THE MILLENIUM INTERNATIONAL CONFERENCE
OF THE
ISRAEL SOCIETY FOR QUALITY

November 28-30, 2000
Jerusalem, Israel

PROCEEDINGS

Volume I
Tuesday, November 28, 2000

Secretariat: ISAS International Seminars, POB 34001, Jerusalem 91340, Israel
Tel: 972-2-6520574, Fax: 972-2-6520558, seminars@isas.co.il

Nested Design for R&R Study in Chemical Industry

Dr. Pavel Grabov
ALD Ltd.
Fraiman, 20, Rishon Le-Zion
03-9678804, fax 03-9667495
e-mail: grabov@ald.co.il

Alex Korenev
Indigo Israel
POB 150 Rehovot 76101
08-9381959, fax 08-9381349
e-mail: koreneva@indigo.co.il

Abstract

Standard R&R study implies repeated testing of some parts by several operators. This 'matrix' procedure is of Factorial Design nature, and works reasonably well in mechanics, electronics and many other applications.
In chemical industry, however, the conventional R&R approach is not applicable due to presence of some preparation stages before the material testing. Since we are interested in total error evaluation, all these stages should be taken into account, because the final result is affected by an uncertainty associated with every stage. The special procedure based on so-called 'Nested Design' should be used in this case. The procedure implies total error decomposition through subsequent purification of upper level error from contribution of error on lower levels.
The results of the procedure application have indicated that the measurement error (EV) is the most critical contributor to overall R&R error. The problem of high EV could be solved either by purchasing more precise testing equipment, or by increasing the number of measurement. Due to budget limitations, the second alternative has been chosen.
Some optimal algorithm has been developed to keep the number of tests at a reasonable level and to enhance the test precision in critical vicinity of specification limits. The algorithm states the dependence of number of measurements on accepted results: 'The closer the test result is to specification limits, the more additional tests are required', and sets the formula for number of measurements calculation.

Introduction

Since measurement systems are used in Statistical Process Control for making decisions about processes, a conclusion about these systems themselves is necessary. The Repeatability and Reproducibility (R&R) Study is used to estimate the ability of a system to produce precise results

[1]. The study is performed by making repeated measurements on the same measurand. The result is evaluated for both repeatability, characterizing variation under identical conditions of measurement (variation due to gage itself), i.e. so-called Equipment Variation (EV), and reproducibility, characterizing variation under changed conditions of measurements (variation due to operator, time, reloading, reset, environment etc.), i.e. so-called Appraisal Variation (AV). The acceptability criterion of R&R results is usually given by the covered percent of the tolerance range (10% or less would be considered as excellent, 11-30% would be considered as acceptable, and greater than 30% would be unacceptable). The decision-making policy proceeding from the R&R Study results is shown in Figure 1.

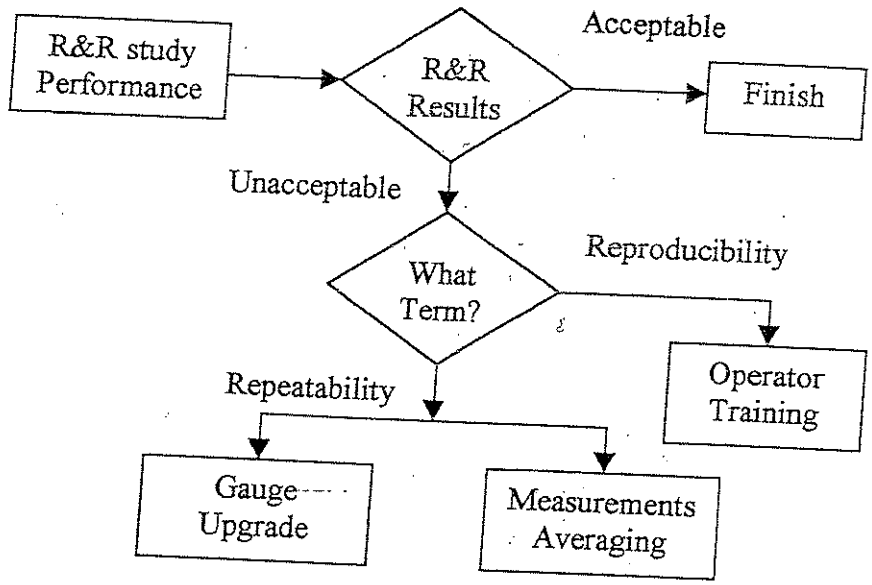

Figure 1. Flowchart of R&R Decision Making

In chemistry, however, the conventional R&R study is not applicable due to complexity of multistage procedure of sampling and sample preparation. Obviously, all these stages represent potential sources of additional variation and corresponding error.
For example, the procedure of viscosity evaluation consists of

- 1st step: 500kg batch sampling (100g)
- 2nd step: sample preparation (dilution using 20g of the material and a solvent)
- 3rd step: viscosity measurement.

Understanding how variation sources (sample preparation and measurement itself) are contributing to the total measurement error is important in directing where work efforts should be concentrated in order to reduce this error. So our purpose is decomposition of total measurement error. Another purpose is to develop some algorithm needed to achieve a reduction in total measurement error.

Nested Design Procedure

The conventional R&R procedure actually represents kind of Factorial Design where all sources of variation in the response variable should be identified and estimated. So most gage studies use rather simple orthogonal structure of experimental design supported by standard procedure of Analysis of Variance (ANOVA).

In chemical industry, however, the situation is more complicated due to the hierarchical structure of sampling and sampling preparation when some samples are taken from each batch and some independent measurements are performed on each sample. Thus the levels of one factor are not identical for different levels of another one.

This particular structure of different levels of one factor within the other is called *nested model* [2], and its statistical model (for simple two-stage nested design) is:

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{(ij)k} \qquad i=1,2,...,a;\ j=1,2,...,b;\ k=1,2,...,n.$$

Here y_{ijk} is the *(ijk)*-th observation, μ is overall observations mean, α_i is the effect of i-th level of factor A (batches), $b_{j(i)}$ is the effect of j-th level of factor B (samples) nested under the i-th level of factor A, and $\varepsilon_{(ij)k}$ is an error term of k-th replicate 'nested' within the combination of A and B. Since every level of factor B does not appear with every level of factor A there is not any interaction between the factors.

The procedure of testing hypotheses about significance of factors *Batch* and *Sample* is summarized in ANOVA table (see Table 1).

Table 1. ANOVA table for two-level nested design.

Source	*Sum of Squares*	*Degrees of Freedom*	*Mean Squares*
Batch - *A*	$bn \cdot \Sigma(\bar{y}_{i..} - \bar{y}_{...})^2$	*a-1*	MS_A
Sample - *B*	$n\Sigma\Sigma(\bar{y}_{ij.} - \bar{y}_{i..})^2$	*a(b-1)*	$MS_{B(A)}$
Error	$\Sigma\Sigma\Sigma(y_{ijk} - \bar{y}_{ij.})^2$	*ab(n-1)*	MS_E
Total	$\Sigma\Sigma\Sigma(y_{ijk} - \bar{y}_{...})^2$	*abn-1*	

The symbols used in the Table 1 are:

$$\bar{y}_{...} = \left(\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}\right)\Big/(a \cdot b \cdot n)$$

$$\bar{y}_{i..} = \left(\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}\right)\Big/(b \cdot n)$$

$$\bar{y}_{ij.} = \left(\sum_{k=1}^{n} y_{ijk}\right)\Big/n$$

The above mentioned approach can be illustrated by the example of viscosity measurements. Viscosity represents one of the critical output parameters of some chemical production process. Expensive material is scrapped if its viscosity falls beyond the specification limits (40-80).
A two-level hierarchy describes the situation with the R&R study of the viscosity testing procedure. To understand the contribution of sampling error 4 samples have been taken. To evaluate the contribution of sample preparation three working dilutions have been prepared from each sample by the same operator. To estimate the gauge error each dilution has been tested three times using the same tool.
The results of the performed R&R study are presented in Table 2.

Table 2. Results of R&R Study.

Sample	*S1*			*S2*			*S3*			*S4*		
Dilution	*D1*	*D2*	*D3*	*D1*	*D2*	*D3*	*D1*	*D2*	*D3*	*D1*	*D2*	*D3*
Test 1	62.1	62.9	64.2	61.4	65.7	64.7	68.3	70.8	62.5	63.5	60.2	64.2
Test 2	67.9	60.2	57.1	66.4	61.5	62.7	58.3	69.4	65.0	60.4	55.9	54.5
Test 3	65.7	61.8	58.5	64.1	71.7	62.3	71.2	71.9	63.1	60.1	63.4	61.1

Mean Squares
MS_A
$MS_{B(A)}$
MS_E

$\sum_{k=1}^{n} y_{ijk} \big)/(a \cdot b \cdot n)$

$_{ijk} \big)/(b \cdot n)$

$'_{ijk} \big)/n$

he example of
critical output
sive material is
s (40-80).
R study of the
n of sampling
tion of sample
m each sample
ution has been

ble 2.

S4	
D2	D3
60.2	64.2
55.9	54.5
63.4	61.1

The statistical model for above data can be presented as follows

$$y_{ijk} = \mu + S_i + D_{j(i)} + \varepsilon_{(ij)k} \qquad i=1,2,3,4;\ j=1,2,3;\ k=1,2,3.$$

Where the number of levels (i and j) of factors S (sample) and D (dilution) is 4 and 3, respectively. Obviously, D is nested within S. Number of repeated measurements (k) is 3. The results of data treatment using the ANOVA procedure are presented in Table 3.

Table 3. ANOVA Table

Source	*Sum of Squares*	*Degrees of Freedom*	*Mean Squares*
Sample	204.45	3	68.15
Dilution	143.39	8	17.92
Error	300.17	24	12.51
Total	648.01	35	

Since both S and D represent known nuisance factors, the general procedure of random effects model should be applied to evaluate their true variance. Table 4 presents the model of variance decomposition.

Table 4. Model of Variance Structure.

E(MS)	*Variance Structure*
MS_S	$\sigma_E^2 + n \cdot \sigma_D^2 + bn \cdot \sigma_S^2$
$MS_{D(S)}$	$\sigma_E^2 + n \cdot \sigma_D^2$
MS_E	σ_E^2

Where σ_E^2, σ_D^2 and σ_S^2 represent variances of error term, dilution and sampling, respectively.

One can see that at each nested level the MS value includes both the given level's error and multiplied errors of the previous ones. Application of the standard purification procedure to the obtained MS values resulted in following estimates

$$\sigma_E^2 = E(MS_E) = 12.51$$

$$\sigma_D^2 = \frac{E(MS_{D(S)}) - E(MS_E)}{n} = \frac{17.92 - 12.51}{3} = 1.80$$

$$\sigma_S^2 = \frac{E(MS_S) - E(MS_{D(S)})}{bn} = \frac{68.15 - 17.92}{9} = 5.58$$

Analysis of the obtained results shows that the gauge is the largest contributor (63%) to total measurement error. Moreover, the EV value is absolutely unacceptable because it covers almost half of the specification width:

$$EV\ /Tolerance\_Ratio = (5.15 \cdot \sigma_E)/(USL - LSL) = 18.22/(80 - 40) = 45.55\%$$

Where 5.15 represents number of standard deviations corresponding to the accepted 99% confidence level.

There are two possible solutions of the measurement error problem:

- gauge upgrade (purchasing the more precise one);
- gauge error reduction by the results averaging, i.e. by increasing number of measurements (n).

Due to budget limitations the second choice has been preferred in our case. Meanwhile taking into account that the performance of additional measurements is also cost consuming, some algorithm of n-value optimization has been proposed.

n -Value Optimization Algorithm

The algorithm idea is based on the general approach of 'Guidelines on Assessment and Reporting of Compliance with Specification' [3]. The approach defines that compliance with the specification can be stated if and only if the specification limits are not breached by the test result, extended by half of expanded uncertainty interval at the accepted confidence level (EV value in our case). Otherwise it is not possible to state compliance.

On one hand, the obtained EV value indicates that there are a lot of situations where the compliance cannot be stated. On another hand, straightforward increase of number of measurements in order to achieve

higher precisio
The proposed a
precision level,
is to the speci
measurements'.
the obtained te
specification lir
number of mea
the specification
The required nu
to keep the pote
C_{pk} value cor
specification targ

$$C_{pk} = \frac{\Delta}{3\sigma_E/\sqrt{n}}$$

From this constra

$$n = \left(\frac{0.5(USL - L}{\Delta} \right.$$

Where Δ is the
specification limit.
Obviously, this re
Δ is close to zero.
coefficient ($K \le 1$
($K = 1$ preserves
specifications).

$$n = \left(\frac{0.5 \cdot K \cdot (USL} {\Delta} \right.$$

Lack of complianc
corresponding valu
distribution. The in
Figure 2.

higher precision by their averaging yields time consuming test procedure. The proposed algorithm sets minimum number of tests providing required precision level, in accordance with the principle: 'The closer the test result is to the specification limits – the larger is the number of additional measurements'. The dependence of number of repeated measurements on the obtained test result provides the higher precision in the vicinity of specification limits through n increase and allows performance of smaller number of measurements when the test results are concentrated around the specification target.

The required number of repeated tests can be derived from the constrain to keep the potential capability level of measurements estimated by given C_{pk} value corresponding to the single test result located at the specification target:

$$C_{pk} = \frac{\Delta}{3\sigma_E / \sqrt{n}} = const$$

From this constrain one can get

$$n = \left(\frac{0.5(USL - LSL)}{\Delta} \right)^2$$

Where Δ is the distance of the observed test result from the nearest specification limit.

Obviously, this reciprocal quadratic function yields infinite n value, when Δ is close to zero. The dependence of n on Δ can be scaled using some coefficient ($K \leq 1$) characterizing the compliant policy of given company ($K = 1$ preserves initial safety in making decisions on compliance to specifications).

$$n = \left(\frac{0.5 \cdot K \cdot (USL - LSL)}{\Delta} \right)^2$$

Lack of compliance safety for small K values can be evaluated using corresponding values of Cumulative Density Function (cdf) of normal distribution. The impact of K value on compliance safety is shown in Figure 2.

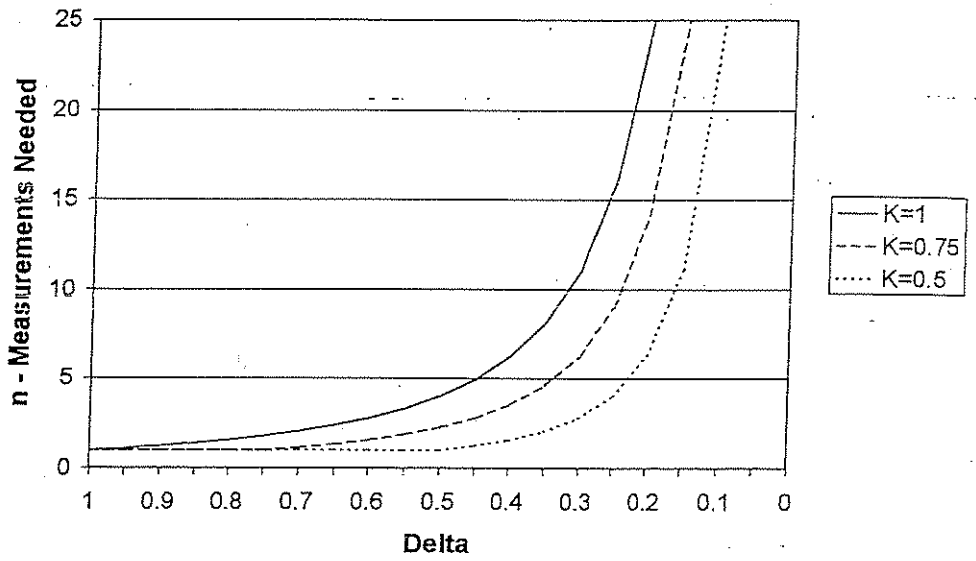

Figure 2. Dependence of n on Δ (the distance is given in terms of half tolerance width).

Summary and Conclusions

Application of the advanced procedure of R&R study for chemical industry has been described. The procedure corresponds to the random model of Factorial Experiments and involves the Nested Design approach. The optimization algorithm of testing procedure has been proposed for the situations with unacceptable EV value. The algorithm yields minimum number of performed measurements providing required level of test precision. This number depends on the obtained test result: 'The closer the test result is to the specification limits – the larger is the number of additional measurements'.

Bibliography

(1) Barrentine, L.B. (1991), Concepts for R&R Studies, ASQ Quality Press.

(2) Montgomery, D.C. (1997), Design and Analysis of Experiments, 4th edition. Wiley, New York.

(3) Guidelines on Assessment and Reporting of Compliance with Specification (1997), ANSI/NCSL Z.540-2-1997.