Assessing Gaussian Assumption of PMU Measurement Error Using Field Data

Shaobu Wang, *Member, IEEE*, Junbo Zhao, Zhenyu Huang, *Fellow, IEEE*, and Ruisheng Diao, *Senior Member, IEEE*

***Abstract*—Gaussian phasor measurement unit (PMU) measurement error has been assumed for many power system applications, such as state estimation, oscillatory modes monitoring, voltage stability analysis, to cite a few. This letter proposes a simple yet effective approach to assess this assumption by using the stability property of a probability distribution and the concept of redundant measurement. Extensive results using field PMU data from WECC system reveal that the Gaussian assumption is questionable.**

***Index Terms*—Non-Gaussian distribution, PMU measurement, stable distribution, power system operation, state estimation.**

I. Introduction

With the wide-area deployment of phasor measurement units (PMUs), many power system online monitoring and control applications become possible, including dynamic state estimation, oscillatory modes monitoring, voltage stability analysis, to name a few [1]–[3]. However, Gaussian measurement error is usually assumed when conducting those applications. It should be noted that if Gaussian assumption is violated, the obtained results could be misleading and harmful for power system security.

In this letter, we propose a simple yet effective approach to assess this assumption by using the field PMU measurements from the Western Electricity Coordinating Council (WECC) system. We utilize the concept of redundant measurement to construct an error vector for Gaussian assumption assessment. The latter is checked by the stability property of a probability distribution and the Shapiro-Wilks test. Extensive results show that PMU measurement error potentially follows thick-tailed non-Gaussian probability distributions.

II. Problem Formulation

Definition 1: PMU measurement error includes the absolute error (or bias), the relative error and noise that is induced by instruments, communication channels, etc. Bias is a constant instead of a random variable, and therefore it does not affect the distribution of the measurement error. It should be noted that measurement noise sometimes has been used with measurement error interchangeably in the literature. However, they are in fact different and measurement error is more general. In addition, it is the measurement error that we care most when implementing measurement-based applications.

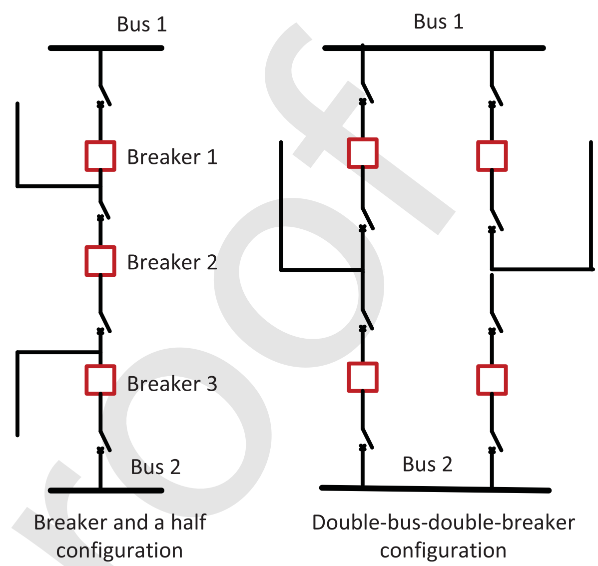

Fig. 1. Different configurations of circuit breakers.

To analyze error statistics, the first step is to extract them from measurements. However, as practical power systems are typically non-stationary, it is very challenging for traditional approaches, such as finite impulse response (FIR) digital low pass filters, median filter [4]–[6]. An alternative way to extract error from field PMU measurements is through a high precision calibrator, where same inputs are fed into the calibrator and the PMU devices under test, then the differences of their outputs are taken as the measurement error. Although this is the best way to obtain almost true error statistics, it is not economically feasible as the high precision calibrator is very expensive. In this letter, we propose a redundant measurement-based approach to check the statistical property of the measurement error.

III. Proposed Approach

A. Concept of Redundant Measurements

Redundant measurements are often set up to guarantee observability for different network topologies or operating modes. For example, Fig. 1 shows the breaker-and-half

Manuscript received June 23, 2017; revised August 25, 2017; accepted October 5, 2017. Date of publication; date of current version. This work was supported in part by the U.S. Department of Energy through its Advanced Grid Modeling Program and its ASCR Applied Mathmematics Program and in part by the U.S. National Science Foundation Award ECCS-1711191. Paper no. PESL-00117-2017. *(Corresponding author: Junbo Zhao.)*

S. Wang, Z. Huang, and R. Diao are with the Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: shaob.wang@pnnl.gov; zhenyu.huang@pnnl.gov; ruisheng.diao@pnnl.gov).

J. Zhao is with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Falls Church, VA 22043 USA (e-mail: zjunbo@vt.edu).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2017.2762927

bus/switching configuration and double-bus-double-breaker configuration cases in the WECC system. In order to ensure the observability of buses 1 and 2 for the breaker-and-half bus/switching configuration case when any one of the three breakers opens, two PMUs are installed, one at each bus. PMUs measure the bus voltage phasor as well as the current phasors of the lines that are adjacent to that bus. Note that we get the information of the breaker status from SCADA data. When all the breakers get closed, the metered voltage magnitudes and angles or line current phasors obtained by these two PMU devices should be identical in absence of errors. Then, we say that PMUs installed at buses 1 and 2 form a redundant case. Furthermore, if these two PMU devices are from the same vendor and operate with almost the same environmental factors, we would expect that the error statistics of them should be approximately the same.

B. *Assessment of Gaussian Assumption*

Let $\mathcal{S}$ be the set of redundant measurements, for $i, j \in \mathcal{S}$, define $\boldsymbol{x}_i = \boldsymbol{x} + \boldsymbol{e}_i$ and $\boldsymbol{x}_j = \boldsymbol{x} + \boldsymbol{e}_j$ as the measured values provided by PMUs, where $\boldsymbol{x}$ is the true value; $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are measurement errors associated with ith and jth PMUs. Note that $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ come from different PMU devices, they are considered as independent. Furthermore, to extract measurement error from the redundant set defined in Section III-A, we define $\Delta\boldsymbol{x}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j = \boldsymbol{e}_i - \boldsymbol{e}_j$. Subsequently, we have the following theorem:

Theorem 1: A necessary condition that both $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ follow a Gaussian distribution is that $\Delta\boldsymbol{x}_{ij}$ follows a Gaussian distribution.

Proof: We will prove this theorem by contradiction.

Case 1: assume that both $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ follow a Gaussian distribution, then by the stability property of the normal/Gaussian distribution, it is straightforward to verify that $\Delta\boldsymbol{x}_{ij}$ follows a Gaussian distribution.

Case 2: assume $\boldsymbol{e}_i$ (or $\boldsymbol{e}_j$) follows a Gaussian distribution while $\boldsymbol{e}_j$ (or $\boldsymbol{e}_i$) has a non-Gaussian distribution, $\Delta\boldsymbol{x}_{ij}$ must follow a non-Gaussian distribution. This is because if $\Delta\boldsymbol{x}_{ij}$ is Gaussian, by subtracting $\boldsymbol{e}_i$ (or $\boldsymbol{e}_j$) from $\Delta\boldsymbol{x}_{ij}$ and using the stability property of the Gaussian distribution, $\boldsymbol{e}_j$ (or $\boldsymbol{e}_i$) must follow a Gaussian distribution. This contradicts the assumption.

Case 3: assume both $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ follow a non-Gaussian distribution, $\Delta\boldsymbol{x}_{ij}$ follows a Gaussian or non-Gaussian distribution.

From cases 1–3, the Theorem follows. ■

Remark 1: According to the central limit theorem, it is true that the difference between two non-Gaussian random variables can yield a Gaussian distribution as the number of measurements tends to infinity. However, this can hardly hold true for practical power systems. This is because the measurement error of PMU depends on the operating conditions of the system and many environmental factors, yielding time varying error statistics. As a result, the independent and identically distributed random variable assumption of the central limit theorem does not hold true.

Remark 2: In our studies, two similar PMUs (from the same vendor) measure the same quantities and are subject to same impacts of environmental factors. We can anticipate that the error statistics of these two PMU devices are quite similar. Thus, according to the stability property of a probability distribution and Theorem 1, if we find that $\Delta\boldsymbol{x}_{ij}$ follows a Gaussian distribution, it is likely that both $\boldsymbol{e}_i$ (or $\boldsymbol{e}_j$) follows a Gaussian distribution; by contrast, if $\Delta\boldsymbol{x}_{ij}$ follows a non-Gaussian distribution, it is of very high probability that both $\boldsymbol{e}_i$ (or $\boldsymbol{e}_j$) follows a non-Gaussian distribution.

Following Theorem 1, validation of the normality of $\Delta\boldsymbol{x}_{ij}$ becomes the key step. To do that, several widely used statistical tests can be applied, including W/S test, Jarque-Bera test, Shapiro-Wilks test, Kolmogorov-Smirnov test [7]. Among them, Shapiro-Wilks test [8] is chosen in this letter because of its well demonstrated performance in many applications. It uses the null hypothesis principle to check whether a sample series is from a normally distributed population. Note that the null hypothesis corresponds to the Gaussian assumption. In the test, two parameters, namely p-value and α, need to be studied. Parameter α represents the significance level of the statistical test and is usually set according to the confidence level of significance. Note that in many existing statistical test approaches for validating normality/Gaussian assumption, anomaly or outlier detection, 95% is the value widely adopted for confidence level of significance. As a result, α is set to be $1 - 0.95 = 0.05$. In this case, if the Shapiro-Wilks test returns a p value and $p \geq \alpha$, we conclude that the random variable follows the Gaussian distribution. The Shapiro-Wilks test can be implemented in Matlab by calling the function SWTEST.

IV. Test Results With Field Data

To test the effectiveness of the proposed approach, field PMU measurements from the WECC system are used. Note that there are many redundant measurement cases in this system, from which nine sets of PMU redundant measurements of 18 buses are investigated. The scan rate of the PMUs is 30 samples/s.

A. *Choice of PMU Data for Testing*

For all hypothesis tests, two important factors need to be investigated: sample size, a number that is directly related to the confidence of the test and the expected false positive rate, and the test power, which represents the chance of false negatives. When the sample size is small and you insist on it with high confidence, the test power gets worse. By contrast, with larger sample size, the test power gets better. Note that many statistical tests will become inaccurate if the sample size is very large (larger than 6000 for Shapiro-Wilks test for example). This is because according to the central limit theorem, the sampling distribution tends to be normal if the number of samples is very large, regardless of the shape of the data. In this letter, we have tested the cases with varying number of measurement samples for bus set 1 and the results are shown in Fig. 2, where p-value is shown at the top of each figure. It is clear that the measurement errors of both voltage magnitudes and angles do not follow a Gaussian distribution as their p values are close to 0. In addition, it is found that the shapes of the histograms remain unchanged for a large number of samples (see the cases with

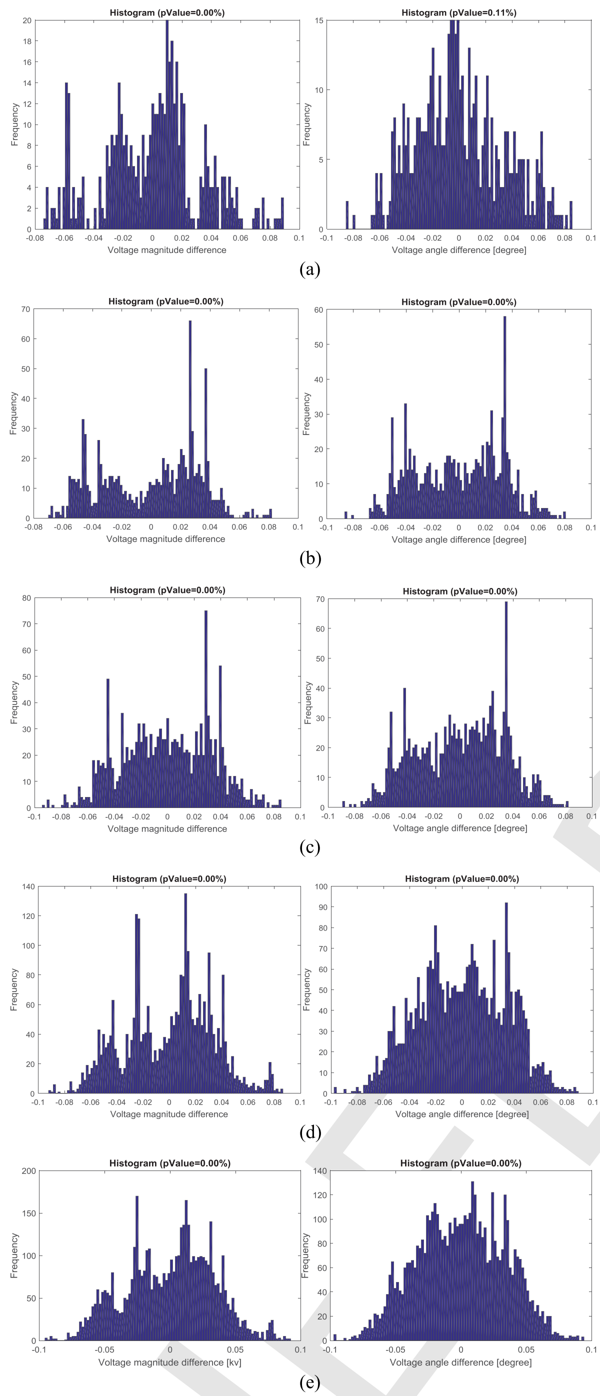

Fig. 2. Results of the Shapiro-Wilks test for bus set 1 with varying number of measurement samples; 500, 1000, 1500, 3000 and 6000 for (a)–(e), respectively; the frequency for y-axis represents the number of times the measurement error fall into each bin.

sample sizes 3000 and 6000 for example). Thus, the number of samples between 3000 and 6000 is a good choice.

B. Results for Multiple Bus Sets and Recommendations

Tests have been carried out on other 8 bus sets that form measurement redundant cases as well. It is found that none of the sets has a normally distributed PMU measurement error. Instead,

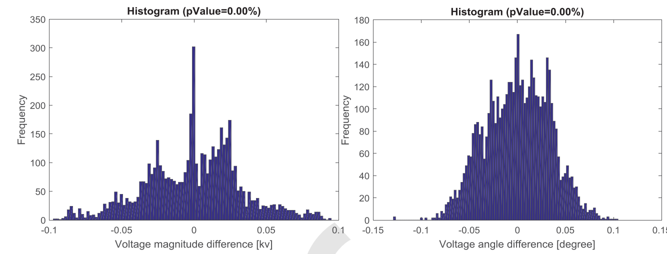

Fig. 3. Results of the Shapiro-Wilks test for bus set 2.

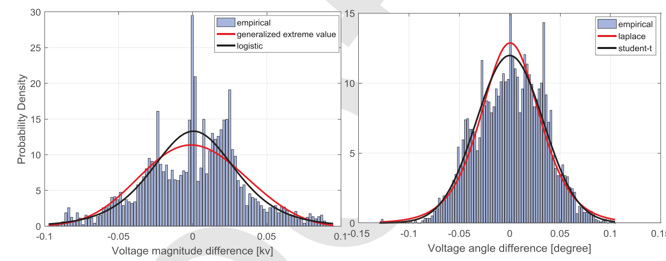

Fig. 4. Estimated probability distributions of the error terms for bus set 2.

they follow a non-Gaussian distribution with long tails, such as student-t, logistic, Laplace distributions. It should be emphasized that due to different system operating conditions, aging process of the PTs and CTs, varying communication channel noises, PMU measurement error can change from time to time. It is thus difficult to suggest a specific probability distribution of the measurement error. On the other hand, it is worth pointing out that in many signal processing problems, student-t and Laplace distributions are the two most widely used ones to model thick-tailed distributions, which can be used to simulate realistic PMU measurement error. For example, the Shapiro-Wilks test results of bus set 2 are displayed in Fig. 3. In the meantime, the estimated probability distributions of the error terms shown in Fig. 3 are displayed in Fig. 4. From these two figures, we observe that the measurement error follows a non-Gaussian distribution with long tails, such as student-t and logistic distributions. Note that these two most probable distributions are selected according to the significance confidence values of the probability distribution fitting test.

V. Conclusion

This letter proposes a simple yet effective approach to investigate the measurement error of the field PMU data. It is found that the realistic PMU measurement errors are unlikely to follow a Gaussian distribution. In future work, we will investigate the impacts of non-Gaussian measurement error on power dynamic state estimation, oscillatory modes monitoring, etc.

References

[1] J. B. Zhao, M. Netto, and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, Jul. 2017.

[2] B. Milosevic and M. Begovic, "Voltage stability protection and control using wide-area network of phasor measurements," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 121–127, Feb. 2003.

[3] Z. Huang, P. Du, D. Kosterev, and S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939–1949, May 2013.

[4] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, "Characterizing and quantifying noise in PMU data," in *Proc. IEEE Power Energy Soc. Gen. Meet.*, 2016, pp. 1–5.

[5] J. Tate and T. Overbye, "Extracting steady state values from phasor measurement unit data using FIR and median filters," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, 2009, pp. 1–8.

[6] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 613–627, May 1995.

[7] H. J. Thode, *Testing for Normality*. New York, NY, USA: Marcel Dekker, 2002.

[8] S. Shapiro and M. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3, pp. 591–611, 1965.