WESTERN POWER DISTRIBUTION

INNOVATION

NEXT GENERATION NETWORKS

Project Entire

OPERATIONAL TRIALS REPORT

Report Title	:	PROJECT ENTIRE OPERATIONAL TRIALS REPORT
Report Status	:	FINAL
Project Reference:		WPD_NIA_017
Date	:	17.05.2019

Document Control		
	Name	Date
Prepared by:	Matt Alderton	25.04.2019
Reviewed by:	Jenny Woodruff	03.05.2019
Approved by:	Matt Watson	17.05.2019

Revision History		
Date	Issue	Status
17.05.2019	1.3	Final

Contents

Glossary

Abbreviation	Term
BaU	Business as Usual
BSP	Bulk Supply Point
CMZ	Constraint Management Zone
CT	Current transformer
DNO	Distribution Network Operator
DSR	Demand Side Response
ESO	Electricity System Operator
FALCON	Flexible Approaches for Low Carbon Optimised Networks
STOR	Short Term Operating Reserve
SYNC	Solar Yield Network Constraints
WPD	Western Power Distribution

Executive Summary

The Entire project aimed to demonstrate the commercial viability of flexibility services for both the Distribution Network Operator (DNO) and the service provider. Building on learning from the FALCON and SYNC projects, the project looked to develop DNO services that could sit alongside existing Electricity System Operator (ESO) services and allow participants to stack revenues, easing access to the service.

The operational trials were limited by the level of participation with only three sites available for the Secure service. Only one site was available for the Restore service. However in total 44 events were logged with interesting learning coming from them. Due to the limited sample size, caution should be taken when interpreting the results. Calls were made to try and replicate potential system needs whilst also testing participant and system performance.

In general data reliability was good with over 97% reliability over the whole trial. Within events this reliability was even better. In addition participant availability declarations were also high, although the uniformity varied by provider. General event performance was good, although this did highlight the minor difference between DNO requirement and the performance driven by the payment mechanics. This highlights the need to over procure flexibility service both for the potential of under delivery but also the potential of no response from a site.

The analysis also highlighted an interesting feedback look within the trial baseline. When highly utilised, the calls in a previous month impacted the baseline in the latter month. This could be improved by altering the baseline to consider previous calls. In addition the challenge around metering positions was also highlighted. Metering as the asset provided clean data and simple baselining, but potentially ignored the effect of other site activities.

Finally the network impact was assessed. The visible impact depended highly on the size and location of the site, as well as its proximity to DNO network monitoring. Where this well matched, response was highly visible. Where not, limited impact could be seen, especially where other dispatchable generation was active on the network.

1 Project Background

Previous trials have shown the technical potential for Demand Side Response (DSR) services to provide value to DNOs. However the roll out of such services has been limited by the commercial complexities of doing so. These are primarily focussed on the challenges customers face when attempting to stack multiple revenue streams. As such project Entire focussed on the development of a simple commercial framework that makes services easily stackable in order to widen the potential market and make the utilisation of DSR as part of Business as Usual (BaU) more viable.

In order to deliver such a framework, the project delivered a wide range of supporting measures such as:

- Network use case development
- Investigations into underlying networks
- Operational system and process development
- Product development
- Stakeholder engagement
- Participant recruitment
- Operational trials

This report focusses on the operational trials that were conducted as part of the project. These aimed to understand the customer behaviours that developed in response to the services. This included what volume was declared available, when it was available, and how well it responded to being triggered. This would then help DNOs understand what volumes would need to be contracted to ensure a reliable service.

Further reports detailing the systems built, the participant recruitment, the results of the operational trials, as well as the project closedown are also available.

1.1 Project Review

The initial project focussed on delivering the simplest customer offering alongside the largest value. As such, alongside the WPD constraint management zone offering (aimed at existing market participants), the intention was to offer a fully managed service, aimed at bringing new entrants into the DSR market. This service would install control equipment in customer premises as well as provide additional value through Triad management and providing Short Term Operating Reserve (STOR) services to National Grid, demonstrating the ability to stack the services. By utilising the customer trust in the Western Power Distribution WPD brand, and the accountability associated with a regulated business, the aim was also to widen the pool of potential participants, boosting the volume available to WPD for the management of network constraints.

However following discussions with Ofgem this element of the trial was de-scoped as Ofgem did not consider the model in which DNO acts as a commercial operator as being in the long term interest of customers. As such the project was redesigned to focus on the delivery of the core WPD network management service. This impacted the services offered, the systems built as well as the customer engagement process. In addition the project was shortened to a single operational season to ensure that any learning was delivered as quickly as possible and at the lowest cost.

2 Services and Process

Flexible Power tested three DSR services to address different operational requirements on the distribution network. These were based on a weekly declaration process with participants declaring capacity by Wednesday at midnight with the DNO accepting or rejecting capacity by 12.00 on the Thursday as shown in Figure 1. This advanced warning would give participants certainty over revenue and allow them to participate in multiple markets.

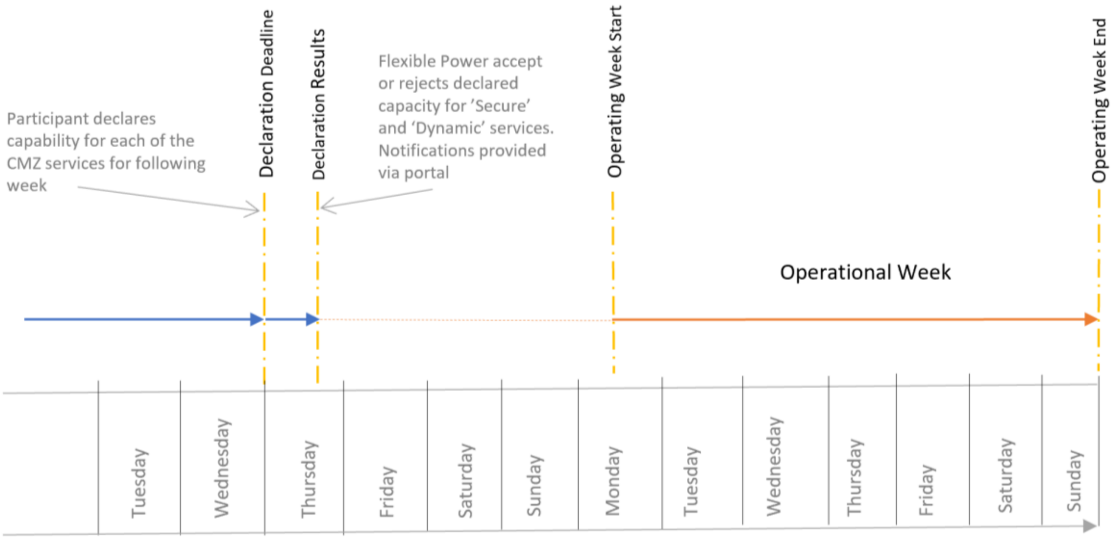


Figure 1: Declarations and Operational Timescale

Within the weekly process, 3 services were designed: Secure, Dynamic and Restore. These were tested in Constraint Management Zones (CMZ) in the East Midlands between April 2018 and March 2019 in the zones are highlighted below (in Figure 2).

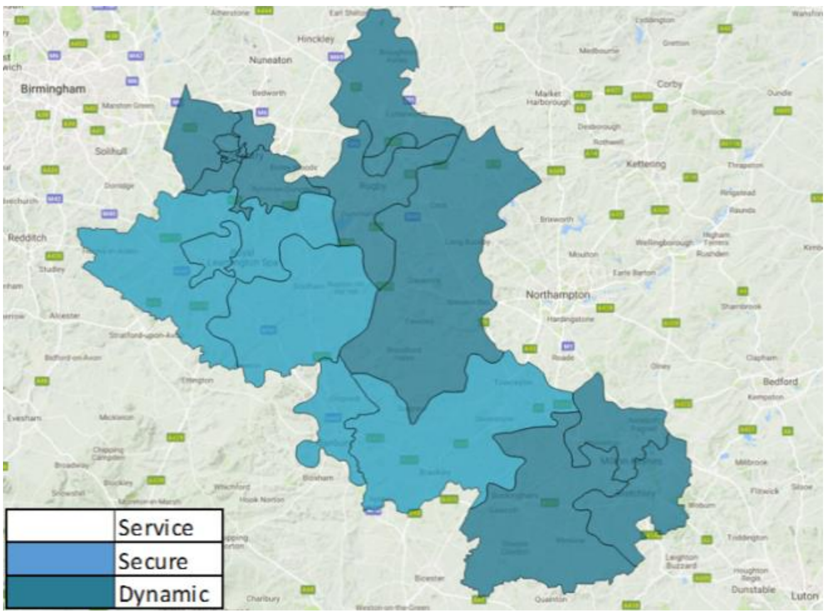


Figure 2: Entire Zones

Secure and Dynamic were designed as main response services to reduce predictable stresses on the distribution network. Each zone operated either a Secure or a Dynamic service. Both services were used to manage known conditions of increased risk on the network. Advance notice was intended to assist participants in assessing their ability to declare capacity whilst minimising conflicts with any other DSR programmes. The Restore service was an additional service and would only be activated in the event of rare faults occurring on the network. The service could help manage an incident and expedite the process of reinstating normal operations. Restore was available across all the zones and was optional. The CMZ services were available to half hourly metered customers in the target area who could increase generation output or reduce demand within 15 minutes of being called and could hold the response for at least 2 hours.

Secure service: The Secure service was used to manage peak demand loading on the network. This service was expected be required on weekday evenings and occur throughout the year due to the seasonal ratings of assets. As the requirements were predictable, Secure requirements were declared each Thursday for the following week (commencing Monday). Payments consisted of an Arming fee which was credited when the service is scheduled and a further utilisation payment awarded on delivery. The week-ahead declarations were scheduled to allow customers to participate in alternative services when not required for the Secure service.

Dynamic service: The Dynamic service has been developed to support the network during maintenance work. This would generally occur during British Summer Time. As the service would be required following a network fault, it consisted of an Availability and Utilisation fee. By accepting an Availability fee, participants were expected to be ready to respond to Utilisation calls within 15 minutes. Dynamic availability windows are declared each Thursday for the following week (commencing Monday). The week-ahead declarations are scheduled to allow customers to participate in alternative services when not required for the Dynamic service.

Restore service:_The Restore service was intended to support the network or help restoration in the occurrence of rare faults. Such events are rare and offer no warning as they depend on failure of equipment. Under such circumstances, response can be used to reduce the stress on the network. As the requirement is inherently unpredictable, Restore was based on a premium ‘utilisation only’ service. This would reward response that aids network restoration, but would pay no arming or availability fees. Participants declared available for the Restore service would be expected to respond to any utilisation calls within 15 minutes and receive an associated utilisation fee.

	Secure	Dynamic	Restore
Advance Payment	Arming	Availability	None
Utilisation	Medium	High	Premium
Customer declaration	Week Ahead	Week Ahead	Week Ahead
FP Accept / Reject	Week Ahead	Week Ahead	Automatic Accept
Dispatch Notice	Week Ahead	15 minutes	15 Minutes
Seasonal Requirement	All	Summer	All
Site Type	Half Hourly Metered	Half Hourly Metered	Half Hourly Metered
Generation	✓	✓	✓
Load Reduction	✓	✓	✓

Table 1: Service Summary

More details are available in the Service Design Report.

3 Trial Design

The trials section of the project was designed to help understand the real world response to the products created within the project. This included a variety of behaviours from the reliability of response to the availability of customers.

From the inception of the project the breadth of the trials, and the potential learning, were always highly dependent on the volume of participation that could be recruited. As highlighted in the Participant Recruitment report, the volume made operational was limited. This has in turn impacted the breadth of trialling available.

In total three sites we made live, all providing the Secure service. Two sites were metered at a site level (meaning other generation and demand were present behind the metering point) and one purely on the generator. One site was made available for the restore service. None were available for the dynamic service.

Calls were made to sites to try and test a variety of different running conditions. These tests were developed to be broadly in line with potential DNO requirements. The calls were also used to test system performance. Data was collected as part of the metering and settlement process and analysis was conducted once the operations phase of the trial had concluded.

4 Trial Observations

Despite the limited sample size, a number of observations could be made from the calls activated. Care should however be taken drawing conclusions due to the sample size. In total there were 43 Secure events and one Restore event. As such the bulk of the following analysis (Sections 4.1 – 4.3) refers solely to the Secure events, with the Restore event covered in Section 4.4. Two Secure events produced no valid data due to incorrect set up at the beginning of the trial period and therefore have not been included in analysis.

Care has also been taken to ensure the anonymity of any data presented.

4.1 System Usage

A basic metric of system reliability is the availability of metering data.

Where missing data can broadly be split into two categories: system metering gaps, where no data was recorded by the Collar system, or customer metering gaps, where zero data points were recorded by the system. In this case there is an issue on the customer side of the metering arrangements.

System Data Gaps

38 events (88%) had no gaps in the minute by minute reporting. Four events had one gap in the data, and only one event had more than a single missing data measurement. This event was the first of the project and was attributed to teething issues within the system. This shows a good level of basic system availability.

Analysis of wider data beyond the event times also showed good reliability with over 97% reliability. Missing data points were concentrated at the start of participant operations and at specific times of issues.

Customer Data Issues

Only one event had a period that suggests that the data reported was faulty, with the reported output dropping from above target, to zero for 5 minutes before returning to above target. This is 5 data points out of 3934 recorded across all event data (0.15%) gathered. This again highlights the good quality of customer data supplied.

4.2 Availability Declarations and Variability

To be called in the services, participants had to make themselves available through the customer portal (see Figure 3). This allowed them to control when there assets were available to be called (via the calendar) as well as the expected volume to be delivered as well as maximum and minimum response times.

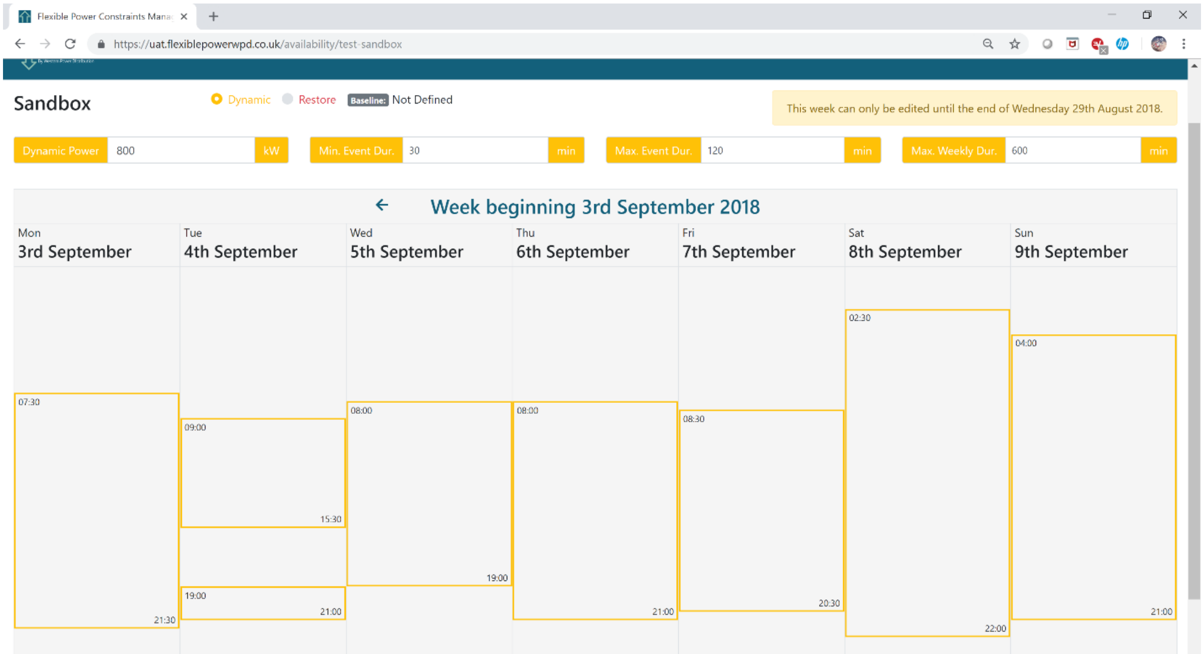


Figure 3: Image of Declaration Calendar

Availability over the trial period was varied between participants. Some participants regularly declared availability for all days of the week for the same period, only changing times declared a couple of times over the trial. Other participants' declarations varied far more between days and weeks, with little predictability to their availability. As the trial period went on there was something of a trend towards sites declaring availability for entire days, instead of just for a period of a day. The variable sites tended to be those monitored on a site level which generated for other purposes.

All sites consistently declared a minimum duration time of 30mins, which was the event length part considered as part of the Secure service. Maximum response times declared varied between sites and days.

Ignoring zero declarations, over the Entire trial period participants made themselves available on 42%, 86% and 100% of the days within the trial (taking into account 5 day working weeks where applicable), on average 74%.

4.3 Secure Event Analysis

There are multiple factors around an event that can be measured or assessed to provide a view on the success. Appendix 1 contains a list of the events run, outcome and assessment criteria and general comments. Different approaches are taken in this section to analyse the results of the Secure events conducted.

Percentages of targets are measured from the baseline to a target, so 0% of target would be a site outputting at baseline level. For the trial a simple baseline was developed as the average of output between 3-8PM for the first three weeks of the previous month. This represented a reasonable option to balance simplicity and establishing a methodology that would be inclusive for all site types. The target was the sum of the baseline and the available volume declared by the participant in the weekly availability declaration. To aid analysis the output for events was graphed, both over the event, and over the whole day of the event. Figures 4-7 highlight two successful events.

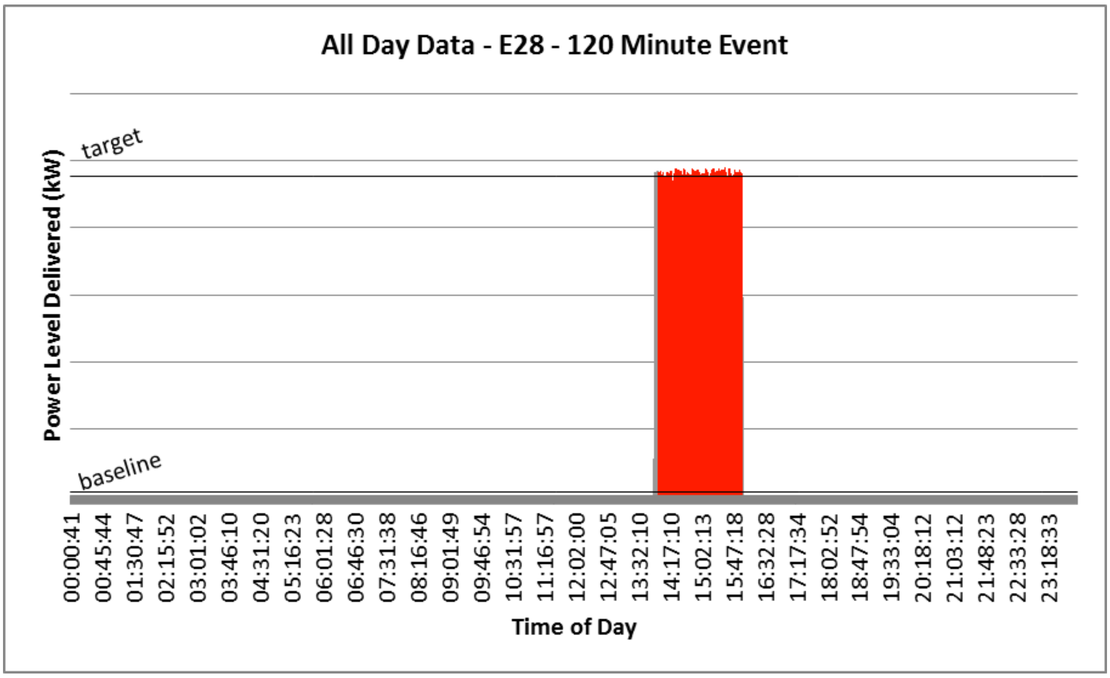


Figure 4: Example Response All Day Data

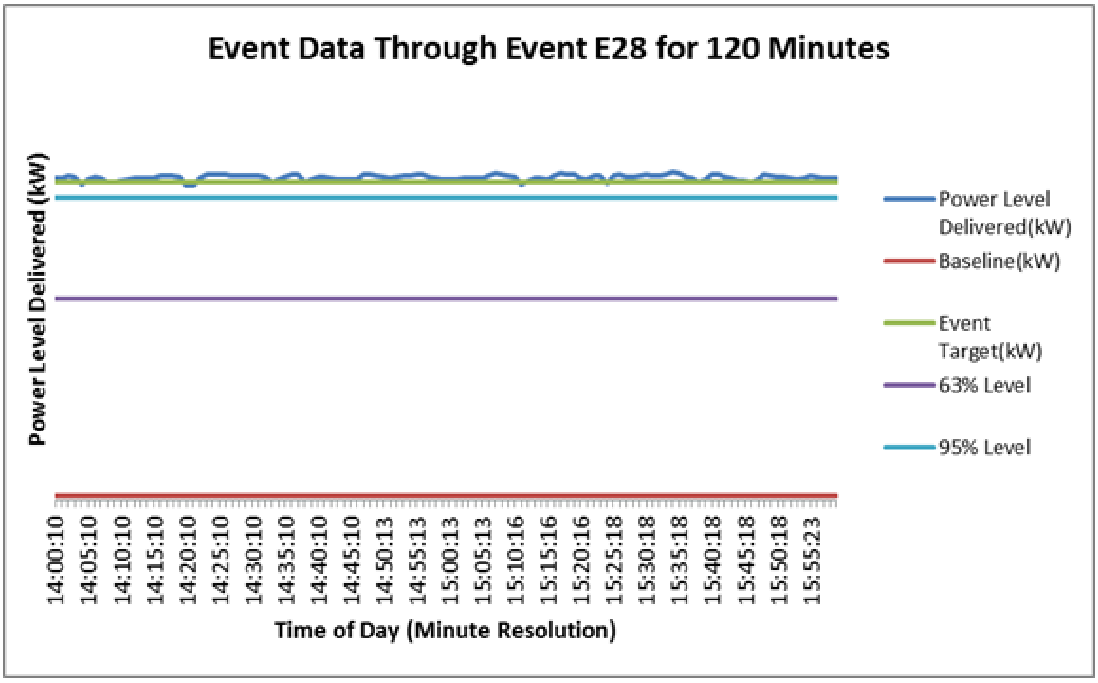


Figure 5: Example Response Event Data

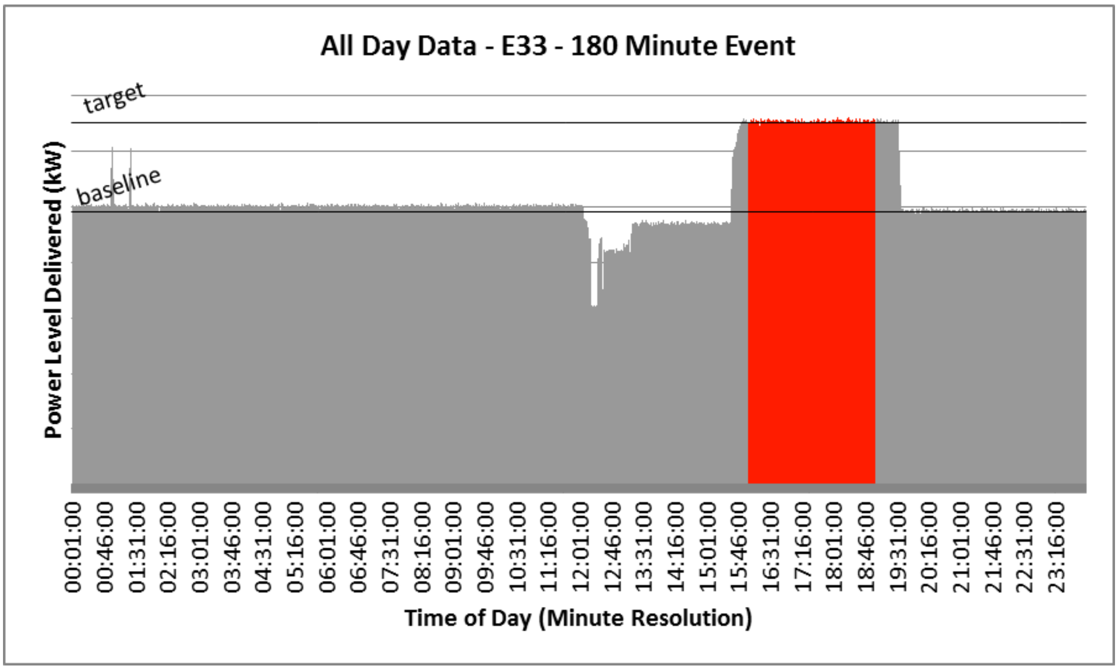


Figure 6: Example Response All Day Data

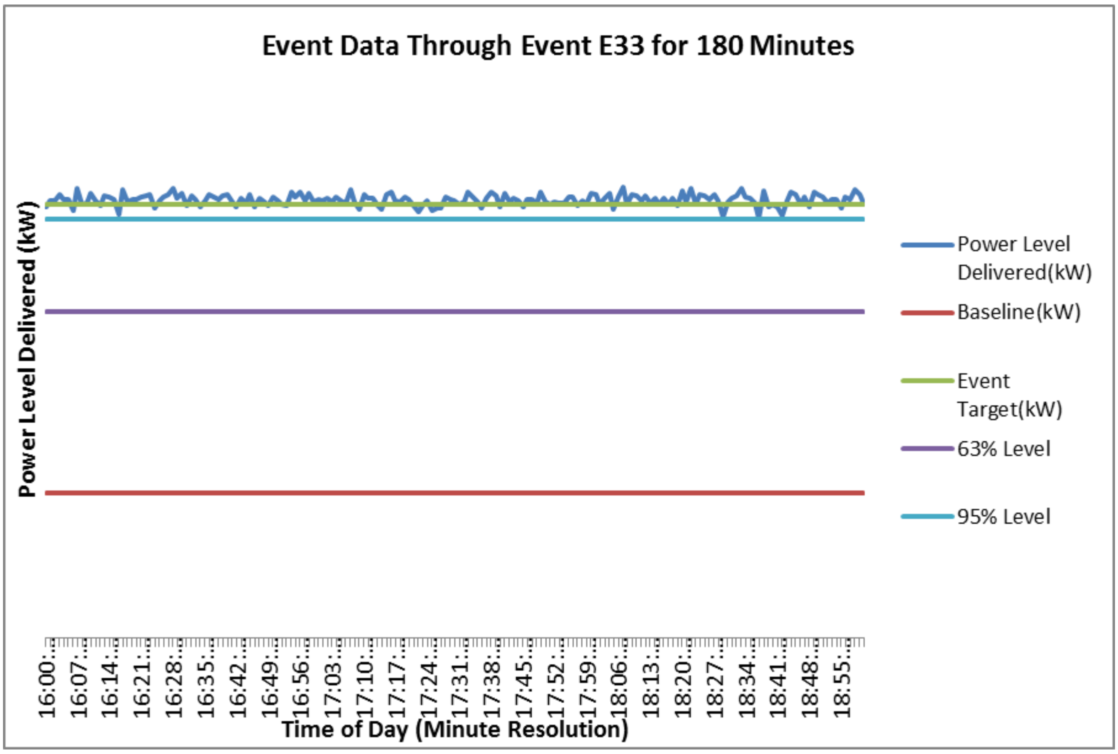


Figure 7: Example Response Event Data

Out of 43 Secure events called, 41 are included in the analysis of call reliability. Two events (E7 and E8) did not produce data due to technical issues as mentioned in section 4.

4.3.1 Service Reliability

An obvious starting point for judging the success of an event is if it provided the service desired. For a DNO, a fully successful event would be one where the requested output was met throughout the event. However dips to below target but above 95% of the target output were deemed acceptable in the commercial contracts. A further categorisation cut off was made at 63% as this was the floor of the sliding scale payment mechanism. Results are shown in Table 2 and Figure 8.

	No. Of events	Percentage of events
Events Continuously Above Target (100%)	9	22%
Events Continuously Above 95%	9	22%
Event Continuously Above 63%	6	15%
Not Continuously above 63%	17	41%

Table 2: Range of Continuous Output Achieved

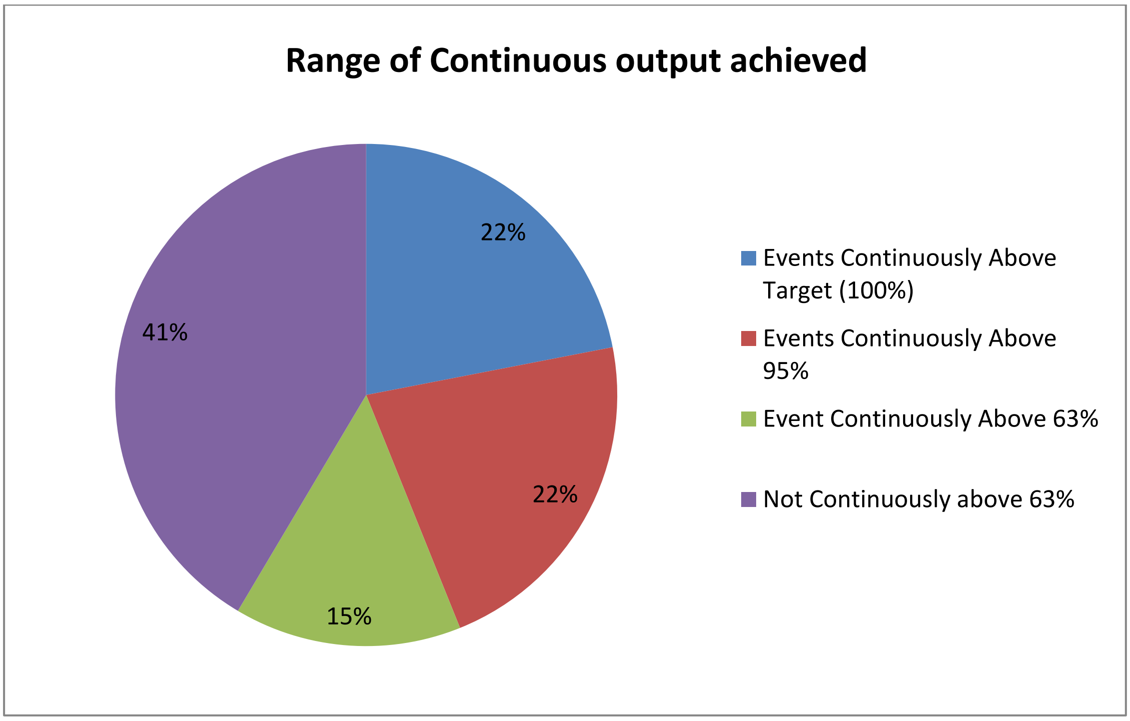


Figure 8: Range of Continuous Output Achieved

Given the commercial contract, an event could be considered successful if output was continuously above 95% of the target, and otherwise was not. In this view 18 of the 41 events were successful, (44%). However this provides a rather uncompromising view and does not provide a full picture of what happened over the events. Several events had brief dips, or were slightly slow to ramp up to output, or tailed off briefly at the end of the event window. Most spent the vast majority of the event window (>95%) above the target output with a small number of issues. As such a scoring metric was developed, giving each event a score out of 100 to try and provide a fairer assessment of the success of an event.

4.3.2 Scoring Metric

Each data point was assigned a score, and the scores for an event were summed, and then divided by the number of data points in an event to provide a normalised score. Data points above the target (100%+) received a score of 1, between 95% and 100% received a score of 0.8, between 63% and 95% a score of 0.1 and below 63% a score of zero. Each events score is shown in Appendix A. Scores have been categorised, with a score of 100 "Excellent", above 90 "Good", above 85 "Below Expectations", below 85 "Poor" and scores of zero "Failed". Results are summarised in Table 3 and Figure 9.

Event Assessment	No. of events
Excellent	9
Good	15
Below Expectations	3
Poor	9
Failed Response	5
Total	41

Table 3: Range of Event Assessments

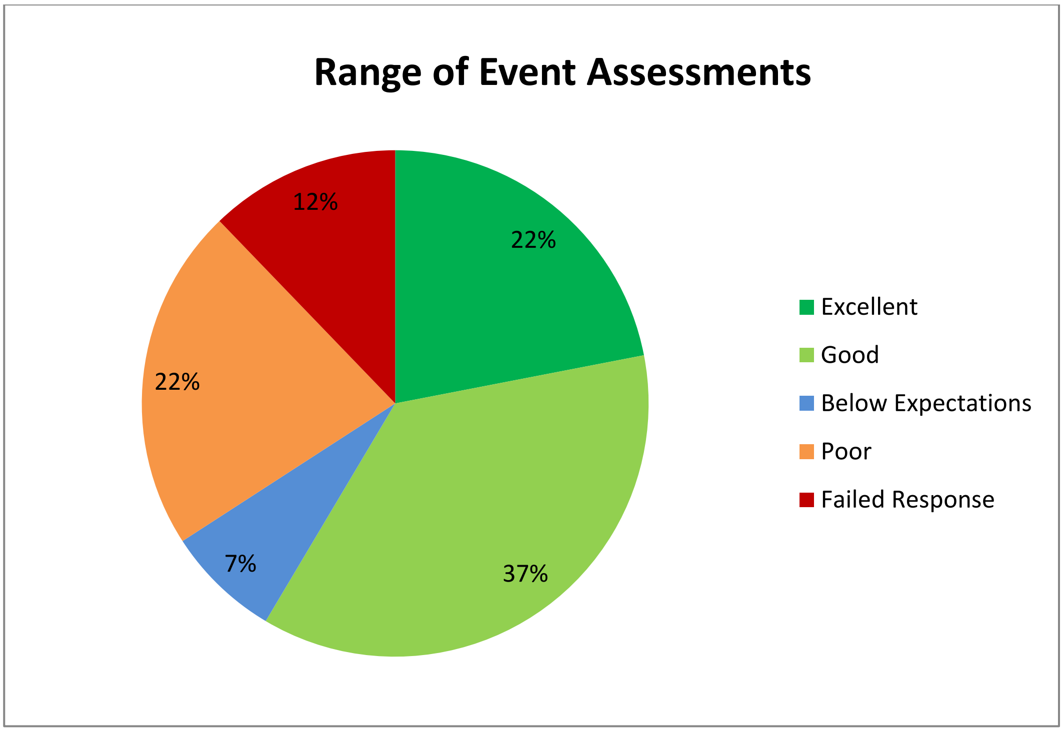


Figure 9: Range of Event Assessments

It is felt this scoring and categorisation mechanism provides a fairer assessment of events in general. The outcome shows a significant proportion of events to have not met expectations. Figure 10 compares scores given with payments received for events (expressed as a percentage of the total potential pay out from an event). There is a slight mismatch between payments in the 100% category, and scores in the 90-99 category the payment mechanism paying in full down to 95% output. This does highlight the slight mismatch between the DNO requirement and the payment mechanism.

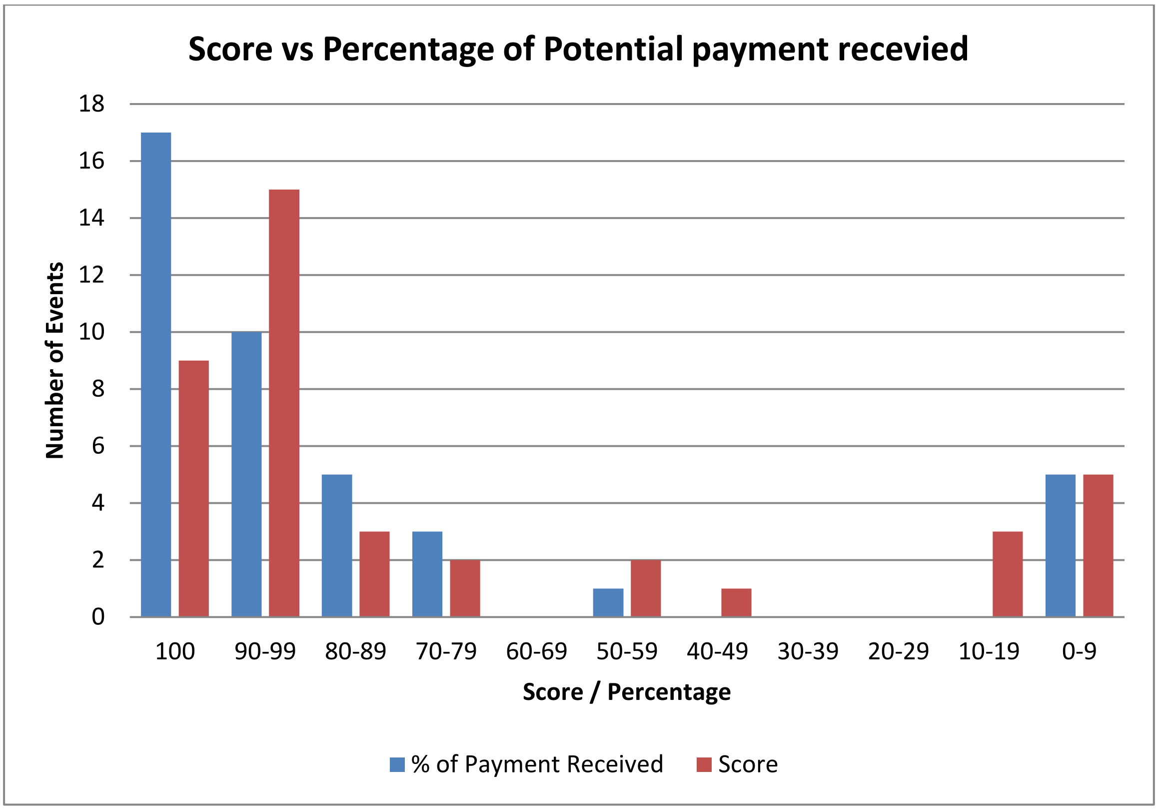


Figure 10: Score vs Percentage of Potential Payment Received

4.3.3 Features of Events

Comments were made on of noteworthy characteristics of events and were categorised in Table 4 to form discussion points to analyse common issues in events and their causes.

No response	Slow ramp /Early drop	Under Response	Mid-event dip(s)
5	7	4	5

Table 4: Event Features

No response

In five events (12%) the participant appears to have failed to respond completely, sitting near or below the baseline with no noticeable response. An Example is shown in Figure 11. No-response events occurred on two of the three participant sites, which were of similar technology types. One of these no response events was due to a technical issue with the generator site causing them to trip just prior to the event. The risk of a no response can be considered a separate risk to that of poor performance in an event, and suggests the need to not just procure services above the level expected necessary, but from multiple suppliers to guard against no responses.

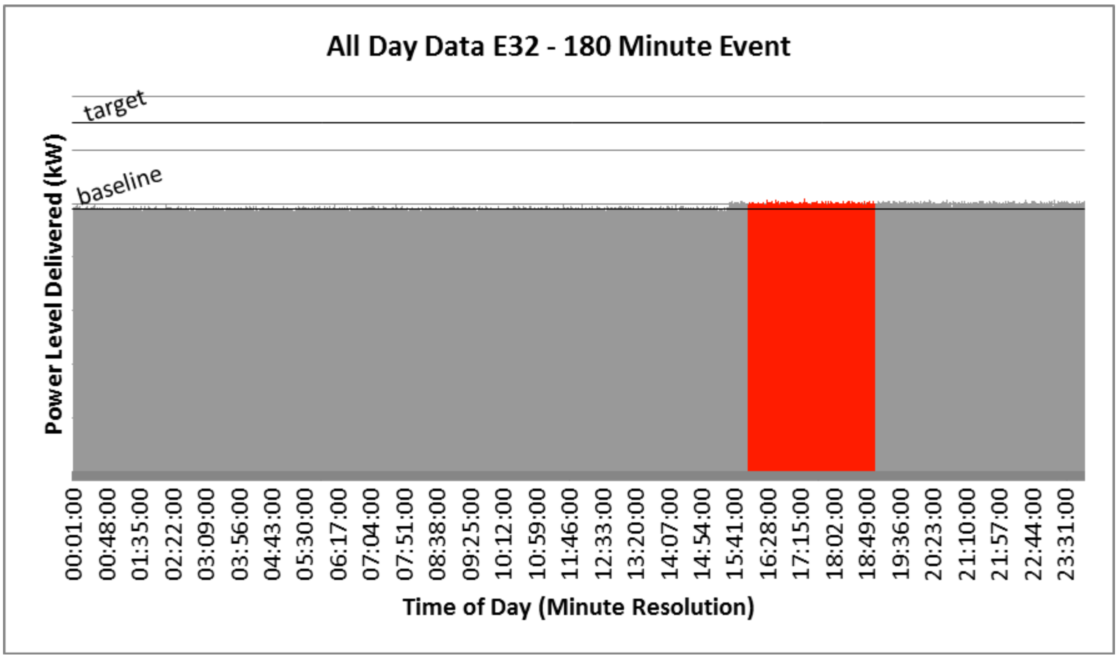


Figure 11: Example of No Response Event

Slow Ramp Up/Early Drop Off

In seven events (17%) the participant sites output was slow to ramp up to the target output, or dropped off from target output slightly ahead of the event window ending. An example is shown in Figure 12. In four events generators were slow to ramp up to expected output, and in three events generator output dropped off before the end of the event window. This suggests that when procuring services in may be prudent to extend event windows by a short period to guard against late ramp up or early drop off cutting into a critical period of time.

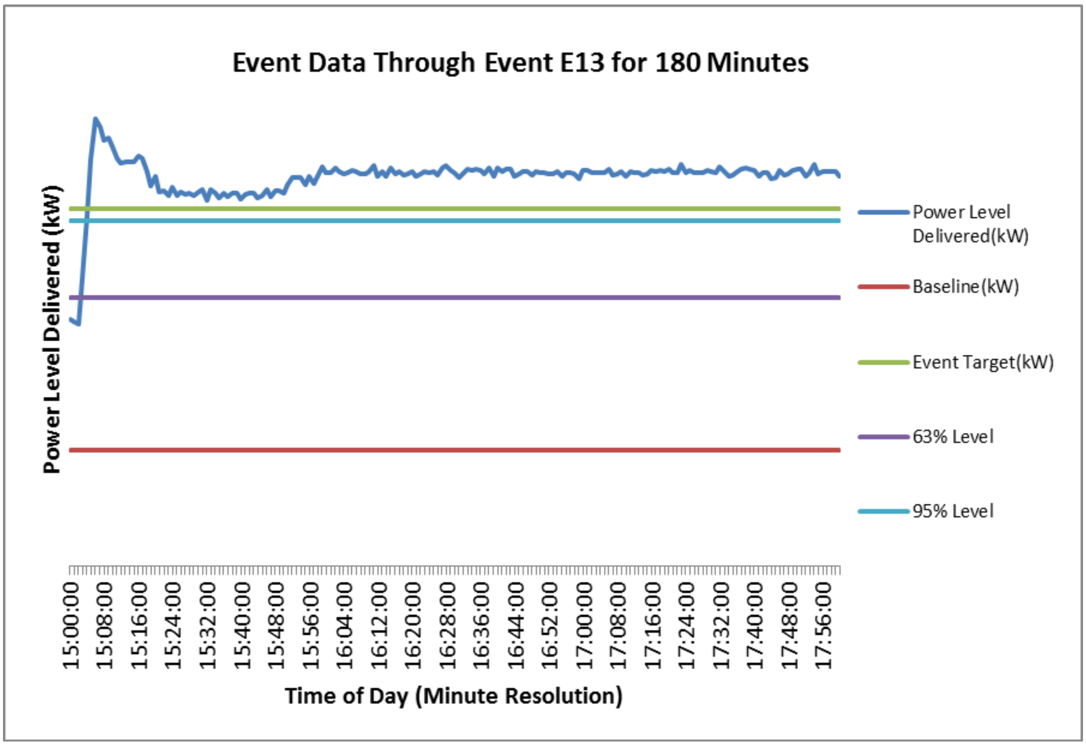


Figure 12: Example of a Slow Ramp up Event

Under Response

In four events the participant site under responded, meaning that there was a noticeable increase in output for the event window, but that output was below target. Of these four events three have been considered due an issue with how the baseline was calculated and is covered in section 4.5. It is unsure why the other event under responded, however the output of the site was below baseline before and after the event window by around the same magnitude as the output was below target within the event window. This suggests that while the generation scheduled to provide flexibility worked as planned, within the site generation was lower or demand higher than expected, resulting in under performance. This sort of event provides an interesting case study to be looked at when considering where it is best to meter for flexibility services. Both the baseline issues highlighted and meter placement are talked about further in section 4.5

Mid- event dips

Mid-event dips were classified as where a sites output during an event dropped significantly in a defined occurrence (as per Figure 13). This happened for 5 events (12%). The dips varied significantly in length and severity with no identifiable common features. 4 out of 5 dips occurred for events monitored at a site level, and 1 at a generator level.

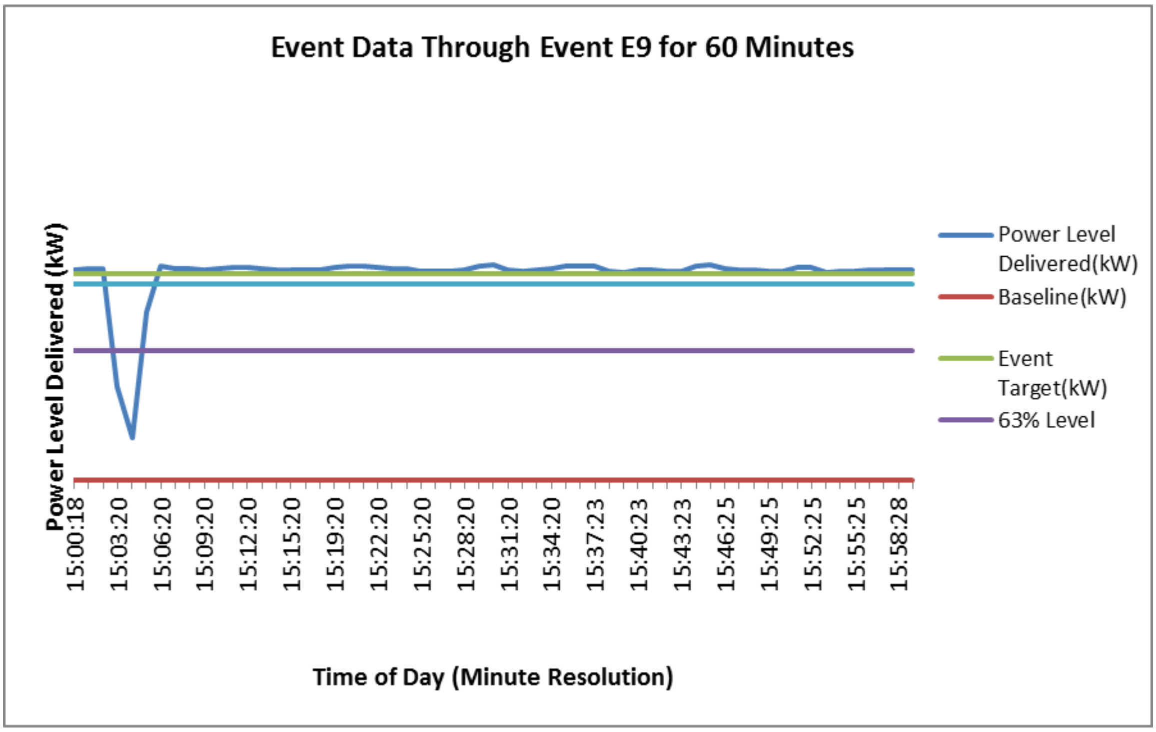


Figure 13: Example of Mid Event Dip

4.3.4 Overall Data Point Analysis

While looking at events provides useful learning and assessment, for more statistical analysis it is perhaps also useful. Data from no response events was felt to distort the data and distract from analysis the quality of service provided when participants did respond. The risk of no response events is clearly visible from the overall event statistics, whereas analysing the data points provides a better view of the average level of service that might be expected when a response occurs.

Figures 14 and 15 show the overall categorisation of data points, and a cropped view of the distribution (the full graph is available in Appendix B)

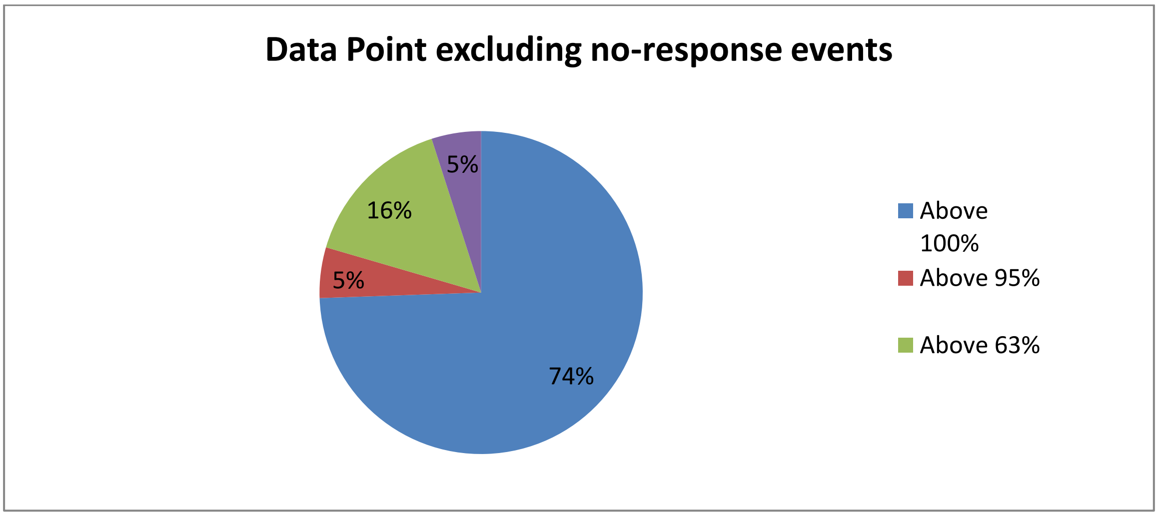


Figure 14: Data Point Excluding No-Response Events

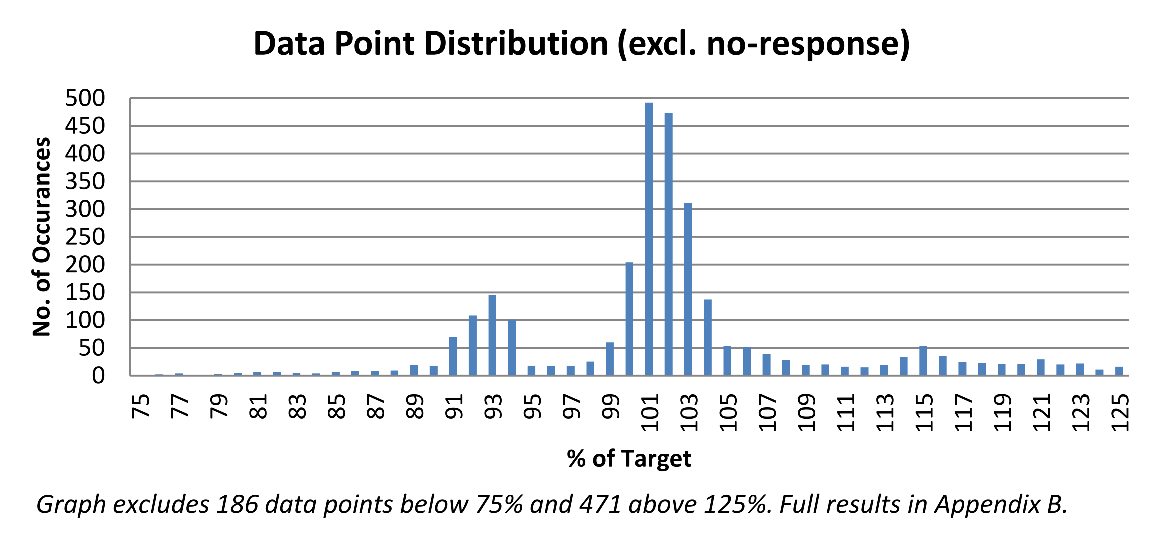


Figure 15: Data Point Distribution Excluding No-Response Events

Overall when considering the events where participants responded, the results are promising. The vast majority of data points were above target, and show a roughly normal distribution peaking just above 100%. The rise around 93% is due to the baseline distortion commented on in the next section. If these are omitted then the small peak does not occur. Using a more probabilistic approach to the chance of suppliers output under/over supplying at any given moment in time seems a likely approach to determining over procurement. The chances of one generator under-supplying at a given moment in time will be greater than the cumulative chance of multiple generators all undersupplying at the same time. Other risks also need to be assessed, as larger generator site may be more reliable and therefore a lower risk option in other respects.

Excluding data from no-response events gives a clearer picture of the level of over procurement that may be necessary, and risk factor of it. The data point distribution graphs provides the beginning of the data needed to develop probabilistic risk analysis approach to the quantity of flexibility that should be procured to reasonably guarantee results.

4.4 Restore Event

The one restore event trailed was completely successful, with the participant responding to the signal sent from WPD more quickly than required and sustaining output above 97% of target for the whole 46 minute event. This proves the technical ability of the systems to call on Restore services via PowerOn, which is a crucial step in interfacing flexibility services into the control room.

4.5 Measuring Flexibility

How much flexibility a site provides can be affected by two main factors, where the metering is located, and how the baseline is calculated.

4.5.1 Baselining Distortion

The baseline in particular has proved to be a problematic concept when designing flexibility services, as it is of key importance to both the defining how much flexibility is being offered, and how much a site operator can earn. This trial has highlighted some areas of concern, and case studies to consider when further developing baselining methods.

The baseline calculation was based on an average of hourly average output/demand for the period between 3pm and 8pm on weekdays for the first 3 weeks of the month. This then determined a baseline to which sites declared available capacity was added to give the target output value. An issue highlighted by this trial, is that WPD's utilisation of the asset affected the asset's baseline for the following month. One generator site submitted a consistent level of available capacity, and performed consistently in output across the 5 months that services were procured for from that site. Figure 16 shows the average output of the site in a month against its target for the month. The baseline started at zero. In Month 1, services were not successfully called in the first three weeks. In Month 2, services were called twice, in Month 3 once, and in Month 4 seven times. As can be seen on in Figure 16, the more frequently a site was called in a month, the higher the target was for the next month. The generator declared the same available capacity as the previous four months, but when called did not increase output over previous months and so was penalised for under-fulfilling despite supplying the same level of actual response in previous months. It can be clearly seen that including generation procured as part of DSR services in calculating the baseline can lead to a feedback loop, increasing the baseline and therefore the target.

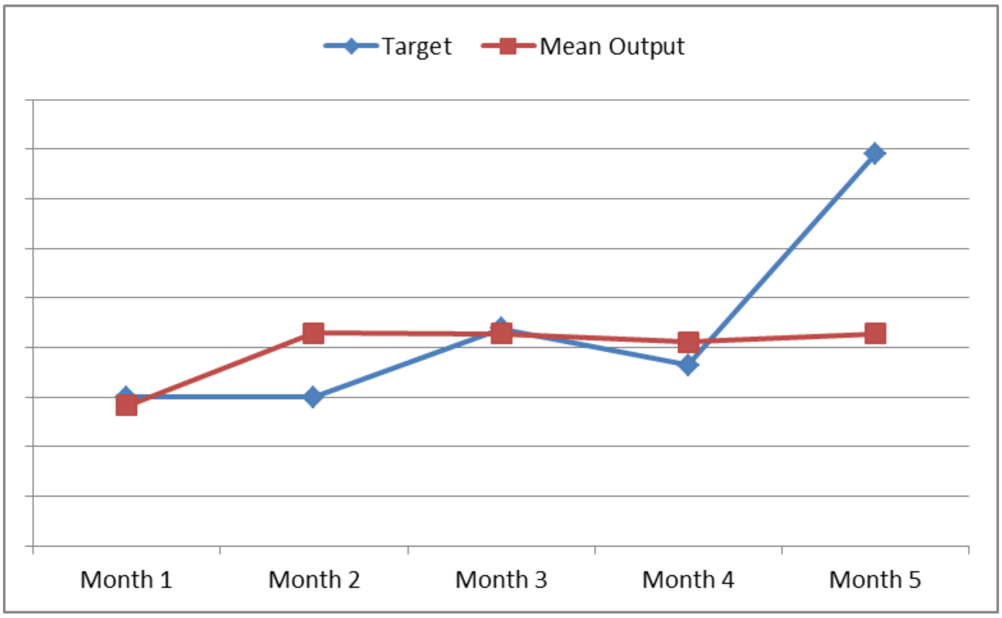


Figure 16: Change in Target Over Time

In line with the arrangements used for this trial, the generator should have reduced their declared availability by the baseline figure, to ensure they were not overcommitted. However this also would have impacted their earnings and reduced their revenue for supplying the same service as in the previous months. Increases in baseline could also distort WPD's view on how much DSR is needed to prevent overloading assets, as it can be assumed that the baseline generation will happen without procuring DSR. This issue could be easily accounted for through a modified baseline. For example all WPD calls could be subtracted from the monthly metering data that feeds into the baseline calculation.

4.5.2 Site or Generator Metering

Another key theme that rose from the trial is the location of the metering point and its associated impact on the output.

There are advantages and disadvantages of metering for flexibility services on a site level, or on an individual generator. In short, measuring on individual generators makes it simpler to provide an obviously fair baseline and observe the action of the generator. However it also requires greater expense to install metering on all generators on a site that may have several, and the sites overall demand could increase and mask the effect of the generator from actually providing the reduction in load for the DNO paying for it.

Figure 17 shows one example from an event monitored at a site level. For the event a generator was obviously set running, and the site output increased by roughly the quantity procured. However, presumably due to other load factors, generation of the site in general was below baseline for most of the day prior to the event, and the output during the event was below target by a similar amount. This event can be viewed through two lenses:

- That the generator provided the desired increase in generation above the level of output at which they had been at all day and therefore in spirit fulfilled their obligation
- That the generator should have adjusted their declared capacity and are shifting less loading off of the critical assets they are being procured to protect

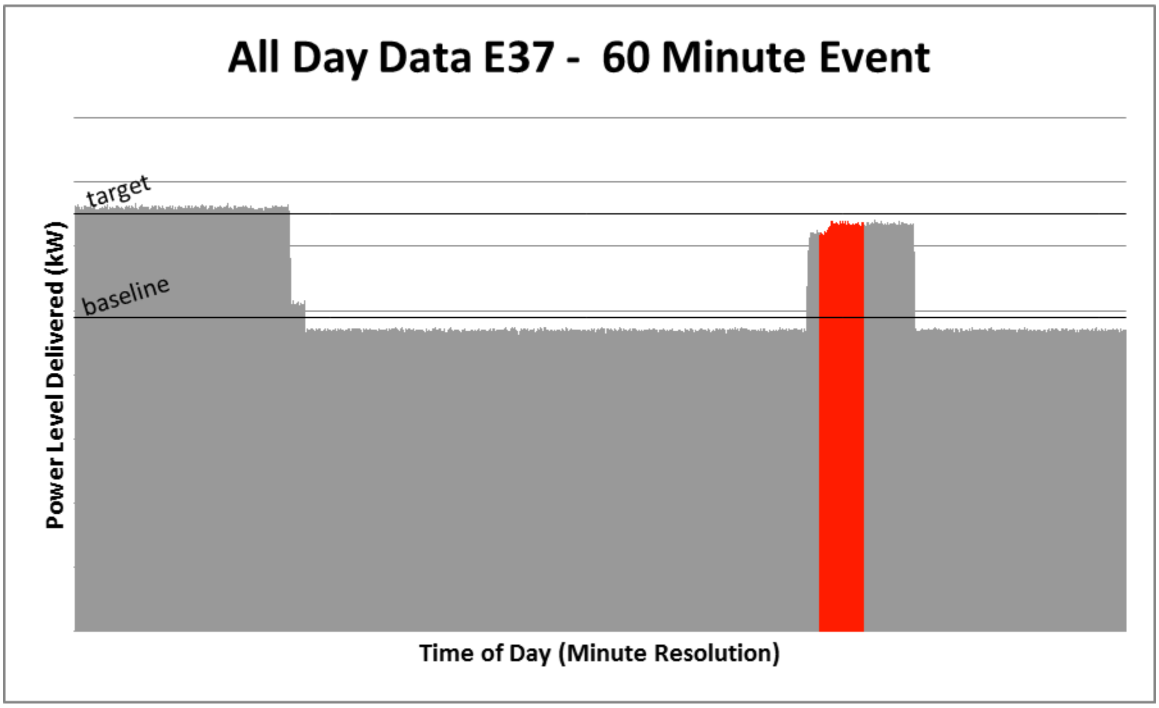


Figure 17: Underperforming Event

This demonstrates the value of monitoring at a site level from the DNO perspective, as if just a single generator had been monitored this would been seen as a completely successful event, when in fact its impact the constrained asset may have been less than expected.

It should be noted that WPD aligned its metering requirements with those from a flexible STOR contract. This gives participants the option to meter at either location; however baselining is a less challenging problem due to the nature of the service. Coordination of metering requirements would provide significant benefits to both participants and service buyers.

4.6 Network Impact

4.6.1 Visibility on SCADA Data

Figure 18 compares two consecutive days, one where the generator operated for an event period and the other where it did not, and the effect of this in decreasing peak loading on transformers at the primary substation.

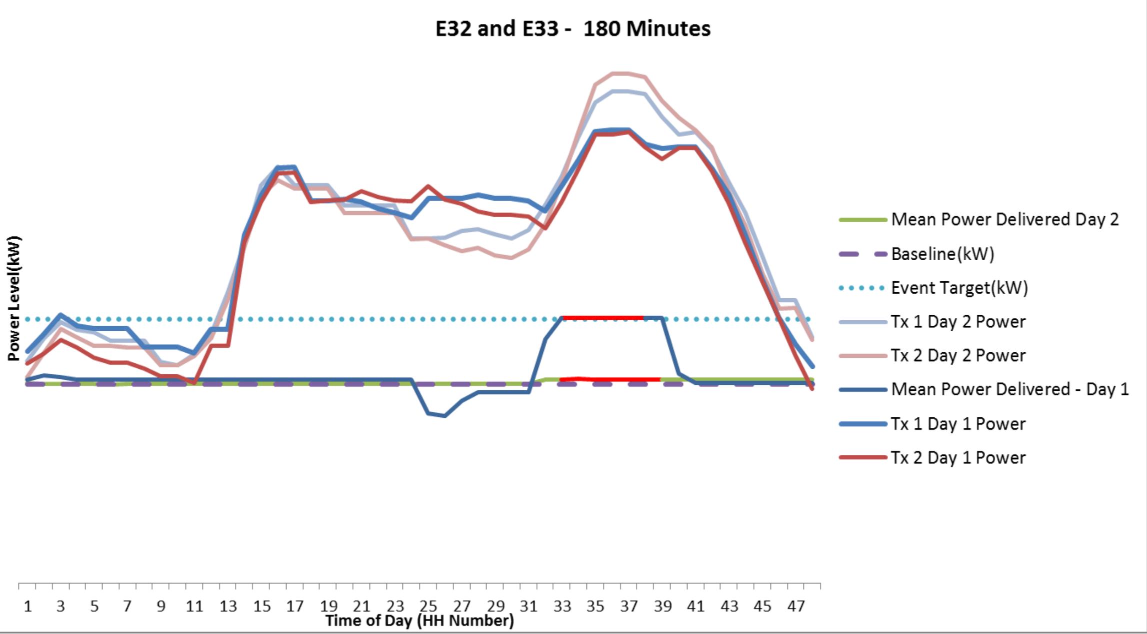


Figure 18: Call Impact on the Wider Network

However the ability to view this depends on sufficient monitoring being in place on the network. In particular one participant's connection was on a circuit to a primary substation that also had a second circuit feeding it. Therefore the drop in current flow should have been split across both feeders as they both fed load there, though more would be expected on the feeder of the connection (circuit 2 on Figure 19). However this is not readily apparent from the data recorded, in fact current on the feeder seems to rise when the generator picks up. The resolution on the Current Transformers (CTs) also hampers gaining an accurate picture, shown by the current flattening at the minimum level the CT can measure at and then dropping to zero repeatedly. This makes drawing any meaningful conclusions from the data on the effect of the flexibility service on the network very difficult.

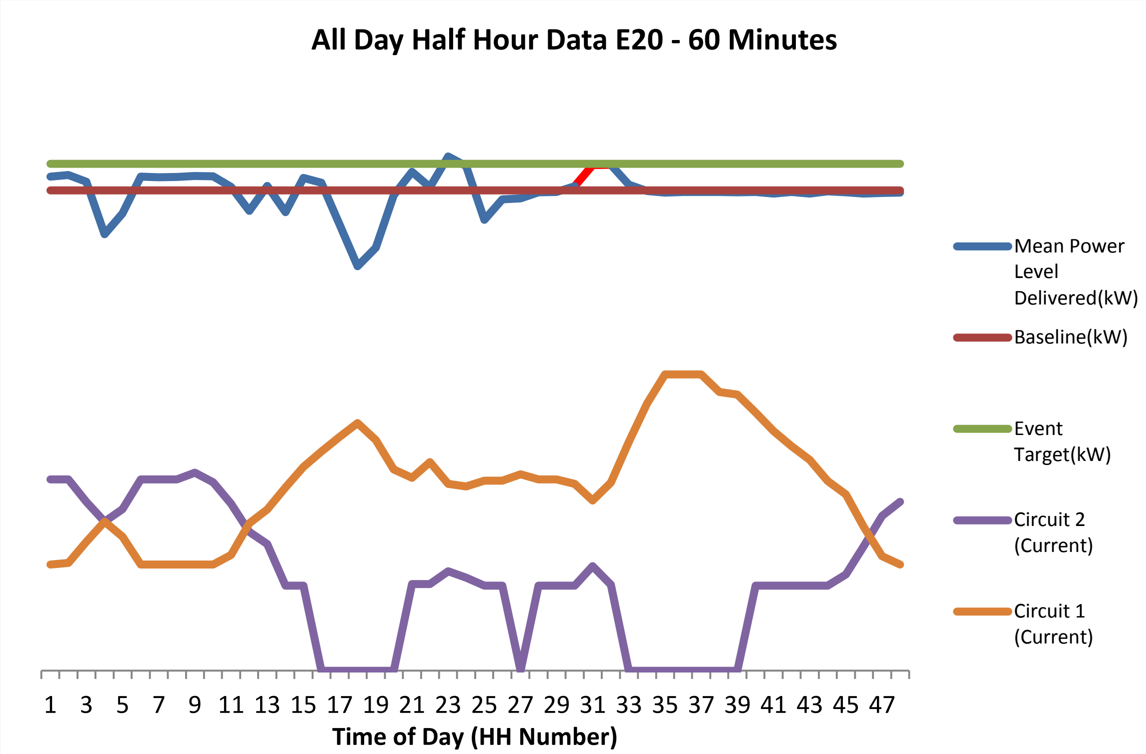


Figure 19: Impact of CT Resolution of Call Visibility

4.6.2 Impact of Other Generation and Network Architecture

When looking for the effects of the services procured in this trial on the modelled constraints (Grid transformers at Busk Supply Points (BSP)) it became apparent that the effect could be difficult to monitor. This was particularly true given the relatively small quantities of services procured. In some cases it was difficult to entirely view the effect of the service procured due to other generators, and large loads on the feeder drowning out any clearly noticeable effect. This was particularly the case on one of the BSP's modelled, which had a large ramping generator connected at the busbars. The effects of this generator on transformer loading were far more significant than that of the services procured in the trial, and frequently coincided with them. Even on a feeder level, other generators or aspects of the network can make it unclear or difficult to correlate what is recorded with what is procured. Figure 20 shows an event, and the power flow on the feeder that site is connected too. While there is a noticeable dip when the generator ramps up during the event, there is a similar dip soon after. This adds uncertainty to attributing readings seen to procured services as other connections on the local network may have a far greater influence than the services procured.

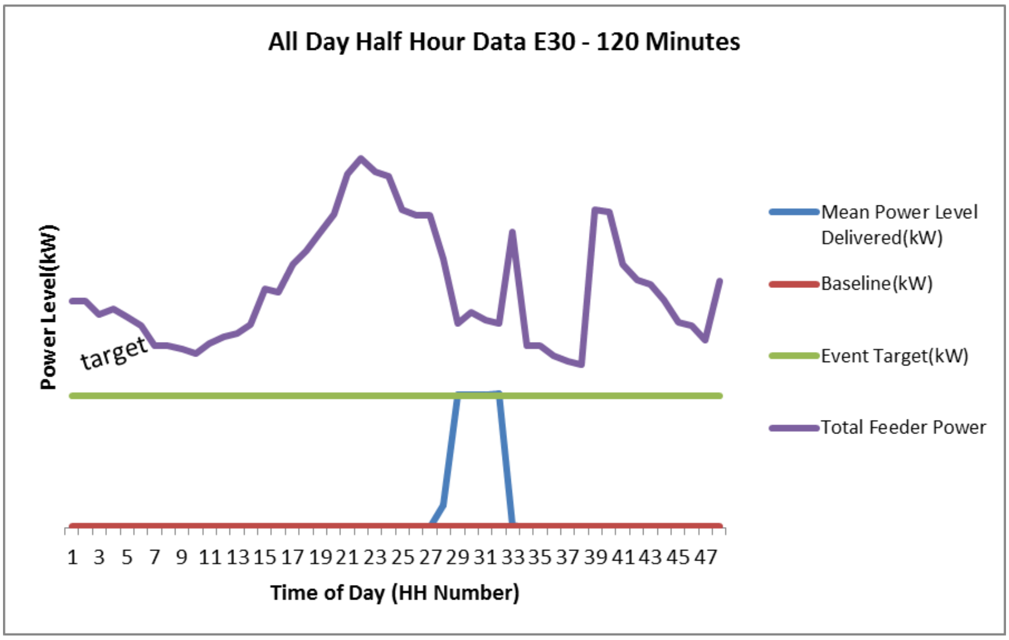


Figure 20: Impact of Other Generation on the Wider Network

4.6.3 Half Hourly Readings

Most monitoring locations on WPD's network take average half hourly readings. This is therefore the best resolution it is possible to view the effects of flexibility services. It would therefore seem most straightforward to apply average half hourly readings to settling flexibility services. However this could be detrimental to all parties involved. It could be detrimental for the DNO procuring services, as it increases risk as 15 minutes of under generating could be masked by 15 minutes of over generating. It could be detrimental to the earnings of providers as brief significant dips or data gaps would affect a whole half an hour, instead of only a few minutes. Several examples of this were found, including from event E9, shown in Figure 21.

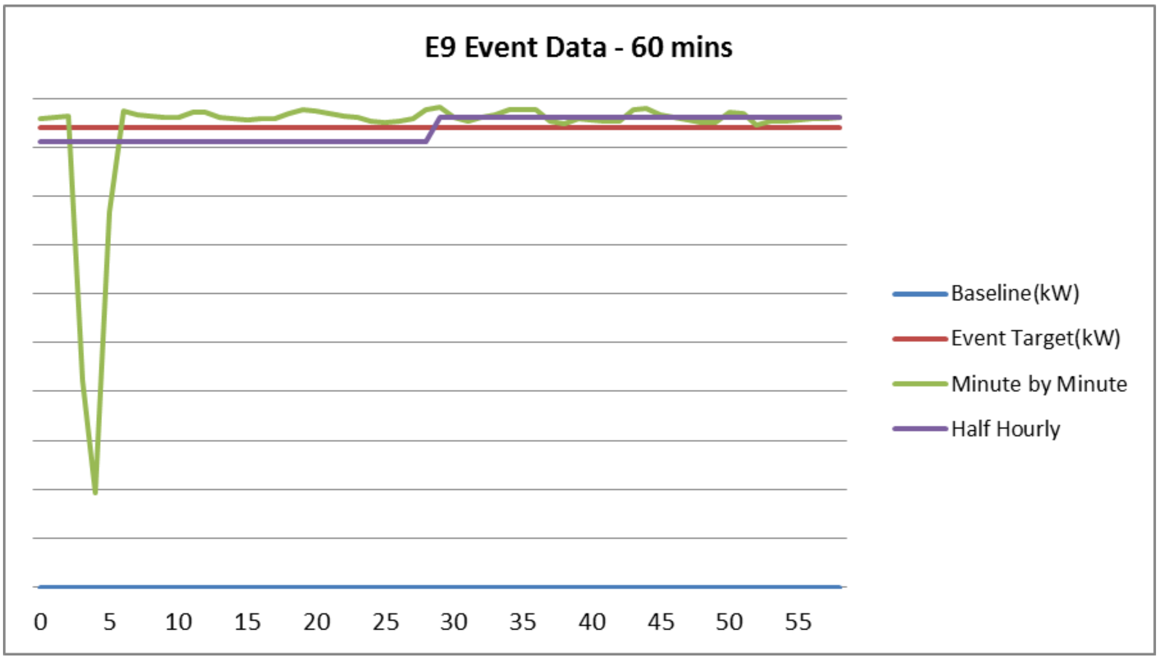


Figure 21: Minute by Minute Event Resolution Versus Half-Hourly Resolution

5 Learning Generated

5.1 Known Limitations

5.1.1 Limited Volume

Limited trial participant number limits the conclusions that can be drawn from the data due to the sample size.

5.1.2 Limited Monitoring

The use of existing network monitoring points limited the impact that could be detected on the network. However this level of monitoring is representative of the network.

5.2 Learning Generated

5.2.1 Baseline Distortion

The inclusion of periods where generators are providing flexibility services in the current baselining calculations can distort the baseline, causing increases over time. This can lead to loss of income for the generator and would seem likely to discourage the long term participation of generators in a DNO flexibility marketplace.

5.2.2 Need for Over-Procurement

An expected learning point, but definitely affirmed by the trial is that to provide a reasonable guarantee of the requested level of delivery, a greater quantity of flexibility services must be procured. The level of over-procurement deemed necessary will depend on the risk appetite of the DNO, and will depend on the criticality of the asset.

5.2.3 Need for Contingencies

This trial saw that in 12% of events participants did not respond. This reinforces that a pragmatic approach must be taken when procuring flexibility services to manage this risk.

5.2.4 Automated Issue Identification

Some events in this trial were not successes due to participants not being aware of changes in baseline, or potentially that they were underperforming. Developing systems to clearly identify changes in baseline, output, and other potential issues would help participants fulfil their obligations and engage in the process, and the DNO identify issues more proactively.

5.3 Potential for Development

5.3.1 More Volume

Due to the limited scale of the trial achieved, the validity of conclusions is limited and going forward with the role out of flexibility services further analysis will be needed to confirm or revise these assumptions. Greater volumes of procurement and dispatch will add to the learning started here to determine the normal reliability and service a DNO can expect when procuring flexibility services.

5.3.2 More in Depth Examination of Technology Types and Other Risk Factors

While it is preferable to keep flexibility service procurement as technologically agnostic as possible to help create a fluid and open marketplace, analysis needs to be done to assess the risk factors involved in different technologies, and greater or lesser diversity of supply.

5.3.3 Standard Flexibility Risk Management Methodology

For the full scale successful roll out of flexibility services it seems necessary for the development of a standard risk management methodology to be developed across the industry, as the ER-P2 documents provide an industry wide standard assessment of risk for the planning of the physical network.

5.3.4 Common Metering Requirement

To simplify the metering requirements, and limit costs to service providers, a common industry approach to the metering of flexibility services would provide significant benefit.

6 Contact

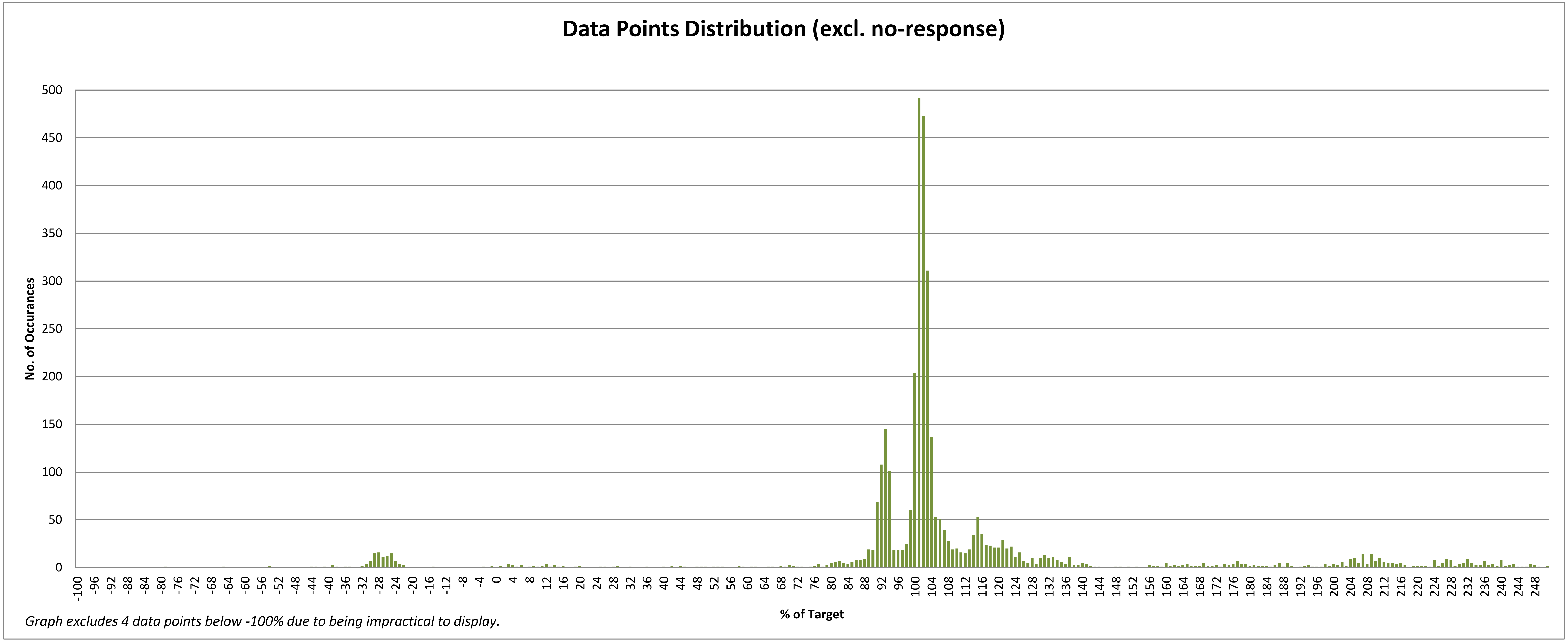
Further details on replicating the project can be made available from the following points of contact:

Innovation Team
Western Power Distribution,
Pegasus Business Park,
Herald Way,
Castle Donington,
Derbyshire
DE74 2TU
Email: wpdinnovation@westernpower.co.uk

Appendix A: Data Point Distribution

New ID	Event Duration (mins)	Event Measurements	No. of Gaps	% above Target	% Above 95%	% Above 63%	Contracted Capacity Delivered	Scoring Mechanism (0,1,8,10)	Score Judgement	% of Possible Revenue Rewarded	Comments
E1	120	79	41	63.29	89.87	98.73	66%	85	Below Expectations	98%	Early drop off (3mins)
E2	60	60	0	75	86.67	95	100%+	85	Below Expectations	90%	Medium dip near start
E3	120	120	0	100	100	100	100%+	100	Excellent	100%	
E4	240	240	0	35	41.67	56.25	48%	42	Poor	54%	Large prolonged dip
E5	120	120	0	91.67	93.33	95.83	100%+	93	Good	95%	Small dip near start
E6	120	121	0	100	100	100	100%+	100	Excellent	100%	
E7	60	60	N/A	**0**	0	0	N/A	N/A	N/A	0%	No data
E8	120	120	N/A	0	0	0	N/A	N/A	N/A	0%	No data
E9	60	59	1	94.92	94.92	96.61	98%	95	Good	96%	Small dip near start
E10	120	119	1	100	100	100	100%+	100	Excellent	99%	
E11	120	120	0	100	100	100	100%+	100	Excellent	100%	
E12	60	59	1	0	0	0	0%	0	Fail	0%	Generator trip immediately before event
E13	180	181	0	97.24	97.79	98.34	100%+	98	Good	96%	Slow to ramp up (5mins)
E14	120	120	0	100	100	100	100%+	100	Excellent	100%	
E15	60	61	0	88.52	100	100	100%+	98	Good	100%	
E16	60	61	0	100	100	100	100%+	100	Excellent	100%	
E17	60	60	0	48.33	100	100	100%+	90	Below Expectations	100%	
E18	60	60	0	90	95	96.67	98%	94	Good	95%	Slow to ramp up (3mins)
E19	60	60	0	96.67	100	100	100%+	99	Good	100%	
E20	60	61	0	37.7	55.74	100	98%	57	Poor	85%	
E21	60	61	0	98.36	98.36	100	100%+	99	Good	99%	Very unstable output
E22	60	61	0	93.44	93.44	93.44	50%	93	Good	93%	Probable data gap for 5mins
E23	60	60	0	100	100	100	100%+	100	Excellent	100%	
E24	60	61	0	98.36	100	100	100%+	100	Good	100%	
E25	60	61	0	95.08	95.08	95.08	100%+	95	Good	95%	Early drop off (3 mins)
E26	60	61	0	0	0	0	0%	0	Fail	0%	No response
E27	60	60	0	100	100	100	100%+	100	Excellent	100%	
E28	120	120	0	95.83	100	100	100%+	99	Good	100%	
E29	120	121	0	78.51	78.51	79.34	100%+	79	Poor	80%	Very slow to ramp up (25 mins)
E30	120	120	0	97.5	100	100	100%+	100	Good	100%	
E31	120	121	0	35.54	54.55	89.26	86%	54	Poor	77%	Large dip and tail off
E32	180	181	0	0	0	0	7%	0	Fail	0%	No response
E33	180	181	0	82.32	100	100	100%+	96	Good	100%	
E34	60	61	0	0	0	0	0%	0	Fail	0%	No response
E35	60	61	0	95.08	100	100	100%+	99	Good	100%	
E36	60	61	0	78.69	78.69	86.89	96%	80	Poor	85%	Slow to ramp up (13mins)
E37	60	61	0	0	0	100	90%	10	Poor	76%	Under-response
E38	60	61	0	0	0	0	8%	0	Fail	0%	No response
E39	120	119	1	95.8	100	100	100%+	99	Good	100%	
E40	120	120	0	100	100	100	100%+	100	Excellent	100%	
E41	120	120	0	0	0.83	100	93%	11	Poor	88%	Benchmarking issue -> Under-Response
E42	120	120	0	0	2.5	100	93%	12	Poor	88%	Benchmarking issue -> Under-Response
E43	120	120	0	0	0.83	100	93%	11	Poor	89%	Benchmarking issue -> Under-Response

Appendix B: Data Point Distribution



	Including no-response events		Excluding no-response events	
Above 100%	2614	66%	2614	74%
Above 95%	178	5%	178	5%
Above 63%	545	14%	545	16%
Below 63%	597	15%	174	5%
Total	3934	100%	3511	100%