A Hybrid Data-Driven and Model-Based Anomaly Detection Scheme for DER Operation

Preprint

Yiyun Yao, Fei Ding, and Weijia Liu

National Renewable Energy Laboratory

Presented at the IEEE PES Innovative Smart Grid Technologies Conference (ISGT NA)
New Orleans, Louisiana
April 24–28, 2022

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-5D00-80628
May 2022

A Hybrid Data-Driven and Model-Based Anomaly Detection Scheme for DER Operation

Preprint

Yiyun Yao, Fei Ding, and Weijia Liu

National Renewable Energy Laboratory

Suggested Citation
Yao, Yiyun, Fei Ding, and Weijia Liu. 2022. *A Hybrid Data-Driven and Model-Based Anomaly Detection Scheme for DER Operation: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-80628.
https://www.nrel.gov/docs/fy22osti/80628.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-5D00-80628
May 2022

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

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. This research was supported by the Grid Modernization Initiative of the DOE as part of its Grid Modernization Laboratory Consortium, a strategic partnership between DOE and the national laboratories to bring together leading experts, technologies, and resources to collaborate on the goal of modernizing the nation's grid. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

A Hybrid Data-Driven and Model-Based Anomaly Detection Scheme for DER Operation

Yiyun Yao, Fei Ding, Weijia Liu

National Renewable Energy Laboratory
Denver, CO, U.S.A
e-mail: {yiyun.yao, fei.ding, weijia.liu}@nrel.gov

***Abstract*—This paper proposes a hybrid data and model-based anomaly detection scheme to secure the operation of distributed energy resources (DERs) in distribution grids. Data-driven autoencoders are set up at the edge device level and they use local DER operational data as inputs. The abnormal statuses are detected by analyzing reconstruction errors. In parallel, model-based state estimation (SE) is set up at the central level and it uses system-wide models and measurements as data inputs. The anomalies are identified by analyzing measurement residuals. The hybrid scheme preserves the benefits of both data-driven and model-based analyses and thus improves the robustness and the accuracy of anomaly detection. Numerical tests based on the model of a real distribution feeder in Southern California highlight the proposed scheme's effectiveness and benefits.**

Keywords—Anomaly detection, state estimation, largest normalized residual, autoencoder

I. Introduction

Because of efforts to decrease carbon emissions, the penetration of distributed energy resources (DERs) has increased consistently during the past decade. This has led to the development of DER management systems (DERMS), which integrate the needs of utility grid operations with the capabilities of demand-side DERs to support multiple objectives related to distribution system operations, end customer value, and market participation. Deployments of DERMS will lead the grid to increasingly depend on the security of the cyberspace infrastructure to provide DER monitoring, protection, and control capabilities.

Recent findings documented in government reports [1], however, indicate that the threat of cyber-based attacks is increasing in both the number and the sophistication of targets toward the U.S. electric grid. A major cyber incident in a DERMS could have severe consequences on the operation of the DERs and could result in blackouts, equipment damage, or market impacts. Therefore, to ensure distribution system reliability, stability, security, and resilience, it is crucial to monitor the operating states of DERs and to detect anomalies quickly to avert disturbances and disruptions.

Most work on anomaly detection can be divided into two categories: model-based and data-driven methods. The key idea of model-based methods is to compare the expected system behavior, estimated by a model, to the actual behavior when the system is in a specific state. These applications range from the diagnosis of switching converters to short-circuit detection [2]. The limitations of using model-based methods in distribution systems include the lack of accurate system models (especially for secondary feeders) and the lack of measurements that support the models' observability into the system.

Data-driven approaches use machine learning methods to conduct statistical inference or decision making based on the available data. The key idea is to extract features (including the voltage and system frequency [3] and its deviation [4]) from the measurements data and to classify those with similar features resulting from the same event type. One limitation of data-driven methods is that the computational complexity generally increases when the data set size increases. Moreover, most classifiers are trained from labeled data (i.e., through supervised learning), whereas DER diagnosis and event-labeling can be expensive and scant [5].

To overcome the above shortcomings, this paper proposes a hybrid data-driven and model-based scheme to detect abnormal DER operation in distribution grids. Data-driven autoencoders are set up at the edge device level and they use local DER operational data as inputs. Auto-encoders detect the anomalous operating status of DERs by analyzing reconstruction errors. Meanwhile, model-based state estimation (SE) is set up at the central level and it uses system-wide models and measurements as data inputs. The anomalous data and their location are identified by analyzing the measurement residuals.

This paper contributes the following:

(1) A dimension reduction-based autoencoder is designed in two parts and deployed in a decentralized manner for anomaly detection. No communication is required among the inverters. No labeling is required for the data.

(2) The proposed hybrid scheme retains the advantages of both data-driven and model-based analyses and thus improves the robustness and the accuracy of anomaly detection.

(3) Without intensive investment in hardware, the proposed hybrid anomaly detection scheme can be established using existing the sensing, communication, and control infrastructures in distribution grids.

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. This research was supported by the Grid Modernization Initiative of the DOE as part of its Grid Modernization Laboratory Consortium, a strategic partnership between DOE and the national laboratories to bring together leading experts, technologies, and resources to collaborate on the goal of modernizing the nation's grid. 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.

II. Background

This section describes the SE and autoencoder approaches that are used in this paper.

A. Weighted Least-Squares-Based State Estimator

Given a network model and system measurements—such as nodal power injection, voltage magnitude, branch power flow, and current magnitude—from supervisory control and data acquisition (SCADA) systems, SE determines the optimal estimate for the system state, which comprises complex bus voltages in the entire power distribution system [6]. The mathematical definition of SE is:

$$\boldsymbol{z} = h(\boldsymbol{x}) + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{e}$ are the system measurement, state, and error vector, respectively; and $h(\cdot)$ is the nonlinear function relating measurement to the state vector. Based on the statistical properties of the errors, two assumptions are commonly made: (1) measurement errors are normally distributed with zero mean, i.e., $E(e_i) = 0$; and (2) measurement errors are independent, i.e., $E(e_i e_j) = 0$. Hence, $Cov(\boldsymbol{e}) = E[\boldsymbol{e} \cdot \boldsymbol{e}^{\mathsf{T}}] = \boldsymbol{R}$.

Different SE techniques have been developed. The most widely used is the weighted least-squares (WLS) approach, which minimizes the following objective function:

$$J(\boldsymbol{x}) = [\boldsymbol{z} - h(\boldsymbol{x})]^{\mathsf{T}} \boldsymbol{R}^{-1} [\boldsymbol{z} - h(\boldsymbol{x})]. \tag{2}$$

At the minimum (local optimal), the first-order optimality conditions will need to be satisfied, expressed as:

$$g(\boldsymbol{x}) = \frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}} = -H^{\mathsf{T}}(\boldsymbol{x}) \boldsymbol{R}^{-1} [\boldsymbol{z} - h(\boldsymbol{x})] = 0, \tag{3}$$

where $H(\boldsymbol{x}) = [\partial h(\boldsymbol{x}) / \partial \boldsymbol{x}]$. To solve these equations, one can expand the nonlinear function $g(\boldsymbol{x})$ into the Taylor series. Neglecting the terms higher than the second order leads to an iterative solution scheme called the Gauss-Newton method [6] to determine the estimated state $\hat{\boldsymbol{x}}$.

B. Autoencoder

Unlike the model-based WLS-SE, an autoencoder is a special type of neural network whose objective is to achieve an identity mapping between its inputs and outputs. It works by compressing the inputs into a latent-space representation (i.e., encoding) and then reconstructing the outputs from this representation (i.e., decoding) [7]. As shown in Fig. 1, an autoencoder consists of (1) a d-dimension vector as the visible layer that collects the input, denoted by $\boldsymbol{u} = [u_1, \dots, u_d]^{\mathsf{T}}$; (2) a reconstructed vector as the reconstruction layer that shares the same dimension as the input vector, denoted by $\hat{\boldsymbol{u}} = [\hat{u}_1, \dots, \hat{u}_d]^{\mathsf{T}}$; and (3) one or more hidden layers, or latent representations, that aim to learn a pattern in the inputs.

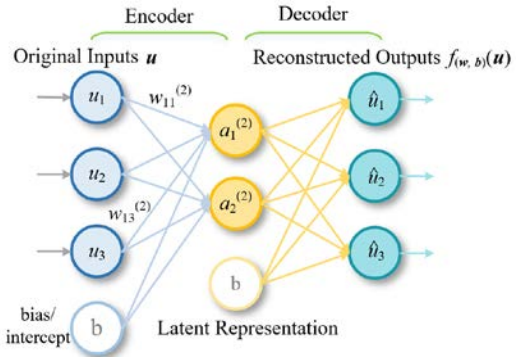

Fig. 1. An autoencoder neural network.

Let us index all the layers in the autoencoder by $l = 1, 2, \dots, L$, and denote the number of neurons in the l-th layer by $n^{(l)}$. The value of the neuron i in the layer l, denoted by $a_i^{(l)}$, is calculated by:

$$a_i^{(l)} = sig\left(\sum_{j=1}^{n^{(l-1)}} w_{ij}^{(l-1)} a_j^{(l-1)} + b_i^{(l-1)}\right), \tag{4}$$

Where $sig(\alpha) = 1/(1 - e^{-\alpha})$ is the sigmoid activation function; $w_{ij}^{(l-1)}$ is the weight associated with the connection from neuron j in layer $l-1$ to neuron i in layer l; and $b_i^{(l-1)}$ is the bias/intercept associated with neuron i in layer l. The final output, $\hat{\boldsymbol{u}}$, is then $\boldsymbol{a}^{(L)}$, and it is denoted as $f_{(\boldsymbol{w},\boldsymbol{b})}(\boldsymbol{u})$.

Given T training samples, $\{\boldsymbol{u}[1], \boldsymbol{u}[2], \dots, \boldsymbol{u}[T]\}$, the objective function of the autoencoder is to minimize the cost:

$$\begin{aligned} Cost(\boldsymbol{w}, \boldsymbol{b}) := & \frac{1}{T} \sum_{i=1}^{T} \left(\frac{1}{2} \left\| f_{(\boldsymbol{w},\boldsymbol{b})}(\boldsymbol{u}[i]) - \boldsymbol{u}[i] \right\|^2\right) \\ & + \frac{\lambda}{2} \sum_{l=1}^{L-1} \sum_{j=1}^{n^{(l)}} \sum_{i=1}^{n^{(l)}} \left(w_{ij}^{(l-1)}\right)^2. \end{aligned} \tag{5}$$

The first term defines the reconstruction error to the inputs, and the second term is a regularization term to prevent overfitting. Minimizing this function can be achieved by using a back-propagation algorithm [8].

Suppose the dimension of the hidden layers is less than that of the visible layer. In that case, the establishment of these hidden layers aims to extract the essential information of the inputs with affordable loss. Further, because the output is set to be as equal as possible to the input, the training samples are automatically obtained by setting $\hat{\boldsymbol{u}} = \boldsymbol{u}$, which is why the autoencoder is considered an unsupervised learning model.

III. Hybrid Data and Model-Driven Anomaly Detection Scheme

This section explains the proposed hybrid scheme and how to integrate the SE-based and autoencoder approaches.

A. Largest Normalized Residual (LNR)

A SE-based approach can be used to detect and identify anomalous measurement/data. When using WLS-SE, anomaly detection is based on the analysis of the residual properties [6]. The estimated value of Δz (denoted as $\Delta \hat{\boldsymbol{z}}$) is given as:

$$\Delta \hat{\boldsymbol{x}} = \boldsymbol{G}^{-1} \boldsymbol{H}^{\mathsf{T}} \boldsymbol{R}^{-1} \Delta \boldsymbol{z} \tag{6}$$

$$\Delta \hat{\boldsymbol{z}} = \boldsymbol{H} \Delta \hat{\boldsymbol{x}} = \boldsymbol{K} \Delta \boldsymbol{z}, \tag{7}$$

where $\boldsymbol{K} = \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{R}^{-1}$, $\boldsymbol{G} = \boldsymbol{H}^{\mathsf{T}}\boldsymbol{R}^{-1}\boldsymbol{H}$ is the gain matrix, $\Delta \hat{\boldsymbol{x}}$ is the last update on $\Delta \boldsymbol{x}$ that reaching the convergence in Gaussian-Newton method. The residual vector, $\boldsymbol{r}$, is defined as the difference between the measured values and the estimated values. Based on the property of the $\boldsymbol{K}$ matrix (i.e., $(\boldsymbol{I} - \boldsymbol{K}) \cdot \boldsymbol{H} = 0$), one can derive that:

$$\boldsymbol{r} = \Delta \boldsymbol{z} - \boldsymbol{H} \Delta \hat{\boldsymbol{x}} = (\boldsymbol{I} - \boldsymbol{K}) \boldsymbol{e} \tag{8}$$

The mean and covariance of the measurement residuals can be derived as $E(\boldsymbol{r}) = (\boldsymbol{I} - \boldsymbol{K}) \cdot E(\boldsymbol{e}) = 0$ and $Cov(\boldsymbol{r}) = \Omega = (\boldsymbol{I} - \boldsymbol{K}) \cdot \boldsymbol{R}$, therefore, $r \sim N(0, \Omega)$. One can use the residual properties to formulate a test to identify bad measurement data [6]. The normalized residual of measurement i can be calculated as:

$$r_i^N = |r_i| / \sqrt{\Omega_{ii}} \tag{9}$$

If $r_i^N > thld$ ($thld$ is a chosen identification threshold, e.g., 3 for a Gaussian distribution), then measurement i will be suspected as anomalous data. The challenge of implementing the LNR for DER situational awareness in distribution grids is twofold: (1) The LNR method relies on knowledge of an accurate system model. But most DERs are connected to the grid edge (e.g., behind the meter), and the secondary models are

usually unknown; thus, it is impractical for WLS-SE and LNR to include states that directly relate to each DER's operational status. (2) Most existing SCADA systems cannot yet obtain DER operational data. The limited measurement redundancy further reduces the capability of LNR to detect anomalies.

B. Dimension Reduction-Based Autoencoder

Different from LNR, using an autoencoder for anomaly detection is a data-driven implementation. We use a dimension reduction-based autoencoder that attempts to find an optimal latent space where the normal and abnormal data appear to be different. Because the normal data in the test data set meet the normal profile, the corresponding error is smaller, whereas the abnormal data will have a relatively higher reconstruction error. As a result, the autoencoder can be applied to detect anomalous data by analyzing the reconstruction error.

The scheme for implementing autoencoders for anomaly detection at the DER inverters is presented in Fig. 2. Each inverter runs one copy of the autoencoder and executes two tasks: (1) performs anomaly detection, which is done independently with local DER data—such as voltage, active and reactive power—and without communications with other inverters or the cloud/server; (2) provides the local DER data, reconstructed copy, and reconstruction error $(\boldsymbol{u}, \hat{\boldsymbol{u}}, \boldsymbol{\xi})$ during the reconstruction to the cloud/server as training data. The cloud/server trains the autoencoder model using the data set $(\boldsymbol{u}, \hat{\boldsymbol{u}})$. Then it sends the updated hyperparameter $(\boldsymbol{w}, \boldsymbol{b})$ back to the inverter to update the autoencoder. If the autoencoder identifies anomalous data by analyzing the reconstruction errors, the inverter can take actions, such as alarming DER owner or utility control center. The implementation consists of two parts—one resides on the inverter, and another resides on the

TABLE I PSEUDO-CODE OF INVERTER ALGORITHM

	Inverter Algorithm: reconstruct local DER data to detect anomaly
1	**for** $t \leftarrow 1$ **to** ∞ **do:**
2	Obtain the inverter readings, $\boldsymbol{u}[t]$;
3	Provide $\boldsymbol{u}[t]$ to autoencoder to reconstruct $\hat{\boldsymbol{u}}[t]$;
4	Obtain the reconstruction error by $\boldsymbol{\xi}[t] = \|\boldsymbol{u}[t] - \hat{\boldsymbol{u}}[t]\|_2$;
5	Determine the anomaly by: $\boldsymbol{\theta}[t] = \begin{cases} \mathbf{0}, & if\ \|\xi[t] - \mu\| \le \rho\sigma \\ \mathbf{1}, & otherwise \end{cases}$, where $\boldsymbol{\rho}$ sets the threshold, e.g., $\boldsymbol{\rho} = \mathbf{3}$ for error in Gaussian;
6	Alarm local agent or control center if $\boldsymbol{\theta}[t] = \mathbf{1}$;
7	**if** t *mod* $T = 0$, **then**:
8	Send $\{\boldsymbol{u}[t: t+T]\}, \{\hat{\boldsymbol{u}}[t: t+T]\}, \{\boldsymbol{\xi}[t: t+T]\}$ to cloud/server;
9	Receive the updated $(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\mu}, \boldsymbol{\sigma})$ from clould/server;
10	**End**

TABLE II PSEUDO-CODE OF CLOUD/SERVER ALGORITHM

	Cloud/Server Algorithm: update the autoencoder model
1	**for** $t \leftarrow 1$ **to** ∞ **do:**
2	**if** t *mod* $T = 0$, **then**:
3	Receive $\{\boldsymbol{u}[t: t+T]\}, \{\hat{\boldsymbol{u}}[t: t+T]\}, \{\boldsymbol{\xi}[t: t+T]\}$ from inverter;
4	Update $(\boldsymbol{w}, \boldsymbol{b})$ by minimizing (5);
5	Calculate $\boldsymbol{\mu} = \frac{1}{T}\sum_t \boldsymbol{\xi}[t]$;
6	Calculate $\boldsymbol{\sigma}^2 = \frac{1}{T}\sum_t (\boldsymbol{\xi}[t] - \boldsymbol{\mu})^2$;
7	Send the updated $(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\mu}, \boldsymbol{\sigma})$ to inverter;
8	**end**

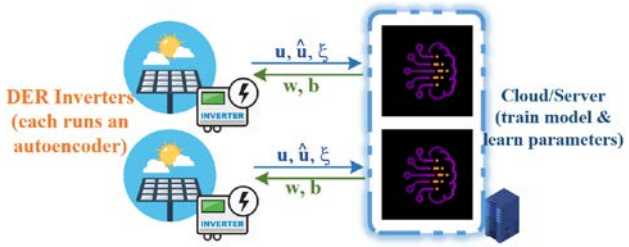

Fig. 2. Using autoencoders for anomaly detection in DER inverters

cloud/server—and it is presented as the pseudo-code in the inverter and cloud/server algorithm in Table I and II.

The autoencoder offers the following advantages:

(1) Decentralized structure. The use of a dimension reduction-based autoencoder for anomaly detection is designed in two parts and deployed in a decentralized structure. No communication is required among inverters.

(2) Low requirements on the data and computational burden. The local DER operational data are not required to be labeled. Both direct DER operational data and their deviations can be used to capture the latent representation and allow anomaly detection [10]. Further, the algorithm on inverters involves only simple matrix dot products to match the limited computational resources at inverters and enable the online application.

(3) Flexibility. Establishing the cloud/server is unnecessary if the local inverter can handle the relatively more computationally intense updates in (5). In addition, separate one-on-one cloud/server connections are also not necessary if the DER data share very similar patterns, e.g., photovoltaics (PV) on the same rooftop. Finally, the communication between the inverters and the cloud/servers can happen at a much lower (and configurable) frequency.

C. Architecture of the Hybrid Anomaly Detection Scheme

The architecture of the hybrid data-driven and model-based anomaly detection scheme for DER operation in distribution grids is illustrated in Fig. 3. Essentially, the proposed hybrid scheme consists of:

(1) A centralized model-based SE is set up at the central level and it uses system-wide models and measurements from existing SCADA points as data inputs. The LNR-based anomaly detection module identifies the anomalous data and their location by analyzing the residuals.

(2) Decentralized data-driven autoencoders are set up at the edge device level and they use local DER operational data (or any local measurements, if available) as inputs. The abnormal

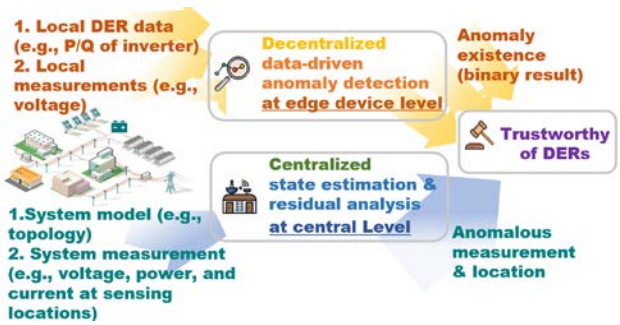

Fig. 3. The architecture of the hybrid anomaly detection scheme.

TABLE III. TRUE AND BAD VALUE AND NORMALIZED RESIDUAL USED BY LNR

	V mag. meas. (p.u.)	P inj. meas. (kW)	Q inj. meas. (kVar)	P fl. meas. (kW)
True value	1.0271	4.43	0.79	3.46
Bad value	1.0098	2.42	2.97	6.42
Normalized residual	8.6914	11.73	9.20	7.58

operations of DERs will be detected by analyzing the reconstruction errors.

The operational status of the DERs will be judged as suspicious when either of the two mechanism alarms, as abnormal when both mechanisms alarm, and normal only when both mechanisms report normal.

The model-based and data-driven anomaly detection mechanisms operate using various types of data, at disparate time frames, on different levels, independently and in parallel. The hybrid scheme builds on the strengths of both the data-driven and model-based methods, and each addresses the shortcomings of the other, thus improving the robustness and the accuracy of the anomaly detection. The proposed hybrid anomaly detection scheme can be established using existing sensing, communication, and control infrastructures in distribution grids, i.e., without intensive investment in hardware.

IV. NUMERICAL TESTS

To demonstrate that the proposed scheme can reliably detect abnormal DER operation in distribution systems, we build a numerical simulation based on a real feeder in Southern California. The network model consists of 7,309 nodes, including both primary and secondary circuits; has a peak load of 16,651 kW; and is operated at 12 kV. To simulate a case with high PV penetration, we randomly select 60% of customers (1,053 of 1,755 load nodes) to be equipped with one or more PV systems, and the size of total PV systems at each node ranges from 0.2 to 0.8 times its nodal peak load value. In total, 1,559 PV systems are modeled in the feeder, with a total capacity of 8,404 kW (50% peak power penetration). To conduct time-series simulation, the real load data of this feeder collected in the field for the year 2019 at 5-min resolution are applied to the load models. Since no PV measurements are available from the field, and the autoencoders need high-resolution data, we use PV inverter data obtained from hardware-in-the-loop experiments at the National Renewable Energy Laboratory [11] to model the time-series power outputs of the PV systems. These data are at 500-ms resolution for 2 days. A two-day time-series simulation with 5-min time step is conducted using OpenDSS [12] to generate power flow data, which are then used as data inputs by the SE. Since the autoencoder measures data locally, we assume that PV power outputs with 500-ms temporal resolution can be directly collected and used by the autoencoder as data inputs.

A. *Effectiveness of LNR Analysis*

The performance of the LNR is validated by adding a gross error to the measurements that are used by the SE and for processing the residual analysis. In the test, the voltage and power measurements are assumed to be available at the substation. Additional meters are assumed to be available at the distribution devices, such as transformers and capacitor banks, and that they can collect voltage and power measurements. The nodes without any load or PV are modeled to provide virtual measurements. The remaining nodes are assumed to deliver pseudo-measurements. Noise following normal distribution is added, and the standard deviation for different types of measurements is obtained using the criteria given in [13]. The SE and LNR analysis are executed for 576 snapshots (2 days with 5-min resolution).

To simulate the anomalous data, a gross error with magnitudes ranging from 5% to 30% (5% step size) is added to a randomly selected 2% of the voltage magnitude (12 of 600), 1% of the power injection (18 of 1,770 pairs), and 2% of the real power flow (12 of 585) measurements at each snapshot.

Table III reports the true value, bad value, and normalized residual for four measurements (real and reactive power injection measurements on a certain node, voltage magnitude on a certain node, and real power flow measurement on a certain branch) at the first snapshot. For instance, the voltage magnitude is manipulated from 1.0271 to 1.0098, and this anomaly is captured with 8.69 as the normalized residual. Similar observations hold for the other measurements. Fig. 4 reports the average normalized residual for all four types of measurement during all snapshots with different magnitudes of the gross errors. Because of the small standard deviation of error in voltage magnitude measurement, even very small anomalies can be identified by the LNR efficiently, e.g., on average, a 5% error leads to 24.4 in the normalized residual. Only 5%–10% of the reactive power injection errors and 5% of the branch real power flow errors were not captured by the LNR module because their corresponding normalized residual is smaller than the threshold that was set at 3. These are hard to detect with the LNR mainly because of the low fidelity of the pseudo-measurements.

B. *Effectiveness of Autoencoder*

The performance of the model-free anomaly detection scheme using the dimension reduction-based autoencoder is validated with the smart inverter data. For each inverter, the terminal real and reactive power data for 30 second at 500-ms resolution (i.e., input dimension $d = 120$) are fed into the reconstruction as a datum. The hyperparameters are updated every 10 hours (i.e., 1,200 datum). A seven-layer autoencoder is built with TensorFlow [14]. The structure is 120-30-16-8-16-30-120, which is determined by k-fold cross-validation.

Synthetic anomalies are created by adding errors generated by the spike and burst models defined in [15]. The magnitude of the spike and burst error is varied and follows a normal distribution. Five spike or burst errors are randomly added to each datum. The performance of the autoencoder is indicated by

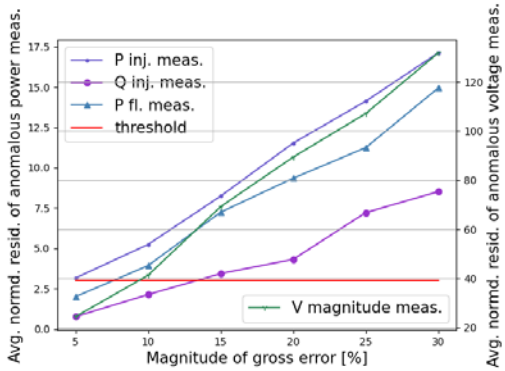

Fig. 4. The average normalized residual for manipulated measurements with different magnitudes of gross errors.

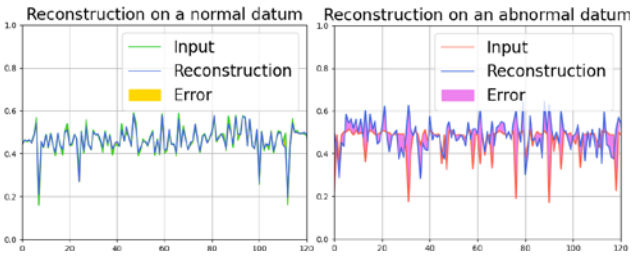

Fig. 5. The reconstruction performance of the autoencoder.

the area under the curve (AUC) of the receiver operating characteristic (ROC) curves [16], which is commonly used to assess the performance of a binary classifier. A good classifier has an AUC of ROC near 1, and a bad classifier has a value near 0, and 0.5 corresponds to a random guessing classifier.

The autoencoder model is prevalidated by checking the reconstruction error. Fig. 5 shows the reconstruction of a normal and an abnormal datum on the terminal data from a certain PV inverter. It can be observed that the reconstructed $\hat{\boldsymbol{u}}$ almost coincides with the original $\boldsymbol{u}$; the area of their difference (marked with yellow) is near zero. But the abnormal datum can be characterized by the large reconstruction error (marked by pink). The heat map of the average AUC of ROC for all inverters in a 2-day test is shown in Fig. 6. It can be observed that AUC is greater than 0.8 for most error means and variances, which indicates a good classifier. The AUC is only low (i.e., 0.5–0.8) when the anomaly is barely notable (when the mean and variance are smaller than 0.02 and 0.1, respectively).

C. Case Study on Malicious Cyberattacks

To validate the benefits of the proposed hybrid anomaly detection scheme, we designed and launched two types of malicious cyberattacks. For these tests, we focused on one load node with three PV systems connected. For the SE (at the central level and on the snapshot of hour 13:00, second day), the nodal power injection (aggregated by three PV power outputs and load power) is obtained as a measurement. The data and model of each PV unit behind the node are unknown to the SE. The autoencoder is deployed to be running on each PV system (at the edge device level and on the first half of the minute including the snapshot) with only inverter terminal data.

Case A: False data injection attack. The real power output of one PV unit is increased by 5 kW with a positive spike error, and another is decreased by 5 kW with a negative spike error. The SE failed to detect the attack because the nodal power injection stays the same and is subject to the power flow principle. However, the autoencoder captured both anomalies because they led to high reconstruction errors in corresponding datum.

Case B: Delay of communication attack. The inverter terminal data of three PV systems are replaced with their data from 3 hours ahead. In this case, the autoencoder failed to detect the attack because all the data follow the latent pattern, and the reconstruction errors are small. Whereas the SE marked the nodal power injection measurement as suspicious because it is not subject to the power flow principle, and the normalized residual is higher than 3.

Both cases would result in the status of these DERs being indicated as suspicious to higher-level controllers and/or system operators.

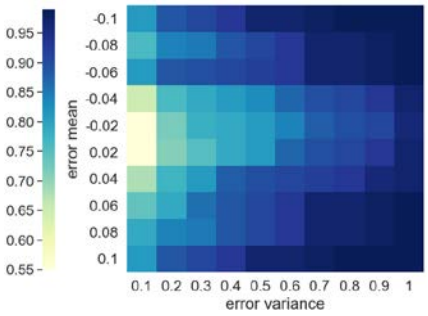

Fig. 6. A heat map of the AUC of ROC with different error mean and variance.

V. Conclusions

This paper proposes a hybrid data-driven and model-based scheme to detect anomalous DER operation in distribution grids. Autoencoders are set up at the edge devices and they use local DER operational data as inputs. The anomalous DER operations are detected by analyzing the reconstruction errors. A model-based SE module is set up at the central level and it uses system models and measurements as data inputs. The anomalous data and their locations are identified by analyzing the measurement residuals. Our preliminary numerical results suggest that the proposed hybrid scheme successfully combines the benefits of both data-driven and model-based analyses and thus can improve the robustness and the accuracy of anomaly detection.

References

[1] U.S. Government Accountability Office, "Electricity Grid Cybersecurity," 2021.

[2] X. Ding, J. Poon, I. Celanovic, and A. D. Dominguez-Garcia, "Fault detection and isolation filters for three-phase AC-DC power electronics systems," *IEEE Trans. Circuits Syst.*, vol. 60, no. 4, pp. 1038–1051, 2013.

[3] O. P. Dahal, S. M. Brahma, and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Trans. Power Deliv.*, vol. 29, no. 3, pp. 1390–1397, 2014.

[4] A. Bykhovsky and J. H. Chow, "Power system disturbance identification from recorded dynamic data at the Northfield substation," *Int. J. Electr. Power Energy Syst.*, vol. 25, no. 10, pp. 787–795, 2003.

[5] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Partial Knowledge Data-driven Event Detection for Power Distribution Networks," *IEEE Trans. Smart Grid*, vol. 3053, no. c, pp. 1–1, 2017.

[6] A. Abur and A. Exposito, *Power system state estimation: theory and implementation*. CRC Press, 2004.

[7] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science (80-.).*, vol. 313, no. 5786, pp. 504–507, 2006.

[8] B. Widrow and M. A. Lehr, "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation," *Proc. IEEE*, vol. 78, no. 9, pp. 1415–1442, 1990.

[9] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Partial Knowledge Data-driven Event Detection for Power Distribution Networks," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 1–1, 2017.

[10] J. Wang, D. Shi, Y. Li, J. Chen, H. Ding, and X. Duan, "Distributed Framework for Detecting PMU Data Manipulation Attacks With Deep Autoencoders," *IEEE Trans. Smart Grid*, vol. PP, no. c, p. 1, 2018.

[11] J. Wang *et al.*, "Performance evaluation of distributed energy resource management via advanced hardware-in-the-loop simulation," *2020 IEEE PES ISGT 2020*, pp. 2–6, 2020.

[12] R. C. Dugan, "OpenDSS manual," 2016.

[13] Y. Yao, X. Liu, D. Zhao, and Z. Li, "Distribution System State Estimation: A Semidefinite Programming Approach," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4369–4378, 2019.

[14] J. Dean and R. Monga, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2015.

[15] S. Reece, S. Roberts, C. Claxton, and D. Nicholson, "Multi-Sensor Fault Recovery in the Presence of Known and Unknown Fault Types," pp. 1695–1703, 2009.

[16] T. Fawcett, "ROC Graphs: Notes and Practical Considerations for Data Mining Researchers ROC Graphs : Notes and Practical Considerations for Data Mining Researchers," *HP Inven.*, p. 27, 2003.