Personalized Probabilistic Load Forecasting using Gaussian Processes

Johannes Ridefelt
Division of Signals and Systems
Dep. of Electrical Engineering, Uppsala University
Uppsala, Sweden
johannes.ridefelt@angstrom.uu.se

Roland Hostettler
Division of Signals and Systems
Dep. of Electrical Engineering, Uppsala University
Uppsala, Sweden
roland.hostettler@angstrom.uu.se

I. Background

With an increasing number of distributed energy resources (DERs) and electric vehicles (EVs), the demands on the electricity grid are becoming higher and more complex. However, these changes also create an opportunity to increase grid efficiency and lower costs for both operators and individual clients. As more private homeowners and enterprises install complex local energy systems that combine electricity production, storage, and flexible loads—e.g., solar panels, battery-energy-storage systems, and EVs — the need for accurate, personalized electrical load forecasting is increasing.

Load forecasting is a well-studied problem, and many different approaches have been proposed. Time series analysis methods such as ARMA and ARIMA, regression based methods, and machine learning methods such as neural networks and support vector machines have all been used for load forecasting [1]. Uncertainty in load forecasting is one of the most important aspects, and various probabilistic forecasting methods have therefore been proposed, such as quantile regression [2] and multitask Bayesian deep learning [3]. These methods mostly rely on large amounts of data to train the models, making them less suitable for personalized load forecasting, where the amount of data and computational resources may be limited.

An alternative approach for probabilistic forecasting is to use Gaussian Processes (GPs), which have been used to model load profiles of enterprises [4] and to residential buildings [5].

GPs are collections of random variables that are jointly Gaussian distributed for any finite subset. The random variables are defined by a mean function and a covariance function. The mean function is often assumed to be zero, and the covariance function (or kernel) defines the relationships between the random variables. Commonly used kernel functions include the radial basis function kernel, the Matérn kernel, and various periodic kernels. Because all kernel functions must be positive definite, the sum or product of multiple kernels is also a valid kernel function.

By using GPs, the variability and heterogeneity in load profiles can be captured while providing probabilistic forecasts that account for uncertainty. This, inherent uncertainty quantification of GPs, is useful for planning the use of flexible loads and battery-energy-storage systems at grid level as well as on local energy management system level.

II. Method

A dataset of half-hourly load values spanning three years (2011–2014) from 300 residential buildings in Australia [6] serves as the basis for analysis. After removing clients with faulty or missing data, a total of 54 clients remained. Using min–max normalization the load values are scaled to the range 0–1.

To model the dataset GPs with a set of different kernels are employed. The kernels including a radial basis function kernel, the Matérn kernel, a periodic kernel, and all possible combinations of these.

As regressors, the time of the week and previous load values are used. The number of values from the previous timesteps, commonly referred to as lags, is varied. Using a single client at a time, we applied a time series variation of k-fold cross-validation to evaluate the performance of different GPs. The predictions are made one step ahead, meaning the prediction is made for the load 30 minutes ahead in time.

III. Results and Discussion

Preliminary results show that the GPs can model individual load profiles using only one previous load value and the time of the week as regressors. The mean absolute error is 0.0464, and the normalized Root mean square error is 0.1041 with a radial-basis-function kernel; similar results were obtained with other kernels.

The preliminary results show that GPs can model individual load profiles with only a few weeks of data and few regressors. Previous studies typically require orders of magnitude more training data, but further comparison and evaluation with probabilistic metrics are still needed. These findings nonetheless indicate that GPs are well-suited to personalized load forecasting when available data are limited.

Making a fair comparison of any results with non-probabilistic models is challenging as they lack any calibrated uncertainties to benchmark against. A similar challenge occurs when comparing with models and studies that use aggregated data, as aggregation generally dampen peaks and smooths variability, masking the very features a individualized model is designed to capture. In theory, individualized GPs preserve - and explicitly quantify - the heterogeneity of load profiles across clients, offering richer information.

Future work will include implementing a federated-learning approach to train GPs on multiple clients without sharing sensitive electrical data. This should enable better modelling at the individual client level, leveraging shared knowledge, in addition to providing accurate forecasting on a aggregated level, e.g. for an entire apartment building or neighbourhood - while using data from just a subset of residents.

References

[1] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustainable Cities and Society*, vol. 35, pp. 257–270, 2017.

[2] S. Ben Taieb, R. Huser, R. J. Hyndman, and M. G. Genton, "Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2448–2455, 2016.

[3] Y. Yang, W. Li, T. A. Gulliver, and S. Li, "Bayesian deep learning-based probabilistic load forecasting in smart grids," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4703–4713, 2020.

[4] P. Kou and F. Gao, "A sparse heteroscedastic model for the probabilistic load forecasting in energy-intensive enterprises," *International Journal of Electrical Power Energy Systems*, vol. 55, pp. 144–154, 2014.

[5] D. van der Meer, M. Shepero, A. Svensson, J. Widén, and J. Munkhammar, "Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using gaussian processes," *Applied Energy*, vol. 213, pp. 195–207, 2018.

[6] E. L. Ratnam, S. R. Weller, C. M. Kellett, and A. T. M. and, "Residential load and rooftop pv generation: an australian distribution network dataset," *International Journal of Sustainable Energy*, vol. 36, no. 8, pp. 787–806, 2017.