependencies are represented by creating notional links between network assets, as there is very little information on the actual physical connection assets (cables, pipes, etc.) between sectors. These links capture the physical (inter)dependencies between most networks and the cyber dependencies with respect to the digital communication (telecoms) network.

In this study we assumed that electricity and telecoms were interdependent networks, by creating directed links from chosen electricity nodes (substations) towards telecoms nodes (exchanges and macro cells), and other sets to direct links from telecoms nodes to all electricity nodes. Water, rail and roads were considered to be dependent on either electricity or telecoms or both networks. In this study we were most interested in modelling instantaneous failure propagations and failure impacts of the order of a few days, not a few weeks. Hence, electricity and telecoms were considered to be the two sectors whose failures would have such short-term failure propagation effects. It was reasonable to exclude longer term dependencies e.g. the dependency of the electricity sector on water supply (in absence of storage) and transport for fuel. These assumptions were validated with sector experts during Quality Assurance (QA) consultations.

In most cases the dependency links from one asset towards another were created by assuming connections based on proximity, i.e., between the nearest selected nodes of the two types of network. Elaborating on the digital communications sector, for all infrastructure assets which rely on digital communications, we assumed (for lack of better data) that all asset connections were routed into the local internet exchange (BT exchange points). This also included both macro and small cells being connected (‘backhauled’) via either fibre, copper or microwave into the nearest exchange. As well as other infrastructure assets (energy, transport etc.) also being connected into a local exchange, via a fibre, copper or microwave connection. We also assumed that each exchange either had an alternative provider operating within it, or did not, based on the cable availability. Unfortunately, we did not have data to determine whether an asset is directly linked into the Internet using a different route which bypasses the exchange. Hence, the aim of this approach was to capture the majority of instances to provide a generalised national understanding of (inter)dependencies with digital communications infrastructure.

Table 3-1 explains all the dependent links created between assets of different networks. We acknowledge that several types of dependencies between systems were not accounted for with the data used for this study. For example, one of the limitations was the assumptions around the electricity (or water) network’s dependency on the limited numbers of telecoms exchanges and macro cells. We did not account for: (1) SCADA systems that would be used for controlling and monitoring operations and failures in networks, especially electricity; (2) several other private telecoms networks that other networks might be using; and (3) removal of telecoms to some nodes would not cause complete failures but might inhibit some activities.

Table 3-1: Dependency data and model assumptions taken in analysis

Dependency edges	Topological modelling assumptions	Assumptions/limitations	Data Privacy Issues
Electricity-rail dependency	• Data collected on electricity point assets along railways network • Electricity traction substations (nodes) connected to rail nodes with known information on route • Other electricity points connected to rail stations and rail tracks based on nearest proximity • Electricity traction substations connected to the rest of the electricity network as well	• No capacity constraints on electricity supply to the assets • Data only shows limited assets and is not along all routes	None - Because all underlying network data is derived from open-source resources and created by us. The dependency links are all synthetically modelled here.
Electricity-water dependency	• Water assets are assumed dependent on their nearest low voltage substation	• No capacity constraints on electricity supply to the assets	
Electricity-telecoms dependency	• Telecom assets are assumed dependent on their nearest low voltage substation		
Electricity-road dependency	• Road tunnels assumed dependent on their nearest low voltage substation		
Telecoms-rail dependency	• Data on telecom masts along existing rail network • Telecoms masts (nodes) connected to nearest rail nodes based on proximity	• Not linked to the rest of the fixed telecom exchanges as they are independently owned and operated by network rail • Data only shows limited assets and is not along all routes	
Telecoms-electricity dependency	• Electricity nodes assumed dependent upon their nearest macro cells and exchanges	• No actual data to inform this dependency	
Telecoms-water dependency	• Water assets connected to their nearest exchanges and macro cells	• No actual data to inform this dependency	

We had some detailed information on the locations and types of rail assets that use other utilities, especially electricity. This was an older dataset, that we had created for a previous study[12], which gave the locations of roughly 9,100 of the following assets: (1) Electrification Switching – Operational; (2) Electrification Substation – Domestic; (3) Electrical Control Room; (4) Remote Monitoring – Critical; (5) NDS National Delivery Service; (6) Telecoms – Domestics; (7) Lighting - Bridge/Navigation; (8) Telecoms – Operational; (9) Signalling/Relay Room – Domestics; (10) Pumps; (11) Lighting - Tunnels/Junction; (12) Lighting – Walkway; (13) Electrification Substation – Operational; (14) Lighting - Yards/Miscellaneous; (15) Signalling/Relay Room – Operational; (16) NR Office Accommodation; (17) Signalling Supply Point – Domestic; (18) Rail traction GSP; (19) Signalling Supply Point – Operational; (20) Signal Box – Power; (21) Signalling Centre/IECC/Route Control; (22) GSM-R/FTN/RETB/CSR/NRN; (23) Level Crossing – CCTV; (24) Electrification Switching – Domestic; (25) Points Heating; (26) Level Crossing – Other; (27) MDU - Accommodation/Storage. While most of these point assets would be providing service to the nearest railway link to them, some of these are used for entire routes sections covering several links. For example, each traction substation would supply electricity to a whole route spanning several stations, and similarly Signalling Centre/IECC/Route Control would be controlling several kilometres of train routes. Such considerations were made in collection information and mapping the influence of each of these point assets on the railway network with dependency edges created from them to several nodes/links where appropriate. All these point assets (nodes) were then assumed dependent upon the electricity network and were connected to their nearest substation of the right voltage.

From the same data source, the information on rail nodes dependent upon telecoms towers along the rail routes was used, but these telecoms towers were not linked to the rest of the telecoms network.

Since the dependencies by proximity mapping were being inferred, they were approximate and it has to be recognised that if the supply point was not correctly identified then the dependent assets might be connected to a network at the incorrect locations, which is very likely to happen.

Translating the above described assumptions into results the numbers of dependent links between network created in the three versions of connectivity mapping we have assumed for this study, are shown in Table 3-2. These are the numbers of links created after removing unwanted connections between nodes that were greater than 10km. The analysis showed that that in 97.5% of the instances the degrees of connections of the nodes increased from 1 to 2, and in further 94.4% of the instances the degrees of connections of the nodes increased from 2 to 3. Which means that in very few instances the distance truncation criteria prevented from adding unwanted redundancies to the network.

Table 3-2: Versions of degrees of connections and the numbers of network links created in the data.

Connections mapping type	Number of links
Single connections	103,624
Two connections	187,457
Three connections	268,766

3.7 Accounting for backup supply

Backup supply signifies the time for which the asset was not disrupted because it has alternative supply for the similar service. Based on discussions with the NIC and sector experts only a small set of assets were assumed to have backup supply for only electricity. No telecoms backup was considered in this study. Table 3-3 shows that the efficacy of the electricity backup supply considered for each sector was interpreted in terms of the duration in hours over which the backup would be able to completely substitute for lost electricity supply. These values were tested with sector experts while doing the QA consultation of the underlying data and assumptions.

Table 3-3: Assumed electricity backup supply duration put in place for different types of dependent assets.

Sector	Node type	Assumed Backup supply of Electricity (hours)
Electricity	All	0
Telecoms	Exchange	24
Telecoms	Macro cell	2
Water	All	72
Road	Road tunnel	24
Rail	All	0

Given that the actual duration of the backup supply at each asset was uncertain and not all assets of the same type would have the same duration of backup supply, we considered a probabilistic distribution for the duration of backup supply. This was based on the following assumptions: (1) Backups survived as per a gamma distribution, which is a very well-known distribution used to model infrastructure reliability for maintenance[49]; and (2) The backup duration for an asset was estimated as the product of this assumed duration (from Table 3-3) and a gamma survival function with value between 0 and 1). This meant that the backup would last anywhere between 0 hours and the assumed duration it was assigned.

3.8 Future network changes

The following drivers of the future scenarios were considered in the study:

- Specific NIA scenarios – which would affect particular sector networks and their interdependencies.
- Spatially disaggregated population changes – which would affect the customer demands for all sectors.
- Spatially disaggregated Gross Value Added (GVA) changes – which would affect the service demands in some sectors.
- Macro-level GDP growth forecast – which would affect the economic losses due to disruptions.

3.8.1 NIA future energy scenarios and changes to the electricity network

The two main NIA future scenarios implemented in this study, explained in Table 2-1, resulted in changes to the supply and demand on the electricity networks for the future. The supply side changes meant adding and removing source nodes in the network, while the demand side changes meant adding more MW loads to the demand nodes.

The energy scenario data from this study was created by Aurora Energy Research[50] previously for the NIC, giving aggregated national-scale supply and demand estimates under the future 70% renewable generation scenarios with high (Elec70) and low (Hydro70) electricity heating demands. We had to disaggregate the values to the network level.

3.8.1.1 Supply side changes to the electricity network

First, we looked that the different energy generation technologies in our data and mapped them to the Aurora data of generation mixes, because there was a difference in the names and types of the technologies in the two datasets. Next, we compared the numbers of the current generation capacities in GW in our network, with the future projected numbers in the Aurora scenarios. Based on the numbers we decided whether to scale up or scale down the current capacities of all nodes of a particular technology in proportion to their current weighted capacities across the whole network. In some instances, we had specific data at the node levels indicated whether we needed to add or remove nodes. Table 3-4 shows the details of the different generation technologies in our network and the comparison with the future Aurora technological changes in the energy generation mix for the electricity network. The implications for the network due to the changing energy mix as also shown in the table.

[49] van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, *94*(1), 2-21.
[50] https://www.auroraer.com

Table 3-4: Comparison of the generation capacities in the current electricity network with the future modelled generation mix and capacities as modelled by the Aurora Energy model. Also shown are the network changes that will result due to future generation mix changes.

Generation capacity (GW) in existing network–2017	Energy types in our data	Aurora energy types	Aurora generation mix (GW)		Scale up/down capacity of current nodes
			Hydro 70	**Elec 70**	
4.08	Bio related	Biomass	6	6	Scale up
		Biomass with CCS	0	0	-
36.19	Gas	CCGT	5.14	24.57	Scale up/down
		OCGT	0.97	0.97	Scale up/down
		Gas recips	17.56	28.36	Scale up/down
16.69	Oil diesel coal	Coal	0.01	0.01	Scale down
		Diesel recips	1.01	1.01	Scale down
		Internal combustion engines	0.13	0.13	Scale down
9.29	Nuclear	Nuclear	11.8	8.22	Remove + Add new
4.52	Hydro related	Pumped Storage	2.81	2.81	scale up
		Hydro	1.87	1.87	scale up
8.10	Solar	Solar	75.5	71.37	Add new and scale up
6.32	Wind offshore	Offshore wind	29.3	49.29	Add new and scale up
13.51	Wind onshore	Onshore wind	24.27	25.43	Add new and scale up
5.00	Interconnectors	Interconnectors	17.9	17.9	Add new
0.41	Waste				
0.05	Ocean related				
		Batteries	12.61	18.05	Add new
		CCS	4.53		Ignore
		DSR	7.43		Ignore
104.16	**Total**	**Total**	**218.84**	**255.99**	

Adding interconnectors

Specific information on locations of future interconnectors was available for the Aurora data and other sources[51] was collated and translated into adding nodes and connecting them at specific locations of the existing network. It was assumed that all interconnectors were to connect to existing substation in the National Grid (NGET) transmission networks. While in reality there might be new substations being built for new interconnectors, we did not have detailed data of planned substation and new connections. The aim here was to approximate to the nearest location where the interconnectors would connect to the existing grid. These changes applied to both future energy mixes.

[51] https://www.4coffshore.com/transmission/interconnectors.aspx

Table 3-5: Description of planned interconnectors and their connections into the Great Britain's electricity network of 2050.

Project	Capacity (GW)	Country link	NGET link
ElecLink	1.0	France – Peuplingues	HVDC converter station at Folkestone + Sellindge 400kV substation
NEMO	1.0	Belgium – Herdersbrug/Gezelle	HVDC Richborough converter station + Richborough 400kV substation
Viking Link	1.4	Denmark – Revsing	HVDC North Ing Drove, Bicker Fen converter stations + Bicker Fen 400 kV substation
IFA 2	1.0	France – Tourbe	HVDC Daedalus converter station +Chilling 400kV substation
FAB	1.4	France – Manuel	HVDC Long Lane converter station + Broadclyst 400kV substation
Gridlink	1.4	France – Dunkerque/Bourbourg	Kingsnorth 400kV substation
Aquind	2.0	France – Barnabos	Lovedean 400kV substation
Neuconnect	1.4	Germany – Conneforde	Greystones 400kV substation
Greenlink	0.5	Ireland – Great Island	HVDC Pembroke converter station + Pembroke 400kV substation
NSL	1.4	Norway – Kvilldal	HVDC East Sleekburn (Blyth) converter + Blyth 400kV substation
NorthConnect	1.4	Norway – Simadalen/Sima	HVDC Fourfields, Boddam, Peterhead converter stations + Peterhead 400kV substation

Adding and removing nuclear sites

The Aurora scenarios gave specific information on decommissioning some nuclear power plants, based on the future energy planning by the UK government. Also, there are plans to build a new Hinkley Point C power plant with 3.34 GW capacity in the future[52,53]. After consultation with the NIC we decided to remove some of the existing nuclear power plants from the future networks, and replace Hinkley Point B with Hinkley Point C while retaining Sizewell B and at least some plants of same capacity as the new Hinkley Point C. Unfortunately adding Hinkley Point C as a new node was not possible because we did not have specific geospatial information about its location and connections to the existing electricity network. The best assumption to make was that Hinkley Point C would be made close to the existing Hinkley Point B.

Table 3-6 shows all the changes made to the future energy mix by removing nodes and upgrading the capacities of existing nodes to match the forecasted capacities of the Aurora/NIC scenarios.

Table 3-6: Network changes made to the nuclear power mix in the future network scenarios.

Plant	Changes made	Capacity (GW)	
		Hydro70	Elec70
Dungeness B	Remove	-	-
Hartlepool	Remove	-	-
Heysham I	Upgrade capacity	2.7	1.55
Heysham II	Upgrade capacity	2.94	1.68
Hinkley Point B	Replace as Hinkley Point C	3.34	3.34
Torness	Remove	-	-
Hunterston B	Remove	-	-
Sizewell B	Upgrade capacity	2.82	1.62

52 https://www.edfenergy.com/energy/nuclear-new-build-projects/hinkley-point-c
53 https://www.gov.uk/government/collections/hinkley-point-c

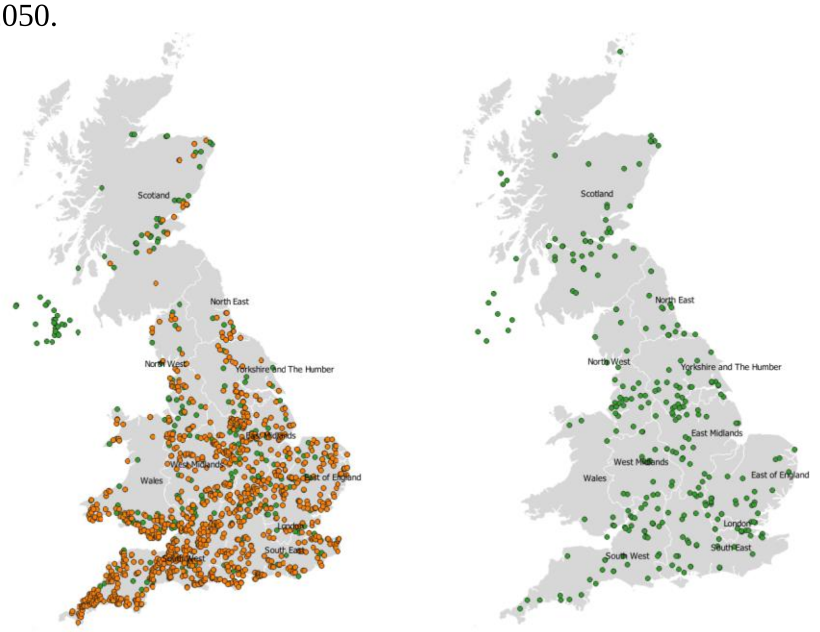
Adding renewables

The more significant change in the future energy mix was in terms of adding more renewables and embedded generation nodes to the existing network. This included adding new nodes of batteries, onshore and offshore wind, and solar sites.

Data from the Renewable Energy Planning Database (REPD)[54] quarterly extract, updated till September 2019, gave the locations, and capacities of planned renewable technologies that were under different stages of development including currently operational, under construction, awaiting construction and application approved, application submitted. We extracted this dataset and mapped out all the new nodes with their capacities to add to the existing network.

Figure 3-9 shows the locations of the existing sites (in orange) and new additional sites in the future (in green) selected from the REPD database, for inclusion to the electricity networks in the future scenarios. Once these sites were selected they were connected to their nearest HV distribution substation/transformer node (33kV to 132kV) with a step-down transformer (< 132kV) leading to the LV networks, as generally that is the level at which embedded generation technologies would mostly connect to the electricity networks[55].

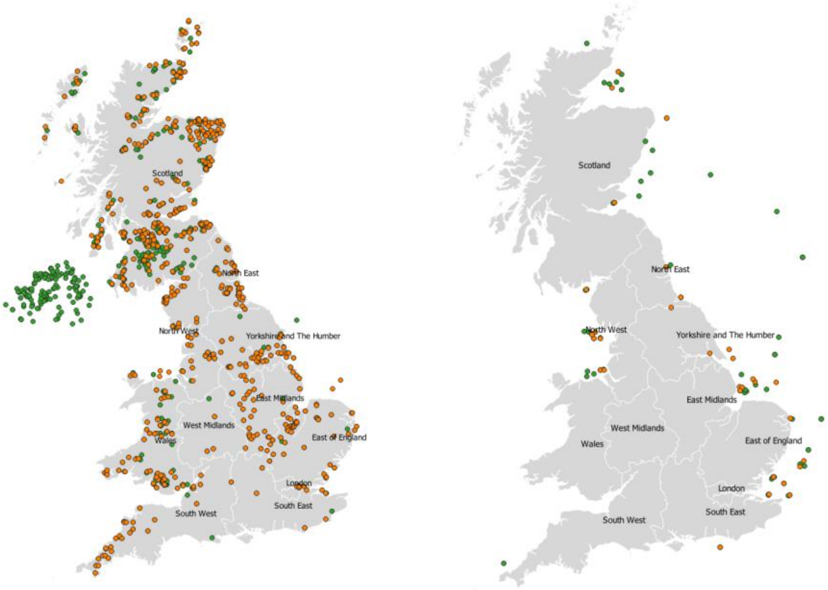
We note that the assembled data did not show as high total cumulative capacities as forecasted in the future Aurora/NIC model energy mix scenarios. Hence, we scaled up all the capacities of the nodes of each technology to match the cumulative Aurora/NIC estimates for that technology in 2050.



(a) Solar (b) Battery

[54] https://www.gov.uk/government/publications/renewable-energy-planning-database-monthly-extract

[55] https://www.energynetworks.org/assets/files/electricity/engineering/distributed%20generation/DG%20Connection%20Guides/July%202014/G59%20Full%20June%202014%20v3_Updated.pdf



(c) Wind Onshore (d) Wind Offshore

Figure 3-9: Map representations showing the locations of existing (in orange) and future (in green) generation sites for (a) Solar; (b) battery; (c) wind onshore; and (d) wind offshore. Note that the sites in Northern Ireland are ignored.

Scaling up and down other technologies

Significant changes will be made in the electricity networks of the future in terms of decommissioning coal, gas and diesel oil technologies. The big difference between the two Aurora/NIC scenarios is the reduction of gas in one (Hydro 70) compared to the increase in gas in the other (Elec 70). Under both scenarios, usage of diesel oil and coal significantly reduces in the future. Hence the approach should have been to remove most of the nodes these technologies in both scenarios and add some more gas nodes in the Elec 70 scenario. Unfortunately, we did not have any data or expert feedback on how to do this. The next best option was to scale up and down existing nodes in the network to match future cumulative capacity projections for these technologies. A similar approach was followed for the nodes using biomass.

3.8.1.2 Demand side changes to the electricity network

The two future Aurora/NIC scenarios differed significantly in terms of the energy demands in TWh being placed on the electricity networks, with the main difference being the demands due to heating. Table 3-7 shows the estimated cumulative energy demands on the electricity network for the two scenarios, with the third column highlighting the difference is mainly due to use of electricity heating. We also note that there are significant demands in the future due to electric vehicles (EVs).

Table 3-7: Electricity energy demands estimated in the future Aurora/NIC scenarios.

Scenario	Annual base electricity demand net of heat and transport (TWh)	Annual electricity demand from heat (TWh)	Annual electricity demand from EVs (TWh)	Total annual electricity demand (TWh)
Hydro 70	357	17.7	91.07	465.4
Elec 70	357	148.4	91.07	596.4

Translating these demands onto the networks required several assumptions as the Aurora/NIC estimates only gave single cumulative estimates at the national scale. The following assumptions were made:

Estimating electricity and heat loads

1. All demand nodes in the future networks were the same as the current network sink (demand) nodes. Due to lack of any data on where new substations would be built, we were not able to add further demand nodes, but rather scale up or down future demands at existing nodes.
2. From the ITRC long-term energy model[26] we obtained hourly energy demand estimates as MW loads at the LAD level over the whole year. We added these up to get the annual energy usage in TWh and accordingly scaled up or down the numbers to match the Aurora/NIC scenario estimates. Subsequently, the hourly MW loads changed by the same scaling factor.

Estimating EV loads

3. Since EV loads on the electricity network originated from transport, we used the spatial transport OD matrix in the future for estimating EV demands.
4. The ITRC long-term transport model[46] gave a future EV demand that translated trips generated into EV demands, which were aggregated from the NTEM zones to the LAD administrative levels. These demands gave the daily energy usage in TWh from EV.
5. Assuming that the EV usage was uniform for the whole year the daily usage was multiplied by 365 to convert to annual usage, as scaled up or down to match the Aurora/NIC scenario estimates.
6. The Aurora/NIC scenarios for EV demands also gave an half-hourly electricity charging profile, which was converted to an hourly profile. The daily/annual EV energy usage was converted to hourly MW load based on this charging profile.

Estimating total network loads

7. From Steps 2-6 we were able to get LAD level hourly load profiles of total electricity demands from electricity plus heating plus EV usage. We were interested in the peak load on the whole electricity network, which we extracted as the maximum hourly load over the whole year. This gave us the LAD level electricity demands to be assigned to the network nodes.
8. From the LAD levels the energy loads were downscaled to the LSOA and then aggregated at the demand nodes, as described in Section 3.1.2.

3.8.2 Changes to network topologies

From the previous section it is clear that the electricity network topology would change in the future, mainly by adding more energy source nodes. In reality all other networks would also witness similar changes. But unfortunately, we did not have any data on other networks so we assumed that there would be no change in their topologies.

Figure 3-10 shows the changes made to the electricity network topology in the future (in red), by adding more links to the current network (in green). Table 3-8 shows the details of the estimated network topologies, demand loads and supply capacities in the current and future scenarios.

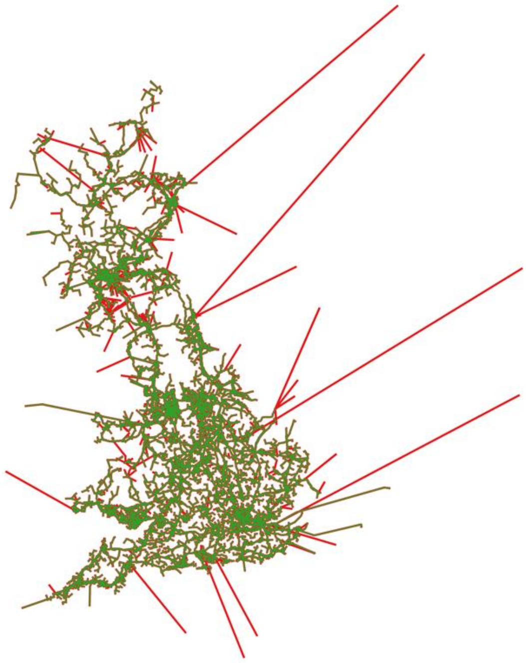


Figure 3-10: Map representation showing the current network topology (in green) and future added links (in red) in the future scenarios for the electricity network of Great Britain.

Table 3-8: Comparison between the current and future electricity networks properties modelled in this study.

Scenario	Topology	Demand				Generation capacity
		Electricity	Heat	EV	Total	
Current	Nodes – 18,062 Links – 13,254	55 GW	0	0	55 GW	104 GW
Hydro70	Nodes – 18,801	56GW	9.1GW	4.9GW	70GW	207 GW
Elec70	Links – 13,993	56GW	75GW	4.9GW	136 GW	260 GW

The overall changes in the electricity network topology result in also creating addition dependency links with the telecoms network assets. Hence, additional links from the telecoms exchanges and macro cells towards the new electricity sources are added to the interdependent network topology.

Table 3-9: Versions of degrees of interdependencies and the numbers of network links created in the data in the current and future network configurations.

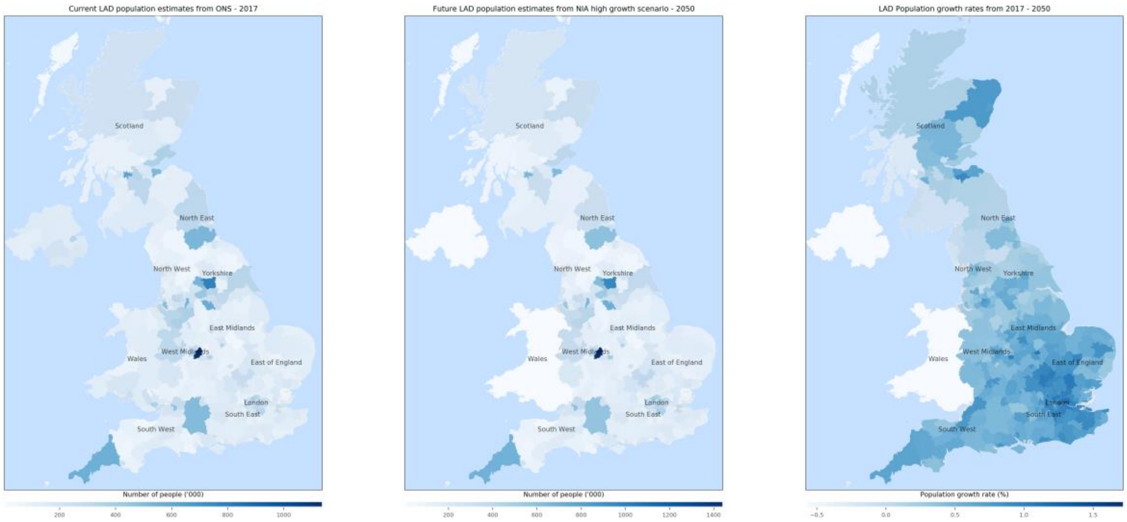
Interdependency mapping type	Current networks - Number of links	Future networks - Number of links
One interdependency	103,624	104,998
Two interdependencies	187,457	109,161
Three interdependencies	268,766	272,731

3.8.3 Changes in customer demands across all networks

All sectors were allocated new demands in 2050 based on population projections at the Local Authority District (LAD) level (380 areas), which were downscaled to thee sector specific admin levels and the service output areas. The future population projections were based on the NIA scenario of *high fertility (or high growth)* which included the following assumptions:

- England - ONS 2014-based high fertility subnational experimental projection.
- Scotland - Scotland Stats 2014-based high fertility subnational projection.
- Wales - Calculated based on ONS 2014-based high fertility national projection.

Figure 3-11 shows map visualisations of the LAD level population estimates from the current 2017 levels and the NIA high growth scenario for 2050 selected for this the study. The figure also shows the annual population growth rate in percentages for each LAD, which show -0.5% to +1.6% growth changes across LADs with some of the highest positive growth rates concentrated around London and the South East.



(a) (b) (c)

Figure 3-11: NIA high growth scenario-based LAD level population estimates of Great Britain showing (a) Current 2017; (b) Future 2050; and (c) Annual growth rate (%) between future and current populations.

GVA data was also considered at the Local Authority District (LAD) level for estimating the future demands, especially for reworking the electricity demand profiles and rail and road OD matrices in the future. GVA data taken from the Office of National Statistics (ONS), included.

- Current ONS estimates of GVA in 2017[56].
- Future GVA growth scenario projections for 2050 derived by Cambridge Econometrics[57] and used for a previous study for the NIC[58].

We note that the GVA values might take some infrastructure failures into account, as they estimate the total output of goods and services less the value of goods and services used in the production process[56]. This means that if the production would have gone down in 2017 due to economic failures then the ONS GVA estimates would reflect that. But as far as we are aware, there are no significant observed or projected infrastructure failures in the GVA estimates. These estimates are just being used to project future transport demand based on a simple GVA elasticity. Though there are many assumptions in the economic estimates and transport projections, we do not believe the possible misrepresentation of infrastructure failure in the GVA data is a significant concern. Here we are only using the proportional change in GVA for understanding how the service demands of some sectors might change in the future (see Section 3.8.3.4 and 3.8.3.5), and not in the failure calculations.

Figure 3-12 shows map visualisations of the LAD level GVA estimates from the current 2017 levels and the ONS projections for 2050 selected for this the study. The figure also shows the

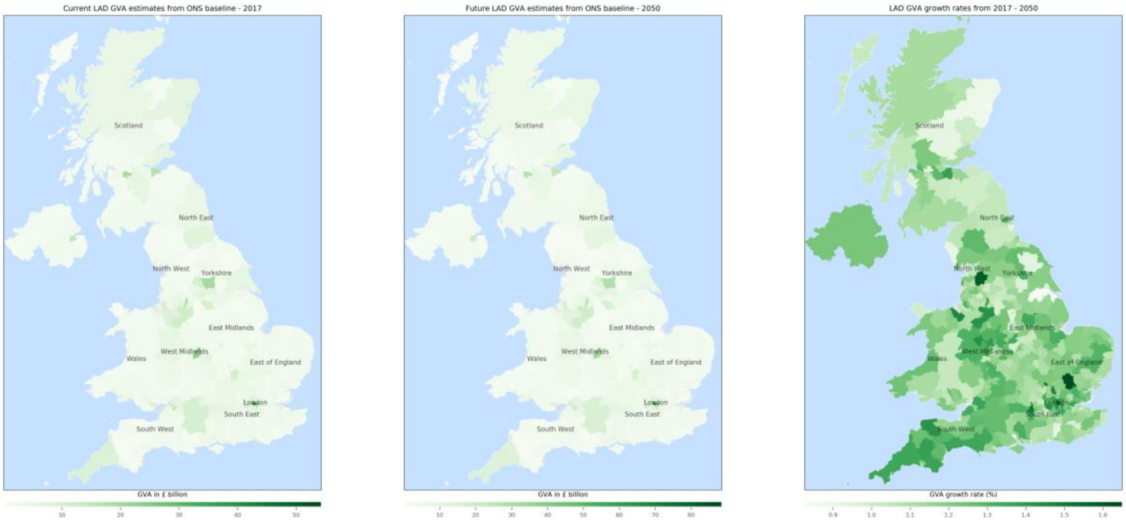
---

56 https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/regionalandsubregionalproductivityintheuk/february2019

57 https://www.camecon.com/how/lefm-model/

58 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/601163/Economic-analysis-Cambridge-Econometrics-SQW-report-for-NIC.PDF

annual GVA growth rate in percentages for each LAD, which show +0.8% to +1.7% growth changes across LADs with some of the highest positive growth rates concentrated around London, East of England and the North West of England.



(a) (b) (c)

Figure 3-12: ONS scenario-based LAD level GVA estimates of Great Britain showing: (a) Current 2017; (b) Future 2050; and (c) Annual growth rate (%) between future and current GVA levels.

Below we discuss the assumptions being made in estimating the customer demands for each sector in 2050.

3.8.3.1 Electricity demand changes

We assumed that the assigned populations to the electricity sink nodes were expected to roughly change in the same proportions as the changes in LAD populations. Hence the 2050 projected population estimates were taken and disaggregated to the LSOA level before aggregating LSOA populations to the sink nodes. We did not have any LSOA future projections of building footprints, so we assumed that the current building footprint areas (weights) would be the same in the future. Hence, the future LAD population estimates were disaggregated the LSOA levels in the same proportion as the present. Changes in GVA were assumed to have no effect on the changing customer demands for the electricity assets.

3.8.3.2 Telecoms demand changes

Similar in concept to the electricity, we assumed that the assigned populations to the exchanges and macro cells in 2050 were disaggregated from the 2050 LAD population estimates. Here we took the LAD estimates and disaggregated them to the postcode levels and then intersected the post code population densities with the services areas of the exchanges and the macro cells. Again, we did not have any post code level address point number for the future, we assumed that the current post code address point numbers (weights) would be the same in the future. So, the future LAD numbers were distributed to the post code level in the same proportion as the present. Changes in GVA were assumed to have no effect on the changing customer demand for the telecoms assets.

3.8.3.3 Water demand changes

Estimating future customer demands was more straightforward. The future 2050 LAD level populations numbers were distributed to the 128 WRZ levels by spatially intersecting the two areas and summing up over the product of the population density and common areas of intersection.

3.8.3.4 Roads demand changes

Future road network flows we estimated by changing the OD matrix estimates. For each NTEM OD zone, future OD flows were derived based on the equation 1 below, derived from a long-term transport planning model study[46]:

$$OD_{future} = OD_{current}\left(\frac{Pop_{future}}{Pop_{current}}\right)\left(\frac{GVA_{future}}{GVA_{current}}\right)^{0.63} \qquad (1)$$

Here the populations and the GVA estimates for each NTEM zones were across all the LAD's polygons intersecting that zone. Following the estimation of a future OD matrix the trip allocation was based on assuming the same traffic volume weights at the road nodes as the current levels, since there was no data on future traffic statistics. The future road speeds were assumed to be the same as the current and the allocation was again based on the least cost (time) path choice.

3.8.3.5 Railways demand changes

Railways OD flows were derived from the station usage statistics and the train timetables. Unfortunately, there were no data sources to incorporate timetable changes, hence they were assumed unchanged. The station usage was assumed to change in a similar manner as the road OD matrix as shown in equation 2[46].

$$Station\ Usage_{future} = Station\ Usage_{current}\left(\frac{Pop_{future}}{Pop_{current}}\right)\left(\frac{GVA_{future}}{GVA_{current}}\right)^{0.63} \qquad (2)$$

Here the population and GVA estimates of the LAD area that contained the station were used. The allocation of passenger flows on the network were done with the existing timetable patterns of travel.

3.9 Implications of future change on failure analysis

Due to the changes in network topology and increased demands in the future there would be some expected changes in the failure outcomes of the networks. This difference would be driven by the changes in mapped source-sink flow paths in the future. For example, we would expect that adding more sources in the electricity network would create several more source-sink paths adding more redundancies in several cases. Table 3-10 summarises the differences in flow paths between current and future networks and their implications on the failure analysis results.

Table 3-10: Flow paths for each network in the current and future scenarios and their implication on the failure outcomes.

Sector	Current	Future	Expected Failure implications
Electricity	1,002,837	1,319,935	Increased source-sink paths would add redundancies and reduce some failure outcomes. Mostly disruptions could increase due to increased population and hence demands in the future
Telecoms	97,992	97,992	No change in failure propagation. Disruptions could increase due to increased population and hence demands in the future
Water	520	520	
Railways	15,038	15,038	
Road	182,528	207,793	Most nodes could have more future flow paths and flows through them increasing their failure impacts

3.10 Economic loss estimations

3.10.1 Input-Output model and data

For this study, a Leontief Input-Output (IO)[59,60] macroeconomic model based on empirical data is used to represent economic losses at the UK-scale (which includes Northern Ireland). Leontief IO model is a very well recognised model in macroeconomics literature[61], with Wassily Leontief being awarded the Nobel Prize in 1973 for IO modelling. The Leontief IO model captures macroeconomic interdependencies across industry sectors at an aggregated region-scale (provincial, national, international), and the most important insight the model provides is to show how individual or groups of sectors influence the rest of the economy[60,61]. The model is very popular because it is supported by empirical data globally, with several countries maintaining and releasing IO accounts[62,63], making the model useful in practice globally[64,65]. In the UK annual Input-Output tables are generated by the Office of National Statistics[66,67]. While the Leontief IO data and model was originally meant for studying macroeconomic growth modelling and structural planning, it has now been extensively used in disaster impact assessment with different extensions and variations to the original model[68,69].

The classical Leontief IO model, which we have used for this study, is based on following guiding principles[70,71]: (1) The macroeconomic system is in equilibrium where each industry sector produces a single homogenous output that is either absorbed by itself and other industries in the economy in further production of their outputs or used for final consumption; (2) The output produced by a sector is used in a fixed proportion by another sector in producing its

---

59 Leontief, W. (Ed.). (1986). *Input-output economics*. Oxford University Press.
60 Leontief, W. (1987). Input-output analysis. *The new palgrave. A dictionary of economics*, *2*(1), 860-64.
61 Miller, R. E., & Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge university press.
62 https://www.bea.gov/industry/input-output-accounts-data
63 http://www.oecd.org/sti/ind/input-outputtables.htm
64 Yamano, N. (2016). OECD Inter-Country Input–Output Model and Policy Implications. In *Uncovering value added in trade: New approaches to analyzing global value chains* (pp. 47-59).
65 Ghosh, P. P., Ghose, A., & Chakraborty, D. (2011). A critical review of the literature on integrated macroeconometric & input-output models. In *The 19th International Input-Output Conference. Alexandria VA, USA.*
66 https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/inputoutputanalyticaltables/methodsandapplicationtouknationalaccounts
67 https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/commentaryonsupplyandusebalancedestimatesofannualgdp/1997to2014
68 Koks, E., Pant, R., Husby, T., Többen, J., & Oosterhaven, J. (2019). Multiregional disaster impact models: Recent advances and comparison of outcomes. In *Advances in Spatial and Economic Modeling of Disaster Impacts* (pp. 191-218). Springer, Cham.
69 Koks, E., Pant, R., Thacker, S., & Hall, J. W. (2019). Understanding Business Disruption and Economic Losses Due to Electricity Failures and Flooding. *International Journal of Disaster Risk Science*, 1-18.
70 West, G. R. (1995). Comparison of input–output, input–output+ econometric and computable general equilibrium impact models at the regional level. *Economic Systems Research*, *7*(2), 209-227.
71 Christ, C. F. (1955). A review of input-output analysis. In *Input-output analysis: An appraisal* (pp. 137-182). Princeton University Press.

outputs. This means that the production technologies are fixed and there is no substitution in the economy; (3) The changes in the economy are driven by changes in the final consumptions (exogenous demands) to which the supply side responds by changing it production to create a new equilibrium in the economic system. This means that there are no supply side constraints in the model; (4) There are no price effects when the economic equilibrium shift and employment is maintained with infinite elasticity in labour supply.

It goes without saying that the classical demand-side Leontief IO model has been critiqued in literature for its overtly simplified representation of a linear non-substitutable economic system with no price and labour effects[70,71]. Over the years several advances have been made to overcome limitations of the IO data, with the main approach now being to create Social Accounting Matrices (SAMs) that provide supply and use tables linking multiple industries to multiple commodities from with the IO accounts are created[72]. Specifically, for disaster impact modelling, several hybrid approaches that build from the Leontief IO model have been proposed to account for supply side disruptions[73], substitution effects across industries and regions[74], and changing production functions with inventory management during disasters[75]. Other approaches of computational general equilibrium (CGE) modelling that also use SAMs have been extensively used for disaster impact modelling, with such models using non-linear product functions with price effects and labour elasticity[70]. While there have been extensive comparisons and critiques of IO and CGE models in literature, it should be noted that all of them only model one out of several possible outcomes of economic disruptions and each model outcome has its limitations[76].

The attraction of using the simplified IO model for study is simply based on the ease of data availability, whereas other hybrid IO and CGE models would require data that was beyond our scope. We look at these disruptive effects in the very short-term (over a day), where we can relax assumptions of changing prices and have a fixed technology for sectors. But on the other hand, over such short timelines of disruptions sectors would be able to substitute for lost production and the economy would most probably not adjust to a new equilibrium, which would be more realistic of the durations of disruption lasted several weeks or months.

The main insight from the IO model we want to get here is to understand the amplification of interdependent (indirect) losses on the rest of the economy produced by infrastructure sector customer disruptions (direct losses). The ability of IO models to quantify the direct and indirect economic losses, has been one of the main reasons why they are extensively used in economic impact assessments[77]. The magnitudes of economic losses here would represent close to worst-case impacts under the assumption of losing a day's worth of economic demand, as the IO model used here is known of give an overestimation of impacts[78].

We now explain the formulation of the IO model. As per the Leontief IO model, in a macroeconomic system comprised of *n* industry sectors the output produced by sector i, x_i, is

[72] Stahmer, C. (2004). *Social accounting matrices and extended input-output tables* (pp. 313-344). Measuring sustainable development: Integrated economic, environmental and social frameworks, Paris.
[73] Steenge, A. E., & Bočkarjova, M. (2007). Thinking about imbalances in post-catastrophe economies: an input–output based proposition. *Economic Systems Research*, *19*(2), 205-223.
[74] Koks, E. E., & Thissen, M. (2016). A multiregional impact assessment model for disaster analysis. *Economic Systems Research*, *28*(4), 429-449.
[75] Hallegatte, S. (2014). Modeling the role of inventories and heterogeneity in the assessment of the economic costs of natural disasters. *Risk analysis*, *34*(1), 152-167.
[76] Okuyama, Y., & Santos, J. R. (2014). Disaster impact and input–output analysis. *Economic Systems Research*, *26*(1), 1-12.
[77] Kelly, S. (2015). Estimating economic loss from cascading infrastructure failure: a perspective on modelling interdependency. *Infrastructure Complexity*, *2*(1), 7.
[78] Okuyama, Y. (2008). Critical review of methodologies on disaster impact estimation. *Background paper for EDRR report*.

used to satisfy the intermediary demands from the rest of the economic sectors $\sum_{j=1}^{n} a_{ij}\, x_j$ and exogenous demands f_i. The Leontief coefficient $a_{ij} < 1$ is based on the assumption of a linear production function where every 1 unit of output from sector j, x_j, requires a_{ij} units of input from sector i. The Leontief IO model of the whole balanced economy is represented as:

Output ($\mathbf{x}$) = Intermediate industry demand ($\mathbf{Ax}$) + Final exogenous demand($\mathbf{f}$) (3)

Where $\mathbf{x}$ is a vector of n sector outputs, $\mathbf{A}$ = the $n \times n$ Leontief coefficient matrix, which captures inter-industry sector linkages, and $\mathbf{f}$ is a vector of n sector exogenous demands. A Leontief IO model represents an economy in equilibrium, which means that there is a unique solution to Equation (3) obtained as following:

$$\mathbf{x = Ax + f} \Rightarrow \mathbf{[I\text{-}A]x=f} \Rightarrow \mathbf{x = [I\text{-}A]^{-1}f} \qquad (4)$$

Furthermore, the exogenous demands can be further split as following:

$\mathbf{f}$ = Household demand ($\mathbf{h}$) + Government demand($\mathbf{g}$) + Exports ($\mathbf{e}$) (5)

Rewriting $\mathbf{f}$ in terms of its components gives

$$\mathbf{x = [I\text{-}A]^{-1}(h + g + e)} \qquad (6)$$

Equation (6) shows that output ($\mathbf{x}$) is driven by demands, and the Leontief Inverse Matrix ($\mathbf{L = [I\text{-}A]^{-1}}$) shows the economic multipliers will magnify the effects of demand driven perturbations. We use this simplified demand-driven model and concept to estimate economic losses.

Assuming the IO structure of the UK economy does not change (i.e. the $\mathbf{A}$ matrix is unchanged), we assume due to infrastructure failures the household demands are affected (due to residential customer disruptions) and some industry demands are reduced to a new level $\mathbf{h}_l < \mathbf{h}$. So, simply the economy reacts by shifting to a new equilibrium

$$\mathbf{x}_l = \mathbf{[I\text{-}A]^{-1}(h}_l \mathbf{+ g + e)} \qquad (7)$$

Consequently, the direct economic losses are = $\mathbf{h} - \mathbf{h}_l$, and the total economic losses (direct + indirect) are = $\mathbf{x} - \mathbf{x}_l$.

For this study, we have used the UK 2015 IO tables[79], which show the balanced accounting of annual supply and demand between 129 macroeconomic private and government industry sectors, households, imports, exports. See Appendix B for the detailed list of 129 sectors included in the IO data for UK.

To translate infrastructure disruptions into economic losses we first matched the infrastructure networks to their represented economic sectors in the IO accounts table, as shown in Table 3-11.

Table 3-11: Mapping of infrastructure networks to the economic sectors in the IO economic structures.

Infrastructure network	Economic sector
Telecoms	61 - Telecommunications services
Electricity	35.1 - Electricity, transmission and distribution
Water	36 - Natural water; water treatment and supply services
Roads	49.3-5 - Land transport services and transport services via pipelines, excluding rail transport

[79] https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/datasets/ukinputoutputanalyticaltablesdetailed

Rail	49.1-2 - Rail transport services

We assumed that household demand losses for a specific sector (s) were proportional to the fraction of numbers of total residential users disrupted due to infrastructure failures. Hence, for electricity, telecoms and water this meant:

$h^s{}_l$ = (numbers of population counts disrupted/total UK population)$\times h^s$ (8)

For road and rail the estimation was based on proportional disruptions of passenger trips:

$h^s{}_l$ = (number of daily trips disrupted /total daily trips considered)$\times h^s$ (9)

3.10.2 Estimating future Input-Output losses

To estimate future economic losses in 2050 we would need data on the future disaggregation of the economy into IO sectors, which would show whether new industry sector classifications are created and how the economic linkages (the **A** matrix) between economic sectors would change in the future. Unfortunately, such data does not exist. The next best alternative was to assume the economic structure remains unchanged, but future losses would grow in relation to future projections in demands and GDP growth. Hence, we estimated future economic losses ($\mathbf{Loss}_{future}$) from Equation (10), where *GDP* is the assumed annual growth rate projection in percentage for the UK, $\mathbf{Loss}_{current}$ are the economic losses estimated with the current economic structure but with the total sector demands and disruptions (from Equations (8) and (9)) based on future projected values, and T = 2050-2017:

$$\mathbf{Loss}_{future} = \left(1 + \frac{GDP}{100}\right)^T \mathbf{Loss}_{current} \qquad (10)$$

We assumed a GDP growth rate of 1.9% for the UK, based on recent studies[80].

[80] https://www.pwc.co.uk/press-room/press-releases/uk-could-remain-top-10global-economy-in-2050.html

4. RESULTS

4.1 Example demonstration of cascading failures and impacts

To demonstrate our failure model and its results we first show an example *failure event*, with the sequences of failures and impacts that follow this event. In this example case we consider single dependencies between networks, where one node of a network connects to only one node of another.

Figure 4-1 shows a failure event initiated in the electricity network at the node location marked by the red star. This initiating failure triggers disruptions of several source-sink flow paths, as a result of which several other nodes are affected. Subsequently in this example, 115 more electricity nodes lose all flow connections and are considered failed. This whole sequence of failures on the electricity network comprises an *Order 0* failure effect.

Due to dependencies directed from electricity towards other networks, the failed electricity nodes disrupt telecoms and railway nodes to trigger the next sequence of failures, which are *Order 1* effects. In the Order 1 effects we see that there are 44 macro cells and 2 exchanges that lose their electricity supply and are considered failed. Also 1 railway utility asset fails due to loss of electricity supply.

The next sequences of failures show how the interdependencies between networks can cause failure feedbacks into the initiating network, thereby triggering further failure cascades. From Figure 4-1 we see that there are *Order 2* failures in the electricity network due to the failures to the telecoms assets on which the electricity nodes were dependent, thus resulting in 18 more electricity nodes losing all flow connections and hence failing. Two water nodes also fail in a similar mechanism to the electricity network failures. These failures are all triggered due to dependencies of these networks on telecoms assets, which failed in Order 1 sequence of events.

The newly failed electricity nodes trigger another set of *Order 3* failure cascade, which result in knocking out the supply to 5 macro cells and 1 more railway utility asset. In this example we did not notice any further feedbacks for the telecoms back to the electricity. But the new railway failure (Order 3) knocks out a whole route section (a link) resulting a several journeys being affected. The final *Order 4* failure sequence demonstrates how widespread the journey disruptions are on the railways network.

Table 4-1 shows the total impacts in terms of the disrupted customers following each Order of failure. This result strongly highlights the significance of considering cascading failures across networks. As shown in the results, an additional 64,000 electricity customers are disrupted due to telecoms failures, while railways is not initially affected by any failures but there is a delayed sequence of events that ultimately disrupt about 82,000 railway passenger journeys.

Table 4-1: Total disruption impact due to the failure event and its triggered failure cascades.

Initiating Network	Order	User Disruptions
Electricity	0	158,801
Telecoms	1	87,885
Rail	1	0
Electricity	2	64,046
Telecoms	3	7,372
Rail	3	0
Rail	4	82,103
Total		**400,207**

Following the estimation of the user disruptions we estimate the infrastructure direct economic losses and total economic losses due to this failure event. Here the only user disruptions are recorded in the electricity, telecoms and railways networks, which result in direct demand losses to the economic sectors with these networks (see Table 3-11). Assuming the disruptions last for 24 hours and the economic losses correspond to losing demand from the equivalent of 24 hours of customers across sectors, the direct and total economic losses estimated for this event are shown Table 4-2. Here the direct demand losses of £131,507/day in the electricity sector correspond to the total customer disrupted (Order 0 + Oder 2), and similarly the telecoms and rail demand losses correspond to their total customer losses. Due to the forward and backward linkages in the economic IO model, there are indirect economic losses to all sectors that use electricity, telecoms and railways outputs, and some of these losses feedback to these infrastructure sectors as well. Here, the indirect losses for electricity are also almost as high as direct losses, which shows electricity has significant feedbacks from the rest of the economic systems. The sector 'Other' corresponds to the total losses added across all 124 non-infrastructure sectors in the UK economy (see Appendix B), which have about £345,000/day indirect economic losses. Overall the economic impact of this event results in about £0.92 million/day total economic losses.

Table 4-2: Total economic losses due to the failure event and its triggered failure cascades.

Network/Sector	Direct economic losses (£/day)	Indirect economic losses (£/day)	Total economic losses (£/day)
Electricity	131,507	98,699	230,206
Telecoms	71,233	4,575	75,808
Rail	260,274	636	260,910
Water	0	286	286
Road	0	6,667	6,667
Others	0	345,069	345,069
Total	**463,014**	**455,932**	**918,946**

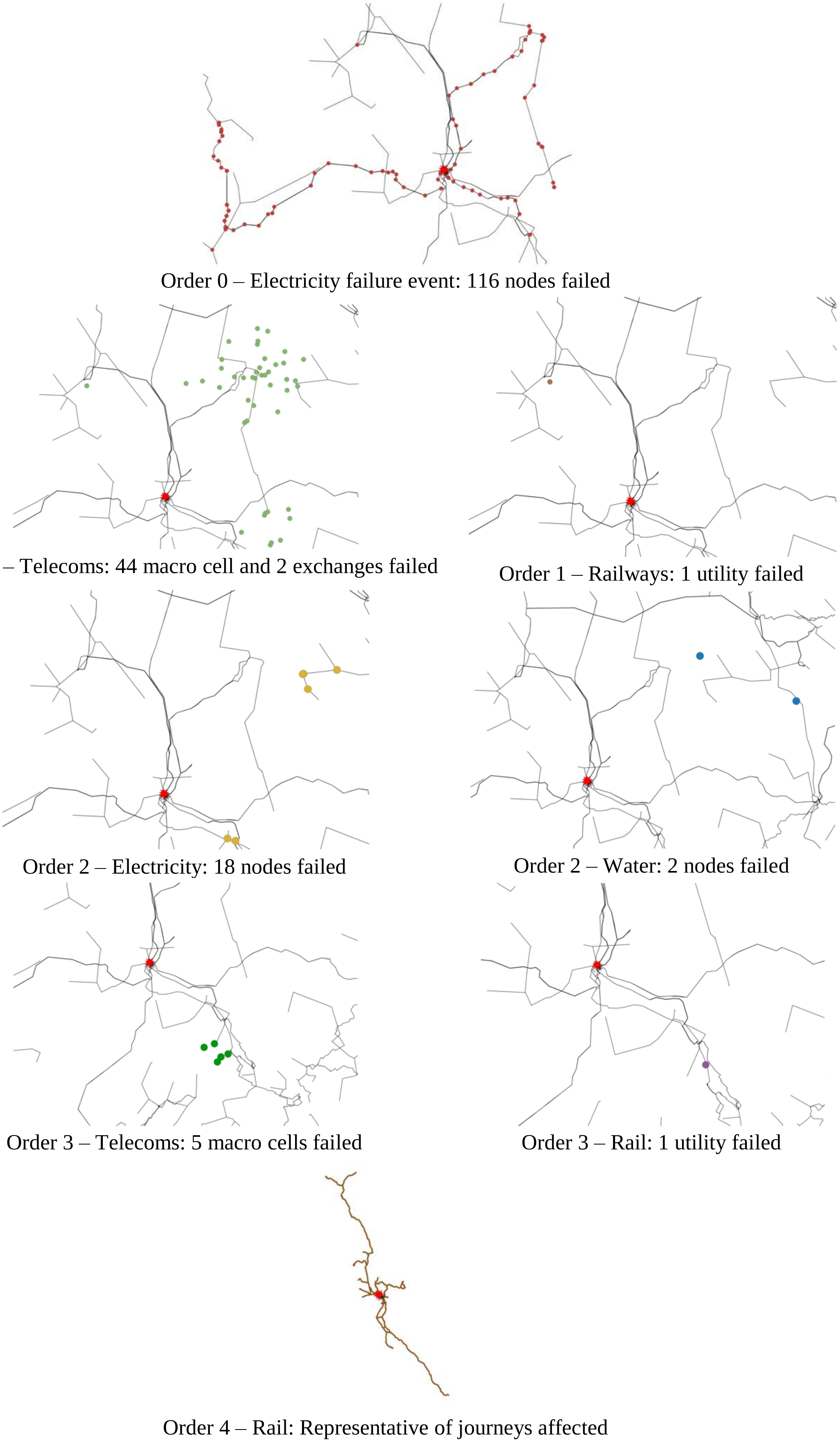


Figure 4-1: Demonstration of example failure cascading event and the sequences of failures it generates across multiple networks.

4.2 Understanding systemic propagation of failures

Systemic assessment of failures involves analysing a large numbers of failure events and inferring some generalised behaviours of networks in terms of the instances and impacts of failure propagations. We conducted such systemic assessment to answer the following two questions:

1. What are the different (inter)dependencies between networks and how do these affect failure propagation?
2. Can we see a difference in the failure propagation if we increase the connections between networks?

To understand the overall role of network interdependencies in failures cascades, we looked at the exhaustive set of all 'single point' initiating failure events in a network. Here single point implied that an individual node from a network was removed and then its failure sequences were estimated by the model. We considered the exhaustive analysis for the electricity and telecoms network nodes, because every other network was dependent on these two networks.

We further looked at the failure propagation effects when the degrees of network connections were increased. This was the done to see whether there were any reductions in cascading failures if more redundancy were added to the networks.

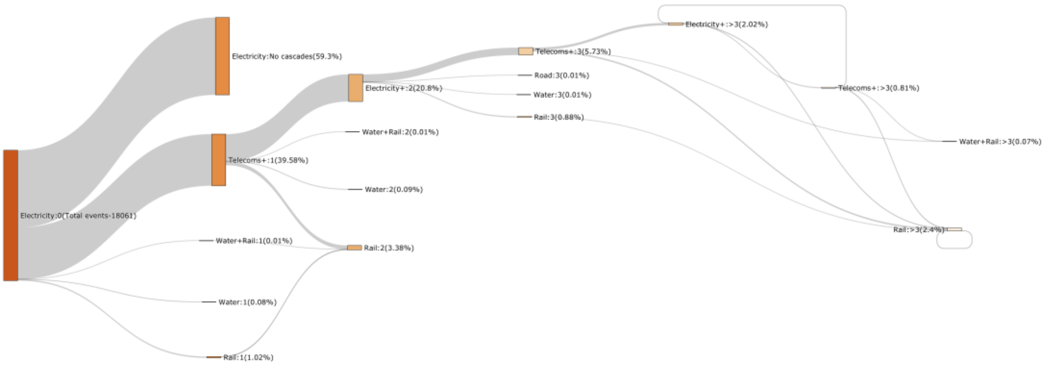
4.2.1 Extent of cascading failures

Figure 4-2 shows Sankey diagrams of the chain of cascading events in the current system state due to failures initiated in the electricity network, by testing all 18,061 individual node failures. We note here that the dimensions of the rectangles and arrows in the three plots are not shown to the same scale, and to avoid confusion we have reported the values next to each arrow. The first rectangle in each plot shows the total number of failure events, which are same in each case. The subsequent rectangles show what percentages of the total failure events correspond to particular sector(s) and order effect – for example Rail:1(1.02%) means 1.02% of all failure events resulted in Order 1 Rail failures only. In the notation Telecoms+ (or Electricity+) implies that Telecoms (or Electricity) is one of the disrupted sectors and there might possibly be other sectors (water, railways, roads) disrupted simultaneously. From the first result in Figure 4-2(a), where we assumed that a selected node from one network was dependent upon only one node of another network, we infer that: (1) The most significant chain of cascading failures is from electricity to telecoms, with about 40% events leading to telecoms and at least one of rail and water disruptions, with further 20% events leading to electricity failures, and 5.7% to another order of telecoms failures; and (2) About 5.2% failure cascades go to Order 4 and above.

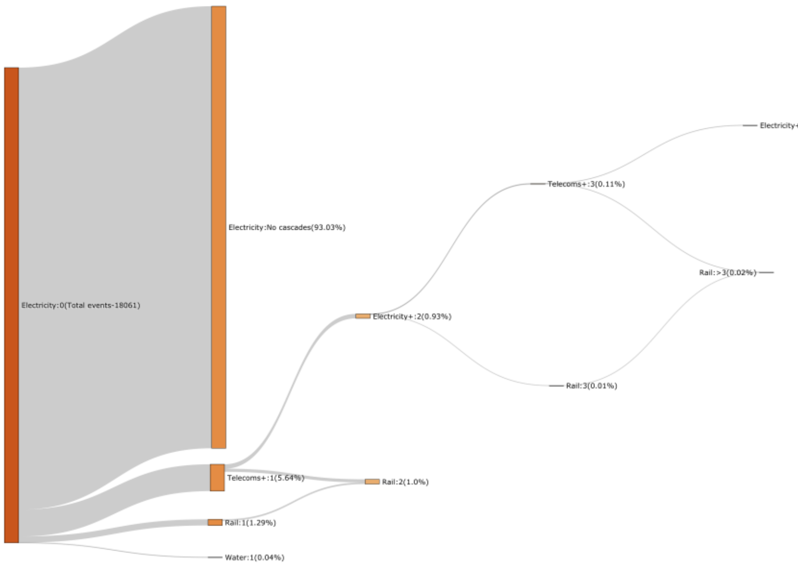
In the case where the degrees of connections are increased to two, by linking each dependent node to two nodes of the supplying network, we see from Figure 4-2(b) that: (1) Cascading failures are reduced significantly, with about 5.6% events leading to telecoms and at least one of rail and water disruptions, with further 0.9% events leading to electricity failures, and 0.11% to another order of telecoms failures; and (2) About 0.02% failure cascades go to Order 4 and above.

Figure 4-2(c) shows the results when the degrees of connections are increased to three, linking each dependent node to three nodes of the supplying network. The results show that: (1)

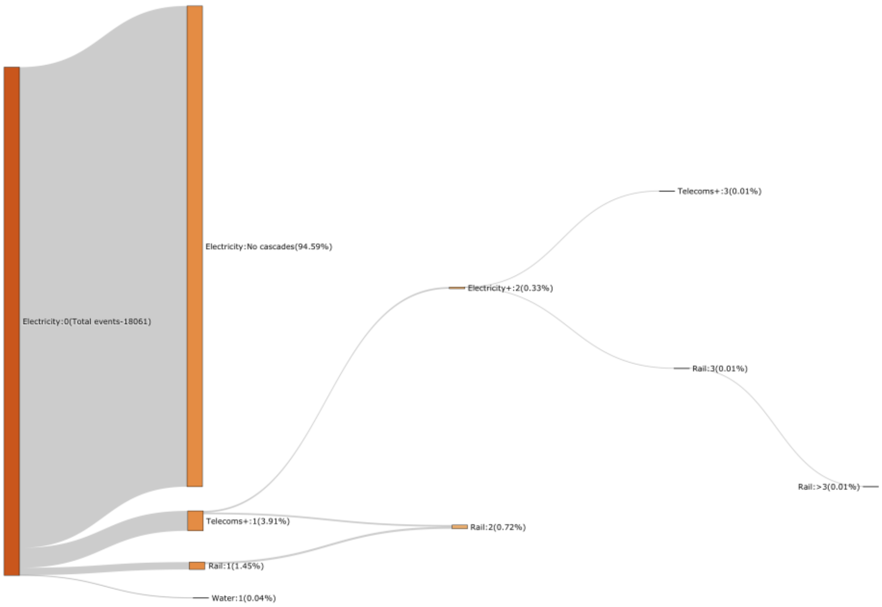
Cascading failures are again reduced significantly, with about 3.9% events leading to telecoms and at least one of rail and water disruptions, with further 0.33% events leading to electricity failures, and 0.01% to another order of telecoms failures; and (2) Order 4 and above cascading failures are avoided.



(a) Single connections



(b) Two connections



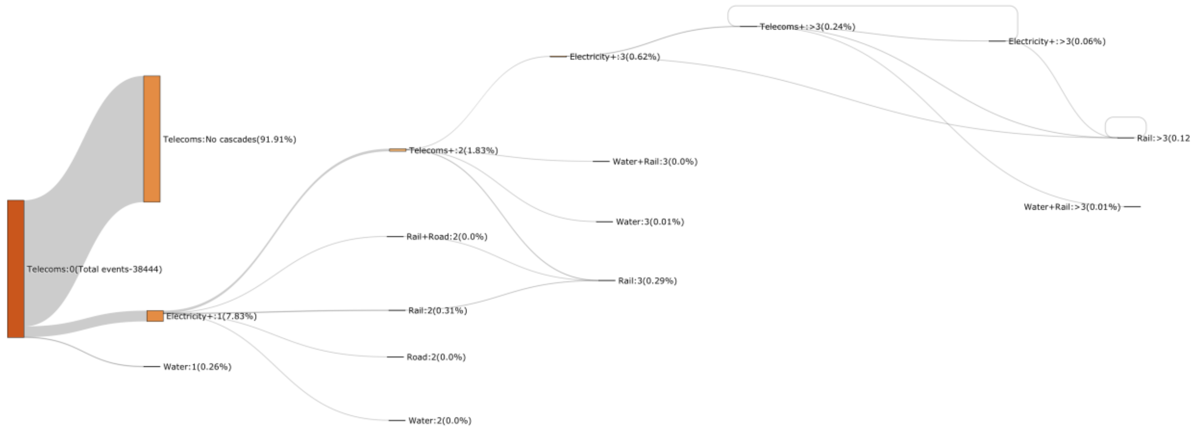
(c) Three connections

Figure 4-2: Failure propagation showing numbers of instances of individual failure events cascading from electricity to other networks and beyond under different degrees of connections between links.

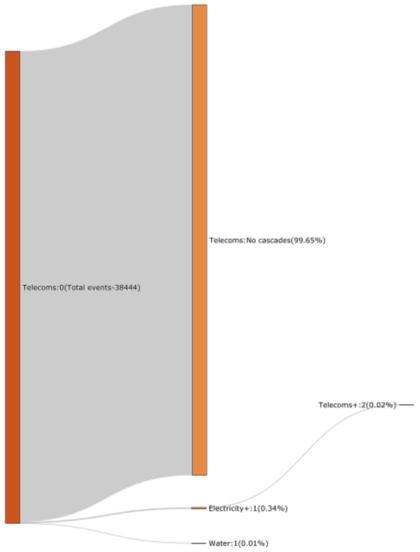
Figure 4-3 shows similar Sankey diagrams of the chain of cascading events in the current system state due to failures initiated in the telecoms network, by testing all 38,444 individual node failures. From the single connections result in Figure 4-3(a) we infer that: (1) In comparison to electricity, there are fewer cascading failures from telecoms, with about 7.8% events leading to electricity and at least one of rail and water disruptions, with further 1.8% events leading to another order of telecoms failures; and (2) About 0.43% failure cascades go to order 4 and above. Telecoms failures have less cascades because we assume that if at least one connection to a working exchange or macro cell still exists then the dependent asset is still functioning. Hence in reality the model accounts for two dependencies on telecoms, but since on most cases the macro cells are dependent on the exchanges, so if the exchange fails then the macro cell would fail as well.

In the two connections case for telecoms we see from Figure 4-3(b) that: (1) Cascading failures are almost gone, with about 0.34% events leading to electricity and at least one of rail and water disruptions, with further 0.02% events leading to another order of telecoms failures; and (2) Order 3 and above failures are eliminated.

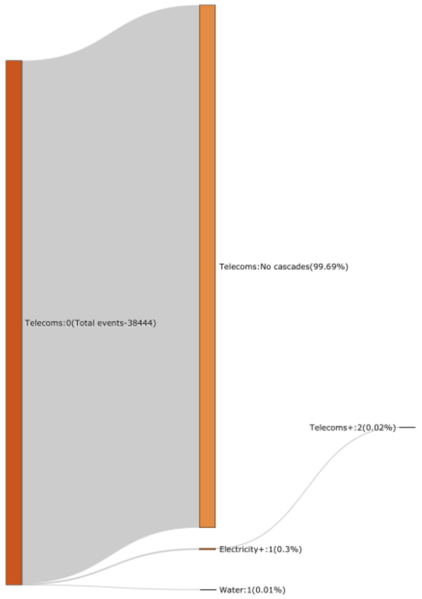
Similarly the three connections case results of Figure 4-3(c) show that: (1) Cascading failures are almost gone, with about 0.3% events leading to electricity and at least one of rail and water disruptions, with further 0.02% events leading to another order of telecoms failures; (2) Order 3 and above failures are eliminated.



(a) Single connections



(b) Two connections



(c) Three connections

Figure 4-3: Failure propagation showing numbers of instances of individual failure events cascading from telecoms to other networks and beyond under different degrees of connections between links.

4.2.2 Failure impacts as user disruptions

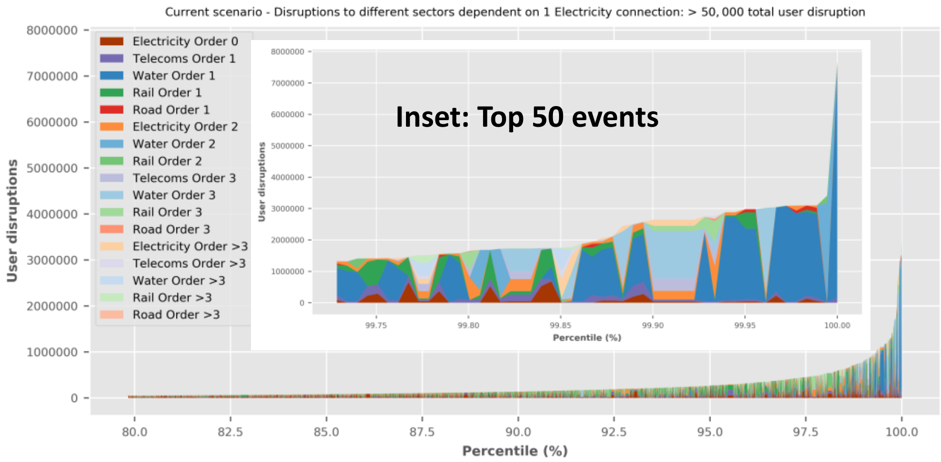
Comparing the failure impacts in terms of the numbers of disrupted users (customers over a day) of each sector, and cumulatively, further shows the failure events whose disruptions create highest impacts and the effect on these disruptions if the degrees of connections are increased.

Figure 4-4 shows all user disruptions, across all current day networks, due to the failures initiated in the electricity network. Only those failure events are shown that led to >50,000 user disruptions, which are reported as a percentile (on the x-axis) of the exhaustive set of events. For visual clarity, each figure also shows the top 50 failure event outcomes.

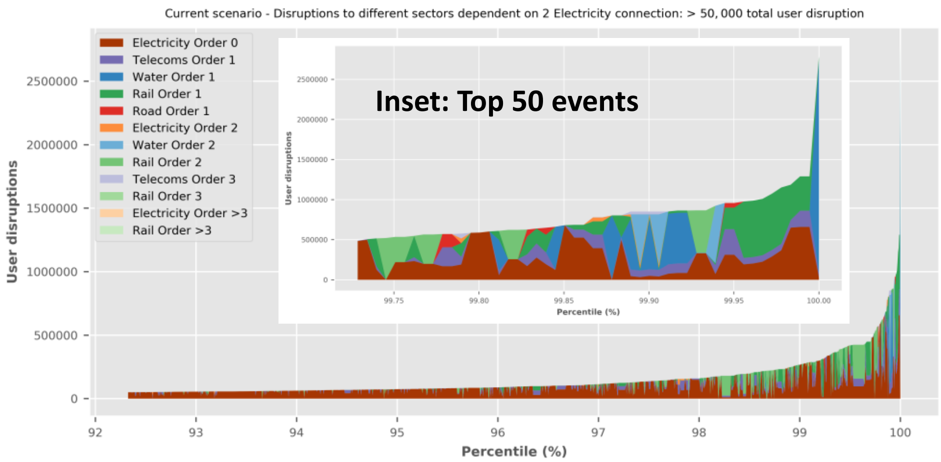
From the first result in Figure 4-4(a), with the single connections, we infer that: (1) There are about 20% of failure events for which the failures are above 50,000 which is a significant number of failures events out of the total of 18,061 events; (2) The highest impacts are recorded due to Order 1 and Order 3 disruptions in the water supply network that has very high demands concentrated at individual nodes, given that it is a high-level network. The largest disruption of about 8 million users is mainly due to a knock-on effect on the water network from an electricity failure; and (3) There are clusters of failure events that produce similar disruptions, which could indicate that these are assets that affect similar flow paths and dependencies. If such clustered failures occurred simultaneously then we might see similar impacts. For example, if there are three nodes close to each other and all cause the same failure impact then there it is very likely that they are all knocking out each other when failed individually. Hence, if all three were to fail at the same time, then it would produce the same failure effect and impact.

In the two connections case we see from Figure 4-4(b) that: (1) There is a significant reduction in the numbers of cases of failures exceeding 50,000 user disruptions, which is now about 7% of total failure events; (2) The highest failure impact is now around 2.6 million users, which is again due to Order 1 water network failures. But most of the high impact failures in the water network are eliminated in comparison to the single connections case. There are some Order 1 railway failures that also contribute to the highest impact events.

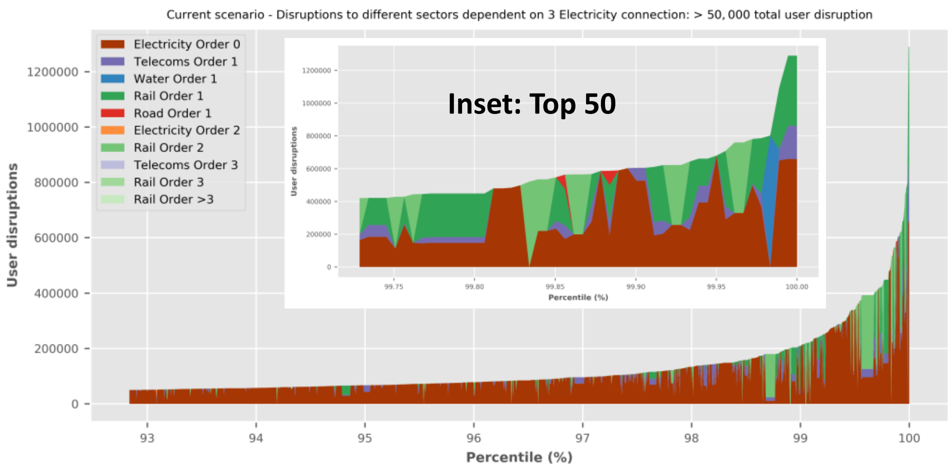
Figure 4-4(c) shows the three connections case results where: (1) The number of failures exceeding 50,000 users does not differ much from the case with two connections case, and is around 7% of total failure events; (2) There is a significant reduction in the highest failure impact event, which now results in 1.3 million user disruptions due to Order 1 telecoms and railway failure initiated from Order 0 electricity failures; (3) Most of the high impact water failure have been eliminated.



(a) Single connections



(b) Two connections



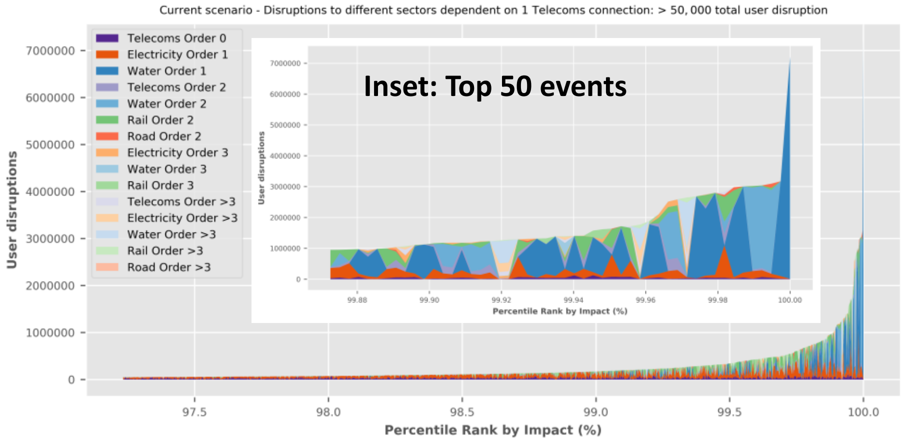
(c) Three connections

Figure 4-4: Magnitudes of customer disruptions due to failures initiated in the electricity network under different degrees of connections between networks.

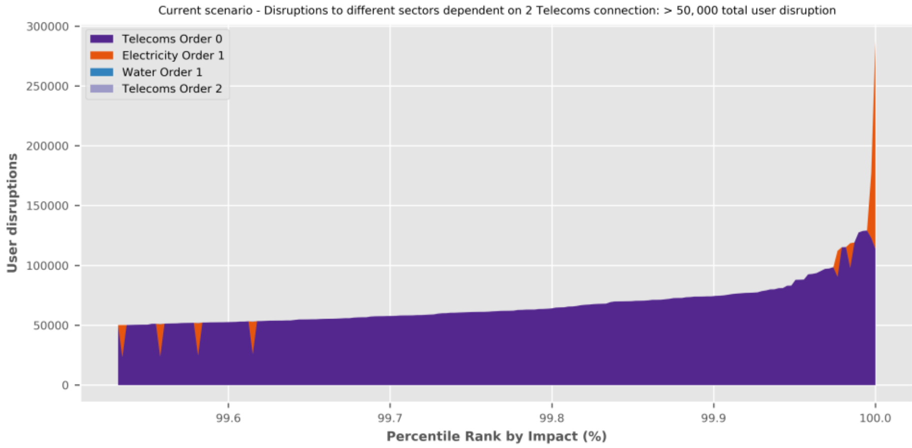
Figure 4-5 shows the impacts for the failure initiated in the telecoms network. The first result in Figure 4-5(a), with the single connections, shows that: (1) There are about 2.7% of failure events for which the failures are above 50,000 which is a small but still significant number of failures events out of the total of 38,444 events; (2) Similar to the case of the electricity network initiated disruptions, the highest impacts are recorded due to Order 1 and Order 3 disruptions in the water supply network that has very high demands concentrated at individual nodes. The largest disruption of about 7 million users is mainly due to a knock-on effect on the water network from telecoms failure; and (3) There are clusters of failure events that produce similar disruptions, which could indicate that these are assets that affect similar flow paths and connections. If such clustered failures occurred simultaneously then we might see similar impacts.

In the two connections case we see from Figure 4-5(b) that: (1) There is a significant reduction in the numbers of cases of failures exceeding 50,000 user disruptions, which is now about 0.5% of total failure events; (2) The highest failure impact is now around 280,000 users, which is due to Order 1 electricity network failures, following a telecoms failure; and (3) Almost all cascading failure have been eliminated, which is mainly because the telecoms provides the extra redundancies from both macro cell and exchange connections, which is effect makes it a case of four degree of connections.

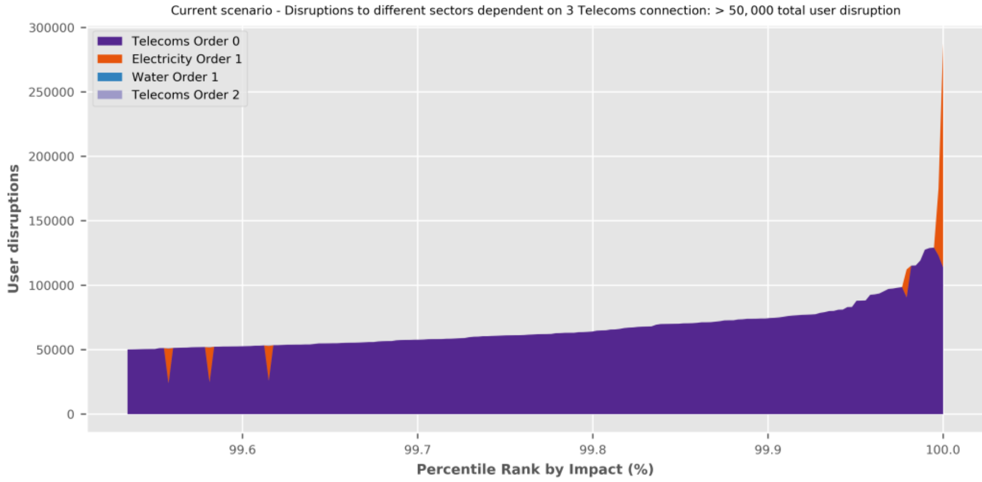
The results with three degrees of connections of Figure 4-5(c) are very similar to the results of the two interdependencies case, with the exception of a few more cascading failures to electricity being eliminated. This shows that there is not much gain in adding further redundancy with respect to controlling telecoms failures.



(a) Single connections



(b) Two connections



(c) Three connections

Figure 4-5: Magnitudes of customer disruptions due to failures initiated in the telecoms network under different degrees of connections between networks.

4.2.3 Failure impacts as macroeconomic losses

The economic losses resulting from the user disruptions are presented next, with specific focus on the 50 worst-case of impacts ranked in terms of the cumulative user disruptions. These economic losses show how the economic flows are first disrupted due to demand perturbations economic sectors causing *direct losses*. The rest of the economy reacts to these losses and adjusts to a new equilibrium resulting in *indirect and total output losses*. We note that the cumulative user disruptions for an individual infrastructure network contribute towards direct economic losses, as the economic effects are considered to follow after all the user disruptions have been accounted for.

As described in Section 3.10 the economic IO model developed for this study is a linear model where the output losses are a linear factor ($\mathbf{L} = [\mathbf{I}\text{-}\mathbf{A}]^{-1}$) times the direct losses. One of the inferences from the IO data is to find the *multiplier effects*, as explained and estimated by the Office of National Statistics from their IO data[81], of each sector's demand losses on the rest of the economy, which show the ratio between the total economic losses and the demand losses in a particular sector. Table 4-3 shows these multiplier effects for the infrastructure network specific economic sectors, where for example we see that for every 1 unit of direct demand losses in the electricity sector the total economic losses will be 2.36. These multiplier effects show which sector has greater interdependencies to the rest of the economic sectors, with electricity being a basic commodity that is used by most sectors so it has the highest multiplier effects.

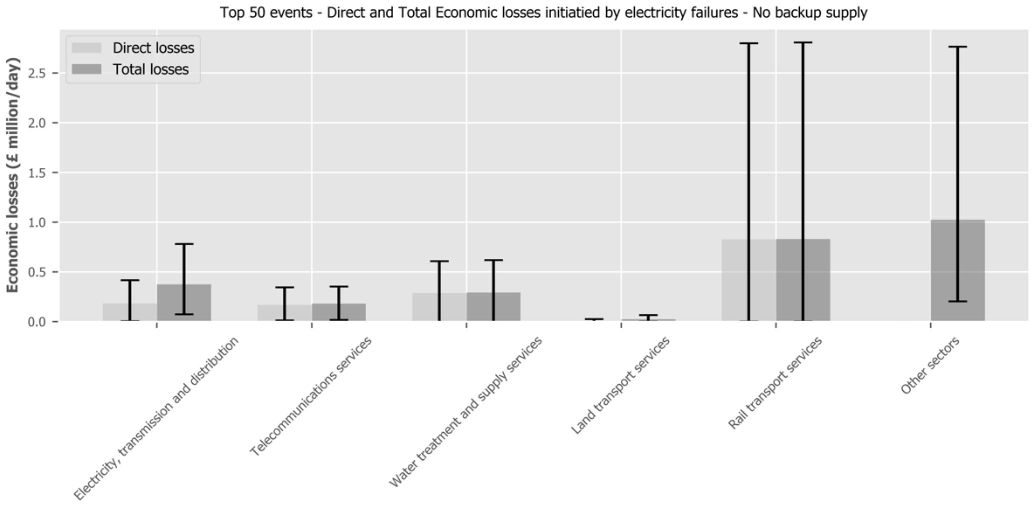
Table 4-3: Infrastructure networks specific economic sectors and their multiplier effects.

Economic sector	Multiplier effect
61 - Telecommunications services	1.41
35.1 - Electricity, transmission and distribution	2.36
36 - Natural water; water treatment and supply services	1.53
49.3-5 - Land transport services and transport services via pipelines, excluding rail transport	1.64
49.1-3 - Rail transport services	1.95

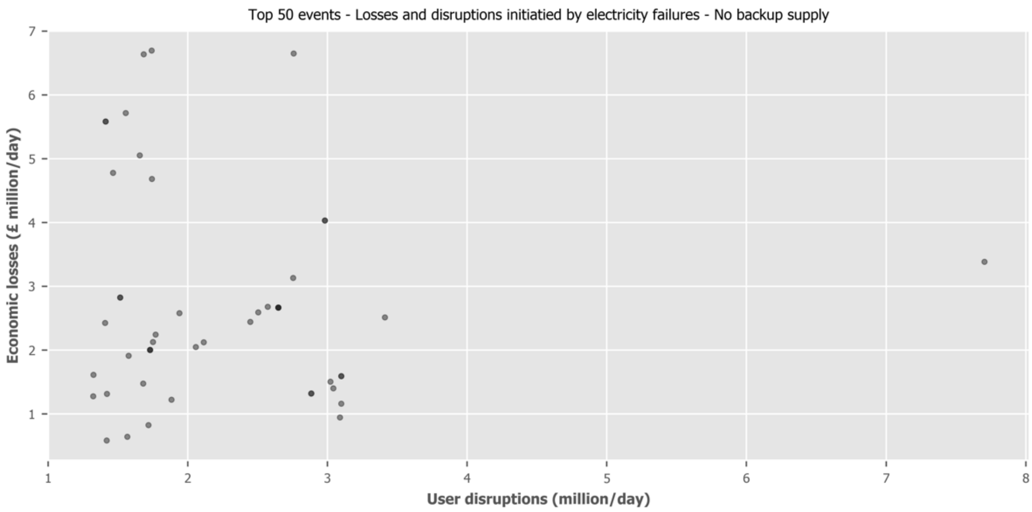
Figure 4-6(a) shows error bar plots with the mean values and 95% confidence intervals for economic losses averaged across all top 50 user disruptions events for failures initiated by the electricity networks and considering only single degrees of connections. The results show the direct and total economic losses for the infrastructures specific sectors and the rest of the economy ('Other' sectors). The important insights gained from this result are that the largest economic losses are recorded in the railways sectors, which are as high as £2.7 million/day. Earlier, from Figure 4-4(a) we saw that user disruptions were highest in the water network. This difference arises because proportionally the railway sector is more impacted in terms of reduced capacity to meet journey demands as compared to the water supply sectors proportional reduction in demands. The analysis shows that direct losses for the top 50 failure events vary between £0.36 million/day – £3.4 million/day and total losses vary between £0.58 million/day – £6.7 million/day, with the event specific total losses being 1.52 – 1.99 times the direct losses.

[81] Howse. J. (2013). Input-output analytical tables: methods and application to UK national accounts. Office of National Statistics, UK. Available online: https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/inputoutputanalyticaltables/methodsandapplicationtouknationalaccounts, Accessed April 2020.

It is also important to note that economic losses and user disruptions might not be similarly ranked for failure events, i.e., the largest user disruptions might not result in the largest economic losses. This is highlighted in Figure 4-6(b) where the largest user disruption event of 7.8 million user disruptions has about £3.2 million/day economic losses but events with less than 3 million user disruptions produce the highest economic impacts. This is again due to the proportional impacts on railway capacity to meet demands which result in highest economic impacts.



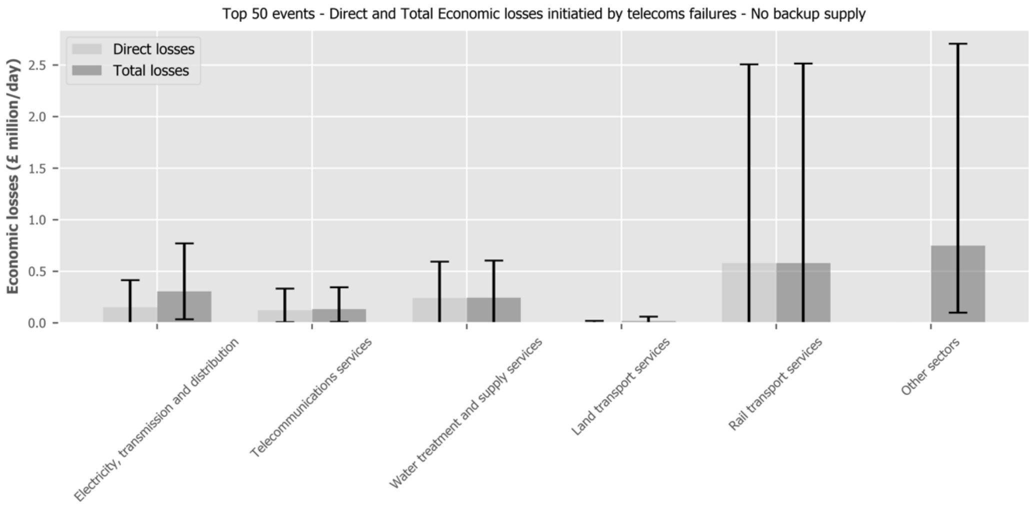
(a) Direct and total macroeconomic losses - Single connections



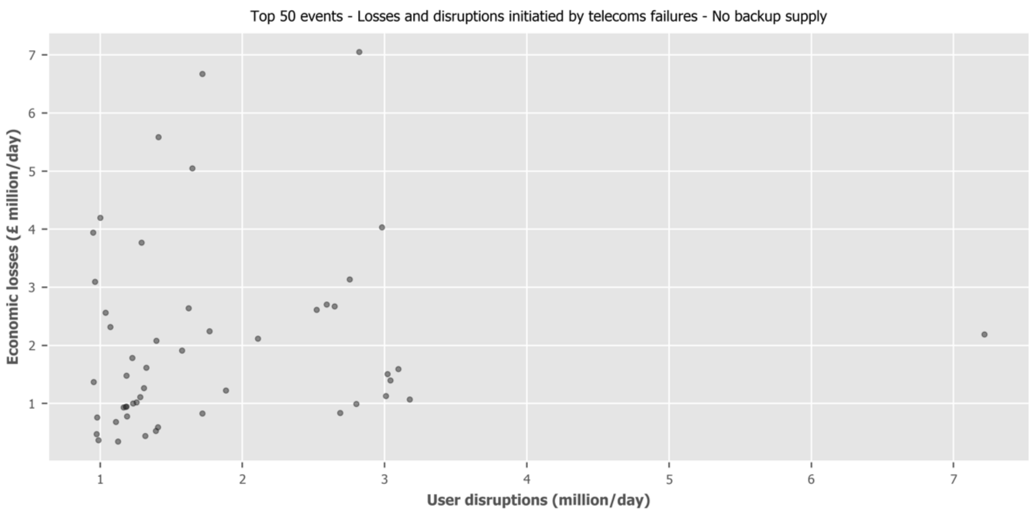
(b) Total economic losses vs User disruptions – Single connections

Figure 4-6: (a) Mean value with 95% CI estimates of direct and total macroeconomic losses across top 50 user disrupted events initiated by electricity failures; (b) scatter plot between the total economic losses and user disruptions.

Figure 4-7 shows the similar results for the failure events initiated in telecoms network with single degrees of connections. Here again the highest economic losses are recorded in the railway sector (Figure 4-7(a)), which can be high as £2.5 million/day. The analysis shows that direct losses for the top 50 events vary between £0.22 – £3.6 million/day and total losses vary between £0.34 – £7.0 million/day, with the event specific total losses being 1.52 – 1.99 times the direct losses. Again, the largest user disruption event of 7.2 million user disruptions has about 2.1 £million/day economic losses but events with less than 3 million user disruptions produce the highest economic impacts. This is again due to the proportional impacts on railway sector demands which result in highest economic impacts.



(a) Direct and total macroeconomic losses - Single connections

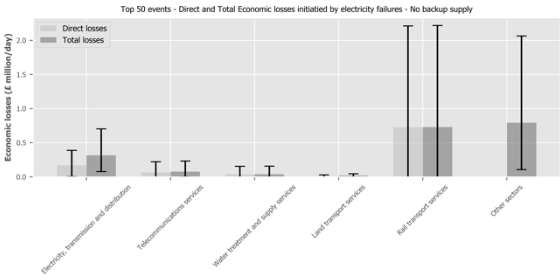


(b) Total economic losses vs User disruptions – Single connections

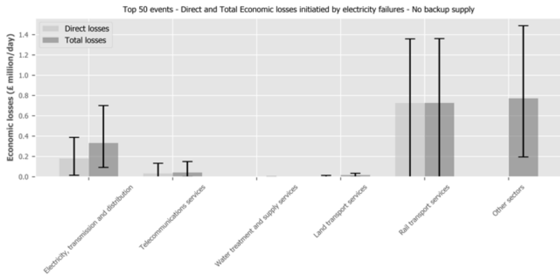
Figure 4-7: (a) Mean value with 95% CI estimates of direct and total macroeconomic losses across top 50 user disrupted events initiated by telecoms failures; (b) scatter plot between the total economic losses and user disruptions.

As the degrees of connections are increased the economic impacts will decrease, and as the network failure cascades decrease the economic impacts will be driven mostly by the failures in the initiating sector. This is very pronounced in the cases where the telecoms network-initiated failures are analysis with two and three degrees of connections. Figure 4-8(a)-(b) shows the direct and total macroeconomic losses for the top 50 user disruption event with electricity-initiated failures with two and three connections linkages. We note that these are not necessarily the same 50 events in each case, as some for some events the failures are significantly reduced when more redundancies are added between networks. From the results of Figure 4-8(a)-(b) economic losses to railways still remain the most dominant but their highest total losses are respectively reduced to about £2.1 million/day and £1.4 million/day. The overall demand losses range from £0.17 million/day – £2.5 million/day and total losses range from £0.26 million/day – £4.92 million/day for the two connections case, while for the three connections case the such losses are in the ranges £0.17 million/day – £1.9 million/day and £0.26 million/day – £3.77 million/day respectively. For both cases the event specific total losses are 1.52 – 2.36 times the direct losses, with values being highest when the direct economic losses are mainly due to electricity disruptions.

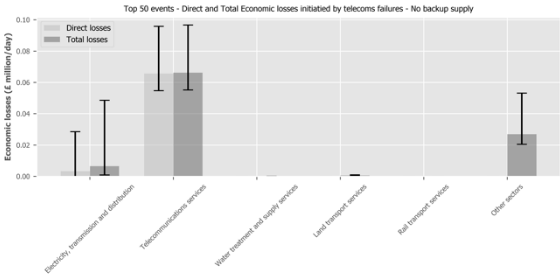
Figure 4-8(c)-(d) shows similar results as the Figure 4-8(a)-(b), but with telecoms-initiated failures with two and three connections respectively. Since the user disruptions for both cases are very similar (see Figure 4-5(b)-(c)) the economic losses show similar results. In both cases the economic losses to telecoms are the most dominant, since most cascading failures are eliminated. The highest direct losses are only about £0.09 million/day in both cases. The overall demand losses range from £0.05 million – £0.19 million/day and total losses range from £0.08 million/day – £0.36 million/day for both cases. The event specific total losses are 1.41 – 1.93 times the direct losses, with lower values occurring when there are telecoms disruptions only while the multiplier effect gets increased when electricity disruptions also contribute to economic losses.



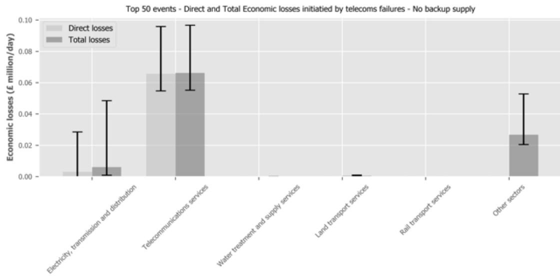
(a) Two connections



(b) Three connections



(c) Two connections



(d) Three connections

Figure 4-8: Mean value with 95% CI estimates of direct and total macroeconomic losses across top 50 user disrupted events initiated by electricity failures with instances of (a) two connections and (b) three connections, and events initiated by telecoms failures with instances of (c) two connections and (d) three connections.

4.3 Role of backups

To understand the role of backups in a systemic way, we re-simulated all single point failure scenarios, with the additional constraint of having backups. Such systemic assessment was done to answer the following two questions:

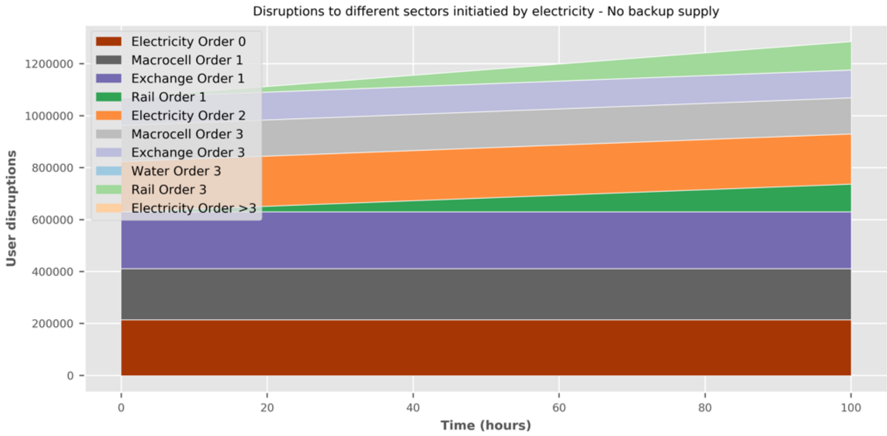
1. What is the effect of adding backups to the different interdependent nodes?
2. What are the failure sequences and over what timeframe do they occur?

We assumed that the disruptions lasted 100 hours, in order to exhaust the backups and see how the disruptions would progress post-backup. Given, that we did not consider any hourly load profiles for any sector we assumed that: (1) For the electricity, telecoms and water sectors once a disruption at some time *t* (<100) was recorded with a certain number of customers it would last till the completion of the 100 hours; and (2) For the transport sectors the daily number of passengers were assumed to be uniformly divided in the hourly intervals, hence the growth progression of the numbers of disrupted passengers would be linear from the time of initial disruptions till the completion of 100 hours.

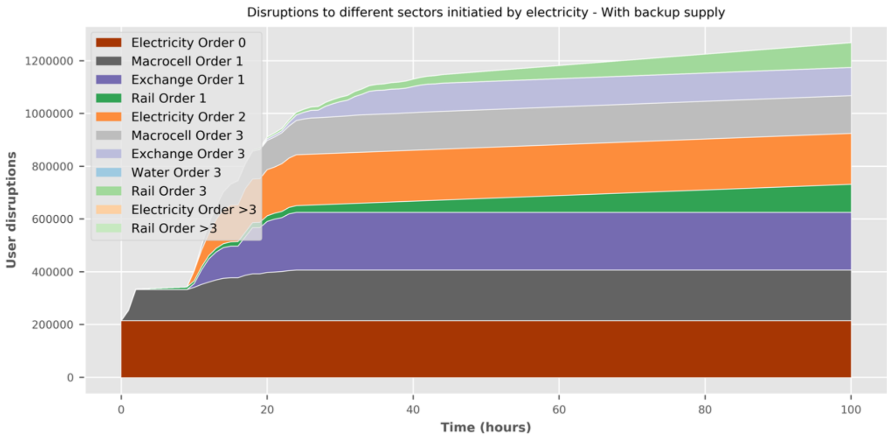
Figure 4-9 shows results for one example event, where we compare the results when there are (a) no backups and (b) backup supply, corresponding to the case of having single connections between networks. From the first result, of Figure 4-9(a) with no backups we see that the disruptions all begin at time t=0, continuing till the 100 hours. Due to the assumptions of linear change in rail disruptions over time, there is a steady growth of the disruptions to around 118 million customer-hours by the end of the over failure event.

When backups are added to the telecoms assets, in this case, there is a delay in disruptions which vary across disrupted telecoms assets due to the assumed gamma probabilistic distribution. The result in Figure 4-9(b) shows the average disruption over time across 20 simulations of the same failure event. After some initial telecoms disruptions in the first 2 hours, mainly of macro cells, there is second sequence of telecoms exchange and macro cell disruptions around 10 hours which triggers the further order effects across sectors. Once the backups have been exhausted at around 24 hours, the disruptions grow to around 104 million customer-hours till the 100 hours.

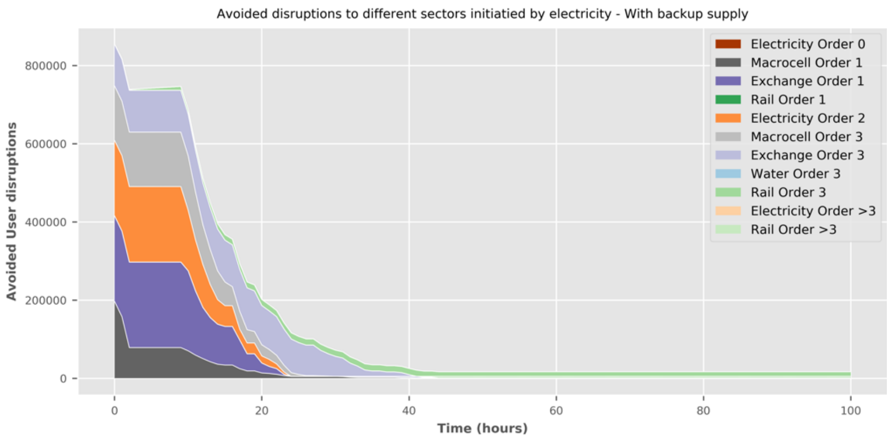
Figure 4-9(c) quantifies the gains made in this example by adding backup supply. Here the difference between the results of Figure 4-9(a) with Figure 4-9(b) are shown as the avoided disruptions. The results highlight that for this event cumulatively 14 million customer-hours of disruptions are avoided due to the backup supply, and 57%-87% of the total avoided disruptions are acquired within the first 10-24 hours. This highlights the importance of having backup supply and crucially also shows that if the original disrupted networks were to be restored then there are several gains that can be made if the repairs occurred within 10-24 hours after the initiating failure event. Especially if the repairs happened closer to 10 hours then most of the Order 2 are greater disruptions could be avoided.



(a) Failure propagation over 100 hours assuming no backups



(b) Failure propagation over 100 hours assuming with backups

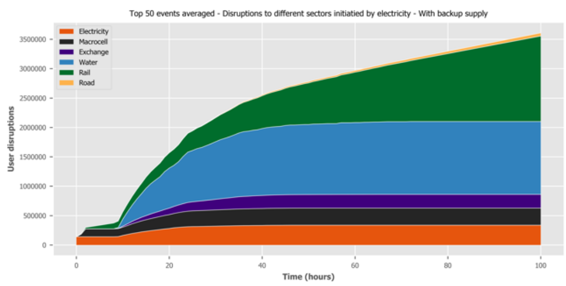


(c) Avoided disruptions over time with backups.

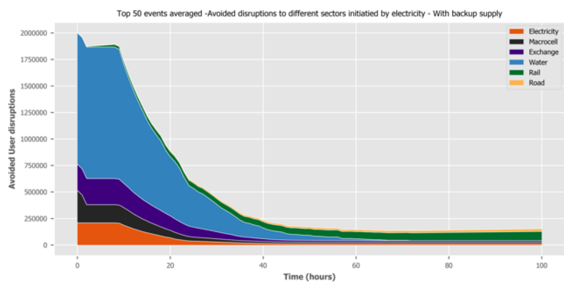
Figure 4-9: Results of example event disruptions showing the progression of failure over time with: (a) no backups; (b) with backups; and (c) difference between the two cases.

To see whether the above hypothesis can be generalised beyond this one event, we look at the time-averages of disruptions across the top 50 worst-case failure events with single degrees of connections. We investigate the top 50 events of cumulative user disruptions for failures initiated by the electricity network, and also the top 50 events of cumulative user disruptions for failures initiated by the telecoms network. These results are shown in Figure 4-10. For the case when the failures are initiated by the electricity network (Figure 4-10(a)-(b)) on average backup supply effects prevent disruptions to grow till around 10 hours after which the impacts grow significantly till around 24 hours and further till up to 42 hours when the electricity backup supply of telecoms exchanges are first exhausted followed by water backups being exhausted. The time-averaged cumulative losses across these events result in about 247 million customer hours of disruptions over 100 hours (Figure 4-10(a)), which is about 51 million customer hours or 17% less (Figure 4-10(b)) than the disruptions if there were no backups. 33%-75% of the total avoided disruptions occur between the first 10-30 hours when most of the backup supply is still working.

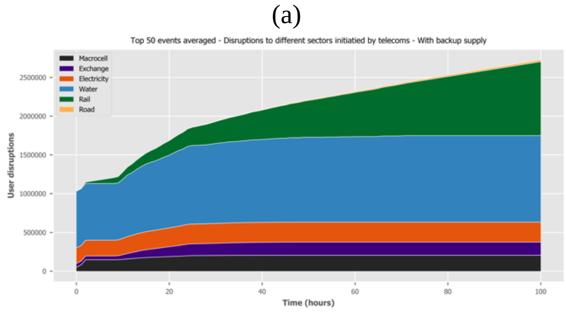
When the failures are initiated by the telecoms network (Figure 4-10(c)-(d)) there are no telecoms backup supply so significant disruptions occur from the start. But later when the electricity network creates further disruptions the electricity backup supply effects prevent disruptions to grow till around 10 hours after which the impacts grow significantly till around 24 hours when the electricity backup supply of telecoms assets are first exhausted. There are some more delayed disruptions when some of the electricity supply of the water assets is exhausted, though this is not very significant. The time-averaged cumulative losses across these events result in about 212 million customer hours of disruptions over 100 hours (Figure 4-10(c)), which is about 16 million customer hours or 7% less (Figure 4-10(d)) than the disruptions if there were no backups. 35%-75% of the total avoided disruptions occur between the first 10-30 hours when most of the backup supply is still working, which is very similar to behaviour for the electricity induced failures.



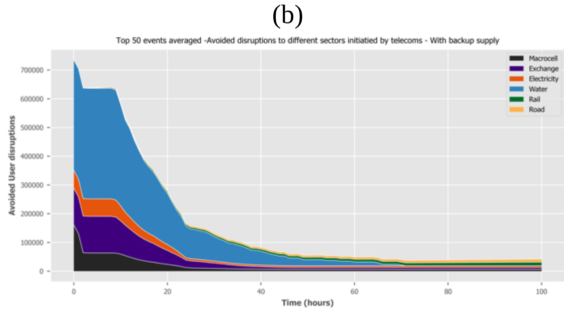
(a)



(b)



(c)



(d)

Figure 4-10: Time-averaged values of top 50 user disruption events for electricity and telecoms initiated failures showing the progression of failure over time with backups (a/c), and the avoided disruptions in comparison to when there was no backup supply (b/d).

4.4 Comparing effectiveness of different options

The two types of resilience options that we have investigated in this study involve: (1) adding more redundancies to connections between networks; and (2) incorporating backup supply for electricity into different assets for a given duration of network inoperability. We now look at the combined effectiveness of these options in preventing disruptions across each network.

We consider the case of 'single connections and no backup supply' as the *baseline case*. From the cumulative user disruptions estimated for this baseline case we select the top 50 most severe events. For the same top 50 events we then estimate the disruptions for the following resilience enhancing options: (1) Two connections (2C); (2) Three connections (3C); (3) Backup supply (B); (4) Two connections and with backup supply (2C+B); and (5) Three connections and with backup supply (3C+B). We find the percentage difference between the user disruptions for each event corresponding to each case and take the average across all events to find the average reduction in disruptions due to the given resilience enhancing option. This is a measure of the average effectiveness of the option, with respect to lowering the worst cases of baseline impacts. We note that we will get similar results if we had chosen economic losses as a metric because the economic losses are a linear function of the user disruptions, as the IO model used is this study is a linear model.

Figure 4-11 shows the results for the case when the disruptions are initiated by the electricity network failures, where the results for the cases (1)-(5) are shown anti-clockwise on each plot. The axis of each plot shows the percentage reduction in average disruptions for each resilience enhancing option. From the results we can see that mostly adding two connections (2C) is very effective by itself in reducing the user disruptions and the gains made by adding another degree of connections (3C) are marginal. With respect to the 2C and 3C options, for the selected 50 worst-case disruptive events in the baseline case, the electricity disruptions are reduced by about 70% in both cases mainly because higher order electricity failures resulting for telecoms networks are eliminated. This is evident when we see that telecoms disruptions are reduced on average by 91%-95%, eliminating further electricity disruptions. Similarly, water and road disruptions are reduced on average by at least 90% and at most 100%. For railways adding three connections (3C) reduce disruptions on average by 93% in comparison to 82% reduction with two connections (2C), showing that there are some gains the adding more redundancy to reduce railway disruptions. The backup supply (B) case is most effective for roads where on average disruptions are reduced by about 40%, and for other networks the gains are between 10%-23%. With the options that include combined backup and increased connections, the biggest gains are made in the electricity networks where the 2C+B option reduces disruptions on average by 78% and the 3C+B option reduces disruptions on average by 81%, a gain of 10%-13% over the options with no backup supply. This shows that adding backup electricity supply to other networks can in turn reduce and delay further cascading impacts on the electricity network and help avoiding disruptions. The effects of all these options in reducing the total cumulative disruptions are quite effective with backup supply by itself reducing impacts by 20% and with increased redundancies and backup supply the disruptions are reduced on average by 89% (2C+B) and 94% (3C+B). Since all these event results in causing cumulative disruptions in excess of 1 million users and £0.5 million/day (see Figure 4-6) such gains are quite significant.

Similar results for failures initiated in the telecoms networks are not shown here because most the cascading disruptions are eliminated with the 2C and 3C options are seen in Figure

4-3 and Figure 4-5, which shows that these options by themselves are most effective in reducing telecoms initiated disruptions.

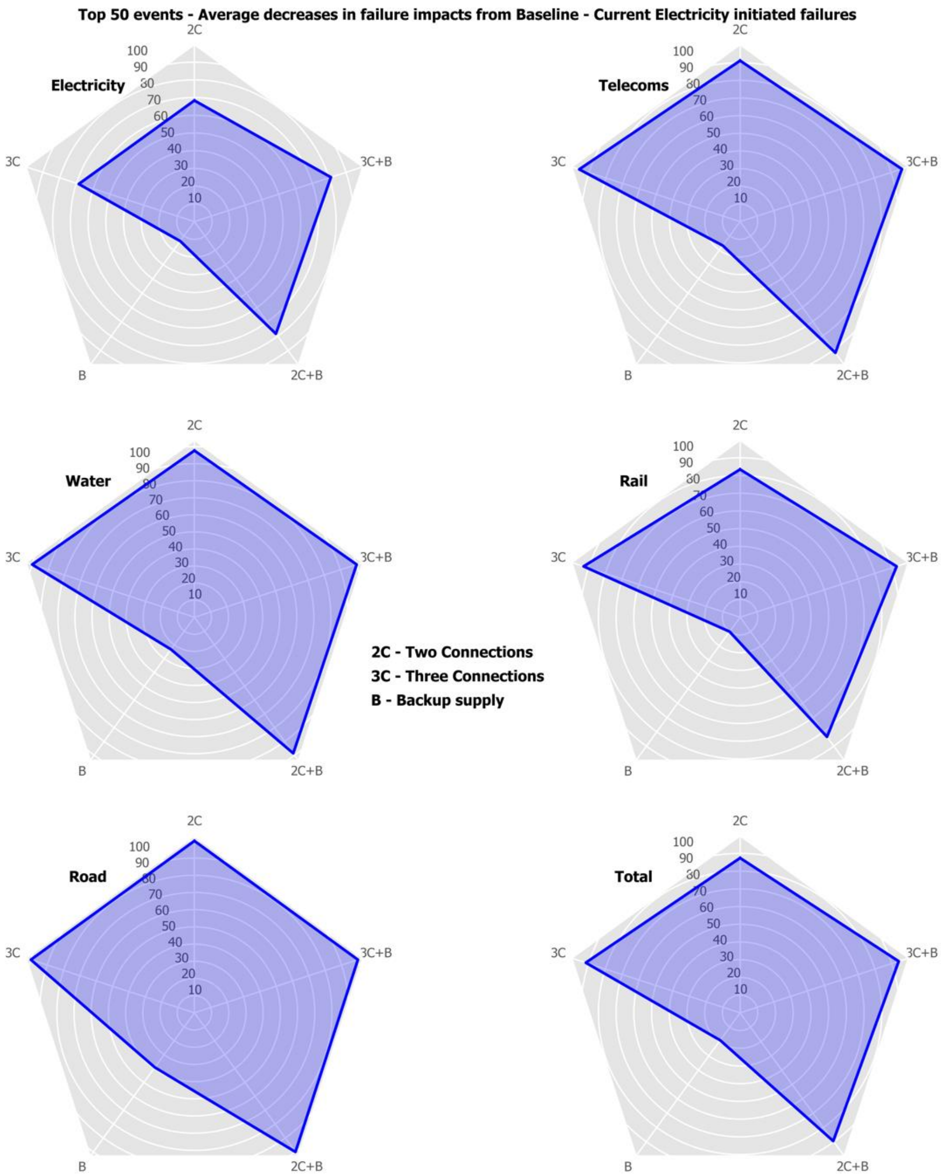


Figure 4-11: Spider plots showing the average percentage decreases in user disruptions for the 50 worst cumulative disruption events for infrastructure networks for different resilience enhancing options in comparison to the baseline option. The failures here are initiated by the electricity networks

4.5 Future networks and failures

4.5.1 Changing network vulnerabilities

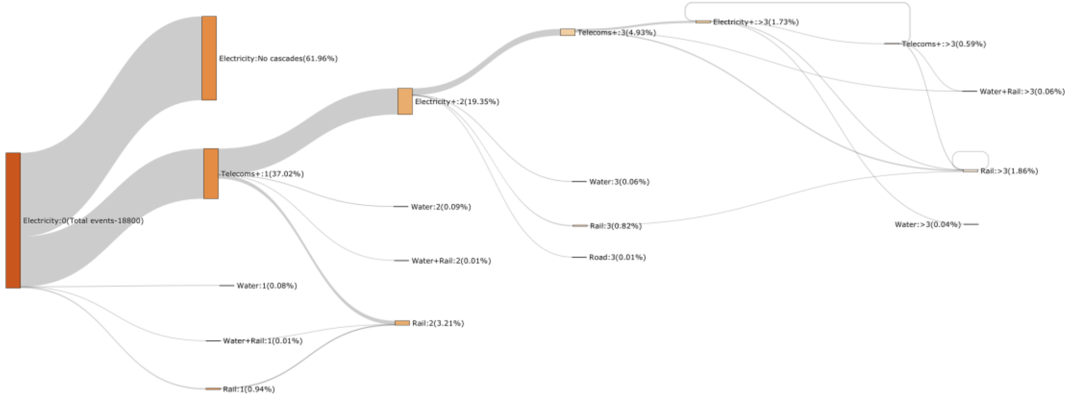
Systemic assessment of the future network failures was done in a similar manner to the current networks, in response to the question:

1. How would the network vulnerabilities change in the future under different planning scenarios?

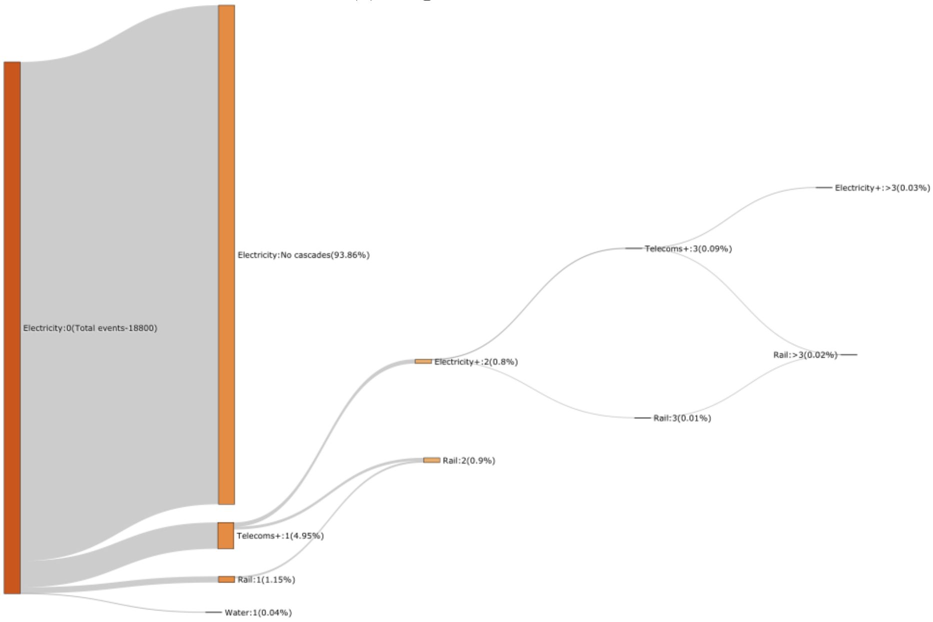
Figure 4-12 shows Sankey diagrams of the chain of cascading events in the future networks state due to failures initiated in the electricity network, by testing all 18,800 individual node failures. From the first result in Figure 4-12(a), with single connections, in comparison to the current network result of Figure 4-2(a) there are about 188 fewer instances of cascading failures in the future networks, which means that some network redundancy has increased by adding new sources. We infer that: (1) The most significant chain of cascading failures is from electricity to telecoms and as further, with about 37% events leading to telecoms and at least one of rail and water disruptions, with further 19% events leading to electricity failures, and 4.9% to another order of telecoms failures; and (2) About 4.2% failure cascades go to Order 4 and above.

In the case where the connections are increased to two we see from Figure 4-12(b) that: (1) Cascading failures are reduced significantly, with about 4.95% events leading to telecoms and at least one of rail and water disruptions, with further 0.8% events leading to electricity failures, and 0.09% to another order of telecoms failures; and (2) About 0.03% failure cascades go to Order 4 and above. In comparison to the current network result of Figure 4-2(b) there are about 88 fewer instances of cascading failures in the future networks.

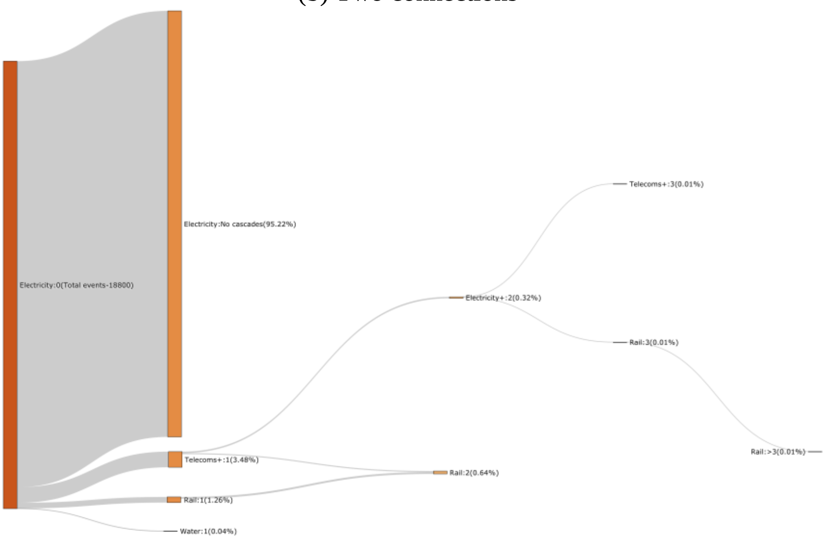
Figure 4-12(c) shows the results when the connections are increased to three the results show that: (1) Cascading failures are again reduced significantly, with about 3.5% events leading to telecoms and at least one of rail and water disruptions, with further 0.32% events leading to electricity failures, and 0.01% to another order of telecoms failures; and (2) Order 4 and above cascading failures are avoided. In comparison to the current network result of Figure 4-2(c) there are about 51 fewer instances of cascading failures in the future networks.



(a) Single connections



(b) Two connections



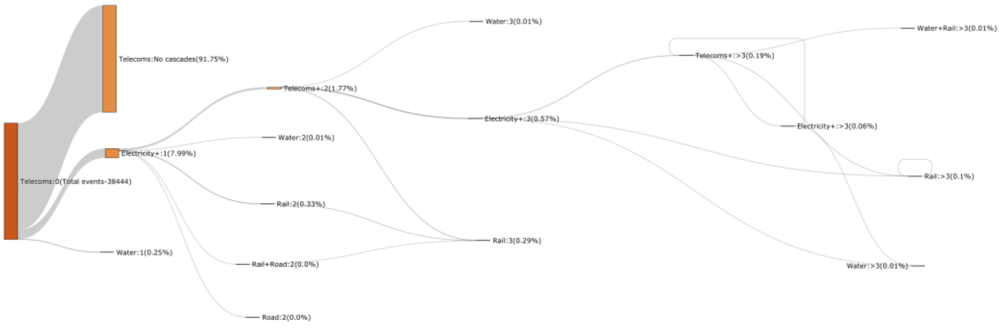
(c) Three connections

Figure 4-12: Failure propagation from electricity to other networks in the future with different degrees of dependencies.

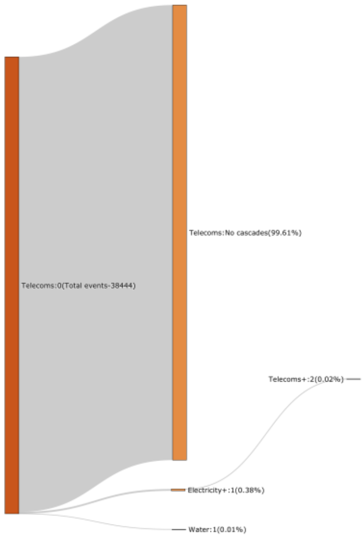
Figure 4-13 shows Sankey diagrams of the chain of cascading events in the future system state due to failures initiated in the telecoms network, by testing all 38,444 individual node failures. From the single connections result in Figure 4-13(a) we infer that: (1) In comparison to electricity, there are fewer cascading failures from telecoms, with about 8% events leading to electricity and at least one of rail and water disruptions, with further 1.8% events leading to another order of telecoms failures; and (2) About 0.28% failure cascades go to order 4 and above. The results are very similar to the current day results of Figure 4-3(a).

In the two connections case for telecoms we see from Figure 4-13(b) that: (1) Cascading failures are almost gone, with about 0.38% events leading to electricity and at least one of rail and water disruptions, with further 0.02% events leading to another order of telecoms failures; and (2) Order 3 and above failures are eliminated. The results are very similar to the current day results of Figure 4-3(b).

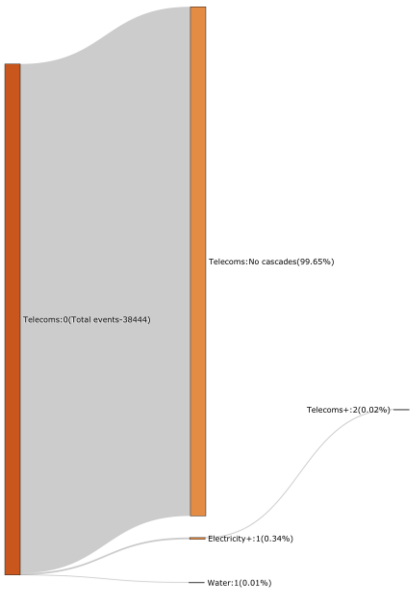
Similarly the three connections case results of Figure 4-13(c) show that: (1) Cascading failures are almost gone, with about 0.3% events leading to electricity and at least one of rail and water disruptions, with further 0.02% events leading to another order of telecoms failures; (2) Order 3 and above failures are eliminated. The results are very similar to the current day results of Figure 4-3(c).



(a) Single connections



(b) Two connections



(c) Three connections

Figure 4-13: Failure propagation from electricity to other networks in the future with different degrees of dependencies.

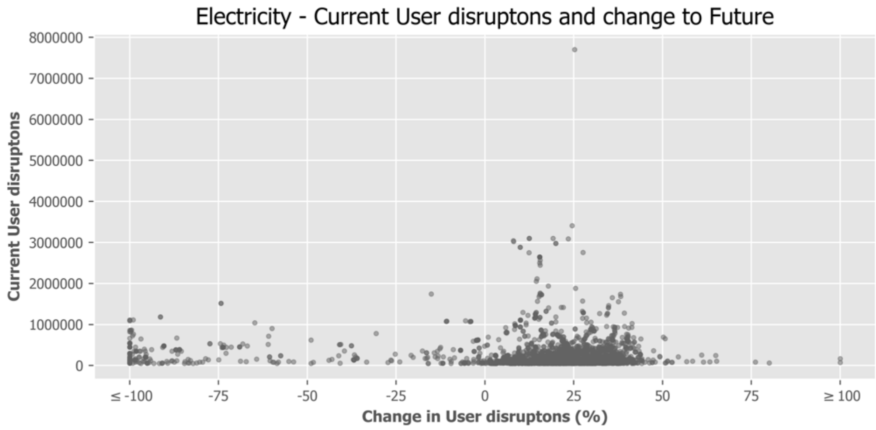
We next look at the changes in failure impacts in the future, in comparison to the current impacts. Figure 4-14 shows the current total user disruptions > 50,000, for the electricity-initiated failures, on the y-axis and the percentage by which these change in the future network configurations on the x-axis. While most failure impacts are expected to increase in the future due to increase in population, there are instances where the failures decrease due to increased network redundancies provided by adding more sources.

Figure 4-14(a) shows the results for the case where one degree of connections was considered. The largest failure event's disruption impact increases by 25%, and similarly most of the highest impact events above 2 million disruption increase by 5%-25% in the future. But there are significant numbers of events clustered around the -100% change values, where the impacts are almost eliminated. These instances are the ones where adding future generation capacities seems to have provided gains in terms of reducing the impacts.

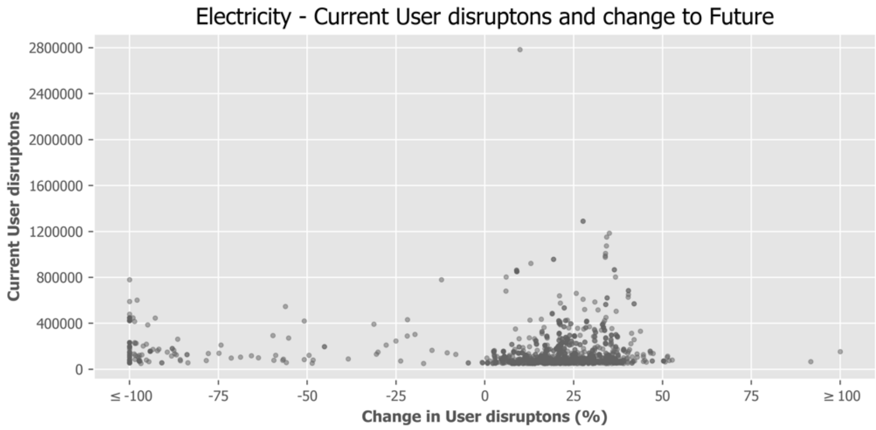
The Figure 4-14(b) case with two degrees of connections also shows that the highest failure event impact increases in the future, though by only about 10%. And the other instances of impacts > 800,000 users also increase in the future by 5%-40%. Here again there are some instances of failures in excess of 400,000 where the future impacts decrease by 100% due to add sources.

The final case with three degrees of connections from Figure 4-14(c) shows that the highest failure event impact increases in the future by about 26%, and most significant failure impacts increase by 5%-45% in the future. There are some instances of failures in excess of 400,000 where the future impacts decrease by 100% due to add sources.

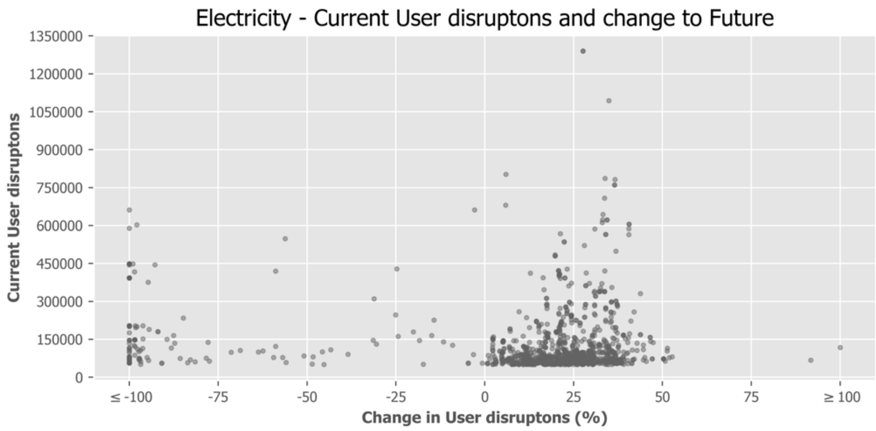
Figure 4-15 shows similar results for the case where the telecoms networks were the initiating network for failures. As we saw in previous results of Figure 4-3, Figure 4-5 and Figure 4-13 that the telecoms network initiated failure propagations in the future do not change much and most cascading failures are eliminated as the degrees of connections are increased from one to two and three. Hence the results of Figure 4-15(a) show that with one degree of connections some instance of failure impacts are reduced by more than 50%, which could be attributed to increased redundancy in the electricity network. However, increasing the degrees of connections to two (Figure 4-15(b)) and three (Figure 4-15(c)) increase impacts because these are all mostly only telecoms impacts that grow due to population increase in the future and with no changes in network topology.



(a) Single connections

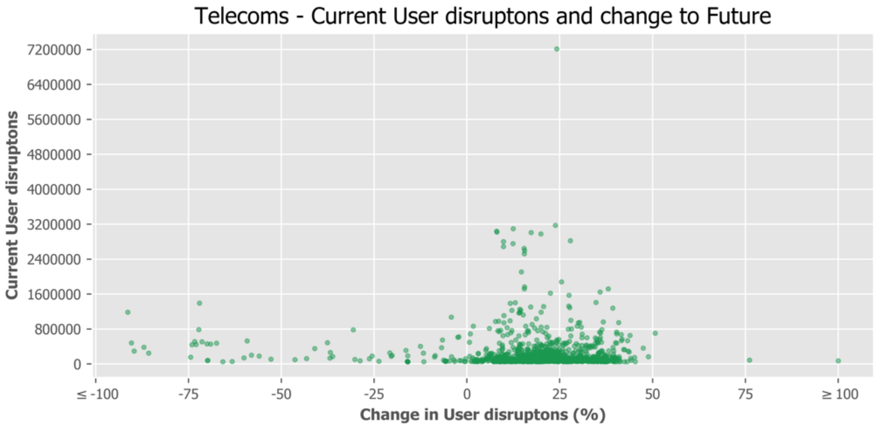


(b) Two connections

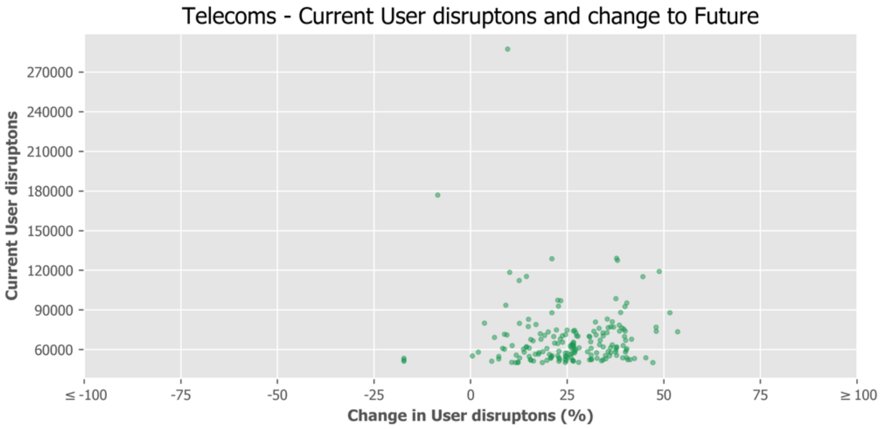


(c) Three connections

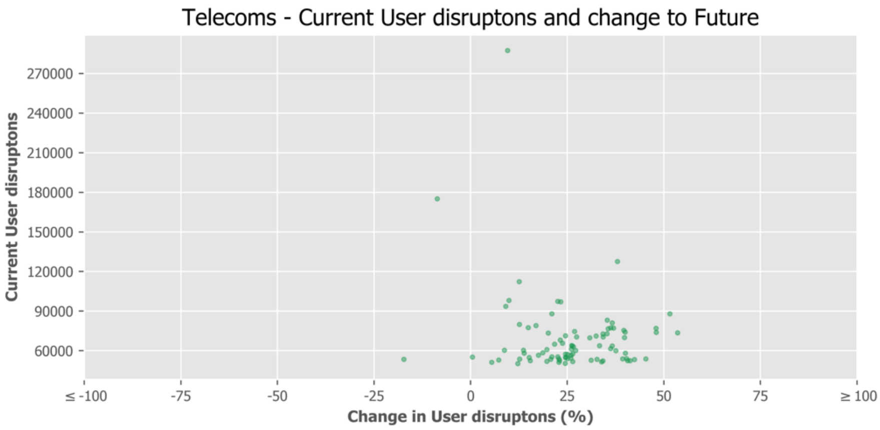
Figure 4-14: Changes in user disruptions in the future networks in comparison to current disruptions, for failures initiated in the electricity network.



(a) Single connections



(b) Two connections

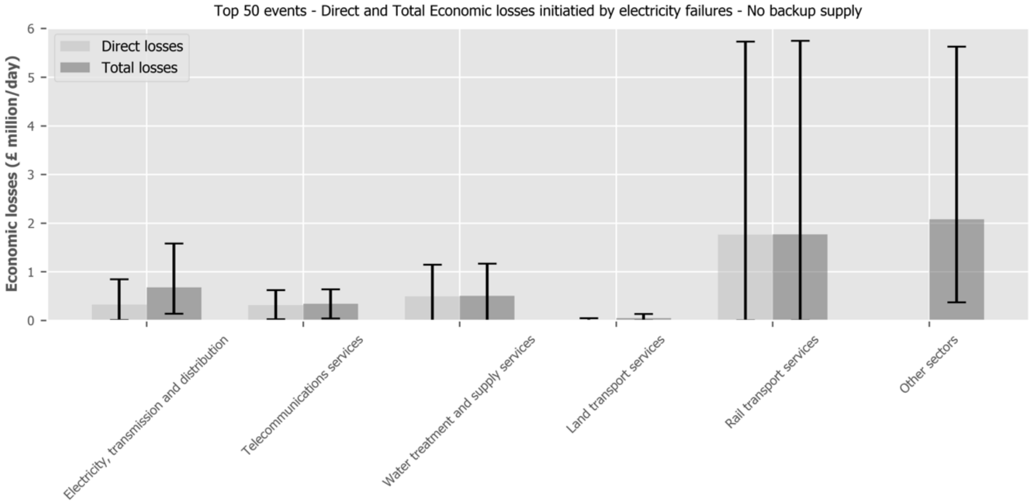


(c) Three connections

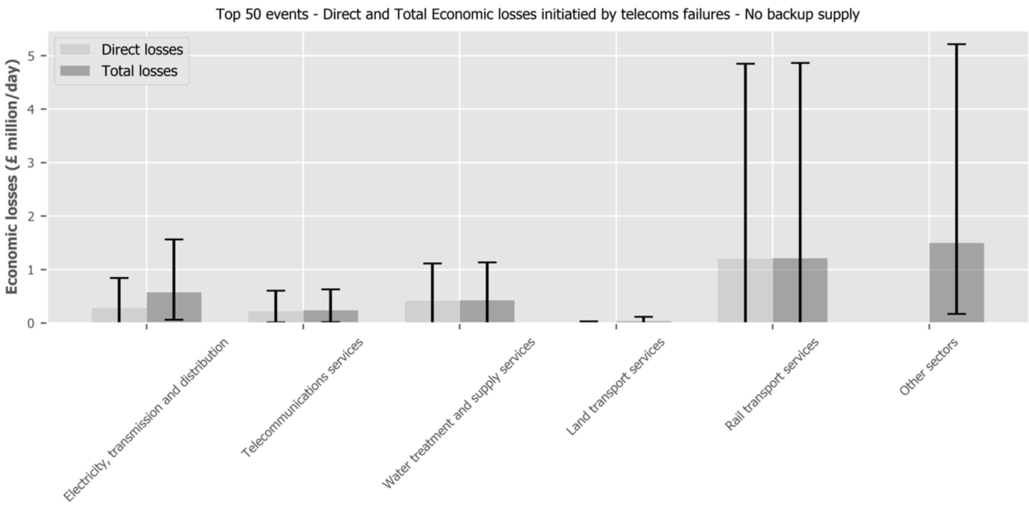
Figure 4-15: Changes in user disruptions in the future networks in comparison to current disruptions, for failures initiated in the telecoms network.

We also estimated the economic losses for the 50 worst-cases of cumulative user disruptions, similar to the analysis presented in Section 4.2.3. The 50 worst-case events in the future had the same initiating failure conditions as the ones in the current, so we get similar cross-sector losses as we saw in Figure 4-6 - Figure 4-8. The differences are seen in the increased losses in the future, accounting for the increased demand disruptions and GDP growth.

Figure 4-16(a)-(b) shows error bar plots with the mean values and 95% confidence intervals for economic losses averaged across all top 50 user disruptions events in the future for failure initiated by the electricity networks and telecoms networks respectively and considering only single connections. The results are similar to the results of Figure 4-6(a) and Figure 4-7(a), with the largest economic losses being recorded in the railways sectors in both instances. In the future, for electricity initiated events (Figure 4-16(a)), the highest economic losses in railways increase to about £5.9 million/day from the current losses of £2.7 million/day. The corresponding increases for the telecoms initiated losses case (Figure 4-16(b)) to about £5.0 million/day from current levels of £2.5 million/day. Overall the cumulative direct economic losses in the future are as high as £7.0 million/day and the total losses are about £13.6 million/day, for both the cases shown in Figure 4-16. Hence. The economic losses in the future increase by a factor of about 1.91 – 2.0 times the losses in the current scenarios, mainly driven by GDP growth as the primary factor and by population growth as the secondary factor. Similar results are seen in the cases with increased connections.



(a)



(b)

Figure 4-16: Mean value with 95% CI estimates of direct and total macroeconomic losses in the future across top 50 user disrupted events initiated by (a) electricity failures; and (b) telecoms failures. Both cases are with single connections.

4.5.2 Exploring options for reducing impacts in the future

Applying the resilience enhancing options, explored in the current scenarios (see Section 4.4), in the future networks shows similar gains averaged over the 50 worst-case use disruption events. Figure 4-17 shows these results for the electricity-initiated failures in the future, which again reinforce the effectiveness of enhancing network redundancy in significantly reducing and in some case eliminating the worst-case disruptive impacts. Here again, the effectiveness of the backup supply is also crucial in delaying and thereby decreasing the disruptions. All these disruptive impacts are in excess of 1 million users/day and 1 £million/day added across all networks and can be as high as 10 million user/day and about 14 £million/day. So, reducing them by 85%-92% in the future with a combination of increased connections and backup supply (2C+B and 3C+B) would be very effective.

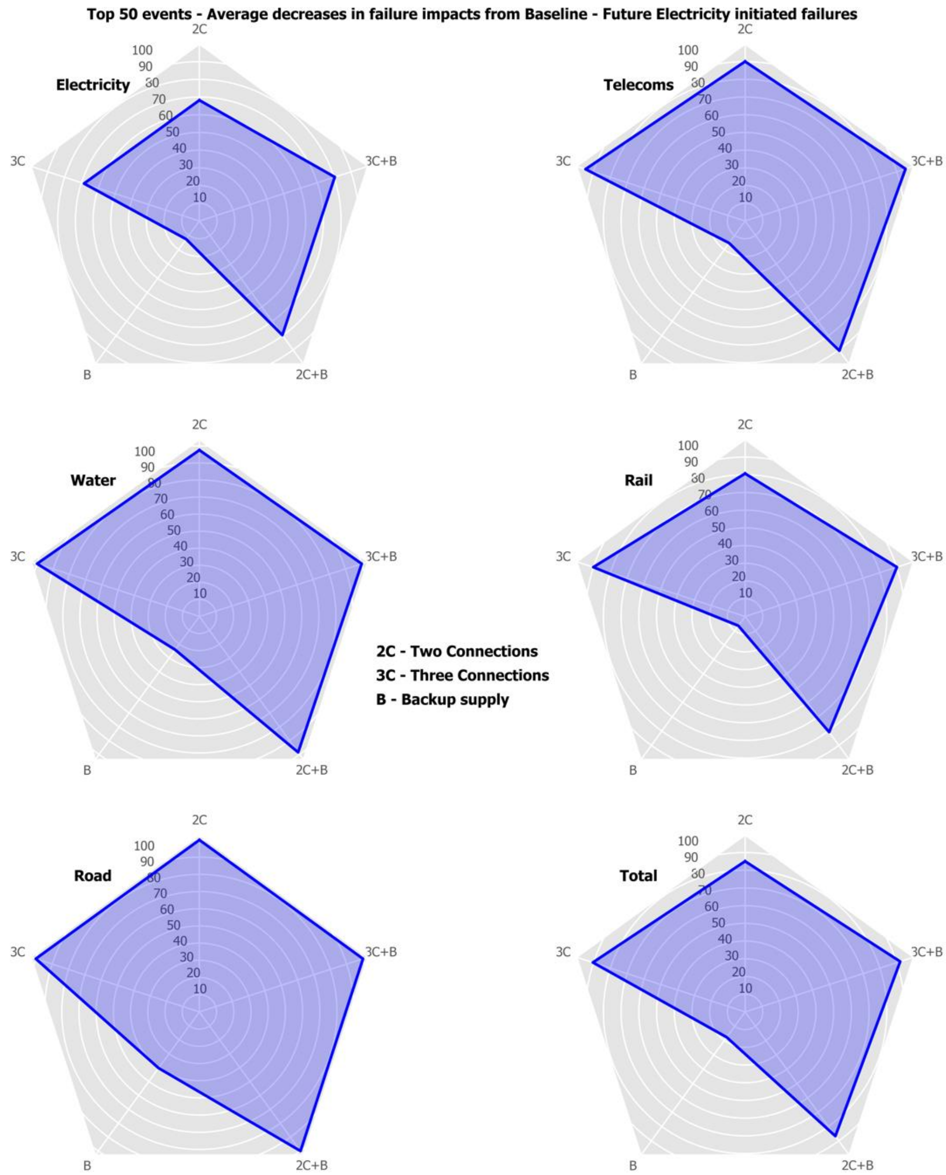
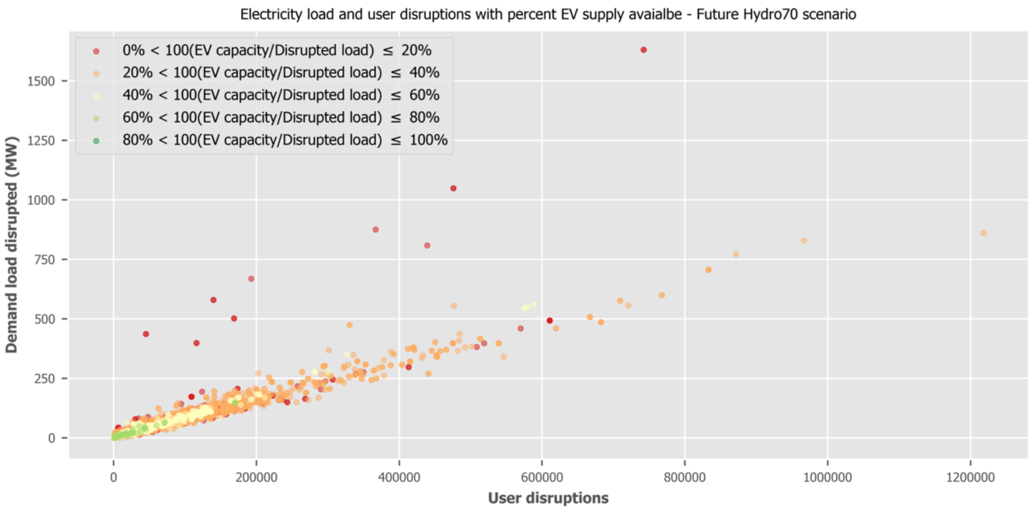


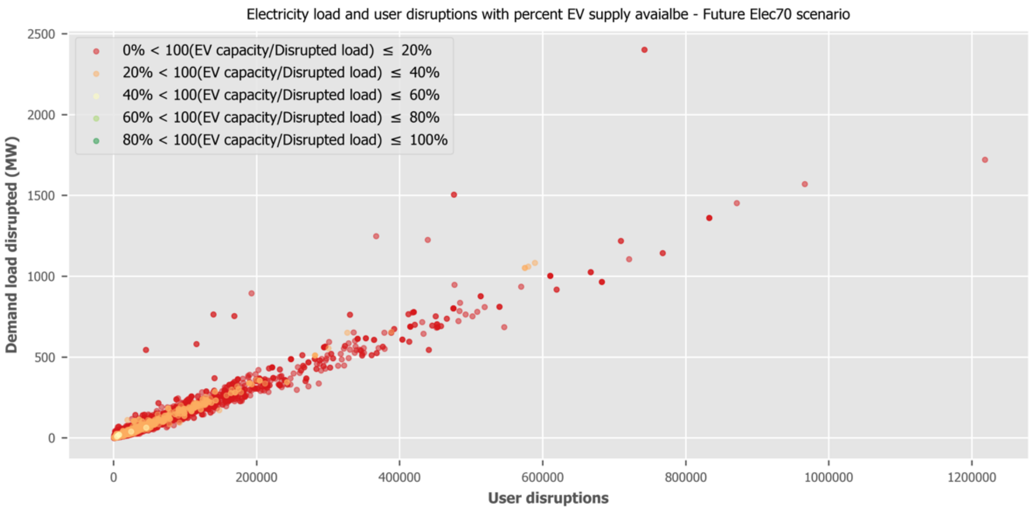
Figure 4-17: Spider plots showing the average percentage decreases in user disruptions in the future for the 50 worst cumulative disruption events for infrastructure networks for different resilience enhancing options in comparison to the baseline option. The failures here are initiated by the electricity networks.

Another possible option for enhancing resilience in the future in the electricity networks is to consider the possibility that Electric vehicles (EV) could be used as a backup supply option for residential consumption, when the grid supply would be disrupted. We explore this option by analysing the total disrupted electricity demand load in MW from each electricity-initiated failure event where there are non-zero disruptions to the network. From the allocation of spatial demands in the electricity network (see Section 3.8) in the future we were able to estimate the EV peak demands on the grid, which we use as a proxy for installed EV capacities at the sink node level, which could be potentially used as a backup supply.

Figure 4-18 shows the scatter plots between the electricity network user disruptions and the demand load disruptions in MW corresponding to the Hydro70 and Elec70 scenarios respectively. Since the assignment of demand loads is based on the geographic spread of building footprint areas, which generally correlate well with population densities, hence the demand disruptions and user disruptions are mostly perfectly correlated but there are a few exceptions in the model result. As expected, the load disruption in the Elec70 scenario are much higher than the Hydro70 scenario because of the increased heating demand in this scenario. For both the future energy scenarios the installed EV capacity is the same, as it comes from the transport sector which has one EV demand in the future. Hence, the effectiveness of the installed EV capacity can be compared between the two scenarios. The Figure 4-18 result show that the installed EV capacity has more potential of being effective as a backup in the Hydro70 future scenario, in comparison of the heat demand intensive Elec70 scenario. For the Elec70 scenario (Figure 4-18(b)) mostly the available EV capacity is only about 0%-20% of what is needed to meet disrupted load MW demand loads, which would probably not be very effective. But for the Hydro70 scenario (Figure 4-18(a)) the available EV capacity is between 20%-40% of the disrupted load for some of the high user disruption events and is even in excess of 60% for instances where user disruptions are as low as 1,300 and as high as 170,000. Generally lower values of user disruptions occur at locations of sparse populations, where the electricity grid connections and accessibility might not be very good. Hence, repairs to restore the electricity supply to such locations might take time, making in worthwhile to explore the EV's as a source of supply to households. We note that in both instances the largest load disruptions do not have enough EV capacity to merit it as a suitable supply backup option.



(a)



(b)

Figure 4-18: Scatter plots showing the disrupted electricity demand load in MW vs the user disruptions with the potential available EV capacity as a percentage of the disrupted electricity demand load corresponding to each failure event in the (a) future Hydro70 scenario; and (b) future Elec70 scenario.

5. CONCLUSIONS OF STUDY AND FURTHER ANALYSIS

The aim of this study was to satisfy the NIC's main requirements[1] to:

1. To pilot an approach to assess the key physical vulnerabilities of the current UK economic infrastructure system.
2. To draw out vulnerabilities that arise from network architecture and how these are likely to change in the future.
3. To inform the development of a framework to identify actions to assess, improve and monitor the resilience of the system.

Through the analysis we have highlighted how interdependencies create disruptions beyond the asset and network where the failure was initiated.

In order to understand how the cascading failures could be controlled we increased the redundancy in connections across networks, which showed that adding two degrees of connections can result in a huge reduction of the cascading failures. Adding a third degree of connections creates further incremental gains, though these depend on the specific asset and network.

We also looked at the role of backup electricity supply in delaying failure impacts and for making a case for prioritising controlled repairs of networks. With an example case we were able to demonstrate that there is a lot of value in fixing disruptions within the first 10-24 hours timeframe when most of the backup supply prevents further failure cascades.

We also looked at future networks during some scenarios of future changes to national infrastructure that were suggested by the NIC. In a scenario in which more supply points were added to the national electricity network there are projected to be some gains in increasing redundancies in networks and reducing failure impacts.

5.1 Strengths and limitations of the analysis

This analysis provides the first national-scale interdependent infrastructure network analysis done in such detail. To our knowledge such extent of data collection and modelling of multiple infrastructure networks and their physical connections has not been done before at a national scale. We have created unique electricity and telecoms network representations with novel data and methods. The water supply network, though high-level is the first detailed representation of the water resource system for England and Wales. Our rail and road networks, built from previous studies, provide a realistic national-scale view of how these systems function. The process of collecting data and modelling connections between the networks is also quite unique and has resulted in novel representations of physically interdependent networks.

This study has also created a first set of representations of future electricity networks, factoring in realistic future network scenarios of increased supply and demand. We have collected best available projections of the location of future network developments and incorporated them into our model.

The failure analysis provides a unique perspective of cascading failures by mapping out the orders in which network disruptions occur and propagate towards other networks. This evidence is very useful for understanding how cascades could be controlled by introducing network redundancy and by adding backup supply options.

Though the study is quite detailed, there are a number of limitations that we acknowledge exist in the current modelling approach. We note that several of these limitations arise due to the limited time and scope of this study, given that it is an initial analysis and focussed on proof of concept. Some of the study limitations we highlight are:

1. We do not have the actual data for the locations of assets and network topology of many systems. In particular for the telecoms asset and networks, we are aware that there are smaller operators that we have not considered and modelled in our study. Similarly, for the water network detailed data on the distribution networks going all the way to households does not exist openly.
2. There is very limited data on network interdependencies, which is mostly assumed in this study.
3. Due to the lack of data within and across networks it is not easy to estimate how much redundancy there is in the systems.
4. The flow assignments on the network has been done in a very simplistic manner, while more dynamic flow assignment models would represent network behaviours more accurately.
5. In the failure analysis we have only tested single points of failures and their resulting impacts. In real-life hazard events multiple network failures are more prone to happen and would provide a more comprehensive picture of failure propagation incidents.

5.2 Future opportunities

In this study we have developed an infrastructure systems resilience model that incorporates interdependent energy, transport, digital and water infrastructure at a national scale. Though there are limitations to the analysis, as listed above, the model development provides a unique capability for exploration of the resilience of national infrastructure systems, so that resilience can be better factored into future NIC work. In this study we have addressed a small number of scenarios of future infrastructure systems, but this model could be used to explore a much wider range of future infrastructure investments and policies that could be considered in the next National Infrastructure Assessment.

There are several opportunities to develop upon the models and analysis built for this study.

1. *Improved data collection* – In order to do a comprehensive national-scale infrastructure network risk and resilience analysis there is a need to collect more data across all sectors. In particular, the quality of analysis would be improved by better data on:
    a. Digital communications networks, including smaller digital providers and connectivity between data processing assets.
    b. Water trunk mains and distribution pipe networks
    c. Interdependencies between infrastructure networks
2. *Analysis of cyber dependencies* – Modern infrastructure is dependent on digital networks for many aspects of system operation and control. Though we have represented some aspects of interdependencies with digital networks, to fully understand the vulnerability of

modern infrastructure networks would require more consideration of how digital technologies are embedded in all other infrastructure, including the implications of software interdependencies as well as hardware networks.

3. *Coverage of missing networks* – The study did not include wastewater, sewage treatment and drainage infrastructure. Nor did it include solid waste processing and recovery assets. These could be incorporated in order to cover the main economic infrastructure sectors considered by the National Infrastructure Commission.
4. *International interdependencies* – UK infrastructure is embedded in global networks. In this study we have considered electricity interconnectors to Europe. There are also significant interdependencies with the rest of the world via shipping, aviation and digital communications. Future developments could consider how UK infrastructure services may be disrupted through interconnections with the rest of the world.
5. *Coverage of supply chains* – The study did not include supply chain disruptions due to infrastructure failures, as they were out of the scope of the study. Supply chain disruptions would significantly affect economic impacts. These could be in considered in future work.
6. *Information sharing* – The main gap in systems research arises due to the lack of information sharing across sectors, which mostly is confined to the high-level of narratives and expert opinions. We are not aware of any instance where asset level information is shared across sectors and factored into their risk and resilience planning. Hence there is a need for some initiative to share data that could be used to provide analytics are the ones developed in this study. Such data could include, among others, location specific information of assets of different networks with connectivity information, the types of services being provided between networks, the demand and capacity limitations of the network interfaces, additional network redundancies and backups in place during disruptions. For continued vulnerability assessments, it is also crucial that such information be updated regularly (at least annually) and changes are made to the information sharing arrangements between assets and networks.
7. *Processed-based network models* – There is a need to develop better processed-based network models at detailed scales, which provide a more dynamic understanding of the progression of failures within and across networks. Such models would also combine performance metrics of service provision with customer disruption and economic losses, which would be more useful for sector long-term and resilience planning.
8. *Analysis of hazards and risks* – The approach taken in this study has been to adopt a 'hazard neutral' approach, which has systematically tested many thousands of scenarios of failure. A complete risk analysis would consider the range of hazards (both natural and mand-made) to which national infrastructure could be exposed, at present and in the future. It would also consider the likelihood of failure of each infrastructure asset that is exposed to a hazard of given severity, i.e. the fragility of each asset. This requires further information and analysis, but full risk analysis provides the basis for prioritisation of investments and other interventions to improve network resilience.
9. *Coping, repair and recovery* – In this study we have examined one approach to enhancing coping capacity during a disaster i.e. the use of back-up storage. There are other strategies that could be adopted to help to avoid disruption and speed up recovery. A more complete analysis of infrastructure network resilience would examine the capacity to restore systems to a functional condition and restart networks.
10. *Empirical validation of failure scenarios* – The failure scenarios that we have tested have been scrutinised by practitioners and domain experts to confirm their realism and validity. More work could be done to collect data on real failures and use that data to validate models of system failure.

11. *Combining long-term planning objectives with resilience planning objectives* – This analysis demonstrated an approach to look at some future planning scenarios for the electricity network, but other networks were not considered. For further analysis planning scenarios for all sectors could be considered and incorporated into the failure estimations. More importantly future analysis might look at how a tool could be used to consider resilience in any long-term planning objectives and make it possible to develop a capability for informed decision making. For example, further analysis could consider how we increase network redundancies in the future and what type of long-term planning would be needed to achieve that.
12. *Harnessing modelling and capabilities for future studies* – This study has created several unique infrastructure network datasets and modelling capabilities that could be useful for the NIC in other studies as well. An initial step of creating a manual documenting the project model codes, written in Python programming language, has been achieved and transferred to the National Infrastructure Commission. The codes and accompanying datasets could next be setup and run on NIC controlled secure computational systems where these important national models will be hosted and can be used for future studies.

APPENDIX A: VULNERABILITY CHARACTERISTICS

A.1 Defining and choosing vulnerability characteristics

In this study we are looked at vulnerability characteristics of networks in response to the two questions below.

1. Can we identify a list of possible characteristics of the UK infrastructure networks that provide indications of the vulnerabilities of the system, as well as its resilience?
2. How do we establish criteria to identify the relative importance of each characteristic in different parts of the system as well as compared to other characteristics?

Though this line of inquiry was limited because we were not able to find any useful insights on the relevance of these characteristics to be able to inform us about network vulnerabilities and their significance in informing us about improving resilience. Further investigation is needed on this topic.

The *characteristics of the UK infrastructure networks that provide indications of the vulnerabilities of the system* are therefore understood in the context of the above types of interdependencies. A *vulnerability characteristic* denotes a metric that can explain the strengths or weaknesses of network interdependencies in influencing the failure propagation and resulting vulnerabilities across networks.

Table A-1 shows the list of network characteristics that have been reviewed and selected to be relevant for this study.

Table A-1: Long list of vulnerability characteristics and their vulnerability implications.

Network metric/characteristic name	**Meaning**	**Implications on vulnerability**	**Infrastructure examples drawn from literature**
1. Degree centrality	Number of linkages that a node or edge has.	Provides information on which nodes/edges could physically knock out most of their surrounding network.	Most well-known network graphs studied include: (1) Scale-free: With node degree centrality following a power law, and are robust to random failures but not targeted; (2) Random (Erdos-Reyni): With binomial node degree centrality, and are robust to targeted failures but not random[82,83].
2. Clustering coefficient	Degree to which connected node triplets of networks cluster together.	Provides information on which groups on nodes would knock out each other.	Barrett et al (2004)[84] - Show that electricity networks have low degree distributions, low clustering coefficients, medium diameters, and so are very less robust. Also, show that wireless ad hoc

82 Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, *67*(2), 026126.
83 Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, *45*(2), 167-256.
84 Barrett, C., Eubank, S., Kumar, V. A., & Marathe, M. V. (2004). Understanding large scale social and infrastructure networks: a simulation based approach. SIAM news, 37(4), 1-5.

<table>
<tr><td></td><td></td><td></td><td>networks have medium degree, high clustering, medium diameter, and so are more robust.</td></tr>
<tr><td>3. Closeness centrality and Diameter</td><td>Average length of the shortest path from a node and all other nodes in the graph. Thus, the more central a node is, the closer it is to all other nodes. Maximum shortest path is called the diameter.</td><td>Provides information on which nodes/edges could most quickly knock out flows.</td><td>Daqing et al (2011)[85] - Have linked this to the node degree distributions, the probabilities of traversing a certain distance on the network, and the distributions of the number of network clusters due to percolation.</td></tr>
<tr><td>4. Betweenness (path) centrality</td><td>The number of times a node/edge acts as a bridge along the shortest path between two other nodes.</td><td>Tells us about the how nodes/edges being knocked out could affect network flows</td><td rowspan="2">Robson et al (2015)[86] - The authors have demonstrated that real infrastructure networks are close to hierarchical networks as they are scale free but also have significant hubs with large connections. The ramifications of this on failures are then analysed by looking at the distributions of numbers of subgraphs as nodes are removed randomly or by selecting based on degree centrality or betweenness.</td></tr>
<tr><td>5. Assortativity</td><td>The likelihood of nodes with similar properties to be connected, e.g. similar degree. Mainly the correlation coefficients of degrees between pair of links nodes.</td><td>Provides information about the connectivity within and between networks. Quick way to infer if two networks are connected at important hubs.</td></tr>
<tr><td>6. Eigenvector centrality</td><td>Measure of how well connected a node is to other well-connected nodes in the network.</td><td>Quick way to accessing the relative contribution of nodes in influencing and spreading failures. High eigen score means a node is connected to other nodes with high connectivity as well. So knocking off high eigen score nodes could knock out other high eigen score nodes as well.</td><td rowspan="5">Rueda et al (2017)[87] - Compared robustness of 15 telecommunications networks for several centrality metrics.</td></tr>
<tr><td>7. Percolation centrality</td><td>Defined for a given node, at a given state, as the proportion of shortest paths between a pair of nodes, where the source node is percolated (e.g., disrupted).</td><td>Tells us about the how source nodes being knocked out could affect network flows.</td></tr>
<tr><td>8. Cross-clique centrality</td><td>Determines the connectivity of a node to different completely connected subgraphs (called cliques).</td><td>Tells us if a node from one network can knock out all nodes in another. A node with high cross-clique connectivity facilitates the disruption of all nodes in the clique.</td></tr>
<tr><td>9. Heterogeneity</td><td>Coefficient of variance in nodal degree (node centrality).</td><td>Tells us if the overall network structure might be well connected or have some significant hubs.</td></tr>
<tr><td>10. Trophic coherence</td><td>Describes how neatly the nodes fall into distinct levels in a</td><td>Tells us how different network hierarchies are organised, which could be</td></tr>
</table>

85 Daqing, L., Kosmidis, K., Bunde, A., & Havlin, S. (2011). Dimension of spatially embedded networks. Nature Physics, 7(6), 481.
86 Robson, C., Barr, S., James, P., & Ford, A. (2015). Resilience of hierarchical critical infrastructure networks. UCL STEaPP.
87 Rueda, D. F., Calle, E., & Marzo, J. L. (2017). Robustness comparison of 15 real telecommunication networks: Structural and centrality measurements. Journal of Network and Systems Management, 25(2), 269-289

	directed network, in terms of their degrees.	useful for understanding failures at different levels.	
11. Motif concentration	Describes the chances of occurrence for a specified network motif - repeated small components within the network.	Provides information on local robustness of network in inferring global robustness. If a locally robust pattern is repeating a lot on the network, then it can be inferred to be robust.	
12. Algebraic connectivity	The second smallest eigenvalue of the Laplacian matrix (i.e. degree matrix minus adjacency matrix) of the graph.	Larger values of algebraic connectivity represent higher robustness against efforts to decouple parts of the network, indicating network robustness and well-connectedness.	
13. Spectral gap	Defined as the difference between the first and second eigenvalues of the adjacency matrix of the graph.	A sufficiently large value of spectral gap is regarded as a necessary condition for the so-called "good expansion" properties, the existence of which, indicates higher structural robustness against node and link failures.	
14. Central point dominance	The mean over the betweenness centrality values of all nodes indexed by the maximum value of betweenness (achieved at the most central-point).	Describes the variance of betweenness centrality of the network. If the variance is low then the network is connected and robust, and if it is high then the network has one dominant connectivity whose failure can make is less robust.	
15. Spectral clustering	Describes clustering of the network from the aspect of graph partition.	Through the identification of a partition of the graph such that the edges between different groups have a very low weight and the edges within a group have high weight, provide information on minimum effort required to cut the network into communities.	
16. Core-periphery	Describes a group of central and densely connected nodes and sparsely connected periphery nodes which governs the overall behaviour of a network.	Shows which nodes are most connected to groups of lesser connected nodes in the network. Knocking out such well-connected nodes will knock out most of the network functionality.	Rombach et al (2014)[88] - Studied the London Tube network of 317 nodes (one for each station) and weighted edges that represent the number of direct, contiguous connections between two stations. They suggest that the London Tube has a core group of (about) 60 stations and a periphery of 257 stations.

88 Rombach, M. P., Porter, M. A., Fowler, J. H., & Mucha, P. J. (2014). Core-periphery structure in networks. SIAM Journal on Applied mathematics, 74(1), 167-190

17. Hotspot centrality	z-scores of network nodes and edges in terms of their spatial clustering within gridded lattices.	Lattices with highest z-scores will show the highest impacts on network vulnerability	Thacker et al. (2018)[13] - Showed hotspot centrality of UK infrastructure creates critical clusters of infrastructures with large customer impacts around big urban centres.
18. MR(D)	In an interdependent network, metric MR(D) denotes the minimum number of node removals from network A which causes the failure of D arbitrary nodes in network B.	If MR(D) is low and D is high then if means network B is highly dependent on network A.	Buldyrev et al (2010)[89] - Application on known degree distribution networks, and demonstration of Italy power-grid failure effect on Internet network. Parandehgheibi & Modiano (2016)[90] - Did a more theoretical presentation of the metrics.
19. MRB(D)	In an interdependent network, metric MRB(D) denotes the minimum number of node removals from both networks which causes the failure of D arbitrary nodes in network B.	If MRB(D) is low and D is high then if means network B is highly dependent on network A and is not very robust itself.	
20. Source-sink centrality - Connectivity loss	Describes the minimum number of sources in the network that are necessary to serve each demand location (sink).	Provides information on the number of sources that you can knock out whilst ensuring that each sink is still connected to a source.	Dueñas-Osorio & Vemuru (2009)[91] - Proposed these metrics for studying cascading failures in electricity networks
21. Cascading susceptibility	Difference between source-sink connectivity loss after considering network cascades with connectivity loss by triggering event	Shows how much cascading effects impact network performance.	

Table A-2 shows the short list of network characteristics, derived from the long-list of metrics proposed in the Inception report, that have been reviewed and selected to be relevant for this study.

The rationale for selecting these metrics was that

1. They represent centrality measures at the asset level, which is more useful for this analysis.
2. There are tested network functions in Python language that we could build and test for these metrics.

From the long-list the following metrics are not included because:

1. Assortativity, Heterogeneity, Motif concentration, Algebraic connectivity, Spectral gap, Central point dominance, Spectral clustering – These are all global network metrics, which give 1 value for a graph. So, they do not apply at the individual nodes or edge level, which is more relevant to the study.
2. Percolation centrality, Hotspot centrality, MR(D), MRB(D), Source-sink centrality - Connectivity loss, Cascading susceptibility – These are impact estimation metrics rather than network topology metrics from which we want to infer the results. So, they are more useful in understanding the results, and are captured in the failure analysis.

89 Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. Nature, 464(7291), 1025.
90 Parandehgheibi, M., & Modiano, E. (2016). Robustness of bidirectional interdependent networks: Analysis and design. arXiv preprint arXiv:1605.01262.
91 Dueñas-Osorio, L., & Vemuru, S. M. (2009). Cascading failures in complex infrastructure systems. Structural safety, 31(2), 157-167

Table A-2: Short list of vulnerability characteristics and their vulnerability implications.

Network metric/characteristic name
1. Degree centrality
2. Clustering coefficient
3. Closeness centrality
4. Betweenness centrality
5. Eigenvector centrality
6. Cross-clique centrality
7. Trophic coherence
8. Path centrality
9. Core-periphery (core number)

A.2 Distributions of characteristics and correlations with failure impacts

To understand the rationale and meaning of the network metrics finalised for this study, we looked at their distributions and correlations with single point failure impacts of individual networks. The aim of this analysis was to answer the following questions:

1. What does each network metric mean in a generalised network graph?
2. What does each metric signify specifically in the GB sector networks built for this study?
3. How much are these network metrics correlated with disruptions estimated from individual asset failures?

We were interested in figuring out whether we can infer anything about assets that are ‘important’ and those which are ‘unimportant’

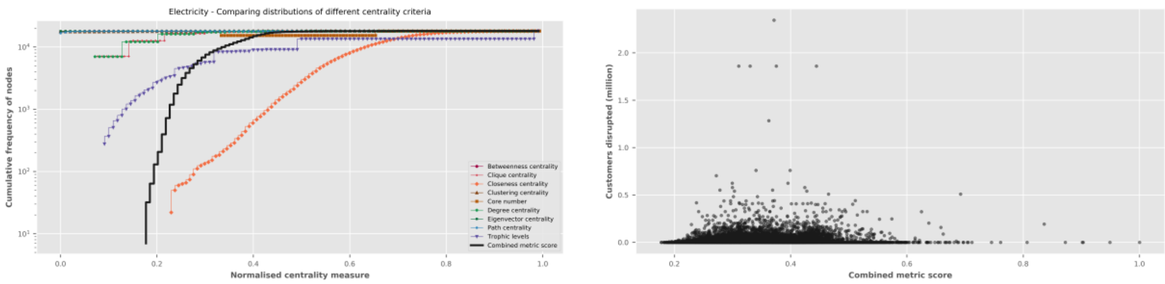
In each of the distribution results presented below we normalised each network metric score of a scale from 0-1, and also created a combined metric score by adding the normalised scores of all metrics giving them equal weightage. Below we discuss the results for the centrality metrics of only the electricity and telecoms networks, because we tested these two networks for failure analysis.

Electricity networks

Figure A-1(a) shows the distributions of the different network metrics, which are explained in Table A-. In Figure A-1(b) we compare how the combined metric score correlates with the failure impacts of nodes, considering only impacts on the electricity network. The reason for this comparison was to understand whether the most central nodes also caused the highest impacts in the network. We see that for the electricity network the most central nodes do not have the highest impacts, and in fact the highest impact nodes have low centrality measures. This is because the most central nodes in the electricity network are located at the transmission levels, where the nodes are all very connected and 1 node failure do not have any impacts due to the N-1 design reliability of the network. At the lower distribution levels (HV and LV) the nodes are not that central as the networks are not that well connected, resulting in single points disruptions that can lead to significant impacts.

Table A-3: Description of the electricity network metrics and their explanations.

Network metric/characteristic name	Explanation of distribution
Degree centrality	A few discreet integer values mostly dominated with transmission level substations
Clustering coefficient	Most of the network has low values = 0. But few clusters at the transmission level can be identified
Closeness centrality	Due to a well-defined network structure values are well distributed, with transmission level nodes having highest values
Betweenness centrality	Most of the network has values = 0. But few nodes at the transmission level have high values as most shortest-paths pass through them
Eigenvector centrality	Most of the network has values close to 0. But few nodes at the transmission level have high values
Cross-clique centrality	Similar behaviour as node degree centrality, but rankings might not be the same
1/Trophic coherence	All sources have a trophic level = 1 and sinks have the lowest values of around 0.1
Path centrality	Values very similar to betweenness centrality
Core number	Values are either =1 (for most nodes), =2 (when HV connects to LV), = 3 (when transmission nodes connect to HV and LV)



(a) (b)

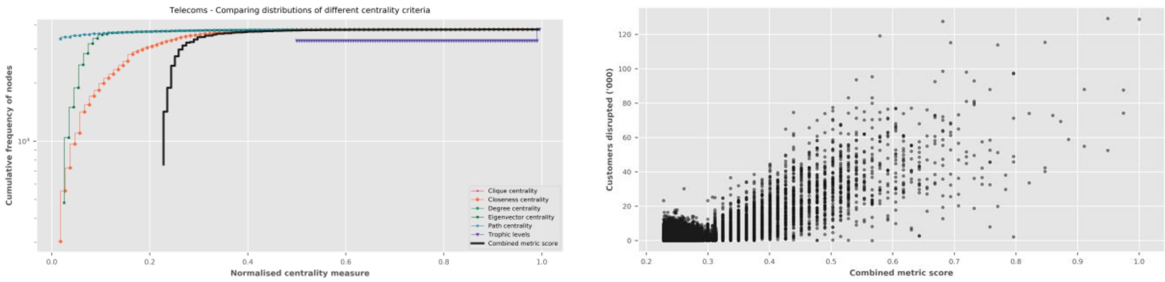
Figure A-1: Electricity network plots showing: (a) the distributions of the different network metrics; and (b) the correlation of the weighted metric score and the user disruptions on the network when nodes are failed individually.

Telecoms

From the distribution of the metrics and their correlations with the impacts shown in Figure A-2 we can see that in the telecoms network the exchanges are the most central nodes and also have the highest failure impacts. Given the radial structure of the network between the exchanges and the macro cells, this outcome is expected. Hence, for the telecoms network the metrics are useful in identifying the most central and high impact nodes simultaneously.

Table A-4: Description of the telecoms network metrics and their explanations.

Network metric/characteristic name	Explanation of distribution
Degree centrality	Ranked by exchanges with most connected macro cells and most customers
Clustering coefficient	All = 0 – As there are no 3 nodes connected to each other
Closeness centrality	Same values as the most degree central nodes, but differences in the lower values.
Betweenness centrality	All = 0 – Because no shortest path between two nodes passes through a third
Eigenvector centrality	Same ranking order as node centrality
Cross-clique centrality	Same values and ranking as node centrality
1/Trophic coherence	All exchanges have a trophic level = 1 and all macro cells have a trophic level = 0.5. Which mainly means that the exchanges are sources and macro cells are connected to 1 source each
Path centrality	Same values and ranking as node centrality
Core number	All = 1 – Because each set of nodes are with 1 core cluster
Combined metric	Same ranking order as node centrality and other metrics that agree with it



(a) (b)

Figure A-2: Telecoms network plots showing: (a) the distributions of the different network metrics; and (b) the correlation of the weighted metric score and the user disruptions on the network when nodes are failed individually.

APPENDIX B

Table B-1: 129 sector IO tables for the UK economy with sector specific multiplier effects.

Sector code	Sector name	Multiplier effect
1	Products of agriculture, hunting and related services	1.82
2	Products of forestry, logging and related services	1.92
3	Fish and other fishing products; aquaculture products; support services to fishing	1.95
5	Coal and lignite	2.01
06&07	Extraction of Crude Petroleum And Natural Gas & Mining Of Metal Ores	1.59
8	Other mining and quarrying products	1.69
9	Mining support services	1.56
10.1	Preserved meat and meat products	2.40
10.2-3	Processed and preserved fish, crustaceans, molluscs, fruit and vegetables	2.04
10.4	Vegetable and animal oils and fats	1.75
10.5	Dairy products	2.30
10.6	Grain mill products, starches and starch products	2.11
10.7	Bakery and farinaceous products	1.93
10.8	Other food products	1.87
10.9	Prepared animal feeds	2.09
11.01-6 and 12	Alcoholic beverages & Tobacco products	1.79
11.07	Soft drinks	2.17
13	Textiles	1.39
14	Wearing apparel	1.56
15	Leather and related products	1.57
16	Wood and of products of wood and cork, except furniture; articles of straw and plaiting materials	1.69
17	Paper and paper products	1.57
18	Printing and recording services	1.71
19	Coke and refined petroleum products	1.35
20A	Industrial gases, inorganics and fertilisers (all inorganic chemicals) - 20.11/13/15	1.64
20B	Petrochemicals - 20.14/16/17/60	1.72
20C	Dyestuffs, agro-chemicals - 20.12/20	1.74
20.3	Paints, varnishes and similar coatings, printing ink and mastics	1.52
20.4	Soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations	1.73
20.5	Other chemical products	1.53
21	Basic pharmaceutical products and pharmaceutical preparations	1.34
22	Rubber and plastic products	1.49
23OTHER	Glass, refractory, clay, other porcelain and ceramic, stone and abrasive products - 23.1-4/7-9	1.78
23.5-6	Cement, lime, plaster and articles of concrete, cement and plaster	1.98
24.1-3	Basic iron and steel	1.75
24.4-5	Other basic metals and casting	1.51
25OTHER	Fabricated metal products, excl. machinery and equipment and weapons & ammunition - 25.1-3/25.5-9	1.54
25.4	Weapons and ammunition	1.45
26	Computer, electronic and optical products	1.53
27	Electrical equipment	1.56
28	Machinery and equipment n.e.c.	1.65
29	Motor vehicles, trailers and semi-trailers	1.61
30.1	Ships and boats	1.87
30.3	Air and spacecraft and related machinery	1.69
30OTHER	Other transport equipment - 30.2/4/9	1.68
31	Furniture	1.63
32	Other manufactured goods	1.53
33.15	Repair and maintenance of ships and boats	1.85
33.16	Repair and maintenance of aircraft and spacecraft	1.87
33OTHER	Rest of repair; Installation - 33.11-14/17/19/20	1.64
35.1	Electricity, transmission and distribution	2.36
35.2-3	Gas; distribution of gaseous fuels through mains; steam and air conditioning supply	2.10
36	Natural water; water treatment and supply services	1.53

37	Sewerage services; sewage sludge	1.58
38	Waste collection, treatment and disposal services; materials recovery services	1.82
39	Remediation services and other waste management services	1.55
41-43	Construction	1.92
45	Wholesale and retail trade and repair services of motor vehicles and motorcycles	1.57
46	Wholesale trade services, except of motor vehicles and motorcycles	1.76
47	Retail trade services, except of motor vehicles and motorcycles	1.63
49.1-2	Rail transport services	1.95
49.3-5	Land transport services and transport services via pipelines, excluding rail transport	1.64
50	Water transport services	1.88
51	Air transport services	1.51
52	Warehousing and support services for transportation	1.95
53	Postal and courier services	1.56
55	Accommodation services	1.58
56	Food and beverage serving services	1.59
58	Publishing services	1.66
59-60	Motion Picture, Video & TV Programme Production, Sound Recording & Music Publishing Activities & Programming And Broadcasting Activities	1.57
61	Telecommunications services	1.41
62	Computer programming, consultancy and related services	1.44
63	Information services	1.45
64	Financial services, except insurance and pension funding	1.56
65	Insurance, reinsurance and pension funding services, except compulsory social security	1.90
66	Services auxiliary to financial services and insurance services	1.47
68.1-2	Real estate services, excluding on a fee or contract basis and imputed rent	1.58
68.2IMP	Owner-Occupiers' Housing Services	1.23
68.3	Real estate services on a fee or contract basis	1.37
69.1	Legal services	1.38
69.2	Accounting, bookkeeping and auditing services; tax consulting services	1.27
70	Services of head offices; management consulting services	1.53
71	Architectural and engineering services; technical testing and analysis services	1.60
72	Scientific research and development services	1.49
73	Advertising and market research services	1.55
74	Other professional, scientific and technical services	1.52
75	Veterinary services	1.31
77	Rental and leasing services	1.56
78	Employment services	1.71
79	Travel agency, tour operator and other reservation services and related services	1.54
80	Security and investigation services	1.50
81	Services to buildings and landscape	1.68
82	Office administrative, office support and other business support services	1.46
84	Public administration and defence services; compulsory social security services	1.82
85	Education services	1.20
86	Human health services	1.16
87-88	Residential Care & Social Work Activities	1.44
90	Creative, arts and entertainment services	1.41
91	Libraries, archives, museums and other cultural services	1.50
92	Gambling and betting services	1.33
93	Sports services and amusement and recreation services	1.64
94	Services furnished by membership organisations	1.24
95	Repair services of computers and personal and household goods	1.48
96	Other personal services	1.29
97	Services of households as employers of domestic personnel	1.00
38g	Waste collection, treatment and disposal services; materials recovery services non-market	1.89
49.3-5g	Land transport services and transport services via pipelines, excluding rail transport non-market	2.16
52g	Warehousing and support services for transportation non-market	1.65
59-60g	Motion Picture, Video & TV Programme Production, Sound Recording & Music Publishing Activities & Programming And Broadcasting Activities non-market	1.54
84g	Public administration and defence services; compulsory social security services non-market	1.48
85g	Education services non-market	1.36
86g	Human health services non-market	1.37
87-88g	Residential Care & Social Work Activities non-market	1.75
90g	Creative, arts and entertainment services non-market	1.73

91g	Libraries, archives, museums and other cultural services non-market	1.50
93g	Sports services and amusement and recreation services non-market	1.73
64n	Financial Services NPISH	1.00
68.1-2n	Real Estate services NPISH	1.92
69.1n	Legal services NPISH	1.04
72n	Scientific research and development services NPISH	1.57
75n	Veterinary services NPISH	1.90
81n	Services to buildings and landscape NPISH	2.09
85n	Education services NPISH	1.27
86n	Human health services NPISH	1.53
87-88n	Residential Care & Social Work Activities NPISH	1.39
90n	Creative, arts and entertainment services NPISH	1.77
91n	Libraries, archives, museums and other cultural services NPISH	1.51
93n	Sports services and amusement and recreation services NPISH	1.83
94n	Services furnished by membership organisations NPISH	1.50

APPENDIX C

Table C-1: Summarised list of assumptions made in this study and their rationale.

Assumptions	Rationale	Limitations/Uncertainty created	Part of code architecture
Methodology			
Nodes were considered to have failed only when they lost all their service. Partial failure states, where nodes might still be operating at below 100% operational levels and providing reduced service were not considered.	The assumption of total loss of service was considered appropriate because we were interested understanding worst-case scenarios of large-scale widespread disruptions.	In reality network nodes might functional at reduced service levels, which might show reduced failure impacts than what are estimated in this study.	Built-in function in the failure analysis code.
For utility networks of electricity, water supply and telecoms nodes service disruption impacts were estimated only for failed nodes. For transport networks we assumed that failures were initiated in a way similar to the utility networks with nodes completely losing their ability to provide service, and we also accounted for disruptions to nodes that lost part of their pre-disruption journeys due to network failure propagation.	For utility networks, as long as there is access to network flows, the service would continue. For transport networks the service is mobility of people, which will be reduced if some flow paths cannot be accessed.	In reality for all networks partial flows along paths with reduced service levels would happen. Due to data and time limitations and no dynamic flow modelling done in this study we were not able to represent such effects.	
Cross-sector dependencies inferred by connecting nodes of dependent network to the geographically closest nodes of the supplying network.	Nearest connection represents the path of least resistance of service flows and is also most cost effective in terms of materials and design of systems.	Lack of any data on how different network assets are actually connected. Difficult to verify across whole country. Is a major source of uncertainty because cascading failures depend on how the cross-sector nodes are connected.	Built-in function in code to join two selected nodes by straight line geometry.
Static representations of flows between source and sink nodes by mapping all shortest distance paths based on network algorithms or known travel routes.	Building dynamic flow representation, which would be a ‘correct’ way, was beyond the time scale of this study, as it is an initial proof of concept exercise. The static flow paths models are a good proxy for showing the relative importance of routes.	In some networks like electricity and water mapping all source-sink flow paths means assigning more redundancies than what might be in reality. While for road and rail only considering known travel routes might under-represent the network redundancies.	Built-in functions in code to estimates flow path allocations for each network.
Only residential customer demands considered for electricity, telecoms and water networks.	No information was available on spatially disaggregated demands from businesses and other non-residential customers	Excluding non-residential/industry demands means we are under-representing the magnitudes of failure impacts in several instances.	Built-in functions in codes for each sector to spatially map census datasets/travel data to service areas of nodes.
Roads and railways demands based on only passenger travel patterns.	No information was available on freight and other commercial travel.		

No network flow rerouting and dynamics considered in the failure analysis.	Building dynamic failure analysis was beyond the time scale of this study, as it is an initial proof of concept exercise.	Rerouting would mean network redundancies have been accounted for properly. At present we might be over accounting for redundancies in the electricity and water networks and under accounting in the transport networks.	Built-in function in the failure analysis.
Economic loss estimations based on a simplified demand-driven Leontief IO model. Losses result from disruptions lasting a day.	Though more complex IO models exist in literature, the Leontief IO model is still the most widely used and is very good in capturing multiplier effects of infrastructure disruptions, which we wanted to represent.	The linear Leontief IO model is an oversimplification of economic productivity. We are not accounting for all demand side disruptions except household losses, and not considering any supply side losses. Neither are we accounting for substitution effects in the economy that would reduce economic impacts. See Section 3.10 for further limitations of the IO model.	Built-in function/code for economic loss analysis.
Increasing redundancies between networks considered as resilience enhancing options.	Due to lack of data we do not know how assets of different network connect with each other and at how many locations. Increasing the connections provides a good sensitivity check for testing the possible ways in which cross-sector network assets might actually connect.	There are large uncertainties in assigning connections properly. So the results will be very sensitive to how redundancies are added and removed.	Built-in function in code to join two selected nodes by straight line geometry.
Backup supply of certain assumed durations considered as a resilience option to absorb and delay initial shock impacts.	Reasonable assumption as many asset owners do keep backup generators in cases of emergency response. Good substitute when we have no information on post-disruption recovery and repairs planning of assets.	Uncertainties are created in the way the backup durations are modelled. See below.	Assumed parameters in the model.
Sector specific data			
Electricity – Only peak annual demand load in MW considered as a single state representation of the network.	The peak load state shows the condition under which the network will be most stressed, which is what we need for failure analysis	Only one realisation of peak demand loads has been considered. Uncertainties in how the peak is estimated would mean that a range of peak loads should be considered in the analysis.	Built in the source data extracted for demand modelling.
Electricity – All possible directed source-sink flow	In agreement with the notion that electricity	Mapping all source-sink flow paths means	Built-in functions in code to

paths mapped. For direction of flow was from transmission network to the high voltage network and then to the low voltage network.	network would work under a N-1 reliability state	assigning more redundancies than what might be in reality. We are not checking whether the source capacity is less than the demand.	estimates flow path allocations.
Telecoms – Only BT exchange network represented based on open data and a model understanding of how the core network nodes should be connected.	No data was available on other telecoms providers	Considering only one provider would mean we cannot account for telecoms redundancies. We are allocating all customers to only one provider here.	Built in the source data extracted for demand modelling.
Telecoms – Mobile network represented as macro cells connected to telecoms exchanges in a radial network structure.	No data was available on actual connectivity between mobile and fixed network, but expert opinion suggests it should be radial.	Underlying asset data is quite old and has not been updated for a while.	Built in the source data extracted for network modelling.
Telecoms – Failures to exchange network only occurred if the whole inner core network failed at once.	This is consistent with the evidence that telecoms core network is a very resilient network and has a lot of redundancies.	This seems to be a reasonable assumption.	Built-in functions in code to estimate telecoms failures.
Water supply – Represented as a high-level public supply network useful for modelling water transfers between water resource zones.	No data was available on a detailed water network.	Due to a very high level and sparse network representation failure analysis will show very high impacts. Which might provide an unrealistic picture that the water network is not very resilient.	Built in the source data extracted for network modelling.
Water supply – All possible directed source-sink flow paths mapped.	Same principle as applied to the electricity network.	Mapping all source-sink flow paths means assigning more redundancies than what might be in reality. We are not checking whether the source capacity is less than the demand load.	Built-in functions in code to estimates flow path allocations.
Rail – Single track representations of geospatial routes on the national railway network.	No data available on multiple tracks.	Flow paths route choices will be limited.	Built in the source data extracted for network modelling.
Rail – Flow paths based on passenger train timetable data, and passenger numbers based on annual station usage statistics.	No data available on other types of travel patterns and actual passenger travel data is not publicly available.	Train timetable information provides a very realistic quantification of travel patterns. But not having passenger travel data means there is a lot of uncertainty in how passengers are assigned on trains and routes.	Built-in functions in code to estimates flow path allocations.
Rail – Failures estimated by assuming all trains along a disrupted route are stopped and all passengers are disrupted.	Possible over estimation of failures, but there have been several instances of total shutdowns of railways during failures.	We are not accounting for rerouting done by passengers who might jump onto other trains or use the road network.	Built-in functions in code to estimate railways failures.

Roads – Only major roads network considered.	No data available on minor roads network, especially on network flows.	Having a more complete road network would mean flow assignments would be more disaggregated. At present all flows are assigned onto the major roads.	Built in the source data extracted for network modelling.
Roads – Flows modelled from a high-level OD matrix by mapping shortest time paths between nodes as the only preferred travel routes.	The purpose of the analysis was to show the relative importance of routes, which is very well captured by showing the most preferred travel routes.	We are not considering multiple routes of travel between a given OD pair, which would be more realistic.	Built-in functions in code to estimates flow path allocations.
Roads – Failures estimated by assuming all cars along a disrupted route are stopped and all passengers are disrupted.	Same as railways.	We are not accounting for rerouting done by passengers who might jump onto other trains or use the road network. At present we are overestimating road failures.	Built-in functions in code to estimate roads failures.
Interdependency mapping			
Electricity and telecoms were assumed to be interdependent networks, by creating directed links from chosen electricity nodes (substations) towards telecoms nodes (exchanges and macro cells), and other sets to direct links from telecoms nodes to all electricity nodes.	We were most interested in modelling instantaneous failure propagations and failure impacts of the order of a few days, not a few weeks. Hence, electricity and telecoms were considered to be the two sectors whose failures would have such short-term failure propagation effects. It was reasonable to exclude longer term dependencies e.g. the dependency of the electricity sector on water supply (in absence of storage) and transport for fuel. These assumptions were validated with sector experts during Quality Assurance (QA) consultations.	Removal of telecoms to a node may not cause any instantaneous failures and be the case and may only impair operation. But due to lack of data we cannot account for this.	Built-in functions in code to estimate network failure cascades.
Water, rail and roads were considered to be dependent on either electricity or telecoms or both networks.		Removal of service to the dependent assets implied total failure of the node (no partial functioning). Including for removal of telecoms service. This is probably an overestimation of the failure state of the assets.	
Link between two nodes created only if they are <10km apart.	It would be irrational to connect nodes that are very far apart.	The creation of cross sector edges is very sensitive to the choice of distance threshold. If we choose a smaller threshold, e.g. 1km, we would expect a smaller number of dependency linkages. This would also have a huge impact on failure cascades.	Distance threshold parameter assumed in data creation.
Backup supply			

<table>
<tr><td>Only electricity backup supply considered, with telecoms macro cells having at most 2 hours supply, telecoms exchanges with at most 24 hours supply, all water assets with at most 72 hours supply and road tunnels with at most 24 hours supply.</td><td>These values were tested with sector experts while doing the QA consultation of the underlying data and assumptions.</td><td>Due to lack of data we are limited in accounting for electricity backup supply in rail network, and also other backup supply (telecoms) for other networks.</td><td>Backup durations value parameters are fixed inputs in the failure code.</td></tr>
<tr><td>Backups assumed to last anywhere between 0 hours and the assumed duration it was assigned, as per a gamma probability distribution-based survival rate.</td><td>Gamma distributions are very well-known distributions used to model infrastructure reliability for repairs.</td><td>Adds uncertainty to the modes and orders of failures in the network. Useful for sensitivity analysis.</td><td>Gamma distribution parameters encoded as fixed inputs within the backup function of the failure code</td></tr>
<tr><td colspan="2">Future network scenarios</td><td></td><td></td></tr>
<tr><td>The future network state representations are chosen for the year 2050.</td><td>Based on NIC feedback.</td><td rowspan="7">• Only one realisation of future states and different scenarios were considered whereas there could be several possible future states.
• All future projection scenarios of population, GPD, GVA, population growth, energy mix were fixed, which means deterministic future outcomes were considered. There should be greater uncertainty in estimating future possible outcomes.</td><td rowspan="7">All future scenarios assumptions and parameters are built in the codes written for extraction and creation of future network and flow modelling.</td></tr>
<tr><td>In 2050 it is assumed that 70% of the generation mix in the electricity supply would be made up of renewables.</td><td>The choice of 70% was based on the NIC’s assessment that these would be the most realistic futures given the current renewable energy trajectory and future nuclear phasing decisions being made in the UK.</td></tr>
<tr><td>Two future electricity scenarios were considered: (1) Hydro70 – Where domestic heating would be predominantly provided through hydrogen gas; and (2) Elec70 – Where demand for heating by electrification would be very high.</td><td>Based on NIC energy modelled work.</td></tr>
<tr><td>By 2050 it assumed that the vehicle fleet would be 100% electric.</td><td>Based on NIC transport modelling, which is in line with the governments targets to have 100% electric vehicles sales by 2040.</td></tr>
<tr><td>Under future scenario assumptions only electricity network topology is assumed to change, while all other network topologies remain the same.</td><td>Only geospatial data on future energy technologies being planned was publicly available or could be inferred from reports. For all other sectors no geospatial data on future network level developments was easily available.</td></tr>
<tr><td>Residential demands of all sectors would increase based on future high population growth rate forecasts.</td><td>Based on NIC population scenarios modelling.</td></tr>
<tr><td>High GVA growth scenario considered for future.</td><td>Based on ITRC scenarios modelling.</td></tr>
</table>

Passenger usage on transport increase with population and GVA which has an elasticity factor of 0.63.	Based on ITRC long-term transport model assumptions.		
The macroeconomic IO structure is assumed to remain unchanged in 2050. Future economic losses would grow based on compounded GDP growth rate of 1.9% forecasts for UK.	No data on future IO data for the economy. Growth rate number Based on latest PwC report.		

APPENDIX D

Table D-1: Explanation and list of data resources used in the modelling.

Description	Source
Energy – Network Topology	The locations of the nodes were collected and verified from several sources[92,93,94,95] and meticulously checked with satellite imagery as best as possible. Several of the substation data at the distribution level were simply scraped from Google Maps and OpenStreetMap. Similar data sources were used for geolocating the link information, which has lesser accuracy in terms of the geometries but more accuracy in terms of connecting the right types of nodes to each other.
Energy - Demand Allocation	The allocations of demands in MW was first done at the 380 Local Authority District (LAD)[96] administrative area levels for Great Britain, using an energy demand model[97] Data on the supply capacities of the generation sites was collected[94] to identify the source nodes and also to check that supply was greater than the demand. The LAD level data was further disaggregated to the Local Super Output Area (LSOA)[98] level of which there were 41,667 polygons in Great Britain. The disaggregation at this finer scale was done by assuming the energy usage within each LSOA was in proportion to its building areas, where the data from building footprints was obtained from the Ordnance Survey (OS) MasterMap[99]. A similar principle was adopted in allocating residential customers to electricity nodes, by disaggregating LAD level population numbers to LSOA levels based on building footprints and then grouping the LSOA estimates to the nodes.
Telecoms – Network topology	OS Codepoint postcode[100] data was also required to map this information into exchange boundary areas. For estimating core locations and other layers of the fixed network, information from Kitz[101] or SamKnows[102] on the BT's 21st Century Network (21CN) was obtained. Core nodes exist in the most urban areas (London, Birmingham, Manchester, Leeds, Glasgow etc.) and Kitz provides a list of the specific core and metro node locations. A total of 85 exchanges were identified as metro nodes, with 12 of these being outer code nodes, and 8 being inner core nodes.

92 http://datasets.wri.org/dataset/globalpowerplantdatabase
93 https://wiki.openmod-initiative.org/wiki/Power_plant_portfolios -
94 https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes
95 https://www.nationalgridgas.com/land-and-assets/network-route-maps
96 https://geoportal.statistics.gov.uk/datasets/local-authority-districts-december-2017-full-clipped-boundaries-in-great-britain
97 Eggimann S, Hall JW, & Eyre N (2019). A high-resolution spatio-temporal energy demand simulation to explore the potential of heating demand side management with large-scale heat pump diffusion. Applied Energy, 236, 997-1010.
98 https://data.gov.uk/dataset/fa883558-22fb-4a1a-8529-cffdee47d500/lower-layer-super-output-area-lsoa-boundaries -
99 https://www.ordnancesurvey.co.uk/business-government/tools-support/open-mastermap-programme
100 Ordnance Survey, 2019. Code-Point - locates every postcode unit in the UK [WWW Document]. URL https://www.ordnancesurvey.co.uk/business-and-government/products/code-point.html (accessed 10.8.19).
101 https://kitz.co.uk/adsl/21cn_network.htm
102 https://availability.samknows.com/broadband/exchanges/21cn_overview

	Cellular asset data was taken from Sitefinder[103] and pre-processed to identify single site macro cell locations by buffering all points by 50 meters[104]. We also assumed that each exchange either had Virgin Media operating within it, or did not, based on the cable availability provided by SamKnows.
Telecoms - Demand	4G information on coverage by local authority was also taken from Ofcom's Connected Nation report (2018)[105]. LAD level population data was intersected with Postcode/exchange boundary areas. Data for the working population at the LAD level was obtained from official labour market statistics[106] and Scottish Census data[107]. This was intersected and merged with the boundary areas of the mobile macro cells, which were created based on Voronoi decomposition[108].
Water – Network Topology	The best available model was a water resource system model of England and Wales (WREW hereafter) developed at the University of Oxford for studying water risks and scarcity[109]. The data from the WREW model was modified and adopted for this study. WREW is the product of an extensive collaboration led by the University of Oxford between a range of stakeholders: England and Wales's environmental agencies, UK-based water consultancies, the Water UK council, and all of England and Wales's water supply companies. The water system formulation in the model was based on communications with, and datasets provided by, the above stakeholders.
Water – Demand	LAD level population census estimates were intersected with WRZs (Water Resource Zones) areas, which were then assigned to demand nodes based on the allocations of WRZs to specific demand nodes as described in the WREW model data.
Rail – Network Topology	The railways model created for this study relied on a previous study we did on vulnerability assessment of Great Britain's railways[110]. This model has been used in several other peer-reviewed studies[111,112] OS Strategi data[113] on the locations of all existing 2,564 railways station was first collected along with the geospatial information on the line geometries of different railway routes in Great Britain. The line geometries showed the single-track routes, which were sufficient for this analysis. The OS data gave very accurate geospatial information

[103] Ofcom, 2012. Sitefinder [WWW Document]. URL https://www.ofcom.org.uk/phones-telecoms-and-internet/coverage/mobile-operational-enquiries (accessed 12.21.16).
[104] Oughton, E.J., Frias, Z., Russell, T., Sicker, D., Cleevely, D.D., 2018. Towards 5G: Scenario-based assessment of the future supply and demand for mobile telecommunications infrastructure. Technological Forecasting and Social Change 133, 141–155. https://doi.org/10.1016/j.techfore.2018.03.016
[105] Ofcom, 2018. Connected nations 2018: UK report. Ofcom, London.
[106] https://www.nomisweb.co.uk/census/2011/workplace_population
[107] https://www.scotlandscensus.gov.uk/news/workplace-population-and-daytime-population-council-areas
[108] Thacker, S., Pant, R., & Hall, J. W. (2017). System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures. *Reliability Engineering & System Safety*, *167*, 30-41.
[109] http://www.mariusdroughtproject.org/
[110] Pant, R. Hall, J.W. and Blainey, S.P. (2016). Vulnerability assessment framework for interdependent critical infrastructures: case study for Great Britain's rail network. *EJTIR*, 16(1): 174-194, ISSN 1567-7141.
[111] Lamb, R., Garside, P., Pant, R., & Hall, J. W. (2019). A Probabilistic Model of the Economic Risk to Britain's Railway Network from Bridge Scour During Floods. *Risk Analysis*, *39*(11), 2457-2478.
[112] Oughton, E. J., Ralph, D., Pant, R., Leverett, E., Copic, J., Thacker, S., ... & Hall, J. W. (2019). Stochastic Counterfactual Risk Analysis for the Vulnerability Assessment of Cyber-Physical Attacks on Electricity Distribution Infrastructure Networks. *Risk Analysis*, *39*(9), 2012-2031.
[113] https://www.ordnancesurvey.co.uk/opendatadownload/products.html

	on the node and route locations, as verified by matching with satellite imagery. But this data set has not been updated since 2016, so the new railway stations and routes were identified through OpenStreetMap data[114], to plug the gaps in the OS data.
Rail – Demand	We created a trip assignment model using openly available train timetable data[115] and annual station-usage statistics from the Office of Rail and Road[116]. For details of the model see Pant et al. 2016[38]
Road - Network Topology	The road network topology was derived from the Department for Transport (DfT) road traffic statistics data[117]. The original DfT data was post-processed to fill all gaps in connectivity between road links, and in some instances, this was done by also adding ferry links over waterways. The DfT also produces traffic statistics of vehicle counts by direction of travel on roads, which was merged with the spatial network topology. We used the OS Open Roads data[118] to identify all major roads with tunnels and matched them to our road network for this study.
Road – Demand	National Trip End Model (NTEM) of the Trip End Model Presentation Program (TEMPRO)[119]. The NTEM provided an OD matrix of vehicle trips between 7,000 geographical area zones in Great Britain. Passenger numbers by assuming an average occupancy factor of 1.6 across all types of vehicles[120,121].
Cross sector dependencies	We had some detailed information on the locations and types of rail assets that use other utilities, especially electricity. This was an older dataset, that we had created for a previous study[38]
Future Energy Networks Topology	Information on locations of future interconnectors was inferred from the Aurora data generated for a previous NIC study [122] and other sources[123]. Data from the Renewable Energy Planning Database (REPD)[124] quarterly extract, updated till September 2019, gave the locations, and capacities of planned renewable technologies. We looked at the plans to build a new Hinkley Point C power plant on 3.34 GW capacity in the future[125,126].
Future energy - Demand	NIC/Aurora projections based on data generated for a previous NIC study.[127]

114 https://download.geofabrik.de/europe/great-britain.html
115 http://data.atoc.org/how-to
116 https://dataportal.orr.gov.uk/statistics/usage/passenger-rail-usage/
117 https://roadtraffic.dft.gov.uk/downloads
118 https://www.ordnancesurvey.co.uk/business-government/products/open-map-roads
119 https://www.gov.uk/government/publications/tempro-downloads
120 https://www.statista.com/statistics/314719/average-car-and-van-occupancy-in-england/
121 https://www.gov.uk/government/statistical-data-sets/nts09-vehicle-mileage-and-occupancy
122 https://www.nic.org.uk/publications/technical-annexes-electricity-system-modelling/
123 https://www.4coffshore.com/transmission/interconnectors.aspx
124 https://www.gov.uk/government/publications/renewable-energy-planning-database-monthly-extract
125 https://www.edfenergy.com/energy/nuclear-new-build-projects/hinkley-point-c
126 https://www.gov.uk/government/collections/hinkley-point-c
127 https://www.nic.org.uk/supporting-documents/aurora-energy-research-july-2018-power-sector-modelling-system-cost-impact-of-renewables/

	All sectors were allocated new demands in 2050 based on population projections at the Local Authority District (LAD) level (380 areas), which were downscaled to thee sector specific admin levels and the service output areas. The future population projections were based on the NIA scenario of *high fertility (or high growth)*[128] which included the following assumptions: • England - ONS 2014-based high fertility subnational experimental projection. • Scotland - Scotland Stats 2014-based high fertility subnational projection. • Wales - Calculated based on ONS 2014-based high fertility national projection. GVA data taken from the Office of National Statistics (ONS), included. • Current ONS estimates of GVA in 2017[129]. • Future GVA growth scenario projections for 2050 derived by Cambridge Econometrics[130] and used for a previous study for the NIC[131].
Economic Input-Output data	In the UK annual Input-Output tables are generated by the Office of National Statistics[132,133]. For future economic growth and losses we assumed a GDP growth rate of 1.9% for the UK, based on recent studies[134].

128 https://www.nic.org.uk/wp-content/uploads/2906064-NIC-Population-and-Demography-Document-v1_1w.pdf

129 https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/regionalandsubregionalproductivityintheuk/february2019

130 https://www.camecon.com/how/lefm-model/

131 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/601163/Economic-analysis-Cambridge-Econometrics-SQW-report-for-NIC.PDF

132 https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/inputoutputanalyticaltables/methodsandapplicationtouknationalaccounts

133 https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/commentaryonsupplyandusebalancedestimatesofannualgdp/1997to2014

134 https://www.pwc.co.uk/press-room/press-releases/uk-could-remain-top-10global-economy-in-2050.html