




SUMMARISED REPORT OF THE EGOLF ROUND-ROBIN NR. TC2 09-1 IN FIRE RESISTANCE TESTING

Prepared by
Fabien Dumont
Quality Manager
Fire testing laboratory
University of Liège

on behalf of EGOLF

CONTENTS

SCOPE	2
PART 1 – DRAWING UP OF THE PROJECT	2
1-1 ORGANIZATION	2
1-2 SCOPE OF THE ROUND-ROBIN	2
1-2.1 Test method	2
1-2.2 Test specimen	2
1-2.3 Scheme of the experience	3
1-2.4 Laboratories experience	3
PART 2 – DATA PROCESSING AND REPORTING	4
2-1 ACCURACY EVALUATION	4
2-1.1 Starting data	4
2-1.2 Repeatability	4
2-1.3 Selection of regular data	4
2-1.4 Final accuracy results	5
2-2 PERFORMANCE EVALUATION	6
2-2.1 Starting data	6
2-1.1 Ranks	8
2-1.2 Z-scores	9
2-1.3 K-scores	10
2-3 VALIDITY PERIOD OF THE RESULTS	10

	Egolf TC2 Fire resistance Round Robin <i>Summarised report</i>	
--	---	---

SCOPE

In 2009 the EGOLF organization conducted a round robin (RR) on resistance to fire tests according to EN 1364-1 with 32 participating labs. This round robin enables the participating laboratories to demonstrate their ability to obtain regular results, to express their trueness and precision and to calculate their uncertainty of measurement.

The present document is a summary of the full report (TC2 09-1 round-robin) from the round robin.

PART 1 – DRAWING UP OF THE PROJECT

1-1 ORGANIZATION

A steering group was formed consisting of 4 laboratories, each represented by one expert familiarized with the standard test method and its applications.

Their main responsibilities were:

1. managing the project (test procedure, test specimens, results templates)
2. processing the test results (data recording, statistical analysis)
3. reporting (edition of a full report, presentation of the conclusions at EGOLF meeting)

1-2 SCOPE OF THE ROUND-ROBIN

1-2.1 TEST METHOD

The test method chosen for the RR was the determination of the fire resistance in accordance with EN 1364-1, EN 1363-1 and EN 1363-2. If applicable, EGOLF technical resolutions and recommendations should be applied. The fire resistance performances to be measured were integrity and insulation. All other parameters were also measured as in accredited tests (deflection, pressure, oxygen rate, ...).

The participating laboratories were required to strictly follow all the requirements from those standards from receiving the test specimens, conditioning the materials, conducting the test and finally to stating the obtained results as if the tests were just another accredited test ordered by a client.

1-2.2 TEST SPECIMEN

The test specimen consisted of 3000 mm x 3000 mm gypsum plasterboard partition with a steel stud frame and one gypsum plasterboard layer on each side of the frame. The test specimen incorporated a free edge in one of the vertical sides, and a horizontal board joint running the full width of the test specimen (see drawing 1). This test specimen was chosen in accordance with the scope of the chosen test standard and is representative of typical test specimens tested according to EN 1364-1.

In order to avoid that heterogeneity of the test specimen became a main component of the accuracy value, attention was paid to reduce the number of material types in the test specimen, to use materials from the same batch (the sampling had been done by a employee of one of the steering group labs) and to produce detailed drawings and an installation manual for the test specimen.



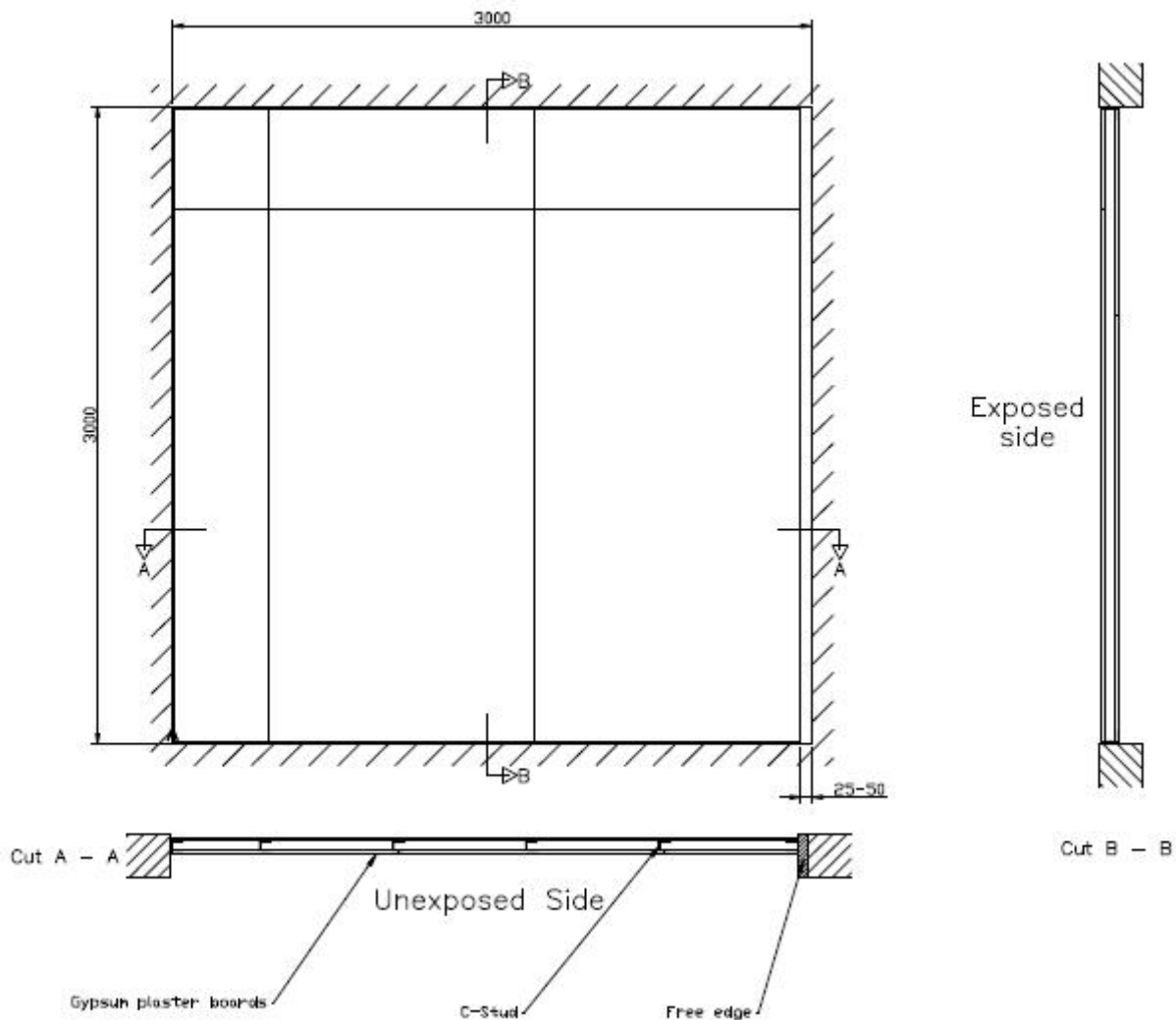
1-2.3 SCHEME OF THE EXPERIENCE

32 laboratories participated in the RR. It can be assumed that the number of laboratories participating in the RR is large enough to be a reasonable cross-section of the population of qualified laboratories since approximately no more than 50 laboratories conducts tests according to EN 1364-1.

Two tests were conducted in each lab on identical replicates of the test specimen, except two labs who preferred to conduct one test instead of two.


Each laboratory was requested to conduct the tests under repeatability conditions.

Drawing 1
 (Test specimen seen from unexposed side)



1-2.4 LABORATORIES EXPERIENCE

All participating laboratories are EGOLF members. The majority are ISO 17025 accredited and have attended the EGOLF harmonization course on EN 1363-1 / EN 1363-2 as exemplified by EN 1364-1.

	Egolf TC2 Fire resistance Round Robin <i>Summarised report</i>	
--	---	---

PART 2 – DATA PROCESSING AND REPORTING

Statistical analysis was processed on the raw data for integrity and insulation as they've been submitted by the participants through the provided forms.

Five tests were stopped before the occurrence of any mode of failure for integrity or insulation, resulting in missing data. One lab has presented its data in such a way that they turn out to be unprocessable regarding the aims of this RR.

2-1 ACCURACY EVALUATION

The test specimen, the laboratories, the number of replicates, the instructions and the protocol of this experiment have been chosen to fully comply with the ISO 5725 prescriptions. As a result, the data processing tools presented in the ISO 5725 could also be implemented.

The first aim of the analysis was to work out as accurate as possible an accepted value for the fire resistance of the specimen (reference value m) and a quantitative measurement of the spread of the lab results (repeatability standard deviation s_r and reproducibility standard deviation s_R). The more correct results are included in these estimations, the more accurate these are. That's why all the necessary corrections of the data have been allowed for the purpose of this aim.

2-1.1 STARTING DATA

First of all, the starting test results were drawn up from the raw data. For each laboratory, the within-cell mean (\bar{y}_i) and within-cell standard deviation (s_i) was calculated.

2-1.2 REPEATABILITY

Repeatability conditions require both tests to be conducted within a short time interval in each laboratory in order to keep constant operating conditions. The delay between the two tests in each lab has proved to be less than one week on average. Repeatability conditions require also both tests to be conducted by the same technical operator and using the same equipment.


These requirements were checked and it can be stated that global repeatability conditions have been kept under control in each lab.

2-1.3 SELECTION OF REGULAR DATA

Since the accuracy results are estimated from the data submitted by the participating laboratories, the presence of irregular data could distort the estimates. A deep analysis was conducted on the starting data to find and discard irregular data, i.e. erroneous data and inconsistent data.

2-1.3.1 Erroneous data

Some raw results submitted by the participating laboratories needed complementary inquiries, clarifications from these labs, and sometimes corrections of the raw results. The following inspections were conducted on original data to detect deviations from standard requirements (EN 1363-1 and EN 1364-1): tolerances on the percentage deviation of the average furnace temperature, specified pressure and tolerances, tolerances on the ambient temperature, measurements and observations, deduction of performance criteria.

	Egolf TC2 Fire resistance Round Robin <i>Summarised report</i>	
--	---	---

Some original data submitted by the participating laboratories turned out to deviate from the standard requirements. When useful, the laboratories producing a deviation were consulted for possible technical error. When it proved impossible to replace the suspect item, the data was rejected as erroneous, other else the necessary corrections have been brought to the data.

The origin of errors regarding deduction of performance criteria turned out to always consist of mistakes in capturing the results from the recorded values on the spreadsheet. As a consequence, results from 11 labs were corrected. Only test results fully complying with the standard requirements were admitted at this step. As a result, the whole data of 13 tests were discarded.

2-1.3.2 Consistency tests for outlier data

Many techniques can be used for testing the consistency of the test results reported from various laboratories. The ISO 5725 recommends both graphical technique (Mandel's k and h statistics) and numerical techniques (Cochran's statistic and Grubbs' statistic).

Principles of the consistency tests

Measurement data frequently contain a proportion of extreme values. It is, however, often difficult to distinguish erroneous values from chance variations, which can also give rise to occasional extreme values. Outlier detection by consistency tests help to distinguish between chance occurrence as part of the normal population of data, and values that can not reasonably arise from random variability. So, a statistical outlier is only unlikely to arise by chance.

If an outlier is found from the consistency tests, it can be useful to consult the laboratory that produced this outlier for possible technical error. If it proves impossible to replace the suspect item, the outlier should be rejected from the study unless there is a good reason to remain it.

An outlier in the k -statistic or Cochran's test within a laboratory may be a strong indication that the laboratory's within-laboratory variance is exceptionally high, while an outlier in the h -statistic or Grubbs' test may be a strong indication that the laboratory's bias is exceptionally large.

Application of the tests

The consistency tests were applied in successive stages following the sequential protocol described in ISO 5725-2, from the starting data from which the erroneous data were discarded. Cochran's C -statistic showed that one lab was undeniably an outlier for insulation. As a consequence, this lab was rejected from the study.

2-1.4 FINAL ACCURACY RESULTS

The only accuracy results that can be calculated are estimates m , s_R and s_r characterising the true values μ , σ_R and σ_r . The expanded uncertainties of these estimates at a confidence level of 95% have also been processed. The regular selected data have been used for calculating the final accuracy results for integrity (see table 1) and insulation (see table 2).

The repeatability value r is the value below which the absolute difference between two single results obtained under repeatability conditions may be expected to lie with a probability of 95 %.

The reproducibility value R is the value below which the absolute difference between two single results obtained under reproducibility conditions may be expected to lie with a probability of 95 %.


	<p>Egolf TC2 Fire resistance</p> <p>Round Robin</p> <p>Summarised report</p>	
--	--	---

Table 1

INTEGRITY				
Number of valid labs	25			U _{95%}
General mean = Reference value	<i>m</i>	1:05:09		0:02:32
Repeatability standard deviation		<i>s_r</i>	0:03:12	0:00:53
Between-laboratory standard deviation		<i>s_L</i>	0:05:37	
Reproducibility standard deviation		<i>s_R</i>	0:06:28	0:01:37
Ratio		$\gamma = s_R/s_r$	2,0	
Repeatability limit		<i>r</i>	0:08:58	0:01:29
Reproducibility limit		<i>R</i>	0:18:07	0:02:42

Table 2

INSULATION				
Number of valid labs	26			U _{95%}
General mean = Reference value	<i>m</i>	0:58:04		0:01:55
Repeatability standard deviation		<i>s_r</i>	0:01:58	0:00:32
Between-laboratory standard deviation		<i>s_L</i>	0:04:30	
Reproducibility standard deviation		<i>s_R</i>	0:04:55	0:01:16
Ratio		$\gamma = s_R/s_r$	2,5	
Repeatability limit		<i>r</i>	0:05:31	0:00:54
Reproducibility limit		<i>R</i>	0:13:46	0:02:06

Note that the first principle of ISO/TS 21748 states that the reproducibility standard deviation obtained under a proficiency testing program represents a valid basis for the general measurement standard uncertainty estimation (i.e. the average standard uncertainty that can be expected during any test conducted by any lab).

2-2 PERFORMANCE EVALUATION

In this part 2, individual accuracy results have been produced from the raw data submitted by the laboratories. These individual accuracy results consist of the bias and the standard deviation for each lab. The performance of each lab could then be deduced by comparison with the general accuracy results produced here above.

This second part implements some simple graphical and numerical criteria to these labs' individual accuracy results. Those methods, presented in the ISO 13528, allow deducing a clear picture of the performances of the laboratories.

2-2.1 STARTING DATA

For each laboratory, the within-cell mean (\bar{y}_i), bias ($\Delta_i = \bar{y}_i - m$) and within-cell standard deviation (s_i) have been calculated from the original test results reported from participating laboratories. Contrary to what has been done in accuracy evaluation, the performance assessment of the labs must be based on the original results submitted by the labs. So, no correction is allowed.

Egolf TC2 Fire resistance

Round Robin

Summarised report



Below is a graphical representation (histogram) of the density distribution of integrity and insulation stated in completed minutes, as requested in EN 1363-1 (see charts 1 and 2).

The closer to the reference value the within-laboratory mean is, the smaller the lab's bias is, the better the lab's trueness is.

The smaller the within-laboratory standard deviation is, the smaller the lab's variability is, the better the lab's precision is. A comparison value is the repeatability standard deviation (s_r) given in tables 1 and 2. Such a comparison has been done below (see k-scores).

The global accuracy of a laboratory results from these two components.

Chart 1

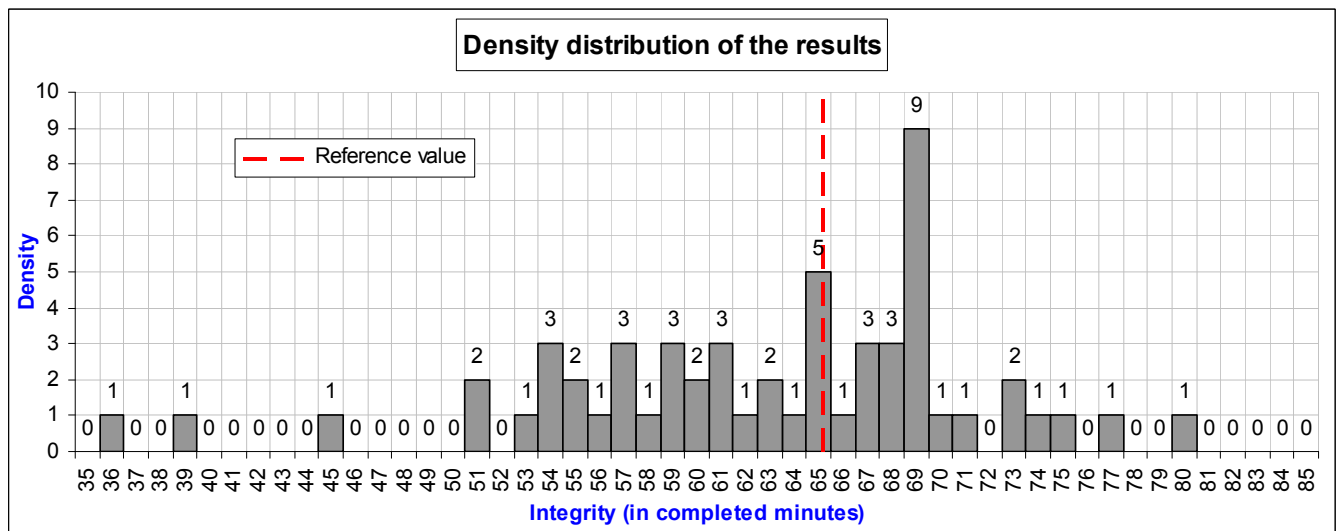
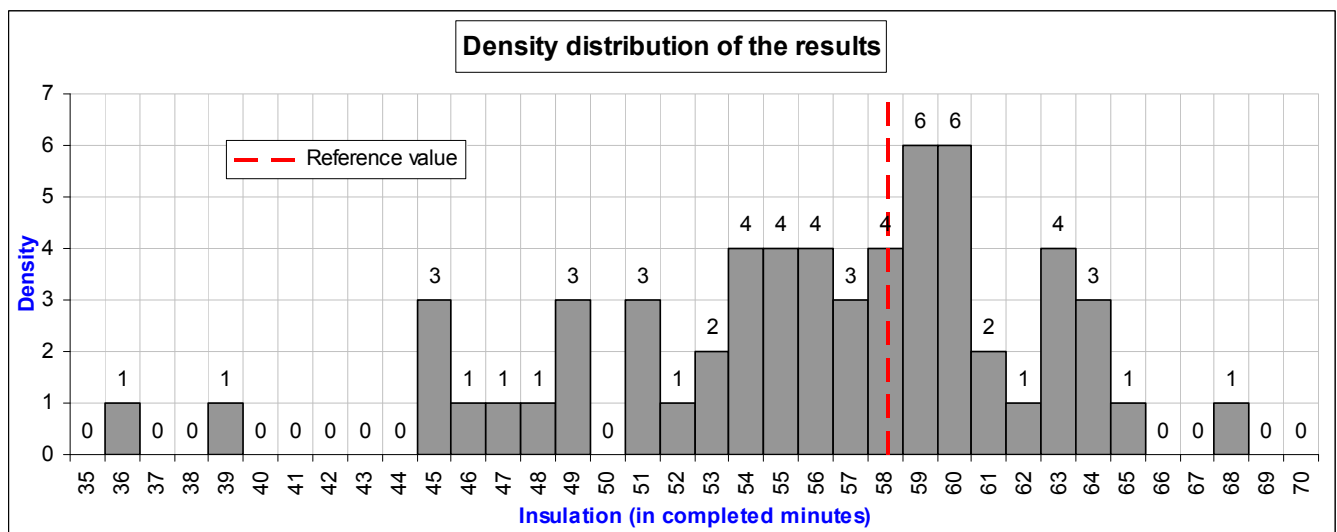


Chart 2



Egolf TC2 Fire resistance
Round Robin
Summarised report



2-2.2 RANKS

Ranks for the means are deduced from the mean result of each laboratory by assigning the rank 1 to the lab having the smallest mean, rank 2 to the lab having the next upper mean, ... up to rank p to the lab having the highest mean. Similarly, ranks for the standard deviations can be deduced from the standard deviation of each laboratory.

Gauss plots of the rank-sorted laboratories are shown below (see charts 3 to 6).

Chart 3

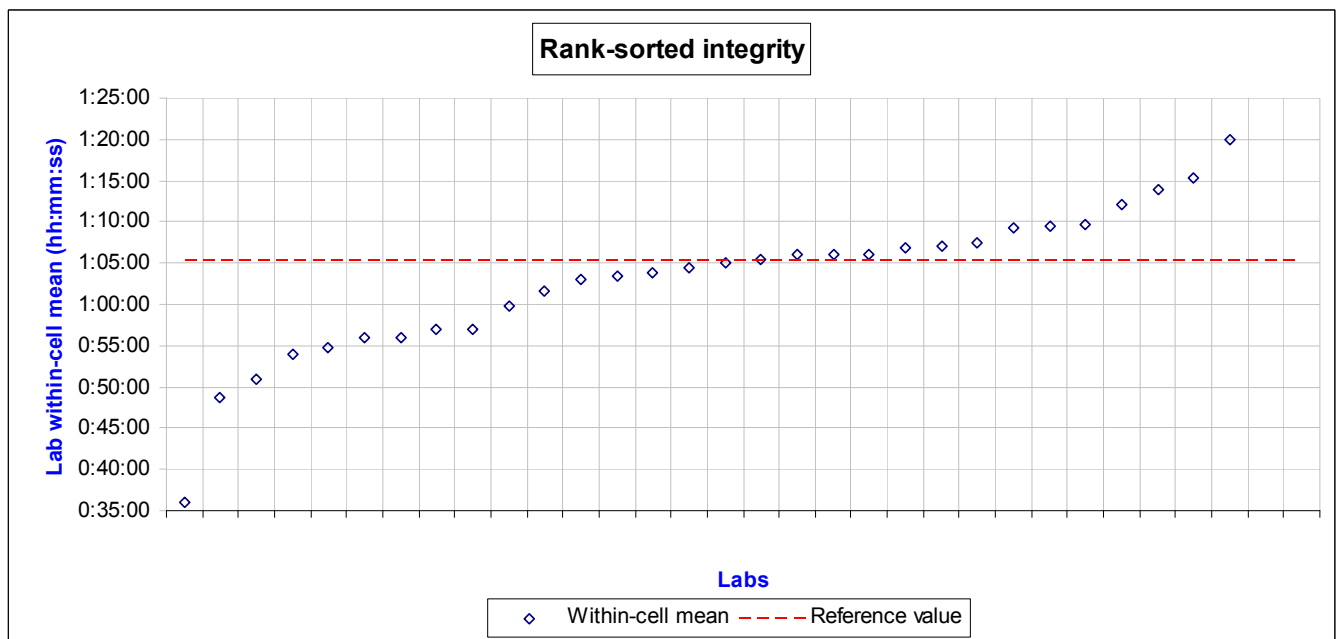
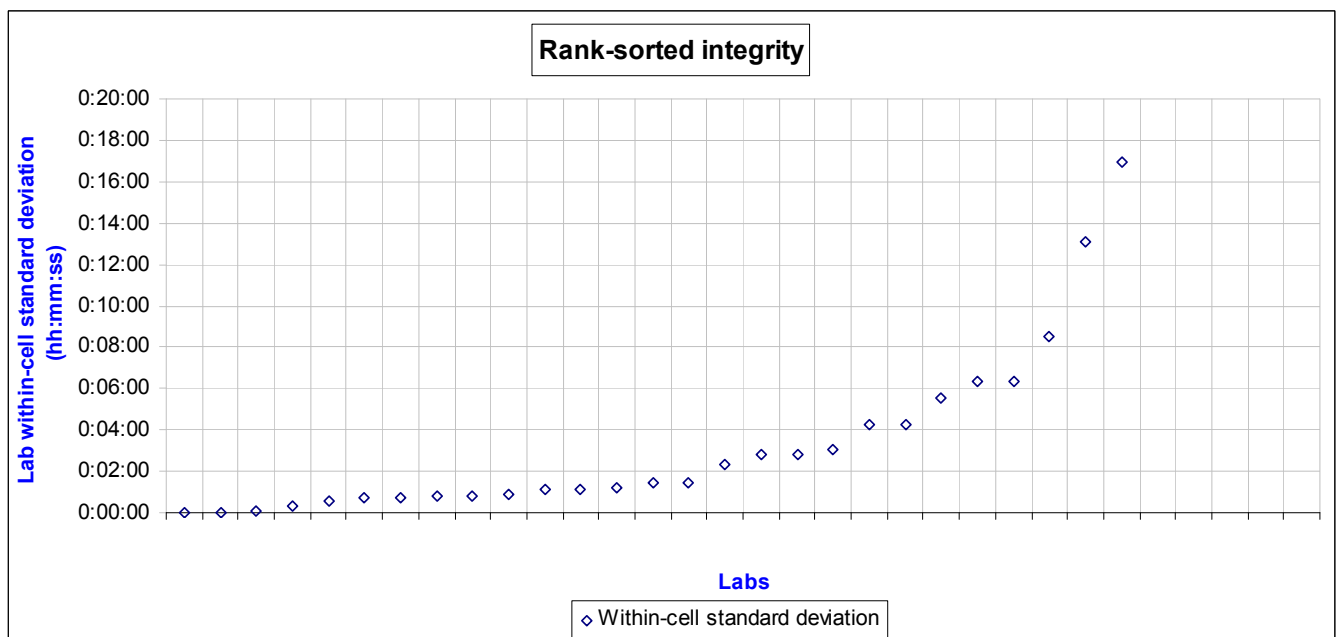


Chart 4



Egolf TC2 Fire resistance

Round Robin

Summarised report



Chart 5

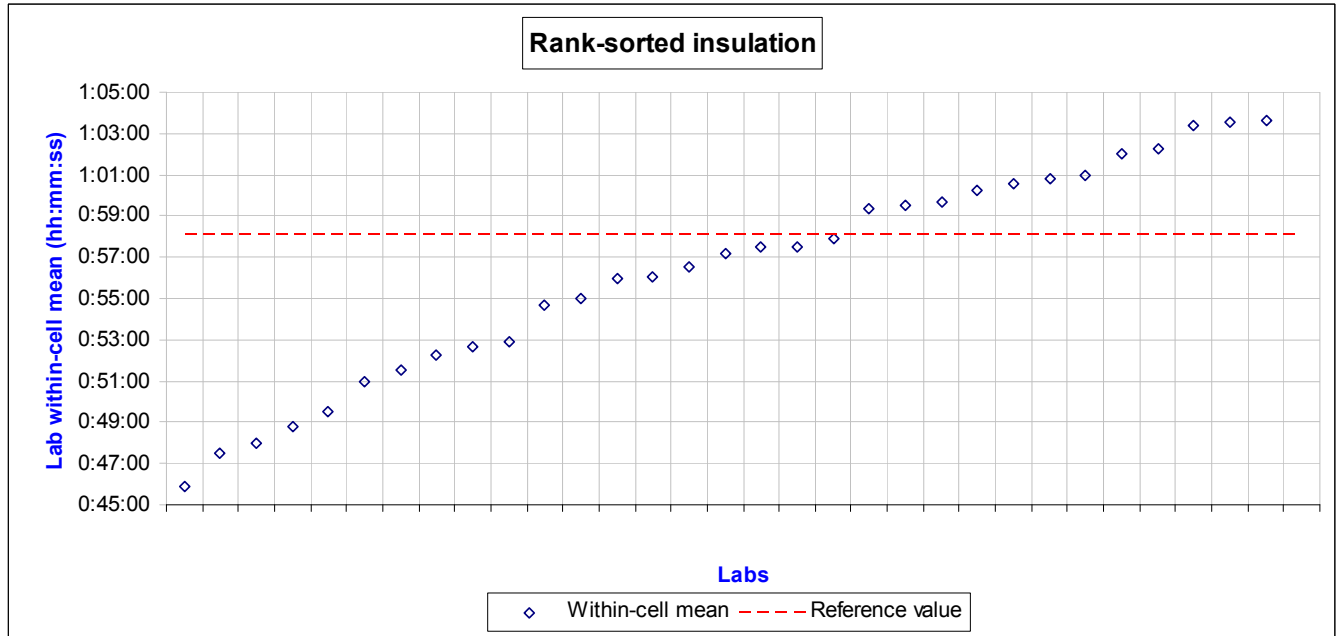
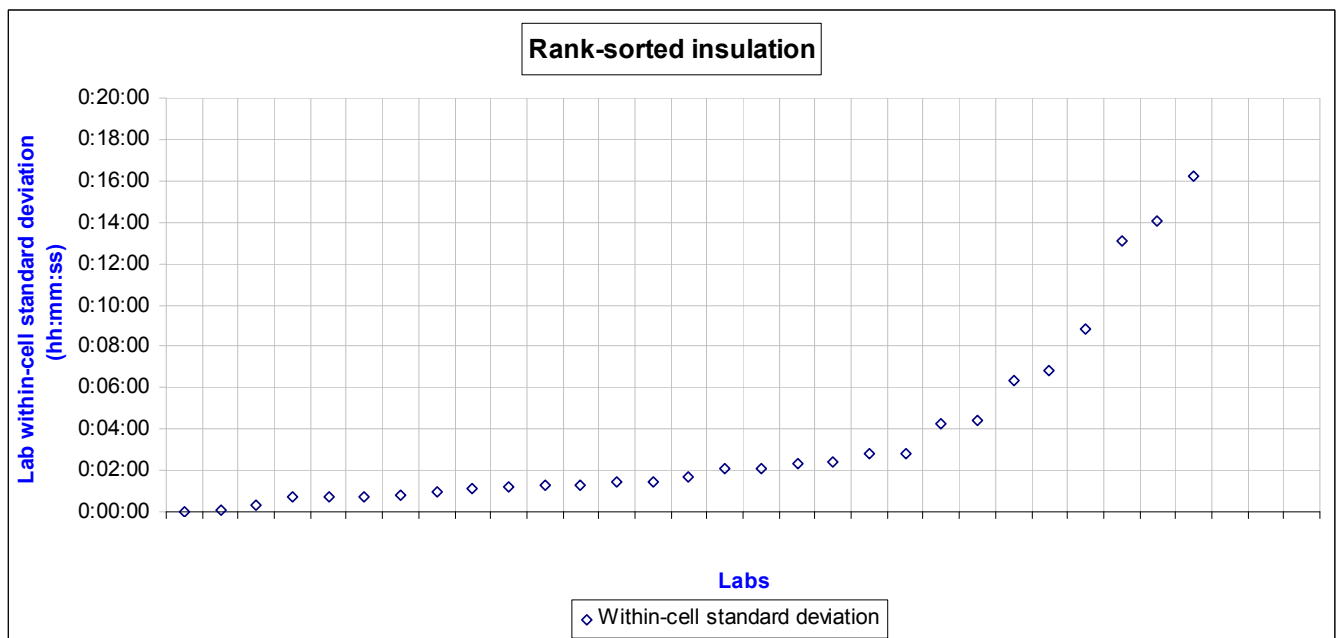



Chart 6



The rank-sorting provides a simple method to identify the laboratories having the most extreme results. They are often used to identify the laboratories that would be the more likely to improve their performance.

2-2.3 Z-SCORES

Z-score characterizes the bias and thus the trueness of a laboratory. It's defined by

	<p>Egolf TC2 Fire resistance</p> <p>Round Robin</p> <p>Summarised report</p>	
--	--	---

$$z_i = \frac{\bar{y}_i - m}{s}$$

where s is the proficiency testing standard deviation (standard deviation of the average results of the labs). Z-score is similar to Mandel's h-statistic. It's the most popular performance parameter.

The interpretation of the z-score follows these simple rules :

- $|z_i| \leq 2$: the trueness performance of the lab is satisfactory (the lab's mean is found to fall in the 95% range of more probability occurrence values),
- $2 < |z_i| \leq 3$: warning signal, the trueness performance of the lab is questionable (the lab's mean is found to fall in the 5% range of less probability occurrence values),
- $3 < |z_i|$: action signal, the trueness performance of the lab is unsatisfactory (the lab's mean is found to fall in the 0,3% range of less probability occurrence values).

(Out of 32 participating labs)	Z-score	
	On integrity	On insulation
Warning signal	3 labs	3 labs
Action signal	1 lab	-

2-2.4 K-SCORES

K-score characterizes the variability and thus the precision of a laboratory. It's defined by

$$k_i = \frac{s_i}{s_r}$$

K-score is similar to Mandel's k-statistic. It also expresses the same than the interpretation of the repeatability value r .

The interpretation of the k-score follows these simple rules :

- $k_i \leq 2$: the precision performance of the lab is satisfactory (the lab's standard deviation is found to fall in the 95% range of more probability occurrence values),
- $2 < k_i \leq 3$: warning signal, the precision performance of the lab is questionable (the lab's standard deviation is found to fall in the 5% range of less probability occurrence values),
- $3 < k_i$: action signal, the precision performance of the lab is unsatisfactory (the lab's standard deviation is found to fall in the 0,3% range of less probability occurrence values).

(Out of 32 participating labs)	K-score	
	On integrity	On insulation
Warning signal	1 lab	2 labs
Action signal	2 labs	6 labs

2-3 VALIDITY PERIOD OF THE RESULTS

The validity period of the result acquired by a laboratory on a single cycle of a RR limits itself to the period of this RR cycle. As a result, if a laboratory acquires a satisfying result on a single cycle, this result should not be used to claim that the laboratory acquires reliable data in any opportunity.

Only laboratories engaged in quality system and recording satisfying results during several cycles of a RR can be entitled to use results as proof that they're able of acquiring reliable data regularly.