Paris 2020

C1-112

Elicitation of Structured Expert Judgement to estimate the probability of a major power system unreliability event

K. BELL T. BEDFORD A. COLSON
University of Strathclyde
UK

M. BARONS S. FRENCH
University of Warwick
UK

SUMMARY

Resilience of a power system is concerned with preventing adverse outcomes from disturbances to the system, containing them when they do occur and recovering from them as quickly and safely as possible.

Major unreliability events worldwide over many years have shown a number of common phenomena, e.g. cascades of outages and frequency or voltage instability. However, the precise pathway is different every time and dependent on complex and uncertain system behaviour. Events also differ based on the characteristics of each system and social factors such as the priorities and judgements of individuals such as key control room staff. Moreover, major unreliability events remain very rare. The use of modelling to estimate the probability of a regional or whole system shutdown is therefore extremely difficult.

This paper concerns an approach to the estimation of the probability of a major power system unreliability event. The general approach – use of Structured Expert Judgement (SEJ) – has been used in many fields in which major risks need to be understood in respect of phenomena that are not readily amenable to modelling. This paper describes what the authors believe to be the world's first application of the approach to understanding and enhancement of power system resilience. It outlines a conceptual model of how a system blackout might happen and describes how it is used with a group of experts to provide estimates of the likelihood of a system disturbance propagating along different paths.

KEYWORDS

Major power system unreliability events; Risk; Resilience; blackouts; power system stability; Structured Expert Judgement.

Keith.bell@strath.ac.uk

1 Introduction

Resilience of a power system is concerned with preventing unacceptable outcomes from disturbances to the system, containing them when they do occur and recovering from them as quickly and safely as possible. Security standards are in place worldwide to address the first of these aspects while defence plans address the second and, where they succeed in containing the impact of an event, help the third. However, one particular question remains: how often might a system collapse and black start be required?

Transmission systems are normally operated to a very high reliability standard (a variant on the widely known "N-1" convention in which single credible unplanned events are secured against) with multiple protection systems in place. Furthermore, when transmission operators anticipate that the system will be subject to adverse conditions, such as extreme weather, they will normally take precautionary measures to make the network more resilient, normally at the cost of running more expensive generation in key locations (although, if unplanned events occur, limited disconnection of some customers may be used, i.e. a precautionary partial shutdown, in order to prevent a larger or total accidental shutdown).

Although major unreliability events worldwide over many years have shown a number of common phenomena, e.g. cascades of outages and frequency or voltage instability, the precise pathway is different every time and dependent on complex and uncertain system behaviour as well as the different characteristics of each system and social factors such as the priorities and judgements of individuals such as key control room staff. Moreover, such major events remain very rare. The use of modelling to estimate the probability of a regional or whole system shutdown is therefore extremely difficult.

This paper describes an alternative approach to the estimation of the probability of a major power system unreliability event. The general approach – use of Structured Expert Judgement (SEJ) – has been used in many fields in which major risks need to be understood in respect of phenomena that are not readily amenable to modelling. The paper describes what the authors believe to be the world's first application of the approach to understanding and enhancement of power system resilience.

The SEJ approach involves the development of a conceptual, high level model of the phenomena of interest that is used as the basis for a structured set of questions to elicit the judgements of experts in the relevant field. Discussion between the experts is used to test buy-in to the overall model and to gain a common understanding of the questions and highlight the range of scenarios included under each event within the high level model. The approach also involves a number of specific steps with the following aims:

(i) to allow the assessment of the different experts' ability not only to estimate a probability but also an uncertainty range within which they believe the true answer lies, and to use this assessment in weighting the experts' answers to the main questions of interest; and

(ii) to elicit each expert's own judgement.

The high level system model is an event tree based on initiating or precursor events which are violations of normal system operating limits: low frequency, low voltage, overloads or low frequency oscillations on the system. The model outlines pathways through which there either a system wide frequency or voltage instability, or there is a transmission network split that leads to a failure of the entire system. This model is described further in section 5.

Choice of the SEJ approach with the particular high level power system model used was based on the team's view that:

a) Experts would find it very difficult to make direct judgements on the probability of total shutdown of a large power system, and would feel more able to make assessments of intermediate events leading to the total shutdown.

b) If operational data were available from the operator of the system under consideration then this should be used where appropriate to provide initiating frequencies of precursor events.

c) A practical limitation on number of intermediate events included in the risk model would be imposed by the time that a set of experts could be expected to make available for the work.

d) Intermediate events would therefore incorporate a number of different scenarios, including diverse environmental conditions, states of equipment on the electricity network, states of knowledge of the operators of the network, management practice in the operational interpretation of standards. Experts would be asked to "fold in" these aspects in their assessment of the uncertainty.

e) A high level risk model of this type would provide extra benefits to stakeholders by providing some understanding of the major contributors to overall risk, and therefore also of broader approaches by which this could be mitigated. It would also be easier to communicate to less technical audiences.

The rest of this paper describes the main foundations of the SEJ approach, the high level power system model, and some lessons from use of the approach in an expert workshop.

2 Structured Expert Judgement

Structured expert Judgement is a transparent and principled method for using diverse expertise across relevant areas to support evidence-based policy analysis and practice.

The theory behind the use of expert judgement in decision and risk analysis has a long and distinguished pedigree [1]. By the early 1990s the Classical method, developed by Cooke [2], was being applied in several case studies. A joint EU and US project on the risks relating to nuclear accident consequence modelling did much to formalise procedures [3]. In the past decades, many further applications have been undertaken: e.g.

- Future rates of antimicrobial resistance in Europe for key pathogen-antibiotic pairs [4].
- An estimation of the economic value of the non-health benefits of breastfeeding in low- and mid-income countries [5].
- A quantification of the effects on pollinator abundance of changes in weather, environment and disease pressure [6].
- Assessment of the uncertainty in future sea level rise from melting of the ice sheets [7].

Many variations of the approach have been proposed and used, each eliciting different experts' judgements and combining them in different ways. Broadly speaking, approaches to combining the judgements of several experts have fallen into two categories: mathematical and behavioural.

Methods using mathematical aggregation ask individual experts to provide their quantitative probability judgements and then proceed to combine these via some mathematical or statistical formula. The earliest methods, known as opinion pools, simply took an average across all the experts. Developments introduced geometric and other averaging mechanisms, and weights which could reflect the different relative expertise of the experts. However, this also raised the issue that there was no clear operational definition of expertise upon which to base the weights.

Cooke [2] developed the Classical Method to address this lack. He argued that it was appropriate to weight experts' judgements differentially, but only on the basis of evidence. He suggested not only asking for the uncertainty judgements on the quantities of interest, but also on a set of calibration questions on seed variables whose true values are known to the analysts but not the experts. This would allow the experts' judgements to be weighted differentially according to how effectively they judge the uncertainties on these variables, assessed in terms of particular measures of statistical performance.

The second category of approaches, behavioural aggregation, essentially allows the experts to discuss the issues and agree on a single set of quantitative values, which represents their best view of what a reasonable observer would think based on all the evidence and experience available [8]. Their discussion is facilitated and statements are continually challenged to reduce the possibility of psychological biases producing inappropriate judgements.

It is important to emphasise that in both mathematical and behavioural aggregation qualitative evidence is gathered alongside the quantitative judgements. Numbers are not just elicited, but the experts are expected to give qualitative arguments to justify their assessments.

Recently, Hanea *et al.* [9] have proposed the IDEA – Investigate, Discuss, Estimate and Aggregate – approach, which seeks to combine the benefits of both mathematical and behavioural aggregation (see also, [10]). Essentially, this involves the experts in a facilitated discussion as they form their judgement, but they then make their quantitative assessments independently and anonymously. The aggregated results are then shared anonymously with the group and discussion re-initiated. These numerical judgements are aggregated via the Classical method.

3 Application of Structured Expert Judgement to investigation of major power system disturbances

The selection of the SEJ approach was based on the team's view that:

a) The events were sufficiently complex that group discussion was required to gain a common understanding of each event and to highlight the different elements of importance, key pathways, impact of different geographies on the network, etc.

b) A range of different types of expertise was required, especially to highlight aspects of operator response, variations in equipment type, the behaviour of power system protection, geography etc.

c) There could be differences of performance in terms of judging uncertainty between the different experts, and we should try to manage this appropriately.

d) The approach should allow the questions to be dealt with in the time available.

The SEJ approach adopted was based on the IDEA protocol and had two phases. During the first phase, the risk model was developed for the assessment of the target uncertainties. This then reviewed by several independent international experts including some from countries with experience of major outages. These experts also reviewed the draft elicitation questions to be used in the workshop. The second phase was a 1.5-day workshop.

In summary, the key steps in the application of the SEJ approach here were as follows.

1. An Event Tree Model was developed to provide a high-level risk model bridging the gap between events in which violations of normal system operating limits have occurred that, by reference to shutdown events worldwide, can be seen as precursors to total shutdown of a large power system, and the total shutdown itself. This high-level risk model (described in section 5) was developed by the team, and critiqued and modified by a

group of international experts and was subsequently challenged again in a workshop of experts held as part of the SEJ process.

2. During Workshop Preparation and Operational Data steps we developed questions for the experts that would allow us to assess their opinions on the likelihood of unwanted events happening as described in the event tree model. Operational data were obtained that could be used to judge expert effectiveness in assessing uncertainty, i.e. to calculate weights for the experts. Data were also sought to quantify the rates of initiating events in the event tree model.
3. The SEJ Workshop primarily assessed failure probabilities and asked experts not only to give their best estimates but also an uncertainty range associated with each step of the event tree model. The workshop followed the IDEA expert judgement protocol to manage the discussions, an approach that draws on the Delphi process and Cooke's Classical model.
4. Data Processing was carried out using Cooke's Classical model of expert judgement. For each event in the event tree model this combines the expert assessments into a single uncertainty range with 5%, 50% and 95% values for the probability of that event. The written comments of experts were reviewed as were the numerical values.

The SEJ workshop was run over 1.5 days and was attended by 11 experts, including transmission and distribution network operators, protection experts, a power station operator and an academic. The risk model was discussed first, as well as a short introduction to uncertainty and probability assessment. Individual questions were discussed in plenary before the experts were asked to provide initial probability assessments (in the form of values that they judged had a 50%, 5% or 95% chance of being exceeded) independently.

The questions included a set of so-called calibration questions related to incidents that have occurred on the system under discussion, which could be used to assess the skill of each expert in encoding his or her uncertainty probabilistically. Overnight, the initial answers were processed by the team. The next morning a group discussion was held around the outcomes with, for each question, the range of estimates from within the group being shown, anonymised. This enabled further elucidation of the pathways, as well as identifying some areas where mistaken interpretations had been made. The experts then independently provided a second round of assessments where they were able to update their initial values if they wished to.

4 Assessment of risk of power system shutdowns

Power systems are large, dynamic, non-linear and complex. Simulation of their behaviour has been commonplace around the world for decades, in particular in the conduct of 'what if' assessments of the potential impacts of disturbances to inform system operator actions to prevent impacts from being significant. However, only a limited set of 'credible' disturbances – 'secured events' – are generally tested. Disturbances beyond this set can and do happen and are the result of many and varied factors. These are subject to many uncertainties arising from the randomness of external influences on the system and the difficulty of modelling the system accurately both in terms of equations that represent power systems' behaviour and the parameters describing a particular system. In addition, given the possibility of measurement errors and differences in when measurements are received, the current condition of a system cannot be known with absolute confidence. Nevertheless, established policies and procedures, supported by practical engineering models, do, for the most part, succeed in ensuring stable operation of power systems.

Simulation models of power systems have been in use around the world for many decades to support decisions related to both planning and operation of a system. However, the scope of each individual model is limited and depends on a number of assumptions. At present, none are in use that incorporate all the significant factors that would influence whether a particular set of circumstances and a particular disturbance would cascade into a whole system blackout [11] [12].

All of the aforementioned issues make it extremely difficult to assess the probability of a large power system suffering a complete shutdown. Instead of creating a truly "bottom-up" model for a major power system failure it was decided to create an event tree model for the top event of a whole system blackout, starting with basic events that are, in fact, intermediate conditions of the system rather than component-level failures.

The use of a fault/event tree approach for the assessment of risk is standard in risk analysis of critical systems. (This approach was developed initially in the nuclear sector with the support of the US Nuclear Regulatory Commission, and has become the world-wide standard approach to assess the risk of engineered systems[1]). That said, the key critique of a fault tree approach is that it does not capture dynamical system behaviour very well, for example feedback loops that affect system evolution.

A host of different dynamic risk modelling approaches have been developed, but all suffer from the problem that they become very complex: although many power system models broadly capture the required dynamic behaviours of the system, many of the details that affect the development of the system's state beyond the first minute or so after a disturbance are omitted. Moreover, assumptions are made regarding many of the phenomena that occur at timescales of less than a few milliseconds. For this reason, and given that we were using expert judgement to support assessment of conditional probabilities, it was decided that such dynamic complexities would most appropriately be captured within expert judgements rather attempting to model them explicitly.

One concrete illustration of the use of judgement is in the way that cascades are dealt with. We simply described an event of "cascades leading to… " without being explicit about the equipment involved in the cascade, how many failures or trips were involved, whether these involved breaches of normal operating limits and correct or inadvertent operation of system protection, and whether they occurred as a result of the system operator not having the correct information about the system state, as a result of unforeseen environmental conditions, or something else. In fact, a large power system could be considered as a Socio-Technical-Environmental system because its behaviour depends on the interaction between the humans operating, maintaining, and using it, the reliability of responses of the technical systems, the responses embedded in the automated systems and the impact of the environment (primarily weather, but also other aspects from the growth over time of local vegetation to the occurrence of electrical storms). The full complexity of this system's behaviour is therefore impossible to model fully using currently available modelling approaches.

5 A power system shutdown event tree

The event tree describes the progression from various intermediate conditions to the complete system shutdown through event sequences. These events are selected with the purpose of describing the main routes or scenarios by which escalation to a black start occurs. The contributing events are still at a high level, but are easier for experts to deal with and closer to events or circumstances for which there might be some empirical data. Note that the possibility of a terrorist or other deliberate adversarial act was *not* taken into account.

---

[1] The Fault Tree Handbook (NUREG-0492) was initially created by the US NRC, and more recently an updated version was produced by NASA for use in the space industry. It is used in the UK for the nuclear and railway industries amongst others.

Review of historic major power system disturbances worldwide suggests that scenarios leading to a full system shutdown will include low system frequency, low voltage, or overloads (where the first is a system-wide phenomenon and the others can be observed at particular locations). These are sustained for some period of time (representing a steady state or quasi steady state that, in principle, is observable by the System Operator (SO)) during which the system is vulnerable to: further disturbances from outside the system or to the operation of automatic controls on the system3. These may have the effect of worsening the system's state still further and, potentially, setting a cascade in train.

5.1 Violations of normal operating limits and requirement for operator action

One key to avoiding a major disturbance is that the SO recognises these violations of limits and takes action either to correct them or to reassure themselves that the violation will be removed by the expected transition of generation and/or demand. An SO's failure to correct a violation may due to lack of situational awareness and the need for action, or failure of the action, e.g. due to an action having been incorrectly specified or failure of equipment to implement it. There is also always the possibility of other external disturbances giving rise to new or exacerbated violations.

Low system frequency, low voltages and overloads are not the only operating limits with which an SO will be concerned. Others include high frequency and high volts. If these violations are not addressed by an SO, there is the possibility of equipment tripping and leading to other violations. To the team's knowledge, they have not been identified as part of the sequence of events that has led to any of the major disturbances observed worldwide. However, if they did, it would likely be via, for example, high system frequency causing the tripping of so much generation that a low system frequency, low voltage or overload condition arises. In other words, high voltage or high system frequency could create a major disturbance but via one of the other types of violation. (Figure 1).

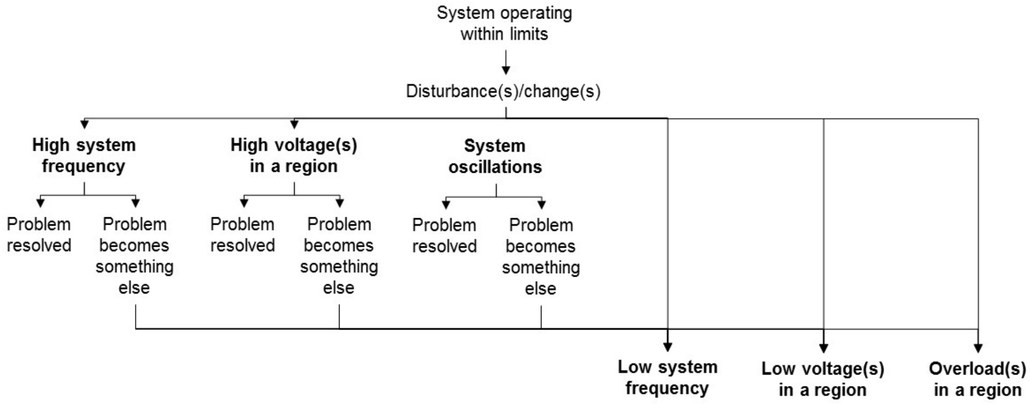


Figure 1: initial conditions in the high level risk model

Periodic rises and falls in voltages, currents and, potentially, system frequency, with the oscillations themselves having frequencies of less than 50 Hz, could lead to transient conditions that are sufficient to trigger the operation of automatic protection equipment on the system and, hence, cause outages that might lead to cascades. Such a mechanism has been highlighted as having been a feature of the South Australia shutdown in September 2016 [13].

In summary, the initiating conditions for the model are low system frequency, low voltage and overload violations, as well as system oscillations. In principle, data should exist for these events, although we also made use of expert judgement as is explained further below.

5.2 *Cascades of outages and propagation through an event tree*

Many major system disturbances worldwide have the feature of cascades of outages (See, for example, [14][15][16]). Although an initial event (such as a violation of system limits – frequency, voltage or overloads – which usually arises only as a consequence of an equipment failure) might be within the set of credible events that an SO should secure against, in many major disturbances the actions of the system operator have failed to re-establish a secure situation before a 'point of no return' is reached.

An uncorrected violation then triggers a new outage which, in turn, causes another violation and another outage. The sequence can happen very quickly, i.e. in some cases within a few seconds. Often, the start of the sequence or part of its propagation involves an item of equipment tripping that, ideally, would not have. The exact path of a cascade varies considerably and is subject to many uncertainties. However, some key features can be observed. For example, the British system has a system, of under frequency load shedding, known as 'Low Frequency Demand Disconnection', (LFDD) as a defence against frequency instability. For the system to collapse, LFDD as a defence must have failed, either because a rapid cascade took place before it was triggered, or because it was triggered but was ineffective in establishing a stable situation. (Figure 2).

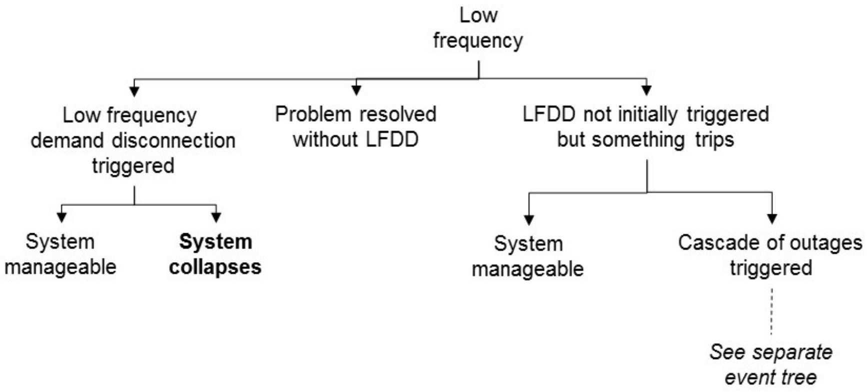


Figure 2: events following occurrence of low system frequency

The system is normally a single interconnected system. However, as a cascade progresses, the further propagation of the disturbance can involve one or more splits of the system, in general due to the action of protection, to remove overloads, to interrupt faults, to separate circuit ends across which there are already significant voltage angle differences or in response to particular combinations of voltage and current that have arisen on electrically weak boundaries between two parts of the system.

Where the whole system or an electrical island collapses, reviews of past events suggest that it finally due to frequency instability, voltage instability or a combination of them. These pathways are summarised in Figure 3.

6 Reflections on the Structured Expert Judgement process as used

6.1 *Expertise and organisation of a workshop*

For reasons of confidentiality, it is, unfortunately, not possible for the results of the SEJ workshop to be reported here. However, some general observations can be made.

The effort in building the event tree model of a major power system was considerable as it had to be developed from scratch and both be accurate in its representation of power system phenomena, and clear and succinct in enabling experts to articulate their judgement and inform the end objective of the exercise. The team was fortunate in having access to people with, on the one hand, power systems knowledge and, on the other, good working and theoretical knowledge of the SEJ process. As an example of inter-disciplinary working, the team would argue that it was very successful. However, given the need to develop the high level system model from scratch, there was a quite heavy dependency on the sole power systems engineer in the team.

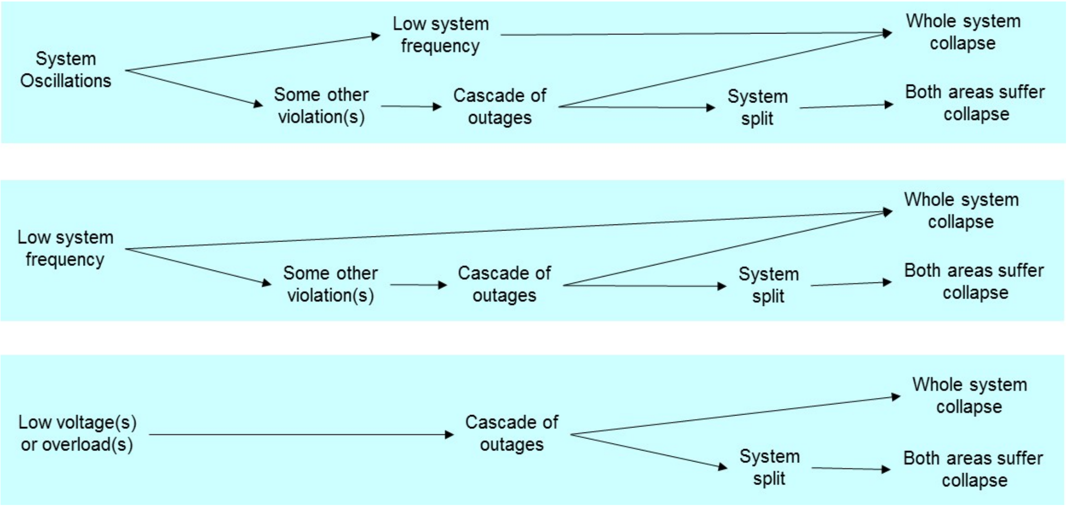


Figure 3: summary of pathways leading to whole system collapse

In the workshop, it quickly became apparent that, in respect of the key questions in the expert elicitation, all experts are equal but ‘some are more equal than others’. Perhaps the single most important judgement in estimating the likelihood of a system disturbance propagating through different paths in the event tree was whether or not a system operator succeeded in making a given system situation less ‘stressed’. A lot therefore depended on experience of system operation. However, in the workshop, other experts’ knowledge – overhead line specialists, protection specialists, engineers studying system dynamic behaviours, etc. – was useful in the discussion in unpacking the phenomena and inter-dependencies.

The judgements being made extended beyond engineering phenomena. If the assumption that violations of voltage or frequency limits, poorly damped oscillations or overloads are preconditions for a whole system shutdown is correct then the key factors influencing the overall probability of a such shutdown are arguably:

1. the frequency with which violations occur;
2. the likelihood that a system operators fails to correct the violations and the system’s state deteriorates, either as a consequence of an initial violation or due to other unplanned, quasi-random events.

Human factors will undoubtedly be involved in both of the above and affect both the likelihood of a violation arising and the likelihood of a system operator correcting it. Reasons why a system operator might fail to correct a violation include a lack of situational awareness – for example, they failed to realise that there was a violation of an operating limit – or a lack of available actions. Building on the assumption that violations of operating limits are features of major disturbances and are present in their early stages, the expert judgement questions relating to violations were structured in terms of

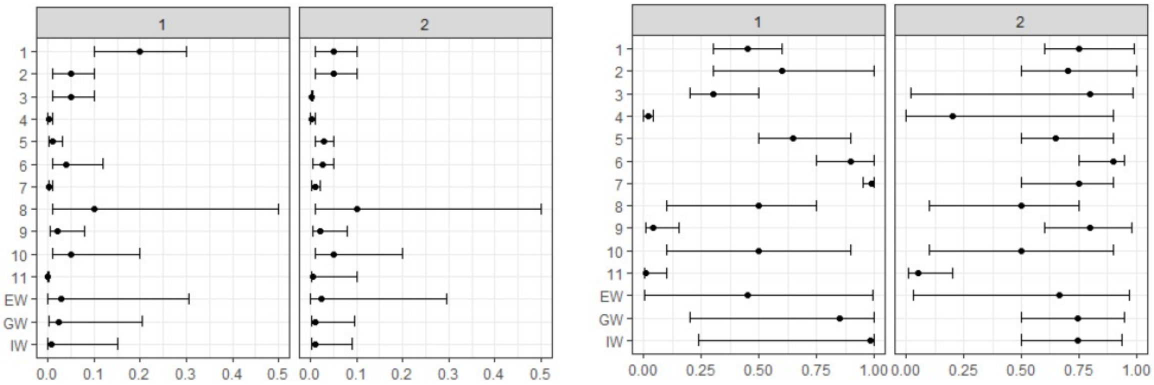
1. there being a violation (whether or not an operator sees it) and
2. whether the system operator succeeds in correcting it.

A lot of judgement was required in order to answer the expert elicitation questions, all wrapped up together. It was difficult for experts who are unfamiliar with probabilities or who have not thought through possible system failure mechanisms to reach a judgement for any one question.

With the benefit of hindsight, insufficient time was granted for the workshop. If the exercise was to be run again, much of the guidance that was given to the attendees only as part of the pack they received at the workshop would have been sent out in advance. In addition, there would have been more opportunity for attendees to raise queries privately during the workshop. To help to facilitate the latter, the project team should probably have comprised more than one power systems specialist.

6.2 The value of discussion and the weighting of answers

The value of the facilitated discussion between the first and second rounds of the IDEA protocol can perhaps be seen most clearly in the changes to experts' estimates of the answers to particular questions. For example. Figure 4(a) shows the different experts' answers to one particular question in each of the two rounds. It can be seen that experts 1, 3 and 5 made significant changes to their best estimates in the second round, following the discussion. Figure 4(b) shows the answers to a different question. It can be seen that, even after discussion, no consensus emerged. After facilitated discussion, experts 4 and 7 increased their uncertainty intervals substantially, suggesting they believe there is inherent uncertainty in the subject of the question.



(a) (b)

Figure 4: eleven experts' answers to two particular questions in each of the two rounds of elicitation, before and after the facilitated discussion. The points represent each expert's best etimate and the whiskers show the ranges within which each expert believes the true answer lies (uncertainty). EW is the equally-weighted aggregation of the experts' estimates. GW and IW are the performance-weighted aggregations of the experts' estimates using global weights and item weights respectively.

Figure 5 shows the answers to another question. It can be seen that expert 8's opinion was significantly different from that of the others, even after discussion. This expert also gave very wide uncertaitnty ranges, suggesting they believe that uncertainty is inherent in the subject of the question. In contrast, expert 4 gave a very wide uncertainty range in the first round but, after the facilitated discussion, provide a much narrower uncertainty range. This suggests that the additional information expert 4 gained during the discussion reduced their own uncertainty about the answer to the question.

Figures 4 and 5 show three methods for using the individual answers to arrive at one aggregated assessment for each question. Equal weights (EW) averages the distributions from each expert for each question; this essentially weights all experts equally regardless of their performance on the calibration questions. We also created two types of performance-weighted aggregated judgements, which are weighted averages of the individual expert assessments based on their ability to provide statistically accurate, informative responses to the calibration questions. Global performance weights (GW) uses a constant set of weights for the experts across all questions, based on their average performance scores. Item-based performance weights (IW), however, use question-specific information scores to determine a unique set of weights for each question. For example, if Expert A and Expert B are both equally statistically accurate across all of the calibration questions, but Expert A provided a more informative uncertainty assessment (i.e., a narrower range) for one question, then the item-based performance weights would assign Expert A a higher weight for that specific question. The different approaches to weighting can also be used to derive aggregated uncertainty ranges. The effect of using either of the performance weightings and the difference from equal weighting can be seen particularly clearly in the 2nd round of answers shown in Figure 4(b) where, in spite of the wide spread of answers among the experts, some clarity does emerge for the aggregated answer.

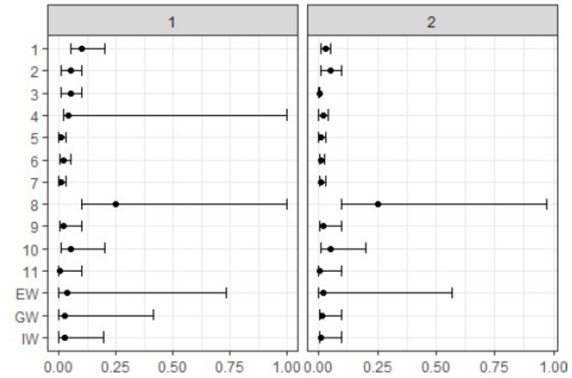


Figure 5: answers to one question showing expert eight's answer outlying the others.

6.3 The need for data

Although a lot of the numbers facilitating an estimate of the probability of system blackout came from expert judgement, there was still a need for data, for both the calibration questions and the frequency of occurrence of the 'precursors'. Access to the relevant data depends on the cooperation of a system operator or, potentially, a regulator. It is also depends on the effectiveness of the systems used by those parties to gather and store data, and whether the values of particular variables are routinely collected. If system operators are to make decisions in operational planning and real-time operation that take more explicit account of risk and the costs and benefits of different actions than they do now, they will need comprehensive plans in place for the collection and management of system data [17]. Furthermore, as has been done in the US, a regulator may oblige the system operator and network owners to collect and publish information such as the frequency of occurrence of system states such as the precursors identified as part of this work, the aim being to identify any trends that might indicate a need for particular interventions [18].

7 Power system issues highlighted by workshop participants

Beyond the basic aim of the exercise – to form an estimate of the probability of a major system unreliability event – it proved very useful as an opportunity to bring a group of experts together and facilitate thinking and exchange of views on existing and emerging system challenges. In this section, a number of the issues that were raised are briefly highlighted followed by a discussion of how the SEJ approach itself might be of further benefit to a system operator or network owner.

7.1 Changes to generation and demand

Power systems around the world have seen major changes to the generation mix with fossil fuelled thermal plant retiring, increased capacity of weather-dependent intermittent generation and an increase in the amount of distributed generation, i.e. generation connected within the distribution network, little of which in some countries is currently actively monitoring or controlled. The closure of thermal plant means losing the characteristics of the associated synchronous generators.

In the short to medium term, there is the possibility of extended outages for replanting of existing combined cycle gas turbines. The use of interconnectors as sources of power is growing in many countries. In Britain, this gives rise to issues associated with the characteristics of high voltage DC technology and the dependability of energy in the exporting market.

Significant changes to demand are either under way or are expected in many places. These include: less industrial demand; reduction in demand for reactive power; charging of electric vehicles (EVs); the possibility of more flexible/controllable demand, e.g. EV charging or home energy management systems; electrification of heat demand currently met by gas; and the possibility of aggregators and end use automation making more centralised decisions about demand and therefore reducing demand diversity.

7.2 System operation and control

Use of more automation, in particular greater reliance on corrective actions and distribution network 'active network management' entail an increasing reliance on remote control and communications with changes to the nature of the telecommunication infrastructure used within a power system. This potentially gives rise to increasing uncertainty about vulnerabilities of public communication systems. There is also high dependence on equipment vendor software, e.g. for control, much of which is 'black box' and little understood. There is sensitivity to software upgrades, in vendors' control systems, Energy Management Systems and offline analysis tools.

7.3 Resilience of electricity network assets

There is uncertain suitability of protection systems against the changing generation background. Network and generator assets are ageing with an uncertain effect on the probability of forced outages and the need for frequent, long duration maintenance outages. There is a risk of common cause failures arising from purchase of new assets from the lowest cost supplier of equipment and the possibility of early-life failures on flimsy new assets. Furthermore, as the climate changes, there is likely to be increased vulnerability of assets to extreme weather, e.g. flooding.

7.4 Electricity market drivers

The workshop attendees judged that the ability to reinforce the network is slower than that for new generation to be developed and that industry standards, e.g. the Grid Code, or ancillary service arrangements are not keeping pace with changes to the generation mix.

There was an expectation that once demand starts changing e.g. for EV charging, the speed of changes in demand will be faster than generation or network development; demand can change quickly (through incentivising electric vehicles, for example, or more automation of end use demand), but changing the system to meet the demand takes more time. Political changes such as increased nationalism were noted with the possibility of adverse impacts on security of supply in systems that depend on power imports either periodically or routinely.

The experts judged that there is regulatory and market pressure to decrease system margins and improve asset utilisation, e.g. to operate closer to ratings where ratings are defined on a real-time basis.

7.5 Other electricity sector issues

As was discussed in section 4, human factors play a big part in the resilience of a power system and how a disturbance is managed. The workshop attendees noted a loss of experienced staff from within the sector and an age gap in current staff: there are good young people, but middle-career people are missing. There is a lack of experience with emergency situations, in particular. The attendees also judged that the lack of experience leads to a reduced ability of staff to think 'outside the box' in unusual situations and a lack of implicit knowledge.

The attendees' experience of the electricity supply industry is that it has become increasingly segregated now with more training for a job, versus broader education for a career.

Trends towards what is referred to the UK as Distribution System Operators (DSO) were noted, i.e. where distribution network operators make many more active decisions to manage flows on their network: we do not yet know what a DSO will look like and how it will work and interact with other industry actors.

7.6 Future uses of the SEJ approach

The SEJ exercise described in this paper was focused on a specific question – what is the probability of a whole system blackout – that is very difficult to answer via conventional engineering modelling. The answer itself might be used to inform the business cases for potential investments in defence measures or system restoration facilities. Looking more closely at the results, the experts' combined judgements on the probabilities of the different individual pathways that can lead to a collapse might be used to target defence measures at particular phenomena or to inform training for operational staff. However, they might also be prompts to further expert workshops to explore more deeply why certain paths are viewed as being more likely than others and the reasons why a system operator might fail to address particular issues. This additional insight might, in turn, inform the development of particular facilities, revised industry standards or enhanced training. As compared to a bottom up reliability assessment approach, this type of exercise may provide better insights into systemic effects and challenges.

The experience gained from this particular exercise could be built upon by running a similar exercise again, applying the lessons learned in this one. This would allow an update to the results after some period of time, and provide a comparison with the results from the first exercise. The SEJ approach might also be suitable for allowing a system operator or network owner to assess other difficult to quantify risks such as those associated with climate change or, something explicitly excluded from the exercise reported here, to assess a power system's different areas of vulnerability to malicious attack.

8 Conclusions

This paper has described an approach to the estimation of the probability of a major power system unreliability event. The general approach – use of Structured Expert Judgement (SEJ) – has been used in many fields in which major risks need to be understood in respect of phenomena that are not readily amenable to modelling. The exercise reported here is what the authors believe to be the world's first application of SEJ to the understanding and enhancement of power system resilience.

A conceptual model, believed to be unique, was developed describing how a system blackout might happen. It was used in a 1.5 day workshop with a group of experts to provide estimates of the likelihood of a system disturbance propagating along different paths.

Experience from running the workshop showed that while individuals with system operation experience have the best feel for the key phenomena and how a system operator deals with them, other experts' knowledge is useful in unpacking the phenomena and inter-dependencies and helping to inform others in the workshops. These experts include specialists in overhead lines, system and generator protection specialists, and system dynamic behaviours.

One and a half days was not really enough time to complete the exercise comfortably. The judgements requested were difficult for experts who are unfamiliar with probabilities or who have not previously thought through possible system failure mechanisms. There was also still a need for good data: for the calibration questions and the frequency of occurrence of the ‘precursors’ on which the event tree was based.

As well as providing an estimate in respect of the main question, the exercise provided a platform for discussion of what the experts believed to be key current and emerging system issues.

The answer itself might be used to inform the business cases for potential investments in defence measures or system restoration facilities. Looking more closely at the results, the experts’ combined judgements on the probabilities of the different individual pathways that can lead to a collapse might provide a platform for deeper analysis of those phenomena or allow the targeting of development of particular facilities or staff training.

Acknowledgements

The authors would like to thank the various experts - the reviewers of the high level model and the participants in the workshop - for their generously given time and insights.

BIBLIOGRAPHY

[1] S. French, “Aggregating Expert Judgement”, *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales*, 105(1): 181–206, 2011.

[2] R.M. Cooke, *Experts in Uncertainty*, Oxford, Oxford University Press, 1991.

[3] L.H.J. Goossens and G. N. Kelly, “Special: Issue: Expert Judgement and Accident Consequence Uncertainty Analysis”, *Radiation Protection Dosimetry*, 90(3): 293-381, 2000.

[4] Colson, A. R., Megiddo, I., Alvarez-Uria, G., Gandra, S., Bedford, T., Morton, A., Cooke, R. M., & Laxminarayan, R., “Quantifying uncertainty about future antimicrobial resistance: Comparing structured expert judgment and statistical forecasting methods”, *PLOS ONE*, 14(7), e0219190, 2019. https://doi.org/10.1371/journal.pone.0219190

[5] Lutter, R., Colson, A. R., & Cooke, R. M., “Effects of Increases in IQ in India on the Present Value of Lifetime Earnings: A Structured Expert Judgment Study”, *Resources for the Future*, 2017 http://www.rff.org/research/publications/effects-increases-iq-india-present-value-lifetime-earnings-structured-expert

[6] Barons, M. J., Hanea, A. M., Wright, S. K., Baldock, K. C. R., Wilfert, L., Chandler, D., Datta, S., Fannon, J., Hartfield, C., Lucas, A., Ollerton, J., Potts, S. G., & Carreck, N. L., “Assessment of the response of pollinator abundance to environmental pressures using structured expert elicitation”, *Journal of Apicultural Research*, 57(5), 593–604, 2018. https://doi.org/10.1080/00218839.2018.1494891

[7] Bamber, Jonathan L., Oppenheimer, M., Kopp, R. E., Aspinall, W. P., & Cooke, R. M., “Ice sheet contributions to future sea-level rise from structured expert judgment”, *Proceedings of the National Academy of Sciences*, 116(23), 11195–11200, 2019. https://doi.org/10.1073/pnas.1817205116

[8] A. O'Hagan, C. E. Buck, A. Daneshkhah, R. Eiser, P. H. Garthwaite, D. Jenkinson, J. E. Oakley and T. Rakow, *Uncertain Judgements: Eliciting Experts' Probabilities*, Chichester, John Wiley and Sons, 2006.

[9] A. Hanea, M. McBride, M. Burgman and B. Wintle, “Classical meets modern in the IDEA protocol for structured expert judgement”, *Journal of Risk Research*, 21(4): 417-433, 2018.

[10] V. Hemming, M. A. Burgman, A. M. Hanea, M. F. McBride and B. C. Wintle, “A practical guide to structured expert elicitation using the IDEA protocol”, *Methods in Ecology and Evolution*, 9(1): 169-180, 2018.

[11] M. Papic, K. Bell, Y. Chen, I. Dobson, L. Fonte, E. Haq, P. Hines, D. Kirschen, X. Luo, S. S. Miller and N. Samaan, “Survey of tools for risk assessment of cascading outages”, *IEEE Power and Energy Society General Meeting*, 2011.

[12] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges", *IEEE Transactions on Power Systems*, 27(2): 631, 2012.

[13] AEMO, *Black System South Australia 28 September 2016: Final Report*, March 2017

[14] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik and J. Sanchez-Gasca, “Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance”, *IEEE Transactions on Power Systems*, 20(4): 1922-1928, 2005.

[15] P. Kundur, C. Taylor and P. Pourbeik, *Blackout experiences and lessons, best practices for system dynamic performance, and the role of new technologies*, IEEE Task Force Report, 2007.

[16] CIGRE Working Group C1.17, *Planning to Manage Power Interruption Events*, Technical Brochure 433, CIGRE, Paris, 2010.

[17] GARPUR Project, *A New Approach to Reliable and Affordable Electricity Supply in Europe: Project Results*, 2017. Available: https://www.sintef.no/projectweb/garpur/

[18] North American Electric Reliability Corporation, *Reliability Assessment and Performance Analysis*, https://www.nerc.com/pa/RAPA/Pages/default.aspx