Event-Based Dynamic Response Modeling of Large Behind-the-Meter Solar Farms: A Data-Driven Method Based on Real-World Data

Parviz Khaledian, Alireza Shahsavari, and Hamed Mohsenian-Rad
Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA

***Abstract*—The dynamic behavior of a 4.3 MW behind-the-meter solar farm in California is modeled in response to grid-induced disturbances using data-driven techniques. The proposed modeling approach is *data-driven* and works solely based on the micro-PMU measurements at the solar farm's point of interconnection with the power. Accordingly, the proposed approach does *not* require any prior knowledge about the physical elements inside the solar farm. This allows applying the proposed methods to existing solar installations by the grid operators. Three methods are proposed based creating libraries of regression models and examining the similarities across recorded events and disturbances. The performance of the proposed methods are demonstrated by using real measurements and case studies.**

***Index Terms*—Dynamic response, solar farm, data-driven model, micro-PMU measurements, model library, practical study.**

I. Introduction

The growing penetration of solar power generation has introduced new challenges in the operation of the power grid, including the *changes in the dynamic behavior* of the power system [1]. Importantly, the dynamic response of the inverter-based resources (IBRs), such as solar farms, to disturbances is often very different from the dynamic response of the synchronous machines in traditional power plants [2]. This calls for introducing new models that can accurately capture the dynamic behavior of different types of solar farms in response to disturbances in the power system.

When a disturbance happens in the power grid, it causes some agitation in the operation of solar farms. Having knowledge of the dynamics of such agitation as a function of the disturbance is necessary for proper planing and operation of the power system with the high penetration of solar generation.

Standards require compliance in the dynamic response of the new installations of inverter-based resources, such as Rule 21 in California [3]. However, the existing solar installations do *not* follow these rules. Understanding the dynamic response of solar farms is important, whether old or new.

A. Technical Approach and Contributions

In this paper, we seek to address the following question: *How to model the dynamic response of solar farms using sensor measurements at the point of common-coupling, i.e., without the need to instrument the solar farm itself?* We answer this question by developing a novel event-based method.

The corresponding author is H. Mohsenian-Rad (hamed@ece.ucr.edu).

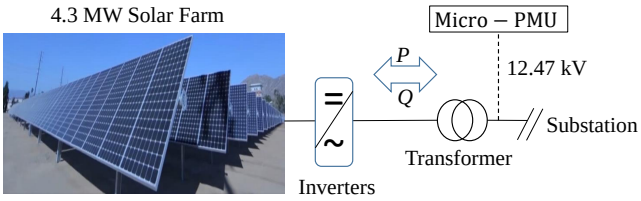

Fig. 1. The 4.3 MW behind-the-meter solar farm in this study is connected to a substation in Riverside, CA and it is monitored with a micro-PMU [4].

The analysis in this paper is based on real-world measurements that are obtained at a large solar farm that is directly integrated into a distribution substation; see Fig. 1. This solar farm is *behind-the-meter* (BTM), i.e., it is *not* owned and operated by the utility. However, it is monitored at the point of interconnection by a micro-PMU at the distribution substation.

The modeling approach in this paper is data-driven. It works by examining the dynamics of the solar farm when it *responds* to external disturbances in the power system. We do *not* need any prior knowledge about the physical components inside the solar farm. We also do *not* need access to the measurements inside the solar farm. The analysis in this paper is solely based on the measurements by the micro-PMU at the point of interconnection that are captured during grid-induced disturbances.

Two different dynamic modeling approaches are proposed. The first design is to develop a *single* dynamic model for the solar farm. This design mainly serves as a baseline for our second design, which is a *multiple* dynamic model. The latter model includes a comprehensive *library* of models that are developed and maintained through an *active learning process*. That is, new models are added to the library as needed.

The resulting models are useful in practice. The proposed method is applicable to the existing solar installations. The models can also be used for future feasibility studies.

B. Literature Review

Two common approaches in dynamic response modeling in power systems are *precise physical modelling* and *equivalent physical modelling* [5] [6]. Precise physics-based models are often heavy in computation complexity. Due to the high-order of the non-linearity of the power electronics and control components in solar farms, it is often not practical to include many internal states and details in physics-based models [7].

As for the equivalent physical models, they were introduced to reduce the complexity of the analysis. In [8], a PV station was made to be the equivalent of a large-scale distributed

PV stations. A simplified version of a boost converter was proposed in [9], the precision was satisfactory but the model ignored its grid-connected dynamic characteristics. Equivalent physical modeling is also considered by the Western Electricity Coordinating Council (WECC) [10]. In general, equivalent physical modeling may *not* reflect an accurate dynamic characteristics due to the need to often ignore several internal states.

A recent alternative to the above physics-based approaches is to use *data-driven* methods. There are methods that are hybrid, i.e., they combine measurements with equivalent physical models, e.g., see [11]. In [12] the distributed PV stations were clustered based on the dynamic affinity propagation.

One key advantage of data-driven dynamic modeling is that, the real-world data that is obtained from the under-study solar farm, naturally reflects the dynamic behavior of the system and its characteristics. Therefore, stand alone data-driven models can capture the real-world dynamic of the system well.

II. Approach and Data Preparation

The approach in this paper is data-driven. It involves the analysis of events that are captured by the micro-PMU at the point of common coupling (PCC) of the solar farm. In this section, we explain how we use the events to establish an input-output relationship for the dynamic behavior of the solar farm. We also discuss how we do a pre-processing on the measurements before we use them to develop the dynamic models.

A. Event-Based Approach

Recall from Section I that the analysis in this paper is based on examining the *events* that are captured by the micro-PMU in Fig. 1. There is already a rich literature to detect an event in micro-PMU measurements, e.g., in [13]. Let $\mathcal{E}$ denote the set of all the events that are captured for the purpose of developing the dynamic model of the solar farm. In general, we have:

$$\mathcal{E} = \mathcal{S} \cup \mathcal{G}, \tag{1}$$

where $\mathcal{S}$ denotes the set of the events whose root cause is the solar farm, and $\mathcal{G}$ denotes the set of the events whose root cause is the rest of the grid. In this regard, the events in set $\mathcal{G}$ act as *external disturbances* to the solar farm. Accordingly, our focus in this paper is on the events in set $\mathcal{G}$; and we seek to model the *dynamic response* of the solar farm to such events.

Importantly, there are ways to distinguish the events that belong to set $\mathcal{G}$ from the events that belong to set $\mathcal{S}$. In this paper, we assume that we use the method in [4].

Fig. 2 shows an example event that belongs to set $\mathcal{G}$, i.e., an event that is an *external disturbance* to the solar farm. Four different *time series* are shown in this figure for this event:

- Voltage at the PCC,
- Frequency at the PCC,
- Active power injection by the solar farm at the PCC,
- Reactive power injection by the solar farm at the PCC.

Throughout this paper, we use V, f, P, and Q to denote the above four different types of measurements.

Since the event in Fig. 2 is an external disturbance, V and f act as *inputs* to the dynamic system of the solar farm. In response to the external disturbance in these two inputs, the solar farm experiences an agitation in its active power injection and reactive power injection to the power grid. Therefore, P and Q act as *outputs* of the dynamic system of the solar farm. In this regard, the input-output model of the solar farm as a dynamic system can be presented as shown in Fig. 3.

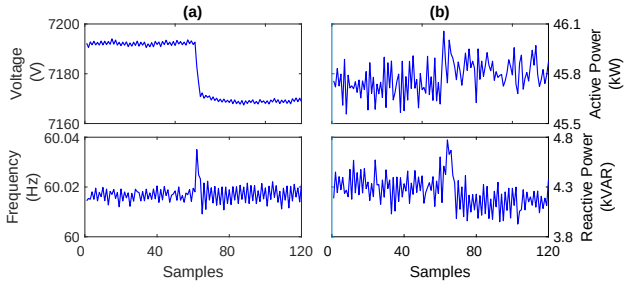

Fig. 2. A real-world example of an event that in an external disturbance to the solar farm: a) the two input time series; b) the two output time series.

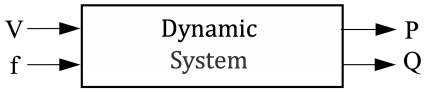

Fig. 3. The input-output relationship in the dynamic system of the farm.

For each event $i \in \mathcal{G}$, i.e., for each external disturbance, we denote the time series corresponding to event i as follows:

$$\left.\begin{aligned} \mathbf{V}_i &= [V_{i,1}, \ldots, V_{i,L_i}]^T \\ \mathbf{f}_i &= [f_{i,1}, \ldots, f_{i,L_i}]^T \end{aligned}\right\} \text{Inputs} \tag{2}$$

and

$$\left.\begin{aligned} \mathbf{P}_i &= [P_{i,1}, \ldots, P_{i,L_i}]^T \\ \mathbf{Q}_i &= [Q_{i,1}, \ldots, Q_{i,L_i}]^T \end{aligned}\right\} \text{Outputs} \tag{3}$$

where L_i denotes the *length* of the time series for event i.

The purpose of this study is to obtain a dynamic model for the solar farm based on the above event-based input-output relationships at the solar farm. Once such dynamic model is obtained, we can predict how the solar farm would respond to any given external disturbance in the power grid.

B. Data Preparation and Preprocessing

In Fig. 2, we can observe two types of patterns in the changes in the time series in this figure. On one hand, we observe some fluctuations in the measurements that are present *even before* the external disturbance occurs. On the other hand, we observe some (more significant) changes in the measurements that are caused *by* the external disturbance. Importantly, only the latter type of changes in the measurements are relevant to the dynamic response of the solar farm. In this regard, the continuous fluctuations that we observe in all the four time series are *not* related to our analysis; and they need to be filtered out. In fact, these high-frequency fluctuations act as noise as far as our analysis is concerned. Therefore, we remove these fluctuations by applying a *low-pass filter* to the measurements. Furthermore, we *normalize* all the four time series to the range between 0 to 1. These two changes in the measurement data are part of our *pre-processing* task.

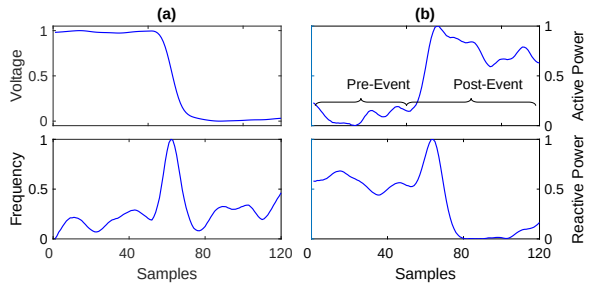

Fig. 4. The time series in Fig. 2 after pre-processing.

Fig. 4 shows the time series of the same event in Fig. 2, but *after applying the pre-processing* steps. Notice that the continuous fluctuations are no longer present. The dynamic behavior of the system is now much more clear.

III. Proposed Dynamic Models

In this section, we present three different data-driven approaches to develop dynamic models to mathematically capture the event-based input-output relationship in Fig. 3.

A. Single Model

In this section, we assume that a single model is capable of capturing the dynamic response of the solar farm under *all* the events in set $\mathcal{G}$, i.e., under *every* external disturbance that we have observed. To construct this single dynamic model, we make use of the measurement time series and linear regression.

In this section, we employ the Autoregressive Moving Average eXogenous (ARMAX) model and estimate its parameters:

$$A(q^{-1})y[k] = B(q^{-1})u[k] + C(q^{-1})e[k], \tag{4}$$

where

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \\ C(q^{-1}) &= c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}. \end{aligned} \tag{5}$$

Here, $u[k]$ denotes the vector of input samples and $y[k]$ denotes the vector of output samples. Here the "samples" are the *combined* measurement samples of *all* the events in set $\mathcal{G}$:

$$u[k] = [\ \mathbf{V}_i\ \mathbf{f}_i\ ], \quad y[k] = [\ \mathbf{P}_i\ \mathbf{Q}_i\ ], \quad \forall i \in \mathcal{G}. \tag{6}$$

The length of vector $u[k]$ is n_u and the length of vector $y[k]$ is n_y. Notation $e[k]$ is the vector of white noise. Notations $A(q^{-1})$, $B(q^{-1})$, and $C(q^{-1})$ are polynomials of delay operator q^{-1}, with polynomial orders n_a, n_b, and n_c, respectively.

We estimate the parameters of the ARMAX model by using the prediction-error method and the polynomial orders specified as $[n_a\ \ n_b\ \ n_c\ \ n_k]$, where n_k is the input-output delay. The model properties include estimation covariances (parameter uncertainties) and the goodness of the fit between the estimation results from the model and the measured data. We denote the resulting ARMAX model as

$$\text{ARMAX}(\mathcal{G}). \tag{7}$$

All the events in set $\mathcal{G}$ are used to develop the above model.

B. Multiple Models with a Model Selection Mechanism

For *each* event $i \in \mathcal{G}$, let us obtain one ARMAX model, denoted by ARMAX(i) that is obtained based on $\mathbf{V}_i$ and $\mathbf{f}_i$ as inputs, and $\mathbf{P}_i$ and $\mathbf{Q}_i$ as outputs. Accordingly, we develop a *library* of $|\mathcal{G}|$ dynamic models, where $|\mathcal{G}|$ denotes the cardinality of set $\mathcal{G}$. Such library can be expressed as

$$\text{ARMAX}(i), \quad \forall i \in \mathcal{G}. \tag{8}$$

Note that, unlike in Section III-A, where we used *all* the events in set $\mathcal{G}$ to develop *one* ARMAX model as in (7), here we use *each* event in set $\mathcal{G}$ to develop a separate ARMAX model; thus, developing *multiple* ARMAX models as in (8).

For *each* event $i \in \mathcal{G}$, let us also obtain a *nonlinear* ARMAX model, denoted by NL-ARMAX(i). A nonlinear ARMAX model uses nonlinear polynomials, see [14] for more details. Accordingly, we can add another $|\mathcal{G}|$ dynamic models to the *library* of dynamic models as follows:

$$\text{NL-ARMAX}(i), \quad \forall i \in \mathcal{G}. \tag{9}$$

For the rest of this section, we develop a mechanism to decide *which* one of the dynamic models in (8) and (9) should be used to estimate the dynamic response of the solar farm.

Suppose, we seek to estimate the dynamic response of the solar farm to a given test event j. We define a *similarity index* between the input time series associated with event j and the input time series associated with an event i in set $\mathcal{G}$:

$$s_{i,j} = f\left([\mathbf{V}_i\ \mathbf{f}_i], [\mathbf{V}_j\ \mathbf{f}_j]\right), \tag{10}$$

where $f(\mathbf{X}, \mathbf{Y})$ is a function that evaluates the similarity between two time series vectors $\mathbf{X}$ and $\mathbf{Y}$.

We consider two choices for function $f(\mathbf{X}, \mathbf{Y})$. The first choice is the Dynamic Time Warping (DTW) function, which is an elastic similarity measure that optimally warps time series vector $\mathbf{X}$ and time series vector $\mathbf{Y}$ in a way that accumulates error of alignments is minimized. This accumulated error is obtained by conducting dynamic programming [15]:

$$f_{\text{DTW}}(\mathbf{X}, \mathbf{Y}) = D(\mathbf{X}, \mathbf{Y}) = D_{m,n}, \tag{11}$$

where we recursively apply the following:

$$\begin{aligned} &D_{v,w} = \\ &\quad (X_v - Y_w)^2 + min\{D_{v,w-1}, D_{v-1,w}, D_{v-1,w-1}\}. \end{aligned} \tag{12}$$

Parameters m and n are the lengths of vectors $\mathbf{X}$ and $\mathbf{Y}$, $D_{v,w}$ is the similarity between entry v of vector $\mathbf{X}$ and entry w of vector $\mathbf{Y}$. The initial condition is $D_{1,1} = (X_1 - Y_1)^2$.

The second choice for function $f(\mathbf{X}, \mathbf{Y})$ is the Pearson correlation (PC). It is obtained as follows [16]:

$$f_{\text{PC}}(\mathbf{X}, \mathbf{Y}) = \rho(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{k=1}^{m}(X_i - \overline{\mathbf{X}})(Y_i - \overline{\mathbf{Y}})}{\sqrt{\sum_{k=1}^{m}(X_i - \overline{\mathbf{X}})^2 (Y_i - \overline{\mathbf{Y}})^2}}, \tag{13}$$

where $\overline{\mathbf{X}}$ and $\overline{\mathbf{Y}}$ are the *mean* over the entries of vectors $\mathbf{X}$ and $\mathbf{Y}$. Pearson correlation is *invariant* under separate changes in location of the entries and the scales of the two time series.

Either one of the above two similarity indexes can be used to identify the event in set $\mathcal{G}$ that is most similar to the testing event of interest. Accordingly, we can obtain the dynamic model corresponding to the select most similar event such that we can estimate the dynamic response of the test event.

In this paper, we use both $f_{\text{DTW}}(\mathbf{X}, \mathbf{Y})$ and $f_{\text{PC}}(\mathbf{X}, \mathbf{Y})$. Suppose i_{DTW} and i_{PC} denote the *most similar* events in the event library, compared to the test event j, based on the DTW and PC similarity indexes, respectively. We have:

$$i_{\text{DTW}} = \underset{i \in \mathcal{G}}{\operatorname{argmax}} \; f_{\text{DTW}}\left([\mathbf{V}_i \; \mathbf{f}_i], [\mathbf{V}_j \; \mathbf{f}_j]\right), \tag{14}$$

$$i_{\text{PC}} = \underset{i \in \mathcal{G}}{\operatorname{argmax}} \; f_{\text{PC}}\left([\mathbf{V}_i \; \mathbf{f}_i], [\mathbf{V}_j \; \mathbf{f}_j]\right). \tag{15}$$

If $i_{\text{DTW}} = i_{\text{PC}}$, then the choice of the event from the library is clear; because we would choose $i^\star = i_{\text{DTW}} = i_{\text{PC}}$. However, if $i_{\text{DTW}} \neq i_{\text{PC}}$, then we would check the output of the test event j that is estimated based on ARMAX(i_{DTW}), NL-ARMAX(i_{DTW}), ARMAX(i_{PC}), and NL-ARMAX(i_{PC}), then we would choose the model that causes the *least maximum violation* of the normalization bounds 0 and 1; which we had discussed as part of the pre-processing task in Section II.B. In this regard, suppose $y_{j,\text{DTW}}[k]$ denotes the estimated output that is obtained by applying the dynamic model ARMAX(i_{DTW}) to the input time series corresponding to the test event j. The violation of the normalization bounds by the estimated response $y_{j,\text{DTW}}[k]$ can be obtained as

$$\Gamma_{j,\text{DTW}} = \max_k \left\{ \left[y_{j,\text{DTW}}[k] - 1\right]^+, \left[0 - y_{j,\text{DTW}}[k]\right]^+ \right\}, \tag{16}$$

where notation $[\cdot]^+$ in (16) is defined as $[x]^+ = \max\{0, x\}$. We similarly define $\Gamma_{j,\text{PC}}$ as the violation of the normalization bound if we use ARMAX(i_{PC}); define $\Gamma_{j,\text{DTW,NL}}$ as the violation of the normalization bound if we use NL-ARMAX(i_{DTW}); and define $\Gamma_{j,\text{PC,NL}}$ as the violation of the normalization bound if we use NL-ARMAX(i_{PC}). In order to obtain the model that we should use, we check the following minimization:

$$\min \left\{ \Gamma_{j,\text{DTW}}, \Gamma_{j,\text{PC}}, \Gamma_{j,\text{DTW,NL}}, \Gamma_{j,\text{PC,NL}} \right\}. \tag{17}$$

For example, if $\Gamma_{j,\text{DTW}}$ is the minimum, then we use ARMAX(i_{DTW}) to estimate the response of the solar farm to test event j. This concludes the process of obtaining the dynamic response when using *multiple* dynamic models.

To improve the modeling accuracy, we also check our estimated response with the observed response of the solar farm once the corresponding measurements are obtained. Accordingly, if the error in the estimation of the dynamic response based on ARMAX($i^\star$) is above an *tolerated threshold*, we can choose to add the test event to the training data set $\mathcal{G}$. This will help us improve our model in the future. The summary of the above steps is given in Fig. 5 in form of a flowchart.

C. Combined Library of Models

Interestingly, there can be a *trade-off* between the *course-grained* single model approach in Section III-A and the *granular* multiple model approach in Section III-B. The multiple

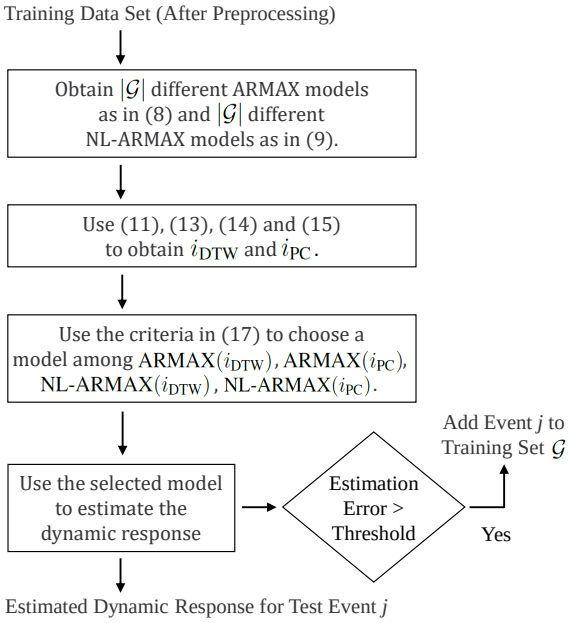

Fig. 5. The flowchart of the process to obtain the dynamic response of the solar farm based on using multiple models; and updating the model library.

model approach works well for the *majority* of the test events. However, *occasionally*, it may demonstrate a less accurate result than the single model. This may happen only when there is *high dissimilarity* between the test event j and *all* the events that are available in the training set $\mathcal{G}$. In such cases, the single model approach may result in a accurate dynamic response.

In this section, we capture the above trade-off by combining the use of the single model approach in Section III-A and the multiple model approach in Section III-B. Suppose $y_{j,\mathcal{G}}[k]$ denotes the estimated output that is obtained by applying the single dynamic model ARMAX($\mathcal{G}$) in (7) to the input time series corresponding to the test event j. Similar to the analysis in the (16), we can obtain the maximum violation of $y_{j,\mathcal{G}}[k]$ with respect to the normalization bounds 0 and 1 as follows:

$$\max_k \left\{ \left[y_{j,\mathcal{G}}[k] - 1\right]^+, \left[0 - y_{j,\mathcal{G}}[k]\right]^+ \right\}. \tag{18}$$

Accordingly, we choose ARMAX($\mathcal{G}$) over the multiple models, if the above violation is less than the minimum in (17).

IV. Case Studies

In this section, we use 450 real-world events, which are all external disturbances, to test the performance of the proposed methods. A total of 400 events are used for *training*, i.e., to form set $\mathcal{G}$. The remaining 50 events, labeled from 401 till 450, are used for *testing*. In all case studies we use $n_a = 2$, $n_b = 3$, and $n_c = 1$. We use the Root Mean Square Error (RMSE) in estimating each output of the solar farm, in terms of active power and active power, by using each model.

1) Results for the Single Model: Fig. 6 shows the results for the single dynamic model. We can see that the estimation error can vary significantly. On average, RMSE is 0.42507 to estimate the time series of the reactive power and 0.44189 to estimate the time series of the active power.

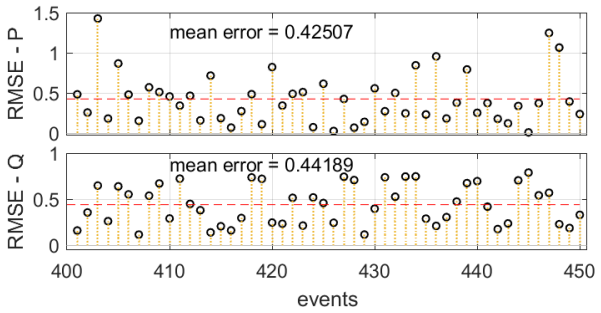

Fig. 6. The estimation errors of the *single* model for the *test* events.

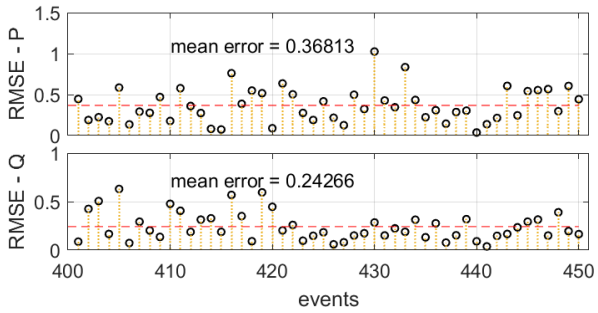

Fig. 7. The estimation errors of the *multiple* models for the *test* events.

2) Results for the Multiple Models: Fig. 7 shows the results for the multiple dynamic models. The estimation error is significantly reduced (i.e., improved) compared to the results in Fig. 6. On average, RMSE is 0.36813 to estimate the time series of the reactive power and 0.24266 to estimate the time series of the active power. Therefore, on average, the performance has improved by 13.4% and 45.1%, respectively.

3) Results for the Combined Models: By comparing the results in Fig. 6 and Fig. 7, we can see that, although the approach based on multiple models is generally much better than the approach based on a single model, there are a number of cases where the single model leads in better results. This supports our argument in Section III-C that it can be beneficial to combine the two approaches. The results are shown in Fig. 8. On average, RMSE is 0.23077 to estimate the time series of the reactive power and 0.18565 to estimate the time series of the active power. Thus, on average, the performance has *further* improved by 37.3% and 23.5%, respectively.

It is worth noting that, the better estimation of the reactive power is due to the strong relationship between the the voltage and the reactive power exchange from the solar farm.

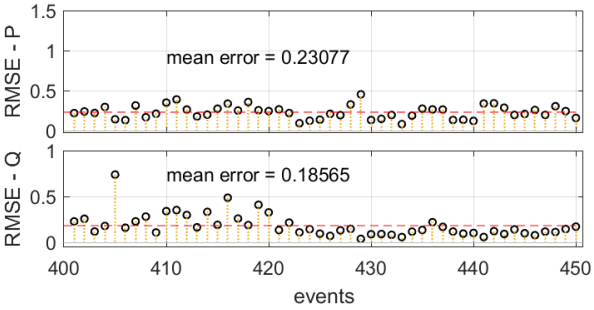

Fig. 8. The estimation errors of the *combined* models for the *test* events.

V. Conclusions

In this study, the dynamic behavior of a real-world behind-the-meter solar farm in California was modeled in response to grid-induced disturbances using data-driven techniques. Three data-driven approaches were proposed: a single dynamic model approach, a multiple dynamic model approach, and a combined approach. The key was to build a *library* of dynamic models that are trained based on real-world events that act as external disturbances to the solar farm's dynamical system. The performance of the proposed method was demonstrated by using real-world case studies. The results in this paper have potential practical applications for the integration of inverter-based resources, including for system-wide planning and operation and for evaluating the interconnection requirements.

References

[1] M. S. Alam, F. S. Al-Ismail, A. Salem, and M. A. Abido, "High-level penetration of renewable energy sources into grid utility: Challenges and solutions," *IEEE Access*, vol. 8, pp. 190277–190299, 2020.

[2] A. Ekic, D. Wu, and J. N. Jiang, "Impact of solar inverter dynamics during grid restoration period on protection schemes based on negative-sequence components," *Energies*, vol. 15, no. 12, p. 4360, 2022.

[3] Pacific Gas and Electric Company, "CA Rule 21 interconnection," tech. rep., California Public Utility Commission, 2020.

[4] P. Khaledian and H. Mohsenian-Rad, "Automated event region identification and its data-driven applications in behind-the-meter solar farms based on Micro-PMU measurements," *IEEE Trans. on Smart Grid*, vol. 13, pp. 2094–2106, May 2022.

[5] M. Abido and M. S. Khalid, "Seven-parameter PV model estimation using differential evolution," *J. of EE Electrical Engineering*, vol. 100, no. 2, pp. 971–981, 2018.

[6] H. Long, S. Xu, X. Lu, Z. Yang, C. Li, J. Jing, and Z. Wu, "Data-driven hybrid equivalent dynamic modeling of multiple photovoltaic power stations based on ensemble gated recurrent unit," *Frontiers in Energy Research*, vol. 8, p. 185, 2020.

[7] H. Cai, J. Xiang, and W. Wei, "Modelling, analysis and control design of a two-stage photovoltaic generation system," *IET Renewable Power Generation*, vol. 10, no. 8, pp. 1195–1203, June 2016.

[8] D. Remon, A. M. Cantarellas, and P. Rodriguez, "Equivalent model of large-scale synchronous photovoltaic power plants," *IEEE Trans. on Industry Applications*, vol. 52, no. 6, pp. 5029–5040, 2016.

[9] M. C. Di Piazza, M. Luna, G. Petrone, and G. Spagnuolo, "Translation of the single-diode PV model parameters identified by using explicit formulas," *IEEE J. of Photovoltaics*, vol. 7, no. 4, pp. 1009–1016, 2017.

[10] D. Ramasubramanian, I. Alvarez-Fernandez, P. Mitra, A. Gaikwad, and J. Boemer, "The new aggregated distributed energy resources (DER_A) model for transmission planning studies: 2019 update," *Electric Power Research Institute (EPRI)*, Washington, DC, USA 2019.

[11] W. Li, P. Chao, X. Liang, J. Ma, D. Xu, and X. Jin, "A practical equivalent method for dfig wind farms," *IEEE Trans. on Sustainable Energy*, vol. 9, no. 2, pp. 610–620, 2017.

[12] P. Li, W. Gu, H. Long, G. Cao, Z. Cao, B. Xu, and J. Pan, "High-precision dynamic modeling of two-staged photovoltaic power station clusters," *IEEE Trans. on Power Systems*, vol. 34, pp. 4393–4407, 2019.

[13] A. Aligholian, A. Shahsavari, E. M. Stewart, E. Cortez, and H. Mohsenian-Rad, "Unsupervised event detection, clustering, and use case exposition in micro-PMU measurements," *IEEE Trans. on Smart Grid*, vol. 12, no. 4, pp. 3624–3636, 2021.

[14] R. B. Mrad, "Non-linear systems representation using ARMAX models with time-dependent coefficients," *Mechanical Systems and Signal Processing*, vol. 16, no. 5, pp. 803–815, 2002.

[15] J. Serra and J. L. Arcos, "An empirical evaluation of similarity measures for time series classification," *Knowledge-Based Syst.*, vol. 67, pp. 305–314, Sep. 2014.

[16] C. Cassisi, P. Montalto, M. Aliotta, A. Cannata, and A. Pulvirenti, "Similarity measures and dimensionality reduction techniques for time series data mining," *Advances in data mining knowledge discovery and applications, InTech*, pp. 71–96, Rijeka, Croatia 2012.