Validating the long-term electricity market model ElecSim using genetic algorithms

Alexander J. M. Kell, A. Stephen McGough, Matthew Forshaw

School of Computing, Newcastle University, Newcastle-upon-Tyne, United Kingdom

Abstract

Keywords: Long-Term Energy Modelling, Model Validation, Machine learning, Optimization, Climate Change

1. Introduction

To limit the effects of climate change, a transition from a fossil-fuel based energy system to one based on low-carbon, renewable energy is required. The report by the Intergovernmental Panel on Climate Change detailed that reaching and sustaining zero global anthropogenic CO_2 would halt anthropogenic global warming on multi-decadal time scales [1].

The Paris Agreement was a declaration signed in 2015 by 195 state parties to plan and regularly report on the contribution made to mitigate global warming [2]. Based on this commitment, policy makers require quantitive advice on interventions to aid in the mitigation of climate change and limit global average temperatures to well below 2°C.

The decarbonization of electricity generation is of strategic importance for this goal due to the fact that low-carbon electricity can enable reductions in CO_2 emissions in industry, transport and building sectors [3].

However, there remain a number of uncertainties in the technological transition to a low-carbon energy supply. Examples of these uncertainties are investor behaviour, future prices of electricity generation and storage, domestic and international policy, energy efficiency and electricity demand. To successfully create effective policies an increase in understanding of these uncertainties and how they interact is required.

Energy modelling is a method that allows policy makers to increase their understanding of policy decision outcomes under a wide range of scenarios. Agent-based modelling (ABM) is a simulation technique that allows for heterogeneous agents to interact and can lead to effects on the aggregated level of the total system, a phenomenon called "emergence" [4]. Traditional models for analysing electricity systems, such as centralised optimisation models do not account for the heterogeneous nature of electricity investors and are, to some extent, based on obsolete assumptions [5].

In this paper we motivate that agent-based models are a valid way of complimenting existing models to provide advice to decision makers. We show that the model ElecSim [6] can be validated over a 5 year period, starting from the year 2013 and ending in the year 2018, with a root mean squared error of $\sim$ 0.045 and a standard deviation of $\sim$ 0.16. Similarly to Nahmmacher *et al.* we demonstrate how clustering of multiple relevant time series such as electricity demand, solar irradiance and wind speed can reduce computational time by selecting representative days [7]. However, distinctly to Nahmacher *et al.* we use a K-means clustering approach [8] as opposed to a hierarchical clustering algorithm described by Ward [9].

We use a genetic algorithm approach to find an optimal set of price curves predicted by generation companies (GenCos) that adequately model observed investment behaviour in the real-life electricity market in the United Kingdom. However, similar techniques can be employed for other countries of various sizes [6]. We are able to model the transitional dynamics of the electricity mix in the United Kingdom as shown in Figure 1, where there was an $\sim$ 88% drop in coal use, $\sim$ 44% increase in Combined Cycle Gas Turbines (CCGT), $\sim$ 111% increase in wind energy and increase in solar from near zero to $\sim$ 1250MW.

There is a desire to validate the ability of energy-models to make long-term predictions. Validation in-

Email address: a.kell2@newcastle.ac.uk (Alexander J. M. Kell)

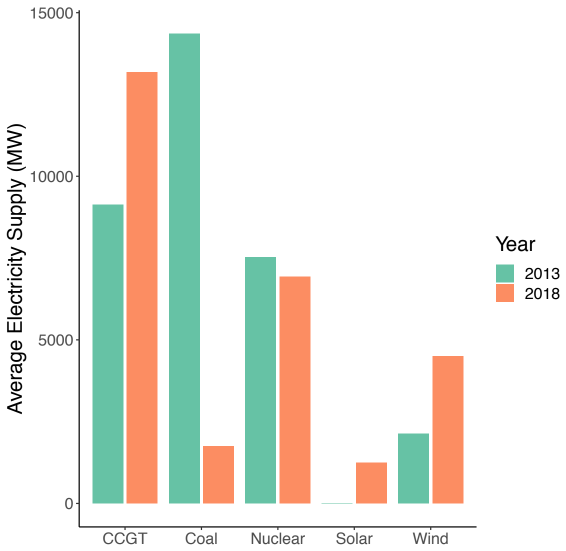

Figure 1: Electricity generation transition from 2013 to 2018 in the United Kingdom.

creases confidence in the outputs of a model and leads to an increase in trust amongst the public and policy makers. However, energy models are frequently criticised for being insufficiently validated, with the performance of models rarely checked against historical outcomes [10].

The model OSeMOSYS [11], however, is validated against the similar model MARKAL/TIMES. Whereas the model PowerACE shows that realistic prices are achieved through modelling, however do not indicate success in modelling investor behaviour [12].

However, under the definition by Hodges *et al.* [13] long-range energy forecasts are not validatable [14]. Under this definition, validatable models must be observable, exhibit constancy of structure in time, exhibit constancy across variations in conditions not specified in the model and it must be possible to collect ample data [13].

Whilst it is possible to collect data for energy models, the data covering important characteristics of energy markets are not always measured. Furthermore, the behaviour of the human population and innovation are neither constant or entirely predictable. This leads to the fact that static models cannot keep pace with global long-term evolution. Assumptions made by the modeller may be challenged in the form of unpredictable events, such as the oil shock of 1973.

This, however, does not mean that energy-modelling is not useful for providing advice in the present. A model may fail at predicting the long-term future because it has forecast an undesirable event, which lead to a change in human behaviour. Thus avoiding the original scenario that was predicted.

A retrospective study published in 2002 by Craig *et al.* focused on the ability for forecasters to accurately predict electricity demand from the 1970s [14]. They found that actual energy usage in 2000 was at the very lowest end of the forecasts, with only one exception. They found that these forecasts underestimated unmodelled shocks such as the oil crises which lead to increased energy efficiency.

Hoffman *et al.* also developed a retrospective validation of a predecessor of the current MARKAL/TIMES model named Reference Energy System [15], and the Brookhaven Energy System Optimization Model [16]. These were studies applied in the 70s and 80s to develop projections to the year 2000 . They found that the models were able to be descriptive, but were not entirely accurate in terms of predictive ability. They found that emergent behaviours in response to policy had a strong impact on forecasting accuracy. They concluded that forecasts must be expressed in highly conditioned terms [17].

Schurr *et al.* argued against predicting too far ahead in energy modelling due to the uncertainties involved [18]. However, they specify that long-term energy forecasting is useful to provide basic information on energy consumption and availability which is helpful in public debate and in guiding policy makers.

Ascher concurs with this view, and states that the most significant factor in model accuracy is the time horizon of the forecast, with the more distant the forecast target, the less accurate, due to unforeseen changes in society as a whole [19].

Work by Koomey *et al.* expresses the importance of conducting retrospective studies to help improve models [20]. For example, a model can be rerun using historical data in order to determine how much of the error in the original forecast resulted from structural problems in the model itself and how much from incorrect specification of the fundamental drivers of the forecast [20].

It is for the reasons previously described that this paper focuses on a shorter-term (5 year) horizon window when validating the model. This enabled us to increase confidence that the dynamics of the model worked without external shocks and could provide descriptive advice to stakeholders.

2. Material and methods

The model, ElecSim, is made up of five distinct sections: power plant data; scenario data; the time-steps of the algorithm; the power exchange and the investment algorithm. ElecSim has been previously published [6], however, amendments have since been made to the model in the form of efficiency improvements as well as increasing the granularity of time-steps from yearly to representative days. In this paper we have used 8 representative days for electricity demand, solar irradiance and offshore and onshore wind speed.

In this section we summarize previously published results and detail the modifications made for this paper. In this paper we initialised the model to a scenario of the United Kingdom, however, the model is generalisable to any country and is dependent on input data.

2.1. Plant Data

The simulation was initialised with every power plant and generation company (GenCo) in the United Kingdom using the Department of Business, Energy and Industrial strategy of the British government for the year 2013 [21]. The individual costs of these power plants were also initialised with data from the Department of Business, Energy and Industrial strategy of the British government [22]. For power plants that were out of the scope of this dataset (pre-2018), historical levelized cost of electricity (LCOE) values were used to infer the granular costs by a scaling factor from the International Energy Agency (IEA) [23]. Historical efficiency was also taken from the Energy Information Administration (EIA) [24]. Each of the initialised power plants were initialized with a scaling factor which modified the operation and maintenance costs stated in the dataset provided [22]. This was done to take into account differences in labor, land and breakages between projects. This was sampled from a uniform distribution between 0.3 and 2.0. As well as varying operation and maintenance costs, each of the GenCos purchased fuel at varying prices. This was done to model the element of chance and differing hedging strategies of each of the GenCos. The distribution that this was sampled from was taken from fitting an ARIMA model and sampling from the standard deviation of the residuals [25].

Financing of the project was provided by stock shares and debt, with nuclear power plants given an average weighted average cost of capital (WACC) of 0.1 and non-nuclear given a WACC of 0.059 [26, 27]. Where WACC is the rate that a company is expected to pay on average for its stock and debt.

2.2. Power Exchange

ElecSim is modelled on a uniform pricing market. This means that bids are sorted with respect to price, and accepted in merit-order. Merit-order in this case indicates the cheapest bids are accepted first. Uniform pricing is a market mechanism in which the highest accepted bid price is paid to all generators irrespective of price bid. This mechanism encourages GenCos to bid their short run marginal cost (SRMC) to ensure that their power plant is dispatched whilst not losing money whilst dispatching.

In the case of ElecSim SRMC is defined as follows:

$$SRMC = O\&M_{var} + CO_{2price} + \left(fuel_{price} \times mod_{fuel}\right) \tag{1}$$

Where $O\&M_{var}$ is the variable operating and maintenance costs, CO_{2price} is the carbon tax, $fuel_{price}$ is the cost of the respective fuel and mod_{fuel} is the fuel price modifier. These are all in the units of $£/MWh$.

Each of the GenCos submit a bid based on the SRMC of each of their plants at the start of every representative day, to model the day-ahead market.

$$Bid = SRMC \times Cap_{fac} \times Avail_{fac} \tag{2}$$

Where Cap_{fac} and $Avail_{fac}$ is the capacity and availability factors of the plants respectively. The capacity factor is defined as the actual electrical energy produced over a given time period divided by the maximum possible electrical energy is could have produced. This can be impacted by regulatory constraints, market forces and resource availability. For example, higher capacity factors are common for photovoltaics in the summer, and lower in winter [6].

$$Cap_{fac} = \frac{energy_{produced}}{energy_{max}} \tag{3}$$

Availability is the percentage of time that a power plant is able to produce electricity. This is typically reduced by outages and breakdowns. We integrate historical data to model improvements in reliability over time.

This uniform pricing market is stepped for every representative day, which, in the case of this model was 8 days.

2.3. Investment Algorithm

Investments in power plants occur at the beginning of each year. Each GenCo sequentially assesses the viability of different power plants. The order of GenCo is randomized every year to prevent certain GenCos having an advantage over others.

Investment in power plants is based upon a net present value (NPV) calculation. NPV is a summation of the present value of a series of present and future cash flow. This metric provides a method for evaluating and comparing investments with cash flows that are spread over many years.

Equation 4 is the calculation of NPV, where t is the year of the cash flow, i is the discount rate, N is total number of periods, or lifetime of power plant, and R_t is the net cash flow at time t.

$$NPV(i, N) = \sum_{t=0}^{N} \frac{R_t}{(1+t)^t} \tag{4}$$

The discount rate set by the GenCo is based upon the WACC [28]. We sample from a Gaussian distribution to adjust to adjust for varying risk profiles, opportunity costs and rates of return. Giving us sufficient variance whilst deviating from the expected price.

Future cash flow is based upon predicted earnings, which is based upon a predicted, exogenous price duration curve (PDC). A PDC is the cost of electricity with respect to hours of the year. A central part of this paper is on estimating a suitable predicted PDC that enables us to validate the model.

The SRMC cost of the power plant is calculated by fitting a linear regression to historical CO_2 and fuel price, referred to in Equation 1.

The plant with the highest NPV is chosen by each of the GenCos.

2.4. Representative days

Due to computational restrictions energy-system models often represent variations in demand, supply, solar irradiance and wind speed by using the data of a limited number of representative historical days [29].

A number of authors have shown that models with an insufficient time-step granularity leads to an underestimation of the variability of intermittent renewable energy resources (IRES). This leads to an overestimation of the uptake of IRES and an underestimation of flexible technologies [30, 31]. We exhibit the same problems in our paper [6] and later in Section 4.

To overcome this problem we used a k-means clustering approach to select representative days of electricity demand, solar irradiance and wind speed for both onshore and offshore in the UK. We used simulated data for wind and solar by Staffell *et al.* [32]. For electricity load data we used data from gridwatch.co.uk [33].

Once we had clustered all of the time series, we found the average value of each of the clusters centers, which we will refer to in this text as the "centroids" technique. In addition to this we chose the day that was closest to the centroid by euclidean distance. This will be referred to in the rest of this text as the "medoids" technique.

To measure the validity of the optimum number of days we used a technique similar to Poncelet *et al.* [34, 29]. We trialled the number of clusters against three different metrics: correlation, normalised root mean squared error and relative energy error (REE_{av}).

Firstly, the average capacity factor over the selected time series should preserve the annual electricity demand load factors. To evaluate the ability for the representative days, the average value over all the considered time series $p \in P$ compared to the actual average value is used as a metric. This is referred to as REE_{av}. Note that in this text $|\cdot|$ refers to the absolute value and $||\cdot||$ refers to the cardinality of a set. Below, the index $t \in T$ refers to a specific time step of the original time series (e.g. hourly interval).

$$REE_{av} = \frac{\sum_{p \in P} \left(\left| \frac{\sum_{t \in T} DC_{p,t} - \sum_{t \in T} \widetilde{DC}_{p,t}}{\sum_{t \in T} DC_{p,t}} \right| \right)}{||P||} \tag{5}$$

Another metric that we wanted to measure was that of the distribution of electricity demand and capacity factors to be similar to that of the actual time series. The distribution can be represented by a duration curve (DC) of the original time series. We therefore used the normalised root mean squared error between the actual duration curve and representative duration curve. In this context, the duration curve can be constructed by sorting the capacity factor and electrical load data from high to low. The $x-$axis for the DC exhibits the proportion of time that each capacity factor represents. The approximation of the duration curve is represented in this text as $\widetilde{DC}_p$.

$$NRMSE_{av} = \frac{\sum_{p \in P} \left(\frac{\sqrt{\frac{1}{||T||} \cdot \sum_{t \in T} (DC_{p,t} - \widetilde{DC}_{p,t})^2}}{max(DC_p) - min(DC_p)} \right)}{||P||} \tag{6}$$

The final metric used is the correlation between the different time series. This is used due to the fact that wind and solar output influences the load within a single region. This is referred to as the average correlation error (CE_{av}).

$$CE_{av} = \frac{2}{||P|| \cdot (||P|| - 1)} \cdot \left(\sum_{p_i \in P} \sum_{p_j \in P, j > i} \left| corr_{p_i, p_j} - \widetilde{corr}_{p_i, p_j} \right| \right) \tag{7}$$

Where $corr_{p1,p2}$ is the Pearson correlation coefficient between two time series $p_1, p_2 \in P$. Here, $V_{p1,t}$ represents the value of time series p_1 at time step t.

$$corr_{p1,p2} = \frac{\sum_{t \in T} \left(\left(V_{p1,t} - \overline{V}_{p1} \right) \cdot \left(V_{p2,t} - \overline{V}_{p2} \right) \right)}{\sqrt{\sum_{t \in T} \left(V_{p_1,t} - \overline{V}_{p1} \right)^2 \cdot \sum_{t \in T} \left(V_{p2,t} - \overline{V}_{p2} \right)^2}} \tag{8}$$

Figure 2 displays the number of clusters plot against each of the error metrics. For this experiment we chose eight clusters, due to eight being the smallest number with the best performing error metrics. After eight clusters $NRMSE_{av}$ and CE_{av} do not improve significantly, whilst REE_{av} deteriorates with number of clusters. Selecting less than eight clusters leads to a significant deterioration for both CE_{av} and $NRMSE_{av}$.

The error metrics do not exhibit a significant difference between using either the centroids or medoids technique. However, we chose to use the medoids technique. This was done due to the fact that the extreme high and low values would not be lost due to averaging [35].

2.5. Problem Formulation

- Reproducible data
- Summarize previously published results
- Modifications of previous results for this paper

3. Calculations

- Practical development

4. Results

Technology	MAE	MASE	RMSE	SD
CCGT	6.42	0.59	7.94	5.27
Coal	9.3	0.4	10.79	10.83
Nuclear	3.05	1.52	3.16	1.38
Solar	0.59	0.43	0.97	1.23
Wind	1.36	0.24	1.58	4.56

Table 1: Error metrics for time series forecast from 2013 to 2018

- Clear and concise results

5. Discussion

- Significance of work
- Avoid discussion of public work

6. Conclusion

- Main conclusions

7. Funding Sources

This work was supported by the Engineering and Physical Sciences Research Council, Centre for Doctoral Training in Cloud Computing for Big Data [grant number EP/L015358/1].

References

[1] V. Masson-Delmotte, P. Zhai, H. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, T. Waterfield, IPCC Special Report 1.5 - Summary for Policymakers, IPCC, 2018.

[2] Paris Agreement, United Nations 21 (2015) 1–23.

[3] P. Salas, The Effects of Uncertainty in the Technological Transitions of the Power Sector (March) (2017). URL `https://www.repository.cam.ac.uk/bitstream/handle/1810/2658`

[4] a. Epstein, Joshua M., Generative Social Science: Studies in Agent-Based Computational Modeling /, course book. Edition, Princeton University Press,, Princeton, N.J. :.

[5] P. Ringler, D. Keles, W. Fichtner, Agent-based modelling and simulation of smart electricity grids and markets - A literature review, Renewable and Sustainable Energy Reviews 57 (2016) 205–215.

[6] A. Kell, M. Forshaw, A. S. Mcgough, ElecSim : Monte-Carlo Open-Source Agent-Based Model to Inform Policy for Long-Term Electricity Planning, The Tenth ACM International Conference on Future Energy Systems (ACM e-Energy) (2019) 556–565.

[7] P. Nahmmacher, E. Schmid, L. Hirth, B. Knopf, Carpe diem: A novel approach to select representative days for long-term power system modeling, Energy 112 (2016) 430–442. doi:10.1016/j.energy.2016.06.081. URL `http://dx.doi.org/10.1016/j.energy.2016.06.081`

[8] E. Forgy, Cluster analysis of multivariate data: Efficiency versus interpretability of classification, Biometrics 21 (3) (1965) 768–769.

[9] J. H. W. Jr., Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association 58 (301) (1963) 236–244. arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1963.10500845, doi:10.1080/01621459.1963.10500845. URL `https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.1`

[10] J. Beckman, T. Hertel, W. Tyner, Validating energy-oriented CGE models, Energy Economics 33 (5) (2011) 799–806. doi:10.1016/j.eneco.2011.01.005. URL `http://dx.doi.org/10.1016/j.eneco.2011.01.005`

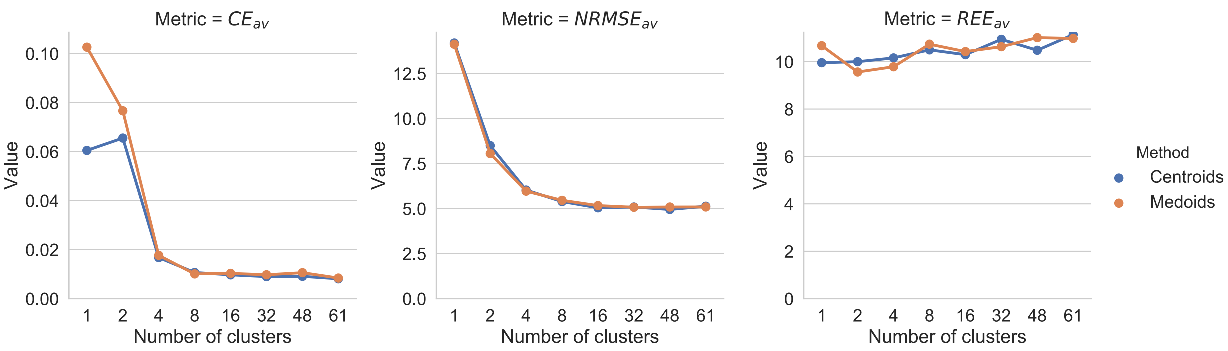

Figure 2: Comparison of number of clusters for accuracy.

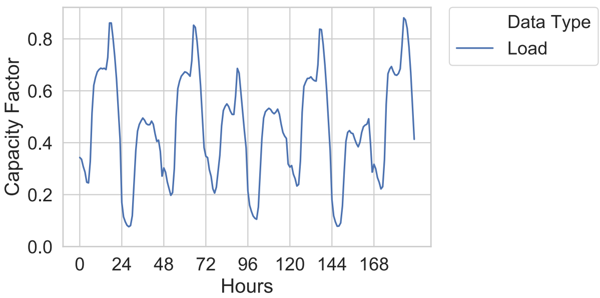

Figure 3: Representative days of electricity demand.

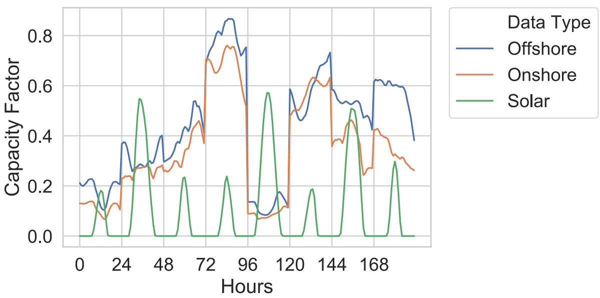

Figure 4: Representative days of renewable resources.

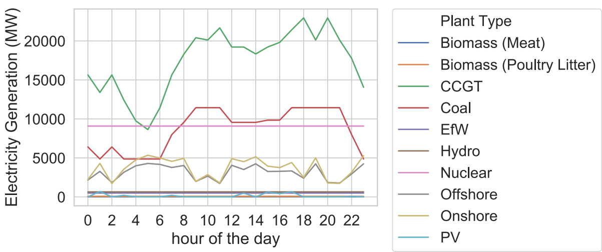

Figure 5: Example of a single day of dispatched supply.

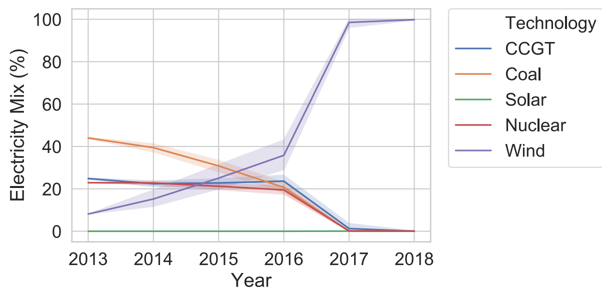

Figure 6: Electricity mix actual vs. simulated for validation scenario with a single representative day.

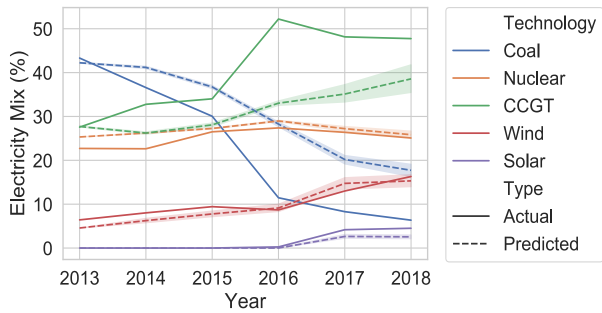

Figure 7: Electricity mix actual vs. simulated for validation scenario with eight representative days.

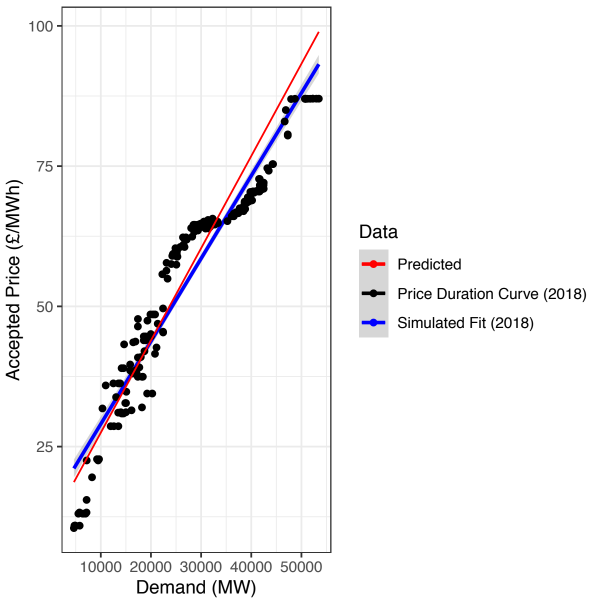

Figure 8: Predicted price curve for investment for most accurate run against simulated run in 2018.

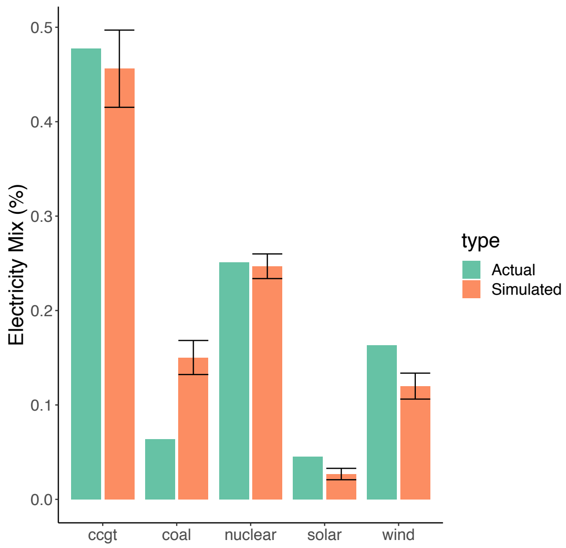

Figure 9: Electricity generation transition from 2013 to 2018 in the United Kingdom.

[11] M. Howells, H. Rogner, N. Strachan, C. Heaps, H. Huntington, S. Kypreos, A. Hughes, S. Silveira, J. DeCarolis, M. Bazillian, A. Roehrl, OSeMOSYS: The Open Source Energy Modeling System, Energy Policy 39 (10) (2011) 5850–5870. doi:10.1016/j.enpol.2011.06.033.
URL `http://dx.doi.org/10.1016/j.enpol.2011.06.033`

[12] P. Ringler, PowerACE Agent-based simulation of electricity markets, MOCAP Workshop on Modelling Carbon Prices - Interacting agent networks & Strategies under risk (2012).

[13] J. S. Hodges, J. A. Dewar, Is it you or your model talking? A framework for Model Validation (1992).

[14] P. P. Craig, A. Gadgil, J. G. Koomey, What can history teach us? A retrospective examination of long-term energy forecasts for the United States, Annual Review of Energy and the Environment 27 (1) (2002) 83–118. doi:10.1146/annurev.energy.27.122001.083425.

[15] K. C. Hoffman, The u.s. energy system: A unified planning framework, energy modeling, resources for the future (1973) 103–43.

[16] National plan for energy research, development, and demonstration: Creating energy choices for the future (1975).

[17] K. Hoffman, Perspectives on the Validation of Energy System Models Motivation, MITRE (February 2006) (2011) 1–13.

[18] A. P. Usher, Energy in the american economy, 18501975. an economic study of its history and prospects., The Journal of Economic History 21 (03) (1961) 418–421.

[19] J. V. Gillespie, Forecasting: An appraisal for policy-makers and planners. by william ascher. (baltimore: Johns hopkins university press, 1978. pp. xiii 239. $15.00, paper.), American Political Science Review 73 (2) (1979) 554–555. doi:10.2307/1954899.

[20] J. Koomey, P. Craig, A. Gadgil, D. Lorenzetti, Improving Long-Range Energy Modeling: A Plea for Historical Retrospectives, Energy Journal 24 (4) (2003) 75–92. doi:10.5547/ISSN0195-6574-EJ-Vol24-No4-4.

[21] D. for Business Energy, I. Strategy, Power stations in the united kingdom, may 2019, Digest of United Kingdom Energy Statistics (DUKES) (2019).

[22] Department for Business Energy & Industrial Strategy, Electricity Generation Costs (November) (2016).

[23] IEA, Projected Costs of Generating Electricity (2015) 215.

[24] U. E. I. Administration, Power plant operations report.

[25] N. W. et al.", "autoregressive integrated moving average" (1930).

[26] KPMG, Cost of Capital Study 2017, KPMG (2017).

[27] I. W. Paper, P. Heptonstall, Cost estimates for nuclear power in the UK, ICEPT Working Paper (August) (2012).

[28] S. C. Kincheloe, The weighted average cost of capital - the correct discount, The Appraisal journal. 58 (1) (1990).

[29] K. Poncelet, H. Hoschle, E. Delarue, A. Virag, W. Drhaeseleer, Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems, IEEE Transactions on Power Systems 32 (3) (2017) 1936–1948. doi:10.1109/TPWRS.2016.2596803.

[30] S. Ludig, M. Haller, E. Schmid, N. Bauer, Fluctuating renewables in a long-term climate change mitigation strategy, Energy 36 (11) (2011) 6674–6685. doi:10.1016/j.energy.2011.08.021.
URL `http://dx.doi.org/10.1016/j.energy.2011.08.021`

[31] G. Haydt, V. Leal, A. Pina, C. A. Silva, The relevance of the energy resource dynamics in the mid/long-term energy planning models, Renewable Energy 36 (11) (2011) 3068–3074. doi:10.1016/j.renene.2011.03.028.
URL `http://dx.doi.org/10.1016/j.renene.2011.03.028`

[32] I. Staffell, S. Pfenninger, Using bias-corrected reanalysis to simulate current and future wind power output, Energy 114 (2016) 1224–1239. doi:10.1016/j.energy.2016.08.068.

URL http://dx.doi.org/10.1016/j.energy.2016.08.068
[33] Gb fuel type power generation production.
URL https://gridwatch.co.uk/
[34] W. D'haeseleer, J. Duerinck, K. Poncelet, D. Six, E. Delarue, Impact of the level of temporal and operational detail in energy-system planning models, Applied Energy 162 (2015) 631–643. doi:10.1016/j.apenergy.2015.10.100.
URL http://dx.doi.org/10.1016/j.apenergy.2015.10.100
[35] A. P. Hilbers, D. J. Brayshaw, A. Gandy, Importance subsampling: improving power system planning under climate-based uncertainty, Applied Energy 251 (May) (2019) 113114. doi:10.1016/j.apenergy.2019.04.110.
URL https://doi.org/10.1016/j.apenergy.2019.04.110