On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers

Preprint

Marcos Netto, Jun Hao, Harsha Padullaparti, and
Venkat Krishnan

National Renewable Energy Laboratory

Presented at the 2021 IEEE Power and Energy Society General Meeting
July 25–29, 2021

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-5D00-78295
July 2021

On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers

Preprint

Marcos Netto, Jun Hao, Harsha Padullaparti, and Venkat Krishnan

National Renewable Energy Laboratory

Suggested Citation
Netto, Marcos, Jun Hao, Harsha Padullaparti, and Venkat Krishnan. 2021. *On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-78295. https://www.nrel.gov/docs/fy21osti/78295.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-5D00-78295
July 2021

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

On the Use of Smart Meter Data to Estimate the Voltage Magnitude on the Primary Side of Distribution Service Transformers

Marcos Netto, *Member, IEEE,* Jun Hao, *Student Member, IEEE,* Harsha Padullaparti, *Member, IEEE*, and Venkat Krishnan, *Senior Member, IEEE*

***Abstract*—This paper develops a novel method to estimate the voltage magnitude on the primary side of distribution service transformers. The proposed method relies exclusively on smart meters, and therefore it is fully data-driven. This is an important feature because electric utilities have detailed models of only the primary network—that is, the network between the distribution substation and the primary side of service transformers that are installed closer to end-customer sites. The network that connects the secondary side of service transformers to end-customer sites, referred to as the secondary network, is simply represented by a lumped load. For each secondary network, the proposed method uses data acquired from only 2 smart meters: the *closest* and the *farthest*—in the sense of electrical distance—from the service transformer. As a reference to this feature, the proposed method is named `SM2Vp`. To our knowledge, this is the first time a method is shown to provide actionable information for real-time operation and control of power distribution grids using only two smart meters per secondary network. This is important because utilities have experienced barriers in managing and using large data sets for real-time operation and control. `SM2Vp` is primarily intended to provide pseudo-measurements for distribution system state estimation, but it can also be used directly for voltage control schemes. The performance of `SM2Vp` is demonstrated by numerical simulations carried out on three secondary network synthetic models and by using field data provided by a utility partner serving customers in southwestern California. A maximum relative error of approximately 3.9% or less is observed for the primary voltage magnitude estimates in all numerical experiments.**

***Index Terms*—Distribution system state estimation, distribution service transformer, pseudo-measurement, smart meter.**

I. Introduction

Electric utilities are undergoing radical changes triggered by the advent of new technologies in power electronics, energy conversion and storage, sensing and measurements, and communications and networking. These technologies have collectively led to the development of small-scale electric power generation and storage systems that are connected to the grid in a distributed fashion, closer to end-customer sites. These systems are commonly referred to as distributed energy resources (DERs) [1]. Following [2], "rooftop solar photovoltaics (PVs) have the highest profile of these resources, but DERs include any generator or energy-storage device connected at distribution voltage levels and characterized by relatively small capacities (e.g., a few kilowatts to a few megawatts)." The dynamics of distribution systems are changing tremendously because of DERs. For example, electrical quantities such as voltage and power are varying more abruptly in distribution secondary networks with significant numbers of installed PV. This poses a threat to the effectiveness of long-established distribution operation and control schemes. In particular, the voltage profile from the distribution substation, along the feeder, down to the end-customer site is no longer monotonically decreasing. Consequently, rules of thumb [3] used in the past to ensure that voltages were within limits (1.00 ± 0.05 per unit) throughout the distribution network are no longer acceptable. It is now widely recognized that the next generation of distribution operation and control schemes [4], [5] will rely on distribution system state estimation (DSSE) [6]. Typically, DSSE is designed to provide estimates of the algebraic states of the distribution system, including the *primary network*. There have been some proposals to include the secondary network as well [7], but this would require a detailed model that is not available. For this reason, in what follows, we assume that DSSE involves only the primary network, and the secondary network is simply modeled by a lumped load. Note that the set of algebraic states is not unique, and it can be defined as i) voltage magnitudes and phase angles or ii) current magnitudes and phase angles, depending on how the DSSE problem is formulated. Despite all the progress in DSSE algorithms, deploying meters in distribution networks [8] is challenging because of the greater number of spatially distributed electrical nodes. As a consequence, observability remains the key bottleneck preventing DSSE from being effectively realized. This is why researchers and practitioners resort heavily to using pseudo-measurements [9], [10] for DSSE.

Meanwhile, legacy energy meters used for customer billing are being progressively replaced by smart meters [13]. It turns out that, other than billing, the data acquired from smart meters have a secondary set of equally important roles. Examples include topology and phase identification and model identification and calibration [14]–[17]. These tasks are accomplished, however, under the assumption that all end-customer sites connected to the secondary network are monitored by smart meters. As another example, data from smart meters can also be used to enhance DSSE [18]. Along these lines, [19] studied the role of smart meter data to enhance the system observability after reformulating the DSSE problem as an optimization problem; [20] addressed the use of nonsynchronized measurements coming from smart meters in DSSE; and [7] proposed a multilevel DSSE that includes both primary and secondary networks. These works assume either that a detailed model of the secondary network is available, or that a large portion of end-customer sites are monitored by smart meters, or both. These assumptions, which are not always true, have been an impediment for the adoption of previous works by electric utilities for real-time operation and control.

In this paper, we propose to estimate the voltage magnitude on the primary side of distribution service transformers by

This work was authored by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, 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 (EERE), and by Electric Program Investment Charge (EPIC) of San Diego Gas & Electric Company under the cooperative research and development agreement CRD-17-712. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government 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. The authors are with the National Renewable Energy Laboratory, Golden, CO 80401, USA. Corresponding author: marcos.netto@nrel.gov.

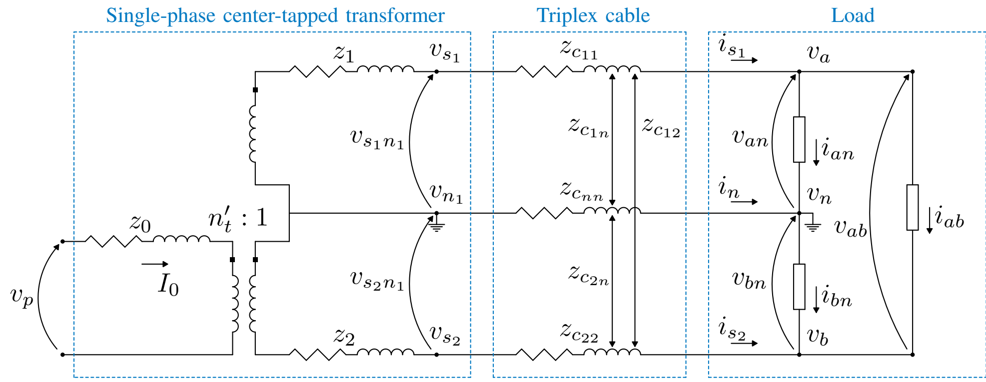

Fig. 1: Single-phase center-tapped transformer with a single connected load [11], [12].

using only a limited number of smart meters—in particular, we consider the extreme case of using data from only two smart meters per secondary network. The obtained voltage magnitude estimates can be readily incorporated into existing DSSE algorithms as pseudo-measurements. To give the reader a sense of the impact of the proposed method, referred to as `SM2Vp`, a single feeder in our utility partner's distribution grid has 345 service transformers; therefore, `SM2Vp` provides 345 pseudo-measurements by using data from 700 smart meters installed across this particular feeder, without any additional information regarding the secondary network. Further, we observe that the accuracy of `SM2Vp` is better than the reported accuracy of other data-driven methods used to generate pseudo-measurements from smart meters [10], [21]. This is possible because of the derived two-stage approach that uses the fact that smart meters are strategically placed on the closest and farthest—in the electrical sense—load from the service transformer. Note that the proposed approach is extendable to include more than two smart meters per secondary network.

This paper proceeds as follows. Section II briefly introduces the modeling of power distribution secondary networks and shows the electrical link between smart meter data and the primary voltage of interest. Section III presents the proposed `SM2Vp` method. Section IV discusses the numerical results, and Section V provides conclusions and future work.

Notation: In what follows, lower-case (upper-case) boldface letters are used to denote column vectors (matrices); $\mathbf{0}$ denotes the all-zero vector of suitable dimension; and $\boldsymbol{I}_m$ and $\mathbf{0}_m$ denote, respectively, the identity matrix and the square matrix whose entries are all zero, both of dimension m. The superscript $^\top$ denotes transposition.

II. Secondary Low-Voltage Circuit Model

This section introduces the basics of modeling power distribution secondary networks. The goal is to establish the link between smart meter data and the primary voltage of interest. The full model of a single-phase, center-tapped transformer with a single load connected to its terminals through a triplex cable is shown in Fig. 1. Note that using center-tapped transformers is common practice in the United States, hence its use in the present work. For simplicity of presentation, Fig. 1 illustrates the connection of a single load. In practice, a secondary network typically supplies 2–10 customers with split-phase service consisting of a 240-V connection that is split into two 120-V circuits [12].

A. Single-Phase Center-Tapped Transformer

Transformer manufacturers usually provide the transformer total losses, no-load losses, and the full-winding short-circuit impedance, $z_{sc} = r_{sc} + jx_{sc}$. From [12], based on z_{sc}, for core-type or shell-type transformers with interlaced secondary winding:

$$z_0 = 0.5r_{sc} + j0.8x_{sc}, \tag{1}$$
$$z_1 = 1.0r_{sc} + j0.4x_{sc}, \tag{2}$$
$$z_2 = 1.0r_{sc} + j0.4x_{sc}. \tag{3}$$

For shell-type transformers with noninterlaced secondary winding:

$$z_0 = 0.25r_{sc} - j0.6x_{sc}, \tag{4}$$
$$z_1 = 1.50r_{sc} + j3.3x_{sc}, \quad \text{(outer winding)} \tag{5}$$
$$z_2 = 1.50r_{sc} + j3.1x_{sc}. \quad \text{(inner winding)} \tag{6}$$

The impedances in (1)–(6) are in per unit and must be converted to ohms relative to the respective sides of the transformer. Let:

$$n_t' = \frac{\text{high-side rated voltage}}{\text{low-side half-winding rated voltage}}, \tag{7}$$

where the low-side half-winding rated voltage is 120 V. The relationship between voltages and currents in the single-phase center-tapped transformer shown in Fig. 1 is given by [11]:

$$\begin{bmatrix} v_{s_1} \\ v_{s_2} \end{bmatrix} = \frac{1}{n_t} \begin{bmatrix} v_p \\ v_p \end{bmatrix} - \boldsymbol{Z}_t \begin{bmatrix} i_{s_1} \\ i_{s_2} \end{bmatrix}, \tag{8}$$

where:

$$\boldsymbol{Z}_t = \begin{bmatrix} z_1 + z_0/n_t^2 & -z_0/n_t^2 \\ z_0/n_t^2 & -z_2 - z_0/n_t^2 \end{bmatrix}. \tag{9}$$

B. Triplex Cable

Typical triplex cables consist of two identical insulated conductors wrapped around a noninsulated neutral conductor.

The π-model is typically adopted. The model is defined by a series of primitive impedance matrices, $\boldsymbol{Z}_{c_{12n}}$, and primitive shunt admittance matrices, $\boldsymbol{Y}_{c_{12n}}$, where:

$$\boldsymbol{Z}_{c_{12n}} = \boldsymbol{R}_{c_{12n}} + j\boldsymbol{X}_{c_{12n}}, \tag{10}$$
$$\boldsymbol{Y}_{c_{12n}} = \boldsymbol{G}_{c_{12n}} + j\boldsymbol{B}_{c_{12n}}. \tag{11}$$

The conductance, $\boldsymbol{G}_{c_{12n}}$, is typically ignored because it is much smaller than the capacitive susceptance, $\boldsymbol{B}_{c_{12n}}$ [22]. For short lines, $\boldsymbol{B}_{c_{12n}}$ is so small that the resulting impedance is much larger than $\boldsymbol{Z}_{c_{12n}}$, and thus $\boldsymbol{B}_{c_{12n}}$ is also neglected; therefore, the relationship between voltages and currents in the triplex cable is given by:

$$\begin{bmatrix} v_{s_1} \\ v_{s_2} \\ v_{n_1} \end{bmatrix} = \begin{bmatrix} v_a \\ v_b \\ v_n \end{bmatrix} + \begin{bmatrix} z_{c_{11}} & z_{c_{12}} & z_{c_{1n}} \\ z_{c_{12}} & z_{c_{22}} & z_{c_{2n}} \\ z_{c_{1n}} & z_{c_{2n}} & z_{c_{nn}} \end{bmatrix} \begin{bmatrix} i_{s_1} \\ i_{s_2} \\ i_n \end{bmatrix} = \boldsymbol{v}_{\ell_{12n}} + \boldsymbol{Z}_{c_{12n}} \boldsymbol{i}_{12n}. \tag{12}$$

Provided that the phase conductors are identical and placed symmetrically with respect to the neutral conductor, $\boldsymbol{Z}_{c_{12n}}$ is always symmetric. The elements of $\boldsymbol{Z}_{c_{12n}}$ can be calculated using, e.g., the modified Carson's equations [22]. If the neutral conductor is effectively grounded, i.e., $v_{n_1} = v_n = 0$, the

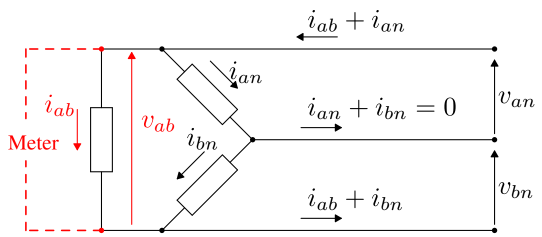

Fig. 2: Smart meter location and acquired measurements.

equation for the neutral conductor in (12) can be eliminated via the Kron reduction, yielding:

$$\begin{bmatrix} v_{s_1} \\ v_{s_2} \end{bmatrix} = \begin{bmatrix} v_a \\ v_b \end{bmatrix} + \begin{bmatrix} z'_{c_{11}} & z'_{c_{12}} \\ z'_{c_{12}} & z'_{c_{22}} \end{bmatrix} \begin{bmatrix} i_{s_1} \\ i_{s_2} \end{bmatrix}. \tag{13}$$

The secondary network synthetic models in this manuscript are built based on the principles introduced in this section.

III. The Proposed SM2Vp Method

From the brief introduction in Section II to distribution grid secondary network modeling, it is evident that an estimate of v_p is challenging to obtain if only 2 of n_ℓ loads connected to the service transformer are equipped with smart meters. We start from the following assumptions:

1) The neutral conductor of the triplex cable is effectively grounded, i.e., $v_{n_1} = v_n = 0$ (see Fig. 1).
2) The loads connected between v_a and v_n, and between v_b and v_n (see Fig. 1) are identical, and thus the current in the neutral conductor is equal to zero (see Fig. 2).
3) The smart meter is connected as shown in Fig. 2, and it provides ideal measurements—noises, gross measurement errors, data dropouts, and time delays are neglected.

Note that, as demonstrated in [12], "split-phase secondary networks [as shown in Fig. 1] can be modeled accurately with single-phase equivalents under perfectly balanced conditions;" the balanced condition is provided by Assumption 2. We propose using the equivalent one-line circuit shown in Fig. 3. To this end, we further assume that:

4) The closest (p_1) and the farthest (p_2)—in the sense of electrical distance—loads from the service transformer are equipped with smart meters; all remaining $n_\ell - 2$ loads are not metered.
5) All remaining $n_\ell - 2$ loads in the secondary are lumped as a single load, p_u.

At any point in time, v_1, p_1, v_2, p_2, and $n_t = 2n'_t$ are known; this is the only information we have access to. Note that circuit reactances and reactive power are not considered because the available smart meter provides measurements of voltage magnitude and kilowatt-hour—from which we obtain a sample measurement of active power per hour.

The `SM2Vp` method has two stages. The first stage performs a linear regression on the latest data window available at the control center—we use a data window of 288 points, which is equivalent to a day for a 5-minute sampling resolution—and it is executed only once. The second stage comprises a Kalman filter and it is used to update the voltage magnitude estimates continuously based on new data points. The formulation of both stages is presented next.

A. First Stage: Linear Regression

From Kirchhoff's voltage law, we have:

$$\begin{aligned} v_p &= v'_p + r_p i_p = n_t v'_s + r_p i_s / n_t \\ &= n_t v'_s + r_p \left(i_1 + i_2 + i_u\right)/n_t. \end{aligned} \tag{14}$$

$$v'_s = \left(r_s + r_1\right)\left(i_1 + i_2 + i_u\right) + v_1, \tag{15}$$

$$\quad = \left(r_s + r_1\right)\left(i_1 + i_2 + i_u\right) + v_2 + r_2 i_2. \tag{16}$$

$$v_1 - v_2 = r_2 i_2. \tag{17}$$

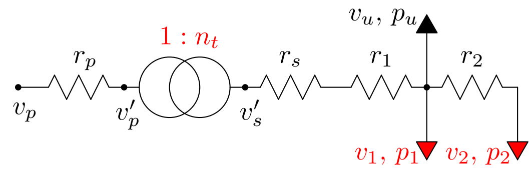

Fig. 3: Equivalent circuit used by the `SM2Vp` method.

where r_p and r_s denote, respectively, the losses in the primary and the secondary winding of the service transformer; v'_p and v'_s denote, respectively, the voltage magnitude at the primary and the secondary of an ideal transformer; n_t is as defined in (7); v_1 and p_1 (v_2 and p_2) denote, respectively, the voltage magnitude and the active power measured at the closest (farthest) load from the service transformer; r_1 and r_2 account for cable impedance; and v_u and p_u are unknown, time-varying quantities that denote the voltage magnitude and the active power of the lumped load. By substituting (15) and (16) into (14) and rearranging:

$$v_1 = -r'_1\left(i_1 + i_2\right) - k_u + v_p/n_t, \tag{18}$$

$$v_2 = -r'_1\left(i_1 + i_2\right) - k_u + v_p/n_t - r_2 i_2, \tag{19}$$

where $r'_1 = r_1 + r_s + \frac{1}{n_t^2} r_p$, and $k_u = r'_1 i_u$. Basically, k_u is the voltage drop through the line caused by the pseudo-node. Now, let $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{i}_1$, and $\boldsymbol{i}_2$ be column vectors containing m measurements. For example, $\boldsymbol{v}_1 = \left[v_1^{(1)} \ \cdots \ v_1^{(m)}\right]^\mathsf{T}$. By using (17)–(19), it follows that:

$$\begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \boldsymbol{v}_1 - \boldsymbol{v}_2 \end{bmatrix} = \left[\begin{array}{c:c:c:c} -\left(\boldsymbol{i}_1 + \boldsymbol{i}_2\right) & \boldsymbol{0} & \frac{1}{n_t}\boldsymbol{I}_m & -\boldsymbol{I}_m \\ -\left(\boldsymbol{i}_1 + \boldsymbol{i}_2\right) & -\boldsymbol{i}_2 & \frac{1}{n_t}\boldsymbol{I}_m & -\boldsymbol{I}_m \\ \boldsymbol{0} & \boldsymbol{i}_2 & \boldsymbol{0}_m & \boldsymbol{0}_m \end{array}\right] \begin{bmatrix} r'_1 \\ r_2 \\ \boldsymbol{v}_p \\ \boldsymbol{k}_u \end{bmatrix},$$

or simply:

$$\boldsymbol{d} = \boldsymbol{C}\boldsymbol{x}. \tag{20}$$

Unfortunately, the matrix $\boldsymbol{C}$ is rank deficient, and (20) has infinitely many solutions; thus, a classic linear least-squares regression-type approach will probably lead to an incorrect solution. Instead, to solve (20), we formulate it as a constrained linear least-squares problem of the form:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \frac{1}{2}\left\|\boldsymbol{C}\boldsymbol{x} - \boldsymbol{d}\right\|_2^2 \\ \text{s.t.} \quad & \boldsymbol{\ell}_b \leq \boldsymbol{x} \leq \boldsymbol{u}_b, \end{aligned} \tag{21}$$

where the lower, $\boldsymbol{\ell}_b$, and upper, $\boldsymbol{u}_b$, bounds play an important role in finding a physically meaningful solution to (20). After solving (21), we have $\hat{\boldsymbol{x}} = \left[\hat{r}'_1 \ \hat{r}_2 \ \hat{\boldsymbol{v}}_p^\mathsf{T} \ \hat{\boldsymbol{k}}_u^\mathsf{T}\right]^\mathsf{T}$ and $\hat{\boldsymbol{i}}_u^\mathsf{T} = \hat{\boldsymbol{k}}_u^\mathsf{T} \big/ \hat{r}'_1$.

B. Second Stage: Kalman Filtering

The first stage in Section III-A is intended to estimate the parameters of the equivalent circuit, r'_1 and r_2, which are constant scalars. Conversely, the vector-valued variables, v_p and k_u, are time-varying. We design a discrete-time Kalman filter to track these variables, as presented next. Let:

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1} \\ \boldsymbol{z}_k &= \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{e}_k, \end{aligned} \tag{22}$$

where $\boldsymbol{x} = \left[v_p \ i_1 \ i_2 \ i_u\right]^\mathsf{T}$, $\boldsymbol{z} = \left[v_1 \ v_2 \ i_1 \ i_2 \ \tilde{i}_u\right]^\mathsf{T}$,

$$\boldsymbol{H} = \begin{bmatrix} 1/n_t & -r_1 & -r_1 & -r_1 \\ 1/n_t & -r_1 & -\left(r_1 + r_2\right) & -r_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

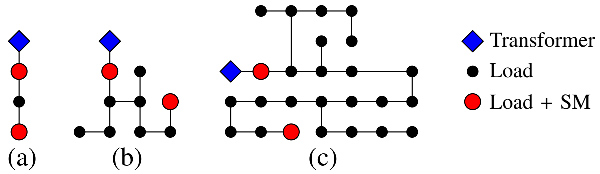

Fig. 4: Topology of the secondary network synthetic models. (a) 3 loads. (b) 9 loads. (c) 25 loads.

and $\tilde{i}_u$ denotes a pseudo-measurement of the current in the unknown lumped load, given by an autoregressive model of Order 2 as follows:

$$i_{u,k} = a_1 i_{u,k-1} + a_2 i_{u,k-2} + e_k, \tag{23}$$

where $k \in \mathbb{Z}$ is the discrete time, e_k is the error at time step k, and a_i denotes the model parameters.

IV. Numerical Results

A. Secondary Network Synthetic Models

To test SM2Vp, we developed three secondary network synthetic models of increasing complexity, as shown in Fig. 4. These models are developed in OpenDSS and consider i) resistance and reactance for the transformer and all cables; and ii) a power factor equal to or less than 0.9 for all loads. This is illustrated in Fig. 5 for the circuit with three loads (case a in 4). One week of synthetic smart meter data is generated for each of the three circuits under different scenarios, which include balanced and unbalanced networks and different levels of PV penetration, as shown in Table I. Note that to be consistent with current smart meter technology, we generate synthetic measurements of voltage magnitude on a 5-minute time resolution and measurements of kilowatt-hour on a 60-minute time resolution. The proposed SM2Vp method currently uses linear interpolation to bring measurements of active power to the same time resolution of measurements of voltage magnitude. The data are separated into training and testing data sets. One day of data—that is, 288 data points on a 5-minute resolution—is used for training, described as the first stage in Section III-A. The remaining six days of data—1,728 data points—are for testing, described as the second stage in Section III-B. We use the relative error of the estimated primary voltage magnitude to evaluate the performance of the proposed methodology. The relative error is defined as $e := \left(100 \cdot |\hat{v}_p - v_{p,\text{true}}|\right) / v_{p,\text{true}}$, where e is the relative error, $\hat{v}_p$ denotes the estimated voltage magnitude, and $v_{p,\text{true}}$ is the true voltage magnitude obtained in OpenDSS.

Table II shows the maximum relative error for all cases in Table I. For the circuits 25a–25d, Fig. 6 shows in detail the relative error in time. The largest relative error among all the cases is 2.5225%, which is the case with an unbalanced network and low PV penetration level. The results indicate that the proposed SM2Vp method provides results with good accuracy under various secondary network topologies and different loading conditions, despite the fact that the mathematical formulation neglects reactances and reactive power. We should mention that 316 of 345 secondary networks in one of our utility partner's feeders have 25 loads or less, thereby corroborating the practical relevance of test cases 25a–25d. We should also point out that all the results have the similar periodic pattern shown in Fig. 6. The periodic oscillation in the relative error is mainly caused by the lumped load representing the part of the secondary network that does not have smart meters. The second-order autoregressive model used to forecast the lumped load is only effective to a certain extent. Table III shows the computational time for all cases in Table I. In all cases, 205 seconds or less is needed to simulate 144 hours—or 6 days worth of data. If we assume that a batch of smart meter data is collected at the control center on a 5-minute time resolution, it is evident that the proposed SM2Vp method is adequate in terms of processing time.

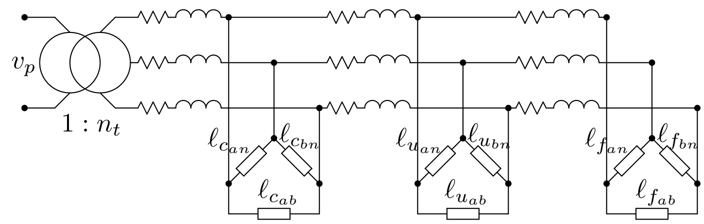

Fig. 5: (a) The three-phase model of the secondary network synthetic with three loads, as built in OpenDSS.

TABLE I: Secondary network synthetic models

Circuit	No. of loads	Circuit loading	No. of loads with PV
3a	3	Balanced	0
9a	9	Balanced	0
9b	9	Unbalanced	0
9c	9	Unbalanced	3
9d	9	Unbalanced	5
25a	25	Balanced	0
25b	25	Unbalanced	0
25c	25	Unbalanced	5
25d	25	Unbalanced	10

TABLE II: Maximum relative error for synthetic models

No. of loads	Case (all values in %)			
	a	b	c	d
3	1.0069	—	—	—
9	1.3863	1.3632	1.3458	1.0362
25	0.6011	0.7269	2.5225	1.3959

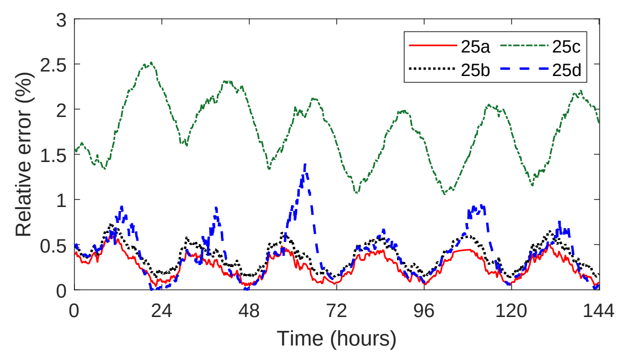

Fig. 6: Relative error in the voltage magnitude estimates at the primary side of the service transformer for the circuit with 25 loads. Note that a relative error of 1% is equivalent to a relative error of 0.01 per unit if $v_{p,\text{true}}$ is the base voltage.

TABLE III: Computational time for synthetic models

No. of loads	Case (all values in seconds)			
	a	b	c	d
3	205	—	—	—
9	189	199	196	187
25	189	187	188	188

TABLE IV: Maximum relative error for utility models

Secondary	No. of loads in the field	Maximum relative error (%)
A	9	3.2520
B	13	1.7592
C	181	3.4023
D	20	3.1706
E	35	3.8517

B. Real Smart Meter Data from a Utility

The original feeder model from the utility partner has only the primary network; service transformers and associated secondary networks are simply modeled by a lumped load. The secondary networks in this particular feeder are, however, fully monitored by smart meters. We therefore used all available smart meter data sets to build secondary network equivalent models for this feeder as in Fig. 5. The utility partner identified

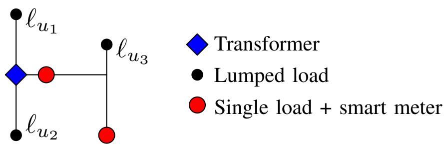

Fig. 7: Utility's equivalent secondary network model.

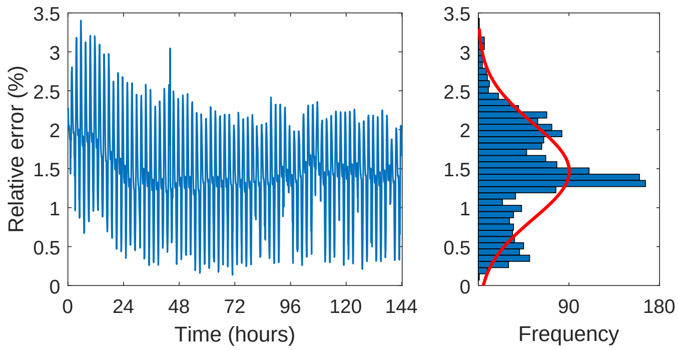

Fig. 8: (Left) Relative error in the voltage magnitude estimates at the primary side of the service transformer for Secondary C in Table IV. (Right) Histogram of the relative error with a Gaussian distribution fit.

the smart meter data sets associated with the closest and farthest loads for each secondary; these data sets are directly used as load profiles. The remaining unknown load is modeled as an aggregated unmeasured load, and its load profile is computed by subtracting the sum of the two smart meter data sets from the service transformer rated power. The service transformer ratings and impedance parameters are selected based on the available field data. This approach is used to model 340 out of the 345 service transformers in this feeder. The remaining 5 service transformers were selected for this study because they have different number of loads, as shown in Table IV, and interesting characteristics of the 5 secondaries designated A to E. The secondary networks associated with these service transformers are represented by a refined equivalent model, using a principled methodology developed at NREL. The description of this methodology is beyond the scope of this paper and will be reported elsewhere. To provide the reader with some details, however, the topology of the refined equivalent model for secondary network C in Table IV is shown in Fig. 7.

The primary voltages of the selected secondaries are estimated using `SM2Vp`, and the maximum relative errors are summarized in Table IV. The maximum relative error is higher than that observed in the case of the synthetic models. This could be because of the increased complexity of the utility feeder compared to the synthetic models. The relative error in time is shown in Fig. 8 for the Secondary C, which has the highest number of customer loads. For the same secondary, a histogram of the error is also shown in Fig. 8.

V. Conclusions and Avenues for Future Research

We propose a novel method, referred to as `SM2VP`, to estimate the voltage magnitude on the primary side of service transformers using only two smart meter data sets. The method comprises two stages and can achieve consistent and promising performance in diverse scenarios tested with synthetic as well as real utility network data with relative error less than 3.9%. Two aspects need to be addressed in future work: i) Currently, bounded linear-least squares is implemented in stage one to estimate the parameters, and other optimization methods—such as multi-objective genetic algorithm—could be investigated in future work to improve the estimation results in the training stage. ii) Advanced forecasting algorithms—such as deep neural networks—could be implemented to replace the autoregressive model in the second stage to improve the prediction results for pseudo-nodes; therefore, the peaks in the results could be reduced, and the periodic estimation error pattern could potentially be eliminated. We envision that such use of smart meter data from the secondary network for the estimation of the primary-side voltage of the service transformer has the potential to enhance DSSE by supplying pseudo-measurement data in practical utility networks with limited primary-side measurements.

References

[1] R. Sherick and R. Yinger, "Modernizing the California grid: Preparing for a future with high penetrations of distributed energy resources," *IEEE Power & Energy Magazine*, vol. 15, no. 2, pp. 20–28, March 2017.

[2] S. P. Burger, J. D. Jenkins, S. C. Huntington, and I. J. Perez-Arriaga, "Why distributed?: A critical review of the tradeoffs between centralized and decentralized resources," *IEEE Power & Energy Magazine*, vol. 17, no. 2, pp. 16–24, March 2019.

[3] C. W. Brice, "Comparison of approximate and exact voltage drop calculations for distribution lines," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-101, no. 11, pp. 4428–4431, Nov 1982.

[4] J. L. Mathieu, S. Koch, and D. S. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 430–440, Feb 2013.

[5] P. Bagheri and W. Xu, "Assessing benefits of volt-var control schemes using AMI data analytics," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1295–1304, May 2017.

[6] A. Primadianto and C. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.

[7] M. Pau *et al.*, "Design and accuracy analysis of multilevel state estimation based on smart metering infrastructure," *IEEE Trans. Instrumentation and Measurement*, pp. 1–13, 2019.

[8] R. Singh, B. C. Pal, and R. B. Vinter, "Measurement placement in distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 668–675, May 2009.

[9] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Gener., Transm. & Dist.*, vol. 4, no. 1, pp. 50–59, January 2010.

[10] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Systems*, vol. 27, no. 4, pp. 1888–1896, 2012.

[11] W. H. Kersting, "Center-tapped transformer and 120-/240-V secondary models," *IEEE Trans. Ind. Applications*, vol. 45, no. 2, pp. 575–581, March 2009.

[12] J. Peppanen, C. Rocha, J. A. Taylor, and R. C. Dugan, "Secondary low-voltage circuit models—How good is good enough?" *IEEE Trans. Ind. Applications*, vol. 54, no. 1, pp. 150–159, Jan 2018.

[13] S. Neumann, F. Wilhoit, M. Goodrich, and V. M. Balijepalli, "Everything's talking to each other: Smart meters generate big data for utilities and customers," *IEEE Power & Energy Magazine*, vol. 14, no. 1, pp. 40–47, Jan 2016.

[14] T. A. Short, "Advanced metering for phase identification, transformer identification, and secondary modeling," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 651–658, 2013.

[15] J. Peppanen *et al.*, "Leveraging AMI data for distribution system model calibration and situational awareness," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2050–2059, July 2015.

[16] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Trans. Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, 2018.

[17] H. Padullaparti *et al.*, "Considerations for AMI-based operations for distribution feeders," in *2019 IEEE Power & Energy Society General Meeting*, 2019, pp. 1–5.

[18] M. Baran and T. E. McDermott, "Distribution system state estimation using AMI data," in *2009 IEEE/PES Power Systems Conference and Exposition*, 2009, pp. 1–3.

[19] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5953–5961, Nov 2018.

[20] A. Alimardani *et al.*, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, Nov 2015.

[21] K. Dehghanpour, Y. Yuan, Z. Wang, and F. Bu, "A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 5942–5951, 2019.

[22] W. H. Kersting, *Distribution System Modeling and Analysis*. CRC Press, 2017.