



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

D.W. van der Meer<sup>a,\*</sup>, M. Shepero<sup>a</sup>, A. Svensson<sup>b</sup>, J. Widén<sup>a</sup>, J. Munkhammar<sup>a</sup>

<sup>a</sup> 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

<sup>b</sup> 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 to 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](mailto:dennis.vandermeer@angstrom.uu.se) (D.W. van der Meer).

(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 calendar 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. Bilonis 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

have significant impact on training time and higher accuracy than the standard GP model, although the performance with respect to HGP is unclear. Dong et al. [35] forecast residential electricity consumption based on data-driven models, e.g., GPs, physical models that describe the houses and a hybrid model. The GP consistently showed poor performance in terms of mean absolute percentage error (MAPE) and root mean square error (RMSE) and unfortunately the authors did not elude as to why that was. An interesting application of GPs was done by McLoughlin et al. [36], who used the model to characterize residential electricity consumption on dwelling level and concluded that it was a suitable technique. Moreover, it was found that GPs performed well in the case of customers that had sharp increments in electricity consumption. Finally, Lloyd [37] utilized GP in combination with GB to perform a hierarchical load forecast for the Global Energy Forecasting Competition 2012. The author admits to not have tested covariance functions but interestingly has used several complicated ones for backcasting and forecasting. More specifically, the author used a linear combination of the squared exponential (SE) and periodic covariance function. As a final step, the author combined the forecasts of the separate models into an ensemble to improve the overall accuracy, whilst mentioning that ensembling is a particularly effective method in case the predictions are uncorrelated.

An important observation from the studies in our literature review is that none of them, except for Bilonis et al. [30], did extensive testing with different covariance functions, even though this encodes the relationship between any of the input variables, which we denote  $\mathbf{x}$ . We return to this issue in Section 2.5.

Apart from the previously mentioned study by Appino et al. [19], several other studies have been performed that clearly presented the usage and benefit of probabilistic forecasts. For example, Ela et al. [38] showed how such probabilistic forecasts of renewable sources could be used in stochastic optimization and dynamic reserve requirements, which benefited cost savings and reliability of the system. Similarly, Sarshar et al. [39] used the uncertainty expressed by a probabilistic wind power forecast to perform multi-objective optimization of power flows in a micro-grid containing PV, fuel cells, wind turbines and energy storage systems. And finally, Al-Sumaiti et al. [40] showed how uncertainty in temperature forecasts could aid an optimal scheduling strategy of electricity supply in developing countries. Concluding, we can observe that probabilistic forecasts can provide significant added value to decision making processes where the incorporation of uncertainty is vital.

## 1.2. Aim and contributions

The aim of this study is threefold. Firstly, rather than assuming the most appropriate covariance function to be the SE, we test several covariance functions and combinations thereof. Secondly, we compare recursive GPs to static GPs and use these to predict residential electricity consumption, rooftop PV power production and net demand. As mentioned before, only few papers have focused on these aspects. And finally, inspired by Shaker et al. [41], we aim to examine the difference in accuracy between a direct and indirect forecasting strategy, where a direct strategy constitutes forecasting net demand, whereas the indirect strategy implies forecasting electricity consumption and PV power generation separately and subtracting these afterwards. The contributions of this study can therefore be summed up as follows:

1. We employ GPs to model and predict residential electricity consumption, rooftop PV power generation and net demand, i.e., of an individual building. This has not been done before to the best of our knowledge.
2. Unlike the previously mentioned studies, we thoroughly test several (combinations of) covariance functions in order to find the most appropriate one.
3. Both static and recursive GPs are tested and compared, where the

latter is updated every  $i$ th iteration in order to account for seasonal variations, which has not been done before.

4. We investigate two strategies in case of net demand forecasting, i.e., a direct and indirect one, on residential scale. This has not been done before to the best of our knowledge.

The rest of the paper is organized as follows. In Section 2, we introduce the data set and give a general introduction to GPs and specifically their application to time series. In Section 3 we present the results, discuss these in Section 4 and finally in Section 5, we conclude this study.

## 2. Methodology

The present section introduces the data set, the GP and the benchmark ARIMA model, various performance metrics and  $k$ -fold cross validation for time series. In order to implement the GP, we utilize the `GPML` package in MATLAB that is publicly available with the book by Rasmussen and Williams [42]. For the ARIMA model we use the `forecast` package [43] in R.

### 2.1. Data

The data used in this study comprises residential electricity consumption and rooftop PV generation from the Sydney metropolitan area, Australia [44]. This data set is publicly available and contains 300 de-identified residential customers, of whom their consumption and production have been measured on half hourly basis from 1 July 2010 until 30 June 2013. After a thorough data cleaning process, described in [45], data of 54 customers for three years remain, of which we have randomly selected customer number 69. More information regarding the data set and the cleaning process can be found in Ratnam et al. [45].

Fig. 1 shows the time series of customer 69 over the course of three years. It is interesting to note that there exists a daily pattern in the electricity consumption data although the times at which peaks occur are highly variable. We can also observe that net demand frequently becomes negative, especially on sunny days. As a final note on the data, we present the variances of the normalized time series, which are 0.005055, 0.08154 and 0.006018 for electricity consumption, PV power generation and net demand, respectively. Although the variance of the net demand time series has slightly increased with respect to electricity consumption due to the addition of PV power, total variance is still substantially less than that of PV power generation alone. It will therefore be interesting to study the effect this has on the direct and indirect strategies.

### 2.2. Gaussian Processes

The definition of the GP states that it is a collection of random variables, and that any subset of these random variables has a joint multivariate Gaussian distribution with mean  $\mu$  and covariance matrix  $\mathbf{K}$ . This brings with it an advantage over other machine learning techniques where computations are analytically intractable and can only be approximated [42]. More intuitively, suppose  $f(x_1)$  and  $f(x_2)$  at  $x_1$  and  $x_2$  are jointly Gaussian distributed according to  $\mathcal{N}(\mu, \mathbf{K})$ . The GP allows us to determine the shape of the underlying function  $f$  in a probabilistic manner. However, this does not have to be limited to two observations and can be extended to any finite number  $N$  of inputs  $\mathbf{x} = \{x_1, \dots, x_N\}$ , and consequently the covariance matrix  $\mathbf{K}$  can be defined as [46]:

$$\mathbf{K}(\mathbf{x}, \mathbf{x}) = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_N) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_N, x_1) & k(x_N, x_2) & \cdots & k(x_N, x_N) \end{pmatrix}, \quad (1)$$

where  $k(x_i, x_j)$  is a covariance function, or kernel, that represents the

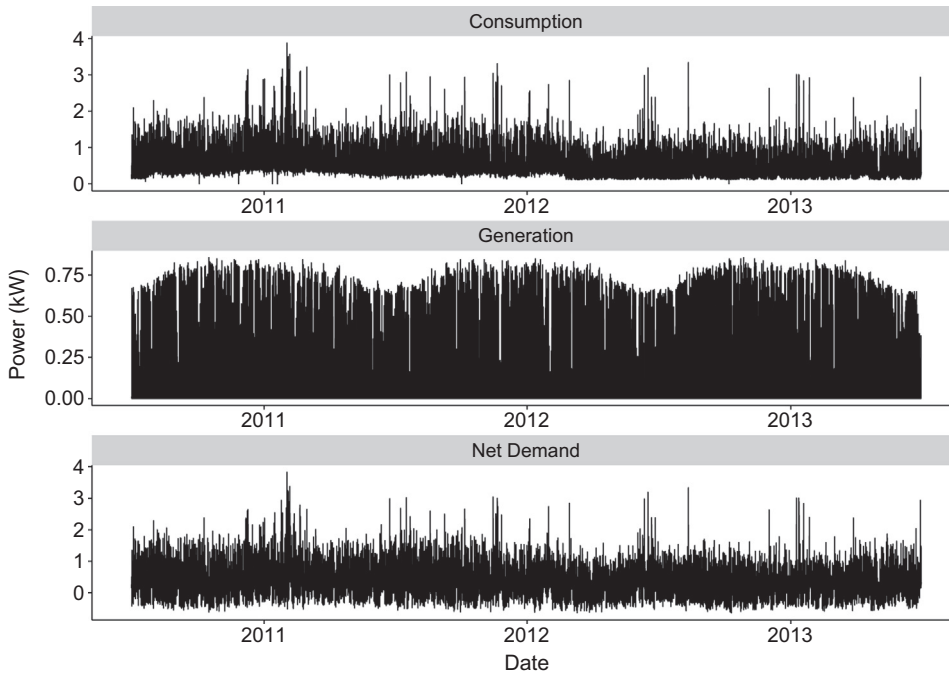


Fig. 1. Electricity consumption, PV power generation and net demand data from customer 69 on half hourly basis, from 1 July 2010 until 30 June 2013.

correlation between any two of the inputs  $\mathbf{x}$ . The hyperparameters defining the behavior of the covariance function are learned by maximizing the log marginal likelihood, defined as [42]:

$$\log p(\mathbf{y}|\mathbf{x},\theta) = -\frac{1}{2}\mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K}| - \frac{N}{2} \log 2\pi, \quad (2)$$

where  $\theta$  represents the vector containing the hyperparameters. As becomes apparent, the choice of covariance function is an important one since it represents the relationship between any of the inputs in  $\mathbf{x}$ . We come back to this issue in Section 2.5. For an in-depth discussion on covariance functions and the hyperparameters, the interested reader is referred to [42]. The multivariate Gaussian distribution amounts to:

$$p(\mathbf{f}(\mathbf{x})) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{x}), \mathbf{K}(\mathbf{x}, \mathbf{x})), \quad (3)$$

where  $\boldsymbol{\mu}(\mathbf{x})$  is the mean function, although this is frequently set to zero [42].

In case a new observation is made, e.g.,  $x_*$ , the covariance between  $x_*$  and  $\mathbf{x}$  can be formulated as follows

$$\mathbf{K}(x_*, \mathbf{x}) = (k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_N)). \quad (4)$$

Following from this, we can compute the posterior distribution by first defining the new joint distribution

$$p\left(\begin{bmatrix} \mathbf{f} \\ f(x_*) \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}(\mathbf{x}) \\ \boldsymbol{\mu}(x_*) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{x}, \mathbf{x}) & \mathbf{K}(\mathbf{x}, x_*) \\ \mathbf{K}(x_*, \mathbf{x}) & k(x_*, x_*) \end{bmatrix}\right). \quad (5)$$

Finally, the posterior distribution can be computed according to

$$\mathbf{m}_* = \boldsymbol{\mu}(x_*) + \mathbf{K}(x_*, \mathbf{x}) \mathbf{K}(\mathbf{x}, \mathbf{x})^{-1} (\mathbf{f} - \boldsymbol{\mu}(\mathbf{x})), \quad (6)$$

$$\sigma_*^2 = k(x_*, x_*) - \mathbf{K}(x_*, \mathbf{x}) \mathbf{K}(\mathbf{x}, \mathbf{x})^{-1} \mathbf{K}(\mathbf{x}, x_*). \quad (7)$$

The functioning of the GP is presented in Fig. 2. It should be noted that in this paper we focus on one-step ahead predictions as a first step and leave multiple-step ahead forecasts for future work. For more information on multiple-step ahead prediction with GPs, the interested reader is referred to Girard et al. [47].

### 2.2.1. GPs for time series modeling

The GP introduced in the previous section is mainly specified for regressive settings where a particular ordering of the data is not always present. This section will therefore introduce the GP model used in this

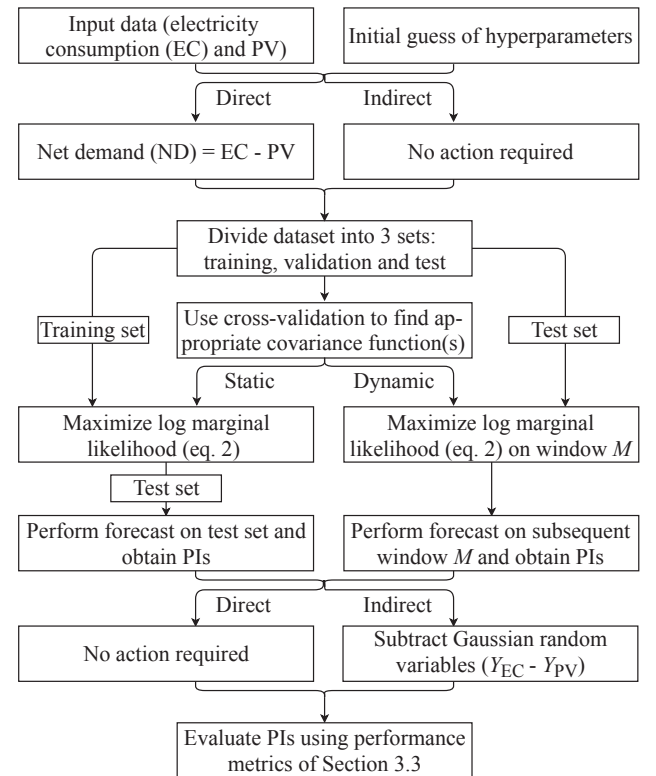


Fig. 2. Flowchart representing the functioning of the GP utilized in this study.

study, which can be applied to time series in an autoregressive setting. Firstly, we define the test input, i.e., latest observation  $x_* = x_t$ , such that a vector of lagged observations would entail  $x_{t-i} = (y_{t-1-i}, y_{t-2-i}, \dots, y_{t-L-i})$ , where  $L$  represents the number of lags we incorporate and  $i$  indicates the moment in the past we are interested in. Then, we define the training inputs  $\mathbf{x}$  as  $\mathbf{x} = (x_{t-1}, x_{t-2}, \dots, x_{t-N})^T$ . In this setting, Eqs. (1), (4) can be reformulated as follows



$$\mathbf{K}(\mathbf{x}, \mathbf{x}) = \begin{pmatrix} k(x_{t-1}, x_{t-1}) & \cdots & k(x_{t-1}, x_{t-N}) \\ \vdots & \ddots & \vdots \\ k(x_{t-N}, x_{t-1}) & \cdots & k(x_{t-N}, x_{t-N}) \end{pmatrix}, \quad (8)$$

$$\mathbf{K}(x_t, \mathbf{x}) = (k(x_t, x_{t-1}), \dots, k(x_t, x_{t-N})). \quad (9)$$

Finally, Eqs. (6) and (7) can then be defined as:

$$\hat{y}_t = \mu(x_t) + \mathbf{K}(x_t, \mathbf{x})\mathbf{K}(\mathbf{x}, \mathbf{x})^{-1}(\mathbf{f} - \mu(\mathbf{x})), \quad (10)$$

$$\hat{\sigma}_t = k(x_t, x_t) - \mathbf{K}(x_t, \mathbf{x})\mathbf{K}(\mathbf{x}, \mathbf{x})^{-1}\mathbf{K}(\mathbf{x}, x_t), \quad (11)$$

where  $\mathbf{f} = (y_{t-1}, \dots, y_{t-N})^T$ . However, as stated before, time to learn the hyperparameters and inference from the GP increases cubic when more data becomes available. Moreover, once the hyperparameters are learned, they remain fixed until they are updated, which is most likely suboptimal due to the seasonalities in the time series presented in Section 2.1. Consequently, we introduce a dynamic GP in order to quantify the difference in accuracy when compared to a static GP.

### 2.2.2. Dynamic GP

The two main reasons for introducing the dynamic GP are to reduce computational burden and to account for seasonal variations in the data set. The computational burden is important to consider, since practical applications of online forecasts require the inference time to be shorter than the forecast horizon in case the forecasts are used by e.g., a smart decision making algorithm. To facilitate this, we define  $\mathbf{x}$  as  $\mathbf{x} = (x_{t-1}, x_{t-2}, \dots, x_{t-M})^T$ , where  $M$  represents the length of the moving training window. In contrast to e.g., ANNs for which one requires an extensive training set, according to Salcedo-Sanz et al. [29], the GP model does not require this and therefore  $M$  can be significantly smaller than  $N$ , such that  $M \ll N$ .

It should be noted that this approach is similar to the one taken by Yan et al. [48], who utilized GPs to forecast wind power production, although their moving training window was small ( $M \in [2, 7]$ ) and preliminary tests revealed that more data was required in case of the time series under investigation here. Another difference lies in the learning of the hyperparameters. Yan et al. [48] employed the sum of squared errors (SSE) instead of the marginal likelihood, although they did not specify the reason for this. However, Rasmussen and Williams concluded that learning hyperparameters through SSE seems unattractive due to the fact that it ignores the variance of the prediction, and only takes the mean into account [42]. Moreover, the authors of [48] did not clarify the choice for the SE covariance function.

### 2.2.3. Explanatory variables

The present study only takes endogenous explanatory variables into account as a first step to assess the performance and viability of GPs in forecasting such stochastic time series. More specifically, in order to predict  $\hat{y}_t$ , we test several alternatives for explanatory variables that are mostly composed of sets of time lags. Based on preliminary testing, a set of explanatory variables is then selected for further comparison. Table 1 presents the compared alternatives of the explanatory variables in every case. It is important to note that the explanatory variables are tested separately for every case. In addition, we select these particular time lags due to their typical correlation to time  $t$ , where it is perhaps helpful to remind that the temporal resolution is 30 min.

It should be noted that exogenous explanatory variables will likely improve the performance of the GP, as was shown for various case studies in an excellent introduction into the GP and time series modeling by Roberts et al. [46].

### 2.3. ARIMA

As a benchmark model we utilize the well known ARIMA model, first proposed by Box and Jenkins in 1970. Mathematically, the ARIMA model can be expressed as follows:

**Table 1**

Description of the explanatory variables compared in the three cases of forecasting.

Alternative	Explanatory variables
Electricity consumption	
Alternative 1	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i}, y_{t-35-i}, \dots, y_{t-48-i}, \Delta(y_{t-1-i}, \dots, y_{t-4-i}))$
Alternative 2	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i}, y_{t-35-i}, \dots, y_{t-48-i})$
Alternative 3	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i}, \Delta(y_{t-1-i}, \dots, y_{t-4-i}))$
Alternative 4	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i})$
Alternative 5	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i}, y_{t-47-i}, y_{t-48-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 6	$x_{t-i} = (y_{t-1-i}, \dots, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 7	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-6-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
PV power generation	
Alternative 1	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-3-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 2	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-4-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 3	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-5-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Net demand	
Alternative 1	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-3-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 2	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-4-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$
Alternative 3	$x_{t-i} = (y_{t-1-i}, \dots, y_{t-5-i}, \Delta(y_{t-1-i}, \dots, y_{t-2-i}))$

$$\phi(B)(1-B)^d X_t = \theta(B)\epsilon_t, \quad (12)$$

where  $B$  is the backward shift operator, such that  $BX_t = X_{t-1}$ . Furthermore,  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  describes the AR(p) process with  $\phi_1, \dots, \phi_p$  as parameters and  $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$  the non-seasonal MA(q) process with  $\theta_1, \dots, \theta_q$  as parameters [49].

The ARIMA model in this study is generated using the `auto.arima` function from the `forecast` package [43] in R. Consequently, we do not need to specify beforehand which features of the time series we take into consideration, since the aforementioned function automatically selects the most appropriate model for each problem.

### 2.4. Performance metrics

In order to assess the performance of the GP and ARIMA models, we employ both deterministic and probabilistic error metrics. In order to make comparisons across time series, independent of e.g., installed capacity, Madsen et al. [50] recommended normalizing using the installed capacity. This was corroborated by Hoff et al. [51], who argued that it would be most appropriate to normalize using the installed capacity if one were to use both the MAE and RMSE. Since normalizing using the installed capacity effectively implies normalizing using the range of observed or measured values, we utilize the latter definition as we also consider electricity consumption and net demand for which there is not one installed capacity. The MAE and MAPE can then be defined as:

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{y}_t - y_t|, \quad (13)$$

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \left| \frac{\hat{y}_t - y_t}{\max(y) - \min(y)} \right|, \quad (14)$$

where  $T$  is the length of the test time series,  $\hat{y}_t$  is the forecast value,  $y_t$  is the measured value and  $\max(y)$  and  $\min(y)$  represent the maximum and minimum value of time series  $y$ . The RMSE and normalized root mean square error (NRMSE) are defined as

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_t - y_t)^2}, \quad (15)$$

$$\text{NRMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T \left( \frac{\hat{y}_t - y_t}{\max(y) - \min(y)} \right)^2}. \quad (16)$$

It should be noted that the results of the cross validation procedure in Fig. 4a to c are presented as non-normalized, whereas the forecast results in Section 3 are presented as percentages.

The aim of any probabilistic forecast is to cover the measurements with the PIs, when considering a certain nominal confidence level  $\eta$ . In order to quantify the aforementioned we calculate the prediction interval coverage probability (PICP), formulated as follows [23]:

$$\text{PICP} = \frac{1}{T} \sum_{t=1}^T \epsilon_t, \quad (17)$$

where  $\epsilon_t$  is defined as:

$$\epsilon_t = \begin{cases} 1 & \text{if } y_t \in [L_t, U_t] \\ 0 & \text{if } y_t \notin [L_t, U_t], \end{cases} \quad (18)$$

where  $L_t$  and  $U_t$  represent the lower and upper bound of the prediction interval, respectively. Reliability of the forecast can be quantified by the PICP, which should be higher than  $\eta$  and otherwise regarded as invalid and discarded [52]. In this study we use  $\eta = 80\%$ .

It is however undesirable to utilize PICP as the sole measure for the quality of the forecast, since wide PIs guarantee high PICP but contain high uncertainty and therefore convey little information for stakeholders. Consequently, it is important to simultaneously assess the width of the PIs, which quantitatively expresses the sharpness, or informativeness, of the forecast. This can be done via the prediction interval normalized average width (PINAW), which is defined as follows [23]:

$$\text{PINAW} = \frac{1}{TR} \sum_{t=1}^T (U_t - L_t), \quad (19)$$

where  $R$  normalizes PI average width and represents the difference between the maximum observed value and minimum observed value.

Finally, we employ the normalized continuous ranked probability score (NCRPS). The CRPS is formulated such that it measures both reliability and sharpness. It can be shown that this score reduces to the MAE in case of a deterministic forecast, offering the opportunity to compare probabilistic and point forecasts directly [53]. The CRPS can be formulated as follows [53]:

$$\text{CRPS}(\hat{F}_t, y_t) = \int_{-\infty}^{\infty} (\hat{F}_t(z) - \mathbb{1}\{z \geq y_t\})^2 dz, \quad (20)$$

where  $\mathbb{1}$  is the Heaviside step function and  $\hat{F}$  is the cumulative distribution function (CDF) forecast. As can be seen from the formulation in Eq. (20), the CRPS considers the entire distribution of forecasts. Another advantage of the CRPS is the fact that it has the same unit as the variable that was forecast [53], which enhances interpretability of the score. Similarly as for MAPE and NRMSE, and deduced from [54], we can calculate NCRPS as follows

$$\text{NCRPS}(\hat{F}_t, y_t) = \frac{\text{CRPS}}{\max(y) - \min(y)}. \quad (21)$$

### 2.5. K-fold cross validation for time series

As stated before, the covariance function represents the relationship between any of the inputs  $\mathbf{x}$  and therefore requires special attention. The studies discussed in Section 1.1 all employed the SE covariance function, except for Kou and Gou [34], who used the automatic relevance detection (ARD) covariance function, and Lloyd [37], who utilized a linear combination of the SE and periodic covariance functions. More importantly, only Bilonis et al. [30] specified to have tested several covariance functions and then decided to utilize the SE. However, Stein [55] argued that this particular covariance function is too smooth, which is due to its infinitely differentiable character, and is therefore an unrealistic assumption for physical processes. As a consequence, he recommended the Matérn covariance functions instead.

**Table 2**

The selection of covariance functions to be used during  $k$ -fold cross validation.  $r = \|\mathbf{x} - \mathbf{x}'\|$  and  $r^* = \|(\sin(\mathbf{x}), \cos(\mathbf{x}))^T - (\sin(\mathbf{x}'), \cos(\mathbf{x}'))^T\|$ .

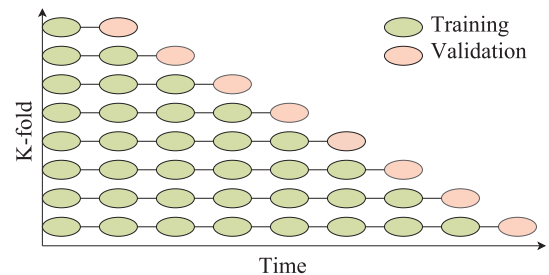
Covariance function	
$k_{\text{SE}}^1(r)$	$= \sigma_f^2 \exp\left(-\frac{r^2}{2\ell^2}\right)$
$k_{\text{M3}}^2(r)$	$= \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right)$
$k_{\text{M5}}^3(r)$	$= \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right)$
$k_{\text{SE-M3}}^4(r)$	$= k_{\text{SE}}^1(r) \cdot k_{\text{M3}}^2(r)$
$k_{\text{SE+M3}}^5(r)$	$= k_{\text{SE}}^1(r) + k_{\text{M3}}^2(r)$
$k_{\text{SE+M5}}^6(r)$	$= k_{\text{SE}}^1(r) + k_{\text{M5}}^3(r)$
$k_{\text{LIN-M3+Per-M3}}^7(r, r^*)$	$= \frac{r}{\sigma_f^2} k_{\text{M3}}^2(r) + k_{\text{M3}}^2(r^*)$

Table 2 presents the covariance functions we therefore examine in this study. Here, we test the SE, Matérn 3 and 5 and combinations thereof. More specifically, covariance functions 3–5 are combinations of the aforementioned ones and are tested because they allow for more freedom since there are more hyperparameters. We also test the periodic covariance function to examine whether this improves accuracy. As a final note, hyperparameters  $\sigma_f^2$  and  $\ell$  are to be learned by maximizing Eq. (2). It should be noted that other covariance functions or combinations thereof, e.g., by multiplication, led to unsatisfactory results and therefore we continue with this selection.

Another important consideration in time series modeling and prediction is the set of explanatory variables to incorporate. Several methods exist for this, such as information criteria, e.g., Akaike information criterion (AIC), or assessing the partial autocorrelation function (PACF). However, preliminary tests gave inconclusive results in this respect and we therefore employ  $k$ -fold cross validation for time series, also referred to as “forward chaining”, to decide on both the covariance function and the alternative of explanatory variables to consider, for more information see e.g., [56]. The core idea is to use a number of points as training data, e.g., 100, and use the learned parameters of the covariance function to predict the following 100 points. Subsequently, these 200 points are used as training data and the following 100 points are predicted, which is repeated until a certain constraint is satisfied. Fig. 3 illustrates the procedure, where the most important aspect is that the autocorrelation is kept in tact.

### 3. Results

The first year of the three year long data set is used to find the most appropriate covariance function and number of lags, using the  $k$ -fold cross validation procedure. The first half of the second year is subsequently used to learn the hyperparameters of the appropriate covariance function, after which we employ the GP to predict one step ahead for the second half of the second year. In addition, we use the moving window approach on the second half of the second year as well. However, this does not require learning the hyperparameters with the preceding 6 months, since these are systematically updated.



**Fig. 3.**  $K$ -fold time series cross validation procedure.

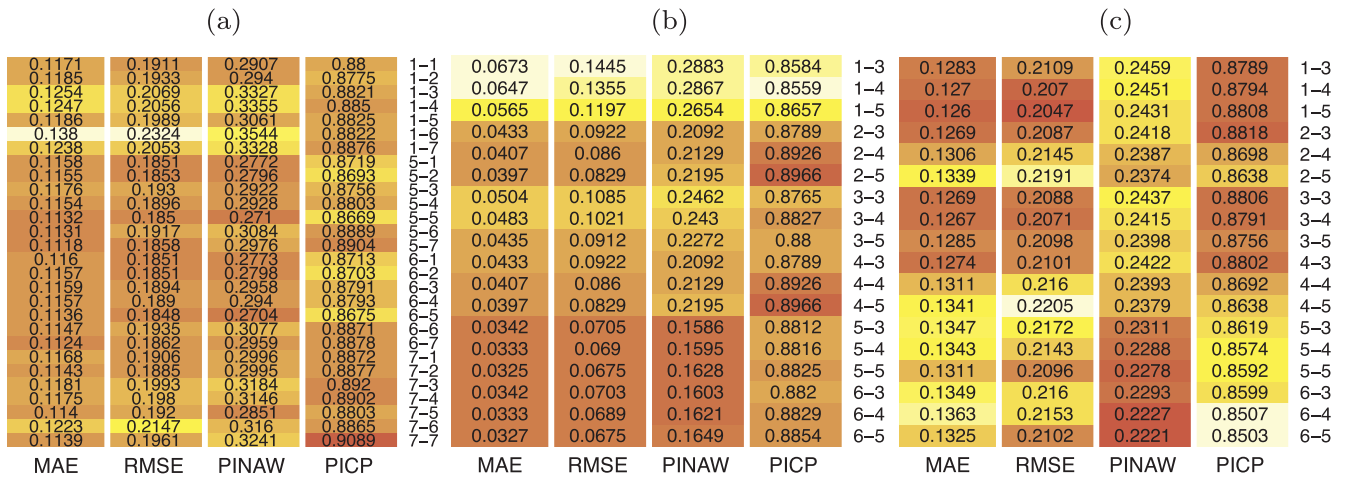


Fig. 4. Heat maps containing the results of  $k$ -fold time series cross validation procedure in case of electricity consumption (a), PV power generation (b) and net demand (c). The numbers on the right represent the covariance function and the alternative, respectively. Both the covariance function and the alternative were presented in Tables 1 and 2, respectively. A darker color represents preferable performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

All simulations are run on an Intel Core i5 3.1 GHz dual core processor with 16 GB RAM.

### 3.1. Forward chaining

The number of training samples for each fold of the forward chaining procedure increases with 800, while beginning with 200 training samples. The reasoning behind this is purely in terms of computation time, however, iteration increments smaller than 800 did not show to have significant effect on the results.

#### 3.1.1. Electricity consumption

Fig. 4a presents the results of the forward chaining procedure in case of electricity consumption. It should be noted that the selection of covariance functions used to create this figure is slightly different than those presented in the following subsections, due to the fact that preliminary tests revealed that the ones not taken into account here led to unsatisfactory results. In addition, we employ here covariance function 7 from Table 2 because it could potentially benefit from the periodic term. The figure shows that while indeed this covariance function achieves high PICP, it does not perform as well in terms of PINAW, the metric that we value most since it represents the informativeness of the PIs. Although the deterministic and probabilistic measures are quite close, we select  $k_{SE+M3}^5(r)$  using alternative 7, i.e., 6 time lags and the difference between the first two time lags, as it provides the most satisfactory result overall.

#### 3.1.2. PV power generation

Fig. 4b presents the results of the forward chaining procedure for PV power generation, where dark red implies favorable performance. We can observe that there appears to be a systematic trade-off between PINAW and PICP, which can be expected since these counterbalance each other. Furthermore, we can observe that the SE covariance function performs poorly, which, as stated before, is likely due to the fact that it is too smooth for this physical process. Since we value informativeness to be most important we select  $k_{SE+M3}^5(r)$  using alternative 1, i.e., 3 time lags, from Table 2, especially since it performs satisfactory in terms of MAE and RMSE as well.

#### 3.1.3. Net demand

Fig. 4c presents the results of the forward chaining procedure in case of net demand. We can clearly observe three trends, the first two being that the first three covariance functions provide both the highest PICP and highest deterministic accuracy. However, as stated before,

PICP in itself is not of much importance since we prefer PIs to have high informativeness. Moreover, it is important to point out that the color scale suggests a larger difference between the covariance functions than there actually is. Similar as for our decision in case of PV power generation we select  $k_{SE+M3}^5(r)$  using alternative 3, i.e., 5 time lags, due to its favorable performance in case of PINAW and still satisfactory PICP.

### 3.2. Electricity consumption forecasting

As per the previous subsection, we employ  $k_{SE+M3}^5(r)$  to encode our belief in the general shape of the time series and utilize the first half of the second year of the data set, i.e., 1 July 2011 until 31 December 2011 to learn the hyperparameters of the aforementioned covariance function in case of the static GP. This implies inverting the covariance matrix presented in Eq. (1), which, as stated before, costs  $\mathcal{O}(N^3)$  in time. In contrast, the dynamic GP does not require as much data and its hyperparameters are learned at each  $i$ th iteration using a moving window with the previous 250 data points, which are used to predict the following 250 points as one-step ahead, after which the hyperparameters are updated.

Figs. 5 and Table 3 present the results of the forecasts of both the GPs and ARIMA models. For clarity of graphical presentation, we select a part of the forecast rather than the entire 6 months. The reason we choose April is the fact that this month is during Autumn on the Southern Hemisphere, which tends to correspond to weather that is challenging to predict. Since weather and electricity consumption are closely related, this should affect electricity consumption too [57]. In addition, Fig. 5b and d present a further reduced timescale in order to assess the performance.

As a first observation, it is interesting to note that the traditional morning and evening peaks are difficult to discern, especially the former. We also observe that the peaks generally are not of the same duration, nor are they of the same scale. These observations are the reasons forecasting on this spatial resolution is such a challenging task. Several interesting observations can further be made. Firstly, Fig. 5a shows that the GPs are able to capture most peaks in the PIs with 80% confidence, except for a few outliers between 2 April and 9 April, which is likely caused by the fact that there does not seem to be any indication for these sudden increases in consumption. It is likely that a higher temporal resolution could improve the GPs in this respect. This observation is supported by Table 3, which shows that PICP for both the static and dynamic GPs are 88.85% and 87.21%, respectively, and are therefore considered to be satisfactory. Interestingly however, when examining Fig. 5b we can observe that the dynamic GP appears to have

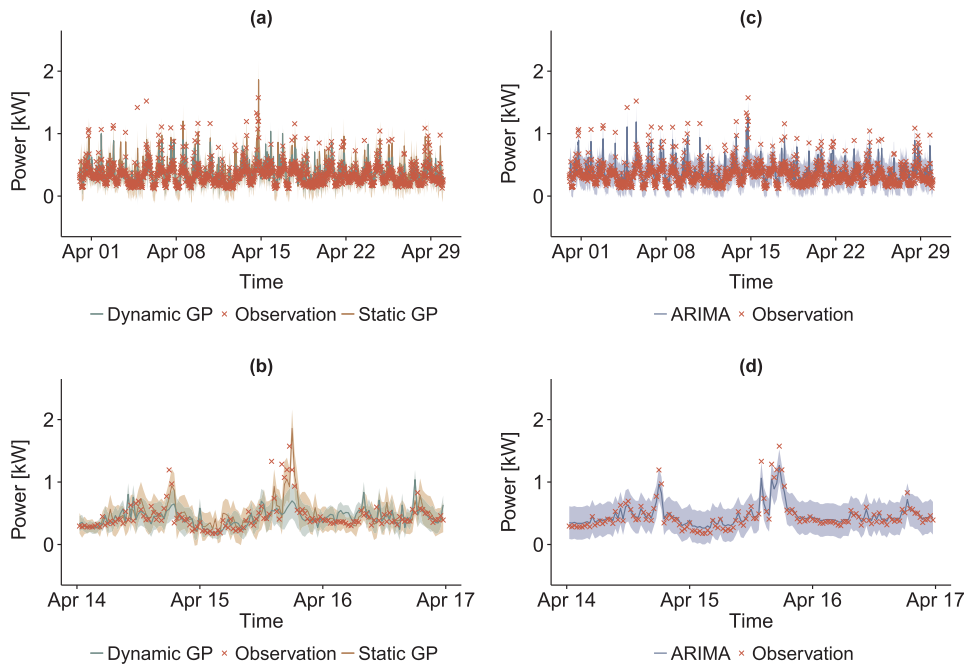


Fig. 5. Probabilistic forecasts of electricity consumption of (a) the GPs during the entire month of April 2012, (b) the GPs during a selection of days in April, (c) the ARIMA during the entire month of April 2012 and (d) the ARIMA during a selection of days in April. The shaded areas represent the 80% prediction intervals.

difficulty following the substantial and sudden peak in consumption on the night of 14 April. Since Table 3 shows that PICP is slightly lower for the dynamic than for the static, this might indicate that this behavior is indicative of the dynamic GP and is most likely caused by a lack of training data. This implies that the dynamic GP has not been trained on sufficient data to be able to predict such peaks accurately, whereas the static has seen similar behavior before. A slightly higher NRMSE supports the reasoning that the dynamic GP has difficulty capturing sudden peaks. Future work will therefore investigate the optimal length of the moving window  $M$  in order to learn sufficiently from the electricity consumption behavior. Furthermore, the static and dynamic GPs attain 3.868% and 3.755%, respectively, in terms of NCRPS and room for improvement exists here. It is important to note however, that regularly updating the hyperparameters does improve PINAW, i.e., 14.89% and 13.03% for the static and dynamic GP respectively, due to its ability to better cope with seasonal variation, e.g., weekly or monthly. This is an interesting result since it implies that the dynamic GP is able to produce sharper PIs whilst not deteriorating significantly with regards to the other measures, and actually improving in terms of NCRPS, with the added benefit of significantly reduced computational cost.

The ARIMA model proves to be highly capable in predicting the electricity consumption time series. Fig. 5c and d display that peaks are generally covered by the PIs, an observation that is supported by a PICP of 99.01% as presented in Table 3. In addition, the MAPE and NRMSE are 1.872% and 2.513%, respectively, and therefore it outperforms both GPs in these respects. Therefore, as a univariate and deterministic

forecast model, the ARIMA remains an interesting choice for such problems. However, Table 3 shows that the sharpness of the PIs is lower than that of the GPs. Although numerically this difference might not be substantial, Fig. 5b and d demonstrate the importance of the predictive density produced by the GPs. More specifically, we can observe that the PIs of the ARIMA prediction remain constant over time, effectively expressing the same certainty for different periods of the day or month. In contrast, the GPs differentiate between periods of high and low uncertainty through wide and narrow PIs, respectively. For example, during midnight on 13 April, both GPs express little uncertainty about their prediction through sharp PIs, whilst the ARIMA model does not make such distinction.

In order to put the results presented in Table 3 into perspective, we briefly compare them to the state of the art. For example, He et al. [58] used quantile regression neural networks (QRNNs) to forecast electricity demand on city-scale and achieved a PICP between 98.81% and 99.85%, and PINAW between 9.610% and 16.30%. In a follow-up publication [59], they utilized support vector quantile regression (SVQR) using various covariance functions to forecast electricity demand on city-scale and achieved a PICP between 82.81% and 100%, and PINAW between 11.62% and 30.65%, where the results varied significantly between the months. Finally, Quan et al. [24] utilized NNs with the LUBE approach to forecast electricity demand on city-scale and achieved a PICP between 90.64% and 91.14% ( $\eta = 90\%$ ), and PINAW between 14.42% and 36.75%, depending on the city. It is important to note that forecasts on city-scale are generally considered to be more

Table 3  
Numerical results for the static and dynamic forecasts.

GPs	Strategy	MAPE (%)	NRMSE (%)	PINAW (%)	PICP (%)	NCRPS (%)
Cons.	Static	3.516	5.798	14.89	88.85	3.868
PV	Static	3.629	7.795	14.07	89.11	3.038
Net	Static	3.335	5.227	12.07	85.95	4.462
Cons.	Dyn.	3.369	5.819	13.03	87.21	3.755
PV	Dyn.	3.856	8.212	12.42	87.57	3.031
Net	Dyn.	3.199	5.238	10.97	83.38	4.396
ARIMA						
Cons.	Static	1.872	2.513	16.96	99.01	1.376
PV	Static	1.209	2.156	21.58	100.0	1.131
Net	Static	1.306	1.869	14.86	99.71	0.9927



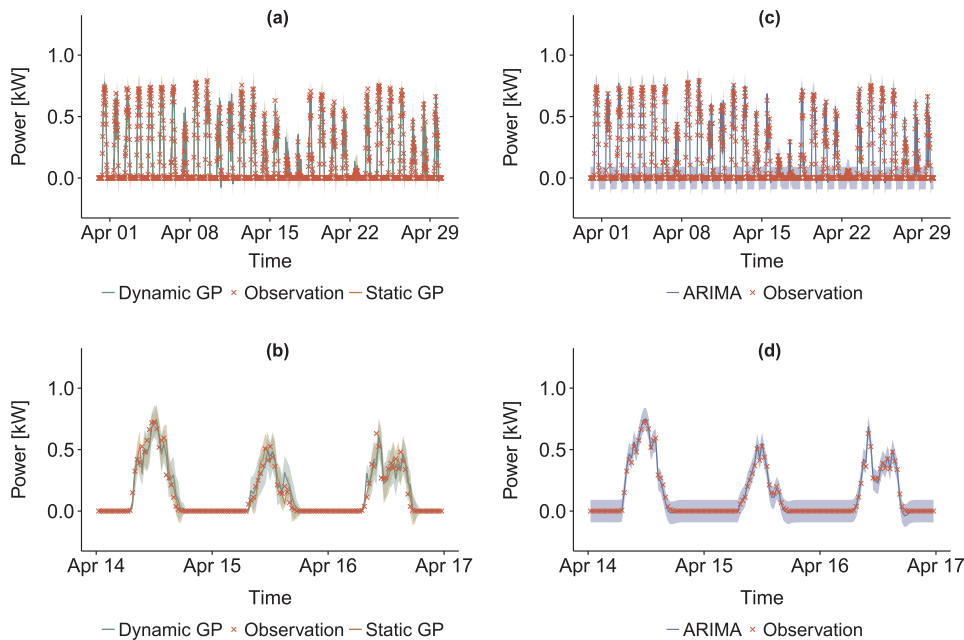


Fig. 6. Probabilistic forecasts of PV power production of (a) the GPs during the entire month of April 2012, (b) the GPs during a selection of days in April, (c) the ARIMA during the entire month of April 2012 and (d) the ARIMA during a selection of days in April. The shaded areas represent the 80% prediction intervals.

straightforward since the aggregated demand displays a smoother pattern [10].

A final note is that on computational time. The static GP requires 8942 s and 89 s to learn the hyperparameters and predict, respectively. In contrast, the dynamic GP needs 109 s to learn and update the hyperparameters, and make predictions, while the ARIMA model requires 162 s to train the model and predict future electricity consumption. The previously mentioned study by Quan et al. [24] noted an average construction time for PI construction between 5.60 and 9.58 ms, which can be compared to our construction times.

### 3.3. PV power generation forecasting

In order to predict PV power generation with GPs we employ an identical strategy as in the case of electricity consumption. Thus, we select  $k_{SE+M3}^5(r)$  as covariance function and subsequently learn the hyperparameters in a static and dynamic manner, as previously explained. Fig. 6a and b present the resulting forecasts during April, whereas Table 3 presents the results over the entire half year.

We can observe from Fig. 6a that both the static and dynamic GPs perform well on sunny days, i.e., when variability is low. On cloudy days the PIs widen due to the increased uncertainty of the forecast. Interestingly, it seems like the static GP produces narrower PIs than the dynamic GP on such days, which is more clearly displayed in Fig. 6b. Overall, PINAW for the static and dynamic GPs are 14.07% and 12.42%, respectively, which therefore contradicts this observation. However, upon closer inspection the dynamic GP expresses less uncertainty during nighttime than the static, which accounts for a large part of the data. This explains the fact that on cloudy days the static GP produces narrower PIs but overall shows lower performance in this respect. The fact that the dynamic GP is less certain during days with increased variability will be studied more thoroughly in future work, where special focus will lie on finding the optimal length of the moving window  $M$ . Another interesting phenomenon that occurs for both GPs is the consistent overshoot at nightfall, when PIs go below zero. This behavior can be expected since variability during the late afternoon tends to increase due to e.g., increased thickness of the atmosphere and consequently more scatter, and potential shading. However, it is worth pointing out that the static GP tends to exacerbate this effect, likely because of the variable day length and it not accounting for these changes. Although dynamically updating the hyperparameters reduces

this effect, introducing additional explanatory variables such as clear sky irradiance might improve accuracy during the late afternoon. In terms of PICP the static GP performs slightly better than the dynamic GP, 89.11% versus 87.57%, respectively. Keeping in mind that the static GP produces sharper PIs during cloudy days than the dynamic GP, we note that the probabilistic results favor the former strategy. The NRMSE corroborates this conclusion, since it indicates that the dynamic GP performs worse with increasing variability. The NCRPS for the strategies are too close to draw any meaningful conclusion.

Similarly as was the case for electricity consumption, the ARIMA shows good performance during forecasting of PV power. This becomes particularly clear when looking at Table 3, from which we can see that MAPE and NRMSE are 1.209% and 2.156%, respectively. Fig. 6c and d present this as well, where it is interesting to note that the deterministic predictions are very accurate, even on cloudy days. In addition, the results show that the PIs cover all observations, albeit at the cost of these being wide as expressed by a PINAW of 21.58%. Due to the fact that the `auto.arima` function of the `forecast` package uses the residuals to produce PIs, the width remains constant as mentioned in the previous subsection. Therefore, we conclude that although the ARIMA model produces high quality deterministic forecasts, it does not express the uncertainty as accurately that inherently follows from a forecast.

Here we briefly compare the results of Table 3 with the state of the art. Ni et al. [14] used an ensemble of ELMs in combination with the LUBE approach and achieved a PICP between 91.55% and 96.17% ( $\eta = 90\%$ ) and PINAW between 17.68% and 20.12%, depending on the ensembling method. Bracale et al. [18] also utilized an ensemble and reported a CRPS of 5.24 kW and 5.53 kW, depending on the ensembling method, for the probabilistic forecast of a 100 kW<sub>p</sub> PV system. Finally, Chu and Coimbra [60] utilized a k-nearest neighbor (kNN) ensemble model to produce intra-hour direct normal irradiance (DNI) forecasts and achieved a PICP between 91% and 96% ( $\eta = 90\%$ ), and PINAW between 22% and 70%, depending on location and forecast horizon.

In terms of time, the static GP requires 20,492 s to learn the hyperparameters and 107 s to predict. In contrast, the dynamic GP and ARIMA require 103 s and 87 s, respectively. This can be compared to the calculation time reported by Sanjari and Gooi [15], which was 47.2 s.

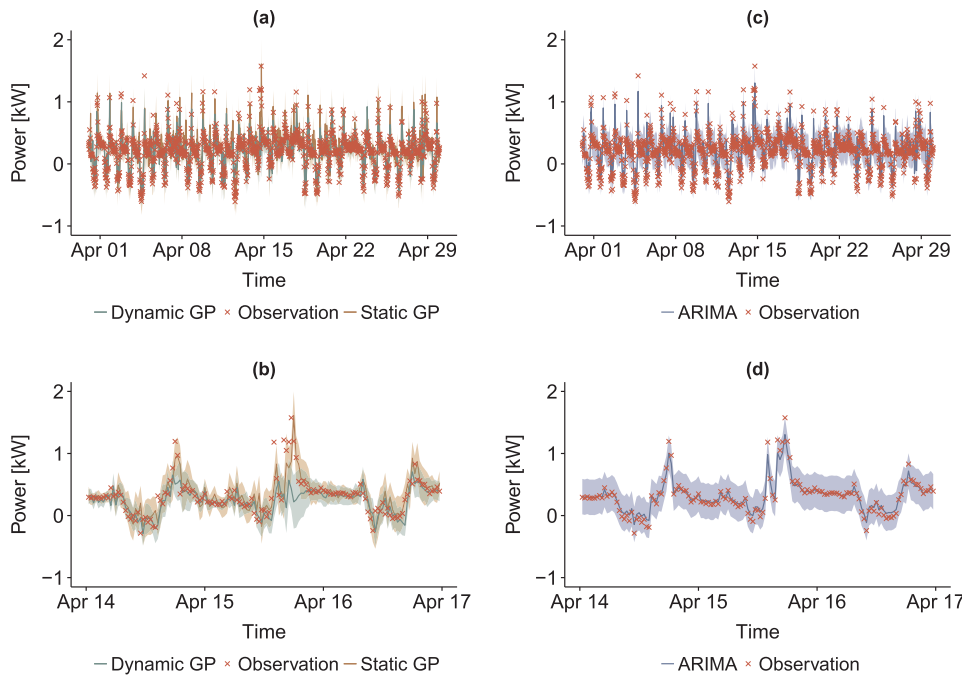


Fig. 7. Probabilistic forecasts of net demand of (a) the GPs during the entire month of April 2012, (b) the GPs during a selection of days in April, (c) the ARIMA during the entire month of April 2012 and (d) the ARIMA during a selection of days in April. The shaded areas represent the 80% prediction intervals.

### 3.4. Net demand forecasting

The net demand time series presented in Fig. 1 represents an interesting case as it combines seasonalities present in both electricity consumption and PV power production. The time series appears to be more stationary than either of the two aforementioned separately, although variance does not remain constant over the entirety of the data. Better stationary behavior could suggest that some of the seasonalities cancel each other out or at least dampen their respective effects.

Fig. 7a and b, and Table 3 present the results graphically and numerically, respectively. From Fig. 7a we can observe that the peaks, both positive and negative, are well captured by the GPs. This is corroborated by the PICP for both cases, where the static and dynamic GPs achieve 85.95% and 83.38%, respectively, and thus remain above the nominal confidence level. What is interesting to note however, is that although PICP is lower than the separate forecasts, the sharpness of the net demand forecasts are also lower. Although this in itself is a recurring phenomenon in Table 3, realizing PINAWs of 12.07% and 10.97% for the static and dynamic strategy, respectively, implies achieving competitive results with the current state-of-the-art in the literature [10]. Therefore, forecasting net demand directly could prove to be valuable when aiming to improve accuracy of PIs. This will be further studied in Section 3.5. Furthermore, we can see from Fig. 7b that here, similarly as in the previous cases, a clear trade-off exists between the static and dynamic GPs. On the one hand, the static GP tends to perform better when sudden peaks in either electricity consumption or PV power production arise, whereas the dynamic GP instead produces wide PIs and regularly fails to capture these, as observed as well in Sections 3.2 and 3.3. On the other hand, the dynamic GP shows that it can better express the reduction in uncertainty during nighttime when electricity consumption is low and PV production is zero. Table 3 supports these visual observations, since the dynamic GP offers sharper PIs at the loss of reliability. It therefore seems there exists some balance between the ability to capture peaks and the expressiveness of uncertainty during periods with higher variability. Future work will further investigate this balance. Interestingly however, NCRPS does improve slightly. This observation was also made in Section 3.2 where the improvement in PINAW led to lower PICP as well, but also to an improvement of NCRPS. This suggests that indeed informativeness is valued higher and as such, special attention should be directed to PINAW. Finally, we

observe from Table 3 that MAPE and NRMSE are relatively high, and although direct comparison between time series on different scales is not possible [61], there is certainly room for improvement here. However, it is important to note that the aim of probabilistic forecasting is not to predict a mean value.

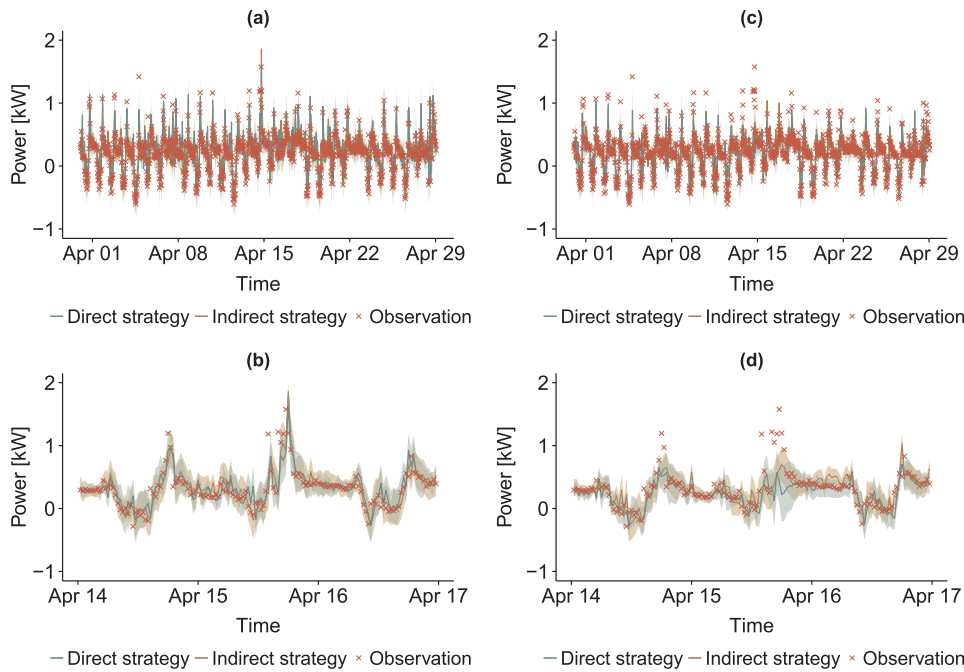
The ARIMA model has slightly lower MAPE and NRMSE when compared to electricity consumption, although this could also partly be ascribed to the different scale. Nevertheless, the ARIMA performs very well in terms of deterministic error metrics, with MAPE and NRMSE being 1.306% and 1.869%, respectively. This can also be observed from Fig. 7c and d, where the mean predictions are highly accurate, except for several positive and negative peak values. These are then captured by the PIs, resulting in a PICP of 99.71%. What is interesting to note is that the improvement of net demand forecasting over the two separately is more substantial than that in case of the GPs, where the better performance in PINAW results in slightly reduced PICP. However, as previously mentioned, the PIs of the ARIMA remain rather wide and moreover, of constant width. It is interesting to note the NCRPS however, that with a value of 0.9927% is relatively low, which suggests that although the PIs are relatively wide, the high PICP compensates PINAW when calculating NCRPS.

The results presented in Table 3 are challenging to compare since few studies aimed to forecast net demand. Wang et al. [25] predicted net demand although they used the pinball score to assess the quality of their probabilistic forecasts, which makes direct comparison not a straightforward task.

In terms of computation time, the static GP requires 19,032 s to learn the hyperparameters and 101 s to predict, whereas the dynamic GP and ARIMA require 93 s and 96 s, respectively.

### 3.5. Comparison between the direct and indirect strategies for net demand

As a final contribution we investigate the effect that the direct and indirect strategies have in case of net demand forecasting, where we define the latter strategy to be subtracting the individual forecasts to produce the net demand forecast. In case of probabilistic forecasting, this brings with it an important aspect one should keep in mind. Namely, when Gaussian and independent random variables  $X$  and  $Y$  are added or subtracted to create random variable  $Z$ , then  $Z \sim \mathcal{N}(\mu_X \pm \mu_Y, \sigma_X^2 + \sigma_Y^2)$ . This implies that the indirect strategy will



**Fig. 8.** Comparison between the direct and indirect strategies in the case of probabilistic net demand forecasts produced by the static (a and b) and dynamic (c and d) GPs. The shaded areas represent the 80% prediction intervals.

lead to an increase in variance.

Fig. 8 and Table 4 present the results of the direct and indirect strategies, as produced by the static (a and b) and dynamic (c and d) GPs. By visual inspection we observe that the PIs in Fig. 8b appear to be more narrow in case of the direct strategy. This observation is supported by the lower PINAW for the direct strategy, i.e., 12.07%, versus the PINAW for the indirect strategy, i.e., 12.56%, as presented in Table 4. Closely related is PICP, which is higher in case of the indirect strategy, likely due to the fact that the addition of variances increases the probability of covering more observations. In terms of NCRPS, the indirect strategy is favored. Overall, selecting the best strategy depends mainly whether one prefers higher informativeness of PIs, or higher coverage probability.

A similar pattern arises when comparing the strategies in case of the dynamic GP. The direct strategy favors PINAW and consequently reduces PICP, which increases NCRPS slightly. As a conclusion regarding selecting the most appropriate strategy we can say that this varies between application. However, the direct strategy requires training of only one model versus training two models in case of the indirect strategy, which definitely offers a computational advantage.

A final remark is that on the comparison between the static and dynamic GPs, since Fig. 8 allows us to point out clearly the difference

between these two strategies, as highlighted in Sections 3.2, 3.3, 3.4. In contrast to e.g., Fig. 7, Figs. 8b and d clearly show that the static GP is generally better able to capture the peaks, in this case of net demand, than the dynamic GP, while the latter produces more narrow PIs. Therefore, optimizing the moving window length is proving to be vital to design competitive GPs in terms of both accuracy and speed.

### 3.5.1. Sensitivity analysis

In this section, we conduct a sensitivity analysis of the proposed method using 4 additional customers from the data set [44]. More specifically, we select additional customers 74, 157, 211 and 273. The selection criterion is variance, i.e., we want to investigate the accuracy of the GPs on different time series, where we assume that higher variance implies a more challenging forecast. Variances of the normalized time series of customers {69,74,157,211,273} are: 0.006018, 0.01310, 0.02504, 0.004959 and 0.01160, respectively. For the sake of conciseness, we do this for the “Static and indirect” and “Dynamic and direct” strategies from Table 4, since these strategies represent the best performance in terms of reliability and informativeness, respectively. Furthermore, we assume the covariance functions found during the  $k$ -fold cross validation procedure remain valid for these cases. Table 4 presents the results of this sensitivity analysis, where it is important to

**Table 4**

Numeric results for the direct and indirect net demand forecasts using the static and dynamic GPs of customer 69. In addition, numeric results for the direct and indirect net demand forecasts using the static and dynamic GPs for several customers. Variances of the normalized time series of customers 69–273 in ascending order are: 0.006018, 0.01310, 0.02504, 0.004959 and 0.01160, respectively.

Customer	Strategy	MAPE (%)	NRMSE (%)	PINAW (%)	PICP (%)	NCRPS (%)
69	Static and direct	3.335	5.227	12.07	85.95	4.462
	Static and indirect	3.138	4.973	12.56	87.91	2.447
	Dyn. and direct	3.199	5.238	10.97	83.38	4.396
	Dyn. and indirect	3.036	5.038	11.23	86.14	2.374
74	Dyn. and direct	3.817	6.899	13.49	84.98	6.695
157	Dyn. and direct	4.324	6.778	11.25	81.89	10.45
211	Dyn. and direct	3.486	5.181	11.61	82.41	5.398
273	Dyn. and direct	3.228	6.317	9.690	81.96	6.022
74	Static and indirect	3.351	6.088	15.46	91.12	6.504
157	Static and indirect	3.819	6.033	15.36	88.34	10.41
211	Static and indirect	3.339	4.884	10.43	81.22	5.418
273	Static and indirect	3.048	5.704	13.53	90.07	5.767

point out that customer 69 is the original customer on which the previous analyses have been performed. The results clearly show the stability of the GP, i.e., higher variance implies a more challenging time series to forecast, which is represented in the performance metrics.

#### 4. Discussion

The high spatial resolution of this study led to a challenging forecasting problem, since the time series contained substantial stochastic behavior and it was difficult to discern patterns. The GP requires selecting a covariance function that encodes the general shape of the time series to be modeled. Therefore, we employed  $k$ -fold time series cross validation in order to find the most appropriate covariance function for the time series under consideration, especially since it has been argued that the commonly utilized squared exponential (SE) covariance function is too smooth for many physical processes [55]. In addition, we solely relied on endogenous variables, and the  $k$ -fold time series cross validation was therefore also used to find the optimal combination of lags and differences to consider. The results of this procedure showed that the SE covariance function did not lead to the best performance, but rather a combination of the SE and Matérn class covariance functions.

We showed that the dynamic GP produced narrower PIs than the static GP. However, the improvement in PINAW was accompanied by a reduction in the PICP in all cases, implying that an improvement of sharpness tends to lead to a reduction in reliability. In order to draw a final conclusion on this, we used the NCRPS that measures both sharpness and reliability. We found that NCRPS improved when changing from the static to the dynamic strategy. However, it was noted that the dynamic strategy had difficulty predicting sudden peaks in the time series, which was likely a consequence of a too narrow moving training window. Furthermore, we compared the GPs to the well-known ARIMA model and found that the GPs were consistently outperformed by the ARIMA model in terms of MAPE and NRMSE. However, the ARIMA produced wide PIs and, moreover, was unable to differentiate between periods with high and low uncertainty, whereas the GPs performed well in this respect. When compared to the state of the art, the GP showed to be competitive in terms of accuracy, especially given the fact that the compared studies generally considered aggregated electricity demand or included exogenous variables.

We examined the difference between direct and indirect net demand forecasting for static and dynamic GPs. The results showed that the direct strategy yielded sharper PIs, but lower PICP than the indirect strategy. Moreover, the direct strategy has the additional benefit that it requires only one GP to train, as opposed to two. These are important findings since the strategies emphasize different, and conflicting, aspects of the probabilistic forecast. This knowledge can be exploited in decision making processes. Assessment of the stability of the GP for different customers showed that its performance was strongly related to the variability of the time series, which is common for prediction using machine learning.

As a concluding discussion point, we would like to mention the benefit of probabilistic forecasts in operational sense. As several studies reviewed in Section 1.1 revealed, such forecasts provide important information that is of value when considering decision making processes such as energy scheduling in micro-grids. Such processes are characterized by uncertainties, however, deterministic forecasts do not provide stakeholders with such information. Consequently, scheduling is sub-optimal, which generally implies the need for expensive power reserves.

#### 5. Conclusion

In this paper we have utilized Gaussian Processes (GPs) as a non-parametric model to produce probabilistic forecasts with a time horizon of 30 min of electricity consumption, photovoltaic (PV) power

generation and net demand, where the latter was defined as the electricity consumption minus PV power generation, of a single household. One important issue of the GP is the time it takes to learn the hyperparameters of the covariance functions and perform inference from it. Since this costs  $\mathcal{O}(N^3)$  time, we compared a static GP to a dynamic GP where the latter used a moving training window to learn the hyperparameters, which significantly reduced the computational burden. However, future study will investigate the optimal window width to improve performance. Overall, the GP was found to be competitive when compared to the state of the art in terms of both reliability and sharpness of the prediction intervals (PIs), which calls for further investigation of this model over various spatial and temporal scales.

#### Acknowledgments

This work was financially supported by SamspeL 2016–2020 in the project “Development and evaluation of forecasting models for solar power and electricity use over space and time”, financed primarily by the Swedish Energy Agency. In addition, we would like to thank the anonymous reviewers and editor for their efforts to improve the paper.

#### References

- [1] IEA. 2016 Snapshot of global photovoltaic markets, Tech. rep., ISBN 978-3-906042-42-8; 2017.
- [2] Wirth H. Recent facts about photovoltaics in Germany, Tech. rep. Fraunhofer ISE < <https://www.ise.fraunhofer.de/> >; 2017.
- [3] Agüero JR, Steffel SJ. Integration challenges of photovoltaic distributed generation on power distribution systems. In: 2011 IEEE power energy society general meeting; 2011. p. 1–6, doi:<http://dx.doi.org/10.1109/PES.2011.6039097>.
- [4] Walling RA, Saint R, Dugan RC, Burke J, Kojovic LA. Summary of distributed resources impact on power delivery systems. IEEE Trans Power Deliv 2008;23(3):1636–44. <http://dx.doi.org/10.1109/TPWRD.2007.909115>.
- [5] Mohammadi P, Mehraeen S. Challenges of PV integration in low-voltage secondary networks. IEEE Trans Power Deliv 2017;32(1):525–35. <http://dx.doi.org/10.1109/TPWRD.2016.2556692>.
- [6] Ropp M, Newmiller J, Whitaker C, Norris B. Review of potential problems and utility concerns arising from high penetration levels of photovoltaics in distribution systems. In: IEEE photovoltaics special conference, doi:<http://dx.doi.org/10.1109/PVSC.2008.4922861>; 2008.
- [7] Inman RH, Pedro HT, Coimbra CF. Solar forecasting methods for renewable energy integration. Prog Energy Combust Sci 2013;39(6):535–76. <http://dx.doi.org/10.1016/j.pecs.2013.06.002>.
- [8] Hong T. Energy forecasting: past, present and future. Foresight Int J Forecast 2014(32):43–9.
- [9] Zamo M, Mestre O, Arbogast P, Pannekoucke O. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. Part II: probabilistic forecast of daily production. Sol Energy 2014;105:804–16. <http://dx.doi.org/10.1016/j.solener.2014.03.026>.
- [10] van der Meer D, Widén J, Munkhammar J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renew Sustain Energy Rev 2017(December 2016):1–29. <http://dx.doi.org/10.1016/j.rser.2017.05.212>.
- [11] Hong T, Fan S. Probabilistic electric load forecasting: a tutorial review. Int J Forecast 2016;32(3):914–38. <http://dx.doi.org/10.1016/j.ijforecast.2015.11.011>.
- [12] Scolari E, Torregrossa D, Boudec JL, Paolone M. Ultra-short-term prediction intervals of photovoltaic AC active power. In: International conference on probabilistic methods applied to power systems PMAPS 2016, doi:<http://dx.doi.org/10.1109/PMAPS.2016.7764064>; 2016.
- [13] Scolari E, Sossan F, Paolone M. Irradiance prediction intervals for PV stochastic generation in microgrid applications. Sol Energy 2016;139:116–29. <http://dx.doi.org/10.1016/j.solener.2016.09.030>.
- [14] Ni Q, Zhuang S, Sheng H, Kang G, Xiao J. An ensemble prediction intervals approach for short-term PV power forecasting. Sol Energy 2017;155:1072–83. <http://dx.doi.org/10.1016/j.solener.2017.07.052>.
- [15] Sanjari MJ, Gooi HB. Probabilistic forecast of PV power generation based on higher order Markov chain. IEEE Trans Power Syst 2017;32(4):2942–52. <http://dx.doi.org/10.1109/TPWRS.2016.2616902>.
- [16] Nagy GI, Barta G, Kazi S, Borbély G, Simon G. GEFCom2014: probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. Int J Forecast 2016;32(3):1087–93. <http://dx.doi.org/10.1016/j.ijforecast.2015.11.013>.
- [17] Wang H, Yi H, Peng J, Wang G, Liu Y, Jiang H, et al. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. Energy Convers Manage 2017;153:409–22. <http://dx.doi.org/10.1016/j.enconman.2017.10.008>.
- [18] Bracale A, Carpinelli G, De Falco P. A probabilistic competitive ensemble method for short-term photovoltaic power forecasting. IEEE Trans Sustain Energy 2017;8(2):551–60. <http://dx.doi.org/10.1109/TSTE.2016.2610523>.
- [19] Appino RR, González Ordiano JA, Mikut R, Faulwasser T, Hagenmeyer V. On the use of probabilistic forecasts in scheduling of renewable energy sources coupled to



- storages. Appl Energy, doi:<http://dx.doi.org/10.1016/j.apenergy.2017.08.133>.
- [20] Bracale A, Carpinelli G, Falco PD. A Bayesian-based approach for the short-term forecasting of electrical loads in smart grids. Part II: numerical applications. In: International symposium on power electronics, electrical drives, and automated motion; 2016. p. 129–36.
- [21] Dordonnat V, Pichavant A, Pierrot A. GEFCom2014 probabilistic electric load forecasting using time series and semi-parametric regression models. Int J Forecast 2016;32(3):1005–11. <http://dx.doi.org/10.1016/j.ijforecast.2015.11.010>.
- [22] Ziel F, Liu B. Lasso estimation for GEFCom2014 probabilistic electric load forecasting. Int J Forecast 2016;32(3):1029–37. <http://dx.doi.org/10.1016/j.ijforecast.2016.01.001>. arXiv:1603.01376v1.
- [23] Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans Neural Netw Learn Syst 2014;25(2):303–15. <http://dx.doi.org/10.1109/TNNLS.2013.2276053>.
- [24] Quan H, Srinivasan D, Khosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. Energy 2014;73:916–25. <http://dx.doi.org/10.1016/j.energy.2014.06.104>.
- [25] Wang Y, Zhang N, Chen Q, Kirschen DS, Li P, Xia Q. Data-driven probabilistic net load forecasting with high penetration of invisible PV. IEEE Trans Power Syst 2017;1–10. <http://dx.doi.org/10.1109/TPWRS.2017.2762599>.
- [26] Cabrera BL, Schulz F. Forecasting generalized quantiles of electricity demand: a functional data approach. J Am Stat Assoc 2017;112(517):127–36. <http://dx.doi.org/10.1080/01621459.2016.1219259>.
- [27] Taieb SB, Huser R, Hyndman RJ, Genton MG. Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression. IEEE Trans Smart Grid 2016;7(5):2448–55. <http://dx.doi.org/10.1109/TSG.2016.2527820>.
- [28] Arora S, Taylor JW. Forecasting electricity smart meter data using conditional kernel density estimation. Omega 2016;59:1–13. <http://dx.doi.org/10.1016/j.omega.2014.08.008>.
- [29] Salcedo-Sanz S, Casanova-Mateo C, Munoz-Mari J, Camps-Valls G. Prediction of daily global solar irradiation using temporal Gaussian Processes. IEEE Geosci Rem Sens Lett 2014;11(11):1936–40.
- [30] Bilonis I, Constantinescu EM, Anitescu M. Data-driven model for solar irradiation based on satellite observations. Sol Energy 2014;110:22–38. <http://dx.doi.org/10.1016/j.solener.2014.09.009>.
- [31] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. ScienceDirect A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 2015;112:446–57. <http://dx.doi.org/10.1016/j.solener.2014.12.014>.
- [32] Sheng H, Xiao J, Cheng Y, Ni Q, Wang S. Short-term solar power forecasting based on weighted Gaussian Process regression. IEEE Trans Ind Electron 2017;0046(c). <http://dx.doi.org/10.1109/TIE.2017.2714127>. pp. 1–1.
- [33] Lauret P, David M, Calogine D. Nonlinear models for short-time load forecasting. In: Energy Proc 2012;14:1404–9, doi:<http://dx.doi.org/10.1016/j.egypro.2011.12.1109>.
- [34] Kou P, Gao F. A sparse heteroscedastic model for the probabilistic load forecasting in energy-intensive enterprises. Int J Electr Power Energy Syst 2014;55:144–54. <http://dx.doi.org/10.1016/j.ijepes.2013.09.002>.
- [35] Dong B, Li Z, Rahman SM, Vega R. A hybrid model approach for forecasting future residential electricity consumption. Energy Build 2016;117:341–51. <http://dx.doi.org/10.1016/j.enbuild.2015.09.033>.
- [36] McLoughlin F, Duffy A, Conlon M. Evaluation of time series techniques to characterise domestic electricity demand. Energy 2013;50(1):120–30. <http://dx.doi.org/10.1016/j.energy.2012.11.048>.
- [37] Lloyd JR. GEFCom2012 hierarchical load forecasting: gradient boosting machines and Gaussian processes. Int J Forecast 2014;30(2):369–74. <http://dx.doi.org/10.1016/j.ijforecast.2013.07.002>.
- [38] Ela E, Tuohy A, Philbrick R, Lannoye E, Philbrick R. Using probabilistic renewable forecasts to determine reserve requirements. In: 7th Solar integration workshop. Berlin; 2017.
- [39] Sarshar J, Moosapour SS, Joorabian M. Multi-objective energy management of a micro-grid considering uncertainty in wind power forecasting. Energy 2017;139:680–93. <http://dx.doi.org/10.1016/j.energy.2017.07.138>.
- [40] Al-Sumaiti AS, Salama MM, El-Moursi M. Enabling electricity access in developing countries: a probabilistic weather driven house based approach. Appl Energy 2017;191:531–48. <http://dx.doi.org/10.1016/j.apenergy.2017.01.075>.
- [41] Shaker H, Chitsaz H, Zareipour H, Wood D. On comparison of two strategies in net demand forecasting using Wavelet Neural Network. In: 2014 North American power symposium, NAPS; 2014, doi:<http://dx.doi.org/10.1109/NAPS.2014.6965360>.
- [42] Rasmussen CE, Williams CKI. Gaussian processes for machine learning. MIT Press; 2006.
- [43] Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R (Version 8.1). J Stat Softw 2008;27(3):1–22. <http://dx.doi.org/10.18637/jss.v027.i03>.
- [44] Solar home electricity data - Ausgrid < <http://www.ausgrid.com.au/Common/About-us/Corporate-information/Data-to-share/Solar-home-electricity-data.aspx#.WafBXa2B10s> > .
- [45] Ratnam EL, Weller SR, Kellett CM, Alan T, Ratnam EL, Weller SR, Kellett CM, Murray AT, et al. Residential load and rooftop PV generation: an Australian distribution network dataset Residential load and rooftop PV generation: an Australian. Int J Sustain Energy 2017;36(8):787–806. <http://dx.doi.org/10.1080/14786451.2015.1100196>.
- [46] Roberts S, Osborne M, Ebdon M, Reece S, Gibson N, Aigrain S. Gaussian processes for time-series modelling. Philos Trans R Soc Lond A Math Phys Eng Sci 1984;371.
- [47] Girard A, Rasmussen CE, Candela JQ, Murray-Smith R. Gaussian process priors with uncertain inputs-application to multiple-step ahead time series forecasting. Adv Neural Inf Process Syst 2003:545–52.
- [48] Yan J, Li K, Bai EW, Deng J, Foley AM. Hybrid probabilistic wind power forecasting using temporally local Gaussian process. IEEE Trans Sustain Energy 2016;7(1):87–95. <http://dx.doi.org/10.1109/TSTE.2015.2472963>.
- [49] van der Meer D, Chandra Mouli GR, Morales-España G, Ramirez-Elizondo L, Bauer P. Energy management system with PV power forecast to optimally charge EVs at the workplace. IEEE Trans Indust Inform 2017;3203(February):1–8. <http://dx.doi.org/10.1109/TII.2016.2634624>.
- [50] Madsen H, Pinson P, Kariniotakis G, Nielsen HA, Nielsen TS. Standardizing the performance evaluation of short-term wind power prediction models. Wind Eng 2005;29(6):475–89. <http://dx.doi.org/10.1260/030952405776234599>.
- [51] Hoff TE, Perez R, Kleissl J, Renne D, Stein J. Reporting of irradiance modeling relative prediction errors. Prog Photovolt Res Appl 2012;21:1514–9, doi:<http://dx.doi.org/10.1002/pip.2225>.
- [52] Khosravi A, Nahavandi S, Creighton D. Prediction intervals for short-term wind farm power generation forecasts. IEEE Trans Sustain Energy 2013;4(3):602–10. <http://dx.doi.org/10.1109/TSTE.2012.2232944>.
- [53] Gneiting T, Katzfuss M. Probabilistic forecasting. Annu Rev Stat Appl 2014;1:125–51. <http://dx.doi.org/10.1146/annurev-statistics-062713-085831>.
- [54] Bracale A, Carpinelli G, De Falco P. A Bayesian-based approach for the short-term forecasting of electrical loads in smart grids.: part I: theoretical aspects. 2016 International symposium on power electronics, electrical drives, and automated motion IEEE; 2016. p. 121–8. <http://dx.doi.org/10.1109/SPEEDAM.2016.7526022>.
- [55] Stein ML. Interpolation of spatial data: some theory for Kriging. New York: Springer; 1999.
- [56] Majidpour M, Qiu C, Chu P, Pota HR, Gadh R. Forecasting the EV charging load based on customer profile or station measurement? Appl Energy 2016;163:134–41. <http://dx.doi.org/10.1016/j.apenergy.2015.10.184>.
- [57] Widén J, Lundh M, Vassileva I, Dahlquist E, Ellegård K, Wäckelgård E. Constructing load profiles for household electricity and hot water from time-use data-modelling approach and validation. Energy Build 2009;41(7):753–68. <http://dx.doi.org/10.1016/j.enbuild.2009.02.013>.
- [58] He Y, Xu Q, Wan J, Yang S. Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function. Energy 2016;114:498–512. <http://dx.doi.org/10.1016/j.energy.2016.08.023>.
- [59] He Y, Liu R, Li H, Wang S, Lu X. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl Energy 2017;185:254–66. <http://dx.doi.org/10.1016/j.apenergy.2016.10.079>.
- [60] Chu Y, Coimbra CFM. Short-term probabilistic forecasts for Direct Normal Irradiance. Renew Energy 2017;101:526–36. <http://dx.doi.org/10.1016/j.renene.2016.09.012>.
- [61] Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. Int J Forecast 2006;22(4):679–88. <http://dx.doi.org/10.1016/j.ijforecast.2006.03.001>.