respect › commitment › teamwork

13 Month AMR Validation Analysis

Spring 2014 Modelling

TWG - 25th March 2014

Background

- Each year the Spring modelling will consider the sample data for the most recent summer and winter period over the previous 12 months. In normal circumstances this would cover the period 1st April to 31st March
- Usually this 12 month period would include a full Easter holiday period (as defined by the modelling system)
- As described in the Spring Approach document, this years modelling data the period 1st April 2013 to 31st March 2014 does not include a complete Easter holiday period. The modelling will therefore be need to be performed over a 13 month period - 1st March 2013 to 31st March 2014
- Prior to modelling, the sample data is processed and cleansed following validation criteria described in Appendix 1 of the NDM report. The last occurrence of this scenario was is in Spring 2009

Background

- In Spring 2009, validation rules for the Small and Large NDM data loggers were adjusted to reflect the fact that 'Summer' was a month longer, i.e. instead of 1st April to 30th September it was extended to 1st March to 30th September. These rules for data loggers will be applied in Spring 2014

- For the sample managed by Xoserve, referred to as AMRs (Band 1 and 2 sites) there has never been the need to use a 13 month period because until recently the 12 month analysis period was 17th March to 16th March which always includes all of the Easter holiday days

- This is the first occurrence of a 13 month validation period being required for AMR's and so a review of its validation rules is necessary

respect › commitment › teamwork

AMR Validation Rules

- The current validation rules for AMR 'Missing Days' are as follows
  - Any AMR with 15 or more missing days of data during Summer (01 Apr - 30 Sep) will be rejected
  - Any AMR with 15 or more missing days of data during Winter (01 Oct – 31 Mar) will be rejected
- The current validation rules for AMR 'Consecutive Zeros' are as follows
  - Any AMR with 33 or more consecutive daily zero consumptions during the Winter (01 Oct – 31 Mar) will be rejected. No impact as Winter only
- The current validation rules for AMR 'Spikes' are as follows:
  - The ratio of the maximum consumption to the average consumption cannot exceed 8:1 during winter (01 Oct – 31 Mar)
  - The ratio of the maximum consumption to the average consumption cannot exceed 15:1 during summer (01 Apr – 30 Sep)

Objective

- As the 'Summer' period will be extended to 1st March 2013 to 30th September 2013 the rules require a review to consider if they need any adjustment

- The purpose of this analysis is to investigate the appropriateness of

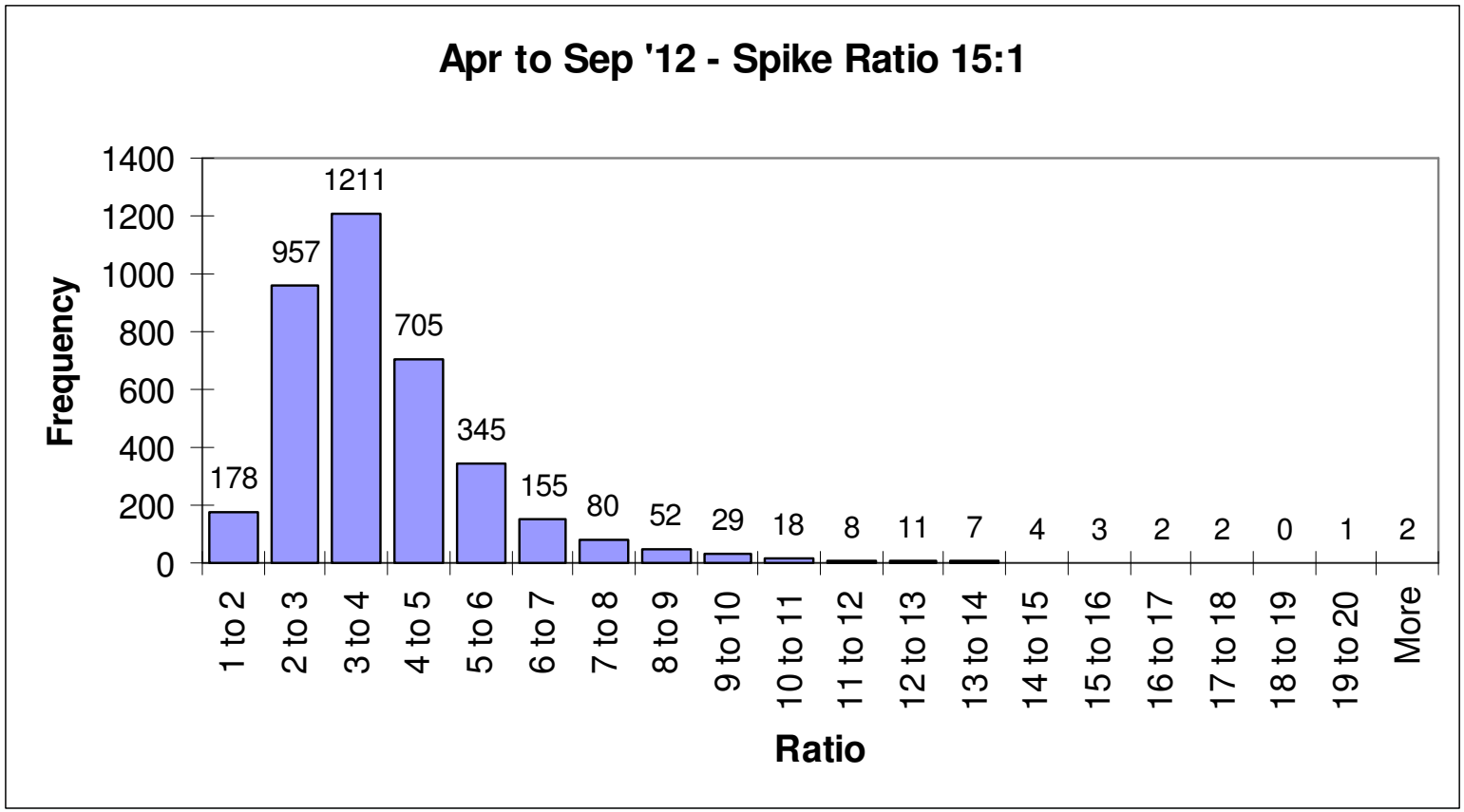
  – a) the current spike validation rules when using an extra month (March) in the summer validation period

  – b) the current missing days validation rules when using an extra month (March) in the summer validation period

Test Data Used in Analysis

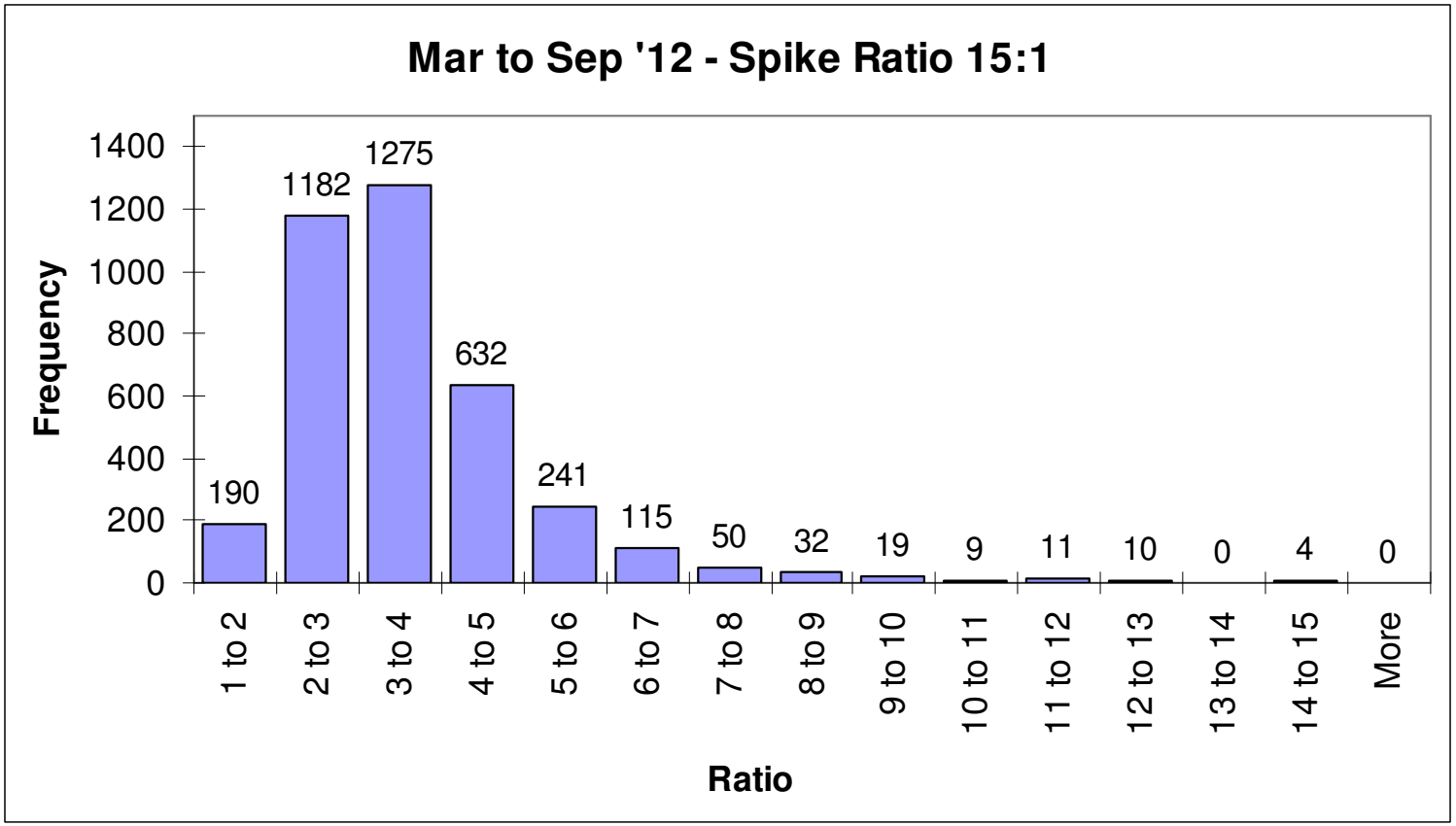
- The data analysed was the most recent equivalent 13 month period
  - Validated sample data for the period 01/03/2012 – 31/03/2013
  - 3770 sites passed the 13 month spike validation analysis
    - Band 1 – domestic sites only (Mon – Thu excluding holidays)
      - 3010 sites passed validation
    - Band 2 – domestic and non domestic sites (Mon – Thu excluding holidays)
      - 760 sites passed validation

Frequency of ratio values (Apr – Sep)



The above Histogram shows the 3760 sites that had a ratio below 15 and passed validation. There are 10 sites with a ratio above 15 that failed the summer validation (Apr-Sep) but passed the summer validation (Mar-Sep)

Frequency of ratio values (Mar – Sep)



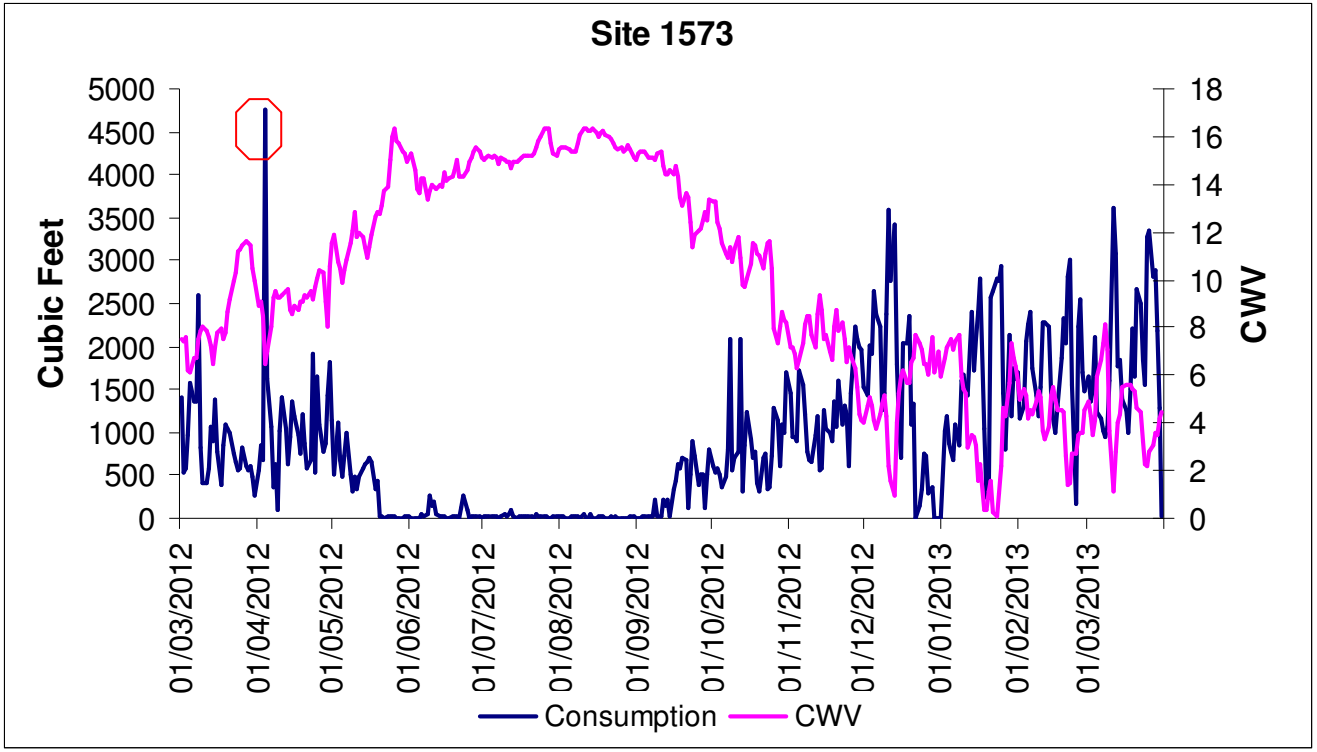
xoserve

respect › commitment › teamwork

Analysis

- The 10 sites that failed the Apr-Sep validation but passed the Mar-Sep validation were then analysed to see if they failed based on errors or if the consumption appeared to be genuine
  - 1 site appeared to contain an error
  - 9 sites appeared to be genuine consumptions
- The following graphs demonstrate both examples:

Analysis



The spike on this chart appears to be an error.

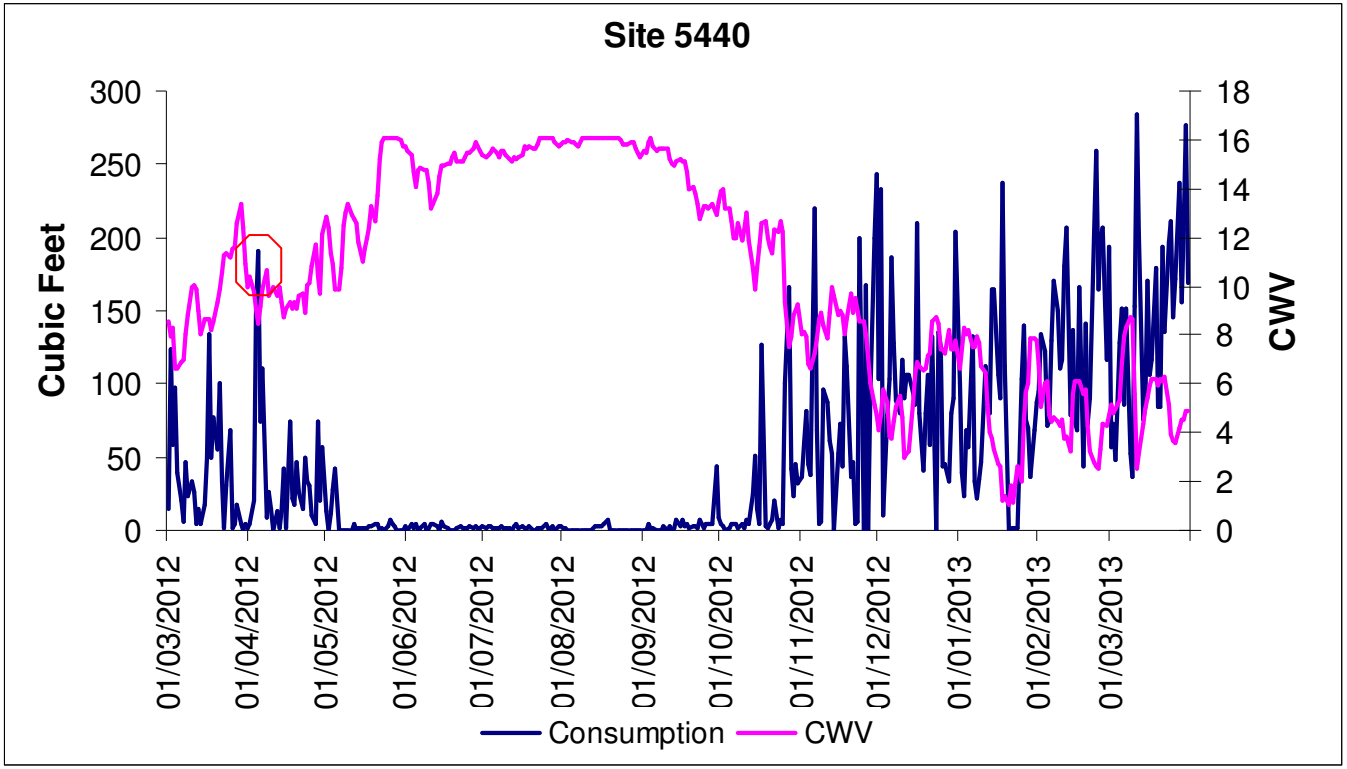
This is a Band 2 non domestic site.

The spike appears on the 04/04/2012

xoserve

respect ❯ commitment ❯ teamwork

Analysis



The spike on this chart appears to be genuine consumption.

This is a Band 1 domestic site.

The spike appears on the 05/04/2012

xoserve

respect › commitment › teamwork

Analysis

Site ID	EUC	AVG_CONSUMPTION APRTOSEP	AVG_CONSUMPTION MARTOSEP	Difference	RATIO APRTOSEP	RATIO MARTOSEP
8940	2	257.61	354.43	96.82	15.0381	10.9304
1573	2	313.85	394.41	80.56	15.1568	12.0612
8400	1	3.35	4.2	0.85	15.5237	12.3782
7708	1	8.28	11.39	3.11	16.5376	12.0254
8004	2	356.45	545.47	189.02	16.7766	12.228
7945	2	343.15	497.96	154.81	17.8026	12.268
2048	2	226.41	285.72	59.31	17.8305	14.1292
4691	1	23.85	37.06	13.21	19.709	12.6819
5440	1	8.61	12.8	4.19	22.1924	14.9175
411	2	197.16	478.06	280.9	22.4989	11.5341

The sites highlighted are the ones displayed in the examples. 1573 is the error and 5440 is the genuine consumption. (All other sites appeared to be genuine consumption too).

Further Analysis

- Regression models were run with the sites that passed validation for bands 1 and 2 (using a 12 month period) to replicate the current modelling process results to use as a benchmark.

- Regression models were then run using a 13 month period to compare against the benchmark results.

- The results are as follows:

Comparison of R^2

LDZ	Band 1 12 month 15:1	Band 1 13 month 15:1	Band 2 12 month 15:1	Band 2 13 month 15:1
EA	98.8%	98.8%	98.4%	98.4%
EM	98.9%	98.9%	98.1%	98.1%
NE	98.2%	98.2%	98.1%	98.1%
NO	98.0%	98.0%	98.0%	97.9%
NT	98.9%	98.9%	98.6%	98.6%
NW/(WN Band 2)	98.7%	98.7%	98.4%	98.4%
SC	97.9%	97.9%	98.4%	98.4%
SE	98.9%	98.9%	98.2%	98.3%
SO	98.6%	98.6%	98.7%	98.7%
SW	98.6%	98.6%	97.8%	97.8%
WM	99.0%	99.0%	98.6%	98.7%
WS	97.6%	97.6%	97.5%	97.7%

Decreasing the Spike Ratio

- We then investigated the option of decreasing the summer spike ratio to analyse if any of the sites that were passing validation happened to be errors (see the following slide).
- By decreasing the ratio to 14 you would dismiss 4 sites from the modelling process
  - From previous analysis we know that sites 2048 and 5440 are displaying genuine consumption
  - From further analysis it also appears that 7458 and 726 are genuine.
- If we decreased the ratio to 12 we would dismiss:
  - 4 sites that would have originally passed validation (Apr-Sep)
  - 5 sites with genuine consumption
  - 1 error

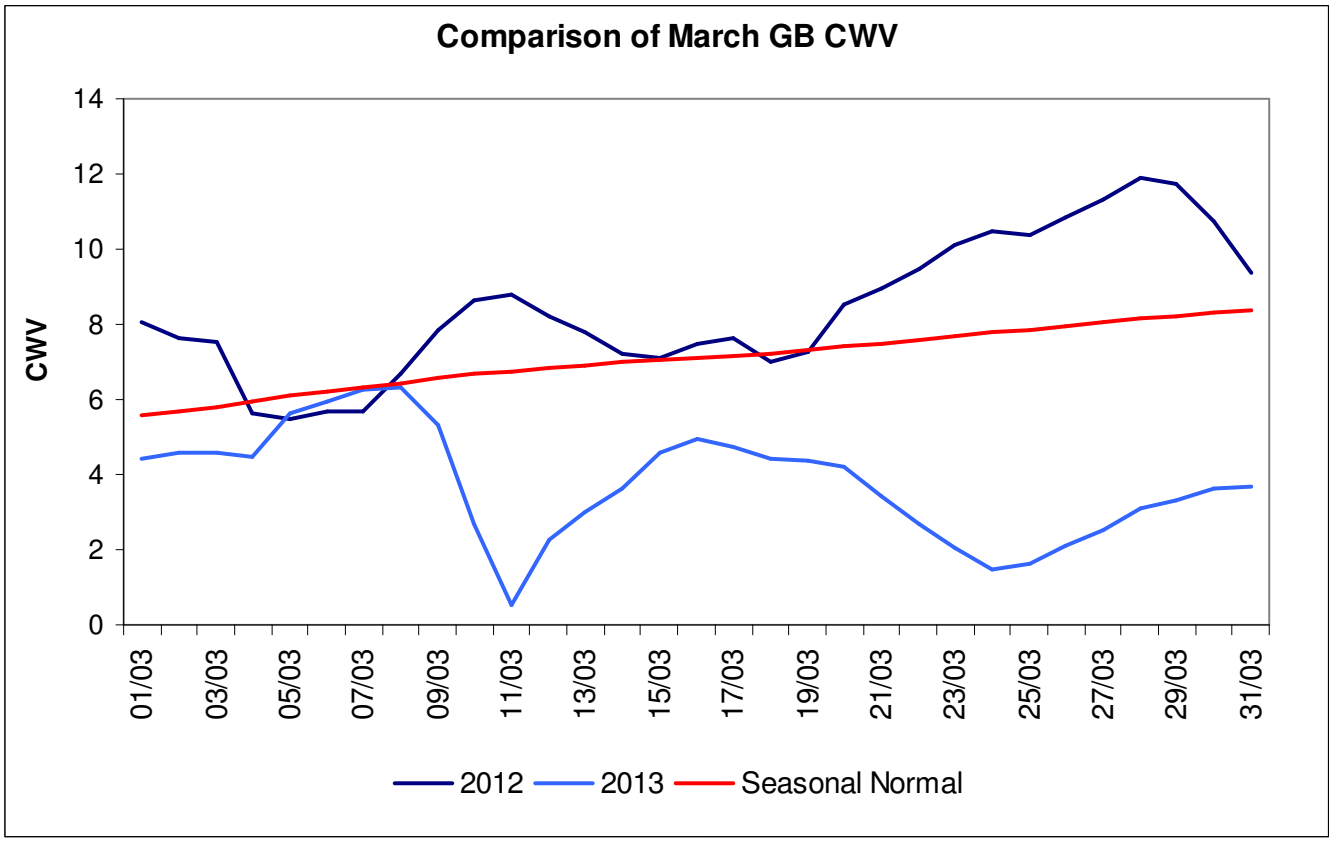
Site ID	EUC	AVG_CONSUMPTION APRTOSEP	AVG_CONSUMPTION MARTOSEP	Difference	RATIO APRTOSEP	RATIO MARTOSEP
7708		8.28	11.39	3.11	16.5376	12.0254
1573		313.85	394.41	80.56	15.1568	12.0612
8004		356.45	545.47	189.02	16.7766	12.228
7748		14.65	15.86	1.21	13.2421	12.2286
7605		10.78	16.96	6.18	13.2703	12.2623
7945		343.15	497.96	154.81	17.8026	12.268
8400		3.35	4.2	0.85	15.5237	12.3782
9078		55.49	60.06	4.57	13.5335	12.5049
1397		12.69	14.31	1.62	14.1072	12.5101
4691		23.85	37.06	13.21	19.709	12.6819
2048		226.41	285.72	59.31	17.8305	14.1292
7458		12.33	12.31	-0.02	14.4324	14.4617
726		185.57	264.3	78.73	4.0954	14.7938
5440		8.61	12.8	4.19	22.1924	14.9175

Further Considerations and Recommendation

- The analysis period has not covered the period of Mar ’13, which is when we experienced colder than normal temperatures (see following slide) – which will increase the average consumption over the summer period when calculating the ratio and **potentially** allowing errors to pass through validation.

- At this present time there does not appear to be any valid reason to decrease the spike ratio from 15.

Comparison of March Temperatures



	Average CWV	Max CWV	Min CWV
Mar-12	8.4	11.9	5.5
Mar-13	3.8	6.3	0.5

xoserve

respect › commitment › teamwork

AMR Validation Analysis (Mar – Sep 2013)

Updated 24th March 2014

TWG - 25th March 2014

respect ❯ commitment ❯ teamwork

Additional Analysis

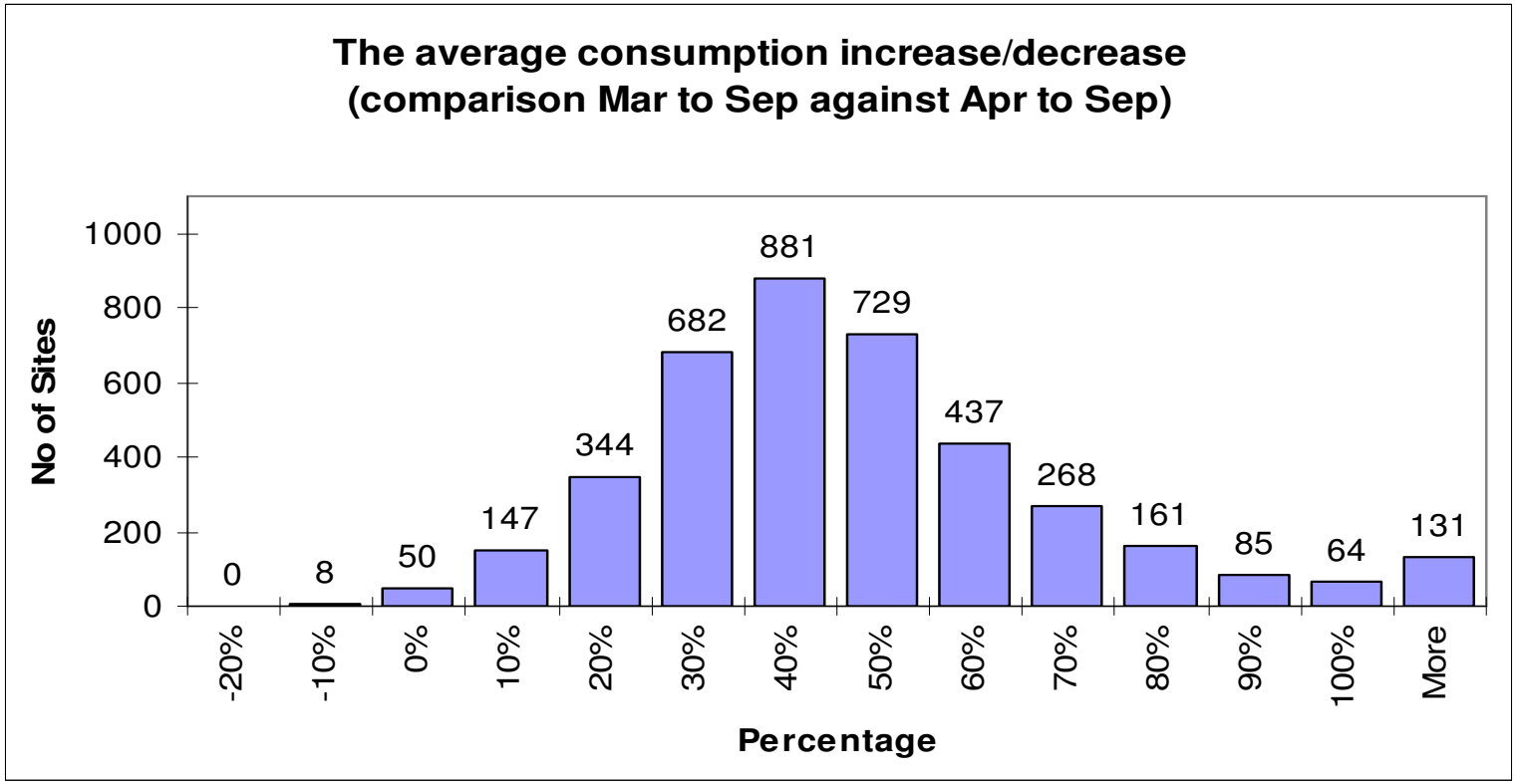
- Some ‘quick’ analysis has been done for the purpose of this meeting to assess the impact of having an additional month in the summer period.
- We have used the actual consumption data from 01/03/2013 – 30/09/2013
- 3988 sites (not all would pass validation – but this cannot be assessed fully until we have the data for the whole 13 months)
  - Apr to Sep ’13
    - 66 sites with a ratio above 15
  - Mar to Sep ‘13
    - 24 sites with a ratio above 15

respect › commitment › teamwork

Additional Analysis

- Apr to Sep ’13
  - 66 sites from this period failed the summer validation rules – 46 of those sites would pass in the Mar to Sep ’13 validation period.
  - From those 46 sites, we know that at least 13 sites would fail the winter validation, missing days or consecutive zeros rule.
  - At the most 33 erroneous sites will go through validation. Further analysis will be carried out on these sites to verify if they look like errors or are showing signs of being genuine consumption
- Mar to Sep ’13
  - 24 sites from this period failed the summer validation rules
  - From those 24 sites, 4 would have originally passed the Apr-Sep summer validation

Additional Analysis



This slide shows the average consumption increase by site – when adding in March 2013 into the summer period.

We can see that the majority of sites have an increased consumption by around 40%