Net Load Forecast based on Behind-the-Meter Disaggregation of Smart Meter Data

Mian Jia*, Kang Pu*, and Yue Zhao†
*Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794, USA
†Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794, USA
Emails: {mian.jia, kang.pu, yue.zhao.2}@stonybrook.edu

***Abstract*—As the penetration of behind-the-meter (BTM) rooftop solar energies continues to increase in power distribution systems, it is of paramount importance for load serving entities and system operators to forecast net loads in the system. In this paper, novel algorithms for training high-performance predictors for day-ahead net load forecasting are developed. Importantly, the overall method only utilizes metered net load data and does not require any monitoring data of solar generation. Methodologically, the net load data trace is disaggregated into estimated BTM solar and load traces, based on which separate predictors are then trained for solar generation and load forecasting exploiting their distinct natures, respectively. For solar generation forecasting, time data, weather forecast, and potentially solar irradiance forecast are used as input features of the predictor. For load forecasting, time data, weather forecast, and judiciously chosen load data in the recent past are used as input features of the predictor. The two predictors' outputs are combined to produce the final net load forecast. The developed method is comprehensively evaluated based on two real-world smart meter data sets from Ithaca, NY and Clifton park, NY, respectively. High accuracy of day-ahead net load forecast is demonstrated.**

I. Introduction

Day-ahead net load forecast, especially at aggregate levels (e.g., feeder, substation, and above), is of paramount importance for efficient and reliable system operations by utilities and independent system operators (ISOs). Net load forecast is crucial for efficient procurement of energy supplies in energy markets: e.g., ISO relies on net load forecast to perform optimal generation scheduling to serve the loads, and load serving entities need such information to do optimal bidding in electricity markets. Net load forecast also provides critical indicators of system reliability risks, e.g., potential overloading in parts of the power transmission and distribution systems.

Traditionally, net load forecasting is mostly about load forecasting. With the rapid increase of behind the meter (BTM) renewable energies, predominantly rooftop solar, the composition of net load changes and thus the problem nature of net load forecasting changes fundamentally. It is important to note that, unlike utility-scale renewable energy sites, BTM solar generation is not monitored by utilities (and hence the term "behind the meter"), and all that utilities can measure are the net loads. While existing data-driven methods designed for load forecasting could be applied to net load forecasting by replacing the load data with the net load data, there tends to be performance degradation because load forecasting methods do not distinguish and exploit the two very different components of net load — solar generation and load.

This work is supported by NYSERDA with Agreement Number 159613.

There have been a large number of studies on the respective topics of solar generation forecasting and electric load forecasting. We refer the readers to the review papers [1] and [2] among all these works. For net load forecasting in the presence of both solar generation and loads, recent works have trained predictors that a) take net loads (as opposed to loads) and informative solar energy related features (such as sky images and separately measured solar generation) as inputs, and b) directly output net load forecast [3], [4]. Different from such "direct" strategies, there have been works that treat solar generation forecasting and load forecasting as two separate tasks and integrate the respective results in the final stage. In [5], a Wavelet Neural Network was developed and it significantly outperformed direct strategies in high solar penetration scenarios. In [6], a Dynamic Gaussian Process and quantile regression were employed, and the impact of aggregating customers was also investigated. In [7], autoencoders and cascade neural networks were employed to perform forecasting across different time horizons. Notably, however, these approaches all rely on *separately measured* solar and load traces which serve as supervision signals in training the separate predictors for solar generation and load, respectively. In practice, however, it is often the case that BTM solar and load data are *not* available to the utilities and ISOs. To address this issue in order to still use separate predictors for solar generation and loads, [8] estimates solar generation and loads from net load data based on maximal information coefficient based correlation analysis and a grid search. It then decomposes net-loads into solar generation, loads, and residuals, where separate predictors are trained for forecasting these three components. However, the disaggregation approach employed therein is relatively simple.

In this work, we develop a novel two-step approach for net load forecasting that first a) disaggregates historical net load data to obtain BTM solar generation and load traces, and b) utilizes the disaggregated traces to separately train predictors for solar generation and load forecast, before combining them to form the eventual net load forecast. Notably, we employ the *state of the art* BTM disaggregation algorithm from our recent work [9]. We show that the ensuing separate training of predictors based on our disaggregated data leads to a net load forecast performance *very close to the ideal performance bound* achieved by training with *ground truth* BTM solar generation and load data. In particular, recognizing the distinct natures of the two sub-problems of solar generation and load forecast, we develop different strategies of training effective predictors for these two sub-problems respectively. Moreover, we develop a novel *customer clustering* component in the load forecasting sub-task based on the disaggregated load traces of *individual* customers. As such, clusters of customers with similar behaviors are identified, and separate predictors are trained to better capture the different customer behavior patterns. In this work, compared with short-term forecasting, we focus on the more challenging *day-ahead* net load forecasting problem.

II. Problem Formulation

For a set of residential energy customers, their total net load at time t is denoted by $n(t)$. We study the problem of *day-ahead* forecast of this total net load. Our objective is to predict, at any given time t, the vector of net loads in a future time window of $[n(t+\Delta_1), n(t+\Delta_2)]$. Specifically, we consider $\Delta_1 = 13$ and $\Delta_2 = 36$ in this paper, where we employ hour as the unit of time. The reason is that this time window aligns with the operation routine of the day-ahead power markets in some independent system operators (ISOs), e.g., New York ISO (NYISO): every day, at 11am, the day-ahead market operation decisions must be determined, and hence *all* 24 *hours'* net loads of the next day need to be predicted. This implies the need for net load predictions from 13-hour ahead to 36-hour ahead. We note that a simplified objective of 24-hour ahead prediction, i.e., predicting $n(t+24)$, is often employed in existing studies. Clearly, this is included in our objective as a special case. We further note that the proposed method in this work can be straightforwardly adapted to net load forecast with different look-ahead times.

With smart meters, residential customers' net loads are directly measured, and a typical time granularity is 15 minutes. Based on massive *historical net load data* collected by smart meters, it is intuitive that some net load patterns of the customers may be captured and useful information could be provided for day-ahead net load predictions. Furthermore, *weather* also has an influence on both energy consumption and solar generation. As such, we aim to develop a predictor that takes a) historical net load measurements from smart meters, b) day-ahead weather forecast, and c) time related information such as time of day and day of year, as the input variables, and outputs the next day's net load forecast.

A. Decomposition of Net Load by BTM Smart Meter Data Disaggregation

Importantly, we have seen a steady increase in rooftop solar panel installation among residential customers, and the trend is only accelerating. Such solar energies are however installed "behind-the-meter", meaning that their generation are not separately metered or monitored in any way by the utility companies. Instead, for a given customer k, its smart meter measures its *net load*, which is equivalent to the difference between load and solar generation, i.e.

$$n_k(t) = l_k(t) - g_k(t), \tag{1}$$

where $n_k(t), l_k(t)$ and $g_k(t)$ are its net load, load and solar generation, respectively.

Hypothetically, let us assume that, not only $n_k(t)$, which is measured by smart meters, but also both the load and solar generation traces $l_k(t)$ and $g_k(t)$ are separately known: *they would fundamentally provide more information than that of just the net load trace* $n_k(t)$, and hence would in principle enable more accurate forecast of future net loads.

Notably, the traces of the load $l_k(t)$ and that of the solar generation $g_k(t)$ have very different natures:

- The load is primarily driven by human activities and needs. While weather clearly has an impact on load, it is just one of the potentially many determining factors.
- In contrast, given the installation parameters of the solar panel, solar generation is directly determined by meteorological conditions.

Consequently, an important implication is as follows: *Conditioned on knowing the weather information of the next day*,

- The historical load data still provide important predictive information on the loads of the next day due to the *intrinsic temporal correlations* of human behaviors. For example, the load pattern of the next day likely bears some similarities to the current day.
- In contrast, the historical solar generation has *no conditional correlation* with the solar generation of the next day, as the latter is fully determined by the corresponding weather conditions.

Accordingly, predicting future loads and predicting future solar generation require fundamentally different strategies. Indeed, in this work, we employ two distinct strategies tailored for day-ahead forecasting of load and for that of solar generation, respectively. The results are then combined to provide the day-ahead net load forecast.

Having said the above, separate data traces of load and solar generation are however not measured by the smart meters and hence are not available to the utilities. To exploit different strategies for solar and load forecast, we address this issue by employing a two-step methodology: We first a) *disaggregate* the net load traces into solar generation traces and load traces in a fully unsupervised fashion (i.e., without utilizing any BTM measurements), and then b) utilize the disaggregated solar and load generation traces to train separate predictors for day-ahead forecast of solar generation and load, respectively.

III. The Proposed Forecasting Method

In this section, we introduce in detail the proposed methodology for net load forecasting.

A. BTM Solar Energy Disaggregation from Net Load

Given the net load data measured by smart meters, the first step of our method is to disaggregate the net load traces into BTM solar generation traces and load traces. Specifically, we employ the method developed in our recent work [9] which represents the state of the art. As a result, for any customer k, its net load trace $n_k(t)$ is disaggregated into its solar generation trace $g_k(t)$ and load trace $l_k(t)$, so that $n_k(t) = l_k(t) - g_k(t)$. In brief, our method is an unsupervised

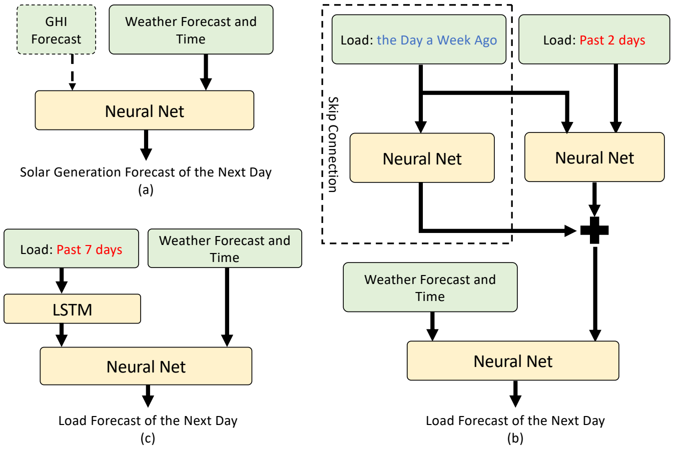


Figure 1: Predictor models: (a) Solar generation forecast (FCNN-based); (b) Load forecast (FCNN-based); (c) Load forecast (LSTM-based).

learning based algorithm that exploits generic physical models of solar panels. It takes a) net load traces and b) solar irradiance, weather and time data as inputs, and outputs the BTM solar generation (and hence load) traces. The main ideas of this disaggregation method are to a) find pairs of time slots in which customers likely have similar loads, and b) estimate the unknown parameters of the solar panel physical model by minimizing a loss constructed using these time pairs with similar loads. With the estimated physical parameters of the solar panels, the solar generation traces $g_k(t)$ can then be computed based on the solar irradiance, weather and time data, and the load $l_k(t) = n_k(t) + g_k(t)$ can also be recovered. Due to space limitations, we refer the readers to [9] for more details. With the disaggregated data, we next design algorithms for forecasting solar generation and load traces, respectively.

B. Solar Generation Forecast

Based on the disaggregated solar generation traces $g_k(t), \forall k$, we compute the total solar generation $g(t) = \sum_k g_k(t)$ of the set of customers. We then train predictors for day-ahead forecasting of the total solar generation $g(t)$.

1) Feature Selection: For a future time slot of interest $t+T$, we employ the following input features for predicting $g(t+T)$:

- Time and location information: time of day $\in [1, 96]$, and solar zenith angle. We note that the 96 time slots are from the fact that the data have 15-minute intervals, and thus there are 4 time slots in each of the 24 hours.
- Weather information: weather forecast including temperature, cloud type, humidity, dew point, precipitable water, pressure, and wind speed.

It is worth mentioning that the above weather forecast is standard practice and widely available and used in many applications. In addition to the above, another piece of information that can be particularly helpful for forecasting solar generation is solar irradiance. Existing works have demonstrated that day-ahead forecasting of solar irradiance can also be performed with reasonable accuracy [10]. As such, in our study, we consider both of the following two problem settings:

- Day-ahead solar generation forecasting *without* using any solar irradiance forecast information.
- Day-ahead solar generation forecasting *with* solar irradiance, in particular, global horizontal irradiance (GHI) forecast as an input feature.

The first setting is more conservative without assuming the knowledge of solar irradiance forecast and uses only standard weather forecast information. In the second setting, we will demonstrate how even better performance can be achieved if solar irradiance forecast is available.

Notably, the historical solar generation data are *not* included as an input feature because of the conditional independence of future solar generation and past solar generation, conditioned on the knowledge of future weather, (cf. Section II-A). This is not to be confused with the fact that the historical solar generation data are used in *training* the predictors. Specifically, the solar generation predictors are trained in a supervised fashion with a) the historical solar generation data as training labels of the predictor outputs, and b) the historical time, weather, and potentially GHI data in the corresponding time slots as the predictor inputs.

2) Predictor Model: We employ neural networks with fully connected layers as the predictor model. A diagram for this architecture is depicted in Fig. 1(a). For the time of day data, due to their circular nature (e.g., hour 1 and hour 24 are in fact very close to each other), we employ a two dimensional transformation using sine and cosine functions to properly encode such circular variables [11]. For the categorical weather variable of cloud type, one-hot encoding is applied.

C. Load Forecast

Based on the disaggregated load traces $l_k(t), \forall k$, we compute the total load $l(t) = \sum_k l_k(t)$ of the set of customers. We then train predictors for day-ahead forecasting of the total load $l(t)$.

1) Feature Selection: In addition to the time and weather information as in Section III-B, we further employ *historical load data* as part of the inputs to the predictors. The way we include the historical load data in the input features depends on the predictor model architecture. In particular, we develop two types of predictor model architectures: a) fully connected neural networks (FCNN), and b) recurrent neural networks, in particular, long short-term memory (LSTM) networks.

For FCNN predictors, at the current time t, for predicting the loads in the future time window $[t+13, t+36]$, we utilize a) the past two days' load data up to the current time, i.e., $l(t-47), \ldots, l(t)$, and b) the load data in the day exactly *one week* before the day to predict, i.e., $l(t-155), \ldots, l(t-132)$. The reasoning is the following: a) the load pattern in the very recent past, represented by the past two days, likely provides useful information for the load pattern of the very near future, and b) the load pattern in the same day a week ago also likely provides useful information due to the weekly activity patterns of the customers. For LSTM predictors, the past 7 days of historical load data trace is utilized as predictor inputs.

2) Predictor Model Architecture: For FCNN-based predictors, a diagram of the neural network architecture is depicted in Fig. 1(b). In this architecture, we employ an idea inspired by "skip connections" in residual neural networks [12]. Specifically, we introduce a "skip connection" branch (cf. the left branch in Fig. 1(b)) of the neural network solely for the loads from the day one week before the predicted day. The output is then added to the general branch (cf. the right branch in Fig. 1(b)) of neural network. The intuition is to exploit the weekly patterns of customer behaviors via this skip connection so that the general branch can focus on learning the more intricate load patterns beyond this weekly pattern. For LSTM-based predictors, we employ two LSTM layers followed by one fully connected layer (cf. Fig. 1(c)).

3) Clustering of Customers: As our objective here is day-ahead forecasting of the *total* loads of a set of customers (e.g., at the feeder level), it is sufficient to train the load forecasting predictor based on just a *single* trace of the *total* load disaggregated from the total net load. Having said this, with the availability of smart meter data from *individual* customers, the *individual* load traces can fundamentally offer more information to improve the forecasting performance of the total load. The intuition is to design predictors specialized for different customers' behaviors that exhibit distinct patterns: Training more nuanced predictors that separately forecast loads of different patterns can perform better than training just a single predictor for the total load in which all different load patterns are aggregated. As such, we perform a step of clustering of customers based on their load traces, and train a separate predictor for each cluster to forecast its corresponding cluster-wise total load. We then combine the cluster-wise predictors to form the forecast of the total load.

Specifically, we employ a hierarchical clustering method – the agglomerative clustering method with Ward's linkage – on the set of load traces of all the customers, In brief, Ward's method is an iterative algorithm: at each iteration, a pair of clusters that leads to the minimum increase in total within-cluster variance are merged [13].

IV. Data-Driven Evaluation

In this section, we evaluate the performance of our methods on two real-world smart meter data sets: a) a Pecan Street Inc. data set [14] of 12 BTM-PV-owning residential customers residing in Ithaca, NY, collected in a six-month period from 05/01/2019 to 10/31/2019, and b) a National Grid data set of 185 BTM-PV-owning residential customers residing in Clifton Park, NY, collected in a 12-month period from 01/01/2019 to 12/31/2019. Both data sets are collected with 15 minutes intervals. In the Ithaca data set, the ground truths of BTM solar generation are also separately measured and available. In the Clifton Park data set, only the net loads measured by smart meters are available which is the typical case in practice for most utilities. The weather data, including GHI data, for the same periods of time and from (approximately) the same locations are collected from the National Solar Radiation Database (NSRDB) [15].

For each data set of net loads, we perform BTM disaggregation of the net load traces (cf. Section III-A). to obtain estimated solar generation traces and a load traces. The solar generation and load predictors are then trained based on these two disaggregated traces, respectively. Notably, for the Ithaca data set, since the ground truth solar and load data are also available, this allows us to compute a performance bound achieved by training with these BTM ground truths, which would typically not be available in practice although available in this particular data set. We will then compare the performance achieved by using the disaggregated data, which is practical, with this performance bound which is based on impractical BTM measurements.

Given that the Ithaca data set consists of only 12 customers, whereas the Clifton Park data set have 185, we will perform customer clustering for load forecasting (cf. Section III-C3) with the latter but not the former. To comprehensively evaluate the proposed method, we perform a 5-fold cross-validation-like testing. Specifically, we evenly divide both the Ithaca and Clifton Park data sets into 5 folds. For each fold as the testing data, we train the predictors on the other 4 folds. The average testing performance across all the 5 folds are computed.

Evaluation Metrics: To evaluate forecasting accuracy, we utilize the Mean Squared Error (MSE) and hourly Normalized Mean Absolute Error (nMAE) as the metrics:

$$\text{MSE} = \frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x_t})^2, \tag{2}$$

$$\text{nMAE} = \frac{1}{N}\sum_{t=1}^{N}\frac{|x_t - \hat{x_t}|}{\text{range}(x)} \times 100\%, \tag{3}$$

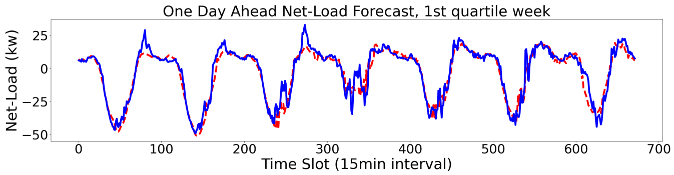
where x_t and $\hat{x_t}$ are the ground truths and predicted values of the variable to forecast, respectively, and range(x) is defined to be the difference between the highest and lowest realizations of x in the period of evaluation.

A. Ithaca Case Study

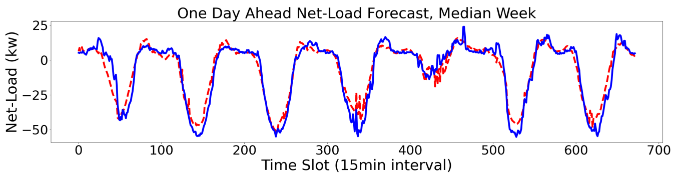
We first present the performance evaluated with the Ithaca data set under the following three dichotomies:

- Train the predictors either a) based on the ground truth BTM solar generation and load traces, or b) based on the disaggregated traces (cf. Section III-A) which are estimation of the ground truths. Clearly, the former is expected to have a better performance.
- Either a) do not utilize GHI forecast information, or b) include GHI forecast as part of the input features. The latter is expected to have a better performance. Specifically, prior research has shown that day-ahead forecast of GHI can achieve an RMSE of as low as 6.6 W/m^2 [10]. In our experiments, we make a conservative assumption of a forecast RMSE of 10.0 W/m^2.
- For load forecast, either a) employ FCNN as the predictor architectures, or b) employ LSTM.

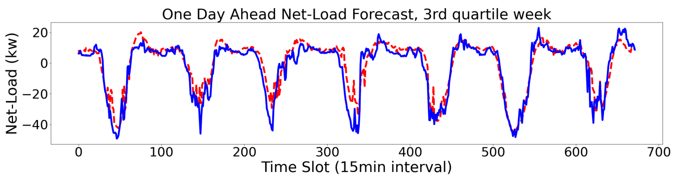
The eight performance evaluation under these dichotomies are summarized in Table I. Furthermore, as a baseline method that does *not* exploit disaggregated solar generation and load traces, we train a FCNN predictor that takes the historical *net load* data (in addition to the other time and weather related data) as inputs, and directly produces the net load forecasts as



(a)



(b)



(c)

Figure 2: Ithaca data set's testing performance (using LSTM without GHI forecast): forecast (red dashed) vs. ground truth (solid blue) traces for three representative weeks.

Table I: Testing MSE and nMAE of Day-Ahead Net Load Forecast, Ithaca, NY

MSE/NMAE	FCNN w/o GHI	FCNN w/ GHI	LSTM w/o GHI	LSTM w/ GHI
Ground Truth	38.145/4.038%	25.148/3.322%	37.519/4.002%	24.939/3.312%
Disaggregation	40.603/4.294%	28.115/3.604%	39.854/4.280%	27.640/3.646%
Baseline	57.358/5.525%			

its outputs. The performance of the baseline for the setting of FCNN without GHI is presented in Table I.

We make the following observations:

1) Compared with the baseline, significant performance gain is achieved by performing the step of disaggregation and exploiting the disaggregated BTM solar generation and load traces for separate training.
2) FCNN and LSTM have similar performance (with LSTM being slightly better in this case).
3) Having day-ahead GHI forecast information can significantly further improve the net load forecast accuracy due to the improved solar generation forecast.
4) *Without* knowing the ground truth BTM solar generation and load traces, based on the disaggregated estimates of these traces, the performance is in fact very close to that achieved *with* the ground truth knowledge. This is great news for applying the developed methods in practice, because a) such BTM ground truth information are typically not available to the utilities, and yet b) the utilities can still achieve almost the same day-ahead net load forecasting performance as if the BTM ground truths are known.

To visualize the performance, we plot the traces of the 24-hour ahead net load forecast vs. the ground truths in three representative testing weeks in Figure 2. Specifically, among all the weeks ranked by their MSEs, we choose the three weeks at the 1st, 2nd (i.e., the median), and 3rd quartile marks. As such, these three weeks exhibit above-average, average, and below-average performance, respectively. We note that these plots are derived based on the results from using LSTM but *without using GHI forecast* as input. We observe that the day-ahead net load forecasting accuracy is reasonably high. As shown in Table I, if GHI forecasts are available as input information to the predictors, the net load forecasting performance will be even better.

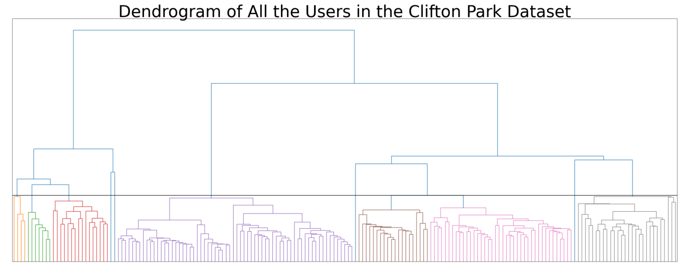


Figure 3: The dendrogram illustrating the result of hierarchical clustering performed on the 185 customers.

B. Clifton Park Case Study

We now present the performance evaluated with the Clifton Park data set. Recall that, in this data set, we only have the measured net loads, and the ground truth BTM solar generation and loads are not available. Thus, we only evaluate the predictors trained based on the disaggregated solar generation and load traces.

With the relatively large number of customers in the Clifton Park data set, we perform a step of customer clustering (cf. Section III-C3) for load forecasting. We plot the clustering results in Fig. 3: All 185 customers were clustered into 8 groups as marked by the black horizontal line.

To verify the performance gain with clustering, we evaluate the load forecasting performance achieved by a) training 8 predictors for each cluster separately and b) combining the 8 predictors to obtain the total load forecast. We also evaluate the performance without clustering. The testing performance comparison is shown in Table II. We note that this is for *load* forecasting which is one of the two critical components (the other being solar generation forecasting) for our net load forecasting method. We observe that, with the help of clustering, the MSE of the day-ahead load forecast can be significantly lowered.

Table II: Testing MSE of Day-Ahead Load Forecast by LSTM

MSE	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean
w/o Clustering	535.094	402.333	558.327	526.258	477.699	501.942
w/ Clustering	504.230	383.526	500.289	451.976	425.112	453.027

Next, utilizing clustering in load forecasting, we evaluate the net load forecasting performance under the following two dichotomies:

- Either a) do not utilize GHI forecast information, or b) include GHI forecast as part of the input features. The latter is expected to have a better performance.

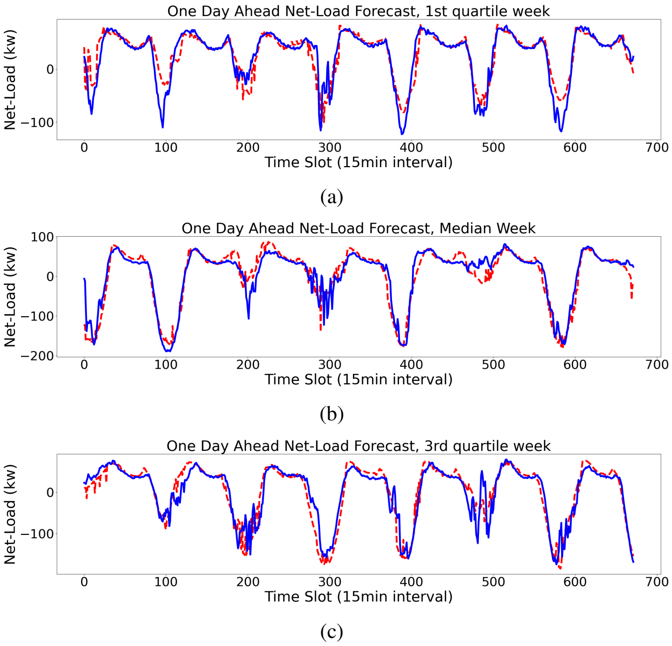


Figure 4: Clifton Park data set's testing performance (using LSTM without GHI forecast): forecast (red dashed) vs. ground truth (solid blue) traces for three representative weeks.

- For load forecast, either a) employ FCNN as the predictor architectures, or b) employ LSTM.

Again, as a baseline method that does not exploit disaggregated solar generation and load traces, we further train a FCNN predictor that takes the historical net load data (in addition to the other time and weather related data) as inputs, and directly produces the net load forecasts as its outputs. The performance comparison is summarized in Table III.

Table III: Testing MSE and nMAE of Day-ahead Net Load Forecast, Clifton Park, NY

MSE/NMAE	FCNN w/o GHI	FCNN w/ GHI	LSTM w/o GHI	LSTM w/ GHI
Disaggregation	538.325/4.006%	502.275/3.840%	513.876/3.864%	475.373/3.704%
Baseline	708.518/5.127%			

We make the following observations:

1) Compared with the baseline, significant performance gain is again achieved by performing the step of disaggregation and exploiting the disaggregated BTM solar generation and load traces for separate training.
2) LSTM has a noticeably higher performance than FCNN.
3) Having day-ahead GHI forecast information can again greatly improve the net load forecast accuracy due to the improved solar generation forecast.

We note that the above observations for the Clifton Park data set are consistent with those for the Ithaca data set.

To visualize the performance, we plot the traces of the 24-hour ahead net load forecast vs. the ground truths in three representative testing weeks (corresponding to the 1st, 2nd, and 3rd quartiles of testing MSEs) in Figure 4. We note that these plots are derived based on the results from using LSTM but *without using GHI forecast* as input. We observe that the day-ahead net load forecasting accuracy is again reasonably high. As shown in Table III, if GHI forecasts are further available as input information to the predictors, the net load forecasting performance will be even better.

V. Conclusion

We developed an effective data-driven method for training day-ahead net load predictors that do not assume any behind-the-meter sensor data. The method disaggregates the net load data traces into estimated BTM solar generation and load data traces, and utilizes such disaggregated data traces to train two separate predictors for solar generation and load forecasting. The distinct natures of these two sub-problems are exploited for effective input feature selection and neural network architecture design. The outputs of the two trained predictors are then combined to produce the net load forecast. Comprehensive evaluations with two real-world data sets collected in New York demonstrated the high forecast accuracy of the method.

References

[1] R. Ahmed, V. Sreeram, Y. Mishra, and M. Arif, "A review and evaluation of the state-of-the-art in pv solar power forecasting: Techniques and optimization," *Renewable and Sustainable Energy Reviews*, vol. 124, pp. 109–792, 2020.

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

[3] Y. Chu, H. T. Pedro, A. Kaur, J. Kleissl, and C. F. Coimbra, "Net load forecasts for solar-integrated operational grid feeders," *Solar Energy*, vol. 158, pp. 236–246, 2017.

[4] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 188–201, 2019.

[5] H. Shaker, H. Chitsaz, H. Zareipour, and D. Wood, "On comparison of two strategies in net demand forecasting using wavelet neural network," in *2014 North American Power Symposium (NAPS)*, 2014, pp. 1–6.

[6] D. van der Meer, J. Munkhammar, and J. Widén, "Probabilistic forecasting of solar power, electricity consumption and net load: Investigating the effect of seasons, aggregation and penetration on prediction intervals," *Solar Energy*, vol. 171, pp. 397–413, sep 2018.

[7] M. Alipour, J. Aghaei, M. Norouzi, T. Niknam, S. Hashemi, and M. Lehtonen, "A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration," *Energy*, vol. 205, pp. 118–106, 2020.

[8] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li, and Q. Xia, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter pv," *IEEE Trans. on Power Sys.*, vol. 33, no. 3, pp. 3255–3264, 2018.

[9] K. Pu and Y. Zhao, "An unsupervised similarity-based method for estimating behind-the-meter solar generation," in *Proceedings of the IEEE Conference on Innovative Smart Grid Technologies (ISGT)*, 2023.

[10] S. Tiwari, R. Sabzehgar, and M. Rasouli, "Short term solar irradiance forecast using numerical weather prediction (nwp) with gradient boost regression," in *9th IEEE International Symposium on Power Electronics for Distributed Generation Systems*. IEEE, 2018, pp. 1–8.

[11] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[13] A. D. Gordon, *Classification*. CRC Press, 1999.

[14] C. Holcomb, "Pecan Street Inc.: A Test-bed for NILM," in *International Workshop on Non-Intrusive Load Monitoring*, 2012.

[15] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, 2018.