A Hybrid State Estimator Based on SCADA and PMU Measurements for Medium Voltage Distribution System

Xiangyu Kong 1,* , Ying Chen 1, Tao Xu 1, Chengshan Wang 1, Chengsi Yong 1, Peng Li 1 and Li Yu 2

1 Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China; chen_ying@tju.edu.cn (Y.C.); taoxu2011@tju.edu.cn (T.X.); cswang@tju.edu.cn (C.W.); chengsi2016@tju.edu.cn (C.Y.); lip@tju.edu.cn (P.L.)
2 Electric Power Research Institute of China Southern Power Grid, Guangzhou 510080, China; yuli@csg.cn
* Correspondence: eekongxy@tju.edu.cn

Received: 25 July 2018; Accepted: 17 August 2018; Published: 1 September 2018

Abstract: With the increasing importance of renewable energy and flexible loads, the operation of the distribution system is becoming more stochastic and complex, and it is necessary to monitor the power system in real-time. Considering the gradual applications of intelligent electronic devices in the distribution systems, a hybrid state estimator based on supervisory control and data acquisition (SCADA) and phasor measurement unit (PMU) measurements is proposed in this paper, which consists of the improved robust estimation and linear state estimation. At the time of SCADA data acquisition, the improved robust estimation combining the SCADA measurements with PMU measurements is performed. To eliminate the effect of bad data, the internal student residual method is introduced, and the robust thresholds are adjusted adaptively. Then the linear state estimation is performed at the time of PMU data acquisition based on the results of the previous estimation time and the PMU measurements, which can quickly correct the robust estimation results and track the changes of the distribution system. Finally, the effectiveness and performance of the proposed method are verified in a modified IEEE 33-bus distribution system and a real distribution system in China.

Keywords: distribution system state estimation; phasor measurement units (PMUs); improved robust estimation; fast state tracking

1. Introduction

As more and more distributed generation (DG) and electric vehicles are connected to the distribution networks, the operational characteristics of the distribution system have been changed [1]. The bidirectional power flow occurs when the generation from distributed units exceeds the local load, and the voltage profile of the distribution network will be influenced strongly. Furthermore, the network configuration of the smart distribution network will be changing dynamically to achieve minimum power loss and voltage deviations. The real-time monitoring of the distribution network is becoming more and more challenging due to the increasing dynamics and changing behavior of actors in distribution systems [2]. The operating condition needs to be perceived first and a key technology obtained to achieve the reliable and accurate system condition awareness is known as the distribution system state estimation (DSSE). The DSSE is used to calculate the real-time measurement data according to some optimal estimation criterion, thus improving the accuracy of the data and obtaining a credible state for the whole system. The traditional real-time measurement data in the
distribution system mainly comes from the supervisory control and data acquisition (SCADA) system, whose sampling period is 2–5 s. The three-phase voltage and power measurements are mainly configured at the high- and low-voltage sides of the distribution substation and some critical buses. Only the current amplitude measurements are configured at the sub section post (SSP), ring main unit (RMU), cable branch boxes, and column switches with no voltage and power measurements [3].

Different from the SCADA system, the sampling period of the phasor measurement units (PMUs) is 30–40 ms so that PMUs can capture the dynamic changes of the distribution system. PMUs can not only measure voltage phasors and current phasors accurately, but also provide synchronized real-time measurements (with reference to the global position system (GPS) time). In the transmission systems, the PMUs and wide area monitoring systems are applied to constantly monitor grid stability, detecting power faults, and supporting power system restoration. While in the distribution systems, with the development of an active distribution network and the effect of a bidirectional power flow, measurement instruments, such as PMUs, have become increasingly useful. The unbalanced and uncertain characteristics of the distribution network and the small angle differences between the bus voltage and line current phasors require higher phase measurement accuracy. As a result, PMUs are indispensable in distribution systems and one of the most known examples used in distribution systems is the micro-PMU [4]. Originally, the PMUs in the distribution system are mainly installed for the advanced applications, such as fault diagnosis and islanding detection, instead of the DSSE [5,6]. Considering the economics and infrastructure of the distribution system, it is not necessary to configure the PMU at each bus. At present, the research on the configuration scheme of PMUs in the distribution system is mainly focused on the complete observability of the network, improvement of the state estimation accuracy and the observability of the system power flow [2,4,7,8].

In the SCADA system, overhead lines, coaxial cables, and power line carriers are fundamental equipment to transmit measurement data, while in the PMUs, the wireless and wireline communication infrastructures, such as the Wi-Fi, Bluetooth, ZigBee, and fiber optics, are used to carry measurement information [9–12]. In Reference [13], a Cloud-based architectural solution is proposed to decouple the handling of PMU measurement data from the underlying physical devices. In DSSE, the measurement data is treated as uncertainty data, and the uncertainty of PMU measurement data is smaller than that of SCADA measurement data. The state estimation is carried out by using probability statistics or fuzzy analysis [14,15]. Furthermore, the interval linear state estimation model is proposed in References [16,17]. A Monte Carlo method is used in weighted least squares (WLS) to deal with the uncertainty of the measurement data [18–26].

Due to the increasing dynamics and changing behavior of actors in distribution systems, it is necessary to study how to effectively utilize the PMU measurement data with millisecond update rates in the DSSE. Some references utilize the high accuracy of the PMU measurements and the increased measurement redundancy to improve the performance of DSSE. In Reference [8], the numerical results show that a high level of accuracy in the state estimation task can be achieved with a much fewer number of PMUs than voltage magnitude meters. In References [18,19], the root node voltage is extended into state variables, which can effectively deal with the voltage phasor and voltage amplitude measurements to improve the estimation accuracy. In Reference [20], an estimator is added for processing the PMU measurements, treating the SCADA measurements as prior information to obtain the final estimation results. However, the state estimation period of the above methods is still second-level, which is the same as the sampling period of the SCADA, and a great quantity of PMU measurement data are ignored.

For the integration of measurements with different sampling periods in the distribution system, the researchers have proposed some different solutions. Reference [21] combines SCADA measurements with advanced metering infrastructure (AMI) measurements, whose sampling period ranges from 15 min to 24 h and converts DSSE into the power flow matching. Reference [22] uses SCADA, PMU, AMI, and load forecast (LF) data to perform DSSE. The AMI information is combined with SCADA, LF, and PMU to perform the static state estimation every 1 h. The unscented Kalman
filter (UKF) is used to dynamically estimate the system states between polling periods, and to predict the estimated values every 15 min using SCADA, LF, and PMU measurements. Reference [23] proposes a hybrid method combined with nonlinear static state estimation, static state estimation, and linear dynamic state estimation. Nonlinear static state estimation is carried out every 1 min with AMI, remote terminal unit (RTU), and PMU measurements, while static state estimation is carried out every 10 s with RTU and PMU measurements. Static state estimation and linear dynamic state estimation is carried out every 10 s. There are some studies focusing on the data fusion of SCADA and PMU measurements in the transmission network such as References [24–26]. However, in the distribution system, there are not much research done on this problem.

In the distribution system, bad data may be caused by communication, transducer failures, or third parties tampering. When it comes to the problem of bad data elimination during iterations, a robust estimation model based on an exponential objective function is adopted in Reference [27]. A weighted least absolute value (WLAV) estimator, Schweppe-type Huber generalized M (SHGM) estimator, and improved WLS are also introduced to the DSSE. Reference [28] compares the performance of the conventional WLS, WLAV, and SHGM in DSSE. The distribution system solves the observability problem by adding the pseudo-measurements with large errors, but the WLAV and SHGM estimators treat the pseudo-measurements as bad data and there are not enough redundant measurements to eliminate these bad data. Therefore, the estimation performance of the WLAV and SHGM estimators in the distribution system is not ideal, while the WLS has a better fitness. However, the WLS estimator is only optimal for the measurement data obeying a Gaussian distribution, while the bad data does not obey a Gaussian distribution. Considering the above shortcoming, References [29,30] adopt the equivalent weight theory to improve the robust performance of the conventional WLS. In Reference [31], WLAV is converted into the WLS, and the denominator of the weight function is modified by an exponential function with the square of the normalized residuals to improve the robustness of the estimator. Furthermore, the residuals are normalized by the internal student residual method.

In the current distribution system, SCADA-based state estimators are not accurate enough to capture the system states because of the second-level refresh rates. Although PMU measurements provide millisecond update rates, the system observability cannot be guaranteed. Meanwhile, there may be bad data in SCADA and PMU measurements. To ensure the accuracy of DSSE and effectively track system state changes, a hybrid state estimator for a medium voltage distribution network is proposed based on SCADA and PMU measurements in this paper. At the time of SCADA data acquisition, the improved robust estimation combining the SCADA and PMU measurements is performed. During each iteration, the internal student residual method is used to unite the accuracy of the residuals at the same scale and the robust thresholds are adjusted by the normalized residuals adaptively to further eliminate the effect of bad data. Then according to the fast refresh frequency characteristics of the PMUs, the linear state estimation (LSE) is performed at the time of PMU data acquisition based on the results of the previous estimation time and the PMU measurements, which can quickly capture the changes of operation states in the distribution system.

The main contributions of this paper include:

(1) To solve the problems of measurements with different refresh frequencies, a hybrid state estimator based on the improved state estimation and LSE is proposed, which makes full use of the PMU measurements to fast track distribution system states and guarantee the timeliness of real-time measurements.

(2) To eliminate the influence of the bad data, the normalized residuals and adaptive thresholds are introduced to improve the conventional robust estimation method. Furthermore, the performance of the improved robust estimation method is compared with other robust estimation methods.

The following parts of this paper are organized as follows: Section 2 introduces the multi-sampling period data fusion in the distribution system. Section 3 introduces the proposed method and Section 4 shows the simulations, followed by conclusions in Section 5.
2. Multi-Sampling Period Data Fusion in the Distribution System

The traditional distribution system shown in Figure 1 uses the telemetry function of the SCADA to measure and transmit measurement data. The function mainly measures the three-phase current ($I_{r}^{m}$) and branch power ($P_{B_{ij}}^{m}$, $Q_{B_{ij}}^{m}$), the power injection ($P_{L_{ij}}^{m}$, $Q_{L_{ij}}^{m}$) and the three-phase voltage ($V_{r}^{m}$), while the PMUs mainly measure the voltage phasor ($V_{r}^{m}$) and some or all the connected branch current phasor ($I_{r}^{m}$). $\phi$ represents a phase in A, B or C.

$$
\begin{align*}
V_{r}^{m} & \rightarrow r \\
I_{r}^{m} & \rightarrow i
\end{align*}
$$

\[ (P_{B_{ij}}^{m} \cdot Q_{B_{ij}}^{m}) \]

\[ (P_{L_{ij}}^{m} \cdot Q_{L_{ij}}^{m}) \]

**Figure 1.** The radial distribution system.

As mentioned above, the SCADA data acquisition period is generally 2–5 s, and the PMU data acquisition period is generally 30–40 ms. If only using the SCADA measurement data, the state estimation interval time is at least 2 s, and the PMU measurement data during the SCADA data acquisition period will not participate in the state estimation, resulting in a waste of PMU data, which are more accurate and have a higher refresh rate. In this paper, the interval time of SCADA data acquisition is $T_{SCADA}$, while the interval time of PMU data acquisition is $T_{PMU}$. Besides, $T_{SCADA}$ is divided into $n$ portions, and $P$ is the PMU data acquisition time during the SCADA data acquisition period.

According to the data acquisition time, the state estimator switches between the robust estimator and the linear estimator to fast track states. As shown in Figure 2, at the SCADA data acquisition time, the SCADA and PMU measurements are combined in the robust estimation with pseudo-measurements, which can filter the bad data in measurements and prepare for the next state estimation. In the distribution system, the observability is guaranteed by pseudo-measurements which are mainly obtained from load forecasting based on AMI measurements [32,33]. Due to a large number of load buses in the distribution system, there is a certain error in the load forecasting, so that the obtained pseudo-measurements have a low precision. At a non-SCADA data acquisition time, namely the PMU data acquisition time $P$, the LSE is carried out based on estimation results of the previous estimation time and the PMU measurement data, which guarantees the system observability and makes the estimation results continue to update during the SCADA data acquisition period. Besides, the estimation accuracy of PMU installation buses can be close to the PMU measurement accuracy through the high-frequency refresh, including the bad data of the PMU measurements, so that combining the robust estimation with a linear estimation can ensure the estimation accuracy and track the distribution system status in real-time.

**Figure 2.** Measurement data fusion diagram of SCADA and PMU.
3. A Hybrid State Estimator Based on SCADA and PMU Measurements

3.1. State Variables and Measurement Equations in Robust State Estimation

At the time of SCADA data acquisition, the available measurements including the bus voltage amplitude, the bus voltage phasor, the current amplitude, the current phasor and power measurements, and the mathematical observability is guaranteed by pseudo-measurements. This paper follows the method of Reference [19] to deal with the voltage amplitude and voltage phasor measurements. Taking the branch currents as the state variables, the root bus voltage is expanded into the state variables, and the state variables can be expressed in the form of rectangular coordinates as Equation (1):

$$\begin{bmatrix} x_h \end{bmatrix} = \begin{bmatrix} v_r^{\text{slack}} & v_i^{\text{slack}} & i_r^l & \ldots & i_r^{N_p} & i_i^l & \ldots & i_i^{N_p} \end{bmatrix}$$

(1)

where, $v_r^{\text{slack}}$ and $v_i^{\text{slack}}$ are the real and imaginary parts of the root bus voltage on phase $\phi$, respectively, and $i_r^l \ldots i_r^{N_p}$ and $i_i^l \ldots i_i^{N_p}$ are the real and imaginary parts of branch currents on phase $\phi$ respectively.

Since the phase angle difference between the two ends of the lines in the distribution network is very small, the voltage phase angles of all buses are not much different from that of the root bus [34], so the measurement of the bus voltage amplitude on the phase $\phi$ can be represented by the state variable as Equation (2):

$$v_{h\phi} \simeq v_r^{\text{slack}} \cos(\delta_{h\phi}^{(k-1)}) + v_i^{\text{slack}} \sin(\delta_{h\phi}^{(k-1)}) - \sum_{l \in T_{h\phi}} \frac{3}{\sum_{\varphi=1}^{N_p}} \left( r_{\varphi \phi} i_r^{\varphi \phi} + x_{\varphi \phi} i_i^{\varphi \phi} \right) \cos(\delta_{h\phi}^{(k-1)})$$

(2)

$$- \sum_{l \in T_{h\phi}} \frac{3}{\sum_{\varphi=1}^{N_p}} \left( r_{\varphi \phi} i_r^{\varphi \phi} + x_{\varphi \phi} i_i^{\varphi \phi} \right) \sin(\delta_{h\phi}^{(k-1)})$$

where $T_{h\phi}$ is the path between the bus $h$ and the root bus. $r_{\varphi \phi}$ and $x_{\varphi \phi}$ are the resistance and reactance of branch $l$ on the path $T_{h\phi}$, respectively, and $\delta_{h\phi}^{(k-1)}$ is the voltage phase angle of the bus $h$ at the $(k-1)$th iteration.

The Jacobian matrix elements corresponding to the bus voltage amplitude are given by Equation (3):

$$\frac{\partial v_{h\phi}}{\partial v_r^{\text{slack}}} = \cos(\delta_{h\phi}^{(k-1)})$$

$$\frac{\partial v_{h\phi}}{\partial v_i^{\text{slack}}} = -r_{\varphi \phi} \cos(\delta_{h\phi}^{(k-1)}) - x_{\varphi \phi} \sin(\delta_{h\phi}^{(k-1)})$$

$$\frac{\partial v_{h\phi}}{\partial i_r^{\varphi \phi}} = x_{\varphi \phi} \cos(\delta_{h\phi}^{(k-1)}) - r_{\varphi \phi} \sin(\delta_{h\phi}^{(k-1)})$$

(3)

The bus voltage phasor measurements can be represented by state variables as Equation (4):

$$\hat{v}_{h\phi} = v_{h\phi} + jv_{i\phi} = v_r^{\text{slack}} - \sum_{l \in T_{h\phi}} \frac{3}{\sum_{\varphi=1}^{N_p}} \left( r_{\varphi \phi} i_r^{\varphi \phi} + x_{\varphi \phi} i_i^{\varphi \phi} \right)$$

(4)

and the Jacobian matrix elements corresponding to the bus voltage phasor are given by Equation (5):

$$\frac{\partial \hat{v}_{h\phi}}{\partial v_r^{\text{slack}}} = 1$$

$$\frac{\partial \hat{v}_{h\phi}}{\partial v_i^{\text{slack}}} = 0$$

$$\frac{\partial \hat{v}_{h\phi}}{\partial i_r^{\varphi \phi}} = x_{\varphi \phi}$$

$$\frac{\partial \hat{v}_{h\phi}}{\partial i_i^{\varphi \phi}} = -x_{\varphi \phi}$$

(5)

From Equations (2)–(5), it can be seen that the bus voltage measurements will affect the real and imaginary parts of the root bus. As a result, the root bus voltage will not be a fixed reference bus, but an estimated value that is closer to the actual situation based on the actual measurements.
When the measurements in polar coordinates obtained from the PMUs are converted into the rectangular coordinates, and the power measurements are converted into the equivalent current measurements, the measurement error transmission is involved. The covariance of equivalent measurements [35] can be obtained through measurement correlation, and the covariance matrix of the branch current measurements in rectangular coordinate can be expressed as Equation (6):

$$
\sum_{i=1}^{n} \begin{bmatrix}
\frac{\partial r_{hk}}{\partial x_{hk}} & \frac{\partial r_{hk}}{\partial \delta_{hk}} \\
\frac{\partial \delta_{hk}}{\partial x_{hk}} & \frac{\partial \delta_{hk}}{\partial \delta_{hk}}
\end{bmatrix} \times \begin{bmatrix}
\sigma_{x_{hk}}^2 & \sigma_{\delta_{hk}}^2 \\
\sigma_{\delta_{hk}}^2 & \sigma_{\delta_{hk}}^2
\end{bmatrix} \times \begin{bmatrix}
\frac{\partial r_{hk}}{\partial x_{hk}} & \frac{\partial r_{hk}}{\partial \delta_{hk}} \\
\frac{\partial \delta_{hk}}{\partial x_{hk}} & \frac{\partial \delta_{hk}}{\partial \delta_{hk}}
\end{bmatrix}^T
$$

(6)

where $I_{hk}$ and $\delta_{hk}$ are the real and imaginary parts of the branch $h - k$ after the conversion, respectively. $l_{hk}$ and $\delta_{hk}$ are the current amplitude and phase angle of branch $h - k$ obtained by PMUs, respectively. $\sigma_{x_{hk}}^2$ and $\sigma_{\delta_{hk}}^2$ are the current covariance of the amplitude and phase angle, respectively. Both the conversion covariance of the PMU voltage measurements and the power measurements can be obtained by Equation (6), and the weights of the measurements can be expressed as the reciprocal of the covariance.

The correction of state variables can be obtained using Equation (7):

$$
H^T WH \Delta x^{(k)} = H^T W \left( z - h (x^{(k-1)}) \right)
$$

(7)

where $H = \partial h / \partial x$ is the Jacobian matrix, $W$ is the measurement weight matrix considering the delay of SCADA data transmission [36], $x^{(k-1)}$ are the state variables obtained by the $(k-1)$th iteration, $h(x)$ is the measurement function, and $z$ is the real-time measurement data of SCADA and PMU or pseudo-measurements.

3.2. The Improved Robust State Estimation with Normalized Residuals and Adaptive Thresholds

In the traditional WLS, the measurement weight matrix contains the covariance of each measurement error. During the measuring process, due to the reliabilities of sensors or communication, the bad data with errors greater than $\pm 3\sigma$ may appear. The traditional WLS does not distinguish normal data from bad data, which will cause a large estimation error, thereby affecting the subsequent decision-making credibility of the distribution system. However, considering the good convergence performance of WLS, this paper uses an improved robust state estimator based on WLS at the time of SCADA data acquisition. The concept of equivalent weight is used to integrate the residuals into the preservation zones, decentralized zones, and elimination zones. Compared with dividing the measurement data into normal, suspicious, and harmful measurements, dividing the measurement data into three segments can make full use of measurement information. The weights are expressed as Equation (8):

$$
\pi_i = \begin{cases}
w_i, & |r_i| \leq k_0 \\
w_i \frac{k_1 - |r_i|}{k_1}, & k_0 < |r_i| \leq k_1 \\
0, & |r_i| > k_1
\end{cases}
$$

(8)

where $w_i$ is the $i$th diagonal elements of the measurement weight matrix $W$ in the traditional WLS, $k_0$ and $k_1$ are the robust thresholds, and $d_i$ is the smooth factor, where $d_i = (k_1 - |r_i|) / (k_1 - k_0)$. When the value of $r_i$ is smaller than $k_0$, its weight maintains the original measurement weight $w_i$. When the value of $r_i$ is between $k_0$ and $k_1$, its weight decreases as $r_i$ increases. When the value of $r_i$ is larger than $k_1$, its weight is 0.
The validity of the estimation in the robust estimation depends largely on the rationality of the equivalent weight. The efficiency of robust estimation is different from statistics and robust thresholds, especially the rationality of decentralized zone weight threshold and elimination zone weight threshold. For all measurement data, if the thresholds of the robust scheme are the constant experience values, the measurements that originally belong to the normal segment may be defined as the harmful segment or suspicious segment, which will lose a certain amount of efficiency in the state estimation. For this problem, this paper refers to Reference [29] and gives a robust estimation method with variable thresholds. The key to the variable thresholds is that during the iteration process, the thresholds gradually change with the residuals, which can overcome the shortage of the traditional thresholds determined by the experience. The robust thresholds are designed in accordance with Equation (9):\[
\begin{align*}
  k_0 &= \min(k_{11}, k_{12}) \\
  k_1 &= \max(k_{11}, k_{12})
\end{align*}
\]
where \(k_{11} = \alpha r_i\) and \(k_{12} = k_{11} + (|r_i| - k_{11})/3\). To make the robust thresholds adapt to the normal condition, \(\alpha\) is set to 1.438 [29, 38].

The residuals can be expressed as Equation (10):
\[
\begin{align*}
  r &= z - h(x) - (H\Delta x) \\
  &= \Delta z - H(H^T\bar{W}H)^{-1}H^T\bar{W}\Delta z \\
  &= \left[I - H(H^T\bar{W}H)^{-1}H^T\bar{W}\right]\Delta z \\
  &= (I - J)\Delta z
\end{align*}
\]
where \(\bar{W}\) represents the modified measurement weight value and \(J = H(H^T\bar{W}H)^{-1}H^T\bar{W}\).

In the case that real-time measurements are not enough, pseudo-measurements are usually added to ensure the observability. The accuracy of the pseudo-measurements is different from that of the real-time measurements; besides, the measurement accuracy of SCADA and PMU are also different, thus the variances of residuals are not the same. In this paper, we use the internal student residual method to standardize the residuals as given in Equation (11):
\[
\begin{align*}
  \begin{cases}
    r_s = \frac{l\sqrt{1 - J_{ii}} + \varepsilon \sqrt{I - 1}}{\sqrt{R}} \\
    r_{si} = \frac{r_i \sqrt{1 - J_{ii}} + \varepsilon \sqrt{I - 1}}{\sqrt{R}}
  \end{cases}
\end{align*}
\]
where \(l\) represents the scale parameter of the standardization and \(l^2 = r^T R^{-1} r / (m - n)\), \(m\) represents the number of measurements, \(n\) represents the number of state variables, and \(\varepsilon\) is a very small number to avoid \(1 - J_{ii} = 0\).

3.3. Fast-State Tracking with PMUs Based on the Linear State Estimation

PMU measurements with millisecond refresh frequency are used to fast track the system states during the SCADA data acquisition period. However, the limited number of PMUs cannot guarantee the observability, so other measurement information at the PMU data acquisition time should be included. In this paper, the estimation results from the previous estimation time (the branch currents) are combined with the PMU measurements (the bus voltage phasors and the branch current phasors) to guarantee the observability. The corresponding measurement weights of the PMU measurements and estimation results of the previous estimation time can be expressed in a matrix block as Equation (12):
\[
W_{\text{linear}} = \sum_{\text{linear}}^{-1} = \begin{bmatrix} 
\sum_{\text{SE}} & \sum_{\text{PMU}} \\
\sum_{\text{PMU}} & \sum_{\text{PMU}}
\end{bmatrix}^{-1}
\]
where the covariance matrix of the voltage phasors $\Sigma_{PMU}$ and the current phasors $\Sigma_{PMU}$ in rectangular coordinate are obtained using Equation (6).

The estimation results from the previous estimation time are used as the prior information whose covariance matrix $\Sigma_{SE}$ [39] can be obtained using Equation (13):

$$\Sigma_{SE} = \left( H^T W H \right)^{-1}$$

The state variables are consistent with that of robust state estimation. Equation (5) can be used to obtain the measurement Jacobian matrix whose elements are constant. Therefore, the LSE can be utilized, and the characteristic of a short execution time can be used to track states with PMU measurements. The LSE results can be obtained using Equation (14):

$$x = \left( H^T W_{linear} H \right)^{-1} H^T W_{linear} z$$

The whole flowchart of a hybrid estimator based on SCADA and PMU hybrid measurements in the distribution system is shown in Figure 3. If the estimation time $T_n$ is the SCADA sampling time, the robust estimation combining the SCADA measurements with PMU measurements is carried out. Also, the specific flowchart of the improved robust estimation is shown in the left part of Figure 3. Otherwise, the LSE combining the estimation results of the previous estimation time with PMU measurements is carried out.

![Flowchart](image)

**Figure 3.** The flowchart of a hybrid estimator based on SCADA and PMU measurements.

4. Case Studies

The modified IEEE 33-bus distribution system and measurement configuration are shown in Figure 4. The network parameters and load values provided by Reference [40] were used as the reference condition. PMUs were installed at buses 12 and 30, assuming that all buses had pseudo-power injection measurements and the network reconfiguration was not considered in the state estimation. Note that the measurement configuration was based on the technical reasoning and without considering any optimal placement technique.
In order to prove the validity of the proposed method, the power flow results were taken as the true values. Also, the measurement data are formed by superimposing the corresponding random measurement errors obeying the normal distribution on the basis of the power flow results as given by Equation (15):

$$X^m = X(1 + \sigma)$$

where, $X^m$ are the measurement values, $X$ are the true values and $\sigma$ is the random measurement error. The maximum error of the current amplitude measurement and power measurement in SCADA were 3%, while the maximum error of voltage amplitude measurement was 1%. The maximum error of amplitude measurement in PMU was 1%, and the maximum error of the phase angle measurement was $10^{-2}$ rad. The maximum error of the pseudo-measurements was 50%. A Monte Carlo method was used to deal with the uncertainty and the number of Monte Carlo trails was 5000 [40].

The dynamic state estimation process of this paper is shown in Figure 5. The period for the improved robust estimation and LSE could be set according to the request of the system and the real refresh frequency of measurement data. In this paper, the period for the improved robust estimation was set to 5 s, while the period for LSE was set to 0.5 s. Taking the SCADA data acquisition time as the starting moment, the improved robust estimation was performed with the SCADA and PMU measurements and pseudo-measurements. When reaching the start time of fast-state tracking with the LSE, the results of the last estimation time were taken as pseudo-measurements, and the estimation results were updated by the PMU measurement data. During the SCADA acquisition data period, a total of nine linear state estimations were performed. When reaching the SCADA data acquisition time again, the above process was repeated. The calculation time of the two state estimators is also shown in Figure 5.
The performance of the proposed hybrid state estimator is verified under the normal condition and bad data condition. At the SCADA data acquisition time, the exponential equivalence robust estimation (EXP) mentioned in Reference [31] and initial IGGIII with no improvement (IGGIII) are used to be compared with the improved IGGIII (IIGGIII). During the dynamic state estimation process, the method in Reference [22] (WLS-UKF) was used to be compared with the proposed hybrid state estimator (IIGGIII-LSE).

4.1. Comparison of State Estimation Results under the Normal Condition

Table 1 shows a comparison of convergence, iteration, execution time, and the mean absolute error (MAE) of voltage amplitude between the different methods at the SCADA data acquisition time. Also, the MAE of the bus voltage amplitude is defined using Equation (16).

$$r_{MAE} = \frac{\sum |\hat{x} - x_{pf}|}{x_{pf} n}$$  \hspace{1cm} (16)

where $\hat{x}$ is the state estimation results and $x_{pf}$ is the power flow results.

Table 1. Performance comparison of five estimation methods.

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Convergence</th>
<th>Iterations</th>
<th>Execution Time(s)</th>
<th>MAE of Bus Voltage Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLS</td>
<td>Yes</td>
<td>22</td>
<td>0.144</td>
<td>$3.17033 \times 10^{-5}$</td>
</tr>
<tr>
<td>EXP</td>
<td>Yes</td>
<td>37</td>
<td>0.496</td>
<td>$6.57133 \times 10^{-5}$</td>
</tr>
<tr>
<td>IIGGIII</td>
<td>Yes</td>
<td>32</td>
<td>0.357</td>
<td>$1.72627 \times 10^{-5}$</td>
</tr>
<tr>
<td>IGGIII</td>
<td>Yes</td>
<td>26</td>
<td>0.182</td>
<td>$9.44840 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

Under the normal measurement condition, the four estimation methods could effectively converge in less than 0.5 s, which would not collide with the set time in Figure 5, and the MAE of each method were all less than $10^{-4}$, which shows the high estimation accuracy. The improved IGGIII has certain advantages in iterations compared with the exponential equivalence robust estimation. Furthermore, it shows that the reasonable measurement weights could speed up the iteration process. However, the execution time of the improved IGGIII was longer than that of WLS and initial IGGIII. This was because the recalculated weights and robust thresholds in each iteration increased the execution time.

Table 2, Figures 6 and 7 show the dynamic state estimation results at bus 6 (without PMU installation) and 30 (with PMU installation). The proposed IIGGIII-LSE could track the changes of the states, whose effect was similar to the method WLS-UKF. While at the bus with the PMU installation, the proposed IIGGIII-LSE could track the changes more accurately compared to the bus without PMU installation, which indicates the effectiveness of using PMUs.

Table 2. State trace at bus 6 and bus 30 of the 33-bus system under the normal condition.

<table>
<thead>
<tr>
<th>Time/s</th>
<th>MAE of Bus Voltage Amplitude</th>
<th>The Relative Error of the Estimated Voltage Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WLS-UKF</td>
<td>IIGGIII-LSE</td>
</tr>
<tr>
<td>20</td>
<td>0.00106</td>
<td>0.00134</td>
</tr>
<tr>
<td>20.5</td>
<td>0.00142</td>
<td>0.00124</td>
</tr>
<tr>
<td>21</td>
<td>0.00133</td>
<td>0.00083</td>
</tr>
<tr>
<td>21.5</td>
<td>0.00127</td>
<td>0.00124</td>
</tr>
<tr>
<td>22</td>
<td>0.03083</td>
<td>0.05652</td>
</tr>
<tr>
<td>22.5</td>
<td>0.03860</td>
<td>0.06421</td>
</tr>
<tr>
<td>23</td>
<td>0.06008</td>
<td>0.03733</td>
</tr>
<tr>
<td>23.5</td>
<td>0.03075</td>
<td>0.03675</td>
</tr>
<tr>
<td>24</td>
<td>0.00116</td>
<td>0.00124</td>
</tr>
<tr>
<td>24.5</td>
<td>0.00132</td>
<td>0.00114</td>
</tr>
</tbody>
</table>
Appl. Sci. 2018, 8, 1527

4.2. Comparison of State Estimation Results under the Bad Data Condition

To verify the robustness of the improved IGGIII at the SCADA data acquisition time, the bad data were assumed to appear in the voltage amplitude at bus 6 and the active power of branches 5–6 (measurement error was 75%).

Table 3 shows comparisons of various estimation methods under the conditions of different measurement errors. The errors of bad data always exceed $6\sim7\sigma$ in the actual situation, and all the four estimation methods could all converge even when the error of bad data reaches 95%. Besides, the execution time was less than 0.5 s, which could ensure that the next state estimation will not be affected and guarantee the timeliness of real-time measurements.

The improved IGGIII was superior to the other three estimation methods in terms of iteration, and MAE of voltage amplitude. And the execution time of the improved IGGIII was shorter than that of WLS and EXP. That was because the robust thresholds can effectively distinguish the bad data from the normal data and the recalculated weights could reduce the effect of bad data on the estimation results. However, the execution time of the improved IGGIII was longer than that of initial IGGIII. That was because the recalculated weights and robust thresholds in each iteration increased the execution time.

Figures 8 and 9 show the comparisons between the two robust estimation methods with a 50% error in voltage amplitude. It can be seen that the voltage amplitude and phasor angle estimated by the EXP deviated from the power flow values in the vicinity of bad data, while the improved IGGIII could make the estimation results closer to the power flow values. Furthermore, it indicates that the normalized residuals and adaptive robust thresholds could eliminate the effect of bad data effectively.
Table 3. Performance comparisons of various estimation methods.

<table>
<thead>
<tr>
<th>Measurement Error of Voltage Amplitude at Bus 6</th>
<th>Estimation Method</th>
<th>Convergence</th>
<th>Iterations</th>
<th>Execution Time</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal condition</td>
<td>WLS</td>
<td>Yes</td>
<td>12</td>
<td>0.236</td>
<td>0.00500</td>
</tr>
<tr>
<td></td>
<td>EXP</td>
<td>Yes</td>
<td>6</td>
<td>0.165</td>
<td>0.00004</td>
</tr>
<tr>
<td></td>
<td>IIGGIII</td>
<td>Yes</td>
<td>9</td>
<td>0.174</td>
<td>0.00004</td>
</tr>
<tr>
<td></td>
<td>IGGIII</td>
<td>Yes</td>
<td>13</td>
<td>0.165</td>
<td>0.00510</td>
</tr>
<tr>
<td>5%</td>
<td>WLS</td>
<td>Yes</td>
<td>12</td>
<td>0.197</td>
<td>0.00452</td>
</tr>
<tr>
<td></td>
<td>EXP</td>
<td>Yes</td>
<td>7</td>
<td>0.167</td>
<td>0.00008</td>
</tr>
<tr>
<td></td>
<td>IIGGIII</td>
<td>Yes</td>
<td>8</td>
<td>0.162</td>
<td>0.00002</td>
</tr>
<tr>
<td></td>
<td>IGGIII</td>
<td>Yes</td>
<td>13</td>
<td>0.175</td>
<td>0.00449</td>
</tr>
<tr>
<td>10%</td>
<td>WLS</td>
<td>Yes</td>
<td>12</td>
<td>0.190</td>
<td>0.01125</td>
</tr>
<tr>
<td></td>
<td>EXP</td>
<td>Yes</td>
<td>7</td>
<td>0.199</td>
<td>0.00062</td>
</tr>
<tr>
<td></td>
<td>IIGGIII</td>
<td>Yes</td>
<td>8</td>
<td>0.193</td>
<td>0.00004</td>
</tr>
<tr>
<td></td>
<td>IGGIII</td>
<td>Yes</td>
<td>13</td>
<td>0.173</td>
<td>0.01123</td>
</tr>
<tr>
<td>50%</td>
<td>WLS</td>
<td>Yes</td>
<td>12</td>
<td>0.190</td>
<td>0.06593</td>
</tr>
<tr>
<td></td>
<td>EXP</td>
<td>Yes</td>
<td>11</td>
<td>0.336</td>
<td>0.00300</td>
</tr>
<tr>
<td></td>
<td>IIGGIII</td>
<td>Yes</td>
<td>8</td>
<td>0.174</td>
<td>0.00044</td>
</tr>
<tr>
<td></td>
<td>IGGIII</td>
<td>Yes</td>
<td>12</td>
<td>0.172</td>
<td>0.06580</td>
</tr>
<tr>
<td>95%</td>
<td>WLS</td>
<td>Yes</td>
<td>12</td>
<td>0.194</td>
<td>0.12745</td>
</tr>
<tr>
<td></td>
<td>EXP</td>
<td>Yes</td>
<td>12</td>
<td>0.201</td>
<td>0.00353</td>
</tr>
<tr>
<td></td>
<td>IIGGIII</td>
<td>Yes</td>
<td>8</td>
<td>0.189</td>
<td>0.00127</td>
</tr>
<tr>
<td></td>
<td>IGGIII</td>
<td>Yes</td>
<td>11</td>
<td>0.163</td>
<td>0.12723</td>
</tr>
</tbody>
</table>

Figure 8. Bus voltage (amplitude and phase angle) profile curves estimated by different methods: (a) voltage amplitude curve, and (b) voltage angle curve.

Figure 9. Bus voltage (amplitude and phase angle) estimation errors under different methods: (a) the estimation error curve of voltage amplitude, and (b) the estimation error curve of voltage angle.
Figure 9 further illustrates the estimation errors of the two robust methods. To characterize the accuracy of state estimation, the absolute error for the estimated voltage amplitude of each bus is defined using Equation (17), and the relative error of the estimated voltage phase of each bus is defined using Equation (18). Because the values of the phase angle were close to zero, to avoid large fluctuation in the absolute error, the relative error was used to reflect the estimation error of the phase angles.

\[
\begin{align*}
    r_{\text{absolute}} &= \frac{\left| \hat{x} - x_{pf} \right|}{x_{pf}} \\
    r_{\text{relative}} &= \frac{\left| \hat{x} - x_{pf} \right|}{\left| x_{pf} \right|}
\end{align*}
\]  

Equation (17)

Equation (18)

From Figure 9a, the estimation errors of voltage amplitude obtained by the EXP were larger than that obtained by the improved IGGIII, especially in the vicinity of bad data. From Figure 9b, the estimation errors of phase angle in the vicinity of bad data obtained by the EXP were more than 1°, while the improved IGGIII could maintain the estimation error to lower than 0.1°, which highlights the better robustness of the improved IGGIII. The contrast results in Figures 8 and 9 also show that reasonable weights of measurements were very important during the iterations, especially under the bad data condition. Furthermore, the good estimation results obtained by the improved IGGIII indicated that the normalized residuals and adaptive robust thresholds could effectively improve the estimation accuracy under the bad condition. To sum up, the improved IGGIII could converge under the bad data condition with good robustness.

When the bad data appeared, the proposed IIGGIII-LSE could also track the changes in the distribution system, and the dynamic estimation results are shown in Table 4, Figures 10 and 11. To emphasize the advantage of the proposed IIGGIII-LSE, the EXP combined with LSE (EXP-LSE) was used. For EXP-LSE, the estimation results of bus voltage and phase angle at bus 6 without PMU, were unsatisfying, as shown in Figure 10, because bus 6 was in the vicinity of bad data and there were no PMU measurements to update the inaccuracy states. The phase angle estimation results at bus 30 equipped with PMU were inaccurate due to the bad robustness of EXP at the SCADA data acquisition time. However, the estimation results could be corrected by the highly accurate PMU measurements, as shown in Figure 11b.

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>MAE of Bus Voltage Amplitude</th>
<th>The Relative Error of the Estimated Voltage Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EXP-LSE</td>
<td>IIGGIII-LSE</td>
</tr>
<tr>
<td>20</td>
<td>0.00120</td>
<td>0.00022</td>
</tr>
<tr>
<td>20.5</td>
<td>0.00317</td>
<td>0.00284</td>
</tr>
<tr>
<td>21</td>
<td>0.00258</td>
<td>0.00114</td>
</tr>
<tr>
<td>21.5</td>
<td>0.00277</td>
<td>0.00101</td>
</tr>
<tr>
<td>22</td>
<td>0.00218</td>
<td>0.00177</td>
</tr>
<tr>
<td>22.5</td>
<td>0.00102</td>
<td>0.00001</td>
</tr>
<tr>
<td>23</td>
<td>0.00339</td>
<td>0.00258</td>
</tr>
<tr>
<td>23.5</td>
<td>0.00204</td>
<td>0.00080</td>
</tr>
<tr>
<td>24</td>
<td>0.00145</td>
<td>0.00007</td>
</tr>
<tr>
<td>24.5</td>
<td>0.00134</td>
<td>0.00129</td>
</tr>
</tbody>
</table>

Table 4. State trace at bus 6 and bus 30 of the 33-bus system under the bad data condition.
For the proposed IIGGIII-LSE, the dynamic estimation results under the bad data condition were as accurate as under the normal condition. While at the bus with the PMU installation, the proposed method could track the changes more accurately compared to the bus without the PMU installation, which indicated the effectiveness of using PMU measurements. To sum up, the robust state estimation was indispensable in the dynamic state estimation and the proposed IIGGIII-LSE performed well under the bad data condition. Besides, the estimation results would be less accurate if the estimated bus or area was far from the installation of the PMUs.

The feasibility of the proposed IIGGIII-LSE was also verified by the actual medium voltage distribution system in the suburb of Guangzhou, China. The system included 62 buses and 108 real-time measurements, and the diagram is shown in Figure A1. Figure 12 shows the application results of the proposed method in this system. The mean absolute error of the bus voltage amplitude was less than 0.6%, while the minimum mean absolute error of the bus voltage amplitude was less than 0.1%. The maximum relative error of the voltage phase was less than 0.4°. At the time of 3 min, 6 min, 7.5 min, and 9 min, the bad data were created by adding a disturbance into the measurements, and the proposed IIGGIII-LSE performed well, which indicated the practicability and effectiveness of the method in this paper.
The feasibility of the proposed IIGGIII-LSE was also verified by the actual medium voltage distribution system in the suburb of Guangzhou, China. The system included 62 buses and 108 real-time measurements, and the diagram is shown in Figure A1.

Figure 12 shows the application results of the proposed method in this system. The mean absolute error of the bus voltage amplitude was less than 0.6%, while the minimum mean absolute error of the bus voltage amplitude was less than 0.1%. The maximum relative error of the voltage phase was less than 0.4°. At the time of 3 min, 6 min, 7.5 min, and 9 min, the bad data were created by adding a disturbance into the measurements, and the proposed IIGGIII-LSE performed well, which indicated the practicability and effectiveness of the method in this paper.

(a) 
(b) 

Figure 12. Performance of the proposed method: (a) the estimation error curve of voltage amplitude, and (b) the estimation error curve of voltage angle.

5. Conclusions

In order to improve the accuracy of state estimation and effectively track system states, this paper proposed a hybrid state estimator considering the integration of a limited number of PMU measurements and bad data in the DSSE. The following conclusions were obtained:

(1) Using the PMU measurements in the DSSE was necessary and fast-state tracking with LSE could guarantee the timeliness of real-time measurements and avoid the waste of the PMU measurements. However, the estimation results would be less accurate if the estimated bus or area was far from the installation of the PMUs.

(2) The normalized residuals and adaptive thresholds could improve the estimation robustness of dealing with the bad data in real-time measurements.

(3) The proposed hybrid state estimator could effectively converge and quickly track the system states while ensuring the estimation accuracy.

Author Contributions: X.K. and C.W. conceived the idea of the study and conducted the research. Y.C. and C.Y. analyzed most of the data and wrote the initial draft of the paper. T.X., P.L., and L.Y. contributed to finalizing this paper.

Funding: This research was funded by the National Key Research and Development Program of China grant number 2017YFB0902900 and 2017YFB0902902.

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

Figure A1. The diagram of an actual medium voltage distribution system in China.

References


