Gaussian Process Regressions for Aggregate Baseline Load Forecasting

Kadir Amasyali
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA
amasyalik@ornl.gov

Mohammed Olama
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA
olamahussemm@ornl.gov

***Abstract*— Demand response (DR) is one of the most effective ways to maintain the reliability and improve the flexibility of power systems. Accurate forecasts of baseline loads are essential for DR programs. In the era of big data, machine learning-based approaches present a unique opportunity for baseline load forecasting. Thus, this paper presents a machine learning-based approach using a relatively less explored algorithm, Gaussian process regression (GPR), to forecast aggregate baseline loads. As such, a dataset was generated using a set of EnergyPlus simulations. Using the generated dataset, a GPR-based forecasting model was developed. In addition, support vector regression (SVR)- and averaging-based models were developed as baseline models for comparison. The prediction performance of the models showed that the GPR-based model is more accurate and reliable than the others. Such high performance shows the potential of the GPR in baseline load forecasting. GPR, therefore, can be used for DR applications.**

***Index Terms*—Aggregate baseline load forecasting, demand response, Gaussian process regression, machine learning, support vector regression.**

I. Introduction

Motivated by the adverse effects of the fossil fuels on the environment, there is a growing trend towards using renewable energy resources for electricity generation. Renewable energy resources (e.g., solar and wind), however, pose some challenges for power system operations due to their variable and intermittent nature [1]. To overcome these challenges, power systems are going through a transformation to be able to adapt more flexible operation and management strategies. Demand side management, energy storage, and fast-acting supply services are crucial in this transformation and already play a key role in the operation and management of energy systems with various programs [2].

Demand response (DR), as one of the demand-side management services, is arguably one of the most effective ways to maintain the reliability and improve the flexibility of power system operation. DR aims to change the electricity usage of end-use customers using price-based or incentive-based mechanisms [3]. Price-based mechanisms motivate the end-use customers through the dynamic changes in electricity price over time. For example, in price-based mechanisms, the operator sets higher prices during peak load periods and encourages end-use customers to shift their loads to off-peak periods. Incentive-based mechanisms encourage end-use consumers to change their load consumption behaviors when needed through a contractual agreement.

In DR programs, DR aggregators collect DR capacity from end-use customers and then trade in the electricity market. For example, an incentive-based DR aggregator rewards participating end-use consumers based on the amount of load that they reduce during peak load periods. To calculate the load reduction amount, the electricity that would have been consumed in the absence of DR (baseline load) must be forecasted. Therefore, an accurate forecast of baseline load is essential for DR programs [4]. For example, if the baseline load is underestimated, DR aggregator may determine an incentive or a rate schedule that is not sufficient to encourage end-user customers to shift their peak loads. On the contrary, if the baseline load is overestimated, the DR aggregator may end up with spending more than needed. In addition, baseline load forecasting helps utilities with planning the operation of their generation and distribution infrastructure in advance to minimize the chance of blackouts or brownouts [5].

Baseline load forecasting can be performed at the end-use customer or aggregate level. For end-use customer level, there exist different methods for residential and commercial customers. On one hand, residential baseline load is subject to many uncertainties such as occupancy, occupant behavior and actions. Hence, many different methods have been developed for residential baseline load forecasting. These methods can broadly be categorized as averaging and regression methods [6]. Averaging methods take the average of the loads in the recent days. In taking averages, several different approaches, such as taking the average of only representative days, and taking the weighted average, can be followed. On the other hand, baseline load forecasting for commercial customers is relatively simpler because the consumption of commercial customer is quite regular and therefore the load of a representative day can be utilized as a baseline load for the future days.

Utilizing data-driven techniques is a way to develop more accurate and therefore more reliable baseline load forecasting approaches. Driven by new technologies, such as Internet-of-Things (IoT) and advanced metering infrastructure (AMI), power systems are entering into a new digital era. The massive amount of data generated through these technologies can help improve the safety, productivity, accessibility and sustainability of power systems [2]. Machine learning-based approaches are key in harnessing these data. Certainly, machine learning can be used for baseline load forecasting as well.

Towards addressing this prospect, this paper presents a machine learning-based approach using a relatively less explored algorithm, Gaussian process regression (GPR), to forecast aggregate baseline loads. In this paper, the machine learning models are trained using a dataset generated through simulating residential and commercial building models. In addition to the proposed GPR-based model, a support vector regression (SVR) based model and a traditional averaging-based model were developed as baseline models to compare with.

The structure of this paper is organized as follows: Section II presents the existing studies in the area of baseline load forecasting. Section III summarizes the GPR. Section IV presents the research methodology, which consists of three primary steps: data generation, model development and model evaluation. Section V presents the simulation results and forecasting performance of the developed models. Finally, Section VI concludes the paper.

II. Relataed Works

Supervised machine learning algorithms are designed to learn a mapping between an input vector x and an output vector y, given that there is an existing labelled dataset that includes a set of input-output pairs $\mathfrak{D} = \{(x_i, y_i)\}_{i=1}^{K}$. This attribute of the supervised machine learning algorithms makes them an ideal choice for baseline load forecasting. For baseline load forecasting, a model, which maps some features (e.g., outdoor weather conditions) into electricity loads, is trained using supervised algorithms. To the date, many supervised machine learning algorithms, including artificial neural networks (ANNs), SVR, and tree-based algorithms, have been used for baseline load forecasting.

SVR is a kernel-based algorithm. It is one of the most popular algorithms in load forecasting. For example, [7] developed an SVR-based model for forecasting energy consumption of a chiller and a supply fan in an air handling unit (AHU) using historical building operation data and weather forecast information. To demonstrate the effectiveness of SVR, [8] developed an SVR model to forecast baseline load for office buildings and compared the developed model to other seven traditional forecasting models.

ANNs are computational models, inspired by the human brain. ANNs have been used in DR primarily for forecasting applications. For example, [9] used a self-organizing map (SOM), a type of feed-forward neural network (FFNN), for residential baseline load forecasting. Similarly, [10] developed an ANN-based model to forecast residential baseline load using features including outdoor temperature, day of week, working day indicator, and previous load values.

Tree-based algorithms use a tree to map features into outputs. These algorithms have also been extensively used for forecasting applications in many DR applications due to their simplicity and relatively lower computational cost. For example, [11] built regression trees-based predictive models using historical data of office buildings. Similarly, [12] used regression tree algorithm for demand forecasting models.

Despite the importance of these studies in baseline load forecasting, there are still less explored algorithms such as GPR in the field of baseline load forecasting. GPR models, unlike other supervised machine learning algorithms, can provide probabilistic forecasts. In many DR applications, probabilistic approaches may lead to better informed decisions and planning, because DR aggregators can take the uncertainties associated with the forecasted baseline load into account [2]. Nevertheless, GPR, was used in only a few studies (e.g., [13, 14]) in the field of baseline load forecasting. Additional studies are still needed to better understand the applicability and limitations of GPR models in baseline load forecasting using different feature sets and datasets.

In addition, the majority of the machine learning-based forecasting studies focus on individual residential and/or commercial end-use customers. However, there is lack of studies in aggregate baseline load forecasting. The forecasting problem is relatively simpler for aggregate level, because the fluctuations and the noise in the individual customers may cancel each other out when considering the aggregate load [5].

III. Gaussian Process Regression

GPR is a non-parametric probabilistic kernel algorithm [15]. GPR can be applied for prediction and can provide the confidence interval for each point in the prediction to quantify the uncertainty of the prediction [16].

A regression model with noise is assumed as follows:

$$Y = f(x) + \varepsilon \tag{1}$$

where Y is the observation and $f(x)$ is an underlying function. It is assumed that the observed values Y differ from the function values $f(x)$ by additive noise ε. Furthermore, it is assumed that the noise ε follows an independent, identically distributed Gaussian distribution with zero mean and variance of σ_n^2 as follows:

$$\varepsilon \sim \mathcal{N}(0, \sigma_n^2) \tag{2}$$

Then, priori distribution of the observation Y and the joint prior distribution of the observed value Y and the prediction value y can be obtained from [15] as follows:

$$Y \sim \mathcal{N}(0, K(x,x) + \sigma_n^2 I_n) \tag{3}$$

$$\begin{bmatrix} Y \\ y \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(x,x) + \sigma_n^2 I_n & K(x, x_*) \\ K(x_* x) & K(x_*, x_*) \end{bmatrix}\right) = \mathcal{N}\left(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix}\right) \tag{4}$$

where $K(x,x) = (\kappa_{ij})$ is a symmetric positive covariance matrix, whose elements κ_{ij} measure the correlation between x_i and x_j through a kernel function κ. $K(x_* x) = K(x, x_*)^T$ is the covariance matrix between the test set x_* and the training set x.

$K(x_*, x_*)$ is the covariance matrix of the test set itself. I_n is an n-dimensional unit matrix. Radial basis function kernel (aka squared exponential kernel), is one of the most widely used kernels and can be calculated as follows:

$$\kappa_{ij} = p_1 \exp\left(-\frac{(x_i - x_j)^2}{2p_2}\right) \tag{5}$$

where p_1 and p_2 are tuning parameters. The posterior distribution of the predicted value y is

$$y|Y \sim \mathcal{N}\left(\overline{y}, \sigma_y^2\right) \tag{6}$$

$$\overline{y} = K_* K^{-1} Y \tag{7}$$

$$\sigma_y^2 = K_{**} - K_* K^{-1} K_*^T \tag{8}$$

Thus, the predictions of GPR is $\overline{y}$ and the prediction interval with 95% confidence level is $[\overline{y} - 1.96\sigma_y, \overline{y} + 1.96\sigma_y]$. The probability density function of the *i*-th predicted value is as follows:

$$p(y_i) = \frac{1}{\sqrt{2\pi}\sigma_{yi}} exp\left(-\frac{(y_i - \overline{y}_i)}{2\sigma_{yi}^2}\right) \tag{9}$$

IV. METHODOLOGY

Three primary steps were followed to develop the proposed GPR-based aggregate baseline load forecasting model. First, a dataset including outdoor weather conditions and resulting electricity load profile were generated. Second, the datasets were preprocessed to make them compatible with machine learning algorithms. Third, the GPR-based forecasting model and a set of baseline models were developed and compared.

A. *Dataset Generation and Preprocessing*

To train the machine learning models, a dataset was generated using EnergyPlus simulations. The dataset included (1) aggregate consumption of 100 residential heating, ventilation, and air conditioning (HVAC) units, 100 commercial HVAC units, and 100 water heater (WH) units, and (2) the corresponding weather conditions.

To create a dataset that represents the overall characteristics of the U.S. building stock to some extent, the EnergyPlus models provided by the U.S. Department of Energy were used [17]. The prototype single-family building model was used for the residential HVAC and WH units. The reference small office building model was used for the commercial HVAC units.

The models were simulated in EnergyPlus using the typical meteorological year 3 (TMY3) weather data of Atlanta, GA. Then, the aggregate baseline load profile of these 300 units was computed by aggregating the hourly cooling and water heater electricity consumption profiles by date and time.

B. *Data Preprocessing and Model Development*

Prior to machine learning model development, non-summer months were filtered out. Only temperature and solar radiation data were extracted from the weather conditions data, and other irrelevant variables were removed from the dataset. Hourly consumption and weather conditions profiles were interpolated into 10-minute intervals. The weather conditions data were normalized using their means and standard deviations. A working day/hour indicator was created using the time stamps to indicate whether a particular data instance is subject to thermostat setpoint or setback hours. Finally, the dataset was split into training and testing datasets with proportions of 75% and 25%, respectively.

In developing the prediction models, in addition to the proposed GPR-based model, an SVR-based model and an averaging-based model were developed as baseline models for comparison. GPR- and SVR-based models take outdoor temperature, solar radiation, and working day/hour indicator values as features. These models were trained on the training dataset. The averaging-based model simply takes the average of past 10 days.

C. *Model Evaluation*

The models were evaluated based on the three criteria suggested by [10]: accuracy, simplicity, and integrity.

To evaluate the model accuracy, the trained models were used for forecasting the aggregate baseline load. The predicted load values were compared to the actual load values in the testing dataset and the root mean square error (RMSE) was calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(y_{predict,\text{t}} - y_{data,\text{t}})^2}{n}} \tag{10}$$

where $y_{predict,\text{t}}$ is the forecasted load value at time t, $y_{data,\text{t}}$ is the actual load value at time *t*, and n is the number of instances in the testing dataset. The lower the RMSE, the less dispersions are between the forecasted and the actual load values.

An aggregate baseline load forecasting model should also be simple so that it can easily be communicated to the concerned stakeholders (e.g., aggregators and customers) [18]. To evaluate the simplicity of the models, the number of features that the models take, and efforts required to develop the models were considered.

In DR programs, because the participating customers are rewarded based on the amount of load that they reduce during peak load periods, they may choose to reflect their baseline load more than their actual baseline load to be rewarded more. Therefore, integrity is also an important factor to consider when evaluating the baseline load forecasting models. A model should make it difficult to “cheat” the system in favor of a stakeholder. To assess this, whether there is a manipulation opportunity by a customer was also checked.

V. RESULTS

A. Generating Aggregate Baseline Loads

The EnergyPlus simulations generated hourly energy consumption levels for residential HVAC, commercial HVAC, and WH units. Figure 1 shows a sample day of these simulation results. Figure 2 shows the aggregate load and corresponding outdoor weather conditions for that day. As shown in the figures, the residential and commercial HVAC load is relatively lower before 6:00 due to the low outdoor temperature and solar radiation levels. Parallel to the increases in outdoor temperature and solar radiation levels, HVAC loads as well as the aggregate load start to increase after 6AM. The load of the water heaters is higher during morning and evening times, but their pattern does not affect aggregate baseline load significantly.

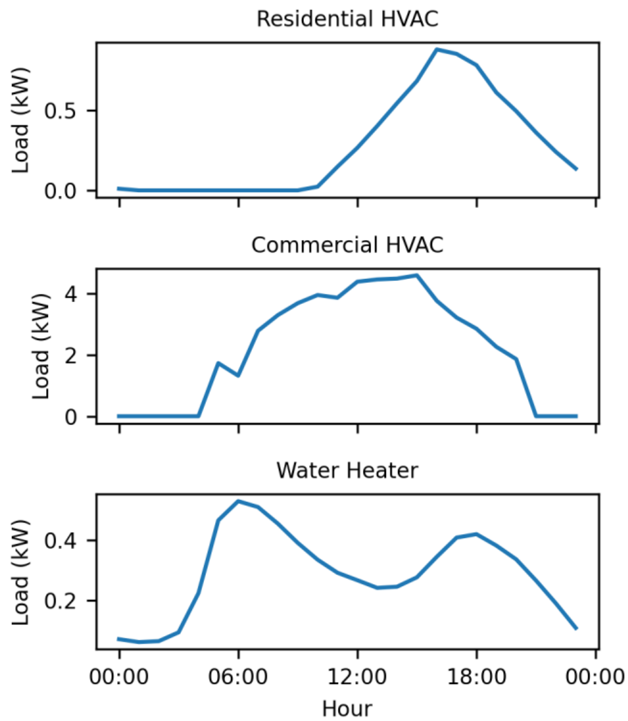

Figure 1. Individual loads of residential HVAC, commercial HVAC, and water heater units.

B. Forecasting Performance

Figure 3 shows the prediction results of the previously described three algorithms on a few days from the testing dataset. The GPR-based model achieved the highest prediction accuracy. The GPR-based, SVR-based, and averaging-based models achieved RMSEs of 85.3 kW, 115.6 kW, and 89.6 kW, respectively. In addition, for almost all testing dataset, the confidence interval determined by the GPR-based model covered the actual load. The 95% confidence interval of the GPR-based model was able to cover the actual load 93% of the time.

In terms of simplicity, the averaging-based model is superior than the GPR-based and SVR-based models due to two reasons. First, unlike machine learning-based models, the averaging-based model does not require any model training. However, machine learning-based models require time consuming model training efforts which involve data collection and preprocessing, and parameter tuning of the selected algorithms. Second, the averaging-based model can forecast the future baseline load using only historical baseline load, whereas the GPR-based and SVR-based models require an accurate weather forecasting.

In terms of integrity, the GPR-based and SVR-based models are better than the averaging-based model. Machine learning-based models reduce the risk of "cheating" to almost zero as they are regression models depending on a long historical dataset. Increasing load on a few days to be rewarded more would not change the forecasts of the GPR-based and SVR-based models. However, the averaging-based model is very sensitive to such artificial changes as it forecasts based on a few recent days only.

Overall, the GPR-based model is better than the SVR-based and averaging-based models due to its higher accuracy and lower chance of being subject to a DR trick. Furthermore, unlike the traditional supervised learning algorithms, GPR can also provide probabilistic forecasts. The only drawback associated with GPR is its relatively lower simplicity. However, this is minor drawback due to the recent advances in computing power. For example, the GPR-based model presented in this paper was trained in 15 seconds only using a typical laptop machine. In addition, the GPR-based model presented in this paper only requires temperature and solar radiation forecasting, which can be easily obtained from weather stations, or from other recent methods (e.g., [19]).

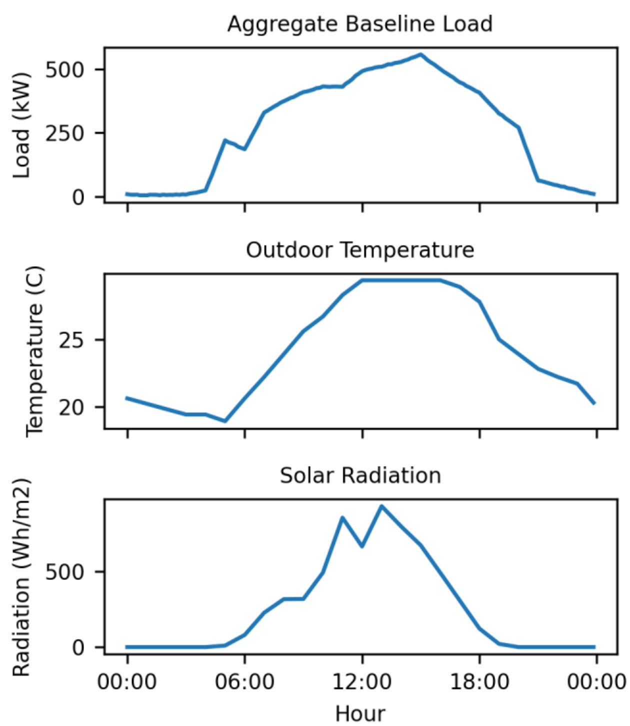

Figure 2. Aggregate baseline load and corresponding outdoor weather conditions.

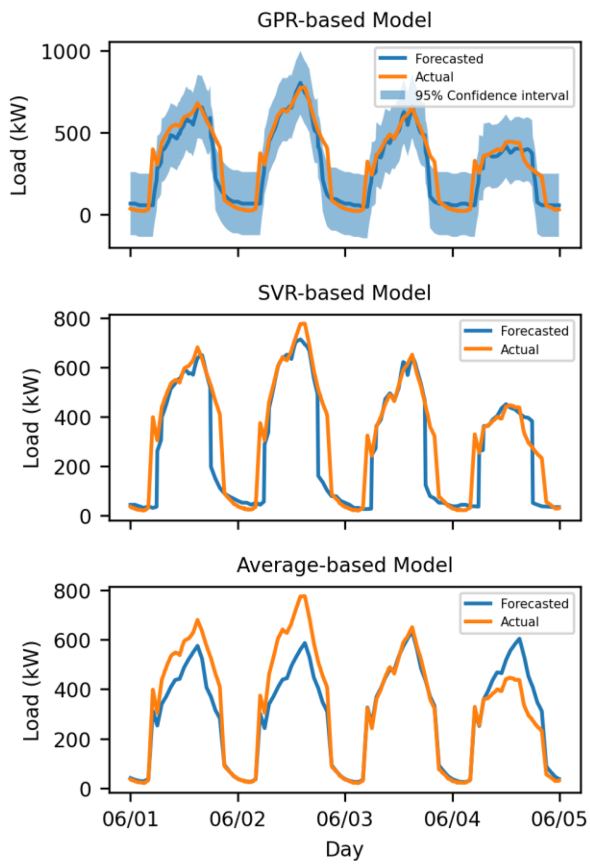

Figure 3. Forecasting performace of the algorithms.

VI. CONCLUSIONS

This paper developed a GRR-based model for aggregate baseline load forecasting using a dataset generated through EnergyPlus simulations. The GPR-based model was compared against two baseline models, including SVR- and averaging-based models, in terms of accuracy, simplicity, and integrity. The results showed that the GPR-based model is superior to the SVR-based and averaging-based models. Such high performance proves the potential of the GPR in aggregate baseline load forecasting. GPR, therefore, can be used for demand response applications.

ACKNOWLEDGMENT

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Building Technologies Office under contract DE-AC05-00OR22725.

REFERENCES

[1] W. Tang, Y. Wu, and H. Yang, "Adaptive segmentation and machine learning based potential DR capacity analysis," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, 13-16 May 2018, pp. 1-5.

[2] I. Antonopoulos *et al.*, "Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review," *Renewable and Sustainable Energy Reviews*, vol. 130, p. 109899, 2020/09/01/ 2020, doi: https://doi.org/10.1016/j.rser.2020.109899.

[3] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Applied Energy*, vol. 253, p. 113595, 2019/11/01/ 2019, doi: https://doi.org/10.1016/j.apenergy.2019.113595.

[4] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalão, "Synchronous pattern matching principle-based residential demand response baseline estimation: mechanism analysis and approach description," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6972-6985, 2018.

[5] S. Bandyopadhyay, T. Ganu, H. Khadilkar, and V. Arya, "Individual and aggregate electrical load forecasting: One for all and all for one," Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems, Bangalore, India, 2015. [Online]. Available: https://doi.org/10.1145/2768510.2768539.

[6] K. Coughlin, M. A. Piette, C. Goldman, and S. Kiliccote, "Statistical analysis of baseline load models for non-residential buildings," *Energy and Buildings*, vol. 41, no. 4, pp. 374-381, 2009/04/01/ 2009, doi: https://doi.org/10.1016/j.enbuild.2008.11.002.

[7] C. Yang, Q. Cheng, P. Lai, J. Liu, and H. Guo, "Data-driven modeling for energy consumption estimation," in *Exergy for A Better Environment and Improved Sustainability 2: Applications*, F. Aloui and I. Dincer Eds. Cham: Springer International Publishing, 2018, pp. 1057-1068.

[8] Y. Chen *et al.*, "Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings," *Applied Energy*, vol. 195, pp. 659-670, 2017/06/01/ 2017, doi: https://doi.org/10.1016/j.apenergy.2017.03.034.

[9] S. Park, S. Ryu, Y. Choi, J. Kim, and H. Kim, "Data-driven baseline estimation of residential buildings for demand response," *Energies*, vol. 8, no. 9, pp. 10239-10259, 2015. [Online]. Available: https://www.mdpi.com/1996-1073/8/9/10239.

[10] J. Jazaeri, T. Alpcan, R. Gordon, M. Brandao, T. Hoban, and C. Seeling, "Baseline methodologies for small scale residential demand response," in *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, 28 Nov.-1 Dec. 2016 2016, pp. 747-752.

[11] M. Behl, F. Smarra, and R. Mangharam, "DR-advisor: A data-driven demand response recommender system," *Applied Energy*, vol. 170, pp. 30-46, 2016/05/15/ 2016, doi: https://doi.org/10.1016/j.apenergy.2016.02.090.

[12] Y. Simmhan *et al.*, "Cloud-based software platform for big data analytics in smart grids," *Computing in Science & Engineering*, vol. 15, no. 4, pp. 38-47, 2013.

[13] Y. Weng and R. Rajagopal, "Probabilistic baseline estimation via gaussian process," in *2015 IEEE Power & Energy Society General Meeting*, 26-30 July 2015, pp. 1-5.

[14] Y. Weng, J. Yu, and R. Rajagopal, "Probabilistic baseline estimation based on load patterns for better residential customer rewards," *International Journal of Electrical Power & Energy Systems*, vol. 100, pp. 508-516, 2018/09/01/ 2018, doi: https://doi.org/10.1016/j.ijepes.2018.02.049.

[15] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[16] Z. Zhang *et al.*, "Wind speed prediction method using shared weight long short-term memory network and gaussian process regression," *Applied Energy*, vol. 247, pp. 270-284, 2019/08/01/ 2019, doi: https://doi.org/10.1016/j.apenergy.2019.04.047.

[17] U. S. D. o. Energy. "Building energy codes program." https://www.energycodes.gov/development (accessed 9/28/2020, 2020).

[18] D. G. Holmberg, D. Hardin, and E. Koch, "Towards demand response measurement and verification standards," in *2012 Grid Interop*, Irving, TX, December 3-6, 2012 2012.

[19] F. M. Lopes, H. G. Silva, R. Salgado, A. Cavaco, P. Canhoto, and M. Collares-Pereira, "Short-term forecasts of GHI and DNI for solar energy systems operation: assessment of the ECMWF integrated forecasting system in southern Portugal," *Solar Energy*, vol. 170, pp. 14-30, 2018/08/01/ 2018, doi: https://doi.org/10.1016/j.solener.2018.05.039.