Scalable Hybrid Classification-Regression Solution for High-Frequency Nonintrusive Load Monitoring

Preprint

Govind Saraswat,[1] Blake Lundstrom,[2] and Murti V. Salapaka[3]

1 National Renewable Energy Laboratory
2 Enphase
3 University of Minnesota

To be presented at the 2023 IEEE Conference on Innovative Smart Grid Technologies North America (ISGT NA)
Washington, D.C.
January 16–19, 2023

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-76389
November 2022

Scalable Hybrid Classification-Regression Solution for High-Frequency Nonintrusive Load Monitoring

Preprint

Govind Saraswat,[1] Blake Lundstrom,[2] and Murti V. Salapaka[3]

1 National Renewable Energy Laboratory
2 Enphase
3 University of Minnesota

Suggested Citation
Saraswat, Govind, Blake Lundstrom, and Murti V. Salapaka. 2022. *Scalable Hybrid Classification-Regression Solution for High-Frequency Nonintrusive Load Monitoring: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-76389. https://www.nrel.gov/docs/fy23osti/76389.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-76389
November 2022

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by Advanced Research Projects Agency-Energy (ARPA-E). The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Scalable Hybrid Classification-Regression Solution for High-Frequency Nonintrusive Load Monitoring

Govind Saraswat[1], Blake Lundstrom[2], and Murti V Salapaka[3]

***Abstract*—Residential buildings with the ability to monitor and control their net-load (sum of load and generation) can provide valuable flexibility to power grid operators. We present a novel multiclass nonintrusive load monitoring (NILM) approach that enables effective net-load monitoring capabilities at high-frequency with minimal additional equipment and cost. The proposed machine learning based solution provides accurate multiclass state predictions while operating at a faster timescale (able to provide a prediction for each 60-Hz ac cycle used in US power grid) without relying on event-detection techniques. We also introduce an innovative hybrid classification-regression method that allows for the prediction of not only load on/off states but also individual load operating power levels. A test bed with eight residential appliances is used for validating the NILM approach. Results show that the overall method has high accuracy, good scaling and generalization properties.**

***Index Terms*—Nonintrusive load monitoring (NILM), multiclass classification, regression, power prediction, feature extraction, smart buildings, grid-interactive, smart grid.**

I. Introduction

With increasing adoption of distributed and renewable energy resources, and reduction of traditional synchronous generation, the electric grid is at a unique junction with both immense opportunities as well as challenges. Further, the changes in generation profiles, along with evolving load types and capabilities, are making behind-the-meter (BTM) net load (i.e., net generation and load) an increasingly important component of system operations. BTM visibility is integral to the emerging concepts and operating platforms for power systems. Smart buildings have the capability to regulate their net load in a way that can provide extra flexibility to grid operators[1] as well as maximize energy cost savings and preferences for their customers. 38.5% of the total electrical energy produced [2] in United States is consumed by residential buildings which are predominantly not equipped with modern sensors and controllers. Thus, if such residential buildings can be transformed into smart buildings, significant resilience can be added to the distribution grid. The simplest approach to achieve such visibility is having circuit-level net-load metering, but this may not be economically viable nor practical due to limitations and cost of communication infrastructure as well as privacy concerns. Nonintrusive load monitoring (NILM) provides a cost effective solution to enable smart metering of residential buildings. It uses a single current measurement at the building's point of common coupling (PCC) with the grid and extracts power information of the individual BTM devices inside the building. Such power information may include on/off status and/or power consumption.

Existing work in NILM can be categorized into two categories based on the type of measurement inputs used. NILM approaches belonging to the first category use steady-state measurement quantities, such as active and reactive power or root mean square (RMS) current, on a macro timescale (generally $>$ 1 minute). Steady-state power measurements have been used with a variety of contemporary machine learning methods, including multilabel k-nearest neighbor [3], binary relevance [3], hidden Markov models [3], [4], and long short-term memory neural networks [5]. In the second category, transient features derived from micro-timescale data (generally $<$ 1 second) are used. Here, micro-timescale data is employed to infer features based on a frequency-domain transformation [6], instantaneous wave shape [7], or wavelet transformations [8]. Some approaches [9] fit into both categories, using micro-timescale data to detect events and macro-timescale data to capture steady-state data throughout the event and build a library of events that can be used later for prediction.

NILM methods that use steady-state measurements can easily misclassify two different combination of loads which have near identical steady-state measurements but differing instantaneous waveforms as shown in [10]. Most of the existing NILM approaches do not provide state predictions faster than 1 s, which is too slow to enable a residential building to provide grid services (e.g., fast frequency response using building net load [11]). In contrast, the approach developed in this article can accurately identify changes in on/off state and operating power level for each individual load on a fast timescale (e.g., each AC line cycle), throughout a transient event, and without reliance on event detection techniques.

The majority of NILM literature has focused on predicting the on/off state of connected loads via classification. Some existing work (e.g., [12], [5]) also considered regression-based approaches, but they rely on availability of long sequence (i.e., multiple hours) of power data. To the best of our knowledge, there is no method currently capable of load power disaggregation at the timescale of <1-s grid ancillary services.

Our initial contributions related to high-speed classification for NILM appear in [10]. That work presented a method for accurate, high-speed (>60-Hz) prediction of the multiload on/off

[1]National Renewable Energy Laboratory,Golden, CO, USA; [2]Enphase, Austin, TX, USA; [3]University of Minnesota, Minneapolis, MN, USA

The authors acknowledge the Advanced Research Projects Agency-Energy (ARPA-E) for supporting this research through the project titled "Rapidly Viable Sustained Grid" via grant no. DE-AR0001016. This work was authored in part by the National Renewable Energy Laboratory, managed and operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes

state of a four-load residential building experiment configuration. In this work, we extend this high-speed classification method to improve scalability, and we apply the improved classification method to a larger eight-load residential configuration. In addition, we introduce a novel hybrid classification-regression model that allows for the prediction of load RMS current consumption levels via regression in addition to the load on/off states via classification while still operating at a high-speed timescale with predictions for each 60-Hz cycle. The hybrid classification-regression approach is validated using an experiment configuration including eight residential building appliances, and it is shown to give high-accuracy RMS current and multiload on/off state predictions while allowing for good generalization. A detailed version of this paper is provided in [13]

II. Experimental Setup and Approach

A. Problem Formulation

A total of N_L loads are connected to the main load center panel of a residential building, and each consumes an instantaneous current $I_i(t) = q_i(t)\bar{I}_i(t)$, where $\bar{I}_i(t)$ is the load signature, $q_i = \{0, 1\}$ is the on/off status (where $q_i = 1$ denotes that the load is on), and I_i is the resulting current consumption of the i^{th} load. The total current consumption of the residential building is $I_{TOT}(t) = \sum_{i=1}^{N_L} q_i(t)\bar{I}_i(t)$. System load on/off state for the building is $y_c = [q_1, q_2, \ldots, q_i, \ldots, q_{N_L}]$. With N_L loads, there are 2^{N_L} possible combinations of y_c. The first objective is to train a classifier that uses an input vector, X, of features calculated for each 60-Hz ac cycle observation of $I_{TOT}(t)$ to predict y_c, representing the building's complete load on/off state. The second objective is to build a regression model that takes an input feature vector X_r (potentially including the result from the classifier) and predicts the RMS current consumption of each load, $Y_{reg} = [\widetilde{I}_1, \widetilde{I}_2, \ldots, \widetilde{I}_i, \ldots, \widetilde{I}_{N_L}]$. Both the classifier and the regression model should provide a prediction for every 60-Hz ac cycle of $I_{TOT}(t)$. For the application of fast-frequency smart building control, the classifier and regression model should predict the state and operating level of all constituent loads within 10 AC cycles or $\tau = 10/60 = 0.167$ s. To this end, a two-step hybrid approach (see Fig. 1) is adopted. The first step uses a classifier block to predict the on/off state of each load. The second step applies Deep Neural Network (DNN) models as a regression block to predict the current consumption of each load i using the output of the classifier block as well as the subset of feature inputs given to the classifier block.

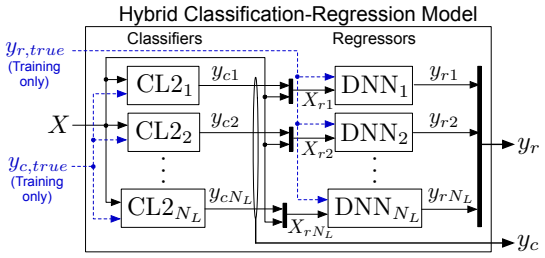

Fig. 1. Hybrid classification-regression modeling approach. It can be seen that output of the classifiers are used by the regressors to predict individual load RMS current.

B. Experiment Configuration

A residential-scale demonstration using eight household appliances in the Energy Systems Integration Facility at the National Renewable Energy Laboratory was completed. The full experimental configuration, shown in Fig. 2, includes eight residential appliances (details in [13]). One voltage sensor and nine current sensors (as depicted in Fig. 2) are used to sample instantaneous measurements with a 200kHz bandwidth. Seven independent 200-s data sets are collected. Loads were randomly perturbed while collecting these data sets. For example, fridge door was opened and closed, oven door was opened and closed, and so on. The operating conditions of all loads for different data sets are described in [13]. Thus, a variety of operating conditions of the loads were captured.

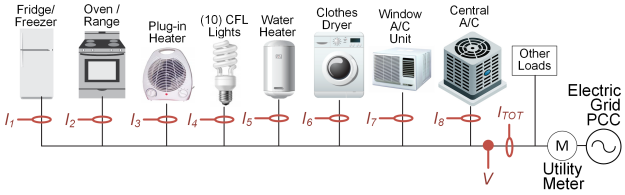

Fig. 2. Experimental configuration with eight residential appliance loads. I_{TOT} measures the total current flowing through the eight loads.

C. Input Data Processing

High-speed instantaneous voltage and current waveforms are the input to this approach (see Fig.4 in [13]). The current waveforms are divided at each 60-Hz ac cycle using zero-crossing detection, then each cycle of the aggregate current measurement (I_{TOT}) becomes an observation. Feature vector X is derived from each observation. Voltage measurement data are not strictly required, but if available, they can be used to provide more accurate detection of zero crossings. Current measurements from each load are used to create the label for each observation. For further increasing the richness of training data, we employ a novel strategy for creating a synthetic data set from the original data that expands each observation into multiple observations by using the superposition principle of electric current. We then employ a reduction strategy to remove samples that are very similar to each other. These two strategies are described in detail as follows:

Synthetic data set: Each observation, j, in the original data set has a total current measurement, I_{TOT}, based on $1 \leq N_{L,j} \leq 8$ superimposed currents. $N_{L,j}$ is determined by thresholding all I_i and counting the number of nonzero I_i. Then, the $N_{L,j}$ nonzero currents can be used to generate $2^{N_{L,j}}$ unique observations by superimposing the corresponding I_i for each combination together and calculating the features.

Data set reduction: Within each synthetic data set, all the observations are first sorted with respect to features, the distance (in terms of the L_1-norm) between successive observations is determined, and new observations are discarded if they are not significantly different from the previous observation. With this reduced data set, computational complexity and memory requirements for training the models are significantly reduced.

D. Labeling (Training Phase)

For classification, observations are generally categorized into multiple classes. When classifying between more than two classes, multiclass classfication is used. Each observation cycle

of I_{TOT} is labeled by examining the instantaneous current data at each circuit, using level detection to derive the on/off state of that circuit, and finally building the multiclass label. This multiclass label could be a multilabel vector or a single integer label with the class encoded using a binary encoding scheme (e.g., Label 12 corresponds to [0,0,0,0,1,1,0,0]), referred to as the label power-set method. For regression, the RMS current of each individual load is used as a label for the corresponding regression model.

Two training methodologies are employed. The first training strategy (strategy 1) follows the traditional approach of randomly shuffling the full data set (all seven independent data sets are used) and then splitting the data into a training set (80%) and test set (20%). The second approach (strategy 2) is to use six of the seven independent data sets for training and testing on the remaining data set. This leads to seven distinct runs. The second approach emphasizes the generalization capabilities of the classifier. As detailed in [13], data sets are quite different because they capture different running conditions of each load.

III. High-frequency, Multiclass Classification and Power Prediction of Residential Appliance Loads

In this work, we use a random forest classifier (RFC) [14] for classification. As any number and combination of loads can be turned on/off at any given time, multiclass classification is required for this problem. The 'label power-set method', which takes each combination of possible load states as one class, was initially used when considering the four-load test configuration [10]. This method (Classifier 1) considers the correlation of different labels and can provide accurate predictions when the number of loads is small; however, the total number of classes in this method increases exponentially with the number of loads. Thus, when applying the approach to the larger eight-load configuration, the label power-set method does not scale well because the class size explodes to 255. This leads to the resultant model to be extremely complex and requires large memory ($>$ 30 GB). To address this, we present a multi-model approach (Classifier 2) wherein we train simple binary classifiers to detect the on/off state of a single load. This approach trains N_L RFC models for N_L loads (see left side of Fig. 1). Here, model CL2$_i$ corresponds to load i. The set of all N_L models is termed as Classifier 2. The output of the classifier is a vector comprising individual predictions of each model. Let y_ci be the prediction of model i. Then the output of Classifier 2 is $y_c = [y_{c1}, \ldots, y_{ci}, \ldots, y_{cN_L}]$. Here, the total memory requirement is ≈ 1 GB for all 8 classifiers. This leads to much better scaling capabilities and easier implementation for parallel processing of multiple load classifications. Further, both the classifiers take similar run time with serial implementation, thus Classifier 2 has a clear advantage as it can be easily implemented in parallel, significantly reducing the total run time.

Deep neural networks (DNNs) are used in this work to predict the RMS currents of each load. A detailed description of DNNs can be found in [14]. We train a single neural network for each load; thus, the number of neurons in the outer layer is one.

A. Model Training and Tuning

Here, we use the reduced synthetic data from the eight-load experiment configuration for training, as described in Section II. To find the best classifier for each load, multiple RFCs using different subsets of the features described in [13] are considered. The final classifier model chosen is an RFC ensemble using bootstrapping and including 200 decision tree classifiers, each with a max depth of 35 and using entropy (information gain) as the split criterion. Further details on RFCs hyperparameters and testing methodologies can be found in [14]. Here, the RFC was implemented in Python using the scikit-learn package [15].

Similar to the tuning of RFCs, multiple DNNs are considered in a search for a model that can accurately predict the RMS current. Details of hyperparameters of the selected DNN can be found in [13]. In contrast with RFCs, neural networks are very susceptible to the scaling of inputs and outputs. Thus, we apply a 'power transformation' to the feature matrix X to make the input data more Gaussian. Specifically, we use a 'Yeo-Johnson' transform [16], which supports both positive or negative data. For scaling the output y, we employ standard scaling, which removes the mean and scales the data to unit variance. The above methodology is implemented in Python using the keras package [17].

IV. Results

Here, we present the results for both classification and regression along with the hybrid model where results from the classifier are used by the regression model.

A. Classification

As mentioned in Section III, we implemented two classifiers. We present prediction accuracy of both the classifiers when using strategy 1 (80-20 split) for training in Fig. 3(a). Accuracy is defined as the ratio of correctly predicted observations to the total number of observations. The median prediction accuracy (across 7 runs, each run has random selection of 80% training and 20% test data) for each load with the 1st and 3rd quantiles as error bars is plotted in Fig. 3(a). It can be clearly seen that the accuracy of both classifiers is within $\pm 0.75\%$ for each load. Overall average accuracy of Classifier 1 (power set) is 98.95% and Classifier 2 (binary) is 98.84%.

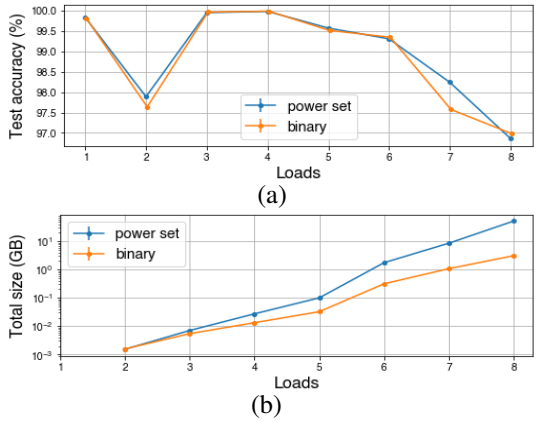

Fig. 3. Test accuracy (a) and total size (b) of Classifier 1 (power set) and Classifier 2 (binary) with number of load when using strategy 1. (b)

The total memory requirement for each classifier as a function of the number of loads is plotted in Fig. 3(b) on

a log scale. As the number of loads increases, the difference in the size of the two classifiers also increases. For an 8 load configuration, there is more than an order of magnitude difference between the memory requirements for two classifiers. Thus, as described in Section III, Classifier 2 (binary) provides similar performance in accuracy when compared to Classifier 1, while being far less memory intensive and more conducive to parallel implementation.

We next present the results for these classifiers when strategy 2 (testing on independent unseen dataset) for training is used. Here, for each run, six of the seven independent data sets are used for training and testing is completed using the remaining data set. The median prediction accuracy (across 7 runs) for each load with the 1st and 3rd quantiles as error bars is plotted in Fig. 4. Here, clearly Classifier 2 (binary) outperforms Classifier for almost all loads. The overall average test accuracy for Classifier 1 (power set) is 92.54% and Classifier 2 (binary) is 93.07%. Both of these classifiers show good generalization capabilities as they have good overall accuracy (93%) when testing is performed on an independent (not seen during the training phase) data set .

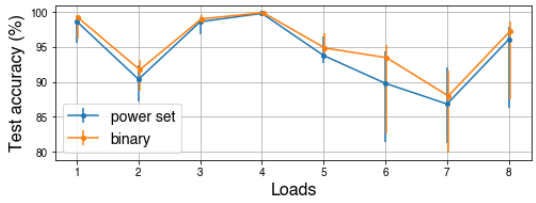

Fig. 4. Test accuracy of Classifier 1 (power set) and Classifier 2 (binary) by load when using strategy 2

B. Transient Performance

Fig. 5 (a) shows I_{TOT} for the entire length of data set 3; it can be seen that significant aggregate variation occurs throughout as the eight loads turn ON and OFF. Fig. 5 (b), (c), (d) and (e) present the classifier (trained using strategy 2 with only 6 of the 7 dataset used) predictions for load 3, 4, 5 and 6, respectively, at multiple demonstrative time periods within data set 3.

The NILM method presented here is able to accurately predict the true state of the loads throughout most transient periods. Fig. 5 (c) shows an example of one such transient period (near t=37 s) where the load 4 (compact fluorescent lights) turn-on event transient lasts nearly 15 ac cycles. During this event, the classifier correctly predicts most cycles, despite the cycles having a variety of wave shapes and magnitudes. This demonstrates the high-accuracy and high-frequency performance of the approach and its advantage over many existing NILM approaches based on event detection, which would not predict a state change until after the entire 15 ac cycle event signature was detected.

C. Regression and Hybrid model

Similar to the classification, we present the results of the regression model on both the training strategies. To evaluate the regression model, we present two sets of results. For the first set of results, we assume feature set X has an accurate load on/off state input, to independently test the accuracy of the

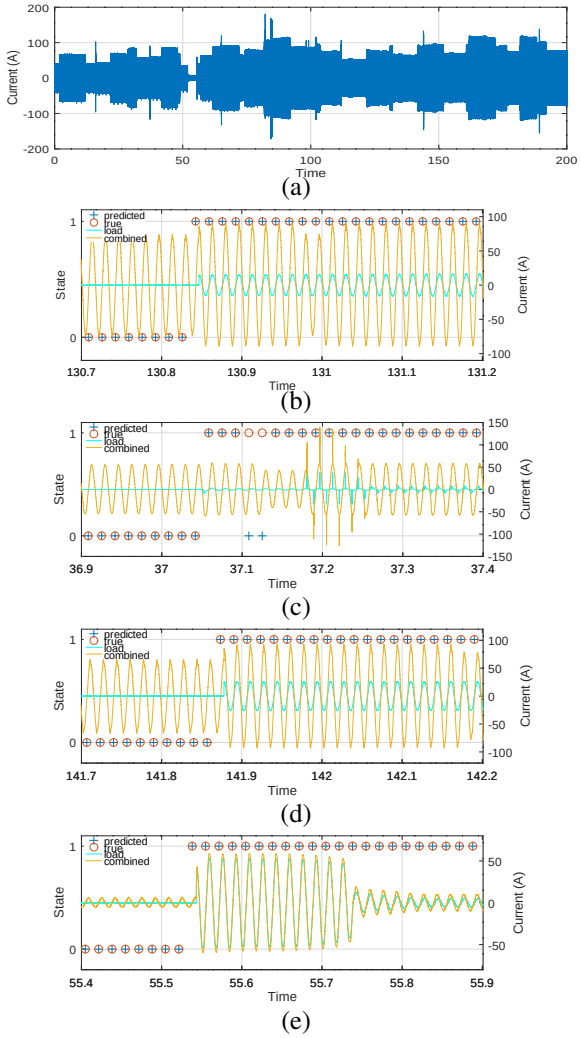

Fig. 5. (a) I_{TOT} for the entirety of data set 3, showing significant variation in aggregate current). Magnified views for load 3, 4, 5 and 6 switching ON events are shown in (b), (c), (d), and (e), respectively. The yellow traces in each plot show I_{TOT} (available to the model), while cyan traces show the individual load current (not available to the model). Red circles denote the true ON/OFF state while blue crosses are the predicted states.

regression (DNN) model, whereas the second set of results is for the hybrid model, where X contains predicted state values from the classifier. To quantify the accuracy of the regression model, a test sample is deemed predicted true if the predicted RMS value from the DNN model is within $\pm 10\%$ of the true RMS value. In the case when a load is OFF, a test sample is deemed predicted true if the predicted RMS value is within ± 0.2A.

The overall test accuracy of the DNN and hybrid models for strategy 1 (80-20 split) is 98.4% and 97.8%, respectively, when taken as an average over 10 runs across all loads. The median accuracy for each load is shown in Fig. 6 (a), with the first and third quantiles plotted as error bars. Similar to Classifier 2 (binary), both the DNN and hybrid models do very well when tested on data for which similar samples have been seen before. With the second approach, when a completely new data set is used for testing, the overall test accuracy of the DNN and hybrid model decreases to 94.8% and 91.8%, respectively, when averaged over 7 separate runs across all

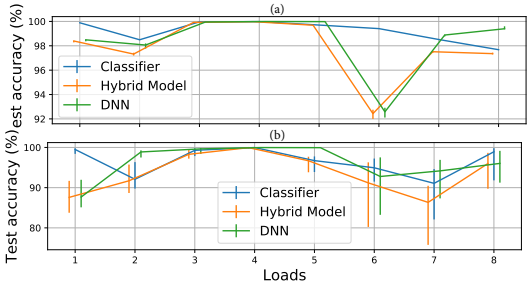

Fig. 6. Median test accuracy with error bars for Classifier 2 (binary), DNN, and the hybrid model evaluated using the (a) first (80/20 splits) and (b) second (independent data set for testing) testing approaches. The hybrid model and DNN curves are shifted slightly left and right, respectively, for visibility.

loads. The median accuracy for each load is shown in Fig. 6 (b), with the first and third quantiles plotted as error bars. Thus, even for the case when regression model does not have the knowledge of exact operating state of loads, accuracy is above 90%.

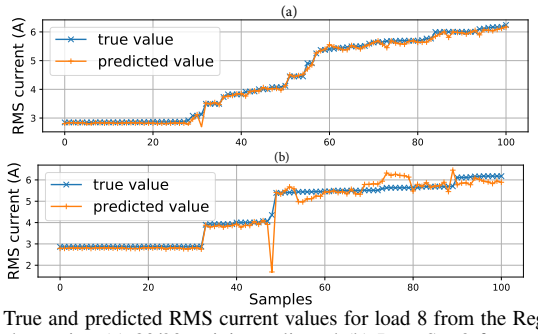

Fig. 7. True and predicted RMS current values for load 8 from the Regression model when using (a) 80/20 training split and (b) Data Set 3 for testing

Fig. 7 (a) shows the true and predicted RMS currents from the hybrid model for load 8 when using strategy 1 (80-20 split). These are 100 random test samples (sorted by RMS) from the test data set. Here, we can see the model predicting the RMS current very accurately across different operating regimes (different power consumption) of this load. Fig. 7 (b) shows the true and predicted RMS current (by the hybrid model) of load 8 using Data Set 3 and strategy 2 (testing on independent unseen dataset) for testing. These are 100 random test samples (sorted by RMS) from the test data set. When compared to Fig. 7 (a), it is clear that for this strategy, the model performs very well when predicting the RMS current for the main operating conditions while being fairly accurate even for the intermediate operating conditions. Similar to classifiers, the DNN model generalizes reasonably well when tested on completely independent data sets.

For this proof-of-concept study, the NILM approach is not implemented in real time, but the total processing time—including zero-crossing detection, feature extraction, and classifier and regression prediction—is measured (averaged over 80,000+ cycles) to be 10.42 ms, which is fast enough to support high-frequency load disaggregation and certainly well within the desired $\tau = 167$ ms response time. This time includes 10.37 ms for feature extraction, 0.016 ms for classification, and 0.036 ms for DNN model prediction. This is measured on a laptop with an Intel i7 processor and Nvidia GeForce GTX 1050 Ti GPU and does not consider any delays in data acquisition. With a response time more than 10 times faster than the desired response time, however, there is sufficient room for data acquisition delays and/or decreased performance if implemented on a slower computational platform (like a RaspberryPi) while still meeting the response time requirement.

V. Conclusion

This paper presented a novel NILM approach which combined classification and regression to predict power consumption along with on/off state of BTM loads thus providing full visibility. The approach was shown to have high accuracy, good scaling and generalization properties, when tested on a test bed consisting of eight residential appliances. Even on a generic PC, approach was fast enough to support building grid-interactive control at fast timescales (e.g., within 10 ac cycles) relevant to the provision of grid frequency support services.

References

[1] S. Patel, B. Lundstrom, G. Saraswat, and M. V. Salapaka, "Distributed power apportioning with early dispatch for ancillary services in renewable grids," *arXiv preprint arXiv:2007.11715*, 2020.

[2] U.S. EIA, "Electricity Explained: Use of Electricity," 2018. [Online]. Available: https://www.eia.gov/energyexplained

[3] K. Basu, V. Debusschere, S. Bacha, U. Maulik, and S. Bondyopadhyay, "Nonintrusive Load Monitoring: A Temporal Multilabel Classification Approach," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 1, pp. 262–270, Feb. 2015.

[4] H. He, Z. Liu, R. Jiao, and G. Yan, "A Novel Nonintrusive Load Monitoring Approach based on Linear-Chain Conditional Random Fields," *Energies*, vol. 12, no. 9, p. 1797, May 2019.

[5] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015, pp. 55–64.

[6] R. A. S. Fernandes, I. N. d. Silva, and M. Oleskovicz, "Load Profile Identification Interface for Consumer Online Monitoring Purposes in Smart Grids," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1507–1517, Aug. 2013.

[7] T. Hassan, F. Javed, and N. Arshad, "An Empirical Investigation of V-I Trajectory based Load Signatures for Non-Intrusive Load Monitoring," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, p. 10, 2014.

[8] J. M. Gillis, S. M. Alshareef, and W. G. Morsi, "Nonintrusive Load Monitoring Using Wavelet Design and Machine Learning," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 320–328, Jan. 2016.

[9] M. Nardello, M. Rossi, and D. Brunelli, "An innovative cost-effective smart meter with embedded non intrusive load monitoring," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Sep. 2017, pp. 1–6.

[10] B. Lundstrom, G. Saraswat, and M. V. Salapaka, "High-frequency, multiclass nonintrusive load monitoring for grid-interactive residential buildings," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2020, pp. 1–5.

[11] B. Lundstrom, S. Patel, S. Attree, and M. V. Salapaka, "Fast primary frequency response using coordinated der and flexible loads: Framework and residential-scale demonstration," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, 2018, pp. 1–5.

[12] M. Kaselimi, N. Doulamis, A. Doulamis, A. Voulodimos, and E. Protopapadakis, "Bayesian-optimized bidirectional lstm regression model for non-intrusive load monitoring," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 2747–2751.

[13] G. Saraswat, B. Lundstrom, and M. V. Salapaka, "Scalable hybrid classification-regression solution for high-frequency nonintrusive load monitoring," 2022.

[14] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2019.

[15] F. e. a. Pedregosa, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[16] I.-K. Yeo and R. A. Johnson, "A new family of power transformations to improve normality or symmetry," *Biometrika*, vol. 87, no. 4, pp. 954–959, 2000.

[17] F. Chollet *et al.*, "Keras," https://keras.io, 2015.