Ensemble-Based Model Selection for Smart Metering Data

Martina Friese, Thomas Bartz-Beielstein, Katya Vladislavleva, Oliver Flasch, Olaf Mersmann, Boris Naujoks, Jörg Stork, Martin Zaefferer

Fakultät für Informatik und Ingenieurwissenschaften, Fachhochschule Köln, 51643 Gummersbach
`vorname.nachname@fh-koeln.de`

In times of accelerating climate change and rising energy costs, increasing energy efficiency becomes a high priority goal for businesses and private households alike. Smart-metering equipment records energy consumption data in regular intervals multiple times per hour, streaming this data to a central system, usually located at a local public utility company. Here, consumption data can be correlated and analyzed to detect anomalies such as unusually high consumption.
This paper describes results from an on-going project with GreenPocket GmbH (`greenpocket.de`), one of the leading players in smart metering and smart home in Germany. GreenPocket develops software solutions that preprocess the data to give consumers an insight to their consumption habits and also provides forecasts for the future. Since the underlying forecast models are relatively simple, it is of great interest to compare existing with sophisticated modeling approaches.

The final goal of our study is to analyze the benefit of ensemble-based approaches, i.e., approaches that evaluate several time-series models in parallel and allow an adaptation of the current model based on new features and trends in the actual data. In theory, model adaptation improves the prediction accuracy. We will present a sound experimental study, which is based on real-world data, to analyze the usability of our approach in a real-world setting.

The experiments are run on set of four energy consumption time series, supplied by GreenPocket GmbH. Each time series is a series of timestamp and meter reading pairs taken quarter hourly over a time period of about 4 months, meter readings are given in kilowatt hours (kWh). The data was recorded by two independent smart metering devices, installed at a commercial customer.

The method that we focus on is implemented in the R forecast package which includes methods like exponential smoothing via state space models and automatic ARIMA modelling. [1] This method employs a basic ensemble since it automatically chooses a model at the beginning.

At the first stage of our study, this simple ensemble method will be compared to the method currently employed by GreenPocket GmbH. The model selection will be dynamically adapted (learned) during the second stage of our study. In addition to GreenPocket's approach and our ensemble approach a state of the art modeling approach based on Genetic Programming (GP), Evolved Analytics Data Modeler will be included in our study as the third modeling tool. [2]

First experiments show promising results as can be seen from Fig.1. Our study is of great value for both industry and academia, because prediction accuracy will be compared on solid ground: (i) models used in practice, (ii) modern statistical methods and (iii) data-driven modeling based on GP.

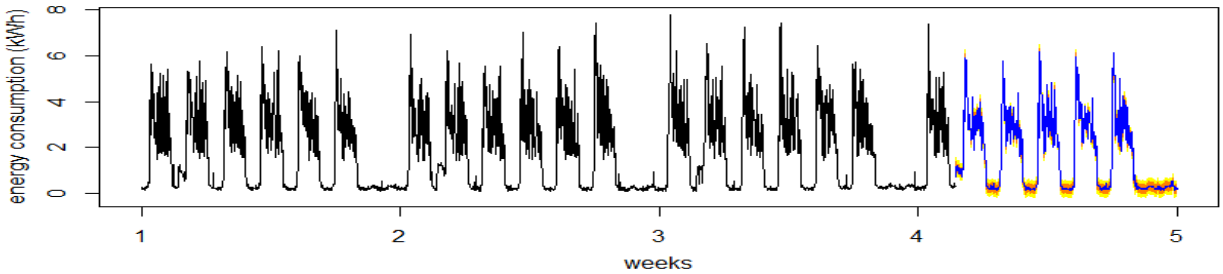

Figure 1: Forecast produced with the forecast package without any tuning of the method's parameters

References

[1] R. J. Hyndman and Y. Khandakar. Automatic Time Series Forecasting. The forecast Package for R. *Journal of Statistical Software*, 27(3):1–22, 7 2008.

[2] A. K. Kordon, G. Smits, and M. E. Kotanchek. Industrial evolutionary computing. In D. Thierens, editor, *GECCO (Companion)*, pages 3297–3322. ACM, 2007.