IUPAC/CITAC Guide: Evaluation of risks of false decisions in conformity assessment of a substance or material with a mass balance constraint (IUPAC Technical Report)

IUPAC Technical Report

Francesca R. Pennecchi, Ilya Kuselman* and D. Brynn Hibbert

IUPAC/CITAC Guide: Evaluation of risks of false decisions in conformity assessment of a substance or material with a mass balance constraint (IUPAC Technical Report)

https://doi.org/10.1515/pac-2022-0801
Received August 9, 2022; accepted August 18, 2023

Abstract: A Bayesian multivariate approach to the evaluation of risks of false decisions on conformity of chemical composition of a substance or material due to measurement uncertainty is adapted to cases for which the composition is subject to a mass balance constraint. The constraint means that sum of the actual ("true") values of the composition component contents under conformity assessment is equal to 1 (or 100 %) or another positive value less than 1 (less than 100 %). As a consequence, the actual values of the component contents are intrinsically correlated. Corresponding measured values of the component contents are correlated also. Any correlation can influence evaluation of risks of false decisions in conformity assessment of the substance or material chemical composition. A technique for appropriate evaluation of the relevant risks, including evaluation of the conformance probability of a subject or material composition, is discussed for different scenarios of the data modeling, taking into account all observed correlations. A Monte Carlo method is applied in the R programming language for the necessary calculations. Examples of evaluation of the risks are provided for conformity assessment of chemical composition of a platinum-rhodium alloy, pure potassium trioxidoiodate, a sausage, and synthetic air.

Keywords: Chemical composition; conformity assessment; correlation; mass balance constraint; measurement uncertainty; risk of false decision.

CONTENTS

Article note: Sponsoring body: IUPAC Analytical Chemistry Division: see more details on page 1230 for IUPAC project number and 1251 for Division membership.

***Corresponding author: Ilya Kuselman**, Independent Consultant on Metrology, 4/6 Yarehim St., 7176419 Modiin, Israel, e-mail: ilya.kuselman@bezeqint.net. https://orcid.org/0000-0002-5813-9051
Francesca R. Pennecchi, Istituto Nazionale di Ricerca Metrologica (INRIM), Strada delle Cacce 91, 10135 Turin, Italy. https://orcid.org/0000-0003-1328-3858
D. Brynn Hibbert, School of Chemistry, UNSW Sydney, Sydney, NSW 2052, Australia. https://orcid.org/0000-0001-9210-2941

1 Introduction

Conformity assessment of the chemical composition of an item (batch or lot of a substance or material, environment compartment, entity, etc.) is performed by comparison of measured values of contents of the item components with their tolerance (specification) limits [1]. Besides the tolerance limits, which are related to the actual ("true") values of the content, narrower acceptance limits, leading to more stringent requirements for the measured values, can be applied taking into account measurement uncertainty [2] to the advantage of the consumer's interests. When the producer's interests are dominant, applicable acceptance limits are wider than tolerance limits due to the same measurement uncertainty. In both cases, the decision rules (does the test item conform or not?) are based on comparing the measured contents with the acceptance limits [3, 4].

In current practice, decision rules are often based on direct comparison of measured values with the tolerance limits. The reason is that these limits already take into account the measurement uncertainty. Thus, the tolerance and the acceptance limits coincide. The measurement uncertainty is constrained for this purpose by a requirement to use a specific standardized or fit-for-purpose validated chemical analytical/measurement method [5].

Risks as probabilities of false decisions on conformity of chemical composition of a multicomponent material or object due to measurement uncertainty have been defined using a Bayesian approach in the IUPAC/CITAC Guide [6]. According to Bayes' theorem, a prior knowledge of the measurand values (actual contents of components in the items) and new information acquired during the measurement are combined in a posterior probability density function (pdf) [1]. Such a posterior pdf accumulates what is known about the measurands in the items and allows evaluation of the risks of false decisions in their conformity assessment, caused by the measurement uncertainty. The probability of accepting the item after comparing a measured value with the acceptance limit of a content of an item component, when it should have been rejected, is called the "consumer's risk," whereas the probability of falsely rejecting the item is the "producer's risk."

It is shown in the Guide [6] that successful conformity assessment for each particular component of a material is not always sufficient for a final conclusion on the material quality, as the total probability of a false decision (total consumer's risk or producer's risk) concerning the material as a whole might still be significant. This relates both to a specific batch or other item of a substance or material (specific consumer's and producer's

risks), and to a population of the items (global consumer's and producer's risks). A model of the total probability of false decisions for cases of independent actual ("true") contents of the components and corresponding measured values was formulated based on the law of total probability. In such a case, total risk was evaluated as a combination of the particular risks in conformity assessment of components of the item. For a more complicated task, *i.e.*, for a greater number of components under control, the total risk increases.

Compositions of an item with actual and measured property values (the components' contents) correlated due to natural physicochemical or technological reasons, were modeled by multivariate distributions. Then, a total global risk of a false decision on conformity of the chemical composition was evaluated by calculation of multivariate integrals of corresponding joint pdf of actual (true) and measured property values. A total specific risk was evaluated using the multivariate integral of the posterior density function of property values on the multivariate specification (tolerance) domain of the item compositions.

When component contents of a substance or material are subject to a mass balance constraint (sum of the actual contents—mass fractions or amount fractions—is 1 (or 100 %)), measured values of the component contents are called "compositional data." These data are intrinsically correlated because of the constraint, and the relevant correlation was named by Karl Pearson in 1897 as "spurious" [7–9]. Spurious correlations are a kind of mathematical property of the data, not related to physicochemical interdependence of true values of component contents of a material. Compositional Data Analysis (CoDA) based on an isometric logratio transformation of the original measured values was developed in the 1980s by John Aitchison [10–12]. Logratios of amalgamations of test results for pairs of components (parts of the data)—pairwise logratios—were recently used for practical simplification of CoDA [13–16]. There are several examples of CoDA applications in geochemistry [17, 18], agriculture and environmental analysis [19–22], bioinformatics [23, 24], forensic science [25, 26], materials [27], and other fields [28].

There is also a strong message in the literature (*e.g.*, in refs. [17, 18, 29]) stressing how traditional statistical techniques may produce inadequate results if applied to raw compositional data without suitable transformation. However, the relevant techniques of CoDA are not implemented widely in metrology in chemistry and conformity assessment. Spurious correlations may influence evaluation of measurement uncertainty of compositional data (*e.g.*, of standard gas mixtures [30, 31]) and quantification of risks of false decisions due to measurement uncertainty in conformity assessment of a substance or material. A special case is characterization of materials based on a mass balance [32–35], for example matrix reference materials, as well as evaluation of purity of substances and corresponding (pure) certified reference materials.

A problem of application of CoDA in this field is that a subset of the actual component contents may not be related directly to the mass balance, *i.e.*, be outside of the multidimensional simplex; for example, when a sum of impurities' actual contents is a member of the mass balance constraint, but contents of individual impurities or a subset of those impurities also have to undergo conformity assessment. Since each actual component content participating in conformity assessment has its tolerance/specification limits, the mass balance constraint forms a multivariate sub-domain of the feasible substance or material compositions in the domain of specification limits. Moreover, spurious correlation coefficients may be comparable or less (in absolute values) than coefficients of correlations caused by natural properties of the materials and technological reasons, and in general be not easily distinguishable from them [36].

In this Guide, evaluation of risks due to measurement uncertainty in conformity assessment of a substance or material with a mass balance constraint is discussed in detail, based on the Bayesian multivariate approach and Monte Carlo method developed in the R programming language. Examples of evaluation of risks provided in Annex A are related to conformity assessment of a platinum-rhodium alloy, pure potassium trioxidoiodate, a sausage, and synthetic air.

1.1 Scope and field of application

This Guide is developed for study of the risks caused by measurement uncertainties in conformity assessment of a multicomponent substance or material with a mass balance constraint. It will be helpful also for correct risk

management in a factory producing multicomponent materials, for environmental monitoring of several substances, and for similar tasks.

The document is intended for quality control, measurement and testing chemical analytical laboratories, metrologists and analytical chemists, specialists involved in the laboratory accreditation activity, laboratory customers, quality managers, and regulators.

1.2 Terms and definitions

Terms and definitions used in this Guide correspond to JCGM 106 [1], JCGM 200 (VIM) [37], ISO/IEC 17000 [38], ISO 3534 [39], IUPAC Gold Book [40], and IUPAC Recommendations [41].

The term "content" is applied in the Guide [6], Sec. 1.2, for a quantity of a component of an item subject to conformity assessment expressed per unit mass of the item. In the present document, "content" is used as a generic term for both extensive and intensive properties [40], *i.e.*, mass, volume, amount of substance, numbers of entities, and their fractions: mass fraction, volume fraction, amount fraction, and fraction of number of entities.

The most relevant definitions relating to the risks in conformity assessment of a multicomponent substance or material due to measurement uncertainty are given in the Guide [6]. The definition of mass balance and related terms are given below.

1.2.1 Mass balance

1.2.1.1 Mass balance constraint

Constraint (limitation, restriction) of compositions of a substance or material, object or system based on invariance of mass of a substance or material, object or system which is closed to transfer of matter and energy, according to the law of conservation of mass.

NOTE 1: The law of conservation of mass has a long history [42, 43].

NOTE 2: Conservation of mass implies conservation of extensive and intensive values of component contents (mass, volume, amount of substance, numbers of entities, and their fractions). The term *mass balance* is used to cover all of these measures.

NOTE 3: Extensive values of component contents can be transformed into corresponding intensive values (fractions) using a closure operation. This operation consists of dividing a component extensive content value by sum of the content values of all the components of a substance or material, object or system subject to the mass balance constraint.

NOTE 4: For any composition, sum of actual values of all component content fractions should be equal to 1 (or 100 %).

NOTE 5: When not all component contents are under consideration, the sum of their actual values can be equal to a positive fraction, less than 1 (or 100 %).

NOTE 6: The sum of measured values of component contents can differ from 1 (or 100 %, or another positive fraction set for actual values of the component contents under consideration) because of the associated measurement uncertainty.

1.2.2 Mass balance approach

1.2.2.1 Mass balance method

Application of *mass balance* for metrology, quality, and conformity assessment purposes.

EXAMPLE 1: Indirect purity evaluation of a substance or material, *i.e.*, evaluation of a content of its main component, by measurement of contents of impurities and subtraction of their sum from 1 (or 100 %); described as a concept in IUPAC Recommendations [41], clause 3.22. In such a case, the purity uncertainty is calculated

taken into account the measurement uncertainties of the impurities' contents at a selected coverage probability.

EXAMPLE 2: Evaluation of a bias of measured values of the component contents in a substance, material object or system by calculation of a deviation of the sum of the measured values from 1 (or 100 %) and comparison of the deviation with its associated measurement uncertainty at a selected coverage probability.

NOTE: There are also applications of *mass balance* in chemical engineering [44, 45].

1.3 Symbols

A	multivariate region (domain) of acceptable measured values
A_i	acceptance interval of measured values $c_{i\mathrm{m}}$ of i-th component content
$A_{\mathrm{L}i}$	lower limit of the acceptance interval of $c_{i\mathrm{m}}$
$A_{\mathrm{U}i}$	upper limit of the acceptance interval of $c_{i\mathrm{m}}$
C	normalizing constant
c	superscript "complementary"
c_i	actual ("true") content of i-th component in an item
$\boldsymbol{c}$	vector of c_i, $i = 1, 2, \ldots, n$
$c_{i\mathrm{m}}$	measured value of content of i-th component
$\boldsymbol{c}_{\mathbf{m}}$	vector of $c_{i\mathrm{m}}$, $i = 1, 2, \ldots, n$
clo	closure operation
cov_{ij}	covariance of contents of components $i \neq j$
cov_{ijclo}	covariance after closure operation
D	subscript meaning "direct"
$\boldsymbol{e}_{\mathbf{m}}$	vector of errors $e_{i\mathrm{m}}$, $i = 1, 2, \ldots, n$
f_i	probability density function of truncated normal distribution of i-th component content
ϕ	volume fraction
ϕ_{norm}	probability density function of normal distribution
Φ_{norm}	cumulative normal distribution function
g	posterior probability density function
g_0	prior probability density function
h	likelihood function
i, j	subscripts of the components in the range from 1 to n
k	mass balance constraint (of a sum of component contents)
m	subscript of measured component content
μ_i	theoretical (prior) mean of i-th component content
$\boldsymbol{\mu}$	vector of the prior mean values μ_i, $i = 1, 2, \ldots, n$
M	number of Monte Carlo simulations
n	number of components
N	number of batches (items)
p	coverage probability
P_{conf}	conformance probability
R^*_{ci}	particular specific risk
$R^*_{ci\,(\mathrm{c})}$	particular specific consumer's risk
$R^*_{ci\,(\mathrm{p})}$	particular specific producer's risk
R_{ci}	particular global risk
$R_{ci(\mathrm{c})}$	particular global consumer's risk
$R_{ci(\mathrm{p})}$	particular global producer's risk
R^*_{total}	total specific risk
$R^*_{\mathrm{total}\,(\mathrm{c})}$	total specific consumer's risk
$R^*_{\mathrm{total}\,(\mathrm{p})}$	total specific producer's risk
R_{total}	total global risk
$R_{\mathrm{total(c)}}$	total global consumer's risk
$R_{\mathrm{total(p)}}$	total global producer's risk
r_{ij}	correlation coefficient of contents of components $i \neq j$
r_{ijclo}	correlation coefficient after closure operation

s_i^2	experimental (sampling) variance of *i*-th component content
σ_i^2	theoretical (population) variance of *i*-th component content
S^n	simplex of compositional data (of *n* component contents)
T	superscript meaning “transpose” (makes the columns of the new matrix the rows of original)
T	multivariate tolerance region (domain)
T_i	tolerance interval of *i*-th component content c_i
T_{Li}	lower limit of the tolerance interval of *i*-th component content c_i
T_{Ui}	upper limit of the tolerance interval of *i*-th component content c_i
u_i	measurement uncertainty of *i*-th component content
u_{ij}	covariance term of measurement uncertainties $i \neq j$
$\boldsymbol{U}$	covariance matrix of measurement uncertainties u_i
v	number of components, for which the measured values are out of their acceptance intervals
$\boldsymbol{V}$	covariance matrix of actual component contents
x_i	amount fraction of component *i*
$\boldsymbol{x}$	vector of amount fractions x_i

1.4 Abbreviations

ANOVA	analysis of variance
CITAC	Cooperation on International Traceability in Analytical Chemistry
CCQM	Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology
CoDA	Compositional Data Analysis
D	dimension
EP	European Pharmacopoeia
Eurachem	A Focus for Analytical Chemistry in Europe
HPLC	high performance liquid chromatography
IEC	International Electrotechnical Commission
ICH	The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
JCGM	Joint Committee for Guides in Metrology
NMI	National Metrology Institute
MC	Monte Carlo
pdf	probability density function
SOP	standard operating procedure
TMN	truncated multivariate normal distribution
TN	truncated (univariate) normal distribution
VIM	International Vocabulary of Metrology

2 Classification of correlations

2.1 Metrologically related correlations

A number of techniques are used during development of a chemical analytical method to overcome possible correlations between measured values $c_{i\mathrm{m}}$ of *i*-th contents of different components, $i = 1, 2, \ldots, n$ (main as well as impurities) of a substance or material, when their actual (“true”) values c_i are not correlated. Some of these techniques include extraction of target components (analytes) from a sample and chromatographic separation of an analyte from other components of the sample [46]. Chemometrics software is applied for separation of spectral signals and multivariate calibrations of spectrometers [47]. Sample digestion [48] and standard additions of an analyte to a sample [49] are used for calibration of a measuring system to overcome multiplicative matrix effects, and so on. There are validation requirements for “analytical selectivity” of a standard operating procedure (SOP), as a performance characteristic, to prove the procedure’s fitness for purpose. IUPAC Recommendations [50] and

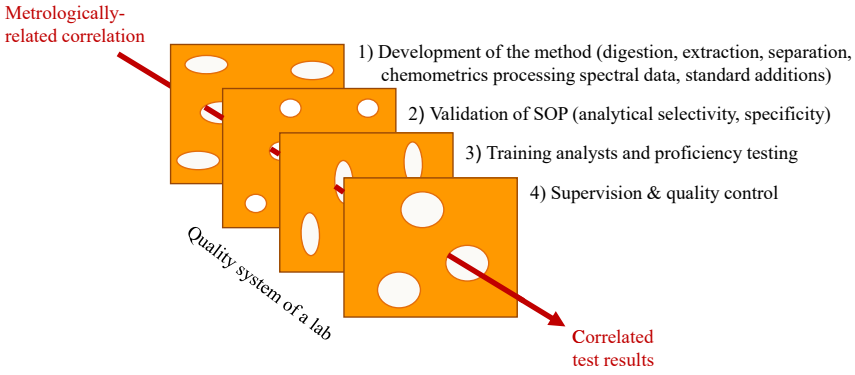

Fig. 1: Metrologically related correlation arising in a multicomponent chemical analysis—a Swiss-cheese model [36].

the Eurachem Guide [5] define that "analytical selectivity relates to the extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behavior". This definition is consistent with "selectivity of a measuring system" in JCGM 200 [37]. The corresponding procedure validation parameter in the pharmaceutical industry is termed in ICH Guideline [51] "specificity." A procedure not able to answer the requirements of selectivity or specificity is not "fit-for-purpose" and cannot be applied to the given task. Validation of the SOP, training analysts and proficiency testing, supervision and quality control are elements of the quality system of an analytical laboratory that should prevent human error causing metrologically related correlation, unless, as shown in Fig. 1, the Swiss-cheese model [52] lines up an error in each element. Therefore, correlations that have arisen in the routine measurement process should be in general negligible and measured content values c_{im} of two or more components of the same item are expected to be metrologically independent. In practice, if statistically significant correlation between measured values of contents of components of the same item is detected, analysis cannot be continued without a thorough chemical analytical inspection of the reason for the correlation in the laboratory.

For example, when a medication is tested routinely with a pharmacopeial HPLC procedure, correlations might be related to the resolution of the chromatography column used, not able to separate the analytes completely, and the column must be replaced by another one. In the IUPAC/CITAC Guide [52], such an event is rated as a skill-based mistake or omission error (lapse).

2.2 Natural and technological correlations

Measured values c_{im} of component contents of a substance or material are inevitably correlated when their actual values c_i are correlated. Correlations of true component contents c_i can be caused by natural physicochemical properties of substances, such as stoichiometry [53–55], and by technological reasons in a material production [56, 57]. These correlations are taken into account in conformity assessment using "conventional" multivariate statistical methods [6].

2.3 Correlations due to mass balance constraint

When n actual component contents c_i of a substance or material are subject to a mass balance constraint ($\sum_{i=1}^{n} c_i = 1$ or 100 %), measured values of the component contents are called "compositional data." Compositional data may be depicted in a multidimensional simplex, as in Fig. 2, in which, in general, Euclidean geometry cannot be blindly applied.

When two component contents are under the mass balance constraint, *e.g.*, content c_1 of the main component and content c_2 of the sum of impurities in a pure substance, $c_1 + c_2 = 100$ % [58]. The constraint means here that larger content c_2 of impurities leads to lower purity c_1 of the substance. The Pearson correlation coefficient,

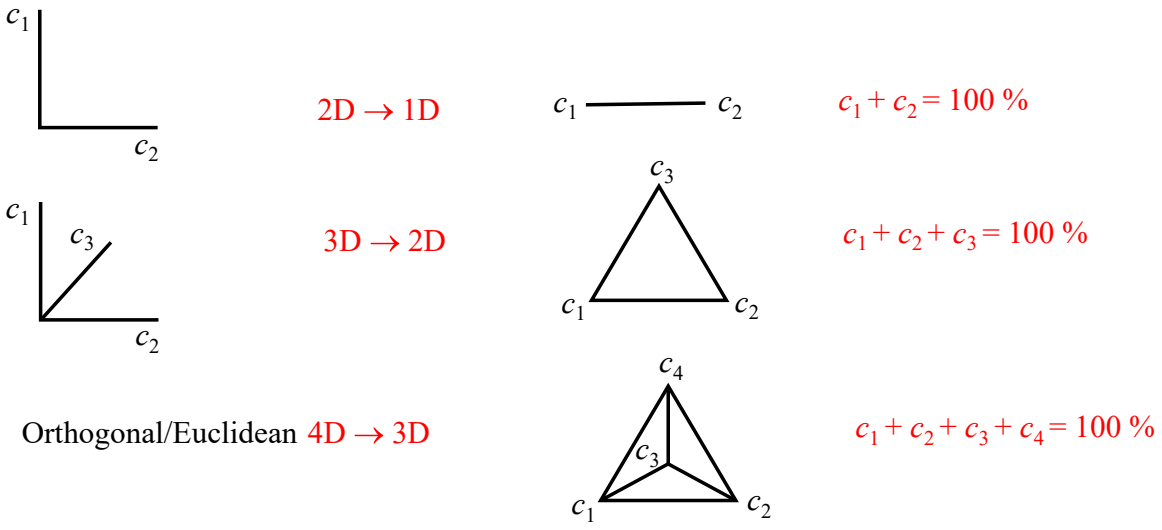

Fig. 2: Orthogonal coordinates of Euclidean space and simplices. The variables c_i (%) are contents of a material components $i = 1, 2, 3$, and 4; each simplex vertex corresponds to $c_i = 100$ %; D is dimension. In each case, mass balance constraint reduces the number of dimensions by one [36].

evaluated analytically for purity vs. content of impurities from this constraint is $(r_{12})_{an} = -s_2^2/s_2^2 = -1$, where s_2 is the standard deviation of c_2 coinciding with s_1, the standard deviation of c_1. Possible metrologically related, natural and technological sources of correlation, as well as measurement uncertainties are able to decrease the absolute value of the correlation coefficient between corresponding measured values, but it usually remains close to 1.

The correlation coefficients evaluated analytically from the constraint $c_1 + c_2 + c_3 = 100$ % were described for a platinum-rhodium (PtRh) alloy in ref. [57], when Pt content c_1 was calculated from the mass balance using Rh content c_2 and sum of the eight impurities' content c_3. For Pt vs. Rh contents, the correlation coefficient was $(r_{12})_{an} = -s_2^2/(s_2^2 + s_3^2) = -0.961$, and for Pt vs. sum of eight impurities' contents it was $(r_{13})_{an} = -s_3^2/(s_2^2 + s_3^2) = -0.276$. The standard deviations s_i were evaluated from a dataset of compositions of a sufficient number of the alloy batches to give useful statistics. The absolute values of $(r_{12})_{an}$ and $(r_{13})_{an}$ were found to be even smaller than those of r_{12} and r_{13}, respectively, calculated directly from the experimental data. Thus, the observed correlations were caused as much by the natural chemical origin of the raw materials used in the alloy production, as by the mass balance constraint.

There are statistical techniques, mostly for applications in geochemistry, that allow evaluation of correlation coefficients related to a natural correlation, separately from "spurious" correlation due to a mass balance constraint, based on logratio transformation and centered logratio transformation of the data [59, 60].

Note that a mass balance constraint $(\sum_{i=1}^{n} c_i = 1$ or 100 %) is applicable to actual components contents c_i, while sum of measured values c_{im} cannot in general be exactly equal to 1 (100 %) or another constant because of measurement uncertainties u_{im} associated with measured values c_{im} [2]. Application of a mass balance constraint to measured values may lead to a distortion of measurement information. This is critically important for evaluation of risks in conformity assessment due to measurement uncertainties. Therefore, measured values c_{im}, being correlated because of correlation of actual values c_i, are not subjected directly to a mass balance constraint.

Summarizing the classification of data according to their correlations, the data "cloud" shown as a Venn diagram on Fig. 3 is divided into a green part of independent (hence, uncorrelated) actual and measured values, and a blue part of values correlated because of metrologically related, natural and/or technological reasons. These data were analyzed in the IUPAC/CITAC Guide [6] as subject to conformity assessment.

The correlated data containing compositional actual values, highlighted by the dark blue core, are studied for the purposes of conformity assessment in this Guide.

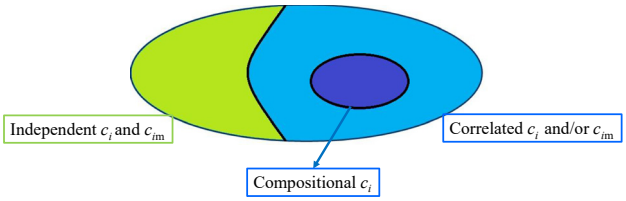

Fig. 3: Venn diagram of the data. Independent (uncorrelated) actual and measured values are shown as the green part of the data "cloud," and correlated values as the blue part of the cloud. The dark blue core represents correlated data containing compositional actual values.

3 A Bayesian multivariate approach

The chemical composition of a multicomponent material or object is considered to conform when the actual content c_i of each i-th component under control in the item, $i = 1, 2, \ldots, n$, is within its tolerance/specification interval $T_i = [T_{Li}, T_{Ui}]$, where T_{Li} and T_{Ui} are lower and upper limits of the interval, respectively. To decide whether the material or object conforms or not, the measured value c_{im} is compared with the limits of the acceptance interval $A_i = [A_{Li}, A_{Ui}]$, where A_{Li} and A_{Ui} are lower and upper limits of this interval, respectively, taking into account the measurement uncertainty associated with c_{im}. Note that the acceptance interval may be narrower than the tolerance interval when defending interests of a consumer and wider than the tolerance interval when defending interests of a producer. The tolerance and acceptance intervals may also coincide as discussed in the Introduction to this Guide.

3.1 Properties of a composition

The vector of actual ("true") component contents $\boldsymbol{c} = [c_1, c_2, \ldots, c_n]$ describes the n-component (n-part) composition of a material or substance. The compositional space is the simplex

$$S^n = \{\boldsymbol{c} = [c_1, c_2, \ldots, c_n] \,|\, c_i > 0, i = 1, 2, \ldots, \mathrm{n}\,;\, \textstyle\sum_{i=1}^{n} c_i = k\}, \tag{1}$$

where k is usually equal to 1 (100 %), but might be any other positive constant less than 1 (or less than 100 %).

As c_i are positive quantity ratios, a vector of content values $\boldsymbol{c}$ multiplied by any positive constant contains the same information as the original one, *i.e.*, represents the same composition and can be considered as an equivalence class. This property is termed "scale invariance." In other words, if $\boldsymbol{c}$ is scaled by a constant, *e.g.*, content values c_i changing from parts-per-unit to percentages, the information which $\boldsymbol{c}$ conveys is completely equivalent. Therefore, it is natural to select a representative of the equivalence class to facilitate data analysis and interpretation of corresponding results. This selection is formalized by the closure operation:

$$clo\,(\boldsymbol{c}) = \left[\frac{k \cdot c_1}{\sum_{i=1}^{n} c_i}, \ldots, \frac{k \cdot c_n}{\sum_{i=1}^{n} c_i}\right]. \tag{2}$$

Compositional data do not depend on the order of contents c_i within vector $\boldsymbol{c}$. The order does not influence conclusions of any analysis on $\boldsymbol{c}$. This is known as the "permutation invariance" property. Another property is "subcompositional coherence," which means that analyses concerning a subset of the component contents should not depend on the remaining part. If a metric is used to compare two compositions on the simplex, the distance between the two should be greater than or equal to that obtained by comparing any couple of corresponding subcompositions, a condition called "subcompositional dominance." This property is used to measure distances between compositions and subcompositions following the rule of a projection that distances become smaller in a projection.

Note, ordinary correlations (metrologically related, natural and/or technological) depend on the subcomposition considered and violate subcompositional coherence. The ordinary Euclidean distance between the vectors cannot be evaluated here, as both scale invariance and subcompositional dominance are violated [10]. Also, logratio transformations applied in CoDA are not helpful in such cases, since they take into account spurious correlations only [36].

3.2 Modeling composition

As vector $\boldsymbol{c} = [c_1, c_2, \ldots, c_n]$ represents the actual ("true") n-component composition of a substance or material, component contents c_i, $i = 1, 2, \ldots, n$, are the measurands in conformity assessment, and $\boldsymbol{c}$ is the vector of the measurands. Using a multivariate Bayesian approach [6], knowledge about a composition $\boldsymbol{c}$ can be modeled by a random multivariate posterior variable and expressed in terms of its probability density function (pdf). Such a pdf combines prior knowledge about the measurands and new information acquired during the measurements:

$$g(\boldsymbol{c}|\boldsymbol{c}_\mathbf{m}) = C g_0(\boldsymbol{c}) h(\boldsymbol{c}_\mathbf{m}|\boldsymbol{c}), \tag{3}$$

where $g(\boldsymbol{c}|\boldsymbol{c}_\mathbf{m})$ is the posterior pdf; C is a normalizing constant; $g_0(\boldsymbol{c})$ is the multivariate prior pdf; and $h(\boldsymbol{c}_\mathbf{m}|\boldsymbol{c})$ is the multivariate likelihood function taking into account the measurement uncertainties and possible covariance terms.

A large enough dataset of results of testing items of the same material produced at the same factory, or results of monitoring the same environmental compartment, can be used for modeling the prior pdf $g_0(\boldsymbol{c})$. The assumption is that the actual content values are approximated by the test/measurement results adequately, since measurement uncertainty is negligible in comparison with item-to-item (batch-to-batch) variations caused by changes of conditions of the material production, environmental conditions, etc. Based on this assumption, theoretical normal and gamma distributions in JCGM 106 [1], as well as normal, lognormal and Weibull distributions in IUPAC/CITAC Guide [61], were considered as prior pdfs in univariate conformity assessment. Multivariate normal and lognormal distributions in IUPAC/CITAC Guide [6] were used in conformity assessment of multicomponent materials and an environmental object. Adequacy of the prior as the theoretical distribution of c_i to the empirical distribution of $c_{i\mathrm{m}}$ from the dataset is proved by their goodness-of-fit testing [62]. When more than one theoretical distribution is adequate, the simplest is preferable as the prior.

If there is no detailed prior knowledge about distribution of the component content in the tested items, the prior pdf is vague. According to the principle of maximum entropy, a multivariate normal distribution is usually considered as the prior pdf for the vector of actual component contents $\boldsymbol{c}$ when the vector of mean values $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{V}$ constitute the only available information about the vector quantity [63]. In a case of a mass balance constraint, the data properties need to be taken into account when assigning a corresponding pdf. Since the vector $\boldsymbol{c}$ is non-negative, and the first two moments of the distribution (the expected value and variance) are known, the maximum entropy distribution is a truncated multivariate normal distribution, not a logarithmic one [64]. This is also the case for an elliptical truncation region [65], when the pdf is supposed to be nonzero-valued just for $(\boldsymbol{c} - \boldsymbol{\mu})^\mathrm{T} \boldsymbol{V}^{-1} (\boldsymbol{c} - \boldsymbol{\mu}) \le k$, and superscript T means "transpose" and makes the columns of the new matrix the rows of the original.

Therefore, in the present Guide, a truncated multivariate normal (TMN) distribution on the nD region $[0, k]^n$ is employed in modeling the prior pdf of $\boldsymbol{c}$. This is TMN($\boldsymbol{\mu}$,$\boldsymbol{V}$), where $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are the location and scale parameters, respectively, of the original normal pdf from which the TMN distribution arises by truncation on $[0, k]^n$. For the univariate case, for example, the TMN pdf of c_i on the interval $[0, k]$ is

$$f_i(c_i) = \frac{\varphi_\mathrm{norm}\left(\frac{c_i - \mu_i}{\sigma_i}\right)}{\sigma_i\left(\Phi_\mathrm{norm}\left(\frac{k - \mu_i}{\sigma_i}\right) - \Phi_\mathrm{norm}\left(\frac{-\mu_i}{\sigma_i}\right)\right)}, \tag{4}$$

where μ_i and σ_i are the mean and standard deviation (the location and scale parameters) of the normal pdf from which this TMN distribution arises; φ_norm is the pdf of the standard normal distribution and Φ_norm is its cumulative distribution function. Outside the interval $[0, k]$, the pdf f_i is equal to zero.

The likelihood function $h(\boldsymbol{c}_{\mathrm{m}}|\boldsymbol{c})$ is recovered from knowledge about the measurement uncertainty (pdf of the measured values $c_{i\mathrm{m}}$ at the actual value c_i) available from the understanding on the measurement process and the measurement procedure. As the prior pdf, the likelihood function $h(\boldsymbol{c}_{\mathrm{m}}|\boldsymbol{c})$ can be also modeled on the base of a TMN of the measured values, having location parameter equal to the vector $\boldsymbol{c}$ of actual ("true") values, *i.e.*, TMN $(\boldsymbol{c},\boldsymbol{U})$, where $\boldsymbol{U}$ is the covariance matrix of measurement uncertainties u_i and covariance terms u_{ij} whose corresponding correlation coefficients are the same as for $\boldsymbol{V}$. Correlation between measured content values $c_{i\mathrm{m}}$ and $c_{j\mathrm{m}}$ is the same as between the corresponding actual ones c_i and c_j, when no other (metrologically related) correlation arises in the measurement process.

The multivariate truncated distribution accommodates the mass balance constraint of the data to lie on the region $[0, k]^n$ and is promising for overcoming the above-mentioned problems of application of CoDA for evaluation of risks of false decisions in conformity assessment.

3.3 Modeling prior distribution

Three scenarios of modeling the prior distribution are considered below for simplicity for the case of a three-component composition described by the vector $\boldsymbol{c}$ and corresponding mass balance constraint $c_1 + c_2 + c_3 = 100$ %:

1) Modeling all the actual values of the components' contents by applying the closure operation $clo(\boldsymbol{c})$ to $\boldsymbol{c}$ which follow a TMN$(\boldsymbol{\mu},\boldsymbol{V})$ distribution on the 3D region $[0, 100]^3$.
2) Modeling actual values of two components' contents and deriving the third: $[c_2, c_3]$ follows a bivariate TMN$(\boldsymbol{\mu}_{23}, \boldsymbol{V}_{23})$ distribution on the 2D region $[0, 100]^2$, where subscript "23" indicates that the applied operation refers just to components 2 and 3. The scale parameter $\boldsymbol{V}_{23}$ involves the correlation coefficient r_{23} and c_1 is deterministically calculated as $c_1 = 100\ \% - c_2 - c_3$, disregarding possible negative values.
3) Sequential modeling, when c_2 follows a TN(μ_2, σ_2), a univariate normal distribution truncated on the interval $[0, 100]$; $c_3|c_2$ has a conditional probability distribution following a TN(μ_3, σ_3) on the interval $[0, 100 - c_2]$; and again deterministically $c_1|(c_2, c_3) = 100\ \% - c_2 - c_3$.

The constraint equal to 100 % is shown in this modeling for example and can be replaced by 1, when components contents are expressed in fractions of 1.

Two parameters are considered to assess suitability of the models: 1) the correlation matrix encompassing correlation coefficients r_{ij} of the multivariate prior pdf, and 2) the pdf coverage probability p, calculated as its integral over the multivariate tolerance/specification region. By definition, p is the probability that the prior values of all the three variables lie within corresponding tolerance intervals. Equivalence or dissimilarity of this parameter's values from model to model indicate that the multivariate pdfs generated by the various models are the same or different, respectively. Results of calculating r_{ij} and p obtained by 10^7 Monte Carlo (MC) simulations for each model, are reported for the PtRh alloy in Annex A to this Guide, Example 1.

In general, modeling 1 is suitable for matrix reference materials, standard gas mixtures, food products, etc., *i.e.*, for materials whose composition is tested completely (all component contents are measured). Modeling 2 is applicable for substances or materials in which the main component content is not measured directly using physicochemical measurement method(s) but calculated based on a mass balance constraint. Sequential modeling 3 is more complicated, since the sequence of c_2 and c_3 can be exchanged and that may lead to different results as the conditional probabilities are different. Moreover, the number of such models increases significantly with the number of components $n > 3$. This modeling does not allow consideration of types of correlation other than spurious and cannot be implemented adequately when natural (stoichiometric) and/or technological correlations are significant. On the other hand, modeling 3 could be a helpful tool for understanding sources of the observed correlation, disentangling its spurious part from the rest.

3.4 Likelihood modeling

The modeling of the likelihood function for measured content values $\boldsymbol{c}_\mathbf{m}$ is based on the idea that an appropriate pdf with zero expectation is chosen for an error vector $\boldsymbol{e}_\mathbf{m}$ and then translated to the vector of actual ("true") content values $\boldsymbol{c}$ generated for the prior. Therefore, vector $\boldsymbol{c}_\mathbf{m}$ is recovered as $\boldsymbol{c}_\mathbf{m} = \boldsymbol{c} + \boldsymbol{e}_\mathbf{m}$. The covariance matrix $\boldsymbol{U}$ associated with $\boldsymbol{c}_\mathbf{m}$ contains the squared measurement uncertainties u_i and the covariance terms u_{ij} whose corresponding correlation coefficients are the same as for $\boldsymbol{V}$. The error $\boldsymbol{e}_\mathbf{m}$ pdf, having zero mean, is the same as that for the likelihood when translated into the vector $\boldsymbol{c}$. The truncation limits are chosen such that each component of the vector $\boldsymbol{c}_\mathbf{m}$ must be between 0 % and 100 %: $\mathbf{0} < \boldsymbol{c}_\mathbf{m} = \boldsymbol{c} + \boldsymbol{e}_\mathbf{m} < \mathbf{100}$, where $\mathbf{0}$ is a vector with each component equal to 0 %, and $\mathbf{100}$ is a vector with each component equal to 100 %. Hence, a vector of errors is extracted between $-\boldsymbol{c}$ and $(\mathbf{100} - \boldsymbol{c})$, where $\boldsymbol{c}$ is drawn from the prior pdf. For ease of computation, the $\boldsymbol{c}$ values in these truncation limits are approximated with corresponding mean values $\boldsymbol{\mu}$. Thus, the modeling of the likelihood for a three-component composition follows that of the prior:

1) Modeling the likelihood of all the measured component contents: $\boldsymbol{e}_\mathbf{m}|\boldsymbol{c}$ is modeled as a TMN$(0, \boldsymbol{U})$ distribution on the region $[-\mu_1, 100 - \mu_1]$ % × $[-\mu_2, 100 - \mu_2]$ % × $[-\mu_3, 100 - \mu_3]$ %, where $\boldsymbol{c}$ is approximated by $\boldsymbol{\mu}$.
2) Modeling the likelihood of two measured component contents and deriving the third: in order to generate $c_{2\mathrm{m}}$ and $c_{3\mathrm{m}}$, $[e_{2\mathrm{m}}, e_{3\mathrm{m}}]|\boldsymbol{c}_{23}$ is taken as a bivariate TMN$(0, \boldsymbol{U}_{23})$ distribution on the region $[-\mu_2, 100 - \mu_2]$ % × $[-\mu_3, 100 - \mu_3]$ %, where $\boldsymbol{c}_{23}$ is approximated by $\boldsymbol{\mu}_{23}$, and $\boldsymbol{U}_{23}$ involves the correlation coefficient r_{23}. Then, $c_{1\mathrm{m}} = 100\ \% - c_{2\mathrm{m}} - c_{3\mathrm{m}}$ is calculated directly, disregarding its possible negative values.
3) Sequential modeling the likelihood: in order to generate $c_{2\mathrm{m}}$, $e_{2\mathrm{m}}|c_2$ is represented by a univariate TN$(0, u_2)$ distribution on $[-\mu_2, 100 - \mu_2]$ % interval, where c_2 is approximated by μ_2, and $e_{3\mathrm{m}}|\boldsymbol{c}_{23}$ by a univariate TN$(0, u_3)$ on the interval $[-\mu_3, 100 - \mu_3 - c_{2\mathrm{m}}]$ %, for the $c_{3\mathrm{m}}$ generation, where c_3 is approximated by μ_3. Then, $c_{1\mathrm{m}} = (100 - c_{2\mathrm{m}} - c_{3\mathrm{m}})$ % is calculated directly, as in the previous modeling.

Thus, the modeling of both prior pdf and likelihood function is based on the experimental data, which contain the initial knowledge about compositions of the substance or material, ahead of the conformity assessment. However, the likelihood function characterizes the measurement process with corresponding measurement uncertainties, mimicked by $\boldsymbol{e}_\mathbf{m} = [e_{1\mathrm{m}}, e_{2\mathrm{m}}, e_{3\mathrm{m}}]$. Therefore, sum of the measured values $\boldsymbol{c}_\mathbf{m} = [c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}]$ may be not equal to 100 %, and so the closure operation is not applied here.

As previously, when components contents are expressed in fractions of 1, the value 100 % above is replaced by 1.

3.5 Modeling posterior probability density function

Once the prior pdf and likelihood function are modeled, the posterior pdf $g(\boldsymbol{c}|\boldsymbol{c}_\mathbf{m})$ of the actual content values $\boldsymbol{c}$ at the measured values of the component contents $\boldsymbol{c}_\mathbf{m}$ can be calculated by eq. 3 as the normalization of the product $g_0(\boldsymbol{c})h(\boldsymbol{c}_\mathbf{m}|\boldsymbol{c})$.

This posterior pdf contains an updated state of knowledge about the composition of the substance or material. Since the posterior pdf predicts further actual content values (after accumulation of the dataset used for modeling the prior pdf) taking into account what may happen during the measurement process *via* the likelihood function, the closure operation is not appropriate for the posterior data. In other words, the sum of the calculated (predicted) actual component content values may differ from 100 % because any predicted value has its associated prediction uncertainty [66–68].

4 Evaluation of risks of false decisions

When the prior pdf, likelihood function and posterior pdf are modeled, evaluation of the risks as probabilities of false decisions on conformity of a chemical composition of a substance or material to the tolerance/specification

limits can be performed by integration of the multivariate posterior or the joint pdf, for calculating total specific and global risks, respectively, as described in the IUPAC/CITAC Guide [6]. If contents of one only particular component is under conformity assessment, eqs. 5 and 6 below reduce to the univariate counterparts described in JCGM 106 [1]. When the contents of the components are independent, eqs. 5 and 6 simply involve the evaluation of particular risks for each component [6].

4.1 Global risks

Total global consumer's risk R_c and the total global producer's risk R_p are, respectively:

$$R_c = \int_{T^c}\int_A g_0(\boldsymbol{c})h(\boldsymbol{c}_m|\boldsymbol{c})\,d\boldsymbol{c}_m d\boldsymbol{c} \text{ and } R_p = \int_T\int_{A^c} g_0(\boldsymbol{c})h(\boldsymbol{c}_m|\boldsymbol{c})\,d\boldsymbol{c}_m d\boldsymbol{c}, \tag{5}$$

where T is the multivariate tolerance/specification domain $T_1 \times T_2 \times \ldots \times T_n$, A is the multivariate acceptance domain $A_1 \times A_2 \times \ldots \times A_n$, and the integral symbols indicate multiple integrals. Superscript "c" of T in the formula for R_c means "complementary" for at least one T_i, whereas the integration with respect to all c_{im} is performed within A. The superscript "c" of A in the formula for R_p means "complementary" for at least one A_i, whereas the integration with respect to all c_i is performed within T.

4.2 Specific risks

The total specific consumer's risk R_c^* and the total specific producer's risk R_p^* are, respectively:

$$\begin{aligned} R_c^* &= 1 - \int_T g(\boldsymbol{c}|\,\boldsymbol{c}_m)d\boldsymbol{c} \text{ when } \boldsymbol{c}_m \text{ is in } A, \text{ and} \\ R_p^* &= \int_{T_1}\cdots\int_{T_v}\int_0^{100}\ldots\int_0^{100} g(\boldsymbol{c}|\boldsymbol{c}_m)d\boldsymbol{c} \text{ when } c_{im}, 1 \le i \le v, \text{ are outside } A. \end{aligned} \tag{6}$$

Here, R_c^* is the probability that at least one of corresponding actual content values c_i is actually outside its tolerance interval T_i, when all the measured content values c_{im} are within their acceptance intervals A_i (false conforming). Thus, R_c^* is equal to one minus the probability that all c_i are within the tolerance domain T when all c_{im} conform, *i.e.*, are within the acceptance domain A. The integration limits from 0 to 100 (%) can be replaced by the limits from 0 to 1, when components contents are expressed in fractions of 1.

Symbol v in eq. 6 for R_p^* indicates the number of those components, $1 \le v \le n$, whose measured content values c_{im} are outside their acceptance intervals A_i. Hence, the vector $\boldsymbol{c}_m$, being out of the acceptance domain A for those v components of the substance or material, is rejected as non-conforming. For simplicity, and without losing generality, the measured values c_{im} outside their acceptance intervals are the first v values. Given that these v measured values do not conform, R_p^* is the probability that all corresponding actual values are inside their tolerance intervals, hence the item (batch or lot of the substance or material) does satisfy its specifications and rejection of this item is a false decision.

5 Implementation

5.1 Computational details

To evaluate the total global risks in the case of variables related to a mass balance, prior pdf, and likelihood function are modeled as in Sec. 3.3 and Sec. 3.4, respectively. Then, a large number M ($M = 10^7$ in the examples in Annex A) of MC simulations of values drawn from the joint pdf $g_0(\boldsymbol{c})h(\boldsymbol{c}_m|\boldsymbol{c})$ are performed simulating

prior values from $g_0(\boldsymbol{c})$ and corresponding likelihood values from $h(\boldsymbol{c}_\mathbf{m}|\boldsymbol{c})$. The three models for the prior and likelihood discussed in the previous sections give three ways to obtain the joint pdf. The function "rtmvnorm" [69] is applied in the first and the second model, and the function "rtnorm" [70] in the third model. Finally, the joint pdf is numerically represented by M vectors consisting of actual and measured values $[c_1, \ldots, c_n, c_{1\mathrm{m}}, \ldots, c_{n\mathrm{m}}]$ in each MC simulation. Also, the evaluation of the coverage probabilities, according to the models of prior pdfs, was performed in the R programming language using the library "compositions" [71].

The total global producer's risk R_p is evaluated considering the fraction of simulated vectors (among the M generated) in which all the actual values c_i are within the corresponding tolerance region T, while at least one of the measured values $c_{i\mathrm{m}}$ is out of its acceptance interval A_i. The total global consumer's risk R_c is estimated as the fraction of vectors in which all the measured values $c_{i\mathrm{m}}$ are within the acceptance region A, while at least one of the actual values c_i is out of its tolerance interval T_i.

For each specified vector of measured values $[c_{1\mathrm{m}}, \ldots, c_{n\mathrm{m}}]$, the integrals of the posterior pdf in eq. 6 for calculation of total specific producer's risk R_p^* and consumer's risk R_c^* involve ratio of multiple integrals of the joint pdf of variables $[c_1, \ldots, c_n, c_{1\mathrm{m}}, \ldots, c_{n\mathrm{m}}]$ with respect to variables c_i ($i = 1, \ldots, n$) over appropriate domains. The numerical evaluation of such integrals is performed by simulation of M random vectors $[c_1, \ldots, c_n, c_{1\mathrm{m}}, \ldots, c_{n\mathrm{m}}]$, generated according to the prior modeling for c_i values and the likelihood modeling for $c_{i\mathrm{m}}$ values.

Codes developed in the R programming language (R codes) for calculations of the risks in Annex A to this Guide, Examples 1–4, have been published in papers [36, 58, 67, 68]. These examples are related to conformity assessment of a PtRh alloy, pure potassium trioxidoiodate, a sausage, and synthetic air, respectively. The specific R codes can be sent upon request to the corresponding author.

5.2 Limitations

This Guide discusses the risks of false decisions as probabilities following JCGM 106 [1] and IUPAC/CITAC Guide [6]. It is noted that some guidelines for risk management [72–74] define risk as the product of probability of an event and severity of its consequences, expressed for each case in a different way, not always quantitatively: examples include financial loss, safety and/or security changes, quality loss, aesthetic and taste worsening in a product. However, in this Guide the consequences of false decisions on conformity of a substance or material and their severity are not considered.

All limitations of the Bayesian approach listed in IUPAC/CITAC Guide [6] are relevant for this Guide also. These are the use of any model as a simplified reflection of reality; the assumption of negligible definitional uncertainty of actual component content c_i (in particular, inhomogeneity and/or instability of an item of the multicomponent substance or material leading to an increase of the standard deviation of the prior pdf and its skewness); adequacy of treatment of a dataset of item-to-item (batch-to-batch) test/measurement results for modeling a prior pdf; goodness-of-fit of experimental and theoretical distributions, etc.

Acknowledgments: The Task Group would like to thank M. Sega (Italy), A. Sobina and E. Sobina (Russia), A.A. Semenova (Russia), F. Rolle (Italy), and V. Altshul (Israel) for the help in preparation of Examples in Annex A of this Guide; Elsevier (www.elsevier.com) for permission to use material from the published papers cited in this Guide as refs. [36, 58, 67, 68].

Research funding: This work was prepared under project 2019-012-1-500 of IUPAC (cfr. **1.4**) (Funder ID: 10.13039/100006987).

Annex A: Examples

Example 1. Risks in conformity assessment of a PtRh alloy

A-1-1 Introduction

The aim of this Example is implementation of modeling and calculation of total *global* risks in conformity assessment of a substance or material with a mass balance constraint, using the PtRh alloy studied earlier [57] and described in IUPAC/CITAC Guide [6] as a multicomponent material with a complex nature of correlation among contents of components. Contents of the following three alloy components expressed in mass fractions, %, are considered here: Pt content c_1, Rh content c_2, and content c_3 of the eight impurities equal to sum of their mass fractions. Note that the content of three precious impurities (a part of the eight impurities) is also under control in conformity assessment. However, this component content is not taken into account as it is not related directly to the mass balance, $c_1 + c_2 + c_3 = 100$ % [36].

A-1-2 Experimental and tolerance limits

When X-ray fluorescence method for measurements of Rh content c_2, and optical atomic emission spectrometry for measurements of mass fractions of the eight impurities (and their sum c_3) were applied, Pt content c_1 was calculated based on the mass balance: $c_1 = 100\ \% - c_2 - c_3$.

The standard specifications for PtRh 92.5-7.5 alloy [75] include the following lower and upper tolerance limits T_{Li} and T_{Ui} of contents c_i, respectively:
$i = 1)$ $T_{L1} = 92.2\ \% \le c_1 \le 92.8\ \% = T_{U1}$;
$i = 2)$ $T_{L2} = 7.3\ \% \le c_2 \le 7.7\ \% = T_{U2}$;
$i = 3)$ $T_{L3} = 0\ \% \le c_3 \le 0.18\ \% = T_{U3}$.

A-1-3 Prior probability density function

The means μ_i and standard deviations σ_i of the marginal theoretical normal distributions, satisfying requirements of the Kolmogorov–Smirnov test of goodness-of-fit to the dataset on composition of $N = 100$ alloy batches produced during about 2 years at the same plant, are presented in Table 1.

The Pearson correlation coefficients r_{ij}, $i \neq j$, are also available in Table 1. The critical value of r_{ij} for $N - 2 = 98$ degrees of freedom and level of confidence 0.95 is 0.197 [57].

To assess the suitability of the models described in Sec. 3.3 of this Guide, two parameters were considered: (1) the correlation matrix of the multivariate prior pdf, containing r_{ij}, and (2) the pdf coverage probability p, calculated as its integral over the multivariate tolerance region, which is [92.2 %, 92.8 %] × [7.3 %, 7.7 %] × [0 %, 0.18 %].

The calculation of p using the Monte Carlo (MC) method was performed in the R programming language using the library “compositions” [71] and a Monte Carlo method with 10^7 simulations.

Table 1: Initial parameters of the prior distribution.

i / j	Component	μ_i, %	σ_i, %	r_{ij}		
				Pt 1	Rh 2	8 impurities 3
1	Pt	92.483	0.081	1	−0.967	−0.467
2	Rh	7.457	0.073	−0.967	1	0.228
3	8 Impurities	0.059	0.021	−0.467	0.228	1

For the first model, where $\boldsymbol{c}$ follows a TMN($\boldsymbol{\mu}$,$\boldsymbol{V}$) distribution on the region $[0, 100]^3$, a prior pdf as a three-variate truncated normal distribution was generated using the function "rtmvnorm" with truncation limits equal to [0 %, 0 %, 0 %] and [100 %, 100 %, 100 %]. A non-singular matrix was inserted as an approximation of matrix $\boldsymbol{V}$. Then, the function "clo" was applied for performing the closure operation with the constraint equal to 100 %.

For the second model the function "rtmvnorm" was applied to generate a bivariate truncated normal distribution with truncation limits [0 %, 0 %] and [100 %, 100 %]. Then, the third variable was computed as $c_1|(c_2, c_3) = 100\ \% - c_2 - c_3$. Rows $[c_1, c_2, c_3]$ of the matrix of the generated values, containing negative values of the third component content, were removed.

In the third model, c_2 is generated using the function "rtnorm" with the truncation lower limit 0 % and the upper limit 100 %. Then, variable $c_3|c_2$ was given by the same function, "rtnorm," but with the truncation limits 0 % and 100 %$-c_2$, *i.e.*, for each value of c_2, hence the corresponding value of c_3 was in the interval [0 %, 100 %$-c_2$]. Finally, the third variable was computed as $c_1|(c_2, c_3) = 100\ \% - c_2 - c_3$.

When the prior pdf was generated in the form of a matrix containing rows of values $[c_1, c_2, c_3]$, the coverage probability can be evaluated using function "length" to count cases where c_1, c_2, and c_3 values are simultaneously within their limits. One can also compute marginal coverage probability, taking into account the variables one by one.

Results of calculating r_{ij} and p obtained by means of 10^7 MC simulations for each model are reported in Table 2. For comparison, $p = 0.979$ was obtained when the ordinary multivariate normal distribution was applied as in ref. [57] and the original/experimental correlation coefficients shown in Table 1 were used.

Correlation coefficients in Table 2 for the prior pdf obtained by models 1 and 2 are equal, within the precision of the MC simulation, to the experimental ones in Table 1, considering that two decimal digits in the coefficient estimates can be taken as reliable when 10^7 MC simulations were performed. For both models 1 and 2, coverage probabilities slightly greater than $p = 0.979$ were obtained, meaning that the prior pdfs generated by these models were narrower than the ordinary multivariate normal pdf.

To show the influence of the closure operation on a TMN pdf, boxplots of the marginal pdfs of the three variables by model 1, before the closure operation (boxplot 1) and after it (boxplot 2), are depicted in Fig. 4.

The closure operation shifts the marginal pdf location toward smaller values, especially for Pt content c_1. More details of this pdf are shown in the histogram of c_1 in Fig. 5.

By model 3, the contents c_2 of Rh and c_3 of the impurities were each drawn from a univariate truncated normal distribution, one of which depends on the other just in its truncation limits. Pt content c_1 values were deterministically generated to satisfy the mass balance constraint. In this model, there is no place for a covariance matrix encompassing the correlation coefficients which reflect different kinds of correlation between the contents. The correlation matrix by model 3 can reveal only the spurious correlations arising from the truncation effect and the mass balance constraint. In the present example, as the sum of Rh and impurities contents is always far away from 100 %, correlation between the two due to the truncation effect is negligible. Thus, the correlation coefficient arising from spurious correlations between Rh and the impurities contents is zero, despite the fact that the experimental correlation coefficient r_{23} is 0.288. In addition, the spurious correlation coefficient for Pt and the impurities contents is different from the observed correlation coefficient, due to both the mass balance constraint

Table 2: Correlation coefficients and coverage probabilities of different prior models.

Parameters	Model		
	1	2	3
r_{12}	−0.968	−0.968	−0.962
r_{13}	−0.464	−0.464	−0.274
r_{23}	0.226	0.226	0.000
p	0.985	0.981	0.981

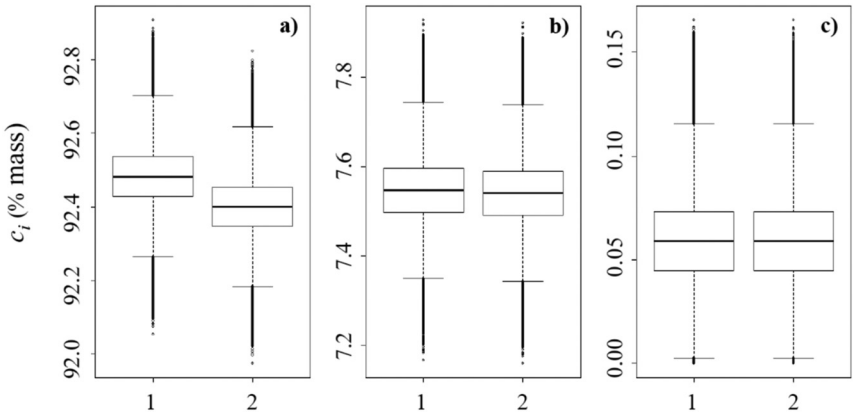

Fig. 4: Boxplots of the marginal prior probability density functions of the components' contents. Boxplot 1 is before closure operation, and boxplot 2 after it; for (a) Pt content c_1; (b) Rh content c_2; and (c) content c_3 of the eight impurities. The band near the middle of each box is the 50 % percentile (the median, equal to the mean when the distribution is symmetrical); the bottom and top of the box corresponds to 25 % and 75 % percentiles, respectively; the distance between upper and/or lower whiskers and the box is equal to the 1.5 box length [36].

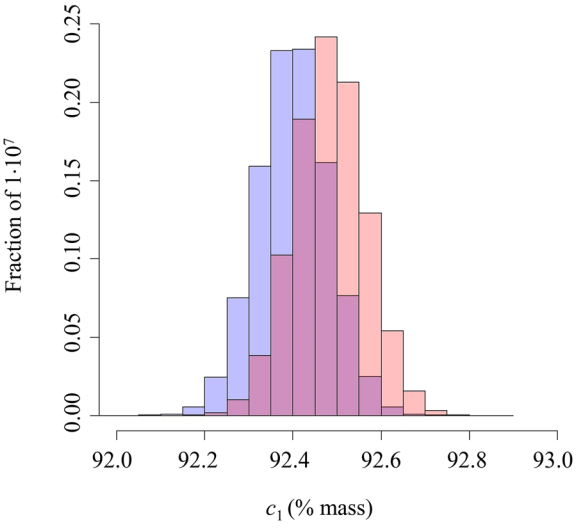

Fig. 5: Histogram of the marginal probability density functions of the Pt content c_1. Pink bars show frequencies of the Pt content (fractions of 1×10^7 MC simulations) before the closure operation, while blue bars present the frequencies after this operation, *i.e.*, after taking into account the mass balance constraint [36]. The magenta color is the result of overlaying pink bars with blue bars.

and the nature of the raw materials. Nonetheless, the coverage probability obtained for model 3 is very close or equal to those by models 1 and 2.

A-1-4 Likelihood functions

According to the propagation of uncertainty for correlated input quantities in JCGM 100 [2], the standard measurement uncertainty u_1 associated with the Pt content c_{1m}, calculated based on the mass balance constraint, was evaluated as $u_1 = (u_2^2 + u_3^2 + 2u_2u_3r_{23})^{0.5} = (0.040^2 + 0.011^2 + 2 \times 0.040 \times 0.011 \times 0.228)^{0.5} = 0.044$ (%). Here $u_2 = 0.040$ % is the standard uncertainty associated with the measured Rh content c_{2m} in its tolerance (specification) interval, and $u_3 = 0.011$ % is the standard measurement uncertainty associated with the measured mean of the impurities content $c_{3m} = 0.059$ % (the relative standard uncertainty being $u_3/c_{3m} = 0.18$) [57]. When correlation between contents of components 2 and 3 is ignored, *e.g.*, in model 3 where $r_{23} = 0$, the measurement uncertainty associated with the calculated c_{1m} is $u_1 = 0.041$ %.

Further modeling the likelihood functions for vector of measured content values $\mathbf{c_m}$ is performed according to the three models described in detail in Sec. 3.4 of this Guide.

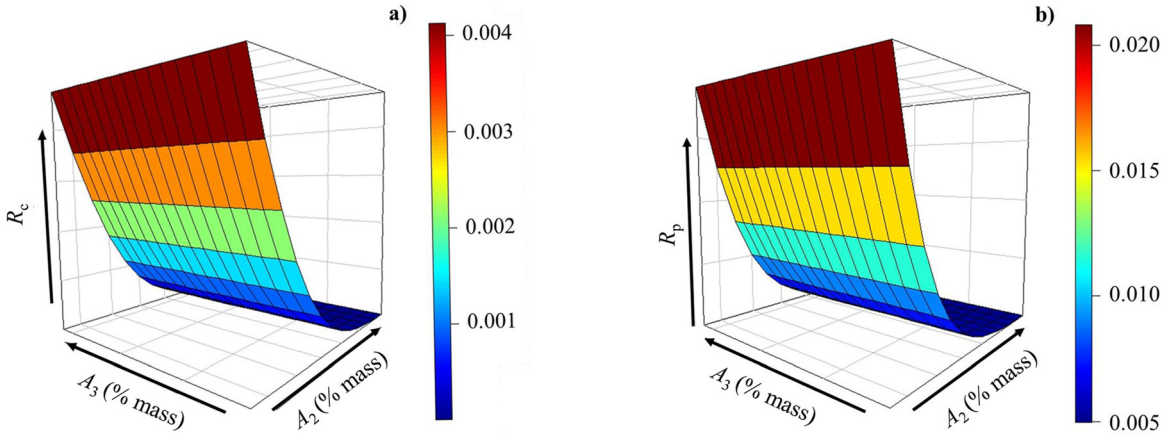

Fig. 6: Surfaces of the total global consumer's risk R_c (plot "a") and producer's risk R_p (plot "b"). The arrows show the acceptance intervals from $A_2 = T_2$ and $A_3 = T_3$, at the axes beginning, to the acceptance intervals with limits differing from the tolerance limits by three times the relevant measurement uncertainties at the axes end, when $A_1 = T_1$ for Pt content. The color column bar in plot 6a gives indication of the R_c values between the minimum and the maximum on the surface, *i.e.*, from 3×10^{-6} to 4.7×10^{-3}; in plot 6b the bar is for the R_p values between 4.9×10^{-3} and 2.4×10^{-2} [36].

A-1-5 Global risks

For the three discussed models, the total global consumer's risk by eq. 5 of this Guide at the acceptance limits equal to the tolerance limits ($A = T$) was the same $R_c = 4.7 \times 10^{-3}$. A similar result $R_c = 5.1 \times 10^{-3}$ was reported in the paper [57] for a scenario when Pt, Rh and the eight impurities contents only were taken into account: the slight difference in the risk values is due to different numerical implementations of the risk calculation. For more details, the R_c dependence on A_2 and A_3 (at $A_1 = T_1$) was also studied, as shown in Fig. 6a for model 1. The arrows show the acceptance intervals from $A_2 = T_2$ and $A_3 = T_3$, at the beginning of the axes, to the acceptance intervals with limits differing from the tolerance limits by three times the relevant measurement uncertainty ($3u_i$), at the axes' ends.

Specifically, in Fig. 6a, these are the lower acceptance limit $A_{L2} = T_{L2} + 3u_2 = 7.3 + 3 \times 0.04 = 7.42$ (%) and the upper acceptance limit $A_{U2} = T_{U2} - 3u_2 = 7.7 - 3 \times 0.04 = 7.58$ (%). Since the upper only tolerance limit T_{U3} is set for the content of the impurities, at the end of the A_3 axis $A_{L3} = 0$ and $A_{U3} = T_{U3} - 3u_3 = 0.18 - 3 \times (0.059 \times 0.18) = 0.15$ (%), where 0.059 % is the value of the content prior mean μ_3 of the impurities (Table 1). The color column bar in the plot gives indication of the risk R_c values between the minimum and the maximum on the surface, from 3×10^{-6} to 4.7×10^{-3}. One can see on the plot that the influence of the impurities' content on the risk is negligible under these conditions. That is because the content prior mean $\mu_3 = 0.059$ % is far enough from the tolerance limit $T_{U3} = 0.18$ %.

The dependences of R_c on the acceptance intervals for the second and the third models are very similar (R_c varying from 4×10^{-6} to 4.7×10^{-3} and R_c from 6×10^{-6} to 4.6×10^{-3}, respectively). These results show that the mass balance influence on R_c in the small sub-domain of feasible alloy compositions is practically negligible.

The total global producer's risk by eq. 5 at $A = T$ was $R_p = 2.4 \times 10^{-2}$ for models 1 and 2, and a bit smaller $R_p = 2.0 \times 10^{-2}$ for model 3. The producer's risk values were not discussed in ref. [57] but have been calculated, for sake of comparison, according to the approach proposed there, leading to the same value $R_p = 2.4 \times 10^{-2}$. The R_p dependence on A_2 and A_3 (at $A_1 = T_1$) is presented for model 1 in Fig. 6b, where, at the end of the axes, $A_{L2} = T_{L2} - 3u_2 = 7.3 - 3 \times 0.04 = 7.18$ (%) and $A_{U2} = T_{U2} + 3u_2 = 7.7 + 3 \times 0.04 = 7.82$ (%), and $A_{L3} = 0$ and $A_{U3} = T_{U3} + 3u_3 = 0.18 + 3 \times (0.059 \times 0.18) = 0.21$ (%), respectively. The color column bar gives indication of R_p values from 4.9×10^{-3} to 2.4×10^{-2}. The influence of the impurities' content on the producer's risk is minor at these conditions, as it is for the consumer's risk for the same reason.

The dependence of R_p on the acceptance intervals for the second model is very similar as well, its values being practically the same: R_p from 4.8×10^{-3} to 2.4×10^{-3}. However, the third model produces smaller values of R_p from 3.7×10^{-4} to 2.0×10^{-3}.

The obtained results show that the mass balance influence on the consumer's and producer's risks, evaluated with models 1 and 2, is not visible in the small sub-domain of feasible alloy compositions. The plots by these two models, like in Fig. 6, allow choice of acceptance limits corresponding to the suitable risks of the consumer(s) and the producer. The departure of results by model 3, with respect to the results obtained with models 1 and 2, is caused by the fact that model 3 does not take into account correlations other than spurious (natural and technological). Therefore, this model is helpful not only to separate spurious correlation from the observed/experimental (cumulative) correlation, but also to indicate how significant the difference between them in a material conformity assessment is.

Example 2. Risks in conformity assessment of pure KIO_3

A-2-1 Introduction

This Example demonstrates an implementation of the Bayesian methodology for evaluation of total *specific* risks of false decisions on conformity of a substance or material due to measurement uncertainty, using Monte Carlo simulations and taking into account a mass balance constraint of the data. Measurement results (measured values and associated measurement uncertainty) obtained for characterization of a specific potassium trioxidoiodate (KIO_3) batch are analyzed and discussed [58].

A-2-2 Experimental and tolerance limits

A purchased batch of synthesized potassium trioxidoiodate, 1.5 kg, was divided in vials of about 10 g and studied for the batch homogeneity and a year-long stability. The KIO_3 content/purity c_1 in the batch, expressed as mass fraction of KIO_3, %, was measured in this study by the direct method based on the coulometric titration of oxidants. The mean measured value was c_{1mD} = 99.966 %. The standard deviation of the purity caused by the batch inhomogeneity, as well as the stability standard deviation, were statistically negligible in comparison with the standard measurement uncertainty of the direct method u_{1D} = 0.007 % (subscript D means "direct").

The indirect method was also used for the determination of purity c_1, based on the mass balance: $c_1 + c_2$ = 100 %, where c_2 is content of the impurities, *i.e.*, sum of their mass fractions, %. The impurities under control were substances insoluble in water; nitrogen-containing compounds; iodide and diiodine (free iodine); sulfate, chloride, bromide and trioxidochlorate(1-) (chlorates) ions; sodium, iron, and other metals. Inductively coupled plasma mass spectrometry was applied for measurement of mass fractions of elements as impurities, and ion chromatography for measurement of mass fractions of ionic impurities. Mass fractions of seventy-one impurities were measured. The potassium excess was evaluated using the concept of electroneutrality of an ion system (the system should maintain equality of positive and negative charges [76]) and the law of conservation of charge (invariance of the total electric charge in an isolated system [77]). Substances insoluble in water were not detected, neither iodide ions and free iodine, chlorates and nitrogen containing dioxidonitrate(1-) and trioxidonitrate(1-) (nitrite and nitrate) ions, were also, and so not taken into further account. The mean measured content of the impurities was c_{2mD} = 0.025 %. The purity calculated based on the mass balance was c_{1mI} = 100 % − c_{2mD} = 99.975 %, where subscript I means "indirect." No correlation was observed between measured values of the impurities' mass fractions. The combined standard measurement uncertainty associated with c_{2mD} and c_{1mI}, was $u_{2D} = u_{1I}$ = 0.005 % [58].

The lower tolerance/specification limit of purity c_1, required for a candidate reference material of this type, was T_{L1} = 99.9 % [78]. The upper tolerance limit of content c_2 of the impurities (the sum of their mass fractions) was, accordingly, T_{U2} = (100 − 99.9) % = 0.1 %.

The acceptance limits for a candidate reference material of potassium trioxidoiodate were not set, and the tolerance/specification limits were used instead. In other words, acceptance limits were coincidental here with tolerance limits, $A = T$.

A-2-3 Prior probability density function

The prior purity pdf was approximated by a univariate truncated normal distribution TN(μ_1, σ_1) on the interval [0 %, 100 %], where location parameter μ_1 = 99.95 % is the mean of the tolerance purity interval [99.9 %, 100 %] and scale parameter σ_1 = 0.015 % is the standard deviation of a batch purity, assumed equal to the target (standard) measurement uncertainty [78, 79]. Note that in this Example the scale parameter does not characterize variability of conditions of a batch production, but reflects the variability of selection of raw material appropriate for development of the reference material.

Because of the mass balance constraint, for each value c_1 of the purity, simulated by randomly sampling from TN(μ_1, σ_1), a corresponding content value c_2 of the impurities is deterministically evaluated as $c_2 = 100 - c_1$. This approach follows Model 2 for prior pdfs described in Sec. 3.3 of this Guide. Incidentally, for just two involved variables, this model degenerates into the sequential Model 3. The c_2 pdf is the specular reflection of the c_1 pdf on the complementary range of the content values. Hence, the probability for c_2 values to be greater than T_{U2} = 0.1 % is the same as that for c_1 values to be smaller than T_{L1} = 99.9 %.

A-2-4 Likelihood functions

There are two tasks of the evaluation of specific risks and corresponding two scenarios of likelihood modeling:

1) Both contents of purity and impurities are under conformity assessment. One of the two quantities is measured directly, while the other is recovered from the former through the mass balance constraint.
   - 1a) Purity c_{1mD} is measured directly, while content of the impurities is evaluated as $c_{2mI} = 100\ \% - c_{1mD}$. In this scenario, the likelihood for c_{1mD} is modeled by a truncated normal distribution on [0 %, 100 %], with location parameter equal to the actual value c_1 and scale parameter equal to measurement uncertainty u_{1D}. This modeling follows Model 2 for likelihood functions described in Sec. 3.4 of this Guide.
   - 1b) Purity c_{1mI} is evaluated indirectly from the measured content c_{2mD} of the impurities, *i.e.*, $c_{1mI} = 100\ \% - c_{2mD}$. The likelihood for c_{2mD} is modeled by a truncated normal distribution on [0 %, 100 %], with location parameter equal to the actual value c_2 and scale parameter equal to measurement uncertainty u_{2D}. This modeling follows also Model 2 as in the scenario 1a) above.
2) Both c_{1mD} and c_{2mD} are measured directly (and independently), and both undergo conformity assessment. For modeling their likelihood functions, the univariate truncated normal distributions TN(c_1, u_{1D}) and TN(c_2, u_{2D}) on the interval [0 %, 100 %] are considered. Note, the measured values c_{1mD} and c_{2mD} are not mandated to satisfy the mass balance constraint, as the sum of c_{1mD} and c_{2mD} may be less than 100 %, because not all the impurities are detected, or greater than 100 % due to measurement uncertainties associated with the measured values.

A scenario when the directly measured purity c_{1mD} is the only quantity under conformity assessment (a univariate task), is also discussed for further comparison.

A-2-5 Specific risks

When modeling prior $g_0(c_1)$ and likelihood $h(c_{1mD}|c_1)$ was based on truncated normal distributions and followed scenario 1a, integrals in eq. 6 were simplified according to JCGM 106 [1] for the univariate case with coinciding lower tolerance and acceptance limits. These integrals were evaluated by means of the R "integrate" function for adaptive quadrature of c_2 [80].

Note that for scenario 1a the total consumer's risk R_c^* coincides with the particular consumer's risk R_{c1}^* relevant to the directly measured purity c_{1mD}. Indeed, when $c_{1mD} > T_{L1}$, corresponding content of impurities

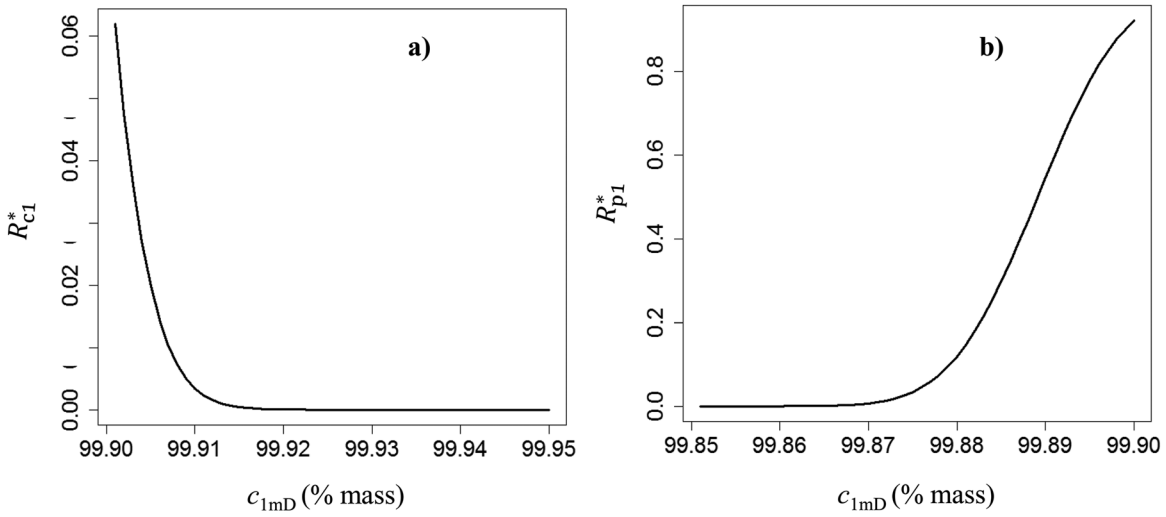

Fig. 7: Dependences of the particular specific consumer's risk R^*_{c1} (a) and producer's risk R^*_{p1} (b) on the measured directly purity value c_{1mD} in the univariate scenario. The ordinates are the risks expressed as probabilities in fractions of one, and the abscissae are measured values, % mass [57].

calculated as $c_{2mI} = 100\ \% - c_{1mD}$ is smaller than T_{U2}, and the candidate reference material is conforming. The total risk that this material is actually not conforming is equal to one minus the probability that both the posterior values $c_1|c_{1mD}$ and $c_2|c_{2mI}$ are within their corresponding tolerance limits. However, $c_1|c_{1mD} > T_{L1}$ when and only when $c_2|c_{2mI} < T_{U2}$ by definition of the prior pdf and the likelihood function for content c_2 of the impurities.

The described reasoning holds as well for the total specific producer's risk R^*_p, hence being equal to R^*_{p1}. The same kind of considerations applies also to scenario 1b. The difference between the risk values in scenarios 1a and 1b is caused just by the difference in uncertainties u_{1D} and u_{2D} associated with the directly measured values of the components' contents c_{1mD} and c_{2mD}, respectively.

The particular specific consumer's risk and producer's risk are plotted in Fig. 7a and b, respectively.

In the present study, "consumer" is the Regulator defending interests of laboratories which will purchase the reference material, applying the regulation [78]. The consumer's risk is the probability of the event when the measured (and certified later) purity value is $c_{1mD} \geq 99.90\ \%$, while the actual ("true") value is $c_1 < 99.90\ \%$. Such actual purity value is not recommended for verification of measuring instruments by the regulation and may cause further problems in a laboratory activity. However, this consumer's risk shown in Fig. 7a is practically zero, when $c_{1mD} > 99.92\ \%$.

The risk of "producer" of the reference material is the probability that the measured purity value is $c_{1mD} < 99.90\ \%$, when the actual value is $c_1 \geq 99.90\ \%$. In such case, a false decision of the producer on quality of the raw material and/or its testing is possible, leading to rejection of the material and unnecessary loss of money and time. The producer's risk illustrated in Fig. 7b decreases to very small values when $c_{1mD} < 99.87\ \%$: the chance that actual purity value is not less than 99.90 % is already negligible.

In scenario 1 (a and b) the dependences of the risks on the impurities' content and the dependences on purity are symmetric by definition. In addition, the uncertainty of the directly measured purity $u_{1D} = 0.007\ \%$ by scenario 1a does not differ significantly from the uncertainty of the indirectly-measured purity by scenario 1b, which is equal to the uncertainty of the directly measured impurities' content $u_{2D} = 0.005\ \%$. Therefore, dependences of the risks on the indirectly measured purity by scenario 1b, are very similar to those shown in Fig. 7. Nevertheless, the risks values may be distinguished: the maximum consumer's risk values on the interval of purity [99.90 %, 99.95 %], measured directly by scenario 1a and indirectly by scenario 1b, are 0.06 vs. 0.10, respectively. The maximum producer's risk values on the interval of purity [99.85 %, 99.90 %], measured directly by scenario 1a and indirectly by scenario 1b, are 0.92 vs. 0.80, respectively.

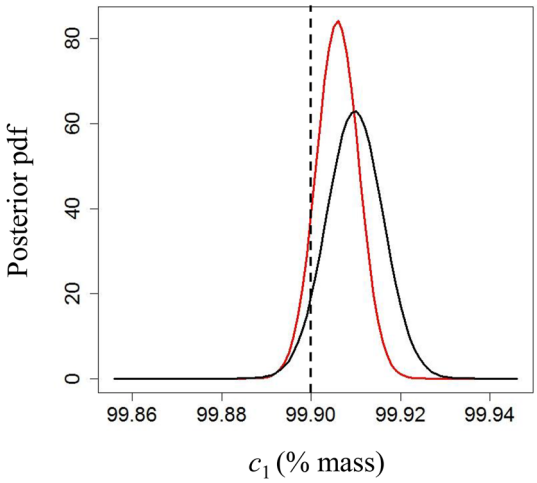

Fig. 8: Posterior probability density function of actual purity values c_1. Black line is the posterior pdf for scenario 1a; red line is the posterior pdf for scenario 1b; dotted line is the lower tolerance purity limit T_{L1} [58].

Note, although the measurement uncertainty u_{2D} is smaller than u_{1D}, the consumer's risk in scenario 1b is greater than in scenario 1a, whereas the producer's risk is smaller. The reason is that the posterior pdf depends on both the likelihood function (in which the measurement uncertainty is involved as the scale parameter) and the prior pdf. The posterior pdf is characterized by an expectation value lying in between the prior mean and the measured value. At the same measured purity value, *e.g.*, at $c_{1mD} = c_{1mI} = 99.901\,\%$ where the consumer's risk values are the greatest, the (univariate) posterior pdf of c_1 obtained for scenario 1b with $u_{2D} = 0.005\,\%$ is narrower than that for scenario 1a with $u_{1D} = 0.007\,\%$, as shown in Fig. 8. However, the posterior pdf of scenario 1b is more shifted toward the measured value 99.901 %, has a larger tail of the distribution to the left of $T_{L1} = 99.90\,\%$, and therefore, the consumer's risk is greater here than in scenario 1a.

For scenario 2, where both purity and content of the impurities are measured directly, the total specific risks are calculated according to eq. 6, as described in Sec. 5.1 of this Guide. For each specified couple of measured values (c_{1mD}, c_{2mD}), for which the total risk has to be evaluated, integrals of the posterior pdf in eq. 6 involve the ratio of multiple integrals of the joint pdf of vector [c_1, c_2, c_{1mD}, c_{2mD}] with respect to variables c_1 and c_2 over appropriate domains. $M = 10^7$ random vectors [c_1, c_2, c_{1mD}, c_{2mD}] were generated according to the prior modeling for c_1 and c_2 in Sec. 3.3 and the likelihood modeling for c_{1mD} and c_{2mD} in Sec. 3.4, Model 2. For a better rendering, risks values related to scenario 2 were smoothed by means of the "loess" [81] and "loess.surf functions" [82].

The total consumer's and producer's specific risks by scenario 2 are demonstrated in Fig. 9, plots a) and b), respectively. Each plot containing the risk surface is a transparent cube. The cube is shown in the perspective where the longest vertical line is its front edge, while the shortest vertical line behind is the back edge (the upper

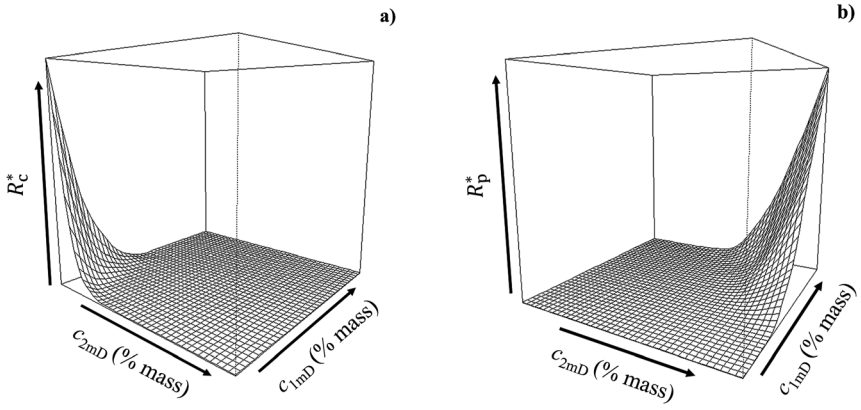

Fig. 9: Total specific consumer's risk R_c^* (a) and producer's risk R_p^* (b) in dependence on measured directly purity c_{1mD} and impurities content c_{2mD}, without closure operation. The interval of R_c^* values in plot (a) is [0, 0.11], while the interval of R_p^* values in plot (b) is [0, 0.76]. The measured purity and impurities content intervals are as Fig. 7 [58].

face is visible from inside the cube). In this case the outcomes of the measurement process, *i.e.*, couples (c_{1mD}, c_{2mD}) in which $c_{1mD} \neq 100 - c_{2mD}$ (whereas $c_1 = 100 - c_2$), are feasible also. The measurement uncertainty u_{2D}, being smaller than u_{1D}, influences the risk values here more than in scenario 1. The maximum value of the consumer's risk is 0.11, while for the producer's risk it is 0.76.

Note, the maximum risks in Fig. 9, as well as in Fig. 7, are related to the displayed intervals of the variables, *i.e.*, measured purity and impurities' content values. Out of these intervals the risks' behavior and their maximum values may change. In particular, the total specific producer's risk in Fig. 8b is zero when $c_{1mD} > 99.90\,\%$ and $c_{2mD} < 0.10\,\%$, by definition. However, this risk may increase when, for example, $c_{1mD} > 99.90\,\%$ and simultaneously $c_{2mD} > 0.10\,\%$, or vice versa, when $c_{1mD} < 99.90\,\%$ and simultaneously $c_{2mD} < 0.10\,\%$.

In general, comparing the discussed univariate and bivariate scenarios of the risks related to determination of potassium trioxidoiodate purity, one can see that practically the same results are obtained in bivariate scenarios, transformed de facto into univariate cases by the mass balance constraint. This means that there is no significant difference between direct and indirect purity test methods if their measurement uncertainties are approximately equal. It is important that with changes of the measurement conditions, the prior pdf and likelihood function might change from those discussed above. Therefore, the risks of false decisions could change accordingly. Note also that the "collapse" of multivariate scenarios into univariate scenarios because of the closure operation, as happens in bivariate scenarios, is impossible for three or more variables.

Example 3. Risks in conformity assessment of a sausage

A-3-1 Introduction

The objective of this Example is evaluation of both *global* and *specific* risks in conformity assessment of chemical composition of a sausage, based on a multivariate Bayesian approach, taking into account measurement uncertainty, correlation, and the mass balance constraint. As a case study, a dataset of compositions of a summer (dry) sausage produced from a mixture of beef, pork, and bacon with addition of salt and some other ingredients, was analyzed [67].

A-3-2 Experimental and tolerance domain

Results of testing the chemical composition of a total number of batches $N = 83$ of the sausage, produced according to standardized technical conditions [83] during about 3 years at two similar factories, were studied as a case study of a dataset for quantification of the total risks. The main components of the tested composition were fat, protein, moisture, and salt. Each factory tested a sausage batch at its laboratory for conformity assessment before the product is placed on the market.

The standard method for measurement of a fat content c_1 in the sausage was based on multiple fat extractions from the dried sample with a solvent (hexane, diethyl ether, or petroleum ether) in a Soxhlet fat extraction apparatus. Then, the solvent is removed and the fat dried to constant weight. The standard measurement uncertainty associated with a measured fat content c_{1m} was $u_1 = 0.05\ c_{1m}$.

Protein content c_2 was measured by the standard Kjeldahl method. Hot-acid digestion of a sample converts protein to ammonia, which is then distilled into standardized acid, after which the acid is back-titrated and the result calculated. The standard measurement uncertainty associated with a measured protein content c_{2m} was $u_2 = 0.04\ c_{2m}$.

The standard measurement method for moisture content c_3 consisted of drying a sample with heat-conducting sand to constant weight at a temperature of $(103 \pm 2)\,°C$. The standard measurement uncertainty associated with a measured moisture content c_{3m} was $u_3 = 0.06\ c_{3m}$.

Salt content c_4 was measured by Mohr's standard titration method. This method determines chloride ions extracted from the sample by titration with silver nitrate. As the silver nitrate solution is slowly added, a precipitate of silver chloride forms. At the end point, additional silver ions react with tetraoxidochromate

(chromate) ions of the indicator (potassium chromate) to form a red-brown precipitate of silver chromate. The standard measurement uncertainty associated with a salt measured content c_{4m} was $u_4 = 0.04\ c_{4m}$.

Note that a metrologically related correlation of the test results (between measured content values of different components of the same batch or sample) is impossible here as the applied chemical analytical methods are based on different principles.

The standard [83] sets the lower and upper tolerance/specification limits, T_{Li} and T_{Ui} respectively, of contents c_i as mass fractions (%) of the four main chemical components in the sausage:
i = 1) fat content $c_1 \le 53.0\ \% = T_{U1}$;
i = 2) protein content $c_2 \ge 15.0\ \% = T_{L2}$;
i = 3) moisture content $c_3 \le 40.0\ \% = T_{U3}$;
i = 4) salt content $c_4 \le 5.0\ \% = T_{U4}$

Sum of these component contents is constrained by the mass balance: $\sum_{i=1}^{4} c_i = 100\ \%$.

Acceptance limits, which take into account measurement uncertainties, were not applied in the factories, and measured values c_{im} were directly compared with specification limits. Therefore, in the present study acceptance limits are taken as to be the same as tolerance/specification limits.

The specification limits of contents of the components, T_{Li} and T_{Ui}, form a multivariate specification domain of permissible sausage compositions. However, there is also the mass balance constraint that leads to a multivariate sub-domain of feasible compositions. For example, for fat content $c_1 = T_{U1} = 53.0\ \%$, moisture content $c_3 = T_{U3} = 40.0\ \%$ and salt content $c_4 = T_{U4} = 5.0\ \%$, the protein content from the mass balance is $c_2 = 100\ \% - (53.0 + 40.0 + 5.0)\ \% = 2.0\ \%$, which is less than $T_{L2} = 15.0\ \%$ and hence not permissible. On the other hand, a composition such as $c_1 = T_{U1}$, $c_2 = T_{L2}$, $c_3 = T_{U3}$ and $c_4 = T_{U4}$ is within the specification domain but cannot be realized because it contradicts the mass balance constraint. Therefore, the multivariate sub-domain of feasible sausage compositions can be imagined as a part of the 3-simplex for the four components, which is the transparent triangular pyramid (tetrahedron), shown in Fig. 10 with c_1, c_2, c_3 and c_4 equal to 100 % in their vertices.

A-3-3 Prior probability density function

It was shown by analysis of variance (ANOVA) that the data of the two factories are practically homogeneous, and test results of all the $N = 83$ batches can be used for further calculations as the unified dataset.

The means μ_i and standard deviations σ_i of the marginal theoretical normal distributions, satisfying requirements of the Kolmogorov–Smirnov test of goodness-of-fit to the dataset on composition of $N = 83$ sausage

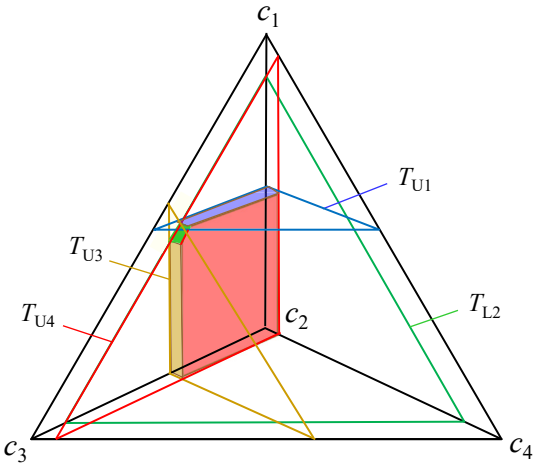

Fig. 10: The sausage compositions. Each vertex of the simplex corresponds to a component content c_i equal to 100 %. The permissible compositions are the part of the simplex delimited by the tolerance/specification limits T_{U1} (the transparent triangle shown by blue borderlines), T_{L2} (the transparent triangle shown by green lines), T_{U3} (the transparent triangle shown by brown lines), and T_{U4} (the transparent triangle shown by red lines). The left green border of T_{L2} is hidden by the red border of T_{U4}. The facets of the sub-domain of feasible compositions are marked with the colors of the tolerance limits. Thus, this sub-domain is seen as the rectangular prism with the left upper corner cut by the T_{L2} plane [67].

Table 3: Initial parameters of the prior distribution.

i	Component	μ_i,%	σ_i,%	r_{ij}			
j				Fat	Protein	Moisture	Salt
				1	2	3	4
1	Fat	40.5	3.66	1	−0.163	−0.318	−0.217
2	Protein	24.6	1.40	−0.163	1	−0.235	0.301
3	Moisture	29.7	4.15	−0.318	−0.235	1	−0.111
4	Salt	4.07	0.38	−0.217	0.301	−0.111	1

batches are presented in Table 3. The Pearson correlation coefficients r_{ij}, $i \neq j$, are also available in Table 3. The critical value of r_{ij} for $N - 2 = 81$ degrees of freedom and level of confidence 0.95 is 0.216 [67].

$M = 10^7$ Monte Carlo (MC) simulations of the actual sausage compositions $\boldsymbol{c} = [c_1, c_2, c_3, c_4]$ were performed using a multivariate truncated normal pdf, according to Model 1 in Sec. 3.3 of this Guide, having as location parameter the mean vector $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_4]$ and as scale parameter the covariance matrix $cov_{ij} = r_{ij}\,\sigma_i\,\sigma_j$ (μ_i, σ_i as well as r_{ij} are in Table 3).

As the sum of actual/"true" values, which have no uncertainties by definition, is to be equal to 100 %, data drawn from the multivariate normal pdf truncated on the domain $[0, 100]^4$ were subjected to the closure operation according to eq. 2. The resulting correlation matrix of r_{ijclo} is given in Table 4 (subscript "*clo*" in r_{ijclo} means after the "closure operation"). Comparing the correlation matrices in Table 3 and Table 4, one can see that the correlation coefficient related to contents of fat and moisture is still negative, but its absolute value is much larger. Also, the positive correlation coefficient between protein and salt contents is increased due to the closure operation.

The probability of conformance of the multivariate prior pdf, calculated as the fraction of M of the events when the obtained (simulated and closed) sausage compositions $\boldsymbol{c} = [c_1, c_2, c_3, c_4]$ were within the tolerance domain T, was $P_{\mathrm{conf}} = 0.972$.

A-3-4 Likelihood functions

Modeling of the multivariate likelihood function for measured values $\boldsymbol{c}_{\mathbf{m}} = [c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ was performed according to Model 1 in Sec. 3.4 of this Guide, using a multivariate truncated normal pdf with zero expectation as a model for error $\boldsymbol{e}_{\mathbf{m}} = [e_{1\mathrm{m}}, e_{2\mathrm{m}}, e_{3\mathrm{m}}, e_{4\mathrm{m}}]$. This error vector was then translated to the vector of actual content values $\boldsymbol{c} = [c_1, c_2, c_3, c_4]$ generated from the multivariate prior pdf. Therefore, $\boldsymbol{c}_{\mathbf{m}}$ is recovered as $\boldsymbol{c}_{\mathbf{m}} = \boldsymbol{c} + \boldsymbol{e}_{\mathbf{m}}$.

The covariance matrix, used as the scale parameter of the truncated pdf associated with vector $\boldsymbol{e}_{\mathbf{m}}$, is given in Table 5. The diagonal elements of this matrix are squared measurement uncertainties u_i^2, obtained from u_i discussed in Sec. A-3-2, and the off-diagonal elements are covariance terms equal to products $cov_{ij\mathrm{lf}} = r_{ij}\,u_i\,u_j$ whose correlation coefficients r_{ij} are in Table 3. Subscript "lf" in $cov_{ij\mathrm{lf}}$ means "likelihood function."

Table 4: Correlation coefficients r_{ijclo} after the closure operation.

i	Component	Fat	Protein	Moisture	Salt
j		1	2	3	4
1	Fat	1	−0.142	−0.823	−0.165
2	Protein	−0.142	1	−0.436	0.511
3	Moisture	−0.823	−0.436	1	−0.230
4	Salt	−0.165	0.511	−0.230	1

Table 5: Covariations cov_{ijlf} the multivariate likelihood function.

i	Component	Fat	Protein	Moisture	Salt
j		1	2	3	4
1	Fat	4.11	−0.32	−1.15	−0.07
2	Protein	−0.32	0.96	−0.41	0.05
3	Moisture	−1.15	−0.41	3.18	−0.03
4	Salt	−0.07	0.05	−0.03	0.03

Thus, modeling of both the likelihood function and the prior pdf was based on the experimental data, which contain the initial knowledge about the sausage compositions. However, the likelihood function characterizes the measurement process with corresponding measurement uncertainties, mimicked by $\boldsymbol{e}_\mathbf{m} = [e_{1\mathrm{m}}, e_{2\mathrm{m}}, e_{3\mathrm{m}}, e_{4\mathrm{m}}]$. Therefore, the measured values $\boldsymbol{c}_\mathbf{m} = [c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ are no longer required to sum to 100 %, and so the closure operation is not applied here.

A-3-5 Global risks

The total global consumer's risk $R_\mathrm{c} = 0.006$ was numerically recovered according to eq. 5 as the fraction of the M generated vectors of true and measured values $[c_1, c_2, c_3, c_4, c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ in which all the measured values $c_{i\mathrm{m}}$ were within their A_i but at least one of the corresponding true value c_i was outside T_i.

Correspondingly, the total global producer's risk $R_\mathrm{p} = 0.017$ was evaluated as the fraction of the M generated vectors $[c_1, c_2, c_3, c_4, c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ in which all the true values c_i were within their T_i, while at least one of the corresponding measured value $c_{i\mathrm{m}}$ was outside A_i.

That means six only from thousand sausage batches may be falsely assessed as corresponding to the specifications, whereas 17 conforming batches have a chance to be falsely rejected. This indicates a clear preference of the consumer's interests over the producer's interests. Obviously, a producer is interested in maintaining the satisfaction of his consumers no less than the consumers themselves.

A-3-6 Specific risks

For each specified vector of measured values $[c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$, the integrals of the posterior pdf in eq. 6 involve multiple integrals of the joint pdf of vector $[c_1, c_2, c_3, c_4, c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ with respect to variables c_i ($i = 1, \ldots, 4$) over appropriate domains. $M = 10^7$ random vectors $[c_1, c_2, c_3, c_4, c_{1\mathrm{m}}, c_{2\mathrm{m}}, c_{3\mathrm{m}}, c_{4\mathrm{m}}]$ were generated, according to the prior modeling for c_i values in Sec. 3.2.1 and the likelihood modeling for $c_{i\mathrm{m}}$ values in Sec. A-3-3.

The results of calculations of the total specific consumer's risk R_c^* were practically zero (less than 0.001) at the vector of measured values $\boldsymbol{c}_\mathbf{m}$ containing $c_{i\mathrm{m}}$ equal to the prior means μ_i of the distributions in Table 3. When $c_{i\mathrm{m}}$ move away from μ_i toward the tolerance limits, the R_c^* values naturally increase. The mean μ_1 of the fat contents is more than three standard deviations (σ_1 in Table 3) away from the tolerance limit $T_{\mathrm{U}1}$. Therefore, the influence of $c_{1\mathrm{m}}$ on R_c^* is minor. As the distance of the prior mean μ_2 of the protein contents from their tolerance limit $T_{\mathrm{L}2}$ is greater than five standard deviations σ_2, the $c_{2\mathrm{m}}$ influence on R_c^* is also very minor. The distances of the mean μ_3 of the moisture content from the tolerance limit $T_{\mathrm{U}3}$, and of the mean μ_4 of the salt content from the tolerance limit $T_{\mathrm{U}4}$, are less than three standard deviations (σ_3 and σ_4, respectively), and as a result do have a non-negligible influence on the risk. For example, $R_\mathrm{c}^* = 0.039$ at the measured fat content $c_{1\mathrm{m}} = \mu_1 = 40.5$ %, protein content $c_{2\mathrm{m}} = \mu_2 = 24.6$ %, moisture content $c_{3\mathrm{m}} = 35.7$ % (about 1.5 standard deviations from μ_3 toward the tolerance limit $T_{\mathrm{U}3}$) and salt content $c_{4\mathrm{m}} = 4.79$ % (about 2 standard deviations from μ_4 toward the tolerance limit $T_{\mathrm{U}4}$).

Note that even for $M = 10^7$ simulations there were certain instabilities (and maybe bias) in the risk values, increasing when working in the tails of the posterior distribution, where only a few random numbers could be

generated. Therefore, in the considered case, the MC method proved to be less reliable for calculation of risks than an analytical approximation method based on normal distributions [67].

Since the marginal distributions of the available data were successfully approximated by normal distributions and there was no evidence of an effect of truncation in the corresponding marginal distributions of the joint prior pdf, this pdf was constructed as a multivariate normal function with mean (Table 3) and covariance cov_{ijclo} matrix estimated from the random values generated as described in Sec. A-3-3. Because of the negligible effect of truncation, the multivariate likelihood function described in Sec. A-3-4 was also approximated by a multivariate normal function having the covariance matrix shown in Table 5. Resorting to such approximations for the prior pdf (cfr. **1.4**) and the likelihood function, a multivariate normal posterior pdf with parameters calculated as in Eq. (34) of the IUPAC/CITAC Guide [6] was applied for R_c^* and R_p^* calculations.

The dependence of total specific consumer's risk on measured values of fat and protein contents is shown in Fig. 11a for cases in which measured contents of moisture and salt are constant and equal to their prior means.

The R_c^* values are practically zero at the majority of combinations of c_{1m} and c_{2m} on the intervals from their prior means to the tolerance limits, and increase to 0.004 only, if simultaneously c_{1m} and c_{2m} are equal to their tolerance limits T_{U1} and T_{L2}, a case not observed in the raw data. In other words, the influence of fat content and protein content in the sausage on the total specific consumer's risk is negligible, as already explained above. The most influential variables here are contents of moisture and salt. The dependence of R_c^* on c_{3m} and c_{4m} on the intervals from their prior means to the tolerance limits is illustrated in Fig. 11b for cases in which measured contents of fat and protein are constant and equal to their prior means. The range of R_c^* values in this plot spreads from practically zero to 0.34. Of these two variables, the measured salt content has the most influence. Again, the risk is greatest when both c_{3m} and c_{4m} are at their tolerance limits.

The dependence of total specific producer's risk R_p^* on measured values of moisture and salt contents is shown in Fig. 12 for cases in which the measured contents of fat and protein are constant and equal to their prior means, as in Fig. 11b.

The interval of the measured moisture contents in Fig. 12a is from the tolerance limit T_{U3} to 1.18 T_{U3}, *i.e.*, to T_{U3} plus three standard measurement uncertainties ($3 \times 0.06\ T_{U3}$), while the interval of the measured salt contents is from the prior mean to the tolerance limit. Hence, in this particular case, measured contents of one component only do not conform. The calculated R_p^* values are from 0.69 to 1. Fig. 12b is the other way round with respect to

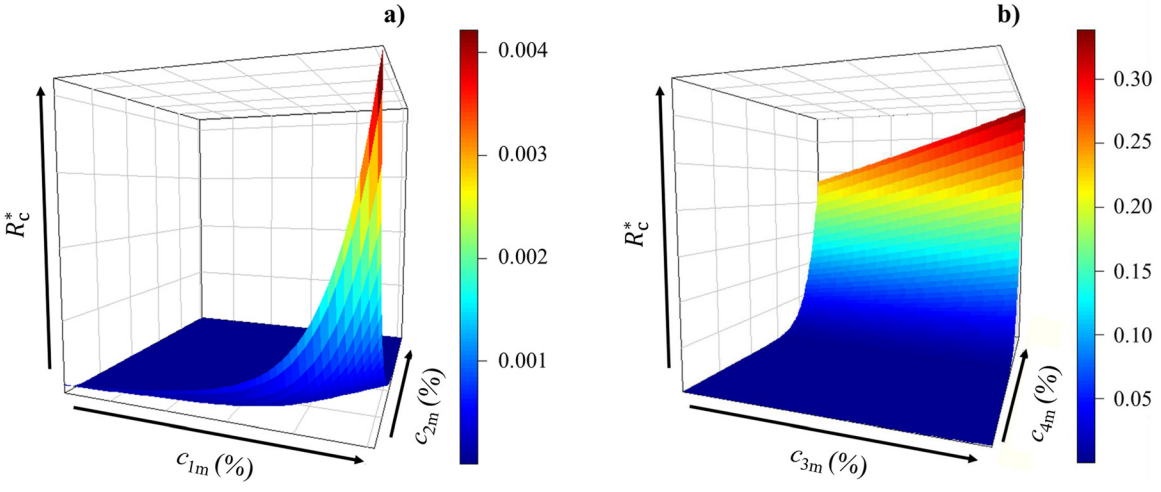

Fig. 11: Surface of total specific consumer's risk R_c^* vs. measured values of the contents of the sausage components. (a) Dependence of the risk on measured contents of fat c_{1m} and protein c_{2m} from their prior means to the tolerance limits, while the measured moisture content and salt content are constant and equal to their prior means. (b) Surface of the risk for cases when measured contents of fat and protein are constant and equal to their prior means, while the contents of moisture c_{3m} and salt c_{4m} are from their prior means to the tolerance limits. The color column bars code risk values between the minimum and maximum of the surface and refers to its plot only [67].

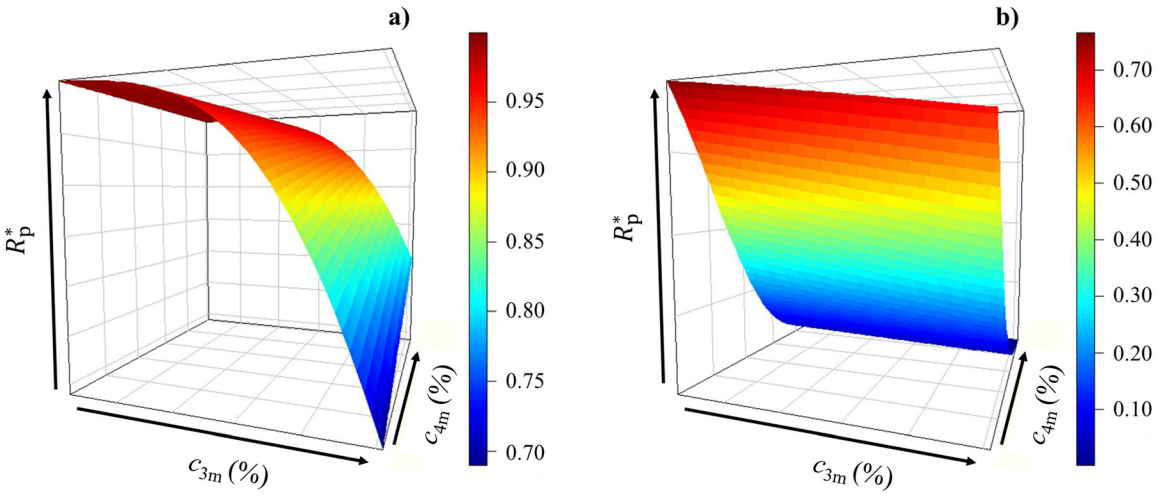

Fig. 12: Surface of total specific producer's risk R_p^* vs. measured values of the contents of the sausage components. Both the plots demonstrate the risk at the measured fat and protein contents constant and equal to the prior means. The measured moisture contents c_{3m} in Fig. 12a are on the interval of the three standard measurement uncertainties starting from the tolerance limit, while the interval of the measured salt contents c_{4m} is from the prior mean to the tolerance limit. In Fig. 12b, the interval of the measured salt contents c_{4m} is of the three measurement standard uncertainties starting from the tolerance limit, whereas the interval of the measured moisture contents c_{3m} is from the prior mean to the tolerance limit. The color bars are as in Fig. 11 [67].

Fig. 12a: here the measured salt contents do not conform, being in the interval from the tolerance limit T_{U4} to 1.12 T_{U4} (three measurement standard uncertainties being $3 \times 0.04\ T_{U4}$), while the interval of the measured moisture contents is from the prior mean to its tolerance limit. The R_p^* values go from practically zero at the $c_{4m} = 1.12\ T_{U4}$, up to 0.77 when $c_{3m} = \mu_3$ and simultaneously $c_{4m} = T_{U4}$. The logic is that there is no producer's risk if a component content in the tested batch substantially exceeded its tolerance limit and the batch is rejected. On the other hand, R_p^* can be extremely high when a measured content value is close to its tolerance limit. Correlations complicate the picture, but do not change it in essence.

In general, although the total global consumer's risk $R_c = 0.006$ and producer's risk $R_p = 0.017$ are small, the total specific risks related to a (specific) sausage batch can be much more significant when measured contents of moisture or salt, or both, are close to, or exceed, their tolerance limits.

Example 4. Risks in conformity assessment of synthetic air

A-4-1 Introduction

The objective of the present Example is evaluation of risks in conformity assessment of synthetic air based on a multivariate Bayesian approach, taking into account measurement uncertainty, correlation and the mass balance constraint.

Synthetic air is a mixture of nitrogen and oxygen (and some other minor components or impurities) that is used as "zero gas" in maintenance and calibration of test equipment for environmental monitoring, and as "balance gas" in calibration mixtures. Other applications are for medicinal purposes as a replacement for atmospheric air. Therefore, it is also important to understand fitness-for-purpose of the evaluated risks [68].

Two datasets related to synthetic air were analyzed: (1) from an industrial producer of medicinal synthetic air working according to the European Pharmacopoeia (EP) [84] and (2) from National Metrology Institutes (NMIs) that participated in the key comparison CCQM-K120 "Carbon dioxide at background and urban level" [85].

A-4-2 Experimental and tolerance domain

Medicinal synthetic air

Test results of $N_{\mathrm{med}} = 316$ lots of the EP medicinal synthetic air produced during a year at an industrial factory were used as a dataset for the study (subscript "med" in N_{med} means "medicinal"). The dataset includes measured values of the components' contents $c_{i\mathrm{m}}$ in the N_{med} lots, expressed as volume fractions of nitrogen $\phi_{1\mathrm{m}}$ cL/L, oxygen $\phi_{2\mathrm{m}}$ cL/L, and water vapor $\phi_{3\mathrm{m}}$ µL/L, where cL is 10^{-2} L and µL is 10^{-6} L. Note that the units of volume fraction,% *V*/*V* and ppm *V*/*V*, used in the European Pharmacopeia, are replaced here by cL/L and µL/L, respectively, to be consistent with IUPAC terminology and SI unit prefixes.

Medicinal synthetic air is defined in the EP as a mixture of nitrogen ($i = 1$) and oxygen ($i = 2$) in which the volume fraction of oxygen is 95.0 %–105.0 % of the nominal value being between 21.0 cL/L to 22.5 cL/L. These specifications can be interpreted as the acceptance limits of measured oxygen volume fraction $\phi_{2\mathrm{m}}$, *i.e.*, $A_{\mathrm{L}}(\phi_{2\mathrm{m}}) = 21.0$ and $A_{\mathrm{U}}(\phi_{2\mathrm{m}}) = 22.5$ cL/L, while the tolerance limits of actual ϕ_2 value are $T_{\mathrm{L}}(\phi_2) = (95.0/100)\cdot 21.0$ cL/L = 20.0 cL/L and $T_{\mathrm{U}}(\phi_2) = (105.0/100)\cdot 22.5$ cL/L = 23.6 cL/L.

Water vapor is specified in the EP as an impurity for which the maximum actual volume fraction ϕ_3 should be 67 µL/L. That means the acceptance limits of measured water vapor volume fraction $\phi_{3\mathrm{m}}$ are $A_{\mathrm{L}}(\phi_{3\mathrm{m}}) = T_{\mathrm{L}}(\phi_3) = 0$ µL/L and $A_{\mathrm{U}}(\phi_{3\mathrm{m}}) = T_{\mathrm{U}}(\phi_3) = 67$ µL/L. Note that even achieving $T_{\mathrm{U}}(\phi_3)$ water volume fraction ϕ_3 is negligible in comparison with the volume fraction ϕ_2 of oxygen, being four orders smaller.

Therefore, the mass balance for medicinal synthetic air $\sum_{i=1}^{3} \phi_i = 100$ cL/L can be simplified to $\phi_1 + \phi_2 = 100$ cL/L, from which nitrogen volume fraction is $\phi_1 = (100 - \phi_2)$ cL/L. Hence, the acceptance limits of measured nitrogen volume fraction $\phi_{1\mathrm{m}}$ are $A_{\mathrm{U}}(\phi_{1\mathrm{m}}) = (100 - 21.0)$ cL/L = 79.0 cL/L and $A_{\mathrm{L}}(\phi_{1\mathrm{m}}) = (100 - 22.5)$ cL/L = 77.5 cL/L. The tolerance limits of actual nitrogen volume fraction ϕ_1 are $T_{\mathrm{L}}(\phi_1) = (100 - 23.6)$ cL/L = 76.4 cL/L and $T_{\mathrm{U}}(\phi_1) = (100 - 20.0)$ cL/L = 80.0 cL/L.

Note also that water volume fraction ϕ_3, although insignificant for the mass balance, is still important for quality of the air as a product and must be taken into account in its conformity assessment.

Oxygen volume fractions are measured with a portable gas analyzer Servomex MiniMP 5200 equipped with a high-performance sensor which is based on the paramagnetic susceptibility of the oxygen molecule. The calculated/assumed standard measurement uncertainty was $u(\phi_{2\mathrm{m}}) = 0.15/\sqrt{3}$ cL/L = 0.09 cL/L.

Water vapor volume fractions are measured with a Shaw Dew Point SADP-Red Spot Meter providing direct indication in dew point temperature, which can be converted into water vapor volume fraction, µL/L. The standard measurement uncertainty was $u(\phi_{3\mathrm{m}}) = 1/\sqrt{3}$ µL/L = 0.6 µL/L.

Standard measurement uncertainty of nitrogen volume fraction calculated as $\phi_{1\mathrm{m}} = (100 - \phi_{2\mathrm{m}})$ cL/L was the same as of oxygen volume fraction, *i.e.*, $u(\phi_{1\mathrm{m}}) = u(\phi_{2\mathrm{m}}) = 0.09$ cL/L.

Synthetic air for CCQM-K120

The CCQM-K120 key comparison was designed to evaluate the level of compatibility of NMIs' preparative capabilities for carbon dioxide in air. Synthetic air was used as the balance gas (dry air matrix) for preparation of the measurement standards of carbon dioxide at background and urban level with the carbon dioxide amount fraction (380–480) µmol/mol in CCQM-K120.a and (480–800) µmol/mol in CCQM-K120.b, respectively. A standard containing carbon dioxide (480 ± 10) µmol/mol was applicable for both parts of CCQM-K120, when requirements/specifications of the synthetic air were satisfied. Thus, in the present work we refer to the dataset of the synthetic air properties of the $N_{\mathrm{CCQM}} = 23$ cylinders prepared by 12 NMIs, and the requirements of CCQM-K120.a for this air.

The air was synthesized at the NMIs gravimetrically by mixing and blending purchased pure gases in the cylinders. The dataset included acronyms of the 12 NMIs and their measured (assigned) values of components' contents $c_{i\mathrm{m}}$ in the N_{CCQM} cylinders, expressed in amount fractions (mol/mol) of nitrogen $x_{1\mathrm{m}}$, oxygen $x_{2\mathrm{m}}$ and argon $x_{3\mathrm{m}}$.

Nitrogen, oxygen and argon were defined in the protocol of the CCQM-K120.a comparison [86] as the main components of the air with the following tolerance limits of the amount fractions, mol/mol:
$i = 1$) nitrogen $T_{\mathrm{L}}(x_1) = 0.7804 \le x_1 \le 0.7814 = T_{\mathrm{U}}(x_1)$;

i = 2) oxygen $T_L(x_2) = 0.2088 \le x_2 \le 0.2098 = T_U(x_2)$;
i = 3) argon $T_L(x_3) = 0.0089 \le x_3 \le 0.0097 = T_U(x_3)$.

Since the tolerance limits were set based on the published measured values of the component amount fractions in ambient (real) air, the acceptance limits $A_L(x_{im})$ and $A_U(x_{im})$ are considered here equal to the respective tolerance limits $T_L(x_i)$ and $T_U(x_i)$. Amount fractions of the main components corresponded to the mass balance $\sum_{i=1}^{3} x_i = 1\,\text{mol/mol}$.

Details of the gas standard preparation and measurement uncertainties evaluation are described in the measurement reports of the corresponding 12 NMIs [86, Annex 5]. The standard measurement uncertainties for nitrogen $u(x_{1m})$ varied widely in the interval from 0.000002 mol/mol. A similar interval of standard measurement uncertainty values $u(x_{2m})$ for oxygen was (0.000002–0.000295) mol/mol, and the interval of $u(x_{3m})$ values for argon was (0.000001–0.000015) mol/mol.

A-4-3 Prior probability density function

Medicinal synthetic air

For both oxygen and water vapor volume fractions, the pdf of the theoretical distribution is assumed to be a mixture $f(\phi_i)$ of normal pdfs, $f_1(\phi_i)$ and $f_2(\phi_i)$ with non-negative weights $w_{i1} = (1 - w_{i2})$ and w_{i2}, respectively, such that the sum of the weights is equal to 1:

$$f(\phi_i) = w_{i1} f_1(\phi_i ; \mu_{i1}, \sigma_{i1}) + w_{i2} f_2(\phi_i ; \mu_{i2}, \sigma_{i2}), \tag{7}$$

where the weights and other pdf parameters for the present case study are shown in Table 6. These theoretical distributions satisfied requirements of the Kolmogorov–Smirnov test of goodness-of-fit to the dataset of compositions of N_{med} = 316 lots of the medicinal air. R codes for calculations on mixtures of normal pdfs are available, for example, in ref. [87].

Note that the pdf of the nitrogen volume fraction, being a mirror of the pdf of the oxygen volume fraction, refers to nitrogen volume fractions that are a complement to 100 cL/L of the corresponding oxygen volume fractions, according to the mass balance constraint. For example, when oxygen volume fractions $\phi_2 = w_{21}\mu_{21} + w_{22}\mu_{22}$ = 0.1·21.1 cL/L + 0.9·21.6 cL/L = 21.6 cL/L, the corresponding volume fraction of nitrogen is $\phi_1 = w_{11}\mu_{11} + w_{12}\mu_{12}$ = 0.9·78.4 cL/L + 0.1·78.9 cL/L = 78.4 cL/L, and their sum is 100 cL/L.

The Pearson correlation coefficient between volume fractions of oxygen and nitrogen is $r_{12} = -1$ by definition, since $\phi_{1m} = (100 - \phi_{2m})$ cL/L. The absolute value of the correlation coefficient between volume fractions of nitrogen or oxygen and water vapor, calculated from the dataset, is $|r_{13}| = |r_{23}| = 0.049$. It is statistically insignificant, as the critical value for such coefficient at the 0.95 level of confidence for a sample size of 316 pairs of variables, is 0.110.

The univariate prior pdfs for the components of the medicinal synthetic air were modeled by the mixture distributions with parameters presented in Table 6. The probability of conformance of the oxygen pdf, calculated as the fraction of $M = 10^7$ MC simulations of the events when the oxygen volume fractions ϕ_2 are within the tolerance interval $T_2(\phi_2)$, was equal to P_{conf} = 0.99997. As expected, this is also equal to the probability of conformance of the prior pdf modeling the nitrogen volume fractions ϕ_1. The probability of conformance for water volume fractions ϕ_3 is equal to one, since they are extremely far from their upper tolerance limit.

As the nitrogen volume fractions are calculated from the oxygen ones based on the mass balance, hence any bivariate pdf modeling for the two components actually degenerates into a univariate one, as in Example 2 of this

Table 6: Parameters of the theoretical mixture of normal distributions for the medicinal air.[a]

i	Component	μ_{i1}	σ_{i1}	w_{i1}	μ_{i2}	σ_{i2}	w_{i2}	ij	r_{ij}
1	Nitrogen	78.4	0.4	0.1	78.9	0.04	0.9	12	−1
2	Oxygen	21.1	0.04	0.9	21.6	0.4	0.1	13	−0.049
3	Water	0.6	0.2	0.6	1.5	0.4	0.4	23	−0.049

[a]Parameters μ and σ are the means and standard deviations of the theoretical distributions and have the same units as the respective parameters, *i.e.*, cL/L for volume fractions of nitrogen ϕ_1 and oxygen ϕ_2, and µL/L for volume fractions ϕ_3 of water vapor; the weights w_{i1} and w_{i2}, as well as the Pearson correlation coefficients, are dimensionless.

Table 7: Parameters of the prior distribution for synthetic air in CCQM-K120.

i	Component	$\mu(x_i)$, mol/mol	$\sigma(x_i)$, mol/mol	r_{ij}		
j				Nitrogen	Oxygen	Argon
				1	2	3
1	Nitrogen	0.7809	0.00046	1	−0.767	−0.348
2	Oxygen	0.2094	0.00036	−0.767	1	−0.162
3	Argon	0.0093	0.00015	−0.348	−0.162	1

Guide. The water vapor volume fractions are not correlated with the oxygen (and nitrogen) volume fractions and considered as independent from those. For these two reasons, the total probability of conformance for the medicinal synthetic air as a multicomponent material is equal to the product of that for oxygen and for water vapor, hence coinciding with the probability of conformance of the oxygen (or nitrogen) volume fractions.

Synthetic air for CCQM-K120

Based on the results of the Kolmogorov–Smirnov test of goodness-of-fit of a theoretical normal distribution and the empirical distribution (the dataset of $N_{\mathrm{CCQM}} = 23$ cylinders), the hypotheses of goodness-of-fit were not rejected at the 0.95 level of confidence. The means $\mu(x_i)$ and standard deviations $\sigma(x_i)$ of the theoretical normal pdfs are presented in Table 7.

The distributions of the components' amount fractions, which parameters presented in Table 7, are considered as the marginal distributions of a multivariate pdf. Its covariance matrix consists of variances σ_i^2 as the diagonal elements, and covariances $cov_{ij} = r_{ij}\,\sigma(x_i)\,\sigma(x_j)$, $i \neq j$, are the off-diagonal elements. The Pearson correlation coefficients r_{ij} between measured components' amount fractions presented in Table 7 are calculated from the dataset. The critical value of the correlation coefficients for $(N_{\mathrm{CCQM}} - 2) = 21$ degrees of freedom and 0.95 level of confidence is 0.413. As $|r_{i3}| \leq 0.413$ ($i = 1, 2$) in Table 7, the correlation of the argon amount fractions with the nitrogen and oxygen amount fractions was considered insignificant. However, correlation of the nitrogen and oxygen amount fractions (having a "spurious" origin caused by the mass balance constraint) is statistically significant, since $|r_{12}| = 0.767 > 0.413$.

The multivariate prior pdf of the synthetic air for CCQM-K120 was modeled using $M = 10^7$ MC simulations of the actual air compositions $\boldsymbol{x} = [x_1, x_2, x_3]$. The x_i values were drawn from a multivariate normal distribution, truncated on the domain $[0, 1]^3$, with the location parameter equal to the mean vector $\boldsymbol{\mu} = [\mu(x_1), \mu(x_2), \mu(x_3)]$, and the scale parameter equal to the covariance matrix of cov_{ij} calculated with the parameters presented in Table 7.

Actual ("true") values have no uncertainties and the sum of actual amount fractions of nitrogen, oxygen and argon prepared in CCQM-K120 must be exactly equal to 1. Therefore, the data drawn from the multivariate truncated normal pdf were subjected to the closure operation by eq. 2. The resulting correlation matrix is given in Table 8.

Table 8: Correlation coefficients r_{ijclo} of the component amount fractions in synthetic air for CCQM-K120 after the closure operation.

i	Component	Nitrogen	Oxygen	Argon
j		1	2	3
1	Nitrogen	1	−0.919	−0.284
2	Oxygen	−0.919	1	−0.118
3	Argon	−0.284	−0.118	1

Comparing correlation matrices in Tables 7 and 8, one can see that the correlation coefficient r_{12clo} related to amount fractions of nitrogen and oxygen was increased after the closure operation. The absolute values of the statistically insignificant correlation coefficients r_{13clo} and r_{23clo} between amount fractions of argon and other gases were even lower.

The probability of conformance of the multivariate prior pdf before the closure operation, calculated as the fraction of M of the events when the simulated synthetic air compositions $\boldsymbol{x} = [x_1, x_2, x_3]$ are within the tolerance domain $T(x_i)$, was $P_{\text{conf}} = 0.635$. After the closure operation and calculation of corresponding covariance matrix of cov_{ijclo}, the probability decreased to $P_{\text{conf}} = 0.475$. These low values of conformance probability are caused by nitrogen contents out of the tolerance interval $T(x_1)$ in five cylinders (OMH54, FB03747, JJ108862, PSM298266, PSM266468) and oxygen contents out of the tolerance interval $T(x_2)$ in four cylinders (1029047, FB03744, PSM298266, PSM266468), in the dataset of $N_{\text{CCQM}} = 23$ cylinders. Note that the correlation between nitrogen and oxygen amount fractions increased (in absolute value) after the closure operation, hence complicating the multivariate prior pdf and decreasing the conformance probability of the model based on the multivariate truncated normal pdf.

A-4-4 Likelihood functions

The univariate oxygen and water vapor likelihood functions for the medicinal synthetic air were modeled using, respectively, a normal distribution and a normal distribution truncated at zero (in order to avoid simulation of negative water vapor values). The location parameters were equal to actual values ϕ_2 and ϕ_3 drawn from the corresponding prior pdfs, and scale parameters were equal to $u(\phi_{2\text{m}})$ and $u(\phi_{3\text{m}})$, respectively, described in Sec. A-4-2.

Modeling the multivariate likelihood function for a vector of amount fraction values $\boldsymbol{x}_{\mathbf{m}} = [x_{1\text{m}}, x_{2\text{m}}, x_{3\text{m}}]$ of the synthetic air for CCQM-K120, was based on $\boldsymbol{x}_{\mathbf{m}}$ recovering as $\boldsymbol{x}_{\mathbf{m}} = \boldsymbol{x} + \boldsymbol{e}_{\mathbf{m}}$, where $\boldsymbol{e}_{\mathbf{m}} = [e_{1\text{m}}, e_{2\text{m}}, e_{3\text{m}}]$ is the vector of measurement errors. The vector $\boldsymbol{e}_{\mathbf{m}}$ was modeled assigning to it a multivariate truncated normal pdf with zero expectation, while the vector of actual amount fraction values $\boldsymbol{x} = [x_1, x_2, x_3]$ was generated from the multivariate prior pdf as explained in Sec. 3.4 of this Guide. The covariance matrix, used as the scale parameter of the truncated pdf associated with the vector $\boldsymbol{e}_{\mathbf{m}}$, contained as the diagonal elements the squared standard measurement uncertainties $u^2(x_{i\text{m}})$ discussed in Sec. A-4-2. The off-diagonal elements of this matrix were the covariance terms equal to products $cov_{ij\text{lf}} = r_{ij}\, u(x_{i\text{m}})\, u(x_{j\text{m}})$ whose correlation coefficients r_{ij} are in Table 7. Subscript "lf" in $cov_{ij\text{lf}}$ means "likelihood function." Since the likelihood function characterizes the measurement process with corresponding measurement uncertainties, the sum of the measured values in the vector $\boldsymbol{x}_{\mathbf{m}} = [x_{1\text{m}}, x_{2\text{m}}, x_{3\text{m}}]$ should not be equal to 1 exactly, and therefore, the closure operation is not applied here.

A-4-5 Risks of false decisions and their fitness-for-purpose

Medicinal synthetic air

Any total risk for oxygen and nitrogen volume fractions in medicinal synthetic air, similar to the probability of conformance, simplifies to a particular risk relevant to either one of the two. Since the volume fractions of oxygen (or nitrogen) and water vapor are not correlated, the total risks relevant to oxygen and water vapor are equal to a combination of their particular risks. The probability of conformance for water vapor volume fractions is equal to one, therefore a total risk for the medicinal synthetic air, equal to the product of the particular risks for oxygen and for water vapor, coincides with the particular risk for oxygen.

The particular global consumer's risk R_{c} related to the volume fractions of oxygen (and, therefore, of nitrogen) is equal to zero, while the particular global producer's risk is $R_{\text{p}} = 0.0926$. Both particular global consumer's and producer's risks related to the water volume fraction are zero. Therefore, the total global consumer's risk related to oxygen (or nitrogen) and water vapor is zero, according to the IUPAC/CITAC Guide [6], whereas the total global producer's risk, coinciding with the particular risk related to oxygen, is $R_{\text{p}} = 0.0926$.

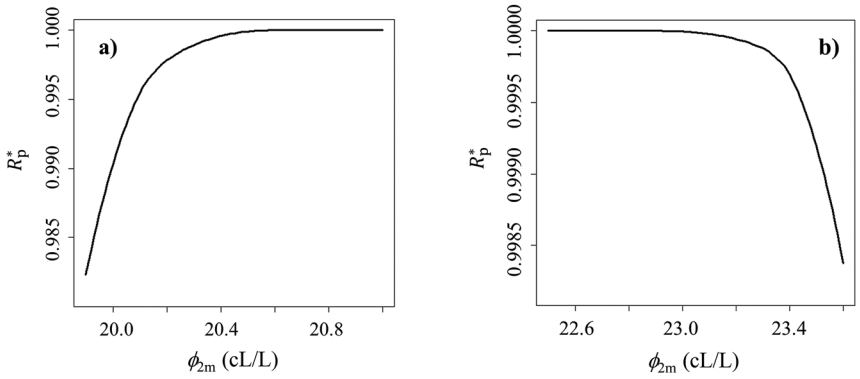

Fig. 13: Dependence of the specific producer's risk R_p^* on the measured oxygen volume fractions ϕ_{2m} in medicinal synthetic air. The interval (20.0−21.0) cL/L, from the lower tolerance to the lower acceptance limits, is on plot (a), while the interval (22.5−23.6) cL/L, from the upper acceptance to the upper tolerance limits is on plot (b) [68].

The particular specific consumer's risks R_c^* related to oxygen and water vapor volume fractions, for ϕ_{2m} and ϕ_{3m} values within the corresponding acceptance intervals, are zero. Hence, also, the total specific consumer's risk, related to oxygen and water vapor, is zero [6].

The specific producer's risks R_p^* as a function of the measured oxygen volume fractions on the intervals from the lower acceptance limit to the corresponding tolerance limit (a) and similarly for the upper limits (b) are shown in Fig. 13.

It is clear from Fig. 13 that the acceptance limits set in the European Pharmacopoeia are extremely reliable: the producer's risk decreases only when the oxygen volume fraction is less than the lower acceptance limit about 0.4 cL/L and greater than the upper acceptance limit about 0.4 cL/L also. In general, the produced medicinal synthetic air is fit-for-purpose (fit for intended use) as it satisfies the EP requirements completely without any statistically significant consumer's risk.

The specific producer's risk R_p^* relevant to water vapor volume fractions greater than 67 ppm, *i.e.*, outside the acceptance interval, is zero as such measured values are not feasible. Hence, the total specific producer's risks can differ from zero only when the measured oxygen volume fractions are outside their acceptance interval, while the water vapor volume fractions are inside their acceptance interval. In that case, the total specific producer's risk coincides with the particular specific risk for oxygen.

Synthetic air for CCQM-K120

The medians of the standard measurement uncertainties obtained from the NMIs were considered as the robust indicators of the NMIs' performance. Their values are for nitrogen $u(x_{1m})$ = 0.0000140 mol/mol, oxygen $u(x_{2m})$ = 0.000009 mol/mol, and for argon $u(x_{3m})$ = 0.000005 mol/mol. Corresponding total global consumer's risk R_c = 0.0079 and total global producer's risk R_p = 0.0081, calculated by eq. 5, are practically equal.

More details are shown in Fig. 14, where plot (a) demonstrates the total global consumer's risk R_c values, and plot (b) shows the total global producer's risk R_p values, as functions of the measurement uncertainties on the ranges for nitrogen $u(x_{1m})$ from 0.000002 mol/mol to 0.000235 mol/mol and oxygen $u(x_{2m})$ from 0.000002 mol/mol to 0.000065 mol/mol. The measurement uncertainty for argon is kept equal to the median $u(x_{3m})$ = 0.000005 mol/mol. These ranges do not include, for simplicity, the outlying values $u(x_{1m})$ = 0.001105 mol/mol and $u(x_{2m})$ = 0.000295 mol/mol of the dataset.

Both risks increase with increasing measurement uncertainties: R_c varies from 0.0014 to 0.0824, and R_p from 0.0015 to 0.1122. The surfaces of the risks are approximate planes twisted by correlations. Each NMI which declared a standard measurement uncertainty in the plotted ranges, can find its own risks R_c and R_p on these surfaces.

For the calculation of total specific risks, the prior pdf and the likelihood function in eq. 6 were approximated by relevant multivariate normal distributions according to the framework of the IUPAC/CITAC

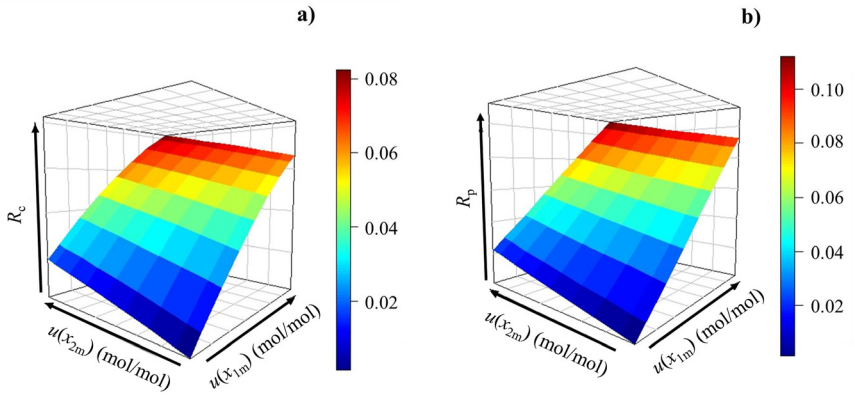

Fig. 14: Dependence of total global risks on standard measurement uncertainties in synthetic air prepared in CCQM-K120. Plot (a) demonstrates the surface of the consumer's risk R_c values vs. measurement uncertainties on the ranges for nitrogen $u(x_{1m})$ from 0.000002 mol/mol to 0.000235 mol/mol and oxygen $u(x_{2m})$ from 0.000002 mol/mol to 0.000065 mol/mol. Plot (b) presents the surface of the producer's risk R_p values vs. $u(x_{1m})$ and $u(x_{2m})$ on the same ranges as in plot (a). The color column bars code the risk values between the minimum and maximum of the surface, each bar referring to its own plot [68].

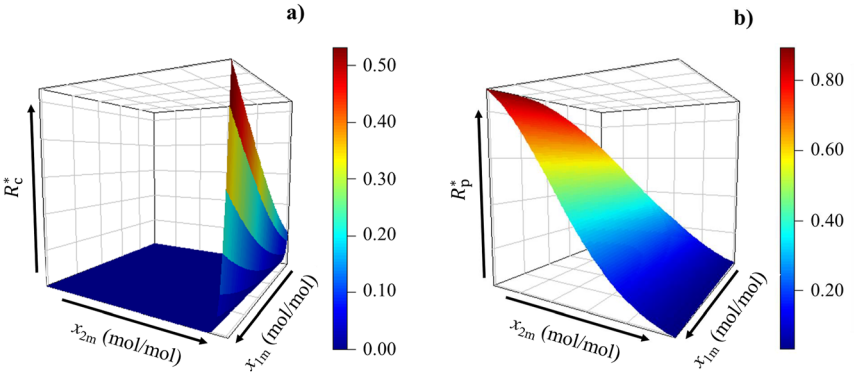

Fig. 15: Dependence of the total specific risks on measured values of the amount fractions of the main components in the synthetic air for CCQM-K120. Plot (a) is the surface of consumer's risk R_c^* vs. the amount fractions of nitrogen x_{1m} from 0.7804 to 0.7809 mol/mol, and oxygen x_{2m} from 0.2094 to 0.2098 mol/mol. Plot (b) is the surface of producer's risk R_p^* vs. the amount fractions of nitrogen x_{1m} from 0.780358 to 0.780400 mol/mol, and oxygen x_{2m} from 0.209800 to 0.209827 mol/mol. The color column bars are as in Fig. 14 [68].

Guide [6, Eq. (34)], in order to obtain a closed (normal) expression for the corresponding multivariate posterior pdf. The approximation of a truncated normal pdf by a normal pdf is sustainable in this study, since the distributions of the relevant quantities are very far away from the truncation limits [0, 1]. Moreover, the matrix in Table 6, rounded to three decimal digits, guarantees a proper full-rank covariance matrix for the prior pdf.

Total specific risks, evaluated by eq. 6 for the medians of the declared standard measurement uncertainties, are plotted against the measured amount fractions of nitrogen x_{1m} and oxygen x_{2m} in Fig. 15 at the argon amount fraction equal to its mean $\mu(x_3)$ = 0.0093 mol/mol in Table 7.

The surface of the total consumer's risk R_c^* in Fig. 15a was evaluated for intervals of x_{1m} and x_{2m} from the means in Table 7 to the tolerance limits, taking into account the negative correlation between the two components: when values of x_{2m} increase toward $T_U(x_2)$, then x_{1m} tends to decrease. Hence, the chosen intervals were for x_{2m} from $\mu(x_2)$ to the upper tolerance limit $T_U(x_2)$, and for x_{1m} from $\mu(x_1)$ to the lower tolerance limit

$T_L(x_1)$. The R_c^* values on this surface vary from 0 to 0.531. The surface of producer's risk R_p^* on Fig. 15b was calculated on the intervals of x_{1m} and x_{2m} out of their tolerance limits, taking into account again the negative correlation between the two components. The chosen intervals were for x_{2m} from the upper tolerance limit $T_U(x_2)$ to $T_U(x_2) + 3u(x_{2m})$, and for x_{1m} from the lower tolerance limit $T_L(x_1)$ to $T_L(x_1) - 3u(x_{1m})$, where $u(x_{1m})$ and $u(x_{2m})$ are the median standard measurement uncertainties. The R_p^* values vary from 0.033 to 0.894.

Note that the risks R_c^* and R_p^* specific for each NMI can be calculated at its own measured values of the air main components and associated measurement uncertainties. These risks are large in some cases, but fit-for-purpose, as the CCQM-K120 comparison was intended for evaluation of NMIs' preparative capabilities for carbon dioxide in air, whereas correspondence of the amount fractions of nitrogen, oxygen and argon to their tolerance limits was less important in this study.

Summarizing the discussion of the results, note that the developed technique can be helpful for both producers and consumers for fit-for-purpose conformity assessment of any substance or material, influenced by measurement uncertainties, correlation, and mass balance constraint. In particular, producers and consumers of medicinal synthetic air, as well as NMIs and calibration laboratories using synthetic air as the "balance gas" in standard gas mixtures for calibration of measuring instruments applied in industry and environment monitoring, can get benefit from application of this technique. Moreover, the proposed approach provides a way for improving the risks by tuning the involved parameters, mainly measurement uncertainties and/or acceptance intervals, taking in account the interests of both the air producer and consumer.

Membership of sponsoring bodies

The present membership of the IUPAC Analytical Chemistry Division (V) is

***President*:** D. Shaw (USA); ***Vice-President*:** D. Craston (UK); ***Secretary*:** L. Torsi (Italy); ***Past President*:** Z. Mester (Canada); ***Titular Members*:** R. Apak (Turkey), V. Baranovskaya (Russia), J. Barek (Czech), I. Kuselman (Israel), T. Takeuchi (Japan), S.K. Wiedmer (Finland); ***Associate Members*:** F. Emmerling (Germany), E.M.M. Flores (Brazil), I. Leito (Estonia), H. Li (China), A. Tintaru (France), W.E. Waghorne (France); ***National Representatives*:** R. Burks (USA), H.R. Byon (South Korea), O. Chaiapakul (Thailand), J. Labuda (Slovakia), C. Lucy (Canada), M.C.F. Magalhães (Portugal), T. Pradeep (India), M. Ramalingam (Malaysia), R.H. Sha'Ato (Nigeria), D. van Oevelen (Netherlands).

The present membership of the IUPAC Subcommittee on Metrology in Chemistry is

Chair*:** I. Kuselman (Israel); ***Members: S.K. Aggrawal (India), V. Baranovskaya (Russia), M.F. Camões (Portugal), A. Botha (South Africa), D. Craston (UK), R.J.N.B. da Silva (Portugal), S.L.R. Ellison (UK), E.M.M. Flores (Brazil), D.B. Hibbert (Australia), I. Leito (Estonia), H. Li (China), J. Meija (Canada), Z. Mester (Canada), F.R. Pennecchi (Italy), M. Ramalingam (Malaysia), D. Shaw (USA), S.K. Wiedmer (Finland).

The present membership of the Cooperation of International Traceability in Analytical Chemistry (CITAC) is

***Chair*:** B. Guetler (Germany); ***Vice-Chair*:** Z. Mester (Canada); ***Secretary*:** R.J.N.B. da Silva (Portugal); ***Past Chair*:** M. Sega (Italy); ***Members*:** A. Squirrell (Australia), M. Horsky (Austria), W. Wegscheider (Austria), F.R. Lourenço (Brazil), O. P. de Oliveira Junior (Brazil), V. Ponçano (Brazil), T.R.L. Dadamos (Brazil), J.E.S. Sarkis (Brazil), H. Li (China), T. Näykki (Finland), P. Fisicaro (France), S.G. Walch (Germany), I. Papadakis (Greece), D.W.M. Sin (Hong Kong, P. R. China), P. K. Gupta (India), M. Nabi (Iran), M. Walsh (Ireland), I. Kuselman (Israel), F.R. Pennecchi (Italy), T. Fujimoto (Japan), R.B. Khousam (Lebanon), O. Zakaria (Malaysia), Y.M. Nakanishi (Mexico), L. Samuel (New Zealand), V. Baranovskaya (Russia), N. Oganyan (Russia), T.L. Teo (Singapore), A. Botha (South Africa), M. Obkircher (Switzerland), S. Wunderli (Switzerland), R. Kaarls (the Netherlands), R.J.C. Brown (UK), S.L.R. Ellison (UK), V. Iyengar (USA), J. D. Messman (USA).

The membership of the Task Group is

Chair*:** I. Kuselman (Israel); ***Members: A. Di Rocco (Italy), D. B. Hibbert (Australia), F.R. Pennecchi (Italy).

References

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML. *Evaluation of Measurement Data. The Role of Measurement Uncertainty in Conformity Assessment*, JCGM 106 (2012), https://www.bipm.org/en/committees/jc/jcgm/publications (accessed Oct. 25, 2022).

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML. *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement (GUM)*, 1st ed., JCGM 100 (2008), https://www.bipm.org/en/committees/jc/jcgm/publications (accessed Oct. 25, 2022).

[3] International Laboratory Accreditation Cooperation. *Guidelines on Decision Rules and Statements of Conformity*. ILAC-G8 (2019), https://ilac.org/publications-and-resources/ilac-guidance-series/ (accessed Oct. 25, 2022).

[4] A. Williams, B. Magnusson (Eds.). Eurachem/CITAC Guide: *Use of Uncertainty Information in Compliance Assessment*, 2nd ed. (2021), https://www.eurachem.org/index.php/publications/guides (accessed Oct. 25, 2022).

[5] B. Magnusson, U. Örnemark (Eds.). Eurachem Guide: *The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*, 2nd ed. (2014), https://www.eurachem.org/index.php/publications/guides (accessed Oct. 25, 2022).

[6] I. Kuselman, F. R. Pennecchi, R. J. N. B. da Silva, D. B. Hibbert. *Pure Appl. Chem.* **93**, 113 (2021), https://doi.org/10.1515/pac-2019-0906.

[7] F. Chayes. *J. Geophys. Res.* **65**, 4185 (1960), https://doi.org/10.1029/JZ065i012p04185.

[8] T. Vigen. *Spurious Correlations*, Hachette Books, New York and Boston (2015).

[9] P. Kynčlová, K. Hron, P. Filmoser. *Math. Geosci.* **49**, 777 (2017), https://doi.org/10.1007/s11004-016-9669-3.

[10] J. Aitchison. *The Statistical Analysis of Compositional Data*, Chapman and Hall, London (1986).

[11] J. Aitchison. In *Multivariate Analysis and Its Applications*, T. W. Anderson, K. T. Fang, I. Olkin (Eds.), pp. 73–81, IMS Lecture Notes. Monograph Series **24**, Hayward, CA: Institute of Mathematical Statistics (1994).

[12] J. Aitchison. *A Concise Guide to Compositional Data Analysis*, ebook (2003), https://www.goodreads.com/book/show/46756704-a-concise-guide-to-compositional-data-analysis (accessed Oct. 26, 2022).

[13] V. Pawlowsky-Glahn, A. Buccianti. *Compositional Data Analysis: Theory and Applications*, Wiley, Chichester (2011).

[14] M. Greenacre, E. C. Grunsky. *The Isometric Logratio Transformation in Compositional Data Analysis: A Practical Evaluation*, Economics Working Paper Series, Working Paper, No. 1627, Universitat Pompeu Fabra, Barcelona (2019), https://econ-papers.upf.edu/papers/1627.pdf (accessed Oct. 25, 2022).

[15] M. Greenacre. *Math. Geosci.* **51**, 649 (2019), https://doi.org/10.1007/s11004-018-9754-x.

[16] M. Greenacre. *Compositional Data Analysis in Practice*, CRC Press, Boca Raton (2019).

[17] A. Buccianti. *Comput. Geosci.* **50**, 165 (2013), https://doi.org/10.1016/j.cageo.2012.06.012.

[18] A. Buccianti, E. Grunsky. *J. Geochem. Explor.* **141**, 1 (2014), https://doi.org/10.1016/j.gexplo.2014.03.022.

[19] P. Filzmoser, K. Hron, C. Reimann. *Sci. Total Environ.* **408**, 4230 (2010), https://doi.org/10.1016/j.scitotenv.2010.05.011.

[20] C. Reimann, P. Filzmoser, K. Fabian, M. Birke, A. Demetriades, E. Dinelli, A. Ladenberger, the GEMAS Project Team. *Sci. Total Environ.* **426**, 196 (2012), https://doi.org/10.1016/j.scitotenv.2012.02.032.

[21] M. E. Edjabou, J. A. Martin-Fernandez, C. Scheutz, T. F. Astrup. *Waste Manage.* **69**, 13 (2017), https://doi.org/10.1016/j.wasman.2017.08.036.

[22] H. A. Souza, S. E. Parent, D. E. Rozane, D. A. Amorim, V. C. Modesto, W. Natale, L. E. Parent. *Front. Plant Sci.* **7**, 1252, https://doi.org/10.3389/fpls.2016.01252.

[23] J. M. Macklaim, G. B. Gloor. In *Microbiome Analysis: Method and Protocols, Methods in Molecular Biology*, R. Beiko, W. Hsiao, J. Parkinson (Eds.), Vol. **1849**, pp. 193–213, Humana Press, New York (2018).

[24] T. P. Quinn, I. Erb, M. F. Richardson, T. M. Crowley. *Bioinformatics* **34**, 2870 (2018), https://doi.org/10.1093/bioinformatics/bty175.

[25] G. P. Campbell, J. M. Curran, G. M. Miskelly, S. Coulson, G. M. Yaxley, E. C. Grunsky, S. C. Cox. *Forensic Sci. Int.* **188**, 81 (2009), https://doi.org/10.1016/j.forsciint.2009.03.018.

[26] T. Neocleous, C. Aitken, G. Zadora. *Chemometr. Itell. Lab. Syst.* **109**, 77 (2011), https://doi.org/10.1016/j.chemolab.2011.08.003.

[27] M. Z. Peterson, S. K. Suram, J. M. Gregoire. *ACS Comb. Sci.* **17**, 130 (2015), https://doi.org/10.1021/co5001458.

[28] Compositional data analysis Website (CODAWEB), http://www.compositionaldata.com/ (accessed Oct. 26, 2022).

[29] J. Aitchison. In *Proceedings of CODAWORK'08*, The 3rd Compositional Data Analysis Workshop, J. Daunis-i-Estadella, J.A. Martín-Fernández (Eds.), University of Girona (2008), CD-ROM, https://dugi-doc.udg.edu//handle/10256/706 (accessed Oct. 26, 2022).

[30] A. M. H. van der Veen, K. Hafner. *Metrologia* **51**, 80 (2014), https://doi.org/10.1088/0026-1394/51/1/80.

[31] M. J. T. Milton, G. M. Vargha, A. S. Brown. *Metrologia* **48**, R1 (2011), https://doi.org/10.1088/0026-1394/48/5/R01.

[32] International Standard Organization. ISO Guide 35, Reference materials. Guidance for characterization and assessment of homogeneity and stability (2017), Available from, https://www.iso.org/standard/60281.html (accessed Oct. 26, 2022).

[33] S. Westwood, T. Choteau, A. Daireaux, R. D. Josephs, R. I. Wiegosz. *ACS Anal. Chem.* **85**, 3118 (2013), https://doi.org/10.1021/ac303329k.

[34] S. R. Davies, M. Alamgir, B. K. H. Chan, T. Dang, K. Jones, M. Krishnaswami, Y. Luo, P. S. R. Mitchell, M. Moawad, H. Swan, G. J. Tarrant. *Anal. Bioanal. Chem.* **407**, 7983 (2015), https://doi.org/10.1007/s00216-015-8971-0.

[35] A. C. P. Osorio, R. C. de Sena, T. O. Araújo, M. D. de Almeida. *Accred. Qual. Assur.* **24**, 387 (2019), https://doi.org/10.1007/s00769-019-01392-w.

[36] F. R. Pennecchi, A. Di Rocco, I. Kuselman, D. B. Hibbert, M. Sega. *Measurement* **163**, 107947 (2020), https://doi.org/10.1016/j.measurement.2020.107947.

[37] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML. *International Vocabulary of Metrology. Basic and General Concepts and Associated Terms* (VIM), 3rd ed., JCGM 200 (2012), https://www.bipm.org/en/committees/jc/jcgm/publications (accessed Oct. 25, 2022).
[38] International Standard Organization, International Electrotechnical Commission. ISO/IEC 17000, *Conformity assessment. Vocabulary and general principles* (2020), https://www.iso.org/standard/73029.html (accessed Oct. 27, 2022).
[39] International Standard Organization, International Electrotechnical Commission. ISO/IEC 3534. *Statistics. Vocabulary and Symbols. Part 1: General Statistical Terms and Terms Used in Probability* (2006), https://www.iso.org/standard/40145.html (accessed Oct. 27, 2022).
[40] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. https://doi.org/10.1351/goldbook.
[41] D. B. Hibbert, E.-H. Korte, U. Örnemark. *Pure Appl. Chem.* **93**, 997 (2021), https://doi.org/10.1515/pac-2019-0819.
[42] R. D. Whitaker. *J. Chem. Educ.* **52**, 658 (1975), https://doi.org/10.1021/ed052p658.
[43] M. F. Camoes, G. D. Christian, D. B. Hibbert. *Pure Appl. Chem.* **90**, 563 (2018), https://doi.org/10.1515/pac-2017-0410.
[44] J. C. Whitwell, R. K. Toner. *Conservation of Mass and Energy*, McGraw-Hill Chemical Engineering Series, New York (1973).
[45] G. Feijoo, J. M. Lema, M. T. Moreira. *Mass Balances for Chemical Engineers*, De Gruyter Textbook Series, Berlin (2020).
[46] J. R. Dean. *Extraction Techniques in Analytical Science*, Willey, Padstow (2009).
[47] J. Dubrovkin. *Mathematical Processing of Spectral Data in Analytical Chemistry: A Guide to Error Analysis*, Cambridge Scholars Publishing, Cambridge (2018).
[48] É.M. D. M. Flores (Ed.). In *Microwave–Assisted Sample Preparation for Trace Element Determination*, Elsevier, Amsterdam (2014).
[49] Bo E. H. Saxberg, B. R. Kowalski. *Anal. Chem.* **51**, 1031 (1979), https://doi.org/10.1021/ac50043a059.
[50] J. Vessman, R. I. Stefan, J. F. van Staden, K. Danzer, W. Lindner, D. T. Burns, A. Fajgelj, H. Müller. *Pure Appl. Chem.* **73**, 1381 (2001), https://doi.org/10.1351/pac200173081381.
[51] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Harmonized Tripartite Guideline Q2(R1), *Validation of Analytical Procedures: Text and Methodology* (2005), https://www.ich.org/page/quality-guidelines (accessed Oct. 26, 2022).
[52] I. Kuselman, F. Pennecchi. *Pure Appl. Chem.* **88**, 477 (2016), https://doi.org/10.1515/ci-2016-0520.
[53] F. Sherman, I. Kuselman. *Accred. Qual. Assur.* **4**, 230 (1999), https://doi.org/10.1007/s007690050357.
[54] K. Schwahn, R. Beleggia, N. Omranian, Z. Nikoloski. *Front. Plant Sci.* **8**, 2152 (2017), https://doi.org/10.3389/fpls.2017.02152.
[55] L. Wang, P. Wang, M. Sheng, J. Tian. *Glob. Ecol. Conserv.* **16**, e00449 (2018), https://doi.org/10.1016/j.gecco.2018.e00449.
[56] I. Kuselman, F. R. Pennecchi, R. J. N. B. da Silva, D. B. Hibbert. *Talanta* **174**, 789 (2017), https://doi.org/10.1016/j.talanta.2017.06.073.
[57] I. Kuselman, F. R. Pennecchi, R. J. N. B. da Silva, D. B. Hibbert, E. Anchutina. *Talanta* **189**, 666 (2018), https://doi.org/10.1016/j.talanta.2018.07.049.
[58] F. R. Pennecchi, I. Kuselman, A. Di Rocco, D. B. Hibbert, A. Sobina, E. Sobina. *Measurement* **173**, 108662 (2021), https://doi.org/10.1016/j.measurement.2021.110542.
[59] P. Filzmoser, K. Hron. *Math. Geosci.* **41**, 905 (2009), https://doi.org/10.1007/s11004-008-9196-y.
[60] W. Long, Q. Wang. *Procedia Comput. Sci.* **18**, 1757 (2013), https://doi.org/10.1016/j.procs.2013.05.344.
[61] I. Kuselman, F. Pennecchi, C. Burns, A. Fajgelj, P. de Zorzi. *Pure Appl. Chem.* **84**, 1939 (2012), https://doi.org/10.1351/PAC-REP-11-10-04.
[62] R. B. D'Agostino, M. A. Stephens (Eds.). *Goodness-of-Fit Techniques. Statistics: A series of textbooks and monographs*, Marcel Dekker Inc, New York (1986).
[63] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML. Evaluation of measurement data. Supplement 1 to the "Guide to the expression of uncertainty in measurement". Propagation of distributions using a Monte Carlo method. JCGM 101 (2008), http://www.bipm.org/en/publications/guides/ (accessed Oct. 26, 2022).
[64] A. Taavitsainen, R. Vanhanen. *Nucl. Instrum. Methods Phys. Res., Sect. A* **854**, 156 (2017), https://doi.org/10.1016/j.nima.2016.11.061.
[65] F. Cosman, E. Krotkov. CMU-RI-TR 94-35, Carnegie Mellon University (1994), https://www.ri.cmu.edu/pub_files/pub1/cozman_fabio_1994_1/cozman_fabio_1994_1.pdf (accessed Oct. 26, 2022).
[66] M. D. McKay. *Evaluating Prediction Uncertainty*, Los Alamos National Laboratory – Technical Report, USA (1995).
[67] F. R. Pennecchi, I. Kuselman, A. Di Rocco, D. B. Hibbert, A. A. Semenova. *Food Control* **125**, 107949 (2021), https://doi.org/10.1016/j.foodcont.2021.107949.
[68] F. R. Pennecchi, I. Kuselman, D. B. Hibbert, M. Sega, F. Rolle, V. Altshul. *Measurement* **188**, 110542 (2022), https://doi.org/10.1016/j.measurement.2021.110542.
[69] S. Wilhelm. *rtmvnorm: Sampling Random Numbers from the Truncated Multivariate Normal Distribution*. RDocumentation, tmvtnorm version 1.4-10, https://www.rdocumentation.org/packages/tmvtnorm/versions/1.4-10/topics/rtmvnorm (accessed Oct. 26, 2022).
[70] J. Hadfield. *rtnorm: Random Generation from a Truncated Normal Distribution*. RDocumentation, MCMCglmm version 2.29, https://www.rdocumentation.org/packages/MCMCglmm/versions/2.29/topics/rtnorm (accessed Oct. 26, 2022).
[71] K. G. van den Boogaart, R. Tolosana-Delgado, M. Bren. *Compositional Data Analysis*. Package 'Composition' version 2.0–4, https://cran.r-project.org/web/packages/compositions/compositions.pdf (accessed Oct. 26, 2022).
[72] International Standard Organization. ISO Guide 73, *Risk management. Vocabulary* (2009), https://www.iso.org/standard/44651.html (accessed Oct. 28, 2022).
[73] International Standard Organization. ISO 31000, *Risk Management. Principles and Guidelines* (2018), https://www.iso.org/standard/65694.html (accessed Oct. 28, 2022).

[74] International Electrotechnical Commission, International Standard Organization. IEC/ISO 31010, *Risk Management. Risk Assessment Techniques* (2019), https://www.iso.org/standard/72140.html (accessed Oct. 28, 2022).
[75] Russian State Standard. GOST 13498-2010, *Platinum and Its Base Alloys*. Marks, Standardinform, Moscow (2012), https://meganorms.com/gost-13498-2010.html (accessed Oct. 26, 2022).
[76] F. Galembeck, T. A. L. Burgo. *Chemical Electrostatics. New Ideas on Electrostatic Charging: Mechanisms and Consequences*, Springer Cham, Switzerland (2017).
[77] E. M. Purcell, D. L. Morin. *Electricity and Magnetism*, Cambridge University Press, New York, 3rd ed. (2013).
[78] Russian Federation Agency for Technical Regulation and Metrology (Rosstandard). *State Verification Scheme of Measuring Instruments for Contents of Inorganic Components in Liquid and Solid Substances and Materials*, Moscow (2018), http://uniim.ru/wp-content/uploads/2017/08/gps_get_176_2017.pdf (accessed Oct. 26, 2022).
[79] R. Bettencourt da Silva, A. Williams (Eds.). Eurachem/CITAC Guide: *Setting and Using Target Uncertainty in Chemical Measurement* (2015), https://www.eurachem.org/index.php/publications/guides (accessed Oct. 25, 2022).
[80] *Integration of one-dimensional functions*. RDocumentation, Package *stats* version 4.1.0, https://stat.ethz.ch/R-manual/R-devel/library/stats/html/integrate.html (accessed Oct. 26, 2022).
[81] NIST/SEMATECH. *e-Handbook of Statistical Methods*, https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd144.htm (accessed Oct. 26, 2022).
[82] K. Aho. *loess.surf: Loess 2D and 3D smooth plots*. RDocumentation, Package asbio version 1.6-5, https://www.rdocumentation.org/packages/asbio/versions/1.6-5/topics/loess.surf (accessed Oct. 26, 2022).
[83] Russian State Standard. GOST 55456-2013, *Dry Sausages. Technical Conditions*, Standardinform, Moscow (2014), https://files.stroyinf.ru/Data2/1/4293774/4293774914.pdf (accessed Oct. 26, 2022).
[84] Air, synthetic medicinal. In *European Pharmacopoeia*, 10th ed., pp. 1770–1771 (2019), https://pheur.edqm.eu/subhome/10-8. (accessed Oct. 28, 2022).
[85] E. Flores, J. Viallon, T. Choteau, P. Moussay, F. Idrees, R. I. Wielgosz, J. Lee, E. Zalewska, G. Nieuwenkamp, A. van der Veen, L. A. Konopelko, Y. A. Kustikov, A. V. Kolobova, Y. K. Chubchenko, O. V. Efremova, B. I. Zhe, Z. Zhou, W. R. Miller, Jr, G. C. Rhoderick, J. T. Hodge, T. Shimosaka, N. Aoki, B. Hall, P. Brewer, D. Cieciora, M. Sega, T. Macé, J. Fükő, Z. N. Szilágyi, T. Büki, M. I. Jozela, N. G. Ntsasa, N. Leshabane, J. Tshilongo, P. Johri, T. Tarhan. *Metrologia* **56**, 08001 (2019), https://doi.org/10.1088/0026-1394/56/1A/08001.
[86] E. Flores, J. Viallon, T. Choteau, P. Moussay, F. Idrees, R. I. Wielgosz, J. Lee, E. Zalewska, G. Nieuwenkamp, A. van der Veen, L. A. Konopelko, Y. A. Kustikov, A. V. Kolobova, Y. K. Chubchenko, O. V. Efremova, B. I. Zhe, Z. Zhou, W. R. Miller, Jr, G. C. Rhoderick, J. T. Hodge, T. Shimosaka, N. Aoki, B. Hall, P. Brewer, D. Cieciora, M. Sega, T. Macé, J. Fükő, Z. N. Szilágyi, T. Büki, M. I. Jozela, N. G. Ntsasa, N. Leshabane, J. Tshilongo, P. Johri, T. Tarhan. *International comparison, CCQM-K120a Carbon dioxide in air at background level (380–480) µmol/mol, CCQM-K120b carbon dioxide in air at urban level (480–800) µmol/mol, final report*, Bureau International des Poids et Mesures (BIPM) and National Institute of Standards and Technology (NIST), USA (2018), https://www.bipm.org/documents/20126/44695456/CCQM-K120.pdf/c76f0923-8590-3d2c-05b1-fc0f5a1df2e6 (accessed Oct. 26, 2022).
[87] S. P. Millard. *An R Package for Environmental Statistics*, Springer, New York, 2nd ed. (2013), [*Mixture of Two Normal Distributions*. RDocumentation, Package *EnvStats* version 2.7.0, https://search.r-project.org/CRAN/refmans/EnvStats/html/NormalMix.html] (accessed Oct. 26, 2022).