



Title	A data driven measurement placement to evaluate the well being of distribution systems operation
Authors(s)	Jafarian, Mohammad, Soroudi, Alireza, Keane, Andrew
Publication date	2021-01-05
Publication information	Jafarian, Mohammad, Alireza Soroudi, and Andrew Keane. "A Data driven Measurement Placement to Evaluate the Well being of Distribution Systems Operation." Institution of Engineering and Technology, January 5, 2021. https://doi.org/10.1049/gtd2.12114 .
Publisher	Institution of Engineering and Technology
Item record/more information	http://hdl.handle.net/10197/12586
Publisher's version (DOI)	10.1049/gtd2.12114

Downloaded 2026-05-01 23:43:27

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

A data-driven measurement placement to evaluate the well-being of distribution systems operation

Mohammad Jafarian  | Alireza Soroudi  | Andrew Keane 

School of Electrical and Electronic Engineering,
University College Dublin, Dublin, Ireland

Correspondence

Mohammad Jafarian, School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland.
Email: mohammad.jafarian@ucd.ie

Funding information

Science Foundation Ireland, Grant/Award Number: SFI/16/1A/4496

Abstract

The widespread integration of intelligent electronic devices has facilitated the employment of data mining methods in evaluating the operating condition of distribution systems. This possibility comes to prominence in active networks, where distributed energy resources can cause unforeseen dynamics that requires an effective monitoring infrastructure and a fast-track procedure to convey the system operating condition in a comprehensible manner to the operator. To this end, a data-driven approach is proposed to assess the status of system operating constraints by presenting each constraint as a classification problem. Afterwards, by exploiting the propounded presentation of the system operating condition, the measurement placement problem in distribution systems is addressed as selecting a set of features that have the most contribution to evaluating the system operating status. To do so, first, the effectiveness of the measurement units is identified through their contribution to the classification process, and then a procedure is proposed to pinpoint the measurement units with redundant information. Monte-Carlo simulations are performed to provide a comprehensive training set. Receiver operating characteristic analysis and time-series power flows demonstrate the effectiveness of the proposed approaches.

1 | INTRODUCTION

To ensure a secure operation, the operating condition of a distribution system should be monitored regularly. Generally, state estimation (SE) techniques are employed for this purpose, by defining a set of variables as the system states and minimizing the variance in the estimation error [1]. While these techniques have been successful in transmission systems, in practice, they have not been widely applied in distribution management systems due to the lack of sufficient measurements [2]. Nowadays, with the widespread integration of intelligent electronic devices in distribution systems and the development of the smart grid, more real-time measurements are available [2], which enables the distribution system operator to benefit from a wide variety of data mining methods in evaluating the well-being of the system. With the employment of these methods, it is possible to provide intelligent but yet fast-track models to evaluate and convey the system operating condition in a comprehensible manner. This advantage is especially beneficial in active networks, considering that the integration

of distributed energy resources (DERs) may cause unforeseen dynamics [3] that requires prompt responses from the operator [4]. Since the secure operating conditions of a distribution system are predefined in the grid code, among the various data mining methods, classification techniques are especially of interest.

Up to now, the application of classification techniques has been widely examined to identify the operating condition of transmission systems, either by mapping this issue onto a classification problem, or employing the classification techniques to model the system security rules. On this subject, discriminant analysis (DA), core vector machine, and support vector machine techniques have been used to recognize the insecure operation of a power system from the transient stability viewpoint in [5–7], respectively, while in [8], the performance of different classification techniques, i.e. decision tree, ensemble decision tree, and support vector machine, have been compared for this purpose. The time-series shapelet method, combined with the decision tree technique, has been used to assess the short-term voltage stability in [9]. In [10], DA together with

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Generation, Transmission & Distribution* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

a neural network, has been examined to evaluate the system security. In [11] and [12], the decision tree technique has been employed to model the power system security rules, while in [13], deep auto-encoders have been deployed to address this problem.

Unlike transmission networks, in distribution systems, the application of data mining methods in evaluating the system operating condition has been mostly confined to increasing the accuracy of the SE, either by providing a better estimation of the loads [14–16], or finding a better initialization point [17]. Only a few studies have been conducted to employ these methods to directly address the condition assessment issue. As an example, in [18], a neural network has been used to directly estimate the voltage profile of the medium voltage (MV) and low voltage (LV) sections, or in [19], six distinct states have been defined for the system and the hierarchical classification technique has been employed to distinguish among these states.

In addition to the evaluation procedure, an effective monitoring infrastructure is necessary to provide the necessary data required to assess the system operating condition. These data could be from the smart devices such as DERs inverters, or the voltage/current/flow measurement units connected by means of voltage/current transformers to the grid. These data could also be the consumption information at the intelligent substations, provided by aggregating the information of the smart meters at the LV level. Here, they are referred to as measurement units. In this regard, introducing an effective measurement placement scheme has been discussed in the literature to provide the optimal location and number of measurement units, together with the type of measurements [20]. Baran et al. [21] are of the pioneers who addressed this issue in distribution systems by proposing a set of rules to increase the SE accuracy. In [22], a robust optimization algorithm is deployed to minimize the total estimation variance of nodal voltages in different network topologies. A Fisher information-based approach has been suggested in [23], to solve the optimal measurement placement problem. In [24], flow measurements are placed to increase the accuracy of the branch current estimation. The optimal measurement placement has been settled in [20, 25], and [26], by translating this problem to bounding the SE uncertainty.

As noted, in most of the existing studies, the optimal measurement placement problem in distribution systems has been addressed as minimizing or confining the SE uncertainty. When data mining approaches are employed to directly evaluate the system operating condition, this problem can be presented as selecting the minimum set of system features that grants an adequate accuracy for the data mining approach, which lays the base for the method adopted here.

In this study, evaluating the distribution system operating condition is addressed by presenting the system operating constraints as a set of binary classification problems. The measurements throughout the network are utilized as the features to differentiate between the binary classes, and DA is applied as the classifier. In this way, a comprehensible measure is provided for the operator to quantify the well-being of the system oper-

ation and identify the susceptible areas in the network. Additionally, based on the propounded condition evaluation technique, a measurement placement plan for distribution systems is proposed involving two steps: singling out the effective measurement units, based on their contribution to the classification process; and pinpointing the units containing redundant information, by deploying the mutual information (MI) and differential entropy (DE) concepts. The main contributions of this work include the following:

- The proposed approach for evaluating the distribution system operating condition provides a simple evaluation process that requires a much lower processing time, compared to the conventional SE approaches. This enables it to be employed in real-time applications.
- This approach can provide higher accuracy than the conventional SE approaches, as it dedicates all the resources only to find the status of important states.
- This approach is applicable for balanced and unbalanced, MV or LV networks and can treat any system operating constraint.
- Since the contribution of each measurement unit is evaluated regardless of its type or configuration, these units could be of any type. This includes single- or three-phase configurations, as well.
- The proposed measurement placement scheme delivers optional solutions, meaning that it provides the accuracy of the estimation that is achievable by each combination of measurement units. This enables to assess the trade-off between the quality of the SE and the investment it requires.

This paper is organized as follows. Section 2 elaborates the proposed method for evaluating the distribution system operating condition and also formulates the problem of measurement placement. Section 3 discusses the results of applying the propounded approaches to the IEEE 123 node test feeder. Section 4 concludes the paper.

2 | METHODOLOGY

Different constraints are established to ensure that the distribution system is operating securely, e.g. over- or under-voltage limits for nodes, thermal limits for components, reverse power flows for lines etc. In this section, assessing the status of each system constraint (or a combination of system constraints) is presented as a binary classification problem with the first class representing the violation of that constraint, and the second class standing for the normal operation of the system, regarding that constraint. Measurements throughout the network are considered as the system features to be utilized by the classifier. In the offline phase, different operating points of the system are simulated to train the classifier in recognizing the pattern leading to the realization of each class. This classifier then could be employed in real time to assess the status of system constraints, based on online measurements.

2.1 | DA to evaluate system operating condition

Among various classification methods, DA is an easy-to-implement technique that has already shown its applicability in settling different issues in power system studies, from load forecasting [27], load modelling [28], and fault detection [29] to assessing the power system security [5, 10]. DA describes each class with a unique multivariate normal distribution and employs the properties of these distributions to distinguish among the classes. Although normal distribution is presumed, DA is robust against the violation of this presumption [30], especially when system features are continuous and bounded [31].

Let $X = (x_{1j})$ contain the values of n measurements in the network. Despite not being investigated in this study, these measurements could include pseudo-measurements (historical data) too. Denote $f_{1r}(X)$ and $f_{2r}(X)$ the probability density functions of the instances in which the r th system constraint (or any of a set of constraints) is violated and the instances in which the system is operating under normal condition, regarding that constraint, respectively. Together, they are denoted as $f_{kr}(X)$, $k = 1, 2$. DA assumes a multivariate normal distribution, $\mathcal{N}_n(M_{kr}, \Sigma_{kr})$, for each class with the probability density function of

$$\begin{aligned} f_{kr}(X) &= \exp(-0.5\psi_{kr}(X)) / (2\pi|\Sigma_{kr}|) \\ \psi_{kr}(X) &= (X - M_{kr})\Sigma_{kr}^{-1}(X - M_{kr})^T \end{aligned} \quad (1)$$

where M_{kr} is the mean and Σ_{kr} is the covariance of the k th normal distribution, for the r th system constraint.

Considering (1), according to the Bayes theorem, the probability of X belonging to the k th class for the r th system constraint ($p_{kr}(X)$) is given by [32]

$$p_{kr}(X) = f_{kr}(X)\rho_{kr} / (f_{1r}(X)\rho_{1r} + f_{2r}(X)\rho_{2r}) \quad (2)$$

where ρ_{kr} denotes the prior probability of associating with the k th class. In DA, the class with the largest probability is selected in order to minimize the expected number of misclassification [32].

If sufficient measurements are available, the proposed approach can be employed as described. In the event that the measurements are deficient, the proposed presentation of the distribution system observability can be used to settle the optimal measurement placement problem. To do so, the measurements from all the nominated units are utilized to form the classifier and then the optimal combination of measurement units is achieved by discarding the ineffective and redundant measurements.

2.2 | Assessing effectiveness of measurement units

The effectiveness of each unit is assessed based on its contribution to the classification problem. To this end, one effective way is to investigate how the variations of each measurement

influence the probability of belonging to each class. A criterion that well demonstrates this sensitivity is the magnitude of the derivative of that probability with respect to that measurement (since $p_{1r} + p_{2r} = 1$, we have $|\partial p_{1r}/\partial x| = |\partial p_{2r}/\partial x|$ for any r and x , and therefore, it does not concern which class is chosen. Here, we consider $|\partial p_{1r}/\partial x|$). This criterion demonstrates to what extent that measurement (x) can change the status of that constraint (r).

As some units may contain multiple measurements, the effectiveness of each unit, i.e. $Y = [y_1, \dots, y_c]$, can be represented by the gradient of p_{1r} with respect to its contained measurements as

$$\nabla p_{1r} = [\partial p_{1r}/\partial y_1, \dots, \partial p_{1r}/\partial y_c]^T \quad (3)$$

where ∇ denotes the gradient operator. It is shown in the Appendix section that ∇p_{1r} can be calculated using (1) and (2) as

$$\nabla p_{1r} = \rho_{1r}\rho_{2r} \frac{(\nabla f_{1r}f_{2r} - \nabla f_{2r}f_{1r})}{(\rho_{1r}f_{1r} + \rho_{2r}f_{2r})^2} \quad (4)$$

Since ∇p_{1r} is a vector and also it depends on the system operating point, the mean of the 2-norm of ∇p_{1r} over all the data points in the training set, denoted by $\|\nabla p_{1r}\|_2$, is employed as the benchmark to assess the effectiveness of each unit regarding each constraint.

While identified as effective, some units may contain redundant information regarding the classification process. In theory, an increase in the information redundancy could decrease the misclassification rate, but it should be investigated whether this is economically reasonable. In this regard, following the concepts of MI and DE are exploited to settle the trade-off between minimizing the number of measurement units and minimizing the evaluation error, by relating the amount of information redundancy in each unit to the amount of accuracy it provides. This relationship is employed to pinpoint the minimum set of units which grants a requested accuracy.

2.3 | Redundancy and the concept of differential entropy

Suppose measurements are divided into groups a and b : group b includes the measurements from a specific unit (or a specific group of units), to be discarded, and group a includes the remainder measurements. In compliance, X in (1) is divided into X^a and X^b . To measure the redundancy between the measurements of groups a and b , one effective index is the MI, defined as [33]

$$I(X^a; X^b) = \int_{X^a} \int_{X^b} f_{(X^a, X^b)} \ln \left(\frac{f_{(X^a, X^b)}}{f_{X^a} f_{X^b}} \right) \quad (5)$$

where $I(X^a; X^b)$ denotes the MI between X^a and X^b , and f_{X^a} , f_{X^b} , and $f_{(X^a, X^b)}$ are the probability density functions of X^a , X^b , and their joint probability density function, respectively. The concept of MI is linked to the DE of random variables.

The DE represents the uncertainty in a continuous random variable and is given by [34]

$$b(X) = - \int_S f_x \ln(f_x) dx \quad (6)$$

where $b(X)$ and S denote the DE of X and its supporting set. Considering (5) and (6), it can be concluded that [33]

$$I(X^a; X^b) = b(X^a) + b(X^b) - b(X^a, X^b) \quad (7)$$

where $b(X^a, X^b)$ is the entropy in the joint probability density function of X^a and X^b .

According to (7), to maximize the MI between X^a and the remainder measurements, X^b should be chosen to maximize $b(X^a) + b(X^b)$. For an n -variate normal distribution, $\mathcal{N}_n(M, \Sigma)$, DE is given by [34]

$$b(X) = n/2 + n/2 \ln(2\pi) + \ln(|\Sigma|)/2 \quad (8)$$

where $|\Sigma|$ represents the determinant of Σ . Considering (7) and (8) together, we have

$$I(X^a; X^b) = (\ln|\Sigma^a| + \ln|\Sigma^b| - \ln|\Sigma|)/2 \quad (9)$$

where Σ^a and Σ^b denote the covariance matrices associated with X^a and X^b , respectively. Σ^a is given by omitting the rows and columns of Σ corresponding to X^b .

In addition to having a large MI between X^a and X^b , the redundant measurements, X^b , should be pinpointed such that the remainders, X^a , preserve most of the information, or in other words, has the largest DE value. According to (8), for a particular number of redundant measurements (equivalently a particular number of remainder measurements), $b(X^a)$ is proportional to $|\Sigma^a|$. Therefore, for a particular number of redundant measurements, to simultaneously preserve most of the information and have a large MI, the best combination to be considered redundant is the one which leads to the largest value for $|\Sigma^a|$. Consequently, since the determinant of a matrix is equal to the product of its singular values, the best practice to discard the redundant measurements is to omit the rows and columns of $|\Sigma|$, such that it leads to the elimination of the lowest singular values of $|\Sigma|$.

Following, based on the derived conclusion, two benchmarks are presented to assess the relative error of discarding the redundant measurements as a function of the number and combination of the selected measurements.

2.4 | Pinpointing redundant measurements

Suppose Σ ($n \times n$) is the covariance matrix for the first class of one of the system operating constraints. As Σ is symmetrical, the singular values of Σ is given by

$$S = U^T \Sigma U \quad (10)$$

where U ($n \times n$) is an orthogonal matrix and S ($n \times n$) is a diagonal matrix that contains the singular values of Σ in descending order [$S = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and λ_i the i th singular value of Σ].

To eliminate the lowest singular values of Σ , let us divide S to S^a ($(n-q) \times (n-q)$) and S^b ($q \times q$), such that S^b contains the q singular values that are relatively small enough to be discarded (later, it will be discussed how small). Accordingly, U is divided into $U^a = (u_{ij}^a)$ and $U^b = (u_{ij}^b)$ as

$$\begin{aligned} \begin{bmatrix} S^a & 0 \\ 0 & S^b \end{bmatrix} &= \begin{bmatrix} U^{aT} \\ U^{bT} \end{bmatrix} \Sigma \begin{bmatrix} U^a & U^b \end{bmatrix} \\ &= \begin{bmatrix} U^{aT} \Sigma U^a & U^{aT} \Sigma U^b \\ U^{bT} \Sigma U^a & U^{bT} \Sigma U^b \end{bmatrix} \end{aligned} \quad (11)$$

It can be concluded from Equation (11) that

$$U^{bT} \Sigma = S^b U^{bT} \quad (12)$$

If Σ_i denotes the i th row of Σ , Equation (12) can be rewritten as

$$\sum_{i=1}^n u_{ij}^b \Sigma_i = \lambda_j^b [u_{1j}^b, u_{2j}^b, \dots, u_{nj}^b], \quad \text{for } j = 1, 2, \dots, q \quad (13)$$

If, as presumed, the values of λ_j^b are small enough, it can be concluded from Equation (13) that

$$\sum_{i=1}^n u_{ij}^b \Sigma_i \approx 0, \quad \text{for } j = 1, 2, \dots, q \quad (14)$$

Equation (14) implies that q rows of Σ are linearly dependent on the other $n-q$ rows. Because each row of Σ is associated with one measurement, Equation (14) relates the problem of finding the redundant measurements to finding the linearly dependent rows of the covariance matrix. Following, the error associated with selecting each combination of the rows of Σ as the dependent rows is presented. Since we are looking for the minimum number of units to be installed, it is desirable to discard as many rows as possible, providing that the error remains below a certain limit. Hence, this error provides a yardstick to decide how many of the rows and which combination should be discarded.

Suppose $\Sigma_m, \Sigma_{m+1}, \dots,$ and Σ_{m+q} are chosen as the linearly dependent rows. Considering Equation (13), we have

$$\begin{aligned} &\left[\Sigma_m^T \dots \Sigma_{m+q}^T \right]^T \\ &= -F^{-1} G \left[\Sigma_1^T \dots \Sigma_{m-1}^T \Sigma_{m+1+q}^T \dots \Sigma_n^T \right]^T \\ &+ F^{-1} S^b U^{bT} \end{aligned} \quad (15)$$

where F and G are defined as (16) and (17), respectively. It should be noted that for the simplicity of presentation, here,

consequent rows of Σ were chosen, while any q out of n combination of the Σ rows could be selected.

$$F = \begin{pmatrix} u_{m1}^b & \dots & u_{(m+q)1}^b \\ \dots & \dots & \dots \\ u_{mq}^b & \dots & u_{(m+q)q}^b \end{pmatrix} \quad (16)$$

$$G = \begin{bmatrix} u_{11}^b & \dots & u_{(m-1)1}^b & u_{(m+1+q)1}^b & \dots & u_{n1}^b \\ \dots & \dots & \dots & \dots & \dots & \dots \\ u_{1q}^b & \dots & u_{(m-1)q}^b & u_{(m+1+q)q}^b & \dots & u_{nq}^b \end{bmatrix} \quad (17)$$

2.5 | Error of discarding redundant measurements

Considering (15), the error of discarding redundant measurements, denoted by E , equals $F^{-1}S^bU^{bT}$, which is a q by n matrix. To present the magnitude of the error as a scalar number, the 2-norm of this matrix is considered. By definition, the 2-norm of a given matrix, A , denoted by $\|A\|_2$, equals to its largest singular value, $\lambda_{\max}(A)$. From the properties of the 2-norm, if A, B, and C are matrices of proper sizes, we have

$$\|ABC\|_2 \leq \|A\|_2 \|B\|_2 \|C\|_2 \quad (18)$$

Based on the above statement, $\|F^{-1}\|_2$ equals $(1/\lambda_{\min}^F)$, where λ_{\min}^F is the smallest singular value of F , $\|U^{bT}\|_2$ equals 1 (since U is unitary), and $\|S^b\|$ equals λ_{n-q+1} . Using Equation (18), an upper limit for $\|E\|_2$ can be provided as

$$\|E\|_2 = \|F^{-1}S^bU^{bT}\|_2 \leq \lambda_{n-q+1}/\lambda_{\min}^F \quad (19)$$

Defining the relative error of discarding the redundant measurements as the ratio of $\|E\|_2$ to $\|\Sigma\|_2$, since $\|\Sigma\|_2$ equals λ_1 , the upper limit for the relative error is given by

$$\begin{aligned} \|E\|_2/\|\Sigma\|_2 &\leq E_1 E_2 \\ E_1 &= \lambda_{n-q+1}/\lambda_1, \quad E_2 = 1/\lambda_{\min}^F \end{aligned} \quad (20)$$

Equation (20) presents two important benchmarks regarding the identification of redundant measurements:

- The first benchmark, E_1 , which relates the number of measurements that are considered as redundant (q) to the cap of the relative error: the more measurements considered redundant, the larger the value of λ_{n-q+1} and, hence, the larger the relative error cap.
- The second benchmark, E_2 , which provides guidance for identifying the reluctant measurements. This benchmark suggests that to minimize the cap of the relative error, a combination of measurements leading to the minimum value for λ_{\min}^F should be selected.

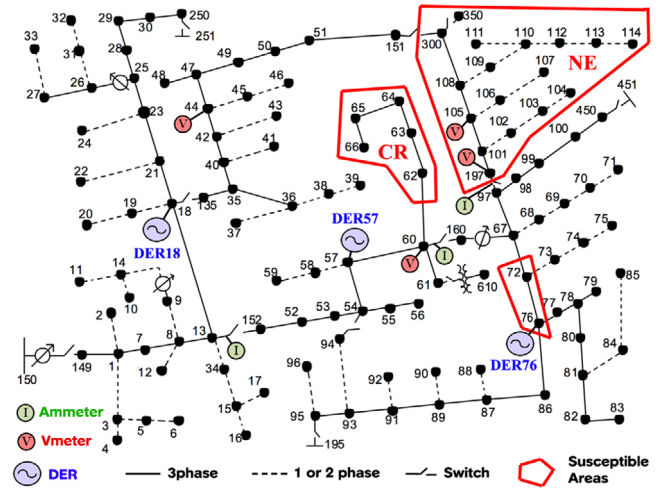


FIGURE 1 Modified IEEE 123 node test feeder

3 | RESULTS AND DISCUSSIONS

The proposed approaches are applied to the IEEE 123 node test feeder. This system is unbalanced, with a nominal voltage of 4.16 kV and contains different types of loads. To study the impact of DERs, the network is modified by integrating a wind power plant, a solar power plant, and a diesel generator, with the rated power of 200, 200, and 400 kW, respectively (DER18, DER57, and DER76, respectively). Figure 1 depicts the schematic diagram of the test system [35].

3.1 | Evaluation of operating condition

To explore different operating conditions, different set points are considered for the DERs by sweeping their active power generation from zero to the rated power, and the reactive power generation from the maximum leading to the maximum lagging power factor. At each set point, 7000 Monte–Carlo simulations are conducted to capture the probable variations of the system loads. To do so, a higher value than one, 1.12, is chosen for the mean of the Monte–Carlo simulations to put the system under further stress. Analyzing the Irish standard load profiles, developed by Electricity Supply Board (an Irish energy utility) for the Irish market in year 2019 [36], shows that the relative standard deviation of these load profiles varies between 0.25 and 0.42, where the relative standard deviation is defined as the ratio of the standard deviation to the mean. Therefore, here, 0.5 is selected as the standard deviation for the Monte–Carlo simulations to include all the load variations. This value corresponds to a relative standard deviation of 0.45, considering that 1.12 is selected as the mean. Note that these load profiles are employed for conducting time-series power flows in the following subsections.

While the proposed approach could be employed to investigate any operating condition, in this study, the focus is on the under- and over-voltage and over-current constraints. Results of

the Monte–Carlo simulations show that while no over-voltage condition (voltages above 1.05 pu) is observed, under-voltage (voltages below 0.95) occurs in 26.6% and 35.2% of the simulations at least at one of the phases of one of the nodes located in areas NE and CR in Figure 1, respectively. In addition, in 15.3% of the simulations, the over-current condition occurs at least at one of the phases of the line that connects bus 76 to 72, L76 (surpasses 140 Amperes). Denote V -NE and V -CR the value of the lowest 1-phase voltage of all the nodes located in NE and CR, respectively, and I -L76 the lowest value of the 1-phase current of L76. Three pairs of classes are formed based on the values of V -NE, V -CR, and I -L76, to represent the system operating constraints. The first classes represent the violation of each constraint and the second classes are standing for the normal operation, regarding each constraint.

In practice, it is not feasible to install measurement units at every desired location. For this purpose, four locations are nominated for the voltage measurements (V44, V60, V97, and V105) and three locations are considered for the current measurements (I13, I60, and I97), while DERs may be equipped with voltage and flow measurements, as well (voltage of the point of common connection and active and reactive power generation). The main substation is also assumed to be equipped with flow measurements (SUB). The measurements from the nominated units are utilized as the system features for the classification. DA is used as the classifier and 90% of the simulations are considered for training purposes. As a result, for each class, a normal distribution is adopted. The remainder 10% are used to form the test set. The proposed DA approach is programmed in Matlab R2019a and simulations are performed on a PC with an Intel(R) Core(TM) i7-8700K 3.7 GHz CPU and 16 GB of RAM. It takes around 22 s to build the DA model for each system constraint and 40 ms to predict using the trained models for each data sample.

Figure 2 presents the results of applying DA to the test set. This figure depicts the probability of the violation of each system constraint, calculated using Equation (2), versus the actual constraints values. As noted, for the values beyond and within the constraints, as expected, the calculated probabilities are higher and lower than 0.5, for most of the data points, respectively. Another noticeable trend is that as the actual values get farther from the constraint values, the probability of the constraint violation grows/shrinks, implying that the calculated probability reflects the intensity of constraint violation and, therefore, it can be used to quantify the well-being of the system operating condition.

3.2 | Identification of effective measurement units

Figure 3 depicts the normalized value of $\overline{\|\nabla p_{1r}\|_2}$ with respect to the measurements of different units for different system constraints. According to the results, V60, V97, V105, and DER57 have high contributions for all the constraints, while SUB, DER18, and I13 have noticeable contributions for V -NE and V -CR. The contribution of I60, V44, and I97 is negli-

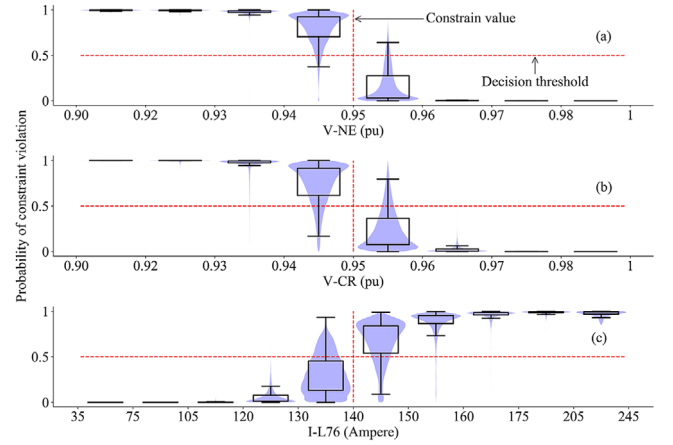


FIGURE 2 Box and whisker, together with Violin plot of the probability of the violation of system constraints versus actual values of the system constraints. The boxes cover 50% of the data points in each range, while the whiskers include 49.3% of the data points (the remainders are considered as outliers). The sides of the violins in Violin plot represent the probability density of the data in each range, smoothed by a kernel density estimator. (a) Constraint V -NE; (b) Constraint V -CR; (c) Constraint I -L76

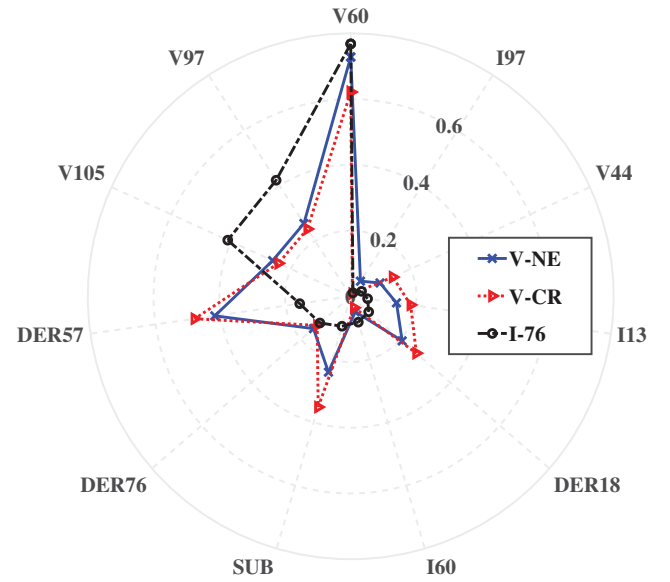


FIGURE 3 Normalized value of $\overline{\|\nabla p_{1r}\|_2}$ over the training set for different system constraints with respect to the measurements of different measurement units

ble for all the system constraints and, hence, they are discarded from the list of candidate units.

3.3 | Identification of redundant measurement units

As mentioned, Equation (14) links the problem of finding redundant measurement units to finding the linearly dependent rows of the covariance matrix. To implement Equation (14), the nine smallest singular values of the covariance matrices of

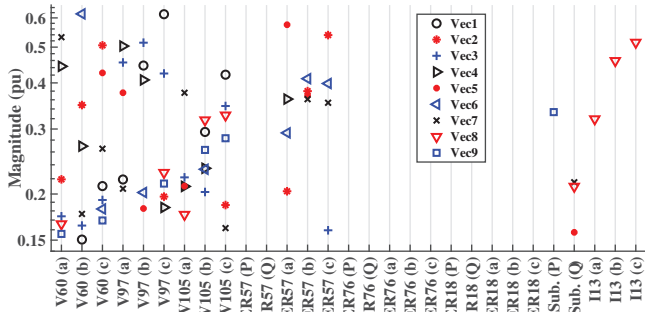


FIGURE 4 Absolute value of the elements of Σ_{1_r} eigenvectors, associated with the nine smallest Σ_{1_r} eigenvalues, for the constraint V-NE

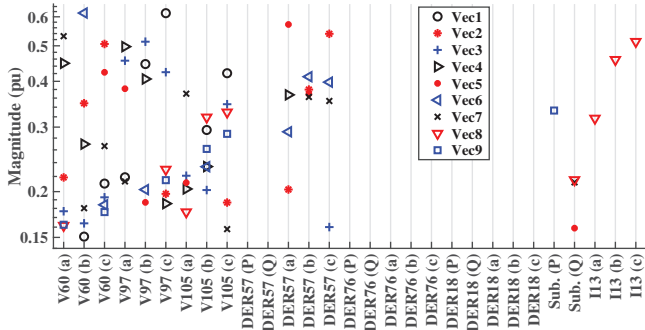


FIGURE 5 Absolute value of the elements of Σ_{1_r} eigenvectors, associated with the nine smallest Σ_{1_r} eigenvalues, for the constraint V-CR

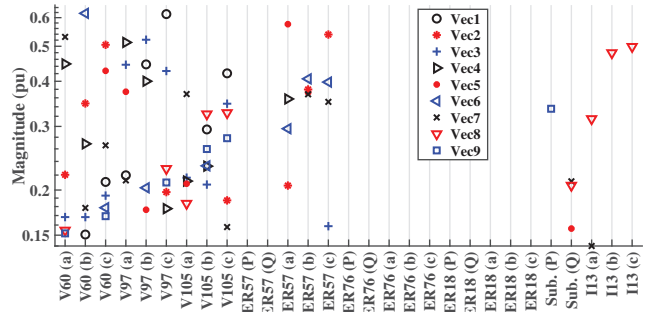


FIGURE 6 Absolute value of the elements of Σ_{1_r} eigenvectors, associated with the nine smallest Σ_{1_r} eigenvalues, for the constraint I-L76

different system constraints, together with their associated eigenvectors, are evaluated. Figures 4–6 show the absolute value of the elements of these eigenvectors for constraints V-NE, V-CR, and I-L76, respectively. Note that in these figures, for the sake of clarity, values beyond 0.15 are not shown. Also, they are plotted on a logarithmic scale. First, a similar pattern is recognizable for all the constraints in these figures. This is mainly because the topology of the network was assumed unchanged and, hence, the violation between each pair of the measurements to a great extent. Second, as the elements associated with the measurements of V60, V97, V105, SUB, and I13 have large values, it can be concluded that the redundancy

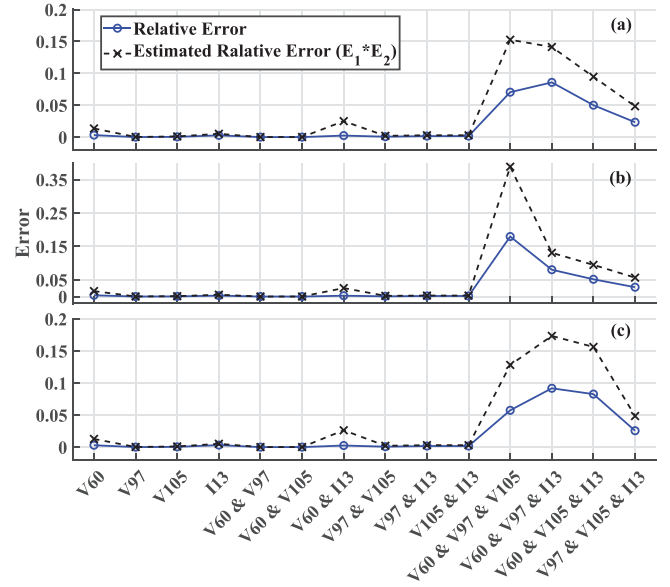


FIGURE 7 Relative error and estimated cap of relative error with discarding redundant units for the system constraints. (a) Constraint V-NE; (b) Constraint V-CR; (c) Constraint I-L76

exists among the measurements of these units. Despite the elements associated with the DER57, voltage measurements have large values; those which associated with the DER57 flow measurements have small values and, hence, DER57 is omitted from the list of redundant units.

Following, at each step, a subset of V60, V97, V105, and I13 is discarded and the related error ($\|E\|_2/\|\Sigma\|_2$) and the estimated cap of relative error (E_1E_2) are evaluated using Equations (19) and (20), respectively. Figure 7 presents the results. It can be remarked that (E_1E_2) provides a well-conservative estimation for the relative error. It is also noticed that for the combinations involving one or two of the units, the relative error is close to zero. On the other hand, for some combinations involving triple units, the relative error can take very large values; however, among them, the one that includes V97, V105, and I13 has the smallest relative error (3%). Therefore, the measurements of these units are considered redundant, leaving the optimal units as SUB, V60, DER57, DER18, and DER76.

In order to validate the provided results, the measurement placement approach of [37], which is based on the conservation voltage reduction concept, is implemented on the test feeder. This approach employs the maximum of the standard deviation of the estimated voltages for system nodes, driven by an SE approach, as a yardstick to evaluate the effectiveness of the measurement placement scheme. This yardstick is denoted by σ_m . Since this paper only focusses on the voltages of the nodes placed in the susceptible areas of Figure 1, σ_m is calculated only over these nodes. Since besides the substation, our approach suggested four of the measurement units, namely, V60, DER57, DER18, and DER76, as the best combination, we calculate σ_m with each four selection out of the effective units, V60, V97, V105, DER57, DER76, DER18, and I13. The least σ_m represents the best measurement placement scheme. Figure 8 depicts

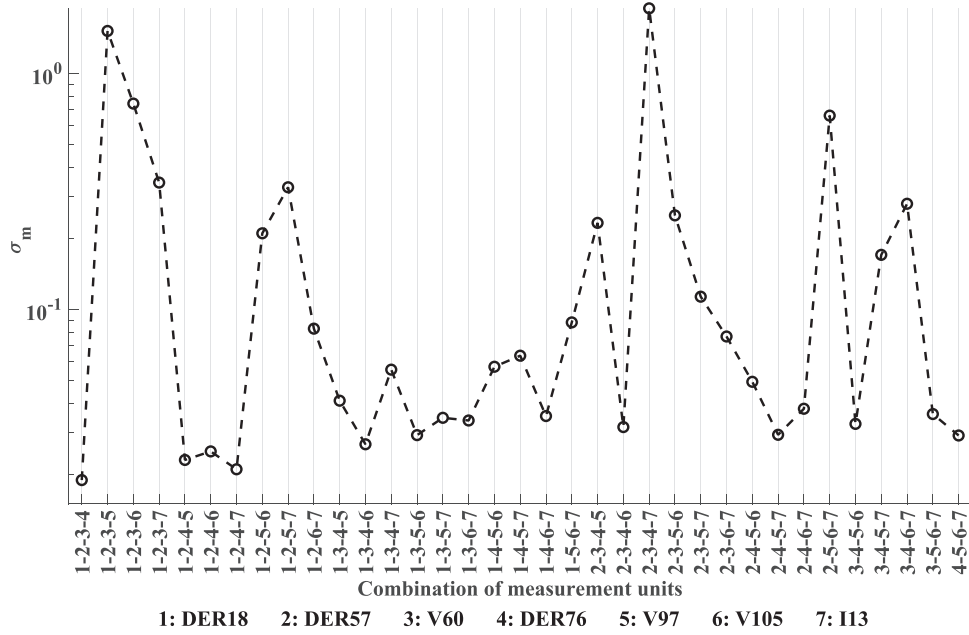


FIGURE 8 σ_m calculated for different combinations of the measurement units

the results. Note that the logarithmic scale is used for clarity. As seen, the results of this figure are consistent with those of Figure 7, as the least value of σ_m is achieved with the combination that includes V60, DER57, DER18, and DER76 (and hence excludes v97, V105, and I13).

3.4 | Receiver operating characteristic analysis

Receiver operating characteristic (ROC) analysis is employed to assess the impact of discarding the ineffective and redundant units on the DA performance in evaluating the system operating condition. This analysis is a prominent tool to assess the performance of classifiers [38]. Classification threshold is the limit that is used to decide whether an observation, X , belongs to a class, k , considering the probability of that observation belonging to that class, $p_k(X)$ in (2). In the binary classification, 0.5 (50%) is generally selected as the classification threshold, which represents an unbiased decision. In ROC analysis, despite 0.5 (50%), other possible values are considered as the probability threshold in (2), from 0 (0%) to 1 (100%), and the true-positive rate (fraction of correct classification) and false-positive rate (fraction of incorrect classification) are calculated for each probability threshold. Afterwards, an ROC curve is formed for each class by depicting the true-positive rate versus the false-positive rate values. In this order, the area under each ROC curve (AUC) reflects the overall performance of the classifier, regarding each class [38]. For a perfect classifier, the ROC curve reaches the top left corner, with the AUC of one, while smaller values show retreating from the ideal classifier, such that AUC of 0.5 represents a total random classifier.

Figure 9 plots the ROC curves for different system constraints with three models: the basic model that involves all the

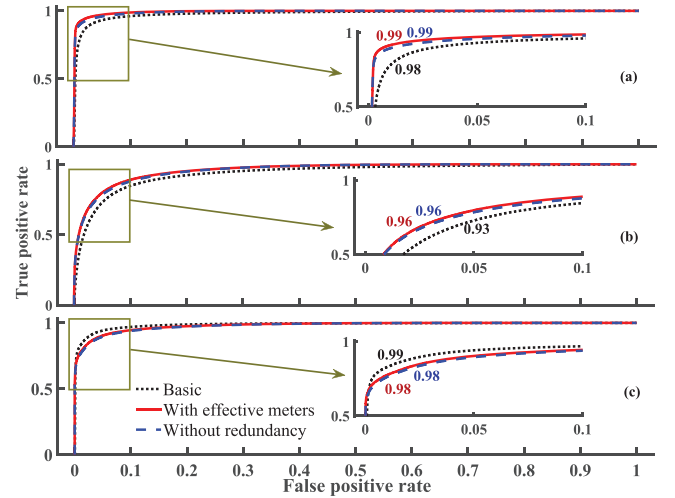


FIGURE 9 ROC curves for the system constraints. The value of AUC for each of the curves has been mentioned beside them. (a) Constraint V-NE; (b) Constraint V-CR; (c) Constraint I-L76

nominated units, the model that includes only effective units, and the model obtained after discarding the ineffective and redundant units. As noted, in all the cases, the AUC is above 0.93, showing that DA works near to ideal for this problem. In addition, discarding the ineffective and redundant units has not noticeably impacted the performance of the classification, which validates the proposed scheme for pinpointing the optimal measurement units. For the constraints V-NE and V-CR, the AUC of the basic model is slightly smaller than the two other ones. This is due to the enhancement in the numerical error of calculating the inverse of the covariance matrix in the latter cases, since with the basic model, the determinant of the covariance matrix is very close to zero (e.g. 3×10^{-135} in the case of

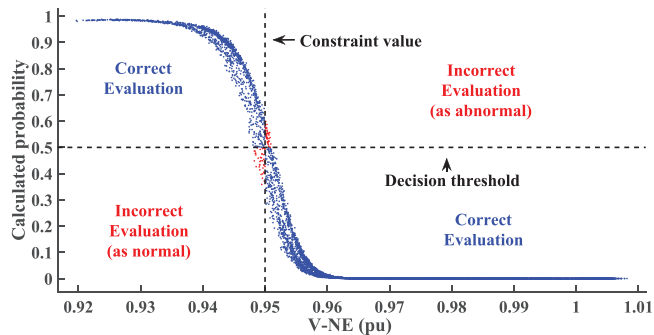


FIGURE 10 Calculated probability of constraint violation versus the actual value of V-NE. The blue hue represents correct evaluation of the system operating condition, while the red hue represents incorrect evaluation (as normal, while the actual condition is abnormal, or as abnormal, while the actual condition is normal)

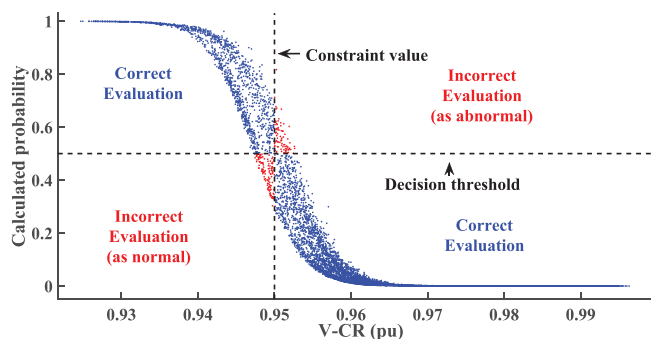


FIGURE 11 Calculated probability of constraint violation versus the actual value of V-CR. The blue hue represents correct evaluation of the system operating condition, while the red hue represents incorrect evaluation (as normal, while the actual condition is abnormal, or as abnormal, while the actual condition is normal)

V-NE, which with discarding the inefficient units improves to 9×10^{-96}).

3.5 | Time-series power flows

Time-series power flows are conducted to investigate the performance of the proposed approaches under a realistic scenario. For this purpose, five different aforementioned Irish standard load profiles [36] are allocated randomly to the system loads, and the generation profiles of wind and solar power plants in a specific year in Ireland are considered for DER18 and DER57, respectively. DER76, a diesel generator, is assumed to be operated for load-flattering purposes. Consequently, its generation is considered to be proportional to the net value of feeder's load in the 24 h ago: equal to zero for values under 3000 kW, equal to its rated power for values upper than 4500 kW, and linear in the middle part.

The model with the optimal units is utilized to evaluate the operating condition of the system. Figures 10–12, associated with the V-NE, V-CR, and I-L76 constraints, respectively, depict the calculated probability of constraint violation versus actual value of the system constraints for the 8760 h of a year. As

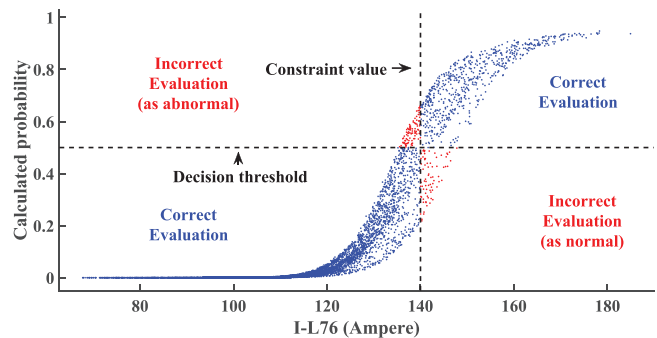


FIGURE 12 Calculated probability of constraint violation versus the actual value of I-L76. The blue hue represents correct evaluation of the system operating condition, while the red hue represents incorrect evaluation (as normal, while the actual condition is abnormal, or as abnormal, while the actual condition is normal)

TABLE 1 Performance of the proposed discriminant analysis approach, regarding the evaluation of the system operating condition in time-series power flows for different operating constraints

Constraint	Incorrect (as normal)	Incorrect (as abnormal)	Correct
V-NE	0.5%	0.5%	99%
V-CR	1.5%	1%	97.5%
I-L76	0.8%	1.1%	98.1%

noticed, while the proposed approach has accurately predicted the condition of the system in most of the hours, the calculated probability is proportional to the severity of the violation, verifying that this measure can be employed as an index to describe the well-being of the system. Table 1 presents the rate of correct evaluation of the system operating condition, together with the rate of incorrect evaluation (as normal, while the actual condition is abnormal, or as abnormal, while the actual condition is normal) for the time-series power flows simulations. As seen, the error rate is below 2.5% for all the constraints. These results are consistent with the results of Figure 9, as the best performance is conceived for V-NE, and then I-L76 (total of 1% and 1.9% error, respectively).

In order to compare the performance of the proposed DA approach, the weighted least square (WLS) SE approach [39] is implemented on the test feeder. In this order, the magnitude and phase of the system nodes voltage are considered as the system state variables. These variables are used afterwards to evaluate the system operating condition, regarding the V-NE, V-CR, and I-L76 constraints. On the same aforementioned environment, it takes around 60 s to estimate all the state variables for each data sample using this approach. Comparison to the 40 ms of the proposed DA approach establishes the superiority of our approach in terms of being fast-track. Table 2 presents the rate of correct and incorrect evaluation of the system operating condition, using the WLS approach on the results of time-series power flows. Compared to the results provided in Table 1, it can be noted that a higher level of accuracy is provided by the proposed DA approach, especially for predicting the status of the I-L76 constraint. This superiority can be justified by

TABLE 2 Performance of the weighted least square state estimation approach [39], regarding the evaluation of the system operating condition in time-series power flows for different operating constraints

Constraint	Incorrect (as normal)	Incorrect (as abnormal)	Correct
V-NE	1.6%	0.7%	97.7%
V-CR	1.5%	1.2%	97.3%
I-L76	19.4%	0%	80.6%

considering that in the proposed DA approach, the concentration is on estimating only the important states, especially when they approach the system constraints, while the conventional SE methods do not prioritize the state variables.

4 | CONCLUSION




A data-driven approach was proposed to deal with the SE in distribution systems, by addressing the system condition through the status of its operating constraints. To this end, each system operating constraint was presented as a binary classification problem and DA was employed as the classifier to evaluate the probability of the violation of each constraint. In this way, a comprehensible measure was provided for the operator to assess the well-being of the system operation and identify the susceptible areas in real time. Comparison showed that this approach is much faster than the conventional SE methods, which makes it a suitable choice to be employed in real-time applications. It also provides higher accuracy, regarding the evaluation of the system operating condition. While in this paper only the low-voltage and over-current conditions were studied, this approach is capable of investigating any operating constraint, including over-voltage, diverse power flow, fault etc.

Following, a scheme for measurement placement in distribution systems was proposed. This was achieved first by assessing the contribution of each measurement unit to the classification problem and then by deploying the MI and DE concepts to pinpoint and discard the units with redundant information. This approach delivers optional solutions which enable to assess the trade-off between the quality of SE and the investment it requires. The conservation voltage reduction meter placement approach validated the results provided by this measurement placement scheme.

ACKNOWLEDGMENTS

This work has emanated from research conducted with the financial support of Science Foundation Ireland under Grant No. SFI/16/IA/4496. The opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Science Foundation Ireland.

ORCID

Mohammad Jafarian  <https://orcid.org/0000-0002-7405-0479>
 Alireza Soroudi  <https://orcid.org/0000-0002-3651-6290>
 Andrew Keane  <https://orcid.org/0000-0001-8527-514X>

REFERENCES

- Baran, M.E., Kelley, A.W.: State estimation for real-time monitoring of distribution systems. *IEEE Trans. Power Syst.* 9(3), 1601–1609 (1994)
- Wang, J., Lu, X., Chen, C.: Guidelines for Implementing Advanced Distribution Management Systems. Argonne National Laboratory, Lemont, Illinois, United States (2015)
- Pau, M., Pegoraro, P.A., Sulis, S.: Efficient branch-current-based distribution system state estimation including synchronized measurements. *IEEE Trans. Instrum. Meas.* 62(9), 2419–2429 (2013)
- Strezoski, L., Stefani, I., Brbakic, B.: “Active Management of Distribution Systems with High Penetration of Distributed Energy Resources,” *IEEE EUROCON 2019 –18th International Conference on Smart Technologies*, Novi Sad, Serbia, 2019, pp. 1–5, doi: <https://doi.org/10.1109/EUROCON.2019.8861748>
- Garcia, G., Bernussou, J., Berbiche, M.: Pattern recognition applied to transient stability analysis of power systems with modelling including voltage and speed regulation. *IEE Proc. B Electr. Power Appl.* UK 139(4), 321 (1992)
- Wang, B., et al.: Power system transient stability assessment based on big data and the core vector machine. *IEEE Trans. Smart Grid* 7(5), 2561–2570 (2016)
- Moulin, L.S., et al.: Support vector machines for transient stability analysis of large-scale power systems. *IEEE Trans. Power Syst.* 19(2), 818–825 (2004)
- Guo, T., Milanovic, J.V.: Online identification of power system dynamic signature using pmu measurements and data mining. *IEEE Trans. Power Syst.* 31(3), 1760–1768 (2016)
- Zhu, L., Lu, C., Sun, Y.: Time series shapelet classification based online short-term voltage stability assessment. *IEEE Trans. Power Syst.* 31(2), 1430–1439 (2016)
- Jensen, C.A., El-Sharkawi, M.A., Marks, R.J.: Power system security assessment using neural networks: Feature selection using Fisher discrimination. *IEEE Trans. Power Syst.* 16(4), 757–763 (2001)
- Cremer, J.L., Konstantelos, I., Strbac, G.: From optimization-based machine learning to interpretable security rules for operation. *IEEE Trans. Power Syst.* 34(5), 3826–3836 (2019)
- He, M., Zhang, J., Vittal, V.: Robust online dynamic security assessment using adaptive ensemble decision-tree learning. *IEEE Trans. Power Syst.* 28(4), 4089–4098 (2013)
- Sun, M., Konstantelos, I., Strbac, G.: A deep learning-based feature extraction framework for system security assessment. *IEEE Trans. Smart Grid* 10(5), 5007–5020 (2019)
- Wu, J., He, Y., Jenkins, N.: A robust state estimator for medium voltage distribution networks. *IEEE Trans. Power Syst.* 28(2), 1008–1016 (2013)
- Dehghanpour, K., et al.: A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation. *IEEE Trans. Smart Grid* 10(6), 5942–5951 (2019)
- Yuan, Y., et al.: A multi-timescale data-driven approach to enhance distribution system observability. *IEEE Trans. Power Syst.* 34(4), 3168–3177 (2019)
- Zamzam, A.S., Fu, X., Sidiropoulos, N.D.: Data-driven learning-based optimization for distribution system state estimation. *IEEE Trans. Power Syst.* 34(6), 4796–4805 (2019)
- Ferdowsi, M., et al.: A scalable data-driven monitoring approach for distribution systems. *IEEE Trans. Instrum. Meas.* 64(5), 1292–1305 (2015)
- Li, C., et al.: Assessment method and indexes of operating states classification for distribution system with distributed generations. *IEEE Trans. Smart Grid* 7(1), 481–490 (2016)
- Singh, R., Pal, B.C., Vinter, R.B.: Measurement placement in distribution system state estimation. *IEEE Trans. Power Syst.* 24(2), 668–675 (2009)
- Baran, M.E., Zhu, J., Kelley, A.W.: Meter placement for real-time monitoring of distribution feeders. *IEEE Trans. Power Syst.* 11(1), 332–337 (1996)
- Damavandi, M.G., Krishnamurthy, V., Marti, J.R.: Robust meter placement for state estimation in active distribution systems. *IEEE Trans. Smart Grid* 6(4), 1972–1982 (2015)
- Xygiakis, T.C., Korres, G.N., Manousakis, N.M.: Fisher information-based meter placement in distribution grids via the d-optimal experimental design. *IEEE Trans. Smart Grid* 9(2), 1452–1461 (2018)

24. Pau, M., et al.: Impact of current and power measurements on distribution system state estimation uncertainty. *IEEE Trans. Instrum. Meas.* 68(10), 3992–4002 (2019)
25. Liu, J., et al.: Trade-offs in pmu deployment for state estimation in active distribution grids. *IEEE Trans. Smart Grid* 3(2), 915–924 (2012)
26. Cataliotti, A., et al.: LV measurement device placement for load flow analysis in MV smart grids. *IEEE Trans. Instrum. Meas.* 65(5), 999–1006 (2016)
27. Chien, N.P., Lautz, L.K.: Discriminant analysis as a decision-making tool for geochemically fingerprinting sources of groundwater salinity. *Sci. Total Environ.* 618, 379–387 (2018)
28. Hubele, N.F., Cheng, C-S.: Identification of seasonal short-term load forecasting models using statistical decision functions. *IEEE Trans. Power Syst.* 5(1), 40–45 (1990)
29. Wang, Y., et al.: Discriminant-analysis-based single-phase earth fault protection using improved PCA in distribution systems. *IEEE Trans. Power Deliv.* 30(4), 1974–1982 (2015)
30. Tabachnick, B.G., Fidell, L.S.: *Using Multivariate Statistics*. Allyn & Bacon/Pearson Education, Boston, MA, USA, p. 207
31. Dillon, W.R.: The performance of the linear discriminant function in nonoptimal situations and the estimation of classification error rates: A review of recent findings. *J. Market. Res.* 16(3), 370 (1979)
32. James, G.M., Hastie, T.J.: Functional linear discriminant analysis for irregularly sampled curves. *J. Roy. Statist. Soc.: Series B (Statist. Methodol.)* 63(3), 533–550 (2001)
33. Batina, L., et al.: Mutual information analysis: A comprehensive study. *J. Cryptol.* 21, 269–291 (2011)
34. Cover, T.M., Thomas, J.A.: *Elements of Information Theory*. Series in Telecommunications and Signal Processing, John Wiley & Sons, Inc., USA (1991)
35. Kersting, W.H.: Radial distribution test feeders. *IEEE Trans. Power Syst.* 6(3), 975–985 (1991)
36. Standard, E.: <https://rmdservice.com/standard-load-profiles/>. ESB Networks Limited, Dublin, Ireland (2020). Accessed 1 April 2020
37. Zamani, V., Baran, M.E.: Meter placement for conservation voltage reduction in distribution systems. *IEEE Trans. Power Syst.* 33(2), 2109–2116 (2018)
38. Chang, C.: Multiparameter receiver operating characteristic analysis for signal detection and classification. *IEEE Sens. J.* 10(3), 423–442 (2010)
39. Abur, A., Expósito, A.G.: *Robust State Estimation*. Power System State Estimation CRC Press, CRC Press, Boca Raton, Florida, United States (2004)

How to cite this article: Jafarian M, Soroudi A, Keane A. A data-driven measurement placement to evaluate the well-being of distribution systems operation. *IET Gener Transm Distrib.* 2021;1–11.
<https://doi.org/10.1049/gtd2.12114>

APPENDIX A

To calculate ∇p_{1r} , suppose X in (1) is rephrased to form $X^s = [X^a, X^b]$, where X^a is the measurements corresponding to a specific measurement unit and X^b is the remainder. Suppose M_{kr} and Σ_{kr} are also rephrased in compliance to form $M_{kr}^s = [M_{kr}^a, M_{kr}^b]$ and Σ_{kr}^s . Let Ψ_{kr} denote the inverse of Σ_{kr}^s as

$$\Psi_{kr} = (\Sigma_{kr}^s)^{-1} = \begin{bmatrix} \Psi_{kr}^{aa} & \Psi_{kr}^{ab} \\ \Psi_{kr}^{ba} & \Psi_{kr}^{bb} \end{bmatrix} \quad (\text{A.1})$$

with Ψ_{kr}^{aa} , Ψ_{kr}^{ab} , Ψ_{kr}^{ba} and Ψ_{kr}^{bb} of proper sizes. Considering (A.1), ψ_{kr} in (1) is given by

$$\psi_{kr}(X^s) = ([X^a \ X^b] - [M_{kr}^a \ M_{kr}^b]) \begin{bmatrix} \Psi_{kr}^{aa} & \Psi_{kr}^{ab} \\ \Psi_{kr}^{ba} & \Psi_{kr}^{bb} \end{bmatrix} ([X^a \ X^b] - [M_{kr}^a \ M_{kr}^b])^T \quad (\text{A.2})$$

Let $\nabla \psi_{kr}$ denote the gradient of ψ_{kr} with respect to X_a . Using (A.2), $\nabla \psi_{kr}$ is calculated as

$$\nabla \psi_{kr} = (\Psi_{kr}^{aa} + \Psi_{kr}^{aaT})(X^a - M_{kr}^a)^T + (\Psi_{kr}^{ab} + \Psi_{kr}^{baT})(X^b - M_{kr}^b)^T \quad (\text{A.3})$$

Considering (1) and (A.3), the gradient of f_{kr} with respect to X_a (∇f_{kr}) is given by

$$\nabla f_{kr} = -0.5 \nabla \psi_{kr} \exp(-0.5 \psi_{kr}(X^s)) / (2\pi |\Sigma_{kr}|) \quad (\text{A.4})$$

Now, Equation (A.4) together with (2) are employed to calculate the gradient of p_{1r} with respect to X_a (∇p_{1r}) as

$$\nabla p_{1r} = \rho_{1r} \rho_{2r} \frac{(\nabla f_{1r} f_{2r} - \nabla f_{2r} f_{1r})}{(\rho_{1r} f_{1r} + \rho_{2r} f_{2r})^2} \quad (\text{A.5})$$