

# *Comparing generator unavailability models with empirical distributions from Open Energy datasets*

Conference or Workshop Item

Accepted Version

Deakin, M., Greenwood, D., Brayshaw, D. J. ORCID: <https://orcid.org/0000-0002-3927-4362> and Bloomfield, H. ORCID: <https://orcid.org/0000-0002-5616-1503> (2022) Comparing generator unavailability models with empirical distributions from Open Energy datasets. In: 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 12-15 June 2022, Online. doi: 10.1109/PMAPS53380.2022.9810629 Available at <https://centaur.reading.ac.uk/105098/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1109/PMAPS53380.2022.9810629>

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Comparing Generator Unavailability Models with Empirical Distributions from Open Energy Datasets

Matthew Deakin,  
David Greenwood  
*School of Engineering*  
*Newcastle University*  
Newcastle-upon-Tyne, UK  
{matthew.deakin, david.greenwood}@ncl.ac.uk

David J. Brayshaw  
*Meteorology Department*  
*University of Reading*  
Reading, UK  
d.j.brayshaw@reading.ac.uk

Hannah Bloomfield  
*School of Geographical Sciences*  
*University of Bristol*  
Bristol, UK  
hannah.bloomfield@bristol.ac.uk

**Abstract**—The modelling of power station outages is an integral part of power system planning. In this work, models of the unavailability of the fleets of eight countries in Northwest Europe are constructed and subsequently compared against empirical distributions derived using data from the open-access ENTSO-e Transparency Platform. Summary statistics of non-sequential models highlight limitations with the empirical modelling, with very variable results across countries. Additionally, analysis of time sequential models suggests a clear need for fleet-specific analytic model parameters. Despite a number of challenges and ambiguities associated with the empirical distributions, it is suggested that a range of valuable qualitative and quantitative insights can be gained by comparing these two complementary approaches for modelling and understanding generator unavailabilities.

**Index Terms**—Generator unavailability, outage modelling, resource adequacy, capacity planning, power system planning.

## I. INTRODUCTION

Power station unavailability modelling is an essential part of power system planning and operations. It is necessary for reserve scheduling to ensure second-by-second stability, hourly unit commitment, or resource adequacy planning several years into the future. Despite this reputation as a textbook problem, insufficient mitigation against supply-side vulnerabilities was found to have been a key contributing factor in the catastrophic Texas blackouts of February 2021 [1]. Furthermore, nascent Digital Twins of energy systems, incorporating highly granular system-specific data, will need to provide a virtual replica of all parts of the energy system so that they can evaluate system security as a part of their core functionality [2].

It is well-known that different generator types (e.g., nuclear, gas) have different availabilities [3]. However, for accurate risk assessment, the overall unavailability of generator fleets also depends on characteristics of the fleet itself, such as the age and condition of units, as well as exogenous factors such as meteorological conditions [4]. Databases containing a large number of generator outage reports are necessary for considering these issues, as there are typically hundreds of large generators in countries such as the UK or Germany.

This work was funded by the Supergen Energy Network (EP/S00078X/2) through the Climate-Energy Modelling for Assessing Resilience: Heat Decarbonisation and Northwest European Supergrid (CLEARHEADS) project. 978-1-6654-1211-7/22/\$31.00 ©2022 IEEE

One such database is provided by the European Network of Transmission System Operators for Electricity (ENTSO-e), who have been providing the Transparency Platform continuously since 2015 [5]. Compared to other data that are presented on the platform, generator unavailability data is more complex [6], from the point of view of the volume of the data, but also its interpretability (e.g., the difference between the self-reported ‘forced’ and ‘planned’ outages). Indeed, previous works that make extensive use of data from the platform stop short of working with the outage data stream directly [7].

Given the challenges with working with this type of data, unavailability modelling and analysis using empirical data directly is not common. In [4], the authors consider the implications of temperature dependence on the PJM system in the USA, finding that temperature dependence of generators significantly increases the capacity that must be procured to achieve a given risk standard. A similar conclusion is found in [8], again focusing on high temperatures and drought in the German system. Further use-cases for consider generator models beyond system adequacy are considered, for example, in [9], where the authors use outage data from the ENTSO-e Transparency Platform for improving forecasting. In [10], the author explores the challenges and possible solutions to mixed data quality in the ENTSO-e unavailability data, considering the Dutch and German systems. The authors of [11] consider Bayesian priors as an alternative and more accurate method of power system outages by modelling plants with missing or incomplete unavailability data with plants that are similar.

To our knowledge, there are no papers which compare analytic fleet unavailability models with empirical data across a wide range of countries. This is an important and timely gap—such unavailability modelling across countries is central to the administration of contemporary capacity markets (such as the ca. £700m GB capacity market [12]). Furthermore, as databases such as the Transparency Platform become increasingly accessible, it becomes even more crucial for the limitations of given datasets to be explored in detail.

The contribution of this work is to address this gap by comparing analytic, aggregate unavailability models for a fleet of generators against equivalent empirical models derived from an open energy database (the ENTSO-e Transparency

Platform). Both time collapsed (non-sequential) and time sequential models are considered for a range of countries, and limitations of the empirical data are described.

This paper outlines analytic unavailability models (Section II); the approach used to download the empirical unavailability data (Section III) and then compares these two approaches (Section IV) before offering conclusions (Section V), to highlight similarities and differences between these complementary approaches of unavailability modelling.

## II. ANALYTIC GENERATOR UNAVAILABILITY MODELS

This work considers two main classes of model: time collapsed (or ‘non-sequential’) models, which study the distribution of plant outages in the long run, and time sequential models, which consider the hour-by-hour time series of generator unavailabilities. In this section, we first highlight models that are most commonly used to model unavailability of an individual generator, before describing how these can then be combined to create the system-wide unavailability.

### A. Time collapsed outage modelling

Time collapsed models are based on the approximation that a unit’s availability is based a model which can be determined by an availability parameter,  $A$ , which determines the likelihood of a generator being available at a given time instant. An individual production unit  $U$  with size  $U_0$  (in MW) and availability  $A$  can then modelled as a Bernoulli random variable,

$$\frac{U}{U_0} \sim \text{Bernoulli}(A). \quad (1)$$

It is well-known that the probability density function of the sum of independent random variables can be determined by the convolution of each of the PDFs of each random variable. Therefore, under the assumption of independent outages, the PDF of total system outages can be determined by calculating the convolution of the PDFs of each individual production unit [13, Ch. 3].

### B. Time Sequential outage modelling

Time sequential outage modelling explicitly considers how a unit’s availability changes in time. The most parsimonious model of the outages of a given generator is via a two-state Markov model: a production unit is modelled as being either in or out of service. In each of these states, there is a transition probability that describes the likelihood of moving to the other state.

These transition probability can be specified directly, or they can be determined by other parameters. For example, it is common to describe the outage states of generator models via an availability parameter  $A$  and Mean Time To Repair, MTTR (in hours). The repair rate  $\mu$  and failure rate  $\lambda$  can be determined from these parameters [14, Ch. 9] as

$$\mu = \frac{1}{\text{MTTR}}, \quad \lambda = \mu \left( \frac{1}{A} - 1 \right). \quad (2)$$

TABLE I  
AVAILABILITY  $A$  AND MEAN TIME TO REPAIR (MTTR) BY FUEL TYPE.

Fuel	Avlby., $A$	MTTR (hrs)	Fuel	Avlby., $A$	MTTR (hrs)
Biomass	0.86	40	Coal	0.86	40
CCGT	0.90	50	Oil	0.91	50
Hydro	0.90	20	Nuclear	0.81	150
CHP	0.90	50	Waste	0.86	40

### C. Generator Availability Models for Northwest Europe

To build total outage rates for the generator fleet of individual balancing zones across Northwest Europe, the total installed capacity for each fuel type is first found from the ‘Installed Production Data’ from ENTSO-e. Individual Production and Generation units sizes for all countries are pooled to create an empirical distribution of reasonable generator sizes for each fuel type. From this, a representative fleet of generators is then created by allocating these generators to each country and fuel type. Generators of a given fuel type are assumed to all have the same availability and MTTR, with the values used summarised in in Table I (MTTRs are from [3] and availabilities are from [15]).

For example: if the UK in 2020 had 10 GW of nuclear power, and across Europe nuclear production units are only 1 GW or 2 GW units with equal probability, then one possible nuclear generation fleet for the UK could be three 2 GW units and four 1 GW units. Each of these would have the same availability and mean time to repair (Table I).

## III. DATA PROCESSING FOR EMPIRICAL UNAVAILABILITY DISTRIBUTIONS

In this section, we describe how the outage data for individual units are accessed from the ENTSO-e platform and subsequently processed to estimate the aggregate unavailability of the whole fleet.

The Transparency Platform reports Planned and Forced outages, thereby covering both ‘planned’ and ‘actual’ unavailability as required by the relevant regulations [16]. The regulation specifically states that plant of capacity greater than 200 MW must report changes of more than 100 MW in availability, both for aggregate generation units or individual production units.

In practise, changes in availability much smaller than 100 MW are reported extensively by many units (see Section IV-A). Users of this data should therefore be cognizant that there could be small outages which are not reported.

### A. Data Access Approach

The data are downloaded using the Transparency Platform API, allowing for automated download, with the data downloaded as XML files for individual days. This means that some reports are duplicated (particularly where outages last many months), but it means that the data can be inspected manually if required. Our focus is on dispatchable generators, and so unavailability of renewable technologies (solar, onshore

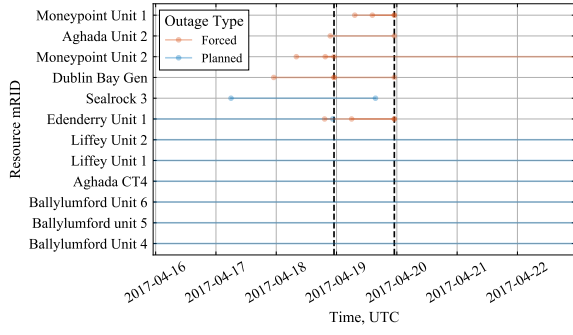


Fig. 1. The outage reports for the IE zone for 19/4/17. The solid dots denote the start and end point of a single outage report for a production unit.

wind, offshore wind, and run-of-river hydro) is neglected. Furthermore, reports with a ‘withdrawn’ flag are also ignored.

A representation of the reports downloaded for one day of the IE zone is shown in Fig. 1. It can be observed that, for some generators, there are multiple overlapping reports for the given day, with some generators having both forced and planned outages. It can also be seen that some of the generators have long-term outages, with their planned outage extending beyond the week window shown here.

1) *Defining Hourly Outages:* Outage reports are given at a temporal resolution of one minute, and so over the course of an hour, the unavailability may change multiple times. For this work, the ‘outage’ for a given hour is the mean outage over the given time period—i.e., if the ‘instantaneous’ (minute-by-minute) outage is denoted by  $o$ , the hourly outage  $O(\tau)$  considered over some time period  $\tau$ ,  $\tau + \delta_\tau$  is

$$O(\tau) = \frac{1}{\delta_\tau} \int_{\tau}^{\tau+\delta_\tau} o(t) dt. \quad (3)$$

For example, if a production unit’s availability is reduced by 50 MW for 12 minutes and 200 MW for 48 minutes, then the corresponding hour’s outage would be calculated to be 170 MW.

In some cases, a given unit will have multiple outage reports which are contradictory at a particular time. In these cases, the conflicting outage reports need to be reconciled. The approach in this work is to combine these multiple conflicting reports  $o_1, \dots, o_n$  and then calculate the minimum, maximum, and mean outage that could occur given all of those reports, as

$$O_{\min}(\tau) = \frac{1}{\delta_\tau} \int_{\tau}^{\tau+\delta_\tau} \min_i \{o_i(t)\} dt, \quad (4)$$

$$O_{\max}(\tau) = \frac{1}{\delta_\tau} \int_{\tau}^{\tau+\delta_\tau} \max_i \{o_i(t)\} dt, \quad (5)$$

$$O(\tau) = \frac{O_{\min}(\tau) + O_{\max}(\tau)}{2}, \quad (6)$$

where the min and max operators calculate the maximum and minimum value over all of the outage reports at each incremental time period.

A further common inconsistency that is found is that, in some cases, the reported reduction in capacity is greater than

the total size of the generator itself (e.g., a reduction in capacity of 1500 MW may be reported for a generator whose size is only 750 MW). For the purposes of this work, outage reports for which the reported outage is 33% greater than the production unit size are ignored (outages slightly above the size of the plant are permitted, as it allows for small inaccuracies in the reporting of unit maximum ratings which may exist).

Forced and planned outages are recorded separately, as well as the total outage level. The latter is determined by calculating the outages irrespective of the type of outage flag. Note that the total outages are always less than or equal to the sum of the forced and planned outages, with conflicting forced and planned outages reconciled to calculate the total outage through (6) (i.e., neglecting their state as forced or planned).

### B. Evaluating Data Quality

Clearly, as highlighted in the previous section, there are a range of reasons as to why the unavailabilities of the aggregated fleet of individual countries may be inaccurate. In the first instance, only shortfalls of more than 100 MW are mandatory to report (although, in many cases shortfalls are reported at a much higher resolution than this). There is also potential for human error in the reports that are submitted to the Transparency Platform, and mothballed or closed plant are at times listed as being on an extended ‘outage’. Furthermore, there may be commercial reasons for generators to attempt to obfuscate the true outage state of their plant (e.g., to give traders within the same organisation as much of a competitive advantage as possible).

One of the uncertainties which can be considered explicitly is the effect that the reconciliation of conflicting reports (6) could have. In particular, we calculate the relative mean absolute reconciliation error  $\epsilon$  over the full time period, i.e.,

$$\epsilon = \frac{\|O(\tau) - O_{\min}(\tau)\|_1}{\|O(\tau)\|_1}. \quad (7)$$

This relative reconciliation error gives a measure as to the amount that the outages could be incorrect given these conflicting outage reports. In other words, by comparing the reconciliation error  $\epsilon$  with the mean and interquartile range (IQR) of the total outage distributions, we can confirm if the reconciliation is likely to have made a significant impact on the final distribution.

## IV. RESULTS

In this section, we look to consider the differences between the empirical and modelled system outages. In total, data was downloaded for five winters and four summers for ten regions. The number of reports per day varies between countries—the smaller Irish system regularly has fewer than ten reports per day, whilst the German and French systems regularly have more than two hundred.

In this work, we focus on modelling outages during the winter period, as this is the season which has peak demands in Northwestern Europe. For the purposes of this, we consider

the winter season to consist of twenty weeks following the first Sunday of November, excluding two low-demand weeks around Christmas [17].

In this section, we first present the unavailability profiles of a number of generators to demonstrate clearly some of the challenges of working with the outage data directly and to consider how this empirical data (Section III) differs qualitatively from the presented analytic models (Section II). The total fleet unavailabilities are then considered for eight Northwest European countries, and the properties of these empirical models compared against both time sequential and time collapsed models. The seasonality of the empirical data is then explored, followed by a short discussion to consider implications of the results for future unavailability modelling.

The aggregated unavailability for each country (and code used to derive these unavailabilities) is available from [18].

#### A. Example Data

First, we consider the unavailability reports for a number of individual periods. Fig. 2 plots the planned, forced and total outages for three production units.

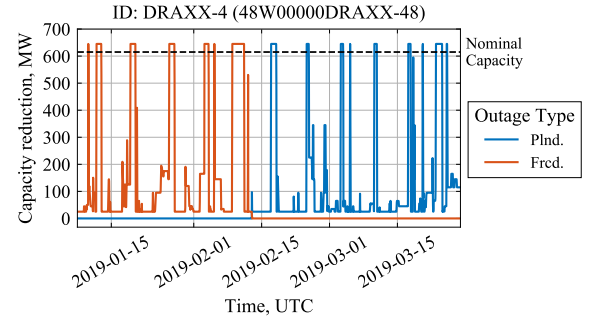
Fig. 2a highlights how the distinction between forced (Frcd.) and planned (Plnd.) outages is not clear-cut. For example, it can be seen that, around mid February, the forced outages from earlier in the winter are instead classified as planned outages. By eye, there appears to be a clear pattern in the outages, and so the reason for the change in classification is not clear. Note also how small capacity reductions (less than 100 MW) are reported regularly.

In Fig. 2b, it can be seen why the Total (Ttl.) fleet outages are not necessarily a linear sum of Forced and Planned outages (as mentioned in Section III-A). In this case, as the nominal capacity of the unit is 400 MW, on the day that there is 400 MW of reduction in both planned and forced outages, the approach outlined in Section III-A returns a total outage of only 400 MW (rather than 800 MW). In other words, it is assumed that these are for the same incident.

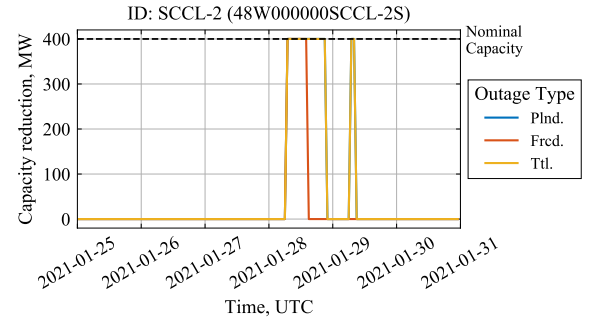
Finally, Fig. 2c shows a more complex example: for this plant, there is an active report that states that unit has a planned outage from 1st October 2016 to 30th September 2017, with capacity of the unit reduced by 485 MW. However, during this time period, there are further active outage reports which clearly contradict this first report (as shown in the figure), and tending to lead the mean outage to drop when these outages are not at the nominal capacity of the unit (due to the approach mentioned in Section III-A). It is beyond the scope of this work to attempt to comprehensively identify and correct apparent errors such as these, although more systematic ways of considering these types of errors could improve accuracy further.

#### B. Time Collapsed Model Comparison

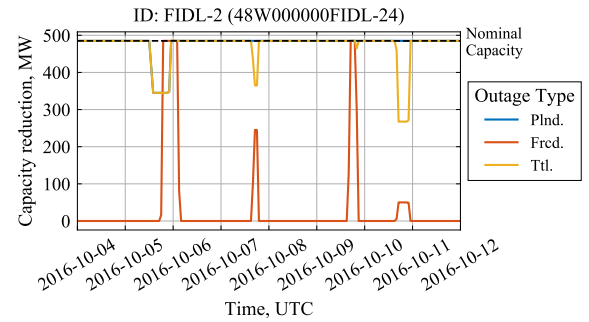
The time collapsed model of total outages for the fleet can be described by the PDF of the total generator unavailabilities. In Fig. 3, histograms of the total and forced outage rates in for five winter seasons are compared against the analytic model



(a) Example outage reports, Jan-Mar 2019.



(b) Example outage reports, Jan 2021.



(c) Example outage reports, Oct. 2016.

Fig. 2. Generator unavailability for three units in the GB system.

based on (1). It can be observed that in this case, the spread of the modelled outages falls roughly between the forced and total outages. At first glance, this appears to show that the analytic model and empirical model show some links, even if there are clear discrepancies.

However, once a full range of countries are considered, it is quite clear that in many cases the empirical data are far from being a good representation of the analytic model—the mean and IQR of the analytic model and empirical data are given in Table II. For example, countries such as Belgium and Norway show means and IQRs that are many times smaller than the modelled outage rates. This suggests a clear discrepancy between the empirical and analytic models; further investigations are clearly required to provide an accurate picture of generator unavailability in those systems. Note that although the reconciliation error (7) is non-trivial, it clearly is relatively small as compared to the bulk errors observed and

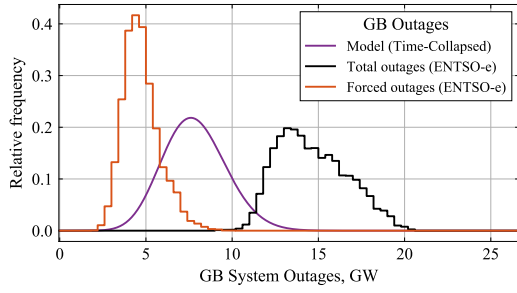


Fig. 3. Histograms of the Total and Forced outages for the GB system from the ENTSO-e platform for five winters (16/17 to 20/21), compared against the PDF of the analytical Time-Collapsed model of the same system.

TABLE II  
MEAN AND INTERQUARTILE RANGE (IQR) OF FORCED (FRCD.), TOTAL (TOT.), AND MODELLED (MLD.) TOTAL SYSTEM OUTAGES FOR NINE COUNTRIES, ALONGSIDE THE MAXIMUM RECONCILIATION ERROR (7).

Ctry. Code	Mean, GW			IQR, GW			Max Recon. Error $\epsilon$ , %
	Frkd.	Ttl.	Mdl.	Frkd.	Ttl.	Mdl.	
GB	4.7	14.7	7.8	1.3	3.0	2.5	0.7%
BE	0.1	0.2	2.2	0.2	0.4	1.4	2.4%
DE	2.8	11.4	14.2	1.5	3.5	3.3	1.6%
DK	0.5	0.8	1.1	0.5	0.5	0.8	3.0%
ES	5.5	6.3	8.1	3.7	3.6	2.2	0.1%
FR	2.7	16.6	15.9	2.3	7.8	4.7	0.1%
IE	0.1	0.6	1.0	0.3	0.5	0.8	1.6%
NL	0.7	3.5	2.7	1.1	1.4	1.3	0.1%
NO	0.2	0.7	3.0	0.3	0.6	1.0	0.0%

so does not explain those discrepancies.

In any case, it is interesting to note that there are marked seasonal characteristics in the outage rates. For the GB system the Forced outages are, to a large extent, constant, whilst planned outages fluctuate on an annual basis, with a clear minimum during the winter months (Fig. 4).

Note that, whilst some seasonality appears to be universal, the extent and period of this seasonality varies considerably. The normalised mean weekly plant Total outages for GB, France and Spain are plotted in Fig. 5 alongside the normalised seasonal demands. Fig. 5a shows how the GB system has less seasonality than either the Spanish or French systems, and

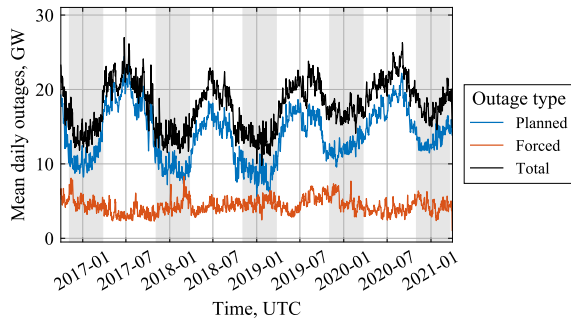
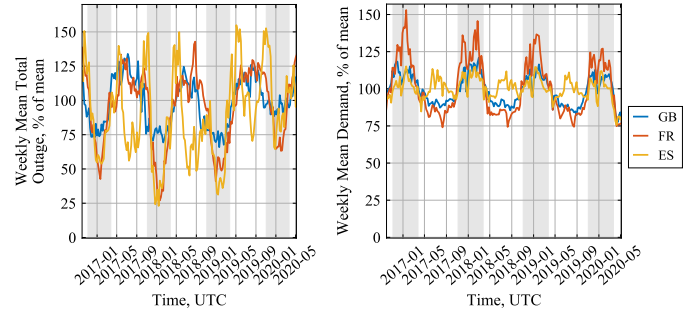


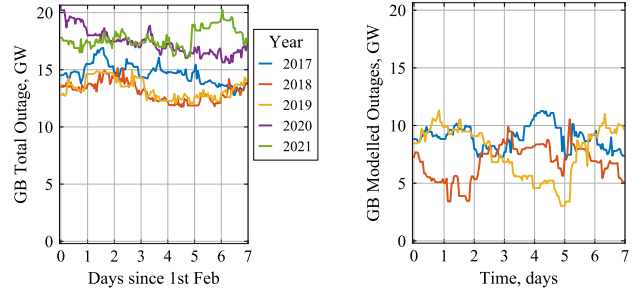
Fig. 4. Forced outages in the GB system are much less seasonal than Planned outages, and make up less than half of all outages in the GB system.



(a) Mean weekly outages

(b) Mean weekly demand

Fig. 5. The seasonality of reported outages varies from system to system, with a clear link between total outages (a) and the demand (b).



(a) Empirical data

(b) Model outputs

Fig. 6. The time series of total outages (Fig. 6a) and modelled outages (Fig. 6b) show that, despite the spread of the model outputs being smaller than that of the Total Outages (Table II), the hour-by-hour changes in model outputs are much greater than the changes from the empirical model.

that the Spanish system has a biannual cycle. This biannual cycle could be explained as being due to generators ensuring maintenance does not occur during demand peaks, of which there are two in the year in Spain, as compared to France or GB which only peak in the winter (Fig. 5b).

### C. Time Sequential Modelling

Time sequential models are much more complex than non-sequential time collapsed models: a complete characterization of the analytic and empirical models is beyond the scope of this work. Here, we compare the data for GB visually, and calculate the signal autocorrelation across a range of countries.

Firstly, the Total outages for the days following Feb. 1st from 2017 to 2021 are compared against three draws of the two-state time sequential model (Section II-B) in Fig. 6. Clearly, the mean of the analytic model is lower than empirical data; this is not unexpected given Fig. 3. However, what is also clear is that, despite the spread of the Total outages being greater than that of the model (as determined using the IQR in Table II), the hour-by-hour changes in the analytic model outputs appear to be greater than those of the reported Total outages from the GB system. Future work could systematically study these changes where they are critical (e.g., for storage scheduling during periods with tight margin).

Finally, in Table III, the autocorrelation is tabulated for a variety of time lags for five countries for which the time-



TABLE III

COMPARING THE MEAN OF THE AUTOCORRELATION OF TOTAL OUTAGES IN WESTERN EUROPEAN COUNTRIES FOR FIVE WINTERS (16/17 - 20/21) AGAINST MODELS, CONSIDERING LAGS FROM ONE HOUR TO ONE WEEK.

Ctry. Code	Autocorrelation at given lag							
	1 hr		6 hrs		1 day		1 week	
	Data	Mdl.	Data	Mdl.	Data	Mdl.	Data	Mdl.
GB	0.97	0.97	0.86	0.85	0.62	0.53	0.27	-0.01
FR	0.99	0.98	0.97	0.87	0.89	0.62	0.69	0.08
DE	0.99	0.97	0.92	0.81	0.72	0.45	0.35	-0.04
ES	0.99	0.97	0.97	0.82	0.93	0.47	0.70	-0.01
NL	0.98	0.97	0.91	0.85	0.77	0.49	0.26	-0.05

collapsed empirical data and analytic models appeared reasonably consistent. As the time series for the five winters are not contiguous, the table reports the calculated mean of the autocorrelation for five winters (and five simulated winters).

It is clear that for GB the autocorrelation is relatively well-calibrated based on the model data (despite the aforementioned disparity between the changes in the data). However, it can be seen that the unavailability data for other countries (particularly France and Spain) show much higher autocorrelation than the modelled data, particularly at time periods longer than one hour. This appears to suggest that the time sequential modelling for those countries is not accurate for longer time periods.

## V. CONCLUSIONS

Accurate models of generator availability will become even more critical to ensure security of supply in energy systems based on a combination of renewable energy technologies and dispatchable peaker plant. This paper has considered how analytic generator unavailability models compare against empirical distributions derived from data sourced from the ENTSO-e Transparency Platform. Nine European systems were considered, centred on Northwestern Europe, with five winters studied based on tens of thousands of individual generator outage reports.

1) *Discussion and Future Recommendations:* Where studies require accurate historical unavailability data, it is suggested that researchers need to work with a relevant TSO directly to access the highest quality data. Even so, ex-post analysis of power system variables linked closely to generator unavailability, such as system margin, may remain elusive without a step-change in data quality and reporting.

The output of time-collapsed unavailability modelling of a fleet of generators is a PDF. For academic studies, using a Bernoulli model for individual generators appears satisfactory, except where a highly accurate distribution is required (in which case, interactions with a TSO are again necessary).

More complex time sequential models appear to show quite different properties than the empirical time series they are intended to represent, considering attributes such as autocorrelation. There therefore appears to a gap in the understanding of the properties of these time series, particularly in uncertainty-aware, secure scheduling of energy storage across daily to

weekly time periods. Future work could use the data presented to compare more detailed generator-based approaches (e.g., multi-level Markov models) and data-driven analysis of the whole fleet (e.g., using Box-Jenkins type models).

Clearly, generator unavailability data brings a range of insights to the nature and workings of modern power and energy systems. It is concluded that, despite its complexities, this data can and should feed into unavailability models required to plan for a secure, Net Zero energy system.

## REFERENCES

- [1] J. W. Busby, K. Baker, M. D. Bazilian, A. Q. Gilbert, E. Grubert, V. Rai, J. D. Rhodes, S. Shidore, C. A. Smith, and M. E. Webber, "Cascading risks: Understanding the 2021 winter blackout in Texas," *Energy Research & Social Science*, vol. 77, p. 102106, 2021.
- [2] National Grid ESO, "Introducing the virtual energy system," <https://www.nationalgrideso.com/news/introducing-virtual-energy-system>, 2021, accessed 15/1/21.
- [3] G. Edwards, S. Sheehy, C. J. Dent, and M. C. Troffaes, "Assessing the contribution of nightly rechargeable grid-scale storage to generation capacity adequacy," *Sustainable Energy, Grids and Networks*, vol. 12, pp. 69–81, 2017.
- [4] S. Murphy, L. Lavin, and J. Apt, "Resource adequacy implications of temperature-dependent electric generator availability," *Applied Energy*, vol. 262, p. 114424, 2020.
- [5] ENTSO-e, "ENTSO-e transparency platform," <https://transparency.entsoe.eu/>, 2021, accessed 20/11/21.
- [6] L. Hirth, J. Mühlenpfordt, and M. Bulkeley, "The ENTSO-e transparency platform—a review of europe's most ambitious electricity data platform," *Applied energy*, vol. 225, pp. 1054–1067, 2018.
- [7] A. Rafiee, M. Qadrdan, and N. Jenkins, "A data-driven approach to study the role of interconnectors in a future low-carbon electricity supply system," *International Journal of Energy Research*, vol. 45, no. 1, pp. 379–395, 2021.
- [8] A. A. Golub, K. Govorukha, P. Mayer, and D. Rübbecke, "How does climate change affect the transition of power systems: the case of Germany (FEEM working paper)." 2020.
- [9] M. Halužan, M. Verbič, and J. Zorič, "Performance of alternative electricity price forecasting methods: Findings from the Greek and Hungarian power exchanges," *Applied Energy*, vol. 277, p. 115599, 2020.
- [10] H. Wels, "Quality in big data for the forced unavailability of power plants," in *Advances in Reliability, Risk and Safety Analysis with Big Data: Proceedings of the 57th ESReDA Seminar*, 2019, p. 33.
- [11] I. Lorencin and M. Pantoš, "Evaluating generating unit unavailability using Bayesian power priors," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2315–2323, 2016.
- [12] National Grid ESO, "Modelling de-rating factors for interconnected countries in the 2021 electricity capacity report," <https://bit.ly/3JDdwc5>, 2021, accessed 14/4/22.
- [13] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems (2nd Ed.)*. Plenum Press, 1996.
- [14] —, *Reliability Evaluation of Engineering Systems*. Springer Science, 1992.
- [15] National Grid ESO, "Electricity capacity report," <https://www.emrdeliverybody.com/CM/Guidance.aspx>, 2020.
- [16] The European Commission, "Commission regulation (EU) no 543/2013 on submission and publication of data in electricity markets and amending annex I to regulation (EC) no 714/2009 of the European parliament," *Official Journal of the European Union*, vol. 163, no. 1, June 2013, <http://data.europa.eu/eli/reg/2013/543/oj>.
- [17] M. Deakin, H. Bloomfield, D. Greenwood, S. Sheehy, S. Walker, and P. C. Taylor, "Impacts of heat decarbonization on system adequacy considering increased meteorological sensitivity," *Applied Energy*, vol. 298, p. 117261, 2021.
- [18] M. Deakin and D. Greenwood, "Dataset: Aggregated generator unavailability data for northwest european countries," <https://doi.org/10.25405/data.ncl.18393971.v1>, January 2022.