IET Smart Grid

Special Issue: Machine Learning in Power Systems

IET Journals
The Institution of Engineering and Technology

Hybrid data-driven physics model-based framework for enhanced cyber-physical smart grid security

eISSN 2515-2947
Received on 1st October 2019
Revised 27th November 2019
Accepted on 6th December 2019
E-First on 15th April 2020
doi: 10.1049/iet-stg.2019.0272
www.ietdl.org

Cody Ruben[1] ✉, Surya Dhulipala[1], Keerthiraj Nagaraj[1], Sheng Zou[1], Allen Starke[1], Arturo Bretas[1], Alina Zare[1], Janise McNair[1]

[1]*Electrical and Computer Engineering, University of Florida, 1064 Center Dr, Gainesville, FL, USA*

✉ *E-mail: cruben31@ufl.edu*

Abstract: This study presents a hybrid data-driven physics model-based framework for real-time monitoring in smart grids. As the power grid transitions to the use of smart grid technology, it's real-time monitoring becomes more vulnerable to cyber-attacks like false data injections (FDIs). Although smart grids cyber-physical security has an extensive scope, this study focuses on FDI attacks, which are modelled as bad data. State-of-the-art strategies for FDI detection in real-time monitoring rely on physics model-based weighted least-squares state estimation solution and statistical tests. This strategy is inherently vulnerable by the linear approximation and the companion statistical modelling error, which means it can be exploited by a coordinated FDI attack. In order to enhance the robustness of FDI detection, this study presents a framework which explores the use of data-driven anomaly detection methods in conjunction with physics model-based bad data detection via data fusion. Multiple anomaly detection methods working at both the system level and distributed local detection level are fused. The fusion takes into consideration the confidence of the various anomaly detection methods to provide the best overall detection results. Validation considers tests on the IEEE 118-bus system.

Nomenclature

$J(\boldsymbol{x})$	WLS cost function – $\mathbb{R}^{1\times 1}$
d	number of measurements – $\mathbb{R}^{1\times 1}$
$\boldsymbol{P}$	projection matrix – $\mathbb{R}^{d\times d}$
α_χ	significance level – $\mathbb{R}^{1\times 1}$
$\boldsymbol{z}$	vector of measurements – $\mathbb{R}^{1\times d}$
$\boldsymbol{z}^*$	operating measurement vector – $\mathbb{R}^{1\times d}$
$\hat{\boldsymbol{z}}$	estimated measurement vector – $\mathbb{R}^{1\times d}$
$\boldsymbol{Z}$	training set – $\mathbb{R}^{d\times K_1}$
$\hat{\boldsymbol{Z}}$	testing set – $\mathbb{R}^{d\times K_2}$
K_1	number of samples in training set – $\mathbb{R}^{1\times 1}$
K_2	number of samples in testing set – $\mathbb{R}^{1\times 1}$
$\boldsymbol{x}$	vector of state variables – $\mathbb{R}^{1\times N}$
$\boldsymbol{x}^*$	operating state vector – $\mathbb{R}^{1\times N}$
$\hat{\boldsymbol{x}}$	estimated state vector – $\mathbb{R}^{1\times N}$
N	number of states – $\mathbb{R}^{1\times 1}$
$\boldsymbol{\Sigma}_{\text{SE}}$	covariance of measurement vector for state estimator – $\mathbb{R}^{d\times d}$
$\boldsymbol{\Sigma}$	covariance of measurement vector – $\mathbb{R}^{d\times d}$
$\boldsymbol{e}$	vector of measurement error – $\mathbb{R}^{1\times d}$
e_{D}	detectable component of error – $\mathbb{R}^{1\times 1}$
e_{U}	undetectable component of error – $\mathbb{R}^{1\times 1}$
II	innovation index of measurement – $\mathbb{R}^{1\times 1}$
σ	standard deviation of e – $\mathbb{R}^{1\times 1}$
$\boldsymbol{h}(\boldsymbol{x})$	vector of measurement estimates – $\mathbb{R}^{d\times d}$
$\boldsymbol{H}$	Jacobian matrix – $\mathbb{R}^{d\times N}$
$\mathscr{B}$	set of buses $j \ni Y_{ij} \neq 0$
$\chi^2_{d,p}$	Chi-squared value – $\mathbb{R}^{1\times 1}$
$\boldsymbol{r}$	vector of measurement residuals – $\mathbb{R}^{1\times d}$
CME	composed measurement error – $\mathbb{R}^{1\times 1}$
W_t	driving noise of O–U process
β	decay-rate of O–U process
σ_n^2	variance of noise of O–U process
$\mu_{o-u}(t)$	long-term mean of O–U process
δ^{CD}	squared Mahalanobis distance value – $\mathbb{R}^{1\times 1}$
μ	mean of measurement vector – $\mathbb{R}^{1\times d}$
τ	threshold value to classify abnormal samples – $\mathbb{R}^{1\times 1}$
$\boldsymbol{T}$	set of τ ($\mathbb{R}^{1\times M}$)
l	label of a sample – $\mathbb{R}^{1\times 1}$
α	weight value – $\mathbb{R}^{1\times 1}$
M	number of buses – $\mathbb{R}^{1\times 1}$
$\boldsymbol{Y}$	label for training samples – $\mathbb{R}^{1\times K_1}$
$\hat{\boldsymbol{Y}}$	label for testing samples – $\mathbb{R}^{1\times K_2}$
ϕ_m	symbol for the mth local CorrDet detector
Φ_E	symbol for ensemble CorrDet detector
Φ_R	symbol for CorrDet detector
$\delta_{Z,m}$	δ^{CD} of all training samples with respect to ϕ_m – $\mathbb{R}^{1\times K_1}$
$\delta_{z_k^*}$	δ^{CD} of the kth testing sample with respect to Φ_E – $\mathbb{R}^{1\times M}$
$\boldsymbol{\Psi}_{\text{ECD}}$	overall decision scores from ensemble CorrDet detector – $\mathbb{R}^{1\times K_2}$
$\boldsymbol{\Psi}_{\text{SE}}$	decision scores from state estimator technique – $\mathbb{R}^{1\times K_2}$
$\boldsymbol{\Psi}_{\text{fusion}}$	fusion decision scores for hybrid data-driven physics model-based framework – $\mathbb{R}^{1\times K_2}$

1 Introduction

The future power grid, or Smart Grid (SG), will integrate control, communication and computation aiming to achieve stability, efficiency and robustness of the physical processes on the system. These advancements bring many challenges to the SG and therefore have drawn much attention from academia, industry and government due to the great impact they will have on society, economics and the environment. While a great amount of research has been done towards these objectives, science and technology related to the cyber-physical security of SGs are still immature. Additionally, much of the critical infrastructure is currently transitioning towards the paradigm of SGs by increasing the dependency of control of physical processes on communication networks, thus becoming exposed to cyber-threats [1].

Guaranteeing the reliable operation of power grids is crucial for today's society and it is done through real-time power system monitoring. Currently, real-time monitoring is done through a process called power system state estimation (PSSE) [2], which provides relevant information on the power grid current operating point based on the measurements throughout the system. These measurements are commonly transmitted to a Supervisory Control and Data Acquisition (SCADA) system, which implements centralised monitoring and control for the electrical grid, where PSSE is performed. One important feature of PSSE is its error processing capability. Measurements that are clearly inconsistent are discarded in the pre-filtering step, which precedes the state estimation step. Following state estimation using pre-filtered data, a post-processing step called bad data analysis is performed. This step aims at detecting bad data or gross errors (GEs), assuming they are statistically large errors.

The established procedure to determine the current operating point for the power grid has been based on iterative numerical linearisation around the incumbent solution, which is determined easily by the least-squares approach to finding solutions within the convex hull. There is always an inherent error in this approximation procedure that is difficult to quantify, but it has been sufficiently small for the current requirements of dispatching, detection of failures, and reliability studies. This linear approximation and the companion statistical modelling error approach alone are not compatible to the new demands of cyber-physical security because its sensitivity and specificity are theoretically lower bounded by the level and dynamics of the approximation error, which can be easily explored by coordinated cyber-attacks.

The increasing dependence on digital monitoring and control of power systems raises concerns with respect to cybersecurity. One common type of cyber-threat is false data injection (FDI) attack, where an attack aims to disrupt the operation of the power grid by modifying a subset of measurement values. While bad data analysis is capable of detecting many instances of gross errors via tests such as the Chi-squared test, largest normalised residual [3] or innovation-based [4] approaches, intelligent cyber-attacks may be engineered to be difficult to detect [5], considering the implicit constraints of physics model-based solutions.

Methods devised to tackle FDI attacks include Generalised Likelihood Ratio Detector with L-1 Norm Regularisation [6], a scheme for protecting a selected set of measurements and verifying the values of a set of state variables independently [7], and the estimation of the normalised composed measurement error for detection of malicious data attacks [8]. These solutions all consider a quasi-static physics measurement model; however, the power grid is a time-varying system with loads and generation constantly changing. Thus grid temporal characteristics are not fully explored for FDI detection.

Considering environment temporal characteristics, machine learning-based solutions have also been explored for anomaly detection. The Correlation-based Detection (CorrDet) algorithm [9] was introduced for landmine detection. The Reed-Xiaoli (RX) Detector [10] was introduced for target detection of remote sensing images. Both methods rely on the approach where an incoming sample is classified as abnormal if its squared Mahalanobis distance with respect to a background statistic is above some threshold.

Some artificial intelligence-related FDI detection methods have been put forward in recent years, which are mainly based on neural network, deep learning and fuzzy clustering [11, 12]. In [13], the authors focus on the FDI detection in smart grids using a deep belief network-based (DBN) method with unsupervised learning to provide the initial weights. The authors of [14] present an artificial neural network (ANN) based approach which identifies the FDI by tracking the measurement data. The uniqueness of the neural network method is a simple infrastructure but uneasy in the parameter adjustment. Numerous tests should be utilised to train the network model. A deep learning method originates from the neural network, which can solve the overfitting problems well but the training method is more complex [15]. Unsupervised learning is performed from the bottom of the restricted Boltzmann machine to provide initial weights for the network. The backpropagation algorithm propagates the error from top to bottom and fine-tunes the model parameters.

Considering hybrid data-driven physics model-based solutions for FDI detection, the literature review will present seldom contributions. It is clear that analytic quasi-static model-based solutions can represent spatial characteristics of the environment, while data-driven solutions can explore temporal characteristics which are inherent to the grid operation. Thus they are complementary solutions. Considering such a rationale, in [16], a previous work of the authors, an extended Chi-squared test using information from PSSE and a data-driven CorrDet algorithm [9] is presented.

In this work, a hybrid data-driven physics model-based framework for FDI detection on system real-time monitoring is presented. Fig. 1 illustrates the presented framework, which explores both temporal and spatial characteristics of the environment in a distributed multi-agent architecture. An ensemble CorrDet (ECD) algorithm is presented to address drifting load scenarios and numerical issues. The CorrDet algorithm is used to learn the background statistics (e.g. mean and covariance for normal samples) over the whole power grid topology, consisting of a series of buses where feature (measurement) values are measured in a single bus or between two buses. ECD detector can be considered as a set of CorrDet detectors for each local environment. Therefore, an ECD detector learns a series of background statistics, one for each bus. There are several advantages for using an ECD detector compared to CorrDet detector. FDI is sparse, thus learning the background statistics on each bus, instead of the whole power system topology, allows for a more sensitive anomaly detection while embedding local environmental spatial characteristics to the data-driven solution. More specifically, spatially neighbouring buses are more highly correlated and easier to be affected by an attack while buses that further away have a lower correlation. Thus, learning a full covariance overall measurements of all buses are unnecessary (nearly sparse covariance), especially when training data is limited. Instead, local, fewer dimensional measurement sets offer a more accurate statistic estimation and a computationally cheaper, more sensitive anomaly detection. ECD detector is also a scalable solution as even when more buses are added to the grid, we could just add a local CorrDet detector to that bus and include the result of this local detector in the ECD algorithm. Second, learning the background statistics on the whole power system topology is usually more challenging than learning on a bus. The latter allows a distributed local environmental model learning on the much smaller dimensional data (usually several measurements) than the former (usually several hundreds of measurements). To avoid the numerical issues due to large dimensions, a much larger training set would be necessary for the CorrDet detector to achieve the same performance; otherwise, estimated background statistics are often ill-posed. Third, ECD detector is more robust and secure than CorrDet detector, since the classification of the anomaly for ECD detector is an aggregated decision of a series of local CorrDet detectors, as a *committee*, allowing for a small number of failed local CorrDet detections, while CorrDet detector is a single detector. Finally, a good feature of ECD detector is FDI localisation. For instance, if an FDI happens on a measurement between two buses and the corresponding two local CorrDet detectors flags the sample as abnormal, it can also be inferred that the FDI is a measurement between the two buses, while CorrDet detector cannot find the location where false data was injected.

In order to create a hybrid between physics model-based and data-driven solutions, a decision level fusion solution is presented. On such, several data-driven and physics-based methods for the FDI attacks detection on SG real-time monitoring are combined producing one output. The fusion considers the confidence of each anomaly detection method, creating the hybrid method that proves to perform better than any of the individual detection methods.

The contributions of this paper towards the state-of-the-art are as follows:

- Decision level fusion is used to create a hybrid physics-based data-driven anomaly detection that considers the confidence of individual anomaly detection methods.
- A data-driven ECD anomaly detection method is presented, which works in a distributed fashion, allowing for more sensitive local detection on the spatial domain.
- Using a physics-based bad data detection method considering the Innovation Concept to improve the data-driven detection method.

The remainder of the paper is organised as follows. In Section 2, background information of the challenges and existing solutions are presented. Section 3 presents the proposed framework, discussing each component and how they are used to build the framework. The results of numerical tests used to evaluate the method's performance are shown in Section 4. Finally, Section 5 presents the conclusions of this work.

2 Background, challenges, existing solutions and their limitations

2.1 SG system

PSSE plays a crucial role in SG real-time monitoring, as it is responsible for estimating unknown system state variables based mostly on measurements of voltage, power and current. As SG evolution exposes the system to cyber-attacks, it becomes increasingly important to develop countermeasures [17–22]. Most research [23–25] in this area focuses on improving bad data detection schemes or improving the security of the communication system. An analysis of the state-of-the-art physics model-based solutions for FDI detection on SGs real-time monitoring will show that these apply residual-based approaches for cyber-attack detection, while ignoring the inherent masked error component [26]. Still, as cited in [27], a stealthy attack requires the corruption of several measurements. This relates to the fact that a stealthy attack must have the attack vector fitting the measurement model, which is equivalent to shifting the result of the state estimation to a physically possible but wrong solution. Cyber-attacks with such characteristics are hard to discover when applying the classical bad data approaches, which may cause the cyber-attack to remain undetected [28].

The authors of [29] proposed an interval forecasting method to predict the possible largest variation bounds of each state variable based on a worst-case analysis based on the forecasting uncertainties of renewable energy sources, and electric loads. Works such as [30] also use extreme learning machine-based (ELM) one-class-one-network (OCON) and prediction methods to improve the resilience of the power system, exploiting the spatial correlation of power data within subnets.

While power systems are ideally in a steady-state, there are constant variations to load and generation. Current bad data analysis techniques in power systems work in a quasi-steady state, only considering a single snapshot of the system. In theory, temporal changes on the system could add more information for bad data detection that current techniques are not using. In [16], the authors begin to make use of this temporal data by developing an extended Chi-squared test that combines the classical PSSE method with a data-driven CorrDet algorithm that takes into account past data. This extended Chi-squared test does not work as well when introduced to more realistic load variation and stealthier FDI attacks.

2.2 CorrDet and ECD detectors

The CorrDet detector is capable of training a data-driven anomaly detector; however it is not very efficient and may suffer from numerical issues when the number of measurements is large. Therefore, the ECD detector is proposed. For the CorrDet detector, only one set of parameters (sample mean, sample covariance matrix and detection threshold) are estimated for the whole power system. These parameters, especially covariance matrix, characterise the correlations among all measurements on all buses. The required number of training samples rapidly grows when the number of measurements increases. Therefore, the detector parameters, especially the covariance matrix, are usually not well-trained, resulting in numerical issues and bad detection performance when the number of training samples is limited. However, learning a full covariance matrix for all measurements is actually not necessary since the covariance matrix is sparse. The reason is that the farther two buses are, the less correlation their measurements will have. In other words, if an FDI happens on one measurement, the most probable affected buses are usually the one or two that the measurement is linked to. Thus, this challenge can be tackled by considering the detection at a smaller spatial scale. More specifically, instead of training the detector parameters at the scale of all measurements of all buses, we can train a set of

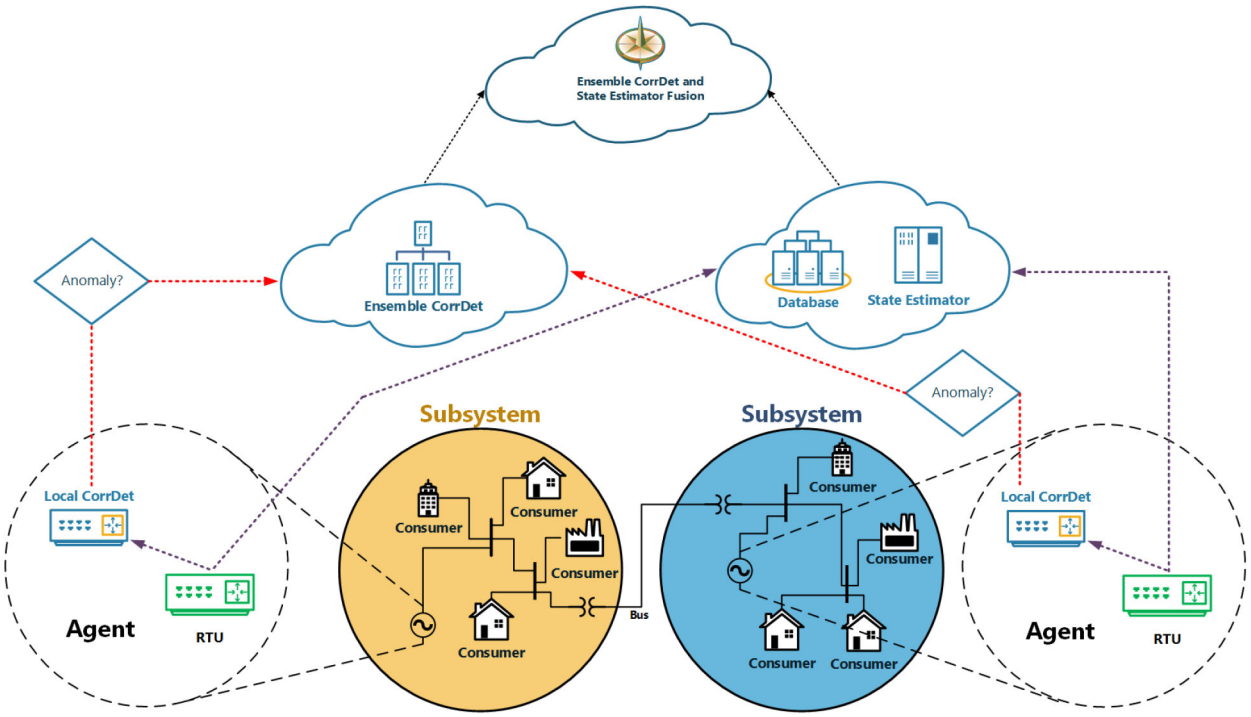

Fig. 1 *Distributed multi-agent architecture*

detectors, which trains measurements associated on every single bus.

Typically, the number of measurements that are linked to each bus is less than ten, while the total number of measurements is several hundred. The reduced dimensionalities for each local CorrDet detector yields more accurate estimation of normal sample means and covariance matrices and finally a more sensitive detection. By looking at which local CorrDet detector(s) reports an anomaly, we can further infer the possible position where the FDI happens.

3 Hybrid physics model-based data-driven framework

3.1 Physics model

In modern Energy Management Systems (EMS), the State Estimation (SE) process is the core process for situational awareness of a power system and is used in many EMS applications, including the detection of bad data. The common approach to SE is using the classical Weighted Least Squares (WLS) method described in [2]. In this approach, the system is modelled as a set of non-linear algebraic equations based on the physics of the system:

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

where $z \in \mathbb{R}^{1\times d}$ is the measurement vector, $\boldsymbol{x} \in \mathbb{R}^{1\times N}$ is the vector of state variables, $\boldsymbol{h}: \mathbb{R}^{1\times N} \to \mathbb{R}^{1\times d}$ is a continuously non-linear differentiable function, and $\boldsymbol{e} \in \mathbb{R}^{1\times d}$ is the measurement error vector. Each measurement error, e_i is assumed to have zero mean, standard deviation σ_i and Gaussian distribution. d is the number of measurements and N is the number of states.

In the classical WLS approach, the best estimate of the state vector in (1) is found by minimising the cost function $\boldsymbol{J}(\boldsymbol{x})$:

$$\boldsymbol{J}(\boldsymbol{x}) = \| z - \boldsymbol{h}(\boldsymbol{x}) \|^2_{\boldsymbol{\Sigma}^{-1}_{\mathrm{SE}}} = [z - \boldsymbol{h}(\boldsymbol{x})]' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} [z - \boldsymbol{h}(\boldsymbol{x})] \tag{2}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of the measurements. In this paper, we consider the standard deviation of each measurement to be 1% of the measurement magnitude, which has been shown to improve the detection of bad data [31]. The covariance submatrix for zero injection measurements is calculated, as shown in (4) and (5). In order to solve this problem, (1) is linearised at a certain point $\boldsymbol{x}^*$ in (3) and the optimal states are found through an iterative process

$$\Delta z = \boldsymbol{H} \Delta \boldsymbol{x} + \boldsymbol{e} \tag{3}$$

$$\sigma(P_i) = \sqrt{\sum_{j \in \mathscr{B}} \sigma^2(P_{ij})} \tag{4}$$

$$\sigma(Q_i) = \sqrt{\sum_{j \in \mathscr{B}} \sigma^2(Q_{ij})} \tag{5}$$

where $\boldsymbol{H} = \delta \boldsymbol{h} / \delta \boldsymbol{x}$ is the Jacobian matrix of $\boldsymbol{h}$ at the current state estimate $\boldsymbol{x}^*$, $\Delta z = z - \boldsymbol{h}(\boldsymbol{x}^*) = z - z^*$ is the correction of the measurement vector and $\Delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^*$ is the correction of the state vector. $\mathscr{B}$ is the set of buses $j \ni Y_{ij} \neq 0$. The WLS solution is the projection of Δz onto the Jacobian space by a linear projection matrix $\boldsymbol{P}$, i.e. $\Delta z = \boldsymbol{P} \Delta \hat{z}$. Letting $\boldsymbol{r} = \Delta z - \Delta \hat{z}$ be the residual vector, the $\boldsymbol{P}$ matrix that minimises $\boldsymbol{J}(\boldsymbol{x})$ will be orthogonal to the Jacobian range space and to $\boldsymbol{r}$; $\Delta \hat{z} = \boldsymbol{H} \Delta \hat{\boldsymbol{x}}$. This is in the form

$$\langle \Delta \hat{z}, \boldsymbol{r} \rangle = (\boldsymbol{H} \Delta \hat{\boldsymbol{x}})' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} (\Delta z - \boldsymbol{H} \Delta \hat{\boldsymbol{x}}) = 0. \tag{6}$$

Solving (6) for $\Delta \hat{\boldsymbol{x}}$:

$$\Delta \hat{\boldsymbol{x}} = (\boldsymbol{H}' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} H)^{-1} \boldsymbol{H}' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} \Delta z \,. \tag{7}$$

At each iteration, a new incumbent solution $\boldsymbol{x}^*_{\mathrm{new}}$ is found and updated following $\boldsymbol{x}^*_{\mathrm{new}} = \boldsymbol{x}^* + \Delta \hat{\boldsymbol{x}}$. Equation (7) is solved each iteration until $\Delta \hat{\boldsymbol{x}}$ is sufficiently small to claim convergence of the solution. Once the SE converges, the Innovation Concept is used for the detection of bad data in the measurement vector z [4, 8]. In the Innovation Concept, the Innovation Index (II) in (8) is defined as the ratio of the detectable and undetectable components of the measurement error and can be calculated using the projection matrix, $\boldsymbol{P}$, of the WLS solution (9)

$$\mathrm{II}_i = \frac{\| e_{\mathrm{D}i} \|_{\boldsymbol{\Sigma}^{-1}_{\mathrm{SE}}}}{\| e_{\mathrm{U}i} \|_{\boldsymbol{\Sigma}^{-1}_{\mathrm{SE}}}} = \frac{\sqrt{1 - P_{ii}}}{\sqrt{P_{ii}}} \,. \tag{8}$$

$$\boldsymbol{P} = (\boldsymbol{H}' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} \boldsymbol{H})^{-1} \boldsymbol{H}' \boldsymbol{\Sigma}^{-1}_{\mathrm{SE}} \,. \tag{9}$$

The II and the measurement residual, r, are used to calculate the Composed Measurement Error (CME) for each measurement, as shown in the following equation:

$$\mathrm{CME}_i = r_i \sqrt{1 + \frac{1}{\mathrm{II}_i^2}} \,. \tag{10}$$

The measurements are considered to be i.i.d, so the statistical Chi-squared test is used, as shown in the following equation:

$$\sum_{i=1}^{d} \left[\frac{\mathrm{CME}_i}{\sigma_i} \right]^2 > \chi^2_{\alpha_\chi, d} \tag{11}$$

for d degrees of freedom and significance level α_χ. In this paper, the common significance level of 0.05 is used [32]. If the sum of normalised CME values is greater than the Chi-squared distribution value, then bad data with $(1 - \alpha_\chi)$ confidence level is detected. We call this statistical test *CME Chi-Square Test* (CMECST).

3.2 Data-driven machine learning

In addition to the spatial information that is used in physics based-model, data-driven machine learning algorithms (CorrDet and ECD algorithms) are proposed to take advantage of both temporal and spatial information. CorrDet algorithm is proposed to estimate the sample statistics globally, while ECD algorithm combines local CorrDet estimates, where spatially remote correlations are ignored, but the spatially neighbouring correlations are reserved.

3.2.1 CorrDet anomaly detection: The machine learning layer of the proposed smart power grid framework uses the knowledge of already verified data to learn the normal state of a properly functioning grid. It is then able to detect any anomalies introduced into the system at any point forward and alerts the cloud layer, as shown in Fig. 1, to identify the anomaly, isolate it from the remainder of the system and take appropriate action to prevent contamination of the system, with regards to both power distribution in other subsystems, and data assimilation by the machine learning system itself.

A machine learning layer is implemented using the CorrDet Anomaly Detection [9, 10, 33] algorithm described in (12), where z is the new incoming data, μ is the mean and $\boldsymbol{\Sigma}^{-1}$ is the inverse covariance matrix of normal samples. Equation (12) calculates the squared Mahalanobis distance, $\delta^{\mathrm{CD}}(z)$, of a given data z, from the mean, μ of the distribution

$$\delta^{\mathrm{CD}}(z) = (z - \mu)^T \boldsymbol{\Sigma}^{-1} (z - \mu) \,. \tag{12}$$

The anomaly detector is trained with the first k number of incoming samples to generate the μ and $\boldsymbol{\Sigma}^{-1}$. It then accepts new data and uses (12) to determine its squared Mahalanobis distance and compares it to a threshold value τ.

$$l(z) = \begin{cases} 1, & \text{if } \delta^{\mathrm{CD}}(z) \geq \tau \\ 0, & \text{if } \delta^{\mathrm{CD}}(z) < \tau \,. \end{cases} \tag{13}$$

If the result is below the threshold, the new data is considered to be normal data (label $l(z) = 0$), but if the result is above the threshold, the new data is flagged as an anomaly (label $l(z) = 1$).

In order to select this threshold τ for Mahalanobis distance, we conduct an experiment to vary τ as a function of standard deviation (σ_{thr}) and mean (μ_{thr}) of Mahalanobis distance values of all the normal samples in training dataset as shown in (14). We use F1-score as the performance metric to choose the value of η of (14) that results in the highest F1-score for training data to decide the optimal value of τ.

$$\tau = \mu_{\mathrm{thr}} + \eta * \sigma_{\mathrm{thr}} \,. \tag{14}$$

The μ and $\mathbf{\Sigma}^{-1}$ statistics can be updated over time with measurement values from normal test samples to make the model more adaptive and learn the behaviour of data over time.

3.2.2 Ensemble CorrDet algorithm: The sample data z_i is usually a set of hundreds of measurements at time i, which can be represented as $z_t = [z_{t1}, z_{t2}, \ldots, z_{td}]$, where d is the number of measurements and each element z_{tj} denotes the jth measurement value at the time t. In this work, the dissimilarity comes from the abnormal behaviours of one or several measurements in the power system.

In other words, an abnormal sample is caused by at least one measurement value that is abnormal. There are a variety of reasons for a measurement to be abnormal, including an FDI. For instance, Fig. 2 shows the normalised real power flow values from bus 5 to bus 3 ($z_{t,340}$) from time $t = 1$ to $t = 10,000$. There are two measurements z_{9161} and z_{9385} in $z_{t,340}$ that are far different than the rest, which can be detected as abnormal measurements. Therefore, the samples z_{9161} and z_{9385}, containing the abnormal measurement $z_{9161,340}$ and $z_{9385,340}$, are detected as abnormal samples.

Since buses are connected via transmission lines, spatially neighbouring buses are more highly correlated and easily affected once being attacked while buses that are farther away have less correlation. Thus, learning a full covariance overall measurements of all buses is unnecessary (nearly sparse covariance) when training data is limited. Instead, it is proposed to use local regions with fewer measurements. These regions have smaller dimensions and offer a more accurate normal statistic estimation while being computationally cheaper and provide more sensitive anomaly detection. Based on the CorrDet algorithm, a spatial-temporal, regional CorrDet anomaly detection method is proposed, named *Ensemble CorrDet* (*ECD*).

ECD detector is defined as a set of local CorrDet detectors on data samples with a few, spatial neighbouring measurements, compared to full measurements in the CorrDet detector. To be more specific, let Φ_E be the ECD detector, Φ_R the CorrDet detector with full dimensionality and ϕ_m the local CorrDet detector with reduced dimensionality where $m = 1{:}M$ and M is the number of buses. So $\Phi_E = \{\phi_m\}_{m=1:M}$. Assume that the total number of measurements is d. There are m_j, $m_j < d$, measurements on each bus m, where each bus is considered as a local, spatial region, corresponding to one local CorrDet detector, ϕ_m.

For Φ_R, the learning process consists of estimating μ and $\mathbf{\Sigma}^{-1}$ from normal training samples z_i ($z_i \in \mathbb{R}^{1\times d}$). A similar strategy is proposed to learn the ECD detector. The learning of Φ_E involves the estimation of a set of local CorrDet detectors, ϕ_m. For each ϕ_m, similarly, the learning process consists of estimating its μ_m and $\mathbf{\Sigma}_m^{-1}$ from the normal training samples with selected measurements $\boldsymbol{z_{i,m}}$ ($\boldsymbol{z_{i,m}}$ is a $1 \times m_j$ vector). The threshold value τ_m for each ϕ_m is estimated using the same strategy of CorrDet detector as shown in (14). For the new incoming samples, a set of squared Mahalanobis distances, Φ_E, are computed and compared with the corresponding set of thresholds, T, where $T = \{\tau_m\}_{m=1:M}$. If at least one squared

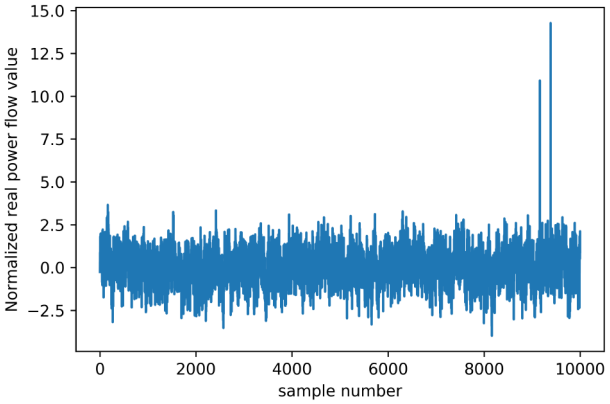

Fig. 2 *Normalised real power flow values from bus 5 to bus 3 for all samples $t = 1$ to $t = 10,000$*

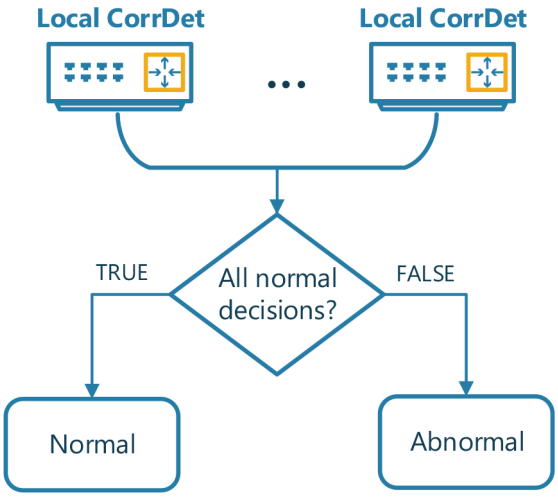

Fig. 3 *Voting strategy for ECD detector*

Mahalanobis distance in Φ_E is greater than its corresponding threshold, this incoming sample is classified as an anomaly. Otherwise, it is classified as normal samples. This voting process is visualised in Fig. 3.

Let K_1 and K_2 the number of training and testing samples, respectively. Let $\boldsymbol{Z}$ ($\boldsymbol{Z} \in \mathbb{R}^{d\times K_1}$) and $\widetilde{\boldsymbol{Z}}$ ($\widetilde{\boldsymbol{Z}} \in \mathbb{R}^{d\times K_2}$) the training and testing samples, respectively. Let $\boldsymbol{Y}$ ($\boldsymbol{Y} \in \mathbb{R}^{1\times K_1}$) and $\widetilde{\boldsymbol{Y}}$ ($\widetilde{\boldsymbol{Y}} \in \mathbb{R}^{1\times K_2}$) the corresponding labels. $\delta_{Z,m}$ ($\delta_{Z,m} \in \mathbb{R}^{1\times K_1}$) denotes the squared Mahalanobis distances of all training samples with respect to mth CorrDet detector, ϕ_m. $\delta_{\tilde{z}_k}$ ($\delta_{\tilde{z}_k} \in \mathbb{R}^{1\times M}$) denotes the squared Mahalanobis distances of the kth testing sample with respect to all CorrDet detectors, Φ_E. The pseudo code for the proposed ECD algorithm is shown in Fig. 4.

3.3 Fusion of physics model-based data-driven detection methods

In order to employ the anomaly detection capabilities of both state estimator solution and the ECD algorithm, we fuse the results from both the methodologies, as shown in Fig. 1. In the ECD algorithm, each sample in the testing set will have M CorrDet distances, one for each region as its decision score. We need an overall ECD decision score for each sample to combine it with the decision score from state estimator ($\mathbf{\Psi}_{\mathrm{SE},\tilde{z}_k}$). SE decision score is calculated as shown in the following equation:

$$\mathbf{\Psi}_{\mathrm{SE}} = \sum_{i=1}^{d} \left[\frac{\mathrm{CME}_i}{\sigma_i}\right]^2 . \tag{15}$$

To find the overall ECD decision score ($\mathbf{\Psi}_{\mathrm{ECD},\tilde{z}_k}$) for each testing sample, we consider the squared Mahalanobis distance of local

1: Train an Ensemble CorrDet detector:
Input: $\mathbf{Z}$, $\mathbf{Y}$, $\widetilde{\mathbf{Z}}$
2: **for** Every Local CorrDet detector $m = 1 : M$ **do**
3: Compute the mean and covariance of normal statistics: μ_m and $\mathbf{\Sigma}_m^{-1}$
4: Compute the squared Mahalanobis distance: compute $\delta_{Z,m}$ using Eq. 12
5: Compute the threshold: τ_m
6: **end for**
7: Test using the Ensemble CorrDet detector:
8: **for** Every test sample $k = 1 : K_2$ **do**
9: Compute the squared Mahalanobis distance: compute $\delta_{\widetilde{z}_k}$ using Eq. 12
10: **if** $\forall m, \delta_{\widetilde{z}_k} < \tau_m$ **then**
11: Classify $\widetilde{z}_k$ as normal sample: $\widetilde{y}_k = 0$
12: **else**
13: Classify $\widetilde{z}_k$ as abnormal sample: $\widetilde{y}_k = 1$
14: **end if**
15: **end for**
Output: $\widetilde{\mathbf{Y}}$

Fig. 4 *Procedure 1: ECD algorithm*

CorrDet detector, which led us to decide whether that sample is anomalous or normal. This is achieved by considering the decision score of the region whose local CorrDet detector detected an anomaly. In the case that there are multiple local CorrDet detectors that detect an anomalous sample, the maximum decision score is selected from these multiple local CorrDet detectors. If there was no anomaly detected from any local CorrDet detectors, the minimum of all decision scores from local CorrDet detectors is considered as the ECD decision score.

We normalise the decision scores obtained from state estimator and ECD detector by subtracting by their corresponding mean value and dividing by their corresponding standard deviation to form $\mathbf{\Psi}_{\text{ECD}_{\text{normalised}}}$ and $\mathbf{\Psi}_{\text{SE}_{\text{normalised}}}$. For each testing sample, we add the normalised decision scores from the state estimator and ECD detector, and create a new decision score termed as fusion decision score ($\mathbf{\Psi}_{\text{fusion},\widetilde{z}_k}$). Fusion decision scores, which are calculated as shown in (16), are compared with ground truth values to show the improvement in fused model performance compared to individual detectors

$$\mathbf{\Psi}_{\text{fusion}} = \mathbf{\Psi}_{\text{ECD}_{\text{normalised}}} + \mathbf{\Psi}_{\text{SE}_{\text{normalised}}}. \tag{16}$$

In order to decide whether an incoming test sample is anomalous or normal, we define a threshold (τ_{fusion}) based on the fusion decision scores of training set using (14). We compare the fusion decision score of the test sample ($\mathbf{\Psi}_{\text{fusion},\widetilde{z}_k}$) with ($\tau_{\text{fusion}}$) as shown in (17) to predict whether it is anomalous or normal

$$\widetilde{z}_k = \begin{cases} \text{anomalous}, & \mathbf{\Psi}_{\text{fusion},\widetilde{z}_k} \geq \tau_{\text{fusion}} \\ \text{normal}, & \mathbf{\Psi}_{\text{fusion},\widetilde{z}_k} < \tau_{\text{fusion}}. \end{cases} \tag{17}$$

4 Case study

4.1 Dataset description

For this work, the IEEE 118-bus system [34] illustrated in Fig. 5 was used to generate the training and testing data. Errors are introduced at random in 5% of the dataset samples and true labels are assigned during this process. Error detection based on physics-based models is performed as a post-processing step to PSSE using CMECST. The dataset consists of 10,000 samples with 712 measurements, of which 21 measurements are zero-injection measurements. Zero-injection measurements are not modelled as equality constraints in the PSSE process. The standard deviation of zero-injection measurements is calculated as shown in (4) and (5) to avoid any gain matrix singularity issues in PSSE process. The types of measurements considered were a combination of standard SCADA measurements, i.e. bus voltage magnitudes, real and reactive power injections, and real and reactive power flows. The measurement set data was generated using the power flow options in the MATPOWER package in MATLAB [35]. PSSE uses all 712 measurements while the ECD detector will only use 691 measurements, excluding the zero-injection measurements. A drifting load profile was considered for the generation of the measurement set. The drift was modelled by the Ornstein–Uhlenbeck (O–U) process – *a mean-reverting process* [36, 37]. This is a stochastic process similar to a random walk, but has a tendency to drift back towards the original load. The mean loading condition is updated periodically to model physical reality as aptly as possible. This presents a greater challenge to the data-driven solution since there will be greater variations from the mean vector. A detailed discussion of O–U process is included in the subsequent section.

4.1.1 O-U process: The O-U process X_t is defined by the following stochastic differential (SDE):

$$\mathrm{d}X_t = -\beta(X_t - \mu_{o-u})\,\mathrm{d}t + \sigma_n \mathrm{d}W_t \tag{18}$$

where X_t is a random variable, W_t is the driving noise, β is the decay-rate, σ_n^2 variance of the noise, and μ_{o-u} is the long term mean.

Equation (18) can be solved using Ito's formula and the solution is given by (19). It can be seen from (19) that $\lim_{t\to\infty} X_t = \mu_{o-u}$.

$$X_t = \mathrm{e}^{-\beta t}X_0 + \mu_{o-u}\left(1 - \mathrm{e}^{-\beta t}\right) + \sigma_n \int_0^t \mathrm{e}^{\beta(t_0 - t)}\,\mathrm{d}W_{t_0} \tag{19}$$

4.2 Experimental results

The anomaly detection problem can be treated as a classification problem, as our goal is to classify testing samples into normal and anomalous classes based on the models trained on training samples. To evaluate the performance of strategies included in this paper, we make use of classification metrics [38] such as Receiver Operating Characteristics (ROC) curves and Area Under Curve (AUC) score, which are defined based on True Positive Rate (TPR) and False Positive Rate (FPR) of a classification model. ROC curves represent the classification performance of a classifier by varying thresholds and calculating TPR and FPR values for each threshold for decision scores. In our analysis, we plot ROC curves using $\mathbf{\Psi}_{\text{fusion}}$ values and ground truth values.

In our dataset, we used 30% of samples for training and 70% of the samples for testing the model's anomaly detection abilities. In order to reduce the bias of our models, we repeated model training and testing for different combinations of samples in the data for ten times. For each combination, we obtained a ROC curve for testing samples to understand the model performance. In Fig. 6, we show

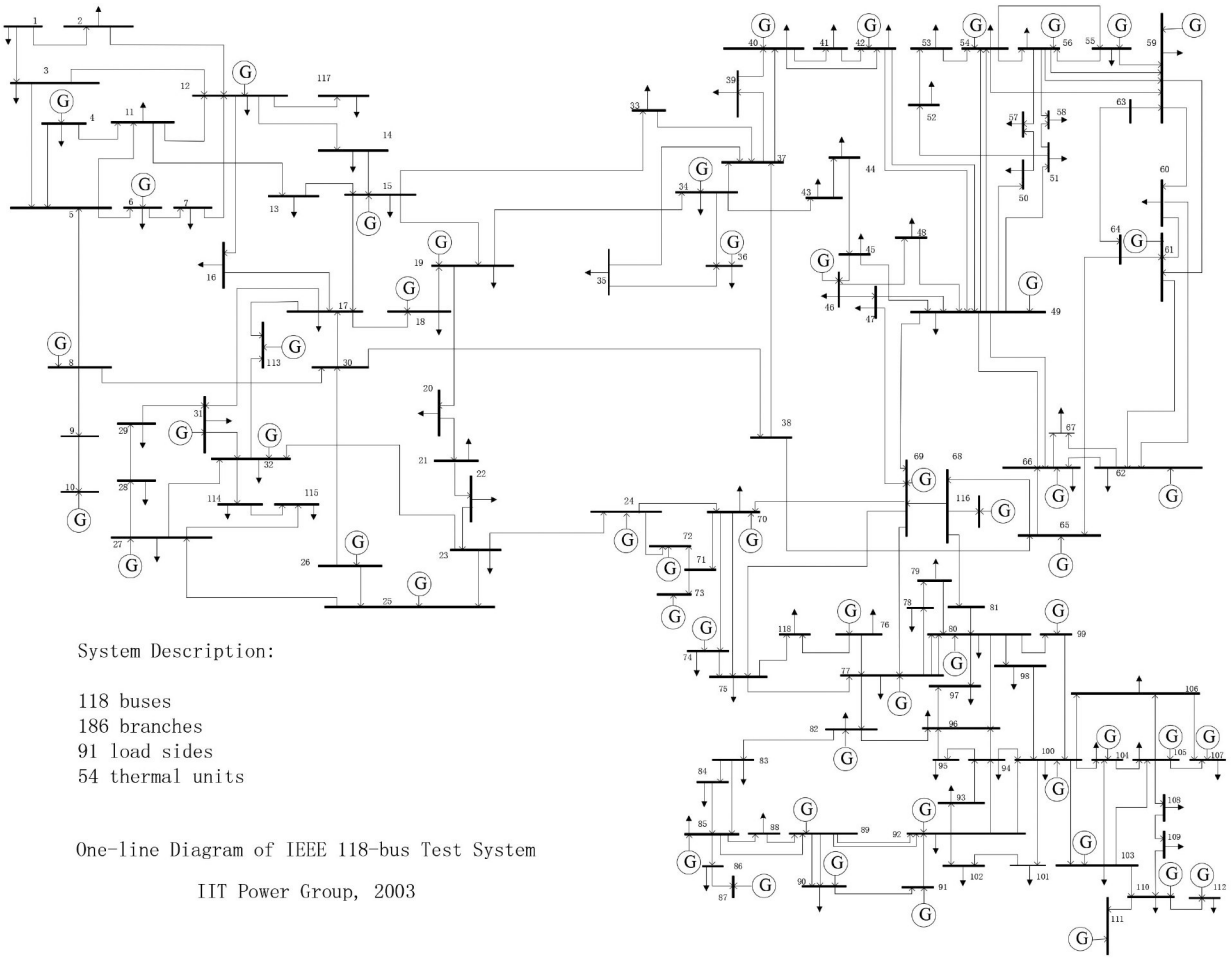

Fig. 5 *IEEE 118-bus system*

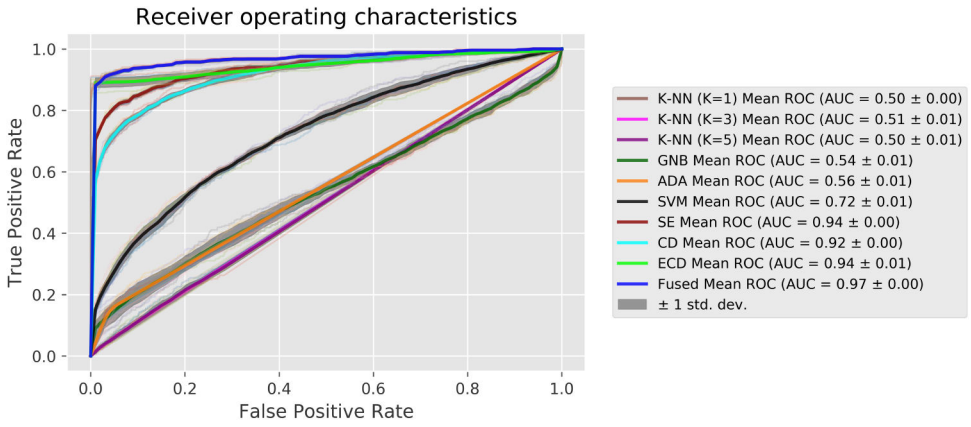

Fig. 6 *ROC curve comparison*

the ROC curves for multiple experiments conducted along with a mean ROC curve (averaged over FPR) and variation of mean AUC score (averaged over FPR) for each of the methodologies proposed in this paper.

In Fig. 6, we can notice that mean AUC score for ECD detector (0.95) is better than mean AUC score for state estimator (0.94), but the mean AUC score for fusion results (0.97) is much better than both ECD algorithm and state estimator method anomaly detection methodologies. The proposed hybrid data-driven physics model-based anomaly detection methodology improved mean AUC score by 3.2% compared to state estimator results.

We can also notice that ECD detector performs much better compared to state estimator solution and CorrDet detector as shown in Fig. 7. One reason is due to the fact that using all the measurements in CorrDet detector incurs numerical issues, but when we look at ECD detector, we are reducing the measurements for individual detectors, thereby reducing the numerical issues which improve the anomaly detection capabilities.

For an anomaly detection problem, FPR should be fairly low, or the model ends up wrongly classifying many normal samples as anomalous. Hence, we also show the performance of the presented anomaly detection methodologies specifically for FPR values less than 0.2 through corresponding mean ROC curves and variation of mean AUC scores for multiple experiments in Fig. 7. The optimal threshold (τ_{fusion}) for fusion decision scores ($\mathbf{\Psi}_{\text{fusion}}$) also lies in this region as to maximise F1-score, FPR value has to be fairly low.

In Fig. 7, instead of absolute AUC score, we show the relative AUC score, which is the ratio of AUC score by 0.2 (area in the curve for FPR values between 0 and 0.2). The presented hybrid data-driven physics model-based anomaly detection methodology improved mean AUC score by 6.75% compared to state estimator results for FPR values <0.2.

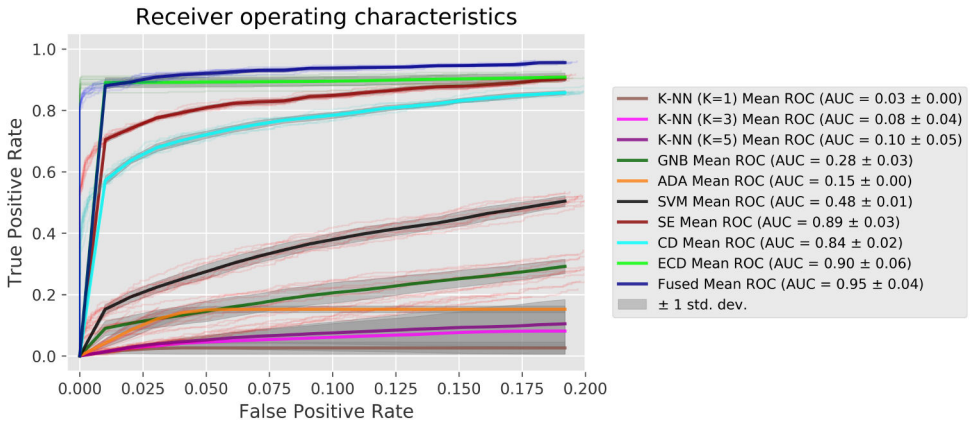

Fig. 7 *ROC curve comparison for an FPR <0.2*

From Figs. 6 and 7, we can observe that the performance improvement from the presented hybrid data-driven physics model-based anomaly detection methodology compared to state estimator prediction is larger in the ROC space where FPR is <0.2 compared to overall ROC space, which is a favourable outcome.

For comparison, the same dataset is evaluated on four other algorithms including *K*-nearest neighbour (KNN) [12], support vector machine (SVM) for FDI detection [12], Gaussian Naive Bayes (GNB) [39] and adaptive boosting with a decision tree (ADA) [40]. The authors in [12] do not consider the effects of concept drift in the data, whereas we have considered drifting load profile in our analysis which also poses more challenges for FDI detection. In order to have a fair comparison, the same training samples and testing samples are used as in the experiments of ECD detector. More specifically, the dataset is split into 30% for training and the rest 70% for testing. The AUC value of ROC curve is used for evaluation by taking the average of AUC values of ten repetitive experiments where each experiment has different training and testing samples.

KNN and SVM are two typical machine learning methods that were introduced for FDI detection in [12]. Since the exact hyperparameter values are not fully disclosed in [12] and the hyperparameter values depend on the dataset as well, a series of experiments with different hyperparameter values (e.g. *K* value in KNN) and other parameter settings (e.g. linear or non-linear SVM) are conducted.

For KNN, the number of neighbours *K* is selected as 1, 3 or 5. The classification results on testing dataset show that almost all testing samples are classified as normal samples, especially when *K* is larger. The reason is that the dataset used in this work has fewer abnormal samples (around 5%), which better models real anomaly detection scenarios, while the number of abnormal samples is more (around 30%) in [12]. Therefore, KNN, a classification algorithm cannot be simply applied for the more challenging anomaly detection problem, since most of the neighbouring training samples are normal samples because of the low abnormal/normal ratio. The scores of detecting using KNN are estimated using weighted KNN, where the scores are the normalised inverse of distances of neighbouring normal or abnormal samples. The ROC curves of KNN are shown in Figs. 6 and 7, where the AUC is much worse than the presented method.

Similarly, the SVM with linear and RBF kernels are used for FDI detection. The SVM with the non-linear kernel (RBF) has better performance than linear SVM. Therefore, the ROC curve of SVM with RBF is shown in Figs. 6 and 7, where the average AUC value is around 0.72, better than the weighted KNN algorithm, but still worse than the presented method. The probable reason for better performance than KNN is that the SVM used for comparison is a class-weighted SVM, where the weights are proportional to the number of samples in each class. Therefore, SVM can address the imbalanced dataset by assigning higher misclassification penalties to training samples of the minority class (abnormal samples).

In addition, two popular machine learning algorithms are picked for additional comparison. The Gaussian Naive Bayes algorithm can address the imbalanced dataset challenge by introducing the class prior, the probability of each class, inferred from the training set. The Adaboost with decision tree uses the decision tree as the base algorithm, with a maximum depth of 500 and a total of 500 estimators. Both algorithms show a slight improvement compared to KNN, but still worse than SVM based on the average AUC scores shown in Fig. 6. Therefore, the presented ECD algorithm and its fusion method both show a much better performance-boosting than these typical machine learning algorithms in the literature.

The novelty and advantage of ECD detector over the four methods in comparison are that ECD detector both has localisation and ensemble attributes. Localisation acts as a feature selection, corresponding to each bus. Ensemble then accumulates the detection results from all local CorrDet detectors. Therefore, both local and overall statistics are fully captured by ECD detector, leading to an improved detection sensitivity through localisation and ensemble attributes, which the four algorithms in comparison lack.

5 Conclusion

This paper presents a hybrid physics model-based data-driven framework for the detection of FDI attacks on SG real-time monitoring. The physics model-based solution uses the state-of-the-art Innovation Concept of bad data detection and the novel, data-driven ECD algorithm is introduced to exploit both spatial and temporal characteristics of the SG. The ECD algorithm uses information from distributed, local CorrDet detectors to make a decision on whether or not there is an FDI in the system. By doing this, the data-driven solution not only uses temporal information, but spatial information as well. Decision level fusion is used to combine the information that each individual anomaly detection method contains, considering the confidence that each method has for a given sample.

The hybrid FDI detection framework was tested on the IEEE 118-bus system. Test results show that the fusion of the individual techniques has the best overall performance, detecting FDI attacks at a high rate without many false alarms, which can cause issues in a similar fashion to an undetected FDI attack. The presented hybrid framework improved mean AUC score for the testing set by 6.75% compared to physics model-based results. The presented framework also outperforms other popular machine learning algorithms mentioned in the literature for FDI attacks. The fusion of physics model-based data-driven solutions and of temporal and spatial information has been shown to be an improvement on the detection of FDI attacks on the SG, opening up opportunities for future research in this area.

6 Acknowledgment

This material is based upon work supported by the National Science Foundation under grant no. 1809739.

7 References

[1] Farag, M., Azab, M., Mokhtar, B.: 'Cross-layer security framework for smart grid: physical security layer'. IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), Istanbul, Turkey, 2014, pp. 1–7

[2] Monticelli, A.: '*State estimation in electric power systems: a generalized approach*', vol. 507 (Springer Science & Business Media, USA, 1999)

[3] Handschin, E., Schweppe, F.C., Kohlas, J., *et al.*: 'Bad data analysis for power system state estimation', *IEEE Trans. Power Appar. Syst.*, 1975, **94**, (2), pp. 329–337

[4] Bretas, N.G., Bretas, A.S., Piereti, S.A.: 'Innovation concept for measurement gross error detection and identification in power system state estimation', *IET Gener. Transm. Distrib.*, 2011, **5**, (6), pp. 603–608

[5] Liu, Y., Ning, P., Reiter, M.K.: 'False data injection attacks against state estimation in electric power grids'. Proc. of the 16th ACM Conf. on Computer and Communications Security, CCS '09, New York, NY, USA: ACM, 2009, pp. 21–32. Available at http://doi.acm.org/10.1145/1653662.1653666

[6] Kosut, O., Jia, L., Thomas, R.J., *et al.*: 'Malicious data attacks on smart grid state estimation: attack strategies and countermeasures'. 2010 First IEEE Int. Conf. on Smart Grid Communications, Gaithersburg, MD, USA, 2010, pp. 220–225

[7] Bobba, R.B., Rogers, K.M., Wang, Q., *et al.*: 'Detecting false data injection attacks on dc state estimation'. Preprints of the First Workshop on Secure Control Systems, CPSWEEK, Stockholm, Sweden, vol. 2010, 2010

[8] Bretas, A.S., Bretas, N.G., Carvalho, B., *et al.*: 'Smart grids cyber-physical security as a malicious data attack: an innovation approach', *Electr. Power Syst. Res.*, 2017, **149**, pp. 210–219. Available at http://www.sciencedirect.com/science/article/pii/S0378779617301657

[9] Ho, K.C., Gader, P.D.: 'Correlation-based land mine detection using gpr'. Proc. Volume 4038, Detection and Remediation Technologies for Mines and Minelike Targets V, (Int. Society for Optics and Photonics), 2000, pp. 1088–1096

[10] Chang, C.I., Chiang, S.S.: 'Anomaly detection and classification for hyperspectral imagery', *IEEE Trans. Geosci. Remote Sens.*, 2002, **40**, (6), pp. 1314–1325

[11] He, Y., Mendis, G.J., Wei, J.: 'Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism', *IEEE Trans. Smart Grid*, 2017, **8**, (5), pp. 2505–2516

[12] Ozay, M., Esnaola, I., Yarman Vural, F.T., *et al.*: 'Machine learning methods for attack detection in the smart grid', *IEEE Trans. Neural Netw. Learn. Syst.*, 2016, **27**, (8), pp. 1773–1786

[13] Wei, L., Gao, D., Luo, C.: 'False data injection attacks detection with deep belief networks in smart grid'. 2018 Chinese Automation Congress (CAC), Xi'an, China, 2018, pp. 2621–2625

[14] Potluri, S., Diedrich, C., Sangala, G.K.R.: 'Identifying false data injection attacks in industrial control systems using artificial neural networks'. 2017 22nd IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 2017, pp. 1–8

[15] Wang, Q., Tai, W., Tang, Y., *et al.*: 'Review of the false data injection attack against the cyber-physical power system', *IET Cyber-Phys. Syst.: Theory Appl.*, 2019, **4**, (2), pp. 101–107

[16] Trevizan, R.D., Ruben, C., Nagaraj, K., *et al.*: 'Data-driven physics-based solution for false data injection diagnosis in smart grids'. 2019 IEEE PES GM, 2019

[17] Liu, Y., Reiter, M.K., Ning, P.: 'False data injection attacks against state estimation in electric power grids', *IEEE Trans. Smart Grid*, 2019, **10**, (3), pp. 2871–2881

[18] Hug, G., Giampapa, J.A.: 'Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks'. *IEEE Trans. Smart Grid*, 2012, **3**, pp. 1362–1370

[19] Sandberg, H., Teixeira, A., Johansson, K.H.: 'On security indices for state estimators in power networks'. Proc. 1st Workshop Secure Control Systems, Stockholm, Sweden, 2010

[20] Esfahani, P.M., Vrakopoulou, M., Margellos, K., *et al.*: 'Cyber-attack in a two-area power system: impact identification using reachability'. Proc. American Control Conf. (ACC), Baltimore, MD, USA, 2010

[21] Esfahani, P.M., Vrakopoulou, M., Margellos, K., *et al.*: 'A robust policy for automatic generation control cyber-attack in two area power network'. Proc. 2010 49th IEEE Conf. Decision Control (CDC), Atlanta, GA, USA, 2010, pp. 5973–5978

[22] Xie, L., Mo, Y., Sinopoli, B.: 'False data injection attacks in electricity markets'. Proc. 1st IEEE Int. Conf. Smart Grid Communications (SmartGridComm), Gaithersburg, MD, USA, 2010

[23] Kosut, O., Liyan, J., Thomas, R.J., *et al.*: 'Malicious data attacks on smart grid state estimation: attack strategies and countermeasures'. Proc. 1st IEEE Int. Conf. Smart Grid Communications, Gaithersburg, MD, USA, 2010

[24] Bobba, R.B., Rogers, K.M., Wang, Q., *et al.*: 'Detecting false data injection attacks on dc state estimation'. Proc. 1st Workshop Secure Control Systems, Stockholm, Sweden, 2010

[25] Kim, T.T., Poor, H.V. 'Strategic protection against data injection attacks on power grids', *IEEE Trans. Smart Grid*, 2011, **2**, pp. 326–333

[26] Bretas, N.G., Bretas, A.S.: 'The extension of the Gauss approach for the solution of an overdetermined set of algebraic non linear equations', *IEEE Trans. Circuits Syst. II, Express Briefs*, 2018, **65**, (9), pp. 1269–1273

[27] Teixeira, A., Amin, S., Sandberg, H., *et al.*: 'Cyber security analysis of state estimators in electric power systems'. Proc. 16th ACM Conf. on Computer and Communications Security, Atlanta, GA, USA, 2010, pp. 5991–5998

[28] Bretas, A.S., Bretas, N.G., Carvalho, B.E.B.: 'Further contributions to smart grids cyber-physical security as a malicious data attack: proof and properties of the parameter error spreading out to the measurements and a relaxed correction model', *Int. J. Electr. Power Energy Syst.*, 2019, **104**, pp. 43–51. Available at http://www.sciencedirect.com/science/article/pii/S0142061518303946

[29] Wang, H., Ruan, J., Zhou, B., *et al.*: 'Dynamic data injection attack detection of cyber-physical power systems with uncertainties', *IEEE Trans. Ind. Inf.*, 2019, **15**, (10), pp. 5505–5518

[30] Xue, D., Jing, X., Liu, H.: 'Detection of false data injection attacks in smart grid utilizing elm-based OCON framework', *IEEE Access*, 2019, **7**, pp. 31762–31773

[31] Bretas, N.G., Bretas, A.S.: 'A two steps procedure in state estimation gross error detection, identification, and correction', *Int. J. Electr. Power Energy Syst.*, 2015, **73**, pp. 484–490. Available at http://www.sciencedirect.com/science/article/pii/S0142061515002495

[32] Bretas, N.G., Bretas, A.S., Martins, A.C.P.: 'Convergence property of the measurement gross error correction in power system state estimation, using geometrical background', *IEEE Trans. Power Syst.*, 2013, **28**, (4), pp. 3729–3736

[33] Ho, K., Gader, P.D.: 'A linear prediction land mine detection algorithm for handheld ground penetrating radar', *IEEE Trans. Geosci. Remote Sens.*, 2002, **40**, (6), pp. 1374–1384

[34] Group, I.P.: 'Index of data Illinois institute of technology', 2001, [accessed 2019-09-30]. Available at http://motor.ece.iit.edu/data/

[35] Zimmerman, R.D., Murillo-Sanchez, C.E., Thomas, R.J.: 'Matpower: steady-state operations, planning, and analysis tools for power systems research and education', *IEEE Trans. Power Syst.*, 2011, **26**, (1), pp. 12–19

[36] Bibbona, E., Panfilo, G., Tavella, P.: 'The Ornstein–Uhlenbeck process as a model of a low pass filtered white noise', *Metrologia*, 2008, **45**, p. S117

[37] Thierfelder, C.: 'The trending Ornstein–Uhlenbeck process and its applications in mathematical finance', Mathematical Finance, University of Oxford Mathematical Institute, 2015

[38] Bradley, A.P.: 'The use of the area under the roc curve in the evaluation of machine learning algorithms', *Pattern Recognit.*, 1997, **30**, (7), pp. 1145–1159. Available at http://dx.doi.org/10.1016/S0031-3203(96)00142-2

[39] Lou, W., Wang, X., Chen, F., *et al.*: 'Sequence based prediction of dna-binding proteins based on hybrid feature selection using random forest and Gaussian Naive Bayes', *PloS one*, 2014, **9**, (1), p. e86703

[40] Freund, Y., Schapire, R.E.: 'A decision-theoretic generalization of on-line learning and an application to boosting', *J. Comput. Syst. Sci.*, 1997, **55**, (1), pp. 119–139