An Empirical Study of Synchrophasor Communication Delay in a Utility TCP/IP Network

Abstract: Although there is a plethora of literature dealing with Phasor Measurement Unit (PMU) communication delay, there has not been any effort made to generalize empirical delay results by identifying the distribution with the best fit. The existing studies typically assume a distribution or simply build on analogies to communication network routing delay. Specifically, this study provides insight into the characterization of the communication delay of both unprocessed PMU data and synchrophasors sorted by a Phasor Data Concentrator (PDC). The results suggest that a bi-modal distribution containing two normal distributions offers the best fit of the delay of the unprocessed data, whereas the delay profile of the sorted synchrophasors resembles a normal distribution based on these results, the possibility of evaluating the reliability of a synchrophasor application with respect to a particular choice of PDC timeout is discussed.

Keywords: communication delay, Phasor Measurement Unit, Phasor Data Concentrator, empirical study, power system communication

1 Introduction

Wide-Area Monitoring and Control (WAMC) systems driven by high-resolution synchrophasors have become an exceedingly important area of research in recent years [1]. Numerous work describing possible new applications have been published by researchers from the academia, for example, Ree et al. [2]. However, until now, the majority of the deployments are only limited to assisting operator decision-making, and there are few synchrophasor-based control applications available other than the wide-area damping application preliminarily tested by Chao et al. [3].

A well-recognized obstacle to deploying synchrophasor-based control application is the insufficient performance of the supporting Information Communication Technology (ICT) infrastructure: see motivations presented by Kamwa et al. [4] and Chaudhuri et al. [5]. This is partly because that the legacy proprietary ICT systems in power utilities are being gradually replaced by off-the-shelf technologies [6]. Such a paradigm shift brings in further challenges, by creating heterogeneous ICT systems that mix multiple solutions from different vendors over a long span of time. This is evidenced by a recent NASPI report stating that 71% of the networked PMUs in the power grid on the east coast of the United States failed to deliver valid data, due to problems that can be associated to ICT systems [7].

In response to the above challenges, efforts have been made to explore the capabilities of the underlying ICT systems via simulations. Ragib et al. [8], Kansal and Bose [9], and Zhu et al. [10] are typical studies carried out on ns-2 platform, while Chenine and Nordström [11], and Deng et al. [12] are examples of work performed using OPNET. Simulations do provide insight into this research question in a cost-efficient manner; however, the results may deviate from reality depending on the assumptions and simplifications made in the experiments [13]. A more convincing and reliable alternative is to draw conclusions over a large amount of empirical evidences. However, these types of results are rare from the power sectors. A possible explanation is that they may be considered too sensitive to share with the public domains. To the author’s best knowledge, the only empirical results available on this topic are from Chao et al. [3], Cyr and Kamwa [14], and Decker et al. [15]. However, the primary contributions of these studies were not reporting empirical data. Consequently, these papers did not provide detailed
descriptions of the researched ICT systems, as well as the involved empirical data.

Moreover, the existing PMU delay results were specific to the studied scenarios. Again, there is again little effort made to generalize PMU communication delay profiles. In the study by Stahlhut et al. [16], a Poisson distribution was used to describe the stochastic delay of PMU measurements. Such a suggestion was made by simplifying the entire underlying ICT system as a single router. In Chenine and Nordström [11] the PMU communication delay was instead approximated as following a normal distribution, based on simplified simulation models of a utility PMU communication network. This was in turn used in Zhu et al. [17], to illustrate the dependency of control over the performance of the supporting ICT systems.

1.1 Purpose

This paper provides an insight into the characteristics of PMU communication delay over a private utility TCP/IP network. Specifically, the delay of the unprocessed PMU data, that is, those synchrophasors received by a PDC but not yet sorted, is examined. Additionally, this paper also deals with the delay of the sorted synchrophasors leaving a PDC. The delay from the unprocessed data can be traced back to the transfer of data over communication networks, whereas the delay of the sorted synchrophasors are due to communication, data sorting and down-sampling performed at the PDC. To determine distributions with best fit, effort was made to analyze more than 3,000,000 delay observations over a period of 2 months. Based on these results, the possibility of evaluating the reliability of a WAMC application with respect to the choice of PDC timeout is also discussed.

1.2 Outline

The rest of the paper is structured as follows: Section 2 briefly describes the data sorting function provided by a PDC. Next, the distributions and the metrics employed in the goodness-of-fit studies are presented in Section 3. Section 4 reports the empirical data used in this study. Results from the goodness-of-fit studies are reported in Section 5, while their possible usages are discussed in Section 6. Finally, in Section 7, conclusions are drawn, and future work is briefly outlined.

2 PDC data aggregation

Generally, the ICT systems supporting WAMC applications include several basic components: PMUs, PDCs, applications (which themselves are composed of different information technologies and underlying algorithms), actuators (for control purposes only), and communication networks that link the aforementioned components together as a WAMC system [10]. PMUs are sensors where synchrophasors are computed and packed into frames according to the C37.118 standard [18]. The PDC is the component where voluminous synchrophasors are sorted based on the timestamp of each frame. Additionally, the frame rate of synchrophasors can be altered at a PDC; this function is referred as down-sampling. Data flows within a PDC are illustrated in Figure 1.

The primary function of a PDC is to filter out data with a notable delay. The rationale behind this is twofold: first, the majority of PMU communications are carried out over wide-area networks, so excessive delay due to communication network irregularity is inevitable; second, WAMC systems are built to exploit advantages of real-time information brought by synchrophasor technology. Therefore, synchrophasor input with excessive delay essentially brings no value to the intended applications or even runs the risk of degrading their performance, Zhu et al. [17]. This filtering function is facilitated by assigning a timeout, which is the amount of time that a PDC input buffer actively waits for synchrophasors with the intended timestamp. When the buffer set is full or the set time has elapsed, a PDC will
forward the data out to the applications. The timeout per set introduces a ceiling in terms of the maximum delay that is experienced. This configuration ensures that a PDC can always forward the available measurements within an expected time range by sacrificing data completeness.

It is intuitive that the less time a PDC waits for the incoming frames, the less time it takes by a PDC to send the sorted data out; however, it may increase the probability of packet loss. On the contrary, a long elapsed time could ensure data completeness at the cost of more notable delay. Suppose that one PMU frame is discarded by a PDC due to timeout, the calculation at the application has to be performed based on the most recently available information, which is one PMU reporting cycle old. For example, a lost frame equals to a 20 milliseconds delay when the PMU frame rate is 50 Hz. Therefore, given a similar influence on the performance of the application, packet loss can also be interpreted as a delay from the user point of view. In Section 6, we shed light on the choice of optimum timeout as a compromise between communication delay and packet loss.

At the PDC, the synchrophasor frame rate can also be altered to meet particular requirements of the applications. This is performed by discarding certain data in a systematic manner from the PDC processing buffer. This function is illustrated at the bottom of Figure 1. This down-sampling function is introduced to ensure the flexibility and scalability of the WAMC systems.

3 Theory

This chapter is divided in two parts: the first part describes the statistical distributions that were tested and the second part describes the employed metrics in goodness-of-fit studies.

3.1 Studied distributions

As presented in Section 1, there are only two distributions available to characterize the synchrophasor communication delay: a Normal distribution (N) from Chenine and Nordström [11] and a Poisson distribution (PO) from Stahlhut et al. [16]. In addition to these two distributions, we also consider other common distributions, i.e., the exponential (EXP), the log-normal (LN), the Weibull (WBL), the Gamma (GAM) and the generalized Pareto (PAR) – in terms of degree of fit to the PMU communication delay, by building analogies to similar research made to characterize delay of Internet traffic. Specifically, Csabai et al. [19] hint that the routing delay in certain segments of the European Internet under heavy loads can be modeled by log-normal distributions. Meanwhile, an empirical study reported by Papagiannaki et al. [20] suggests that single-hop queuing delay distributions are of a Weibull-like shape. Moreover, a Gamma distribution is identified as a suitable candidate to model the majority of traffic delay on fixed Internet paths in Hooghiemstra and Mieghem [21] and Bovy et al. [22]. Exponential and generalized Pareto distributions have also been considered as possible candidates to model traffic patterns in the above studies. Moreover, a bi-modal distribution containing two normal distributions is also tested, given the shape of the histogram of the delay of unprocessed PMU data. There are five parameters for such a bi-modal distribution: two means, two variances, and a mixing factor, p, which defines the significance of one normal distribution with respect to the other.

3.2 Metrics for the goodness-of-fit study

There are various metrics that can be applied to measure how well a statistical distribution matches empirical observations. The Akaike Information Criterion (AIC), a canonical technique for ranking alternative models, is applied in this research to compare the relative goodness-of-fit for the distributions. The AIC was introduced to extend the method of maximum likelihood estimation to the situation of multi-model choice by Akaike [23]. The major advantage of AIC over the other goodness-of-fit metrics is that it links Boltzmann’s entropy, Kullback-Leibler information and maximum likelihood: therefore it ties together information theory with statistics [24]. The AIC is defined as:

$$AIC = -2\ln(L) + 2k$$  \[1\]

where k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model.

A distribution receiving a lower AIC is a better fit for the data. There are many rules of thumb, all expressing essentially the same verdict. For example, Sakamoto et al. [25] maintain that if the difference of AIC between models is larger than 1~2 then the difference is considered to be significant. According to Burnham and Anderson [24], a difference of 10 is considered very significant.

The second metric employed in this study is the Quantile–Quantile plot (QQ plot), which is a graphical method for comparing two probability distributions by
plotting their quantiles against each other [26]. It differs from the AIC method by providing fine-grained visualizations with details about goodness-of-fit at each particular location of the empirical data instead of one numerical value indicating the overall goodness-of-fit. In this paper, QQ plot is employed to complement the AIC metric.

4 Empirical data

This chapter is divided into three parts. The first part gives a high level overview of the test PMU network where the statistics were captured. The second part reports the delay of the unprocessed PMU data. The delay of the sorted data is presented in the third part.

4.1 Information about the studied utility PMU network

The ICT system supporting power system operation is generally considered as one of the critical infrastructures for modern society. Therefore, there are restrictions to disclosing its details to the public. In this paper, our description is limited to the same level as a Cigré report, summarizing the common practices of power utility communication networks [27]. The purpose is to allow practitioners to interpret and use our results without disclosing sensitive information.

The studied statistics were collected from a PMU network including several PMUs from multiple vendors and an OpenPDC test platform [28]. This study deals with unprocessed data from eight PMUs, as well as two sorted data streams from two PDCs. The PMUs are located at five substations that are geographically dispersed within the power grid utility’s responsible regions. Specifically, PMU-4 and PMU-5 are deployed at the same substation, while PMU-6, PMU-7 and PMU-8 are placed at another substation. Moreover, their approximated communication distances are presented in Table 1.

The PMUs are compliant with the C37.118 standard [18], and the synchrophasors are transmitted at a unified frame rate of 50 Hz. All of the synchrophasors are aggregated by PDCs where a large collections of statistics, e.g., delay, packet loss, and data quality error etc., are captured. The primary purpose of this PDC platform is to provide a means to test synchrophasor driven applications [29]. Taking advantage of the time synchronization of all the PMUs and PDCs, this system actually offers an opportunity to actively monitor network behaviors. The time-tagged PMU measurements can be viewed as probe packets that are actively injected into the network, and communication delay characteristics can be studied by examining the properties of the received probe streams. Such a technology is referred to as network tomography by Csaba et al. [19]. It is a canonical approach to studying network congestions [20, 22].

As explained previously, power system communication infrastructures are, in most cases, heterogeneous systems mixing multiple proprietary and off-the-shelf technologies. Such a characteristic distinguishes the delay patterns of traffic over power utility communication networks from those over Internet. The core of this private communication infrastructure is built on meshed fiber optics with capacity of 34 Megabit/second. The local-area network of a substation is 100 Megabit/second. Data communication between the substation and the core network occurs over a 2 Megabit/second link.

Other than PMUs, this communication network also provides services to Remote Terminal Units (RTUs), video for surveillance purposes, and IP phone calls. Moreover, the transport layer protocol of the PMU traffic is TCP. The network supports quality of service, implemented as weighted fair queuing, which allows different scheduling priorities for the multiplexed data flows. RTU and PMU traffic have the highest priority among the data flows. Traffic due to IP phone calls has the third priority and video traffic is down-prioritized in this network. Logically, this network consists of two trust zones, which are separated by a firewall. Furthermore, the PMU traffic is not encrypted.

The purpose of this test network is to gain practical experience with PMU communication. As a response, this infrastructure is constructed as a general purpose PMU communication network based on state-of-the-art technologies rather than being specialized to provide service for any particular type of synchrophasor application. In general, the characteristics of this network correspond well to the current practices of employing an IP network for power grid operations documented in Cigré [27]. Therefore, the results of this study can be used to generalize the properties of PMU communication delay.

4.2 Delay of the unprocessed PMU data

The purpose of the study performed in this phase is to determine a statistical distribution that offers the best fit to the delay of the unprocessed PMU data. The statistics were captured in a time window of 10 seconds, between 9:00:00 AM, Nov 10th, 2012 and 9:00:00 AM, Jan 10th, 2013. During this period, no major disturbance occurred at the power system. The delay characteristics from eight PMUs are summarized in Table 1.
The size of every dataset is smaller than the expected value under ideal conditions, which is, 527,040, given our observation interval (every 10 seconds) and the study period of 61 days. This is because that all of the studied PMUs at the same substation were disconnected on an irregular basis for various reasons, e.g., for device maintenance or by network interruptions. For those PMUs deployed at the same substation, their empirical observations are always of the same sizes. This can be explained by the fact that the presented delay statistics are average values captured from a data window of 10 seconds, and the chance of losing all of the measurements from the same data window due to congestion is slim.

The statistics contain a small amount of data that deviates significantly from the mean value, for example, the maximum delay of PMU-1 reaches 1,158 milliseconds. The percentage of the delay above 100 milliseconds is presented in Table 1. Statistically, the observations falling beyond 100 milliseconds account for a very small percentage of the total data, less than 0.15% of the empirical observations from any studied PMU. This can be explained by the fact that this communication network has sufficient resources to support the above traffic types when considering the total payload of the traffic. Nevertheless, it could be argued that these excessive values are biased observations that could be traced back to network irregularities, such as interruptions or re-transmissions due to packet loss. In this paper, delays of more than 100 milliseconds are considered to be outliers and are removed from the empirical study.

The existence of these excessive delays contributes to notable deviations, $\sigma$, compared to the delay mean, $\mu$. In such a PMU network, the delay mean, $\mu$, falls into a range between 9 and 15 milliseconds. Our results match the conclusion made by Chao et al. [3] that the PMU communication delay over fiber optic networks is less than 15 milliseconds. Additionally, our results show that the distance is not a sensitive parameter to PMU communication delay in the studied network. It can be explained by the fact that synchrophasors are transferred at a rate close to the speed of light in a fiber optic network. In this study, the maximum difference among the communication distances is about 370 kilometers, between PMU-1 and PMU-4/PMU-5. Such a distance only contributes to a delay difference of about one millisecond.

The histograms of delay from the eight PMUs resemble a similar shape. Due to space constraints, only that from PMU-1 is illustrated in Figure 2. The histogram presents two spikes: one between 5 and 10 milliseconds and one between 10 and 30 milliseconds, both of which manifest as normal distributions. In Section 5, goodness-of-fit studies are employed to provide quantified evidences for this preliminary conclusion.

### Table 1 Overview of the delay of the unprocessed PMU data (in milliseconds).

<table>
<thead>
<tr>
<th>PMU</th>
<th>Observation</th>
<th>Distance (km)</th>
<th>Min/Max</th>
<th>$\mu/\sigma$</th>
<th>Outliers (&gt;100 ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>467,819</td>
<td>130</td>
<td>1/1,158</td>
<td>12.66/8.56</td>
<td>0.0618%</td>
</tr>
<tr>
<td>2</td>
<td>447,580</td>
<td>480</td>
<td>3/3,019</td>
<td>9.84/10.48</td>
<td>0.1007%</td>
</tr>
<tr>
<td>3</td>
<td>526,181</td>
<td>450</td>
<td>1/2,390</td>
<td>10.48/10.10</td>
<td>0.0650%</td>
</tr>
<tr>
<td>4</td>
<td>467,773</td>
<td>500</td>
<td>1/6,215</td>
<td>14.13/13.95</td>
<td>0.0767%</td>
</tr>
<tr>
<td>5</td>
<td>467,773</td>
<td>500</td>
<td>1/5,509</td>
<td>14.11/18.53</td>
<td>0.0808%</td>
</tr>
<tr>
<td>6</td>
<td>77,311</td>
<td>400</td>
<td>8/1,073</td>
<td>11.44/6.45</td>
<td>0.1033%</td>
</tr>
<tr>
<td>7</td>
<td>77,311</td>
<td>400</td>
<td>9/1,525</td>
<td>11.54/15.13</td>
<td>0.1122%</td>
</tr>
<tr>
<td>8</td>
<td>77,311</td>
<td>400</td>
<td>9/1,458</td>
<td>11.32/12.67</td>
<td>0.1277%</td>
</tr>
</tbody>
</table>

The sorted synchrophasor data

The sorted synchrophasors are produced at a PDC by aggregating multiple incoming frames from the connected PMUs. In this paper, two sorted PMU data streams...
from two PDCs are examined. The elapsed timeout is set to 23 milliseconds for PDC-1 and to 29 milliseconds (PDC-2). These choices are made based on a previous study (details are provided later in Section 6). The delay results were captured from a time window of 10 seconds, between 9:00:00 AM, Dec 10th, 2012 and 9:00:00 AM, Jan 10th, 2013. The delay characteristics from the two PDCs studied are summarized in Table 2.

The actual sizes of the PDC output datasets are very close to their expectations, that is, 267,840, given the observation interval of every 10 seconds, and the study period of 30 days. This is because that these two PDCs are actually redundant software copies deployed on an Open PDC platform with different timeout settings. Given identical input, it is reasonable for them to produce exactly the same amount of statistics. The delay of the PDC output traffic rises notably as compared to their counterparts from the PDC input data streams, due to processing and waiting in PDC buffers. Similar patterns could be found from simulations studies reported by Zhu et al. [30]. On the other hand, the delay results are less dispersed compared to those of the unprocessed data, this is evidenced by smaller deviations. PDC-1 in general has slightly smaller delay values comparing to PDC-2 due to its lower timeout setting.

The histograms of the delay of the PMU data sorted by PDCs have a similar shape. Due to space constraints, only the PMU data sorted by PDC-1 are illustrated in Figure 3. The histogram presents a bilaterally symmetrical bell shape, which resembles a normal-distribution-like shape. In Section 5, goodness-of-fit studies are employed to validate this preliminary conclusion.

5 Results

This chapter describes results from this study and it is divided into two parts, along with the two types of PMU data studied.

5.1 Delay of the unprocessed PMU data

AIC scores for the tested distributions are given in Table 3. A bi-modal distribution composed of two normal distributions provides the best fit for the delay of the unprocessed PMU data streams, given its lowest AIC scores for all of the eight PMUs studied. Additionally, our study provides no empirical support for Poisson distributions or normal distributions, because the difference between their AIC scores and that for the bi-modal distribution is significantly larger than 10. Maximum likelihood parameter estimates for bi-modal distributions of the delay profile of the unprocessed PMU data are presented in Table 4.

Table 2 Overview of the delay from the sorted synchrophasors (in milliseconds).

<table>
<thead>
<tr>
<th>PDC</th>
<th>Observations</th>
<th>timeout</th>
<th>Min/Max</th>
<th>μ/σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>267,614</td>
<td>23</td>
<td>19/67</td>
<td>40.00/3.11</td>
</tr>
<tr>
<td>2</td>
<td>267,614</td>
<td>29</td>
<td>19/67</td>
<td>40.14/3.15</td>
</tr>
</tbody>
</table>

Table 3 AIC scores of the delay from the unprocessed PMU data.

<table>
<thead>
<tr>
<th>PMU</th>
<th>EXP</th>
<th>LN</th>
<th>WBL</th>
<th>GAM</th>
<th>PAR</th>
<th>PO</th>
<th>N</th>
<th>Bi-modal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3,309,300</td>
<td>2,856,500</td>
<td>3,037,400</td>
<td>2,872,800</td>
<td>3,307,800</td>
<td>3,094,100</td>
<td>3,335,300</td>
<td>2,739,700</td>
</tr>
<tr>
<td>2</td>
<td>1,625,500</td>
<td>1,013,300</td>
<td>1,481,100</td>
<td>1,128,900</td>
<td>1,625,500</td>
<td>1,134,500</td>
<td>1,864,900</td>
<td>1,010,900</td>
</tr>
<tr>
<td>3</td>
<td>3,523,400</td>
<td>2,247,500</td>
<td>3,188,600</td>
<td>2,423,600</td>
<td>3,523,400</td>
<td>2,435,300</td>
<td>3,925,200</td>
<td>2,136,700</td>
</tr>
<tr>
<td>4</td>
<td>3,411,700</td>
<td>2,039,000</td>
<td>3,098,300</td>
<td>2,273,400</td>
<td>3,411,700</td>
<td>2,292,800</td>
<td>3,794,600</td>
<td>1,969,400</td>
</tr>
<tr>
<td>5</td>
<td>3,410,400</td>
<td>2,027,000</td>
<td>3,152,500</td>
<td>2,337,600</td>
<td>3,410,000</td>
<td>2,273,600</td>
<td>4,056,700</td>
<td>1,877,200</td>
</tr>
<tr>
<td>6</td>
<td>4,821,000</td>
<td>2,214,300</td>
<td>4,190,200</td>
<td>2,673,600</td>
<td>4,818,300</td>
<td>3,109,300</td>
<td>4,599,200</td>
<td>2,102,100</td>
</tr>
<tr>
<td>7</td>
<td>4,803,700</td>
<td>2,332,800</td>
<td>4,454,700</td>
<td>3,210,300</td>
<td>4,802,700</td>
<td>3,121,700</td>
<td>5,764,900</td>
<td>2,224,400</td>
</tr>
<tr>
<td>8</td>
<td>4,805,500</td>
<td>2,305,200</td>
<td>4,393,700</td>
<td>3,064,500</td>
<td>4,804,400</td>
<td>3,123,000</td>
<td>5,515,500</td>
<td>2,252,700</td>
</tr>
</tbody>
</table>
As mentioned in the previous section, the delay profile of un-processed data from PMU 1 is representative for the delay from all of the PMUs studied. Therefore, only QQ plots for the delay profile of PMU-1 along with the tested distributions are presented in Figure 4. Among all of the candidate distributions, the empirical data in general follows the bi-modal distribution best. The only exception is that the identified distribution returns liberal estimations when the delay is close to 40 milliseconds. This is consistent with the results from the AIC results presented above. Therefore, our study suggests that a bi-modal distribution offers the best fit to the delay profiles of unprocessed PMU data. Specifically, a bi-modal distribution with \( \mu_1 \) between 7.5 and 13, \( \mu_2 \) between 16 and 18, \( \sigma_1 \) between 0.5 and 2.5, and \( \sigma_2 \) between 1.6 and 4.0 can be used to model the delay of the unprocessed PMU data.

### 5.2 Delay of the sorted synchrophasor data

AIC scores for the tested distributions with respect to the two studied sorted synchrophasor streams studied are presented in Table 5. Among all the candidate distributions, a normal distribution provides the best fit for all of the PDC delay statistics. Maximum likelihood parameter estimates for normal distributions that could model the delay of the sorted PMU data are reported in the last column in Table 5.

As mentioned in the previous section, the delay from PDC 1 is representative for the delay from all the studied PDCs. Therefore, only QQ plots of the delay profile of

<table>
<thead>
<tr>
<th>PMU</th>
<th>Mixing factor p</th>
<th>( \mu_1 )</th>
<th>( \mu_2 )</th>
<th>( \sigma_1 )</th>
<th>( \sigma_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4375</td>
<td>7.6368</td>
<td>16.4227</td>
<td>1.4794</td>
<td>4.0148</td>
</tr>
<tr>
<td>2</td>
<td>0.5680</td>
<td>7.9142</td>
<td>10.1201</td>
<td>0.6342</td>
<td>3.3541</td>
</tr>
<tr>
<td>3</td>
<td>0.3935</td>
<td>7.7890</td>
<td>12.1232</td>
<td>1.4213</td>
<td>2.5111</td>
</tr>
<tr>
<td>4</td>
<td>0.2333</td>
<td>12.5431</td>
<td>18.0002</td>
<td>2.4321</td>
<td>3.0001</td>
</tr>
<tr>
<td>5</td>
<td>0.7032</td>
<td>12.8636</td>
<td>15.3219</td>
<td>1.0312</td>
<td>2.4480</td>
</tr>
<tr>
<td>6</td>
<td>0.6211</td>
<td>10.1232</td>
<td>13.0982</td>
<td>1.0999</td>
<td>1.6812</td>
</tr>
<tr>
<td>7</td>
<td>0.3754</td>
<td>11.0232</td>
<td>13.5198</td>
<td>0.8591</td>
<td>2.3008</td>
</tr>
<tr>
<td>8</td>
<td>0.2915</td>
<td>12.1321</td>
<td>15.3291</td>
<td>1.0891</td>
<td>1.6799</td>
</tr>
</tbody>
</table>

Table 4 Maximum likelihood parameter estimates for bi-modal distributions.

![Figure 4](image-url) QQ plots of the delay of unprocessed data from PMU-1.
PMU data sorted by PDC-1 are presented in Figure 5. Though a normal distribution returns conservative estimates for delay in the range between 25 and 30 milliseconds, it is evident that it provides the best fit for the empirical data in general. Therefore, the QQ plots again suggest that a normal distribution with delay mean, $\mu$, between 3.68 and 3.70 and standard deviation, $\sigma$, between 0.006 and 0.007 is a suitable candidate to model the PMU communication delay including PDC sorting. The simulation and the empirical study reported by Hooghiemstra and Mieghem [21] reveal that the processing delay of several routers together can as well be approximated as following a normal distribution, also. Such a similarity hints on that a PDC could exert a similar influence on PMU communication delay as much as a cluster of routers does.

6 Discussion

While there is a plethora of literature dealing with the reliability of power grids, as well as that of each individual component, the reliability of a WAMC system is still an open topic. The results from this study provide a means to evaluate the reliability margin of an intended application considering its input delay and packet loss with respect to a particular choice of PDC timeout.

Let $T_d$ denote the maximum delay that can be tolerated by an application. It is a known value from the application design. The delay probability density function, $P_d$ is estimated by performing studies as described in Section 5. By referring $T_d$ to $P_d$, the probability of this

<table>
<thead>
<tr>
<th>PDC</th>
<th>EXP</th>
<th>LN</th>
<th>WBL</th>
<th>GAM</th>
<th>PAR</th>
<th>PO</th>
<th>N</th>
<th>$N \mu/\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,509,600</td>
<td>1,373,000</td>
<td>1,431,800</td>
<td>1,370,100</td>
<td>2,248,500</td>
<td>1,544,500</td>
<td>1,366,900</td>
<td>3.6858/0.0062</td>
</tr>
<tr>
<td>2</td>
<td>2,511,500</td>
<td>1,379,000</td>
<td>1,443,800</td>
<td>1,376,000</td>
<td>2,248,700</td>
<td>1,546,600</td>
<td>1,373,000</td>
<td>3.6893/0.0063</td>
</tr>
</tbody>
</table>

Table 5: AIC scores of the delay of the sorted PMU data, and the parameters from distributions with best fit.
application that fails to deliver its intended functionality can be estimated as, \( P_d(T_d) \).

The reliability margin calculation can also be expanded to include the impact of input delay and packet loss due to the PDC timeout. Because irregularity of communication networks is inevitable, the maximum delay due to communication, \( T_{max} \), needs to be defined by referring to a confidence interval, for example, 99.9%. Moreover, packet loss poses a similar threat to the applications as delayed input does because whenever a packet is lost, the computation (performed by the application) has to be performed based on the most recently available measurements (in this case, old measurements). This forms a theoretical foundation to transform the tolerable delay into tolerable number of consecutive packet losses. The maximum tolerable delay, \( T_d \), can be transformed to the number of tolerable consecutive packet losses by \( T_d \times f_s \), where \( f_s \) is the PMU frame rate. Following the same rationale, \( T_{max} \) is converted to the corresponding number of packet losses as \( T_{max} \times f_s \). Using a liberal assumption that the maximum delay always occurs at the first packet available after consecutive packet losses, a linear equation can be derived as:

\[
N_{\text{pl}} = \text{round}(T_d \times f_s - T_{max} \times f_s) \tag{2}
\]

where \( N_{\text{pl}} \) is the tolerable consecutive packet loss due to the PDC timeout. An assumption here is that there is no packet loss due to communication, because, over a long span of time the packet loss is negligible in a TCP network (this is justified by our empirical data that will be published in another paper). The probability of packet loss due to PDC timeout from a single PMU, \( P_d(T_{TO}) \) can be calculated by referring a specific \( T_{TO} \) to the known delay distribution, \( P_d \). The probability of application failure (due to consecutive packet loss from the PDC sorting) is defined as \( P_d(T_{TO})^{N_{\text{pl}}} \). For applications that are dependent on multiple PMU measurements, the probability of the function failure should be expressed as a product of the above probability from each involved PMU. The reliability margin of a synchrophasor driven WAMC system, \( R_C \) can be written as a function of the PDC timeout, \( T_{TO} \) as:

\[
R_C = 1 - \prod_{j=1}^{n} P_d(T_{TO})^{N_{\text{pl}}} \tag{3}
\]

where \( n \) is the total number of PMU measurements utilized by this synchrophasor application.

7 Conclusion and future work

This paper brings important insight into the characterization of the delay of synchrophasors by performing large-scale empirical studies. Our goodness-of-fit study results suggest that the delay of the unprocessed PMU data could be approximated by a bi-modal distribution containing two normal distributions, while the delay of the sorted PMU data follows a normal distribution best, among all of the studied distributions. These results offer a prospect for better predictions of PMU communication delay, as well as other information that is relevant to the research consortium.

The Open PDC platform provides a comprehensive set of statistics, that paves the way for many future studies. In this paper, the impact of sorting and down-sampling algorithms on PMU communication delay are not examined. A tempting next step is to perform a sensitivity analysis on the sorted PMU data delay with respect to different sorting and down-sampling algorithms.

Acknowledgement: The authors would like to acknowledge Hannes Holm and Ulrik Franke for their valuable input.

References