Report Covering

May 2010 to April 2011 Inclusive

# Report on the Accuracy of the System Management Action Flagging Methodology

nationalgrid

The power of action.

## Executive Summary

The P217A – Revised Tagging Process and Calculation of Cash Out Prices was implemented from November 2009. The modification aims to remove pollution from the imbalance price caused by actions taken to resolve transmission constraints. Under this modification the System Operator determines which actions are taken to resolve constraints and flag these actions. These flags are then sent to the BSC Systems and used in the imbalance price calculation methodology.

National Grid developed a System Management Action Flagging Methodology Statement (SMAF) which outlines the methodology used by the System Operator in determining what actions should be flagged as constraints.

To ensure that the flagging methodology is operating as intended, National Grid committed to produce a report after the first six months, and thereafter on an annual basis, looking at the accuracy of the methodology and considering any materiality. This is the second of such reports, covering the 12 months between May 2010 – April 2011 inclusive.

The initial report covering the first 6 months found that in almost all cases P217A flagging methodology was correctly applied in the spirit of the Balancing Services Code, with only 2.6% of overall actions being subject to potential error, and where error may have occurred these had a limited impact on imbalance prices. It also noted that a more robust assessment process was desirable.

This report finds that as a result of the success of a range of actions to increase the robustness of the P217A flagging process the number of potential errors has fallen to 0.88% of overall actions and that those few inaccuracies which may have slipped through had a negligible impact on imbalance prices.

If you have any comments or queries on this report, please contact National Grid on:

**balancingservices@uk.ngrid.com**

nationalgrid
The power of action.

## Contents



nationalgrid
The power of action.

# Section 1
# Introduction

*This section provides an introduction as to the rationale behind the development of this report.*

## 1.1 Purpose of the report

The objective of this report is to present a review of the accuracy of the P217A flagging process that took place between 1st May 2010 – 30th April 2011. Under the SMAF Methodology Statement, National Grid is required to report on the accuracy of the flagging methodology 6 months after implementation and thereafter, on an annual basis. This is the second report and so spans 12 months of P217A operation and National Grid's flagging of constraint actions.

## 1.2 Outline of P217A SO Flagging

The rationale behind the development of this report is discussed in the initial report document covering November 2009 – April 2010. Its objective is to remove distortive pollution from cash out caused by bids Offer Acceptances (BOAs) taken to resolve transmission constraints. This follows from a P217A review begun in 2007, which so far as concerns this report states that from the 5th November 2009, under the Balancing Settlement Code (BSC) section Q5.3.1(d) and section Q6.3.2(b) National Grid is required to assess whether an action is wholly or partly taken to resolve a transmission constraint. Such actions are 'SO-Flagged' for the purposes of the BSC Systems who then determine the cash prices using the P217A cash out price methodology.

In practice SO-Flagging of BOA actions occurs when National Grid identifies specific Balancing Mechanism Units (BMUs) that, in the event of an active transmission constraint, would be utilised to resolve the constraint, and actions on these units are subsequently flagged by National Grid Control Room in real time for the duration of resolving a constraint . When the Control Room is satisfied that the transmission constraint is no longer active the BMUs are de-flagged and therefore, any actions taken thereafter are not flagged as resolving a constraint. The accuracy with which this flagging takes place is the subject of this report.

### 1.3 P217A Flagging Assessment Methodology

National Grid uses a number of different processes to assess the accuracy of the Control Room Flagging process and identify potential periods where errors may have occurred. Below we outline the 3 main processes used in determining the accuracy.

*Data Inquiry Report.*

Used in the event of the Control Room becoming aware that the flagging of constraint BOAs has been incorrectly set in real time. The Control Room can raise a Data Inquiry report (DIR) to note the discrepancy. 78 DIRs were recorded in this annual review period (53 in initial 6 month report).

*Post Event Cross Reference (Working Day +1)*

Since the introduction of flagging in November 2009, a procedure has been in place to review the accuracy of flags that works by cross-checking the units identified by P217A flags against the manual process that allocates Constraint Costs undertaken for BSIS reporting. This process takes place on a working-day +1 basis, in which BOA actions are analysed against various operational reports and if taken to resolve a constraint they are 'tagged' with a constraint cost marker ('BSIS SUPERBAAR Constraint Cost Tagging'). Apparent differences between the P217A flagging and SUPERBAAR tags are reviewed with the Control Room as necessary to better determine the correct P217A flags & BSIS tags.

A high correlation between the P217 Flagging and the SUPERBAAR Constraint Tagging is expected but it should be noted that differences between the two mechanisms do exist due to the different criteria that apply for flagging under SMAF and tagging under BSIS SUPERBAAR: - in particular relating to;

- The treatment of actions that resolve both constraint and margin issues, these being flagged under P217A but not seen as an additional cost under BSIS as they are required for margin; in which case they would carry a P217A flag but no SUPERBAAR tag.
- Differences due to other anomalies such as the running of units for Black Start security; such actions being neither a balancing issue nor a constraint issue and so would carry a P217A flag for cash out but no SUPERBAAR tag.

*Post Event Periodic Review*

A period-by-period analysis of P217A performance is done on a weekly basis at week +1, in which P217A flagging & SUPERBAAR tagging is cross-matched so as to give an indication of incorrect, under/ over-tagging and missing flagging/tagging issues. This picks up on any data which may have been missing

or late at the time of the Post Event Cross Reference above. This review is written up and is shared with Control staff for any learning points that may arise.

Since the initial report both the Post Event Cross Reference and Post Event Periodic Review processes have been strengthened by the creation of tools that provide clarity on bid and offer actions by BMU and half hour periods. These enable errors to be identified more easily. The process is further strengthened by a monthly performance report for control staff.

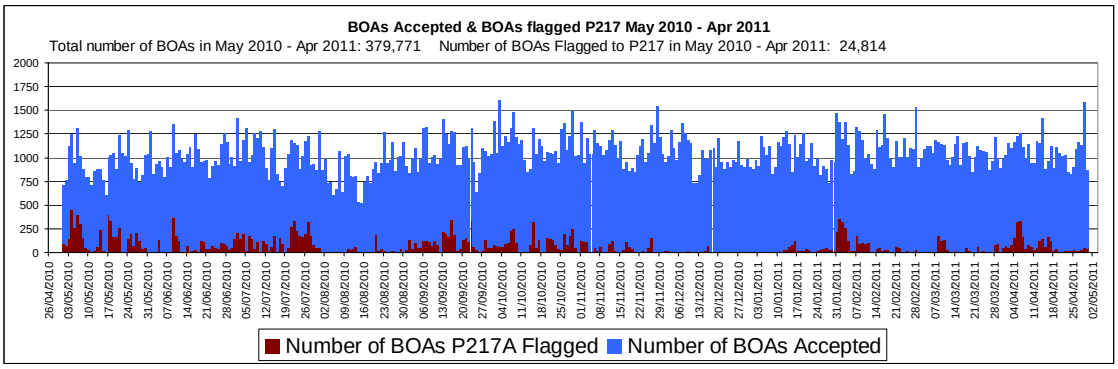
# Section 2
# Flagging Accuracy

## 2.1 Overall Statistics

During the 12 months May 2010 – April 2011, 24,814 BOAs were flagged under the P217A criteria out of the total 378,913 BOAs accepted, equating to approximately 6.55% of the total actions. The distribution of these actions are tabulated and charted below. The number of half-hour periods in the twelve-month review was 17,520, of which 7,554 periods had BOA actions that were P217A flagged (43%).

| ***Month*** | ***Total Number of BOAs*** | ***Number of BOAs P217A Flagged*** | ***% Flagged to P217*** |
|---|---|---|---|
| May - 2010 | 29,060 | 4,478 | 15.41% |
| Jun-10 | 30,364 | 1,765 | 5.81% |
| Jul-10 | 32,652 | 4,055 | 12.42% |
| Aug-10 | 27,019 | 553 | 2.05% |
| Sep-10 | 31,671 | 3,307 | 10.44% |
| Oct-10 | 35,887 | 3,027 | 8.43% |
| Nov-10 | 33,126 | 1,233 | 3.72% |
| Dec-10 | 31,578 | 179 | 0.57% |
| Jan-11 | 32,135 | 946 | 2.94% |
| Feb-11 | 31,557 | 2,011 | 6.37% |
| Mar-11 | 32,502 | 890 | 2.74% |
| Apr-11 | 31,362 | 2,370 | 7.56% |
| Number of BOAs Flagged to P217 in May 2010 - Apr 2011: | | 24,814 | 6.55% |
| All BOAs accepted | 378,913 | | |

The chart below illustrates the days in which actions were P217A flagged. The flagged actions are shown in red with the overall count of actions shown in blue. It can be seen that constraint actions (red) generally occur across a number of days due to the constraint being active over an outage period that may last for a week or possibly longer.

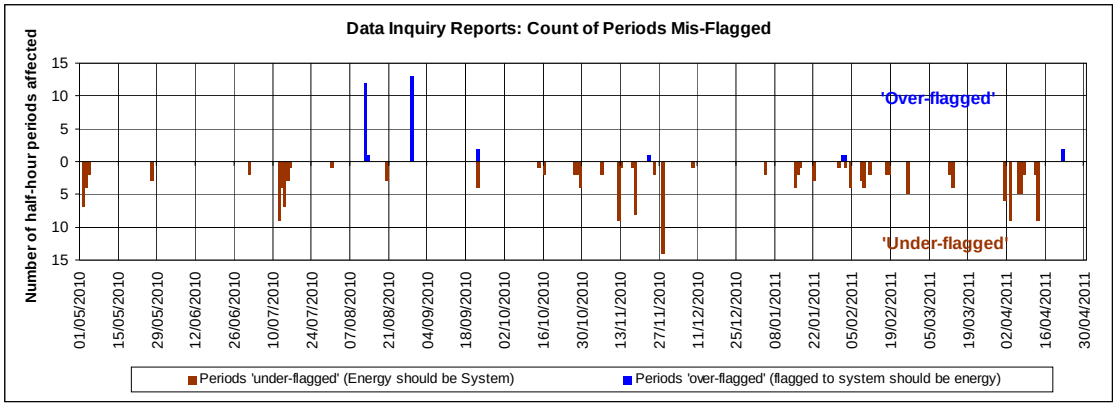


## 2.2 Flagging Errors Known in Real Time

As mentioned in Section 1.2 above, P217A flags are applied by Control staff in real time while balancing the system. This is a manual task and occasionally flags are misapplied, often reflecting higher levels of workload in Control at the time. When such an error is realised within Control timescales it is logged though a Data Inquiry Report (DIR). 78 DIRs were raised in the 12 month. These reports may cover several BOA actions on one or more BMU generator units.

| Month | Number of Data Inquiry reports raised due to P217 errors |
|---|---|
| May-10 | 5 |
| Jun-10 | 0 |
| Jul-10 | 12 |
| Aug-10 | 5 |
| Sep-10 | 4 |
| Oct-10 | 5 |
| Nov-10 | 10 |
| Dec-10 | 1 |
| Jan-11 | 7 |
| Feb-11 | 11 |
| Mar-11 | 3 |
| Apr-11 | 15 |
| Total | 78 |

The chart below shows the number of half-hour periods in which BMU generator units were reported as mis-flagged.




Most errors resulted from not applying P217 flags – 'under-flagging' (as opposed to 'over-flagging' by leaving flags on or incorrect actions). The greatest occurrence was on 28/11/2010 when the flagging of Fiddlers Ferry units 3 & 4 was missed during the early morning and may have had an impact on prices in 11 periods. This incident is further investigated in Section 3 below, together with those of 27/5/2010, 12/08/2010, 20/08/2010 & 12/11/2010. However, neither the number DIRs, nor the count of periods affected are indicators of the magnitude of potential impact on Cashout. Impact on

Cashout will be dependent on prices and volumes involved in context of the prices and volumes of all BOAS at the time.

### 2.3 Assessment of P217A Flagging Accuracy by Cross Reference to SUPERBAAR Constraint Tagging

As in the Initial report, the primary method for assessing accuracy is by cross - referencing the P217A flagging against the BSIS SUPERBAAR constraint cost tagging process (see section 2 above). This considers how BOAs are spread across their respective half-hour periods; 'BOA.Period Actions' are defined as a BOA, which may spread over several half hour periods, and the periods that they affect. Four categories result:

1. 'Energy' – Periods where there was no P217A flagging or actions tagged under the BSIS SUPERBAAR process.

2. BOA.Period actions that tally under both P217A flagging and the BSIS SUPERBAAR tagging process (P217A = SUPERBAAR 'Constraint' actions)

3. BOA.Period action where P217A flags have no corresponding SUPERBAAR tag (legitimate system/margin actions with both P217A flags and SUPERBAAR tags correctly set, or possible P217A over-flagging errors or SUPERBAAR under-tagging errors)

4. BOA.Period actions tagged by SUPERBAAR but with no P217A flag (possible P217A under-flagging errors / SUPERBAAR over-tagging errors)

The cross-reference of these gives the results tabulated below:

| Summary of BOA.Period Action Flagging May 2010 to end Apr 2011 | Totals | As % of all BOA.Period Actions | As % of BOA.Periods Flagged or Tagged |
|---|---|---|---|
| Number of BOA.Period Actions | 782,887 | 100.00% | - |
| BOA.Period actions assigned to Energy (not P217A 'system' nor SUPERBAAR 'constraint') | 715,400 | 91.38% | - |
| BOA.Period actions that tally under both P217A flagging and the BSIS SUPERBAAR tagging process (Constraint actions) | 59,987 | 7.66% | 89.71% |
| BOA.Period action with P217A flags, but no SUPERBAAR tag (legitimate system / margin / possible P217A over-flagging / SUPERBAAR under- | 4,245 | 0.54% | 6.35% |
| BOA.Period actions tagged by SUPERBAAR but with no P217A flag (Possible P217A under-flagging / SUPERBAAR over-tagging) | 2,636 | 0.34% | 3.94% |
| Total BOA.Periods with P217A Flag or SUPERBAAR tag (Flagged or Tagged) | 66,868 | 9.35% | 100.00% |

The table shows that of the 782,887 BOA.Period actions within the assessment period, 64,232 had P217A flags (59,987 + 4,245 , 8.20%).

Of the total number of BOAs.Period Actions taken:

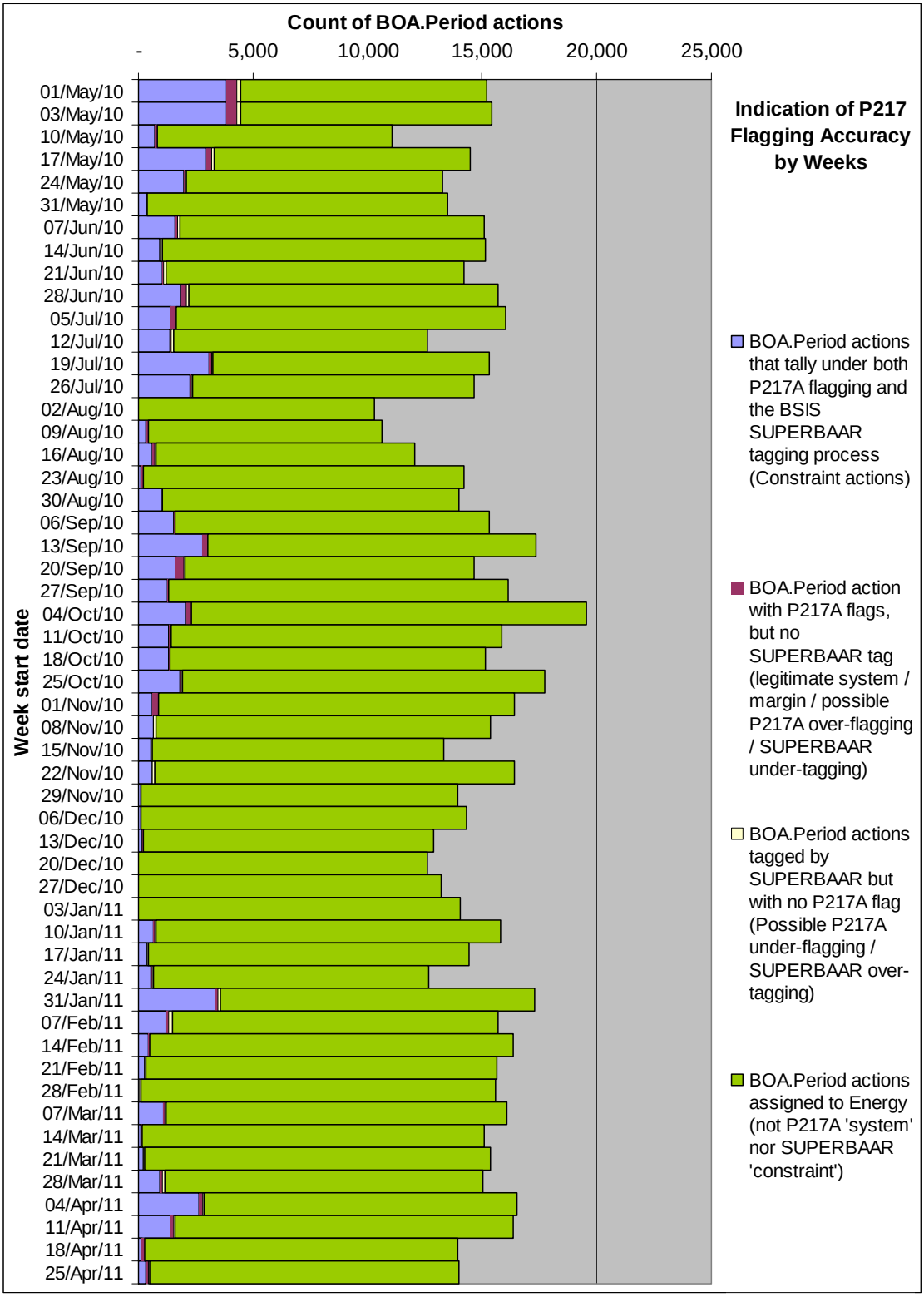
- 91.38% were allocated as Energy actions [Initial report 88.82%]
- 7.66% were allocated as Constraint actions [Initial report 8.56%]
- 0.54% of all actions had P217A flags but no corresponding SUPERBAAR tag. [Initial report 1.61%]. This figure is after adjustment to remove occasions of known legitimate mismatches arising from black start instructions on 25 days in the review period, accounting for 274 BOAs and 619 BOA.Periods.
- 0.34% of overall actions had a SUPERBAAR tag with no corresponding P217A flag [Initial report 1.00%]

**Overall potential inaccuracy** is a maximum of 0.88% of overall actions (0.54% P217A flags no SUPERBAAR tags + 0.34% SUPERBAAR tags no P217A flags) [Initial report 2.6%]

Above is the accuracy of P217A actions as a percentage of overall actions. The accuracy of the actions taken for constraints based on the set of 66,868 actions that could be considered appropriate system actions under either P217A or SUPERBAAR:

- 89.71% were in agreement as assigned to Constraint [Initial 76%]
- 6.35% had an element of difference due to treatment as Margin by SUPERBAAR, possible P217A over-flagging or SUPERBAAR under-tagging [Initial 14%]
- 3.94% had an element of difference due to possible P217A under-flagging, or SUPERBAAR over-tagging [Initial 9%]

A breakdown of the summary figures on a week by week basis is set out in the Indication of P217A Accuracy chart below. The profiles are in reasonable agreement with the chart of BOAs Accepted & BOAs Flagged in 2.1 above.

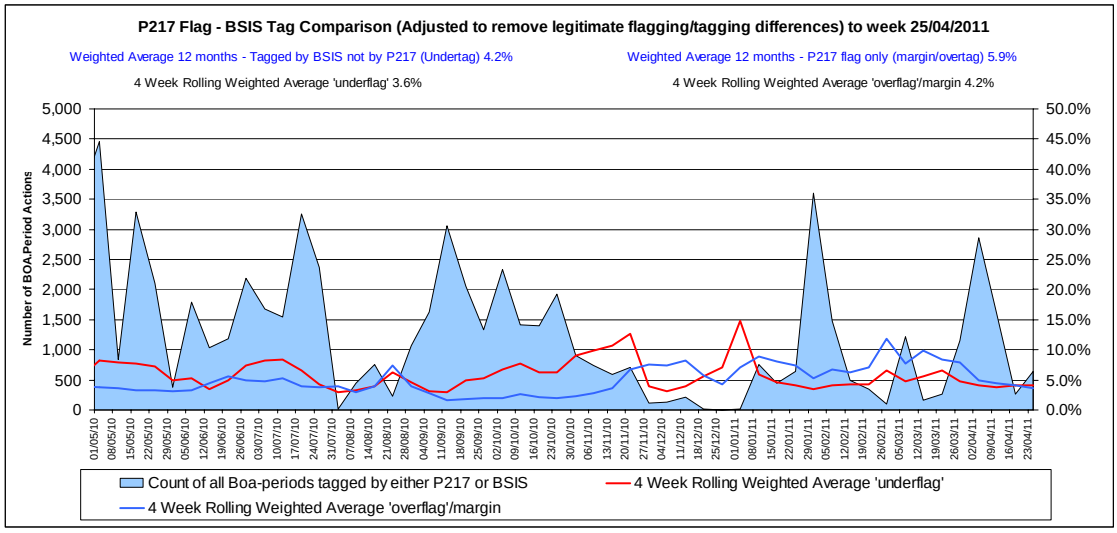

Count of BOA.Period actions
-
5,000
10,000
15,000
20,000
25,000
Week start date
01/May/10
03/May/10
10/May/10
17/May/10
24/May/10
31/May/10
07/Jun/10
14/Jun/10
21/Jun/10
28/Jun/10
05/Jul/10
12/Jul/10
19/Jul/10
26/Jul/10
02/Aug/10
09/Aug/10
16/Aug/10
23/Aug/10
30/Aug/10
06/Sep/10
13/Sep/10
20/Sep/10
27/Sep/10
04/Oct/10
11/Oct/10
18/Oct/10
25/Oct/10
01/Nov/10
08/Nov/10
15/Nov/10
22/Nov/10
29/Nov/10
06/Dec/10
13/Dec/10
20/Dec/10
27/Dec/10
03/Jan/11
10/Jan/11
17/Jan/11
24/Jan/11
31/Jan/11
07/Feb/11
14/Feb/11
21/Feb/11
28/Feb/11
07/Mar/11
14/Mar/11
21/Mar/11
28/Mar/11
04/Apr/11
11/Apr/11
18/Apr/11
25/Apr/11
Indication of P217 Flagging Accuracy by Weeks
BOA.Period actions that tally under both P217A flagging and the BSIS SUPERBAAR tagging process (Constraint actions)
BOA.Period action with P217A flags, but no SUPERBAAR tag (legitimate system / margin / possible P217A over-flagging / SUPERBAAR under-tagging)
BOA.Period actions tagged by SUPERBAAR but with no P217A flag (Possible P217A under-flagging / SUPERBAAR over-tagging)
BOA.Period actions assigned to Energy (not P217A 'system' nor SUPERBAAR 'constraint')

## 2.4 Discrepancies Interpretation Discussion

Behind the figures presented above lie:

- Legitimate discrepancies (System reasons that are not Constraint reasons and not already accounted for in the data)
- 'Straggler mismatch' noise error because P217A data flags the discrete BOAs whereas the SUPERBAAR system tags are in the form of discrete half hour periods and so an action fitting into one time pigeonhole in one system may fit a slightly different one in the other.
- Flagging & Tagging errors on behalf of Control or BSIS.

The chart below illustrates the number of BOA.Period Actions in which P217A flags and/or BSIS tags were applied in each week of the year in review (the blue, red & yellow areas in chart above). The percentage error is given in the form of 'volume weighted average' because a simple percentage of number of BOA actions gives a distorted view. (The 'error' difference can be greatest on days when the underlying volume of BOA.Period Actions is small, simply because the 'straggler mismatch' noise tends to dominate).




It can be seen that the 4 week rolling average of under-flagging and over-flagging errors generally lie below 10% and frequently below 5% of any actions flagged or tagged. The volume weighted average over the entire 12 month period is:

| | |
|---|---|
| Tagged by SUPERBAAR not by P217 flags (under-flag) | 4.2% |
| Flagged by P217 not by SUPERBAAR (over-flag) | 5.9% |

These are believed to be good performance figures when one considers that they contain a background 'noise' and that flagging is applied in real time in an active control

room via a manual process, under circumstances where System conditions can be changing rapidly.

# Section 3
# Materiality

## 3.1 Materiality of Inaccuracies

An initial inspection of the discrepancies around P217 flag setting during the 12 months revealed few occurrences where there would have been anything but a limited or zero effect on cash out prices.

Incidents where P217A flagging may have had a material effect on Cashout were subjected to further scrutiny and re-run through the Elexon price calculation system with flags adjusted to determine what difference, if any, resulted. The incidents for re-run were selected from the DIRs in Section 2 together with other mis-flaggings that were revealed by this review where the plant and actions were judged to have potential impact. These are tabulated below.

| Date | Issue | Period numbers where difference occurs | Original system Buy Price £/MWh | Revised System Buy Price £/MWh | Difference in System Buy Price £/MWh | Original system Sell Price £/MWh | Revised System Sell Price £/MWh | Difference in System Sell Price £/MWh |
|---|---|---|---|---|---|---|---|---|
| 27/05/2010 | COTPS1,2,4 & CDCL-1 should have been flagged. DIR. Apply flags Periods 16-18 | - | - | - | - | - | - | - |
| 12/08/2010 | Fidl-4 control flagged to P217 in error periods 29-41. Should be energy - remove flags 29-41 | 30<br>34<br>36<br>37<br>38 | 37.66<br>41.80<br>41.49<br>41.41<br>40.34 | 37.66<br>41.80<br>41.49<br>41.41<br>40.34 | -<br>-<br>-<br>-<br>- | 32.72<br>31.90<br>32.03<br>30.85<br>31.28 | 31.76<br>31.68<br>30.97<br>29.69<br>30.86 | - 0.97<br>- 0.22<br>- 1.06<br>- 1.16<br>- 0.42 |
| 20/08/2010 | COCK-1, LOAN-1 should have been flagged. DIR. Apply flags Periods 1-3 | - | - | - | - | - | - | - |
| 07/10/2010 | Periods 2 - 11 heavily polluted by incorrectly flagged Offers on Scottish plant that picked up P217 flags which had been put in place for SCOTEX. Remove flags on FIFE-1, LOAN-1,2,3,4, PEHE-1 offers periods 2-11 | - | - | - | - | - | - | - |
| 08/10/2010 | Periods 2 - 11 heavily polluted by incorrectly flagged Offers on Scottish plant that picked up P217 flags which had been put in place for SCOTEX. Remove flags on LOAN-2, 3, 4, PEHE-1 offers periods 2-11 | - | - | - | - | - | - | - |
| 03/11/2010 | Stressful day in Control due high wind volatility BOAs on CRUA 1-4, FOYE-2, COCK-4 periods 34-37 flagged in error. Remove flags on these actions | - | - | - | - | - | - | - |
| 12/11/2010 | FIDL-1, 2, 3 ,4 units over periods 40-46 should have been flagged for GMERSYP2. DIR. Apply flags to these BOAs periods 40-46 | - | - | - | - | - | - | - |
| 28/11/2010 | FIDL-3 & 4 flags missed should have been flagged system for FIDFEX periods 1-11. DIR | - | - | - | - | - | - | - |
| 28/02/2011 | Offers on Sloy flagged in export constraint periods 32-36 Remove flags | - | - | - | - | - | - | - |

The table shows that of the nine sets of actions identified as having a potential impact on Cashout only those relating to the over-flagging of FIDL-4 during five half-hour periods in the afternoon of 12 August had a material effect on Cashout; the P217A flagging had the effect of distorting the System Sell Price by circa +£1 / MWh at that time. These are of the same order as those in the Initial Report. The missed P217A flags on 28/11/2010 for Fiddlers Ferry units 1 & 3 which were the subject of a DIR (Section 2.2 above) had no impact.

### 3.2 Materiality & Flagging Accuracy Conclusion

Section 2 concluded that over the last 12 months the average potential error in under-tagging or over-tagging of those actions flagged of tagged was below 6%, and the potential inaccuracy as a percentage of all actions was 0.88%. Where errors have been identified they have been found to be of small order, they are largely the result of over-flagging rather than under-flagging, and only impacted Cashout on five periods.

## 4 Discussion of Performance Year-on-Year

This is the first report of P217A flagging to span a whole year. The main indications are summarised on the table below.

| | Nov 2009 - April 2010 | May 2010 - Apr 2011 |
|---|---|---|
| | In 6 months | In 12 months |
| **Number of DIRs raised** | 53 | 78 |
| **Number of BOAs Flagged to P217** | 15,345 | 24,814 |
| % flagged to P217A | 8.14% | 6.55% |
| **Number of BOA.Period Actions** | 377,574 | 782,887 |
| % assigned to energy | 88.82% | 91.38% |
| % P217A & SuperBaar agree | 8.56% | 7.66% |
| % P217 flags, not SuperBaar tags | 1.61% | 0.54% |
| % SuperBaar tags, not P217 flags | 1.00% | 0.34% |
| **Overall potential inacuracy** | 2.60% | 0.88% |

It can be seen that the number of Bid Offer acceptances (BOAs) flagged to P217 in the full year was around 60% more than in the first six months of the scheme. Such differences can be expected as the dynamics of managing the system are changeable, for example; resolving constraints by other methods than BOAs (e.g. by contracts or intertrips), taking fewer BOAs but each over longer periods, the different stresses on the system over winter and summer with the impact that system outages may then make and variation of the level of system outages in one year to the next as driven by the Grid's system maintenance and investment plan.

By all measures P217A Flagging accuracy performance over the 12 months in this review has improved over that of the initial six months. This has been due to the bedding in of various processes, diligence and the introduction of various measures to increase the robustness of P217A monitoring and assessment.

# Section 5
# Conclusions

## 5.1 Flagging Performance Conclusions

This report finds that P217A flagging performance in the 12 months of this review is a marked improvement on that of the already good performance of the initial six months. Our analysis indicates that the P217A flagging methodology is performing as expected.