University of Strathclyde Engineering

Royal Charter since 1964
Useful Learning since 1796

Defining the Simulation Scope for Extreme Events

Completion Date 2024/03/17

Magnus R Jamieson, Keith Bell, Panagiotis Papadopoulos, University of Strathclyde

Abstract

This report investigates extreme events through the means of using three large, recent reports into different types of extreme event to determine what constitutes an extreme event, what data is likely to be necessary to model it, and what compromises are likely necessary to make such modelling possible. The discussions consider the use of weather data in power system analysis, the modelling of cascading outages within the power system, and the potential of attacks on the cyber-physical system which constitutes the power system. Recommendations are then made in terms of how weather data should be used, what types of power system simulations are necessary, and what metrics are likely to be useful in such analyses. It is suggested that an extreme event be defined as **"any event that, without suitable mitigating actions, would cause, as a result of conditions arising from that event: interruptions to a large number of end users' supply of energy, beyond those that could be expected due an outage of any single item of energy system plant; extraordinary energy market conditions, or; interruptions of energy supply to significant elements of critical national infrastructure "**. This captures all potential extreme events that are likely to be simulated; those associated with correlated weather events; those associated with an extended abnormal weather or operational conditions; those incurred by cyber-physical attacks on the power system. Different scenarios are proposed which would be appropriate to simulate to help define a clearer scope of what simulations to undertake, for example modelling an extreme windstorm similar to Storm Arwen.

THE QUEEN'S ANNIVERSARY PRIZES 1996, 2019, 2021 & 2023
For Higher and Further Education

UNIVERSITY OF THE YEAR 2012 & 2019
Times Higher Education

UNIVERSITY OF THE YEAR 2024 RUNNER-UP
Daily Mail University of the Year Awards

SCOTTISH UNIVERSITY OF THE YEAR 2024
Daily Mail University of the Year Awards

EUROPEAN ENTREPRENEURIAL UNIVERSITY OF THE YEAR 2023
Triple E Awards

The University of Strathclyde is a charitable body, registered in Scotland, number SC015263.

Contents

1 Introduction

Among the primary objectives of the Scenarios for Extreme Events SIF project (SfEE [1]) are to determine a series of feasible extreme events which could affect the GB energy system and determine modelling frameworks with which to quantify the impacts these events could have, and the resilience of the system to these events.

This brings with it a wide range of challenges which, in order to be addressed, first need to be set out explicitly so they can be better understood. This document seeks to describe various aspects of weather and extreme event simulation, the data challenges thereof, and how this might affect the choice of extreme events used in simulations and resilience modelling on GB. This will be through examining different potential sources of extreme events and recent work related to it, and recent attempts to study cascading outages on interdependent infrastructures.

For the purposes of this report, it is important to explicitly define some terms at the outset to avoid misunderstanding or ambiguity. These are as follows:

Risk – Risk associated with a particular event is a mathematical combination of the *likelihood* of an event and its assessed *impact*. This can via be a qualitative assessment of impact or quantitative. In the latter case, the risk metric would be product of likelihood and impact. An example of a risk metric is Expected Energy Not Served (EENS) – an *expectation* is an average, and in this case the impact is *energy not served*- a function of time and electrical energy[1].

Resilience – Resilience relates to the properties of a system which define its ability to ***anticipate***, ***resist, contain, survive, recover from, and adapt to extreme or unpredicted events***. This can be quantified via concepts such as the resilience trapezoid, described in [2, 3] but the specific metrics used have not been standardised yet. Some metrics such as Value at Risk or conditional Value at Risk [4] have been proposed [5]. This definition is a synthesis of definitions used in, for example, [6-8], though there is no single universally acknowledged definition.

Reliability – This, in general, describes the performance of a power system in ***predictable, typical operational conditions*** and is commonly measured by metrics such as Loss of Load Probability (LOLP), Customer Interruptions (CI) or Customer Minutes Lost (CML). Reliability is much more widely understood and characterised, though whether resilience is an expression of system reliability or should be conceptualised as a sector of power system security in its own right remains a subject of debate. The metrics used to quantify reliability are unlikely to be appropriate in all contexts for quantifying resilience, however, as they rely on averaging across large sample sizes and predictable events, whereas resilience is generally concerned with extreme events or events whose probabilities and impacts carry significant uncertainty. Even if extreme events are included in an estimate of system reliability, their rarity tends to mean that the values of reliability metrics are relatively unaffected by such events.

Risk assessment - Any discussion pertaining to *risk analysis* should be just that **– inclusive of both *impact* and *probability***. If modelling is, in fact, an *impact* assessment, it should be clearly labelled as such so as not to cause confusion between those more versed in risk analysis and those using the terms more liberally and colloquially.

The categorisation and use of language is important in these contexts because there may be different usages deployed by different sectors of expertise and a first point should be to ensure consistency across these groups.

[1] EENS is typically quantified for a system, summing the event-related expected energy not supplied across many different events.

1.1 Quantifying and mitigating risk

In actuarial terms, risk is relatively straightforward to calculate – a combination of impact and probability across all realised scenarios. This is standard and acceptable practise for reliability analysis, which is based on well-described systems with well-understood probabilities and impacts involved, but this simplistic approach can flatten out the impacts of HILP events, as a single event of extreme impact but low probability can equate, in risk, to multiple smaller risks with higher probabilities. This theme is discussed in [9] in the context of "choosing" which reliability-related events to defend against and can be broadly described mathematically here.

A consistent theme across all the investigated scenarios mentioned is that, for the heavy tail of events, probability can be difficult to ascertain, therefore accommodating error or uncertainty in these estimates is essential.

In mathematical terms, a standard formulation for calculating risk can be illustrated as

$$risk = \pi_{A1} c_{RA1} + \pi_{A2} c_{RA2} + \pi_{A3} c_{RA3} = \pi_B c_B \quad (1)$$

in which π_n is the probability of a given outcome of event n, and c_n is the value of the associated impact which may be, for example, load curtailment, customers disconnected, or a monetary impact. In this instance an example is given where the risk of multiple smaller incidents is equivalent to one larger incident.

In reality, an operator may explicitly wish to avoid a single extreme event even if that means an increased probability of less severe events. A suggested amendment to this is discussed in [10] and involves adding an exponent x to more heavily weigh high impact events in risk calculations, i.e. such that

$$risk = (\pi_{A1} c_{RA1})^x + (\pi_{A2} c_{RA2})^x + (\pi_{A3} c_{RA3})^x < (\pi_B c_B)^x \quad (2)$$

In these instances, also, probability may be a suitably generated estimate.

The ultimate aim of any resilience or risk assessment is to manage risk. That does not necessarily mean eliminating it altogether. This is illustrated in Figure 1 [11].

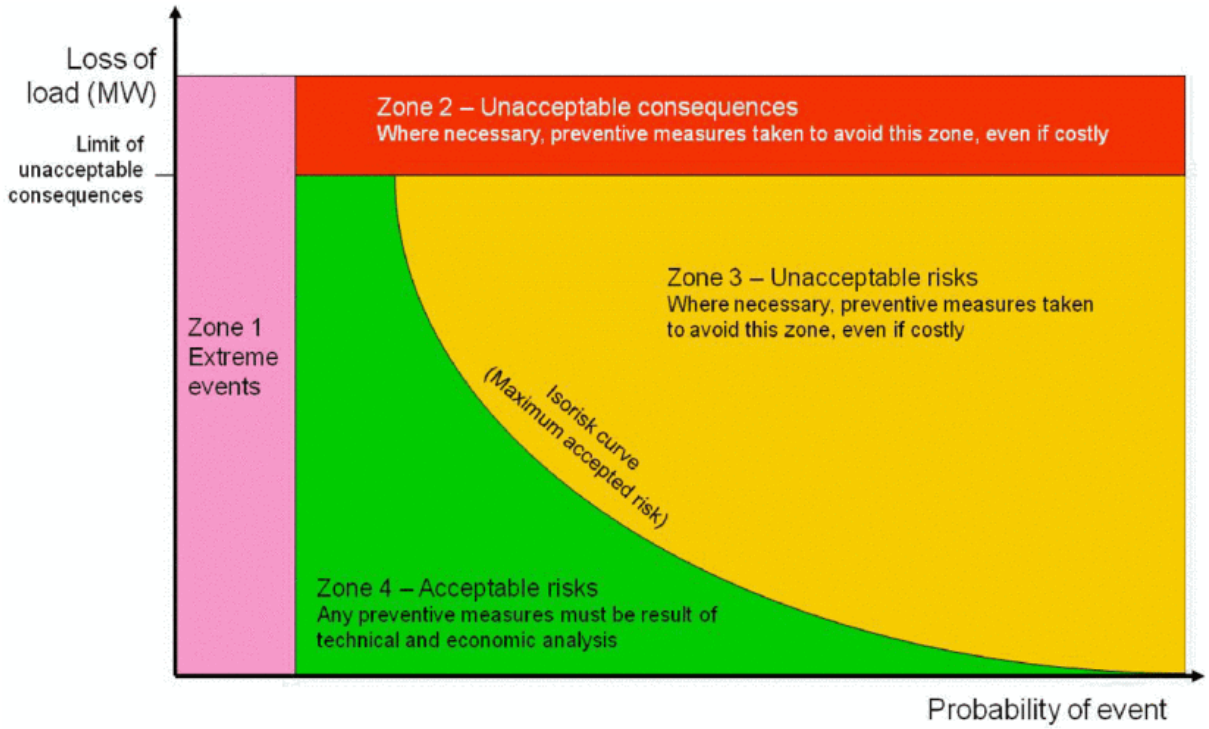

Figure 1 - classification of risk associated with power system operation [11]

Moving a risk "down" a region is always a desirable outcome, even if not eliminating the risk altogether, e.g. moving from Zone 2 to Zone 3 may still be desirable even if it cannot eliminate the risk. This can be done by either reducing the probability of a given series of outcomes or mitigating their impacts. This should be the aim of modelling extreme events – can either the probabilities or impacts be reduced such that unacceptable risks be rendered tolerable? The role of modelling is to answer that question, or to establish the costs thereof. As is argued in [12], the answers and evidence are not always straightforward to present to stakeholders but it is still essential to have a robust evidence base on which to make the arguments.

Bearing these aspects in mind, therefore, the metrics should reflect the impacts being assessed within the power system and be informative, without obfuscating what they are actually representing. Some suggested metrics are:

- Number of customers disconnected
- Average duration of disconnection
- Aggregated duration of disconnections
- Cost of remedial generation (that is, cost of additional generation that must be dispatched to mitigate the given incident versus nominal operation)
- Cost of planned load deferral (i.e. if customers are paid to not use electricity at a given time to reduce remand during acute periods) or curtailment
- Estimated cost of disconnections (VoLL × MWh lost for a given event)
- Number of critical customers disconnected
- Critical customers' average disconnection durations
- Aggregated critical customers disconnected durations
- Generation disconnected, either output at the time of disconnection or capacity
- Expected number of component/infrastructure failures (e.g. expected number of tower failures)
- Energy not served

This list is not exhaustive, and more may present themselves as useful as the SIF project on extreme event progresses. More typical metrics should also be considered for stochastic simulations as a baseline, such as Loss of Load Probability (LOLP) and Expected Energy Not Served (EENS). Increasingly popular risk-based metrics such as VaR and cVaR could also be used to demonstrate their utility.

2 Weather Data – ESIG Report and Related Literature

ESIG (Energy Systems Integration Group) – a consortium dedicated to research of large power systems - has produced an extensive report into using weather datasets for power system simulations [13]. This is a significant piece of research which touches many themes which will be discussed herein as to how they impact resilience studies, using the ESIG report as a backbone to discuss wider weather-related resilience modelling issues.

2.1 Linking weather data to power system modelling

The report is primarily concerned with forecasting features such as demand and generation and how these are affected by weather conditions, but there are a significant number of features which are also relevant to more generalised resilience modelling. Figures and quotations used within this section should be presumed to originate from this report unless explicitly referenced otherwise.

Weather data within power systems, it is argued, is primarily used for 3 general purposes:

- Development of weather-related generation and assessment of the performance of that generation
- System planning and adequacy studies
- Operations planning

However, it is also increasingly used for scenario analysis and study of extreme weather events, such as in [14, 15]. This will be discussed throughout this section as well as broader discussions of the themes raised in the ESIG report to contextualise the weather-related discussions to power system resilience studies.

The key outcomes and findings of this report, and the most consequential figures, will be discussed here. The over-arching theme of the report is that there has been, to date, insufficient co-operation between the meteorology sector and power system planners and operators. The report itself is relatively US-centric, however its findings are generalisable and still useful in a GB context, but key differences exist.

Historically, the primary interaction between the power system and weather has been temperature-related, with temperature driving winter peak demands associated with heating and summer demands in the USA associated with air-conditioning and cooling, This is a finding corroborated by [16]. Reports such as [17] have also investigated the impacts of climate change more specifically within the UK in an increasingly renewables-heavy system, which should be referenced for more specific discussions of different aspects of power system modelling associated with climate change.

If we are concerned with modelling extreme events, we must first define what we categorise as “events” before we can say whether climate change as a factor is a feature of the events we wish to model.

A major problem with climate change and its interaction with the power system is the nature of these effects to be nonlinear – changes will likely not be gradual but may well accelerate in unanticipated ways. This has consequences for modelling the impacts of extreme weather on power systems as not only are the statistical relationships between weather and outages difficult to establish, but the probabilities of the kinds of weather events which lead to power outages are likely to shift.

Different weather events and parameters will have different interactions with different aspects of the power system. This is illustrated within the report which presents the following figure, in Figure 2 [13].

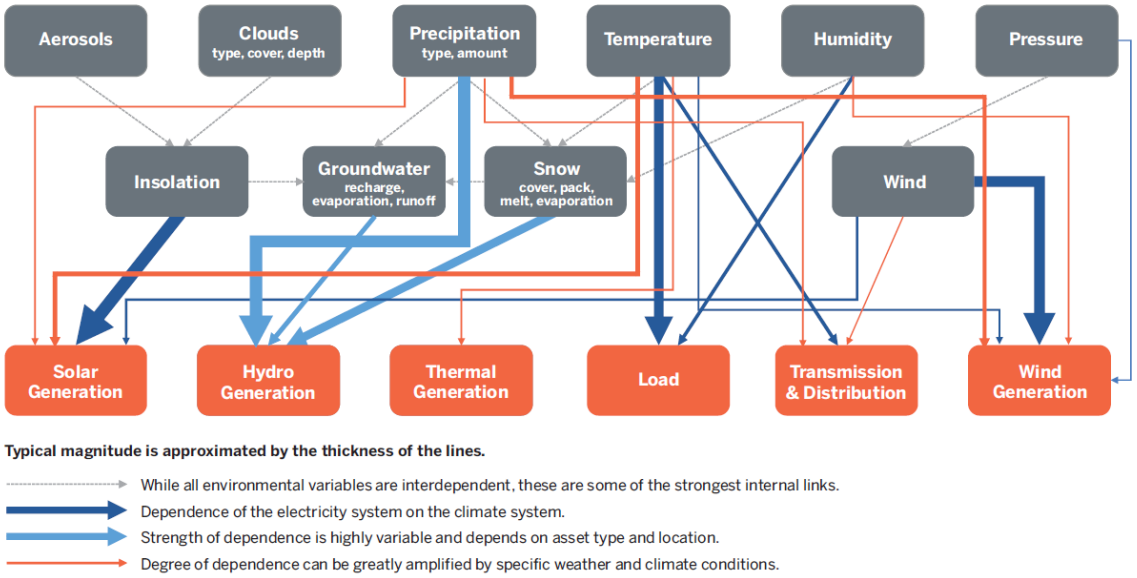

Figure 2 - The primary linkages between variables in the weather and climate system (gray) and the electricity system (orange) [13]

The strongest links in the figure are shown in dashed grey lines. Dark blue lines indicate direct dependencies that are most important in typical operation of the electricity system, while orange lines indicate dependencies that do not typically have a large impact on a daily basis but can have a profound impact in particular circumstances or combinations.

These linkages could be used to refine the scope of events we wish to simulate and help define the weather datasets needed to perform them – but care should still be taken. For instance, “wind” is itself an umbrella term. Different features of wind have different impacts on different aspects of the power system and in different timescales. For example, a long period of low wind will not generate risk related with overhead lines but in combination with a low-solar period (also known as *dunkelflaute) this* could result in acute energy supply challenges.

Snow and ice are not shown here as having an impact on transmission and distribution infrastructure, but historically in the GB system they absolutely have, especially on exposed assets in places like the Highlands of Scotland. It can also affect solar and wind generation through icing of turbines at low temperatures, and through snow coverage of panels and roofs.

Further, it would be more useful to separate out “high” and “low” temperatures as independent phenomena affecting the power system. While heating will lead to forced derating of assets or emergency shutdowns of generation associated with high cooling water temperatures and water shortages, icing will have different impacts across critical infrastructure. Groundwater can also have an impact on transmission and distribution infrastructure.

There is also no mention of space weather, though this is not as unsurprising in and of itself due to the complexities associated with understanding the associated phenomena and their relative infrequency compared to the effects of, for example, extreme wind on power systems.

Another figure is produced showing specific tasks and how weather data is used for these, shown in Figure 3 [13].

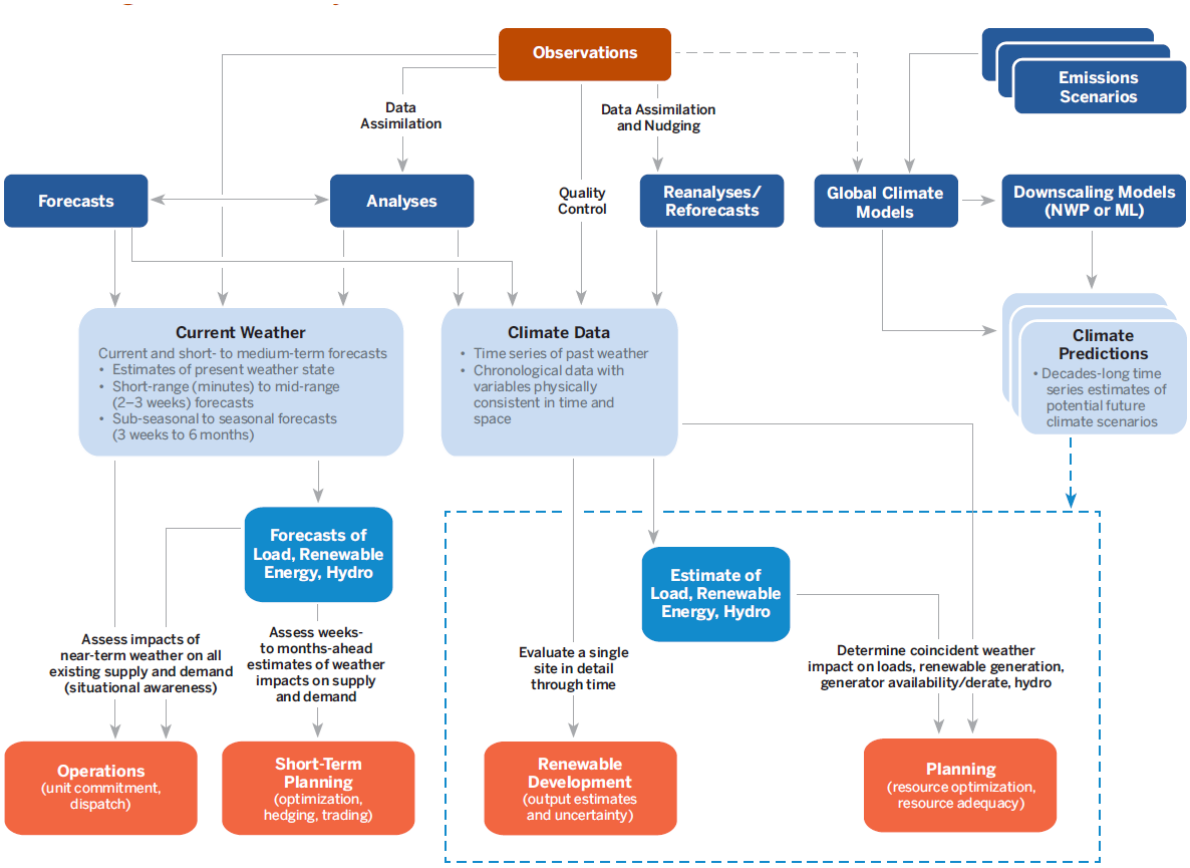

Figure 3 – Flow of Weather Data into the Primary Power Systems Processes That Require Meteorological Data for Analysis Tasks [13]

Weather data (light blue and dark blue boxes) flow into the primary power systems processes (orange boxes) that require meteorological data for analysis tasks. In some cases, this includes tasks that estimate states of the electricity system (medium blue boxes) that are impacted by the state of the atmosphere. Solid lines indicate the flow of data that are output from one process and become an input to another. The dashed line from observations to global climate models indicates that the data are used in validation of average global climate model outputs versus as an input to their production.

Another important factor in modelling high-risk weather events is, as demonstrated in Figure 2, different features of weather can act at the same time to impact the power system. Wind can occur concurrently with snow, sleet, and blizzard (SSB) conditions, which is discussed in [18]. Categorising faults associated with weather carries significant challenge in its own right [19].

One means of associating weather parameters with failure probabilities are *fragility curves* which relate a hazard of some sort to a failure or damage probability. A review was carried out into the deployment of these which can be found in [20], with a brief discussion of their conceptualisation and principles also found in [21]. A drawback with these approaches is that isolating an individual feature which has caused an outage is extremely difficult because of the number of ways which weather impacts the power system.

The incident weather upon an asset ”when” it fails, for example, may not be the event which has caused the damage leading to its failure, but may rather be, to use an analogy, the straw that broke the camel’s back. Modelling the failure as a binary representation is therefore almost always a simplification of both the effects of damage to the asset and the causes associated with the damage to that asset.

Extreme wind speeds can also result in wind farms curtailing their output as a defence mechanism to prevent damage to turbines. High wind events will be correlated with high failure probability on overhead lines and the probability of curtailed wind generation [22]. Further, localised wind impacts will vary across wind generation sites, and different heights will experience different wind conditions based on local geographic conditions which can make projections of wind power challenging [23], [24].

The evidence base surrounding the impact of climate change on wind resources is also not strong, with mixed findings. Some imply a “southward shift” [25] but other papers find little change, to slight growth [26]. This implies that scenario-based analysis, longer-term, is likely to be important (e.g. scenarios where there is significant growth in wind resource availability versus status quo or falls). Post hoc, probabilities can be assigned to these scenarios to establish risk.

The lack of universal observational data requires methods such as reanalysis to create weather datasets which can be used to generate fragility curves. That is, it is not possible to have observational data for every point on the surface of the Earth (at the time of writing). Power system analysis can use observational data for specific sites, such as windfarms, but will typically rely on reanalysis datasets such as MERRA-2 [27] for the analysis necessary to generate fragility curves which can be used to generate scenarios used in resilience analysis. This is necessary where observations do not exist or are not of satisfactory temporal or spatial granularity. For forecasting, analysis may use Numerical Weather Prediction (NWP) methods from a wide range of sources.

Any weather models used for scenario generation should, ideally, “capture the physical and dynamical relationships between weather variables and produce weather states that are physically plausible, evolve realistically in time and space, and produce distributions of conditions like those that are observed”.

Though sources of data such as MERRA-2 are useful, they should not be treated as a “black box”, and efforts should be made to ensure that the data used in the generation of any fragility curves deployed is the same as, or at least comparable to, the data used in the scenario generation itself.

Alternatively, appropriate numerical corrections should be applied to this data to make it appropriate. In the cases where this is not fully possible, full acknowledgement of the limitations of the data should be present – even if the absolute values which are generated from such implementations carry significant error or uncertainty, the relative outcomes of related scenarios using consistent datasets is still useful or could still be useful to generate plausible scenarios for analysis.

Data from different sources may not be appropriate to “mix and match” when making projections, as the presumptions underlying the data regarding the physical processes involved may not be consistent. Similarly, there is a difference between data which is *physics-driven* and data which is *statistical*. One is derived from an initial condition and extrapolated forward based on physical rules, and the latter can be formed from retrospective analysis from the measurements taken and the data related to it. An example of this is the cumulative probability distributions generated in [18] versus fragility curves generated from analytically produced mechanical models of towers and power lines such as those used in [28].

Weather datasets such as MERRA-2 may be of a form that is too large in scale to be appropriate for some types of network, e.g. some data from MERRA-2 is ~50kmx50km in horizontal spatial resolution, which is too large to represent distribution networks. Features such as wind shadow, shading from mountains, and wind tunnels cannot be accurately captured at these resolutions. The following Figure 4 [13] is presented to illustrate this concept.

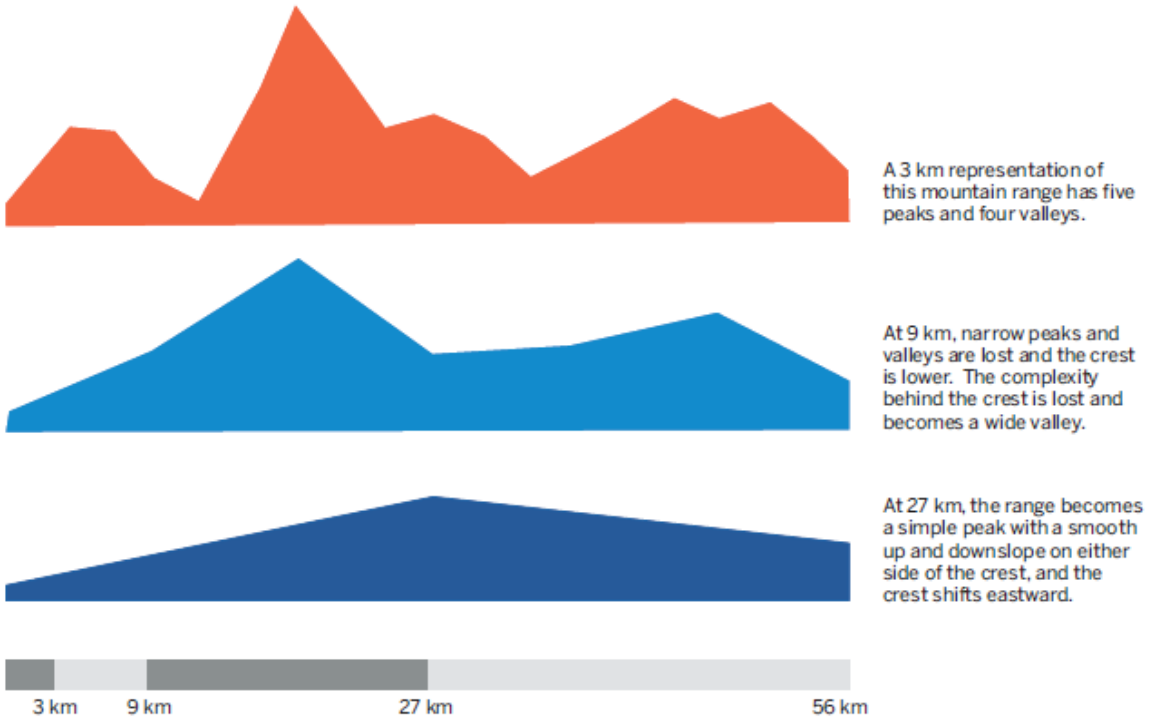

Figure 4 – The top plot shows a cross-section of hypothetical complex topography represented at 3 km grid spacing. The middle plot uses the average of sets of three 3 km points for each 9 km point. In the bottom plot, three 9 km points were averaged to get to each 27 km point [13].

For modelling of geographically diverse regions, such as the Appalachian range in the USA or the Highlands of Scotland, such large resolution pixels (in the scale of 50km) may not always be appropriate for accurate representation of the weather parameters across these ranges. Downscaling and interpolation may be deployed retrospectively but this still cannot directly and accurately capture these effects. But merely allows more granular representation of networks subject to the weather represented by this data.

Given the cubic relationship between wind speed and power generated from a turbine, the “averaging” effect of these large resolution datasets could have significant impacts on the projections of wind generation. This is illustrated in Figure 5 [13].

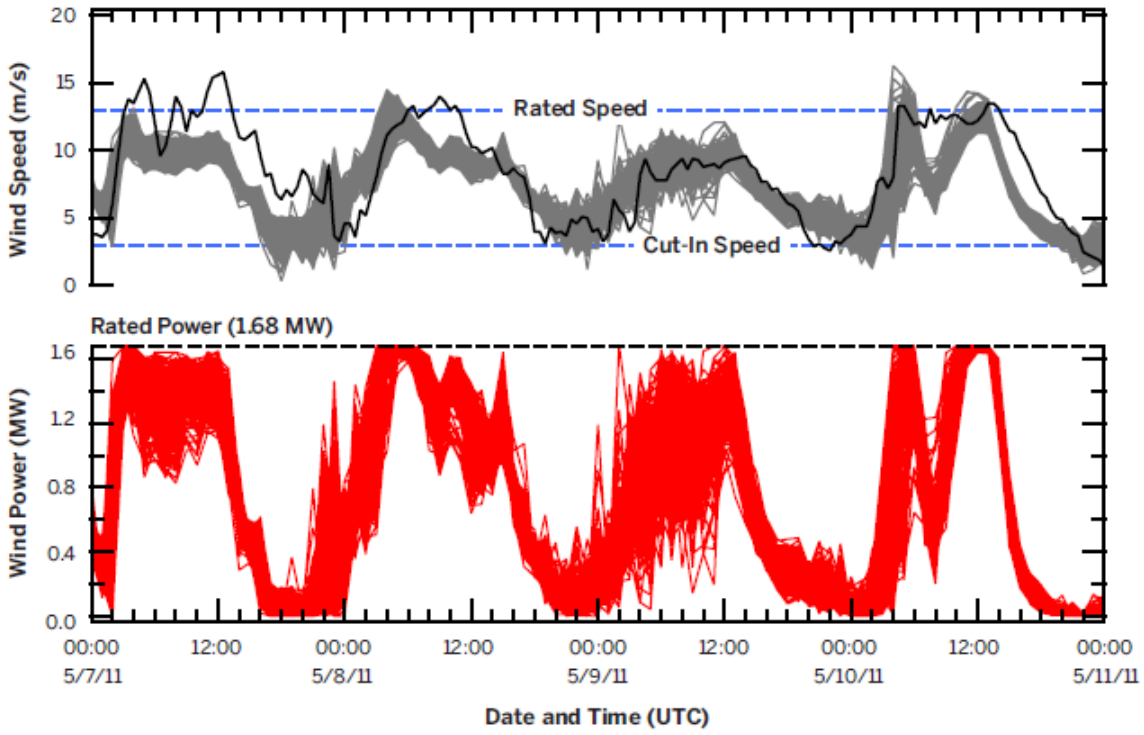

Figure 5 – Traces of wind speed and wind power for many different iterations of a model run with everything held constant except parameters related to turbulence and surface roughness [13]

The upper plot in this figure represents a range of wind speeds generated by numerous model runs. The lower plot translates those wind speeds to estimates of wind generation. The wider spread seen in the lower plot shows the effect of parameter choice when the cubic relationship between wind speed and power output amplifies these differences.

The reason for the scale of the resolution used is, of course, related to the computational expense of simulations on a global scale. As noted, “the computational resources needed to increase horizontal resolution scale by at least the third power because the number of required time steps increases by the same factor as the resolution change to keep the model computationally stable.” Hence, simulations based on 1km take 27 times the computational resources of a 3km simulation, and 27,000 times the resource of a 30km simulation. All simulations based on geospatial data therefore require a balance between data granularity and computational and temporal expense.

Mountain ranges and valleys will affect both precipitation and wind patterns in local areas, and heavy precipitation itself carries other risks such as landslips and flooding, which can affect large areas of network for significant amounts of time. An examination of the social and economic impacts of flooding-related outages is also considered in [29].

There is some discussion regarding the use of machine learning techniques and Generative Adversarial Networks (GANs), a machine learning technique, to generate weather scenarios and climate models, but this work is not fully developed. Machine learning has been used in the past as a “classifier” for simulating many perturbation events quickly and efficiently, however [30]. A problem with such techniques and the use of “convolutional networks” is that the data can exhibit regression to the mean and can “blur” out extremes of data, underestimating values of e.g. rainfall in specific regions. This is particularly problematic in the case of resilience simulations because it is those extremes which are of specific concern. Methods and approaches which work in one region may not work in another precisely because of these geographic effects on the underlying data.

ESIG (used here to be understood as to refer interchangeably to the authors of the report) therefore recommends that validation of the underlying data used for retrospective analysis is of significant importance, while also recommending specific tools for doing so (e.g. the WIND Toolkit [31], though it is USA-specific). To this end, wind datasets should not be used as a “black box”.

ESIG categorise 3 different types of renewable-energy-related power system studies. These are:

1) Renewable integration studies
   These simulate the power system with varying levels of renewable resources
2) Integrated resource plans
   These are optimization-based models to evaluate long-term costs and benefits of various resource mixes
3) Resource adequacy studies
   These typically require hourly data and are concerned with evaluating the capacity of the system to meet energy demand based on reliability metrics

Parts of these concepts will overlap with resiliency and security studies, but for our purposes a fourth category explicitly studying extreme events should be categorized.

The approaches used for traditional reliability studies in less weather-dependent systems are likely to be inappropriate for resiliency studies, particularly those concerning extreme weather. The specific approaches resilience studies take will be considered in the following section, but within this section it is worth observing that conventional approaches of Monte Carlo Markov Chain (MCMC) simulation are unlikely to be able to capture the events used for resilience scenarios without either specialised sampling algorithms or large sample sizes.

Similarly, given the interdependence of weather variables, generation, and demand, traditional methods which would involve sampling from demand curves and dispatching generation to meet this demand cannot be reconciled with increasing penetration of weather-dependent generation. Such data relies on physical relationships that affects multiple aspects of the power system and can no longer be treated as independent. This changes the necessary approaches for modelling extreme weather events and implies a need for more physics-driven models, while ensuring consistency in the datasets used for scenario generation.

Demand has always, to varying degrees across the world, been dependent on weather, but increasing interdependence between weather and generation availability is fundamentally changing this relationship. Relatedly, features such as ambient temperature, precipitation and humidity can impact generation through e.g. covering solar panels in snow, icing of wind turbines, etc, as well as their impacts on overhead line (OHL) networks at transmission and distribution level.

Further, if behind-the-meter (BTM) generation is observed by operators as a reduction in demand, this complicates the modelling of distribution networks yet further particularly if e.g. high wind and high solar power days intersect during frequency excursion events, which can see disconnected demand during Low Frequency Demand Disconnection also disconnect significant amounts of generation, in turn exacerbating whatever event is underway.

Given the lack of certain types of spatio-temporally complete and comprehensive datasets for all modelling contexts, compromises may have to be made, but there should be a level of coherence maintained between datasets and, where necessary, compromises must be acknowledged even if ideal validation is not possible.

Further, historical models used to produce estimates of wind generation also have to be updated for improvements in technology and for scale of turbines; e.g. turbines in the North Sea will be larger and have much greater capacities than those onshore due to e.g. availability of space on the sea vs land, and the consistency of wind on flat seas compared to geographically diverse regions which are more

complex to model. Similarly, the height at which observation data is measured is not always the height at which the hub of the turbine is, thus there may be discrepancies between the actual speed incident on the wind turbine (or any elevated asset) – this is also discussed in [22]. There may well be "reasonable" estimates, but given the cubic relationship between wind speed and wind power, and the exponentially increasing failure rate on OHL assets above certain thresholds (typically ~25ms^{-1} [22]), these small discrepancies can combine significantly.

Wind and solar, of course, are not the only sources of power affected by weather. Hydroelectricity – pumped and storage – can also be affected by long-and-short-term weather conditions. Pumped storage will also not only be affected by weather patterns, both seasonal and over shorter periods, but by the deployment of pumped-storage hydro by system operators, adding further complexity to modelling these interactions.

These factors all combine during, in particular concern for GB, winter. Common mode failures associated with wind, SSB, and icing can occur during periods when there is maximum demand associated with heating. Further, extreme cold, if systems are not designed for it, can even affect thermal generation such as gas turbines, as demonstrated in the ERCOT blackouts most recently where cascading outages and outages were in part associated with freezing of gas infrastructure [32].

Weather conditions such as this also aid the system in specific conditions – high wind that is not so high as to pose a risk of outage can facilitate enhanced cooling of lines, and thus enable techniques such as Dynamic Line Ratings (DLR) to enhance system resilience and enhance system capacity. After all, the role of the transmission system is to transport power, and the impact of line outages is to reduce the capacity of that system to move power. Dynamic line ratings can counteract that effect [33].

Counter to this, of course, is high wind scenarios in areas at high risk of wildfires – high wind speeds and dry conditions can precipitate extreme risk of wildfires which leads to operators having to shut down power lines to mitigate this risk in a process known as public safety power shutoff (PSPS), which is widely deployed in California [34]. The use of microgrids has been explored to improve power system resilience related to these phenomena, as well as the wider implications of wildfires on power systems [35].

2.2 Data requirements

As is evident, there are clearly a wide range of factors dictating what data is required in power system analysis and how it should be applied. These requirements are summarized in the following figure [13].

Including the necessary variables	Include the necessary variables at sufficient spatio-temporal resolution and accuracy to reflect actual conditions that define the generation potential at current and future wind/solar sites and temperature at load centers
Covering multiple decades with ongoing extension	Cover multiple decades with consistent methodology and be extended on an ongoing basis to capture the most recent conditions and allow climate trends to be identified
Coincident and physically consistent	Are coincident and physically consistent, in space and time, across weather variables
Validated	Are validated against real conditions with uncertainty quantified
Documented	Are documented transparently and in detail, including limitations and a guide for usage
Periodically refreshed	Are periodically refreshed to account for scientific and technological advancements
Available and accessible	Publicly available, expertly curated, and easily accessible

Figure 6 – data requirements for power system analysis [13]

These attributes are explored in more explicit detail within the report, but these fundamental principles should be considered and adhered to as much as possible in modelling extreme weather events, though naturally some principles do not apply to all contexts (for example, if modelling a single scenario with a known set of parameters, it is not necessary to consider decades of data). The temporal granularity of data for simulations will also, naturally, be dictated by the nature of the power system simulations being undertaken.

Some generalised observations can be taken, however, from the more specific properties of these attributes, which shall be discussed within this section.

It is necessary to know temperature in sufficient detail to determine its impact on loads on the system. Similarly, variables which can impact generation should all be considered to reasonably approximate generation from various sources – e.g. if line icing or galloping is likely to be a threat, it may be necessary to include data relating to ambient temperature and precipitation as well as wind speed and direction.

When extrapolating wind temperature to specific heights to determine impacts on e.g. connectors on OHL, the correct number of vertical levels should be used to be able to correctly extrapolate these wind speeds.

A typical wind speed power curve for an individual wind turbine is shown. These will be well understood by windfarm owners and manufacturers [13].

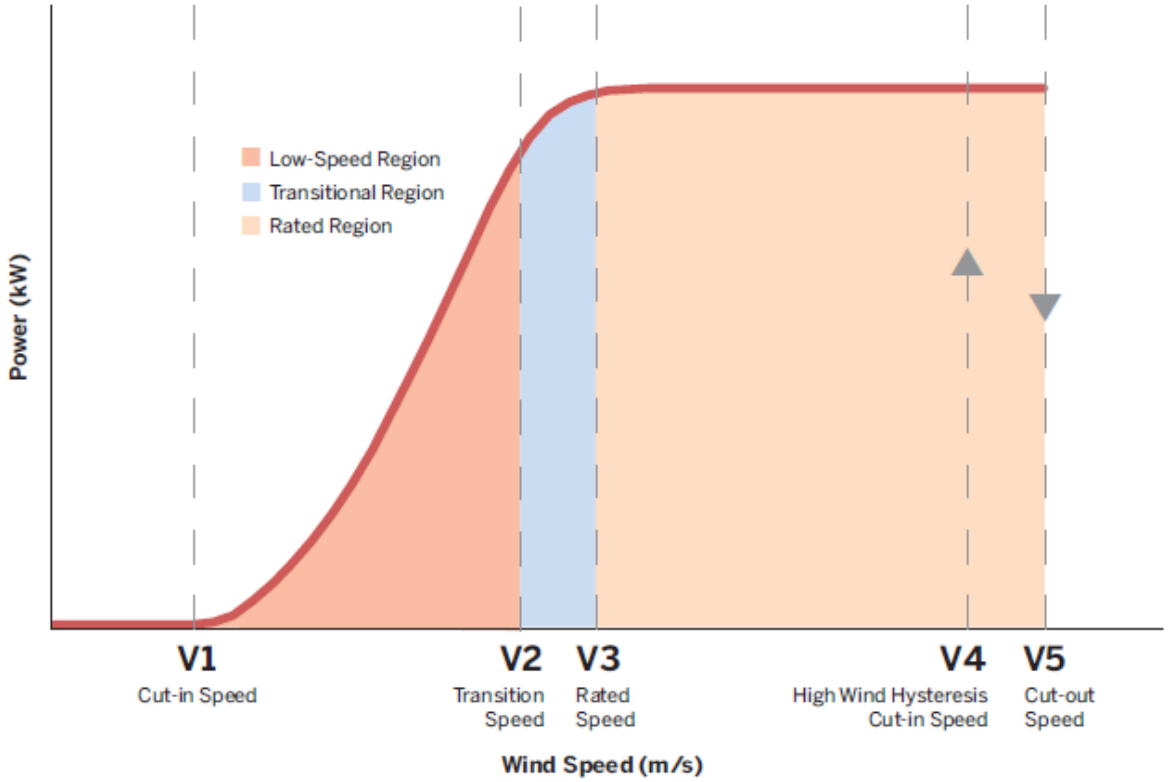

Figure 7 – typical wind turbine power curve [13]

The various critical points are described.

- V1 is the cut-in wind speed, the speed above which a turbine begins generating power.
- V3 is the rated wind speed, the speed at which the turbine reaches its rated power output; at speeds higher than this, no additional power available in the wind is captured, as the generator cannot further increase its output.
- V5 is the cut-out wind speed, the speed at which the pitch of the turbine blades reduces the output to 0 to protect the turbine.
- Operation of the turbine is suspended until the wind speed has slowed to V4 before it goes back up again, cycling between V4 and V5.

This can accurately describe the behaviour of an individual turbine, but for an aggregated wind farm the behaviour will not be linear as wind speeds will vary across the site, resulting in a complementary sigmoidal-type curve around where the cut-in and cut-out speeds occur as wind turbines drop in and out of service. This is explored in [23], [22], and [36].

Features such as ambient temperature also impact solar generation, as it can reduce the efficiency. This can be difficult to measure because local effects can be significant, e.g due to wind, shade, or cloud cover. The temporal variability for individual panels may be significant, as well as at site-level. Any model of solar generation will either have to acknowledge abstracting these effects out of it or take direct consideration. Effects such as the Urban Heat Island effect may also affect urban panels and thus have greater thermal effects than solar farms in more rural regions. The same is true for extrapolating thermal cooling or heating demand in cities.

Projections of hydroelectricity are particularly difficult because there is inherent hysteresis in both run-of-river and pumped storage hydroelectricity. Both can be affected by precipitation, flash floods, ice

melt, and ambient temperature- though pumped storage can more directly be controlled by human operators.

Though short-term dispatch problems and optimal-power-flow problems can involve sampling of e.g. power curves based on instantaneously determined readings of wind, solar, etc, longer term unit commitment problems are made more challenging by the time-delayed impacts of hydroelectricity which can make them challenging to model and incurs significant uncertainty (or should, if model accuracy of hydroelectricity is a priority).

To summarize these data recommendations:

- Data should be of appropriate horizontal granularity (2km is suggested)
- Appropriate vertical levels should be deployed to ensure accurate modelling of wind speed at different heights
- Splitting a country or analysis region into sub-regions is not recommended by ESIG as it can introduce problems such as dataset seams; however, from a power systems perspective, specific analysis of, for example, the Highlands versus the Central Belt may be of value
- If modelling unit-commitment problems, the data used should contain all the information to project hourly load and wind/solar at any generator as if it were being modelled in real-time
- While we want to use the longest datasets possible for long-term projections, poor quality data should be avoided which, generally, means going back no further than 1990; it is not evident why 1990 specifically is chosen but it does coincide with the earliest date available from the BARRA dataset [37]
- Observation data should be used in preference to statistical models wherever possible, but statistical or reanalysis methods may be useful to help clean up, complete or summarise observations.
- It may be worthwhile to utilise different datasets through the same model to generate multiple realisations of the atmospheric states to generate an “envelope of truth”, thus giving an indicator of uncertainty
- Datasets which accurately predict annual capacity factors but not outlier events may be appropriate for developers, but not for extreme event analysis
- Documentation describing the datasets used and how they are applied will be essential for reproducibility and validation of work
- Standardization of results and availability of data used will be important if work is to be developed further or disseminated

2.3 Grading of specific data sources

From the wide variety of data sources appropriate for energy-related studies available, a table is produced to grade these as necessary, which is presented in Figure 8 [13]. This grades different weather sets based on whether they meet the specific attributes and requirements they describe.

	Spatial Resolution	Temporal Resolution	Length	Continuously Extended	Correct Variables/ Levels	Coincident and Coherent	Validated/Uncertainty Quantified for Power System Use	Detailed Documentation	Future-Proofed	Availability/ Ease of Access	Curation and Advice	Region Covered
MERRA-2[a]	~60 km	60 min	1980–present	Yes	Yes/No	Yes	No		Probably		Basic	Global
ERA5[b]	~30 km	60 min	1940–present	Yes	Yes/No	Yes	Some		Yes		Good	Global
HRRR[c]	3 km	15 min	2014–present	Yes	Yes/No	Yes/No	No		Unideal		Basic	U.S.
WIND Toolkit[d]	2 km	5 min	2007–2014	No	Yes/Yes	Yes	Yes		No		Basic	Various
WTK-LED[e]	2 km/4 km	5 min	3 year/ 20 year	No	Yes/Yes	Yes	Not yet	Not yet	No	Unknown, dataset not yet available		Various
NSRDB[f]	4 km/ 60 km	30 min	1998–present	Yes	Yes/No	Solar only	Yes		Yes		Basic	Most of globe
CERRA[g]	11 km/5.5 km	60 min	1980–present		No/Yes	No solar	Yes		Possibly		Basic	Europe
CONUS404[h]	4 km	60 min/ 15 min (precip)	1980–2020	No	Unknown/ Probably	Yes	Not the intended use					Continental U.S.
BARRA[i]	12 km/ 1.5 km	60 min	1990–2019	No	Yes/ Probably	Yes				Fee-based		Australia/ New Zealand
Public Observing Networks[j]	Non-uniform, variable density	1 hr or less	Variable	Yes	Yes/No	Mostly	Varies. Not for power systems	Varies	Usually	Usually easy	Varies	Global
Renewable Energy Project Data[k]	Non-uniform, variable density	Usually minutes	Variable but rarely more than 10 years	Varies	Yes/ Usually	Yes	Usually	Varies, but usually poor	Varies	Usually poor	Usually none	Very limited
Proprietary Statistically Derived VRE Shapes[l]	Non-uniform, variable density	Usually hourly	Variable. Rarely reliable long records.	Varies	Usually incomplete	No	Partial	See note	No		None	Very limited

Fully Met | Close to Being Met | Partially Met | Met in a Very Limited Way | Not Met at All | Not Enough Info. for Determination

Figure 8 - Summary of Current Power System Modelling Weather Input Data Sources

MERRA-2 has been discussed already, and is used in [22]. Of note is that MERRA-2 is of significant utility, but “must be downscaled first”. ERA-5 is deemed “unquestionably the best global reanalysis dataset currently available”, though it is “not a panacea”. Therefore, it is recommended to use this dataset going forward if attempting to reproduce historic events or generate credible extreme weather scenarios.

Of note, also, is that though tools such as the WIND Toolkit [31] have significant use also, they have a tendency to over-predict windspeeds. This may be acceptable for more conservative resilience analyses (i.e. you want higher failure rates on lines to establish upper bounds of risk), when projecting wind power it resulted in a 5-10% too high capacity factors. Improvements are currently being researched, however (e.g. the WIND Toolkit Lon-term Ensemble Dataset (WTK-LED).

A similar toolset has been generated for Solar – the NSRDB [38]. This covers the USA and selected international locations dependent on satellite coverage. Unfortunately, as can be seen in Figure 9 [38], this does not include Scotland.

Figure 9 - METEOSAT IODC coverage, which excludes Scotland [38]

There is also a readily accessible data repository generated by researchers at Imperial College known as renewables.ninja, which references different turbine manufacturers, solar panel types, weather repositories etc to provide relatively accessible data for weather studies [39]. This is not mentioned in the ESIG report, though the scale of corrections used in the data may mean it is not necessarily appropriate for high fidelity localised studies.

ERA-5 and the WIND Toolkit and the NSRDB used in tandem are compared in the following figure to summarize the key attributes of these weather sources [13].

Attribute	WIND Toolkit/NSRDB Combination		ERA5
	For Wind/Load	For Solar	Wind/Solar/Load
Has required temporal resolution[a]	5-min produced	5-min since 2019	Hourly
Has required spatial resolution	2 km	4 km; 2 km since 2019	30 km
Includes multiple heights above the surface		N/A	
Available for several decades	8 years[b]	Since 1998	Yes
Has regular updates	Nothing formal	Annual	Daily (7-day lag)
Is future-proofed	Ad hoc	Yes	Yes
Is long enough to detect climate signals	Unlikely	Possibly	Yes
Models are adequately validated			
Accuracy assessed, including for risk periods	Against tall meteorology towers	Limited	Limited
Variability assessed, against reality	Limited	Limited	Several studies
Assessed power system modeling applicability?	Designed for this	No studies found	No studies found
Provides companion "forecasts"[c]	Produced	No, but possible	No
Is based on consistent input observations and/or models	Yes, except 2014	Yes	Yes (single modeling system)
Physical consistency between wind/solar	No; impact should be investigated		
Well documented and easy to use			
Limitations are clearly specified			

Fully Met | Close to Being Met | Partially Met | Met in a Very Limited Way | Not Met at All

Figure 10 - Summary of Best Available Public Datasets to Estimate Site-Level Generation at All Current and Future Wind and Solar Assets in All Regions of the United States [13]

It is also asserted that observations should almost always be preferred over e.g. reanalysis data "as observations are always better than model data". This is not strictly true in GB, as observation data can be inconsistently collected and is not always available as readily as the repositories suggested here.

The following properties are proposed for an open-access, comprehensive public dataset to meet the proposed attributes and requirements. It is proposed it would have to meet the following conditions:

- Having sufficient spatial and temporal resolution for analyses
- Including all necessary variables in one dataset in space and time
- Coverage of multiple decades with a consistently applied methodology and being extended on an ongoing basis
- Coincident and coherent across all variables
- Validated with quantified uncertainty
- Comprehensively and transparently documented, including an account of limitations of the data and a user guide

- Future proofed
- Publicly and easily accessible

While the data already referenced adheres to these principles to varying extents, it is not generally true for network data for representative models of e.g. distribution systems and generator models on GB. This data tends to be incomplete or spread across multiple sources or requires time-costly processes involving non-disclosure agreements (NDAs) across multiple stakeholders, which can make academic publication of results awkward and more time consuming than it necessarily needs to be.

IEEE test models do, of course, exist and are widely used in academia [40], but are not typically appropriate for real-world case studies and can be difficult to contextualise for those not familiar with the test systems being deployed.

Similarly, failure and fault data, and its relationship to weather data, does not have a generalised central repository which can be easily accessed with such curves. There are resources which reply reliability-related data, e.g. the Transmission Availability Data System (TADS) [41] in the USA, but this is not associated with particular weather events which would be useful for modelling resilience events retrospectively.

Finally, it is suggested that when analysing HILP events retrospectively, the reanalysis datasets should be compared with observational data from the time to determine how well the extremes were captured. Similarly, regions with high penetration of renewable energy resources and potential energy resources should be targeted for observational data readings as these often correlate with regions with low population density and low penetration of observational data recording sites.

2.4 Suggested improvements or adaptations for deployment of weather datasets

The datasets that end up being used are likely to be imperfect. Therefore, it is almost certain that corrections or adaptations will be necessary to improve the data and its usage. This was the case throughout [22] which both interpolated wind data and extrapolated it using the Wind Power Law to correct wind speeds to the appropriate height for quantifying OHL failure probability and aggregated wind power – this is not to mention the adaptations necessary for the electrical power system models used. These can include:

- Using surface measurements in addition to synthetic datasets (but not mixing and matching synthetic datasets)
- Extending data sources to represent higher levels of renewable resources
- Extrapolating existing datasets to include more years of consistent data
- Extending existing datasets to evaluate the impact of extreme weather

It is noted that just because a data source has the required spatio-temporal fidelity and covers the appropriate geographical region, that it will be appropriate for the given task at hand. A key observation within this report is **“If a model is simple and easy to understand by a non-meteorologist, it likely is not very accurate”.**

Emphasis is placed on collaboration between power system engineers and meteorologists throughout – power system engineers will understand best what operational states lead to stress within the system, while meteorologists are best placed to understand the physical mechanisms which can create the scenarios under study and the uncertainties in any data used to model these scenarios. It is also emphasised that power system experts should not tamper with the underlying data “unless they are working closely with a qualified meteorologist”. In practical terms, this is simply not always possible because there may simply not be enough meteorologists available to provide insight into every research project. Nonetheless, this does not mean efforts should not be made.

It is possible that such uncertainties in understanding the weather data being used to make forecasts of resource availability can lead to adverse operational scenarios, e.g. if there is an overestimate in the models to forecast wind generation availability and reserve generation is accordingly under-provided, which could lead to significant operational risk.

Suggestions for methods to generate plausible scenarios are provided. This includes one specific example where a Markov Chain was used to wander between bins of data representative of real daily weather transitions and according to weather parameters were derived from associated weather patterns from within those bins. The code and procedures for this method are provided in the links associated with [42]. This is an attractive method because it can generate plausible events from historical data reflective of realistic conditions, based on recent historical events.

Of note are that two “stress tests” are defined for future analysis, based on longer term phenomena. One involved early gas plant retirement, low hydro availability, and coal retirements which impacted on the ability to import power into California over multiple years. The other related to transmission bottlenecks on the Eastern seaboard. We should be clear, if we are modelling “extreme events” what these events are and the temporal range of those events, because long term events are not necessarily less harmful than short term shocks.

2.5 Weather data do’s and don’ts

The insights throughout the section are summarized into the following “do’s and don’ts”, which are all relevant to the resilience analysis we wish to undertake:

- **Do** consult meteorologists: “don’t go it alone”
- **Do** model stressors to all resource types
- **Don’t** assume extreme weather only impacts renewables
- **Do** stress-test systems against as many future weather realizations as possible
- **Don’t** make investment decisions using single-weather years
- **Don’t** just evaluate doomsday scenarios
- **Do** use data reflecting likely correlations among stressors
- **Do** consider weather in neighbouring grids
- **Don’t** assume each power system is an island

2.6 Impacts of climate change

If we are modelling extreme events, climate change itself will be an incidental causal factor, rather than a specific subject of analysis. That is, it will change the frequency and intensity of events under study, but if our concern is on the heavy tail of weather events, we are already concerned with the extremes.

What climate change can contribute towards is creating combinations of extreme events as more wind and solar are added to the energy mix, and it is these combined – or “compound” – events which are likely to be of greatest risk to the power system and should be considered as key subjects of study, e.g. low wind coinciding with extreme cold and restrictions on gas supply, as is increasingly likely following the Ukraine war.

Some variables are known with greater certainty than others. For example, near-surface temperatures are “increasing at a rate proportional to anthropogenic greenhouse gas emissions”, with events such as multi-day heatwaves and “tropical nights”, “when temperatures do not drop below 20°C during the night”, also increasing in frequency. This will in turn increase electrical demand associated with cooling in summers, though it is difficult to predict the scale to which this occurs. That is, policy decisions can have a direct impact on this type of demand with factors such as retrofitting of insulation and cooling devices or enhanced building stock reducing electrical demand passively.

The following events are identified as being necessary for modelling with “moderate certainty” that they will occur:

- Extreme temperatures leading to more frequent outages and derating of system assets, coincident with high load scenarios
- Periods of low resource availability for renewables
- Drought and shifting timing and location of precipitation, contributing to increased wildfire-related risks

Not noted within this section directly is the fact that changing precipitation patterns also affect run-of-river cooling of nuclear plant, which can run dry during extreme heat or, in periods of acute heat, result in water too warm to cool generators reliant on it. Controls on thermal emissions from such plants during extreme heat events can also force plant to shut down for environmental reasons to prevent further heating of watercourses.

Further, extremes of temperature associated with cold causing spikes in demand, or extremes of wind causing high line failure probability coincident with ramping events on wind generation. A summary of proposed events and case studies will be proposed in the concluding section.

For real-time estimates of conditions on the system, if such simulations are to be considered, properties of the weather data used which should be considered also include:

- Timeliness
- Accuracy
- Can facilitate a general estimate of net load, including both demand and BTM generation

2.7 Summary of weather-related literature and ESIG report

There are a number of useful observations and case studies provided within the ESIG report but they required contextualisation and further analysis to make them relevant to resilience studies for GB. Clearly, the ways in which weather has interacted with the power system in the past are not the ways in which they will do so in the future.

Historically, time-of-day and season have been the two primary drivers of generation and demand on the system. Demands such as heating and lighting, and energy for cooking, were predictable with shift patterns, and heating and cooling demand were forecastable with relative confidence. Ramping events could be predicted to be linked to events such as sporting events or mass television events. However, penetration of renewables, “smart” devices, streaming, and to an unknown extent increasing shifts in working patterns are all changing how we use electricity.

The transmission system observes demand no longer as simply a function of how much energy customers are using, but also as a function of how much generation on distribution networks is providing the system.

The transmission and distribution systems themselves have always been subject to weather, but the effects of climate change are likely to exacerbate already-known issues on power systems globally. It should be noted that “high-risk events do not have to be “extreme” in the classical sense to pose risks”. The war in Ukraine was not an “extreme” weather event, but it had an acute effect on energy markets that was felt profoundly by consumers. Similarly, when weather and climate patterns change, what we now understand as “extreme” may no longer be so.

There is a need for better weather data to quantify these risks. The desirable attributes for this data have been discussed and illustrated at length. Generating better weather data is not simply a useful intellectual exercise – it is also likely to be significantly less costly than “blindly building trillions of dollars of infrastructure without the basic tools to cost-effectively optimise it and assess its reliability”.

In the context of this project, the data used will help determine what events are chosen for analysis, and thus will help guide what events are deemed highest risk. It is important, then, that this data is as robust as possible.

The ESIG report did not itself discuss in great detail the consequences of extreme-weather related outages and events, but nonetheless understanding the data which is behind analysis of these events can aid in understanding how likely events which can affect the power system are, as well as helping to describe what an “extreme” or “stress event” actually is.

2.8 Suggested weather-related case studies

Defining an “extreme” weather event is clearly of importance as a basic starting point, as well as what a “stress test” for the power system might realistically be. Based on the discussion within this section, then, the following table is used to describe what are believed to be the key potential **weather-related** case studies for examination. It is assumed that system recovery is also to be modelled, and associated factors relevant to modelling these events which should be considered. These events are deemed to be probable within GB based on recent historical events or likely developments in the GB system in the near future.

Table 1 - proposed weather-related stress tests for the system

Scenario	**Weather parameters likely necessary**	**Example incident (if present)**	**Affected infrastructure**	**Consequences on infrastructure**	**Exacerbating features**	**Temporal range**
Extreme wind in winter (no precipitation)	Wind magnitude (multiple levels), solar irradiance	Storm Arwen	Wind generation, overhead lines, , gas demand and networks	Damage to OHL, ramping events on wind generation, high wind speed shutdown (HWSS)	Demand, spatial extent, asset ages, cumulative damage over multiple events	Seconds – days
Extreme wind in winter (snow, sleet, blizzard, (i.e. SSB)	Wind magnitude (multiple levels), precipitation, ambient temperature, solar irradiance	Winter storms 09/10, “Beast from the East”	OHL, wind generation, exposed substations, solar generation, hydro generation, gas demand and networks	Damage to OHL, ramping events on wind generation, icing, line collapse due to weighting, flashovers and infrastructure being snowed-in, HWSS, solar panels being covered in snow, transport inaccessibility	Demand, spatial extent, asset ages, cumulative damage over multiple events, accumulated precipitation or icing, post freeze melt	Seconds - days
Extreme cold (no precipitation, low wind)	Wind magnitude (multiple levels), ambient temperature, solar irradiance	“Beast from the East”	Demand, OHL, exposed substations, gas demand and networks, hydro generation	Line icing, turbine icing, high demand, damage to pipes and other underground infrastructure	Duration, availability of generation through maintenance and random	Hours - Weeks

					outages, gas supply	
Extreme heat (low wind)	Wind magnitude (multiple levels), ambient temperature, humidity, solar irradiance	Summer 2018, 2022	Demand, constrained distribution infrastructure, generation, solar generation, thermal generation, hydro generation, gas demand and networks	OHL derating, transformer derating, hydro generation curtailment, thermal generation curtailment, solar generation inefficiency	Hydro availability, spatial extent, duration, cumulative heat effects on system	Hours - weeks
Dunkelflaute	Wind magnitude (multiple levels), solar irradiance, ambient temperature	Not yet experienced in a GB system with a high dependency on wind	Demand, wind generation, solar generation, gas demand and networks	Energy market price spikes, inability to perform maintenance on generation to maintain supply	Hydro availability, generator maintenance schedules, gas availability, storage, duration	Hours - Weeks
Extreme rain (no/low wind)	Ambient temperature, rainfall, wind magnitude (multiple levels), solar irradiance	Storm Desmond	Low-lying substations and generation facilities, solar generation, demand, hydro generation, gas	Inundation of assets, damage to hydro run-of-river infrastructure, landslips, coastal asset damage from storm surges	Previous temperatures and permeability of land, cumulative rain, duration, spatial extent	Minutes - Days

			demand and networks			
Extreme rain (high wind)	Ambient temperature, rainfall, wind magnitude (multiple levels), solar irradiance, rainfall	Storm Babet	Low-lying substations and generation facilities, solar generation, demand, hydro generation, OHL, gas demand and networks	Inundation of assets, damage to hydro run-of-river infrastructure, landslips, coastal asset damage from storm surges, damage to coastal assets, OHL failures, HWSS	Previous temperatures and permeability of land, cumulative rain, cumulative damage to assets, duration, spatial extent	Seconds - Days
High renewables output during low demand (low precipitation)	Ambient temperature, solar irradiance, wind magnitude (multiple levels)	Summer day in Future Energy Scenario (FES) from the ESO	Solar generation, wind generation, hydro generation, gas demand and networks	Energy price volatility, potential for system instability	Vulnerability to outages, low system inertia, duration	Seconds - Days
High renewables output during low demand (high precipitation)	Ambient temperature, solar irradiance, wind magnitude (multiple levels), rainfall	Autumn day in FES	Solar generation, wind generation, hydro generation, low-lying generation and transmission infrastructure, gas demand and networks	Inundation of assets, damage to hydro run-of-river infrastructure, landslips, potential for system instability	Land permeability, cumulative damage to assets, cumulative precipitation,	Seconds - Days

Wildfires	Ambient temperature, solar irradiance, wind magnitude and direction (multiple levels), rainfall, Convective Available Potential Energy (CAPE)	The Camp Fire	Solar generation, wind generation, hydro generation, low-lying generation and transmission infrastructure, gas demand and networks	Fire damage to transmission and distribution infrastructure, widespread property damage, damage to gas networks, damage to demand and generation	Dry land can exacerbate heavy rainfall and cause flash flooding afterwards, smog, pollution, widespread damage	Hours - weeks
Lightning storms (without precipitation)	Ambient temperature, CAPE, Rainfall, wind magnitude (multiple levels), solar irradiance	N/a	Solar generation, wind generation, exposed transmission and distribution infrastructure	Can cause transient outages and damage to Transmission and Distribution Infrastructure (TDI)	Can spark wildfires following dry periods, can be exacerbated by extreme winds	Milliseconds - Hours
Lightning storms (with precipitation)	Ambient temperature, CAPE, Rainfall, wind magnitude (multiple levels), solar irradiance, precipitation	“Thundersnow”	Solar generation, wind generation, exposed transmission and distribution infrastructure, low lying substations and infrastructure, gas networks and demand	Inundation of assets, damage to hydro run-of-river infrastructure, landslips, coastal asset damage from storm surges, damage to TDI, transient outages, line icing in SSB conditions	Can be exacerbated by extreme winds with associated impacts, if dry ground beforehand can cause flash flooding	Milliseconds - Hours

This list is not intended to be exhaustive but should reflect a broad spectrum of potential scenarios and the kinds of events likely to affect the GB system now and into the future. For some features (e.g. communications networks) the assumption of effects may be more indirect. For example, telecommunications towers in towns, villages and cities may be affected by extreme winds; communications systems will be affected by any power outage. If models of these networks are unlikely to be simulated (e.g. communications or water networks) they are not directly included as model features but may be affected. Particularly challenging are compound events where there is a coincidence of extreme events either at the same time or before the system has had the capacity to recover from the previous event, or where subsequent events exacerbate each other. For example, extreme rain following a protracted dry period can lead to flash flooding as the ground is less permeable and less able to absorb precipitation.

Though efforts have been made to include modelling of transportation networks and features such as visibility [43], at this point in time it is recommended to be treated as a low priority for extreme event analysis of the kind undertaken here, but should be a future consideration. This is simply due to management of feasibility of models at this point. Water networks were deemed marginal and should be only be included if an appropriate proxy or representation can be determined.

There are complex relationships and interdependencies between the water, transportation, gas, and communications networks and outages will percolate in difficult to anticipate manners. These “network effects” are discussed in the following section, which considers “cascading outages”. A general principle is that a parameter is included in this table if 1) it is reasonable to expect that it will be directly impacted by the weather event being studied; 2) it cannot reasonably be abstracted out of the model; and 3) it is possible and reasonable to form a representative model. Point 3 in particular will be key to defining the scope of any model used. It should be noted that even if complete quantitative modelling of a given scenario or threat cannot be conducted within the scope of the project, a qualitative assessment would still be of value to determine future research directions or take account of threats to the system. Such analysis can be useful as a platform for further investigation of risk and need for quantitative modelling, judgements on risk mitigation actions, or awareness raising.

3 Cascading Outages Modelling – NIC/Oxford University Group Report

Unlike the ESIG report, this piece of research by the NIC and a research group at the University of Oxford [44] attempted to perform some quantitative analysis of how outages may propagate through a system. The work is of varying strength, and the assumptions made in developing some of the network models used can be questioned.

The network models used linked electrical, water, communications, and transport networks together in a single node, including both road and rail. The aim of the research was to identify a "range of vulnerabilities characteristics that arise from the architecture of the UK economic infrastructure network"; model likely changes to how these systems interact in the future; use the model to produce an assessment of these characteristics and their relative importance; identify resilience enhancement options.

Infrastructure systems are defined in [44] as "the collection and interconnection of all physical facilities and human systems that operate in a coordinated way to provide infrastructure services"- this is an appropriate definition both for what they sought to achieve and what this research project aims for.

The over-arching approach will be described here, and the validity of the approach and the results discussed. As with the previous section, quotations, unless otherwise explicitly stated, will originate from [44] – tables and figures are referenced accordingly.

3.1 High-level methodology summary

The network models themselves area modelled as interlinked node-branch models overlain on each other, as shown in Figure 11.

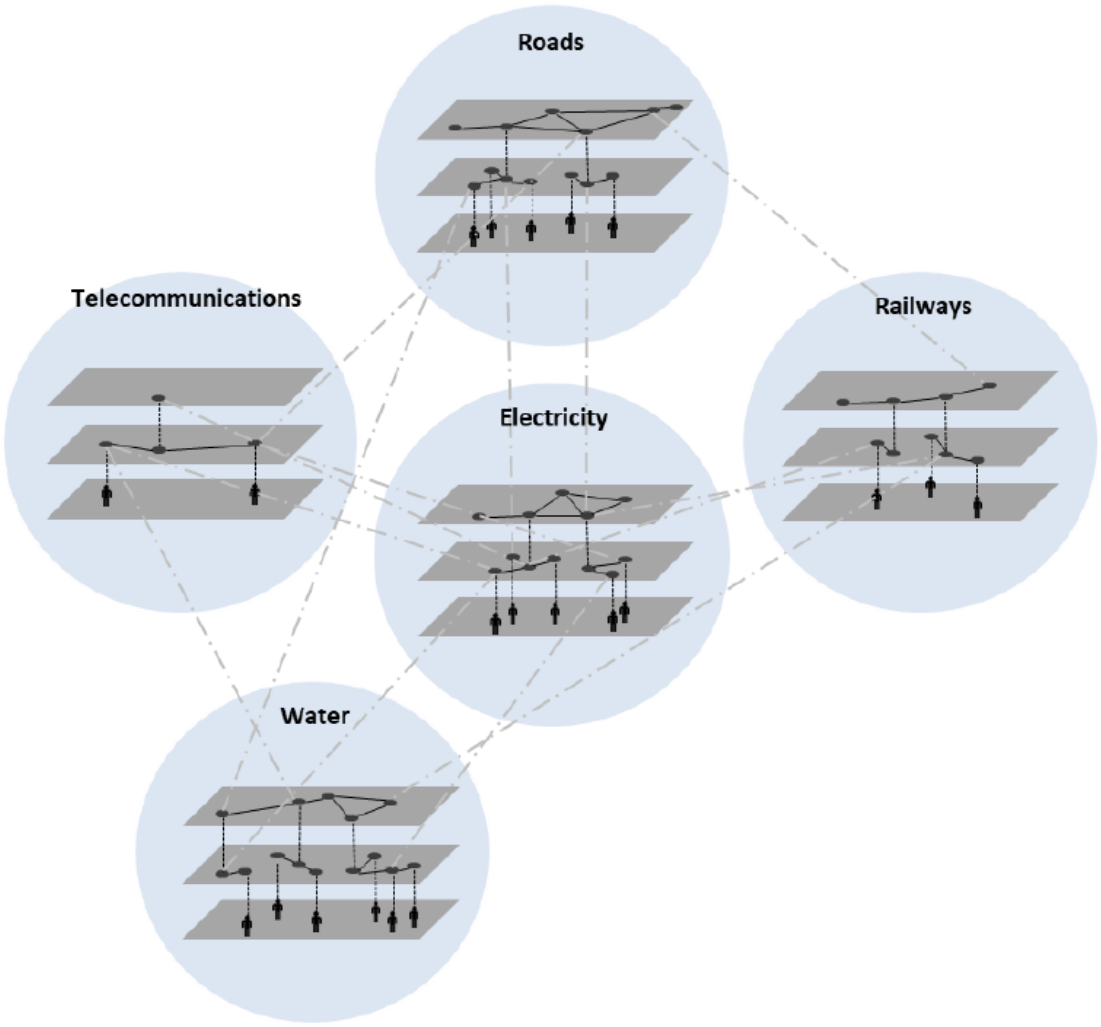

Figure 11 - System-of-systems conceptualisation of infrastructure networks and their interdependencies [44]

This is a logical approach to take in the first instance, because clearly these networks are interlinked in materially significant ways – railways are used to transport coal to power stations, roads are used to transport staff, water is used in thermal plants of all kinds to varying extents, and telecommunications networks are used to co-ordinate responses. The model is an attempt to model three different FES, generated not by the ESO itself, but by *Aurora* and the National Infrastructure Assessment (NIA) via the NIC, envisioning different levels of Hydrogen penetration; and one envisioning a 100% EV fleet. These are described as shown in the table below.

Table 2 - Future scenarios from the NIA and their translation in network topology, flow, and failure models [44]

Future scenarios	Network topology modifications	Flow/demand modifications	Implications on failure analysis
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	• 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	• 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	• 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
2. *Elec70* – Electricity generation is mainly driven by increased renewables supported by gas and demand for heating by electrification is very high			
3. Preparing for 100 per cent electric vehicle sales	• No changes to road or electricity topology • All other networks remain the same	• Added transport EV demand will add more load onto the electricity network	• Will increase electricity service demand losses • EV demands will be tested as alternative backup supply options

The network analysis fundamentally uses topological searches and path searches to establish if network nodes are still connected. If they are not, they are taken out of service and this effect "cascades" down to interlinked services. The approach of modelling outages is shown in Figure 12 [44].

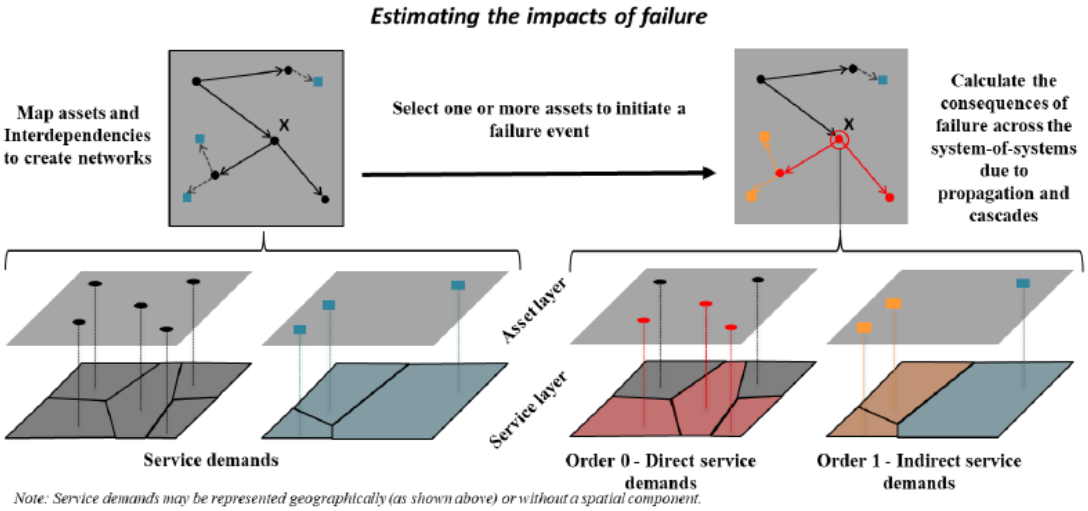

Figure 12 - Representation of direct and indirect service disruptions across interdependent networks [44]

3.2 Weaknesses in consortium approach

There are various weaknesses to the approach taken within this modelling in how it relates to modelling cascading outages on the power system.

Firstly, there is no consideration of gas networks in the model. Even if hydrogen gas replaces methane in the gas transmission system, this still needs transported. As evidenced by the ERCOT cascading outages referenced in the previous section [32], faults in the gas system can cascade into faults across the whole energy system. It is therefore counter-intuitive to model rail, transport, and water networks but to not consider gas networks which are responsible for meeting considerable heat demand.

Secondly, there is no consideration of network *capacity.* In distribution networks which are typically radial, full simulations of network flows, voltages and dynamics may not be necessary simply due to the fact that there is little or no redundancy in the system, so if a connection is lost it is almost always the case that any demand or generation on the opposite side of the line outage from the feeder will be lost. Therefore, branch capacity is not an issue – simply the binary of whether the network is still connected or not. This assumption does not hold for transportation or meshed power networks, whereby loss of connectivity will lead to congestion in transport networks, and in power systems may lead to cascading outages, system destabilisation, sympathetic tripping, thus further outages.

Thirdly, there is no consideration of frequency or voltage in the modelling of “cascading” outages. What the paper refers to as a “cascading” outage and what power system engineers understand as “cascading outages” are materially different. The outages referred to within the research paper which the consortium deploy can better be thought of as “connected” or “percolating” outages – outages with a direct causal relationship due to network effects, but not due to cascading electro-mechanical or electro-magnetic effects such as frequency or voltage collapse.

Simulation of cascading outages is a significant area of study for electrical engineers both in terms of quantifying the causes of these events and the effects thereof, as well as the analysis of the kinds of events which can cause them.

For example [45] uses a link between bespoke software and an optimal power flow (OPF) solver to determine voltage and power flows within a software loop with the optimisation function being used as proxy for what a system operator or the electricity market would do under those circumstances. This approach has been deployed in various forms for some time, for example in the “Manchester Model” from 2002/03 [46], illustrated in Figure 13, and more recently in [47].

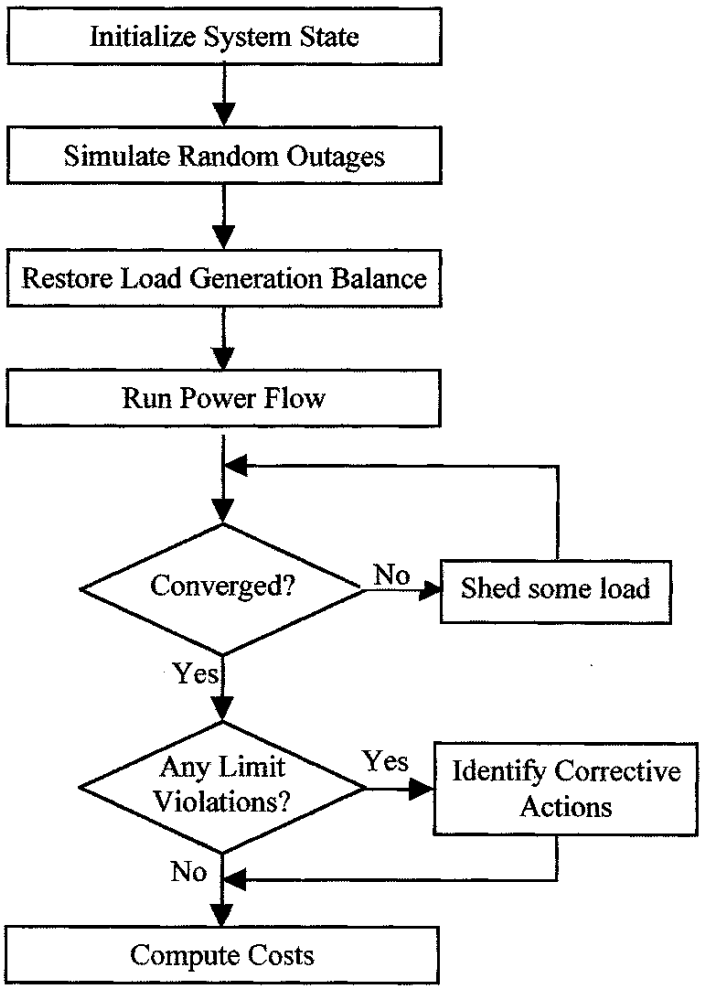

Figure 13 - so-called "Manchester Model" of representing cascading outages [46]

Using such models to compute the value of system security has also been discussed within [48]. The market for corrective actions and preventive dispatch has diversified significantly in recent years to reflect the changing nature of power grid operation, with services segregated into, for example “primary frequency response”, “secondary frequency response”, “fast frequency response”, and “dynamic containment” [49]. These refer to different types of power injection to the power system to maintain stability when the balance between supply and demand rapidly changes.

The dispatch of systems to prevent cascading outages, codified in the Security and Quality of Supply Standards (SQSS) [50], has always been implicit in the dispatch and planning of day-to-day power system operations and can be abstracted and calculated via methods such as Security-Constrained OPFs [51] and Unit Commitment (UC) problems [52]. It is common to abstract out frequency-and-voltage-related features in large-scale resilience studies and use, for example, OPFs [15], but these still rely in some part on modelling the ability of the network to actually transport power. Recently it is also increasingly the case that abstractions or representations of frequency response are included in resilience simulations [53]. It is not reasonable or realistic to ignore branch capacity altogether as well as frequency response on power networks as has been conducted in the NIC study.

Another potential weakness in the modelling of networks relates to how they generate water networks. Because there were insufficient data on water networks to create a network for GB, the Oxford-led consortium synthesised a model from another project which they linked to in the report, but the link provided to the data did not work at the time of writing. The base model used to synthesise the water network was also only based on England and Wales. The model is shown in Figure 14.

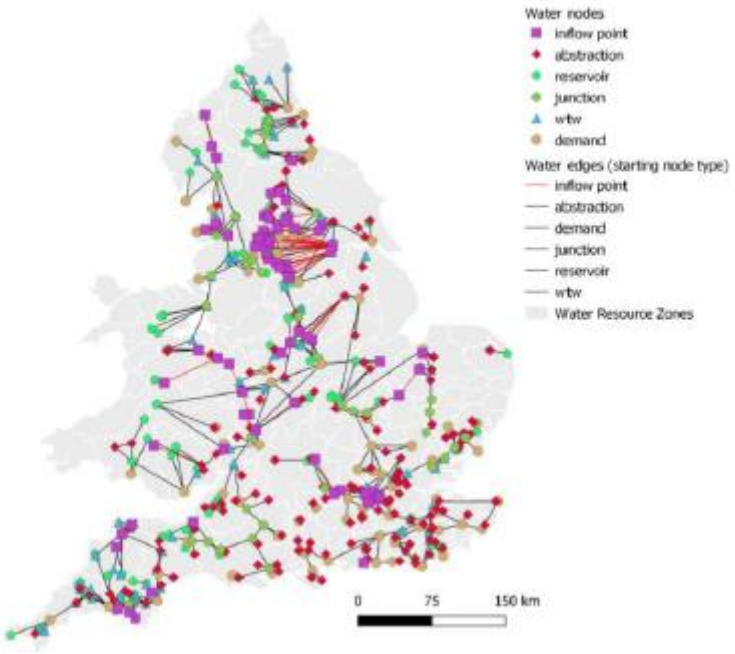

Figure 14 - water network represented within [44]

Notwithstanding that the majority of England's water comes from ground sources, compared to Scotland with far more significant surface freshwater resources, if there is insufficient representative data to formulate a model it is unclear as to the benefit of including such a model within the analysis when it could still be represented simply via point demands for water-related infrastructure (e.g. pumps). This could still communicate and imply impacts on the water networks without needing to model the networks themselves- especially if the robustness of the water networks is impossible to verify.

Further, not modelling capacity on those synthesised networks carries many of the same issues as those of power networks; water network capacities are limited by the pressure the pipes can survive and their physical size, therefore they are clearly not able to limitlessly supply any node which happens to be connected to a supply point, especially if localised drought conditions affect supply routes or water begins to freeze in pipes.

They also use an economic model to replicate the effects of power systems and related outages, while acknowledging the weaknesses of the economic model they use. Typically, in power system analysis, Value of Lost Load (VoLL) is used as a proxy for this, where a currency value is assigned to each MWh not "served" compared to expected values. This value is typically in the range of £17,000/MWh. Though analysing the impacts of disruption to other infrastructures will necessitate more involved economic analysis, there is already a standard value that can be used as a benchmark. The estimates of economic disruption provided within the report for various cascading outages, shown in Table 3 [44], seem low when compared to estimates of VoLL, and are unlikely to be realistic.

Table 3 - estimates of economic costs caused by infrastructure disruption [44]

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**

3.3 Summary of issues related to cascading outages

Cascading outages do not always emerge only from natural hazards. Events such as the power cuts of August 2019 [54] illustrate that concurrent outages with very little or no association with each other can combine to create wide-scale system disruption with very little exogenous perturbation to the power system itself. Further, this is also why transmission-scale systems and meshed networks cannot simply be analysed through topological searches, even if simpler networks can.

Combining disparate network models does not necessarily improve the overall standard of modelling conducted if the representation of those models is not appropriate for the task being undertaken but may make any analysis less than the sum of its parts. Any modelling conducted does not necessarily need to model the full dynamics of a perturbation scenario but does need to at the very least consider factors such as capacity constraints on the networks being modelled when they are as materially significant as they are on electrical power networks. Given features such as frequency response, supply/generation balance, generator availability, and protection action, which are not captured in simple topological searches, can directly contribute to large-scale power system disruption, if the aim is to model extreme events these must be considered in some way eventually.

Furthermore, as discussed in the previous section, in future networks weather will continue to impact not only the transmission networks, but also demand and generation through changes in temperature and wind/solar conditions. Therefore, cascading outage models must take consideration of these factors as they will directly affect the ability of the power system to respond to perturbations.

Increases in renewable generation reduces system inertia. Without suitable countermeasures by the system operator, such as use of fast dynamic frequency containment services, the ability of the power system to resist changes to frequency (measured in terms of Rate of Change of Frequency, or RoCoF) is reduced with increasing levels of wind generation [55]. A resilience enhancement approach proposed by the Oxford-led consortium in the modelling performed is to increase redundancy and provide backup storage – but these do not address many of the causes of large blackouts in and of themselves. If the cause of a blackout is an extreme windstorm damaging significant amounts of distribution or transmission infrastructure, this will also carry a risk of destroying any additional network infrastructure constructed, meaning there may be negligible improvements in resilience.

The next subsection provides recommendations and suggestions on the requirements of simulating cascading outages.

3.4 Suggested cascading outage-related case studies

A number of recommendations for modelling of cascading outages under different circumstances are summarised in

Table 4. It should be noted that the events which *cause* these cascading outage events may be weather driven, but if assumptions are made about generator setpoints, system inertia, frequency

response etc. the specific simulations used in isolation may be viewed as a “black box” depending on the simulation approach deployed.

Further, in a very broad sense, simulation types within power systems can be separated into Steady State, Quasi-Steady-State (QSS), and dynamic models. The last of these is sub-divided according to which kinds of transients are of interest: studies that consider electro-mechanical dynamics using models based on phasor representations of root mean squared voltages and currents (“RMS models”); or simulations of electro-magnetic transients, i.e. using “EMT models”. A QSS simulation could be, for example, a simplified simulation for a system frequency response (SFR) simulation as deployed in [53]. In each of these approaches, a different set of equations is solved to describe the physical phenomena of interest with the more complex models used only where judged to be necessary. Very often, the more temporal granularity is needed, the more limited is the practical spatial scope of a model[2]. .

In this sense the examples given here may be extensions of the scenarios described in

[2] A ”model” can generally be understood as the combination of the set equations to be solved and the parameters describing a particular physical system. Very often, in a power systems context, there are quite standardised sets of equations – for steady state, QSS, RMS or EMT studies – to be solved and commercial software to solve them. For many power systems engineers, a “model” is then a particular set of parameters describing a particular physical system under particular conditions.

Table 1, but the simulations of the electrical responses of the system could be cause-agnostic if the data fed into the simulations is corrected appropriately based on the principles decided here.

In this sense the examples given here may be extensions of the scenarios described in Table 1, but the simulations of the electrical responses of the system could be cause-agnostic if the data fed into the simulations is corrected appropriately based on the principles decided here.

Table 4 - example scenarios and requirements for analysis of event cascading from the electricity system to other infrastructures (cause-agnostic)

Scenario	**Electrical parameters necessary**	**Example incident (if present)**	**Affected infrastructure**	**Models needed (simulation and features)**	**Temporal range**
Frequency excursions (no network damage)	Inertia constants, dispatchable generation set points, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities, ramping rates	August 2019	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response, storage	Generator tripping models, demand response, storage, renewables capacities, SFR	Milliseconds - days
Outages including network damage	Inertia constants, dispatchable generation set points, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities, network topology and parameters, ramping rates	Storm Arwen	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response, storage, protection equipment	Generator tripping models, demand response, storage, renewables capacities, weather models, SFR	Milliseconds - days

Protection failures	Inertia constants, dispatchable generation, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities, network topology and parameters, protection equipment locations and settings, ramping rates	August 2019	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response, storage, protection equipment	RMS dynamic models and, potentially, EMT models, generator tripping models, demand response, storage, renewables capacities	Milliseconds - hours
Resource adequacy studies/ Integrated Resource Plans (long term stress tests)	Demand, storage, demand response, dispatchable generation, DER, renewable capacities	ESO FES	Dispatchable generation, solar generation, hydro generation, wind generation, storage, demand response	OPFs, UCs, demand, storage, renewables capacities, weather/climate models	Hours - years
Renewables integration studies	Demand, storage, demand response, dispatchable generation, DER, renewable capacities, frequency response, ramping rates	ESO FES	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response,	Demand, storage, renewables capacities, weather/climate models, SFR	Seconds-years

			storage, protection equipment		
System restoration	Inertia constants, dispatchable generation set points, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities, network topology and parameters, ramping rates,	N/A	Dispatchable generators, renewable generators, gas networks and demand, storage, TDI, protection equipment	Demand, storage, renewables capacities, weather models, potentially also load flow, models of network switching, RMS dynamic models, UCs, OPFs, potentially EMTs if interested in specific phenomena such as energisation	Hours - days

The wide temporal range of these studies illustrates the complexities associated with power system analysis of extreme events. Further, frequency variations are normal within the power system but can trigger cascading outages and severe system outages if they are not controlled, e.g. as occurred in August 2019 in GB [56]. Analysis of these events is complex as the response of individual generators and localised protection and inertia properties of systems can cause localised impacts of system-wide frequency-related phenomena. It is a significant challenge to cover every potential factor in an event which causes power system disruption with the same simulation package, which is why simulation packages used in power systems tend to be special purpose and tailored to specific use-cases, be that real-time simulations, EMT simulations, or simple load flow calculations.

The work in [53] investigated the problems associated with the wide range of temporal impacts by exporting data from a real-time dynamic simulation (RTDS) via an IP connection, and solving the frequency response and power flow calculation using simplified representations in the *python* programming language. The work in [57] linked multiple unit-commitment SCOPFs with a frequency response simulation and different types of optimal power flows to model largescale outages on a representation of the GB system caused by an extreme wind storm.

In that sense, linking different models together to capture different events is possible, but it can be a complex and time-consuming effort to ensure consistency of results. This also complicates how outcomes are measured and quantified, particularly if there are disparities across the system as to how long it takes customers to restore, and to what level interdependent systems are modelled. For example, system restoration following an extreme windstorm alone compared to system restoration following extreme blizzards and the subsequent snow melts will be very different. If high wind penetration on a system means lower inertia, and a series of cascading outages leads to a system blackout, the control responses and frequency response of the system will need to be simulated to fully capture the eventual consequences for the system as this will not be captured adequately by just using an OPF, even if that OPF includes load curtailment.

Features such as Low Frequency Demand Disconnection (LFDD), also known as Under Frequency Load Shedding (UFLS), must either be captured directly by performing an appropriate system frequency response (SFR) simulation or through appropriate proxies within the optimal power flow used to model system behaviour [58]. This is particularly consequential in GB as falling system inertia makes the system increasingly vulnerable to high Rate of Change of Frequency (RoCoF) events [59].

Generators which trip due to frequency excursions will typically have to re-synchronise with the system and recover after a period of time, and in the event of a total blackout system restorations becomes even more complex and Black Start or system restoration procedures will have to be enacted. Whether or not these matters are considered should be decided before any modelling is conducted due to the complexities associated with modelling system restoration – it is possible, for example, to model restoration just via an OPF for each step and presume that the system maintains stability, but features such as unit commitment also come into effect once the simulation extends into the range of hours, as generators will be scheduled to come in and out of service and weather patterns will affect both customers and generators, as well as, potentially, the integrity of the transmission system and the capacity to get power where it needs to go.

Finally, it is important to remember in any complex power system simulation the concepts of *aleatory* and *epistemic* uncertainty. That is, respectively, the uncertainty associated with the abstractions used in the development of any model used, and the uncertainties associated with the inherent, fundamental randomness of the world.

To summarize, any modelling of weather events that involves power systems should:

- Clearly define the boundaries of the simulation problem spatially and temporally so that the types of simulation to be used can be chosen appropriately

- Define clearly interactions with other infrastructures so that either the effects on those systems can be modelled appropriately or such that potential challenges can at the very least be indicated
- Only model systems and events you have confidence in and which are based on robust, appropriate datasets and models
- Not try to model "too much"; overcomplexity can obfuscate model results and undermine any findings by introducing potentially unnecessary error.

In addition, it must be decided what approach to the modelling should be taken, and whether that is:

- **Stochastic**: based on performing many simulations and resolving fully each scenario in situ, before averaging the results, using datasets representative of a presumed scenario or event (i.e. Markov Chain Monte Carlo simulation, linked with fragility curves). That is, for *n* samples, at least *n* simulations will be performed, dependent on how many types of simulations are performed in each sample (e.g., if only a load flow is performed for every sample, then *n* loadflows will be performed, but if a loadflow and a frequency response is performed for every sample, *2n* simulations will be performed, meaning simulation cost increases linearly for sample size).
- **Scenario-based**: predetermined weather and outage events are chosen and performed by the simulation framework to calculate outcomes; these events would have to be chosen by expert elicitation or pre-established conventions; an example of this could be N-1 simulations

Thus far, the focus has been on natural hazards and their impacts on the power system, but another threat comes from intentional or malicious attacks on the system. These are the subject of discussion in the next section.

4 Cybersecurity-related Events

Compared to natural hazard modelling and power system outages more generally, there is relatively little real-world evidence of disruption to power systems associated with intentional, malicious attacks to power system integrity- at least on a comparable scale. Two events are noteworthy in recent years both because of the scale of damage they incurred when they were successful, and because of how unusual it was for them to succeed. These attacks shall be discussed here as well as more general cybersecurity and cyber-physical system security and how it may prove to be relevant to the research undertaken in this project.

4.1 Example attacks

Two attacks in recent years are of particular significance; the *Stuxnet* attacks on Iranian Uranium ore enrichment facilities, and the extensive cyberattacks conducted against Ukraine's electrical power systems in 2015. Even more recently it has been claimed there has been a cyberattack exposed at Sellafield, but details are still emerging and specifics are not yet known, so a review cannot yet be conducted [60]. Sellafield disputes it has been attacked at all, but it is claimed that adversaries have gained access to monitoring systems on the site itself.

4.1.1 Ukraine and Black Energy 3

A review in [61] is conducted to understand the implications of the Ukraine cyberattack on bulk power systems. This attack affected 225,000. The malware "Black Energy 3" was used to steal VPN (virtual private network) credentials, in turn allowing access to privileged systems. Afterwards, KillDIsk firmware was used to wipe backup hard-drives to disrupt and delay system restoration.

State estimation was disrupted using False Data Injections, flooding workstations with inaccurate information to further disrupt operations.

Such attacks can take a significant amount of planning and require in-depth knowledge of the system under attack. For example, they rely on the attackers:

- Having knowledge of victims' systems
- Being able to manipulate meter measurements
- Having knowledge of control and operations
    - Network topology
    - Electrical parameters
    - Data detection schemes and mitigations
    - SCADA device specifications

Using this information and access to control systems, attackers could perform actions such as:

- Maliciously disrupting measurements
- Switching circuit breakers
- Disrupt communications infrastructure

There are different approaches to maliciously altering data packets, or adversarial attacks, which will depend on the type of attack being undertaken. Some examples are:

- Disrupting meters at a local level or physical disruption
- Intercepting and altering data packets at a communications protocol level
- Forging data packets
- Altering data within the control centre itself

In the context of the current research project being undertaken, we are less interested in the specific mechanisms by which an attacker might be able to achieve access to these systems, rather what they might do once they do.

The probability of success affects any aggregated risk assessment of such cyberattacks, but it is unlikely to be within the scope of the project to be able to realistically quantify this risk – probability could however be qualified or assumed e.g. through a weighting from 0-5 based on expert elicitation which could in turn be used to determine risk. It still might be worth evaluating the extent of the impact following such events and the modelling approach followed by the project could be able to potentially play out such scenarios for impact assessment.

4.1.2 Iran and Stuxnet

Stuxnet was an attack on nuclear-related infrastructure in Iran linked to Israel as the source [62]. This report used a systems analysis report to analyse the vulnerabilities which contributed to the attack being possible as well as providing an analytical method for how to analyse similar events more generally.

To summarize, Stuxnet was discovered by “VirusBlockAda” in June 2010. It primarily affected computers in Iran but did spread to systems across the globe. The attack process is illustrated in the following figure.

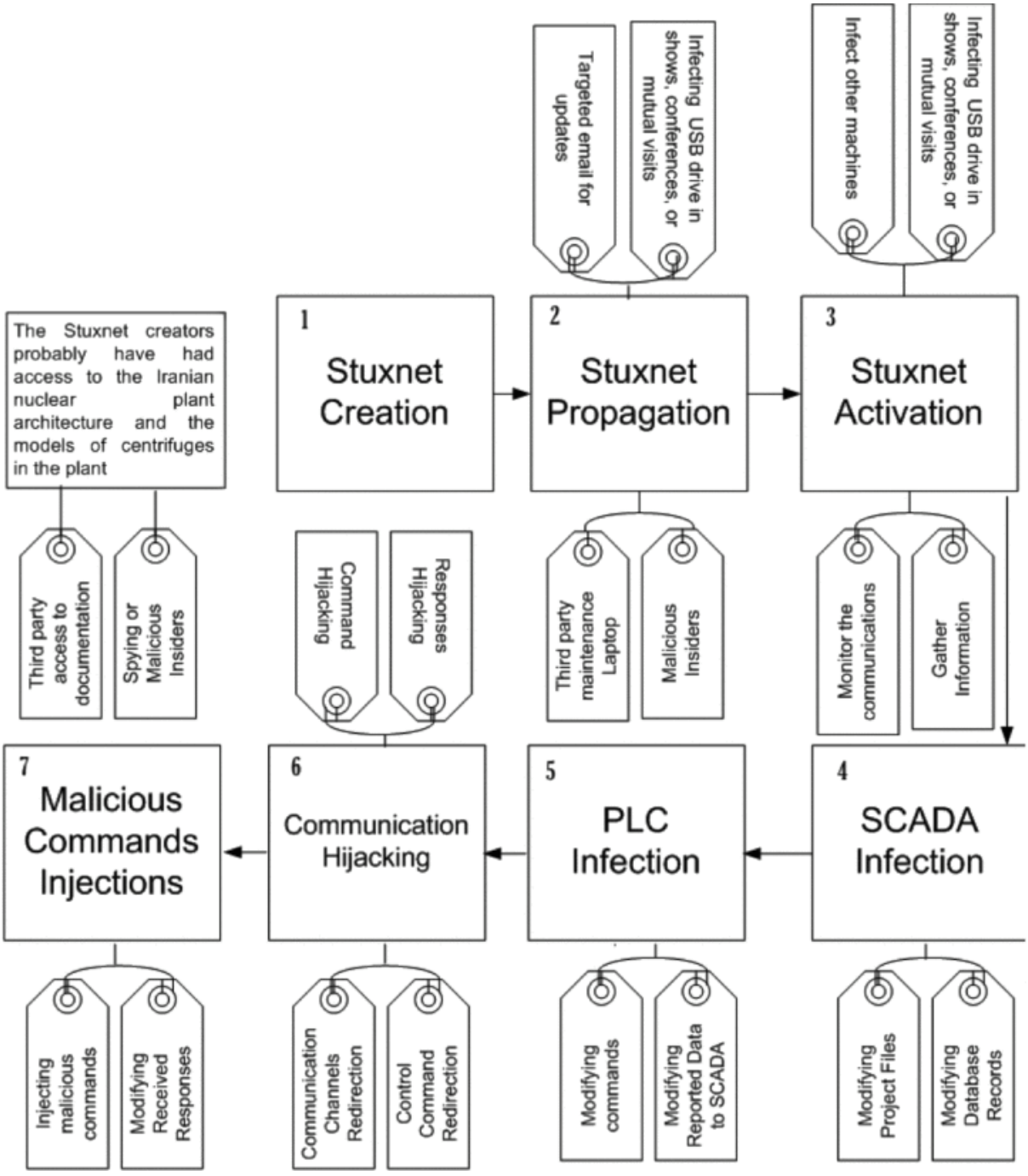

Figure 15 – Studxnet attack process, described in [62]

The attack itself was likely tested in isolation to prevent impacts on other infrastructures (i.e. those within the attacking state). It was aimed at program logic controllers (PLCs) used for controlling uranium centrifuges, which itself would have been challenging as each PLC would have a unique configuration. This would have required access to manufacturers, contractors, or snooping malware.

Given such infrastructure was likely to be airgapped, it is also probably that the malware was physically introduced to the system, e.g. by an infected USB device. Once the infection was within the system it was able to self-proliferate and affected other systems, such as SCADA.

It utilised known flaws in the Windows operating system to proliferate – using legacy operating systems is not uncommon in infrastructures as the flaws are well known and can be designed around, but the onus is still to ensure those vulnerabilities are addressed before they can be exploited.

Stuxnet lurked on systems and mimicked normal operation, first by recording what “normal” operation looked like, then replicating normal command messages and sending them to appear legitimate. Registers of infected devices were kept to monitor the proliferation of the malware. Eventually, these access routes were used to attack the system. The primary targets were SCADA, web-servers, network adapters, central data repositories, and database servers. Legitimate nodes on the system were subsequently replaced with illegitimate ones to aid in seizing control of systems. The paper itself then proposes a methodology to identify specific threats to the system, but this is outside of the scope of this report.

Both these attacks have in common that they targeted not the physical components of the power system directly, but the mechanisms which controlled it – the “cyber” component of the “cyber-physical system” (CPS). In [62] it is argued that the main components of a CPS are:

- SCADA
    - Used to gather and control geographically dispersed assets
- Distributed Control Systems (DCS)
    - Controllers which are grouped together to carry out specific tasks at a specific location
- PLC
    - Device-level controllers used by DCS and SCADA

Security in power systems initially emerged from air-gaps and a lack of digitization, and centralisation of control. “Smart grid” concepts and increasing digitisation undermine this.

Attacks can undermine systems not just by attacking components themselves but by attacking communications and interactions between components – e.g. delays, inaccurate information, distributed denial-of-service (DDoS) attacks.

Diversity of devices means that homogenous security profiles are unlikely to work, especially if all devices on a system are “trusted”. More devices which communicate with parts of the system mean a larger attack surface. Every trusted node is a potential attack vector, thus the attack surface increases with connectivity and scale. As we are moving to systems with much more software and control and much higher number of controllable devices by an increasing number of entities (e.g. not just the system operator but aggregators or even individuals), this can also manifest as an increasing risk. This is an aspect that is worth considering as we are potentially moving towards systems with more software and control of a larger number of devices (e.g. distributed energy resources, batteries, EVs) and potentially by a larger number of actors (e.g. aggregators or even individual consumers).

Prioritisation of securing “obvious” targets like SCADA can overshadow security of lower-level devices which can still cause harm to the wider system. Different approaches are identified to model safety analysis of CPS:

- Fault tree analysis
- Failure mode and effects analysis
- Hazard analysis and critical control points
- Hazard and operability study

Risk assessments can then be performed to identify hazards, that is any event or situation with a potential to cause damage. Hazards can be related to

- Actions
    - Undesirable actions (e.g. tripping of devices, incorrect setpoints, etc.)
    - Desirable actions not taken (e.g. failed or partial delivery of an expected service)
- Timing
    - Desirable actions taken too soon or too late

- Sequence
    - Desired action in sequence of actions skipped, or actions in a sequence performed out of order
- Amounts
    - Desired action performed too much or too little

Quantifying the risk can then be performed via methods already discussed within this report, which in turn can be used to design safety and resilience measures.

Systems can be “safe” but not “secure” – parameters can be altered to create adverse operational conditions that technically speaking are still within the “safe” range, but might cause uneconomical operation. Cyberattacks can also use this principle to malicious ends by manipulating the system to still operate within operational bounds but might not be secure against further disturbances. For example, a system might be driven to operate in a condition which is not N-1 secure or according to SQSS. This could potentially be envisaged to be exploited even further to deliberately cause instability, for example through enforcing disturbances that will drive the system to an unstable operating condition, although this would potentially require much more effort and access to critical systems (see example of the Ukraine cyberattack). [62];

- Failure mode and effects analysis (FMEA):
    - Identify individual failure models of a system/its components and effects on reliability
    - Performed at start of development phase, after design
    - Uses “risk priority number” (RPN) in quantitative analysis to identify “reliability rates” for each failure mode. This could potentially be useful in prioritising scenarios to further investigate. Deciding how to prioritise those scenarios is an of itself an additional challenge
    - RPN = severity x probability x detection ranking
    - Unclear what “detection ranking” is
    - Can be useful for single points of failure but some failures can occur even if all components are operating normally
    - Fails to consider combinations of failures

It is argued these “traditional” methodologies view safety as reliability issue – failures being a result of linear chain of undesired events. CPS security threats however can occur without any evident failures. E.g. an attacker could modify settings within a “safe” range but still cause disruption.

A method, labelled STAMP was proposed – i.e. Systems Theoretical Accident Model and Process. It is designed to consider interactions among components as opponents to design safe systems.

- Looks at systems as dynamic systems rather than static
- Consider safety and security as control issues rather than reliability issue
- Components within a system require control through constraints, inadequate enforcement leading to failures
- Models how these systems interact to determine a “safe” and “secure” state

From this, a means of investigating failures retrospectively is proposed, “Causal Analysis based on STAMP” or CAST. It uses the following procedure:

Defines a means of investigating failures – “Causal Analysis based on STAMP” (CAST)

- Define hazards
- Find safety constraints and requirements
- Define control structure
- Find possible events causing failure or accident

- Navigate through system control structure to find vulnerabilities and how they can cause failure, unsafe states
- Analyse interactions and finding potential factors that can lead to failure
- Finding external and dynamic factors that can affect overall safety structure
- Produce design recommendations for improvements to safety design

This could be a useful approach for designing events for analysis more generally from a “Bottom Up” approach, as opposed to a top-down approach of inducing a causal event and investigating its impacts. However, given the significant number of ways in which power system collapse can occur it is likely not possible to produce a complete set of pathways within the scope of this project, and is likely a significant body of research in its own right. Furthermore, it should be noted that such methods might end up requiring large computational effort if detailed models of energy systems are needed to be employed.

4.2 Review of general cyber-physical security principles

A systemic literature review of cybersecurity and cyber-physical system resilience was undertaken in [63]. In this case it understand cyber-physical systems (CPS) as “architectures that incorporate digital, [analogue], and physical components”. Cyber-physical energy systems (CPES) are those explicitly related to the delivery of energy to customers.

It is important to demarcate the digital (“cyber”) and physical components of the system because these present as different attack surfaces for adversaries, though the line is not always obvious. The power system itself is an enormous and complex system integrating physical machinery (generators, transmission), monitoring and control (Supervisory Control and Data Acquisition – SCADA), and digital components. These present multiple attack avenues for adversaries. However, there is little practical experience of hackers or malicious attackers successfully deploying widescale attacks against infrastructure in states such as the UK – which is not to say they have never happened, with events such as the WannaCry ransomware attack affecting companies of all kinds globally. Most vulnerabilities on the power system, however, are already known – a figure is quoted suggesting that “99% of vulnerabilities exploited in 2020 were known to security officials”. More recent events such as the reporting of the Sellafield attack suggest there may also be attacks that have occurred or are in progress which we simply do not yet know about.

There are multiple different labs utilising different approaches specifically designed to investigate cyber-physical threats to systems, outlain in Table 5 [63].

Table 5 - Cyber-physical testbed architectures, accuracy, repeatability, cost characteristics, and example testbeds with their simulation resources. [63]

Testbed Architecture	Accuracy	Repeatability	Cost	Example Testbed	Resources*	Year
Hardware Assisted	High	Low	High	INL	PS: *RTDS* NS: not applicable	2015
				NREL	PS: *RTDS*, *Opal-RT* NS: not applicable	2015
Software Assisted	Medium	High	Low	Texas A&M	PS: *RTDS* NS: *OPNET*	2014
				TU Dortmund	PS: *Opal-RT* NS: *OPNET*	2014
Hybrid: Physical Hardware & Simulated Cyber	Medium	Medium	Medium	FSU-CAPS	PS: *RTDS*, *Opal-RT* NS: *OPNET*, *EXataCPS*	2013
				PNNL	PS: *Opal-RT* NS:*OPNET*, *ns-2*, *ns-3*	2017
Hybrid: Simulated Hardware & Physical Cyber	Medium	Medium	Low	HELICS	PS: *RTDS*, *Opal-RT* NS: *OPNET*, *OMNeT*	2017

*PS, and NS stand for the corresponding CPS tetbeds' power and network simulators.

What should be immediately observed is that these all entail real-time dynamic simulations and network model simulations. This therefore places the simulations or replications outside of the scope of this research project as it is unlikely that a single solution can be provided that incorporates extreme weather simulations with real-time simulations of this nature, though this depends on what types of attack the users actually wish to investigate. It could still be worthwhile to "play out" scenarios which could be caused by cyber-attack related failures, however.

It is of course desirable to be able to model such attacks in testbeds such as these safely without risking a lab-based experiment or cyberattack leaching into the MITS, though a consequence of this is that then the realism of such experiments has to be carefully considered given the difference between any lab-based work and a real physical system. It should also be noted that this is essentially PHiL testing, which the Power Networks Demonstration Centre, a research centre affiliated with Strathclyde, conducts.

Four different categories of cyberattack studies are proposed as general concepts, illustrated in Table 6 [63]. This is purely to illustrate the range and scale of different types of study conducted.

Table 6 - CPES security study categories and research examples. [63]

Security Study Category	Example Literature
Attacks exploiting CPES vulnerabilities	Low-budget GPS spoofing attacks
	Load redistribution attacks
	Coordinated DoS attacks
	Data integrity attacks on state estimation
	FDIAs with limited resources
	Hall sensor spoofing attack
Evaluation of attack impact on CPES	Graph-theoretic quantitative security assessment
	Unsupervised learning-based evaluation
	Simulation-based impact evaluation
	Deep reinforcement learning analysis
	Stochastic modeling assessment
	Quantitative attack impact evaluation on substations
	Markov-process based reliability analysis
Attack detection algorithms in CPES	Unsupervised learning-based FDIA detection
	Wavelet transformation-based FDIAs in AC state estimation
	Historical data-assisted FDIA detection
	Distributed collaborative FDIA detector
	Machine-learning assisted FDIA detection
	Unsupervised learning anomaly detector
	Sensor- and process noise-based attack detection
	Autoencoder-based anomaly detection
Attack mitigations and defenses in CPES	Semi-supervised method for malware
	Markov-process based reliability analysis
	Data-driven and compressive sensing resilient state estimator
	Battery-based hardware security authentication
	Battery-power assisted risk mitigation
	Robust control-based defense mechanism
	Control flow integrity validation method
	Hardware security-based communication protocol extension

The specific nature of each of these studies is provided in the reference within the report and will not be repeated here. The key point to note is that it is clearly an area of intense ongoing study, and that there are a significant range of studies being conducted. What should concern us is whether 1) it is appropriate to replicate these studies within the modelling framework of “extreme events” or 2) whether it is practical and reasonable to do so.

Typically cyberattacks, within the cases described within the review, as modelling an “adversary” and a “defender” or “red team” and “blue team”. This requires modelling of both the objective function of the attacker and the defender who will have opposing objectives; one to inflict harm and the other to prevent it. The capability of this attacker will depend on the information it has on the system under attack and the resources they have to exploit it.

The impacts of the attack will also depend on what section of the system is attacked. To these ends, the following categorizations are offered. Firstly, the categorization of assets under attack, described in Table 7 [63].

Table 7 - ICS functional levels, equipment categories, and their corresponding components [63]

Functional Level	Category	Components
Level 2	Supervisory Equipment	Supervisory control functions, site monitoring, local displays
Level 1	Control Equipment	Protection devices, local control devices
Level 0	Process Control	Sensors, actuators

Attackers are classed as Class I and Class II. Class I attackers are described as those who lack the resources or ability to attack without being detected, whereas Class II attackers can be organized crime-related organisations, or state actors.

Correspondingly, attacks are also categorized as:

- **Level 0 –** attacks on CPS processes and operational equipment (e.g. sensors, actuators)
- **Level 1 –** attacks on the control network (e.g. PLCs, controllers)
- **Level 2 –** attacks that target SCADA, monitoring devices

A threat model is proposed in the following Figure 16 as a general conceptualization [63].

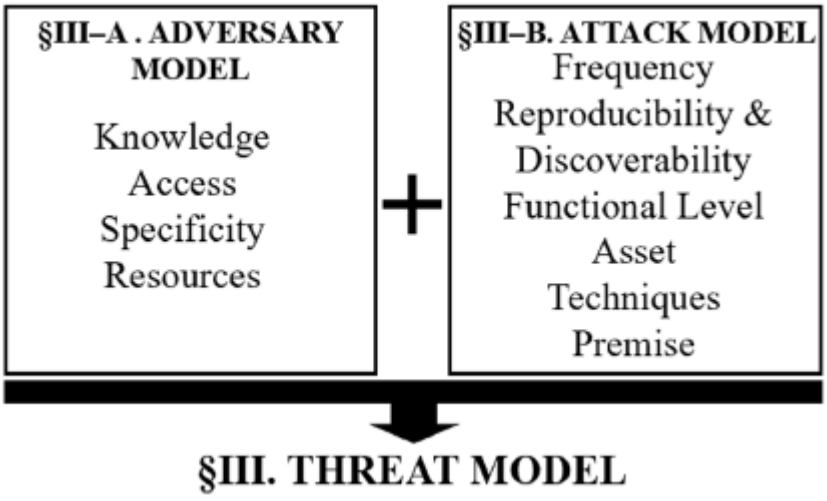

Figure 16 - Adversary model and attack model components comprising the comprehensive threat model architecture [63].

The relevance of these categorizations is whether they can be replicated appropriately within an Extreme Event case study to an acceptable degree within the time frames required in the context of the other modelling work being considered, and whether such simulations add any useful information that can be acted upon to improve resilience.

It is noteworthy that within the review, quantification of the probabilities of different attacks is not attempted, but rather qualifications of probability are provided (from Low to High), these are assigned a numerical value and used to compute a “damage score” to relatively rate different attacks on the system. This sidesteps the need for quantifying attack probability which, in this case, is almost impossible to do, but instead allows for expert judgment.

It is useful, therefore, to examine the case studies actually performed within the review as to whether they are likely to be necessary to emulate within the simulation platform being proposed. Four are proposed, with the following parameters as shown in Table 8 [63]:

Table 8 - Threat model of the attack case studies [63].

Threat Model	**Cross-layer Firmware Attacks**	**Load-changing Attacks**	**Time-delay Attacks**	**Propagating Attacks in Integrated T&D CPES**
Knowledge	Oblivious	Limited or Oblivious	Oblivious	Strong
Access	Possession	Non-possession	Non-possession	Possession
Specificity	Non-targeted	Targeted	Targeted	Targeted
Resources	Class I or II	Class II	Class I or II	Class II
Frequency	Iterative	Iterative	Iterative	Non-Iterative
Reproducibility	Multiple-time	Multiple-times	Multiple-time	One-time
Functional Level	L1	L1 or L2	L1	L2
Asset	PLC	PLC or HMI	Control Server	Engineering workstation
Technique	Modify control logic	Modify control logic or wireless compromise	Wireless compromise, MitM, Spoofing, DoS	Engineering workstation compromise
Premise	Physical: Invasive, or non-invasive Cyber: Asset control commands	Cyber: Communications and protocols or Asset control commands	Cyber: Communications and protocols	Cyber: Asset control commands

The details and results of these studies can be found within the report itself, but the major findings are useful to note. For example, in the Load-changing attack case, despite changing the attack on a load point by as much as 50% in a microgrid on multiple buses, this only causes frequency swings of ~0.2Hz positively and negatively from a 60Hz nominal frequency. This could be manageable in standard operational cases with adequate frequency response provision. In addition, there is usually some awareness of network devices and additional capacity in the system, and such an attack would necessitate coordination and access to a number of devices (or entities/actors) which might be challenging to achieve. However, if an adversary has the level of technology and capacity to carry out an attack of this magnitude, it’s also possible that they could attack infrastructure such as EVs or flexible demand. Further, seizing control of e.g. an aggregator being relied upon to deliver critical services could cause severe system disruption should an ordinary outage occur at a time of acute need.

A directed attack at a power converter maximum point power tracking (MPPT) component demonstrates much more significant impacts on the system frequency (nearly a 1Hz drop and significant instability in associated system parameters), however this relies on a direct attack on the firmware of a controller. This may require physical access to such a component which could be difficult to achieve without detection (although not necessarily impossible). In such cases, to avoid system instability, a mitigation could be to disconnect the affected device and rely on N-1 security design of the system to withstand such an attack, as long as the attack is limited to a single plant. In addition, a straightforward mitigation in such cases could also be physical security limiting access to the PLC/controllers or alerts if the firmware itself is modified. Modifying the firmware of devices is of course also a risk in its own right and could introduce unforeseen consequences across multiple sites if there are errors or unwanted features of the code – a faulty software update could have indistinguishable consequences to those of an intentional attack, after all.

A far more concerning modelling result emerges from the Time-delay attack. In this case, a relatively oblivious attacker delays a load shedding command to a microgrid controller, causing a severe frequency excursion before remedial action can taken, potentially leading to a system collapse. The scenario itself postulates that a microgrid (MG) disconnects from a main grid through intentional islanding, and to mitigate the loss of infeed a simultaneous signal is sent to perform load shedding. The time delay attack delays the sending of the load shedding signal, and thus frequency collapses.

Such a scenario is not necessarily directly transferable in a GB context. Firstly, intentional islanding is not a common action taken. Secondly, intentionally islanding a section of network and issuing a command to perform remedial action to sustain that then-islanded grid is not necessarily a common design approach for such a system to take because it is vulnerable to precisely the attack listed here; a reasonable and straightforward preventive measure for this type of attack could be as straightforward as ensuring the grid connection is pushed to float *before* islanding occurs so that both sides of the point of coupling remains stable after separation.

Further, LFDD arrays at 11kV or below would perform this role automatically in the event of a major frequency excursion. The attack itself, should this sequence of design features actually be implemented, is of course relatively straightforward because an attacker does not need to know anything about the topology or parameters of the system under attack, it simply needs to congest the network.

There has been research into the impacts of frequency response and the timeliness of responding to frequency excursions due to the impact of high RoCoF scenarios [64] for precisely these reasons, because the speed of frequency response is an essential factor of its efficacy and if that can be delayed then it naturally follows that system disruption is a consequence.

In noting this, of course, even an attack which simply triggered the LFDD could still be considered a success as this would still, for a period of time, cause significant disruption to the power system.

The final case study investigates the consequences of a more direct attack, manipulating circuit breakers (CBs) maliciously to affect power system stability. This does demonstrate more severe consequences and there are real life examples of this kind of attack in action (e.g. the Ukrainian cyberattacks), however they require a significant amount of knowledge about both the physical system (in order to access areas which can significantly impact the system) and of the control systems involved (to be able to manipulate those systems). The demonstrator system used, however, is of a modified IEEE 9-Bus network model with an expanded node to represent a distribution system. While these cases are useful to study to demonstrate the kinds of threats the system could potentially face, the complexity of actual MITS might require more detailed and realistic representations which might be challenging.

4.3 Summary of issues related to cyber-physical attacks

Unlike incidents related to natural hazards, the scale of data available to model and reproduce CPES attacks is particularly limited, at least when it pertains to the scale of event which we may deem to be an “extreme” event. Therefore, determining the probabilities of these events is particularly challenging.

Less challenging is determining the impacts – at a certain point, whether an action has been taken through malice, unfortunate circumstance, or incompetence does not matter if the electrical components on the system perceive it the same way. Which is to say, load suddenly dropping on the system will cause a frequency excursion whether that is due to a transformer failing or whether a hacker has sent out a signal to an aggregator to suddenly disconnect a certain amount of demand at an inopportune moment.

If CPES attacks are to be modelled, therefore, it should be based on a qualified assessment of what the most likely dangers are to be via modelling their impacts on the power system, e.g. rapid variations in demand on the system, malicious control of power converters, etc. Other potential impacts of malicious attacks, e.g. intentional destruction of assets, can be modelled using identical methods relating to natural hazards, only the attack surfaces and the probabilities are different. For example, if a storm hits the power system, we can predict where that storm will impact the power system based on weather forecasts. If an attacker wishes to attack SCADA systems, their access will instead be determined by factors such as ownership of those assets, manufacturer, type, and network design. Ironically, if a lack of modelling of this has been conducted because of the complexity of attaining this information, it also means that a CPES attack utilising this information is unlikely, through security via obscurity, with the diversity and obscurity of systems unintentionally hedging against the ability of attackers to utilise a consistent attack pattern against them.

This is not to suggest that such attacks should not be studied or mitigated against or to foment complacency; CPES attacks are clearly a potential for major disruption across the power system. Even attacks which do not cause blackouts could cause damage to system assets or incur significant costs through causing uneconomical market behaviours. Nonetheless, further study is almost certainly needed to determine what data and modelling are needed in order to incorporate such modelling into the framework satisfactorily.

There are also competing factors in terms of the types of devices which may be attacked and the effects of diversity versus standardisation. With standardisation, weaknesses could be well understood ahead of attacks and planned around, but this also means that if an attack does succeed it increases the probability of a common mode failure affecting all devices of that kind. It is not known within the review to what extent standardization could be used to mitigate or exacerbate such threats and further research into this should be conducted if possible but it may be outside of the scope of this project.

Conversely, introducing too many devices from different operators, standards, and manufacturers can make maintenance of these systems overcomplicated, introduce additional costs, and increases the probability of stranded and abandoned software and components with companies no longer maintaining devices either through commercial decisions or from going out of business.

As noted, at a certain point whether an attack is deliberate or a result of human error or poor design is indistinguishable to the power system itself, but will affect the restoration and recovery of the system from that incident and the attack surface available. Therefore, it is not unreasonable to assume that modelling approaches used for standard perturbation scenarios should still, in many ways, be appropriate here.

4.4 Suggested cybersecurity-related case studies

Some example cyberattacks which could be modelled are suggested, but it is likely significantly more information will be needed to determine the probability of these attacks. Modelling e.g. the communications network facilitating these outages will almost certainly be out of scope at this stage, but replicating the potential impact of these attacks by e.g. adversely changing controllers or demand at nodes could be used to emulate the consequences of those attacks. These are described in Table 9. It is notable that these generally are very similar to the cascading event simulations already suggested and analysis at a detailed level will likely involve dynamic simulations.

Table 9 - suggested cyber-resilience related attacks

Scenario	*Electrical parameters necessary*	*Example incident (if present)*	*Affected infrastructure*	*Models needed (simulation and features)*	*Temporal range*
Malicious manipulation of demand and network topology	Inertia constants, dispatchable generation set points, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities	Ukraine cyberattacks	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response, storage, communications networks	EMTs, generator tripping models, demand response, storage, renewables capacities, RMS modelling, controller models, renewable generation	Milliseconds - days
Malicious manipulation of power converter controllers (e.g. changing setpoints and measurements)	Inertia constants, dispatchable generation set points, primary, secondary, dynamic containment, frequency response capacities, demand, storage capacities, demand response capacities, DER, renewable capacities, controller models for power converters	N/a	Dispatchable generators, transmission infrastructure, distributed generation, gas networks and demand, demand, demand response, storage, communications networks, renewable generation	EMTs, generator tripping models, demand response, storage, renewables capacities, RMS modelling, controller models, renewable generation	Milliseconds-days

5 Conclusion and Recommendations

This report has attempted to understand the various issues arising from modelling "extreme events" through examining recent, directly related literature on the matter. To this end, various issues have become evident relating not only to how to classify extreme events, but also how to simulate them and understand them in the context of wider power system reliability and resilience.

5.1 Defining an "extreme event"

Typical definitions of resilience tend to relate to HILP as a categorization of resilience-related extreme events. That is, related to both the probability and impact of those events. A heuristic for this could be to view anything N-1 related as "reliability" focussed, and anything more extreme than that as "resilience" oriented. However, an event can have major consequences even if nothing on the system breaks, for example in an extended heatwave, during a *dunkelflaute,* or during a cold snap.

Weather-specific categorizations of "extremity" should be left to the meteorologists, as suggested by ESIG, because that is their domain of expertise. The duty of translating these weather events into power system consequences should be charged to those with expertise in modelling the power system itself with understanding of its working.

A suggested approach for defining an "extreme" power system event is

- **"any event that, without suitable mitigating actions, would cause, as a result of conditions arising from that event:**
    - **Interruptions to a large number of end users' supply of energy, beyond those that could be expected due an outage of any single item of energy system plant;**
    - **extraordinary energy market conditions, or**
    - **interruptions of energy supply to significant elements of critical national infrastructure "**

This definition could apply to any N-1-k scenario, to interconnected outages across systems, to a heatwave which forces derating of power system assets to maintain stable operation, or even to a cyberattack which results in market operators having to change dispatch because suboptimal operation is induced by cyberattacks. It is also agnostic to probability, which would allow expert elicitation to address extreme events of which the probability is difficult or impossible to quantify and does not define explicitly a time horizon to allow longer term events to be considered.

5.2 Data concerns and model priorities

Access to data, be that weather or electrical power system, is an endemic challenge for modelling power system resilience. There is a conflicting challenge between the need to ensure model accuracy and robust representation of weather-related phenomena, without models that imply false realism and modelling bloat if the parametric relationships used within any model are not statistically significant or robust.

All models should be "as simple as they can be, but never too simple" [58], and the specific requirements for test system parameters has long been a subject of debate and discussion [65]. A major weakness of the Oxford/NIC work was attempting to do too much all at the same time to the extent it weakened significantly the ultimate findings. The nature of the modelling conducted will depend not only on what data is available but also on what questions we want to have answered by the modelling which is conducted – are we looking to mitigate outage times, mitigate aggregated scale of customer disruption, or prevent incidents altogether?

It is highly unlikely that it is possible for a single model to be able to capture all of the features discussed in this report (e.g. from frequency response to water demand), though a multi-agent model that considers different features and combines their inputs into an aggregated form, with qualified assessments of various model inputs, could be a reasonable approximation. These models do not all have to be developed at the same time by the same people. Different models can be 'soft-linked' given definitions of suitable interfaces. A modular model which allows for additional modules to be added to improve realism (e.g. to correct restoration times on assets if a transportation model is added) could be one avenue of enabling ever more realistic modelling of extreme events.

Reanalysis weather data, such as from ERA-5, should be used, as per the recommendations in the ESIG report, but this might not capture the effects of climate change. Different climate models might be used with expert advice from meteorologists.

In stochastic models it is important that accurate data are used for failure and, depending on the nature of the study, repair rates. Because failure rates, especially for overhead lines, change by many orders of magnitudes between 'normal' and 'adverse; conditions, these should be appropriate for the weather conditions that are being modelled. Utility statistics gathered over many years for large populations of key asset types should be used wherever possible. If fragility curves for overhead lines are to be used, the data sources linked in [20] can be referenced. In the absence of anything better, data from other countries might be used, such as from TADS (Transmission Availability Data System) [41], albeit with clearly presented caveats. As a last resort, failure rates from sources such as the 1996-Reliability Test System might be used as stand-in sources [40].

For the immediate future, however, the focus should be on modelling or representing the fundamentals of power system resilience studies. That is:

- Topological (node-branch) models with capture of connections to other critical infrastructure
- Power flows
- Frequency response
- Renewable generation projections
- Demand projection
- LFDD regimes
- Basic restoration modelling

These represent the most versatile and immediately useful models and allow for a wide range of modelling to be performed. It ought to be relatively straightforward to obtain representative models from project partners. In their absence and for the purpose of testing prototype modelling approaches, IEEE standard networks might be used as placeholders until those models are made available in a limited fashion. A next step, to model more specific phenomena, could be to incorporate features which include:

- Electro-mechanical dynamics in "RMS" simulations, including controller models
- Weather modelling
- Dependency of asset failure probability on weather conditions such as represented by fragility curves
- Markets
- Distributed generation
- Demand response and load prioritisation

More detailed modelling could then incorporate:

- In order to fully assess the impact of behaviour of inverter-based resources (such as wind farms and HVDC interconnections and embedded links) and their control, "EMT" simulations.
- Communications networks
- Transportation networks
- Water networks
- Bus-breaker representation
- Optimised restoration

The first group of models and data would allow basic studies to be conducted and could be cause-agnostic so could cover a wide array of scenarios. Projections of e.g. wind power can be based on historical data as a placeholder and later updated. Further refinements then allow more complex modelling to be performed with more detailed scenarios. Complex behaviours that can render models of large systems or interconnected systems unwieldy or impractical may need to abstractions that are broadly representative of key behaviours, with the caveat that there is some degree of 'state of knowledge' uncertainty, i.e. epistemic uncertainty.

Some research has been conducted into modelling restoration times subject to extreme weather and restoring power systems subject to storms, such as in [43] and [66], so abstractions are possible to improve restoration

time modelling. If data or models are not available, it is still worthwhile to categorise what these data and models are to direct future research or to describe it in qualitative terms. The ESO and academia have a wide range of tools at their disposal, but these are not universal; a clear categorisation and inventory of such tools would be of great use.

The key priority should be facilitating models which can represent a range of initial conditions of the system, and simulate the system being perturbed, capturing the degree to which adverse impacts can be contained and recovered from, and allowing the nature of adverse impacts to be better understood thus informing potential mitigation actions.

5.3 Model purpose and value

A power system simulation model is a tool which should be used to answer a question and is not in and of itself an answer to a question. Choice of a modelling approach should be made having in mind what the answers it is providing should actually be used for. There are multiple potential purposes which will all require subtly different approaches:

- To inform the deployment and design of “last resort” containment ‘defence plans’ such as LFDD/UFLS or under-voltage load shedding
- To mitigate events before they come to rely on “last resort” interventions
- To plan investment strategies for infrastructure
- To plan dispatch policies to manage resource shortages
- To inform revised system security standards
- To justify ‘asset hardening’, either targeted at specific locations or more generally through revised design standards or policies

Modelling such as OPFs are a proxy for human decision-making, and can be used to better inform those decisions, but there is no “one-size-fits-all” approach to power system models. Dynamic simulations such as “RMS” or “EMT” models are limited by computational expense and cannot be used within the same simulation model as, for example, a unit commitment problem, but can be used within a framework which uses network simplifications or reductions as part of a trade-off between spatial coverage and temporal detail.

The choice of events against which to evaluate system resilience or test the need for new interventions will be key to what questions are able to be answered. Events such as the August 2019 outages were not directly caused solely by a significant exogenous threat, but by compounding of independent, flawed protection or control actions – “hidden failures” that were already present on the system and were revealed by the initial short circuit event within the system which caused a cascading of outages. The “last resort” interventions, though causing inconvenience, functioned to prevent a more severe cascading outage or blackout scenario.

As is discussed within Section 4, it is the case that for a given set of inappropriate actions it is often indistinguishable whether a bad or wrong action taken is malice or misfortune. This can also be true of control or protection actions. While an individual component may react “correctly” and according to its settings, if those settings are incorrect, the action could exacerbate or, in and of itself, cause a series of events leading to a cascading outage. This is also true of human attempts to restore the system following a significant outage.

If there is a serious attempt to model cascading outages on some level at least basic consideration of protection actions such as generator tripping, RoCoF actions, loss-of-mains protection, or overload protection on lines will have to be considered at some point. But this, in turn, depends on the temporal granularity of the model, or models, and how they interact, as well as an understanding of the initial conditions of the model and the settings of these protection devices.

Similarly, credible modelling of restoration can only be done if the circumstances under which restoration is needed are modelled. For example, restoration of electricity supplies on August 9th 2019 was relatively straightforward as there was no equipment damage- though restoration of rail services was another matter. Restoration of disconnected supplies during Storm Arwen was highly challenging due to damage to network assets and the storm’s hindering of access by repair teams. In order to replicate an event like August 2019 you have to be able to simulate the events which caused it *as well as* the cascading events within the power system

which resulted. Having this infrastructure within the model then allows you to model mitigations to prevent LFDD deployment (which, again, can only be incorporated in the model if you model if you put it there).

Extreme events can emerge both from severe exogenous threats and from endogenous weaknesses within the system, even if multiple systems independently operate appropriately the sum total of their actions can accumulate to an adverse overall outcome.

If we are looking to model mitigations and preventive actions, it needs to be clear what models are to be used and when. Different phases of an outage and impacts will happen at different speeds – inverter control respond within milliseconds; frequency excursions and associated outages can happen in a matter of seconds. Outages associated with line damage can occur over a period of hours, and can take days to repair, but the loss of demand or generation which can occur with damage to the network can have immediate consequences. Similarly, features such as Loss-of-Mains protection can cause common mode outages from DER which have common settings. Resource adequacy and actions to ensure it are typically assessed some years in advance and involve modelling, as a minimum, of annual peak demand conditions or, for greater confidence, whole years of operation. Each of these different phenomena, to the extent that are modelled today, are assessed with different models using different sets of equations and sets of parameters.

Setting the initial conditions of these simulations will therefore be a key proxy for how an organisation would actually seek to prepare the system for an extreme event. Presumptions about protection settings and generator tripping behaviour can be asserted and refined later; but they still need to be made. The system operator will need to deploy frequency response correctly in accordance with prevailing system conditions, including the level of demand, the availability of power from renewables, the unconstrained level of system inertia and the likelihoods of different levels of loss of infeed. Modelling of what would happen to system frequency under different conditions must take such decisions into account.

Model complexity should match the likely phenomena which can occur. In a distribution system, which is typically radial, it is unlikely to be necessary to model the full dynamics relating to, for example, system frequency and angle stability – a lost connection will simply result in loss of service to those customers, usually with a limited impact on the transmission system. However, disruptions to the Main Interconnected Transmission System (MITS) will, particularly in high-renewables scenarios, almost always require some level of dynamic simulation, even if it is not contained within the same simulation package.

In summary, defining clearly the scope and scale of modelling will be essential to determining its value and purpose. A suggested action is to have a workshop across sectors with power system modelling experts to determine a “toolkit" or inventory of models and to match these models to different scenarios which will be modelled.

6 References

[1] Energy Networks Association. "Scenarios for Extreme Events." https://smarter.energynetworks.org/projects/10060460/ (accessed 17/03/2024, 2024).

[2] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, "Metrics and Quantification of Operational and Infrastructure Resilience in Power Systems," *IEEE Transactions on Power Systems,* vol. 32, no. 6, pp. 4732-4742, 2017, doi: 10.1109/TPWRS.2017.2664141.

[3] Energy Research Partnership, "Future Resilience of the UK Electricity System," United Kingdom, 2018. Accessed: 10/06/2022. [Online]. Available: http://erpuk.org/project/future-resilience-of-the-uk-electricity-system/

[4] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of risk,* vol. 2, pp. 21-42, 2000.

[5] T. Lagos *et al.*, "Identifying Optimal Portfolios of Resilient Network Investments Against Natural Hazards, With Applications to Earthquakes," *IEEE Transactions on Power Systems,* vol. 35, no. 2, pp. 1411-1421, 2020, doi: 10.1109/TPWRS.2019.2945316.

[6] CIGRE Working Group C4.47. "Defining power system resilience." CIGRE. https://www.cigre.org/article/GB/news/the_latest_news/defining-power-system-resilience (accessed 10/06/2022, 2022).

[7] United States National Academies, "Enhancing the Resilience of the Nation's Electricity System," United States Academies of Science, Engineering, and Medicine,, United States, 0309463106, 2017.

[8] National Infrastructure Commission, "Anticipate, React, Recover: Resilient Infrastructure Systems," National Infrastructure Commission,, 2020. Accessed: 19/01/23. [Online]. Available: https://nic.org.uk/studies-reports/resilience/

[9] E. Karangelos and L. Wehenkel, "Probabilistic Reliability Management Approach and Criteria for power system real-time operation," in *2016 Power Systems Computation Conference (PSCC)*, 20-24 June 2016 2016, pp. 1-9, doi: 10.1109/PSCC.2016.7540871.

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

[11] K. R. W. Bell, "Issues in integration of risk of cascading outages into utility reliability standards," in *2011 IEEE Power and Energy Society General Meeting*, 24-28 July 2011 2011, pp. 1-7, doi: 10.1109/PES.2011.6039576.

[12] A. Stirling, "Keep it complex," *Nature,* vol. 468, no. 7327, pp. 1029-1031, 2010.

[13] J. Sharp, M. Miligan, and H. Bloomfield, "Weather Dataset Needs for Planning and Analyzing Modern Power Systems," ESIG,, October 2023 2023. Accessed: 07/11/2023. [Online]. Available: https://www.esig.energy/wp-content/uploads/2023/10/ESIG-Weather-Datasets-summary-report-2023.pdf

[14] R. Entriken and R. Lordan, "Impacts of extreme events on transmission and distribution systems," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, USA, 22-26 July 2012 2012, pp. 1-10, doi: 10.1109/PESGM.2012.6345755.

[15] M. Panteli, P. Mancarella, S. Wilkinson, R. Dawson, and C. Pickering, "Assessment of the resilience of transmission networks to extreme wind events," in *2015 IEEE Eindhoven PowerTech*, Eindhoven, Netherlands, June 29 2015-July 2 2015 2015, pp. 1-6, doi: 10.1109/PTC.2015.7232484.

[16] M. Auffhammer, P. Baylis, and C. H. Hausman, "Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States," *Proceedings of the National Academy of Sciences,* vol. 114, no. 8, p. 1886, 2017, doi: 10.1073/pnas.1613193114.

[17] L. C. Dawkins, "Weather and Climate Related Sensitivities and Risks in a Highly Renewable UK Energy System: A Literature Review," The Met Office, United Kingdom, 02/07/19 2019. Accessed: 13/03/2024. [Online]. Available: https://nic.org.uk/app/uploads/MetOffice_NIC_LiteratureReview_2019.pdf

[18] K. Murray and K. R. W. Bell, "Wind related faults on the GB transmission network," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on*, Durham, UK, 7-10 July 2014 2014, pp. 1-6, doi: 10.1109/pmaps.2014.6960641.

[19] E. A. Morris, K. R. W. Bell, and I. M. Elders, "Spatial and temporal clustering of fault events on the GB transmission network," in *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 16-20 Oct. 2016 2016, pp. 1-9, doi: 10.1109/PMAPS.2016.7764087.

[20] A. Serrano-Fontova *et al.*, "A comprehensive review and comparison of the fragility curves used for resilience assessments in power systems," *IEEE Access,* 2023.

[21] S. Dunn, S. Wilkinson, D. Alderson, H. Fowler, and C. Galasso, "Fragility curves for assessing the resilience of electricity networks constructed from an extensive fault database," *Natural Hazards Review,* vol. 19, no. 1, p. 04017019, 2017.

[22] M. R. Jamieson, G. Strbac, and K. R. Bell, "Quantification and visualisation of extreme wind effects on transmission network outage probability and wind generation output," *IET Smart Grid,* vol. 3, no. 2, pp. 112-122, 2019.

[23] D. J. Cannon, D. J. Brayshaw, J. Methven, P. J. Coker, and D. Lenaghan, "Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain," *Renewable Energy,* Article vol. 75, pp. 767-778, 2015, doi: 10.1016/j.renene.2014.10.024.
[24] H. Macdonald, G. Hawker, and K. Bell, "Analysis of wide-area availability of wind generators during storm events," in *2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, Durham, UK, 7-10 July 2014 2014, pp. 1-6, doi: 10.1109/PMAPS.2014.6960619.
[25] K. B. Karnauskas, J. K. Lundquist, and L. Zhang, "Southward shift of the global wind energy resource under high carbon dioxide emissions," *Nature Geoscience,* vol. 11, no. 1, pp. 38-43, 2018/01/01 2018, doi: 10.1038/s41561-017-0029-9.
[26] J. S. Hosking *et al.*, "Changes in European wind energy generation potential within a 1.5 °C warmer world," *Environmental Research Letters,* vol. 13, no. 5, p. 054032, 2018/05/01 2018, doi: 10.1088/1748-9326/aabf78.
[27] Nasa. "Modern-Era Retrospective analysis for Research and Applications, Version 2." NASA. https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/ (accessed 04/10/2022, 2022).
[28] S. Lee and Y. Ham, "Probabilistic framework for assessing the vulnerability of power distribution infrastructures under extreme wind conditions," *Sustainable Cities and Society,* vol. 65, p. 102587, 2021.
[29] R. Kemp, "Living without electricity - One city's experience of coping with loss of power," Royal Academy of Engineering, London, 2016. [Online]. Available: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/engineering/RAEngLivingwithoutelectricity.pdf
[30] J. L. Cremer and G. Strbac, "A machine-learning based probabilistic perspective on dynamic security assessment," *International Journal of Electrical Power & Energy Systems,* vol. 128, p. 106571, 2021/06/01/ 2021, doi: https://doi.org/10.1016/j.ijepes.2020.106571.
[31] NREL. "Wind Integration National Dataset Toolkit." https://www.nrel.gov/grid/wind-toolkit.html (accessed 09/11/2023, 2023).
[32] J. W. Busby *et al.*, "Cascading risks: Understanding the 2021 winter blackout in Texas," *Energy Research & Social Science,* vol. 77, p. 102106, 2021/07/01/ 2021, doi: https://doi.org/10.1016/j.erss.2021.102106.
[33] F. Teng, R. Dupin, A. Michiorri, G. Kariniotakis, Y. Chen, and G. Strbac, "Understanding the Benefits of Dynamic Line Rating Under Multiple Sources of Uncertainty," *IEEE Transactions on Power Systems,* vol. 33, no. 3, pp. 3306-3314, 2018, doi: 10.1109/TPWRS.2017.2786470.
[34] J. Carlton and E. Ailworth. "PG&E Shuts Power to California Resort Area to Prevent Wildfires." The Wall Street Journal. https://www.wsj.com/articles/pg-e-shuts-power-to-california-resort-area-to-prevent-wildfires-11560046802 (accessed 10/08/2023, 2023).
[35] R. Moreno *et al.*, "Microgrids Against Wildfires: Distributed Energy Resources Enhance System Resilience," *IEEE Power and Energy Magazine,* vol. 20, no. 1, pp. 78-89, 2022.
[36] D. J. Brayshaw, A. Troccoli, R. Fordham, and J. Methven, "The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the UK," *Renewable Energy,* vol. 36, no. 8, pp. 2087-2096, 8// 2011, doi: http://dx.doi.org/10.1016/j.renene.2011.01.025.
[37] Australian Government Bureau of Meteorology. "Bureau's Atmospheric high-resolution Regional Reanalysis for Australia (BARRA)." (accessed 05/12/2023, 2023).
[38] NREL. "NSRDB: National Solar Radiation Database." https://nsrdb.nrel.gov/data-sets/us-data (accessed 09/11/23, 2023).
[39] I. Staffell and S. Pfenninger, "The increasing impact of weather on electricity supply and demand," *Energy,* vol. 145, pp. 65-78, 2018/02/15/ 2018, doi: https://doi.org/10.1016/j.energy.2017.12.051.
[40] C. Grigg *et al.*, "The IEEE Reliability Test System-1996. A report prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee," *IEEE Transactions on Power Systems,* vol. 14, no. 3, pp. 1010-1020, 1999, doi: 10.1109/59.780914.
[41] Nerc. "Transmission Availability Data System (TADS)." North American Electric Reliability Corporation. http://www.nerc.com/pa/RAPA/tads/Pages/default.aspx (accessed 13/11/2023, 2023).
[42] E. Hart and A. Mileva, "Advancing Resource Adequacy Analysis with the GridPathRA Toolkit: A Case Study of the Western US," *Energy Systems Integration Groups ESIG) Webinar Presentation,* vol. 6, 2022.
[43] E. Ciapessoni, D. Cirio, A. Pitto, and M. Sforna, "A quantitative methodology to assess the process of service and infrastructure recovery in power systems," *Electric Power Systems Research,* vol. 189, p. 106735, 2020.
[44] R. Pant, T. Russell, C. Zorn, E. Oughton, and J. Hall, "Resilience study research for NIC–systems analysis of interdependent network vulnerabilities. Environmental Change Institute," ed: Oxford University, UK. , 2020.
[45] M. Noebels, R. Preece, and M. Panteli, "AC Cascading Failure Model for Resilience Analysis in Power Networks," *IEEE Systems Journal,* 2020.
[46] M. A. Rios, D. S. Kirschen, D. Jayaweera, D. P. Nedic, and R. N. Allan, "Value of security: modeling time-dependent phenomena and weather conditions," *IEEE Transactions on Power Systems,* vol. 17, no. 3, pp. 543-548, 2002, doi: 10.1109/TPWRS.2002.800872.
[47] W. Ju, K. Sun, and R. Yao, "Simulation of Cascading Outages Using a Power-Flow Model Considering Frequency," *IEEE Access,* vol. 6, pp. 37784-37795, 2018, doi: 10.1109/ACCESS.2018.2851022.

[48] D. Kirschen, K. Bell, D. Nedic, D. Jayaweera, and R. Allan, "Computing the value of security," *IEE Proceedings-Generation, Transmission and Distribution,* vol. 150, no. 6, pp. 673-678, 2003.
[49] National Grid ESO. "Frequency response services." https://www.nationalgrideso.com/industry-information/balancing-services/frequency-response-services (accessed 28/02/2024, 2024).
[50] National Grid ESO, "Security and Quality of Supply Standard (SQSS)," 2022, vol. 2022.
[51] A. Monticelli, M. V. F. Pereira, and S. Granville, "Security-Constrained Optimal Power Flow with Post-Contingency Corrective Rescheduling," *IEEE Transactions on Power Systems,* vol. 2, no. 1, pp. 175-180, 1987, doi: 10.1109/TPWRS.1987.4335095.
[52] D. N. Trakas and N. D. Hatziargyriou, "Resilience Constrained Day-Ahead Unit Commitment Under Extreme Weather Events," *IEEE Transactions on Power Systems,* pp. 1-1, 2019, doi: 10.1109/TPWRS.2019.2945107.
[53] M. R. Jamieson, Q. Hong, J. Han, S. Paladhi, and C. Booth, "Digital twin-based real-time assessment of resilience in microgrids," *Renewable Power Generation,* 2022.
[54] Office of Gas and Elecricity Markets. "Investigation into 9 August 2019 Power Outage." Ofgem,. https://www.ofgem.gov.uk/publications-and-updates/investigation-9-august-2019-power-outage (accessed 30/01/2024, 2024).
[55] P. J. C. Vogler-Finck and W.-G. Früh, "Evolution of primary frequency control requirements in Great Britain with increasing wind generation," *International Journal of Electrical Power & Energy Systems,* vol. 73, pp. 377-388, 2015.
[56] C. MacIver, K. Bell, and M. Nedd, "An analysis of the August 9th 2019 GB transmission system frequency incident," *Electric Power Systems Research,* vol. 199, p. 107444, 2021.
[57] M. R. Jamieson, "Quantification and mitigation of the impacts of extreme weather on power system resilience and reliability," Doctor of Philosophy (PhD) Doctoral, Electrical and Electronic Engineering, Imperial College London, London, UK, 2020.
[58] S. Gordon, C. McGarry, J. Tait, and K. Bell, "Impact of Low Inertia and High Distributed Generation on the Effectiveness of Under Frequency Load Shedding Schemes," *IEEE Transactions on Power Delivery,* vol. 37, no. 5, pp. 3752-3761, 2022, doi: 10.1109/TPWRD.2021.3137079.
[59] NGESO. "Low Frequency Demand Disconnection." National Grid ESO. https://www.nationalgrideso.com/document/87836/download (accessed 21/11/2023, 2023).
[60] BBC News. "Sellafield: Minister wants answers on alleged cyber hack." BBC News,. https://www.bbc.co.uk/news/uk-england-cumbria-67623880 (accessed 17/03/2024, 2024).
[61] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine Blackout: Implications for False Data Injection Attacks," *IEEE Transactions on Power Systems,* vol. 32, no. 4, pp. 3317-3318, 2017, doi: 10.1109/TPWRS.2016.2631891.
[62] A. Nourian and S. Madnick, "A Systems Theoretic Approach to the Security Threats in Cyber Physical Systems Applied to Stuxnet," *IEEE Transactions on Dependable and Secure Computing,* vol. 15, no. 1, pp. 2-13, 2018, doi: 10.1109/TDSC.2015.2509994.
[63] I. Zografopoulos, J. Ospina, X. Liu, and C. Konstantinou, "Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies," *IEEE Access,* vol. 9, pp. 29775-29818, 2021, doi: 10.1109/ACCESS.2021.3058403.
[64] Q. Hong *et al.*, "Fast frequency response for effective frequency control in power systems with low inertia," *The Journal of Engineering,* vol. 2019, no. 16, pp. 1696-1702, 2019.
[65] K. R. W. Bell and A. N. D. Tleis, "Test system requirements for modelling future power systems," in *IEEE PES General Meeting*, Minneapolis, USA, 25-29 July 2010 2010, pp. 1-8, doi: 10.1109/PES.2010.5589807.
[66] D. Clements and P. Mancarella, "Fragility curve based storm modelling of distribution networks with staff constraints," in *IET International Conference on Resilience of Transmission and Distribution Networks (RTDN 2017)*, Birmingham, UK, 26-28 Sept. 2017 2017, pp. 1-5, doi: 10.1049/cp.2017.0342.