



Proficiency Testing: LGC Schemes and participant performance in the analysis of environmental samples

Dr Matthew Whetton, LGC Standards, Proficiency Testing

Introduction

- LGC Standards, Proficiency testing
- What is proficiency testing?
- Assessment of performance
- LGC Standards, PT schemes (CONTEST and Aquacheck)
- Participant performance in the CONTEST scheme

LGC Standards, Proficiency Testing



- Part of the LGC Group, in the 'Standards' division
- Operating a total of 39 PT schemes
- More than 7,000 active participants
- 52 employees
- More than 100,000 samples produced and distributed to more than 140 countries per annum
- Approximately 3 million data points assessed each year

History of LGCPT

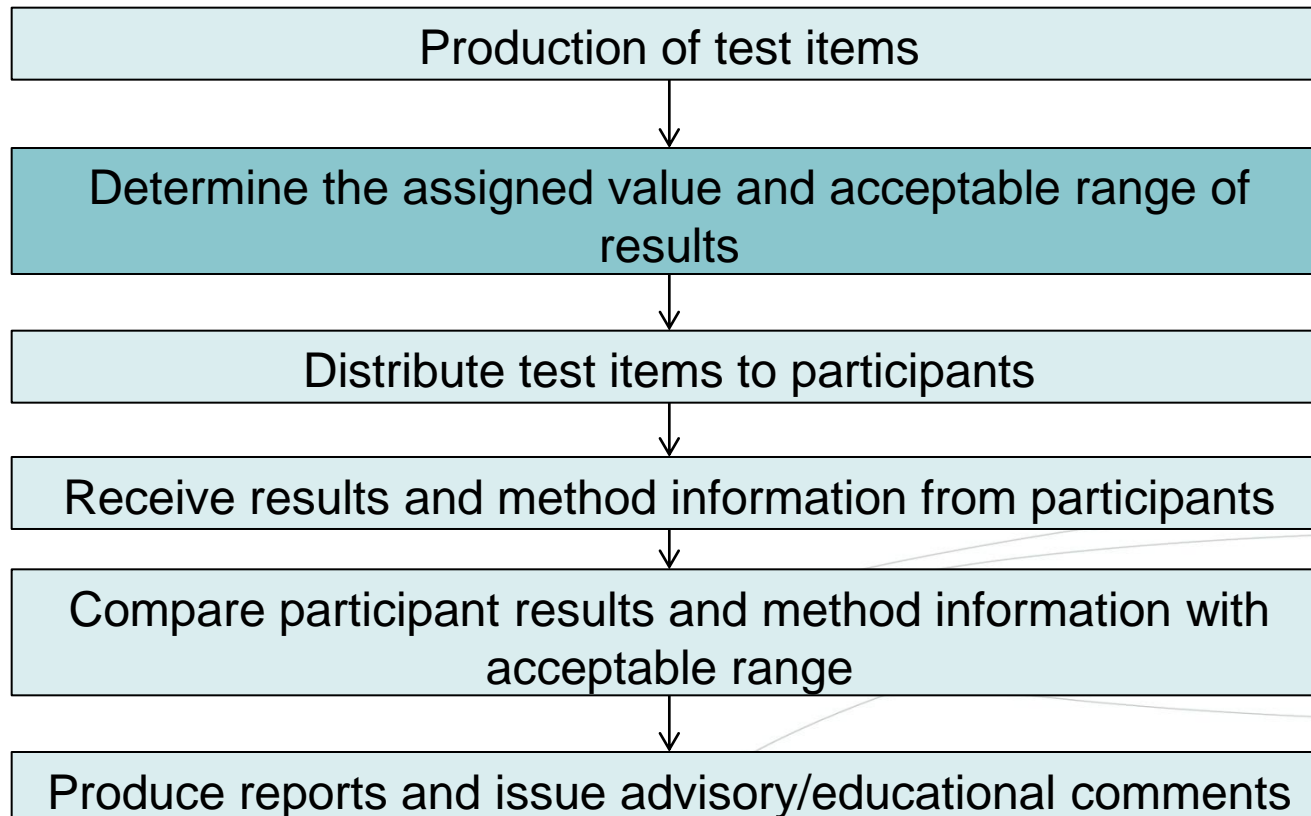
- Regular PT started in clinical chemistry late 1940's
- LGC started their first scheme in 1984
- Aquacheck scheme started by WRc in 1985
- Quality Management food microbiology scheme (QMS) started 1991
- LGC Standards Division created 2007

What is Proficiency Testing?

- PT provides an independent means of testing and comparing individual laboratory/technician results of analysis with a peer group.
- Proficiency Testing (PT) involves the regular distribution of stable homogenous test materials to participants for the analysis of defined parameters. Participants results of analysis are then compared with those of the other participants.

Scheme structure

- According to the international standard ISO/IEC 17043
- Proficiency Testing (PT) consists of:



Performance assessment

- Proficiency testing statistics are based on ISO 13528
- Most PT schemes use an objective performance score
 - The most common is the Z score

$$Z \text{ score} = \frac{x - X}{SDPA}$$

- Where x = the participant result
 X = the assigned value or 'true' result
SDPA = standard deviation for proficiency assessment

Performance assessment

- Setting the Assigned value
 - From formulation
 - From a reference value
 - Relative to a CRM
 - Consensus of expert laboratories
 - Consensus of participant results
- Defining the SDPA
 - Prescribed value (Legislation etc.)
 - Perception (What labs/provider would like to achieve)
 - From a general model (Horwitz)
 - From a precision experiment
 - From data obtained from a round of proficiency testing

Performance Scoring

- The assessment of performance is based on:
 - Satisfactory $|Z \text{ score}| \leq 2$
 - Questionable $2 < |Z \text{ score}| < 3$
 - Unsatisfactory $|Z \text{ score}| \geq 3$
- Statistical methods for the assessment of proficiency testing assume a normal distribution of results

Performance scoring

Sample: 3b - Group B Soil

Analyte: Loss on Ignition

Lab ID	Method	Result (%)	z/z' score*
CN0002	450°C	4.71	-1.04
CN0005	450°C	4.89	-0.71
CN0009	450°C	5.77	0.86
CN0009	550°C	6.73	-0.01
CN0013	450°C	5.29	0.00
CN0013	550°C	6.47	-0.40
CN0020	450°C	5.31	0.04
CN0022	450°C	4.91	-0.68
CN0025	450°C	4.84	-0.80
CN0035	450°C	5.20	-0.16
CN0036	450°C	5.86	1.02
CN0044	450°C	5.26	-0.05
CN0057	550°C	6.47	-0.40
CN0062	450°C	5.48	0.34
CN0077	450°C	5.99	1.25
CN0093	550°C	6.74	0.00
CN0109	450°C	2.25	-5.43
CN0109	550°C	3.09	-5.45
CN0112	450°C	4.49	-1.43
CN0123	450°C	5.60	0.55
CN0129	450°C	4.78	-0.91
CN0133	450°C	5.29	0.00
CN0165	550°C	7.00	0.39
CN0174	550°C	7.40	0.99
CN0177	550°C	6.90	0.24

Performance scoring

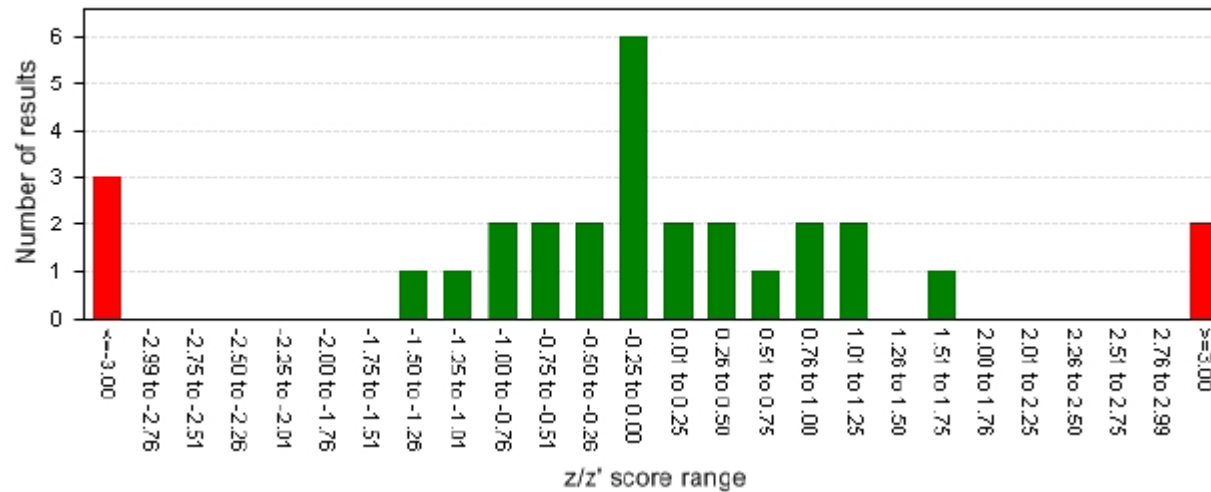
Data Statistics

	450°C	550°C
Number of Results	18	11
Number of Excluded Results	1	2
Mean	5.29	6.78
Median	5.29	6.74
Standard Deviation	0.48	1.23
Robust Standard Deviation	0.59	0.40
Result Range	4.49 to 6.20	4.27 to 9.07

Performance Statistics

	450°C	550°C
Assigned Value	5.29	6.74
Uncertainty of Assigned Value	0.18	0.17
SDPA	0.53	0.67
Satisfactory Range	4.23 to 6.35	5.40 to 8.08
Satisfactory z/z' scores	94.4%	63.6%
Questionable z/z' scores	0.0%	0.0%
Unsatisfactory z/z' scores	5.6%	36.4%

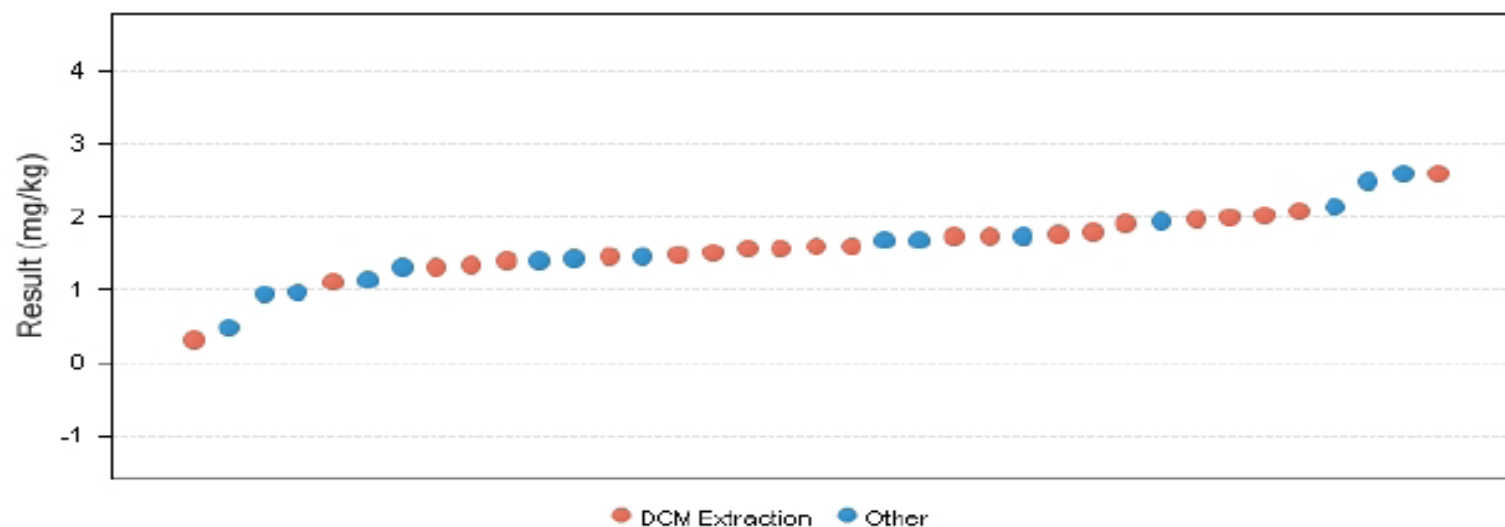
z/z' score Histogram



Assessment of method performance

Analyte: Acenaphthene

Distribution Graph



Methodology Summary

Method	Number of Results	Excluded Results	% of Total	Median	Robust SD	Range	Sat.
				mg/kg			%
DCM Extraction	22	0	59.46	1.61	0.30	0.31 to 2.61	90.9
Other	15	0	40.54	1.47	0.47	0.49 to 2.60	73.3
All	37	0	100	1.61	0.37	0.31 to 2.61	83.8

How can I tell a 'Good' laboratory

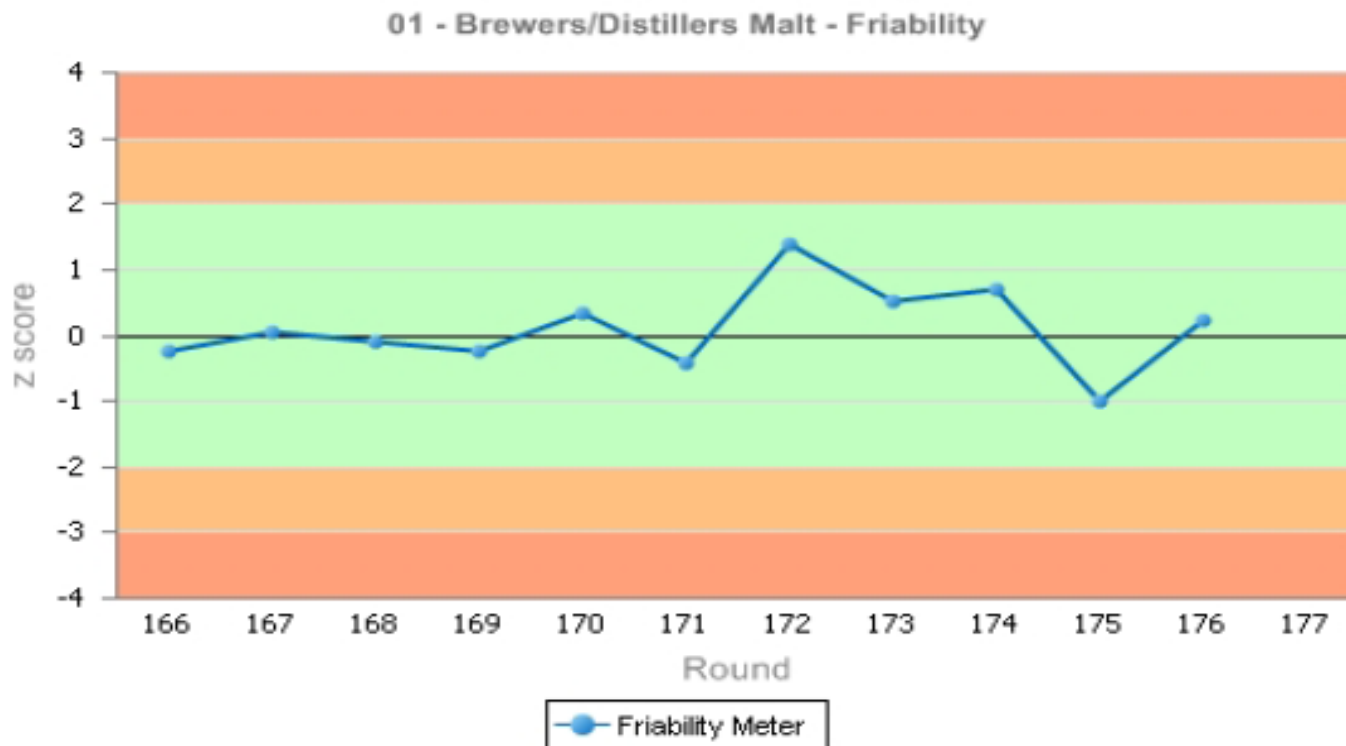
- Participation in PT? A good starting point!
- Other questions you may ask:
 - How frequently do they participate?
 - How broad is their participation?
 - How is their participation planned?
 - Do the PT samples match regular work?
 - How do they assess their own performance?

Assessing performance

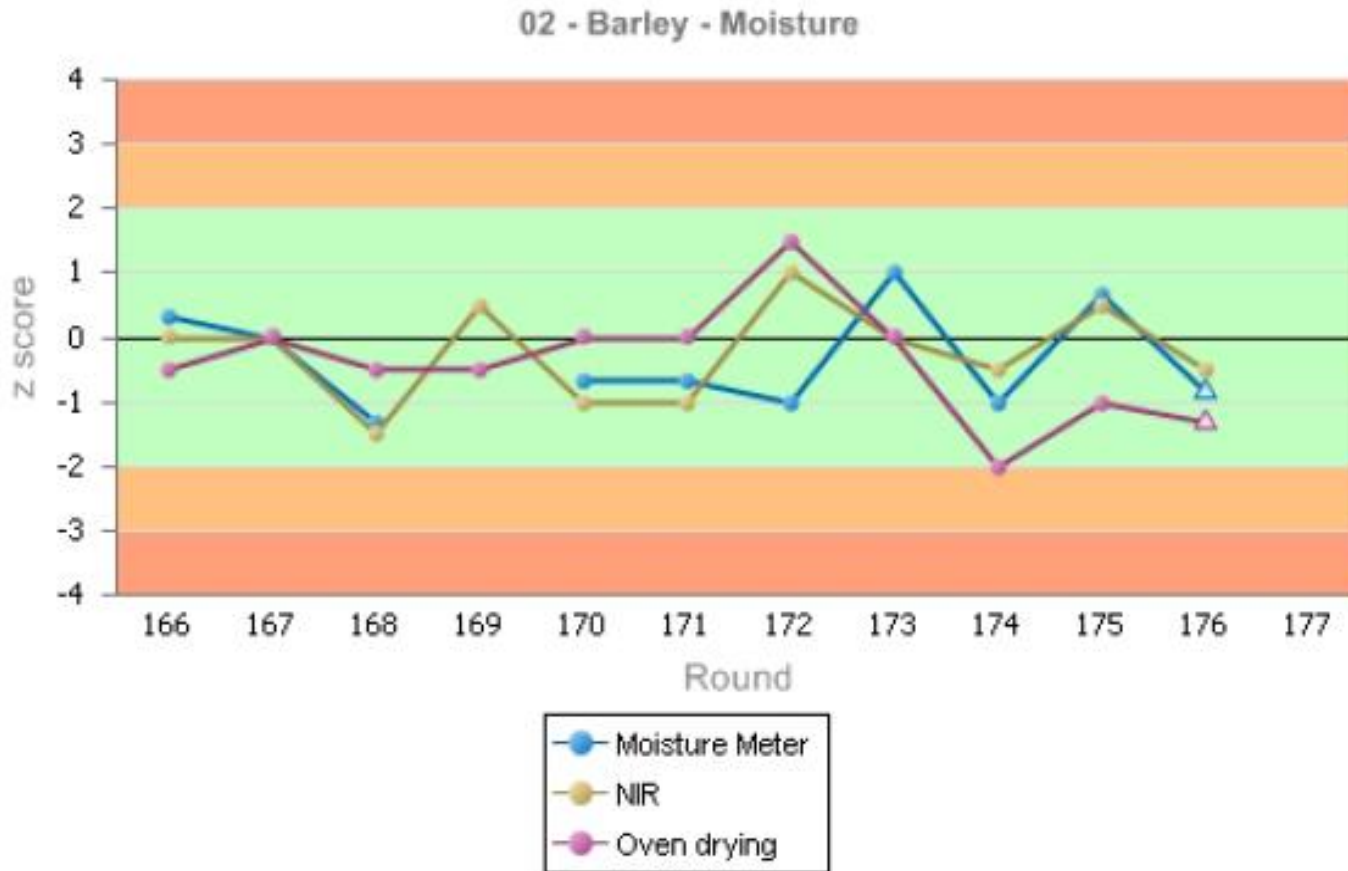
- Good performance
 - A good z score for one round a year?
 - A long term trend which meets specific criteria?
 - A z score less than 3, 2, 1?
 - Something else?

Monitoring performance

- ISO/IEC 17043 recommends graphical methods of monitoring performance over time



Monitoring performance

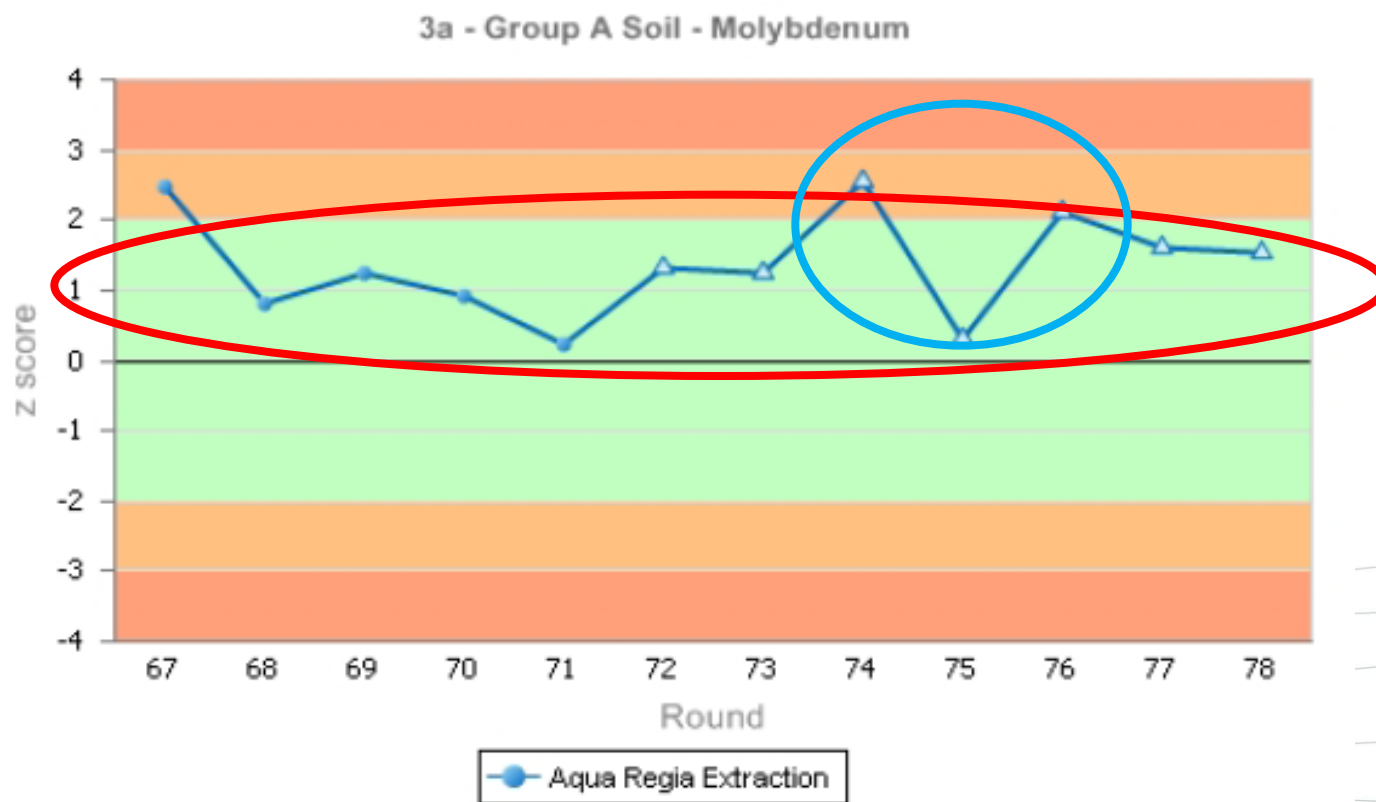


Monitoring performance: A simple approach

- ‘Satisfactory’ performance is easily interpreted for a single round ($z \leq |2|$)
- Over a longer period, standard Shewart chart rules can be applied:
 - Single result $>|3|$
 - Nine on the same side of the centre line
 - Six results increasing or decreasing
 - Fourteen alternating up and down
 - Two out of three greater than 2 (+ve or -ve)
 - Four out of five results $>|1|$

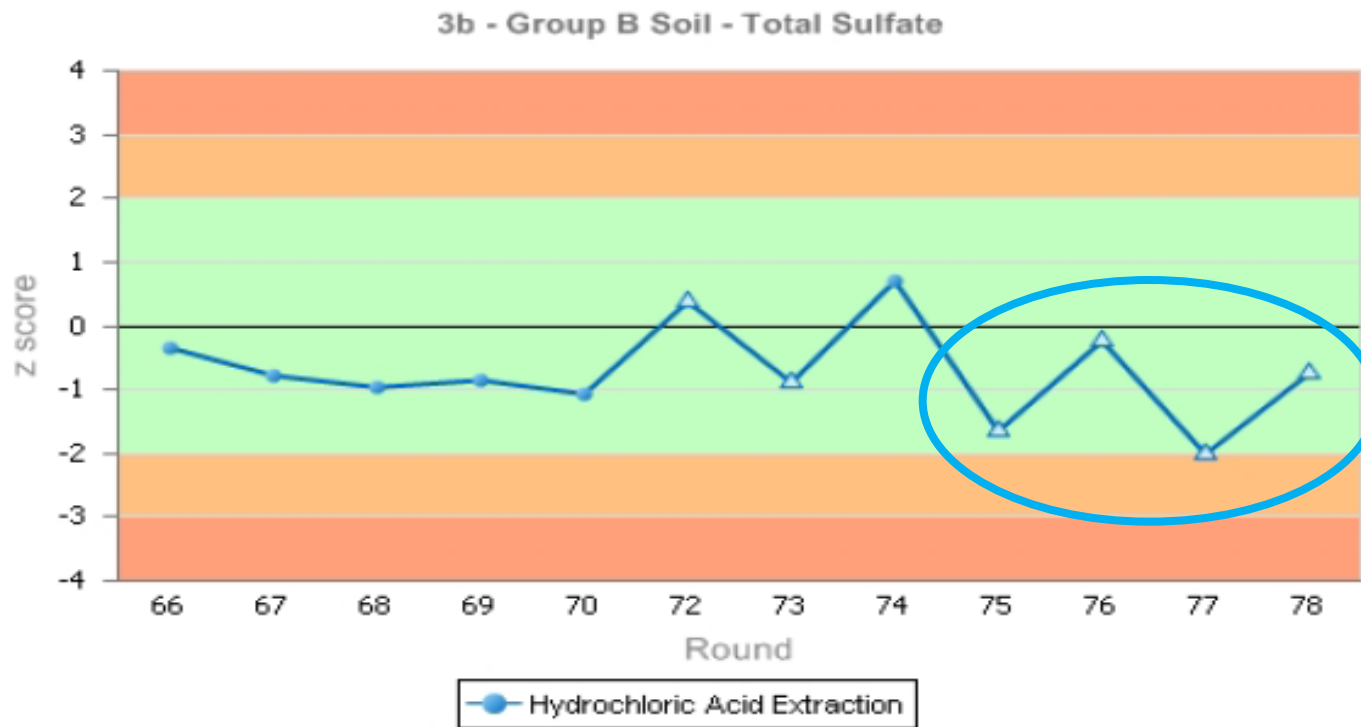
Monitoring performance

- Identifying long and short term bias



Monitoring performance

- Early warning of loss of method control



'Bad' performance or (Something else?)

- Questionable, $|2|$ to $|3|$, is not wrong
- However, ≥ 3 is statistically very unlikely

- When you receive a unsatisfactory z score
 - Check for simple errors
 - Investigate the root cause
 - Implement corrective/preventative actions
 - Check the effectiveness of the actions
 - More PT!

STILL BAD PERFORMANCE ???

LGC Environmental Schemes

- Aquacheck – Chemical analysis of water, wastewater, soil and sludge
- Quality in Water Analysis Scheme (QWAS)
- CONTEST – Analysis of contaminated land
- STACKS – Analysis of solutions and materials from Stack Emissions



The CONTEST Proficiency Scheme

- Four groups
 - Elements, inorganics, organics, leaching tests
- Soils and standard solutions
- Frequency based on sample group
 - Up to five times per year to enable performance trending
- Over 100 registered participants
- Located in >30 countries
 - The majority of labs are from UK and Europe

CONTEST: Scheme structure

- A range of important matrices are provided for 2011/12

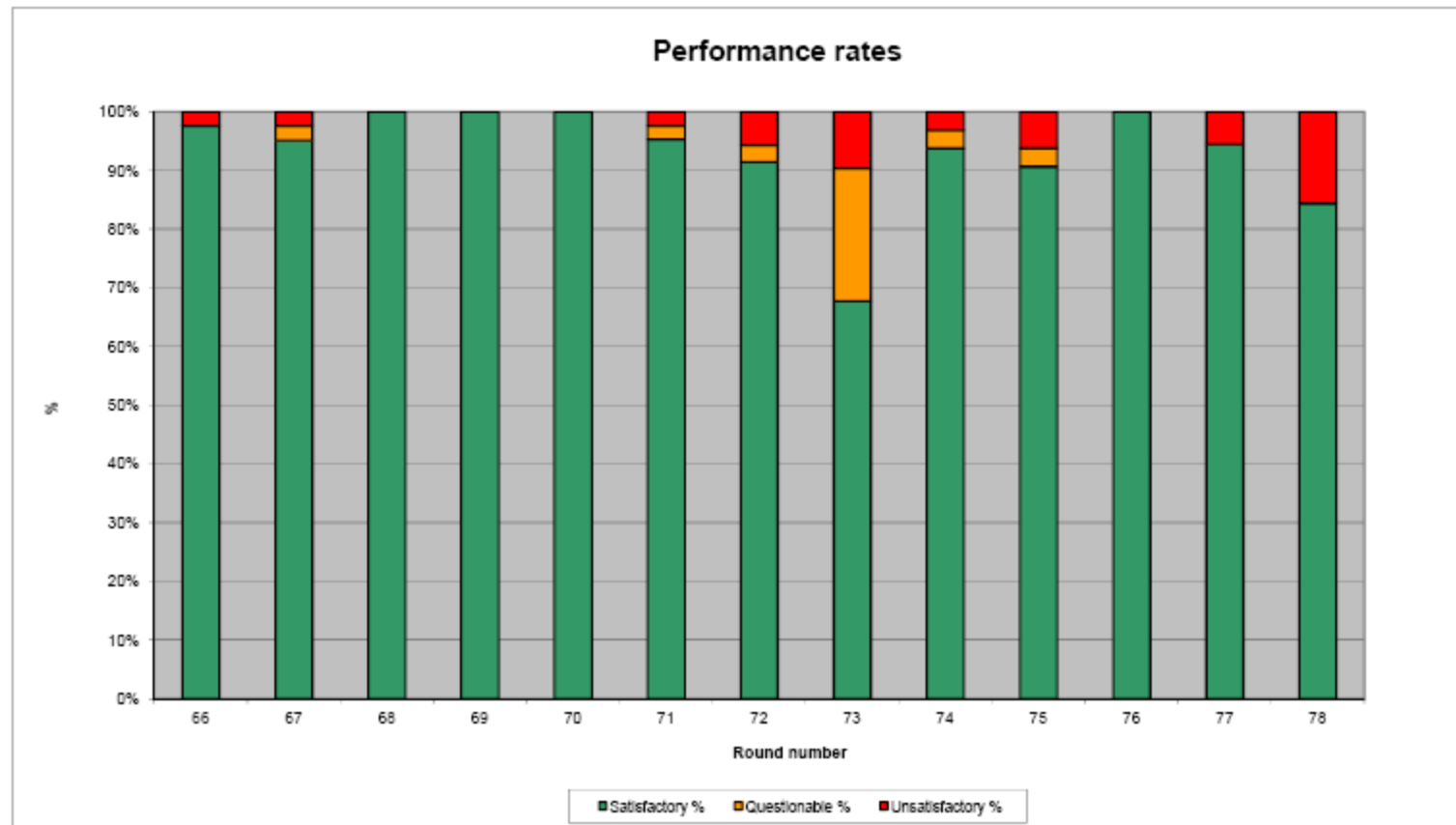
Round	Despatch Date	Reporting Deadline	Level of Participation	Groups					
				A Metals	B Inorganics	C Organics	D Leachate	E WAC	TRIAL*
76	26 Apr 2011	16 May 2011	Full						
			Reduced (Soils)						
77	13 Jun 2011	04 Jul 2011	Full						
			Reduced (Soils)						
78	22 Aug 2011	12 Sep 2011	Full						
			Reduced (Soils)						
79	31 Oct 2011	21 Nov 2011	Full						
			Reduced (Soils)						
80	23 Jan 2012	13 Feb 2012	Full						
			Reduced (Soils)						

The Aquacheck Proficiency Scheme

- Four 'groups'
 - Clean/Waste inorganics, clean/waste organics
- More than 30 different samples, including agricultural soil, sludge and radiochemistry
- Frequency based on sample group
 - Up to four times per year to enable performance trending
- Over 1000 registered participants
- Located in >50 countries
 - The majority of labs are from UK and Europe

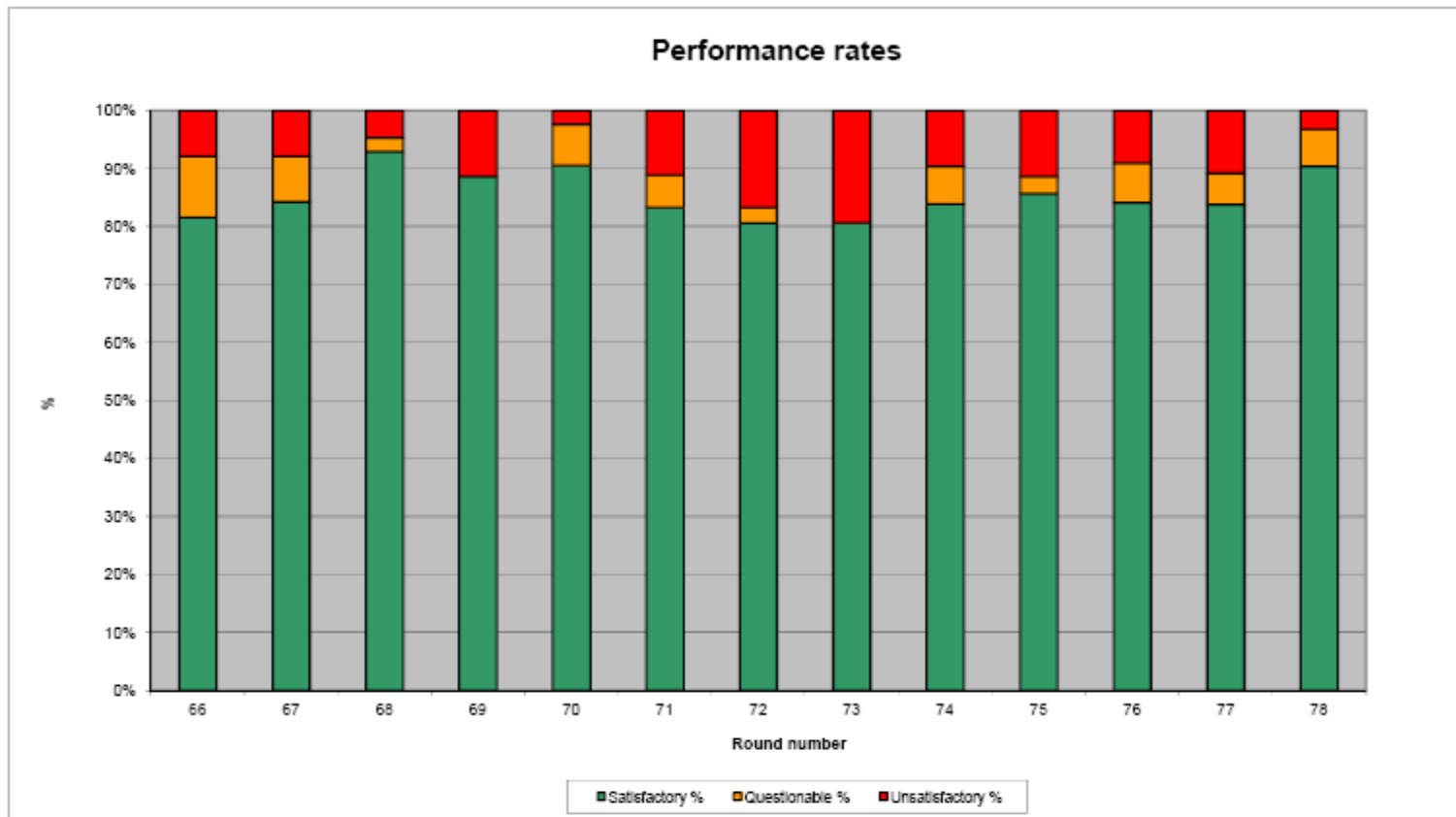
CONTEST: Analysis of Performance

- Percentage of participants receiving satisfactory z scores - Arsenic



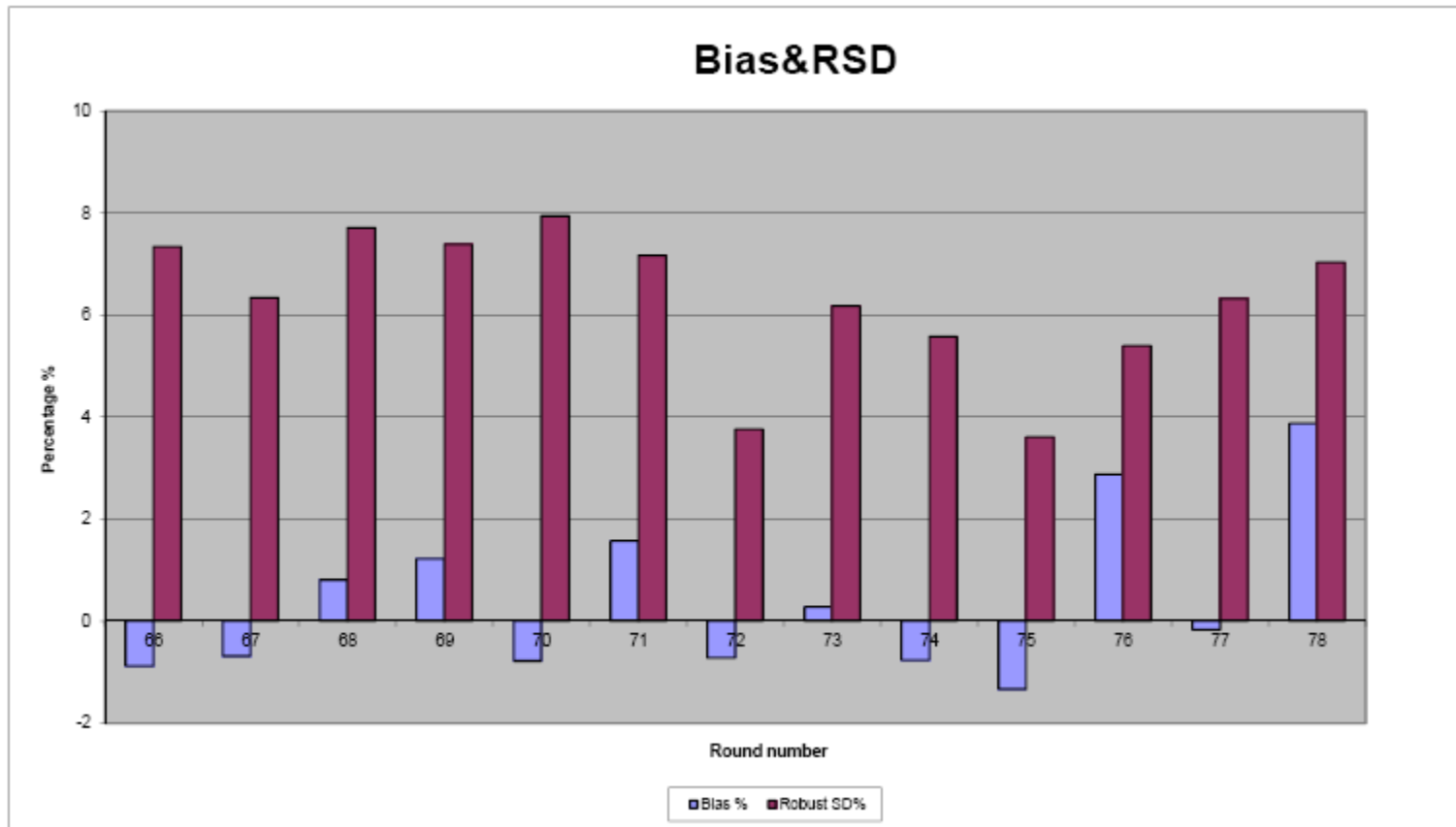
CONTEST: Analysis of Performance

- Percentage of participants receiving satisfactory z scores - Acenaphthene

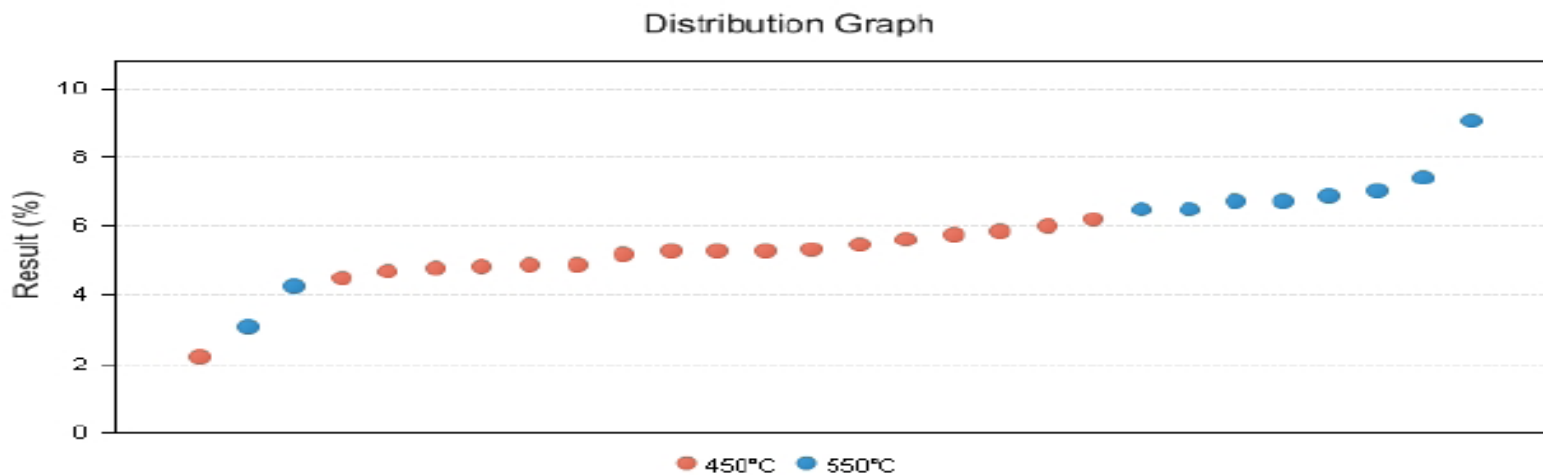
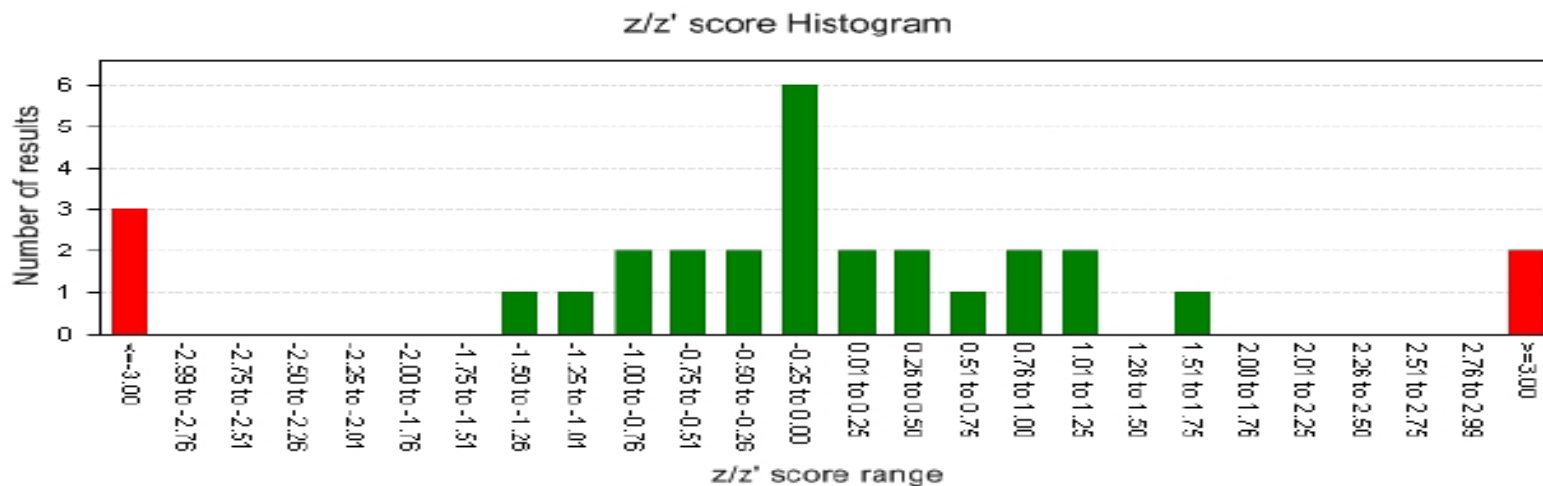


CONTEST: Analysis of Performance

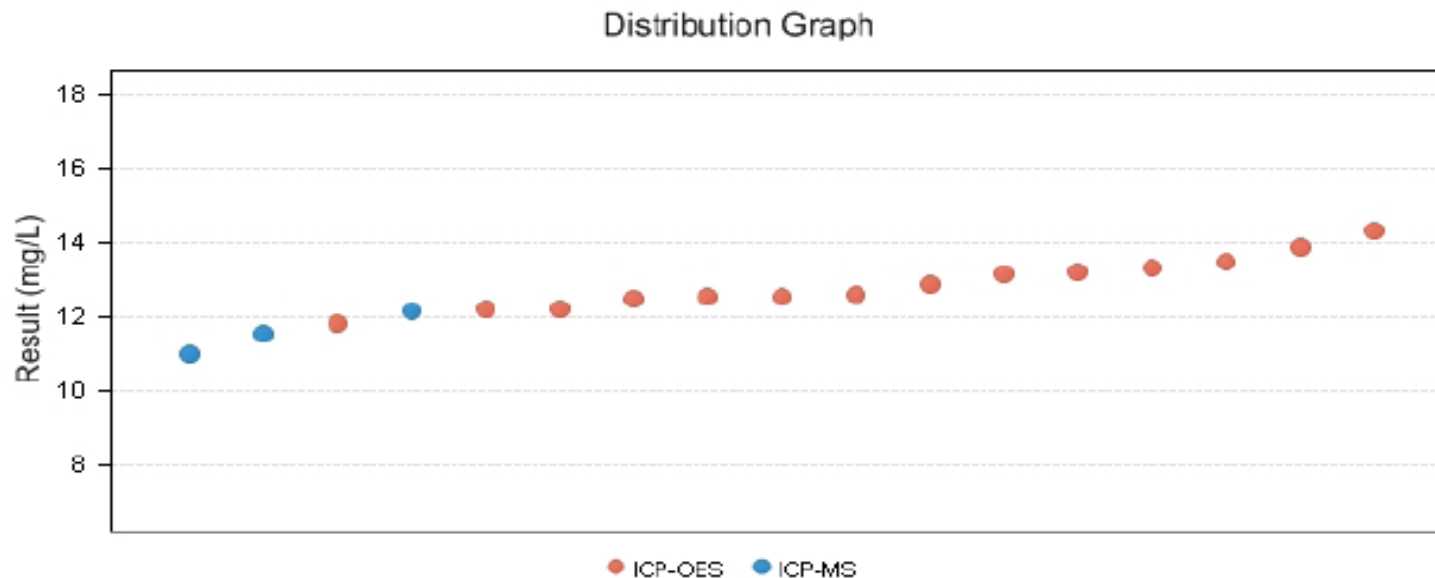
- Bias and robust standard deviation for Cr (VI) analysis



Performance comparison of commonly used methods



Performance comparison of commonly used methods



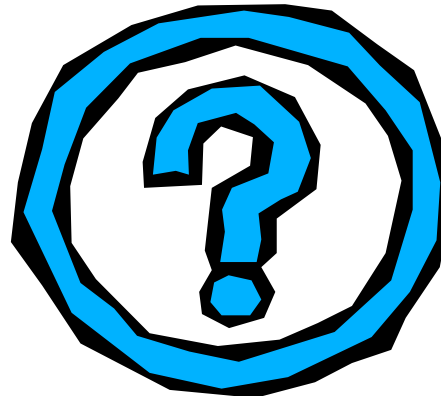
Methodology Summary

Method	Number of Results	Excluded Results	% of Total	Median	Robust SD	Range	Sat.
				mg/L			%
ICP-OES	14	0	82.35	12.74	0.73	11.84 to 14.32	85.7
ICP-MS	3	0	17.65	11.57	0.85	11.00 to 12.16	66.7
All	17	0	100	12.56	0.87	11.00 to 14.32	82.4

Summary

- PT is a very important and versatile QC tool
- Monitoring laboratory performance in PT can:
 - provide peer comparison
 - identify current or future analytical problems
 - be a mechanism for driving laboratory improvement
 - evaluate the performance of different methods

Thank-you for listening



Any questions!