Fast FCAS Sampling Verification in Support of Market Ancillary Services Specification (MASS) consultation

- *Phase 3*

Prepared for the Australian Energy Market Operator

Han Wang, Pierluigi Mancarella

The University of Melbourne

December 2021

Executive Summary

The Australian Energy Market Operator (AEMO) initiated a consultation on the amendment of the Market Ancillary Service Specification (MASS), with regards to the requirements on the measurements for Distributed Energy Resource (DER) to participate in the Contingency Frequency Control Ancillary Services (FCAS) markets. Two stages of the MASS consultation have already been completed, whilst the third stage of the MASS consultation is still in progress. During the second stage of the MASS consultation, the University of Melbourne (UoM) was commissioned by AEMO to assess the impact of the sampling rates on the FCAS verification, during which high-resolution data (20ms/50ms) for aggregated responses from different VPP under various events were provided by AEMO. UoM was again commissioned by AEMO during the third stage of the MASS consultation to assess a wide range of factors that can affect the verification error and to establish a methodology to identify potential oscillatory responses. In that second study, UoM was provided with high-resolution data for various synchronous generators, as well as calculated NMI-level data for 1000 sites that had been made available by one market stakeholder and were characterised by a typical (droop-like) type of response for a variable/proportional FCAS controller. During the first and second studies, UoM identified seven key factors of relevance for the request analysis, namely, sampling rates, integration rules, frequency disturbance time, site aggregation, inertial response, compensation factor, and power measurement error.

In November 2021, UoM was again commissioned by AEMO to further study the impact of aggregations on the verification error on the basis of actual/measured NMI-level data (for about 1600 sites) that had been made available by another market stakeholder. Overall, this was a more diverse dataset containing NMI-level data (i.e., frequency and power measurements) from sites with devices with different characteristics, and including, in particular, step-like responses, for two different contingency events.

The study on site aggregation was repeated with the new dataset, based on which it is possible to draw the following conclusions, also in comparison with the results previously found:

- Similar trends as in the previous report may be observed, i.e., the range of verification errors decreases when increasing the sampling rate and the number of sites in the aggregation.
- From the analysis of the two events analysed, it emerged that the definition of "disturbance time" used in the current methodology may need to be revisited, in order to be better aligned with the actual time of local contingency detection (which is when the device is supposed to start responding). In fact, particularly in one of the two events analysed, the frequency was found to be "hovering" and oscillating around the normal operating frequency band (NOFB) for a few seconds prior to the actual contingency event detection. By applying the current definitions and without a local proxy of the contingency event time, which would require some adjustments even for data captured at 50ms intervals, the relevant frequency measurements might create a misalignment between the frequency disturbance time as currently defined and the actual contingency event time as locally detected.
- In particular, when determining the assessment window in line with the requirements of the MASS, the current definition of the frequency disturbance time without a local proxy to identify the contingency event time may create relatively large errors even for relatively high sampling rates (e.g., 100ms and 200ms). It appears that this may be again essentially due to a misalignment of the defined frequency disturbance time and the time when the relevant contingency event is actually locally detected.

- A local proxy of the contingency event time should be properly defined and implemented in the FCAS verification tool, to better represent the whole VPP's performance in the six-second window after the locally identified contingency event. This is because the local devices making up the VPP aggregate should provide response only when the contingency is *locally* detected, which may be different for devices at different locations. If such a local proxy of the contingency event can be well identified, the verification errors may decrease significantly.
- In this report, we have used a heuristic approach to define such a proxy, which we called "contingency time identifier". However, this should be considered only for the purposes of this work, and more and systematic work is required to develop a robust methodology to identify the local proxy of the contingency event.
- Considering all the datasets analysed and the three studies performed so far, it appears that higher sampling rates and larger number of sites may significantly decrease verification errors. However, it also emerges that the specific numerical impact of these two factors is event-specific and may also depend on the type of devices and their response characteristics. It is therefore advised that the results of this work should be revisited at a later stage as "work in progress", possibly using a wide range of several datasets for different contingency events, technologies, locations, etc. This applies to all levels of aggregation, and is particularly true for VPPs that have a relatively small number of sites (e.g., less than 100 sites), for which more data and studies are required to come up with more robust results.

Table of Contents

1 Introduction

In January 2021, the Australian Energy Market Operator (AEMO) launched a consultation on the amendment of the Market Ancillary Service Specification (MASS) with regards to measurements requirements for Distributed Energy Resource (DER) to participate in the Contingency Frequency Control Ancillary Services (FCAS) markets. Two stages of the MASS consultation have already been completed, while the third stage of the MASS consultation is ongoing. The University of Melbourne (UoM) was commissioned in the second stage and third stage of the MASS consultation to perform independent analysis on the fast FCAS verification process. The studies aimed to identify the impact of different factors on the verification errors, including sampling rates, integration rules, frequency disturbance time, site aggregation, inertial response, compensation factor, and power measurement error.

The first study generally concluded that 1s sampling rate might introduce relatively significant verification errors in FCAS contribution assessment and a trapezoid interpolation rule generally outperform Reiman methods [1]. Then, UoM discussed how the FCAS verification tool was effectively using what was then called "relative-window" method, which defines the start of the six-second assessment window on the basis of the frequency disturbance time, i.e., the first recorded point that is outside the normal operating frequency band (NOFB) [1]. However, this method was inherently designed for high-speed data (i.e., 20ms and 50ms) and would not generally be suitable for much lower sampling rates as it would introduce errors in the assessment of the frequency disturbance time. A theoretical method was then introduced, for reference, in the first report, namely, the "universal window" method, which assumed that the NOFB is crossed simultaneously for all providers in each event. However, the application of this reference method would be hard to achieve in practice [1].

In the second study, a novel "RoCoF-based" method was then proposed, which has similar performance as the "universal window" one and could be more readily implemented in the FCAS verification tool [2]. In this second study, the verification errors for synchronous generators were also analysed. The results indicated that lower sampling rates generally result in inaccurate estimation of the inertial response and compensation factor; hence, it was recommended that the current requirement of 50ms sampling rate should remain for synchronous generators [2]. Moreover, relaxing power measurement errors from 2% to 4% might introduce significant verification errors, and may thus not be suitable [2]. Another question that was specifically addressed in the second study is site aggregation. The results indicated that using NMI-level data instead of aggregated response profiles can substantially reduce the verification errors, and the range of verification errors decreases when the number of sites being aggregated increases [2]. However, due to the data availability, the second report pointed that this matter should be studied in more detail, particularly in conjunction with the effect of sampling rate, by making use of more and possibly more diverse data.

In November 2021, UoM was again commissioned by AEMO to perform further analyses on the site aggregation, whilst new datasets with more and more diverse responses were provided. More specifically, locally measured (at 50ms sampling rate) frequency and power data for individual NMIs of over 1600 sites and for two contingency events were given. The main focus of this study is put on the impact of the site aggregation and sampling rate on the verification errors based on the new dataset provided. This report aims to bring further insights and recommendations based on the new NMI-level data provided and takes into account the analyses previously conducted.

2 Methodology

2.1 FCAS response verification process

This report focuses on the verification of fast FCAS response of VPP using NMI-level data. The FCAS response verification process that is used follows the methodology defined in the second report [2], which is in line with the FCAS verification tool methodology. This methodology is summarised in the flowchart in Figure 1. The steps that are used in this report are highlighted as red blocks.

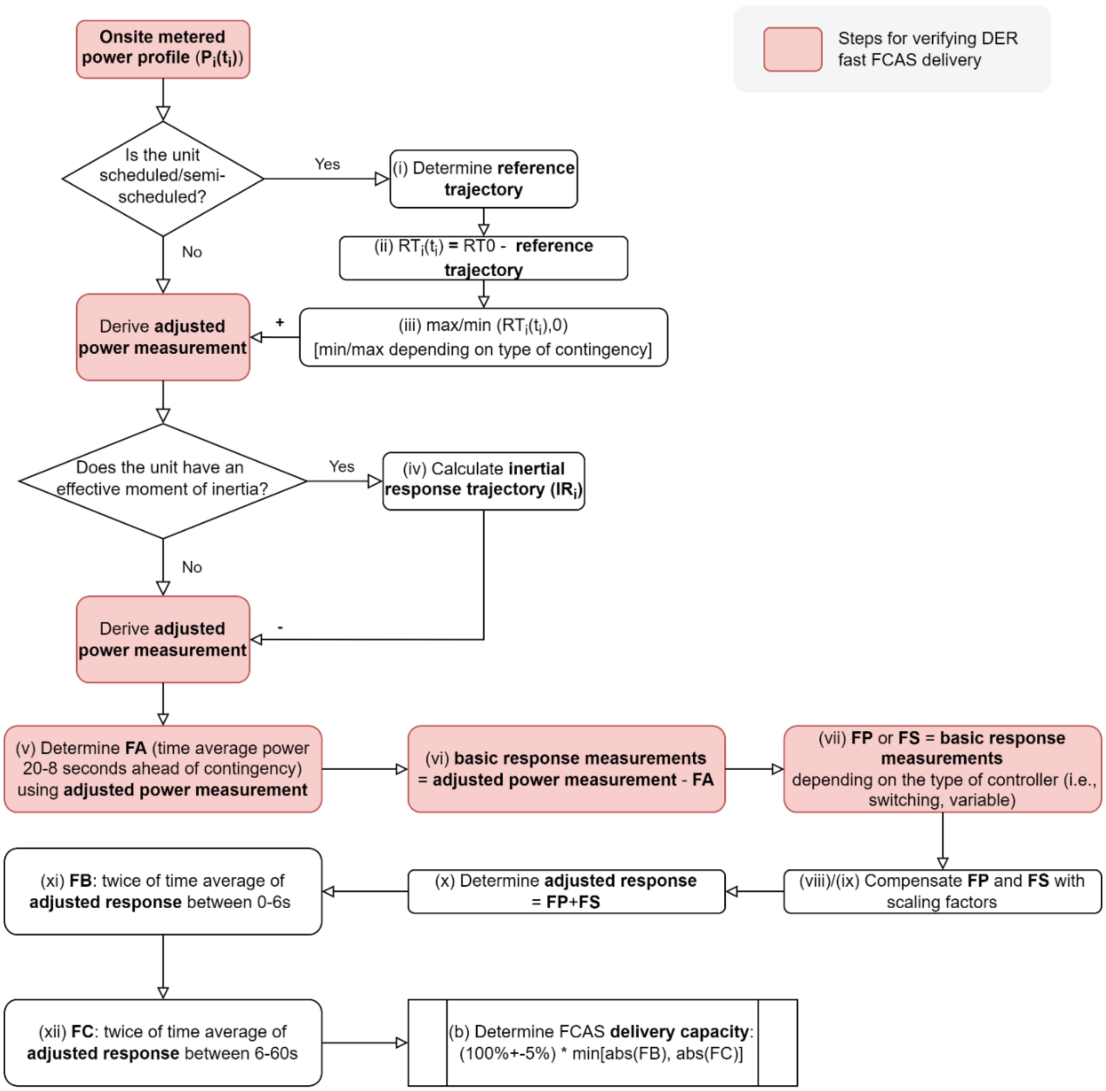


Figure 1. Fast FCAS performance verification methodology (redrafted from [1])

More specifically, the following steps are implemented to verify the fast FCAS delivery for each site:

- Derived adjusted power measurement: it is assumed that the VPP is not classified as scheduled or semi-scheduled units; thus, the power measurement remains unchanged.
- Determine baseline (indicated as *FA*): calculate the average power of the profile before the *frequency disturbance time*. The frequency disturbance time is the time at which the local frequency falls or rises outside the NOFB, during a frequency disturbance [3]. Instead of using the

time average power of the 20s to 8s before frequency disturbance time, the time average power of 3s to 0s is used, as the high-speed data is only available for a 5s prior to the frequency disturbance time, in line with the requirement for fast FCAS verification [3].

- Basic response measurements: the basic response measurements are calculated by subtracting the baseline from the adjusted power.
- Compensation factor: it is assumed that the compensation factor is 1 for the “step-change” response.
- Calculate fast FCAS delivery: calculate the power delivery in the assessment window using the trapezoid rule.

The calculated fast FCAS delivery is then used to calculate verification errors.

2.2 Assessment window

In the FCAS verification tool, the “assessment window”, which is used to evaluate the fast FCAS delivery in contingency events, is defined as the six-second time interval starting from the “frequency disturbance time”. The frequency disturbance time is defined as the time at which local frequency falls or rises outside the NOFB [4] and is intended to be used as a local measurement proxy following the start of the relevant contingency event. This refers to “t=0” in the MASS. The frequency disturbance time has been discussed thoroughly in the first and second reports, whilst different approaches to determine frequency disturbance time were tested. Among all alternative methods to identify frequency disturbance time, the “RoCoF-based” method exhibits performance that is comparable with the “universal window” [2].

Ideally, the frequency disturbance time should be close to the contingency time, and the assessment window should adequately capture the response during the first six-second after the contingency. However, the definition of frequency disturbance time is not fully unambiguous, which under specific circumstances might even lead to verification errors. In fact, for the purpose of response verification, FCAS service providers are required to provide data that is around the frequency disturbance time (i.e., 5s before and 60s after contingency [3]) identified by AEMO. Whilst the contingency event time is somehow universal, the frequency disturbance is assessed locally, on the basis of its definition as the time at which the local frequency falls or rises outside the NOFB. However, in specific cases, e.g., when before the contingency the frequency fluctuates around the NOFB boundary, an application of this definition might lead to having a response verification assessment starting before the local contingency is detected (and thus before the device starts actually responding). An example is illustrated in Figure 2. It can be seen that, by applying the current definition, the frequency disturbance time might be considered as happening at t_0, when the frequency crosses (even if just for one or two cycles) the NOFB for the first time, whilst the contingency that is detected locally actually happens around t_C (this time is also intuitively closer to the actual contingency time). The frequency “hovers” and oscillates around the NOFB between t_0 and t_C. Therefore, to properly verify the fast FCAS performance of the device, the assessment window should start at t_C, as this is the time when the device is indeed supposed to provide fast FCAS response. Thus, a more unambiguous definition of the frequency disturbance time is needed, so that it can be used as a local proxy for the contingency event time.

To address this issue, and only for the purposes of the studies in this report, we propose a new definition of the assessment window, which starts at the local proxy contingency time (i.e., t_C) and end at 6s after t_C. In this way, the assessment window should be able to more clearly capture the device’s response within the 6s after the contingency event.

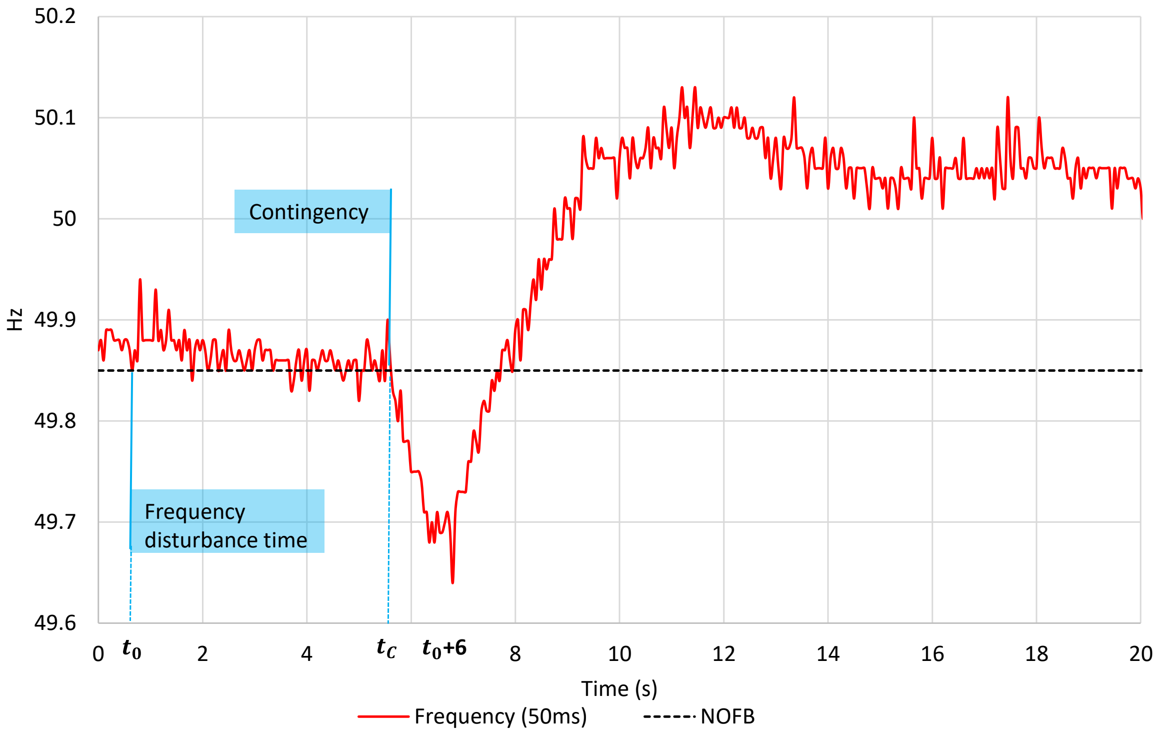


Figure 2. Example of local frequency measurement of one NMI before, during and after contingency event

In order to assess the impact of the assessment window on the verification error, we therefore consider the following methodologies to define the assessment window when calculating fast FCAS delivery:

a. Defining the assessment window as the six-second window from the *frequency disturbance time* as currently defined and used in the FCAS verification tool.
b. Defining the assessment window as the six-second window from the locally detected contingency, which is assessed through a heuristic method. This method is proposed in this report to properly capture the fast FCAS response of individual NMIs for the specific events considered here.

In particular, we introduce the concept of "contingency event identifier", to approximate the contingency event time using locally measured frequency.

2.2.1 Contingency event identifier

The concept of contingency event identifier is used to approximate the contingency event time using local measurements of the frequency. As mentioned above, in this report two events were studied. Based on the behaviour of the frequency measured at NMI-level for these events, and only for the purposes of this report, we proposed a heuristic rule to better identify the local proxy of the contingency event for these two particular events, that is, using a rolling window of 50ms to find the first recorded point when the frequency consistently lies outside the NOFB for at least 250ms. Once the contingency event is located in such a way, the frequency fluctuations prior to the contingency event are ignored, and the assessment window starts at the newly defined proxy of the contingency event time. It is worth noting that this was done only for the scope of this work and the specific events analysed and is not a general rule, as it highly depends on the characteristics of the contingency events and will be different for other events. In the context of this work, the proposed heuristic rule is only to provide an approximation of the contingency event time based on the local measurement and to assess the numerical value of verification errors when considering contingency event time.

2.3 Verification error

The term verification error is used to measure the performance of potential verification configurations. The verification error indicates the relative difference of FCAS delivery when changing the underlying assumptions, e.g., sampling rate, frequency disturbance time, integration rule, etc. In this report, the benchmark is represented by the FCAS delivery (in kW·s) calculated using 50ms data (the highest resolution available for the data provided) with the "universal window" method and trapezoid rule[1].

2.4 Further insights into error analysis via probabilistic assessment

In this report, Monte Carlo simulation and probabilistic metrics are adopted to provide further insights into error analysis. Monte Carlo simulation is a mathematical technique that is used to estimate the possible outcomes and the outcomes' distribution of single/compound uncertain events or stochastic processes that cannot be easily predicted due to the intervention of random variables. In Monte Carlo simulation, uncertain inputs in a model are represented using probability distributions or scenarios. During a Monte Carlo simulation, values are sampled randomly from the input probability distributions or scenarios. Each set of samples is called an iteration, and the resulting outcome from that sample is recorded. Then, Monte Carlo simulation repeats this process, and the result is a probability distribution of possible outcomes. In particular, specific metrics can be introduced to assess the likelihood of extreme events (e.g., worst and best cases) in probabilistic terms (e.g., by making use of percentiles and distribution tail analysis), besides expected values. Compared to single-point estimates (also called "deterministic" risk analysis) for, e.g., worst case, best case, and most likely case, Monte Carlo simulation provides more valuable information, as it does not only show what could happen but also show how likely it is to happen and the shape of the probability distribution. It is widely considered as a state-of-art methodology for risk analysis and is suitable to provide a better picture of possible verification errors given a certain dataset.

Monte Carlo simulation has been used here to capture the uncertainty of device responses in different levels of aggregation. In fact, a VPP may contain a wide range of devices with similar or distinctive characteristics, which may lead to diverse verification errors. In this report, Monte Carlo simulations are conducted with 10,000 iterations[2] to build the probability distribution functions of verification errors. The adopted Monte Carlo simulation process is shown in Figure 3. For each iteration, we randomly select NMI's data from the entire dataset for the assessment, assuming each NMI is equally likely to be selected. We then calculate the fast FCAS delivery for individual NMI for the test case, i.e., using lower sampling rate (i.e., 100ms, 200ms, and 1s) data and "RoCoF-based" method, as well as for the benchmark, i.e., using 50ms data and "universal window" method. In the next step, the fast FCAS delivery of each individual NMI

---

[1] It is worth noting that a small verification error does not necessarily indicate that the provider would have an acceptable performance in terms of FCAS delivery as recognised by AEMO, but rather that the fast FCAS delivery calculated with the given settings is close to the benchmark response (calculated with 50ms data, "universal window" method and trapezoid rule).

[2] An analysis on the accuracy level of the Monte Carlo simulation was carried out and the results showed that, with the datasets provided,10,000 iterations are generally sufficient to achieve satisfactory accuracy levels for a relatively large number of sites (e.g., 100 sites or more). For aggregations with fewer sites, i.e., less than 100 sites, more data and studies would generally be needed to identify the error distribution more robustly.

is summed up to calculate the aggregated fast FCAS delivery, for both test case and benchmark. Finally, the verification error is calculated in the form of percentage error, using the following formula:

$$percentage\ error\ (\%) = \frac{E-E^{base}}{E^{base}} * 100 \tag{1}$$

where E and E^{base} is the aggregated fast FCAS delivery calculated for the test case and benchmark, respectively, whilst n is the number of sites in the aggregation.

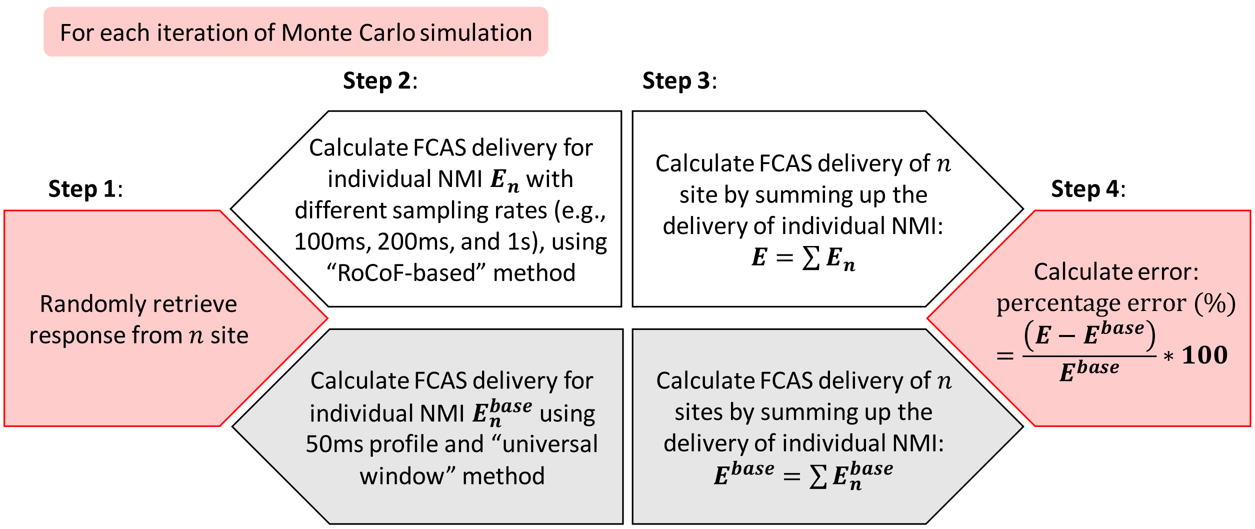


Figure 3. Process of Monte Carlo simulation

Monte Carlo simulation can also provide an estimate of the verification error confidence interval. In particular, besides the maximum and minimum value of the verification errors, percentile scores may also be used to measure, for example, the "risk" of having too high errors. The percentile score $P_{p\%}$ is a score below which falls a given percentage p of scores in the distribution. For example, the 50th percentile (the median) is the score below which lie 50% of the samples in the distribution. Different percentile scores are used considering both the upper-tail and lower-tail of the distribution, e.g., 50th percentile ($P_{50\%}$), 2.5th/97.5th percentile ($P_{2.5\%}/P_{97.5\%}$), 2th/98th percentile ($P_{2\%}/P_{98\%}$), 1.5th/98.5th percentile ($P_{1.5\%}/P_{98.5\%}$), 1th/99th percentile ($P_{1\%}/P_{99\%}$). For example, we can assume that 95% of the verification errors fall between the 2.5th percentile score and 97.5th percentile score. Overall, different metrics may be considered together in order to have a better picture of the error distribution shape.

2.5 Site aggregation

In the VPP Demonstrations programme, AEMO verifies the performance of a VPP based on the aggregated response to delivery contingency FCAS during a frequency disturbance. The aggregators are required to measure frequency and power response at the relevant sites. As the clock associated with individual meters at the relevant sites may record slightly different times, the data is aligned with the frequency disturbance time, i.e., the time at which the local frequency falls or rises outside the NOFB.

However, as indicated in section 2.1, using frequency disturbance times may create significant misalignment with the actual contingency time; hence, the aggregated response calculated according to the current FCAS assessment tool may not be representative of the actual response that the VPP fleet provided after the contingency. Thus, the methodology for aggregating response profiles needs to be revisited.

3 Case studies

In this report, UoM was provided with high-speed (i.e., 50ms sampling rate) NMI-level data of over 1600 sites for two events, namely, event 1 and event 2[3]. The impact of site aggregation and sampling rate on the verification errors is carried out using the following methodologies to identify assessment windows:

a. Defining the assessment window as the six-second window from the frequency disturbance time as currently defined.
b. Defining the assessment window as the six-second window from the local proxy of the contingency event described above.

Note that the time of contingency that is locally detected is only an estimation using the methodology that is proposed in 2.2.1 for the purpose of the analysis of this report. More work is needed to identify a methodology that could be generally applied to any event. The purpose of the local contingency identifier is to estimate the verification error when the assessment window is defined based on the local proxy of the contingency event. Monte Carlo simulation is used with 10,000 iterations to model the distribution of the verification errors.

3.1 Verification errors

In this section, the FCAS delivery verification errors for different numbers of aggregated sites and different sampling rates are analysed for two definitions of assessment window, i.e., the six-second window after the currently defined frequency disturbance time, and the six-second window after the contingency is locally detected using the proxy described above.

3.1.1 Assessment window defined as six-second window after frequency disturbance time

This section analyses the verification errors when using the frequency disturbance time to define the start of the six-second assessment window, which represents the methodology that is currently used in the FCAS verification tool. The fast FCAS delivery is calculated for 100ms and 200ms sampling rate, using the "RoCoF-based" method, for the two events. The benchmark is calculated using 50ms data and the "universal window" method. An analysis for 1s sampling rate was also performed, and the results are shown in the Appendix, Table 9-Table 10.

The percentage errors for the event 1 using 100ms and 200ms sampling rate are illustrated in Table 1-Table 2. The analysis is done for different numbers of aggregated sites. As main outcome, similar trends are observed in this case study and the previous report, i.e., the range of verification errors[4] generally decreases when the number of sites in the aggregation increases. This is basically because, when considering a larger number of sites, when aggregating the FCAS delivery potential over-estimation and under-estimation of the error from individual NMI responses tend to cancel each other, leading to a narrower distribution.

[3] Event 1 refers to a 25 May 2021 event, while event 2 refers to a 25 August 2021 event.

[4] Note that the range of verification errors is not to be confused with the absolute value of average and/or mean of the verification errors. The range of verification errors refer to the distance between the upper tail and the lower tail of the errors.

It can be seen that the verification errors are substantial even for 100ms and 200ms sampling rates, and even when the number of sites increases to 1500 sites. However, these large errors may primarily be a result of the misalignment of the frequency disturbance time and contingency event time. Such misalignment is demonstrated with the example that is illustrated in Figure 4 and Figure 5.

Table 1. Verification errors (%) for event 1, using 100ms sampling rate and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-246965%	597148%	21.99%	-506%	584%	-623%	699%	-810%	922%	-1278%	1351%
25	-57168%	31285%	32.01%	-220%	374%	-306%	444%	-428%	562%	-683%	801%
50	-36379%	22895%	34.95%	-39%	200%	-54%	229%	-87%	281%	-165%	353%
100	-1110%	1637%	35.95%	-4.07%	114%	-6.23%	120%	-8.94%	130%	-14.04%	149%
200	-27.43%	214%	36.07%	8.08%	78.35%	6.99%	81.08%	5.28%	85.60%	3.24%	91.21%
500	4.23%	87.73%	36.14%	18.18%	59.00%	17.22%	60.47%	16.25%	61.98%	14.64%	64.33%
1000	8.39%	67.55%	35.98%	22.95%	51.25%	22.44%	52.32%	21.91%	53.36%	21.11%	54.52%
1500	14.43%	61.98%	36.05%	25.45%	48.49%	25.05%	49.13%	24.31%	49.83%	23.48%	50.95%

Table 2. Verification errors (%) for event 1, using 200ms sampling rate and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-278701%	1014544%	44.61%	-836%	1008%	-1089%	1275%	-1424%	1624%	-2266%	2425%
25	-75466%	118185%	56.19%	-328%	614%	-456%	719%	-643%	928%	-1164%	1274%
50	-60704%	42791%	61.65%	-60.32%	336%	-82.85%	389%	-128%	441%	-251%	581%
100	-2362%	8143%	63.29%	-4.01%	193.79%	-7.66%	207%	-12.11%	227%	-18.15%	255%
200	-37.98%	452.33%	62.92%	16.60%	135.86%	13.96%	141%	10.60%	147%	6.71%	155%
500	2.57%	147.46%	63.37%	32.46%	103.30%	31.31%	105%	29.71%	109%	27.17%	113%
1000	20.48%	115.18%	63.29%	40.78%	89.67%	39.70%	90.84%	38.63%	92.67%	36.61%	95.39%
1500	28.30%	111.22%	63.34%	45.11%	84.39%	44.37%	85.31%	43.45%	86.51%	41.79%	88.48%

From Figure 4 it can be seen that although similar frequency measurements are observed for 50ms and 100ms sampling rates, the frequency disturbance time can vary even substantially when using different sampling rates. In fact, before the contingency happened, the frequency hovered and fluctuated around the NOFB for a few seconds. This creates about 1s misalignment in the frequency disturbance time for 50ms and 100 sampling rates. The frequency measurements for the six-second assessment window for 50ms sampling rate and 100ms sampling rate are shown in Figure 5. The devices only started responding when the local contingency was detected, which is close to 5s in the assessment window for 50ms sampling rate, as shown on the left of Figure 5. Under 50ms sampling rates, the majority of the response may thus be basically "missed". On the other hand, with a 100ms sampling rate, the response may be partially captured. This example also shows how such large errors may really originate from the misalignment of the frequency disturbance and contingency times, rather than from other components of the methodology (e.g., the "RoCoF-based" method or trapezoid rule). Moreover, in this example, the 50ms sampling does not lead to a more accurate result relative to a slower sampling rate, as most of the fast FCAS response would be missed due to the event time identification issue. This is somehow a paradox that again demonstrates that it may be inappropriate to use the current definition of the frequency disturbance time to identify the start point of the six-second assessment window when using local frequency measurement of individual NMI. Instead, a better-defined time of locally detected contingency should be used in order to capture more appropriately the actual device (and aggregated) response after the contingency occurs. For events like the one analysed here, some adjustments would need to be performed in any case to achieve this alignment between assessment window and actual contingency time, even for data captured at 50ms intervals. Moving forward, this process should be automated in some way.

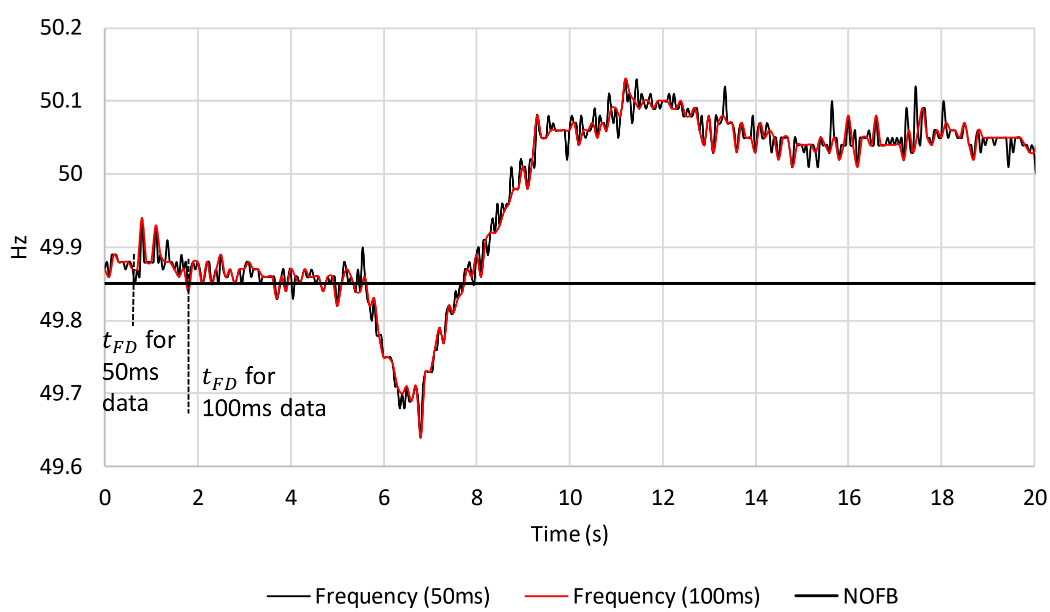
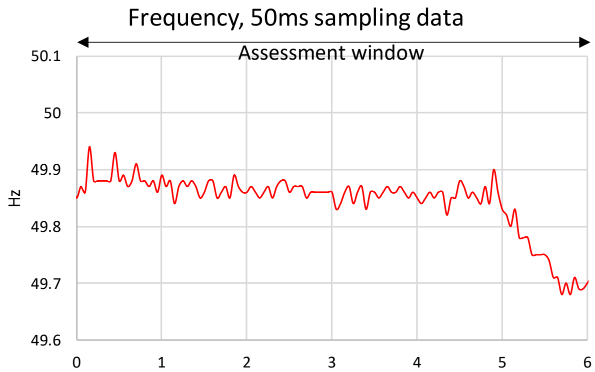


Figure 4. Example of local frequency measurement and identification of the frequency disturbance time of an NMI with 50ms and 100ms sampling rates, for event 1



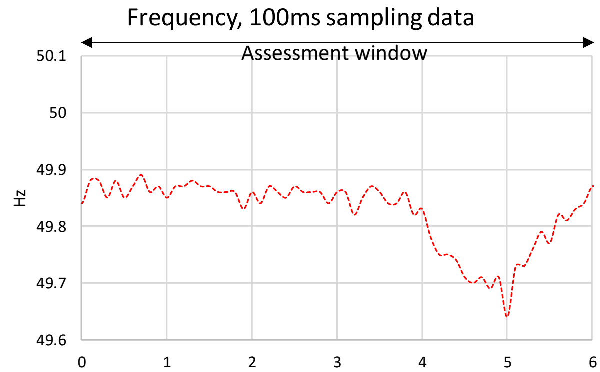


Figure 5. Example of local frequency measurement for the six-second assessment window, with 50ms (left) and 100ms (right) sampling rates, for event 1

The percentage errors for event 2 using 100ms and 200ms sampling rate are illustrated in Table 3 and Table 4, respectively. Similar trends as for the event 1 can be observed, i.e., the range of verification errors decreases when the number of sites increases. However, it can be seen that compared to the event 1 the verification errors decrease more substantially with the number of sites. This is, essentially, because for event 2 the frequency disturbance time is closer to the locally detected contingency, compared with event 1. However, the errors are still very noticeable for the 200ms sampling rate, even when the number of sites increases to 1500 sites. The main reason why such verification errors are observed may be still be attributed the misalignment of the frequency disturbance time and locally detected contingency.

Table 3. Verification errors (%) for event 2, using 100ms sampling rate and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-11092%	1270%	5.70%	-0.89%	21.65%	-1.09%	22.90%	-1.43%	24.97%	-1.96%	28.03%
25	-5.53%	39.01%	6.08%	1.13%	14.32%	0.92%	14.87%	0.62%	15.72%	0.29%	16.59%
50	-0.55%	19.71%	6.19%	2.40%	11.52%	2.28%	11.83%	2.08%	12.19%	1.87%	12.69%
100	1.60%	14.33%	6.26%	3.45%	9.82%	3.35%	10.03%	3.21%	10.25%	3.02%	10.61%
200	2.26%	12.62%	6.33%	4.25%	8.74%	4.14%	8.87%	4.02%	9.08%	3.88%	9.26%
500	3.37%	9.27%	6.34%	5.00%	7.82%	4.94%	7.90%	4.87%	7.97%	4.76%	8.10%
1000	4.59%	8.32%	6.35%	5.37%	7.37%	5.31%	7.42%	5.26%	7.48%	5.19%	7.56%
1500	4.81%	8.03%	6.35%	5.55%	7.18%	5.52%	7.21%	5.47%	7.27%	5.40%	7.34%

Table 4. Verification errors (%) for event 2, using 200ms sampling rate, and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-75255%	1643%	14.46%	2.08%	37.99%	1.56%	39.93%	0.97%	42.12%	0.27%	46.76%
25	0.60%	48.63%	14.67%	6.40%	26.59%	6.09%	27.30%	5.69%	28.16%	5.10%	29.59%
50	3.67%	30.26%	14.76%	8.48%	22.58%	8.22%	22.97%	8.00%	23.41%	7.53%	24.39%
100	6.44%	26.51%	14.82%	10.44%	20.16%	10.26%	20.42%	10.00%	20.74%	9.69%	21.24%
200	8.05%	22.04%	14.83%	11.58%	18.56%	11.43%	18.76%	11.23%	18.97%	10.98%	19.30%
500	10.47%	19.43%	14.86%	12.74%	17.13%	12.63%	17.21%	12.53%	17.34%	12.40%	17.55%
1000	11.87%	17.75%	14.86%	13.37%	16.43%	13.30%	16.49%	13.22%	16.58%	13.14%	16.72%
1500	12.15%	17.54%	14.86%	13.65%	16.16%	13.59%	16.22%	13.51%	16.32%	13.41%	16.41%

An example of the misalignment of frequency disturbance time for 50ms and 100ms sampling rates is shown in Figure 6 and Figure 7. It can be seen that for event 2, the contingency event time is much closer to frequency disturbance time, compared with the event 1 that is illustrated in Figure 4. However, there is still a misalignment of frequency disturbance times when using different sampling rates, as demonstrated in Figure 6. This leads to relatively different six-second assessment windows when using 50ms sampling rate and 100ms sampling rate, as seen in Figure 7.

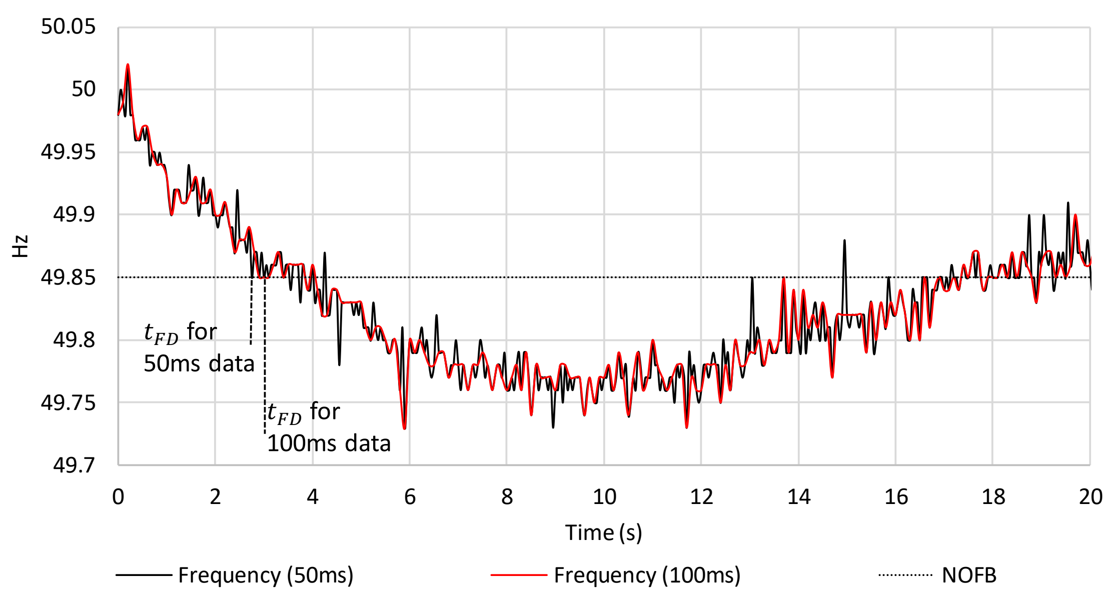


Figure 6. Example of local frequency measurement and identification of the frequency disturbance time of an NMI with 50ms and 100ms sampling rates, for event 2

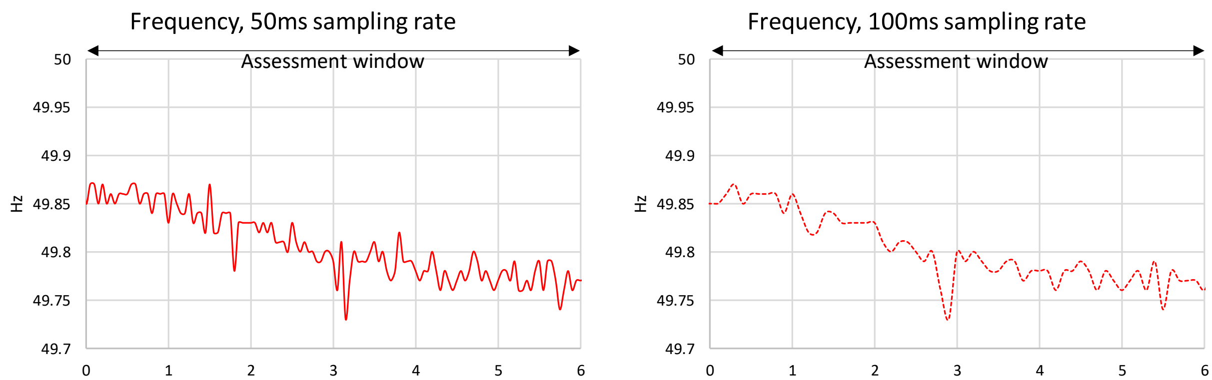


Figure 7. Example of local frequency measurement for the six-second assessment window, with 50ms (left) and 100ms (right) sampling rates, for event 2

3.1.2 Assessment window defined as six-second after the local proxy of the contingency

Instead of using the frequency disturbance time to define the start point of the six-second assessment window, a more accurate proxy for the contingency event time is sought here by using the proposed contingency event identifier methodology with the proposed heuristic rule for the two events that are analysed. The percentage errors for the two events using 100ms and 200ms data are illustrated in Table 5-Table 8. The results for 1s sampling rate are illustrated in the Appendix, Table 11-Table 12.

It can be seen that when using the proposed contingency event identifier methodology to define the start point of the six-second assessment window, the verification errors decrease dramatically. On the other hand, the same trend, i.e., when the number of sites in the aggregation increases the range of verification errors decreases, is still observed for both events. It may be appreciated that when using the local proxy of contingency event time to define the start time of the assessment window, verification errors that are caused by the misalignment of the frequency disturbance time can be largely avoided. This further demonstrates the importance of a well-defined methodology to identify the local proxy of contingency event time, when making use of local data and the importance of implementing a robust methodology in the FCAS verification tool in order to adopt local measurements from individual NMIs.

Table 5. Verification errors (%) for event 1, using 100ms sampling rate and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-17737%	19946%	-0.48%	-25.00%	25.84%	-31.51%	32.03%	-43.23%	41.63%	-66.06%	65.89%
25	-87163%	54340%	-0.43%	-20.72%	21.12%	-25.55%	26.63%	-34.25%	33.69%	-51.82%	49.36%
50	-96567%	25628%	-0.34%	-14.92%	14.24%	-18.30%	17.45%	-24.09%	22.83%	-36.24%	35.71%
100	-712%	12054%	-0.29%	-6.36%	7.38%	-7.82%	8.99%	-10.21%	11.37%	-15.21%	16.00%
200	-4024%	382%	-0.27%	-2.73%	3.07%	-2.97%	3.50%	-3.34%	4.36%	-4.12%	5.80%
500	-27.12%	475%	-0.27%	-1.48%	1.14%	-1.56%	1.25%	-1.67%	1.40%	-1.80%	1.58%
1000	-2.03%	3.61%	-0.27%	-1.08%	0.61%	-1.13%	0.67%	-1.19%	0.73%	-1.26%	0.81%
1500	-1.68%	1.56%	-0.27%	-0.95%	0.43%	-0.98%	0.47%	-1.04%	0.53%	-1.08%	0.60%

Table 6. Verification errors (%) for event 1, using 200ms sampling rate and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-27227%	34579%	-0.49%	-51.11%	59.56%	-62.34%	73.05%	-83.76%	96.69%	-130%	148%
25	-49285%	365840%	-0.40%	-44.44%	42.15%	-55.16%	53.54%	-75.47%	72.34%	-112.53%	110.97%
50	-324536%	23330 %	-0.19%	-30.32%	28.85%	-37.53%	34.60%	-50.26%	44.89%	-77.07%	71.48%
100	-2744%	20612%	-0.10%	-13.24%	15.59%	-16.00%	19.33%	-21.00%	24.40%	-30.85%	34.77%
200	-353%	3540%	-0.09%	-5.47%	6.75%	-6.05%	7.70%	-7.08%	9.24%	-8.86%	12.73%
500	-23.65%	458.49%	-0.07%	-2.66%	2.96%	-2.81%	3.20%	-2.97%	3.46%	-3.30%	3.87%
1000	-4.88%	4.28%	-0.07%	-1.83%	1.79%	-1.92%	1.89%	-2.05%	2.06%	-2.22%	2.24%
1500	-3.51%	4.00%	-0.08%	-1.47%	1.41%	-1.54%	1.49%	-1.63%	1.59%	-1.75%	1.75%

Table 7. Verification errors (%) for event 2, using 100ms sampling rate and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-116%	43.25%	-0.69%	-2.40%	1.11%	-2.50%	1.22%	-2.63%	1.42%	-2.85%	1.70%
25	-3.49%	2.12%	-0.69%	-1.68%	0.36%	-1.74%	0.42%	-1.79%	0.50%	-1.89%	0.64%
50	-2.34%	0.81%	-0.69%	-1.39%	0.05%	-1.43%	0.08%	-1.47%	0.13%	-1.53%	0.22%
100	-1.68%	0.56%	-0.69%	-1.18%	-0.18%	-1.21%	-0.15%	-1.23%	-0.13%	-1.27%	-0.09%
200	-1.34%	0.06%	-0.69%	-1.03%	-0.33%	-1.05%	-0.31%	-1.07%	-0.29%	-1.10%	-0.26%
500	-1.12%	-0.24%	-0.69%	-0.91%	-0.47%	-0.92%	-0.45%	-0.93%	-0.44%	-0.95%	-0.42%
1000	-0.98%	-0.35%	-0.69%	-0.84%	-0.53%	-0.85%	-0.52%	-0.86%	-0.51%	-0.87%	-0.50%
1500	-0.93%	-0.45%	-0.69%	-0.82%	-0.56%	-0.82%	-0.56%	-0.83%	-0.55%	-0.84%	-0.54%

Table 8. Verification errors (%) for event 2, using 200ms sampling rate, and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-124%	114%	-0.89%	-4.69%	2.80%	-4.96%	3.06%	-5.36%	3.41%	-5.88%	3.97%
25	-7.26%	7.46%	-0.89%	-3.17%	1.37%	-3.27%	1.52%	-3.45%	1.69%	-3.73%	1.93%
50	-4.70%	3.00%	-0.89%	-2.47%	0.67%	-2.56%	0.75%	-2.66%	0.88%	-2.79%	1.01%
100	-3.36%	1.43%	-0.88%	-1.98%	0.21%	-2.04%	0.27%	-2.10%	0.34%	-2.19%	0.44%
200	-2.77%	0.87%	-0.89%	-1.67%	-0.11%	-1.70%	-0.08%	-1.74%	-0.03%	-1.81%	0.04%
500	-2.01%	0.17%	-0.88%	-1.37%	-0.39%	-1.39%	-0.36%	-1.42%	-0.34%	-1.46%	-0.30%
1000	-1.62%	-0.23%	-0.89%	-1.22%	-0.54%	-1.24%	-0.53%	-1.26%	-0.50%	-1.29%	-0.47%
1500	-1.46%	-0.33%	-0.88%	-1.16%	-0.60%	-1.18%	-0.59%	-1.20%	-0.57%	-1.22%	-0.55%

4 Conclusions

In the context of the consultations on the amendment of the MASS, further studies on the impact of site aggregation on the fast FCAS verification errors were performed in this report based on a new dataset provided. The main outcomes of the analysis can be summarised as follows.

- Similar trends as in the previous report may be observed [2], i.e., the range of verification errors decreases when increasing the sampling rate and the number of sites in the aggregation.
- From the analysis of the two events analysed, it emerged that the definition of "disturbance time" used in the current methodology may need to be revisited, in order to be better aligned with the actual time of local contingency detection (which is when the device is supposed to start responding). In fact, particularly in one of the two events analysed, the frequency was found to be "hovering" and oscillating around the normal operating frequency band (NOFB) for a few seconds prior to the actual contingency event detection. By applying the current definitions and without a local proxy of the contingency event time, which would require some adjustments even for data captured at 50ms intervals, the relevant frequency measurements might create a misalignment between the frequency disturbance time as currently defined and the actual contingency event time as locally detected.
- In particular, when determining the assessment window in line with the requirements of the MASS, the current definition of the frequency disturbance time without a local proxy to identify the contingency event time may create relatively large errors even for relatively high sampling rates (e.g., 100ms and 200ms). It appears that this may be again essentially due to a misalignment of the defined frequency disturbance time and the time when the contingency event is actually locally detected.
- A local proxy of the contingency event time should be properly defined and implemented in the FCAS verification tool, to better represent the whole VPP's performance in the six-second window after the locally identified contingency event. This is because the local devices making up the VPP aggregate should provide response only when the contingency is *locally* detected, which may be different for devices at different locations. If such a local proxy of the contingency event can be well identified, the verification errors may decrease significantly.
- In this report, we have used a heuristic approach to define such a proxy, which we called "contingency time identifier". However, this should be considered only for the purposes of this work, and more and systematic work is required to develop a robust methodology to identify the local proxy of the contingency event.
- Considering all the datasets analysed and the three studies performed so far, it appears that higher sampling rates and larger number of sites may significantly decrease verification errors. However, it also emerges that the specific numerical impact of these two factors is event-specific and may also depend on the type of devices and their response characteristics. It is therefore advised that the results of this work should be revisited at a later stage as "work in progress", possibly using a wide range of several datasets for different contingency events, technologies, locations, etc. This applies to all levels of aggregation, and is particularly true for VPPs that have a relatively small number of sites (e.g., less than 100 sites), for which more data and studies are required to come up with more robust results.

5 References

[1] P. Mancarella, L. Zhang, and H. Wang, "Fast FCAS Sampling Verification in Support of Market Ancillary Services Specification (MASS) consultation," 2021.

[2] L. Zhang, H. Wang, and P. Mancarella, "Fast FCAS Sampling Verification in Support of Market Ancillary Services Specification (MASS) consultation - Phase 2," 2021.

[3] Australian Energy Market Operator (AEMO), "Market Ancillary Service Specification v6.0," 2020.

[4] Australian Energy Market Operator (AEMO), "FCAS Verification Tool User Guide," 2020.

Appendix – 1s sampling rate studies

Assessment window defined according to the current definition of frequency disturbance time

Table 9. Verification errors (%) for event 1, using 1s sampling rate and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-500025%	717912%	78.38%	-1645%	1839%	-2142%	2321%	-2786%	3180%	-4450%	4926%
25	-221944%	86681%	96.19%	-593%	1023%	-767%	1197%	-1035%	1490%	-1612%	2038%
50	-80670%	61915%	101.93%	-147%	552%	-197%	612%	-275%	747%	-432%	918%
100	-3617%	5101%	105.77%	-21.49%	324.09%	-29.20%	349%	-38.40%	383%	-53.72%	432%
200	-118%	741%	104.88%	16.53%	224.76%	11.10%	235%	6.22%	245%	-1.58%	262%
500	-5.46%	288%	104.64%	49.89%	172.66%	47.15%	177%	44.37%	181%	39.43%	188%
1000	32.28%	192%	105.16%	65.38%	149.53%	63.63%	152%	61.75%	155%	58.53%	160%
1500	41.73%	190%	105.07%	72.87%	140.70%	71.60%	143%	69.98%	146%	67.60%	149%

Table 10. Verification errors (%) for event 2, using 1s sampling rate and using the current frequency disturbance time definition to define the assessment window

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-71948%	3815%	32.51%	10.93%	73.96%	10.15%	78.42%	8.54%	84.63%	7.01%	93.79%
25	6.96%	88.30%	32.59%	18.21%	52.78%	17.57%	53.94%	16.91%	55.41%	15.69%	57.48%
50	11.96%	64.95%	32.48%	22.24%	45.60%	21.80%	46.40%	21.41%	47.45%	20.58%	48.61%
100	18.65%	51.49%	32.45%	24.88%	41.32%	24.49%	41.76%	24.17%	42.34%	23.65%	43.16%
200	22.25%	47.71%	32.42%	27.05%	38.53%	26.80%	38.80%	26.46%	39.30%	26.09%	39.82%
500	25.75%	38.90%	32.38%	28.89%	36.15%	28.72%	36.32%	28.50%	36.54%	28.21%	36.86%
1000	27.78%	38.60%	32.36%	29.91%	34.99%	29.80%	35.13%	29.66%	35.27%	29.45%	35.48%
1500	28.46%	37.61%	32.38%	30.36%	34.54%	30.27%	34.63%	30.14%	34.75%	29.97%	34.93%

Assessment window defined via identification of a local proxy of the contingency time

Table 11. Verification errors (%) for event 1, using 1s sampling rate and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-386612%	1009986%	3.25%	-371%	408%	-475%	507%	-658%	674%	-984%	1017%
25	-111033%	832135%	12.41%	-404%	369%	-517%	467%	-651%	634%	-976%	966%
50	-95334%	2146751%	20.17%	-317%	370%	-411%	460%	-576%	576%	-858%	840%
100	-275278%	50647%	25.53%	-154%	263%	-217%	312%	-324%	421%	-526%	568%
200	-155097%	247470%	27.46%	4.43%	144.09%	3.44%	166.22%	2.14%	196%	-0.76%	260%
500	-15842%	2114%	28.04%	12.06%	68.85%	11.63%	72.61%	10.96%	77.62%	10.10%	86.31%
1000	7.36%	127.27%	28.13%	15.79%	50.27%	15.37%	51.75%	14.89%	53.67%	14.04%	56.39%
1500	10.45%	74.84%	28.03%	17.65%	44.83%	17.22%	46.13%	16.75%	47.40%	16.06%	49.14%

Table 12. Verification errors (%) for event 2, using 1s sampling rate and assessment window defined as six-second window after the local proxy of the contingency

No of sites	Min	Max	Median	2.5/97.5 Percentile		2/98 Percentile		1.5/98.5 Percentile		1/99 Percentile	
				$P_{2.5\%}$	$P_{97.5\%}$	$P_{2\%}$	$P_{98\%}$	$P_{1.5\%}$	$P_{98.5\%}$	$P_{1\%}$	$P_{99\%}$
10	-258%	324%	-1.91%	-15.35%	13.99%	-16.21%	14.83%	-17.34%	16.46%	-18.80%	18.27%
25	-20.57%	28.46%	-1.91%	-10.03%	6.92%	-10.45%	7.48%	-11.02%	8.02%	-11.74%	8.87%
50	-14.80%	10.36%	-1.88%	-7.61%	4.14%	-7.91%	4.46%	-8.22%	4.80%	-8.71%	5.44%
100	-9.93%	7.59%	-1.89%	-5.96%	2.35%	-6.18%	2.56%	-6.41%	2.87%	-6.72%	3.20%
200	-7.49%	4.14%	-1.88%	-4.75%	1.05%	-4.90%	1.20%	-5.08%	1.35%	-5.33%	1.56%
500	-6.50%	1.92%	-1.88%	-3.75%	-0.06%	-3.84%	0.03%	-3.96%	0.14%	-4.12%	0.28%
1000	-4.72%	0.99%	-1.88%	-3.19%	-0.58%	-3.25%	-0.52%	-3.33%	-0.46%	-3.41%	-0.35%
1500	-4.07%	0.08%	-1.88%	-2.93%	-0.83%	-2.99%	-0.78%	-3.05%	-0.71%	-3.14%	-0.62%

Contact: Prof Pierluigi Mancarella, pierluigi.mancarella@unimelb.edu.au