The impact of online machine-learning methods on long-term investment decisions and generator utilization in electricity markets

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, Online learning, Machine learning, Market investment, Climate Change

1. Introduction

The integration of higher proportions of intermittent renewable energy sources (IRES) in the smart grid will mean that the forecasting of electricity demand will become increasingly important. This is due to the fact that supply must mean demand at all times and that IRES are less predictable than dispatchable energy sources such as coal and combined-cycle gas turbines (CCGTs) [1].

Typically, peaker plants, such as reciprocal gas engines, are used to fill fluctuations in demand, that had not been previously planned for. These peaker plants are typically expensive to run and have higher greenhouse gas emissions than their non-peaker counterparts [2].

To reduce reliance on peaker plants, it is helpful to know how much electricity demand there will be in the future so that more efficient plants can be used to meet this demand. To aid in this, machine learning and statistical techniques have been used to accurately predict demand based on several different factors and data sources, such as weather, day of the week and holidays [3, 4, 5, 6].

While various studies have looked at how to predict electricity demand at various time horizons with the highest accuracy [7], the impact that the differing methods used have on a long-term electricity market has been studied to a lesser degree. In this paper, we compare several machine learning and statistical techniques to forecast 24 hours ahead to simulate a day-ahead market.

In addition to this, we use the long-term agent-based model, ElecSim [8, 9], to simulate the impact of different forecasting methods on long-term investments, power plant usage and carbon emissions from the year 2018 to 2035 in the United Kingdom.

To compare the impact of different methods, we trial both online and offline machine learning and statistical techniques. Online learning methods can learn from novel data while maintaining what was learnt from previous data. This type of statistical modelling is useful for non-stationary datasets, and time-series data where recalculation of a model would take a prohibitive amount of time. We trial different algorithms and split times of the year into seasons, weekdays, weekends and holidays.

Using these methods, we take the error distributions, or residuals, and fit a variety of distributions. The distribution that best fits the respective residuals is then used and sampled from to adjust the demand in the model ElecSim. We then observe the differences in carbon emissions, and power plants invested in and utilised with each of the different statistical and machine learning methods to a base case.

We show that online learning has a significant impact on reducing the error for predicting electricity consumption a day ahead when compared to traditional offline learning techniques.

We show that the forecasting algorithm has a non-negligible effect on carbon emissions and use of coal, onshore, photovoltaics, reciprocal gas engines and CCGT. Specifically, the amount of coal, photovoltaics, and reciprocal gas used from 2018 to 2035 was proportional to the median absolute error, while both onshore and offshore wind are inversely proportional to the median absolute error.

Total investments in coal, offshore and photovoltaics are proportional to the median absolute error, while in-

vestments in CCGT, onshore and reciprocal gas engines inversely proportional.

In Section 2 we review the literature. We introduce the model ElecSim, and the methods used in Section 3. The Methods section is shown in Section 4, results in Section 5. We discuss and conclude our work in Sections 6 and 7 respectively.

2. Literature Review

Multiple papers have looked at demand-side forecasting [7]. These include both artificial intelligence [10, 11, 12] and statistical techniques [?, 13, 14]. However, the impact of online learning has been discussed with less frequency. In addition to this, our research shows the impact of different algorithms on investments made, electricity sources contribution and carbon emissions over a 17 year period using the long-term electricity market agent-based model, ElecSim.

The have been several studies in diverse applications on the use of online machine learning to predict time-series data. Johansson *et al.* apply online machine learning algorithms for heat demand forecasting [15]. They find that their demand predictions display robust behaviour within acceptable error margins. They find that artificial neural networks (ANNs) provide the best forecasting ability of the standard algorithms and can handle data outside of the training set.

Baram *et al.* combine an ensemble of active learners by developing an active-learning master algorithm [16]. To achieve this, they propose a simple maximum entropy criterion that provides effective estimates in realistic settings. Their active-learning master algorithm is empirically shown to, in some cases, outperform the best algorithm in the ensemble on a range of classification problems.

An extension of the FLORA algorithm is proposed by Schmitt *et al* in [17]. Their FLORA-MC enhances the FLORA algorithm for multi-classification and numerical input values. They use this algorithm for an ambient computing application. Ambient computing is where computing and communication merges into every day life. They find that their model outperforms traditional offline learners by orders of magnitude.

Pindoriya *et al.* trial a number of different machine learning methods such an adaptive wavelet neural network (AWNN). They find that AWNN has good prediction properties when compared to other forecasting techniques such as wavelet-ARIMA, multi-layer perceptron (MLP) and radial basis function (RBF) neural networks as well as the fuzzy neural network (FNN).

Goncalves Da Silva *et al.* show the effect of prediction accuracy on local electricity markets [18]. To this end, they compare forecasting of groups of consumers in comparison to single individuals. They trial the use of the Seasonal-Nave and Holt-Winters algorithms and look at the effect that he errors have on trading in an intra-day electricity market of consumers and prosumers. They found that with a photovoltaic penetration of 50%, over 10% of the total generation capacity was uncapitalised and roughly 10, 25 and 28% of the total traded volume were unnecessary buys, demand imbalances and unnecessary sells respectively. This represents energy that the participant has no control.

3. Material

3.1. Machine Learning

Machine learning is a methodology for finding and describing structural patterns in data [19]. Typically machine learning algorithms are algorithms which are trained using data. These generate models which can make predictions or decisions with regards to classification of data [15]. Machine learning algorithms can be used for both static datasets as well as streaming data.

3.2. Online learning

Online machine learning is an extension of machine learning algorithms, that can be used on dynamic datasets, such as time-series data. In traditional machine learning algorithms, when new data is obtained, the historical data must be used to retrain the entire model with new model. This can be costly both in terms of time and in computation power [20]. Online algorithms therefore avoids the repeated retraining of data and improves the learning efficiency [21].

Examples of such algorithms are Passive Aggressive Regressor [22], Linear Regression, Box-Cox Regressor [23], K-Neighbors Regressor [24], Softmax Regression, Multilayer perceptron regressor [25]. For our work, we trial the stated algorithms, in addition to a host of offline learning techniques. The online techniques trialled were Lasso regression [26], ridge regression [27], Elastic Net [28], least angle regression [29], Extra Trees Regressor [29], Random Forest Regressor [30], AdaBoost Regressor [31], Gradient Boosting Regressor [32], Support vector regression [33], and offline versions of Multilayer perceptron regressor, K-Neighbors regressor, linear regression.

We trial the previously mentioned statistical and machine learning algorithms, and vary the parameters using grid search and cross-validation in the python package, scikit learn [34].

3.3. ElecSim model

- Introduce ElecSim.

4. Methods

- Use of hyperparameter tuning, talk about time taken to train/query models.
- Talk about ML methods used
- Talk about residuals
- Talk about sampling from residuals and placing these errors on the day-ahead market.

5. Results

- Results of offline learning, online machine learning shown. Include residuals and MAE,MAPE,MASE etc

- Results of the residuals on the output of ElecSim until 2035.

6. Discussion

- Discuss the impact of this on the electricity market and global economy. Make suggestions.

7. Conclusion

- Summary of work and future work.

8. Funding Sources

References

[1] C. Lu, H. T. Wu, S. Vemuri, Neural network based short term load forecasting, IEEE Transactions on Power Systems 8 (1) (1993) 336–342. doi:10.1080/02533839.1995.9677697.

[2] A. Mahmood, M. N. Ullah, S. Razzaq, A. Basit, U. Mustafa, M. Naeem, N. Javaid, A new scheme for demand side management in future smart grid networks, Procedia Computer Science 32 (2014) 477–484. doi:10.1016/j.procs.2014.05.450. URL http://dx.doi.org/10.1016/j.procs.2014.05.450

[3] A. Kell, A. McGough, M. Forshaw, Segmenting residential smart meter data for short-Term load forecasting, in: e-Energy 2018 - Proceedings of the 9th ACM International Conference on Future Energy Systems, 2018. doi:10.1145/3208903.3208923.

[4] M. Al-Musaylh, R. Deo, J. Adamowski, Y. Li, Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia, Advanced Engineering Informatics 35 (November 2017) (2018) 1–16. doi:10.1016/j.aei.2017.11.002. URL https://doi.org/10.1016/j.aei.2017.11.002

[5] P. Vrablecová, A. Bou Ezzeddine, V. Rozinajová, S. Šárik, A. K. Sangaiah, Smart grid load forecasting using online support vector regression, Computers & Electrical Engineering 0 (2017) 1–16. doi:10.1016/j.compeleceng.2017.07.006.

[6] T. Hong, J. Wilson, J. Xie, A. Member, Long Term Probabilistic Load Forecasting and Normalization With Hourly Information 5 (1) (2014) 456–462.

[7] A. K. Singh, Ibraheem, S. Khatoon, M. Muazzam, D. K. Chaturvedi, Load forecasting techniques and methodologies: A review, ICPCES 2012 - 2012 2nd International Conference on Power, Control and Embedded Systems (2012). doi:10.1109/ICPCES.2012.6508132.

[8] 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.

[9] A. J. M. Kell, M. Forshaw, A. S. McGough, Long-Term Electricity Market Agent Based Model Validation using Genetic Algorithm based Optimization, The Eleventh ACM International Conference on Future Energy Systems (e-Energy'20) (2020).

[10] K.-h. Kim, H.-s. Youn, S. Member, Y.-c. Kang, Short-term load forecasting for special days in anomalous load conditions using neural networks, IEEE Transactions on Power Systems 15 (2) (2000) 559–565. doi:10.1109/59.867141.

[11] J. Nagi, S. K. Yap, S. K. Tiong, S. K. Ahmed, Electrical Power Load Forecasting using Hybrid Self-Organizing Maps and Support Vector Machines, The 2nd International Power Engineering optimization Conference (PEOCO) (June) (2008) 51 – 56.

[12] F. L. Quilumba, W.-j. Lee, H. Huang, D. Y. Wang, S. Member, R. L. Szabados, Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities (2014) 1–8.

[13] S.-j. Huang, S. Member, K.-r. Shih, Short-Term Load Forecasting Via ARMA Model Identification Including Non-Gaussian 18 (2) (2003) 673–679.

[14] H. Nguyen, C. K. Hansen, Short-term electricity load forecasting with Time Series Analysis, 2017 IEEE International Conference on Prognostics and Health Management (ICPHM) (2017) 214–221doi:10.1109/ICPHM.2017.7998331. URL http://ieeexplore.ieee.org/document/7998331/

[15] C. Johansson, M. Bergkvist, D. Geysen, O. D. Somer, N. Lavesson, D. Vanhoudt, Operational Demand Forecasting in District Heating Systems Using Ensembles of Online Machine Learning Algorithms, Energy Procedia 116 (2017) 208–216.

doi:10.1016/j.egypro.2017.05.068.
URL http://dx.doi.org/10.1016/j.egypro.2017.05.068

[16] Y. Baram, R. El-Yaniv, K. Luz, Online Choice of Active Learning Algorithms, Proceedings, Twentieth International Conference on Machine Learning 1 (2003) 19–26.

[17] J. Schmitt, M. Hollick, C. Roos, R. Steinmetz, Adapting the user context in realtime: Tailoring online machine learning algorithms to ambient computing, Mobile Networks and Applications 13 (6) (2008) 583–598. doi:10.1007/s11036-008-0095-8.

[18] P. Goncalves Da Silva, D. Ilic, S. Karnouskos, The Impact of Smart Grid Prosumer Grouping on Forecasting Accuracy and Its Benefits for Local Electricity Market Trading, IEEE Transactions on Smart Grid 5 (1) (2014) 402–410. doi:10.1109/TSG.2013.2278868.

[19] I. H. Witten, E. Frank, M. a. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 2011.
URL http://books.google.com/books?id=bDtLM8CODsQC&pgis=1

[20] Y. Li, P. Guo, X. Li, Short-Term Load Forecasting Based on the Analysis of User Electricity Behavior, Algorithms 9 (4) (2016) 1–14. doi:10.3390/a9040080.

[21] H. J. Rong, G. B. Huang, N. Sundararajan, P. Saratchandran, H. J. Rong, Online Sequential Fuzzy Extreme Learning Machine for Function Approximation and Classification Problems, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 39 (4) (2009) 1067–1072. doi:10.1109/TSMCB.2008.2010506.

[22] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, Y. Singer, Online Passive-Aggressive Algorithms, Journal of Machine Learning Research (2006). doi:10.1201/b15810-63.

[23] G. E. P. Box, D. Cox, An Analysis of Transformations Author (s): G . E . P . Box and D . R . Cox Published by : Wiley for the Royal Statistical Society Stable URL : http://www.jstor.org/stable/2984418 Accessed : 08-03-2016 06 : 53 UTC Your use of the JSTOR archive indicates, Journal, Source Statistical, Royal Series, Society 26 (2) (1964) 211–252.

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

[25] G. E. Hinton, Connectionist learning procedures, Artificial Intelligence 40 (1-3) (1989) 185–234. doi:10.1016/0004-3702(89)90049-0.

[26] R. Tibshirani, Regression Shrinkage and Selection Via the Lasso, Journal of the Royal Statistical Society: Series B (Methodological) 58 (1) (1996) 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.

[27] P. Geladi, Measurement, regression and calibration, philip brown, oxford statistical science series. vol. 12, oxford science publications, oxford, 1993, no of pages: 224. price: £27.50. isbn 0 19-852245-2, Journal of chemometrics 8 (5) (1994) 371–372.

[28] J. Friedman, T. Hastie, R. Tibshirani, Regularization Paths for Generalized Linear Models via Coordinate Descent, Journal of Statistical Software 33 (1) (2010) 1–22. arXiv:0908.3817, doi:10.1016/j.expneurol.2008.01.011.

[29] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, The Annals of Statistics 32 (2) (1988) 440–444. doi:10.1109/glocom.1988.25879.

[30] L. Breiman, Random forests, Machine Learning 45 (1) (2001) 5–32. arXiv:dx.doi.org/10.1023%2FA%3A1010933404324, doi:10.1023/A:1010933404324.

[31] Y. Freund, R. E. Schapire, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of Computer and System Sciences 55 (1) (1997) 119–139. doi:10.1006/jcss.1997.1504.

[32] J. H. Friedman, Greedy Function Approximation: A Gradient Boosting Machine (316) 400.

[33] C. Cortes, V. Vapnik, Support-Vector Networks, Machine Learning 20 (3) (1995) 273–297. arXiv:arXiv:1011.1669v3, doi:10.1023/A:1022627411411.

[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.