Household Electricity Demand Forecasting - Benchmarking State-of-the-Art Methods

Andreas Veit, Christoph Goebel, Rohit Tidke,
Christoph Doblander and Hans-Arno Jacobsen
Department of Computer Science, Technische Universität München
andreas.veit@tum.de, christoph.goebel@tum.de,
rohittidke@gmail.com, doblande@in.tum.de

ABSTRACT

We benchmark state-of-the-art methods for forecasting electricity demand on the household level. Our evaluation is based on two data sets containing the power usage on the individual appliance level. Our results indicate that without further refinement the considered advanced state-of-the-art forecasting methods rarely beat corresponding persistence forecasts. Therefore, we also provide an exploration of promising directions for future research.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning

Keywords

Smart Grid; Smart Home; Load Forecasting

1. INTRODUCTION

In this study, we evaluate state-of-the-art forecasting methods for their applicability for household load forecasting. Accurate load forecasts can greatly enhance the micro-balancing capabilities of smart grids, if they are utilized for control operations and decisions like dispatch, unit commitment, fuel allocation and off-line network analysis [1]. Further, accurate load forecasts can help utilities to select customers that are suitable for demand response programs like proposed by [6]. First studies have analyzed the potential of consumption forecasts for individual households [7, 8]. However, most work focuses on disaggregation of electricity consumption (e.g., [2, 3, 4]). Our results show the forecasting methods provide little value, if not embedded into a framework that adapts to individual household attributes, motivating an exploration of promising directions for future research.

2. EXPERIMENT

The technical details of the data pre-processing and experimental setup are explained in the technical report [5].

e-Energy'14, Jun 11-13, 2014, Cambridge, United Kingdom.
ACM 978-1-4503-2819-7/14/06.
http://dx.doi.org/10.1145/2602044.2602082.

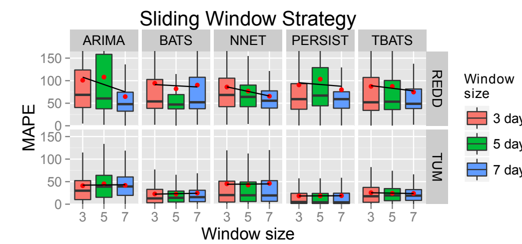

Figure 1: MAPE for varying window sizes.

We use two data sets measuring the power consumption of individual appliances of a household at intervals of 1 to 3 seconds. The TUM data set covers nine month and has a very stable consumption pattern. The Reference Energy Disaggregation Data Set (REDD) [4] covers 18 days and has more frequent and higher fluctuations in consumption. We use different forecasting methods that are all provided by the R `forecast` package: As benchmark, we use the persistence method, where forecasts equal the last observation. Further, we use Autoregressive Integrated Moving Average (ARIMA), i.e., `auto.arima()`, exponential smoothing state space models, i.e., `bats()` and `tbats()` and feed-forward neural networks with a single hidden layer, i.e., `nnetar()`. We used three sampling strategies: The *sliding window strategy* divides the data set into windows, moving forward on the data, after a model has been trained and tested. The *day type strategy* joins each day of the week of consecutive weeks into separate data sets. The *hierarchical day type strategy* first forecasts individual appliances to then compute the aggregated forecast. We use data granularities from 15 to 60 minutes intervals, forecasting horizons from 15 minutes to 24 hours and window sizes from 3 to 7 days. We measure the model quality by the Mean Absolute Percentage Error (MAPE), because it is a relative measure and can be used to compare the performance on different data sets.

3. EXPERIMENTAL RESULTS

Overall we observe that the considered advanced state-of-the-art *forecasting methods rarely beat corresponding persistence forecasts.* This is especially true for the TUM data set. Further, our *results differ largely between the two data sets*, i.e., the accuracy on the TUM data set is almost constantly higher than on the REDD data set. This could be due to the more stable consumption pattern in the TUM data set, which is easier to predict. In addition, Figure 1 shows boxplots and linear trend lines of MAPE for different

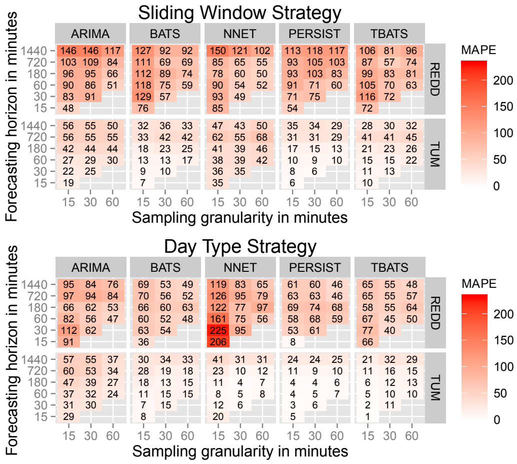

Figure 2: MAPE for varying horizons and granularities.

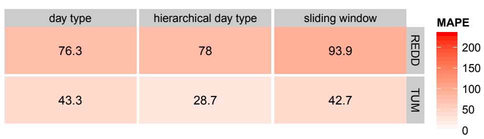

Figure 3: Mean MAPE for ARIMA with different strategies.

window sizes in the sliding window strategy. The results indicate that *increasing window sizes improve the results* of the ARIMA, NNET and TBATS methods on the REDD data set, but not on the TUM data set. Because the days in the TUM data are so similar, additional training data might not provide new important information. Further, Figure 2 shows heatmaps of MAPE from the sliding window and day type strategies for different granularities and forecasting horizons. The results indicate that for almost every method, *longer forecasting horizons lead to lower accuracy*. However, the exponential smoothing methods BATS and TBATS seem more robust against increasing horizons than the other methods. Further, especially on the REDD data set, *lower granularities lead to better accuracy*. In particular, the exponential smoothing strategies BATS and TBATS and the neural network outperform the persistence method for granularities of 30 and 60 minutes. Furthermore, Figure 2 shows that for almost every method *a division of the data into day type windows improves the forecast accuracy*. In addition, Figure 3 compares all three strategies for the ARIMA method indicating that the *hierarchical strategy can greatly improve accuracy* on the TUM data set. This is a surprising result, as generally the prediction of aggregated loads tend to result in higher precision.

4. DISCUSSION AND FUTURE WORK

We have evaluated a wide range of state-of-the-art methods and strategies for short-term forecasting of household electricity consumption based on actual data. Although our current data is limited, we were able to gain useful insights. Overall, the considered advanced forecasting methods only rarely beat the accuracy of persistence forecasts. Further, most of the methods benefit from larger training sets, splitting the data into sets of particular day types and predicting based on disaggregated data from individual appliances. Furthermore, the achievable accuracy in terms of average MAPE is surprisingly low, ranging between 5 and 150%. Thus, our work motivates more research investigating how accuracy can be increased. First, introducing further features, e.g., from occupancy, temperature or brightness sensors, could improve prediction accuracy, because when a device is switched on/off it takes time until the average wattage accounts for the change. Second, many devices have a very predictable consumption pattern once switched on. Thus, it could be beneficial to detect concrete events (e.g., on/off) and based on these events derive a future consumption pattern. Third, we only considered consistent data sets. However, in real world settings load forecasts need to be performed even in situations with missing data. Thus, future work should investigate how to handle temporary sensor outages, which could distract the prediction algorithms. Last, our results differ largely between the two data sets. It is unclear, how common the characteristics of these data sets are. However, the necessary data for carrying out more representative studies is currently missing. Future work will focus on the design of such frameworks.

5. REFERENCES

[1] D. Bunn and E. Farmer. Review of short-term forecasting methods in the electric power industry. *Comparative models for electrical load forecasting*, 1985.

[2] S. Humeau, T. K. Wijaya, M. Vasirani, and K. Aberer. Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. In *Sustainable Internet and ICT for Sustainability (SustainIT), 2013*. IEEE, 2013.

[3] W. Kleiminger, C. Beckel, T. Staake, and S. Santini. Occupancy Detection from Electricity Consumption Data. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*. ACM, 2013.

[4] J. Z. Kolter and M. J. Johnson. Redd: A public data set for energy disaggregation research. In *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, 2011.

[5] A. Veit, C. Goebel, R. Tidke, C. Doblander, and H.-A. Jacobsen. Household electricity demand forecasting - benchmarking state-of-the-art methods. Technical Report arXiv:1404.0200, arXiv.org, 2014.

[6] A. Veit, Y. Xu, R. Zheng, N. Chakraborty, and K. Sycara. Multiagent coordination for energy consumption scheduling in consumer cooperatives. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*. AAAI, July 2013.

[7] H. Ziekow, C. Doblander, C. Goebel, and H.-A. Jacobsen. Forecasting household electricity demand with complex event processing: insights from a prototypical solution. In *Proceedings of the Industrial Track of the 13th ACM/IFIP/USENIX International Middleware Conference*. ACM, 2013.

[8] H. Ziekow, C. Goebel, J. Struker, and H.-A. Jacobsen. The potential of smart home sensors in forecasting household electricity demand. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2013.