The Bartlett School of Environment, Energy and Resources

MSc ESDA Coursework Title Page

UCL Candidate Code: Weasley (HSWW0, HSDN8, FDQX2)

Module Code: BENV0091

Module Title: Energy Data Analysis

Coursework Title: Forecasting the BSUoS Charge: A Machine Learning Approach

Module Leader: Dr Aidan O'Sullivan

Date: 17/01/2020

Word Count: 4991

By submitting this document, you are agreeing to the Statement of Authorship:

I/We certify that the attached coursework exercise has been completed by me/us and that each and every quotation, diagram or other piece of exposition which is copied from or based upon the work of other has its source clearly acknowledged in the text at the place where it appears.

I/We certify that all field work and/or laboratory work has been carried out by me/us with no more assistance from the members of the department than has been specified.

I/We certify that all additional assistance which I/we have received is indicated and referenced in the report.

Please note that penalties will be applied to coursework which is submitted late or which exceeds the maximum word count. Information about penalties can be found in the MSc ESDA Course Handbook which is available on Moodle:

- ***Penalties for late submission:*** *https://moodle-1819.ucl.ac.uk/course/view.php?id=9967&section=30*
- ***Penalties for going over the word count:*** *https://moodle-1819.ucl.ac.uk/course/view.php?id=9967&section=30*

In the case of coursework that is submitted late and is also over length, then the greater of the two penalties shall apply. This includes research projects, dissertations and final reports.

i. Table of Contents

ii. List of Figures

iii. List of Tables

1. Introduction

Anthropogenic climate change has been the crux of geopolitics in recent years, with the UK becoming the first major economy to ratify statutory targets of net zero emissions by 2050 (BEIS, 2019). Energy supply is the second largest contributor to carbon dioxide emissions in the country (>98 $MtCO_2$ in 2018), behind the transport sector, thus climate change mitigation efforts have been largely concentrated towards the energy industry (BEIS, 2018; IEA, 2013). The *Electricity Market Reform* has been a pivotal policy mechanism in the UK, facilitating decarbonisation by supporting the development of renewable generation, nurturing niche innovations towards economic viability (BEIS, 2015).

Over the last decade, renewable energy penetration in the electricity generation portfolio has risen from 2% to 31% (figure 1.1), virtually displacing coal (REF, 2020). In fact, the third quarter of 2019 saw, for the first time, more electricity generated by renewable energy sources that fossil fuels (Ali, 2019). Whilst continuing this energy transition is essential to uphold climate targets, it has resulted in a progressively complex energy system (Panahian, et al., 2017). The intermittency of renewable energy sources has resulted in the convoluted task of balancing a stochastic supply of electricity with an inelastic and volatile demand (Maciejowska, et al., 2019).

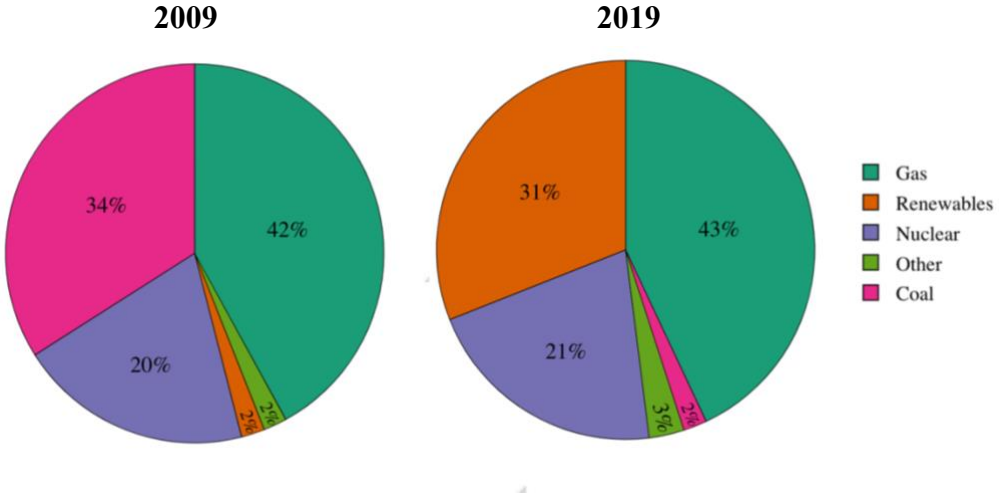

Figure 1.1: UK electricity generation (balancing mechanism) fuel mix (REF, 2020).

The Balancing Services Use of System (BSUoS) charge recovers the costs of operating the transmission network and is levied to all generators and suppliers within the system in the form of a £/MWh tariff, ultimately cascading down to consumers (Ofgem, 2017). The balancing

market is the primary contributor to the BSUoS charge, which includes costs of constraining generation (National Grid, 2015). With enhanced renewable energy capacity, limited foresight of supply coupled with lack of suitable storage mechanisms puts strain on fossil fuel generators to adapt quickly when demand cannot be met by renewables (APS, 2011). Furthermore, with most of the UK wind energy capacity located in remote regions, far from high demand areas, there is an increasing need for generation constraints to avoid grid congestion. Constraint payments now comprise over 8% of the BSUoS charge, reaching a staggering £108 million in 2018 (REF, 2018). These elements have not only seen a significant increase in the volatility of the BSUoS charge, but have more than doubled the average annual tariff, from 1.14 to 2.57 £/MWh, over the last decade (figure 1.2) (Brabben, 2019).

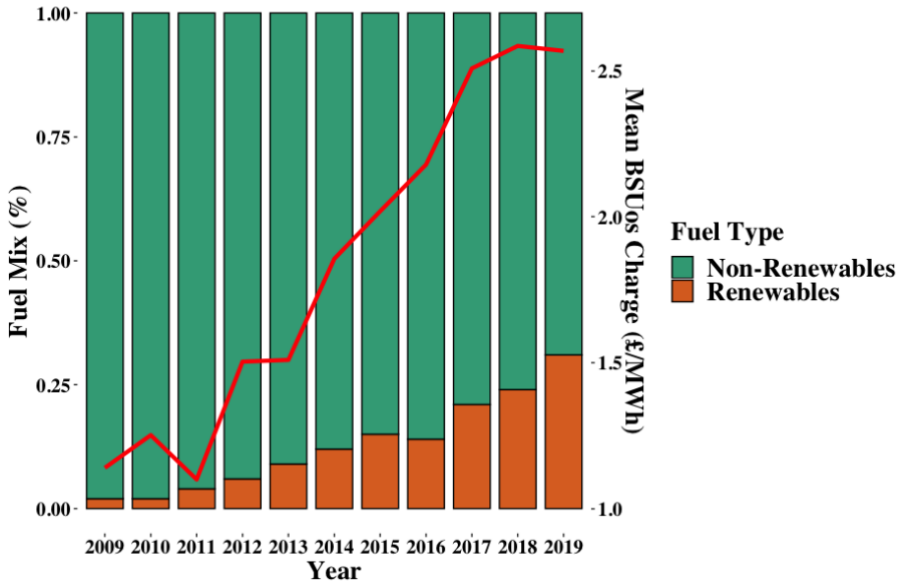

Figure 1.2: BSUoS charge from 2009 to 2019 (National Grid ESO, 2019; REF, 2020).

With many electricity suppliers offering flexible price tariffs, the increasing volatility of the BSUOS charge makes it difficult to price these contracts appropriately (Haven Power, 2015). Furthermore, the BSUoS charge tends to remove, on average, 5% of the hourly revenue of large power plants, with this value fluctuating considerably (Mclean, 2018). As the BSUoS is charged half-hourly, accurate intra-day forecasting of the charge is beneficial to generators, suppliers and consumers, to gain better understanding of revenue fluctuations whilst targeting generation more efficiency (Energy UK, 2016).

Consequently, this essay aims to develop a machine learning (ML) approach for predicting the intra-day BSUoS charge, at an hour-ahead forecast horizon, using the outputs from generation units (GUs) across the UK. First, un-supervised exploratory analysis is employed to gain insights into the ways in which different fuel types contribute to supply, as well the fluctuating characteristics of the BSUoS charge. This is followed by a performance comparison of classical and deep learning ML methods for predicting intra-day BSUoS charges.

2. Literature Review

Political discourse concerning the impact of renewable energy penetration in the transmission network gained traction when, in 2009, Ofgem introduced the *Connect and Manage* regime – a mechanism enabling National Grid to grant early grid access to renewable energy projects (BEIS, 2013). Whilst this enabled rapid diffusion of renewable technologies, un-reinforced transmission infrastructure meant rising constraint costs were to be distributed via the BSUoS charge (Woolf & Babington, 2009). By 2015, new projects amounting nearly 30 GW of installed renewable energy capacity had been granted early grid access, with estimated emissions savings of 5.9 $MtCO_2$ (Ofgem, 2015). However, these installations attributed an increase of over £50 million in annual constraint costs, whilst the average annual BSUoS charge doubled (ibid.).

Recent literature concerning system integration costs of renewable energy generators has been motivated partly through scepticism surrounding the apparent declining levelized costs of electricity for low-carbon technologies (Joos & Staffell, 2018). Quantitative modelling of wind energy penetration resulted in estimated additional system integration costs of approximately £25/MW when the relevant market share surpassed 40% (Hirth, et al., 2015). Whilst these estimates are subject to uncertainty, primarily because true integration costs are system-specific, in general, other studies concur that enhanced renewable penetration into the grid will continue to increase operating and balancing costs of the electricity network – raising concerns about the economic feasibility of renewable energy dominated systems (UKERC, 2017; Miligan, et al., 2011).

Whilst economic theory supposes that the wholesale price of electricity decreases with increasing renewable energy penetration, owed to the merit-order effect, electricity prices in

the retail market tend to also reflect system integration and balancing costs, such as the BSUoS charge, paradoxically raising the consumer price of electricity (Murray, 2019). Although limited literature exists concerning the BSUoS charge directly, the field of electricity price forecasting is extensive (Lago, et al., 2018). Traditionally, empirical techniques for predicting electricity prices largely comprised economic and game theoretic modelling, involving multi-agent simulation of heterogenous markets, as well as computational modelling of stochastic price dynamics – incorporating, for example, the inelasticity of electricity demand and arbitrary fluctuations in weather conditions (Weron, 2014; Ventosa, et al., 2005; González, et al., 2012). However, convoluted market-based simulations require strict assumptions, and so, with the evolution of big data analytics, statistical and ML approaches to electricity price forecasting have gained traction in recent literature (Ugurlu, et al., 2018).

Statistical models for time-series analysis of electricity spot prices have proved effective for modelling day-ahead to long-term forecast horizons (Ziel & Weron, 2018). Forecasting day-ahead electricity spot prices for the Spanish market using an Auto Regressive Integrated Moving Average (ARIMA) model produced, when controlling for seasonality, a mean average percentage error (MAPE) of 12.4% (Cruz, et al., 2011). Garcia et al. (2005) found that Generalised Auto Regressive Conditionally Heteroskedastic (GARCH) models outperformed ARIMA when using lagged price variables to forecast day-ahead electricity prices in the Californian market, with a MAPE of 9.2% and 11.8% respectively (Garcia, et al., 2005). Whilst statistical time-series modelling has been used extensively for day-ahead (or greater) forecast horizons, statistical models tend to fall short when employed with high resolution data (Weron, 2014; Keles, et al., 2016). This is partly due to the linearity of many statistical models, which fail to address the multifarious nature of the electricity market, with non-linear load behaviours and stochastic price signals (Amjady & Hemmati, 2006). Comparing the performance of both statistical and ML methods for forecasting intra-day Belgian electricity prices found, in general, that ML models significantly outperformed statistical models, with the MAPE for ARIMA and GARCH models reaching over 20%, but less than 14% for Support Vector Machines (SVMs) and Neural Networks (NNs) (Lago, et al., 2018).

The popularity of adopting ML methods for ex ante analysis of electricity market prices with short-term forecast horizons stems from the enhanced ability to capture intrinsic volatility in price dynamics (McGlynn, et al., 2018). Employing a black-box training process for predicting hour-ahead spot prices in the New England power market found NNs to outperform Regression

Tree (RT) and SVM algorithms, with root-mean-square errors (RMSEs) of 4.5, 5.7 and 5.5 respectively (Zhu, et al., 2018). Similar results were demonstrated using lagged price variables coupled with weather data, in the sense that NNs were again found to outperform SVMs, with MAPEs of 8.4% and 12.43% (Foruzan, et al., 2015). Performance evaluation of multi-layer NNs, using data from the Turkish day-ahead market, found that as the number of hidden layers increased from 1 to 3, the mean average error decreased from 9.5 to 6.0, but stagnated from this point on (Ugurlu, et al., 2018).

Utilising hourly resolution price data from the Italian market, Harasheh et. al (2016) illustrated the benefits of bootstrap aggregating (bagging), with bagged RTs producing a lower MAPE than NNs – 3.87% and 5.10% respectively (Harasheh, 2016). Building on the use of ensembling methods, Random Forest (RF) algorithms have also been found to outperform NNs, with a lower MAPE (12.03% compared to 12.83%), whilst an AdaBoost (a meta-algorithm used for weighted boosting) model, encompassing 50 Extra-Tree base learners, also outperformed NNs (with MAPEs of 11.59% and 13.99% respectively) using lagged price variables as input features, coupled with load, wind and solar generation (McGlynn, et al., 2018; Mei, et al., 2014).

3. Methodology

3.1. Data

System charging data for the UK electricity network is made available to the public via the National Grid ESO website, thus historical BSUoS charges could be retrieved directly (National Grid ESO, 2019). System data for the balancing mechanism was accessed through the Balancing Market Reporting Service (BMRS) API (ELEXON, 2020).

Data from 2018 was selected for analysis, as at the time of extraction, the 2019 dataset was not yet complete, thus the full seasonality characteristics of the data could not have been assessed. The API service was used to extract the actual generation output per GU, as well as system operator (SO) flag count and rolling demand.

The data was split into both a training and test set – the training set comprised data from the first three quarters (January – September), whilst the fourth quarter (October – December) was

held for testing. The training set was modified to smooth outliers; akin to prior literature, BSUoS charge outliers were remedied by substituting them with the mean charge, plus thrice the standard deviation in the appropriate direction (figure 3.1) (Ketterer, 2015).

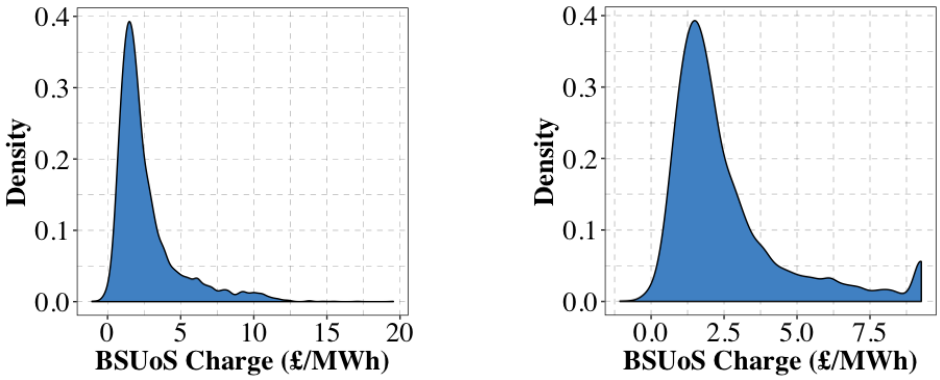

Figure 3.1: Training set before (left) and after (right) smoothing of outliers.

3.2. Feature Description

The outputs of individual generators were deemed relevant to the forecasting of the BSUoS charge, as each GU tends to impart different system integration and transmission costs, depending on their output, fuel type and location etc (National Grid, 2015). A total of 236 generators contributed to the electricity supply in 2018, however not all generators supplied electricity consistently, thus missing values for GU outputs were assumed zero. Individual characteristics of each GU were retrieved by matching the relevant Energy Identification Code (EIC) with public GU information (National Grid ESO, 2020).

The SO flag identifies instances where notable action was taken to balance the system, usually as a result of transmission constraints required to alleviate grid congestion (ELEXON, 2019). Multiple SO flags can occur during the same settlement period (SP), thus the variable was engineered to summarise the total number of flags per SP, and, where missing values were present, the total number of flags was assumed zero. This variable was selected for analysis owing to its significant correlation with the BSUoS charge – the number of balancing actions taken by the SO gives a significant indication of the magnitude of the charge, as, in general, the BSUoS charge increases with increasing number of SO flags (figure 3.2).

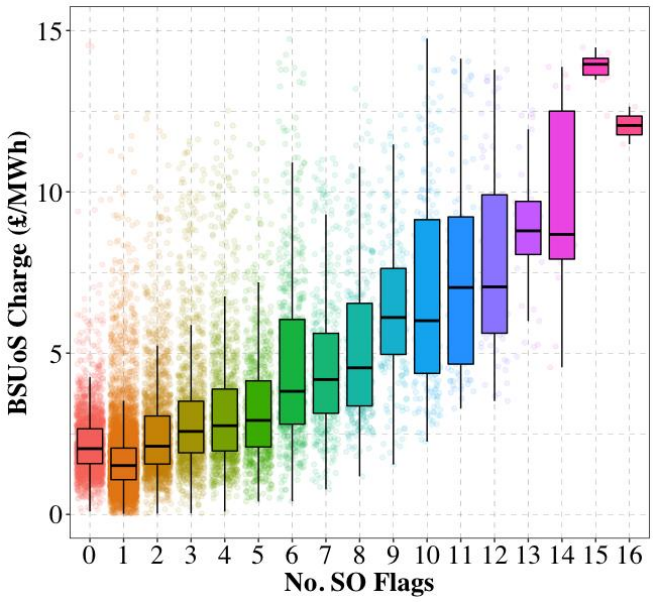

Figure 3.2: The BSUoS charge relative to the number of SO flags.

Rolling demand data was retrieved at five-minute resolution; the data was averaged over 30-minute intervals to reflect each settlement period, to replicate the resolution of the other datasets. System load was selected as an input feature due to its influence over BSUoS charge dynamics – figure 3.3 illustrates that, in general, periods of low demand correspond with periods of high BSUoS charge (and vice versa).

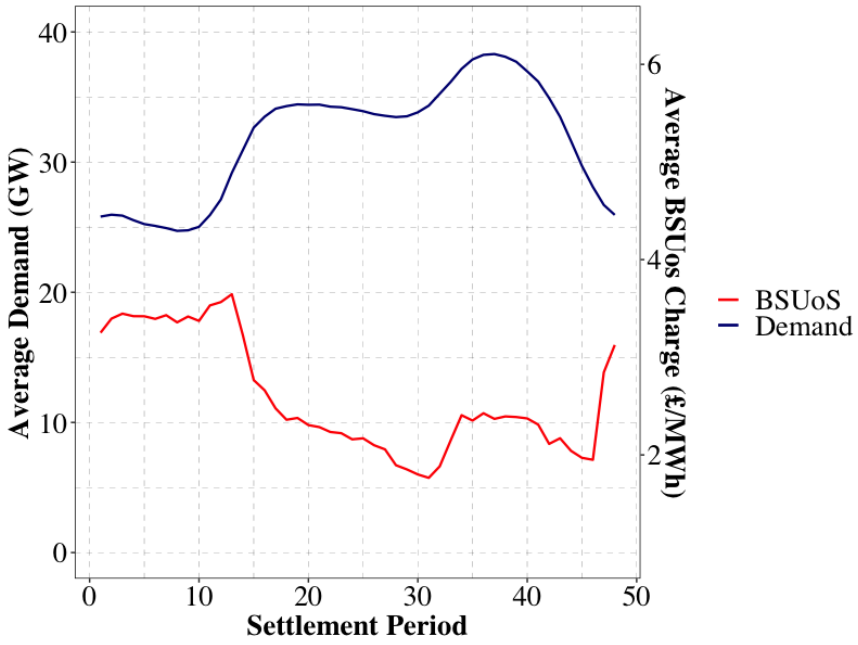

Figure 3.3: Average demand and BSUoS for each settlement period.

This phenomenon is vindicated by the need for greater management of grid inertia when demand is low, as well as the provision of system constraints when over-supply is causing congestion in the transmission network (National Grid ESO, 2018). This is further demonstrated in figure 3.4, which illustrates that, in general, as demand declines, the number of SO flags decreases. This is because in periods of high demand, over-supply is less of an issue, thus fewer balancing actions are required to maintain equilibrium between supply and demand.

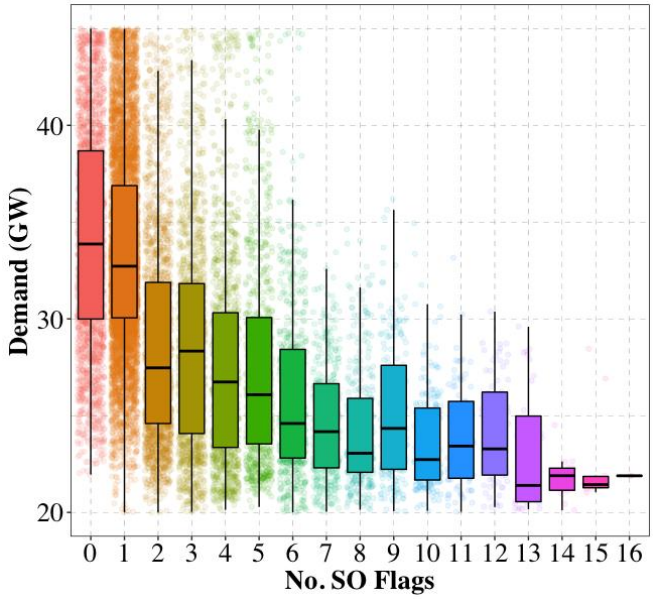

Figure 3.4: Electricity demand relative to the number of SO flags.

3.3. Exploratory Analysis

Unsupervised cluster analysis was utilised to gain insight into the structural characteristics of the variables. The k-means clustering algorithm was applied to the BSUoS charge, as well as several generator fuel types, namely coal, nuclear wind and gas. The k-means algorithm was selected owing to its simplicity and computational efficiency – beneficial due to the large number of features.

3.3.1. BSUoS Charge

Figure 3.5 presents the results of k-means clustering applied to the BSUoS charge over each SP throughout the day.

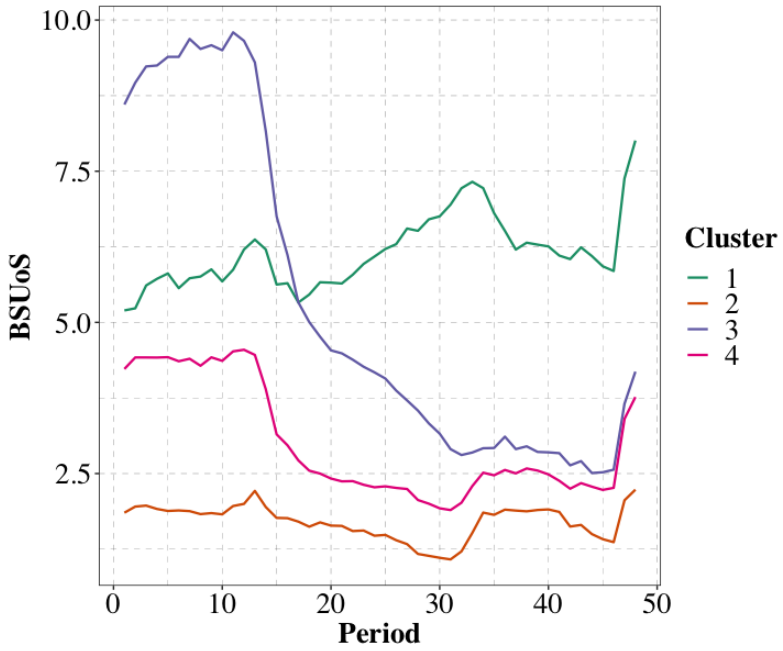

Figure 3.5: k-Means clustering of the BSUoS charge.

Upon visual inspection, cluster two appears to correspond to a consistently below average BSUoS charge, whilst the first cluster illustrates the contrary. Clusters three and four adopt similar shapes, demonstrating a peak in the early hours of the day, before decreasing throughout. With reference to figure 3.3, these two clusters appear to resemble a typical BSUoS profile, fluctuating with demand – the charge is higher during the night when demand is low, but decreases as electricity demand rises.

The second cluster was found to contribute most during winter months, as well as primarily on working days, vindicated by the relationship between electricity demand and the BSUoS

charge – demand for electricity is highest during the week, especially in winter when the temperature is low, thereby lowering the BSUoS charge, as less balancing actions are required to manage over-supply.

In contrast, cluster one contributed most during the months of June, July and August, and primarily on weekends. This, again, substantiates the relationship between system balancing costs and electricity demand, as during periods of lower demand, which occur most frequently in summer, the BSUoS charge appears to be highest – in fact, this cluster demonstrated a value consistently greater than the yearly average.

The seasonality of clusters three and four were less distinct, most likely because these clusters simply reflect the rolling average BSUoS charge profile for the year (figure 3.3). Cluster three did appear to contribute slightly more on weekends and during summer, which explains the higher average charge for this cluster compared with cluster two.

3.3.2. Generator Fuel Type

Figure 3.6 presents the results of k-means clustering applied to total generation of individual fuel types.

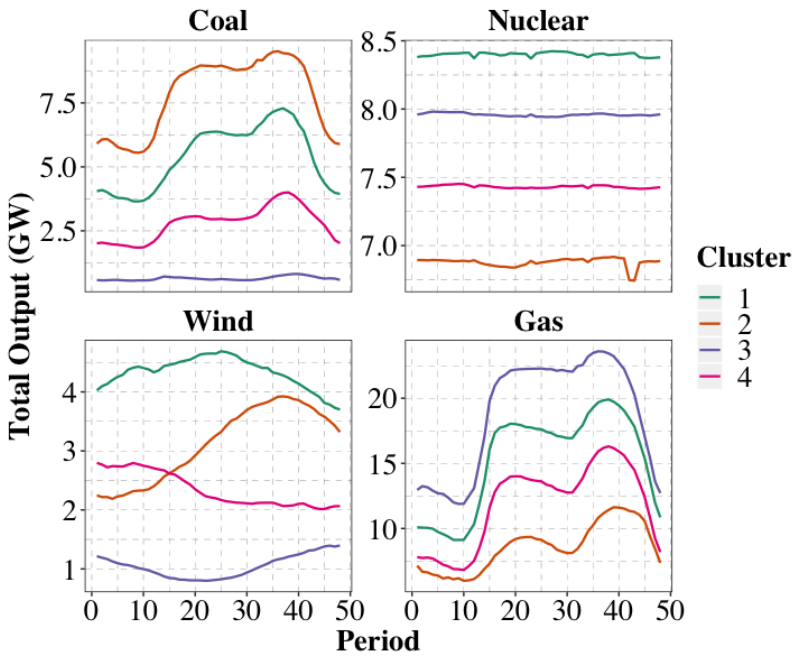

Figure 3.6: k-Means clustering of generator fuel types.

For coal plants, clusters one, two and four reflect the conventional demand profile (figure 3.3), with a local peak during the morning and a higher peak in the evening. The third cluster,

appears to contribute primarily to base load, having a relatively stable profile. Cluster two was found to contribute significantly during winter months, especially on working days, justified by its higher average output compared to the other clusters. In contrast, cluster four was found to be most prominent during periods of low demand, especially during summer. Finally, there were no distinct seasonality profiles observed for clusters one and three, owing to their contribution to relatively average generation and provision of consistent base load.

Clustering of nuclear power plants highlighted their utilisation for base load generation. All clusters illustrated consistent generation throughout the day, with very little seasonality distinction between them, differing only slightly throughout the year as electricity demand changes. This result was expected as nuclear plants have considerably low variable costs with substantial fixed costs, thus base load generation is the most economically viable mode of operation.

Gas (CCGT) plants presented a very similar clustering profile to that of coal plants, directly reflecting the load profile, and being able to fluctuate seasonally with changing levels of electricity demand. It can therefore be concluded that gas plants are extremely flexible, with the ability to ramp up and down with changing market conditions – a valuable characteristic needed to support the increasing penetration of renewable energy sources (Clemente, 2017).

Wind farms presented rather arbitrary clusters; the first cluster appears to gradually rise and fall throughout the day, contrasted by cluster three. The second cluster generates more electricity in the evening, whilst cluster four has higher generation in the morning. Electricity generation from wind is extremely volatile, depending considerably on external factors such as climate and wind speed. Thus, these clusters could indicate variable weather conditions, with cluster one relating to high wind speeds and cluster three the contrary. In this sense, cluster two could reflect days with high wind speeds, but with system constraints when demand is low. The distinction between clusters could also relate to the distance of the plants relative to the shore.

3.4. Forecasting Methodology

The adopted forecasting methodology comprised the following three stages: (1) the application of Principle Component Analysis (PCA) to subset the GUs with the greatest influence over total supply variance, thereby reducing the dimensionality of feature space, (2) building and tuning several ML models to predict hour-ahead BSUoS charge, and (3) comparing the performance of each model based on conventional metrics.

3.4.1. PCA

In 2018, 236 GUs contributed to the electricity supply, resulting in a rather high dimensional dataset – considering them all as input features would likely result in multicollinearity between variables, as well as low computational efficiency.

PCA is commonly used for feature extraction, in which the input variables are projected down to a lower dimensional subspace of linearly uncorrelated components, so as to maximise the variance in the transformed data space. However, this study chose to exploit the inherent characteristics of the leading principle component (PC), using feature selection to subset the GUs with greatest influence over the variance, indicated by the highest loading factors in the first PC. Through eigen-decomposition of the correlation matrix, the percentage variance captured by each PC was estimated and the subsequent loading factors were evaluated, selecting the top 20 GUs as input features.

3.4.2. Regression Algorithms

$$\mathrm{BSUOS_t} = \beta_1 + \beta_2(\mathrm{BSUOS_{t-2}}) + \beta_3(\mathrm{DUM_t}) + \beta_4(\mathrm{DEMAND_{t-2}}) + \beta_i \sum_{i=5}^{24} \mathrm{GU}_{i,t-2} \qquad \text{Eq. 3.1}$$

The hour-ahead BSUoS charge was initially predicted using the above multivariate OLS regression algorithm, in order to provide a basis for model comparison. The seasonality of the BSUoS charge was incorporated using a dummy variable vector (DUM), used to flag to relevant month, day and SP. Rolling demand, BSUoS charge and GU outputs were incorporated in the model as lag variables, representing the SPs from an hour before (at t-2). In order to account for intrinsic non-linearities in the target variable, Polynomial Ridge Regression (PRR) was also used, applying an L1 penalty term to the coefficients (eq.2) to mitigate overfitting.

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - x_i^T \beta) + \lambda \sum_{j=1}^{m} |\beta_j| \qquad \text{Eq. 3.2}$$

Next, tree-based models were assessed, which function by recursively partitioning feature space into binary sub-regions, with predictions made based on the mean value of training data in the appropriate region. Initially, two regression trees (RTs) were evaluated, the first using cost complexity (CP) pruning to find the optimal subtree and the second using maximum tree depth as the termination criterion for the algorithm. Akin to previous literature, this study explored base model ensembling, using data sampling with replacement to train several regression trees on different subsets of training data (Mei, et al., 2014). In order to alleviate correlation of predictions, a Random Forest (RF) algorithm was employed, using random feature sampling to capture as much unique information in the data as possible.

Finally, a neural network (NN) model was formulated, consisting of an input layer, hidden layers and an output layer. Deep learning models were found to be extremely prevalent throughout relevant literature, owing to their ability to describe complex data structures (Ugurlu, et al., 2018). Each layer in the network consisted of interconnected neurons, in which a weighted sum of inputs was transformed through an activation function. A tanh activation function was used to introduce non-linearities in the model, in attempt to capture the stochastic nature of the BSUoS charge. The NN is structurally parallel, enabling simultaneous hierarchical processing, facilitating enhanced processing speeds with unseen data, leading to a more efficient model.

3.4.3. Evaluation Metrics

Throughout literature, the most commonly observed performance evaluation metric used in electricity price forecasting was MAPE, and so this metric was employed to allow for comparison with literature findings. This metric (eq.3) evaluates predictive accuracy by expressing the average loss as a percentage of the true values. Thus, a higher MAPE indicates a lesser performing model.

$$\text{MAPE} = 100\% * \frac{1}{\text{n}} \sum \left| \frac{\text{Prediction} - \text{Actual}}{\text{Actual}} \right| \qquad \text{Eq. 3.3}$$

R-squared (eq. 4) was also used as a supplementary metric to compare the models, to measure to goodness-of-fit of the regression analyses.

$$R^2 = \frac{\text{var}(\text{Actual} - \text{Prediction})}{\text{var}(\text{Actual})} \qquad \text{Eq. 3.4}$$

4. Results

4.1. Principle Component Analysis

PCA was applied to the GU outputs using the correlation matrix as opposed to the covariance matrix to account for scale. The first PC was found to capture 52% of the variance in the data (figure 4.1), and, whilst this is not particularly high, it was significantly greater than the subsequent PCs.

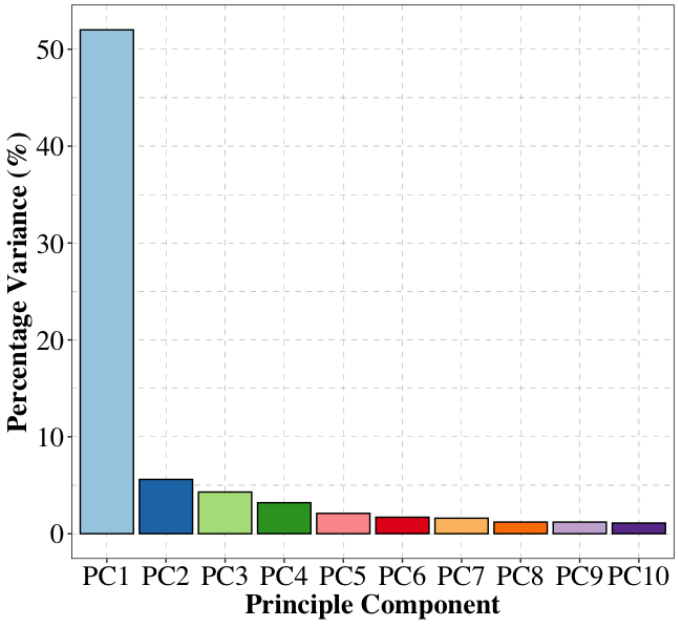

Figure 4.1: Percentage variances of first 10 principle components.

Each of the 236 generators were assigned the appropriate loading factor for PC1; the dataset was filtered to contain only the generators with the 20 highest loading factors, as these were deemed most influential toward supply. Gas power plants were most prevalent in the subset, justified by their considerable role in not only meeting general demand, but fluctuating generation to accommodate the volatility of renewable energy generation. Nuclear reactors

were found to be the second most influential fuel type, most likely due to their substantial contribution to provision of base load.

4.2. BSUoS Forecasting

4.2.1. Linear Regression

The multivariate linear regression (LR) base model was trained without penalisation, to provide a basis for comparison. When evaluated against the test set, this model performed rather poorly, with a MAPE of 62.42% and R-squared of 0.45 (table 4.1). Figure 4.2 demonstrates how this model does indeed capture the general profile of the BSUoS charge, however severely underfits the true signal, owed to the lack of flexibility in the algorithm.

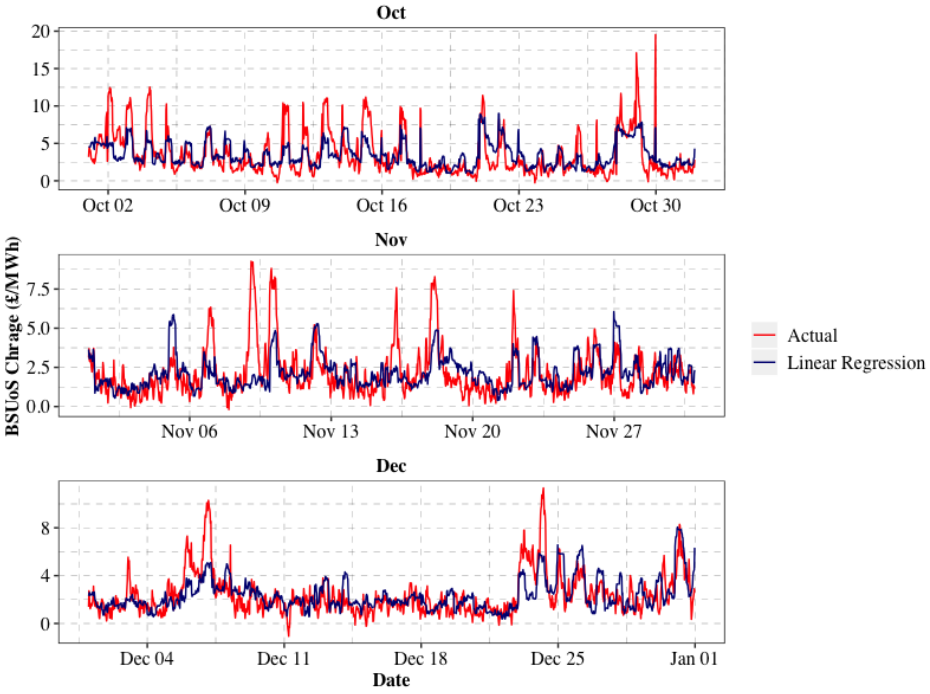

Figure 4.2: Visualising test set performance: Linear Regression.

4.2.2. Polynomial Ridge Regression

In order to account for the temporal dependences of the time-series data and avoid leakage, forward chaining cross-validation was used to tune the PRR model, to establish the optimal L1 penalty weighting (lambda). Successively using each day as the validation set, the RMSE was calculated for a range of values for lambda to determine the mean cross-validation error (figure 4.3). The optimum model was found to be a polynomial regression of degree three, with a lambda value of 0.19.

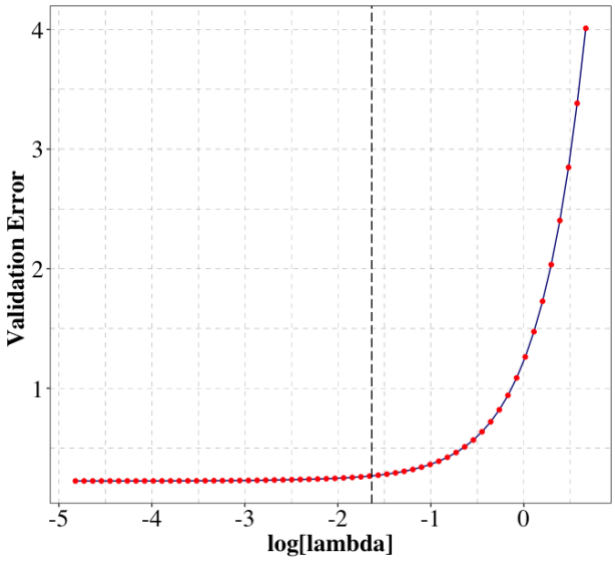

Figure 4.3: Polynomial Ridge Regression hyperparameter tuning: lambda.

The algorithm was trained using the appropriate parameters and evaluated against the test set. The PRR model performed surprisingly well, with a MAPE of 23.81% and R-squared of 0.86 (table 4.1). Figure 4.4 illustrates that, compared with the previous model, increasing the polynomial degree alleviated bias, enabling the model to better describe the signal.

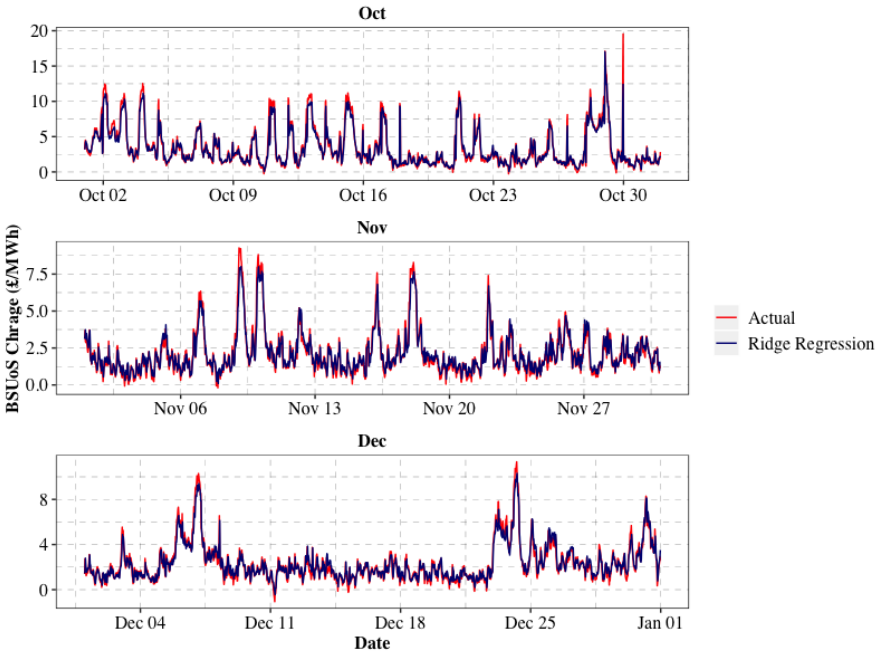

Figure 4.4: Visualising test set performance: Polynomial Ridge Regression.

4.2.3. Regression Tree

Two regression trees were trained using forward chaining cross-validation, the first with CP pruning and the second with a maximum depth stopping criterion. The RMSE was found to increase proportionally to the weight of the CP pruning parameter (alpha), with an optimum value of 0.01 and minimum average RMSE of 1.42 (figure 4.5). Next, as the maximum depth of the tree increased, the RMSE decreased, until a depth of 9, where the performance stagnated, with an RMSE of 1.43.

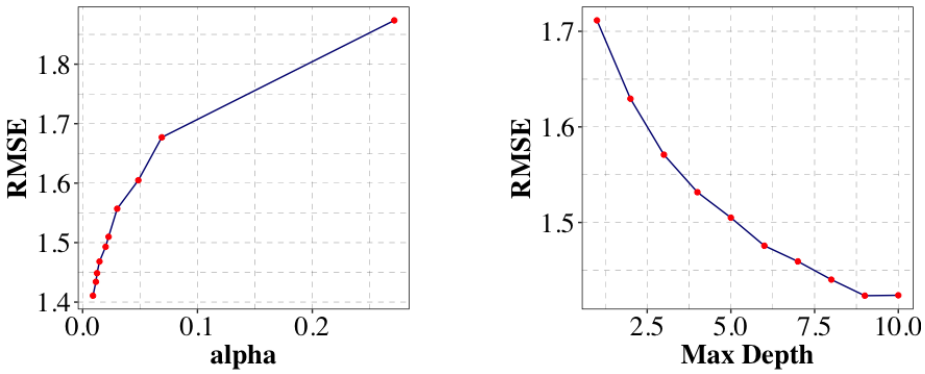

Figure 4.5: Regression Tree hyperparameter tuning: alpha (left) and maximum depth (right).

As the CP model resulted in the lowest validation error, the final RT model was trained using CP pruning, with alpha = 0.01. This model performed quite poorly on the test, with a MAPE of 60.85% and R-squared of 0.47 (table 4.1). Figure 4.6 demonstrates how the model failed to describe the volatility in the true signal, resulting in severe underfitting.

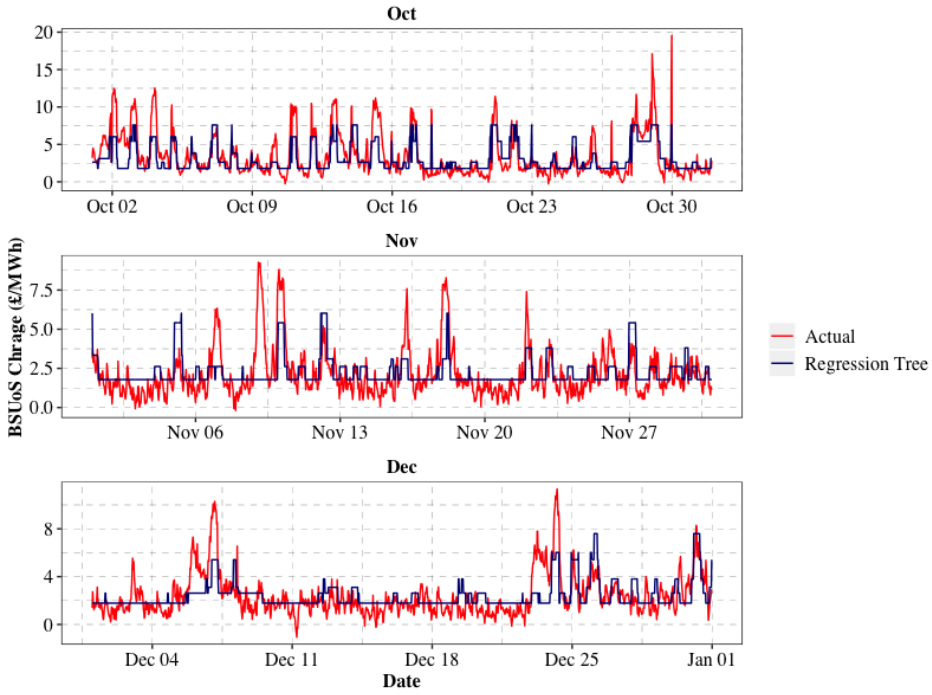

Figure 4.6: Visualising test set performance: Regression Tree.

4.2.4. Random Forest

Two primary parameters were considered in the construction of the RF algorithm: (1) the number of regression trees comprising the ensemble and (2) the number of randomly sampled features at each split. Figure 4.7 demonstrates that as the number of regression trees within the forest increased, so did the performance, indicated by a reduction in the MAPE. The performance stagnated at around 500 trees, with additional trees only adding to the computational cost, without providing significant performance enhancement. The number of randomly sampled features was chosen to be 5, as again, beyond this point performance stagnated.

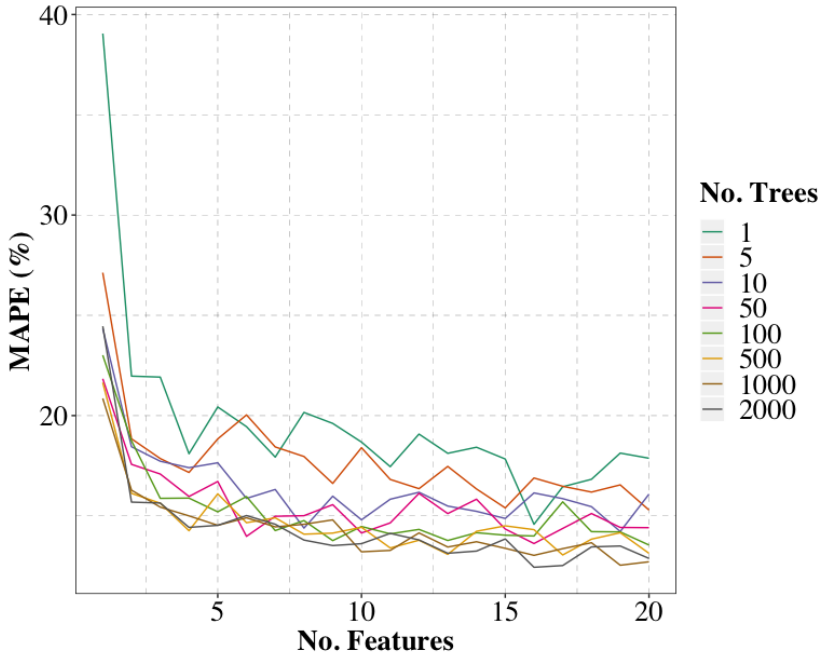

Figure 4.7: Random Forest hyperparameter tuning: no. trees and no. features.

When evaluated against the test set, the RF model performed exceptionally well, with a MAPE of 11.53% and R-squared of 0.98 (table 4.1). Figure 4.8 illustrates the flexibility of this model, virtually aligning with the true signal apart from during severe peaks.

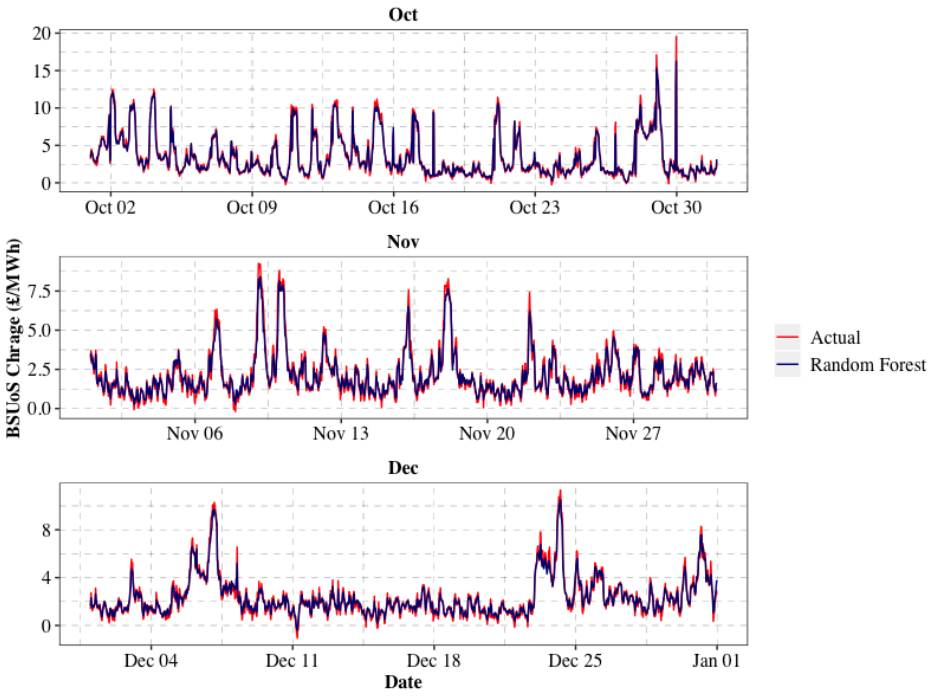

Figure 4.8: Visualising test set performance: Random Forest.

4.2.5. Neural Network

Several NNs with tanh activation functions were trained using the keras package, in order to evaluate the optimum parameters for the model. Looping through various values for the number of neurons per layer, as well as the number of hidden layers, the performance of the NN tended to stagnate around 10 neurons per layer. Using this number, the optimum number of hidden layers was determined to be three (figure 4.9), as beyond this point, performance started to decrease, potentially due to overfitting the training set.

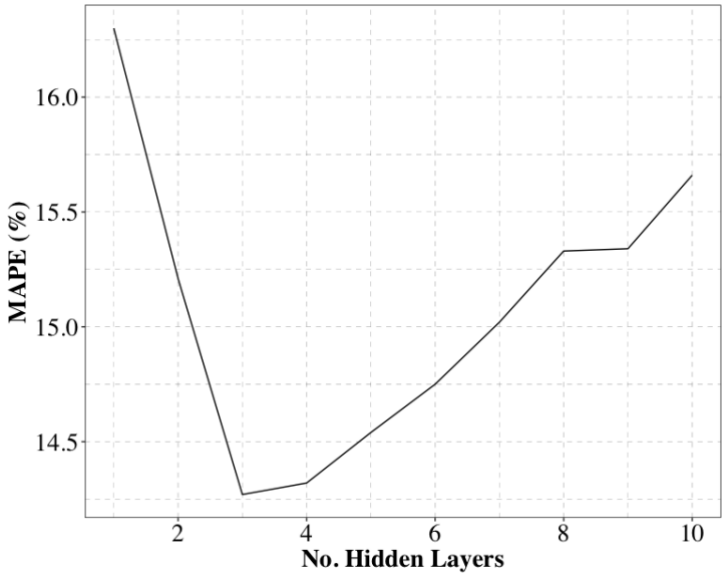

Figure 4.9: MAPE vs. number of hidden layers in the neural network.

A NN with three hidden layers (10 neurons in each) performed well on the test set, with a MAPE of 14.27% and R-squared of 0.93 (table 4.1). Figure 4.10 highlights that although the NN was capable of describing the signal to a high degree, it tended to fall short in predicting sudden peaks, which again could be resultant of overfitting the training data.

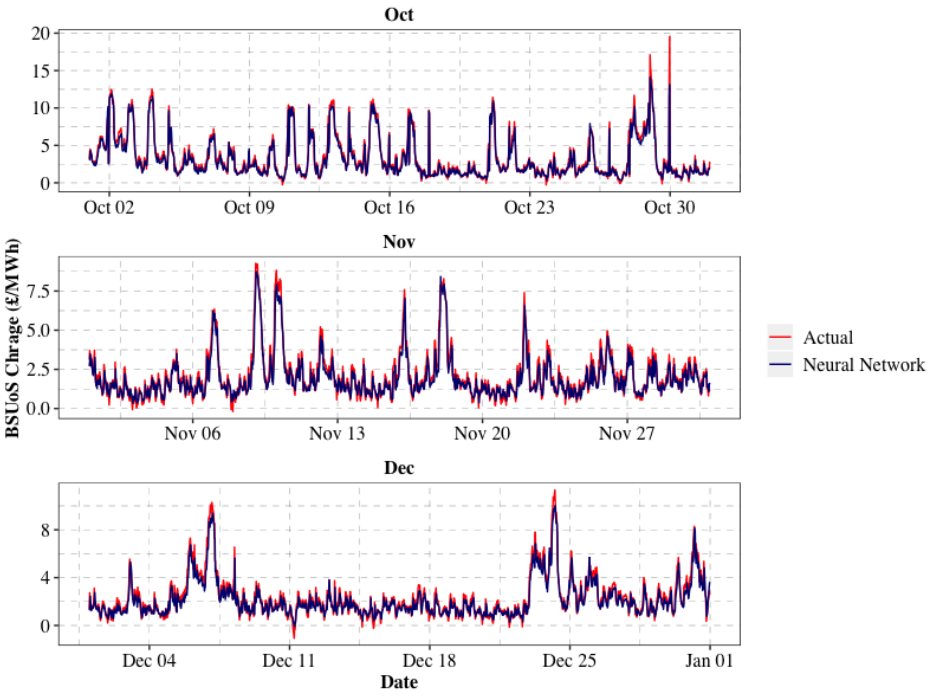

Figure 4.10: Visualising test set performance: Neural Network.

4.2.6. Overall Performance

Overall, the RF presented the greatest performance on the test set, based on the both MAPE and R-squared metrics.

Table 4.1: Comparison of performance metrics.

Model	**MAPE**	**R-squared**
Linear Regression	62.42	0.45
Polynomial Ridge Regression	23.81	0.86
Regression Tree	60.85	0.47
Random Forest	11.53	0.98
Neural Network	14.27	0.93

5. Discussion

This work offers a comprehensive evaluation of the UK balancing mechanism system costs, using unsupervised ML methods to assess the dynamic characteristics of the BSUoS charge and the contributions to national supply of different fuel types, as well as a performance comparison of several classical and deep learning algorithms for predicting the hour-ahead BSUoS charge, based on historical data from National Grid ESO and BMRS.

Prior to forecasting analysis, a k-means clustering algorithm was applied to the rolling BSUoS charge, which accentuated the relationship between the balancing mechanism and system load. A consistently high BSUoS charge was present during seasonal periods of generally low demand (and vice versa), justified by the relative frequency of balancing actions required during these stages, as generation constraints are more likely to occur when demand is low.

Applying the k-means algorithm to individual fuel types demonstrated the flexibility of fossil fuel generators, with coal and gas plants being able to adjust supply in accordance to market demand. In contrast, nuclear plants proved to be highly inflexible, contributing only to base load. Finally, wind generators illustrated rather arbitrary cluster profiles, with variable supply likely resultant from climate, technology type and presence of system constraints.

The top 20 GUs, based on their contribution to the variance in supply, were selected using PCA to determine their relative loading factor in PC1. These GUs were used, along with rolling demand and SO flag count, to forecast the hour-ahead BSUoS charge. The LR algorithm presented the poorest performance, justified by the non-linear characteristics of this regression problem. The added flexibility of the PRR model alleviated the bias present in the first order model, rendering a much higher performance. A notable limitation to the methodology was that the subset of GUs, whilst most influential over supply, may not have been the most correlated with the BSUoS charge, improvements could be made by comparing the PCA results with the correlation matrix of features.

The RT model was limited by the size of the data set, as predictions are based on the mean values of training data in the relevant region, resulting in lower performance. However, axis aligned partitioning of input space, enabling the ability to capture non-linear data structures, coupled with bagging and random feature sampling, led to the RF algorithm performing better

than the NN; analogous to the conclusions of Mei et al. (2014) and Harasheh (2016). The optimum NN was found to comprise three hidden layers (akin to findings of Ugurlu et al. (2018)), with 10 neurons in each. Whilst both the NN and RF presented powerful forecasting abilities, both models fell short in capturing extreme values, potentially due to limited data size leading to overfitting. Furthermore, data points dispersed near the interface between peaks and troughs are significantly affected by information generalisation, resulting in neutralised fluctuations, limiting overall performance. Overall, forecasting the BSUoS charge presented many of the same difficulties found in literature for electricity price prediction, with relatively similar model performances (Foruzan, et al., 2015; Zhu, et al., 2018).

6. Conclusion

Ultimately, this work has demonstrated the utility of ML algorithms for intra-day BSUoS charge prediction, with a RF model presenting the highest performance on this data set, based on MAPE and R-squared metrics. Future work may seek to explore other input features, evaluating autocorrelation to determine the optimum lag-length, or finding other means of dimensionality reduction, better identifying features most correlated with the BSUoS charge. It is encouraged to evaluate the performance of more sophisticated methods better suited to time-series analysis, such as LSTM NNs, Hidden Markov Models, or the application of extreme value theory to better understand the tails in the BSUoS charge distribution.

An online interactive representation of the ML model performances, as well as a visualisation of the electricity generation fuel mix and rolling BSUoS charge is available through the following link: https://weasley.shinyapps.io/esda_project/

7. References

Ali, U., 2019. *UK renewable sources outcompete fossils fuels for the first time.* [Online]
Available at: https://www.power-technology.com/news/renewable-energy-outcompete-fossil-fuels/
[Accessed 7 January 2020].

Amjady, N. & Hemmati, M., 2006. Energy price forecasting. *IEEE Power and Energy Magazine*, pp. 20-29.

APS, 2011. Integrating Renewable Electricity on the Grid. *APS Panel on Public Affairs.*
BEIS, 2013. *Electricity network delivery and access.* [Online]
Available at: https://www.gov.uk/guidance/electricity-network-delivery-and-access#connect-and-manage
[Accessed 5 January 2020].

BEIS, 2015. *Electricity Market Reform: Contracts for Difference.* [Online]
Available at: https://www.gov.uk/government/collections/electricity-market-reform-contracts-for-difference

BEIS, 2018. *2018 UK Greenhouse Gas Emissions: Provisional Figures.* [Online]
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/790626/2018-provisional-emissions-statistics-report.pdf
[Accessed 7 January 2020].

BEIS, 2019. *UK becomes first major economy to pass net zero emissions law.* [Online]
Available at: https://www.gov.uk/government/news/uk-becomes-first-major-economy-to-pass-net-zero-emissions-law
[Accessed 7 January 2020].

Brabben, J., 2019. *Constrained development: Scottish wind and the issues of network charging.* [Online]
Available at: https://www.cornwall-insight.com/newsroom/all-news/constrained-development-scottish-wind-and-the-issues-of-network-charging
[Accessed 8 January 2020].

Clemente, J., 2017. *Natural Gas Is The Flexibility Needed For More Wind And Solar.* [Online]
Available at: https://www.forbes.com/sites/judeclemente/2017/12/31/natural-gas-is-the-flexibility-needed-for-more-wind-and-solar/#1a5b6d05777f
[Accessed 8 January 2020].

Cruz, A., Munoz, A., Zamora, J. L. & Espinola, R., 2011. The effect of wind generation and weekday on Spanish electricity spot price forecasting. *Electric Power Systems Research,* Volume 81, pp. 1924-1935.

ELEXON, 2019. *Imbalance Pricing Guidance.* [Online]
Available at: https://www.elexon.co.uk/documents/training-guidance/bsc-guidance-

notes/imbalance-pricing/
[Accessed 11 January 2020].

ELEXON, 2020. *Balancing Mechanism Reporting Service (BMRS).* [Online]
Available at: https://www.bmreports.com/bmrs/?q=help/about-us
[Accessed 7 January 2020].

Energy UK, 2016. *Electricity charging arrangements report.* [Online]
Available at: https://www.energy-uk.org.uk/publication.html?task=file.download&id=5903
[Accessed 10 January 2020].

Foruzan, E., Scott, S. & Lin, J., 2015. A Comparative Study of Different Machine Learning Methods for Electricity Prices Forecasting of an Electricity Market. *North American Power Symposium.*

Garcia, R., Contreras, J., van Akkeren, M. & Garcia, J., 2005. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Transactions on Power Systems,* 20(2).

González, V., Contreras, J. & Bunn, D., 2012. Forecasting Power Prices Using a Hybrid Fundamental-Econometric Model. *IEEE Transactions on Power Systems,* Volume 27.

Harasheh, M., 2016. Forecasting Wholesale Electricity Prices With Artificial Intelligence Models: The Italian Case. *Preprints.*

Haven Power, 2015. *Haven Power Limited's response to the Consultation on the Final Proposals for Electricity System Operator incentives 2015-17.* [Online]
Available at:
https://www.ofgem.gov.uk/sites/default/files/docs/2015/06/haven_power_response.pdf
[Accessed 8 January 2020].

Hirth, L., Ueckerdt, F. & Edenhofer, O., 2015. Integration costs revisitedeAn economic framework for wind andsolar variability. *Renewable Energy,* Volume 74, pp. 925-939.

IEA, 2013. World Energy Outlook Special Report: Redrawing the Energy-Climate Map. *International Energy Agency (IEA).*

Joos, M. & Staffell, I., 2018. Short-term integration costs of variable renewable energy: Wind curtailment and balancing in Britain and Germany. *Renewable and Sustainable Energy Reviews,* Volume 86, pp. 45-65.

Keles, D., Scelle, J., Paraschiv, F. & Fichtner, W., 2016. Extended forecast methods for day-ahead electricity spot prices applyingartificial neural networks. *Applied Energy,* Volume 162, pp. 218-230.

Ketterer, J., 2015. The Impact of Wind Power Generation on the Electricity Price in Germany. *Energy Economics,* Volume 44, pp. 270-280.
Lago, J., De Ridder, F. & De Schutter, B., 2018. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy,* Volume 221, pp. 386-405.

Maciejowska, K., Nitka, W. & Weron, T., 2019. Day-Ahead vs. Intraday—Forecasting the Price Spread to Maximize Economic Benefits. *Energies*, Volume 12, p. 631.

McGlynn, D., Coleman, S., Kerr, D. & McHugh, C., 2018. Day-Ahead Price Forecasting in Great Britain's BETTA Electricity Market. *Symposium Series on Computational Intelligence.*

Mclean, S., 2018. *ELASTACLOUD.* [Online]
Available at: https://customers.microsoft.com/en-au/story/elastacloud-partner-professional-services-azure-machine-learning
[Accessed 4 January 2020].

Mei, J. et al., 2014. A random forest method for real-time price forecasting in New York electricity market. *Power & Energy Society: IEEE General Meeting.*

Miligan, M. et al., 2011. Integration of Variable Generation, Cost-Causation, and Integration Cost. *The Electricity Journal*, 24(9), pp. 51-63.

Murray, B., 2019. *The Paradox of Declining Renewable Costs and Rising Electricity Prices.* [Online]
Available at: https://www.forbes.com/sites/brianmurray1/2019/06/17/the-paradox-of-declining-renewable-costs-and-rising-electricity-prices/#81ae8d61d5b4
[Accessed 8 January 2020].

National Grid ESO, 2018. *CMP308: Removal of BSUoS charges from Generation.* [Online]
Available at: https://www.nationalgrideso.com/document/141486/download
[Accessed 10 January 2020].

National Grid ESO, 2019. *Balancing Services Use of System (BSUoS) charges.* [Online]
Available at: https://www.nationalgrideso.com/charging/balancing-services-use-system-bsuos-charges
[Accessed 6 January 2020].

National Grid ESO, 2020. *LIO for EIC codes.* [Online]
Available at: https://www.nationalgrideso.com/balancing-services/lio-eic-codes
[Accessed 4 January 2020].

National Grid, 2015. *Introduction to Charging: Which Parties Pay Which Charges?.* [Online]
Available at: https://www.nationalgrid.com/sites/default/files/documents/44939-TNUoS%2C%20BSUoS%20and%20Connection%20Charging%20Information.pdf
[Accessed 7 January 2020].

Ofgem, 2015. *Monitoring the 'Connect and Manage' electricity grid access regime.* [Online]
Available at:
https://www.ofgem.gov.uk/sites/default/files/docs/monitoring_the_connect_and_manage_electricity_grid_access_regime_sixth_report_from_ofgem_0.pdf
[Accessed 7 January 2020].
Ofgem, 2017. *Charging.* [Online]
Available at: https://www.ofgem.gov.uk/electricity/transmission-networks/charging
[Accessed 5 January 2020].

Panahian, M., Ghosh, S. & Ding, G., 2017. Assessing potential for reduction in carbon emissions in a multi-unit of residential development in Sydney. *Procedia Engineering,* Volume 180, pp. 591-600.

REF, 2018. *UK Wind Constraint Payments Reach New and Exceptional Levels.* [Online]
Available at: https://www.ref.org.uk/ref-blog/345-uk-wind-constraint-payments-reach-new-and-exceptional-levels
[Accessed 7 January 2020].

REF, 2020. *British Electricity Generation (Balancing Mechanism) Fuel Mix.* [Online]
Available at: https://www.ref.org.uk/fuel/index.php?tab=yr&share=N
[Accessed 6 January 2020].

Ugurlu, U., Oksuz, I. & Tas, O., 2018. Electricity Price Forecasting Using Recurrent Neural Networks. *Energies*, 11(5), p. 1255.

UKERC, 2017. *The costs and impacts of intermittency.* [Online]
Available at: http://www.ukerc.ac.uk/publications/the-costs-and-impacts-of-intermittency-2016-update.html
[Accessed 7 January 2020].

Ventosa, M., Baillo, A., Ramos, A. & Rivier, M., 2005. Electricity market modeling trends. *Energy Policy,* 33(7), pp. 897-913.

Weron, R., 2014. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting,* Volume 30, pp. 1030-1081.

Woolf, F. & Babington, H., 2009. Greening the grid. *Utility Week*, 31(18), pp. 20-21.

Zhu, Y. et al., 2018. Power Market Price Forecasting via Deep Learning. *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society.*

Ziel, F. & Weron, R., 2018. Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics,* Volume 70, pp. 396-420.