energies

Review

Anomaly Detection in Power System State Estimation: Review and New Directions

Austin Cooper [1], Arturo Bretas [2,3,*] and Sean Meyn [1]

[1] Electrical and Computer Engineering Department, University of Florida, Gainesville, FL 32603, USA; austin.cooper@ufl.edu (A.C.); meyn@ece.ufl.edu (S.M.)
[2] Distributed Systems Group, Pacific Northwest National Laboratory, Richland, WA 99354, USA
[3] G2Elab, Grenoble INP, CNRS, Université Grenoble Alpes, 38000 Grenoble, France
* Correspondence: arturo.bretas@pnnl.gov

Abstract: Foundational and state-of-the-art anomaly-detection methods through power system state estimation are reviewed. Traditional components for bad data detection, such as chi-square testing, residual-based methods, and hypothesis testing, are discussed to explain the motivations for recent anomaly-detection methods given the increasing complexity of power grids, energy management systems, and cyber-threats. In particular, state estimation anomaly detection based on data-driven quickest-change detection and artificial intelligence are discussed, and directions for research are suggested with particular emphasis on considerations of the future smart grid.

Keywords: anomaly detection; cyber-security; false data injection; hypothesis testing; machine learning; power system monitoring; quickest-change detection; state estimation

check for updates

Citation: Cooper, A.; Bretas, A.; Meyn, S. Anomaly Detection in Power System State Estimation: Review and New Directions. *Energies* **2023**, *16*, 6678. https://doi.org/10.3390/en16186678

Academic Editors: Boštjan Polajžer, Davood Khodadad, Younes Mohammadi and Aleksey Paltsev

Received: 24 August 2023
Revised: 13 September 2023
Accepted: 14 September 2023
Published: 18 September 2023

1. Introduction

Since its introduction by Schweppe in the late 1960s [1,2], power system state estimation has proved an integral component of Energy Management Systems (EMSs). Schweppe's proposed nonlinear static state estimation (SSE) provides estimates of the actual network status, which could then be leveraged for subsequent analysis, including contingency evaluation and power flow studies [3]. Soon after, strategies for mitigating erroneous measurement data [4,5] were developed to ensure the fidelity of the power system state estimates. SSE and dynamic state estimation (DSE) both share a rich history of research [6–8]; however, SSE has seen more real-world implementation. Nevertheless, DSE shows great promise in having an enhancing role in legacy SSE-based EMS [9], especially with the increased adoption of synchrophasor measurements [10], and thus, anomaly-detection methods using both approaches are surveyed.

Numerous sources of state estimation error have been identified and formulated in the literature, including measurement, parameter, and topology discrepancies with respect to the system model. More recently, with the integration of EMS into sophisticated computer networks, the potential for cyber-security vulnerabilities became apparent. What new considerations must be made when bad data are malicious? Stealthy false data injection attacks [11], for example, were formulated as an exercise in fooling legacy bad-data-detection schemes. That said, attacks on cyber-physical systems have yielded very real consequences, including equipment damage and rolling blackouts [12]. Anomaly-detection techniques that can properly handle these manufactured instances of bad data, and thus improve bad data processing in state estimation generally, are surveyed in this review. This review also hopes to highlight some considerations for future approaches to anomaly detection in state estimation, including implementation-based research in the face of increasingly dynamic load and generation profiles, the complexity of distributed cyber-physical infrastructure, and pushes for combined SSE and DSE approaches for higher-fidelity EMS information to improve control, efficiency, and stability in the future smart grid. Because the field of anomaly detection covers a wide range of approaches, this survey limits its scope to power system state estimation, which is a central component of EMS and is expected to remain as such well into the future [9].

Articles selected for this review were chosen based on their impact on the power state estimation anomaly detection field. For earlier foundational works, the authors sought to include papers with lasting influence and citation impact for bad data detection generally. Particular emphasis was placed on real-world implementation in modern EMSs. More recent works required consideration of cyber-attacks and/or error types designed to circumvent the approaches of older works. Because many of these approaches have yet to be implemented in EMSs, selected papers required notable metrics of improvement compared to legacy detection methods.

The contributions of this work include:

- Providing a history of legacy bad data detection and error types in power system state estimation and the connection to newer detection approaches and cyber-attack types.
- Surveying various sources of state estimation cyber-threats and the challenges they pose to anomaly detection schemes.
- An overview of newer approaches for anomaly detection based on quickest-change detection and AI.
- Considerations for future research, including the incorporation of dynamic load profiles, autocorrelated data, and asynchronous measurements.

This review is organized as follows. Section 2 provides a brief theory of static and dynamic state estimation generally and the components used for bad data detection. Section 3 describes the theory and physical meaning behind three main types of error in state estimation: measurement, parameter, and topology. Section 4 outlines the traditional methodologies developed for bad data detection and identification, which often serve as a basis for many modern approaches. Section 5 discusses malicious data attacks designed specifically to circumvent traditional bad data detection. Section 6 describes more modern approaches that aim to overcome these pitfalls. Section 7 provides a summary and considerations for future work.

2. Power System State Estimation

2.1. Static State Estimation

One of the most used models to perform power system SE is the Weighted Least Squares (WLS) estimator [7]. A power system with n buses and d measurements can be modeled through a set of nonlinear algebraic equations in the measurement model:

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

where $\mathbf{z} \in \mathbb{R}^{1\times d}$ is the measurement vector, $\mathbf{x} \in \mathbb{R}^{1\times N}$ the state variables vector, $h : \mathbb{R}^{1\times N} \to \mathbb{R}^{1\times d}$ is a continuous nonlinear differentiable function, and $\mathbf{e} \in \mathbb{R}^{1\times d}$ is the measurement error vector. Each measurement error e_i is assumed to follow a zero mean Gaussian distribution. $N = 2n - 1$ is the number of unknown state variables, i.e., the complex voltages at each bus.

In the traditional WLS approach, the state vector estimate in (1) is determined by minimizing the weighted norm of the residual [13], represented with the cost function $J(\mathbf{x})$:

$$J(\mathbf{x}) = \|\mathbf{z} - h(\mathbf{x})\|^2_{\mathbf{W}} = [\mathbf{z} - h(\mathbf{x})]^T \mathbf{W} [\mathbf{z} - h(\mathbf{x})] \tag{2}$$

where $\mathbf{W} = \mathbf{R}^{-1}$ is the inverse covariance matrix of the measurements, otherwise known as the weight matrix.

Linearizing the measurement model (1) yields

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

where $\mathbf{H} = \frac{\partial h}{\partial \mathbf{x}}$ is the Jacobian matrix of h at the current state estimate. The estimate of the linearized state vector is then given by

$$\Delta\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\Delta\mathbf{z}. \tag{4}$$

The estimated value of the measurement vector mismatch $\Delta\mathbf{z}$ is given by

$$\Delta\hat{\mathbf{z}} = \mathbf{H}\Delta\hat{\mathbf{x}} = \mathbf{P}\Delta z. \tag{5}$$

where $\mathbf{P} = \mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}$ denotes the linear projection or "hat" matrix. The idempotent matrix $\mathbf{P}$ also has the following properties [7]:

$$\mathbf{P}\mathbf{H} = \mathbf{H} \tag{6a}$$

$$(\mathbf{I} - \mathbf{P})\mathbf{H} = 0. \tag{6b}$$

These properties facilitate an expression for the measurement residuals [8]:

$$\begin{aligned}
\mathbf{r} &= \Delta\mathbf{z} - \Delta\hat{\mathbf{z}} && \text{(7a)}\\
&= (\mathbf{I} - \mathbf{P})\Delta\mathbf{z} && \text{(7b)}\\
&= (\mathbf{I} - \mathbf{P})(\mathbf{H}\Delta\mathbf{x} + \mathbf{e}) && \text{(7c)}\\
&= (\mathbf{I} - \mathbf{P})\mathbf{e} \qquad \text{[Using Equation (6b)]} && \text{(7d)}\\
&= \mathbf{S}\mathbf{e} && \text{(7e)}
\end{aligned}$$

where $\mathbf{S}$ is known as the residual sensitivity matrix, which was first recognized in [5] for representing the sensitivity of the measurement residual to the measurement error during bad data processing. Also useful is the residual covariance matrix $\mathbf{\Omega}$ [7]:

$$\begin{aligned}
[\mathbf{r}] &= [\mathbf{S}\mathbf{e}] = 0 && \text{(8a)}\\
Cov[\mathbf{r}] &= \left[\mathbf{r}\mathbf{r}^T\right] = \mathbf{S}\left[\mathbf{e}\mathbf{e}^T\right]\mathbf{S}^T && \text{(8b)}\\
&= \mathbf{S}\mathbf{R} = \mathbf{\Omega}. && \text{(8c)}
\end{aligned}$$

The residual covariance matrix is used for the detection and identification of bad data, as well as providing insight into the degree of interaction; these concepts will be elaborated upon further in Section 3.

2.2. Dynamic State Estimation

SSE does not consider any history of the measurement vector $\mathbf{z}$, but instead provides a snapshot of the system. This "memoryless" assumption of SSE proved sufficient for real-time monitoring in early EMS. For one, power networks were not as regimented at the distribution level, with far fewer microgrids, distributed energy resources, and net load dynamics compared to today's systems. Secondly, the measurement data fed to the state estimator almost always came from measurement devices with slow sampling rates, such as the 2–4 s range of SCADA. One might argue, then, that the true bottleneck for capturing dynamic behavior in state estimation was slow metering rates. That said, Schweppe's formulation arrived just shortly after the introduction of the Kalman filter in 1961 [14], which inspired power researchers to seek formulations beyond the still-developing SSE. The practical hangup of slow meter sampling rates would be relieved somewhat with the introduction of synchronized phasor measurements in the 1980s [10]. Phasor Measurement Units (PMUs) provide higher sampling rates compared to SCADA but also GPS coordination to avoid the uncertainty associated with asynchronicity.

Like SSE, dynamic state estimation (DSE) encompasses a wide range of methods. Early DSE formulations considered the same set of measurements and state variables as those used in SSE: active and/or reactive power flow and injections and complex bus voltages. Other approaches seek to better capture load dynamics by considering generator rotor angle and speed as differential-algebraic state variables [9,15,16]; however, this review will primarily consider DSE-based anomaly-detection implementations that use algebraic state variables.

DSE can be accomplished by modeling the power system as a discrete-time dynamic system. The Kalman filter is used [17] to estimate the state variables at time k through prediction and measurement update steps upon each iteration:

Predict:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_k \hat{\mathbf{x}}_{k-1|k-1} \tag{9}$$

$$\mathbf{F}_{k|k-1} = \mathbf{A}_k \mathbf{F}_{k-1|k-1} \mathbf{A}_k^T + \mathbf{Q}_k. \tag{10}$$

Update:

$$\mathbf{K}_k = \mathbf{F}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{F}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \tag{11}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right) \tag{12}$$

$$\mathbf{F}_{k|k} = \mathbf{F}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{F}_{k|k-1} \tag{13}$$

where, at time k, $\mathbf{A}_k$ is the state transition matrix, $\mathbf{K}_k$ is the Kalman gain matrix, and $\mathbf{H}_k$ is the measurement matrix. $\mathbf{F}_{k|k}$ and $\mathbf{F}_{k|k-1}$ denote the state covariance matrix estimates based on measurements up to times k and $k-1$. $\mathbf{Q}_k$ and $\mathbf{R}_k$ are the process and observation noise covariance matrices, respectively.

The authors of the first Kalman filter power system DSE approach [18] hinted at its compatibility with anomaly-detection methods, which, at the time, were being studied for SSE. Early work soon after [19,20] formulated bad data detection by analyzing the innovation process:

$$\mathbf{v}_k = \mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}). \tag{14}$$

Additional approaches for bad data processing in DSE include asymmetry analysis based on the skewness of the normalized estimation error [17,21]. DSE anomaly detection research remains an active field [16,22], especially since dynamic load and generation profiles are commonplace in microgrid systems with distributed energy resources (DERs).

3. Bad Data Types and Considerations

Bad data can be classified as either single or multiple. For single bad data, one measurement in the system is corrupted with a large error. Multiple bad data describe more than one measurement being in error and can be further classified by the degree of interaction and conformity [7]. Multiple bad data are said to interact when the residuals are highly correlated, whereas conformity describes the degree to which gross errors are "masked" in the residual (i.e., nonconforming errors present as highly normalized residuals) [8]. Another illustration of how error is not always fully reflected in the residual is the concept of leverage points [23–26], which can hinder the effectiveness of the largest residual methods. Leverage points arise as a consequence of system topology, parameter values, and measurement placement and are usually caused by the following: (i) injection and flow measurements near branches with a small X/R ratio; (ii) injection measurements near buses with a large number of incident branches; and (iii) a measurement with a large weight [6,27]. Even a single leverage point can compromise bad data detectability.

Gross errors that exist beyond the acceptable noise limit of the state estimation model can be categorized into three types: measurement, parameter, and topology. Each of these errors suggests a discrepancy between the measurement data and model and are described further in the following.

3.1. Measurement Error

Measurement error is inevitable given the limitations of metering equipment accuracy. Meters can fail or degrade, introducing bias and compromising both accuracy and Gaussian error assumption: empirical studies of synchrophasor errors have yielded heavy-tailed error distributions such as Cauchy, Student's t, logistic, and Laplace [28,29]. Further, the communications infrastructure itself may contribute to measurement error in the case of failure or interference [7]. Particularly egregious measurement errors that suggest physically impossible grid conditions, such as negative bus voltage magnitudes or magnitudes several times larger or shorter than nominal values, are filtered through pre-processing [8], but more "agreeable" measurement errors can nevertheless affect the accuracy of state estimates.

3.2. Parameter Error

Parameter errors suggest discrepancies between measurement data and the system model. While Schweppe in his original formulation [1] did recognize the impact of erroneous model parameters, such errors were not considered in the network model. For example, a parameter error might arise when the variability in a line-impedance value due to extreme weather conditions is not taken into account. The mismatch between the measurement data and the line impedance database value, which is used in the Y-admittance matrix for power flow calculations, would be reflected in the state estimation result.

A simple alteration of (1) yields an augmented model [30] and linearization:

$$z_i = h_i(x, p_0) + e_i \approx h_i(x, p) + \frac{\partial h_i}{\partial p}\Delta p + e_i \tag{15}$$

where p is the true parameter value, p_0 is the erroneous parameter value, and $\Delta p = p_0 - p$ is the parameter error.

Stuart and Herget [31] investigated the impact of parameter errors on SSE by simulating erroneous values for line impedance, measurement error variance, and transformer tap settings. Of particular note was an observed relationship between the severity of error and lightly loaded lines.

Parameter errors can be thought of as a special case of multiple bad data in which only the measurements pertaining to the erroneous model parameter are in error. As such, studies have been performed with the goal of differentiating between the two. In [32], it was shown through analysis of the state estimation error distribution that parameter errors are reflected only in the measurement functions with erroneous parameter values. Parameter estimation itself has been treated as a process separate from state estimation. A practical implementation of this was first developed in [33], in which a sensitivity-based WLS estimation approach is used to both identify and estimate parameter error.

3.3. Topology Error

Like parameter errors, topology errors suggest discrepancies in the measurement model. System topology describes the bus-branch network configuration at the time of state estimation. Topology processing, which precedes state estimation, normally determines the correct status of manual switching and the circuit-breaking apparatus. A topological discrepancy, such as a branch outage unaccounted for by the topology processor, would be reflected in the Jacobian measurement matrix **H**, which requires accurate bus-branch connection logic for the calculation of power flow. Topology errors can significantly compromise state estimation accuracy through multiple conforming bad data [7]. Early work showed that such topology errors can be reflected in the state estimation error [34,35] and that normalized residual methods could be used for detection. Other approaches suggest incorporating the statuses of switching devices themselves as additional state variables [36], aiding in the *identification* of topology errors as such.

4. Bad Data Detection

To preserve the accuracy of state variable estimates, bad data must be detected, identified, and either eliminated or corrected. Whether the source of the bad data is measurement-, parameter-, or topology-based, detection is the first step. The classical components of bad data detection can be broadly categorized into three main branches and are often used in conjunction with one another: chi-square χ^2 testing, residual-based methods, and hypothesis testing.

4.1. Chi-Squared χ^2 Test

For a set of d random variables $\{X_i,\ i = 1, 2, \ldots, d\}$ with unit Gaussian distribution $X_i \sim \mathcal{N}(0, 1)$, a new random variable with χ^2 distribution is defined as $Y = \sum_{i=1}^{d} X_i^2$ [6]. This follows the form of the cost function defined in (2) and can be written as the performance index [8]

$$J(\hat{\mathbf{x}}) = \sum_{i=1}^{d} \left(\frac{z_i - h_i(\hat{x})}{\sigma_i} \right)^2 \tag{16}$$

assuming that the measurement errors are independent and distributed $e_i \sim \mathcal{N}(0, \sigma^2)$. $J(\hat{\mathbf{x}})$ then follows a χ^2 distribution with $d - N$ degrees of freedom, where d is the number of measurements and N is the number of unknown state variables.

A critical value $C = \chi^2_{(d-N),p}$ can then be obtained based on the degrees of freedom $d - N$ and the desired detection confidence with probability $p = 1 - \alpha$, where α is a constraint on false probability. If $J(\hat{\mathbf{x}}) \geq C$, then bad data are suspected; otherwise, the measurements are assumed to be free of bad data. χ^2 testing has proved valuable for the detection of bad data even in the early history of SSE [5], where it was quickly realized that χ^2 and normalized residual methods can outperform one another generally, but that χ^2 often proved better for multiple bad data.

4.2. Residual-Based Methods

The χ^2 test soon became commonplace for the *detection* of bad data detection in WLS SSE for a specified constraint on false probability α, after which residual analysis could be performed for the *identification* of the measurement(s) in error [37]. However, in the case of single bad data in larger networks, the analysis of both the weighted and normalized residuals also proved viable for detection due to a more pronounced response in the presence of gross errors when compared to χ^2 testing. The use of normalized residuals for bad data detection was introduced in [5]. Using the residual covariance matrix $\Omega_{ii} = diag(\mathbf{\Omega})$, the normalized residuals can be defined

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} \tag{17}$$

It was shown in [5] that, after bad data had been detected through means such as the χ^2 test, a list of the normalized residuals in descending order could be obtained. The *largest* normalized residual could be used to identify the measurement in error, after which the measurement was removed and the state estimation re-run. If bad data were still detected, the procedure would repeat until all erroneous measurements were eliminated. Further techniques were developed to correct measurements contaminated with bad data, rather than eliminating them [8]. Correction keeps the measurement structure intact, which is especially important in cases of limited redundancy.

Both the detection and identification of bad data can be achieved without χ^2 testing by comparing the largest normalized residual to a statistical threshold depending on the desired sensitivity [7]. The case studies in [5] demonstrated that, in the case of multiple bad data, either interacting or noninteracting, no consensus could be developed as to whether χ^2 testing or the largest normalized residual test proved superior for bad data detection. A geometric interpretation of the normalized residuals was developed in [38], significantly improving the generalizability of multiple interacting bad data detection. The residual difference between estimated and actual measurements continues to be a vital component in state estimation anomaly detection, including in newer formulations to be expanded upon in Section 6.

4.3. Hypothesis Testing

Hypothesis testing is a statistical method for deciding between accepting a null hypothesis H_0 or an alternative hypothesis H_1 based on available observations. In power system state estimation, the hypotheses are formulated as follows:

H_0: z_i is a valid measurement.
H_1: z_i is a measurement in error.

The first work to use hypothesis testing identification (HTI) for bad data in power system state estimation [39] developed regions of acceptance between H_0 and H_1 by comparing the estimation error to a threshold dependent on the measurement standard deviation and a pre-selected constraint on false probability α. New results of this HTI method were presented in [40], where the optimality of the linear estimator is established along with a decision strategy based on a constraint for missed detection β. In [41], the authors bridge the gaps between theory and practice by implementing the HTI on eight test systems, showcasing its strengths in detecting multiple interacting bad data. For bad data identification, HTI methods show significant advantages over methods based on normalized residuals, which may be strongly correlated [7]. HTI techniques also demonstrated potential for discerning error type, such as in topology error identification [42,43].

5. When Bad Data Become Malicious

The introduction of the concept of false data injection attacks (FDIAs) [11] helped to highlight the limitations of classical bad-data-detection methods. What if bad data are malicious and/or statistically derived to avoid conventional detection? The basic idea of FDIAs is that an attacker can design an injection of multiple interacting bad data, which is then applied to the measurement vector $\mathbf{z}$. Consider the representation $\mathbf{z_a} = \mathbf{z} + \mathbf{a}$, where $\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$ is a vector of malicious data. The attacker's goal is to design $\mathbf{a}$ to alter the state estimates, which EMSs use to make operating decisions, but without triggering bad data detection. Ramifications of undetected attacks include compromised system stability [12] and negative economic impact [44]. The success of such attacks is largely dependent on the information available to the attacker, such as the number of meters compromised, state estimates, system topology, and Jacobian structure, to name a few.

Denial-of-service (DoS) attacks are another source of mismatch between the measurement data fed to the state estimator and the true power system state. Causes for DoS attacks are numerous [45], including communication channel jamming, packet flooding, and compromising of metering devices such as SCADA and PMUs so that data are not updated for that region of the power grid. For state estimation, DoS attacks are typically modeled as a set of measurements that are no longer available, which can negatively impact state variable accuracy. If stealthiness is desired, care would need to be taken on the attacker's part so as not to render the system unobservable. FDIAs can also be designed to create a topology error attack [46–48], in which a conventionally nondetectable mismatch between measurement data and topology processing can lead to compromised system stability and cost-effective operation.

The authors of [49] present FDIA strategies from the attacker and defender perspectives. For the attacker, it is typically assumed that there is a cost associated with the information obtained. With this in mind, an algorithm is presented to find the minimal set of measuring devices required to manufacture an unobservable attack. In [50], a comparative analysis of the FDIA impact between so-called DC and AC SSE is conducted. DC SSE considers active power measurements only, with bus voltage angles as the state variables. In contrast, the complete AC SSE considers both active and reactive power measurements, with bus voltage magnitudes and angles as the state variables. Such a study was important due to the DC model warranting far more attention in the FDIA research space at the time, despite the full nonlinear AC model finding use in real-world EMS applications [51,52].

Impacts of FDIAs on Kalman filter DSE approaches were studied in [53], where it was found that the unscented Kalman filter (UKF) [54] yielded better performance compared to the extended Kalman filter (EKF) [55] and the enhanced EKF [56]. Further, an online nonparametric cumulative sum (CUSUM) approach was proposed to detect anomalies based on distribution changes in the state estimation error. This is related to quickest-change detection approaches, which will be elaborated upon further in Section 6.1. A Kalman filter state estimation approach was proposed in [57], where a Euclidean detector was used to overcome the shortcomings of the χ^2 test for detecting statistically derived FDIAs as well as DoS attacks.

The FDIA formulation highlighted a need for improved bad data detection. The classification of bad data as such would also need improvement. Common confusion matrix metrics like false negatives and false positives become harder to minimize when stealth FDIAs can closely resemble power system events like transients, switching, and sudden load changes. Further, with the increasing push towards the cyber-physical operation of the smart grid [58], many new points of entry for cyber-attack became apparent, such as Internet of Things (IoT) infrastructure [59], communication channels [60], and distributed computing [61]. The intersection of model-based and data-driven solutions should grow to better handle the bad data detection limitations posed by FDIAs. With state estimation anticipated to remain a vital component of EMSs, new formulations based on quickest-change detection and AI should be developed for improved anomaly detection.

6. Recent Approaches

6.1. Quickest-Change Detection

Quickest-change detection (QCD) is concerned with detecting a possible change in the distribution of a monitored observation sequence [62], which is indicative of an anomaly in a stochastic environment. The general goal of QCD theory is to design algorithms to detect these changes with the smallest detection delay possible, subject to a constraint on false alarms.

Three main ingredients are needed in the QCD problem [63]: an observed stochastic process $\{X_n,\ n = 1, 2, \ldots\}$, a change time τ^{a} at which the statistical properties of the process undergo change, and a decision maker that declares a change time τ^{s} based on observations of the stochastic process. A false alarm is defined as an instance of the decision maker declaring a change before the change occurs: $I\{\tau^{\mathrm{s}} < \tau^{\mathrm{a}}\}$. The constraint on false alarm follows from the Neyman–Pearson hypothesis testing formulation [64], which is foundational to the QCD problem.

The Neyman–Pearson Lemma [65] establishes the optimal test for binary hypothesis testing, involving the null (H_0) and alternate (H_1) hypotheses. For a single observation X:

H_0: X has pdf p.
H_1: X has pdf q.

Then, comparing the *likelihood ratio* $q(X)/p(X)$ to a threshold value is the most powerful test in terms of deciding which hypothesis is true while minimizing missed detection subject to a constraint on false alarms [66]. The likelihood ratio plays a fundamental role in recursive sequential-change-detection algorithms such as Page's CUSUM [67] and the Shiryaev–Roberts procedure [68], each of which enjoys optimality properties in terms of minimizing false alarm and detection delay $(\tau^{\mathrm{s}} - \tau^{\mathrm{a}})_+ \max(0, \tau^{\mathrm{s}} - \tau^{\mathrm{a}})$. These properties are given proper discussion in [62].

QCD approaches have shown great promise for power system anomaly detection applications, such as line outage detection and identification [69–71]. QCD has further application in detecting changes in the state estimation error, which has been proposed for fault and FDIA detection. The first QCD approach for state estimation FDIA detection implemented an adaptive approach using the CUSUM statistic:

$$S_n = \max\{0, S_{n-1} + L(Z_n)\}\,, \qquad n \geq 1\,,\ \textit{with}\ S_0 = 0. \tag{18}$$

where $\{Z_n,\ n = 1,2,\ldots\}$ is the observed stochastic process and L is the log-likelihood ratio. Sample plots of a subtle change in a Gaussian observation process, along with the corresponding CUSUM statistic, are included in Figure 1.

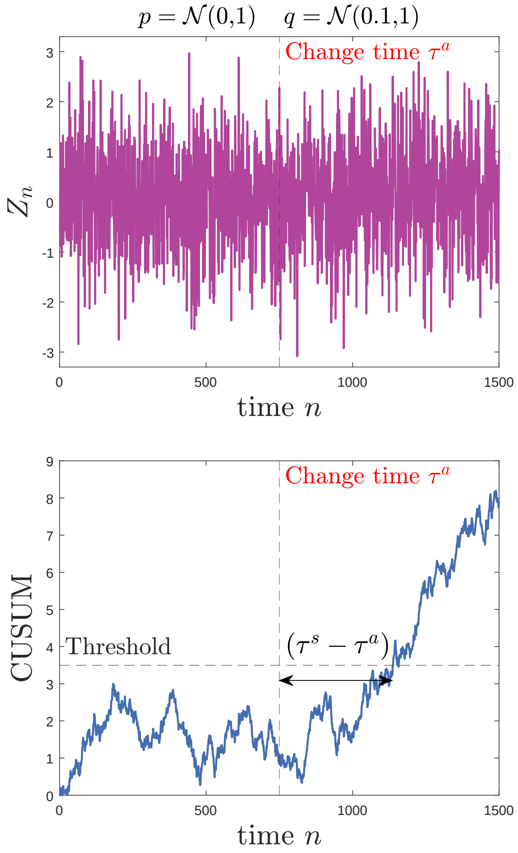

Figure 1. Example of a small mean shift observation sequence with the corresponding CUSUM evolution.

Because the exact form of the post-change distribution q is not known, the authors in [72,73] used a Rao test-based approximation [74] of the generalized likelihood ratio test for CUSUM-based FDIA detection. A low-complexity Orthogonal Matching Pursuit CUSUM (OMP-CUSUM) approach in [75] accounts for the unknown change distribution by maximizing the cumulative log-likelihood ratio to detect FDIAs that are sparse (i.e., only a small number of meters are assumed accessible to the attacker).

Both centralized and distributed CUSUM-based approaches for FDIA detection are proposed in [76], replacing the unknown parameters of the post-change distribution with their maximum likelihood estimates (MLEs). For the centralized case, the observed stochastic process of interest is the projection of the measurement vector on the orthogonal Jacobian space component $\mathcal{R}^{\perp}(\mathbf{H})$. This is expressed as $\tilde{\mathbf{y}}_n \triangleq \mathbf{P}_n\mathbf{y}_n$, where $\mathbf{P}$ is the previously defined linear projection matrix. The distributed case partitions the power system into areas and estimates the state variables through the alternating direction method of multipliers (ADMM) [77], where each area i has its own observed process $\{\tilde{\mathbf{y}}_n^i,\ n = 1,2,\ldots\}$. These approaches outperformed the adaptive-CUSUM approach in [72,73], due in part to the improved detection of FDIAs with negative and larger elements of the attack vector $\mathbf{a}$.

The work in [78] incorporates a Kalman filter approach and separately evaluates DoS attacks and FDIAs. Better detection performance was observed for stealth FDIAs in particular, in which perfect system topology knowledge allows an attacker to inject false data along the column space of **H**. Four Kalman filtering techniques in [53] were evaluated using nonparametric CUSUM, in which both pre- and post-change distributions p and q are unknown. Hybrid FDIA/jamming attacks are assessed for the Kalman filter CUSUM-based detector in [79]. The distinction between persistent and non-persistent attacks was made as well. Most CUSUM-based detectors assume persistence in the change in the observed stochastic process, and so an intermittent attack series could be designed to increase the detection delay. Thus, the Generalized Shewhart Test, which can detect significant increases in L, is presented as a countermeasure against stealthy, non-persistent FDIAs. A relaxed generalized CUSUM (RGCUSUM) algorithm is presented in [80] for FDIA detection. A relaxation on maximizing the post-change likelihood over the unknown parameters yielded a more computationally efficient algorithm than its generalized CUSUM counterpart. A normalized Rao CUSUM-based detector with a time-varying dynamic model was employed in [81] to better distinguish between FDIA and sudden load changes.

The work in [82] also assesses the Shiryaev–Roberts (SR) procedure, along with CUSUM for change detection. In contrast to CUSUM, the optimality of the SR procedure pertains to detecting τ at a distant time horizon [83,84]. The SR procedure is defined recursively as

$$T_n = \exp\big(L(Z_n)\big)\left[T_{n-1} + 1\right], \qquad n \geq 1\,, \ \ \textit{with}\ T_0 = 0. \tag{19}$$

Further, the modified CUSUM and SR procedure algorithms [85] are employed in the same work as evaluation benchmarks for a so-called DeepQCD algorithm for online cyber-attack detection, which uses deep recurrent neural networks to detect changes in transient cases and with autocorrelated observations.

6.2. AI Approaches

FDIA detection can be framed as a binary classification problem in which the measurement vector **z** is determined to be either normal (negative class) or anomalous (positive class). One of the first to use semi-supervised and supervised learning for FDIA detection [86] explored perceptron, support vector machine (SVM), k-nearest neighbors (k-NN), and sparse logistic regression algorithms for supervised learning. Semi-supervised learning, in which unlabelled test data are incorporated in training, was explored with semi-supervised SVMs. Many valuable takeaways were garnered from this work, including considerations of power system size and and computational complexity; however, stealthy FDIAs were not considered. An Extended Nearest Neighbors (ENN) algorithm was proposed in [87] to better handle the imbalanced data problem (i.e., cases in which the number of negative class samples greatly exceeds or is significantly less than the number of positive class samples). Classification performance was then compared to SVM and k-NN algorithms. The work in [88] used a method based on the margin-setting algorithm, typically used in image processing applications, in which hypersphere decision boundaries were formed through labeled PMU time-series data. The MSA approach yielded superior classification performance compared to standard artificial neural networks (ANNs) and SVM.

Unsupervised principal component analysis (PCA) showed utility in the construction of stealthy and blind FDIAs, as well as in developing robust detection methods [89,90]. PCA is again employed in [91] as a preprocessing step to project higher-dimensional correlated measurement data to a lower dimension, removing the correlation between data and magnifying the distance between normal and anomalous measurements. For performance comparison, the authors implemented a supervised distributed ADMM-based SVM, which could only outperform the PCA-based anomaly detection when the training set was large. Mahalanobis distance-based ensemble detection methods demonstrated success for FDIA detection in [92–95], including in high-fidelity real-time simulation.

Reinforcement learning (RL)-based QCD approaches are explored in [82,96]. The QCD problem can be formulated as a case of optimal stopping, in which a decision to exercise must be made to minimize cost [97,98]. In QCD, this is understood as declaring a stop time τ^s at a cost relative to the actual stop time τ^a. For the Markov Decision Process (MDP) component of RL, one can either seek to maximize reward or minimize cost [99]. Two components for the cost are constructed [97]: one for continuing (associated with missed detection) and one for stopping (associated with false alarm). The authors in [96] use a model-free state–action–reward–state–action (SARSA) approach to learn the expected future cost for each state–action pair in a Q-table. The authors opt for a quantization scheme for learning when faced with the continuous observation space. Because the actual change time τ^a is a hidden state, a partially observable Markov decision process (POMDP) formulation is used. This RL approach significantly outperformed the Euclidean [57] and cosine-similarity metric [100]-based detectors in terms of minimizing the mean probability of false alarm and detection delay for various cyber-attack types, including hybridFDI/jamming, DoS, and network topology attacks.

Neural network and deep learning approaches also show promise for malicious and standard bad data detection. A Deep-Belief-Network-based classifier is proposed in [101] using Conditional Gaussian–Bernoulli Restricted Boltzmann Machines in the hopes of revealing higher-dimensional temporal features of stealthy FDIAs. The temporal correlation between measurements with the state estimator is analyzed through Recurrent Neural Networks (RNNs) for FDIA detection in [102]. A nonlinear autoregressive exogenous (NARX) model configuration for ANNs is explored in [103] for stealthy optimized FDIA detection. The authors in [104] consider a limited set of target labels for attacked measurement data, an example of semi-supervised learning. Autoencoders, used for dimensionality reduction and feature extraction, are integrated into a generative adversarial network. The framework compensates for the limited labeled data set by using two neural networks: one generative, responsible for creating fake samples, and the other discriminative, responsible for distinguishing between real and generated samples.

7. Conclusions and Suggestions for Future Work

A survey of legacy bad-data-detection procedures has been presented along with limitations with respect to malicious bad data. Cyber-attack formulations such as FDIA highlight the need for better data detection by pointing out the theoretical manipulation of grid-operating procedures by bad actors. Even if one argues that the FDIA formulation is more of a theoretical exercise than a practical concern, it still points to shortcomings in legacy bad data detection. Standard bad data and physical line faults under the leverage point conditions discussed earlier are difficult to detect for similar reasons as statistically derived stealth FDIAs. Newer methods such as QCD and AI seek to overcome legacy bad-data-detection techniques by leveraging features such as measurement data temporal patterns and probability density changes in the state estimation error.

Increased access to real state estimation measurement data would aid greatly in accessing the practicality of QCD and AI anomaly-detection formulations. For example, a QCD formulation assuming independent and identically distributed (i.i.d.) observations may be compromised under dynamic load and generation profiles, in which case the measurement data exhibit complicating factors like autocorrelation, as investigated in [82]. The robustness of newer anomaly detection strategies to asynchronous measurement data should also be investigated. Until synchronized measurement data for state estimation become standard, uncertainty quantification of this type should considered so as not to be considered a false-positive source of anomalous behavior. The availability of time-series data such as SCADA and/or PMU measurements for multi-bus systems would aid state estimation researchers in quantifying uncertainty and measurement correlation. It is also recommended that future work incorporate dynamic load and generation profiles to better reflect the future directions of the modern smart grid. This was a motivation in the work [81], which highlighted the importance of discerning anomalies from dynamic behavior such as large load shifts. Such conditions are expected to increase with more DER penetration in the future smart grid and should be included when evaluating detection and identification metrics.

Author Contributions: Conceptualization, A.C. and A.B.; methodology, A.C.; software, A.C.; validation, A.C., A.B. and S.M.; formal analysis, A.C.; investigation, A.C.; resources, A.C.; data curation, A.C.; writing—original draft preparation, A.C.; writing—review and editing, A.B.; visualization, A.C.; supervision, A.B.; project administration, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the U.S. Department of Energy under Contract DE-AC05-76RL01830.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schweppe, F.C.; Wildes, J. Power System Static-State Estimation, Part I: Exact Model. *IEEE Trans. Power Appar. Syst.* **1970**, *PAS-89*, 120–125. [CrossRef]
2. Filho, M.; da Silva, A.; Falcao, D. Bibliography on power system state estimation (1968–1989). *IEEE Trans. Power Syst.* **1990**, *5*, 950–961. [CrossRef]
3. Schellstede, G.; Beissler, G. A Software Package for Security Assessment Functions. In *Power Systems and Power Plant Control*; Pingyang, W., Ed.; IFAC Symposia Series; Pergamon: Oxford, UK, 1987; pp. 277–284. [CrossRef]
4. Merrill, H.M.; Schweppe, F.C. Bad Data Suppression in Power System Static State Estimation. *IEEE Trans. Power Appar. Syst.* **1971**, *PAS-90*, 2718–2725. [CrossRef]
5. Handschin, E.; Schweppe, F.; Kohlas, J.; Fiechter, A. Bad data analysis for power system state estimation. *IEEE Trans. Power Appar. Syst.* **1975**, *94*, 329–337. [CrossRef]
6. Monticelli, A. *State Estimation in Electric Power Systems: A Generalized Approach*; Springer: Berlin, Germany, 2012.
7. Abur, A.; Expósito, A.G. *Power System State Estimation: Theory and Implementation*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2004. [CrossRef]
8. Bretas, A.; Bretas, N.; London, J.B., Jr.; Carvalho, B. *Cyber-Physical Power Systems State Estimation*; Elsevier: Amsterdam, The Netherlands, 2021.
9. Zhao, J.; Gómez-Expósito, A.; Netto, M.; Mili, L.; Abur, A.; Terzija, V.; Kamwa, I.; Pal, B.; Singh, A.K.; Qi, J.; et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. *IEEE Trans. Power Syst.* **2019**, *34*, 3188–3198. [CrossRef]
10. Phadke, A. Synchronized phasor measurements-a historical overview. In Proceedings of the IEEE/PES Transmission and Distribution Conference and Exhibition, Yokohama, Japan, 6–10 October 2002; Volume 1, pp. 476–479. [CrossRef]
11. Liu, Y.; Ning, P.; Reiter, M.K. False Data Injection Attacks against State Estimation in Electric Power Grids. *ACM Trans. Inf. Syst. Secur.* **2011**, *14*, 1–33. [CrossRef]
12. Musleh, A.S.; Chen, G.; Dong, Z.Y. A Survey on the Detection Algorithms for False Data Injection Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2020**, *11*, 2218–2234. [CrossRef]
13. 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*, 1269–1273. [CrossRef]
14. Kalman, R.E.; Bucy, R.S. New Results in Linear Filtering and Prediction Theory. *J. Basic Eng.* **1961**, *83*, 95–108. [CrossRef]
15. Zhao, J.; Singh, A.K.; Mir, A.S.; Taha, A.; Rouhani, A.; Gomez-Exposito, A.; Meliopoulos, A.; Pal, B.; Kamwa, I.; Qi, J.; et al. *Power System Dynamic State and Parameter Estimation-Transition to Power Electronics-Dominated Clean Energy Systems: IEEE Task Force on Power System Dynamic State and Parameter Estimation*; IEEE: Piscataway, NJ, USA, 2021.
16. Liu, Y.; Singh, A.K.; Zhao, J.; Meliopoulos, A.P.S.; Pal, B.; Ariff, M.A.b.M.; Van Cutsem, T.; Glavic, M.; Huang, Z.; Kamwa, I.; et al. Dynamic State Estimation for Power System Control and Protection. *IEEE Trans. Power Syst.* **2021**, *36*, 5909–5921. [CrossRef]
17. Bretas, N. An iterative dynamic state estimation and bad data processing. *Int. J. Electr. Power Energy Syst.* **1989**, *11*, 70–74. [CrossRef]
18. Debs, A.S.; Larson, R.E. A Dynamic Estimator for Tracking the State of a Power System. *IEEE Trans. Power Appar. Syst.* **1970**, *PAS-89*, 1670–1678. [CrossRef]
19. Nishiya, K.I.; Takagi, H.; Hasegawa, J.; Koike, T. Dynamic state estimation for electric power systems—introduction of a trend factor and detection of innovation processes. *Electr. Eng. Jpn.* **1976**, *96*, 79–87. [CrossRef]
20. Nishiya, K.; Hasegawa, J.; Koike, T. Dynamic state estimation including anomaly detection and identification for power systems. *IEE Proc. Gener. Transm. Distrib.* **1982**, *129*, 192–198. [CrossRef]
21. Bretas, A.S.; Bretas, N.G.; Massignan, J.A.D.; London Junior, J.B.A. Hybrid Physics-Based Adaptive Kalman Filter State Estimation Framework. *Energies* **2021**, *14*, 6787. [CrossRef]

22. Jin, Z.; Zhao, J.; Ding, L.; Chakrabarti, S.; Gryazina, E.; Terzija, V. Power system anomaly detection using innovation reduction properties of iterated extended kalman filter. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107613. [CrossRef]
23. Mili, L.; Phaniraj, V.; Rousseeuw, P. Least median of squares estimation in power systems. *IEEE Trans. Power Syst.* **1991**, *6*, 511–523. [CrossRef] [PubMed]
24. Celik, M.; Abur, A. A robust WLAV state estimator using transformations. *IEEE Trans. Power Syst.* **1992**, *7*, 106–113. [CrossRef]
25. Majumdar, A.; Pal, B.C. Bad Data Detection in the Context of Leverage Point Attacks in Modern Power Networks. *IEEE Trans. Smart Grid* **2018**, *9*, 2042–2054. [CrossRef]
26. Mili, L.; Cheniae, M.; Vichare, N.; Rousseeuw, P. Robust state estimation based on projection statistics [of power systems]. *IEEE Trans. Power Syst.* **1996**, *11*, 1118–1127. [CrossRef]
27. Zhao, J.; Mili, L. Vulnerability of the Largest Normalized Residual Statistical Test to Leverage Points. *IEEE Trans. Power Syst.* **2018**, *33*, 4643–4646. [CrossRef]
28. Wang, S.; Zhao, J.; Huang, Z.; Diao, R. Assessing Gaussian Assumption of PMU Measurement Error Using Field Data. *IEEE Trans. Power Deliv.* **2018**, *33*, 3233–3236. [CrossRef]
29. Huang, C.; Thimmisetty, C.; Chen, X.; Stewart, E.; Top, P.; Korkali, M.; Donde, V.; Tong, C.; Min, L. Power Distribution System Synchrophasor Measurements with Non-Gaussian Noises: Real-World Data Testing and Analysis. *IEEE Open Access J. Power Energy* **2021**, *8*, 223–228. [CrossRef]
30. Zarco, P.; Exposito, A. Power system parameter estimation: A survey. *IEEE Trans. Power Syst.* **2000**, *15*, 216–222. [CrossRef]
31. Stuart, T.A.; Herczet, C.J. A Sensitivity Analysis of Weighted Least Squares State Estimation for Power Systems. *IEEE Trans. Power Appar. Syst.* **1973**, *PAS-92*, 1696–1701. [CrossRef]
32. Bretas, A.S.; Bretas, N.G.; Carvalho, B.E. 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*, 43–51. [CrossRef]
33. Liu, W.H.E.; Lim, S.L. Parameter error identification and estimation in power system state estimation. *IEEE Trans. Power Syst.* **1995**, *10*, 200–209. [CrossRef]
34. Costa, I.; Leao, J. Identification of topology errors in power system state estimation. *IEEE Trans. Power Syst.* **1993**, *8*, 1531–1538. [CrossRef]
35. Wu, F.; Liu, W.H. Detection of topology errors by state estimation (power systems). *IEEE Trans. Power Syst.* **1989**, *4*, 176–183. [CrossRef]
36. Korres, G.N.; Manousakis, N.M. A state estimation algorithm for monitoring topology changes in distribution systems. In Proceedings of the 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, USA, 22–26 July 2012; pp. 1–8. [CrossRef]
37. Koglin, H.J.; Neisius, T.; Beiβler, G.; Schmitt, K. Bad data detection and identification. *Int. J. Electr. Power Energy Syst.* **1990**, *12*, 94–103. [CrossRef]
38. Clements, K.A.; Davis, P.W. Multiple Bad Data Detectability and Identifiability: A Geometric Approach. *IEEE Trans. Power Deliv.* **1986**, *1*, 355–360. [CrossRef]
39. Cutsem, T.V.; Ribbens-Pavella, M.; Mili, L. Hypothesis Testing Identification: A New Method For Bad Data Analysis In Power System State Estimation. *IEEE Trans. Power Appar. Syst.* **1984**, *PAS-103*, 3239–3252. [CrossRef]
40. Mili, L.; Van Cutsem, T.; Ribbens-Pavella, M. Decision Theory Applied to Bad Data Identification in Power System State Estimation. In Proceedings of the 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, York, UK, 3–7 July 1985; Volume 18, pp. 945–950. [CrossRef]
41. Mili, L.; Van Cutsem, T. Implementation of the hypothesis testing identification in power system state estimation. *IEEE Trans. Power Syst.* **1988**, *3*, 887–893. [CrossRef]
42. Lourenco, E.; Costa, A.; Clements, K. Bayesian-based hypothesis testing for topology error identification in generalized state estimation. *IEEE Trans. Power Syst.* **2004**, *19*, 1206–1215. [CrossRef]
43. Wu, W.B.; Cheng, M.X.; Gou, B. A Hypothesis Testing Approach for Topology Error Detection in Power Grids. *IEEE Int. Things J.* **2016**, *3*, 979–985. [CrossRef]
44. Xie, L.; Mo, Y.; Sinopoli, B. Integrity Data Attacks in Power Market Operations. *IEEE Trans. Smart Grid* **2011**, *2*, 659–666. [CrossRef]
45. Amin, S.; Cárdenas, A.A.; Sastry, S.S. Safe and Secure Networked Control Systems under Denial-of-Service Attacks. In Proceedings of the Hybrid Systems: Computation and Control, San Francisco, CA, USA, 13–15 April 2009; Majumdar, R., Tabuada, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 31–45.
46. Kim, J.; Tong, L. On Topology Attack of a Smart Grid: Undetectable Attacks and Countermeasures. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 1294–1305. [CrossRef]
47. Liu, X.; Li, Z. Local Topology Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2017**, *8*, 2617–2626. [CrossRef]
48. Liang, G.; Weller, S.R.; Luo, F.; Zhao, J.; Dong, Z.Y. Generalized FDIA-Based Cyber Topology Attack with Application to the Australian Electricity Market Trading Mechanism. *IEEE Trans. Smart Grid* **2018**, *9*, 3820–3829. [CrossRef]
49. Kosut, O.; Jia, L.; Thomas, R.J.; Tong, L. Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures. In Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications, Gaithersburg, MD, USA, 4–6 October 2010; pp. 220–225. [CrossRef]

50. 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*, 1362–1370. [CrossRef]
51. Nuthalapati, S. State Estimation Performance Monitoring. 2015. Available online: https://www.nerc.com/pa/rrm/Resources/Monitoring_and_Situational_Awareness_Conference1/10 (accessed on 3 June 2023).
52. ETAP State Estimation Software. 2015. Available online: https://etap.com/product/state-estimation-software (accessed on 3 June 2023).
53. Yang, Q.; Chang, L.; Yu, W. On false data injection attacks against Kalman filtering in power system dynamic state estimation. *Secur. Commun. Netw.* **2016**, *9*, 833–849. [CrossRef]
54. Valverde, G.; Terzija, V. Unscented Kalman filter for power system dynamic state estimation. *Iet Gener. Transm. Distrib.* **2011**, *5*, 29–37. [CrossRef]
55. Ghahremani, E.; Kamwa, I. Dynamic State Estimation in Power System by Applying the Extended Kalman Filter with Unknown Inputs to Phasor Measurements. *IEEE Trans. Power Syst.* **2011**, *26*, 2556–2566. [CrossRef]
56. Shih, K.R.; Huang, S.J. Application of a robust algorithm for dynamic state estimation of a power system. *IEEE Trans. Power Syst.* **2002**, *17*, 141–147. [CrossRef]
57. Manandhar, K.; Cao, X.; Hu, F.; Liu, Y. Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using Kalman Filter. *IEEE Trans. Control. Netw. Syst.* **2014**, *1*, 370–379. [CrossRef]
58. Faheem, M.; Shah, S.; Butt, R.; Raza, B.; Anwar, M.; Ashraf, M.; Ngadi, M.; Gungor, V. Smart grid communication and information technologies in the perspective of Industry 4.0: Opportunities and challenges. *Comput. Sci. Rev.* **2018**, *30*, 1–30. [CrossRef]
59. Yilmaz, Y.; Uludag, S. Mitigating IoT-based Cyberattacks on the Smart Grid. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; pp. 517–522. [CrossRef]
60. Yan, Y.; Qian, Y.; Sharif, H.; Tipper, D. A Survey on Cyber Security for Smart Grid Communications. *IEEE Commun. Surv. Tutor.* **2012**, *14*, 998–1010. [CrossRef]
61. Kurt, M.N.; Yılmaz, Y.; Wang, X. Secure Distributed Dynamic State Estimation in Wide-Area Smart Grids. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 800–815. [CrossRef]
62. Xie, L.; Zou, S.; Xie, Y.; Veeravalli, V.V. Sequential (Quickest) Change Detection: Classical Results and New Directions. *IEEE J. Sel. Areas Inf. Theory* **2021**, *2*, 494–514. [CrossRef]
63. Veeravalli, V.V.; Banerjee, T. Quickest-change detection. In *Academic Press Library in Signal Processing*; Elsevier: Amsterdam, The Netherlands, 2014; Volume 3, pp. 209–255. [CrossRef]
64. Poor, H.V. *An Introduction to Signal Detection and Estimation*, 2nd ed.; Springer Texts in Electrical Engineering; Springer: New York, NY, USA, 1994.
65. Neyman, J.; Pearson, E.S. IX. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. London Ser. Contain. Pap. Math. Phys. Character* **1933**, *231*, 289–337. [CrossRef]
66. Moulin, P.; Veeravalli, V.V. *Statistical Inference for Engineers and Data Scientists*; Cambridge University Press: Cambridge, UK, 2018. [CrossRef]
67. Page, E.S. Continuous Inspection Schemes. *Biometrika* **1954**, *41*, 100–115. [CrossRef]
68. Polunchenko, A.S.; Tartakovsky, A.G. On Optimality of the Shiryaev-Roberrts Procedure for Detecting a Change in Distribution. *Ann. Stat.* **2010**, *38*, 3445–3457. [CrossRef]
69. Banerjee, T.; Chen, Y.C.; Dominguez-Garcia, A.D.; Veeravalli, V.V. Power system line outage detection and identification—A quickest change detection approach. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 3450–3454. [CrossRef]
70. Rovatsos, G.; Jiang, X.; Domínguez-García, A.D.; Veeravalli, V.V. Comparison of statistical algorithms for power system line outage detection. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2946–2950. [CrossRef]
71. Yang, X.; Chen, N.; Zhai, C. A Control Chart Approach to Power System Line Outage Detection Under Transient Dynamics. *IEEE Trans. Power Syst.* **2021**, *36*, 127–135. [CrossRef]
72. Huang, Y.; Li, H.; Campbell, K.A.; Han, Z. Defending false data injection attack on smart grid network using adaptive CUSUM test. In Proceedings of the 2011 45th Annual Conference on Information Sciences and Systems, Baltimore, MD, USA, 23–25 March 2011; pp. 1–6. [CrossRef]
73. Huang, Y.; Tang, J.; Cheng, Y.; Li, H.; Campbell, K.A.; Han, Z. Real-Time Detection of False Data Injection in Smart Grid Networks: An Adaptive CUSUM Method and Analysis. *IEEE Syst. J.* **2016**, *10*, 532–543. [CrossRef]
74. De Maio, A. Rao Test for Adaptive Detection in Gaussian Interference with Unknown Covariance Matrix. *IEEE Trans. Signal Process.* **2007**, *55*, 3577–3584. [CrossRef]
75. Akingeneye, I.; Wu, J. Low Latency Detection of Sparse False Data Injections in Smart Grids. *IEEE Access* **2018**, *6*, 58564–58573. [CrossRef]
76. Li, S.; Yılmaz, Y.; Wang, X. Quickest Detection of False Data Injection Attack in Wide-Area Smart Grids. *IEEE Trans. Smart Grid* **2015**, *6*, 2725–2735. [CrossRef]
77. Kekatos, V.; Giannakis, G.B. Distributed Robust Power System State Estimation. *IEEE Trans. Power Syst.* **2013**, *28*, 1617–1626. [CrossRef]
78. Kurt, M.N.; Yılmaz, Y.; Wang, X. Distributed Quickest Detection of Cyber-Attacks in Smart Grid. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 2015–2030. [CrossRef]

79. Kurt, M.N.; Yılmaz, Y.; Wang, X. Real-Time Detection of Hybrid and Stealthy Cyber-Attacks in Smart Grid. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 498–513. [CrossRef]
80. Zhang, J.; Wang, X. Low-Complexity quickest-change detection in Linear Systems with Unknown Time-Varying Pre- and Post-Change Distributions. *IEEE Trans. Inf. Theory* **2021**, *67*, 1804–1824. [CrossRef]
81. Nath, S.; Akingeneye, I.; Wu, J.; Han, Z. Quickest Detection of False Data Injection Attacks in Smart Grid with Dynamic Models. *IEEE J. Emerg. Sel. Top. Power Electron.* **2022**, *10*, 1292–1302. [CrossRef]
82. Kurt, M.N. Data-Driven Quickest-Change Detection. Ph.D. Thesis, Columbia University, New York, NY, USA, 2020. [CrossRef]
83. Moustakides, G.V.; Polunchenko, A.S.; Tartakovsky, A.G. Numerical Comparison of CUSUM and Shiryaev–Roberts Procedures for Detecting Changes in Distributions. *Commun. Stat. Theory Methods* **2009**, *38*, 3225–3239. [CrossRef]
84. Pollak, M.; Tartakovsky, A.G. Exact optimality of the Shiryaev-Roberts procedure for detecting changes in distributions. In Proceedings of the 2008 International Symposium on Information Theory and Its Applications, Auckland, New Zealand, 7–10 December 2008; pp. 1–6. [CrossRef]
85. Polunchenko, A.S.; Raghavan, V. Comparative performance analysis of the Cumulative Sum chart and the Shiryaev-Roberts procedure for detecting changes in autocorrelated data. *Appl. Stoch. Model. Bus. Ind.* **2018**, *34*, 922–948. [CrossRef]
86. Ozay, M.; Esnaola, I.; Yarman Vural, F.T.; Kulkarni, S.R.; Poor, H.V. Machine Learning Methods for Attack Detection in the Smart Grid. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 1773–1786. [CrossRef]
87. Yan, J.; Tang, B.; He, H. Detection of false data attacks in smart grid with supervised learning. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 1395–1402. [CrossRef]
88. Wang, Y.; Amin, M.M.; Fu, J.; Moussa, H.B. A Novel Data Analytical Approach for False Data Injection Cyber-Physical Attack Mitigation in Smart Grids. *IEEE Access* **2017**, *5*, 26022–26033. [CrossRef]
89. Yu, Z.H.; Chin, W.L. Blind False Data Injection Attack Using PCA Approximation Method in Smart Grid. *IEEE Trans. Smart Grid* **2015**, *6*, 1219–1226. [CrossRef]
90. Hao, J.; Piechocki, R.J.; Kaleshi, D.; Chin, W.H.; Fan, Z. Sparse Malicious False Data Injection Attacks and Defense Mechanisms in Smart Grids. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1–12. [CrossRef]
91. Esmalifalak, M.; Liu, L.; Nguyen, N.; Zheng, R.; Han, Z. Detecting Stealthy False Data Injection Using Machine Learning in Smart Grid. *IEEE Syst. J.* **2017**, *11*, 1644–1652. [CrossRef]
92. Trevizan, R.D.; Ruben, C.; Nagaraj, K.; Ibukun, L.L.; Starke, A.C.; Bretas, A.S.; McNair, J.; Zare, A. Data-driven Physics-based Solution for False Data Injection Diagnosis in Smart Grids. In Proceedings of the 2019 IEEE Power & Energy Society General Meeting (PESGM), Atlanta, GA, USA, 4–8 August 2019; pp. 1–5. [CrossRef]
93. Ruben, C.; Dhulipala, S.; Nagaraj, K.; Zou, S.; Starke, A.; Bretas, A.; Zare, A.; McNair, J. Hybrid data-driven physics model-based framework for enhanced cyber-physical smart grid security. *IET Smart Grid* **2020**, *3*, 445–453. [CrossRef]
94. Nagaraj, K.; Zou, S.; Ruben, C.; Dhulipala, S.; Starke, A.; Bretas, A.; Zare, A.; McNair, J. Ensemble CorrDet with adaptive statistics for bad data detection. *IET Smart Grid* **2020**, *3*, 572–580. [CrossRef]
95. Vega-Martinez, V.; Cooper, A.; Vera, B.; Aljohani, N.; Bretas, A. Hybrid Data-Driven Physics-Based Model Framework Implementation: Towards a Secure Cyber-Physical Operation of the Smart Grid. In Proceedings of the 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Prague, Czech Republic, 28 June–1 July 2022; pp. 1–5. [CrossRef]
96. Kurt, M.N.; Ogundijo, O.; Li, C.; Wang, X. Online Cyber-Attack Detection in Smart Grid: A Reinforcement Learning Approach. *IEEE Trans. Smart Grid* **2019**, *10*, 5174–5185. [CrossRef]
97. Tsitsiklis, J.; van Roy, B. Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Trans. Autom. Control.* **1999**, *44*, 1840–1851. [CrossRef]
98. Chen, S.; Devraj, A.M.; Bušić, A.; Meyn, S. Zap Q-Learning for Optimal Stopping. In Proceedings of the 2020 American Control Conference (ACC), Denver, CO, USA, 1–3 July 2020; pp. 3920–3925. [CrossRef]
99. Meyn, S. *Control Systems and Reinforcement Learning*; Cambridge University Press: Cambridge, UK, 2022. [CrossRef]
100. Chen, P.Y.; Yang, S.; McCann, J.A.; Lin, J.; Yang, X. Detection of false data injection attacks in smart-grid systems. *IEEE Commun. Mag.* **2015**, *53*, 206–213. [CrossRef]
101. 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*, 2505–2516. [CrossRef]
102. Ayad, A.; Farag, H.E.Z.; Youssef, A.; El-Saadany, E.F. Detection of false data injection attacks in smart grids using Recurrent Neural Networks. In Proceedings of the 2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 19–22 February 2018; pp. 1–5. [CrossRef]
103. Ganjkhani, M.; Fallah, S.N.; Badakhshan, S.; Shamshirband, S.; Chau, K.w. A Novel Detection Algorithm to Identify False Data Injection Attacks on Power System State Estimation. *Energies* **2019**, *12*, 2209. [CrossRef]
104. Zhang, Y.; Wang, J.; Chen, B. Detecting False Data Injection Attacks in Smart Grids: A Semi-Supervised Deep Learning Approach. *IEEE Trans. Smart Grid* **2021**, *12*, 623–634. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.