tnei

A specialist energy consultancy

Final Report

QWID FLEXER

National Grid ESO

15929-03-R0
10 April 2024

COMMERCIAL IN CONFIDENCE

nationalgridESO

Quality Assurance

TNEI Services Ltd and TNEI Africa (PTY) Ltd operate an Integrated Management System and is registered with The British Assessment Bureau as being compliant with ISO 9001(Quality), ISO 14001 (Environmental) and ISO 45001 (Health and Safety).

Disclaimer

This document is issued for the sole use of the Customer as detailed on the front page of this document to whom the document is addressed and who entered into a written agreement with TNEI. All other use of this document is strictly prohibited and no other person or entity is permitted to use this report unless it has otherwise been agreed in writing by TNEI. This document must be read in its entirety and statements made within may be based on assumptions or the best information available at the time of producing the document and these may be subject to material change with either actual amounts differing substantially from those used in this document or other assumptions changing significantly. TNEI hereby expressly disclaims any and all liability for the consequences of any such changes. TNEI also accept no liability or responsibility for the consequences of this document being relied upon or being used for anything other than the specific purpose for which it is intended, or containing any error or omission which is due to an error or omission in data used in the document that has been provided by a third party.

This document is protected by copyright and may only be reproduced and circulated in accordance with the Document Classification and associated conditions stipulated or referred to in this document and/or in TNEI's written agreement with the Customer. No part of this document may be disclosed in any public offering memorandum, prospectus or stock exchange listing, circular or announcement without the express and prior written consent of TNEI. A Document Classification permitting the Customer to redistribute this document shall not thereby imply that TNEI has any liability to any recipient other than the Customer.

Any information provided by third parties that is included in this report has not been independently verified by TNEI and as such TNEI accept no responsibility for its accuracy and completeness. The Customer should take appropriate steps to verify this information before placing any reliance on it.

Document Control

Revision	Status	Prepared by	Checked by	Approved by	Date
D0	Draft	RT, SS, OP, GM	SS	GM	13/11/2023
D1	Draft	RT, SS, OP, GM	SS, GM		20/02/2024
R0	First	RT, SS, OP, GM	SS, GM	GM	10/04/2024

TNEI Services Ltd
Company Registration Number: 03891836 **VAT Registration Number: 239 0146 20**

Registered Address		
Bainbridge House	7th Floor West One	7th Floor
86-90 London Road	Forth Banks	80 St. Vincent Street
Manchester	Newcastle upon Tyne	Glasgow
M1 2PW	NE1 3PA	G2 5UB
Tel: +44 (0)161 233 4800	Tel: +44 (0)191 211 1400	Tel: +44 (0)141 428 3180

TNEI Ireland Ltd
Registered Address: 104 Lower Baggot Street, Dublin 2, DO2 Y940
Company Registration Number: 662195 **VAT Registration Number: 3662952IH**

Unit S12, Synergy Centre
TU Dublin Tallaght Campus
Tallaght
D24 A386
Tel: +353 (0)190 36445

TNEI Africa (Pty) Ltd
Registered: Mazars House, Rialto Rd, Grand Moorings Precinct,7441 Century City, South Africa
Company Number: 2016/088929/07

Unit 514 Tyger Lake
Niagara Rd & Tyger Falls Blvd
Bellville, Cape Town
South Africa, 7530

Contents

Document Control 3
Contents 4
1 Introduction 9
2 Methodology 11
2.1 Modelling of long-run variability 13
2.2 Inputs 15
2.2.1 Weather-rescaling of time series 15
2.2.2 Weather rescaling of low-carbon technology profiles 35
2.3 Calculations 40
2.3.1 FES scaling 40
2.3.2 Calculating within-day flexibility 41
2.3.3 Extreme value theory 42
2.4 Output 44
2.4.1 Metric definitions 44
2.5 Key assumptions 47
2.6 Alternative methodology options 48
3 Results 50
3.1 Types of plots 50
3.2 Results 51
3.2.1 Daily peak power 51
3.2.2 Ramp rates 56
3.2.3 Energy metrics 59
3.2.4 Crossing points 63
3.2.5 Duration of energy retention 66
3.3 Summary of all metrics 68
3.4 Disaggregation of WDFN 73
3.5 Change in WDFN over time 76
4 Sensitivity Analysis 79
4.1 Electric Vehicles 79
4.2 GB-Ireland interconnectors 81
5 Discussion 86
5.1 Key results 86
5.2 Future work 86

6 Glossary 88
Appendix A: Model hyperparameters 89
Appendix B: Feature Importance 90
Appendix C: Data sources 95
Appendix D: Code Structure 97
Appendix E: Python Packages 98

TABLES

Table 1: Terms used in formulation of GAM model 20
Table 2: GAM Model Performance 22
Table 3: Residual (quantile regression) model features 23
Table 4: BMU wind model features 28
Table 5: Embedded Solar model features 32
Table 6: Description of BMU hydro model features 34
Table 7: Summary values at key quantiles for all numeric metrics. 68
Table 8: Summary of Boolean metrics 72
Table 9: Hyperparameters used in lightGBM fitting 89
Table 10: Core python packages 98

FIGURES

Figure 1: Methodology flowchart showing the process of calculating WDFN and its metrics 12
Figure 2: Original (blue) and rescaled (red) demand for three different target years[1] 14
Figure 3: Example probabilistic prediction of wind load factor through time. The different colour bands represent different probability intervals of the prediction, from the lightest blue (q45 to q55) to the darkest (q5 to q95) 16
Figure 4: Converting load factor from historic data (for year a) to an out-turn probability, and then from the out-turn probability to a load factor representing a different year (year b, with different weather and time features) 17
Figure 5: Time series of out-turn probabilities 17
Figure 6: Map of Local Authority areas. Black points show the centre points of areas used to create population-weighted weather. 18
Figure 7: GAM prediction and residuals for an example weather and demand scenario day. The probabilistic model is fitted to the residuals. 22
Figure 8: Distribution of predicted quantiles before applying tail modelling 24
Figure 9: Histogram of lower tail residuals (Blue) against fitted GPD distribution. 24
Figure 10: Histogram of upper tail residuals (Blue) against fitted GPD distribution (Orange). 25
Figure 11: Distribution of predicted residual quantiles after applying GPD tails. 25

Figure 12: Reliability plot for 4 CV folds of lightGBM Quantile Regression Model of demand residuals26

Figure 13: National onshore and offshore load factors from 2015-2023 27

Figure 14: Locations of onshore (blue) and offshore (red) wind farms used in calculation of national load factors. Area of circle represents wind farm capacity. 27

Figure 15: Reliability of wind model. Times of high curtailment show bias. 29

Figure 16: National embedded wind load factors 30

Figure 17: Reliability of embedded wind model. 30

Figure 18: Sharpness of wind models. 'Width' denotes the interval width for the corresponding probability interval, i.e. there is an average difference in predicted load factor of 0.05 for an interval of 0.5 (between q25 and q75) for the onshore wind model. Lower widths indicate a sharper, more confident prediction. 31

Figure 19: National embedded solar load factors 31

Figure 20: Reliability of solar predictions, split by day and nighttime 33

Figure 21: National load factors for non-pumped hydro BMUs from 2015-2023. 33

Figure 22: Reliability of BMU hydro model 35

Figure 23: Sharpness of BMU hydro model 35

Figure 24: Average electric heat demand profile from demonstration project data 37

Figure 25: Average electric heat demand profile after rescaling to reanalysis weather data. 37

Figure 26: Example days from Saturday-February residential archetype from EE EV charging profiles.39

Figure 27: Profile of Saturday-February Residential profile with temperature scaling removed 39

Figure 28: Temperature Scaling Factor of Monday-February Residential Archetype profile used by Element Energy 40

Figure 29: EV demand profile on February 4th, rescaled to average temperature in 2019. 40

Figure 30: Comparison of demand and generation elements for a typical day under different FES ... 41

Figure 31: Example day showing stacked demand and generation, and the calculated daily imbalance42

Figure 32: Example of a central distribution below a certain probability threshold, and a tail distribution beyond. 43

Figure 33: Example GPD tail distributions with varying scale and shape 44

Figure 34: Example cumulative distribution function 50

Figure 35: Example of a violin plot 51

Figure 36: Distribution of peak upwards and downwards WDF need by season, for Leading the Way 2035 52

Figure 37: Distribution of peak WDFN across seasons for Leading the Way 2035 52

Figure 38: Proportion of the time peak WDF need is in the upwards direction, for Leading the Way 2035. 53

Figure 39: Distribution of daily difference in magnitude of peak upwards and downwards need by season, for Leading the Way in 2035 53

Figure 40: Variation and spread of peak WDFN throughout the year, for 2035 54

Figure 41: Distribution of peak upwards and downwards WDF need across all FES for 2035 54

Figure 42: Distribution of peak upwards and downwards WDFN across all FES for 2035 55

Figure 43: Proportion of the time peak WDFN is in the upwards direction, by scenario 55

Figure 44: Difference in magnitude of peak upwards and downwards peaks, as a proportion of daily WDFN range 56

Figure 45: Maximum daily half-hourly ramp rates by season for Leading the Way 2035 56

Figure 46: Maximum daily 2-hour ramp rates for Leading the Way 2035 57

Figure 47: Peak 4-hour ramp rates for Leading the Way 2035 57

Figure 48: Comparison of half-hour, 2-hour and 4-hour upwards ramp rates for Leading the Way, 2035 58

Figure 49: Distribution of maximum half-hourly upwards and downwards ramp rates under FES for 2035 58

Figure 50: Distribution of maximum 2-hourly upwards and downwards ramp rates under FES for 2035 59

Figure 51: Distribution of maximum 4-hourly upwards and downwards ramp rates under FES for 2035 59

Figure 52: Seasonal distributions of total energy (area above zero) for Leading the Way 2035 60

Figure 53: Distributions of total energy (area above zero) by FES, for 2035 60

Figure 54: Seasonal distributions of equivalent battery energy capacity for Leading the Way 2035, to meet the largest single WDFN cycle within the day 61

Figure 55: Distributions of equivalent battery energy capacity by FES 61

Figure 56: Distribution of cumulative energy capacity needed if one energy store was to meet all the WDFN by charging and discharging throughout the day, by season for Leading the Way 2035 62

Figure 57: Distribution of cumulative energy capacity needed if one energy store was to meet all the WDFN by charging and discharging throughout the day, by FES scenario for 2035 63

Figure 58: Distributions of zero-crossings per season for Leading the Way 2035 64

Figure 59: Direction of WDF need at the start of the day (midnight), for Leading the Way 2035 64

Figure 60: Distribution of time of longest and shortest cycles within the day, for Leading the Way 2035 65

Figure 61: Distributions of zero-crossings per FES in 2035 65

Figure 62: Direction of WDFN at the start of the day (midnight), for FES 2035 65

Figure 63: Distribution of time of longest and shortest cycles within the day, for FES 2035 66

Figure 64: Distribution of duration between peak upwards and downwards need for Leading the Way 2035 66

Figure 65: Violin plots of distribution of duration between peak upwards and downwards need for FES 2035 67

Figure 66: Average daily WDFN for individual generation and demand profiles, in Leading the Way 2035, indicating the average (solid), 10th and 90th quantiles (dashed), and 1st and 99th quantiles (shaded). 73

Figure 67: WDFN summary statistics for individual generation and demand profiles, upwards and downwards power, under Leading the Way 2035 74

Figure 68: WDFN summary statistics for individual generation and demand profiles, upwards and downwards ramp rate, under Leading the Way 2035 75

Figure 69: WDFN summary statistics for individual generation and demand profiles, upwards and downwards normalised power, under Leading the Way 2035 76

Figure 70: WDFN summary statistics for individual generation and demand profiles, upwards and downwards normalised ramp rates, under Leading the Way 2035 76

Figure 71: Year-on-year change in WDFN for max upward power under different FES, until 2035 77

Figure 72: Year-on-year change in WDFN for max downward power under different FES, until 2035 77

Figure 73: Year-on-year change in WDFN for max cycle capacity under different FES, until 2035 78

Figure 74: Comparison of the distribution of WDFN with and without EVs for Leading the Way 2035 80

Figure 75: Distribution of peak WDF need with and without naively charging EVs for Leading the Way 2035 80

Figure 76: Distribution of time between global peaks with and without naively charging EVs 80

Figure 77: An example day showing WDFN profiles for GB and Ireland, and the resulting interconnector flows. 82

Figure 78: Diagram of interconnector flow depending on WDFN direction in both GB and Ireland ... 83

Figure 79: Proportion of the time GB and Ireland WDFN are in each direction 83

Figure 80: Distribution of the interconnector power flow to Ireland across the times used in the sensitivity analysis 84

Figure 81: Summary probability levels for change in max upwards and downwards power metrics when GB-Ireland interconnection is included (compared to the values of these metrics without interconnection) 85

Figure 82: Summary probability levels for change in total energy and battery cycle capacity metrics when GB-Ireland interconnection is included (compared to the values of these metrics without interconnection) 85

Figure 83: Average feature importance across all quantiles for the BMU wind model 90

Figure 84 Average feature importance across all quantiles for embedded wind model 91

Figure 85: Average feature importance across all quantiles for solar model 92

Figure 86: Average feature importance across all quantiles for non-pumped storage BMU hydro. ... 93

Figure 87: Feature importance in terms of accuracy gain for Quantile regressor for probability level 0.6. 94

Figure 88: Feature importance in terms of accuracy gain for quantile regressor for probability level 0.02. 94

1 Introduction

The decarbonisation of the GB electricity network is happening at pace, introducing many changes to the way the system operates. There is a huge amount of work going on across industry to enable flexibility in the energy system and ensure the grid continues to serve its core function of reliably delivering electricity to customers in the face of changing patterns, increased variability and changing levels of dispatchability in both generation and demand. This necessitates the development of new analyses and processes that are fit for use under both current and future system conditions.

One obstacle in achieving a net zero-carbon network is the need to replace within day flexibility (WDF) from fossil fuels with a system where zero-carbon resources can deliver WDF. Determining how much total WDF is required is an essential step in moving forward with this transition and understanding how this might evolve under different scenarios.

Within-day flexibility is used to match demand and generation across daily peaks and troughs, usually lasting for a couple of hours. A deficit in generation would be resolved through "upward-flexibility", either increasing the amount of energy available to the network or decreasing energy demand, and vice-versa for downward flexibility.

Increasing penetration of renewables is likely to increase the number and size of generation surpluses at certain times, for example wind generation overnight when demand is low. As the contribution to flexibility from traditional dispatchable generation decreases, it is also envisaged that demand-side solutions will increasingly be employed for flexibility, as well as other sources of generation and storage. Zero-carbon WDF will include both generation and demand suppliers to resolve both deficits and surpluses in generation.

NGESO has commissioned TNEI to undertake analysis into the need for WDF as part of the QWID FLEXER NIA project. This project addresses key questions for determining the need for WDF, including:

- **What metrics are most appropriate for measuring the need for WDF?** While remaining solution-agnostic, it is important to provide metrics that enable answers to questions such as 'how much of the WDF need could be covered by EVs?'. As such, we have implemented a range of metrics covering not only the peak needs within a day, but also the required capacity for a flexibility source behaving as a battery and metrics describing the shape of the need across the day. Extreme value modelling will also be used to model low-probability, high-need events over a threshold. This will be implemented in the latter half of the project.
- **What sources of demand and generation should be included in the calculation of WDF need?** Broadly, the WDF needs calculation should be based on 'unmodified' behaviour, e.g., unconstrained renewable generation and demand without any flexibility. This is discussed further in section 0 for each element. Any behaviour deemed controllable, whether by ESO or any other party, and whether directly or through price signals or incentives, is classed as 'flexible' behaviour. This includes, for example, BMU curtailments, EV smart charging and ANM by DNOs. One challenge for defining this is that some types of generation and demand might be *partially* flexible and *partially* inflexible. For example, some heat pumps might be controllable whereas others aren't, or demands might be flexible *up to a point* (e.g., an EV might be flexible about whether it charges at 7pm, 8pm, or 9pm, but it absolutely must charge at some point for an hour between 7pm and 10pm). For this work, we have modelled what we consider to be completely inflexible behaviours (and made assumptions that all EVs and heat pumps are inflexible). It would be possible to extend the modelling to consider the profiles of flexible assets separately in future. From the data and information

currently available, any estimate of flexibility would be judgement based and subject to uncertainty.

- **What impact do different assumptions have?** All assumptions have been documented in the assumptions log, noting the expected effect on the end WDF need for each. Assumptions are discussed in the relevant report section, and the assumptions log will be delivered alongside the final report.
- **What uncertainties and variabilities should be considered in the methodology, and do we expect the nature or levels of these to change in the future?** To consider uncertainties appropriately, a probabilistic methodology has been implemented. The variability of inputs based on historic data (weather, demand, generation) is modelled explicitly and the uncertainty in projections for future technology uptakes and behaviours taken from the Future Energy Scenarios (FES) modelling is considered through sensitivity analyses where it is deemed appropriate.
- **How would inclusion of locational and network constraints affect the need for WDF?** We understand this is beyond the scope of this work and would add significant complexity, but this has been kept in mind in designing the approach as this is likely relevant for future work.
    - We would expect locational constraints to increase the need for WDF overall and introduce differences in WDF need at different points in the network. However, it may also lead to useful coincident WDF requirements at different levels of the system.
    - The sensitivity analysis for GB-Ireland interconnection presented in section 4.2 demonstrates the approach for treating interconnected networks and their impact on within-day flexibility – locational constraints between regions of Great Britain and interconnectors between European countries should, in principle, be treated similarly.

2 Methodology

The overall methodology for calculating WDF need, and any derived metrics is set out in Figure 1 overleaf. While the central calculation of WDF need, $WDFN_{d,t}$, for each half-hour t on day d is a simple difference between demand $D_{d,t}$ and generation $P_{d,t}$ at that time, the preparation and modelling of the respective demand and generation time series required careful treatment.

The basic process is:

- Inputs specific to each Future Energy Scenario (FES) were taken from public FES documentation. These are generally *annual* variables (e.g., total energy consumption in a year, or installed capacity of generation).
- Inputs which are agnostic of the specific FES, like half-hourly renewable load factors, were assembled based on modelling and rescaling of historic time series data. These are time series, the majority of which have a half-hourly resolution.
- Rescaled time series were multiplicatively scaled to the corresponding values from FES, for example time series of generation load factors were multiplied by FES installed capacity. This resulted in MW time series for demand and generation for each component, for any year of any specific FES.
- The daily and half-hourly imbalances between demand and generation were calculated to determine the WDFN for each half-hour.
- From this time series of WDFN, various metrics for each day were calculated, such as the peak need, total energy to fulfil all upwards need across the day, ramp rates, or the size of theoretical battery required to provide flexibility to cover the need.
- The distributions of these daily metrics across the complete set of days are returned, with Extreme Value Theory (EVT) used to model the tails beyond a chosen probability threshold. This is used to determine both very infrequent WDF needs (e.g., levels of need that might only be needed for one half hour per year, or once every ten years), and the most extreme levels of the summary metrics, for metrics where this is applicable.

It is important to note that while results can be generated for any future year in the FES, the main immediate interest is assessing the WDF need out to 2035 so the methodology is based on some assumptions that may not hold far into the future, for example electrification of shipping and aviation has been neglected as this is not expected to reach significant levels by 2035. All assumptions made across all stages of the modelling and calculations are detailed in the assumptions log, which is included as an accompanying Annex, and discussed in Section 2.5.

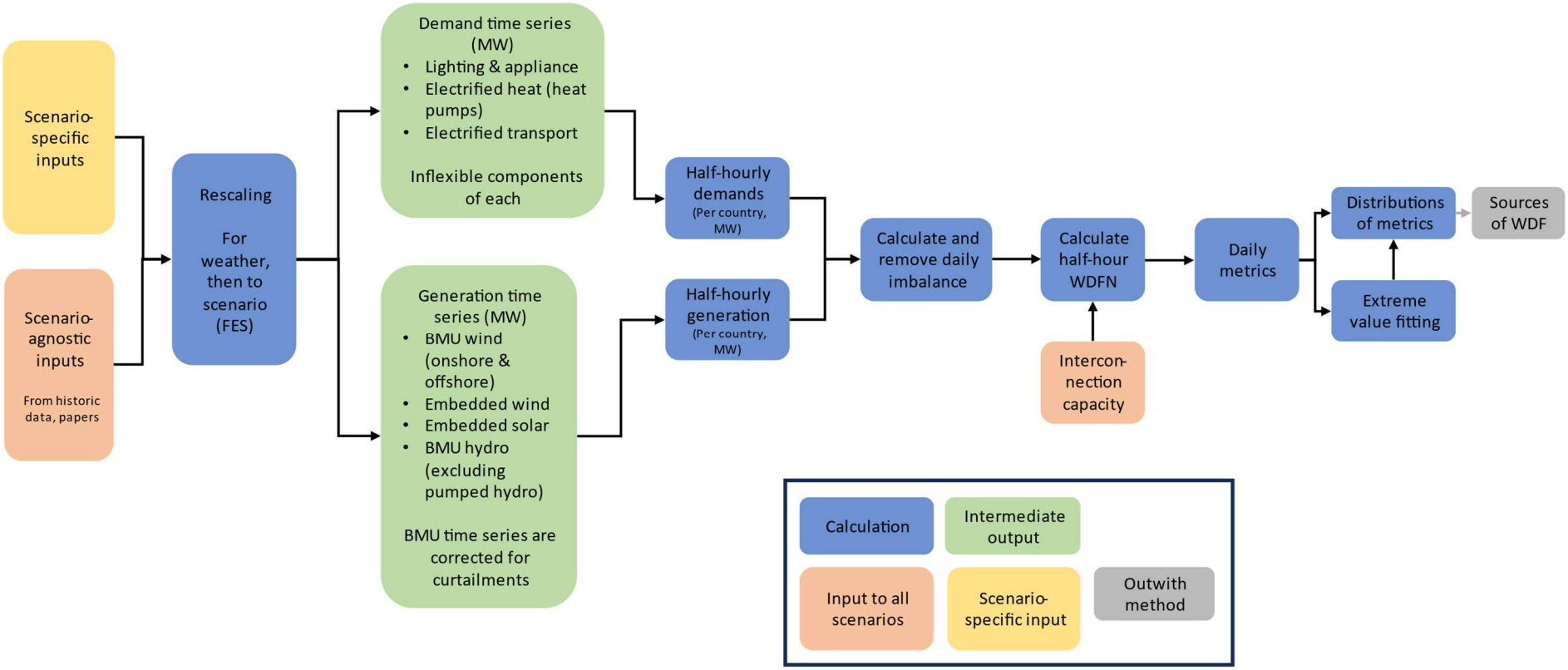


Figure 1: Methodology flowchart showing the process of calculating WDFN and its metrics

2.1 Modelling of long-run variability

One aim of this analysis is to provide insight as to what levels of WDFN might be typical, and what levels might only be required in extreme cases (hence the inclusion of EVT in our methodology). This means it is important to capture the way that electricity supply and demand might vary in the long-term (e.g., between one year and the next) as well as over the course of a day.

Capturing this long-run variability requires analysis of long-spanning time series. However, the availability of measured quantities for demand and generation is limited, particularly for newer technologies such as electric vehicles and heat pumps. Additionally, the relevance of older time series is debatable, particularly for demand, where more recent data reflects changes in behaviour which older datasets will not reflect. For generation, newer data might reflect trends in technology (e.g., taller turbines which lead to higher load factors).

However, high quality (estimated) data is available about how weather conditions have varied over timescales of decades and, given that many aspects of energy supply and demand (particularly wind and solar production) are weather dependent, this can be a very useful resource. Therefore, methods are required to generate realistic long time series of demand and generation quantities based on the short time series of measured values (or profiles in the case of heat pumps and EVs), and longer time series of weather data. It is important that these methods produce individual time series that are realistic while also providing relationships between time series which are realistic.

One area within energy systems where such time series and relationships are modelled frequently is the assessment of capacity adequacy for the Capacity Market. NGESO has well defined processes and methodologies that are embedded within business-as-usual processes; however, there are also many energy systems researchers and statisticians actively working in this area. The approach to generating time series adopted by TNEI within this project is therefore heavily inspired by a 2022 conference paper from academics at the University of Edinburgh[1], produced (largely) in the context of capacity adequacy analysis.

In that paper, the authors present a methodology for rescaling multiple years of historic demand data so that the original data from any specific measurement year can be used in simulation as if it had been measured in any other target year. Examples from the paper are shown in Figure 2. The authors used 25 years of time series data and can then use of those years of measurement data to describe any target year from that data set (1990, 2002, 2015 etc).

The authors' rescaling process involves fitting a regression model to daily peak electricity demand as a function of various calendar and weather variables. When rescaling to any specific target year, the weather is also "swapped", so that the rescaled data reflects the weather from the target year rather than the measured year.

[1] Wheatcroft, E., Dent, C., and Wilson, A. (2022) Rescaling of historic electricity demand series for forward-looking risk calculations.

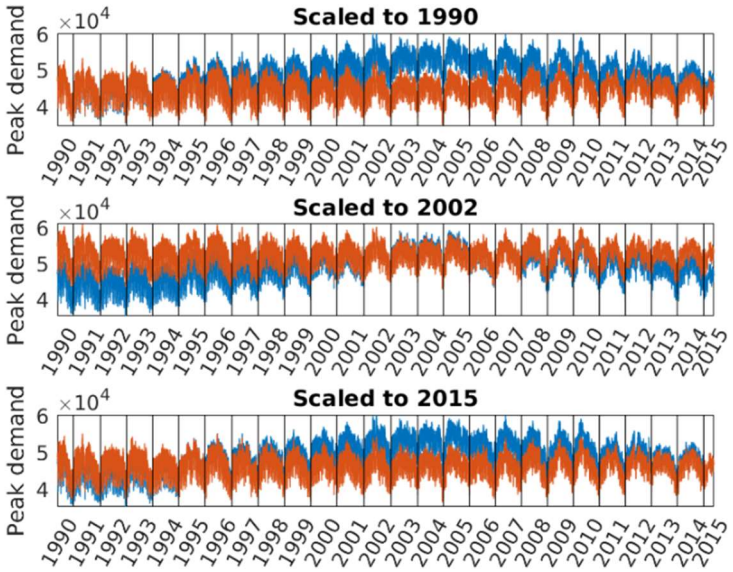


Figure 2: Original (blue) and rescaled (red) demand for three different target years[1]

The approach adopted in this innovation project extends this approach in several ways:

- Firstly, this project does not restrict the analysis to considering *daily* peaks, and instead considers half-hourly values.
- To achieve this, the simple linear models used within the original paper[1] are replaced with more flexible machine learning approaches currently used at the cutting-edge for energy forecasting, including Generalised Additive Models (GAMs) and Gradient Boosting Machines (GBMs).
- These methods permit much greater flexibility in how *probability densities* are modelled. In the paper[1], the authors implicitly assumed that the *residual* of daily peak electricity demand (i.e., any difference from its *expected value*) is normally distributed, and that the spread of the distribution is constant within a year but varies from one year to the next. The methods used in this innovation project do not require any assumptions to be made about probability distributions and allow for the possible spread of the residual to vary from one half-hour to the next.
- The method presented in the paper[1] does not mix and match between weather years and target years. In other words, when rescaling measurements to a specific target year, the authors only apply the weather data from that target year. In this project, we permit mixing and matching between weather and target years – in other words, for data collected in any specific year, we produce dozens of versions of how that data would change based on multiple different weather patterns.

This final step assumes that, after fitting the machine-learning regression models, there is no relationship between the *conditional* outcome and the input weather data. This is necessary because, otherwise, we would be restricted to the length of the shortest available measured demand and generation time series, with fewer than ten years of data. Applying this method to this project also assumes that the sequences of conditional out-turn probabilities that are possible in the future will align with sequences of conditional out-turn probabilities that are observed within the past behaviour of this probabilistic modelling. This is necessary when considering daily *sequences*. This may be a strong assumption since models rarely fit exact relationships and there may be some remaining

correlations with weather data or other features in the predictions. However, this approach allows generation of a large set of weather time series that we believe are still useful and appropriate for the application.

2.2 Inputs

All weather rescaling was conducted using UTC timestamps to ensure dates and times align between years, as clock changes fall on different days in different years. Once rescaled time series of the demand and generation elements were calculated, local time was used to define the midnight cutoff point between days, which are then treated independently for calculating WDF need.

Years have been defined as running from 1st Oct – 30th September, with the labelling convention that 1st Oct 2022-30th September 2023 is labelled as 2023. This ensures each whole winter season falls within one labelled year, rather than being split across years if calendar years were used. While not used in the current model forms, this is also an important prerequisite for adding any per-year random effects to the models which could be a future extension.

2.2.1 Weather-rescaling of time series

Without weather rescaling, we would be limited to the length of historic data available for all component time series, i.e. 2015-2023. This length of dataset would only be expected to include 1 1-in-9-year event, meaning there would be very little data to base analysis of extreme events on. Where a time series of measured historic generation was available, and the variable is dependent on the weather in some way, a weather rescaling methodology has been applied to generate longer plausible time series of the variable of interest based on long historic weather time series. Using 43 years of weather data in combination with the 9 available years of observations allows 43×9=387 rescaled years based on all possible combinations of weather and observation years, greatly augmenting the set of historic days from which to calculate WDFN metrics. In addition, the rescaling process allowed rescaling of any time-based changes in behaviour to present day, for example trends and the effect of COVID on demands, giving (approximately) stationary rescaled time series that represent demand and generation components that could have been expected, if the historic weather occurred under the present (2023) energy system.

Weather data from the European Centre for Medium-range Weather Forecasting's (ECMWF) ERA5 (ECMWF ReAnalysis 5) product was used[2]. This is the highest resolution reanalysis product available for Europe and is widely considered to have good accuracy for the region. Reanalysis data is generated from a large set of weather measurements which are fed into a numerical weather model to determine the atmospheric state. This is then used to output coherent estimates of weather at uniform grid points across the surface of the globe, and at different atmospheric heights. This is not producing forecasts, but rather providing the best estimate of historic weather at points across a region. The ERA5 dataset spans from 1940 to present, and for this work we have used data from 1980 to 2023. This length of data was chosen to minimise data volumes while still ensuring a sufficiently large set of weather years were available for rescaling. Furthermore, the older years of ERA5 are less likely to be representative of current climate than more recent years and are based on lower-resolution weather inputs and so may not maintain the same accuracy as more recent data.

In general, the rescaling method takes sequences of out-turn probabilities, which are derived from observations, and overlays different weather and time features. The general structure of the weather rescaling method consists of two steps:

[2] ECMWF Re-analysis v5 (ERA5) data summary: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5

- Fit a probabilistic model relating weather and time features to the variable of interest. The form of this model is tailored to the variable and could be parametric (predicting the parameters of a distribution) or nonparametric (directly predicting a set of quantiles). All the components where weather rescaling has been applied use nonparametric quantile models, an example of which is shown in Figure 3.
- To generate new year-long time series scaled to levels in year k, based on the weather in year j and the realisations from year i, probabilistic predictions using the weather from year j and the time features from year k are combined with out-turn probabilities from year i.

Out-turn probabilities for year i are derived by combining observations and probabilistic model predictions. The example below shows the predicted distribution of wind load factors in the blue intervals, with the measured national load factor for the same time overlaid (the black trace). These blue intervals provide a probabilistic prediction of what the historic wind load factor was, based on knowledge of historic weather but without any historic measurement of wind farm output.

For each time point, we can effectively take a vertical 'slice' through Figure 3 to get the cumulative distribution function and a single realisation for that time and weather conditions. The observed value at this time (from the black line in Figure 3) is 0.08, so we look up the cumulative probability that a load factor of 0.08 corresponds to in the red predicted distribution (Figure 4), giving an out-turn probability of 0.26 for this time point. This is repeated for all time points to generate time series of out-turn probabilities as shown in Figure 5. These time-series display autocorrelation, where high values are more likely to be followed by further high values, and low values by further low values etc. Retaining this autocorrelation through the sequence of out-turn probabilities allows generation of new (rescaled) time series of demand/generation with the same relationships between time points as the original measured time series.

The final step is to convert the out-turn probabilities from year i to a time series of values that are representative of year k, if the weather of year j had occurred at that time. To do this, predictions are generated based on the time features from year k and the weather from year j. The out-turn probabilities can then be converted to values for the variable of interest by the reverse transformation through the cumulative distribution function (CDF) as in the blue CDF in Figure 4. The time series of out-turn probabilities are treated in year-long blocks, so that an out-turn probability based on a certain day of the year and time of day will only be combined with predicted distributions matching the same time, but from a different weather year and with different trends. This means that not only daily correlations, but also seasonal patterns are preserved. This assumes that these conditional correlation patterns will not change in the future.

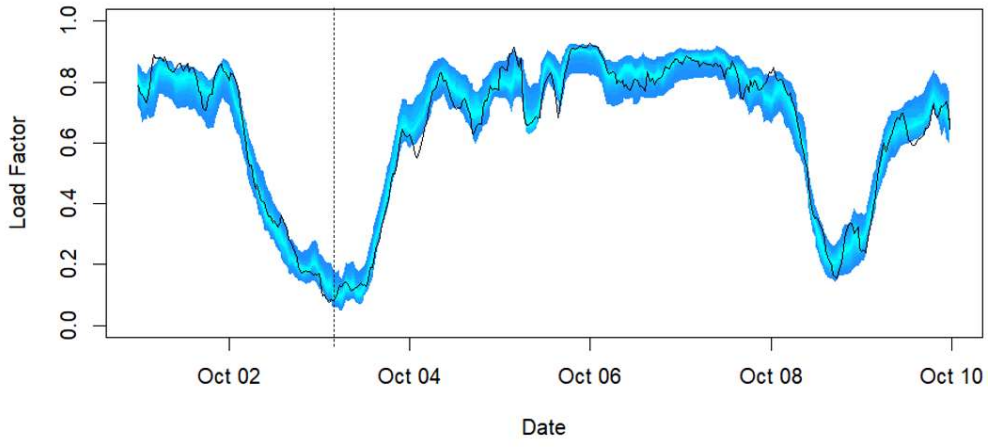


Figure 3: Example probabilistic prediction of wind load factor through time. The different colour bands represent different probability intervals of the prediction, from the lightest blue (q45 to q55) to the darkest (q5 to q95).

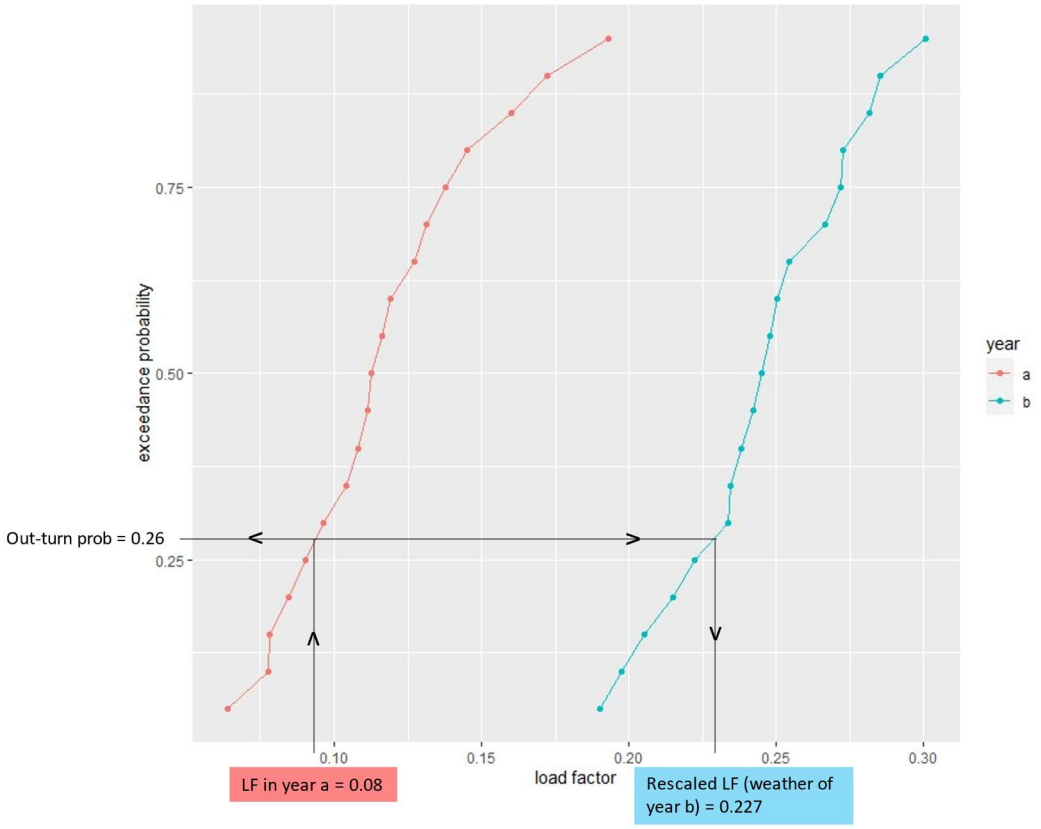


Figure 4: Converting load factor from historic data (for year a) to an out-turn probability, and then from the out-turn probability to a load factor representing a different year (year b, with different weather and time features)

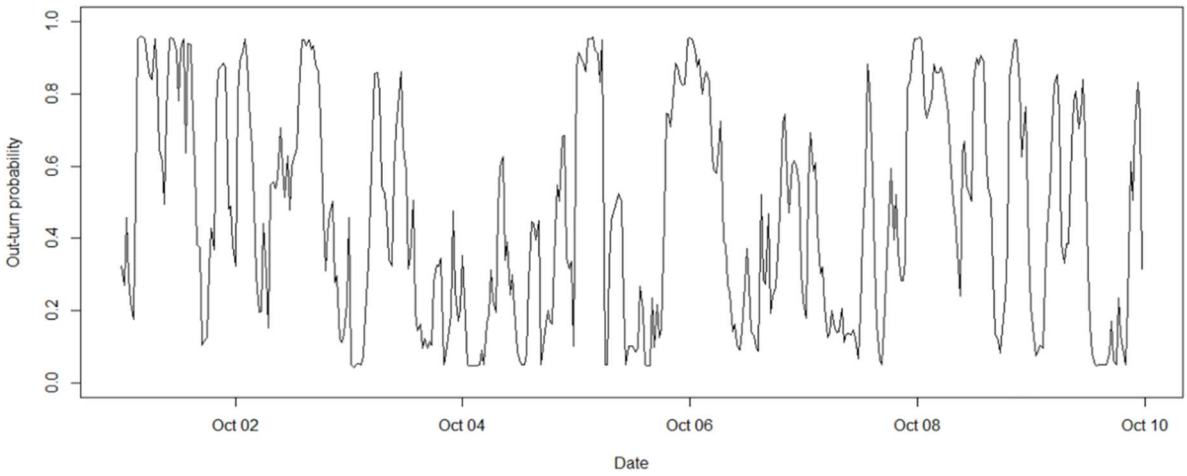


Figure 5: Time series of out-turn probabilities

For this work, we have scaled all time-series to the most recent historic data year of 2023 (1st October 2022-30th September 2023). With 9 years of historic measurements of BMU generation (2015-2023) and 43 years of weather data (1981-2023), this gives 9×43 = 387 years' worth of rescaled data for each variable this is applied to. This forms the basis of our scenario agnostic inputs.

2.2.1.1 Appliance, lighting, and resistive heating demand

Lighting, appliance, resistive space heating and hot water demand is a representation of any demand which has not been included in each other modelled form of demand (including industrial and commercial demand), i.e., any demand not classified as heat pumps or electrified transport. National Grid ESO's data portal provides aggregated demand data from 2009 through to the present data. Only the years 2015-2023 are used in our modelling, to align with available generation data as we need the same set of measurement years and weather years across all time-series.

ERA5 weather data was used to generate national average time series of weather which have been weighted for local population density, providing 43 years of weather data relevant to the main sources of demand. This population weighting has been carried out by local authority area; since these can be quite small in area, especially for the most densely populated regions, a given local authority area may not contain any points on the grid of ERA5 values. Because of this, weather was interpolated to the centre of the local authority area rather than averaging all grid points contained within an area. The 50% of areas with the densest populations, shown in Figure 6, were then used to calculate weighted average weather. The 50% lowest density areas are neglected both to reduce the number of data points to process, and because for these generally larger and less population-dense areas the central point is less representative of the locations of population.

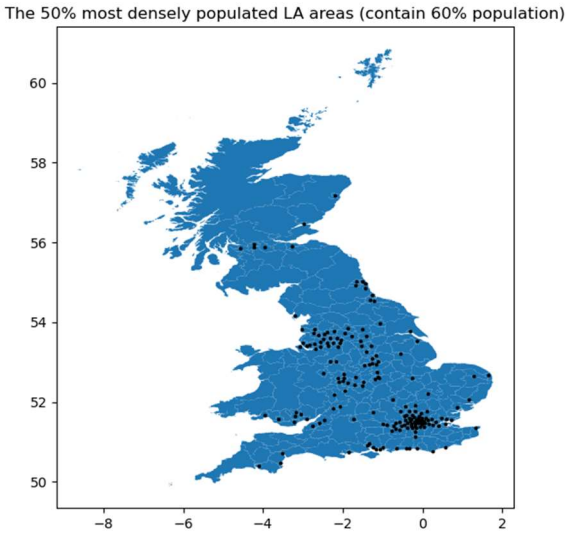


Figure 6: Map of Local Authority areas. Black points show the centre points of areas used to create population-weighted weather.

First, a GAM (implemented using the pyGAM package) is fitted to predict expected national demand, using both calendar variables and national average weather terms as predictors of its form. That model is trained on the entire demand dataset (and coincident weather inputs), and then used in-sample to produce deterministic mean predictions of demand for each day of historical data. Subtracting this demand prediction from the realisation yields a time series of residuals the same length as the historic demand data.

A probabilistic quantile regression model is then fitted to these residuals. This two-stage modelling method allows use of a more explainable model with complex smooth relationships and interactions between variables for the mean demand (from the GAM), and then a nonparametric model for the distribution around that mean, such that no assumptions about the shape of the distribution of residuals are needed. The quantile model estimates quantiles for a set of probability levels which together describe the distribution of the demand residuals at each timestep. To perform this task, a

Gradient-Boosting Machine (GBM) was selected over a GAM. GBMs fit a set of decision trees, with multiple splits per tree on the input features. Each subsequent tree is trained on the residuals of the last, forming a 'boosting' method. While the parameters of a GBM are less explainable than those of a GAM, it can capture complex and unexpected relationships between forms, is not limited to some parametric form for the forecast distribution and has a level of in-built feature selection. By using a GAM to first predict demand, explain-ability of the central location of the overall predicted distribution is maintained. The GBM model is implemented in the python lightGBM package due to its excellent computational performance and well written interface.

A downside of a GBM is their reduced effectiveness when predicting values outside the range of values provided in the original training set, and therefore they are unable to capture what might happen in the most extreme scenarios that may be expected to exceed the range of the training data. To overcome this shortcoming, we have opted to implement static Generalised Pareto Distributions (GPD) to model the residual quantiles after pre-determined probability levels that mark the tails of the predicted residual quantile distribution. The GPD models are considered static because their shape and scale parameters are fitted to all residuals within the tails of the training dataset, so that single values for shape and scale are used for all time points. The location parameter varies per time point and is determined by the quantile models' prediction at the corresponding threshold probability level. With the static shape and predicted location, the GPD is then used to mask the originally predicted residual quantiles. This provides the benefits of the nonparametric GBM as well as effective modelling of the extremes and events not seen within the original training set through the GPD. For the full details of extreme value theory, see section 2.3.3.

For each time step of the historic demand data, a predicted mean value and set of probability levels for its residuals have been calculated. The rescaling process described in Section 2.2.1 is performed on the residuals, which are then added on to a prediction of mean demand from the GAM model using the corresponding time and weather features as inputs. For example, if a set of residuals were generated by scaling to year k, with weather from year j and based on measured residuals from year i, this would be combined with demand predictions generated with time features from year k and weather features from year j.

For scaling residuals to year k, based on weather from year j and historic residuals from year i, the process follows these steps:

1) Combine probabilistic predictions of the residual for year i with the actual residuals for year i that are based on the historic demand data, to get a year-long time series of out-turn probabilities.
2) Generate predictions of the distribution of residuals using time (trend) features for year k and weather features for year j.
3) For each time point, use the distribution from step 2 to convert the out-turn probabilities from step 1 into a rescaled residual demand value.
4) Add the rescaled residual demand to the mean demand prediction from the GAM using the weather data from year j, with datetime features of year k.

Each of these demand-scenario, weather-scenario combinations may now be scaled using data collected from FES building blocks to represent the forecasted growth/decline of demand aggregated across each demand technology type.

Efforts have been made to ensure the FES building block data and original demand data used is co-ordinated with other models to avoid 'double counting'. To avoid 'double counting' when applying FES growth in appliance demand, EV and heat pump demand have been excluded from the aggregated FES scaling factor as these have been modelled in separate models.

Mean Expected value form

As a GAM, the features of the point-based model must be well defined and add skill in predicting mean demand. The currently proposed model uses the features outlined in Table 1. This table also includes an 'Indicative Importance' column which categorises each feature by its significance in producing a prediction (High Medium, and Low). This significance level is a reflection the trained GAM's weighting and expert judgement.

Table 1: Terms used in formulation of GAM model

Feature	Term	Reasoning	Estimated importance
Hour	Spline	Define a relationship between demand and the current time in any day; Learn general profile shape	High
Hour, by day of week	Spline	Capture changing profile across days of week	Medium
Hour, by bank holidays & weekends	Spline	Capture changing profile on non-working days	Low
Christmas	Spline	Flag to represent the significantly different demand profile exhibited on Christmas day	Medium
Population weighted UK average temperature	Spline	Behavioural change due to colder days; Increased heating dependence	High
48-hour Average Population weighted UK temperature	Spline	Changing behaviour due to sustained weather conditions; long periods of cold weather increase likelihood to rely on heating	High
Trend (Year)	Spline	Capture the trend in demand growth across years	High
Year, by weekend	Spline	Capture trend with respect to behaviours on weekends over time	Medium
Day of Year	Spline	How demand relates to the day of the year from 1st October to 30th September	High
English Holidays/weekends	Linear	A flag to indicate if the day is a bank holiday or weekend. Helps the model identify relationship between working and non-working days; An indicator variable with 2 possible values	High

Feature	Term	Reasoning	Estimated importance
Summer day of month	Spline	Represent changing trends through summer days; generally shrinking from Spring then increasing as Autumn approaches	Medium
Population weighted mean precipitation	Spline	Capture changing behaviour due to rainfall; Staying Indoors/ dangerous weather	High/Medium
Population weighted mean solar irradiation	Spline	Changing behaviour due to sunshine; Spending time outdoors or reducing length of days	High/Medium
Population weighted UK average 100 m Wind speed	Spline	Capture the effects of wind on behaviour; wind chill or dangerous weather	Medium/High
Hour and Day of Year by holiday/weekend	Tensor	Capture change in profile throughout year with consideration of working and non-working days	Low
Average temperature and 10-meter wind speed	Tensor	Represent effects of temperature and wind; Cold and windy weather will have a more significant impact on demand compared to both individually	Medium
Hour and average temperature	Tensor	Represent effect of temperature throughout the day; Reaction to temperature will change depending on time i.e., temperature at night is less important than day	Medium
Hour and average temperature by weekend/Holiday	Tensor	Like hour and temperature but considers changing relationship during working and non-working days	Medium
Average Solar and hour	Tensor	Represent effect of sunlight levels throughout the day; Reactions to sunlight will change depending on the time	Medium
Covid Stringency Index and hour	Tensor	Represent impact of COVID lockdown procedures throughout the day; Daytime lockdown will be considerably different from non-lockdown whilst nighttime's will be similar	Low

GAM model Performance

Using this form, the model performs consistently across several cross-validation (CV) folds. Its performance is satisfactory for this purpose. Table 2 outlines the model performance in terms of normalised root mean square error (nRMSE), normalised mean absolute error (nMAE), and R^2 score of the model prediction for normalised demand of the entire national demand dataset. As the R^2 suggests, these features have very powerful explanatory power for demand. The relatively low nRMSE and nMAE suggests sufficient deviation from which the residuals may then be predicted and rescaled.

Table 2: GAM Model Performance

Fold	Normalized RMSE	Normalized MAE	R^2
1	4.0283%	3.116%	96.506%
2	4.285%	3.109%	96.538%
3	4.295%	3.124%	96.502%

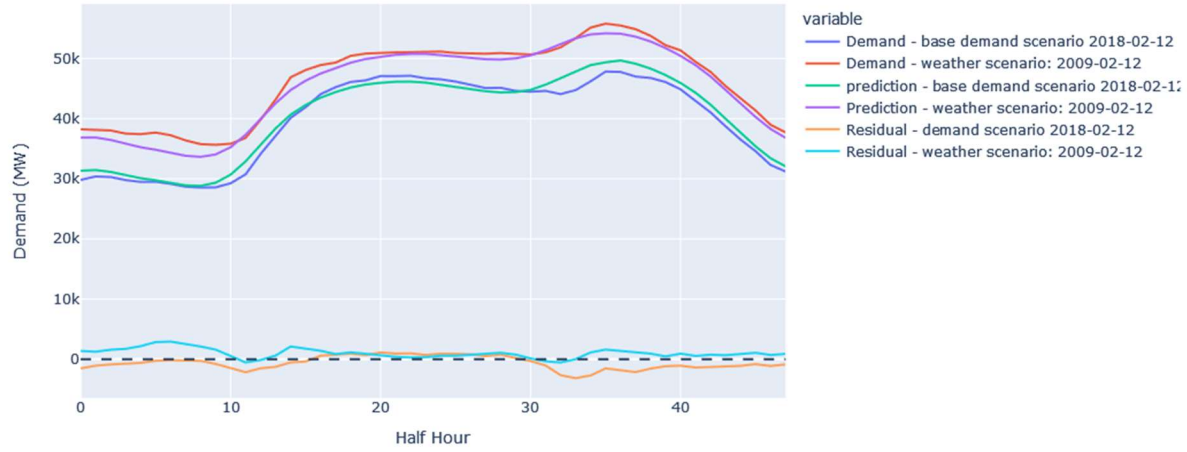


Figure 7: GAM prediction and residuals for an example weather and demand scenario day. The probabilistic model is fitted to the residuals.

Quantile Regression Residual model form

Unlike the GAM based model, a GBM does not require specific formulae for input features to be defined as it performs feature selection and determines relationships automatically by setting tree splits on the most informative features supplied. Table 3 outlines the model features initially provided along with our expectation of their assumed importance for all models. It is important to note that the significance of any feature can change between models trained for different probability levels, therefore any statements regarding any single feature's importance requires consideration. This difference is shown in in Appendix B which compares the ranked significance of features at probability levels 0.02 and 0.6 (Figure 87 & Figure 88 respectively).

Table 3: Residual (quantile regression) model features

Feature	Estimated Importance
Mean population-weighted temperature	Medium
Temperature standard deviation across UK	Low
Mean population weighted solar irradiation	Medium
population weighted solar irradiation standard deviation	Medium
Mean population weighted 100 metre wind speed	Low
100 metre wind speed standard deviation	Low
Mean population weighted 10 metre wind speed	Low
10 metre wind speed standard deviation	Low
Mean population weighted precipitation	Low
population weighted precipitation standard deviation	Low
Hour	High
Day of the year	Medium
Day of Week	Medium
English bank holidays/weekend	Medium
COVID Stringency Index	Low
Christmas	Low
Trend	High

Reasonable values for the lightGBM model hyperparameters were used to prevent overfitting; these are specified in Table 9 of Appendix A.

The overall fitting of the residuals is generally good, demonstrating strong fitting within the body of the distribution. However as previously mentioned, the tails exhibit unusual behaviour due to the GBM models lack of extrapolation capability. Figure 8 is a histogram of the out-turn probabilities from the residual predictions, which should be completely uniform for a perfectly reliable prediction. The disproportionally high probability density near the tails of the predictions indicates that the model is struggling to accurately extrapolate demand at the limits.

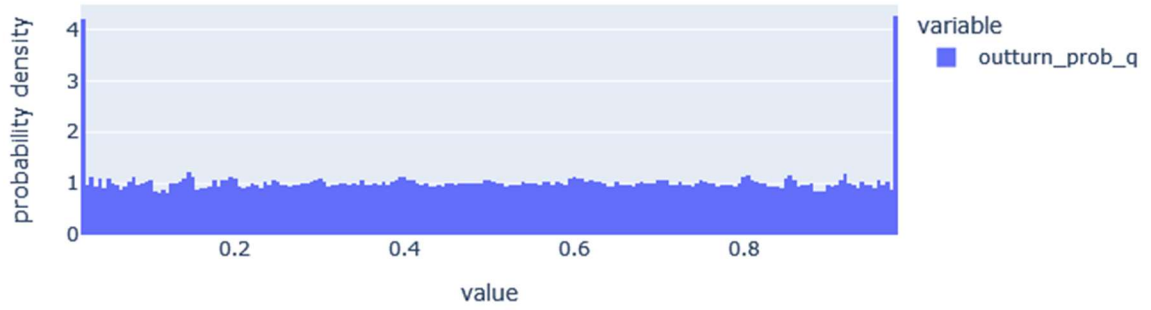


Figure 8: Distribution of predicted quantiles before applying tail modelling.

To mitigate this, we use static GPD distributions to model the tails parametrically and enable extrapolation of predictions into the extremes [3]. By visual inspection of the shape of residuals at the upper and lower tails and resulting calibration of the model, the probability levels of 0.02 and 0.98 were selected as the thresholds for the lower and upper extreme tails of distribution respectively. Figure 9 and Figure 10 compare the residuals of the full training set against the fitted GPD distributions in the lower and upper distributions. Note these plots conclude at the largest empirical residual, but the GPD tail distribution will continue beyond this.

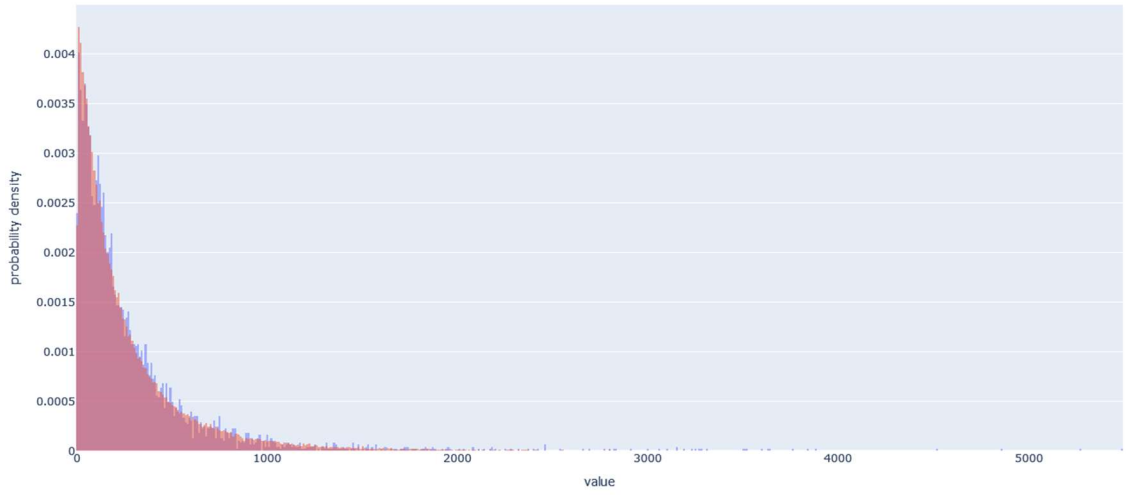


Figure 9: Histogram of lower tail residuals (Blue) against fitted GPD distribution.

[3] Jethro Browell, Matteo Fasiolo : Probabilistic Forecasting of Regional Net-load with Conditional Extremes and Gridded NWP; last revised June 25th 2021

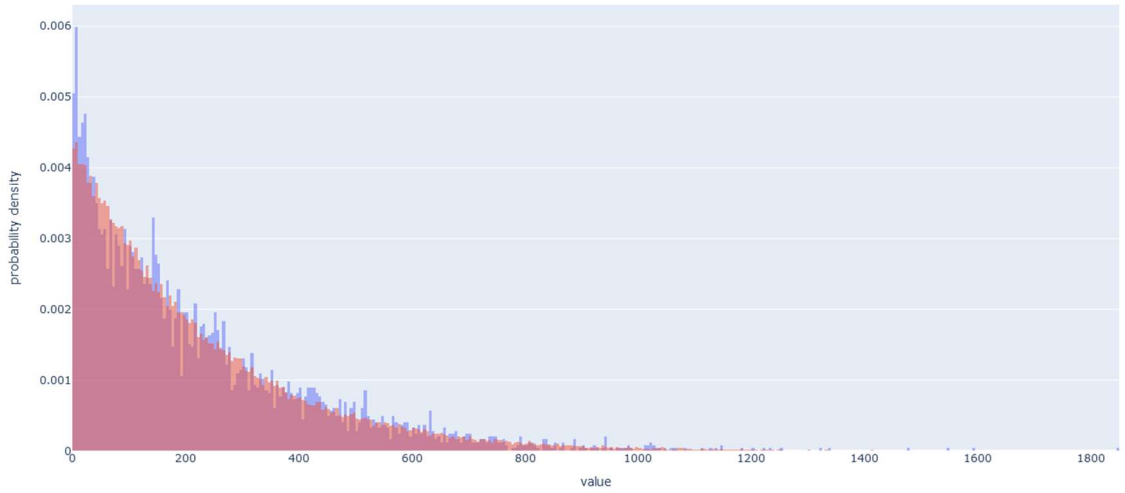


Figure 10: Histogram of upper tail residuals (Blue) against fitted GPD distribution (Orange).

By masking the predicted quantiles past the upper and lower bounds with these distributions (setting the location to predicted quantile at their respective probability level), we achieve a much more uniform distribution of residual quantile predictions. Figure 11 demonstrates the resulting distribution of out-turn probabilities across all quantile models, when the GPD tail models are used.

Figure 11: Distribution of predicted residual quantiles after applying GPD tails.

Quantile regression residual model performance

The model was fitted on all available time points with demand measurements. It is important to check the models produce a good fit; in the case of probabilistic models, we primarily evaluate calibration (reliability), and then sharpness subject to calibration.

A calibrated model has values for each probability level that match their true probability of occurring, so for example we would expect the q5 (5^{th} quantile) prediction to be exceeded 95% of the time across the historic data, the q10 (10^{th} quantile) prediction to be exceeded 90% of the time, and so on. A reliability diagram shows these observed (empirical) exceedances against their nominal probabilities for the whole distribution, where a straight diagonal line represents a perfectly calibrated model.

The model appears to fit well to the data. Figure 12 shows the reliability of the full model across 4 cross-validation (CV) folds. Each cross-fold is a different partition of the training data set, which is fed into the model for training, with one fold held back for out-of-sample predictions. This allows us to examine the 'true' out-of-sample reliability, robustness, and consistency of the model.

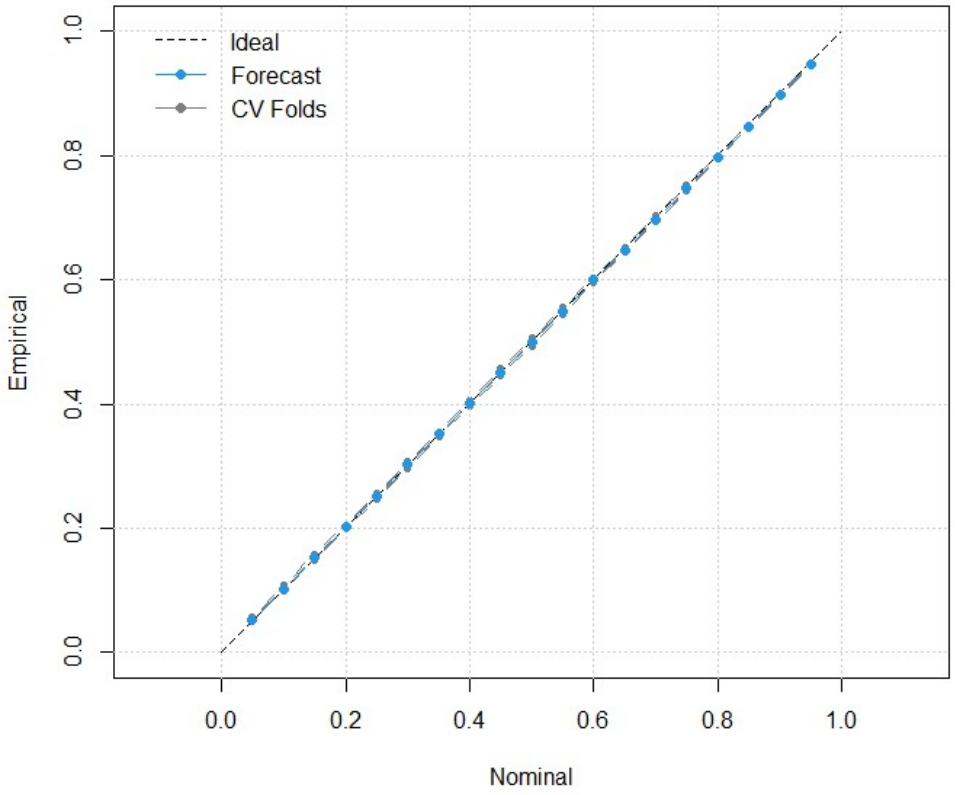


Figure 12: Reliability plot for 4 CV folds of lightGBM Quantile Regression Model of demand residuals

2.2.1.2 BMU wind generation

Load factors for onshore and offshore wind BMUs are modelled independently. Time series of national load factors were generated from Elexon metered generation. To model unconstrained wind, metered values were corrected for Bid-Offer Acceptance (BOA) volumes, also from Elexon. Periods of de-rating when a wind farm first came online were also scaled to the wind BMU's capacity, with capacities taken from Elexon's registered BM units database where available and supplemented with values from the Renewable Energy Planning Database[4]. Load factors were calculated from power values using the instantaneous national capacity (the capacity of the wind BMUs providing data in that settlement period), and the number of wind BMUs with data available at that time was also recorded.

This results in load factors calculated as follows:

$$LF_t = \frac{1}{N_t}\sum_{i=1}^{N_t}(P_{i,t} - BOA_{i,t})/C_i$$

Where N_t is the number of wind BMUs with data for time t, $P_{i,t}$ is the metered power production from wind BMU i at time t, $BOA_{i,t}$ is the corresponding BOA volume and C_i the capacity of that wind BMU.

This resulted in national load factor time series from 2014 to October 2023, as shown in Figure 13.

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

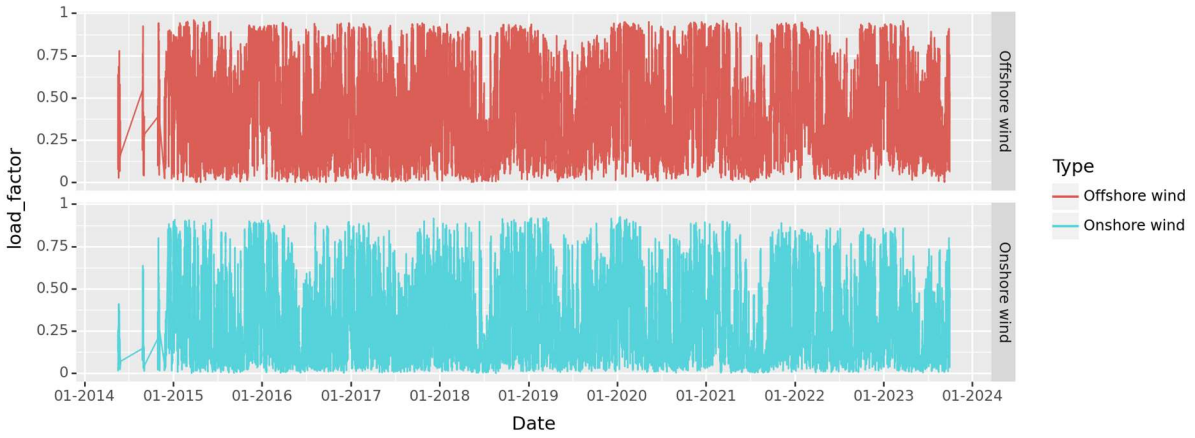


Figure 13: National onshore and offshore load factors from 2015-2023

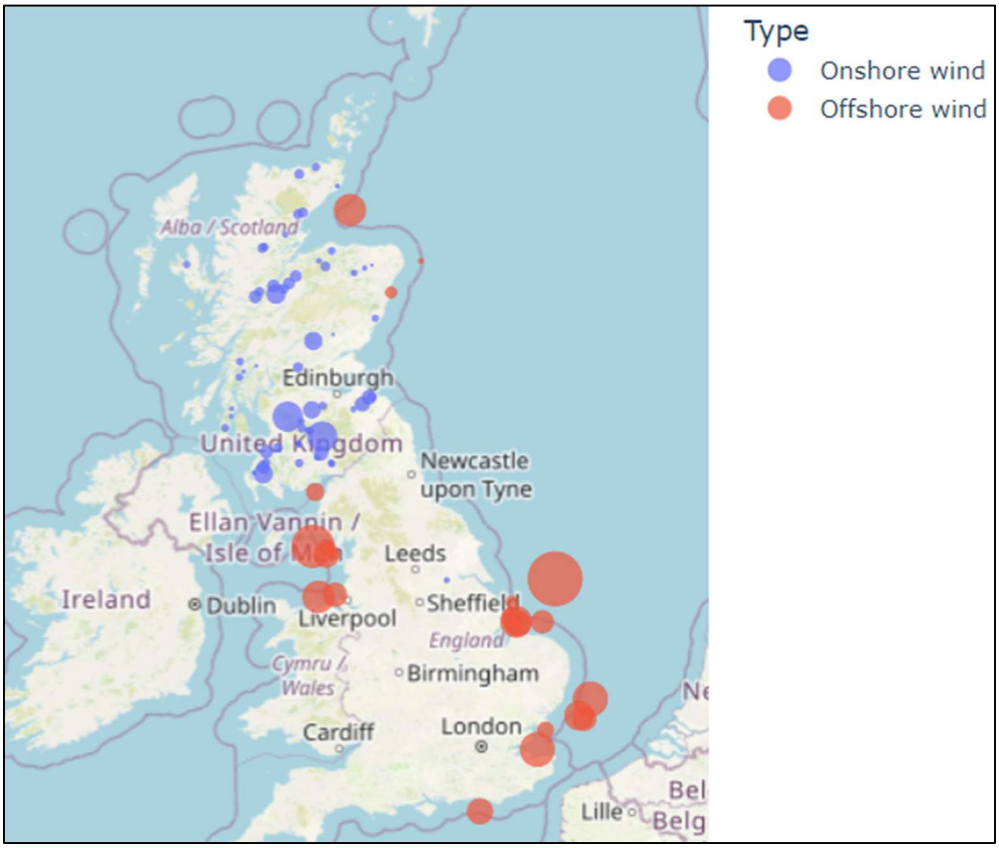


Figure 14: Locations of onshore (blue) and offshore (red) wind farms used in calculation of national load factors. Area of circle represents wind farm capacity.

A Gradient-Boosting Machine (GBM) was used as the model of choice for wind generation. This is due to its ability to capture relationships between features, and the nonparametric form of the models meaning no assumption about the shape of the distribution is needed, while also allowing different relationships between weather variables for each probability level modelled. It will also only fit 'splits' on features that are informative and improve the final prediction, resulting in inbuilt feature selection. The lightGBM implementation of a GBM was used for its speed.

Model features

A linear trend throughout the time series was the only time-based feature used. This feature aims to capture changes in turbine technology and efficiency through time. This feature was aligned with the year naming convention so that, for example, 1st October 2014 would have a trend value of 2015.0 and 2nd October 2014 would have a trend value of 2015.0 + 1/365 = 2014.0027, and so on.

All other model features were based on weather since unconstrained wind generation depends only on the weather and not on human behaviour or other factors. These features were generated from ERA5 reanalysis wind speeds at both 10 and 100m. National-level wind speed features were created by capacity-weighting wind speeds at the individual wind BMU locations. For model training, the capacities of only the sites with measured generation at a given time step were used for this weighting. For generating predictions based on the long historic weather time series, the capacities at wind BMU locations for the most recent time was used, so that the weather features used in the rescaling step represent the most recent spatial configuration and capacities of wind BMUs.

The model features are listed in Just over half of the 20 most important features were for single largest-capacity sites for both the onshore and offshore wind model, with the remaining most important features including the trend and features derived from 10m wind.

Table 4Table 4 with the hyperparameters used to specify the GBM fitting given in Table 9 of Appendix A, and feature importance shown in Figure 83 of Appendix B. Just over half of the 20 most important features were for single largest-capacity sites for both the onshore and offshore wind model, with the remaining most important features including the trend and features derived from 10m wind.

Table 4: BMU wind model features

Weather variable	Description of features created
10m wind speed	Capacity-weighted mean, capacity-weighted standard deviation, min, max of conditions at BMU locations
10m wind direction	Circular mean and standard deviation. Not weighted by capacity.
Wind at largest sites	Wind speed and direction at both 10 and 100m for the 7 largest-capacity wind farms (separately for onshore and offshore)
Wind veer	Site-level difference between 10 and 100m wind direction, averaged across all sites without any weighting applied.
Rolling (smoothed) features	Rolling averages of 10m mean and standard deviation of (capacity-weighted) wind speed. Centred windows of 2.5, 4.5 and 6.5 hours are used.
Lagged and leading features	1,2, and 3-hour lagged and leading 10m (capacity-weighted) mean wind speed
Trend	Linearly increasing in time from 2015.0 (1st October 2014) to 2024.0 (1st October 2023)

Models were fitted for a set of 21 quantiles: from $q5$ to $q95$ in 5% intervals, plus the $q2$ and $q98$ quantiles. The limits of $q0 = 0$ and $q100 = 1$ were also used when interpolating between quantiles

to get the corresponding out-turn probabilities for realisations in the rescaling process. The additional quantiles close to the tails were added to improve the accuracy of this process for values close to the limits.

Model performance

Figure 15 shows the reliability of the wind model. Times with low overall curtailment (red) are very well calibrated, but times of high overall curtailment show an under-forecasting bias where predicted load factors are lower than observed[5] for these times. This is likely due to the inexact nature of the BOA volume correction applied for these times: BOA volumes are based on the FPNs submitted by the wind BMU, which are an estimate of future generation and are not exact – and may be biased towards overestimated production[6]. However, correcting for curtailments in this way is still expected to give a better estimate of generation at that time than not correcting for BOAs or excluding sites under curtailment, and these results show the model is correctly predicting times of normal generation (i.e. low or no curtailment) well, and thus is suitable for predicting uncurtailed generation which is the intended purpose in this modelling.

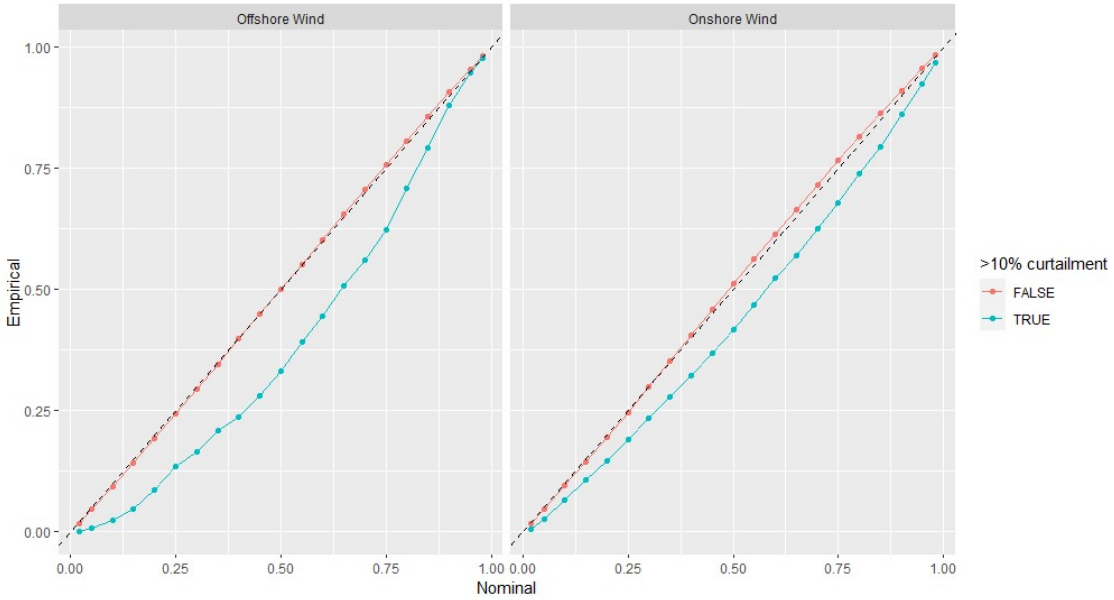


Figure 15: Reliability of wind model. Times of high curtailment show bias.

2.2.1.3 Embedded wind generation

Embedded non-BMU wind was modelled in a similar way to BMU wind, but using different datasets: NGESO estimates for embedded wind generation and capacity[7] have been converted into load factors (plotted in Figure 16), and nationally averaged ERA5 was used as the weather inputs, as the locations of embedded wind are not known (although it is assumed all embedded wind is situated onshore).

[5] Noting that, here, observation means the actual metered output *minus* the BOA volume.
[6] Intini, M and Waterson, M. (2023) Strategic behaviour by wind generators: An empirical investigation, International Journal of Industrial Organization, vol. 89 https://www.sciencedirect.com/science/article/pii/S0167718723000292
[7] NGESO Forecast and Estimated Capacities of Embedded Wind available from: https://www.nationalgrideso.com/data-portal/embedded-wind-and-solar-forecasts/embedded_solar_and_wind_forecast

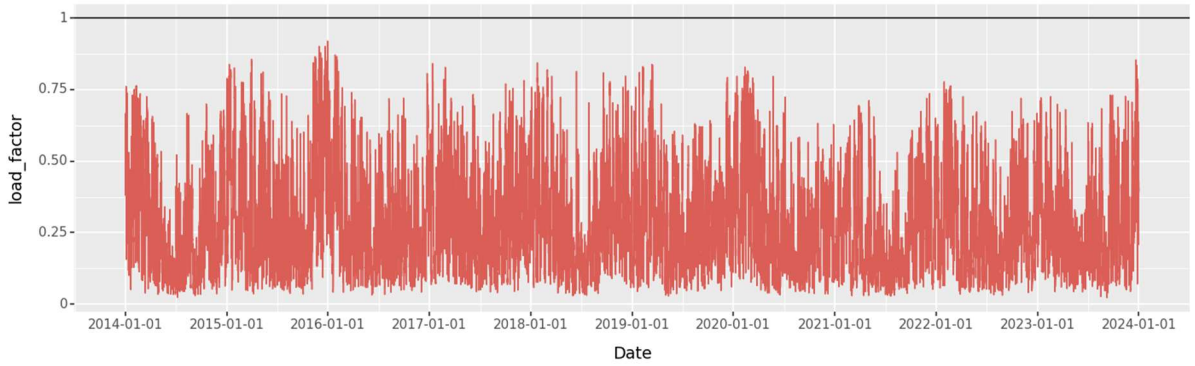


Figure 16: National embedded wind load factors

Model features

A GBM is used as for BMU wind, but with slightly fewer features, as the largest-capacity sites are not known so these features are omitted. The feature importance is shown in Figure 84 of Appendix B, with the trend having the highest feature importance, followed by the overall summary statistics of wind speed and direction for the target hour, with rolling, lagging and leading features of lesser importance but still contributing to the end predictions.

Model performance

Model reliability is very good for embedded wind, as shown in Figure 17. The sharpness of predictions for each of the onshore, offshore and embedded wind are plotted in Figure 18: sharpness quantifies how narrow or broad probabilistic predictions are. Interestingly, predictions are sharpest (most confident) for embedded wind, then onshore wind, and then offshore wind. This could be investigated in future work and may be related to the increasing average load factor of these different components, the differing spatial distributions of generators or size of individual BM units. However, while sharp models are generally desirable, for the purposes of weather rescaling the crucial property is reliability to produce out-turn probabilities that are 'true', and sharpness is a lesser consideration.

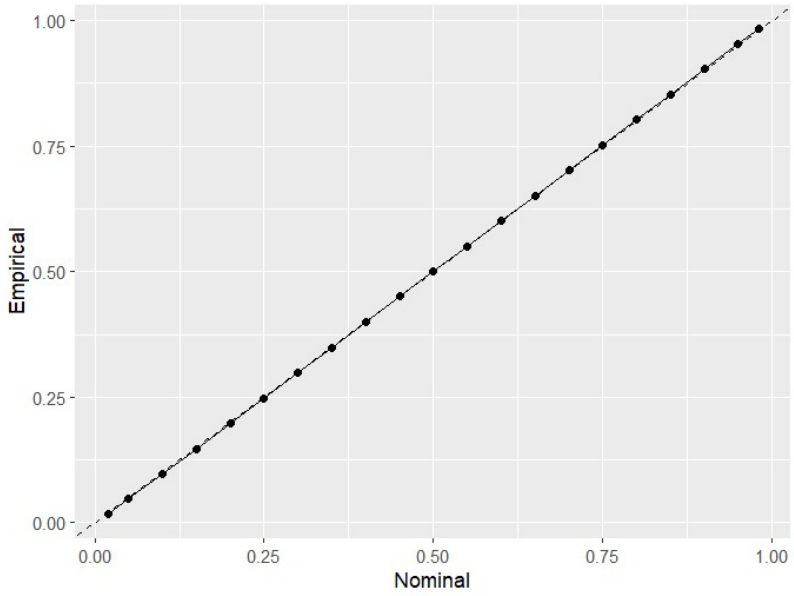


Figure 17: Reliability of embedded wind model.

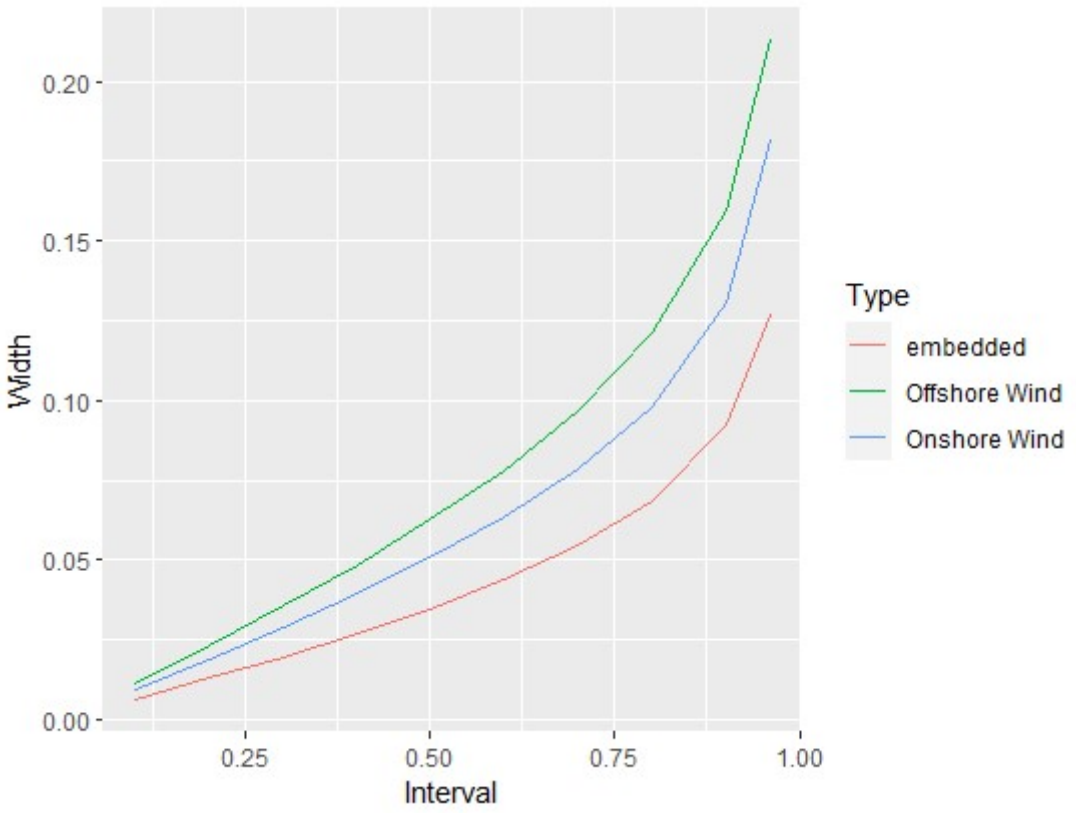


Figure 18: Sharpness of wind models. 'Width' denotes the interval width for the corresponding probability interval, i.e. there is an average difference in predicted load factor of 0.05 for an interval of 0.5 (between q25 and q75) for the onshore wind model. Lower widths indicate a sharper, more confident prediction.

2.2.1.4 Embedded solar generation

Embedded solar generation has been modelled using estimates from the NGESO Data Portal (which is based on the Sheffield Solar modelling). This data provides solar output and installed capacity, which have been used to calculate the load factors plotted in Figure 19. As the locations of embedded solar units are unknown, nationally averaged solar variables from ERA5 have been used as features, along with some calendar variables. These are detailed in Table 5. As with the wind models, a GBM model has been used for solar generation using the lightGBM implementation in Python.

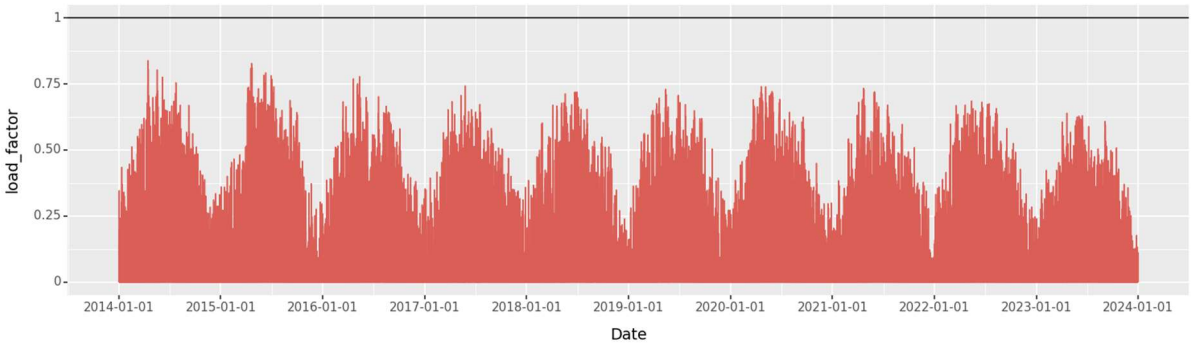


Figure 19: National embedded solar load factors

Model features

Table 5 below details the features used in the embedded solar model. Again, smoothing, lagging, and leading features of the weather variables have been created. In addition to the weather features, several calendar features were also used including the time of day, day of year, and a trend feature which is linearly increasing in time (from 2015.0 (1st October 2014) to 2024.0 (1st October 2023)). A further feature recording whether each time point was within the day or nighttime was also added, based on sunrise and sunset times for Birmingham (a rough central location).

Table 5: Embedded Solar model features

Variable	Description of features created
Surface solar radiation downwards (SSRD)	Mean, standard deviation, minimum and maximum values.
Low Cloud Coverage (LCC)	Mean, standard deviation, minimum and maximum values.
MCC (MCC)	Mean, standard deviation, minimum and maximum values.
High Cloud Coverage (HCC)	Mean, standard deviation, minimum and maximum values.
Rolling (smoothed) features	Rolling averages of mean and standard deviation of SSRD. Centred windows of 2.5, 4.5 and 6.5 hours are used.
Lagged and leading features	1,2, and 3-hour lagged and leading SSRD
Trend	Linearly increasing in time from 2015.0 (1st October 2014) to 2024.0 (1st October 2023)
Time features	Time of day and day of year
Daytime indicator	Whether the time point falls between sunrise and sunset

Figure 85 in Appendix B illustrates the importance of each of these features in the model.

Model performance

One intuitively important feature for solar is whether or not it is daytime. As shown in Figure 17, the reliability of the model for daytime periods is generally good with only slight deviations in some quantiles, whereas reliability of the model for night (is-day = FALSE) is poor. However, as the output is zero overnight this poor reliability is not an issue for the method in practice and will not affect the rescaling process.

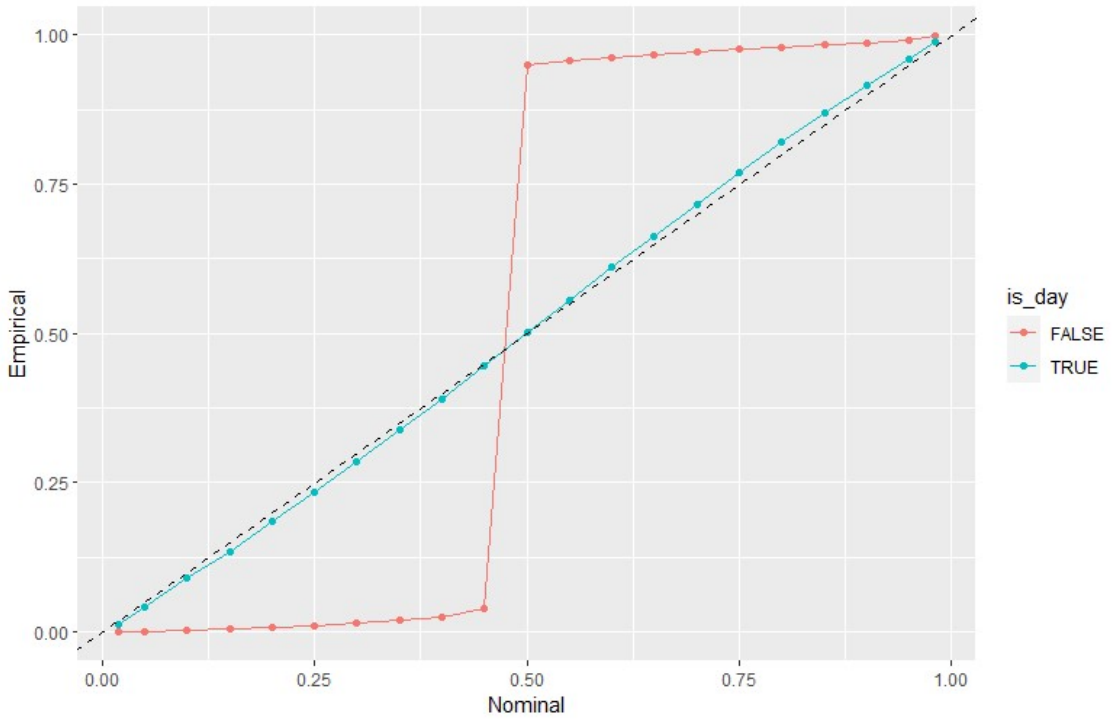


Figure 20: Reliability of solar predictions, split by day and nighttime

The weather variables used in the solar model are nationally averaged as the location of embedded PV assets are unknown. However, a judgement could be made in future as to whether this could be updated to reflect that more PV is installed in the south of GB.

2.2.1.5 BMU hydro generation

BMU hydro load factors were modelled using metered generation data from Elexon corrected for curtailments as the target, in the same way as for BMU wind. Pumped hydro units were excluded from the set of BMUs as they are classed as flexible generation for the purposes of calculating WDFN. The resulting time series of national hydro load factors is plotted in Figure 21.

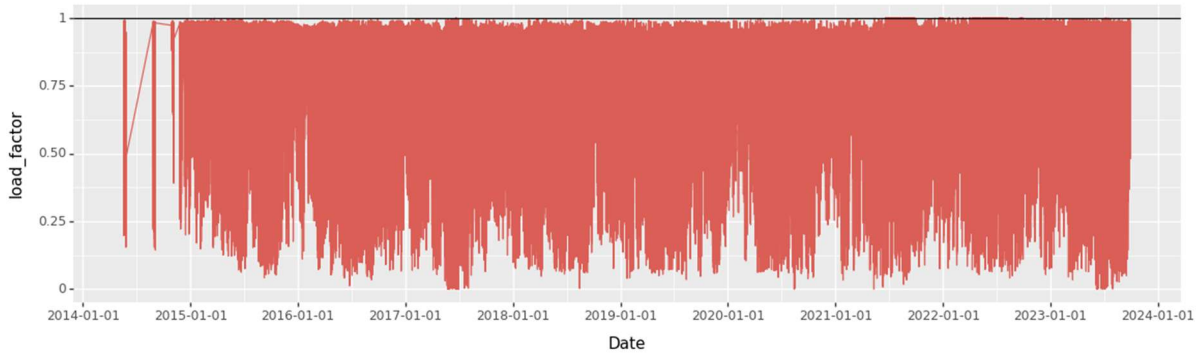


Figure 21: National load factors for non-pumped hydro BMUs from 2015-2023.

Model features

Non-pumped hydro BMUs generally depend on water flow down rivers to generate, so the rainfall across the whole catchment area of the river will influence generation. Therefore, instead of weighting point locations of weather as for wind BMUs, average weather across a box encompassing the catchment areas of all rivers with hydro BMUs on was used as model inputs. For this model, the only

weather variables used were total precipitation and 2m temperature, although rolling averages were used to engineer smoothed features of precipitation, to capture the time lag between rainfall and water flowing into and down rivers. Time features of time of day and day of year were also included to capture any diurnal and seasonal patterns. The full set of features is listed in

Table 6.

Table 6: Description of BMU hydro model features

Weather variable	Description of features created
Total precipitation	Mean, standard deviation, min, max across area
2m temperature	Mean, standard deviation, min, max across area
Rolling (smoothed) features	Rolling mean and rolling standard deviation of all total precipitation features. Windows of the previous 1, 2, 7, and 14 days are used.
Lagged features	1,2, and 3-day lagged values of the previous 24 hours' rolling mean precipitation
Diurnal and seasonal features	Day-of-year and time-of-day
Trend	Linearly increasing in time from 2015.0 (1st October 2014) to 2024.0 (1st October 2023)

Future work could refine this model by using shapefiles for river catchments to capture the exact relevant area for weather data averaging, and use of further weather variables such as soil saturation and surface runoff that are also linked to river flow. A GBM was again chosen for modelling BMU hydro, so no assumptions about the shape of the distribution or the exact form of the relationship between weather and load factors were made. The same set of quantiles as for BMU wind were modelled, using the same set of GBM hyperparameters detailed in Table 9 of Appendix A. The average feature importance across all quantiles is plotted in Figure 86 of Appendix B, showing the largest importance placed on time features, and the smoother features with long (7- and 14-day) rolling windows.

Model performance

As shown in Figure 22, the BMU hydro model is generally reliable, with only slight deviations from perfect reliability for some quantiles. However, the model predicts a much broader (more uncertain) distribution for hydro load factors than for any of the wind models – Figure 23 shows the central 50% interval spans a difference in load factor of 0.19 on average for hydro, whereas the equivalent value for the broadest wind model is 0.06. Further refinement of the model features and weather processing may help to improve this; however, the much smaller value of hydro capacity compared to wind means the confidence of predictions in this component is perhaps less important to the final WDFN calculation, and the requirement for reliability is already satisfied.

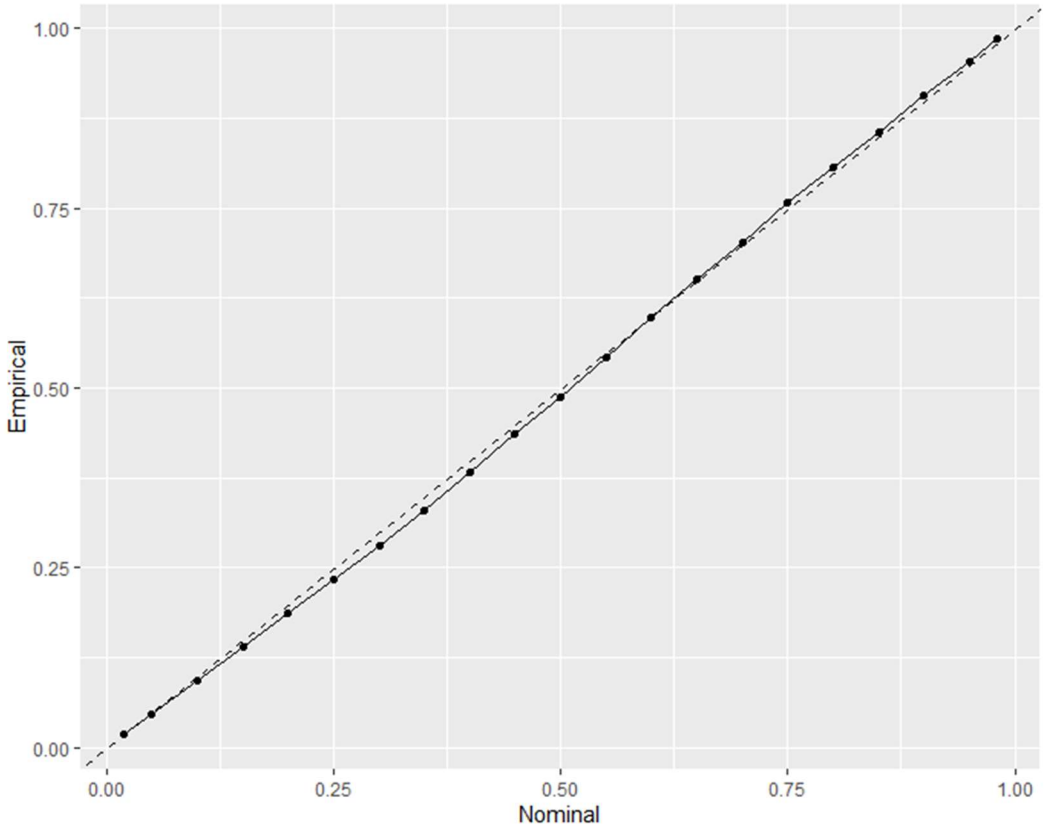


Figure 22: Reliability of BMU hydro model

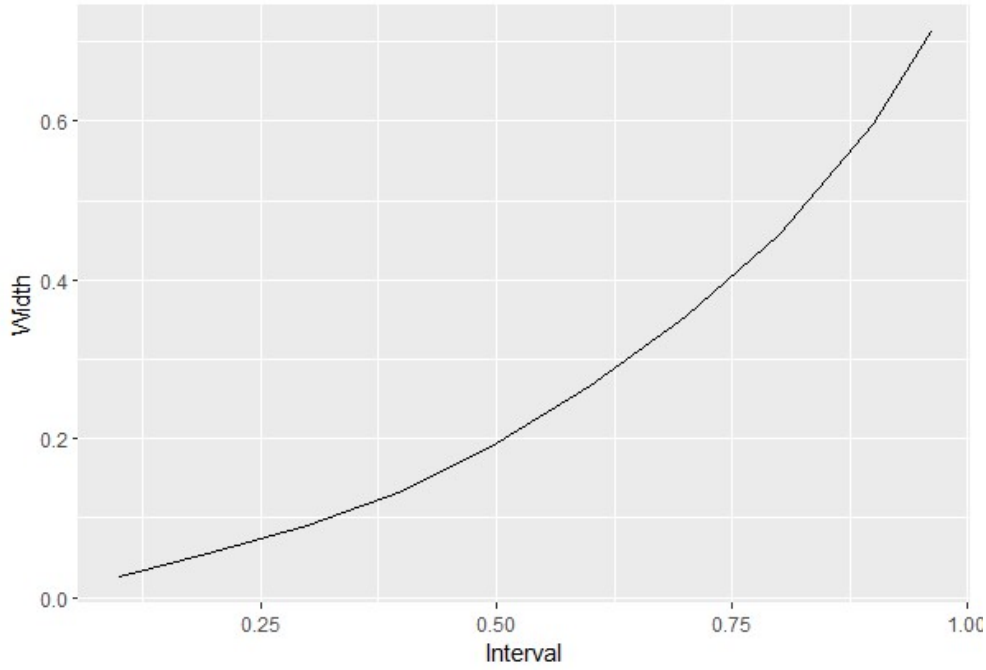


Figure 23: Sharpness of BMU hydro model

2.2.2 Weather rescaling of low-carbon technology profiles

Long time series which describe aggregate, mass-market usage of various low-carbon technologies are not yet available, particularly for electrified heat and transport demand which are widely anticipated to pose significant challenges for future electricity system operation and are central to this project.

Instead, the approach taken in this project has been to rely on data that has been collected from trials or from early adopters, assuming that this data represents inflexible demand since there have been limited opportunities for flexible demand historically – and that this data is a reasonable proxy for future inflexible demands when flexibility becomes more widespread. While we expect significant portions of future heat pump and electrified transport demands to be flexible, the differences in profiles for flexible vs inflexible behaviours are not confidently known and estimates of how much the inflexible component will shrink are rough at this point in time. Simply neglecting a proportion of these demands as flexible would assume that proportion is *completely* flexible with no restraints on the shape of profile across the day, which we believe to be unrealistic. Because of this, we have modelled all heat pump and EV demands as inflexible in this work. This likely provides an overestimate of the WDFN need and could also reduce the peaks and change the shapes of daily profiles. The effect of this assumption will also vary between scenarios, since the FES each include different levels of flexible behaviour.

Projecting the rescaled profiles to future FES years by a multiplicative scaling assumes 100% of all Low Carbon Technologies (LCTs) are behaving inflexibly, for the purposes of calculating WDFN. While this could include some behaviour linked to economy 7 or scheduled EV charging within a time window, it is assumed this is not behaviour that would change in response to WDFN and is likely a small proportion of the overall profile. Any other treatment of these technologies would require separate profiles for both flexible and inflexible behaviours, and knowledge of the relative proportions of flexible vs inflexible units.

For electrified heat demand we have used profiles from individual heat pumps, collected during a recent innovation trial. For electrified transport demand, the EV charging behaviour profiles already produced by NGESO within a previous innovation project have been used as a starting point.

Because these profiles are much shorter than a set of multi-year historic time series – typically spanning only a year – and consist of a subset of users rather than measured national demands, the fitting of a full probabilistic model and application of the weather rescaling method isn't appropriate. Instead, simpler deterministic models that predict the mean heat pump and EV demands for a given set of weather conditions are used.

2.2.2.1 Electrified heat demand

Data from the recent Electrification of Heat Demonstration Project has been used to characterise electrified heat demand from heat pumps. A clean version of this data is published on the UK Data Service[8]. This dataset provides very granular electric heat demand measurements for 740 properties, of which 353 have high enough quality to be used in the analysis (this is the subset identified within that project for use within their own analysis). The data covers the period from 2020 to 2022, although only the final year of data is retained as the earlier records have considerably more missing data.

A significant shortcoming of this dataset is its short duration. Heating demand is very strongly affected by weather conditions, and weather conditions can change considerably from one year to the next. Electric heat demand measurements from heat pumps collected over multiple years would almost certainly exhibit more variation than those collected over a single year and demands in a cold year would almost certainly be higher than those in a warmer year.

To overcome this limitation, we adopt similar weather-based rescaling methods to those set out in the previous subsections. We assume that the electric heat demand varies with time-of-day, day-of-year, and day-of-week, but also that it depends on both the current temperature as well as the temperature over the previous day. These assumptions are used to fit a simple Generalised Additive model, which can then isolate the impacts of temperature on electric heat demand. More weather

[8] https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=9050#!/access-data

variables such as wind speed, and more complex interactions between these weather variables, could be added to the model but increased model complexity would require a larger historic dataset.

The basic principles of this approach are illustrated in the figures below. Figure 24 shows the average electric heat demand from the heat pump trial data as it varies through the days within the year. The tendency for higher demands in colder months and lower demands in warmer months is clear, although there is an unusual jump in demand in June 2022. This anomalous time step is due to a large proportion of meters failing simultaneously and should be excluded from the training data.

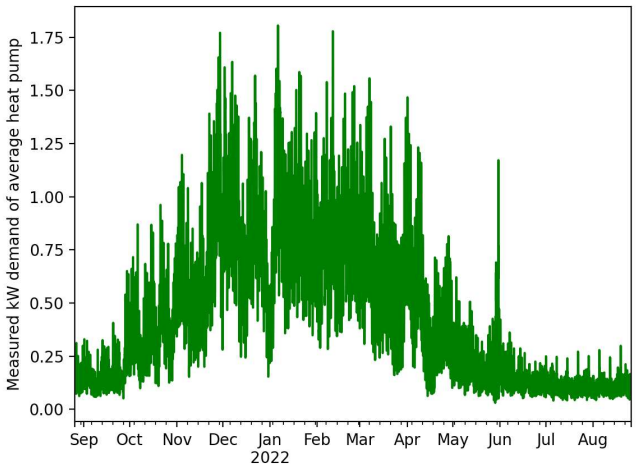


Figure 24: Average electric heat demand profile from demonstration project data

Figure 25 then shows the outputs of applying the rescaling method using the interim data[9], with daily variation in heat pump demand over 40 years of temperature data alongside the 2-metre temperature. The method results in higher electric heat demands during the winters which have the coldest temperatures within this dataset (e.g., 1987, 1982, 1995, 2010 etc).

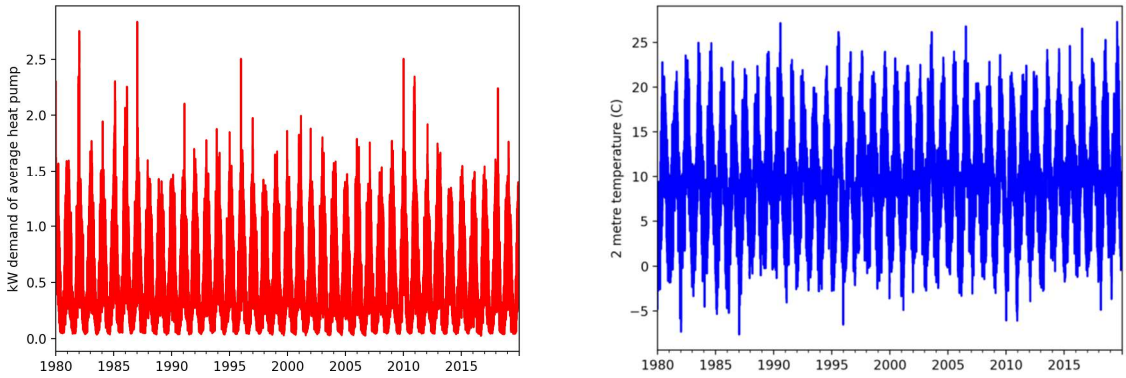


Figure 25: Average electric heat demand profile after rescaling to reanalysis weather data

[9] Bloomfield, H., Brayshaw, D. and Charlton-Perez, A (2020) ERA5 derived time series of European country-aggregate electricity demand, wind power generation and solar power generation: hourly data from 1979-2019. https://doi.org/10.17864/1947.272

2.2.2.2 Electrified transport demand

Electric vehicle (EV) demand is modelled and rescaled using the EV charging behaviour profiles produced by Element Energy for National Grid ESO through a Network Innovation Allowance project[10]. The output of this project includes profiles for 3 types of EV chargers (residential, work and slow/fast public[11]) which have been detrended for EV usage growth which was observed in the study period, and weather corrected for observed changes in behaviour due to average daily temperature. Each profile spans for a single year and are normalized by the estimated total annual energy consumption used by each charger type. At the time of conducting this work, this dataset is the largest and most comprehensive publicly available summary/estimation of GB EV charger usage, making it the most valuable available data source and is assumed to be a proxy for future inflexible demands, meaning any changes in the profile of inflexible EV demands (perhaps due to changing technologies, charging facilities and more widespread uptake of EVs than the early adopters represented in this data) are not reflected in the modelling or results.

A national EV charging profile can be obtained by multiplying each of these profiles by the contribution from the respective charger type to the total energy consumption of all EVs in any given year and summing them together. The projected total energy consumption has been extracted from FES for each year. Due to lack of credible information, the percentage at which each charger contributes to the total projected GB EV energy usage will remain static from the research already conducted and reported by Element Energy in the same NIA project.

In the last step before producing the complete profiles, Element Energy calculated and used a linear weather scaling factor of each day's mean energy demand by the average daily temperature (using data from weather station ENGX0, located near Nottingham) on the dates the raw charging data was originally collected (2017). Element Energy did not include this scaling factor in their outputs, but it is possible recalculate it by performing this step backwards to obtain the underlying non-weather corrected profile as well as the factors originally used to scale each day. This makes it possible to rescale the profiles to the average temperature of the same date in another year to produce new and equally plausible profiles. This allows us to rescale the profile to any of the 43 available weather years provided in ERA5.

We can calculate this underlying profile and factor by extracting the profiles that form 'day archetypes' constructed by Element Energy as part of their modelling process. A 'day archetype' is a representation of the average profile for any day of the week in each month i.e., every Tuesday of January. 'Day archetypes' also include outliers and bank holidays (i.e., Christmas and Easter Day), which consist of a single unique profile. Figure 26 demonstrates the 4 available profiles for the February-Saturday day archetype profile.

[10] Development of GB electric vehicle charging profiles NIA, June 2019
https://smarter.energynetworks.org/projects/nia_ngso0021

[11] Rapid public charging data is not available for public or third party use due to the small sample size which poses risk of breaking anonymity. However, if this dataset increases in future, it could be readily incorporated into WDF calculation process.

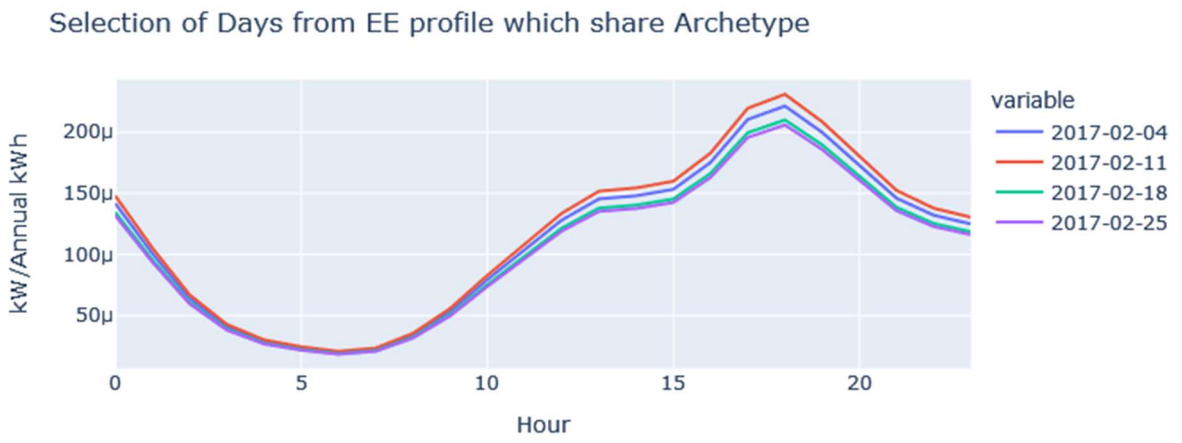


Figure 26: Example days from Saturday-February residential archetype from EE EV charging profiles.

The exact specification of these categories has not been provided, but they can be calculated by removing the original factor. However, as each profile is only scaled, we may normalise every day then identify these archetypes by matching duplicate profiles. Figure 27 show the base underlying February-Saturday archetype profile from which the previous 4 profiles were scaled from.

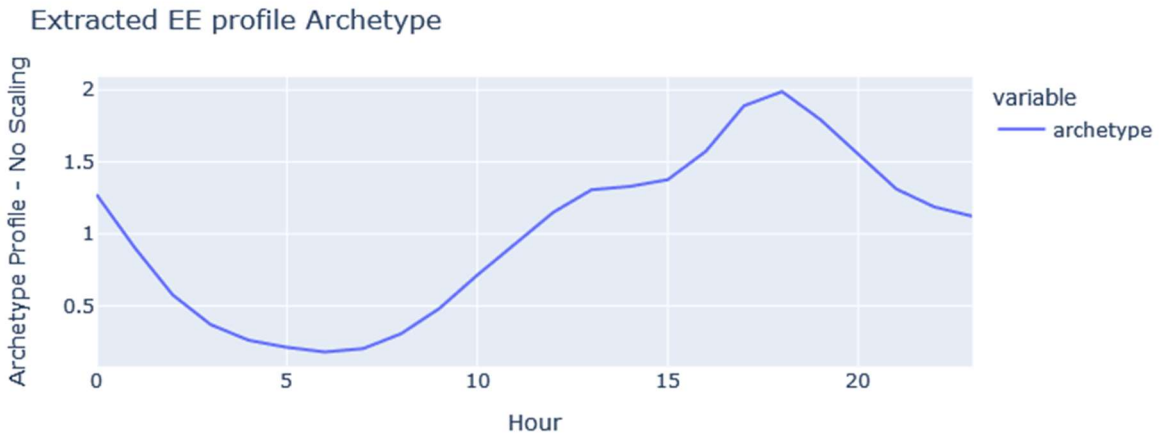


Figure 27: Profile of Saturday-February Residential profile with temperature scaling removed

The scaling factor formulation used for each 'day archetype' can be determined by mapping the mean energy for each day profile to the average temperature of that data. As Element Energy has asserted the relationship is linear (with lower temperatures resulting in increased usage) we can fit a line with intercept (the base demand scale at 0oC) and gradient (the temperature scaling factor) to calculate the mean energy given an average temperature. As this requires at least 2 samples, the holiday and outlier profiles scaling factor is impossible to calculate. In these rare cases, the scaling factor is taken as a constant, asserting that weather has minimal effect on the expected behaviour. An example of this scaling factor for Residential February-Saturday profile is shown in Figure 28, and an example of that profile being rescaled to weather in 2019 is shown in Figure 29.

Whilst Rapid Public charger use profiles are not available in this profile, there are some records available through OLEV's data portal and could be used to create a rough model. As their total contribution is expected to be very small, modelling this data is not required to be particularly accurate. A potential solution could be fitting a GAM model to any given data to capture its shape and relationship to weather and trend in scale/variance.

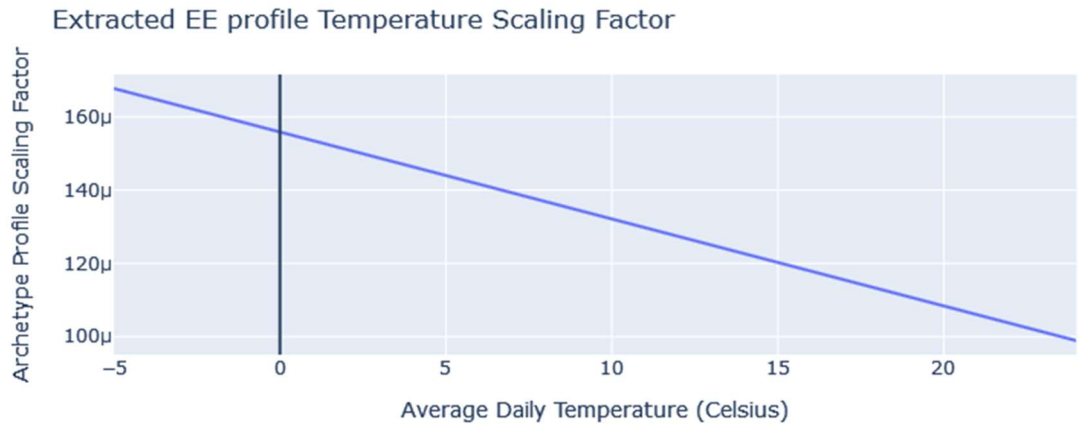


Figure 28: Temperature Scaling Factor of Monday-February Residential Archetype profile used by Element Energy

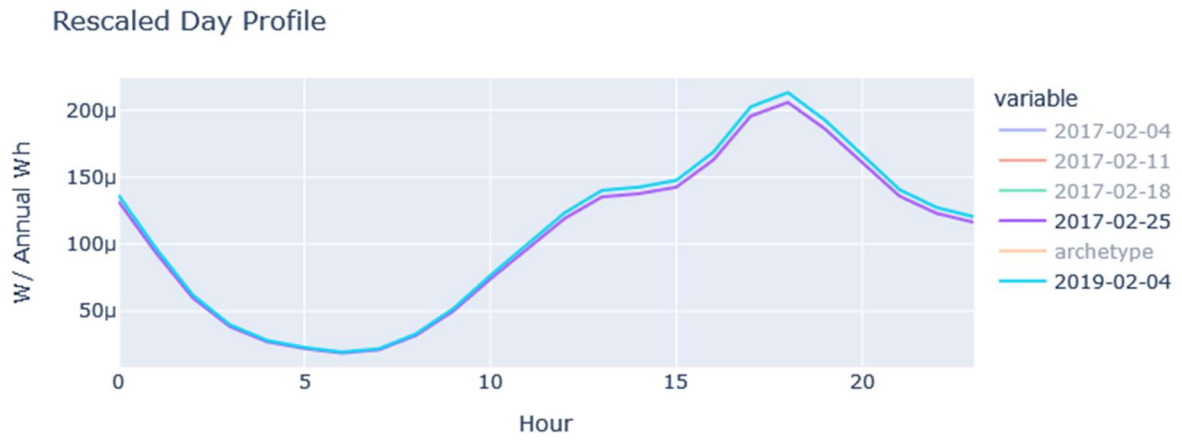


Figure 29: EV demand profile on February 4th, rescaled to average temperature in 2019.

With this model, we can produce 43 plausible annual EV profiles based on weather scenarios which are scalable using building blocks/ summaries of FES.

2.3 Calculations

The weather rescaling process results in time series of plausible 'current' demand and generation, representing the most recent year in the available historic data, in this case 2023.

2.3.1 FES scaling

To assess future WDF needs, all elements of inflexible demand (heat, transport, and other demand) and inflexible generation (wind, solar and hydro) must also be scaled to projected future capacities or technology penetrations. This is a simple multiplicative scaling to produce time series for scenario s (where s is different for each FES pathway and future year), $x^{(s)}$, from values x_k previously representing year k. For a scaling value $A^{(s)}$ in scenario year s and the equivalent value A_k for year k,

$$x^{(s)} = x_k \times \frac{A^{(s)}}{A_k}$$

Peak demand was used for both A scaling factors in the demand modelling, while installed capacities were used to scale load factors for each generation type. Demand could plausibly be scaled on both

peak and average demand, controlling both the spread and level of the time series. However, this can result in nonsensical negative values for large spread which would require extra treatment, and we understand from workshop discussions that this simple multiplicative scaling is in line with NGESO's modelling and insights team.

Plotting the components of demand and generation for an example of a typical day shows differences between scenarios: while the general shape of each component is the same, the absolute sizes vary based on projected technology uptakes in the given scenarios. For example, LCTs make up a much higher proportion of demand in the most ambitious Leading the Way scenario, and the evening peak is exacerbated by EV charging to a much greater extent in this scenario compared to Falling Short. This effect is caused by the assumption that all LCT demands are inflexible: in fact, Leading the Way is predicted to also involve the highest levels of smart charging and flexibility, which would be expected to 'cancel out' some of the EV and heat demand portions shown here.

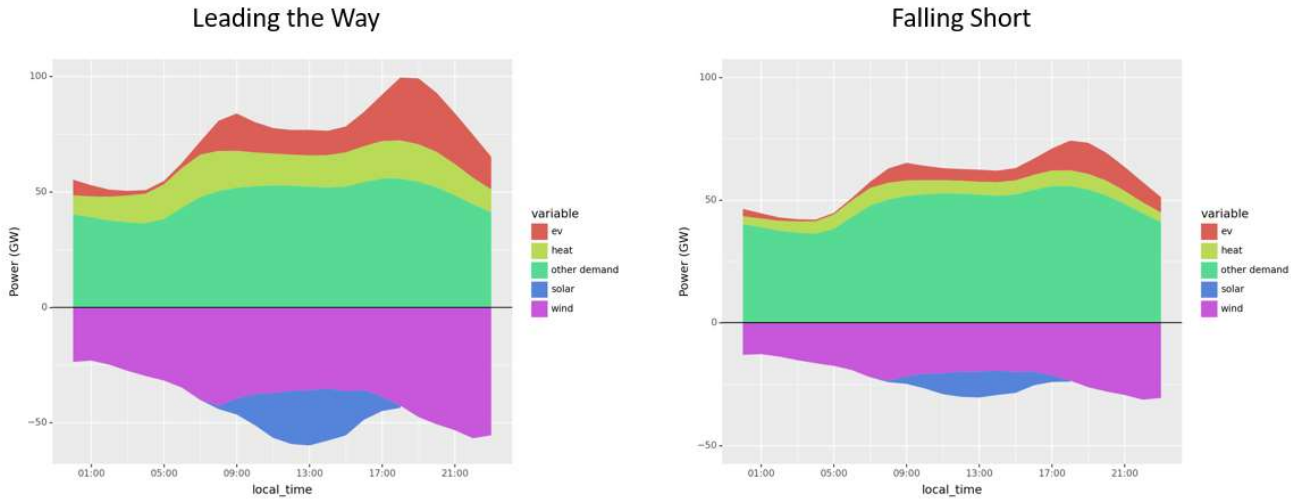


Figure 30: Comparison of demand and generation elements for a typical day under different FES

The final step to calculate WDF need for each settlement period in the time series is the summing of all demands to a total demand time series, and likewise the summing of all generation elements.

2.3.2 Calculating within-day flexibility

Management of the overall daily imbalance of demand and generation is considered a separate problem, so for the calculation of WDFN it is assumed that other actions have been taken to balance total daily energy. To do this, we find the mean difference between demand D and generation P across all settlement periods within the day and apply this as a constant shift in power to each settlement period (labelled as the day imbalance on Figure 31):

$$I_d = \frac{1}{n}\sum_{t=1}^{n} D_{d,t} - P_{d,t}$$

It is important to note that the value of this day imbalance could be positive or negative depending on whether there is a net excess of demand or generation across the day.

We have adopted the convention of positive demand, so that an excess of demand (or a lack of generation) results in a positive (upwards) value of WDFN:

$$WDFN_{d,t} = D_{d,t} - P_{d,t} - I_d$$

Days are treated independently in the calculation of daily metrics, with a cutoff at midnight.

We also align on the weather years and out-turn probability years used to generate that data. Including these in the core calculation of WDFN gives:

$$WDFN^{(s)}_{d,t,\{i\_j\_k\}} = D^{(s)}_{d,t,\{i\_j\_k\}} - P^{(s)}_{d,t,\{i\_j\_k\}} - I^{(s)}_{d,\{i\_j\_k\}}$$

Where d, t define the time of day and day of year, s defines the FES scaled to and $\{i\_j\_k\}$ defines:

- i, the 'year from': the year of observations from which out-turn probabilities are derived
- j, the 'weather year': the year of historic weather used in the rescaling
- k, the year scaled to: time-based (trend) features from this year are used in the rescaling.

In every case, k is set to the most recent year in the historic data (2023).

The daily imbalance and resulting WDFN for the typical Leading The Way day from Figure 30 are shown below.

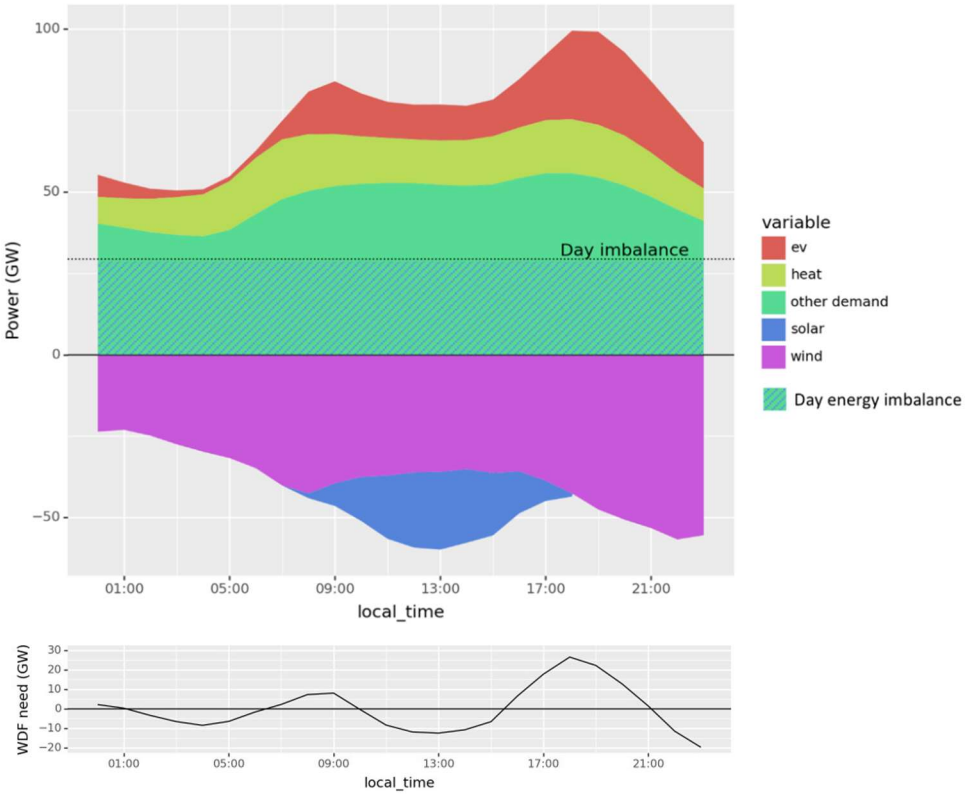


Figure 31: Example day showing stacked demand and generation, and the calculated daily imbalance

2.3.3 Extreme value theory

One of the questions this work aims to answer is what are the rare, but most extreme WDF needs? This is important for assessing the potential impact of these types of events, ensuring there is enough flexibility to meet the needs, and planning for and managing unusual events that may need different approaches. Even with the use of multiple weather and observation years to maximise the length of data, the most extreme events will still be rare. Taking empirical quantiles for the extremes from the historic data would lead to a reliance on a few historic events, which are not guaranteed to represent the true values for those extremes. Extreme value theory (EVT) offers a robust way to model the tails of a distribution, by fitting a parametric distribution for the tails to the available data. This distribution is called the Generalised Pareto distribution, or GPD. The 'peak over threshold' variation of EVT is used, underpinned by limit theorems for threshold exceedance. This uses a separate distribution for

the tails beyond a specified probability level and allows extrapolation beyond previous experience and description of the dependency structure of extremes.

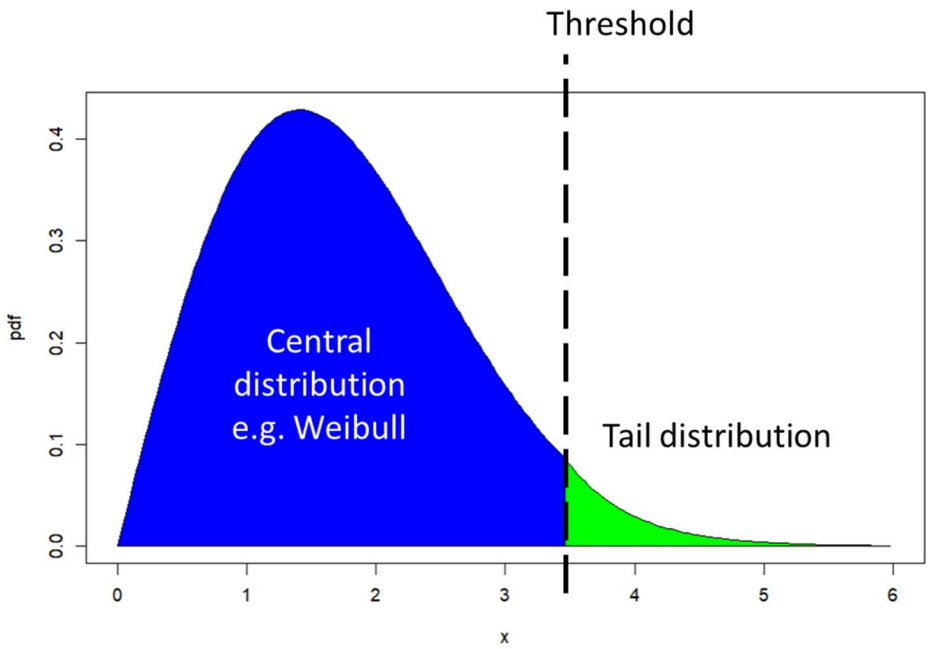


Figure 32: Example of a central distribution below a certain probability threshold, and a tail distribution beyond.

A separate tail distribution is fitted for each FES year and scenario, fitting the GPD distribution to all the historic data that exceeds the specified threshold. We have chosen a threshold of 95% for the upper tail (or 5% for the lower tail), as we believe this gives a reasonable trade-off between the amount of data available above the threshold for GPD parameter fitting, and ensuring the fitted parameters still give an unbiased estimate of the most extreme values. The GPD distribution has 3 parameters: location (a), scale (b) and shape (s). The location parameter is defined by the probability threshold, and then parameters b and s are optimised by fitting to the historic data above this threshold. We define $z = (x - a)/b$, so that the CDF of the fitted tail is given by

$$G(z) = \begin{cases} 1 - [1 + sz]^{-1/s} & \text{for } s \neq 0 \\ 1 - \exp(-z) & \text{for } s = 0 \end{cases}$$

The type of tail can be categorized into three descriptions: if $s = 0$, this is a 'light tail'; if $s > 0$, we have 'heavy tails' that have no upper limit, and if $s < 0$, the tails are bounded. Examples of tails with varying scale and shape are plotted in Figure 33. Our implementation of EVT results in a function where you supply the set of historic values, whether this is a time series of WDFN or of a numeric daily metric, and the probability level you are interested in. If the probability level specified is within the central 5%-95%, the empirical quantile from the historic data will be returned, but if the probability level falls within either of the tails, a GPD will be fitted and the corresponding value for the supplied probability returned from this tail distribution.

In principle, it would be possible to fit different extreme value distributions to each weather / historic year combination, rather than a single distribution to all the pooled values. However, this would mean fitting models which each only have at most 365 (or 366 for leap years) observations, which may not be very robust. However, it would be useful to see how extreme probabilities may vary from one year to the next. Methods for partially pooling might be able to overcome this shortcoming – these are discussed in Section 5.

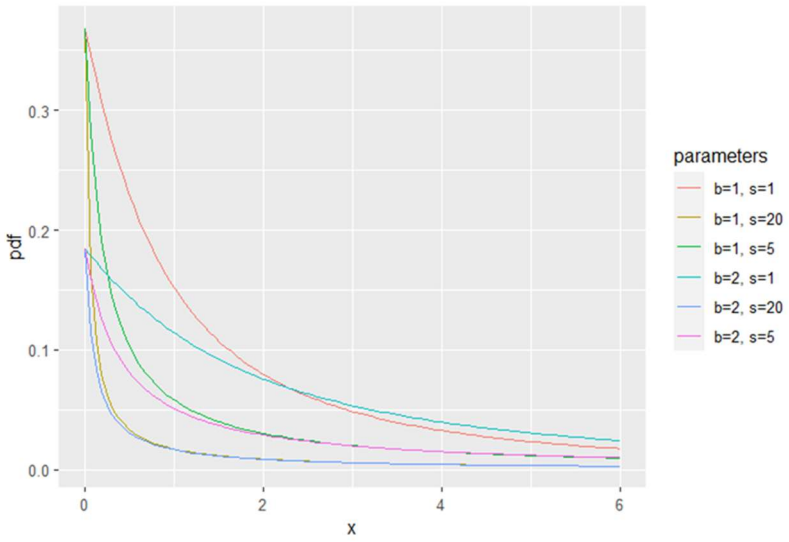


Figure 33: Example GPD tail distributions with varying scale and shape

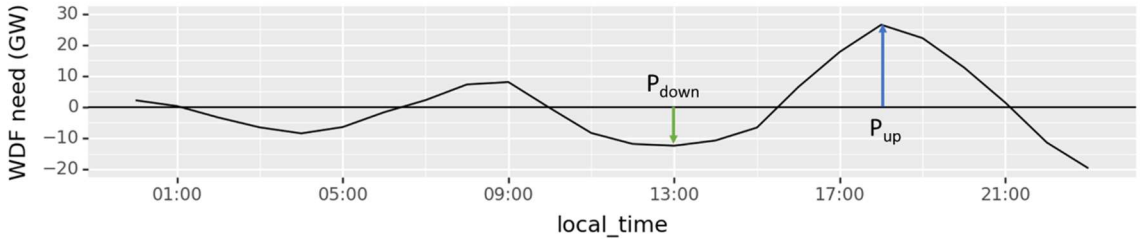
2.4 Output

Several output metrics are produced to describe the WDF need. The definition and calculation of these output metrics is set out in this section.

2.4.1 Metric definitions

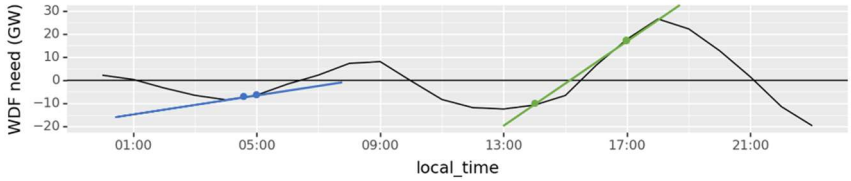
A set of metrics is produced, covering a range of quantities. The set of metrics implemented so far includes:

- Peak upwards (blue arrow) and downwards (green arrow) power requirement.

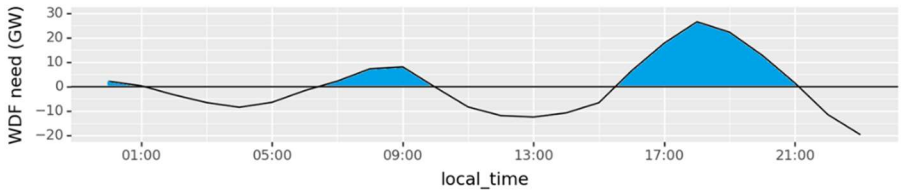


- Based on this, we also calculate:
  - Distribution of peak upwards and downwards power, which can be broken down by season.
  - Proportion of the time the absolute peak need is in the upwards direction, across all days. The above example would count towards the number of days where the peak is in the upwards direction.
  - Difference in magnitude of peak upwards and downwards need, $\left| P_{up} - |P_{down}| \right|$, to give information about the relationship between peak upwards and downwards need.
- Ramp rates:
  - Half-hourly, 2-hourly, and 4-hourly upwards and downwards peak ramps over the day. The half-hourly (blue) and 4-hourly (green) ramps are the gradient of the corresponding lines, associated with the time marked as x. This is calculated for

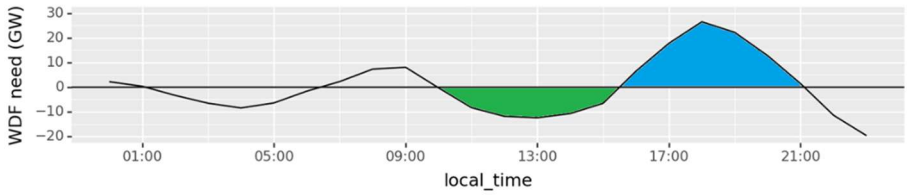
each set of points within the day, and the peak upwards (maximum) and peak downwards (minimum) values reported. This does mean the set of 4-hour ramp rates within the day is smaller than the set of half-hour ramp rates, but this is a consequence of treating days independently with different day energy imbalances so that ramps can't be calculated across the day cutoff time.



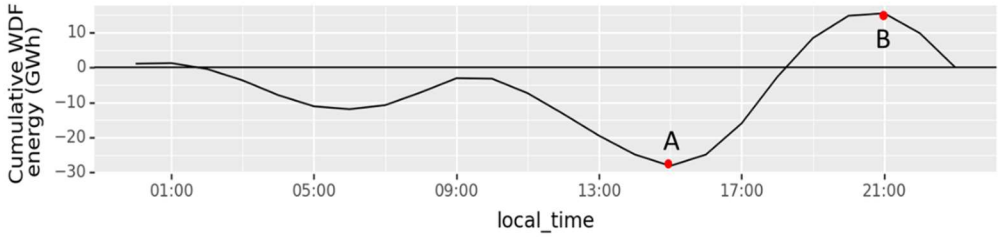
- Energy: These metrics are useful for answering questions about the quantity of storage-based solutions that would be needed to meet the WDFN, for example what proportion of days could WDFN be met by EVs?
    - Total area above zero: Total energy needed across the day. Total downwards energy will also be equal to this as we have defined the net daily WDF need to be zero.



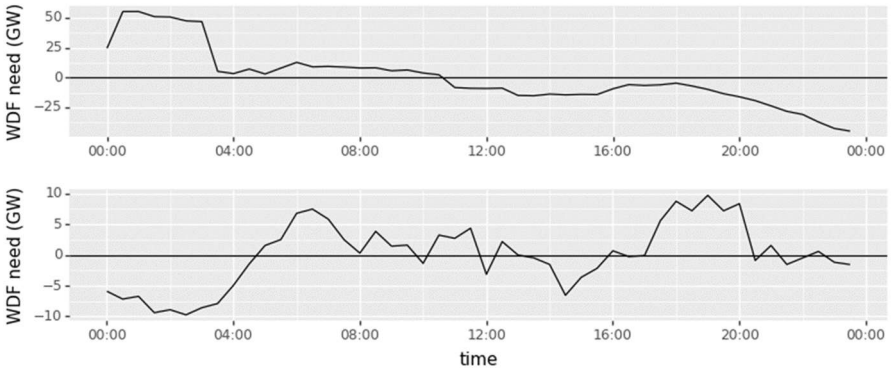
    - Cycle capacity above (below) zero: The capacity needed to provide energy for the single largest upwards (downwards) block of energy in the day. Cycle capacity above zero is shown in blue.



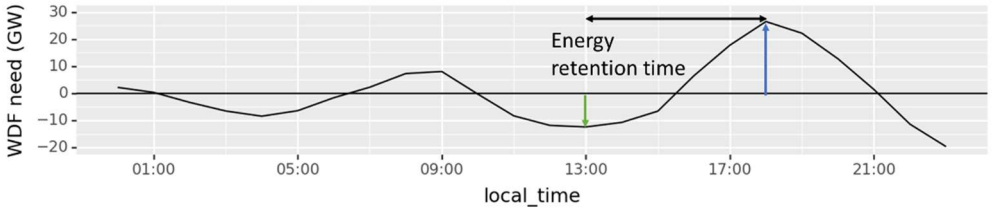
    - Cycle capacity as a battery: Maximum of the upwards and downwards cycle capacities, i.e., the largest amount of energy required upwards or downwards during a single half cycle (between two crossing points). In the example above this would be the blue area, since that is larger than the green area.
    - Cumulative capacity above (below) zero: Maximum (minimum) of cumulative energy throughout the day: considering net energy that would be stored/discharged across all cycles by a theoretical storage asset acting to meet the WDFN. Point A in the diagram shows the time where most energy would be stored, and point B shows the time when most amount of energy stored would be lowest. By definition, cumulative energy will be 0 at the end of each day as there is overall energy balance.



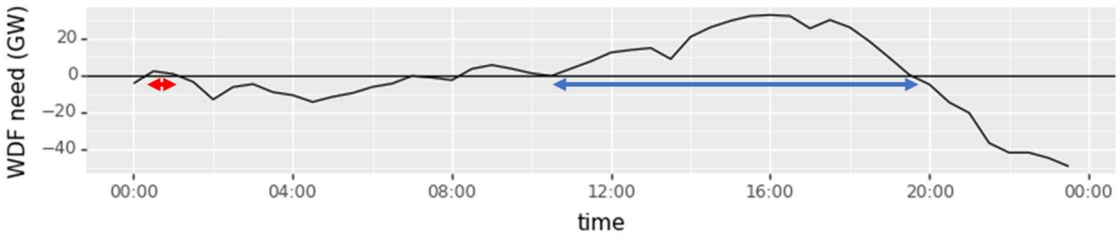
- Cycling: describing the shape of the within-day need
    - Number of crossing points and probability of each number of crossing points: below are examples of days with 1 (top) and 14 (bottom) crossing points.
    - Sign of WDF need at the start of the day upwards (top example) or downwards (bottom example day).



- Energy retention time: time between peak upward and downward power



- Number of hours between crossing points: the minimum (red) and maximum (blue) time the need is sustained in one direction (one cycle).



- Shifting of energy forwards vs backwards: the relative timings of the peak upward vs downward need. This is quantified by the proportion of days where the peak upwards occurs before the peak downwards need.

The calculated values for all metrics for each day are returned by the method and can be used to investigate seasonal patterns, links to weather variables etc. The overall cumulative distributions of metrics are also presented, which give information on the variation in values across the whole set of days and allow probabilistic questions to be answered, such as what level of WDF need has a given probability of being exceeded. This is particularly important for rare but extreme events, where a very large amount of within-day flexibility may be needed. For probability levels beyond 5% and 95%, GPD tails have been fitted to return robust estimates of the most extreme values. It should be noted that this is currently implemented in a separate function, and these extreme value tails are not plotted on the overall CDFs presented below. The function fitting a GPD distribution to the tails can be used on the half-hourly WDF need itself, or on any of the daily metrics that return a numerical value (just not the metrics that return a single proportion of days value).

2.5 Key assumptions

The assumptions made in this work have been logged in an Assumptions Log, which is included as an accompanying Annex to this report. The Assumptions Log has been used to document any assumption made during the WDF calculation process. The log covers the identification of the assumption, the stage at which it is made and the area of the modelling process it applies to. This is followed by a discussion of the need and as well as a justification for the assumption. Finally, the impact of the assumption is stated (high, medium, low or very low), along with whether the assumption will lead to an over or underestimation of the WDFN.

Firstly, there are two fundamental assumptions that define WDFN and underpin the process outlined in this report – that the system can be treated as a single node (locational constraints can be ignored – a high impact assumption that will likely underestimate the WDFN), and that both the sub-half hourly and intra-day imbalances are resolved through other mechanisms. Any change to these assumptions would necessitate a review and update to the WDFN calculation process as outlined in Section 2.3.2.

Additional assumptions have been made during the project. Out of these, there are five key assumptions that have been given an impact rating of medium.

- Examining a single interconnector to Ireland.
    - This assumption arises from the initial definition of WDFN (that calculations are for GB with locational constraints neglected). However, it is also a practical decision due to the considerable volume of additional data that would need to be collected to examine all interconnected countries.
    - Both weather data and a similar type of FES pathway data are available for the island of Ireland, and readily accessible, and setting up the methodology for one interconnector would enable the process to be repeated for others as and when the data is available.
    - It is expected that interconnectors will broadly reduce the WDFN, therefore this assumption will likely overestimate the WDFN. However, there may be certain circumstances where they exacerbate the need, for example coincident periods of high demand and low renewable generation in the interconnected countries.
- The scope is limited to within-day flexibility therefore individual days are independent.
    - This means that any large periods of WDFN at the very start and end of the day may be artificially “clipped” at midnight.
    - It is necessary to define a cut-off point to calculate WDFN, but if this time was adjusted, for example to the operational day of 5am to 5am then the results calculated may be different.

    - Furthermore, the length of a “block” of WDFN occurring over midnight may be underestimated.

- EV charging and heat pump demands are inflexible.
    - Although it can be argued that some flexibility is inherent in EV charging already, the actual volume of this is unknown and for the purposes of the WDFN calculation it is assumed that all EVs are inflexible.
    - EV demand is still defined as inflexible to quantify the need for flexibility. Any flexibility in EVs is still a source of WDF, and therefore out with the scope of this project.
    - Even if flexibility from EVs is expected to be readily available at scale in future scenarios, it is still a source of WDF and so its contribution should be assessed following the calculation of WDF, not within it.
    - For heat pumps, there is limited trial data available and therefore greater uncertainty on their future behaviour. Therefore, assuming inflexibility is a reasonable assumption for this work and can be updated when the technology becomes more widely established.
- For non-domestic EVs, the number of vehicles using depot or private charging is entirely dependent on the underlying FES assumptions.
    - This proportion is judgement based and subject to uncertainty, and introduces uncertainty not captured in the modelling.
    - This will need to be reviewed as the FES scenarios are updated.
- Finally, that all the results are dependent on FES inputs.
    - This is an intuitive and necessary assumption meaning that all results and metrics are conditional on both the FES and future year under consideration. If any of the underlying models driving FES or the definition of the pathways change this will impact all the calculation results, and so the process would need to be re-run with updated inputs.

The effect of both the single interconnector to Ireland and the inflexibility of EV charging profiles are examined via a sensitivity analysis, as discussed later in Section 4. Assumptions rated as low or very low impact are discussed in the Annex only.

2.6 Alternative methodology options

Methodological decisions were taken throughout the project, ranging from the overarching approach to feature engineering for individual models. Some of the main alternative approaches are discussed here.

One alternative approach could have been to set up a system level simulation (e.g., a wholesale market model) with different sources of flexibility, provide the underlying profiles (for demand and renewable generation) and then calculate the volumes of flexibility that are dispatched. However, this was felt to be unnecessarily complex for calculating purely the WDF need, which is primarily driven by the inflexible (and generally weather and time dependent) components of demand and generation and can (within the approach taken here) by derived analytically.

We have chosen to use nonparametric probabilistic models in the weather rescaling of most demand and generation components. This is a flexible approach, allowing modelling of multiple variables with very similarly structured models. The suitability of a parametric model depends on specifying an appropriate distribution for the probabilistic prediction, which can be limited by the set of distributions easily available in Python and can require judgement and application-specific knowledge. Fitting a set of quantiles in a nonparametric formulation allowed a good fit across all probability levels,

and use of lower intervals between probability levels at the tails where necessary to improve the model fit near the boundaries.

In addition, we have deliberately avoided the use of parametric auto- and cross-correlation models for modelling sequences of demands and generation. With five models and half-hourly granularity, this would require a 240-dimension covariance matrix for modelling daily correlations, with the distinct possibility that this needs to vary seasonally.

The assumption of completely inflexible EV and heat pump demands could be addressed by introducing separate profiles for inflexible and flexible behaviours. While a significant assumption, This was deemed the best approach due to the lack of data on and knowledge about the future shape of flexible behaviours (and the relative flexible and inflexible proportions). This is discussed further in sections 2.2.2 and 2.5.

Simpler models were used for electric vehicles and heat pumps, as these models are based on a much shorter historic dataset from trials, rather than national historic values. Since these technologies are not yet as widely established, there is greater uncertainty on their future behaviour than components like wind and solar that are purely weather-dependent. Because of this, a simpler weather correction using deterministic models was applied rather than the more detailed probabilistic weather rescaling approach. An alternative would be to use models that simulate consumer behaviour (e.g., how to users heat their homes in response to different weather patterns).

3 Results

The results presented below are based on the demand and generation models described in section 0. The method in general allows calculation of WDFN and metrics for any FES, or indeed a custom scenario with a different technology mix. These results focus on the year 2035 to limit the number of results presented, but any year where there are FES values (i.e., out to 2050) could be used. Metrics have been compared to their equivalent values for 2023 to help understand the magnitude and significance of the changing needs for system operation and security.

3.1 Types of plots

Two main types of graphs have been used to display the results of WDF metrics. Both display the distribution of the set of daily metrics across all time points in the data, spanning the 40 weather years' worth of varying inputs.

Plots of the Cumulative Distribution Function (CDF) of a variable give the probability that any threshold (plotted on the x-axis) is exceeded. In the example CDF in Figure 34, the peak WDF need exceeds 36.5 GW 10% of the time. Note a logarithmic scale has been used on the x axis to emphasise the long tails in the data, and that the lower limit on this axis indicates the minimum value observed for that metric.

Figure 35 shows an example of a violin plot, which is another way of visualising a probability distribution and is used for some of the results. The smooth envelope shows the estimated shape of the continuous distribution of the data, and the box plot indicates key probability levels, like the median.

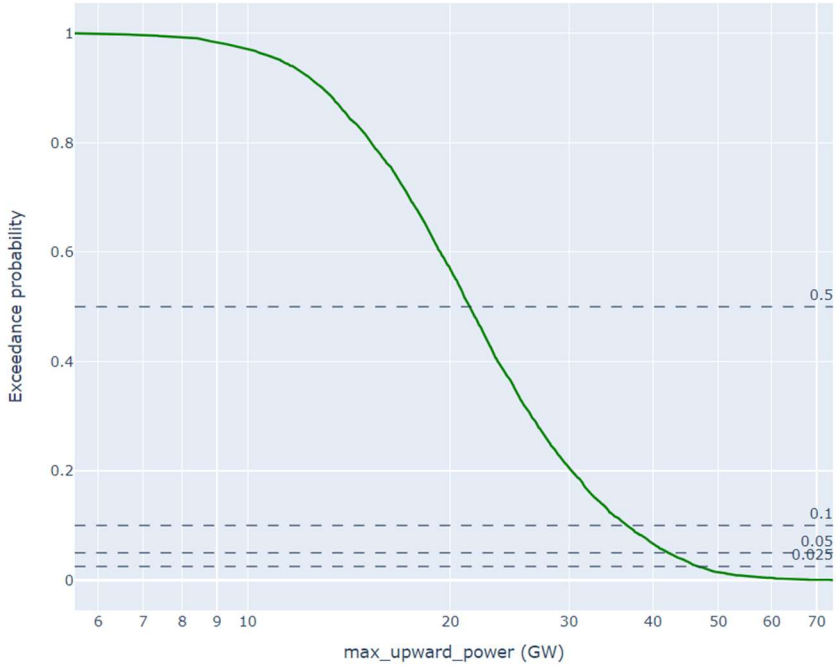


Figure 34: Example cumulative distribution function

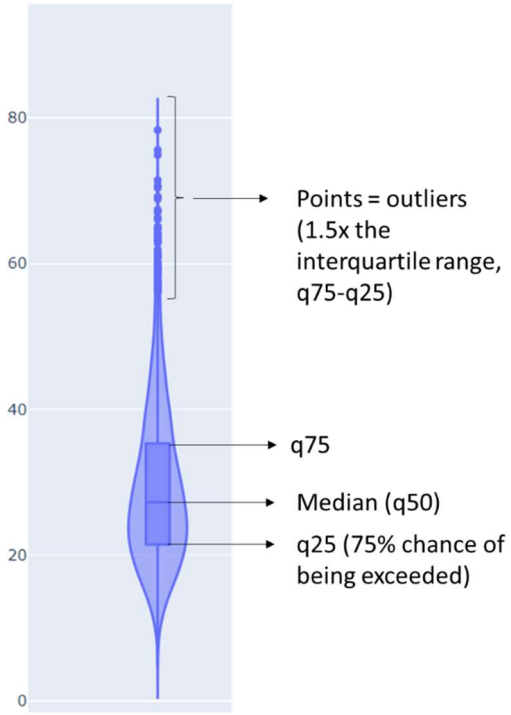


Figure 35: Example of a violin plot

3.2 Results

The following subsections discuss the final results in detail. These show the distributions of metrics calculated on all days with a complete set of 48 half-hour values in the 387-year dataset. For brevity, only 2035 has been used (and 2023 for comparison to current levels), but similar results could be generated for all years in the FES to show how the WDFN changes through the years. We would recommend only using these results out to 2035 since some assumptions are less likely to hold further out.

3.2.1 Daily peak power

The distribution of daily peak upwards and downwards WDFN varies with the season as shown in Figure 36 and Figure 37, for Leading the Way in 2035. These plots suggest that the needs for both upwards and downwards WDF are greatest in winter and lowest in summer, with slightly more seasonal variation in peak downward WDFN than peak upward. Both peak upward and downward WDFN display long tails, with the maximum observed values being around three times higher than the median value. This suggests a requirement for WDF solutions that can meet a need higher than the median, but that will not necessarily be utilised regularly.

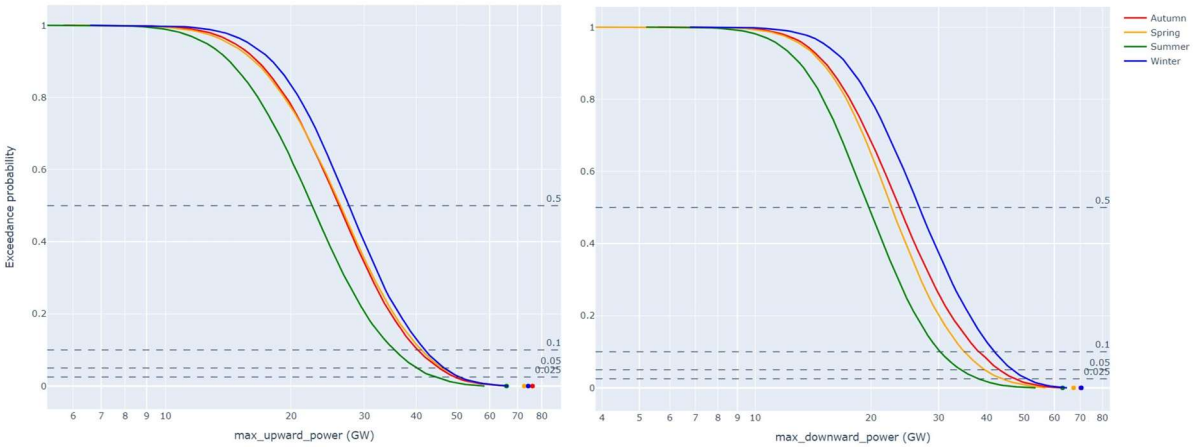


Figure 36: Distribution of peak upwards and downwards WDF need by season, for Leading the Way 2035

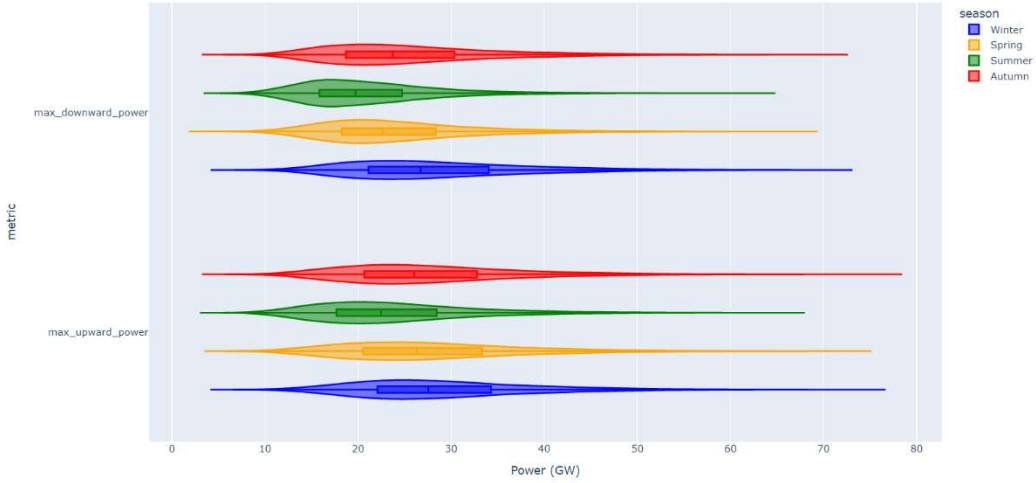


Figure 37: Distribution of peak WDFN across seasons for Leading the Way 2035

Figure 38 summarises the frequency with which the upwards WDFN in a day is greater than the downwards WDFN, by season (for Leading the Way 2035). Upwards WDFN is higher than downwards 57% - 70% of the time across the seasons, with the difference smallest in the winter. This asymmetry is notable given the assumption made in the calculation that the total *energy* needed for upwards and downwards WDFN in the whole day is balanced. This means that meeting WDF requirements will need a greater MW capacity upwards in a single half-hour (increased generation or demand reduction) than it will downwards. This also suggests that ramps in demand may be greater than ramps in generation: the *average* imbalance is zero so taller, slightly narrower peaks in demand could lead to this asymmetry.

However, it is important to bear in mind that this is heavily dependent on the data sources used to represent demand and generation, where the shorter datasets from trials for LCTs like EVs and electric heating from heat pumps may be more uncertain than the equivalent data used to represent renewable load factors.

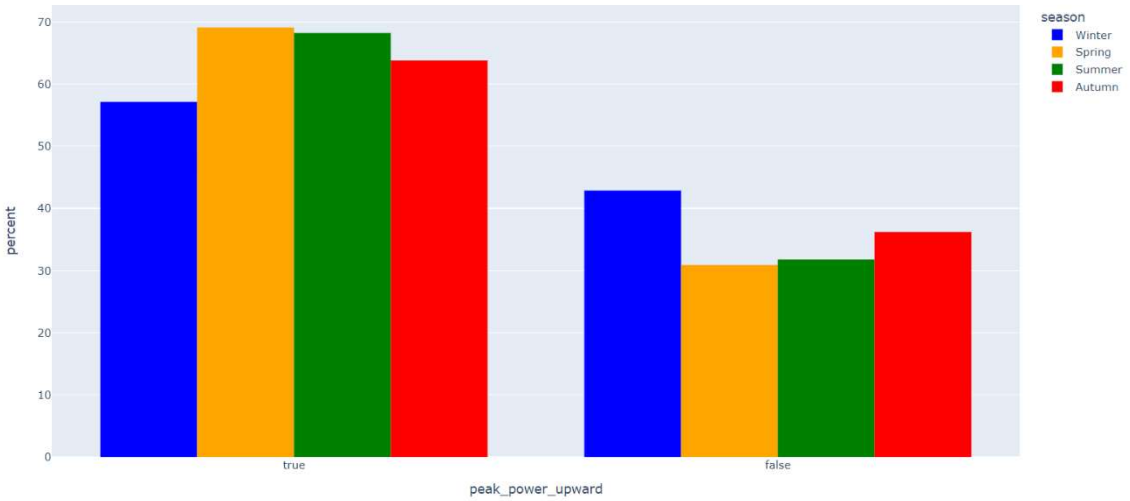


Figure 38: Proportion of the time peak WDF need is in the upwards direction, for Leading the Way 2035.

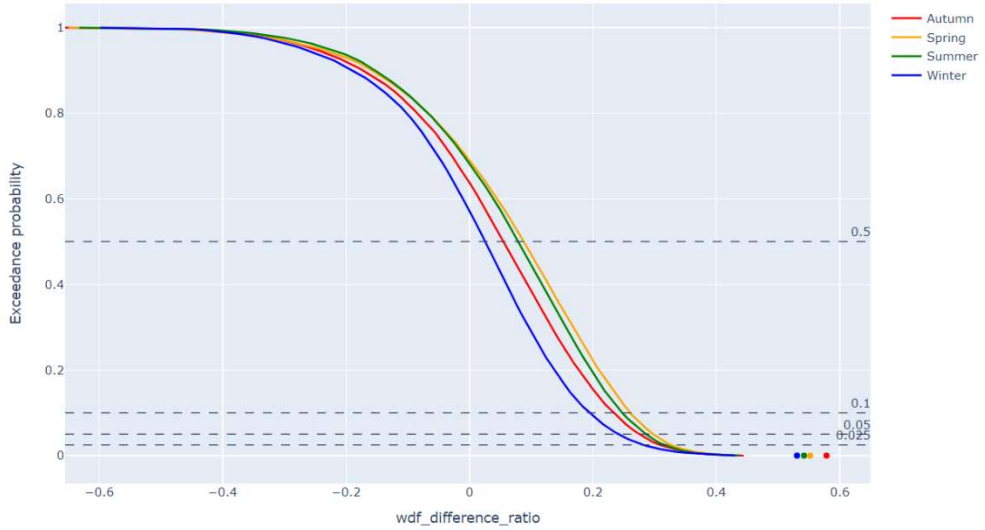


Figure 39: Distribution of daily difference in magnitude of peak upwards and downwards need by season, for Leading the Way in 2035

Figure 40 shows the variation in upwards and downwards peak WDFN through the year, in the Falling Short and Leading the Way 2035 scenario. These plots show the average across all 387 rescaled years in the central line, but also the 90% probability interval and the interval between the minimum and maximum observed for that day of year. As in the previous plots, the greater need in winter than in the summer is evident. This plot also shows the significant difference in requirements between the scenarios, with Leading the Way requiring higher, and more variable, peak WDFNs than Falling Short.

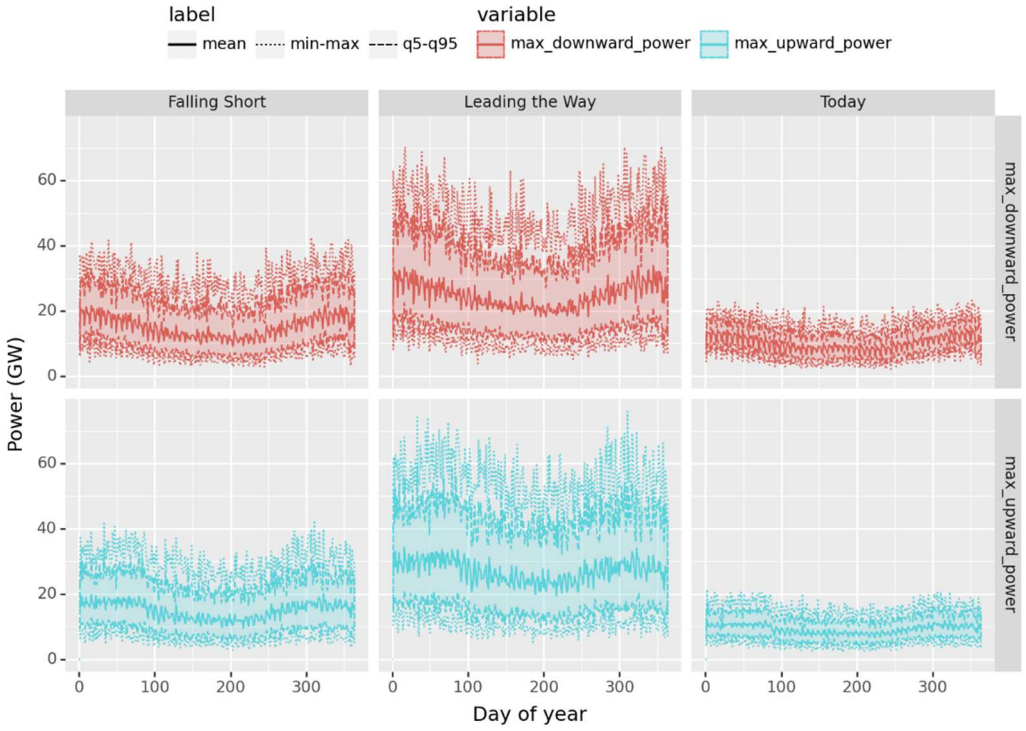


Figure 40: Variation and spread of peak WDFN throughout the year, for 2035

Figure 41 and Figure 42 and show the distributions of peak upwards and downwards WDFN under all four contrasting FES. The distribution for 'Today' is also shown, which is derived by scaling to the 2023 Leading the Way FES values. As might be expected (given the extent of renewable energy and demand electrification), peak WDFN both upwards and downwards is generally highest under Leading the Way. In addition, the spread of daily peak power is much greater for Leading the Way, with the most extreme modelled values nearly double those seen under Falling Short. Leading the Way 2035 includes 83.3 GW of dispatchable electricity supply, including gas, hydrogen, biomass, BECCS, interconnectors and storage. Assuming all capacity was always available, this would cover the peak upwards WDFN for all the modelled historic days in that scenario, and flexibility from LCTs would also be expected in this scenario and is discussed further in section 4.1.

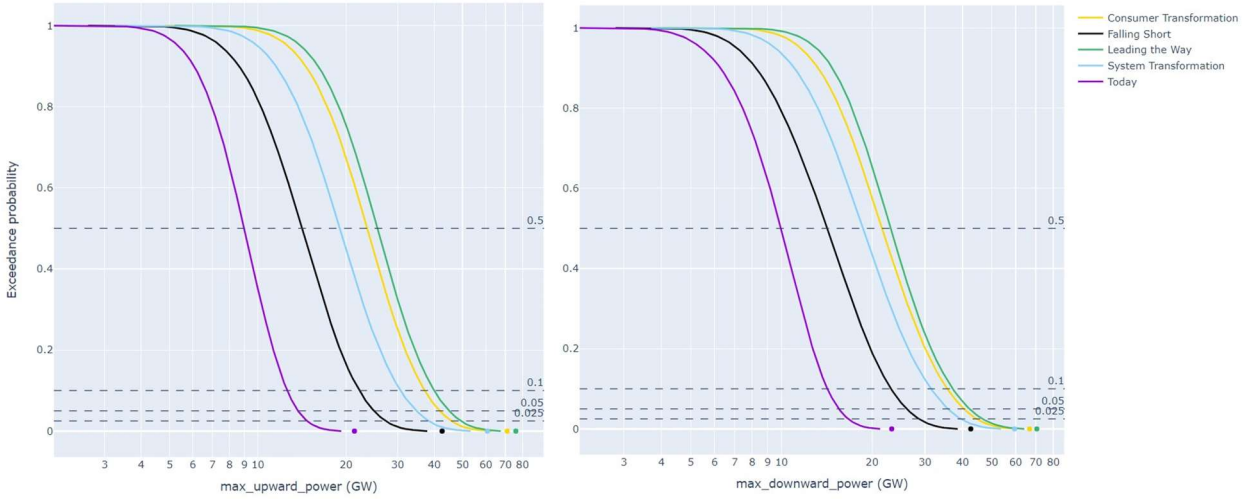


Figure 41: Distribution of peak upwards and downwards WDF need across all FES for 2035

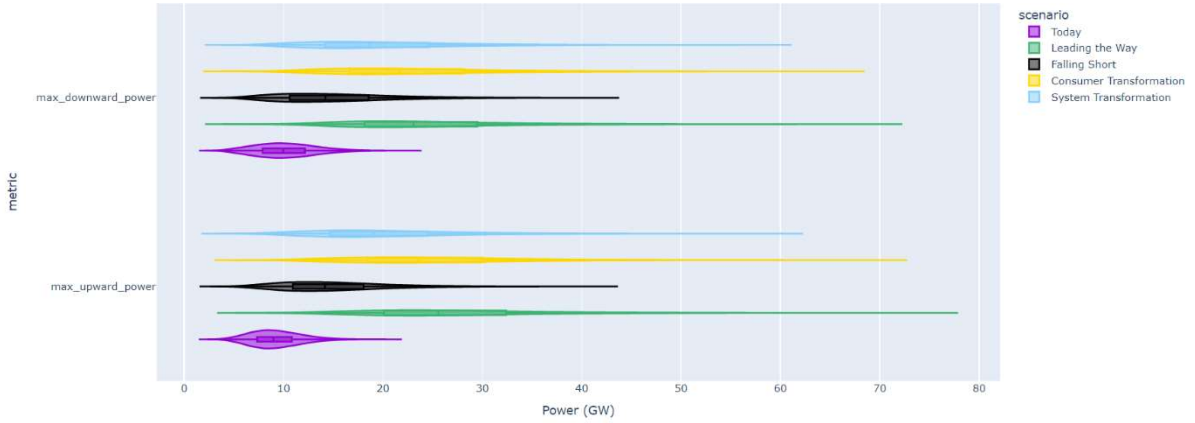


Figure 42: Distribution of peak upwards and downwards WDFN across all FES for 2035

Figure 43 summarises the frequency with which the need for upwards WDFN in a day is greater than the need for downwards WDFN, by FES. Peak WDF is more frequently upwards than downwards for Leading the Way, but the reverse is true for today and Falling Short, suggesting the more ambitious decarbonisation scenarios are more likely to have peak WDFN in the upwards direction (requiring an increase in generation) than scenarios more like the current energy system. However, this plot does not contain information on the *size* of the difference between peak upwards and downwards WDFN.

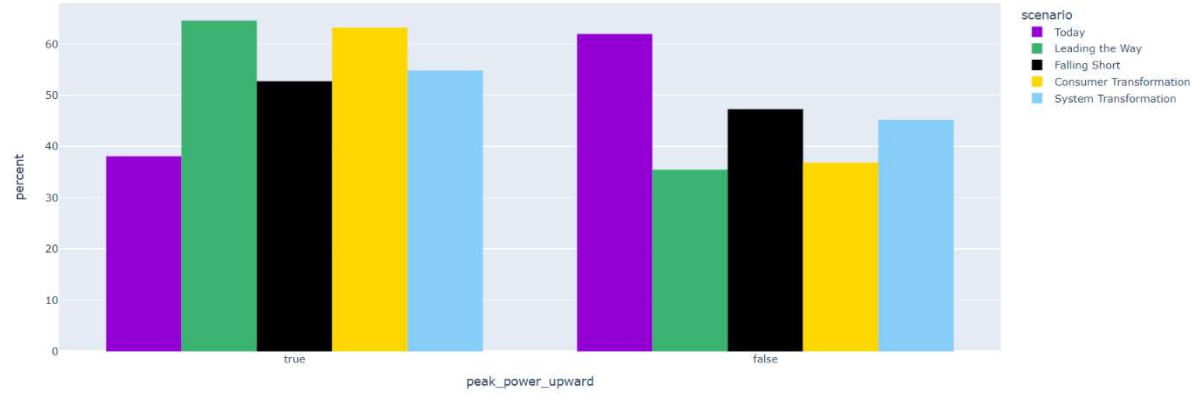


Figure 43: Proportion of the time peak WDFN is in the upwards direction, by scenario

The relationship between peak upwards and downwards WDFN can also be understood by looking at the average difference in the magnitude of upwards and downwards needs, normalised by the overall range in WDF, as plotted in Figure 44. In general, scenarios with the most progress towards zero-carbon technologies show a tendency for peak upwards WDFN to be larger than peak downwards. All 2035 FES scenarios show distributions where peak upwards exceeds peak downwards more often, and by a greater proportion, than today. Conversely, in 2035, on the days where peak downwards exceeds peak upwards, the difference is smaller than today. For Leading the Way in 2035, the difference in peak upwards and downwards WDFN is within 20% of the range, 75% of the time.

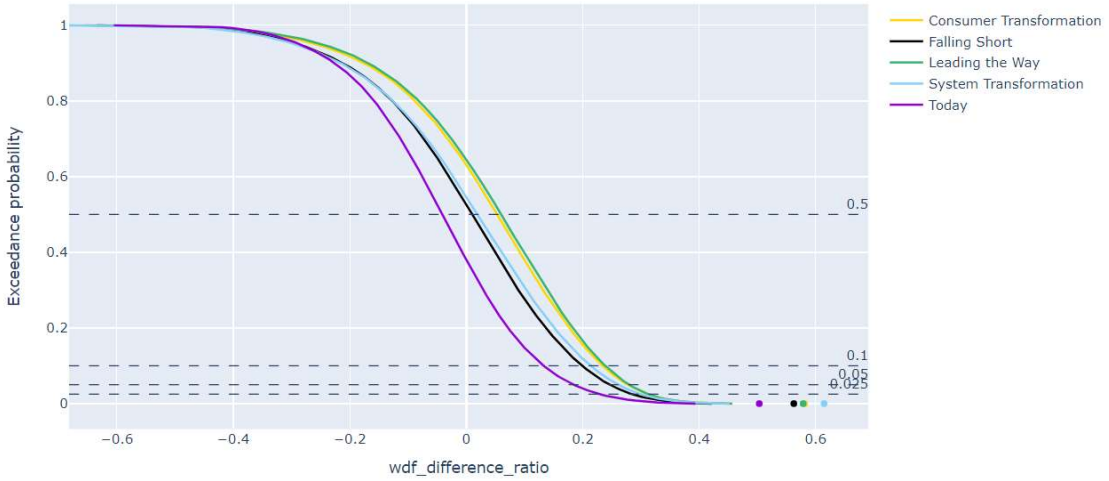


Figure 44: Difference in magnitude of peak upwards and downwards peaks, as a proportion of daily WDFN range

3.2.2 Ramp rates

Ramp rates of WDFN have been examined for three time periods: half-hourly, i.e. the ramp rate between neighbouring half-hour settlement periods, and the average ramp across both 2- and 4-hour windows. The maximum ramp rate across all windows for the day is then retained as the daily peak ramp for this metric.

Figure 45 summarises the variability within and between seasons in the peak half-hourly ramp rate of WDFN (upwards and downwards) for Leading the Way in 2035. These distributions suggest that a minimum system ramp capability of 0.1 GW/min is fairly consistent across all seasons for upwards and downwards ramps, with a median value of between 0.25 and 0.3 GW / min. The largest-observed ramp both upwards and downwards occurs in spring, with ramps well over 1 GW / min possible, even if this is rare. Both upwards and downwards peak ramp rates tend to be lower in summer.

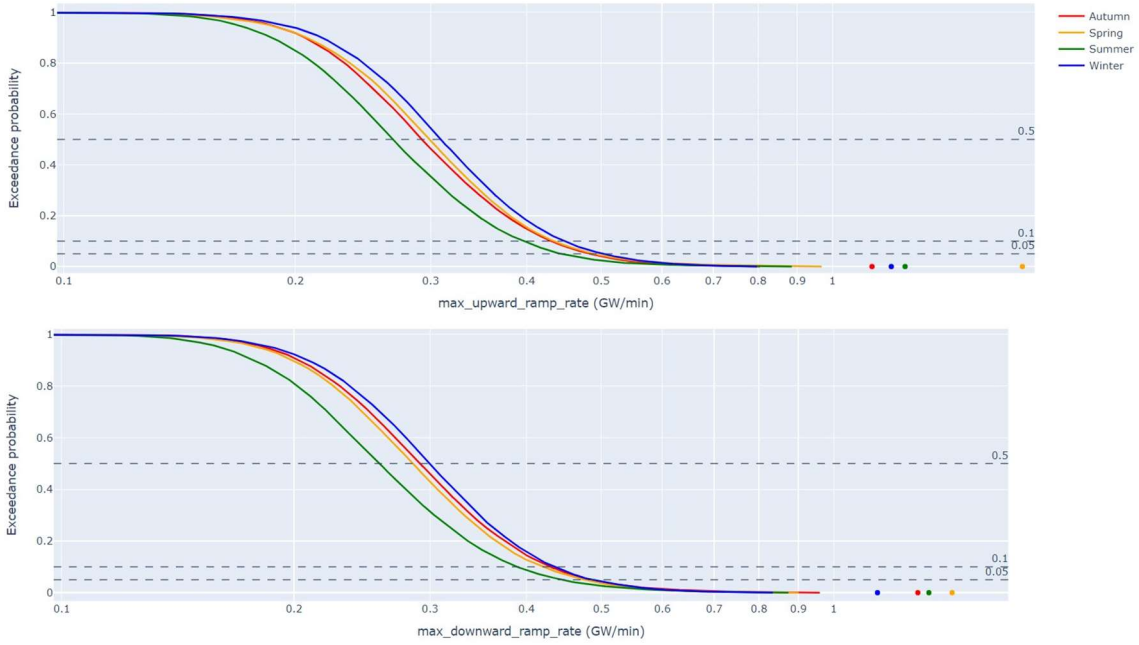


Figure 45: Maximum daily half-hourly ramp rates by season for Leading the Way 2035

Longer sustained ramps of 2 and 4 hours are on average lower but still significant, especially when considering the sustained large change in WDFN they represent. The seasonal differences in the more sustained ramps are like the peak half-hourly ramps, with lowest ramp rates in summer.

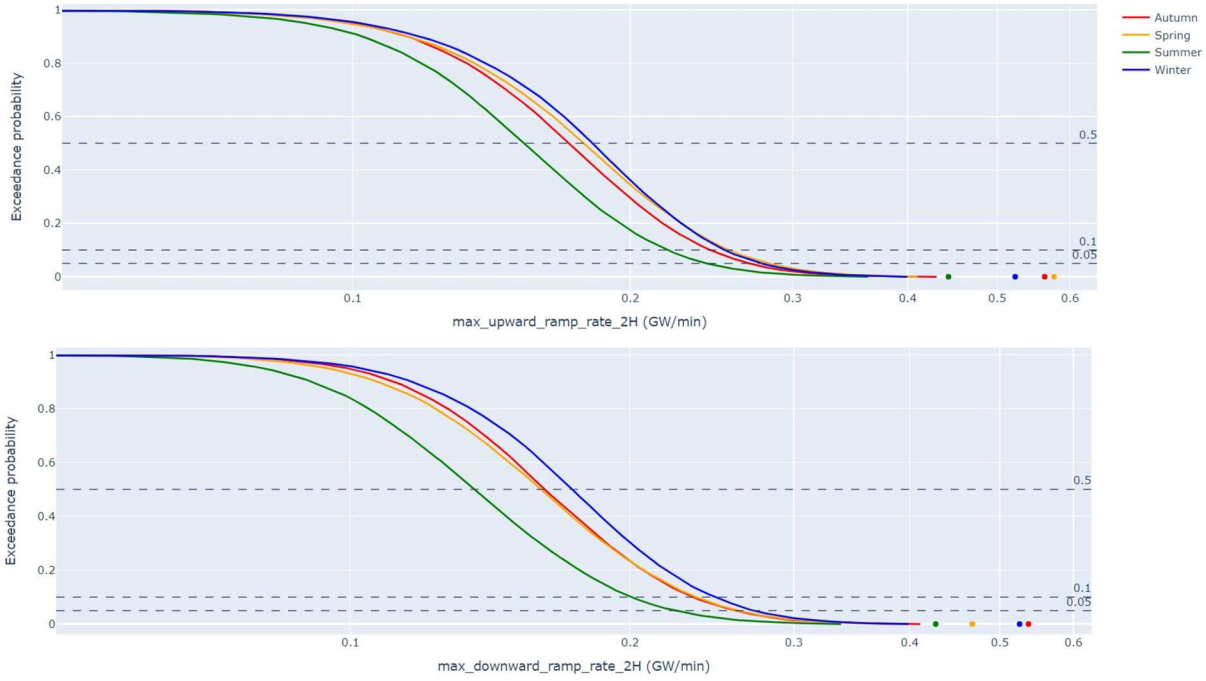


Figure 46: Maximum daily 2-hour ramp rates for Leading the Way 2035

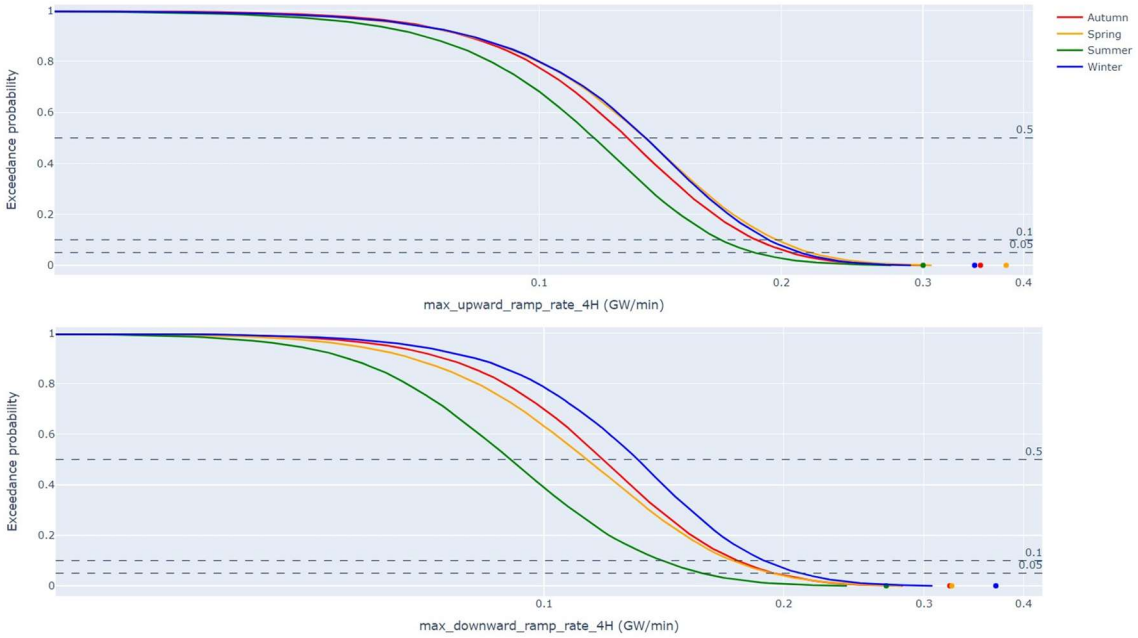


Figure 47: Peak 4-hour ramp rates for Leading the Way 2035

Figure 48 directly compares the upward ramp rates over the three time windows. As might be expected, ramp rates tend to be lower when sustained across longer time windows, with the most extreme observed half-hour upwards ramps around six times greater than the most extreme 4-hour ramps.

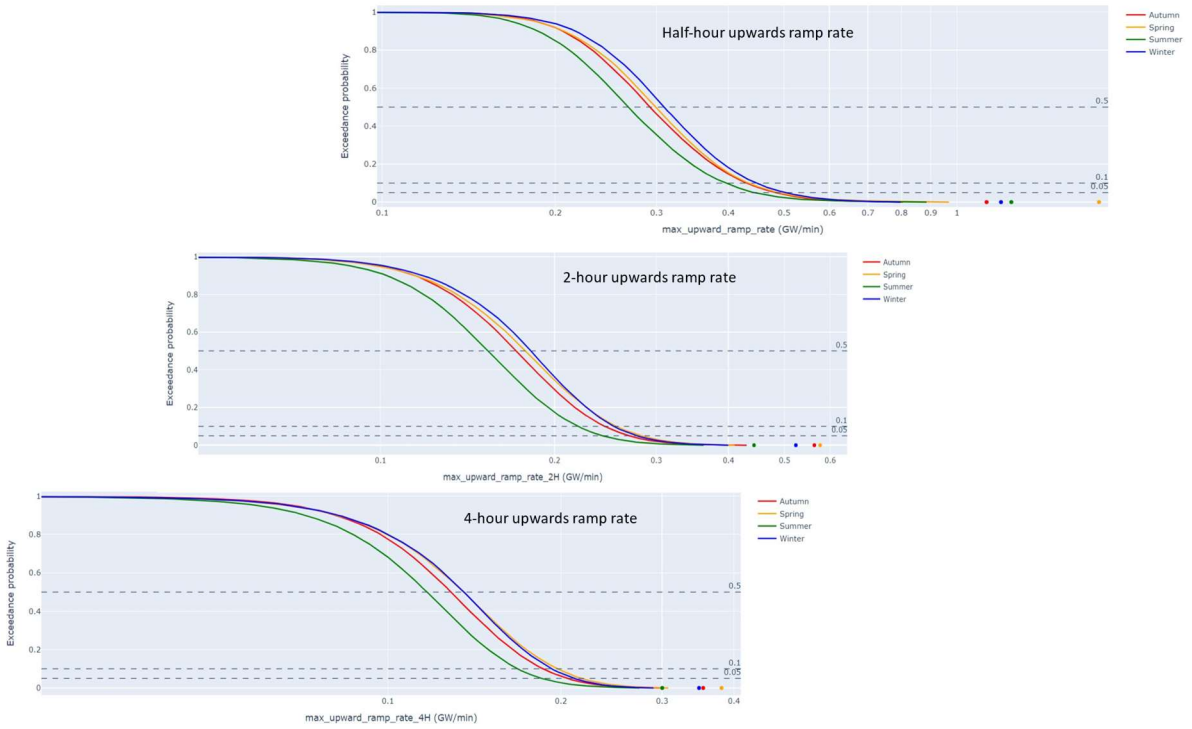


Figure 48: Comparison of half-hour, 2-hour and 4-hour upwards ramp rates for Leading the Way, 2035

Figure 49 compares the distribution of half-hourly peak upwards and downwards WDFN ramp rates across all four 2035 FES, and current levels. As might be expected, current WDFN ramps are lower than any scenario for 2035, with Leading the Way displaying the largest ramps.

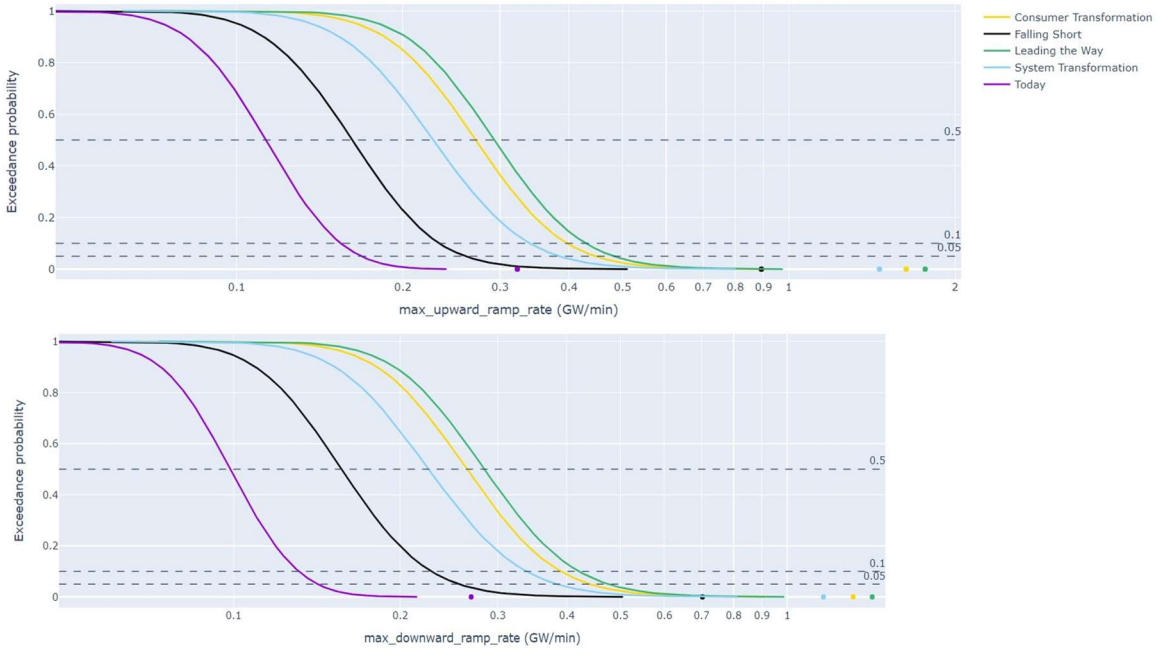


Figure 49: Distribution of maximum half-hourly upwards and downwards ramp rates under FES for 2035

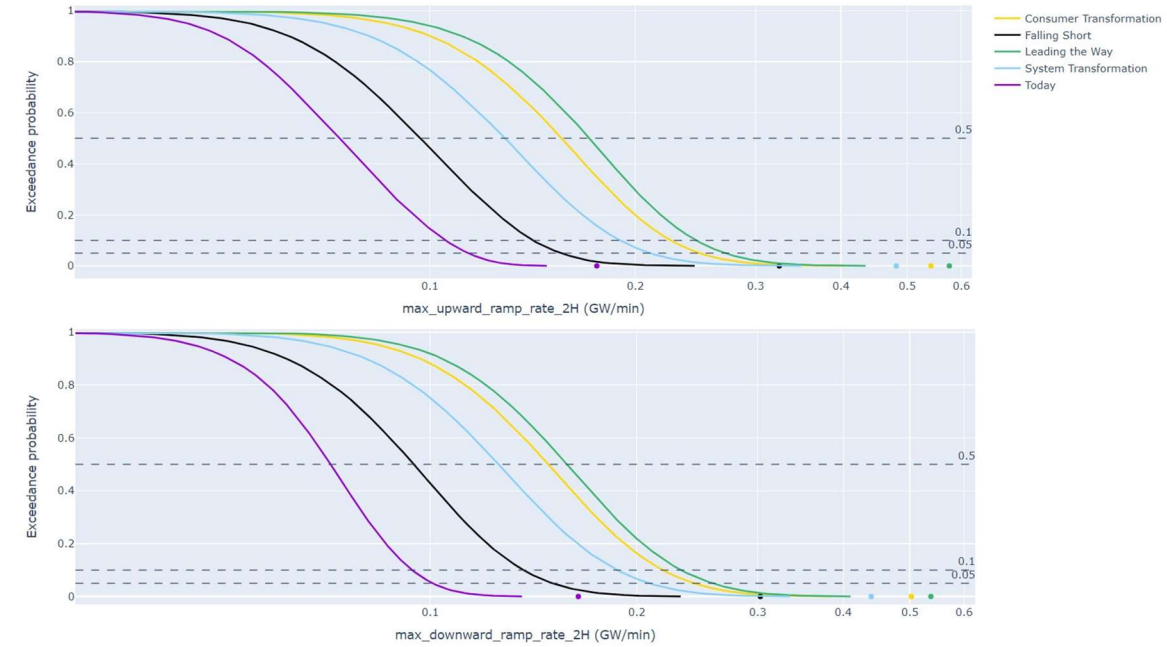


Figure 50: Distribution of maximum 2-hourly upwards and downwards ramp rates under FES for 2035

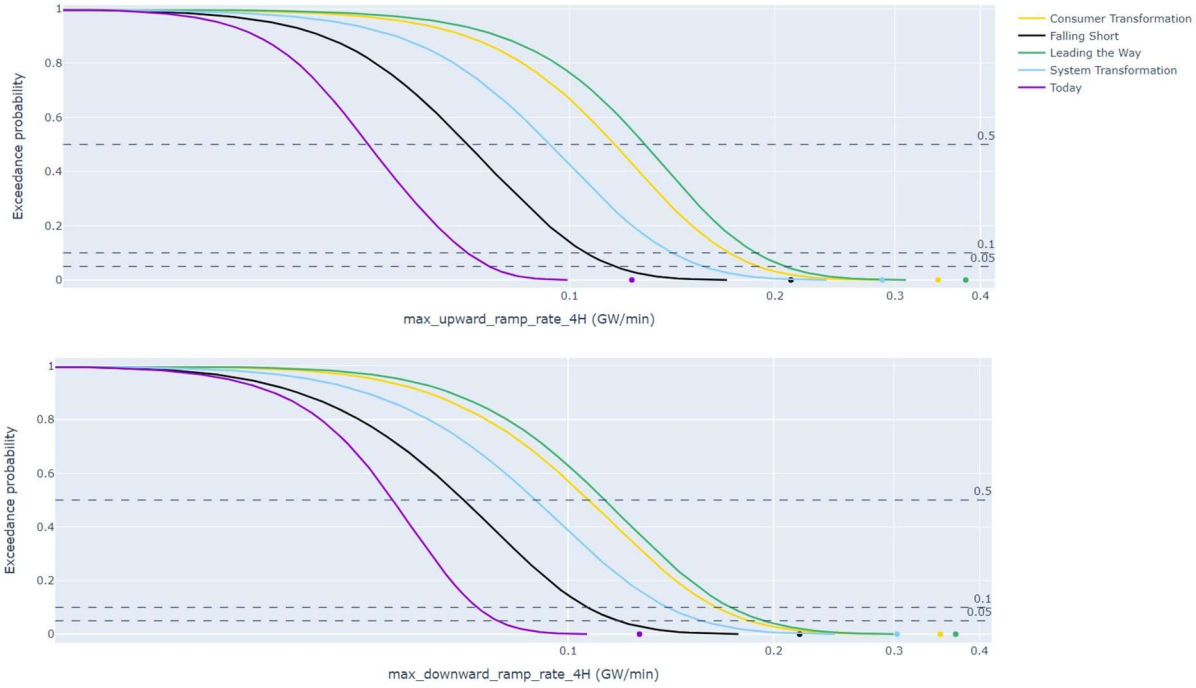


Figure 51: Distribution of maximum 4-hourly upwards and downwards ramp rates under FES for 2035

3.2.3 Energy metrics

These metrics consider the capacity of storage required to meet the need for that day. For context, Leading the Way is expected to have 118 GWh of storage in 2030, excluding V2G (values are not given for 2035).

Figure 52 shows the distribution of total energy that would need to be delivered by upwards *or* downwards WDF over the course of a day, that is, the total upwards (and downwards, as days are balanced) WDFN. This indicates the total daily within-day flexibility requirement that would have to be met through demand reduction and generation increases or storage discharging for upwards

WDFN, and vice versa for downwards. As with other metrics, requirements are highest in winter, lowest in summer, and similar in autumn and spring.

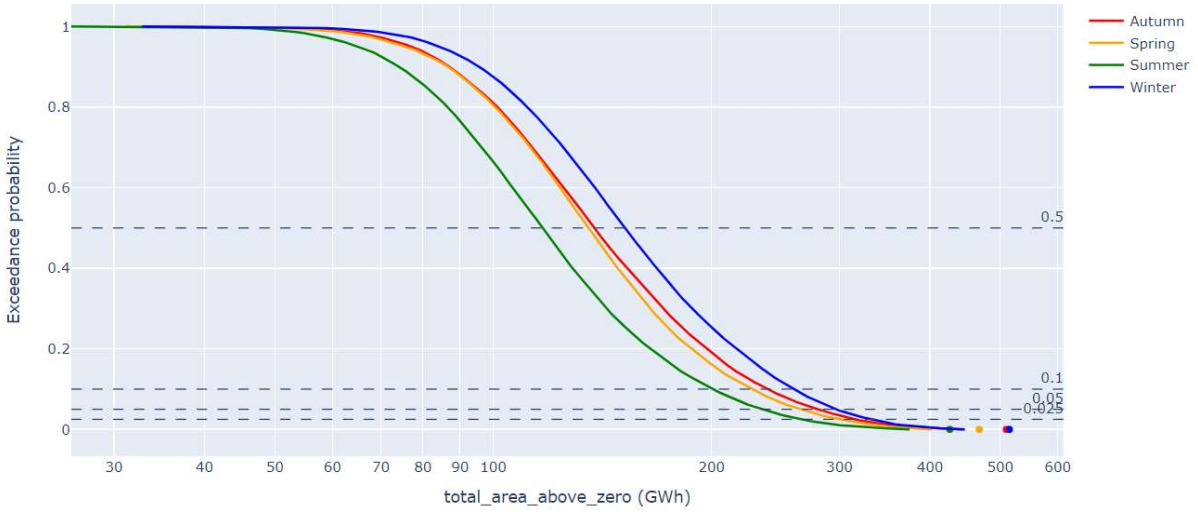


Figure 52: Seasonal distributions of total energy (area above zero) for Leading the Way 2035

Figure 53 then compares the year-round figures for all four FES. Again, Leading the Way shows the greatest requirements, and Falling Short the smallest, with all requiring a substantial increase from today: the level that is only exceeded 10% of the time today would be expected to be exceeded more than 50% of the time under Falling Short 2035, and over 90% of the time under Leading the Way in 2035.

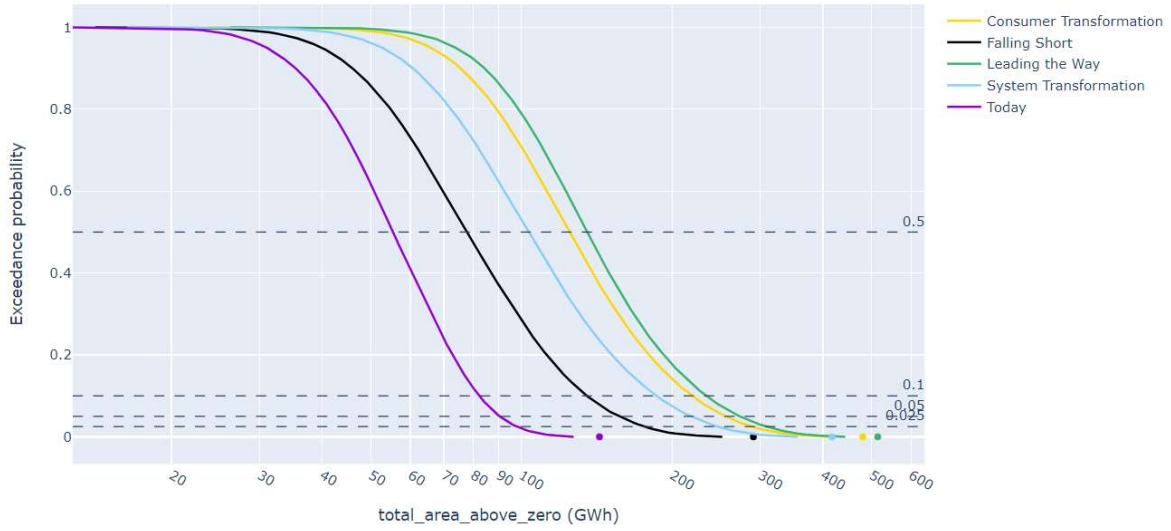


Figure 53: Distributions of total energy (area above zero) by FES, for 2035

Considering WDF solutions as general energy stores allows calculation of the equivalent capacity needed. Figure 54 shows a variant of this, which is how much capacity would be required in a perfect lossless battery to meet the energy need for the largest-energy half cycle (sustained upwards or downwards need between two crossing points) within the day. We can see a seasonal variation of 33 GWh in the median capacity required. More analysis of storage requirements would be possible to consider additional factors such as the state of charge requirements at the start of each day.

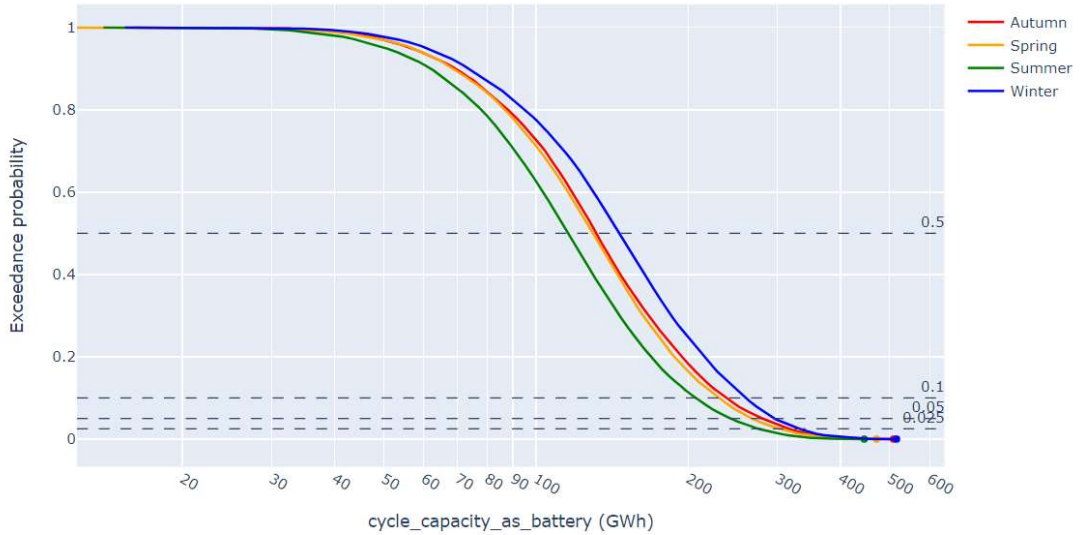


Figure 54: Seasonal distributions of equivalent battery energy capacity for Leading the Way 2035, to meet the largest single WDFN cycle within the day

Figure 55 then shows the equivalent figure for all four FES. The energy requirement represented by this metric differs from the total area above zero in cases where there are multiple cycles and peaks in the WDFN through the day and will always be equal to or lower than the total area above zero. For example, the median battery cycle capacity for Leading the Way 2035 is 127 GWh, but the median value of total upwards WDFN for the whole day is closer to 135 GWh because there is generally more than one cycle per day (and battery cycle capacity represents only the largest cycle).

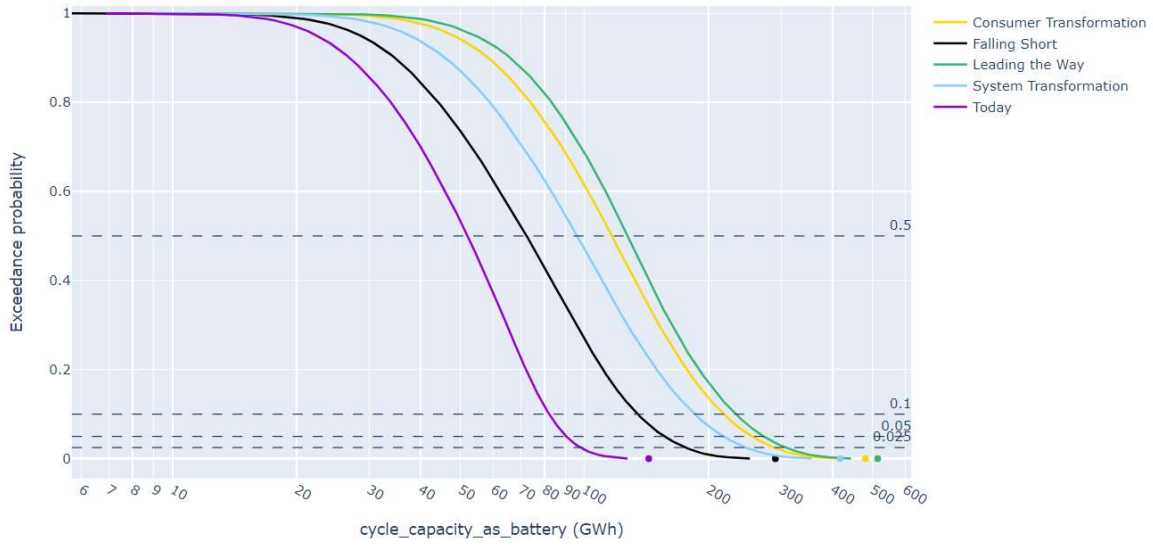


Figure 55: Distributions of equivalent battery energy capacity by FES

If one single lossless energy store was to provide all WDFN (both upwards and downwards) for the day, the capacity needed would be equal to the largest magnitude of the cumulative WDFN (in energy) over all time points. This is lower than the total area above zero, as charging and discharging are assumed to cancel out if from the same source. This maximum of the cumulative capacity both upwards and downwards is plotted in Figure 56, showing large spread out to 350 GWh upwards, or 500 GWh downwards. The larger cumulative capacity downwards is due to most days having downward need earlier in the day than upward need.

The difference between 2035 FES scenarios is more pronounced, with Leading the Way having a much longer tail than Falling Short, which in turn has a higher tail than current system conditions.

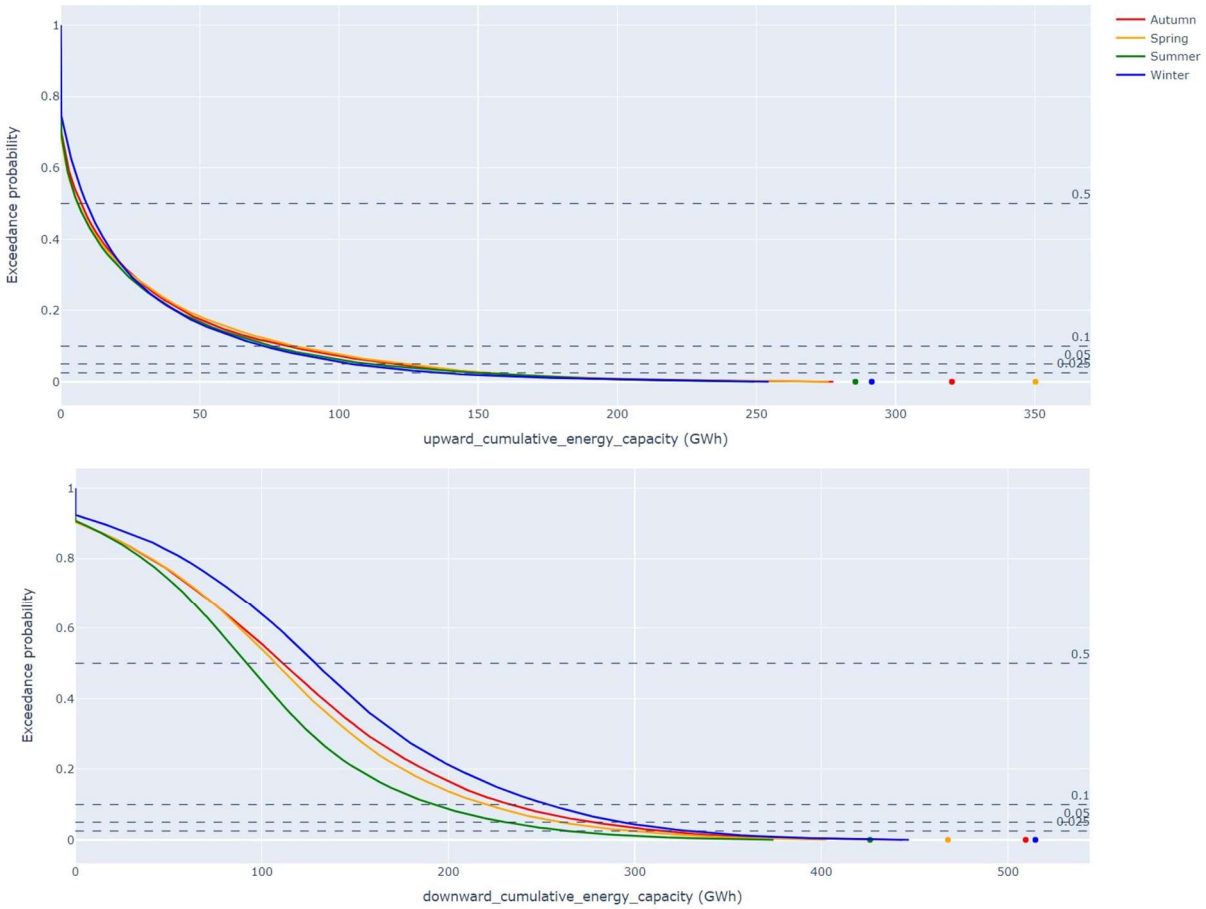


Figure 56: Distribution of cumulative energy capacity needed if one energy store was to meet all the WDFN by charging and discharging throughout the day, by season for Leading the Way 2035

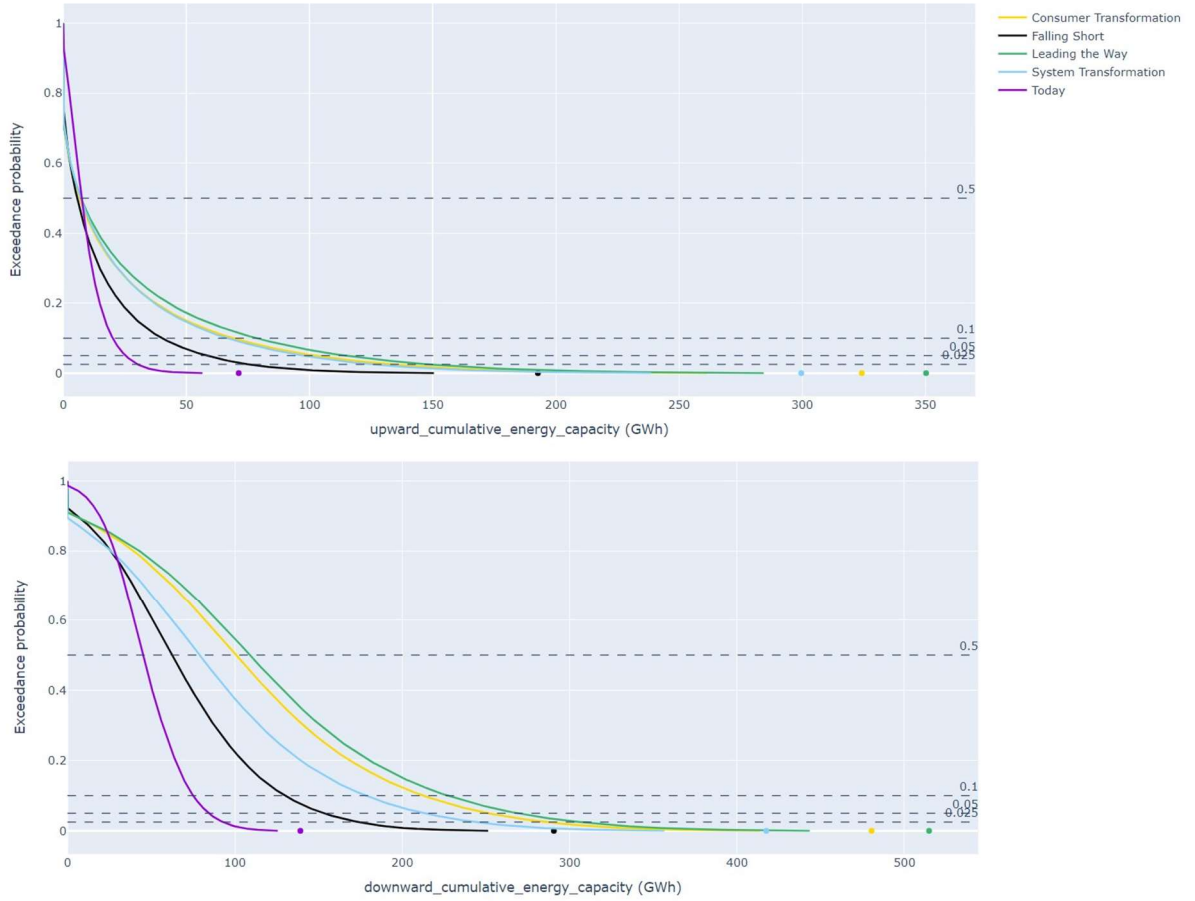


Figure 57: Distribution of cumulative energy capacity needed if one energy store was to meet all the WDFN by charging and discharging throughout the day, by FES scenario for 2035

3.2.4 Crossing points

Figure 58 displays histograms of the number of zero-crossing points for the daily WDFN, by season. The number of crossing points gives an indication of the shape and 'wiggliness' of the WDFN profile, and whether the WDFN tends to be sustained in one direction for multiple hours, or alternates between upwards and downwards need throughout the day. A linear profile would have one crossing point, with the WDF requirement only switching from positive to negative (or vice-versa). The high frequency of three, four, and even greater numbers of zero-crossings within these results is notable and aligns with the results already presented for equivalent ideal store capacities. This suggests that it is common for the WDF requirement to have multiple local peaks in a day which could be due to changing behavioural patterns in the day, for example increased heating demand in the mornings and evenings, and peak solar generation at midday. These distributions of zero-crossings appear similar across seasons, although further interrogation of the causes of the profile shape would be needed to determine how changes in demand and generation components through the day are linked to crossing points, and whether there are other differences between profile shapes between seasons not captured in this metric.

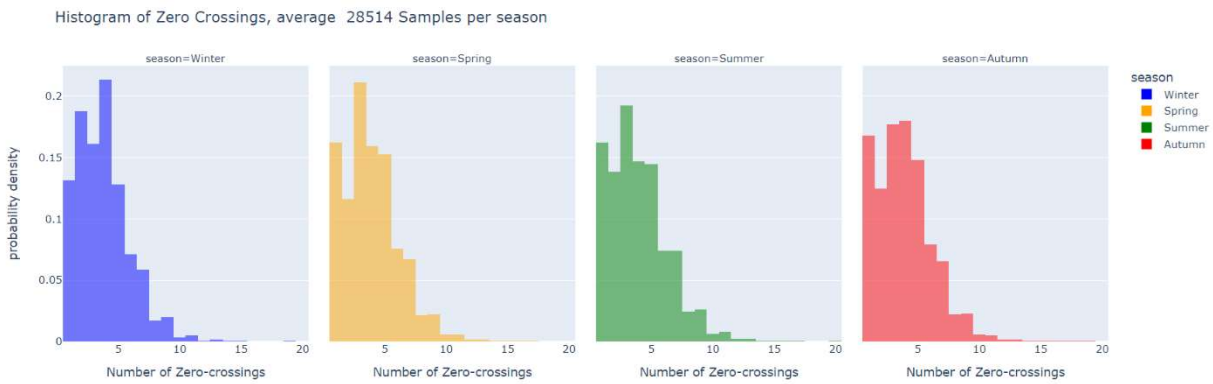


Figure 58: Distributions of zero-crossings per season for Leading the Way 2035[12].

Related to this, Figure 59 summarises the *initial* direction of the WDFN at midnight, in terms of whether the initial requirement is for downwards WDF (in the left-hand group of bars) or for upwards WDF (in the right-hand group). Winter, spring, and autumn are more likely to start the day with a shortfall in demand, while summer more likely to start the day with a shortfall of generation. These reported initial shortfalls are relative to the rest of the day due to the overall day energy balancing applied.

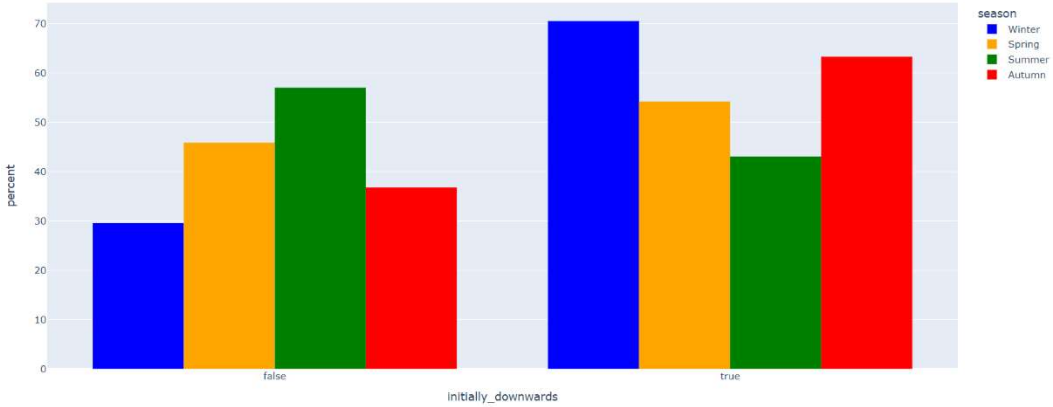


Figure 59: Direction of WDF need at the start of the day (midnight), for Leading the Way 2035

In the context of multiple cycles per day, Figure 60 summarises the length of cycles, showing how the *minimum* and *maximum* daily cycle length are distributed by season. In all seasons, the longest cycle of the day is typically nine to ten hours, but there is considerable variation, with cycles that last over 15 hours possible. The shortest cycles are typically only half an hour, but again, there is a possibility for a lot of variation. The flat portion of the minimum cycle time CDF between 4 and 8 hours suggests minimum cycles in this range are rare. The distributions of both the minimum and maximum cycle times are generally consistent across seasons, with only winter showing a slightly lower probability of minimum cycle times over 2.5 hours compared to the other seasons. Minimum and maximum cycle times are calculated from the lengths of the cycles between all adjacent zero crossing points and the start and end of the day.

[12] The weather rescaling process provides 387 years of modelled data which would equate to 35313 days per season: the slightly lower sample size here is due to dropping any days with missing half-hours.

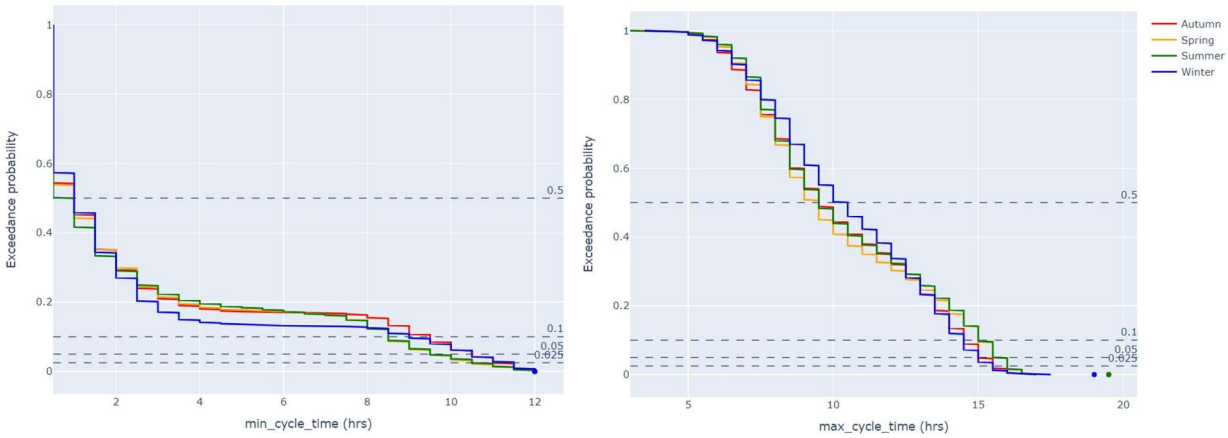


Figure 60: Distribution of time of longest and shortest cycles within the day, for Leading the Way 2035

Figure 61, Figure 62 and Figure 63 present similar results per FES for 2035 rather than per season. The distribution of zero crossings in 2023 shows a dominance of even-numbered crossings, indicating days where the sign of WDFN at the end of the day is the same as the start of the day. This effect can be seen to a lesser extent in 2035 for Falling Short, but is much weaker for the other FES. Figure 62 indicates a more even distribution of positive and negative WDFN at the start of the day in 2035 than 2023, so while the majority of days will still be expected to start with downwards WDFN, the proportion is expected to reduce. Maximum cycle times are expected to are still bimodal in 2035, but less strongly so than in 2023 while there is a higher probability of lower minimum cycle times for all 2035 scenarios compared to today.

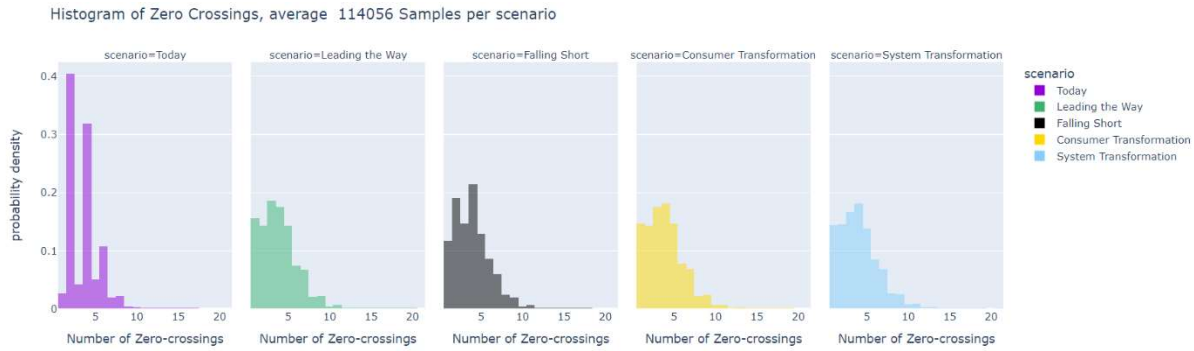


Figure 61: Distributions of zero-crossings per FES in 2035

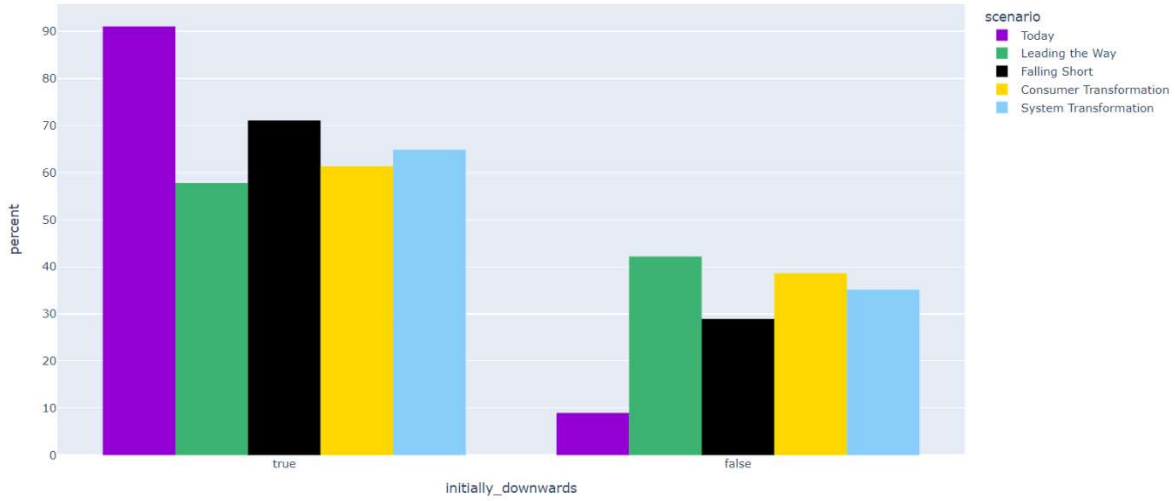


Figure 62: Direction of WDFN at the start of the day (midnight), for FES 2035

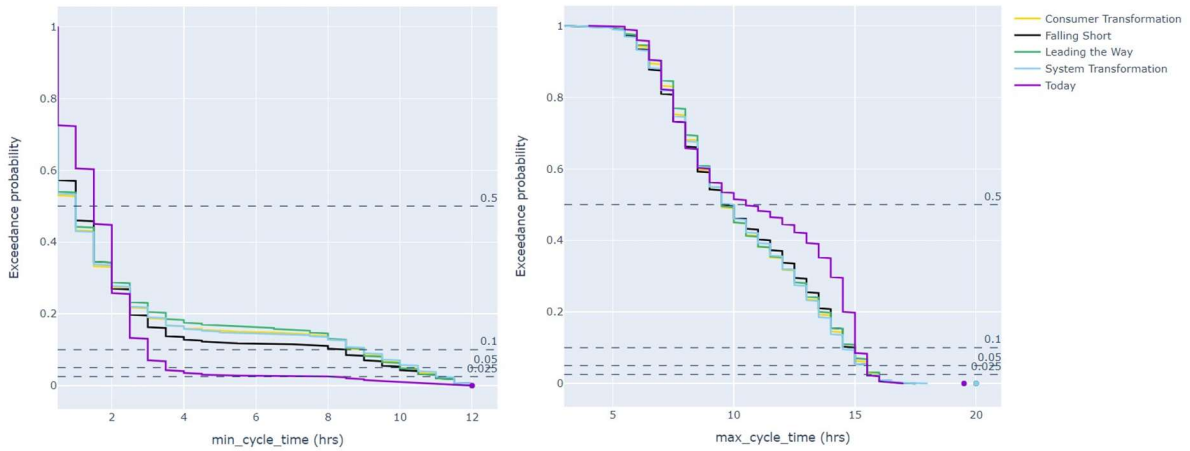


Figure 63: Distribution of time of longest and shortest cycles within the day, for FES 2035

3.2.5 Duration of energy retention

The next set of results relate to duration of energy retention, calculated here as the amount of time between the peak upwards and downwards WDF requirement and representing the length of energy storage or length of delay of behaviours needed by flexibility providers.

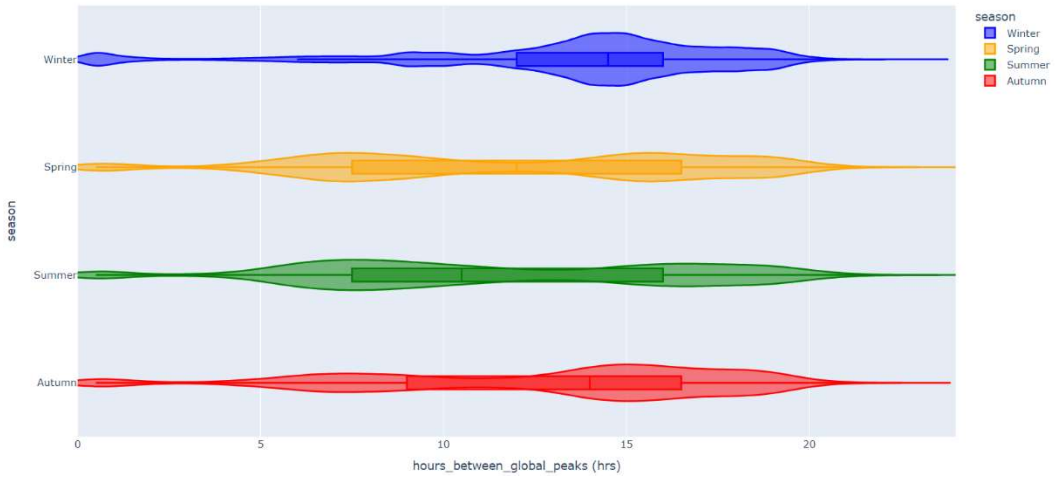


Figure 64: Distribution of duration between peak upwards and downwards need for Leading the Way 2035

The violin plots in Figure 64 and Figure 65 show the distribution of the elapsed time between the daily peak upwards and downwards WDFN, as a proxy for duration of storage or length of time shift of behaviours such as charging an EV or turning a heat pump on/off that would be needed to match the WDFN profile. Most days have more than 10 hours between global peaks, a length of time it is unlikely that flexible LCTs would be able to shift their demand by. Note these distributions are bounded at 0 and 24 (the apparent overspill is due to the way the smooth densities are calculated).

Figure 64 shows how this varies between seasons for Leading the Way 2035. Spring and summer show similar bimodal behaviour, while in winter the median duration between global peak is much higher, with the main probability mass centred around 14 hours. It may also be helpful to examine the order of the peaks, to understand the type of flexible solutions that would be required to mitigate them: for example, to show whether this represents a need for pre-emptive or delayed behaviours, as some

might be more readily available than others, i.e., pre-heating a residence vs delaying putting the heating on.

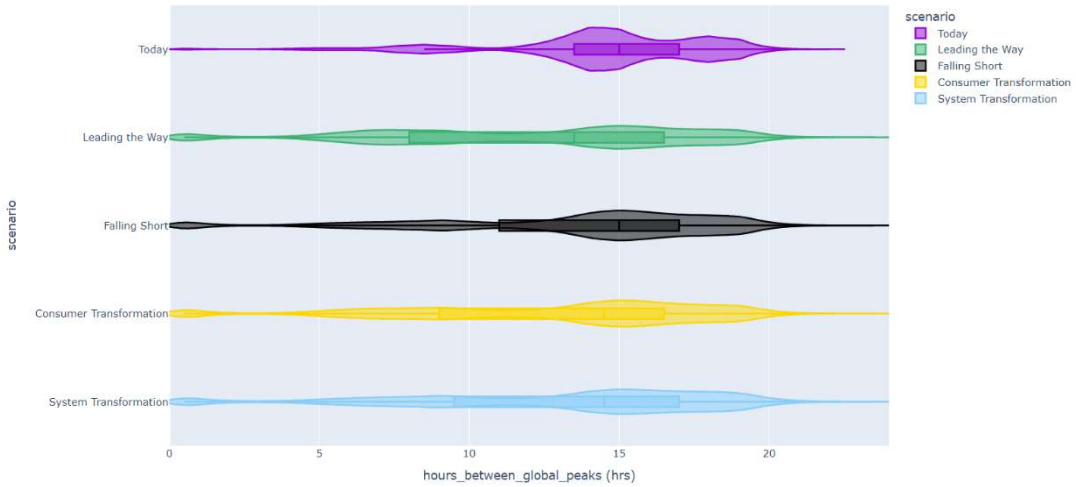


Figure 65: Violin plots of distribution of duration between peak upwards and downwards need for FES 2035

Figure 65 shows how long energy is retained for in each FES in 2035, with similar distributions for all future scenarios.

3.3 Summary of all metrics

Table 7 presents summary statistics for key quantiles (including extreme values), for numeric metrics and for key scenarios. The key extreme quantiles of interest identified as part of the project were 1 day per year, 3 hours per year and 1 in 20 years. The metrics reported below are all daily statistics, so while an extreme value with the same numeric probability as a 3 hours per year event can be calculated, is has been excluded as this way of thinking in hours per year does not directly translate to daily data. However, this metric has still been implemented and can be calculated for the half-hourly WDFN itself.

Table 8 presents the mean values (overall proportions) for Boolean metrics, such as whether the initial need at the start of the day is upwards.

Table 7: Summary values at key quantiles for all numeric metrics.

Metric	Scenario														
	2023					2035 - Leading the Way					2035 - Falling Short				
	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs
Peak upwards WDFN (GW)	5.36	8.99	13.72	17.42	20.2	14.04	25.59	45.05	60.36	71.6	7.46	14.16	24.92	33.48	39.87
Peak downwards WDFN (GW)	5.42	9.97	19.47	19.47	22.09	13.06	23.08	42.14	57.16	66.54	7.12	14.21	26.09	34.66	40.05
Difference in magnitude peak upwards & downwards (GW)	0.16	1.77	5.83	9.81	12.35	0.53	5.68	17.82	30.51	41.79	0.27	2.92	10.07	18.11	24.76

Metric	Scenario														
	2023					2035 - Leading the Way					2035 - Falling Short				
	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs
Half-hourly peak upwards ramp (MW/min)	73.4	112.8	168.2	215.8	260.6	181.1	292.8	482.6	741.4	1167	100.6	162.1	259.9	389.1	594.5
Half-hourly peak downwards ramp (MW/min)	67.84	98.55	142.0	183.2	233.3	175.3	283.8	474.9	743.3	1165	99.05	156.7	255.5	385.0	583.5
2-hourly peak upwards ramp (MW/min)	44.16	73.79	113.8	137.2	154.2	96.56	171.2	271.0	366.1	469.7	54.46	96.85	155.7	206.8	262.8
2-hourly peak downwards ramp (MW/min)	45.40	71.56	100.8	123.7	144.2	91.31	158.2	258.2	351.8	447.2	54.13	94.66	150.9	201.7	250
4-hourly peak upwards ramp (MW/min)	30.63	50.72	76.31	92.21	104.5	65.47	128.9	206.8	273.1	332.2	36.6	70.91	116.43	151.3	179.9
4-hourly peak downwards ramp (MW/min)	31.87	55.42	78.84	96.78	112.2	56.97	113.1	193.3	264.0	323.2	33.79	70.36	118.4	158.2	191.4

Metric	Scenario														
	2023					2035 - Leading the Way					2035 - Falling Short				
	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs
Total energy above zero (GWh)	31.0	55.40	90.48	116.5	134.3	73.86	135.24	272.8	392.6	470.4	39.93	77.78	156.8	222.6	266.4
Cycle capacity above zero (GWh)	21.0	51.06	90.17	116.4	134.3	47.06	121.2	271.2	392.5	470.1	25.36	69.98	156.0	222.6	266.3
Cycle capacity below zero (GWh)	17.66	44.43	84.22	113.0	131.4	43.1	111.2	268	392.2	470.3	22.8	62.76	152.7	221.8	266
Cycle capacity as a battery (GWh)	22.39	52.04	90.25	116.5	134.2	54.24	127.0	272.0	392.5	470.3	28.54	72.4	156.4	222.6	266.3
Cumulative capacity above zero (GWh)	0	7.54	25.32	44.43	62.15	0	7.29	114.8	231.5	310.5	0	5.66	58.4	120.3	167.5
Cumulative capacity below zero (GWh)	11.9	45.06	84.53	113.1	131.6	0	10.91	268.9	392.4	470.1	0	62.28	153.1	222.1	265.7

Metric	Scenario														
	2023					2035 - Leading the Way					2035 - Falling Short				
	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs	q5	q50	q95	1 day per year	Once in 20yrs
Number of crossing points	2	4	7	7	7	1	4	8	8	8	1	4	8	8	8
Time between peak upwards & downwards (hrs)	7.5	15	19	19	19	3.5	13.5	19	19	19	4.5	15	19	19	19
Minimum time between crossing points (hrs)	0.5	1.5	3.5	3.5	3.5	0.5	1	10	11.83	12.05	0.5	1	10	11.95	12
Maximum time between crossing points (hrs)	6.5	10.5	15.5	15.5	15.5	6	9.5	15.5	15.5	15.5	6	9.5	15.5	15.5	15.5

Table 8: Summary of Boolean metrics

metric	2023	2035 - Leading the Way	2035 - Falling Short
Proportion of time peak WDFN is upwards	0.381	0.645	0.527
Proportion of time WDFN is downwards at the start of the day	0.910	0.578	0.711
Proportion of time peak upwards need occurs before peak downwards	0.290	0.263	0.284
Proportion of time max cycle capacity upwards is larger than downwards	0.798	0.687	0.737

3.4 Disaggregation of WDFN

This subsection presents results to indicate which components of energy supply and demand are driving the WDFN. This has been done by repeating the calculation process described in Section 2.3.2 for individual generation and demand profile types to give a profile of the WDFN required just for that type of generation or demand. The full set of metrics for these individual profile-level WDFNs can also be calculated.

It is important to note that while separate analysis of these profiles can be insightful, it is important not to overstate the importance of these results. In particular, the WDFNs for the individual profiles should not be added together, as this will likely overstate the WDFN for the national aggregated profile. This is due to the way in which WDFNs for different profiles will cancel each other at some points in time.

Figure 66 shows the daily profiles of the WDFN per technology, with the solid line showing the mean, the dashed lines showing the 10th and 90th percentile, and the light shaded area indicating the 98th interquantile range (between the 1st percentile and the 99th percentile). The generation profiles have been scaled by -1 for consistency with the definitions of WDFN.

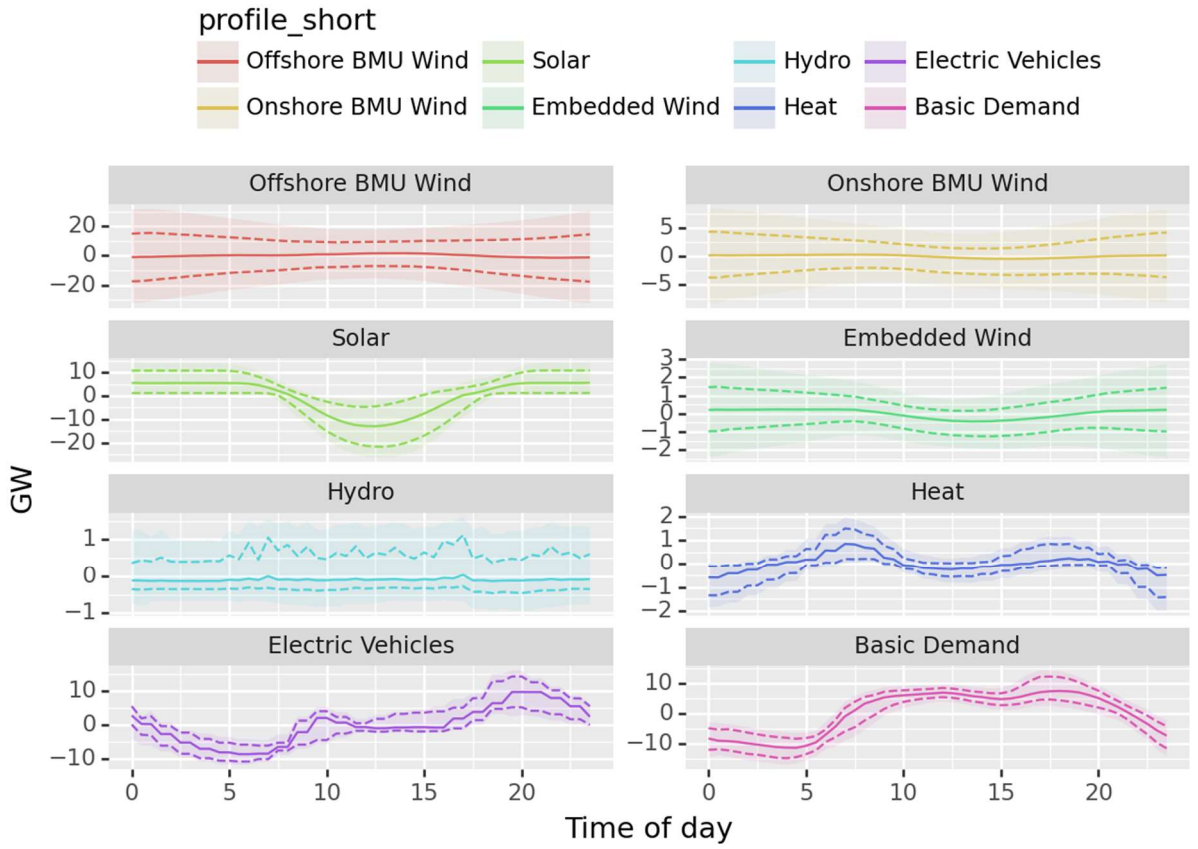


Figure 66: Average daily WDFN for individual generation and demand profiles, in Leading the Way 2035, indicating the average (solid), 10th and 90th quantiles (dashed), and 1st and 99th quantiles (shaded).

For each profile, the plots below show several summary statistics of the daily WDFN (from the average value to the 1-in-20-year extreme value), for maximum downwards and upwards power metrics in GW, and the maximum downwards and upwards ramp rates in GW per minute. These have been calculated for the 2035 Leading the Way scenario.

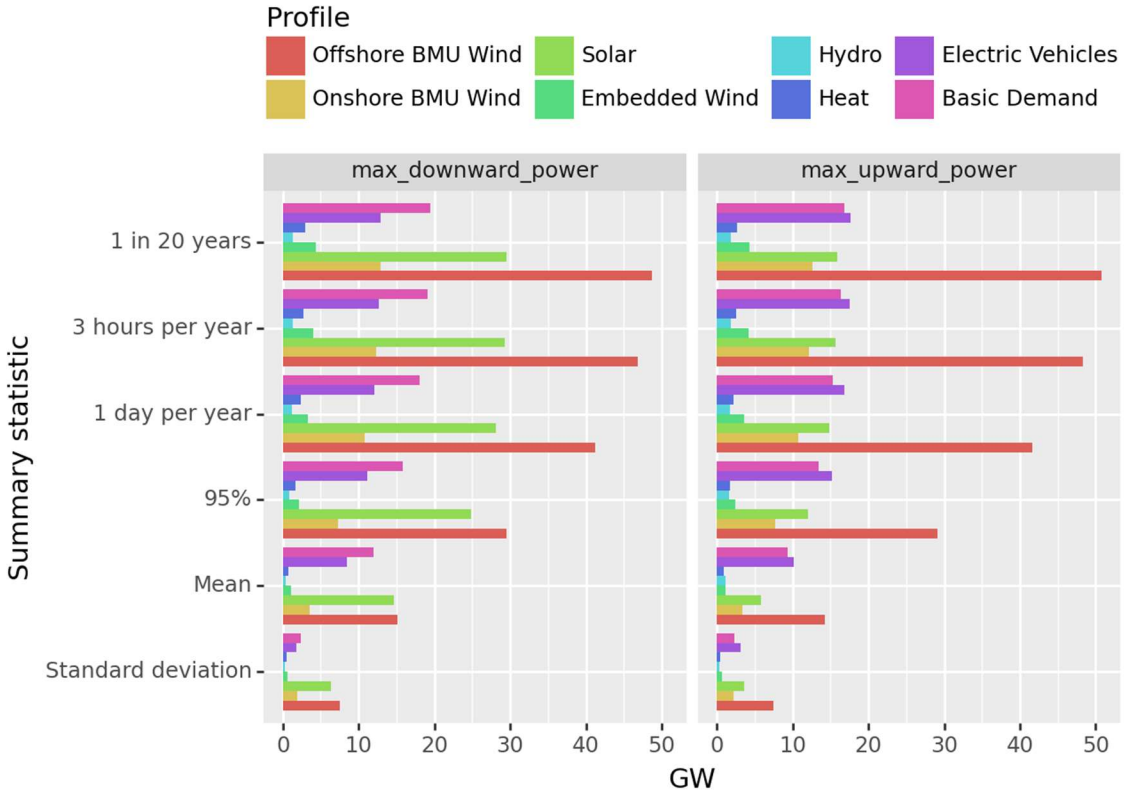


Figure 67: WDFN summary statistics for individual generation and demand profiles, upwards and downwards power, under Leading the Way 2035

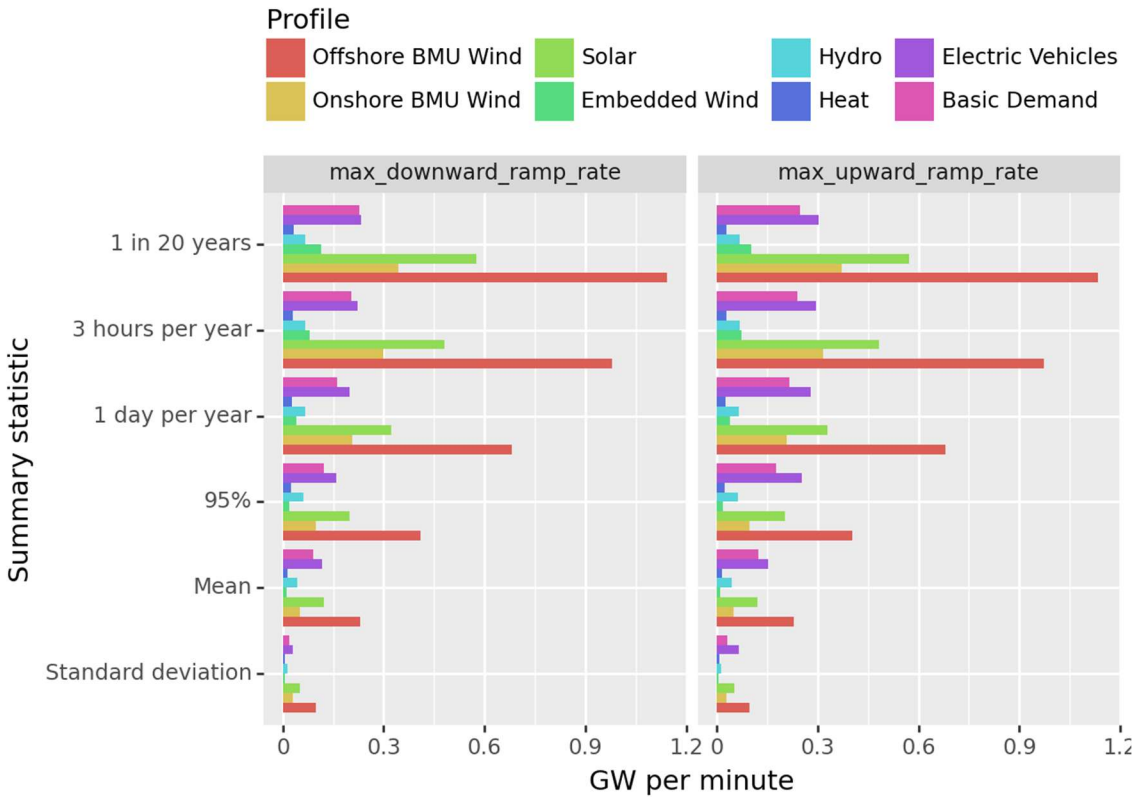


Figure 68: WDFN summary statistics for individual generation and demand profiles, upwards and downwards ramp rate, under Leading the Way 2035

These plots suggest that, for this scenario and year, offshore wind is a key driver of the WDFN. This is to be expected, given the variability of the wind resource and the very high capacity of installed offshore wind relative to other technologies in this scenario. The WDFN for maximum upwards and downwards power for offshore wind is similar in both directions, and so are its ramping requirements.

After offshore wind, the next most significant drivers of the WDFN seem to be basic demand for appliances, lighting and resistive heating, electric vehicle demand, BMU onshore wind generation, and solar PV generation. There are some interesting asymmetries evident for these profiles, particularly for solar, which contributes more to the downward WDFN than to the upward WDFN. This is because, due to zero production in overnight darkness hours, the average daily solar profile is quite close to zero, but the peak solar output can be much higher.

To illustrate the *relative* impact on WDFN, the next pair of plots divide these GW and GW per minute values by the maximum GW value of each profile (which is calculated from each profile). This shows that, for example, hydro has a big impact on the upward WDFN relative to the amount of hydro capacity – but because there isn't much hydro capacity on the system, this doesn't have a big impact at the system level.

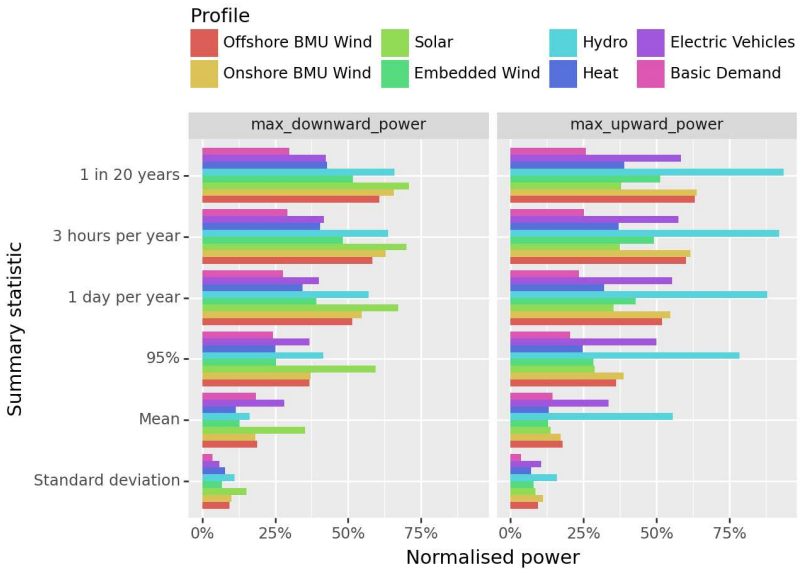


Figure 69: WDFN summary statistics for individual generation and demand profiles, upwards and downwards normalised power, under Leading the Way 2035

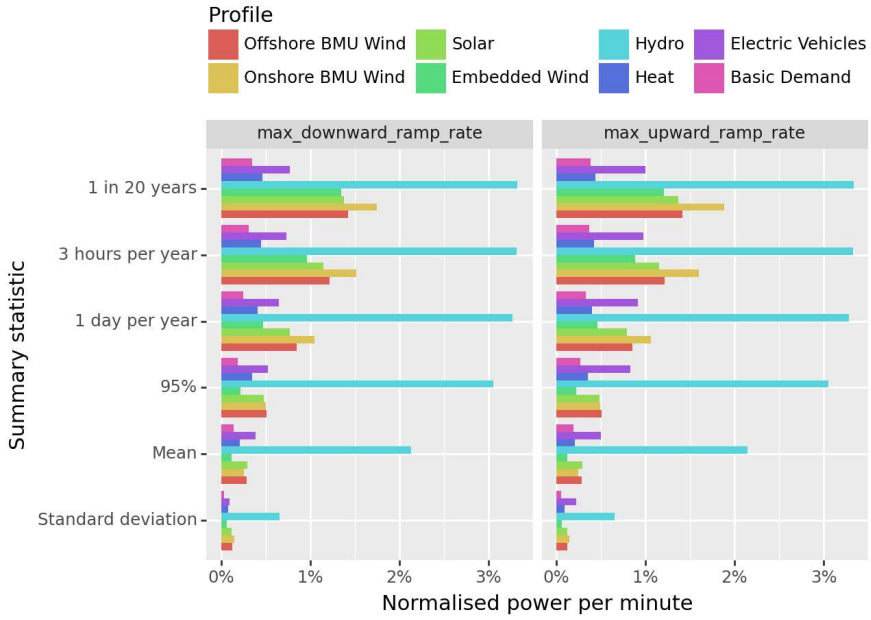


Figure 70: WDFN summary statistics for individual generation and demand profiles, upwards and downwards normalised ramp rates, under Leading the Way 2035

3.5 Change in WDFN over time

This subsection illustrates how WDFN for a subset of metrics may evolve over time, under different FES. The following plots show values for three metrics (max upward power, max downward power, and max cycle capacity), with values plotted for the average value as a solid line, and the 1-in-20-year extreme value as a dashed line. Each FES is indicated with a different colour.

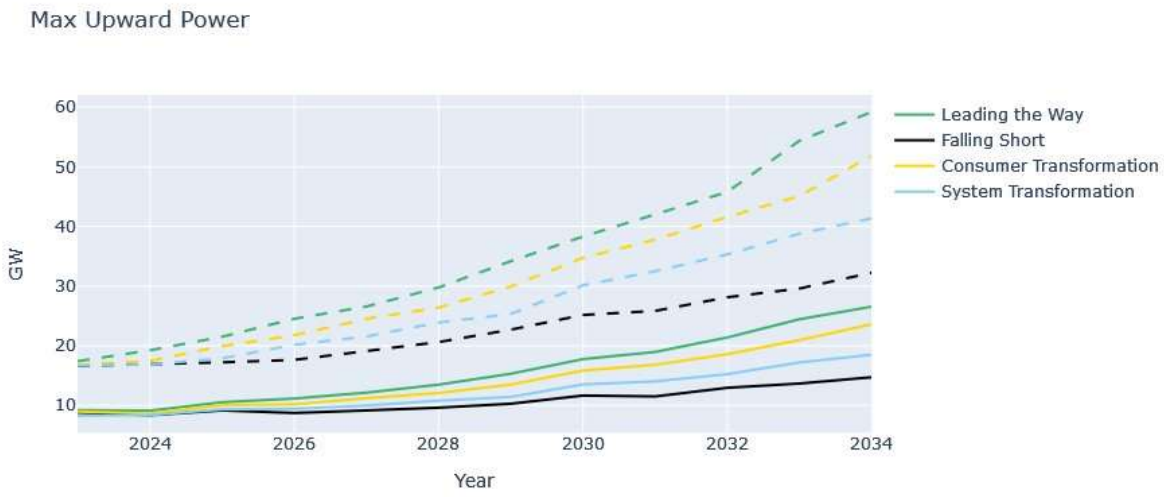


Figure 71: Year-on-year change in WDFN for max upward power under different FES, until 2035

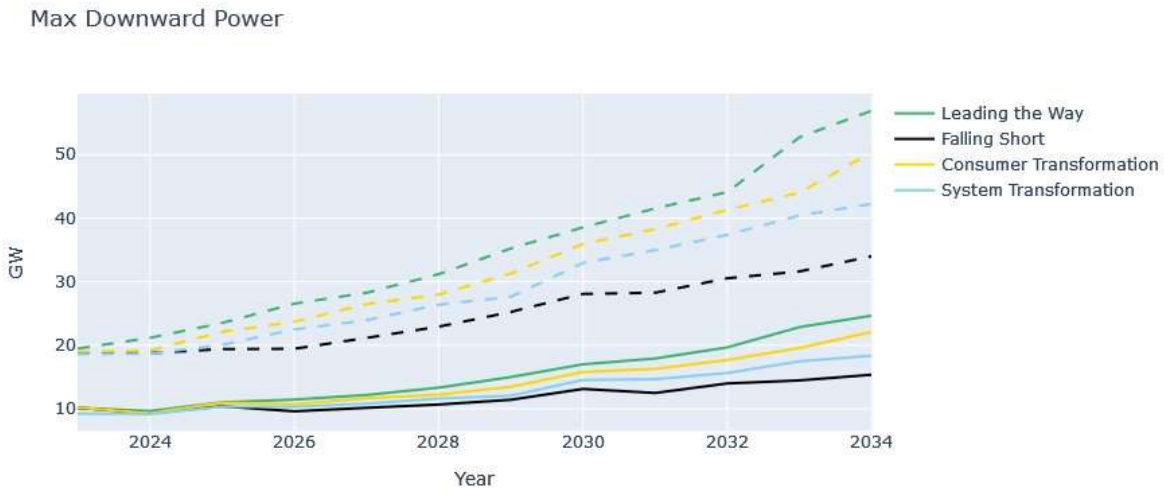


Figure 72: Year-on-year change in WDFN for max downward power under different FES, until 2035

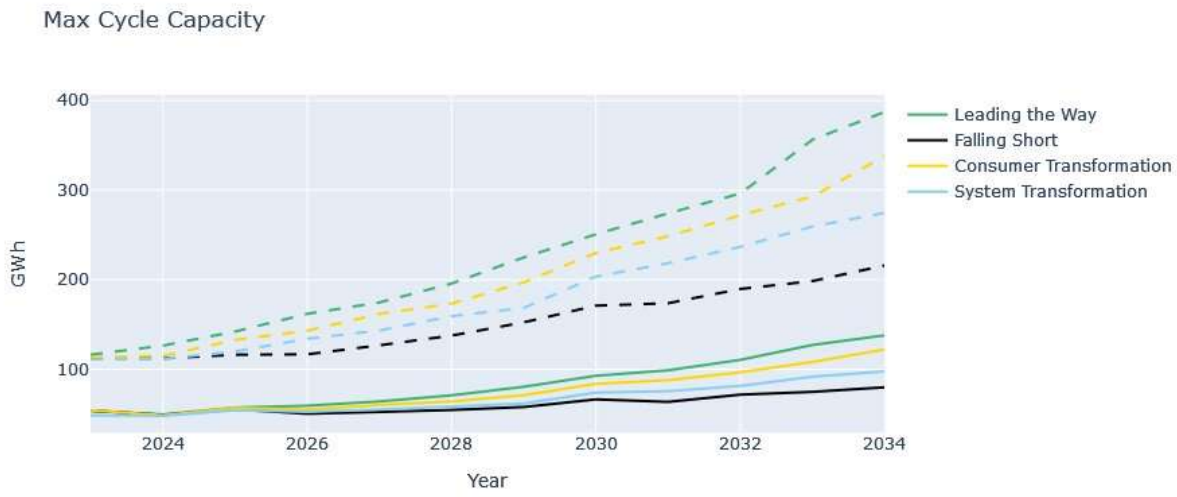


Figure 73: Year-on-year change in WDFN for max cycle capacity under different FES, until 2035

For all three of the metrics considered here, both the average daily values and the extremes diverge across the four scenarios by 2035, although the extremes diverge more and more quickly than the average values do. In every case shown here, the WDFN is greatest for the Leading the Way scenario, then the Consumer Transformation scenario, and then System Transformation. Falling Short has the smallest WDFNs by 2035, but these are still larger than the system's needs today. The change through the years is close to linear for all scenarios, metrics and probability levels, except in the very last years (from 2032) where there appears to be a small increase in the rate of change of the extremes, for the most ambitious scenarios.

These figures also show that the WDFNs will become more variable over time, as the 1-in-20-year values increase at a faster rate than the average. This is particularly true in the more ambitious decarbonisation and decentralisation scenarios (e.g., the 1-in-20-year max downward power WDFN is 10 GW greater than the mean under today's scenario assumptions, but this will be almost 30 GW higher under Leading the Way 2035 scenario assumptions). This suggests a larger volume of WDF capability that is being used less frequently, which has implications for how that WDF is delivered to the system.

4 Sensitivity Analysis

Two sensitivity analyses have been carried out, to show the effect of two main assumptions and simplifications of the method.

- Firstly, the effect of considering all EVs as completely inflexible, and an estimation of the proportion of the time where WDFN could be met by EVs, is presented.
- Secondly, an extension to model interconnector flows to connected countries is shown for the GB-Ireland flows. This method could be repeated for other interconnectors, although this increases the number of models and volume of data required.

4.1 Electric Vehicles

A critical assumption of the modelling process is that all EVs are inflexible. This assumption becomes much weaker the further into the future we look, and for more ambitious FES. The treatment of all EVs as inflexible was chosen as it was deemed preferable to the alternative, where EVs are modelled as *completely* flexible and able to charge by any amount at any time of day in response to the WDFN. Ideally, a profile or profiles for different flexible behaviours would be known and could be included in the modelling; however, this is not currently possible.

Instead, we present a sensitivity analysis on the EV component of the WDFN calculations to indicate the possible magnitude of their contribution. It should be noted that these results are rough estimates of the effect to indicate the order of magnitude of inflexible EVs' contribution to the need. The Leading the Way 2035 scenario has been used for this demonstration.

First, we have calculated WDFN for the complete set of times, without inclusion of any EVs. This represents the case where EVs are not contributing to the need at all, *but also not actively working to reduce the need.* In this case, either there are no EVs, or they are all charging at a constant rate through the day. Figure 74 shows that the overall distribution with no EVs is more symmetrical, while inclusion of EVs introduces a right skew (even though the daily mean WDFN is still 0) and greater spread of WDFN. Broadly, the inclusion of naïvely charging EVs increases the peak WDF need, both upwards and downwards as can be seen in Figure 75: the median value for peak WDFN is 35% higher upwards with naïve EVs, and 15% higher downwards. The distribution of time between peak upwards and downwards WDFN is multimodal both with and without naively charging EVs, but including naively charging EVs increases the median value of this time.

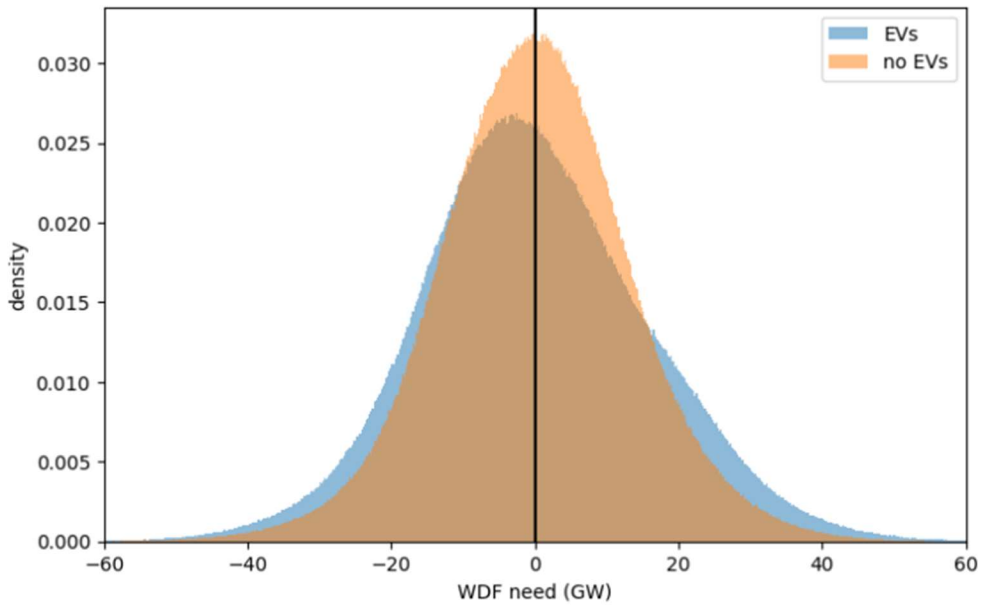


Figure 74: Comparison of the distribution of WDFN with and without EVs for Leading the Way 2035

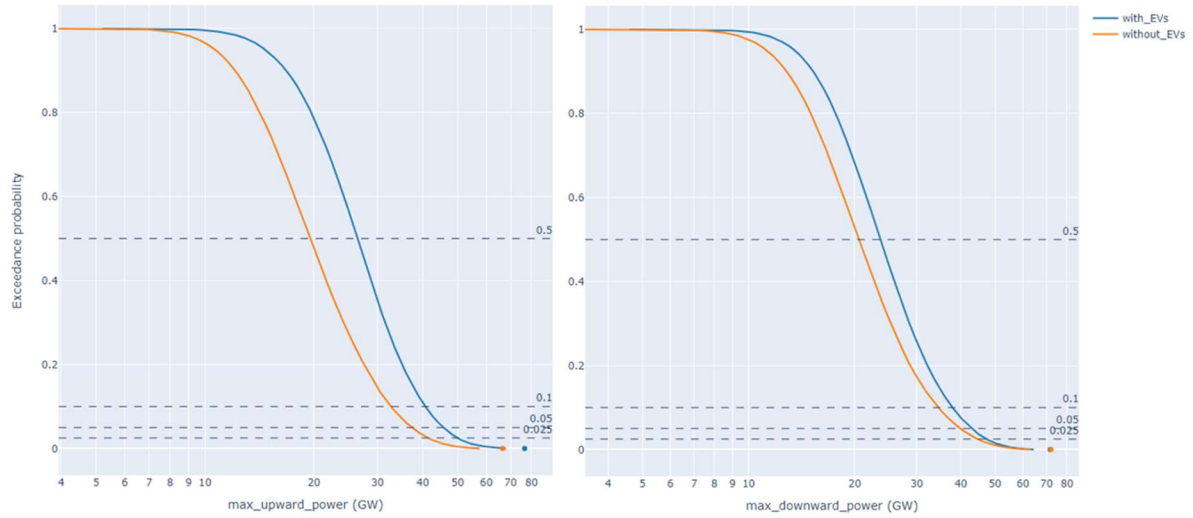


Figure 75: Distribution of peak WDF need with and without naively charging EVs for Leading the Way 2035

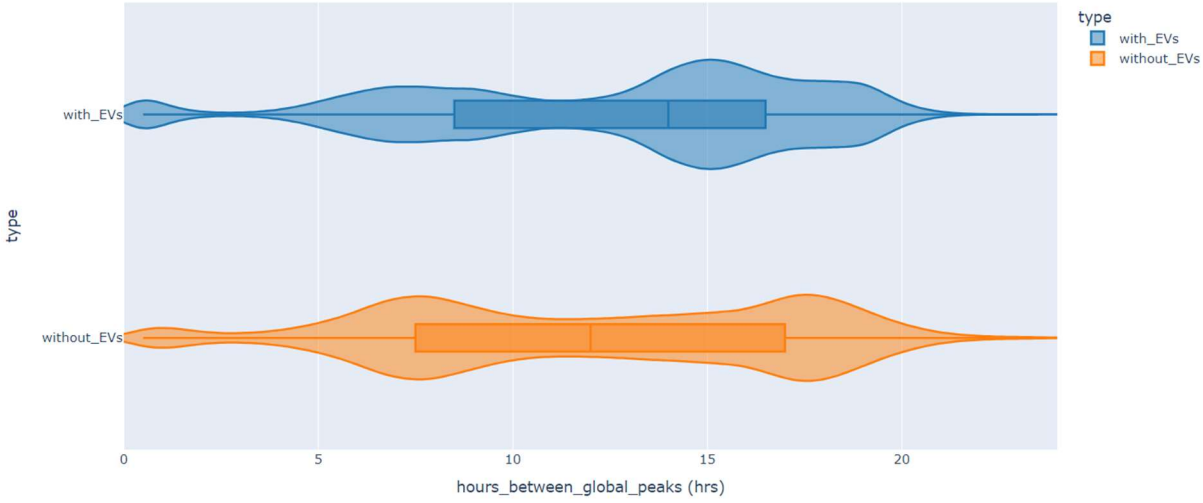


Figure 76: Distribution of time between global peaks with and without naively charging EVs

While this work has focussed on calculating the WDF need, EVs are likely to also be an important WDF solution as well as driver of need. If individual EV plugin times and trip charging demands were available, the question of how much of the WDFN flexible EVs could cover (relatively) easily could be answered through a simple optimisation problem, with the goal of flattening the WDFN profile. However, having the data to conduct this is a barrier to this more detailed analysis at this stage.

4.2 GB-Ireland interconnectors

The analysis presented so far has only considered the drivers of the need for within-day flexibility from demand and generation in Great Britain. However, the GB energy system is heavily interconnected with other European energy systems and is widely expected to become more heavily interconnected in the future. These interconnectors can be used by sources of within-day flexibility from other European energy systems to provide their flexibility to GB, or by baseload generators to provide baseload to GB. However, there is also the prospect that the needs for within-day flexibility in neighbouring energy systems can actually cancel each other out.

This has been examined through a sensitivity which considers future interconnection between GB and Ireland's All Island energy system in the year 2035, under a Leading the Way scenario. The rescaling models have been re-run for historic Irish wind output and total system demand (sourced from EirGrid) but, for simplicity, the GB solar, EV and heat pump profiles have been used for the Irish system. ERA5 data back to 1996 has been sourced and processed for Ireland, meaning there are somewhat fewer year combinations used for this sensitivity analysis. The total capacity of wind in Ireland has been sourced from the FES. For solar PV, the *additional* capacity added since 2024 is taken from the FES, since the EirGrid data on demand already reflects the netting off of existing solar generation. Scenarios for the number of EVs on the road and the annual TWh demand of heat pumps have been sourced from EirGrid's Tomorrow's Energy Scenarios[13] document. For simplicity, it is assumed that the annual MWh consumption per EV is the same in Ireland as it is in GB and that data centre load is constant throughout the day, meaning it does not affect the need for within-day flexibility.

[13] https://cms.eirgrid.ie/sites/default/files/publications/Tomorrows-Energy-Scenarios-2023-Consultation-Report-November-2023_0.pdf

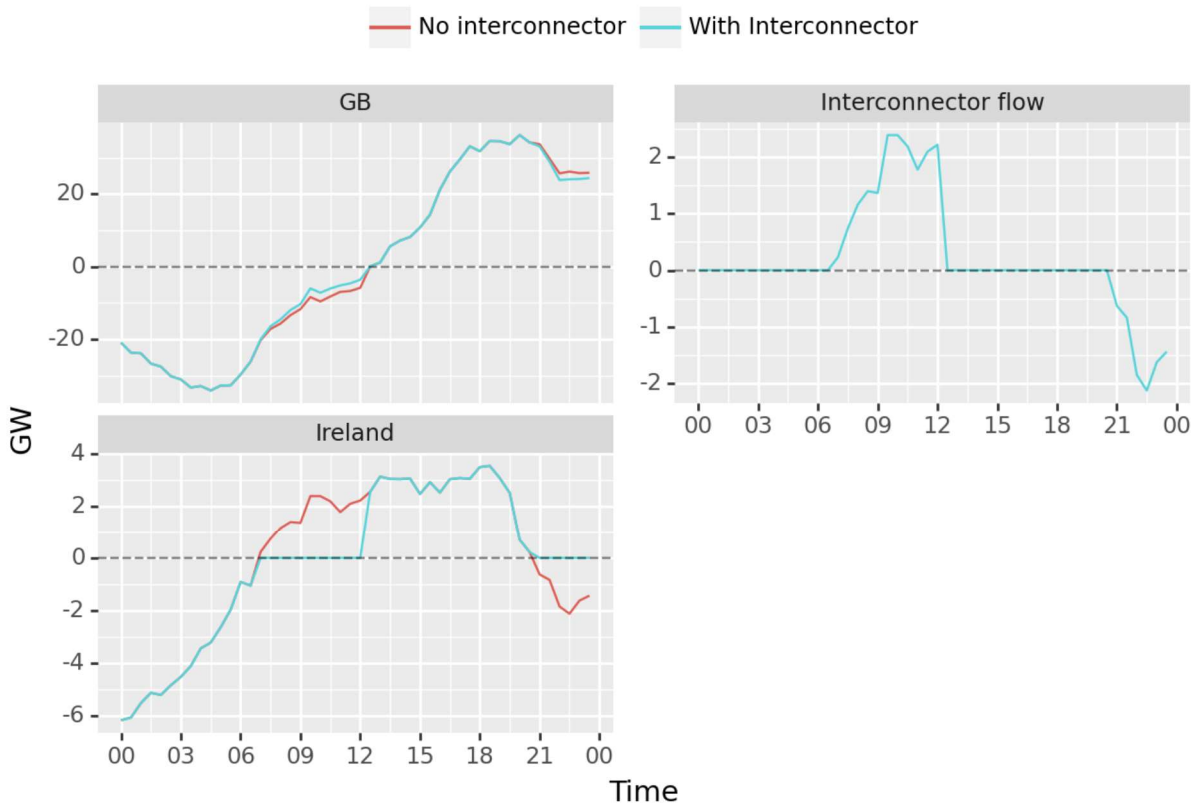


Figure 77: An example day showing WDFN profiles for GB and Ireland, and the resulting interconnector flows.

These profiles are combined to produce total inflexible demand and total inflexible renewable generation profiles for Ireland, from which the requirement for within-day flexibility can be estimated. The interconnection capacity between GB and Northern Ireland and Ireland is assumed to be 3GW. For simplicity, all inefficiencies and losses are ignored.

Figure 77 provides an example of how the interconnectors are assumed to operate. On this example day, the GB system has a downwards WDFN (and so requires additional demand or reduced generation to counteract this) overnight and into the morning, and vice versa in the afternoon and evening. The Irish system's need is in the same direction as GB's from midnight to 6am and from midday to 9pm, but there are windows from 6am-midday and 9pm-midnight where the two systems have opposite needs, resulting in interconnector flows. This plot shows how the WDF need for GB changes due to the interconnector, but Ireland's WDF need will also reduce at times with interconnector flow, as non-zero flows only occur in this analysis when they would reduce the WDFN in both countries.

When both systems have the same direction of WDF need, no assumption is made about how the interconnectors will operate. This is what happens between midnight and 6am, and then again between midday and 9pm. Between 6am and around Midday, Ireland has a need for *upwards* WDF (too much demand), while GB has a need for *downwards* WDF (too much generation). These needs can offset each other, although this is limited by the relative magnitude of both requirements and the size of the interconnector. The interconnector is therefore used to export generation from GB into Ireland. After 9pm, the opposite situation occurs, where there is insufficient generation in GB and too much in Ireland. The flow direction is opposite, with generation being exported into GB.

Ireland WDF Direction			
	Upward	Export excess generation from Ireland to GB	No flow
	Downward	No flow	Import excess generation into GB from Ireland
		Downward	Upward
		GB WDF Direction	

Figure 78: Diagram of interconnector flow depending on WDFN direction in both GB and Ireland

Of course, in practice, flexible sources in one system might use the interconnector to provide flexibility to another, but this is beyond the scope of what can be considered in this project. There is an *implicit* assumption made that interconnectors are not used to schedule baseload, meaning their entire capacity can be used for offsetting WDF needs (if possible). Alternatively, this assumption could be interpreted as meaning that any other flows scheduled over the interconnectors are fully flexible within day.

Figure 79 shows the proportion of time for which the GB and Ireland systems have different directions of WDFN. 61% of the time, the GB and Irish systems have the same direction of need. Around 22% of the time, GB needs upwards WDF while Ireland needs downward WDF, allowing the two needs to balance by exporting generation from GB to Ireland. The reverse is true about 17% of the time.

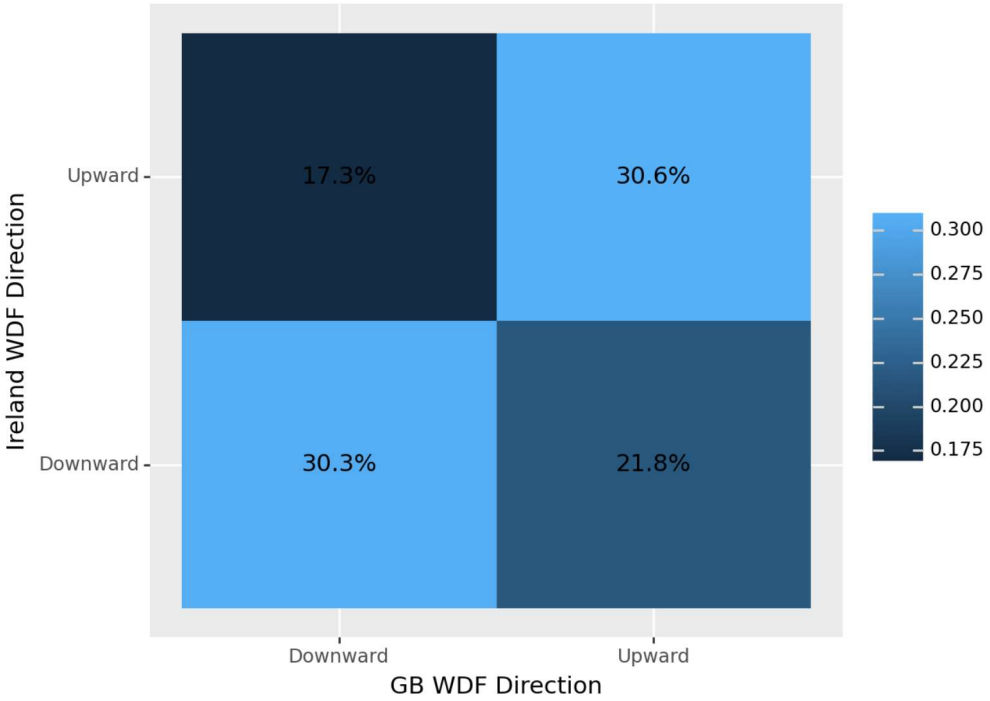


Figure 79: Proportion of the time GB and Ireland WDFN are in each direction

This approach is used to determine the flow over the interconnectors for every timestep in the analysis. 61% of the time, there is no flow (or, at least, there is no scheduled change in the flow during the day, or any flow would be due to sources of within-day flexibility). The flow for the remaining 39%

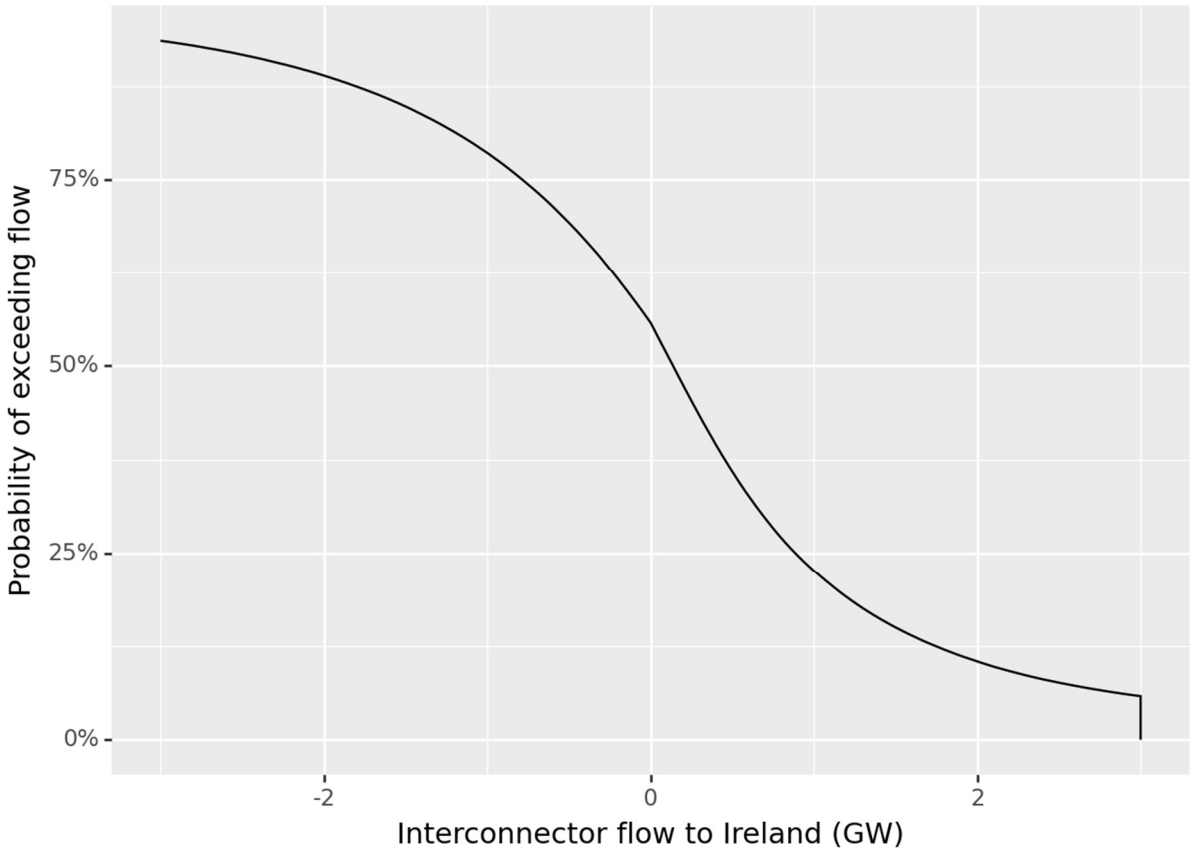


Figure 80: Distribution of the interconnector power flow to Ireland across the times used in the sensitivity analysis

Daily WDF metrics are calculated, and then means, 95th quantile and extreme values of these as well as standard deviations. In every case, the presence of the interconnectors reduces the value of the metric at all probabilities, but somewhat increases the variability of the metric (as reflected in the standard deviation). The typical changes are modest compared to the capacity of the interconnectors (perhaps reflecting the fact that 61% of the time they are not employed), and very modest compared to the size of the GB requirements (which are on the order of 10s of GW). Apart from the most extreme probability (of 1-in-20 years), the impact of the interconnector for upwards flexibility is greater than for downwards, which perhaps reflects the fact that Ireland has a downwards need while GB has an upwards need *more often* than the other way around.

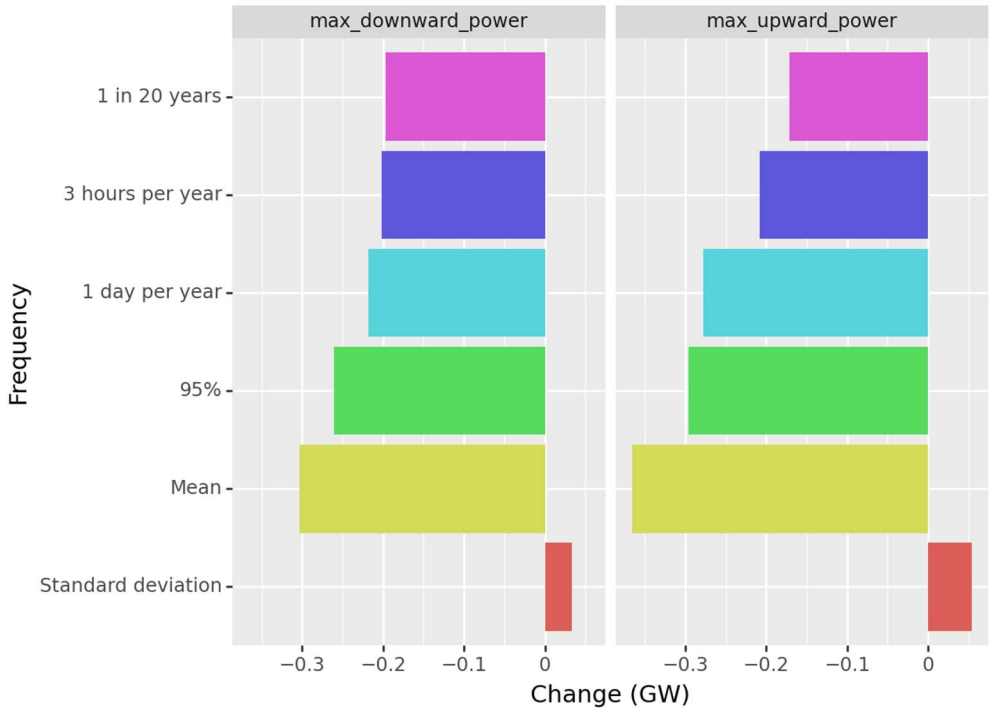


Figure 81: Summary probability levels for change in max upwards and downwards power metrics when GB-Ireland interconnection is included (compared to the values of these metrics without interconnection)

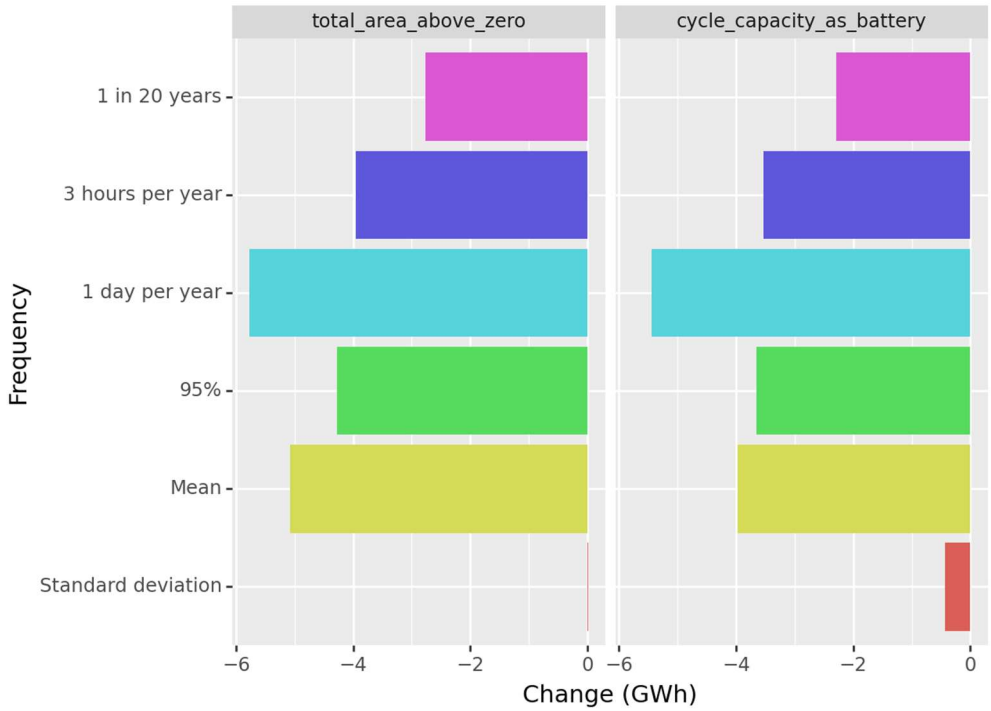


Figure 82: Summary probability levels for change in total energy and battery cycle capacity metrics when GB-Ireland interconnection is included (compared to the values of these metrics without interconnection)

5 Discussion

5.1 Key results

We have presented a methodology for modelling a large set of years of demands and generation to capture weather and behavioural variability. From this we calculated the WDFN as the half-hourly imbalance between demand and generation, assuming overall energy balance for the day. A set of metrics describing properties of the daily WDFN have been presented, including peak WDFN, ramp rates, indication of total energy need across multiple time points and a description of daily profile shape. Returning probabilistic results through distributions of daily metrics allows consideration of not only the average value but also the uncertainty and variability present across days, and quantification of WDFN levels corresponding to defined frequency levels including extremes.

The results presented show some clear trends across all metrics. In general:

- Leading the Way shows the most extreme levels of WDFN of all FES (although this is assuming all EVs and heat pumps are inflexible), including the highest peaks, ramp rates and energy requirements.
- All FES in 2035 show increased levels of WDFN across metrics compared to the present (2023).
- Differences between FES are greater than seasonal differences, suggesting differences in future technology uptake and installed renewable capacities have greater effect on WDFN than weather and seasonal behaviour.
- Many of the numeric metrics have long tails, suggesting the most infrequent but largest events result in a significantly higher WDFN than typical times.
- Peak upwards and downwards WDFN for the same day are generally of similar magnitude, with a difference less than 4% of the range 95% of the time. Peak WDFN is slightly more frequently upwards for all FES in 2035 except Falling Short.

This work has focussed on calculating the *need* for within-day flexibility, as a first step in considering current and future levels. One aim of this work is to allow future development and research in within-day flexibility, including consideration of technologies able to provide *solutions* to the WDFN.

5.2 Future work

These calculations were developed as a first-pass model chain to quantify WDFN and were built with the intention of allowing further development in the future. During this project, several avenues for future work have been identified, including but not limited to:

- Modelling spatially disaggregated need (not just national level, but by, for example, Scotland vs rest of GB, GSP group). This would increase the computational complexity as a separate set of rescaled time series would have to be fitted per area. Data on offshore wind connection points and other locational information necessary may not be publicly available.
- In this project, the scope was limited to a calculation of the WDF need. While there has been some consideration of how to treat components that might also provide WDF solutions, further work on how this matches up to flexibility solutions is necessary and would itself require research on expected profiles and behaviour of flexible assets.

- This work is based on historic weather data, albeit rescaled to represent current demand and generation levels and patterns. This could be extended to consider future climate projects data, although this would mainly benefit results looking further into the future than 2035, and the results of this could depend heavily on the emissions pathways considered.
- Similarly, there is ongoing research looking at correcting historic extreme events to reflect the current or future climate. This work is not mature enough to incorporate yet but may be of future interest.
- Refinement and further investigation of the behaviour of individual component models, in particular heat and transport demands. As more complete datasets before available for new technologies, these component models may be updated, either by replacing the input profiles or by introducing more complex modelling if the data allows.
- Consideration of the most meaningful and intuitive extreme value statistics, and the distribution through time of these extremes. For example, for a 1-in-10-year threshold, does it matter whether all 4 events are clustered in the same year in a 40-year period? The answer to this question determines whether one tail model should be fitted to all data at once, or to individual years separately: our implementation allows both approaches, but only the overall extreme quantiles have been given in the report. Fitting separate models to each year might require different approaches for partially pooling data (such as random effects models or Bayesian hierarchical models) to overcome the small sample size when looking at daily metrics (where there are only 18 days beyond the 95^{th} quantile).
- Similarity and clustering of days of similar WDFN. This work has treated all days as completely independent from each other, but it is likely there are relationships between adjacent days (correlations) that could affect the ability of flexibility sources to deliver consistently.
- Persistence of levels of WDF. By defining a fixed cutoff for each day, some persistently high or low periods of WDFN that run from the end of one day into the next could be missed. This is natural consequence of calculating the energy imbalance of each day independently, but its effect should be considered.

6 Glossary

BOA	Bid-Offer Acceptance (volume)
CDF	Cumulative Distribution Function
CV	Cross-validation
ECMWF	European Centre for Medium-range Weather Forecasting
ERA5	ECMWF ReAnalysis 5
EVT	Extreme Value Theory
FES	Future Energy Scenarios
GBM	Gradient-Boosting Machine
GPD	Generalised Pareto Distribution
LCT	Low-Carbon Technology
V2G	Vehicle-to-grid
WDFN	Within-day Flexibility Need

Appendix A: Model hyperparameters

All LGBM models share the same set of reasonable hyperparameters which have been calibrated for the WDF problem, as detailed in Table 9 below.

Table 9: Hyperparameters used in lightGBM fitting

LightGBM parameter	Description	Value
objective	Specifies quantile regression. Used alongside alpha	'quantile'
alpha	Probability level to predict for	Separate models fitted for probability levels from 0.05 to 0.95 in increments of 0.05
n_estimators	Number of trees fitted for each model	1000
learning_rate	Steepness of gradient between trees.	0.01
min_data_in_leaf	Minimum allowed number of observations in the final split of each tree; prevents overfitting.	30
bagging_freq and bagging_fraction	Bagging is the process of using only a random subset of the whole training dataset to fit each tree. This is also designed to prevent overfitting.	bagging_freq=1 (always used); bagging_fraction=0.9
force_col_wise	Recommended to reduce memory, increase fitting speed, and avoid instabilities when many features are supplied.	True

Appendix B: Feature Importance

The following graphs illustrate, for each model, how important the various features are.

BMU Wind Model

Figure 83 depicts average feature importance of the BMU models across all probability levels. Features labelled '_n' are those variables at the nth largest capacity site, and e.g. 'ws10_6_lag' is the 6-hour lagging value for 10m wind speed. 'ws' denotes wind speed, and 'wd' wind direction.

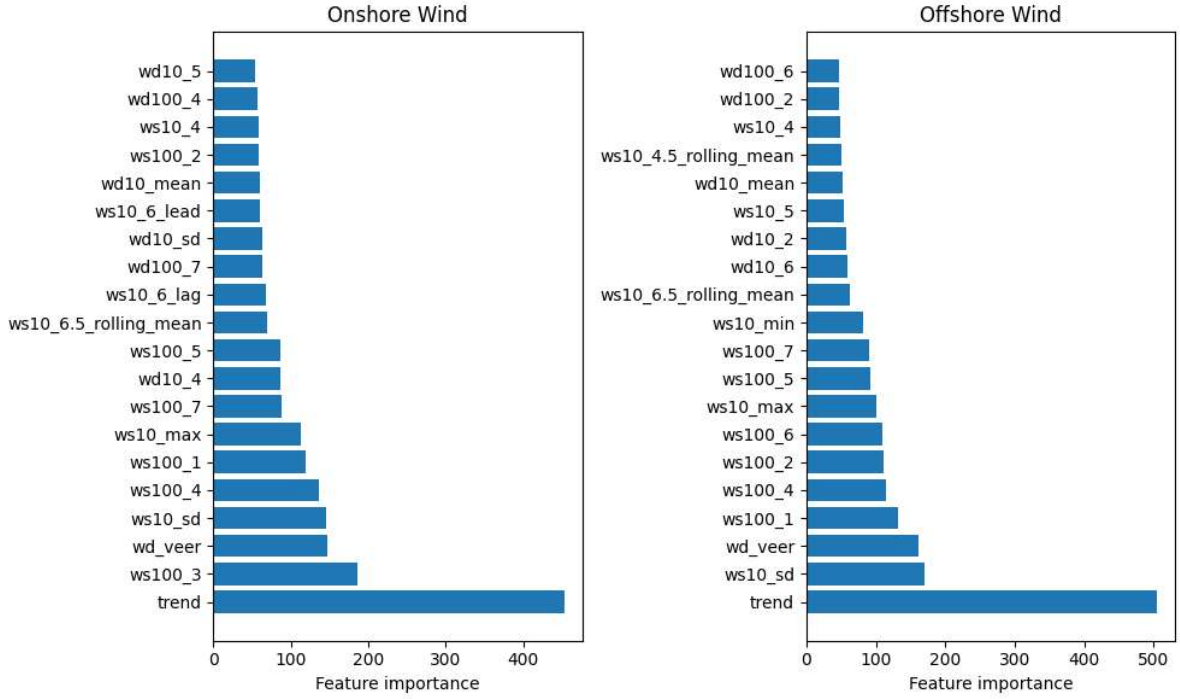


Figure 83: Average feature importance across all quantiles for the BMU wind model

Embedded wind

Figure 84 depicts average feature importance of the Embedded wind LGBM model across all probability levels. The term 'ws' denotes wind speed, and 'wd' wind direction, and e.g. 'ws10_6_lag' is the 6-hour lagging value for 10m wind speed.

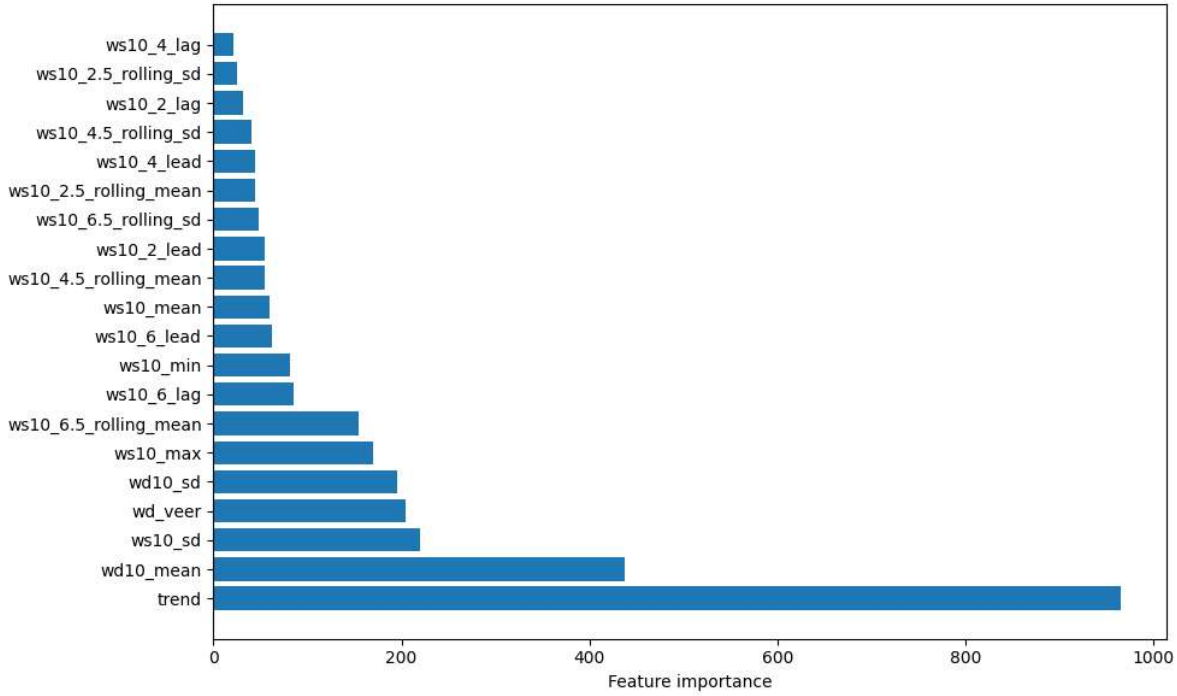


Figure 84 Average feature importance across all quantiles for embedded wind model

Solar

Figure 85 depicts average feature importance of the Embedded solar LGBM model across all probability levels. The terms 'hcc'/'mcc'/'lcc' denotes high/medium/low-altitude cloud cover, and 'ssrd' denotes surface solar radiation downwards. The acronym 'doy' is day of year, 'tod' is time of day and e.g. 'ssrd_mean_6_lag' is the 6-hour lagged value of national average ssrd.

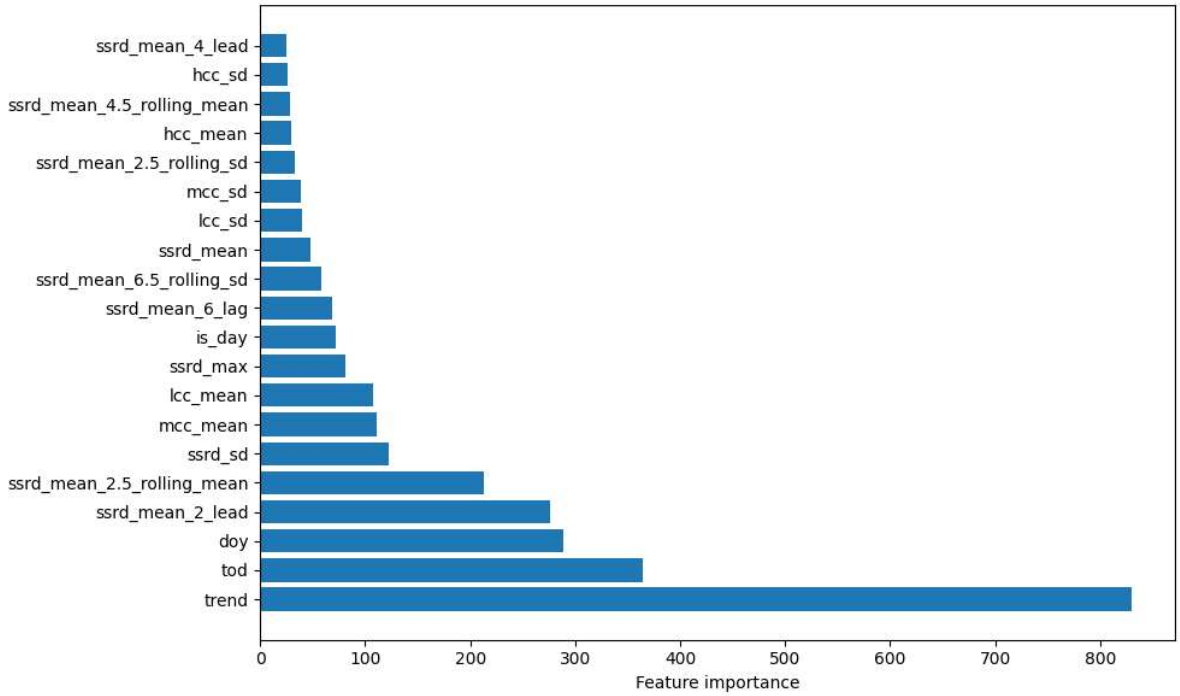


Figure 85: Average feature importance across all quantiles for solar model

BMU Hydro

Figure 86 depicts average feature importance of the BMU hydro GBM model across all probability levels. The term 'tp' is total precipitation, 't2m' is 2m temperature, 'doy' is day of year, 'woy' week of year, and 'tod' is time of day. In addition, 168- and 336-hour (7 and 14 days) rolling mean (smoothed) features make up most informative features.

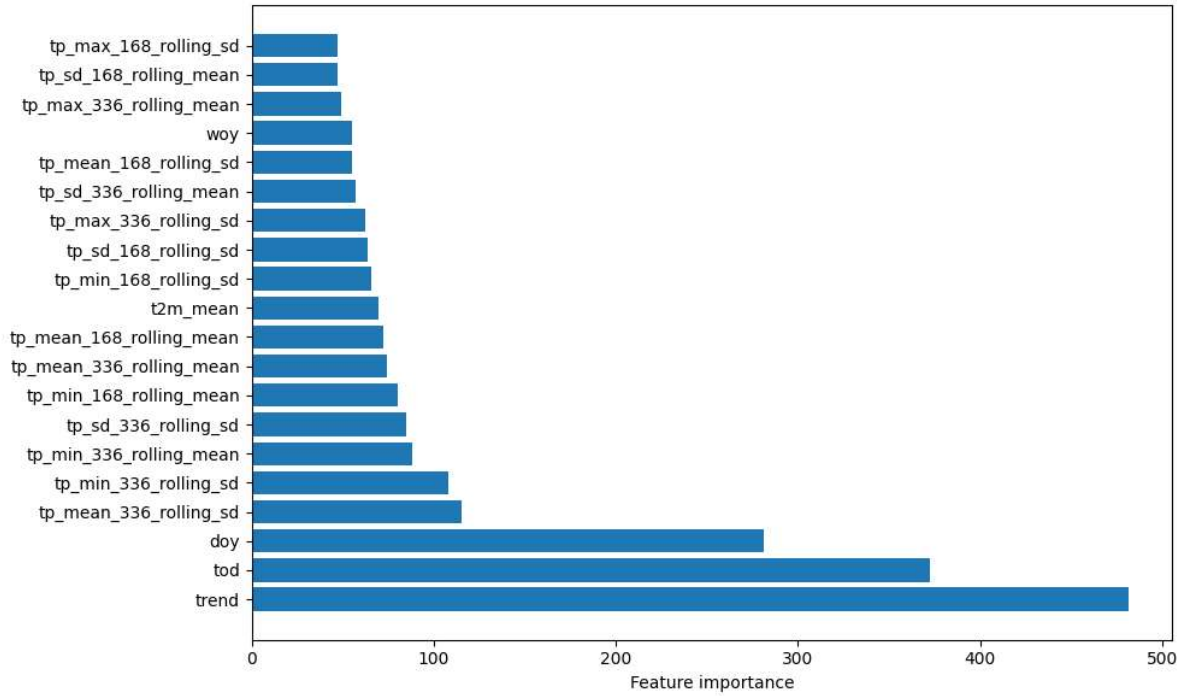


Figure 86: Average feature importance across all quantiles for non-pumped storage BMU hydro.

Demand

Figure 87 and Figure 88 demonstrate the feature importance of the demand residuals' GBM model at probability levels 0.02, and 0.6. This selection has been taken to highlight how feature importance is not uniform across probabilities.

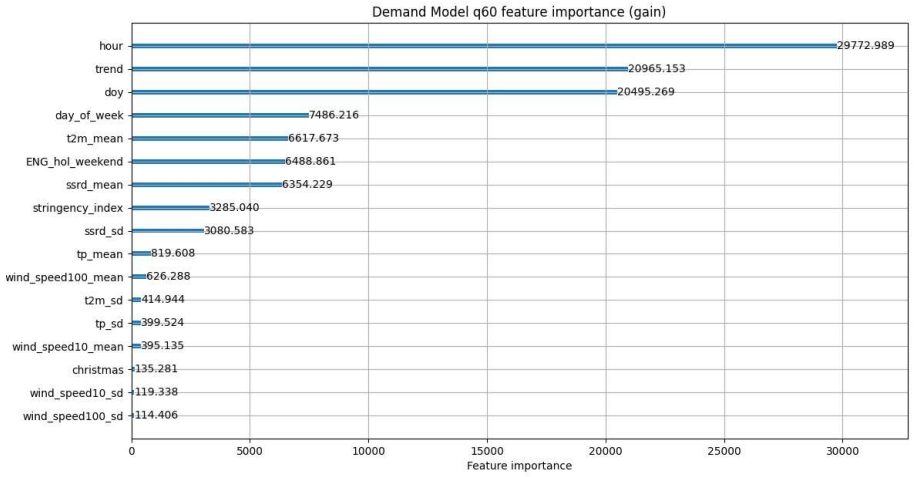


Figure 87: Feature importance in terms of accuracy gain for Quantile regressor for probability level 0.6.

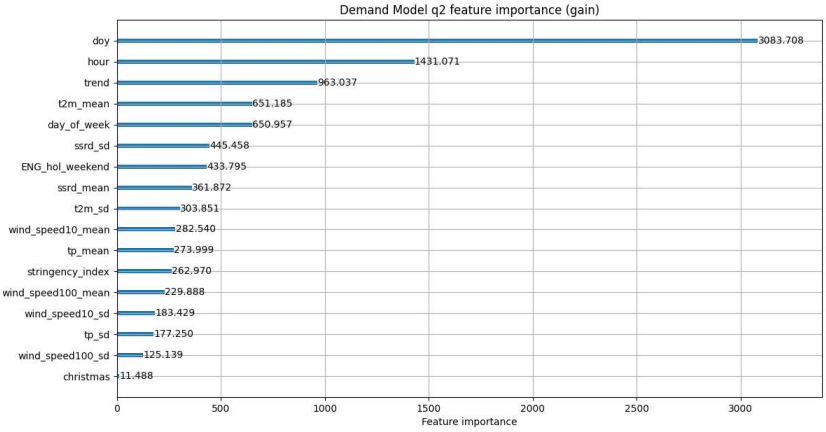


Figure 88: Feature importance in terms of accuracy gain for quantile regressor for probability level 0.02.

Appendix C: Data sources

The full WDF analysis conducted in this project requires various data sources. All data sources used in this project are openly available for public use. A description and use cases of each data source and their current provider is provided below:

All Regions

COVID Stringency Index – Oxford Coronavirus Government Response Tracker (OxCGRT)

> A composite measure of the 'strictness' of governmental response to the COVID-19 pandemic i.e., lockdowns and other measures Used as feature of demand rescaling methods to learn and isolate behavioural changes that might occurred.

ECMWF ReAnalysis 5 (ERA5) – ECMWF

> European weather dataset created using historic satellite imagery. Is compiled into region specific weighted averages. Used as features in rescaling all profiles

GB EV charging profiles – NGESO / Element Energy

> Output from ENA NIA project. Profiles of various EV charger types normalised by total annual energy consumption. Used to produce weather rescaled EV charging profiles.

Population Densities and Geolocations -

> Data describing geolocations and populations across regions. Used in producing population weighted weather data and deciding BMU asset locations.
>
> GB border and local authority district shapefiles are from the open geography portal[14]; equivalents for Ireland are from data.gov.ie.
>
> GB population densities are from the ONS[15]. Population densities for the island of Ireland combine data from the ONS for Northern Ireland, with values for the Republic of Ireland taken from projections created in 2017[16]

Great Britain Only

National Demand Data – NGESO data portal

> Half hourly report of national demand including embedded wind and solar generation capacities. Used to generate demand profile to rescale demand model from

Future Energy Scenarios – NGESO

> Basis for modelling growth of profiles using forecasted data across several years and scenarios

[14] https://open-geography-portalx-ons.hub.arcgis.com/datasets/ons::nuts-level-1-january-2018-generalised-clipped-boundaries-in-the-united-kingdom-1/explore and https://open-geography-portalx-ons.hub.arcgis.com/datasets/ons::local-authority-districts-december-2022-boundaries-uk-bfe-2/explore

[15] https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

[16] https://data.gov.ie/dataset/esri-population-projections-by-local-authority

BMU wind and solar load factors – ELEXON

Profiles of national BMU load factors across GB. Used to train wind and solar rescaling models (after adjusting for curtailment events)

BMU wind and solar derived BM (BOA) data - ELEXON

Used to adjust national BMU load factors to remove any curtailment before being used with wind and solar rescaling models.

Heat pump performance profiles – UK Data Service

Profiles of UK heat pump performance. Used to produce rescaled heat pump demand profiles[17].

The island of Ireland Only

Irish Demand and Wind Generation - EirGrid

Historical demand and wind generation data from EirGrid' open data portal. Used in demand rescaling models and to form wind load factors for wind rescaling model[18].

Irish Embedded Wind Capacities - EirGrid

Historical record of Wind capacities in Ireland. Used to generate wind load factors for rescaling Irish BMU wind profile.

Tomorrow's Energy Scenarios – EirGrid

Analysis carried about by EirGrid to forecast likely future energy scenarios, like FES. Used to scale Irish energy profiles to align with GB profiles grown with FES.

[17] https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=9050#!/details
[18] https://www.smartgriddashboard.com/#all

Appendix D: Code Structure

The code is structured as follows:

- **sensitivity analysis** – Provides functions which result in the data required to perform sensitivity analysis.
- **tests** – Provides scripts which isolates and run sub-processes to test and assess performance
- **QWDF** – Provides all processing and utility code.
    - **calcWDF** – provides scripts used to calculate within day flexibility and other descriptions from a given profile.
    - **data** – Provides structure for all data sources and supporting data processing functions for parsing data to system compatible format.
    - **io_** - utility functions for performing input/output operations.
    - **metrics** – provides functions for calculating and plotting daily metrics.
    - **modelling** – provides Classes used for modelling rescaling of each profile type, also contains supporting functions used each rescaling class.
    - **util** – provides all shared generic utility functions.
    - **conf.py** – exposes variables used by the system which may be changed before running an analysis.
    - **const.py** – exposes constants used throughout the codebase.
    - **main.py** – Provides the key functions for co-ordinating an analysis process.
    - **__init__.py** – exposes each of the main functions which external packages may require.
- **requirements.txt** – provides list of all requirements required to run code in pip format.
- **main.py** – entry point of the software if used in isolation.

There are 4 main functions exposed by the main 'QWDF' package:

- **run_analysis()**: Runs the full analysis including calculating metrics and plotting results using given (or configured in conf.py) arguments
- **create_wdf_profile()**: runs only the process of rescaling and combining profiles to generate a single WDF profile.
- **preprocess_GB()**: Performs processing of raw ERA5 weather data to produce geospatially weighted profiles for Great Britain.
- **preprocess_IRE()**: Performs processing of raw ERA5 weather data to produce geospatially weighted profiles for Ireland.

It is important to note that data processing is separate from the QWDF calculation package, as the data processing code is specific to the data sources we have used, whereas the code within the QWDF package is intended to work with any supplied input data (if it is in the right format).

Appendix E: Python Packages

Table 10 describes each of the core python packages required to run the developed codes core features. It does not include any tertiary packages used for testing, debugging, or required by one of the packages.

Table 10: Core python packages

Name	Purpose	Version
Pandas	Data analysis and manipulation. Provides 'Dataframe' data structure which used throughout the code	2.0.3
joblib	Provides low level system for caching function results, improves performance at cost to required storage space	1.3.2
Scipy	Provides implementation of all used statistical models i.e., Generalized pareto Distributions	1.11.4
Pygam	Provides framework for developing generative additive models	0.9.0
Lightgbm	Provides for developing gradient boosted machines	4.1.0
Geopandas	Extension of pandas which integrates powerful geographic information processing	0.13.1