DWT-Based Aggregated Load Modeling and Evaluation for Quasi-Static Time-Series Simulation on Distribution Feeders

Preprint

Xiangqi Zhu and Barry Mather
National Renewable Energy Laboratory

Presented at the IEEE Power and Energy Society General Meeting
Portland, Oregon
August 5–9, 2018

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.

Conference Paper
NREL/CP-5D00-70975
August 2018

Contract No. DE-AC36-08GO28308

NOTICE

This work was authored 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 the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE), Solar Energy Technologies Office. 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.

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

DWT-Based Aggregated Load Modeling and Evaluation for Quasi-Static Time-Series Simulation on Distribution Feeders

Xiangqi Zhu and Barry Mather
Power Systems Engineering Center
National Renewable Energy Laboratory
Golden, CO, USA
xiangqi.zhu@nrel.gov, barry.mather@nrel.gov

***Abstract*—This paper presents an extension of a previously reported discrete wavelet transform-based load modeling methodology that targets modeling time-series load profiles to enable more effective and realistic quasi-static time-series (QSTS) simulations of distribution feeders. The time-series load model is composed of two major parts: 1) the low-resolution field-measured load data, which usually have resolutions of 30 or 15 min; and 2) the high-resolution load variability model data extracted from the established variability database, which can have resolutions up to 1 s. A load aggregation methodology is developed to aggregate the load profiles so that they can be effectively modeled at different transformer ratings. Validation, evaluation, and analysis of the developed load modeling approach has been performed on the IEEE 123-node test feeder and an actual utility feeder from California. The analysis completed on the two feeders has demonstrated the effectiveness and revealed the value of the developed model for distribution feeder QSTS simulations at high temporal resolutions.**

***Keywords*—Load modeling; quasi-static time-series; QSTS; high-resolution load models; distribution system simulation**

I. Introduction

Load modeling has always played a critical role in distribution system simulation, but in recent years load models that can enable quasi-static time-series (QSTS) simulations are in high demand. This is because utilities and grid operators desire effective load models, at high temporal resolutions, to develop and test new control algorithms and applications that are used to inform the challenges brought by the increasing integration of rooftop photovoltaic (PV) systems and other distributed energy resources (DERs) onto the electric power grid [1]–[-2]. Traditional load modeling and load modeling aggregation methods are discussed in [3], but such methods typically characterize loads based on assumed load classes and dominant behaviors. This work focuses on an empirical, data-driven method that leverages utility data.

Previously, when performing distribution system simulations or conducting new control algorithm testing, all the sub-node profiles were typically populated with a scaled version of the substation load because only substation load profiles were available; however, actual profiles of individual distribution transformer loads are usually "spikier" and contain more variability. Nodes with higher aggregations of load would have smoother profiles. Therefore, representing all the sub-node load profiles with only one substation level profile is not able to capture the variability and diversity of the load profiles along a feeder. Now, with the advent of advanced metering infrastructure (AMI), data with increased spatial resolution are often available, but the data are still of relatively low temporal resolution (e.g., 30 min), which is not sufficient to capture the variability of the feeder loads when performing the feeder simulation or conducting algorithm testing at high temporal resolutions. Such input data, when used in a study, would not be able to accurately reflect the parameter results, including tap changes, voltage ramping, line loss, etc. [4]. In addition, even if the high-resolution data could be obtained, the original user load data could not be used to develop algorithms and test the feeder because the pattern of the original load data would reveal the living/energy usage pattern of the customers, thus causing privacy concerns.

To tackle the challenge of obtaining appropriate load models for conducting QSTS simulations, we developed a load modeling approach to build load profiles that could not only represent the variability of the realistic sub-node load but also effectively protect customer privacy. The developed load model consists of two major parts: 1) one is the low-resolution field-measured load data (i.e., AMI-type data), and 2) the other is the high-resolution load variability model data extracted from the established variability database, which has a resolution up to 1 s. A load aggregation methodology has also been developed to aggregate the load profiles so that load profiles at different transformer ratings could be modeled. The developed load model has been tested on the standardized IEEE 123-node test feeder and an actual utility feeder from California. The testing results on these two feeders demonstrated the effectiveness of the proposed load model.

This paper is organized as follows: Section II introduces the load modeling methodology. Section III presents the feeder testing, including the test on the IEEE 123-node test feeder and an actual utility feeder. Section IV concludes the paper and discusses future work.

II. Load Modeling Methodology

This section introduces the load modeling approach, the load database used to build the variability library, and the methodology of establishing the library. In addition, this section presents the load aggregation approach that is used to develop load models for different load levels.

The database we used in this paper is for residential load profiles because the data obtained from a utility partner primarily consisted of residential transformer data. A database of commercial or industrial loads could be established if such data were available. The modeling methodology we developed here could be applied to the commercial and industrial loads as well.

A. Modeling Approach

A discrete wavelet transform (DWT)-based variability extraction methodology has been developed to establish a variability library using previously measured high-resolution transformer data [5]. As shown in Fig. 1, base low-resolution load profiles of hourly or 30 min-resolution are used along with high-resolution data extracted from a developed variability library. In this way, load variability characteristics for all loads can be modeled while protecting the privacy of customers.

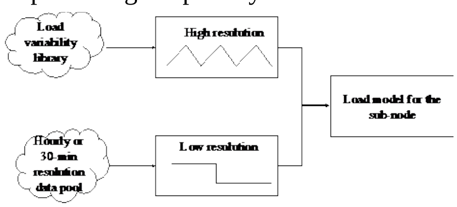


Fig. 1. Load modeling approach

B. Load Database

Currently, 1-s resolution data for 3 years of a transformer with approximately 10–12 houses connected are used to establish the database on which the variability library is built. The data have been divided into four seasons, and the variability library is built for each season. A sample synthetic profile of a day in fall is shown in Fig. 2. The field-measured data in this database also contain the rooftop PV data (i.e., generation), therefore the profile shown in Fig. 2 has some negative data points. Because the measured area usually has a clear sky and solar is present only during daytime, it is reasonable to assume that the variability in this database is mostly caused by load changes instead of the solar intermittency for short time frames.

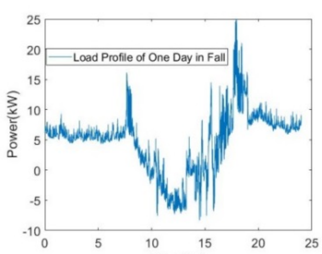


Fig. 2. One-day sample from fall in realistic load database

C. Variability Library

The variability library is built for each season separately. Multi-resolution DWTs are used to decompose the load data into detailed coefficients and approximate coefficients at different levels [6]–[7]. An example of three-level decompostion is shown in Fig. 3; input data *s* will be decomposed into detailed and approximate coefficients. The values of a_1, a_2, and a_3 represent approximate coefficients, which are obtained by low-pass filtering and contain the low-frequency signal. The values of d_1, d_2, and d_3 represent detailed coeffcients, which are obtained by high-pass filtering and contain the high-frequency signal. As shown in (1), the summation of all the detailed coeffcicents and the approximate coefficient from the last level of decompostion will reconstruct the origial signal.

To extract the load variability characteristics in the resolution range from 1-s to 30-min, 12-level DWTs have been performed for the daily load profiles in the database. The summation of the first 12 detailed coefficients define the load variability model, and the 12th approximate coefficient will be kept as the baseload. Each load profile in the database will have one corresponding load variabilty model and a baseload term.

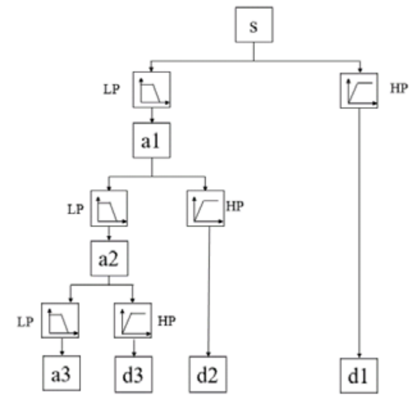


Fig. 3. Wavelet decompostion illustration

$$s = d1 + d2 + d3 + a3 \quad (1)$$

After obtaining the variability models and baseloads, the next step is to construct the variability library. Using the fall season as an example, for all the 273 1-s field-measured daily load profiles in fall, the variability of each load profile will be quantified by calculating the ramp profile. As shown in (2), the ramp is calcuated as the difference between the load magnitude at the next time step (t_1) and the current time step (t_0). In this way, the 273 ramp profiles for fall are obtained:

$$Ramp = P(t_1) - P(t_0) \quad (2)$$

Then the correlation between one ramp profile and the other 272 load profiles is calcuated. When calculating the correlation coefficent of two ramp profiles *A* and *B*, each profile will be divided into *N* bins, and the number of data points in each bin are recorded as A_i and B_i. As shown in (3), the correlation coefficient of the two ramp profiles is calculated using the Pearson correlation coefficient [8], in which μ_A and $\sigma_A$represent the mean and standard deviation of *A*, respectively; and μ_B and $\sigma_B$represents the mean and standard deviation of *B*, respectively.

$$\rho(A, B) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{A_i-\mu_A}{\sigma_A}\right)\left(\frac{B_i-\mu_B}{\sigma_B}\right) \quad (3)$$

If the correlation coefficient of ramp profiles *A* and *B* is greater than 0.99, then these two profiles are recorded as correlated. The ramp profiles are ranked by the correlation record. If one profile has the highest number of correlated profiles, then this profile is ranked top of the list. The ramp profiles that are correlated with at least half the total number of profiles are ranked on the list.

The load variability models associated with the ramp profiles on the list are recognized as variability classes and are included in the variability library for the fall season. The varaibility library of the other three seasons are established in the same way.

D. Load Aggregation Approach

Different sub-nodes should have different load profiles. If we have the peak load for each node, then a load aggregation approach could be used to model the load profiles at various load levels [9].

As shown in the flowchart in Fig. 4, first, the maximum noncoincident load is calculated by multiplying the peak load with the diversity factor (*DF*). Second, the number of loads *N* is calculated by dividing the maximum noncoincident load by the average peak load. The number *N* will be rounded to the closest integer. Third, if *N* is greater than 1, *N* load profiles will be randomly selected from the realistic load database. Then the selected profiles will be summed to get the total load profile. The difference between the peak total load and the given peak load will be compared with the error tolerance. The third step is repeated until the difference is smaller than the tolerance. After the tolerance is satisfied, the corresponding load model is developed by adding the baseloads of the selected realistic loads and the variability models extracted from the variability library. Note that if *N* is smaller than 1, the load profile with the peak load closest to the node peak is directly determined from the load database.

DF is calculated as shown in (4), where $P_1 \sim P_n$ represents the profiles in the load database, and *n* is the total number of profiles:

$$DF = \frac{maximum\ noncoincident\ demand}{maximum\ diversified\ demand} \quad (4)$$

$$= \frac{\max(P_1)+\max(P_2)+\cdots \max(P_n)}{\max(P_1+P_2+\cdots P_n)}$$

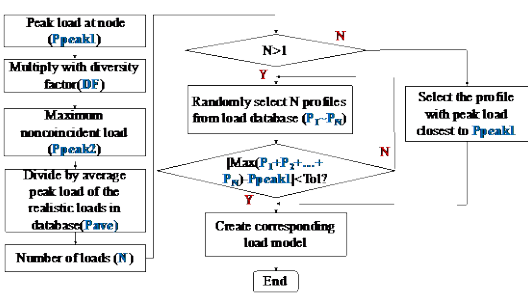


Fig. 4. Load aggreagation flowchart

In this work, because we have only the node peak load information of the feeder, we use the above practical and simplified load aggregation approach to model the load at different levels. We are also developing an advanced load aggregation methodology that could capture salient points of the load profile so that the key characteristics, such as PV impact, could be modeled. We will present this methodology in a follow-up paper.

III. Feeder Testing

The proposed load modeling approach has been tested on two feeders: 1) the standardized IEEE 123-node test feeder [10], and 2) a realistic utility feeder from California. Realistic load data and the modeled load data built from the variability database and baseload pool were used to perform tests on the feeders. A series of evaluation criteria have been selected to evaluate the effectiveness of the modeled load profiles on feeder testing.

A. Evaluation Criteria

As shown in Table I, four metrics have been used to evaluate the performance of the developed load models on feeder testing. Maximum and minimum voltage are compared between the testing results of the original load and modeled load to check if the voltage range remains the same if performing the testing using the modeled load. The number of regulator operations and the voltage ramp distribution are also compared to check if the modeled load effectively reflects the variability characteristics of the original load. The error rate for the comparison of voltage regulator operations and the voltage ramp is calculated in (5) and (6), respectively.

TABLE I. Performance Evaluation Metrics

1	Voltage ramp distribution	3	Minimum voltage
2	Maximum voltage	4	Number of regulator tap change

$$Error\ Rate = \frac{Regulator\ Moves\ (orignal) - Regulator\ Moves\ (Modeled)}{Regulator\ Moves\ (orignal)} \quad (5)$$

$$Voltage\ Ramp = V(t1) - V(t0) \quad (6)$$

B. Testing on IEEE 123-Node Test Feeder

The topology of the IEEE 123-node test feeder is shown in Fig. 5. A total of 91 nodes have load. Four voltage regulators are installed along the feeder.

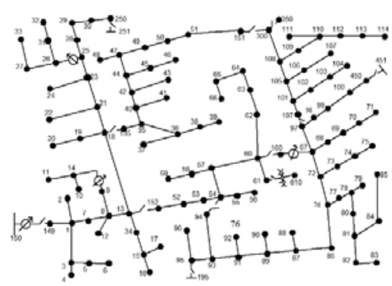
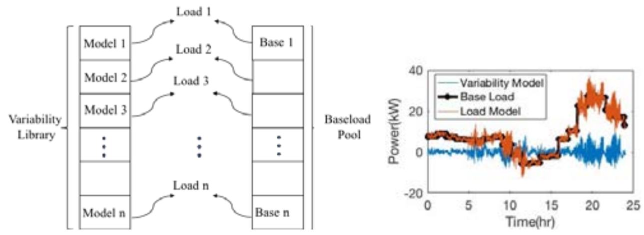


Fig. 5. IEEE 123-node test feeder

Load profiles for *n* days were randomly selected from the orignal load database to perform the testing. Then the corresponding *n* approximate coefficients at Level 12 from the baseload database were extracted as the low-resolution baseloads. As shown in Fig. 6(a), the *n* load variability models, which are ranked on the top of the list from the variability library, were selected and added onto the top of the baseloads to construct the high-resolution load profiles. The value of *n* can be a random integer that is greater than the number of load nodes. Fig. 6(b) shows a sample load profile with constituent components from the descibed approach. The blue line is the load variability model extracted from the library, and the black line represents the baseload from the baseload pool. Adding the above two components gets the modeled load profile, which is the red line.

Then the load aggregation approach is presented in Section II. D is used to obtain the 91 load profiles at the 91 load nodes of the IEEE 123-node test feeder according to the node peak load information.



(a) (b)

Fig. 6. Load modeling for IEEE 123-node test feeder

Evaluations have been completed for various temporal resolutions, including 1 s, 1 min, 5 min, 10 min, 15 min, and 30 min. The data with lower resolutions are obtained by averaging the data with higher resolutions. Fig. 7 shows one sample of the original load data and modeled load data, which has a resolution of 1 min.

Comparisons between the results of the original data testing and the modeled data testing were analyzed via the aforementioned metrics. The number of regulator operation comparisons is shown in Table II. The minimum and maximum voltage comparison is shown in Table III. Table IV shows the correlation coefficients of the voltage ramp profiles. The coeffcients were also calcuated using the Pearson correlation coefficient, as shown in (3). Fig. 8 shows the comparison of the histograms for the voltage ramp profiles.

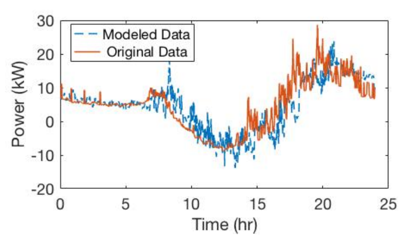


Fig. 7. Load sample of 1-minute resolution

From Table II, we can see that the modeled load has captured most of the impact on voltage regulators caused by load variability. For the test cases of resolution ranging from 5 min to 30 min, the number of regulator operations for the modeled data testing case is almost the same as that of the original data testing case. Also, it is observered that the higher resolution data has captured more regulator operations than the lower resolution data. This reveals the importance of modeling using high-resolution data for feeder testing, particularly for voltage regulation equipment impacts. The error rates for 1-s and 1-min data testing are more than 20%, which might be caused by the size of the load variability library. Because only the fall data were used for this evaluation, there might not be enough unique variability models to capture the system-level diversity characteristics. This could be improved in future work.

Table III shows that the minimum and maximum voltages match well for the original case and modeled case. The high values for the correlation coefficient shown in Table IV demonstrate that the voltage ramp distribution of the two cases for different resolutions align well. This agreement is also shown by the the histograms in Fig. 8, which cover the histograms for the original and modeled cases with resolutions from 1 s to 10 min.

TABLE II. COMPARISON OF NUMBER OF REGULATOR OPERATIONS

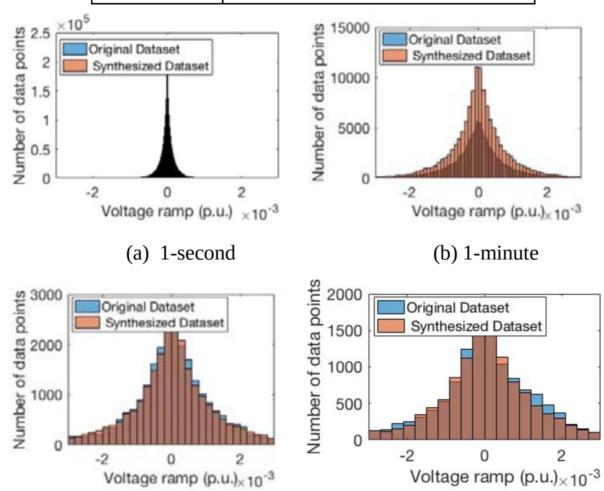
	Voltage Regulator Operations		
Resolution	Original Load	Modeled Load	Error Rate
1 second	126	91	27.78%
1 minute	114	87	23.68%
5 minutes	59	60	1.69%
10 minutes	51	52	1.95%
15 minutes	53	50	5.66%
30 minutes	49	48	2.04%

TABLE III. COMPARISON OF MAXIMUM AND MINIMUM VOLTAGE

	Original Load		Modeled Load	
Resolution	Maximum Voltage (p.u.)	Minimum Voltage (p.u.)	Maximum Voltage (p.u.)	Minimum Voltage (p.u.)
1 second	1.0518	0.9773	1.0516	0.9768
1 minute	1.0509	0.9798	1.0513	0.9787
5 minutes	1.0520	0.9777	1.0524	0.9797
10 minutes	1.0504	0.9758	1.0531	0.9764
15 minutes	1.0507	0.9809	1.0532	0.9742
30 minutes	1.0523	0.9844	1.0521	0.9797

TABLE IV. CORRELATION COEFFICIENT OF VOLTAGE RAMP DISTRIBUTION

Resolution	Correlation Coefficient
1 second	0.9882
1 minute	0.9999
5 minutes	0.9985
10 minutes	0.9926
15 minutes	0.9950
30 minutes	0.9897



(a) 1-second (b) 1-minute

(c) 5-minute (d) 10-minute

Fig. 8. Histogram of voltage ramp for various resolutions

C. Testing on Realistic Feeder

A realsitic utility feeder from California has also been used to test the modeled load. There are 619 load nodes in total and 3 regulators installed on this feeder. The load profies for this feeder are constructed in a way that is similar to the IEEE 123-node test feeder.

The tests have also been done for various resolutions, including 1 s, 1 min, 5 min, 10 min, 15 min, and 30 min. Metric comparisons between the results of the original data testing and the modeled data testing are summarized in Table V, Table VI, and Table VII, repesectively. Fig. 8 shows the comparison of the histograms for the voltage ramp profiles.

Table V shows that the modeled load profiles have successfully captured most of the influence on voltage regulators caused by the the load variability. The error rates and the absolute difference between the number of regulator operations for the original and modeled cases are both small.

Table VI shows that the minimum and maximum voltages match well for the original case and modeled case. Table VII shows that the voltage ramp distribution of the two cases with different resolutions all have very strong correlation. The histograms in Fig. 9 also demonstrate that the voltalge ramp distributions of the orignal and modeled cases align very well. These results demonstrate that the modeled loads could effectively represent the load variability influence on the feeder and could capture most of the charactersictics brought by the original loads.

TABLE V. COMPARISON OF NUMBER OF REGULATOR OPERATIONS

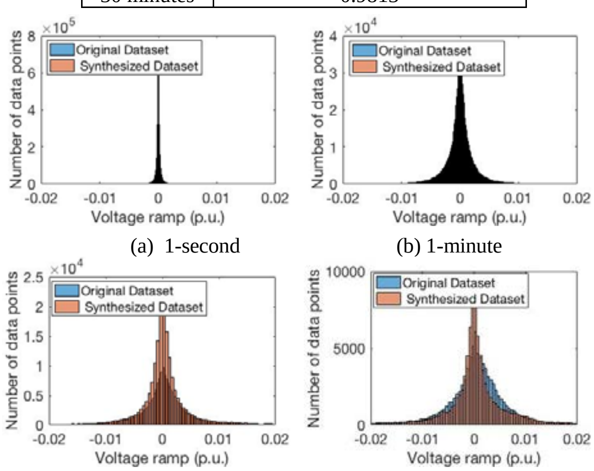
	Voltage Regulator Operations		
Resolution	Original Load	Modeled Load	Error Rate
1 second	130	123	5.38%
1 minute	124	111	10.48%
5 minutes	103	98	4.85%
10 minutes	101	102	0.99%
15 minutes	95	98	3.16%
30 minutes	77	75	2.6%

TABLE VI. COMPARISON OF MAXIMUM AND MINIMUM VOLTAGE

	Original Load		Modeled Load	
Resolution	Maximum Voltage (p.u.)	Minimum Voltage (p.u.)	Maximum Voltage (p.u.)	Minimum Voltage (p.u.)
1 second	1.0544	0.9533	1.0519	0.9539
1 minute	1.0600	0.9748	1.0530	0.9637
5 minutes	1.0547	0.9713	1.0517	0.9691
10 minutes	1.0534	0.9721	1.0514	0.9768
15 minutes	1.0528	0.9751	1.0510	0.9776
30 minutes	1.0525	0.9511	1.0508	0.9501

TABLE VII. CORRELATION COEFFICIENT OF VOLTAGE RAMP DISTRIBUTION

Resolution	Correlation Coefficient
1 second	0.9882
1 minute	0.9984
5 minutes	0.9985
10 minutes	0.9968
15 minutes	0.9700
30 minutes	0.9813



(a) 1-second (b) 1-minute

(c) 5-minute (d) 10-minute

Fig. 9. Histogram of voltage ramp for various resolutions

IV. CONCLUSION AND FUTURE WORK

This paper presents a DWT-based load modeling methodology to model the time-series loads to enable effectual QSTS simulations of distribution feeders. The load model is constructed from two major parts: the low-resolution field-measured load data, which usually have resolutions of 15 or 30 min; and the high-resolution load variability model extracted from a previously developed variability database, which could have resolutions up to 1 sec. A load aggregation methodology is also developed to aggregate the load profiles so that load profiles at different transformer ratings can be modeled effectively. The developed load modeling approach has been tested on the IEEE 123-node test feeder and an actual utility feeder from California. The analysis completed on the two feeders demonstrated the value of the developed model for distribution feeder QSTS simulations. Future work will focus on refining the load variability library and developing models with more effective customer diversity characteristics.

ACKNOWLEDGMENT

This work was authored by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies. 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.

REFERENCES

[1] A. Payne, R. Duke, R. H. Williams. "Accelerating residential PV expansion: supply analysis for competitive electricity markets." *Energy Policy* 29.10 (2001): 787-800.

[2] F. Katiraei, J. R. Agüero. "Solar PV integration challenges." *IEEE Power and Energy Magazine*, 9.3 (2011): 62-71.

[3] A. Arif, et al., "Load Modeling – A Review," *preprint – IEEE Trans. on Smart Grid*, pp. 1-15, May, 2017.

[4] M. Lave, et al. "High temporal resolution load variability compared to PV variability." *IEEE Photovolt. Spec. Conf.* (*PVSC*), 2016.

[5] M. Chamana, B. Mather. "Variability extraction and synthesis via multi-resolution analysis using distribution transformer high-speed power data." *IEEE Intelligent System Application to Power Systems* (*ISAP*), 2017.

[6] R. Li, F. Li, N. D. Smith, "Load Characterization and Low-order Approximation for Smart Metering Data in the Spectral Domain," *IEEE Trans. on Ind. Informatics*, vol. PP, pp. 1-1, 2016.

[7] S. M. Kelo, S. V. Dudul, "Short-term load prediction with a special emphasis on weather compensation using a novel committee of wavelet recurrent neural networks and regression methods," *in 2010 Joint International Conference on Power Electronics, Drives and Energy Systems & 2010 Power India*, 2010, pp. 1-6.

[8] Sir R.A. Fisher, "Statistical Methods for Research Workers.(13-th edition.) Oliver and Boyd." Edinburgh, London (1958).

[9] X. Zhu,, et al. "Voltage-Load Sensitivity Matrix Based Demand Response for Voltage Control in High Solar Penetration Distribution Feeders." *IEEE PES Gener. Meet.* (*PESGM*), 2017.

[10] IEEE 123-bus Feeder, available online: https://ewh.ieee.org/soc/pes/dsacom/testfeeders/