Statistical modelling of dependence between net demands and deficits in two area power systems

Statistical modelling of dependence between net demands and deficits in two area power systems

Nestor Sanchez[a], Chris J. Dent[a], Amy Wilson[a]

[a]*School of Mathematics, University of Edinburgh, Edinburgh, U.K.*

Abstract

Power system resource adequacy risks is dominated by the extremes of the relevant distributions or processes: the upper tail of demand and the lower tail of available generation. Because of this, relevant data in the historic record are sparse. Moreover, for interconnected systems, the degree of statistical dependence in these extremes between systems can also have a sizeable impact on the level of risk in each system. This paper uses results from statistical extreme value theory (EVT) to fit smoothed joint distributions for the values of (demand minus available renewables) and of surplus/deficit dependence in a two-area system, using data from the Irish and Great Britain power system for examples.

As well as the statistical smoothing mitigating the consequences of limited volumes of data, the concept of asymptotic dependence provides a useful explanation of the strength of dependence. There is strong evidence that the deficits in the GB and Ireland systems are asymptotically independent, whereas there is evidence for asymptotic dependence in the distributions of (demand minus available wind capacity). This is consistent with the intuition that, when independent distributions of conventional capacity are convoluted with those of demand and wind, the dependence is weakened. The consequent use of a Gaussian copula to describe the dependence of deficits provides a convenient means of carrying out sentivity analysis to the strength of relationship.

Keywords: resource adequacy, two area systems, asymptotic dependence, extreme value theory, renewables, interconnection

Email address: Chris.Dent@ed.ac.uk (Chris J. Dent)

1. Introduction

Security of supply is one of the three pillars of energy policy, along with affordability and sustainability. In Great Britain (GB), annual capacity procurement auctions take place to ensure that capacity meets reliability standards based on risk estimates for future years [1]. For this reason, risk calculations and their statistical methodology are of considerable interest for policy makers. Such estimates need to consider the existence of interconnectors to other power systems, which potentially represent an additional capacity source; this is particularly relevant in the context of decarbonisation, as the aggregation of renewable generation over wider geographical regions is expected to result in less variable power output. On top of this, there are other benefits regarding market integration [2] like economies of scale from shared planning over wide regions.

Multi-area reliability has been studied since at least the 1960s; in [3] calculations are shown for a two-area system with a no-shortfall-sharing interconnection policy and an empirical model of daily peak demand. In [4] a model is proposed for a general multi-area system, possibly with loops, assuming statistical independence across areas and considering both shortfall-sharing and non-sharing policies. Methodology can vary considerably and depend on factors such as the system's topology and desired level of system detail; methods range from failure states partitioning and aggregation to simulation-based methods, or a combination of multiple approaches [5]. This variety is reflected in the range of practices found across different system operators; in [6] an international survey was conducted by National Grid and it was found that some system operators such as those in France and Belgium as well as some parts of the US, have detailed models of neighbouring systems and use them as input in their reliability studies, while others take interconnector imports as fixed. In the case of Great Britain, an off-line market model is used to specify a distribution of the interconnector's flow conditioned on different load percentiles, and this distribution is then convolved with the overall supply balance thereafter. In this work, we use results from the statistical theory of extremes to propose a principled two-area probability model for events that determine system reliability, e.g. high load or low capacity surplus events. Models for extreme value theory (EVT) have been used before in a single-area system model using data for GB [7]; we build on this work by developing appropriate statistical dependence models for extremal dependence between two areas.

Most of the risk regarding capacity shortfalls come from rare events either at the demand side, which can be unusually high perhaps due to cold weather in the case of Great Britain (GB), or at the generation side, where conventional generators can fail unexpectedly and renewable generation can experience sustained drops in generation output due to certain weather patterns [8]. Interconnected systems are usually geographically close, and so their demand and renewable generation output is likely to be statistically dependent through the action of weather systems and similar energy use patterns. Simultaneous occurrence of extreme demand and renewable generation in both areas can have a large effect on the value of interconnection in terms of security of supply; under strong dependence, it is more likely that both systems experience stress at the same time.

It is a common practice in reliability studies both in industry and in the literature to use the empirical distribution of historic load as input, either suitably rescaled or as is [6, 9, 10]. Empirical distributions are sparse in the tails, with all of the probability mass in these regions concentrated in only a handful of points representing previously observed extreme levels. As risk comes almost entirely from the tails, a consequence of this is that risk indices can be essentially determined by a very small set of past observations [11], potentially making risk estimates unreliable. A similar argument holds when it comes to multi-area studies as post-interconnection risk are driven by an even smaller number of points when using hindcast models (see Figure 1), in particular those in which extreme values co-occur. Models from EVT provide a sound methodological basis for finding smooth parametric alternatives to the sparse tail regimes of hindcast models, and to characterise the tails of the involved distribution and the dependence between them (i.e., the extremal dependence) with a small number of interpretable parameters. This in turn makes it straightforward to perform a sensitivity analysis of interconnection value.

The outline of the paper is as follows: in Section 2, we give a brief review of standard power system risk models and indices that we use in this work. In Section 3, motivation for the use of extreme value theory is given, and the most relevant results for this work are briefly discussed, referencing more comprehensive material on the subject; Section 4 describes the data used. In Section 5 we describe the methodology for the two-area net demand model, and present results comparing this to a hindcast net-demand model; in Section 6 we then analyse statistical dependence in capacity surpluses, characterising dependence between systems and using this to perform

sensitivity analysis of different risk indices to teh strength of statistical dependence. Finally, Section 7 provides conclusions.

2. System model

In this section, we give a brief overview of some of the models from the literature of security of supply that will be using in this work. We denote random variables with uppercase letters and constants with lowercase letters; vectors will use bold letters. We use X, Y and D to denote random variables corresponding to available conventional generation, renewable generation and demand respectively. Furthermore, we define net demand, or demand net of renewables, as $D' = D - Y$, and capacity surplus as $Z = X - (D - Y)$.

2.1. Single area system

Let the period of interest (say, a peak season) be divided into n hourly segments. We will consider two risk indices, the loss of load expectation (LOLE), which is the expected number of hours[1] in which there is a shortfall, defined as

$$\text{LOLE} = \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{I}(Z_t < 0)\right] = \sum_{t=1}^{n} \mathbb{P}(X_t < D'_t), \tag{1}$$

and the expected energy unserved (EEU), which is the expected amount of energy not supplied, defined as

$$\text{EEU} = \mathbb{E}\left[\sum_{t=1}^{n} \max\{0, -Z_t\}\right]. \tag{2}$$

Here t indexes times in the future season or year under study. For the purposes of statistical modelling, as in this paper, it is often more convenient to work in a *time-collapsed* picture with the time-collapsed variable Z representing surplus at a randomly chosen point in time in the peak season under

[1]For simplicity, all results presented here are for an hourly time step, as per the available data from GB – in N American terminology this corresponds to LOLH. The results generalise in a straightforward way to calculations for higher time resolutions. More detail of definitions of LOLE for different time resolutions may be found in the IEEE Working Group paper [12]

study. For a system such as that considered here, which does not have storage or other technologies which link time periods in a similar way, the LOLE is then calculated as

$$\text{LOLE} = n\mathbb{P}(Z < 0) = n\mathbb{P}(X < D'), \tag{3}$$

and an analogous formula applies for EEU.

The most common means of estimating the distribution of net demand is to use the empirical historic distribution as the distribution for predictive risk calculations, sometimes referred to as *hindcast* [13, 14]. The distribution of net demand is then given by

$$\mathbb{P}(D - Y \leq w) = \frac{1}{T} \sum_{\tau=1}^{T} \mathbb{I}(d_\tau - y_\tau \leq w). \tag{4}$$

where T is the number of observations in the historic record, τ indexes the historic records, and historic demand and wind resource are rescaled appropriately to the future system scenario under study. This approach may also be interpreted as estimating the risk conditional on a repeat of historic conditions in one or more years.

Available conventional generation at time t, X_t, is modelled as the sum of individual generating units' availabilities; conventional generators are assumed to be either fully available or unavailable with a given probability derived from historic data, and availabilities of different units are assumed to be independent of each other and of all other variables. Within the examples presented, risk indices that depend on the distribution of Z_t may then be calculated by convolving the distributions of X_t and D'_t. However, a time-sequential model would in general be needed to consider technologies such as storage that link time periods in the calculations, even when only expected value indices such as EEU and LOLE are evaluated.

2.2. 2-area system

In a 2-area system, interconnection can make imports available to a system under stress, up to the interconnection capacity, thereby reducing risk for both systems. The impact of interconnection on the risk indices depends not only on interconnection capacity and statistical dependence across areas, but also on the shortfall-sharing policy, understood here as an agreement of how the interconnector is used when one or both systems are under stress.

Various policies have been studied in the literature; for instance, [15] discusses so-called *veto* and *share* policies, the former being that in which a power system exports only spare available capacity, if any, while in the latter case shortfalls are shared across areas according to an agreed rule. In this work, we assume a *veto* policy between both areas, as this will suffice to illustrate the statistical approaches developed. If $\mathbf{Z} \in \mathbb{R}^2$ is the bivariate pre-interconnection vector of surpluses across both systems, and there is an interconnector with capacity $c \geq 0$ between them, the post-interconnection shortfall region for area 1 under a veto policy can be divided in two subregions:

$$\mathcal{R}_1 = \{\mathbf{Z} \in \mathbb{R}^2 \,|\, \mathbf{Z}_1 < -c\} \tag{5}$$

$$\mathcal{R}_2 = \{\mathbf{Z} \in \mathbb{R}^2 \,|\, -c \leq \mathbf{Z}_1 \leq 0, \mathbf{Z}_2 < -\mathbf{Z}_1\} \tag{6}$$

which represent the cases where area 1 has a shortfall larger than interconnection capacity, and the case of area 2 not being able to cover a shortfall in area 1, despite the shortfall being smaller than interconnection capacity.

Under a snapshot model the two risk indices may be calculated as follows for area 1:

$$\text{LOLE} = n \cdot \mathbb{P}(\mathbf{Z} \in \mathcal{R}_1 \cup \mathcal{R}_2) \tag{7}$$

$$\text{EEU} = n \cdot \mathbb{E}[-\mathbf{Z}_1 \cdot \mathbb{I}(\mathbf{Z} \in \mathcal{R}_1 \cup \mathcal{R}_2)] \tag{8}$$

where $\mathbb{I}$ is an indicator function. This is, there is only a shortfall in area 1 when the pre-interconnection surplus vector $\mathbf{Z}$ is in $\mathcal{R}_1 \cup \mathcal{R}_2$. For area 2, the reasoning is analogous with reversed indices.

3. Extreme value theory

In terms of LOLE and EEU, virtually all of the risk is concentrated in the tails of net demand, that is, it comes from the highest net demand values. As mentioned before, this concentration is especially severe in empirical hindcast models of net demand, where risk indices can be determined by a very small number of historic records with the highest net demand observations. In order to alleviate this problem, we turn to EVT-based models which provide a smooth parametric alternative to hindcast models.

EVT is a branch of statistics that provides mathematically principled methods for making inferences about statistical properties of extreme events,

including those rare enough to be outside the historic record's range; these models arise as limiting behaviour of sample maxima or exceedances above progressibly large levels. EVT makes only mild regularity assumptions on the data while providing general results on extreme occurrences of random variables. It is widely applicable and routinely used in fields from insurance to environmental sciences [16, 17]. In this work we are interested in the threshold-exceedance framework which we outline below. Then, we briefly discuss relevant results in multivariate EVT.

3.1. Univariate exceedances

A key result from EVT states that under mild assumptions on the distribution of a random variable $X \sim F(x)$, exceedances over a threshold u, given by the conditional distribution $X\,|X > u$, follow a Generalised Pareto distribution (GPD) in the limit, as $u \to \infty$ [17]. For the purpose of this work, this result means that for an appropriately large threshold u, net demand exceedances over said threshold, conditioned on $D - Y > u$, are well approximated by a GPD, whose cumulative function is given by

$$\mathbb{P}(W \leq w) = 1 - \left(1 + \xi \left(\frac{w-u}{\sigma}\right)\right)^{-1/\xi}, w > u \tag{9}$$

where $\xi \in \mathbb{R}$ and $\sigma > 0$, and reduces to an exponential distribution if $\xi = 0$.

A semiparametric model for the full data range can then be constructed by using the fitted GPD model for tail exceedances above the chosen threshold u, and the empirical data distribution $\hat{F}$ below it. The full model can thus be written as

$$\bar{F}(Z \leq z) = \begin{cases} \hat{F}(z) & z \leq u \\ \hat{F}(u) + (1 - \hat{F}(u)) * F_{GP}(z) & z > u \end{cases} \tag{10}$$

where F_{GP} is the fitted GPD. A model like this has already been applied to a single-area power system using data from GB [7].

3.2. Asymptotic dependence and multivariate EVT

When dealing with extremes of multiple random variables, as is the case in a two-area system, EVT also provides a framework to measure the degree of association between different components at extreme levels, which in our problem directly influences the utility of interconnection in terms of security of supply.

A central concept in this context is asymptotic dependence, which quantifies the degree to which extreme values at different components occur together. Assuming without loss of generality (as the choice of marginal distributions do not change the dependence structure) that Y_1, Y_2 are standard Frechet random variables with marginal CDF $F(y) = \exp(-1/y), y > 0$, the statistic χ is defined as

$$\chi = \lim_{y \to \infty} \mathbb{P}\left(Y_2 > t \,|\, Y_1 > y\right) \tag{11}$$

The variables are asymptotically dependent if $\chi > 0$, and asymptotically independent otherwise. Asymptotically independent variables can still exhibit strong dependence at non-extreme levels, with dependence vanishing only in the tails; an example of this is the bivariate normal distribution with correlation $-1 < \rho < 1$ whose components can be proven to be asymptotically independent, regardless of ρ [16, p. 285]. Determining the presence of asymptotic dependence is important in devising an appropriate model for the data; there are many well studied parametric models for asymptotically dependent data, while asymptotically independent data may require semi-parametric or non-parametric modelling approaches instead, as done in [18].

Although a useful theoretical concept, precise estimation of χ is not straightforward, and typically visual inspection of empirical approximations are used to assess asymptotic dependence. A related statistic that is more amenable to numerical estimation is the coefficient of tail dependence η [16, p. 345]; this is bounded by $0 \leq \eta \leq 1$, and is defined by making an additional assumption in the context of (11), namely

$$\mathbb{P}(Y_1 > y, Y_2 > y) = \mathcal{L}(y) \cdot \mathbb{P}(Y_1 > y)^{1/\eta}, y > 0 \tag{12}$$

where $\mathcal{L}(x)$ is a slowly varying function, that is, $\mathcal{L}(yz)/\mathcal{L}(y) \to 1$ as $y \to \infty$ for all $z > 0$. Intuitively, slowly varying functions become flat rapidly (in a precise sense) as $y \to \infty$ (note that $\mathcal{L}$ does not necessarily converge, e.g. $\mathcal{L}(z) = \ln(z)$). This class of functions arise in many important results from the theory of extreme values, and (12) has been shown to be valid for a broad range of conditions and models [19, 20]; here, η describes the type of dependence and $\mathcal{L}(y)$ its strength within the dependence type given by η. We have $\eta = 1$ whenever $\chi > 0$, and $0 \leq \eta < 1$ otherwise, thus characterising the presence of asymptotic dependence; furthermore, we can estimate η by maximum likelihood as the shape parameter ξ in (9) from a sample of $Z = \min\{Y_1, Y_2\}$. This in turn allows us to estimate η for our

data by transforming it to approximate standard Frechet margins using the transformation $v_i = 1/(1 - \bar{F}(y_i)), i = 1, ..., n$ for the net demand samples of each individual system, with $\bar{F}$ as in (10), and using the component-wise minima of the transformed data $z_i = \min\{v_i^{(1)}, v_i^{(2)}\}, i = 1, ..., n$, where superscripts correspond to each of the systems.

3.3. Models of extremal dependence

Models of statistical dependence can be described using copula functions. A copula is a multivariate distribution with uniform marginals in $[\mathbf{0}, \mathbf{1}]$, and can be constructed for any given distribution simply by transforming its marginal distributions to standard uniform. Any multivariate distribution can be written in terms of its univariate marginal distribution functions and a copula function, see [16, p. 272] (which references [21] as the original source of this result).

Copulas arising from dependence between extremes are called extreme value copulas. More specifically, $C^*(\mathbf{u})$ is an extreme value copula if there is a copula $C(\mathbf{u})$ such that

$$C^*(\mathbf{u}) = \lim_{n\to\infty} \left(C(\mathbf{u}^{1/n})\right)^n \tag{13}$$

This means the copula between component-wise sample maxima from $C(\cdot)$ converges to $C^*(\cdot)$ as the sample size goes to infinity, and we say that $C(\cdot)$ is in the *domain of attraction* of $C^*(\cdot)$. Sometimes it is more convenient to characterise extreme value copulas through their *Pickands dependence function* $A(t), 0 \leq t \leq 1$ [16, p. 285]. This is a convex function bounded by $\max\{t, 1-t\} \leq A(t) \leq 1$, and has a one to one relationship with $C^*(\mathbf{u})$. In the two-dimensional case any such function induces an extreme value copula, and we can write

$$C(u, v) = (uv)^{A(\log(v)/\log(uv))} \tag{14}$$

From the above we can see that $A(t) = 1$ produces independent components. Conversely, $A(t) = \max\{t, 1-t\}$ produces perfectly dependent components.

One of the simplest parametric model of extremal dependence is the so called *logistic* model, defined by

$$A(t) = \left(t^{1/\alpha} + (1-t)^{1/\alpha}\right)^{\alpha}, \ \alpha \in (0, 1] \tag{15}$$

This Pickands function induces a Gumbel-Hougaard copula, given by

$$C_(u, v) = \exp\left(-\left((-\log u)^{1/\alpha} + (-\log v)^{1/\alpha}\right)^{\alpha}\right), \ \alpha \in (0, 1] \tag{16}$$

which has been shown to arise as an extremal dependence structure from a wide family of copulas in the underlying distributions [22]. However, from (15) we see that this model entails symmetry in $A(t)$ in the sense that $A(1/2-t) = A(1/2+t)$ for all $0 \leq t \leq 1/2$, which is not always appropriate. A more flexible generalisation of this model is the *asymmetric logistic model* whose Pickands function in the bivariate case is given by

$$A(t) = (\psi_2 - \psi_1)t - \psi_2 + 1 + \left((\psi_1 t)^{1/\alpha} + (\psi_2(1-t))^{1/\alpha}\right)^{\alpha} \qquad (17)$$

with $0 \leq \psi_1, \psi_2, \alpha \leq 1$. This reduces to the logistic model if $\psi_1 = \psi_2 = 1$, and to independence if $\alpha = 1$ or $\psi_1 = 0$ or $\psi_2 = 0$.

Another advantage of working with Pickands functions is that they provide a way to visually inspect the goodness of fit of a model for extremal dependence by comparing the fitted model's Pickands function to the empirical approximation induced by the data, as in Fig. 4.

Note that most if not all parametric models of extremal dependence make the asumption of asymptotic dependence. This is because the only possible limiting behaviour from asymptotically independent copulas is full independence; this means that for, say, normally distributed data with correlation $-1 < \rho < 1$, the dependence between components weakens at progressively more extreme levels and disappears completely in the limit, reducing to completely independent components, and hence limiting parametric dependence models are not very useful in this case. However, some useful semiparametric approaches have been explored [18].

4. Data

Wind data were obtained from [23]. It consists of hourly wind capacity factors based on atmospheric reanalysis data and the locations of installed generation on January 2015, and for this work it has been rescaled to a total wind capacity of 3 GW in IRL and 15 GW in GB. Different installed wind capacities can be obtained for purpose of numerical experimentation by rescaling by a constant factor.

Demand data consist of hourly measurements for the peak seasons of 2007 to 2013 for both systems; GB demand data was obtained from [24] while data for IRL was provided by Baringa Ltd; the data has been standardised by rescaling each peak season by their corresponding Average Cold Spell estimates as in reference [25] to correct for external factors such as economic

growth but preserving variation due to weather patterns [7]. Subsequently, all normalised peak seasons were again rescaled to fix the average seven-year period LOLE to three hours per year.

Capacity and failure probability data for conventional generating units in Ireland were developed by Baringa Ltd for the Irish Single Electricity Market Committee in 2016, and are publicly available at [26]; in the case of Great Britain, data were provided by National Grid, and were anonimased to protect its sensitivity; we take the anonymised data as representative of the real system.

Finally, in this work, we refer to each historic peak season in the data by the year at which the season started, so for example we refer to the 2007-2008 winter just by 2007. We illustrate results on a subset of years in the body of the paper, in the interests of space; the full series of plots for all years is shown in the online supplement.

5. Modelling net demand

As mentioned above, we assume that available conventional generation is independent between the systems, and so dependence between the systems' capacity surpluses or deficits comes entirely from net demand. The main motivation in looking for a smooth alternative to a hindcast net demand model is its tendency to produce risk estimates that are almost entirely determined by a very small number of points. Figure 1 shows the concentration of LOLE in the highest net demand observations for each season under a hindcast model. For instance, for 2011 in GB roughly 80% of the estimated pre-interconnection LOLE comes from just eight observations. Moreover, this concentration is exacerbated in post-interconnector calculations, particularly for IRL which is smaller relative to interconnection size. The use of smooth, parametric models in the regions of interest could offer a more balanced alternative.

As the influence of net demand on shortfall risks is almost completely limited to the upper tail of net demand data in both systems, i.e. from the highest values of (demand minus available wind), this section describes the application of methodology from the theory of extremes to model net demand tail events in the two-area system, modelling each peak season separately but using the same methodology for all of them. Then, we compare risk estimates of this model to those of an empirical (hindcast) model of net demand. All of

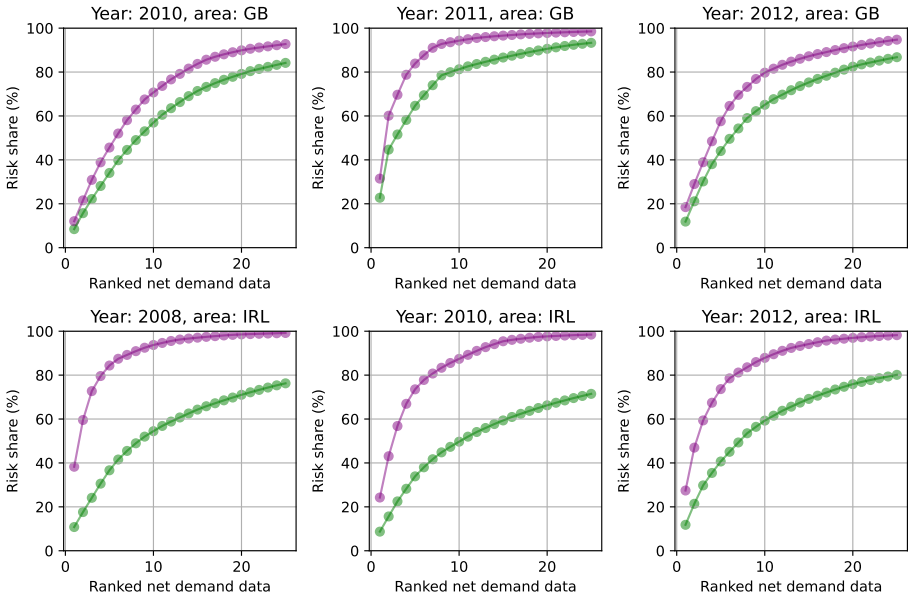


Figure 1: Attribution of pre-interconnection (green) and post-interconnection(purple) LOLE proportion for the highest net demand measurements in each peak season under a hindcast net demand model. For instance, just three demand measurements virtually determine post-interconnection LOLE in 2012 for IRL (lower right corner).

this was done using a bespoke Python package `riskmodels` which is publicly available through the PYPI repository [27].

5.1. Parametric models for net demand extremes

5.1.1. Fitting GPD tail models

We first fit univariate Genralised Pareto distributions for each system using the largest net demand observations in the peak season under consideration; in order to set exceedance thresholds, we follow the approach described in [17] using mean residual life plots, concluding that 95% quantile thresholds are appropriate for both areas in all individual peak seasons; then, using the exceedances above said thresholds we fit generalised Pareto models. For both areas and all peak seasons we observe negative fitted shape parameters (ξ in (9)); this means net demand data has light tails in the sense that the fitted models have a finite upper endpoint. Fig. 2 shows Q-Q plots for the fitted tail models in both areas; we observed similar goodness of fit for net demand in all years. Finally, fitted parameters are shown in Table 1.

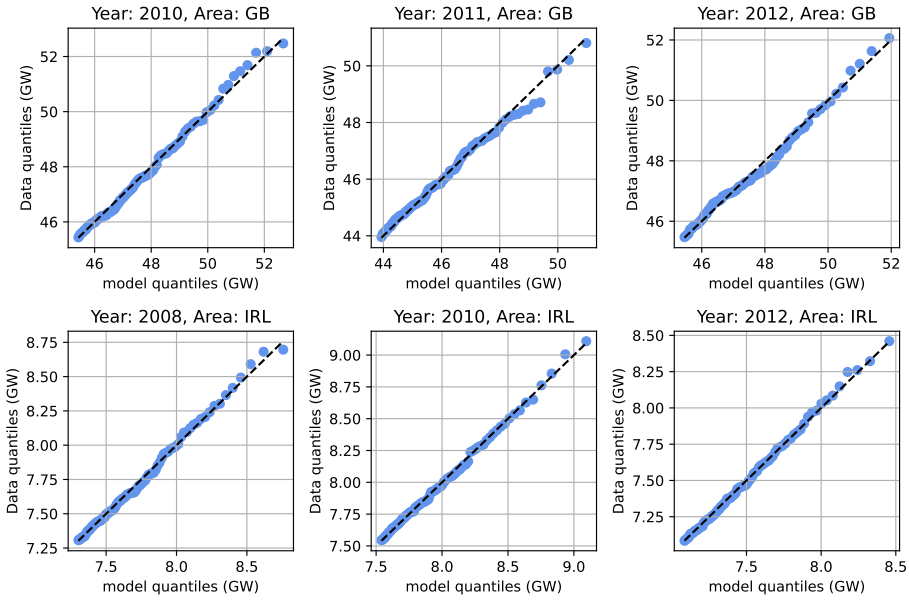
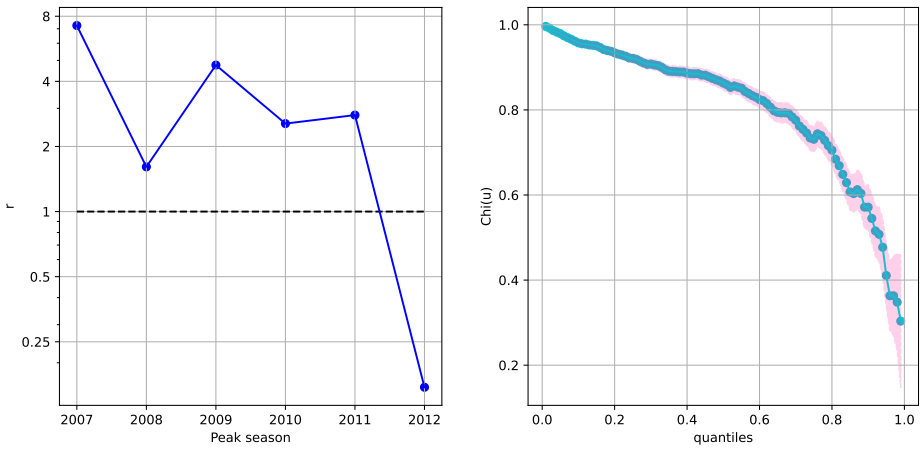


Figure 2: Q-Q plots for tail models in both areas. Data shown are exceedances above the 95% quantile threshold for each peak season, and the fitted models are generalised Pareto distribution.

5.1.2. Testing for extremal dependence in net demand

Empirical approximations for χ were consistent with the hypothesis of asymptotic dependence for most years, as the empirical realisation of (11) keep consistently away from zero for the largest percentiles (see Fig. 3b), hence making the hypothesis of having $\chi > 0$ plausible. However a more quantitative answer can be obtained performing a Bayesian ratio test to test both hypothesis, namely asymptotic dependence and asymptotic independence. We do this through estimation of the coefficient of tail dependence η as described in Section 3.2 using the 95% quantile as threshold for both systems; as the hypothesis of asymptotic dependence consists of a single point in the hypothesis space, namely $\mathcal{H}_0 : \eta = 1$ vs $\mathcal{H}_1 : \eta \neq 1$, $\eta \in [0, 1]$, we calculate the Savage-Dickey ratio r, defined as the ratio between the posterior and prior densities for the value of interest [28, 29], in this case $\eta = 1$. This can be thought of as the limiting value of a Bayes ratio test when the null hypothesis subset shrinks to a single point; the test value itself, r, equals the ratio of posterior to prior odds for $\mathcal{H}_0$ [30].

To perform the test, we set $\mathbb{P}(\mathcal{H}_0) = \mathbb{P}(\mathcal{H}_1) = 0.5$ and use a Jeffreys prior



(a) Savage-dickey ratios for the Bayesian ratio tests in Section 5.1.2. Values larger than one favor the hypothesis of asymptotic dependence in net demand data for the corresponding peak season.

(b) Empirical estimates of asymptotic dependence (11) calculated with increasingly larger empirical quantiles. Data from 2010 was used but the figure is representative for what we observed in other years.

Figure 3: Diagnostics for asymptotic dependence in net demand data.

on both parameters η and σ; Jeffreys priors are invariant under reparametrisation [31], which makes them a robust choice when no additional information is available. We restrict η to $[0, 1]$ and treat $\sigma > 0$ as a nuisance parameter. Under this setting, r values larger than 1 make $\mathcal{H}_0$ more credible than the alternative, and we find this this to be the case for all individual years except 2012 (Fig. 3a). Note that even though some values might appear to provide only weak evidence for $\mathcal{H}_0$, the fact that it is at the very edge of the hypothesis space might make the posterior converge slowly to $\mathcal{H}_0$ even when it is true, and indeed we observed similar values for r when using synthetic data for which $\mathcal{H}_0$ was the correct choice; moreover, the use of uniform priors did not alter the conclusions of the tests. We thus conclude asymptotically dependent extreme value models are appropriate for net demand data.

Lastly, as asymptotic dependence was the favored hypothesis for all other years, for the sake of simplicity we use the same model for 2012 too, rather than alterantives such as [18]. There is further discussion of the validity of this model of asymptotic dependence in Section 6.1 (which considers dependences between surpluses, as opposed to the net demands studied here). The very strong evidence seen there against asymptotic dependence in sur-

plus/deficit provides further support for using a different model with asymptotic dependence for the case of net demands.

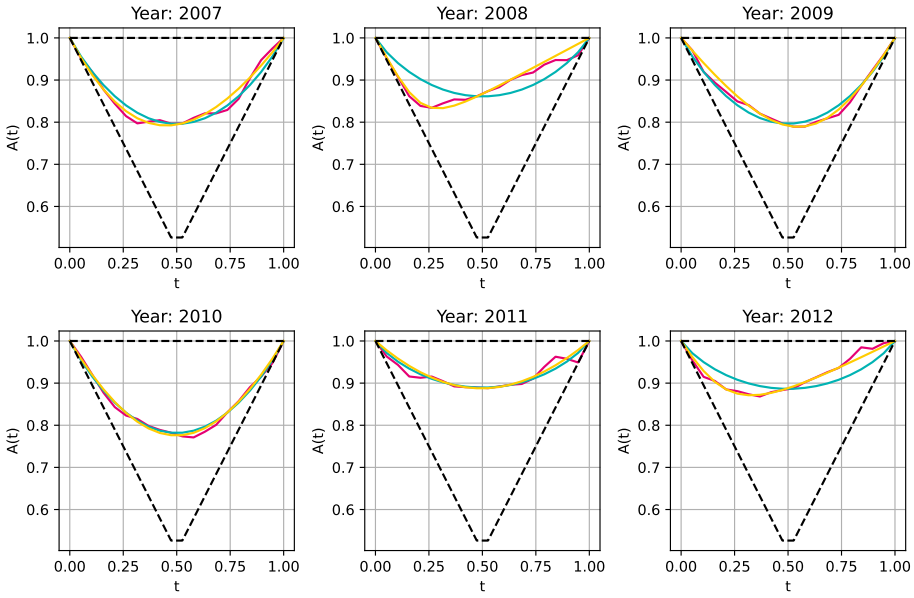


Figure 4: Comparison of the fitted logistic (teal) and asymmetric logistic (yellow) Pickands functions to the empirical Pickands function (red) for each peak season.

5.1.3. Fitting the dependence model

To fit the dependence models we use the same modelling thresholds as for the marginal exceedance models, and we consider logistic and asymmetric logistic dependence (Section 3.3). A visual comparison of the empirical Pickands approximation induced by joint exceedance data (as calculated in [32]) against those of the models fitted on joint exceedances (Fig. 4) suggests that both models provide an appropriate descriptions of the data, except for 2008 where the asymmetric logistic produces a better fit. However, the effect of the choice between these two models turns out to be minimal in the consequent results, and so we choose a logistic model for parsimony.

Lastly, the model is fitted in all of the exceedance region, i.e. that in which at least one exceedance in the two areas occur. Fitted dependence parameters are shown in Table 1.

5.1.4. The fully fitted model

The fully fitted model is semiparametric and comprises the empirical data distribution for the non-extreme region $\mathbf{Y} \leq \mathbf{u}$, with threshold vector $\mathbf{u}$ given in Table 1, and an exceedance model defined on the region $\mathbf{Y} \not\leq \mathbf{u}$. The exceedance model's marginal distributions are as in (10) with fitted tail parameters as in the referenced table, and dependence in this model is governed by the fitted extreme value copula, restricted to the corresponding exceedance region in copula space.

season	GB (u, σ, ξ)	IRL (u, σ, ξ)	dependence (α)
2007	(43.18, 2.78, -0.33)	(6.96, 0.37, -0.08)	0.53
2008	(44.7, 2.58, -0.31)	(7.3, 0.5, -0.21)	0.51
2009	(45.42, 2.05, -0.23)	(7.44, 0.48, -0.28)	0.47
2010	(45.4, 2.9, -0.3)	(7.54, 0.5, -0.18)	0.46
2011	(43.91, 2.66, -0.26)	(6.7, 0.52, -0.2)	0.56
2012	(45.43, 2.41, -0.26)	(7.08, 0.47, -0.21)	0.6
2013	(42.45, 2.48, -0.36)	(6.85, 0.53, -0.27)	0.5

Table 1: Table with fitted model parameters for all peak seasons. GB and IRL columns show the fitted parameters for the respective univariate exceedance distributions; u, σ are in GW, and u correspond to the 95% quantile in all cases.

5.2. Comparison to hindcast net demand models

In this subsection, we compare LOLE estimates from the logistic model fitted in previous subsections to those from the hindcast model (4).

To make this comparison more relevant in the context of growing renewables penetrations and market integration, we perform numerical experiments using an interconnection capacity of 2 GW and a wind generation capacity of 2.5 times that installed in 2014 (i.e. installed wind capacities of 7.5 GW for IRL and 38 GW for GB). The distribution of surplus in each system is then shifted to keep the pre-interconnector value of LOLE averaged across all seven years at the initial level of 3 h/y.

5.2.1. Comparison of LOLE estimates

Fig. 5 shows post-interconnection LOLE estimates for both areas; because the integral in (7) cannot be calculated exactly for the fitted tail models, Monte Carlo estimates are used instead, and corresponding confidence

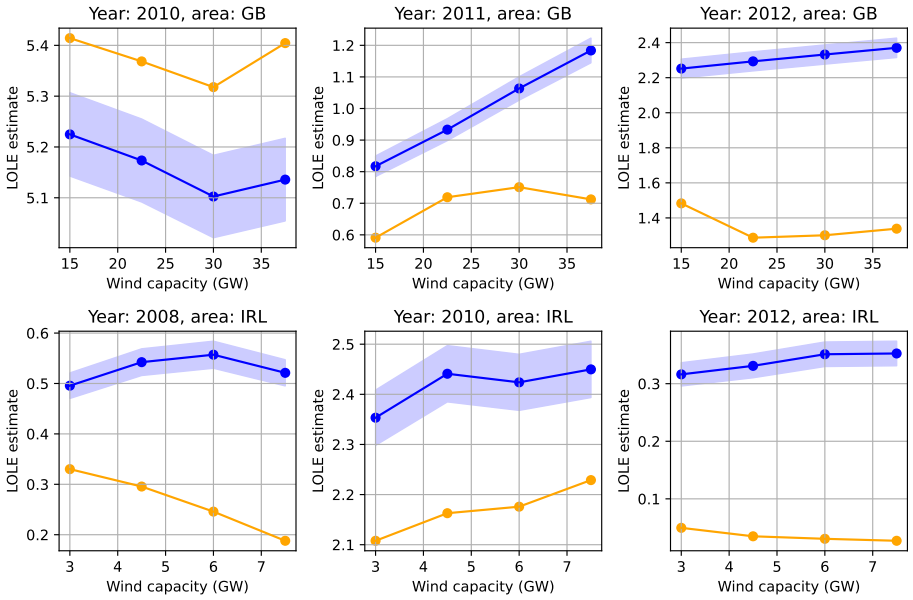


Figure 5: Comparison of LOLE estimates from hindcast (orange) and logisitc (blue) models. Because estimation from a logistic model requires Monte Carlo simulation, 95% confidence bands for its central estimate are shown.

bands for the central estimate are shown. As these results consider the contribution of interconnection, LOLE values are generally low, particularly for Ireland where the system is smaller relative to the capacity of interconnection. However we can observe large relative differences between model estimates in both areas, e.g. years 2011, 2012 for GB and 2008, 2012 for IRL; furthermore, for 2011 in GB and 2008 in IRL, both models suggest diverging risk estimates as more wind is installed. While the difference between the hindcast and EVT results are not very great for the highest risk year (2010), there are substantial differences in some of the higher risk years (2012 in GB, 2008 in Ireland), demonstrating how the effect of smoothing the tail can be material in practical risk calculations.

There are two main theoretical reasons to prefer EVT estimates over hindcast estimates. Firstly, EVT smoothes the tail of the distribution, reducing the influence of outliers and small variations in the tail data. This is particularly important because we know that much of the risk in hindcast estimates is concentrated in a small number of extreme observations. Secondly, EVT allows us to extrapolate beyond the most extreme observation seen in the

dataset – one side effect of this is that EVT can be used to calculate probabilities for events that have not previously been observed. If we were able to simulate multiple realisations of a given year and compute both EVT and hindcast estimates of the LOLE for each realisation, we would expect the variance of the EVT estimates to be smaller than the variance of the hindcast estimates (because of the smoothing effect described above). However, producing such a set of realisations to test this is challenging because net demand is a time series with serial correlation as well as complex seasonal trends that would need to be replicated in any simulation.

6. Statistical dependence in capacity deficits

Having developed a model for extreme net demand events in the two-area system in the previous section, we now turn our attention to statistical dependence in capacity deficits, i.e. between negative values of (available supply minus demand) in the two areas. The analysis in the previous section demonstrated a strong association in extreme *net demand* co-occurrences, however each of these net demands is convolved with the available conventional capacity in the relevant area, and the available conventional capacities are assumed independent. It is thus natural to think that in a relevant sense the dependence may be weaker for *capacity deficits* than for *net demand* – this section will investigate this hypothesis, again by assessing the strength of tail dependence and using this to fit an appropriate statistical model.

We work with the right tails of (demand minus available supply), thus treating capacity shortfalls as maxima instead of minima to make results from EVT immediately applicable. To avoid any confusion we call this the *shortfall distribution*, and negative values simply indicate the non-occurrence of a capacity shortfall.

6.1. Charactersing statistical dependence

We proceed in a manner similar to the analysis of net demand in Section 5, performing the same Bayesian ratio test as in Section 5.1.2 to determine whether asymptotic dependence is present, resulting in a Savage-Dickey ratio of $r < 10^{-3}$ for all years and a wide range of quantile thresholds between 80% and 99.99%. This provides strong direct evidence of capacity shortfalls being asymptotically independent.

The evidence for asymptotic dependence between the two net demands was not as definitive as this evidence against for the shortfalls, but the differ-

ence in r for the two cases clearly demonstrates stronger tail dependence for the net demands as compared to the capacity shortfalls. This would appear to confirm the intuition that the convolution with the (independent) conventional plant distributions should weaken the dependence between the systems – and provides some further justification for the use of the logistic model for the earlier cases even where the evidence for asymptotic dependence was relatively weak.

Extreme value copulas like the ones used on net demand data are not appropriate to model asymptotically independent data. However, as stated in Section 3.2, Gaussian copulas provide a simple parametric example of an asymptotically independent copula, and we test this as a dependence model for joint exceedances of the shortfall distribution, finding that it is a good fit for all peak seasons and tested thresholds between 80% and 99.99%. Fig. 6 shows the comparison between contour lines for each decile of the empirical and fitted Gaussian copula for a threshold of 80%.

We note that Gaussian copulas are not typically used to model extremal dependence in the EVT literature because they are not an extreme value copula in the sense of Equation (13). Our choice here is instead based on practicality, as it provides a simple parametric model that accurately describes tail dependence in the capacity shortfall distribution.

Based on this evidence, we proceed as with the net demand model and use a Gaussian copula to describe dependence between exceedances at both components, defining an exceedance as a shortfall in the corresponding component; the modelling region is illustrated in Fig. 7b. The quantile threshold in this case was much higher than for net demand, as shortfalls occur with a probability of approximately 0.001% in each area, due to the LOLE normalisation to three hours per year over the whole seven-year period.

The difference between the regimes of tail dependence in net demand and the shortfall distributions can be more clearly seen in Fig. 7: the net demand distribution's tails are much more concentrated around the diagonal of the graph than those of capacity shortfalls. A consequence of this, is that simultaneous shortfalls occur rarely relative to shortfalls in just one region, whereas there is a much stronger tendency for extremes of net demand to occur in both areas at the the same time. Fitted dependence parameters for this model are given in Table 2.

We note that these observations about dependence between deficits in the two areas follows in substantial part from the assumption of independence between available conventional capacities in the two areas. This assumption

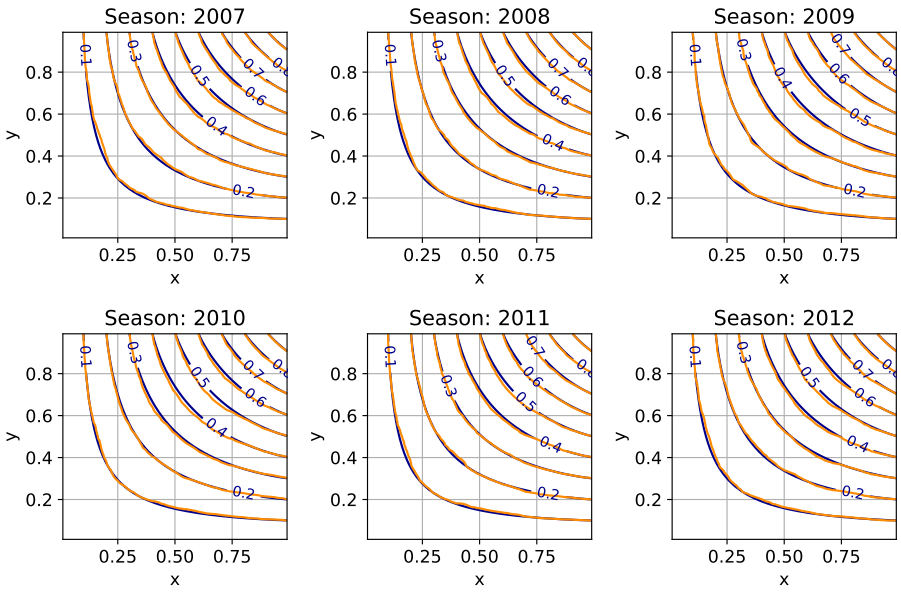
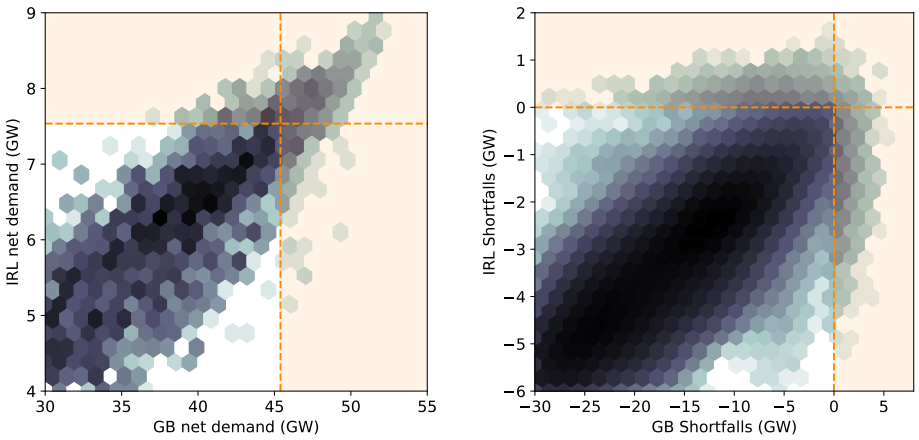


Figure 6: Contour lines for cumulative probability functions of empirical (orange) and fitted Gaussian copulas (blue); the first nine deciles are shown. The data consisted of 2500 simulated joint shortfalls.

season	ρ
2007	0.68
2008	0.7
2009	0.69
2010	0.76
2011	0.7
2012	0.66

Table 2: Fitted values for Gaussian exceedance dependence model.

is usually made in practical calculations, though it must be caveated if there is a possibility of a common cause event affecting units in both systems – for instance restriction on primary fuel supply, or elevated failure rates in certain weather conditions as in [33, 34]. Nevertheless, these results around asymptotic dependence are helpful in understanding outputs of present risk calculations, and it seems reasonable to think that even where common cause events are relevant the influence of conventional plant will weaken the tail



(a) Net demand data.

(b) Simulated shortfalls (negative surpluses).

Figure 7: Hex-binned net demand and shortfall distribution scatterplots for 2010. The modelling thresholds, which delimit the modelling region (shaded orange) are shown as dashed lines. Points in the modelling regions were simulated from the fitted models.

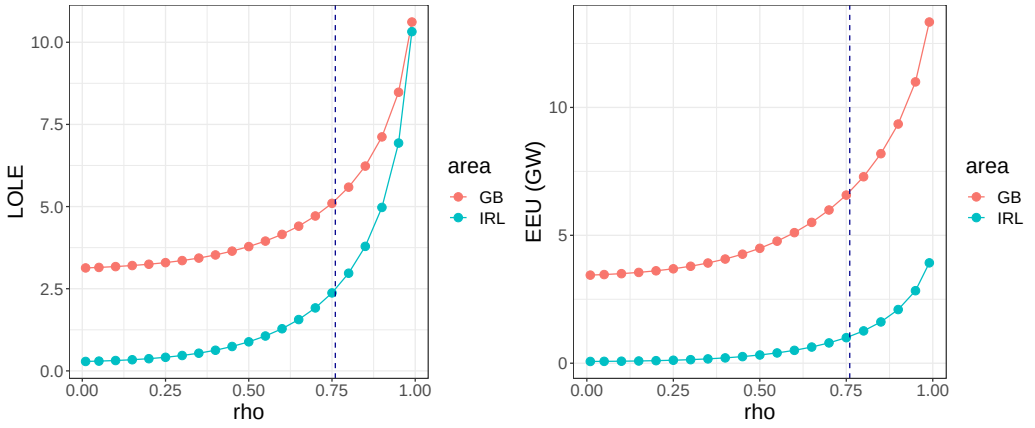
dependence between the distributions of deficit in the two areas as compared to the dependence between net demands.

6.2. Sensitivity of risk metrics to dependence strength

Having a single-parameter dependence model for the shortfall region in the shortfall distributions makes it straightforward to perform a sensitivity analysis of LOLE and EEU to statistical dependence strength between the capacity shortfalls at both areas. Fig. 8 shows the results of the sensitivity analysis using data from 2010, with a similar pattern seen using data from other years. The dependence of the LOLE level is quite weak at low values of ρ, with the estimated risk level only increasing significantly above the 'independence' limiting case forρ above about 0.5.

7. Conclusion

This paper has presented approaches to statistical modelling of net demands and of shortfalls in two area power systems. The general approach is to assess whether the quantities of interest are asymptotically dependent, and then to choose an appropriate bivariate model based on this assessment.



(a) LOLE sensitivity to dependence strength

(b) EEU sensitivity to dependence strength

Figure 8: Sensitivity of each metric (LOLE and EEU) to dependence strength, in both areas and for both policies; the dotted blue line represents the estimated dependence strength in the GB-IRL system.

The methods have been developed for two area systems, but can be extended to multi-area systems.

In the case of net demands, the tests are consistent with asymptotic dependence, though this conclusion is not very definitive. Other diagnostics do also suggest that an extreme value copula with asymptotic dependence provides a good fit to the empirical dependence structure. The resulting smoothing of the sparse extreme region of the empirical distribution of net demand can change risk calculation results substantially, with a possible explanation being that the calculation results involving the empirical distribution are driven by a very small number of historic records.

We also demonstrate how the dependence structure of the shortfalls in the two areas may be modelled directly. In this case there is very strong evidence against asymptotic dependence, and we find that a Gaussian copula describes the behaviour appropriately in this case. This contrast in the assessment of asymptotic dependence for the two cases confirms the intuition that dependence between shortfalls in the two systems might be weakened by convolution with the distributions of available conventional capacity (which are assumed independent between the systems). This also provides a means of performing sensitivity analysis of the risk model outputs to the strength of dependence between the shortfalls, by varying the correlation coefficient of the Gaussian copula.

References

[1] Winter Outlook 2019/20, Tech. rep., National Grid, available at https://www.nationalgrideso.com/document/154166/download (2019).

[2] Final Report of the Sector Inquiry on Capacity Mechanisms, Tech. rep., European Comission, available at https://ec.europa.eu/energy/sites/ener/files/documents/com2016752.en_.pdf (2016).

[3] V. M. Cook, M. J. Steinberg, C. D. Galloway, A. J. Wood, Determination or Reserve Requirements or Two Interconnected Systems, IEEE Transactions on Power Apparatus and Systems 82 (65) (1963) 18–33.

[4] C. Pang, A. Wood, Multi-area generation system reliability calculations, IEEE Transactions on Power Apparatus and Systems 94 (2) (1975) 508–517.

[5] C. Singh, P. Jirutitijaroen, J. Mitra, Electric Power Grid Reliability Evaluation: Models and Methods, Wiley, 2018.

[6] National Grid, Security of Supply - International Review of Standards and Implementation, available at http://site.ieee.org/pes-rrpasc/files/2019/04/National-Grid-Security-of-Supply-International-Review-Final-IEEE-v2.pdf (2017).

[7] A. L. Wilson, S. Zachary, Using extreme value theory for the estimation of risk metrics for capacity adequacy assessment, arXiv preprint arXiv:1907.13050 (2019).

[8] H. E. Thornton, A. A. Scaife, B. J. Hoskins, D. J. Brayshaw, The Relationship Between Wind Power, Electricity Demand and Winter Weather Patterns in Great Britain, Environmental Research Letters 12 (6) (2017) 064017.

[9] S. Sheehy, G. Edwards, C. J. Dent, B. Kazemtabrizi, M. Troffaes, S. Tindemans, Impact of High Wind Penetration on Variability of Unserved Energy in Power System Adequacy, in: International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2016.

[10] R. Billinton, R. N. Allan, Reliability Evaluation of Power Systems, Plenum Press, 1996.

[11] N. Sanchez, C. Dent, A. Wilson, Quantifying The Reliability Contribution of Interconnectors in the Britain - Ireland Power System Using a Hindcast Approach, in: International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2020.

[12] G. Stephen, S. H. Tindemans, J. Fazio, C. J. Dent, A. Figueroa Acevedo, B. Bagen, A. Crawford, A. Klaube, D. Logan, D. Burke, Clarifying the Interpretation and Use of the LOLE Resource Adequacy Metric, in: International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2022.

[13] C. Dent, S. Zachary, Estimation of Joint Distribution of Demand and Available Renewables for Generation Adequacy Assessment, Preprint available at http://arxiv. org/abs/1412.1786 (2014).

[14] A. Keane, M. Milligan, C. J. Dent, B. Hasche, C. D'Annunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Soder, M. O'Malley, Capacity Value of Wind Power, IEEE Transactions on Power Systems 26 (2) (2011) 564–572.

[15] S. H. Tindemans, M. Woolf, G. Strbac, Capacity Value of Interconnection Between Two Systems, in: IEEE Power and Energy Society General Meeting (PES GM), 2019.

[16] J. Beirlant, Y. Goegebeur, J. Segers, Statistics of Extremes, John Wiley and Sons, 2006.

[17] S. Coles, An Introduction to Statistical Modeling of Extreme Values, Springer London, 2013.

[18] J. E. Heffernan, J. A. Tawn, A conditional approach for multivariate extreme values (with discussion), Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66 (3) (2004) 497–546.

[19] A. W. Ledford, J. A. Tawn, Modelling Dependence Within Joint Tail Regions, Journal of the Royal Statistical Society. Series B (Methodological) 59 (2) (1997) 475–499.

[20] J. Heffernan, A Directory of Coefficients of Tail Dependence, Extremes 3 (2000) 279–290.

[21] M. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publications del'Institut de Statistique de l'Université de Paris (8) (1959) 229–231.

[22] P. Jaworski, F. Durante, W. Hardle, T. Rychlik, Copula Theory and Its Applications, Berlin: Springer, 2010.

[23] I. Staffell, S. Pfenninger, Using bias-corrected reanalysis to simulate current and future wind power output, Energy 114 (2016) 1224–1239.

[24] I. Staffell, S. Pfenninger, The increasing impact of weather on electricity supply and demand, Energy 145 (2018) 65 – 78.

[25] National Grid, Modelling Methods, Tech. rep., available at http://fes.nationalgrid.com/media/1246/modelling-methods-v10.pdf (2017).

[26] Baringa Ltd, Baringa SEM PLEXOS Forecast Model for 2016-17, Tech. rep., available at https://www.semcommittee.com/news-centre/baringa-sem-plexos-forecast-model-2016-17 (2016).

[27] N. Sanchez, `riskmodels`: a library for univariate and bivariate extreme value analysis (and applications to energy procurement), https://pypi.org/project/riskmodels/2.2.0/.

[28] M.-H. Chen, Bayesian Computation: From Posterior Densities to Bayes Factors, Marginal Likelihoods, and Posterior Model Probabilities, in: D. Dey, C. R. Rao (Eds.), Bayesian Thinking, Modeling and Computation, Elsevier Science, 2005, Ch. 15, pp. 449–450.

[29] E.-J. Wagenmakers, T. Lodewyckx, H. Kuriyal, R. Grasman, Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method, Cognitive Psychology 60 (3) (2010) 158–189.

[30] J. M. Dickey, B. P. Lientz, The Weighted Likelihood Ratio, Sharp Hypotheses about Chances, the Order of a Markov Chain, The Annals of Mathematical Statistics 41 (1) (1970) 214 – 226.

[31] M. Castellanos, S. Cabras, A default Bayesian procedure for the generalized Pareto distribution, Journal of Statistical Planning and Inference 2 (02 2007).

[32] P. Hall, N. Tajvidi, Distribution and dependence-function estimation for bivariate extreme-value distributions, Bernoulli 6 (5) (2000) 835 – 844.

[33] S. Murphy, J. Apt, J. Moura, F. Sowell, Resource adequacy risks to the bulk power system in North America, Applied Energy 212 (2018) 1360–1376.

[34] S. Murphy, L. Lavin, J. Apt, Resource adequacy implications of temperature-dependent electric generator availability, Applied Energy 262 (2020) 114424.