VALIDATION OF A HIDDEN MARKOV MODEL BASED PROBABILISTIC FORECASTING METHOD ON A HOUSEHOLD'S ELECTRIC POWER LOAD

Florian Schimek[1]*

Abstract

Probabilistic forecasting, unlike point forecasting, is able to quantify the uncertain characteristics of future events and is therefore well suited for highly volatile data like the electricity demand of a household. While probabilistic approaches typically involve higher implementation effort and computational complexity, a simple and efficient method based on Hidden Markov Models (HMM) was developed in a previous work [1]. This novel approach for predicting the distribution of the future electricity demand showed promising results and has the potential to become a competitive alternative to existing probabilistic forecasting methods.

This paper validates these findings at scale (>200 households) by evaluating the model across multiple large datasets and benchmarking it against advanced forecasting methods. In addition, its performance is analysed across different forecasting horizons to establish an understanding of its limitations.

Methodology

HMM Forecasting Method

The proposed forecasting approach is based on a Hidden Markov Model trained on historical load time series using the Baum–Welch algorithm [3]. During forecasting, the probability distribution of future observations is computed using the HMM forecasting algorithm described by [1]. As the model relies on discrete hidden states and observations, appropriate pre- and postprocessing is required, including an equal-mass discretization method that has been shown to improve predictive accuracy compared to equidistant binning.

Case Study

The case study is designed to ensure statistical robustness by repeating the experiments across more than 200 households from multiple data sources. For each test time step, predictive distributions are generated for several forecasting horizons and evaluated using the Continuous Ranked Probability Score (CRPS), see [4]. Further, we compute the mean-CRPS per household and horizon separately. Results are averaged by data source and benchmarked against state-of-the-art probabilistic methods, including two heuristic baselines, a long short-term memory model [5], and quantile regression averaging approach [2].

Preliminary Results

Initial results indicate that the proposed HMM-based forecasting method significantly outperforms heuristic approaches for short forecasting horizons, achieving improvements of up to 30%. These findings underline the predictive power of the proposed method and verify the results of [1].

However, predictive accuracy deteriorates rapidly as the forecasting horizon increases, and beyond six time steps - corresponding to 1.5 hours at a 15-minute resolution - the method is outperformed by a daytime-based sampling approach. This behaviour is consistent with the convergence properties of Markov processes. Benchmarking against machine learning methods is currently ongoing.

[1] TU Wien, florian.schimek@gmx.at

Referenzen

[1] F. Schimek, “Probabilistic forecasting using Hidden Markov models: A use-case for household electrical power load forecasting” Diploma thesis, TU Wien, 2025
doi:

[2] Liu B, Nowotarski J, Hong T, Weron R. Probabilistic load forecasting via quantile regression averaging on sister forecasts. IEEE Trans Smart Grid 2017;8(2):730–7.
doi: https://ieeexplore.ieee.org/document/7137662

[3] L. Rabiner. „A tutorial on hidden Markov models and selected applications in speech recognition“. In: Proceedings of the IEEE 77.2 (1989), pp. 257–286.

[4] T. Gneiting and M. Katzfuss. „Probabilistic Forecasting“. In: Annual Review of Statistics and Its Application 1.Volume 1, 2014 (2014), pp. 125–151.
doi: https://doi.org/10.1146/annurev-statistics-062713-085831.

[5] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang. „Probabilistic individual load forecasting using pinball loss guided LSTM“. In: Applied Energy 235 (2019), pp. 10–20.
doi: https://doi.org/10.1016/j.apenergy.2018.10.078.