

# *False decisions on conformity of a multicomponent material or object and quality of chemical analytical results*

**Ilya Kuselman<sup>1</sup> · Francesca Pennechi<sup>2</sup> ·  
Ricardo J. N. B. da Silva<sup>3</sup> · D. Brynn Hibbert<sup>4</sup>**

<sup>1</sup> Independent Consultant on Metrology, Israel

<sup>2</sup> Istituto Nazionale di Ricerca Metrologica, Italy

<sup>3</sup> Universidade de Lisboa, Portugal

<sup>4</sup> School of Chemistry, UNSW, Australia

# *Conformity assessment*

[https://www.youtube.com/watch?v=\\_Sjp58GUJwk&app=desktop](https://www.youtube.com/watch?v=_Sjp58GUJwk&app=desktop)

JCGM 106 provides a methodology for CA of a component concentration in a material or object with the specified requirements. The Bayesian approach is that knowledge about the component concentration (the **measurand**) can be treated as a random variable and expressed in terms of a pdf.

Such pdf combines prior knowledge of the measurand and new information acquired during the measurement/testing. The posterior pdf allows to estimate the measurand value and the associated measurement uncertainty (MU) as **mean and STD** of the distribution.

# *Comparison of a result with tolerance limits*

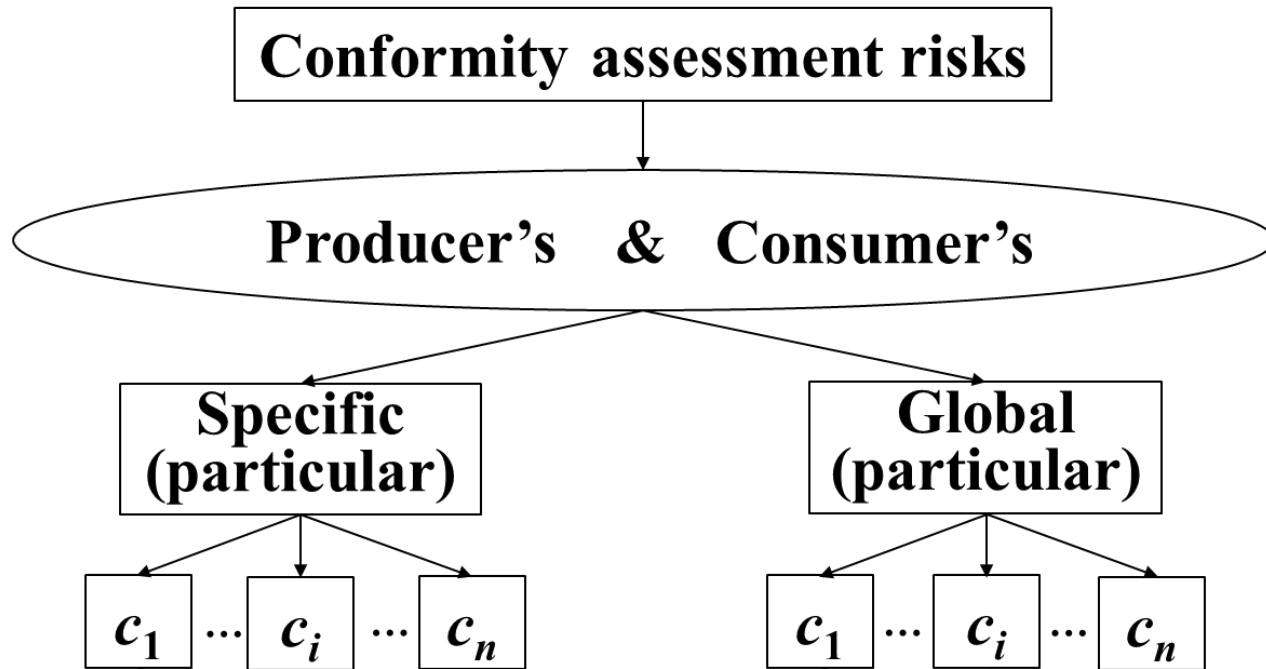
Comparing a measurement/test result with the specification, regulation or legal (tolerance) limits of the material, one should decide whether the tested concentration conforms or not.



# *The risks due to measurement uncertainty*

MU influences the decision and causes risks of two types. The probability of accepting the material batch or lot, when it should have been rejected, is named ‘**consumer’s** risk’, whereas the probability of falsely rejecting the batch is the ‘**producer’s** risk’. For a specified batch (or lot), they are the ‘**specific** consumer’s risk’ and the ‘**specific** producer’s risk’.

The risks of CA of a batch randomly drawn from a population of such batches are the ‘**global** consumer’s risk’ and the ‘**global** producer’s risk’, since they characterize the material production globally.



Joint Committee for Guides in Metrology. JCGM 106:2012  
Evaluation of Measurement Data – The Role of MU in CA.

M. Darstmardi, M. Mohammadi, B. Naderi (2018) Optimizing  
MU to reduce the risk and cost in CA. ACQUAL 23:19-28

*See not only the trees but also the forest*



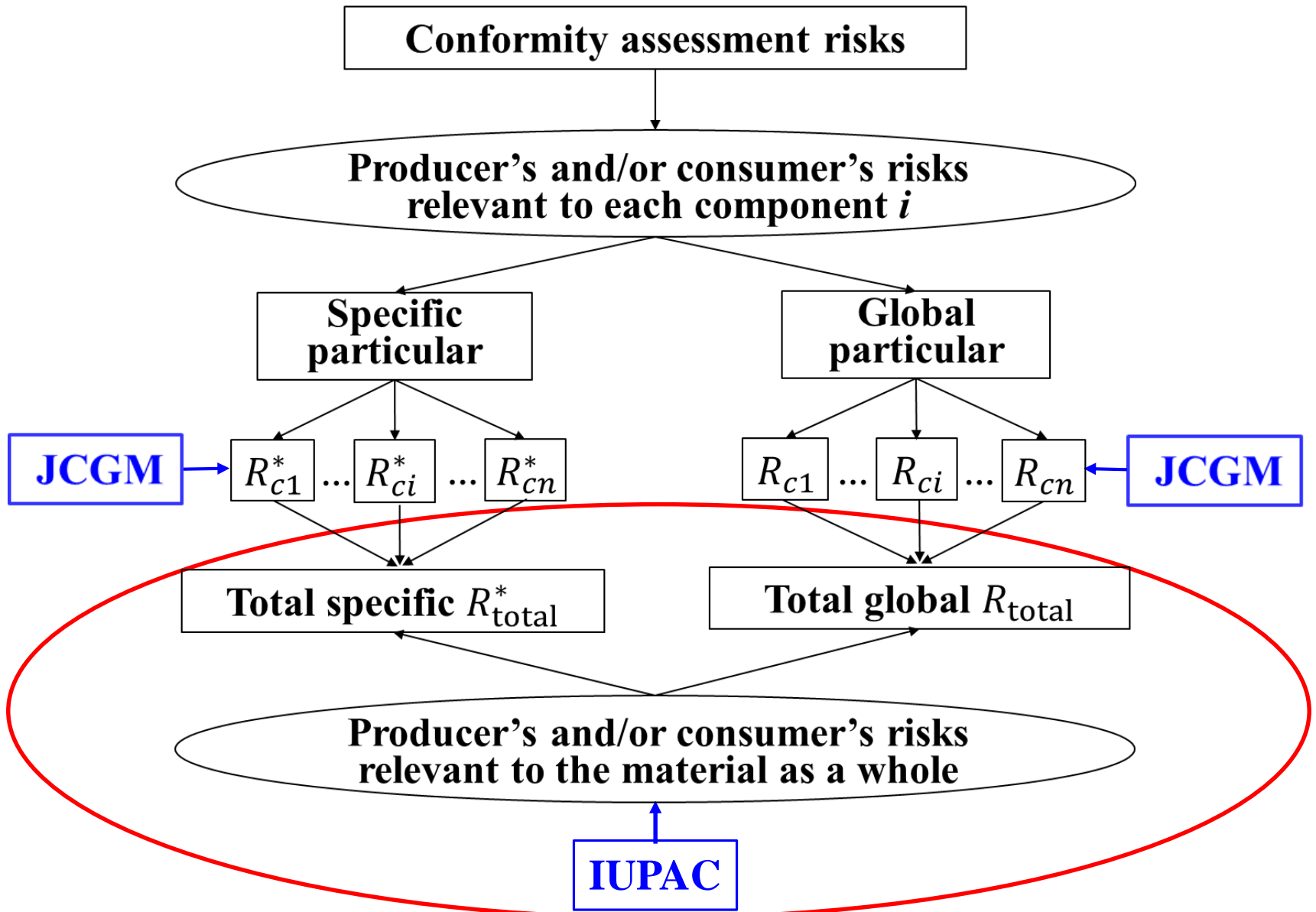
The forest is more than the trees

# *Total risk*

When **particular** CA risks are acceptable, the total probability of a false decision (**total** consumer's risk or producer's risk) on the conformity of the material as a whole might still be significant.

L. Pendrill, H. Karlson, N. Fischer, S. Demeyer, A. Allard. EURAMET: A guide to decision-making and CA - A report of the EMRP joint project NEW04 "Novel mathematical and statistical approaches to uncertainty evaluation", 2015

IUPAC project 2016-007-1-500. I. Kuselman, F. Pennechi, R. da Silva, D.B. Hibbert (2017) Talanta 164C:189-195





# *Acceptance limits*

Besides the tolerance interval, a narrower **acceptance interval** for test results can be applied with the purpose of decreasing the consumer's risk by taking into account the MU. In such a case, the decision rules (is the material conforming or not?) are based on comparing the measured concentration values with the acceptance limits.

In current practice, the decision rules are often based on **direct comparison** of the measured concentration values with the specification or regulatory limits. The reason is that these limits have already taken into account the MU, and so the tolerance and the acceptance limits coincide.

# *Total global risk: events and probabilities*

- $C_1$ : the test result  $c_{1m}$  for component 1 is in its acceptance interval  $A_1$ ;  $P(C_1)$ .
- $C_2$ : the test result  $c_{2m}$  for component 2 is in its acceptance interval  $A_2$ ;  $P(C_2)$ .
- $C$ : the batch as a whole is accepted,  $C = C_1 \cap C_2$ ;  $P(C) = P(C_1) P(C_2)$ , if  $C_1$  and  $C_2$  are mutually independent.
- $B_1$ : the actual ('true') concentration  $c_1$  of component 1 is **not** within its tolerance interval  $T_1$ ;  $P(B_1)$ .
- $B_2$ : the actual concentration  $c_2$  of component 2 is not within its tolerance interval  $T_2$ ;  $P(B_2)$ .
- $B$ : the material as a whole is **not conforming**,  $B = B_1 \cup B_2$ ;  $P = P(B_1) + P(B_2) - P(B_1) P(B_2)$ .

## *Total global consumer's risks*

Particular risks are  $R_{c_1} = P(C_1 \cap B_1)$ ,  $R_{c_2} = P(C_2 \cap B_2)$ .

$$R_{\text{total}} = P(C_2)R_{c_1} + P(C_1)R_{c_2} - R_{c_1}R_{c_2}.$$

The  $i$ -th particular global risk  $R_{c_i}$  can be evaluated as an integral of the joint prior and likelihood (posterior) pdf, described in JCGM 106. The probability  $P(C_i)$  of acceptance of a measurement/test results for  $i$ -th component is calculated by marginalization of this joint pdf.

For example, for  $R_{c_i} = 0.05$  and  $P(C_i) = 0.90$ ,  $i = 1, 2$ ,  
 $R_{\text{total}} = 2 \times (0.90 \times 0.05) - 0.05^2 = 0.09$ .

## *Total specific consumer's risk*

When a specified batch is tested, total specific risk  $R_{\text{total}}^*$  is  $P(B|c_{1m}, c_{2m})$  that the actual concentration of one or both the components in this batch **are not within** the tolerance intervals, whereas the test results  $c_{1m}$  and  $c_{2m}$  of both the components are **within** their acceptance limits.

Since particular specific consumer's risks for the  $i$ -th component,  $i = 1, 2$ , are  $R_{ci}^* = P(B_i|c_{im})$  described in JCGM 106, the total specific risk is:

$$R_{\text{total}}^* = R_{c1}^* + R_{c2}^* - R_{c1}^* R_{c2}^*.$$

E.g. for  $R_{ci}^* = 0.05$ ,  $R_{\text{total}}^* = 2 \times 0.05 - 0.05^2 = 0.10$ .

# Customs control of denatured alcohols

**Denatured Alcohol**

**DANGER**

Highly flammable liquid and vapor. Causes serious eye irritation. May cause drowsiness or dizziness.

**PREVENTION**

Keep away from heat, sparks, and open flames. — No smoking. Keep container tightly closed.

Avoid breathing vapors. Use only outdoors or in a well-ventilated area. Wear eye protection.

**RESPONSE**

**If inhaled:** Remove person to fresh air and keep comfortable for breathing. Call a doctor if you feel unwell.

**If in eyes:** Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. If eye irritation persists: get medical attention.

**In case of fire:** Use water spray, alcohol-resistant foam, dry chemical or carbon dioxide for extinction.

≠



www.shutterstock.com · 533932654

By EU Regulation 162/2013, a procedure for completely denaturing alcohol (CDA) consists of addition of **3 L** of isopropyl alcohol (IPA), **3 L** of methyl ethyl ketone (MEK) and **1 g** of denatonium benzoate (DB) to **1 hL** of EtOH.

Concentrations of EtOH, IPA and MEK are tested using **GC-FID**, and DB - using **HPLC-UV**.

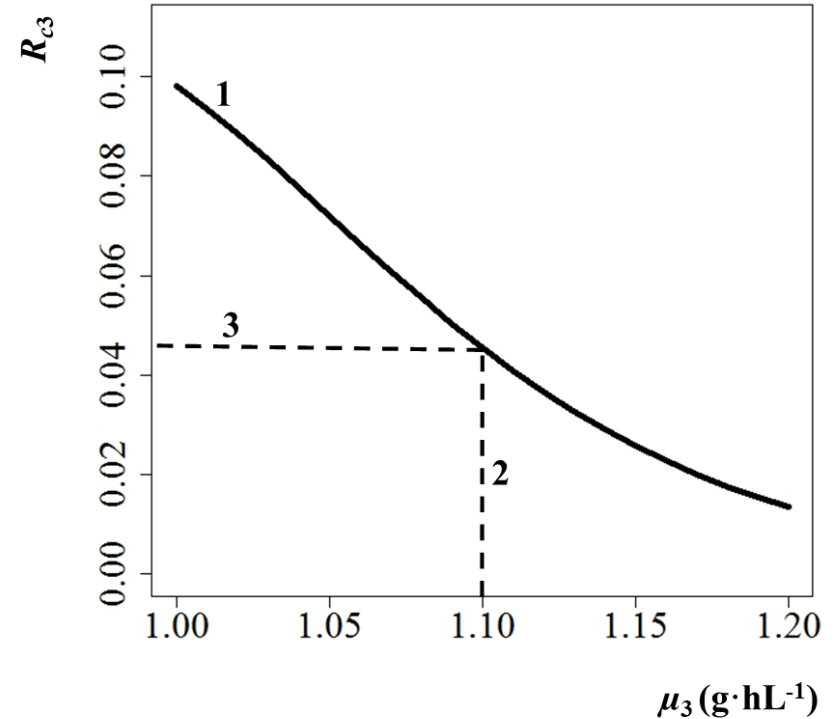
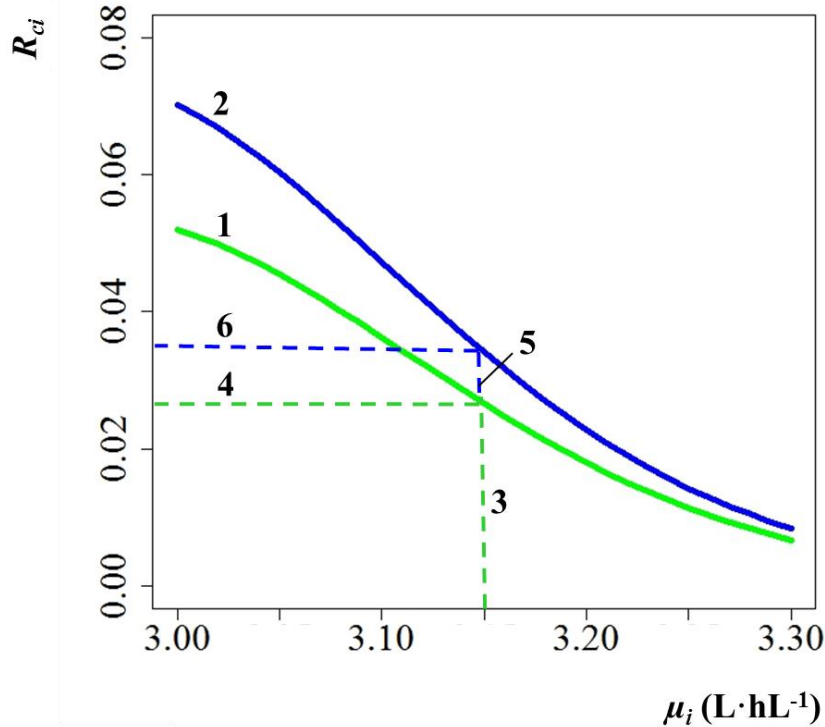
# *Distribution of measurement results – the likelihood functions*

The standard MU is  $u_1 = 0.05 \text{ L}\cdot\text{hL}^{-1}$  for IPA,  $u_2 = 0.07 \text{ L}\cdot\text{hL}^{-1}$  for MEK, and  $u_3 = 0.07 \text{ g}\cdot\text{hL}^{-1}$  for DB as in the methods validation report by [E. Aries et al. Euro-denaturant project – Phase II, Administrative Arrangement TAXUD-2014-DE-317, Geel, Belgium, IRMM \(2016\)](#)

The  $u_i$  value was applied as STD of the distribution of  $c_{im}$ , taken as normal one based on the validation data:

$$f(c_{im}|c_i) = \frac{1}{u_i\sqrt{2\pi}} \exp\left[-\frac{(c_{im} - c_i)^2}{2u_i^2}\right].$$

# Particular global risks $R_{ci}$



The standard MU is  $u_1 = 0.05 \text{ L}\cdot\text{hL}^{-1}$  for IPA,  $u_2 = 0.07 \text{ L}\cdot\text{hL}^{-1}$  for MEK, and  $u_3 = 0.07 \text{ g}\cdot\text{hL}^{-1}$  for DB. Greater MU leads to greater risk.

# *Total global customs risk*

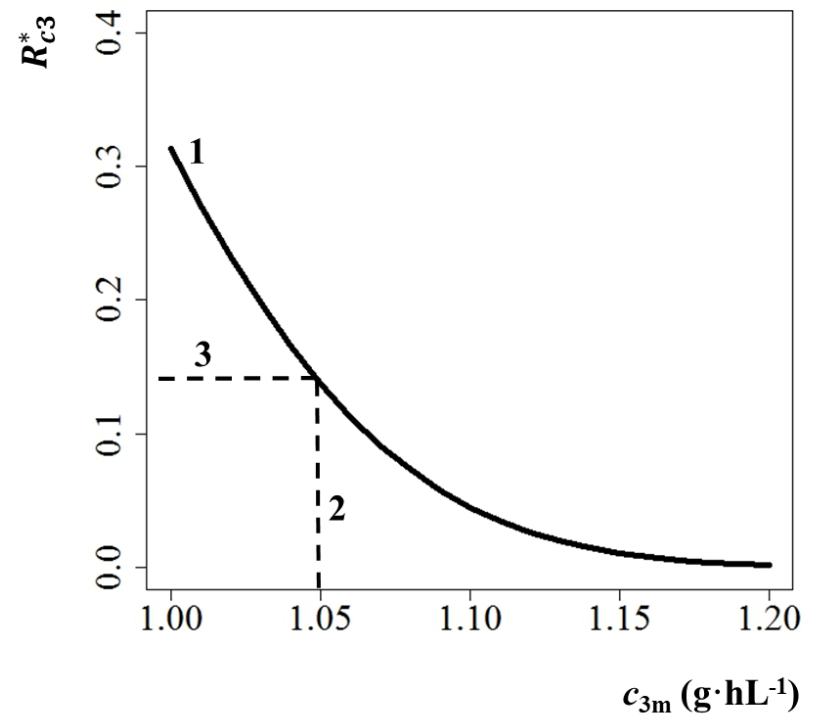
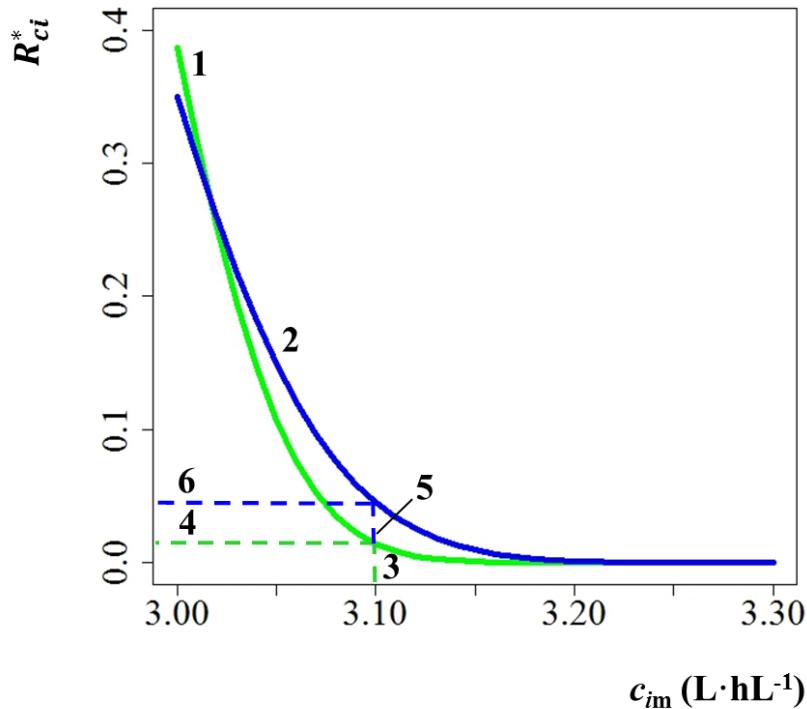
1) In the case of control of IPA and MEK at the above mentioned conditions  $R_{\text{total}} = 0.808 \times 0.027 + 0.818 \times 0.034 - 0.027 \times 0.034 = 0.048$ . It is greater than each particular risk.

2) When all the denaturants (IPA, MEK and DB) are under control at the same conditions,  $R_{\text{total}} = 0.808 \times 0.778 \times 0.027 + 0.818 \times 0.778 \times 0.034 + 0.818 \times 0.808 \times 0.046 - 0.778 \times 0.027 \times 0.034 - 0.808 \times 0.027 \times 0.046 - 0.818 \times 0.034 \times 0.046 + 0.027 \times 0.034 \times 0.046 = 0.066$ .

This value is greater than that calculated in the case of control of just IPA and MEK.



# Particular specific risks $R_{ci}^*$



By dotted lines an example is shown when  $c_{1m} = c_{2m} = 3.10 \text{ L}\cdot\text{hL}^{-1}$  for IPA and MEK, and  $c_{3m} = 1.05 \text{ g}\cdot\text{hL}^{-1}$  for DB - the batch should be recognized as properly denatured.

# *Total specific customs risk*

In the example above, there are still the particular specific customs risks  $R_{c_1}^* = 0.014$ ,  $R_{c_2}^* = 0.045$ , and  $R_{c_3}^* = 0.138$ .

1) If IPA and MEK only influence the decision on the batch conformity,  $R_{\text{total}}^* = 0.014 + 0.045 - 0.014 \times 0.045 = 0.059$ .

2) When all the denaturants are taken into account,  $R_{\text{total}}^* = 0.014 + 0.045 + 0.138 - 0.014 \times 0.045 - 0.014 \times 0.138 - 0.045 \times 0.138 + 0.014 \times 0.045 \times 0.138 = 0.188$ .

This value is caused mostly by DB, since  $R_{c_3}^*$  is larger than  $R_{c_1}^*$  and  $R_{c_2}^*$ . At the same time, DB is the bitterest compound known and some  $c_3$  variations do not change the terrible bitter feeling of a person trying to drink CDA.

# *Conformity assessment of concentration of TSPM in ambient air from three stone quarries*



F. Pennechi, I. Kuselman, R. da Silva, D.B. Hibbert  
(2018) Chemosphere 202:165-176

# *Method of testing and national regulations*

A measured TSPM concentration  $c_{im}$ ,  $\text{mg m}^{-3}$ , is an averaged mass of particles with diameters of  $100 \mu\text{m}$  or less collected from the air drawn through a filter in a high-volume sampler over the sampling period in proximity to the  $i$ -th quarry.

The testing was at a distance of (1-3) km from a quarry. Each test lasted 24 hours for collection of particles from about  $2000 \text{ m}^3$  of air (EPA IO-2.1, 1999).  $T_{Ui} = 0.200 \text{ mg m}^{-3}$ .

The distribution of the test/measurement results  $c_{im}$  at the actual concentration  $c_i$  was found to be normal with STD equal to SMU  $u_i = 0.07c_{im}$ , and mean equal to  $c_i$ .

# *Likelihood, prior and posterior pdfs*

The likelihood functions (of the test/measurement results  $c_{im}$  at actual  $c_i$ ) are **normal**:

$$f(c_{im}|c_i) = \frac{1}{u_i\sqrt{2\pi}} \exp\left[-\frac{(c_{im} - c_i)^2}{2u_i^2}\right].$$

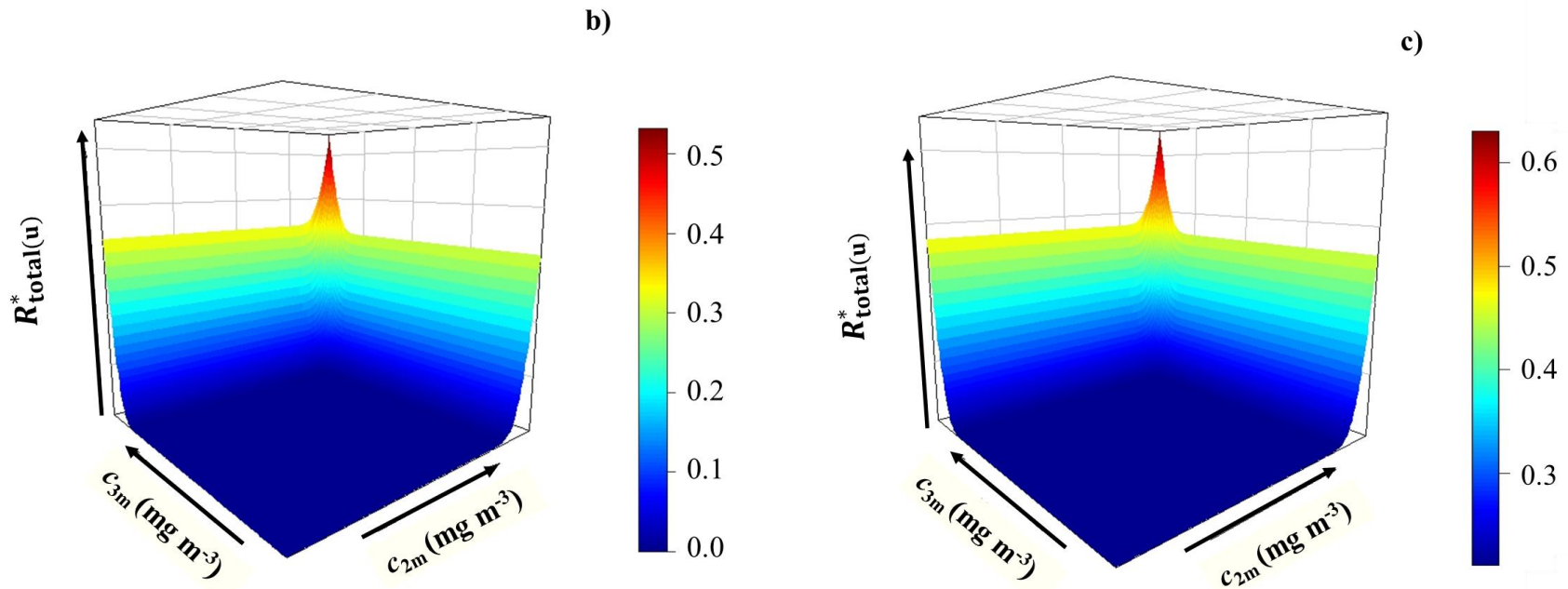
496 test results obtained during a year were fitted successfully by **lognormal** distributions and used as prior pdfs:

$$f(c_i) = \frac{1}{c_i\sigma_i\sqrt{2\pi}} \exp\left[-\frac{(\ln c_i - \mu_i)^2}{2\sigma_i^2}\right].$$

The posterior pdf is

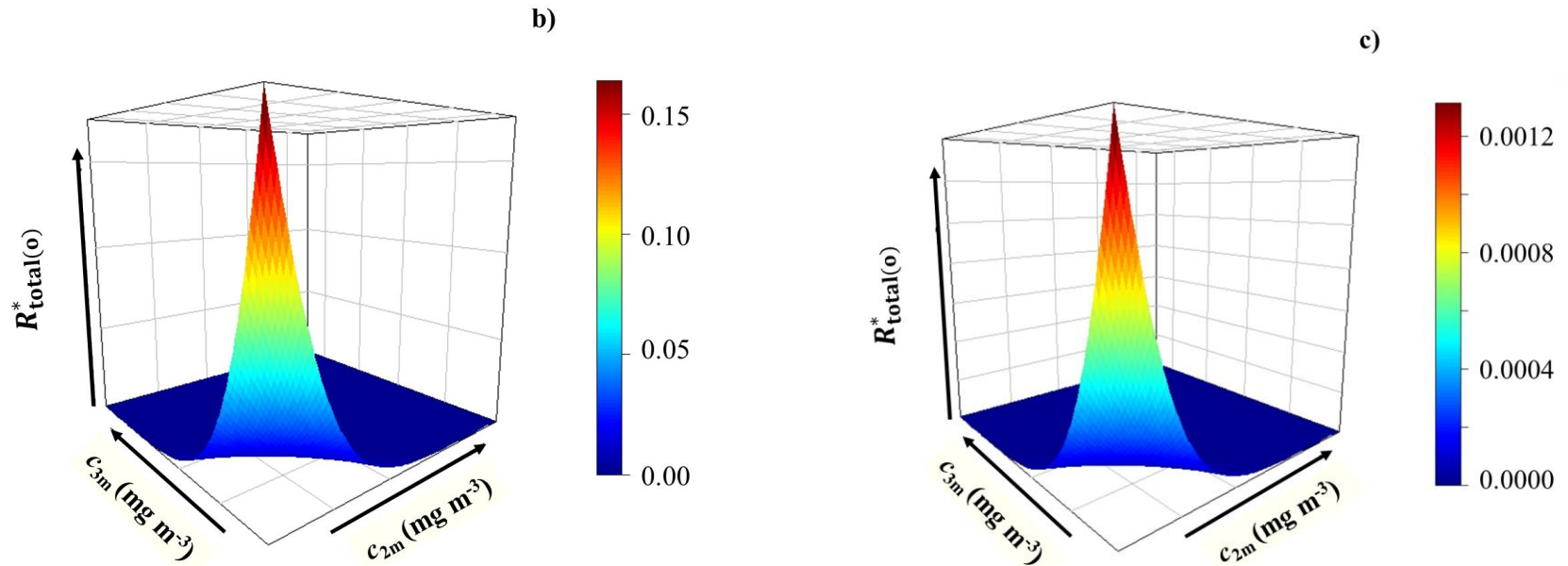
$$f(c_i|c_{im}) = f(c_{im}|c_i)f(c_i) / \int_{-\infty}^{\infty} f(c_{im}|c_i)f(c_i) dc_i$$

# *Dependence of the total specific risk of under-estimation of the TSPM concentration on $c_{im}$*



**b)** quarries  $i = 2$  and  $3$  are active only:  $c_{2m}$  and  $c_{3m}$  are in the range  $[0.010, 0.200] \text{ mg m}^{-3}$ ; **c)** all the three quarries are active:  $c_{2m}$  and  $c_{3m}$  are as on the left plot, while  $c_{1m} = 0.194 \text{ mg m}^{-3}$ .

# *Dependence of the total specific risk of over-estimation of the TSPM concentration on $c_{im}$*



**b)** quarries  $i = 2$  and  $3$  are active only:  $c_{2m}$  and  $c_{3m}$  are in the range  $[0.210, 0.300] \text{ mg m}^{-3}$ ; **c)** all the three quarries are active:  $c_{2m}$  and  $c_{3m}$  are as on the left plot, while  $c_{1m} = 0.250 \text{ mg m}^{-3}$ .



# *Correlation of measurement/test results*



I. Kuselman, F. Pennechi, R. da Silva, D.B. Hibbert  
(2017) *Talanta* 174:789-796



# *Interdependence of the events*

A number of techniques are used to overcome **correlations between measurement/test results**. There are extraction of analytes; chromatographic separation; chemometrics separation of spectral signals; sample digestion and standard additions; and so on. Still something may happen in practice, but in general this kind of correlation should be negligible.

**Correlation of actual ('true') values** of concentrations of different components may be caused by stoichiometry, the law of conservation of mass, and technological constraints. When sum of mass fractions is 100 %, the data are named **'compositional'** and their correlation - **'spurious'**.

# *A case study of test results of NyQuil tablets*

This cold/flu medication contains four active components: 1) acetaminophen (**APAP**) as a pain reliever and fever reducer; 2) dextromethorphan hydrobromide (**DEX**) as a cough suppressant; 3) doxylamine succinate (**DOX**) as an antihistamine and hypnotic; and 4) phenylephrine hydrochloride (**PE**) as a nasal decongestant.

The specification limits for each component  $i = 1$  to 4 are 95.0 - 105.0 % of the labeled amount. Testing is performed using HPLC system with UV detector according to USP <621>. Relative MU was  $u_{\text{rel}} = 2.8 \%$ . Therefore,  $u_i = (u_{\text{rel}}/100 \%) c_{im} = 0.028 c_{im}$ , % of labeled amount.

# *Pearson's correlation coefficients $r_{ij}$ of the test results*

Component	Index	APAP	DEX	DOX	PE
	<i>i</i> \ <i>j</i>	1	2	3	4
APAP	1	1	0.107	0.125	0.177
DEX	2		1	0.311	0.404
DOX	3			1	0.539
PE	4				1

The two-sided critical values of  $r_{crit}$  are 0.195 for the level of confidence  $P = 0.95$ , and 0.254 for  $P = 0.99$ .

# *Correlation levels*

There is no indication for systematic errors which could cause correlation in the chemical analysis/testing. Probably the root cause is in the technological conditions.

• Note, the observed  $r_{ij}$  are positive only: treated quantities are expressed in % of labeled amount of the component in a tablet. There is no limitation of the sum of such values.

To assess influence of the **observed correlation** ( $r_{ij}$ ) on the total risk values, they are compared with those calculated for **independent** test results ( $r_{ij}=0$ ), and also with the values obtained supposing much **stronger correlation** ( $r_{ij}=0.7$ ). Thus, three levels of the correlation are studied.

# *Multivariate likelihood function*

The likelihood function is modelled by a multivariate **normal** pdf. Then, the likelihood covariance matrix for test results  $c_{im}$  equal, for example, to the prior means  $\mu_i = m_i$ , was:

$$S_{cm1} = \begin{pmatrix} 7.7120 & 0.8129 & 0.9655 & 1.3617 \\ 0.8129 & 7.4835 & 2.3662 & 3.0617 \\ 0.9655 & 2.3662 & 7.7353 & 4.1530 \\ 1.3617 & 3.0617 & 4.1530 & 7.6747 \end{pmatrix},$$

where the diagonal elements are variances  $u_i^2 = (0.028 c_{im})^2$ ; and the covariances are  $cov_{ijm} = r_{ij} \cdot u_i \cdot u_j$ ,  $i \neq j$ . The values  $u_i^2$  and  $cov_{ijm}$  are in squared % of labeled amount.

# *Joint posterior function*

The joint posterior function was calculated as a multivariate **normal** pdf having the following parameters:

$$S_{\text{post}k} = (S_{\text{ck}}^{-1} + n_{\text{rep}} S_{\text{cm}k}^{-1})^{-1} \text{ and}$$

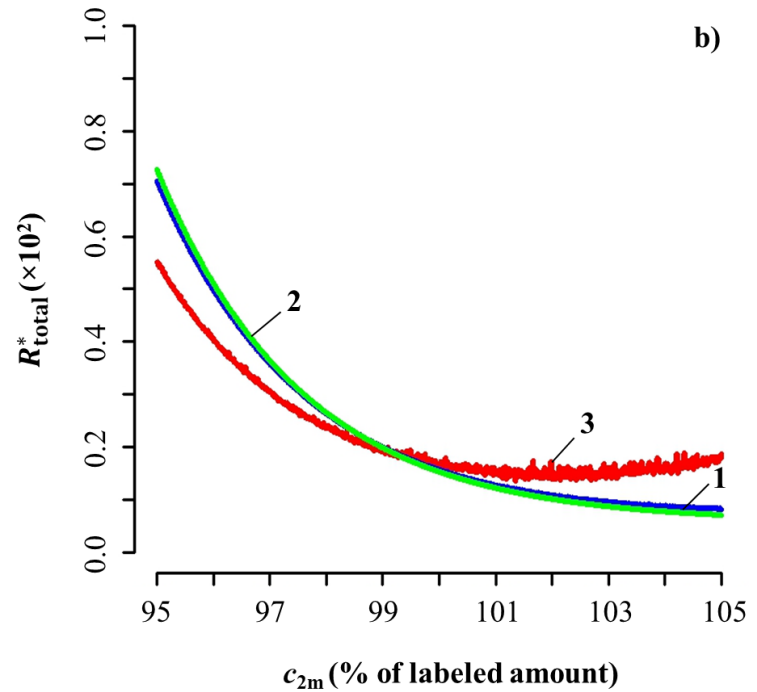
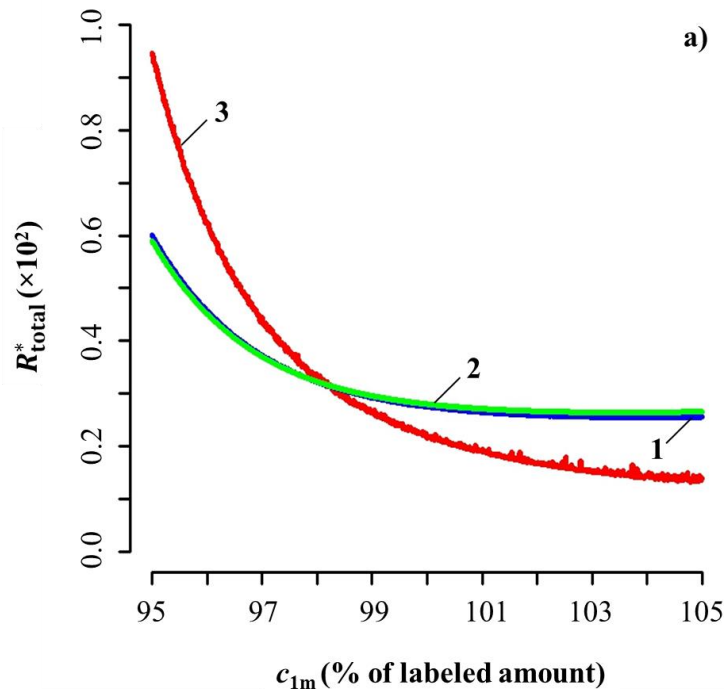
$$c_{\text{post}k} = S_{\text{post}k} (S_{\text{ck}}^{-1} c + n_{\text{rep}} S_{\text{cm}k}^{-1} \bar{c}_m)^{-1},$$

where  $S_{\text{post}k}$  and  $c_{\text{post}k}$  are the posterior covariance matrix and the vector of the posterior means, respectively;  $c$  is the vector of the prior mean values  $[\mu_1, \mu_2, \mu_3, \mu_4]$ .

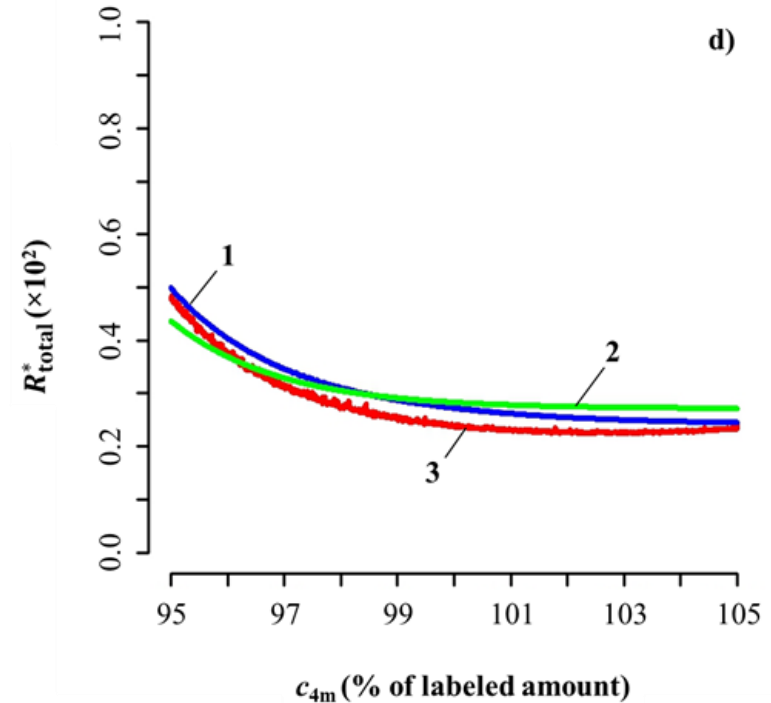
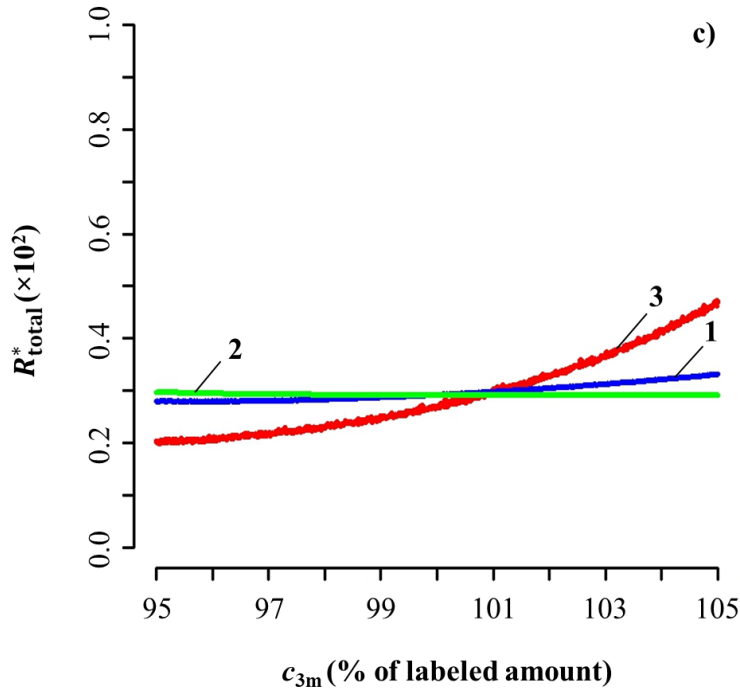
A. Gelman et al. *Bayesian Data Analysis*, 3<sup>d</sup> edn., Charman & Hall/CRC, Boca Raton, 2014

# Total specific risks $R_{\text{total}}^*$

$R_{\text{total}}^*$  was evaluated as the joint posterior function of actual values  $c_i$  of a specific lot lying outside the multivariate specification domain, when the vector of test results  $c_{im}$ , obtained for the lot, is inside this domain.



# Total specific risks $R_{\text{total}}^*$ - continuation



The correlation influence on the total risk values is not easily predictable. When correlation is strong, it may lead either to decreasing or increasing of the total risks, depending on the actual values of the test results.



# *Conformity assessment of chemical composition of a Pt-Rh alloy*



I. Kuselman, F. Pennechi, R. da Silva, D.B. Hibbert  
(2018) ACQUAL 23: ...

# *A dataset and tolerance limits*

Test results of  $N = 100$  batches of PtRh 92.5-7.5 alloy, produced during two years at the same plant, were used.

The standard of the alloy marks sets  $T_{Li}$  and  $T_{Ui}$  of contents  $c_i$  of the four following alloy components:

$i = 1$ ) Pt mass fraction,  $T_{L1} = 92.2 \% \leq c_1 \leq 92.8 \% = T_{U1}$ ;

$i = 2$ ) Rh mass fraction,  $T_{L2} = 7.3 \% \leq c_2 \leq 7.7 \% = T_{U2}$ ;

$i = 3$ ) content of three precious impurities - Au, Ir and Pd - as sum of mass fractions,  $c_3 \leq 0.12 \% = T_{U3}$ ;

$i = 4$ ) content of 8 impurities, both the three precious and non-precious (Fe, Pb, Si, Sn and Zn), as sum of mass fractions,  $c_4 \leq 0.18 \% = T_{U4}$ .

# *Sub-domain of feasible alloy compositions*

The limits  $T_{Li}$  and  $T_{Ui}$ , form a multivariate specification domain of permissible alloy compositions. However, there are also two constraints of the mass balance to be satisfied:

- 1) sum of the contents of the base components and the 8 impurities should be  $c_1 + c_2 + c_4 = 100\%$ , and
- 2) the content of the 3 precious impurities cannot exceed the content of the 8 precious and non-precious impurities in the same alloy, i.e.  $c_3 \leq c_4$ .

These constraints lead to a multivariate sub-domain of feasible alloy compositions.

# *Test methods*

Samples are cut down from an alloy ingot as a strip for preparation of two disks for wavelength dispersive X-ray fluorescence (XRF) determination of Rh content. For  $c_{2m}$  in the specification interval  $[7.3 - 7.7] \%$ ,  $u_2 = 0.04 \%$ .

Samples in form of a band from the same ingot are prepared for optical atomic emission spectrometry (AES) for determination of contents of the impurities;  $u_3 = 0.18c_{3m}$  and  $u_4 = 0.18 c_{4m}$ .

A test result of the Pt content is calculated as a difference between 100 % and test results:  $c_{1m} = (100 - c_{2m} - c_{4m}) \%$ ;

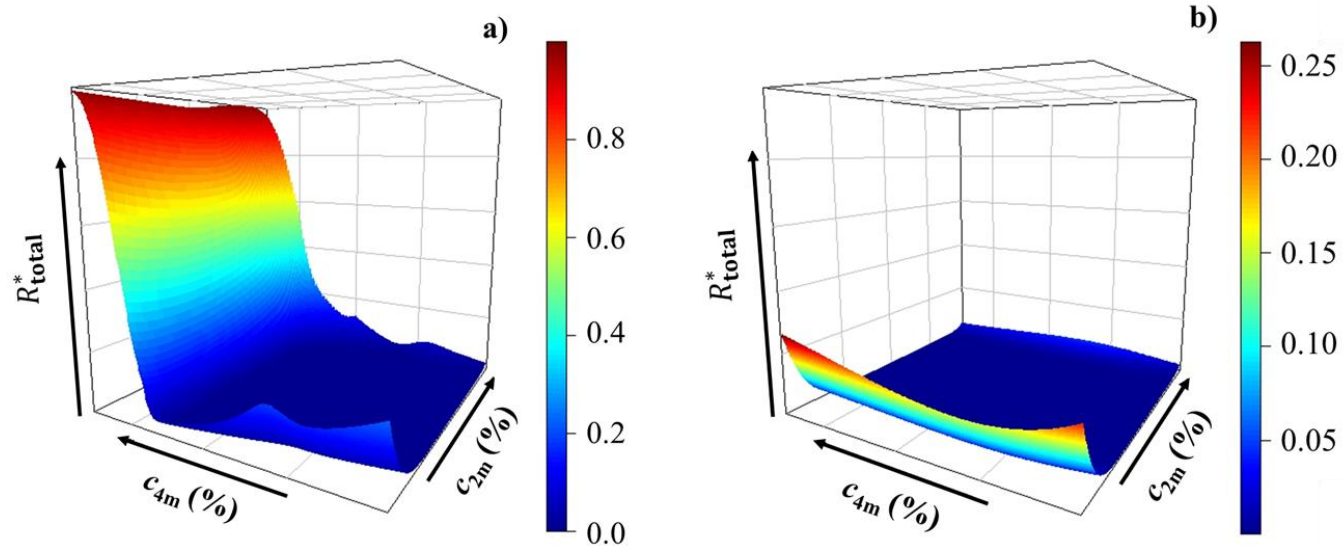
$$u_1 = \sqrt{u_2^2 + u_4^2} = \sqrt{0.04^2 + (0.18 c_{4m})^2}.$$

# *Pearson's correlation coefficients $r_{ij}$ of the test results*

Component	Index	Pt	Rh	Three impurities	Eight impurities
	$i \backslash j$	1	2	3	4
Pt	1	1	-0.967	-0.469	-0.467
Rh	2		1	0.239	0.228
Three impurities	3			1	0.970
Eight impurities	4				1

The critical values of the coefficient  $r_{\text{crit}}$  (when the correlation sign is known) for  $N - 2 = 98$  degrees of freedom are 0.197 for level of confidence  $P = 0.95$ , and 0.256 for  $P = 0.99$ .

# *Total specific consumer's risk $R_{total}^*$*



- a)** the four-component scenario at  $92.2 \% \leq c_{1m} = 100 - c_{2m} - c_{4m} \leq 92.8 \%$  and  $c_{3m} = c_{4m} / 1.16 \leq 0.12 \%$ ;
- b)** the two-component scenario:  $c_{1m}$  and  $c_{3m}$  are not taken into account as correlated with  $c_{2m}$  and  $c_{4m}$ , respectively.

# *Total global consumer's risk $R_{\text{total}}$*

The risk  $R_{\text{total}}$  is calculated on the base of integrals of the product of the prior pdf and the likelihood function. The obtained small value  $R_{\text{total}} = 5.6 \times 10^{-7}$  is an indication of a reliable system of the alloy quality assurance.

To assess the influence of correlation on  $R_{\text{total}}$ , the risk was estimated for a simulated case of uncorrelated contents of the components. This simulation was carried out by setting all correlation coefficients  $r_{ij} = 0$  ( $i \neq j$ ), hence transforming  $S_{C(k)}$  and  $S_{cm(k)}$  into diagonal matrices. The result  $R_{\text{total}} = 6.2 \times 10^{-3}$  was four orders of magnitude greater than for the correlated contents.

# *Conclusions*

When particular CA for each component of a material or object is successful, the total probability of a false decision concerning conformity of **the material as a whole** may still be significant.

The total probability of a false decision is depending on the **two quality parameters** of chemical analytical/test results:

- 1) associated measurement uncertainty, and
- 2) correlation among test results for different components.