Applied Energy 213 (2018) 195–207

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes

D.W. van der Meer[a,*], M. Shepero[a], A. Svensson[b], J. Widén[a], J. Munkhammar[a]

[a] *Built Environment Energy Systems Group (BEESG), Division of Solid State Physics, Department of Engineering Sciences, Uppsala University, P.O. Box 534, SE-751 21 Uppsala, Sweden*

[b] *Division of Systems and Control, Department of Information Technology, Uppsala University, P.O. Box 337, SE-751 05 Uppsala, Sweden*

HIGHLIGHTS

- Probabilistic forecasting of an individual building using Gaussian Processes.
- We assess the performance of multiple covariance functions.
- We examine the difference between a static and dynamic Gaussian Process.
- We explore two strategies for net demand forecasting.

ARTICLE INFO

Keywords:
Gaussian Processes
PV
Residential electricity consumption
Net demand
Probabilistic
Forecasting

ABSTRACT

This paper presents a study into the utilization of Gaussian Processes (GPs) for probabilistic forecasting of residential electricity consumption, photovoltaic (PV) power generation and net demand of a single household. The covariance function that encodes prior belief on the general shape of the time series plays a vital role in the performance of GPs and a common choice is the squared exponential (SE), although it has been argued that the SE is likely suboptimal for physical processes. Therefore, we thoroughly test various (combinations of) covariance functions. Furthermore, in order bypass the substantial learning and inference time accompanied with GPs, we investigate the potential of dynamically updating the hyperparameters using a moving training window and assess the consequences on predictive accuracy. We show that the dynamic GP produces sharper prediction intervals (PIs) than the static GP with significant lower computational burden, but at the cost of the ability to capture sharp peaks. In addition, we examine the difference in accuracy between a direct and indirect forecasting strategy in case of net demand forecasting and show that the latter is prone to producing wider PIs with higher coverage probability.

1. Introduction

Renewable energy sources (RESs), such as solar and wind power, are steadily underway to become substantial shares of the energy mix. In 2016, 75 GW of photovoltaic (PV) power capacity was installed, up from 50 GW installed PV power capacity in 2015, bringing the cumulative installed capacity to at least 303 GW, amounting to 1.8% of worldwide electricity production [1]. However, the consumption of PV power is significantly higher in some countries. A common example is Germany, where during 2016 7.4% of the electricity consumption was covered by PV power, which increased to 35% on sunny weekdays [2]. Increasing penetration of PV power into the electricity mix brings with it challenges such as grid losses, feeder loading and voltage fluctuations [3–5]. It is estimated that penetration levels of 40–50% can cause severe voltage fluctuations caused by e.g., variability due to cloud cover [6]. Accurate forecasts of PV power are generally viewed as a cost-efficient way to mitigate the aforementioned issues because they allow for e.g., unit commitment or curtailment, although these forecasts are challenging due to the stochastic nature of cloud cover and weather phenomena in general [7].

Electricity consumption on the aggregated scale is less uncertain and can currently be predicted with a mean absolute error (MAE) of around 3% in the day-ahead market [8]. However, increasingly stochastic behavior of residential consumers due to e.g., electric vehicles

* Corresponding author.
E-mail address: dennis.vandermeer@angstrom.uu.se (D.W. van der Meer).

https://doi.org/10.1016/j.apenergy.2017.12.104
Received 11 October 2017; Received in revised form 5 December 2017; Accepted 30 December 2017

(EVs) and electronic devices, and increasing PV power penetration can create local issues that are challenging to forecast using an aggregated approach. Moreover, Zamo et al. found that a bottom-up approach can improve the MAE by 3% [9]. Following from this is the opportunity to investigate net demand forecasting, defined as electricity consumption minus PV power generation, on a disaggregated level.

In order to properly quantify the variability of PV power, electricity consumption and net demand, and to express the uncertainty of predicting these quantities, Gaussian Processes (GPs) will be used in this study.

1.1. Previous work

Probabilistic solar power forecasting (PSPF) and probabilistic load forecasting (PLF) have been extensively reviewed in [10,11]. From these studies it becomes apparent that a vast variety of models have been employed on different temporal and spatial resolutions. For example, Scolari et al. [12,13] proposed a nonparametric model to predict irradiance with a lead time of 100–500 ms, based on the correlation between forecast errors and the derivative of irradiance. An interesting approach was taken by Ni et al. [14], who used an ensemble of extreme learning machines (ELMs) in combination with the lower upper bound estimation (LUBE) approach to predict one step ahead, which was 5 min. The authors clearly showed the benefit of the ensemble approach, especially since the ELMs, although faster to train than artificial neural networks (ANNs) are less stable. Since the performance metrics are similar to those used in the present study, we will compare our results with those achieved by Ni et al. On a similar timescale is the study performed by Sanjari and Gooi [15], who used a higher order Markov Chain (HMC) to forecast PV power generation at the next time step. The authors combined several HMCs at different PV system operating points to form an ensemble prediction and then utilized the Gaussian mixture method (GMM) to create a non-parametric density. They outperformed the benchmarks and achieved a continuous ranked probability score (CRPS) of 2.16, although units were not specified. On a lower temporal resolution of one hour was the study performed by Nagy et al. [16], in which they used a voted ensemble of quantile regression forests (QRFs) and stacked random forests (RFs) - gradient boosted decision trees (GBDTs). In order to forecast irradiance with an intra-day lead time, the authors utilized a substantial number of explanatory variables, extracted from a numerical weather prediction (NWP) forecast model. Wang et al. [17] utilized a deep convolutional neural network (DCNN) to produce a deterministic forecast and used spline quantile regression (QR) to produce probabilistic forecasts of the production of PV farms in Belgium. A common issue with ANNs is the large number of parameters to be optimized, and therefore the authors selected DCNN, since this particular model had relatively fewer parameters to be optimized. They achieved a CRPS ranging from 0.23 during winter for a 15 min forecast horizon, to 19.97 during summer with a 2 h forecast horizon. No units for CRPS were specified. Bracale et al. [18] created an ensemble with QR, Bayesian model (BM) and a Markov chain model (MC). They aggregated the base predictors by applying a linear pool ensemble model and used multi-objective optimization to create both sharp and reliable probabilistic forecasts, since these are conflicting. The CRPS of the proposed model was 5.24 kW in February 2014, and the ensemble method showed substantial improvement over the three models separately. An excellent example of what probabilistic forecasts can be used for was given by Appino et al. [19], who used such forecasts in a stochastic optimization framework for reliable power scheduling. More specifically, they incorporated a security factor into their problem formulation, which was directly related to the uncertainty of the probabilistic forecast. The authors found that by adjusting the security level, they could improve the operating cost of the micro-grid when compared to using a deterministic approach.

With respect to load forecasting, similar observations can be made in terms of model variety. For example, autoregressive models such as autoregressive integrated moving average (ARIMA), vector AR (VAR) or AR were employed on various time scales in [20–22]. ANNs have also been employed to create prediction intervals (PIs) using the lower upper bound estimate technique in [23,24]. An interesting approach to predict the net load was taken by Wang et al. [25], who used GBDTs with high penetration of behind-the-meter (BtM) PV. The main idea of their study was to decompose the time series into PV output, electricity usage and residuals under the assumption that these would be more straightforward to forecast separately. However, no verification of this assumption was carried out in the paper. As a final step, the forecasts were aggregated using the dependent discrete convolution method and showed that their proposed method outperformed the benchmarks, which were based on QR. Carbera and Schulz [26] forecast electricity demand of a transmission system operator (TSO) in Germany using a vector autoregressive model with exogenous inputs (VARX). Unfortunately however, the authors did not specify the performance of their model in terms of probabilistic error metrics. Only a handful of studies were aimed at forecasting residential electricity consumption. One with half hourly resolution was performed by Taieb et al. [27], in which QR combined with gradient boosting (GB) was used. Another example is the study performed by Arora and Taylor [28], in which they used conditional kernel density (CKD) estimation to forecast residential demand. Both the aforementioned studies utilized additional explanatory variables such as temperature or calender variables.

The review studies [10,11] showed that GPs have not received much attention in recent years. More specifically, four studies have been found that utilized GPs in case of irradiance forecasting. Salcedo-Sanz et al. [29] employed GPs to forecast daily global solar irradiation using the periodic covariance function for time as an explanatory variable and the squared exponential (SE) covariance function for the remaining meteorological variables such as ozone, water vapor and the presence of clouds using NWP outputs. The time-based GP showed to outperform other data-driven methods such as support vector regression (SVR), GBDTs and ELMs, and additionally showed to have greater robustness over 100 random splits of the training and test data set. Bilionis et al. [30] employed GPs to predict solar irradiation using satellite images as input, after first using factor analysis (FA) to reduce the number of dimensions of the images. When compared with a GP based solely on ground observations, the proposed model produced narrower PIs and lower CRPS, with an average of 0.18, although the units were not specified. Lauret et al. [31] benchmarked several machine learning techniques to forecast clear sky index (CSI) on three different islands. No probabilistic error measures were used to assess the performance, but the GP model showed the overall best performance as it led to the best results most of the time. Finally, Sheng et al. [32] utilized weighted GP regression that incorporated outlier detection to predict one step ahead with 5 min resolution. In their study, the authors aimed at improving the quality of the data set by giving low weight to data samples with high outlier potential. Unfortunately, the authors did not apply any of the standard probabilistic performance metrics described in [10], but did calculate the mean PI width, which varied between 287 W and 697 W.

With respect to load forecasting, Lauret et al. [33] compared ANNs and Bayesian ANNs to the GP and showed that the GP outperformed the other two methods on their data set, unfortunately without using any probabilistic performance metrics. Kou and Gao [34] utilized a sparse heteroscedastic GP (HGP) model in energy-intensive enterprises, due to the fact that the load level does not remain constant and therefore a second GP was trained on the empirical noise levels of the training data. One issue with GPs is the high computational burden due to inversion of the covariance matrix, which costs $\mathcal{O}(N^3)$ for training and inference. This is an important consideration in case GPs would be introduced in an online fashion, since the inference time should not exceed the forecast horizon. Therefore, the authors aimed to sparsify the data set via regularization in order to reduce computation time. It showed to