



**UNIVERSITAT ROVIRA I VIRGILI**

**Facultat de Química**

**INCERTIDUMBRE EN  
MÉTODOS ANALÍTICOS  
DE RUTINA**

**ALICIA MAROTO SÁNCHEZ**

**Tesis Doctoral  
Tarragona, 2002**



*Incertidumbre en Métodos Analíticos de Rutina*

**Tesis Doctoral**

**UNIVERSITAT ROVIRA I VIRGILI**







**UNIVERSITAT ROVIRA I VIRGILI**

**Departament de Química Analítica i Química Orgànica**

Àrea de Química Analítica

**INCERTIDUMBRE EN MÉTODOS  
ANALÍTICOS DE RUTINA**

Memoria presentada por

**ALICIA MAROTO SÁNCHEZ**

para conseguir el grado de

**Doctor en Química**

Tarragona, octubre de 2002



*Quizás pueda haber gente que piense que hacer una tesis es sinónimo de intensos años de trabajo en solitario. Afortunadamente, éste no ha sido mi caso ya que no habría podido acabar esta tesis sin la ayuda de mucha gente.*

*En primer lugar, quiero agradecer a mis directores de tesis, Ricard y Jordi, su constante dedicación y los ánimos que me han dado siempre que los he necesitado. Al final de la tesis, puedo decir que han sido unos grandes directores, pero que todavía han sido mucho mejor amigos.*

*En segundo lugar, agradezco al Prof. F. Xavier Rius, director del Grupo de Quimiometría y Cualimetría, que me haya dado la oportunidad de pertenecer a su grupo de investigación y la confianza que siempre ha depositado en mí.*

*Y por último, agradezco a toda la gente que ha compartido durante todo este tiempo los buenos momentos y los no tan buenos. En especial a mis amigos, a mis compañeros del Grupo de Quimiometría y Cualimetría y a la gente del Grupo de Enología que siempre ha estado dispuesta a proporcionarme su ayuda.*

*Esta tesis esta dedicada con todo mi cariño a mi familia y a Javi por haber estado a mi lado este último año.*

*Gracias a todos.*





*A mi familia y a Javi*



El Dr. RICARD BOQUÉ i MARTÍ, Profesor Titular de Universidad del Departament de Química Analítica i Química Orgànica de la Facultat de Química de la Universitat Rovira i Virgili, y el Dr. JORDI RIU i RUSELL, Investigador Ramón y Cajal del mismo Departamento,

CERTIFICAN: Que la presente memoria que tiene por título: **“INCERTIDUMBRE EN MÉTODOS ANALÍTICOS DE RUTINA”**, ha sido realizada por ALICIA MAROTO SÁNCHEZ bajo nuestra dirección en el Área de Química Analítica del Departament de Química Analítica i Química Orgànica de esta Universidad y que todos los resultados presentados son fruto de las experiencias realizadas por dicho doctorando.

Tarragona, octubre de 2002

Dr. Ricard Boqué i Martí

Dr. Jordi Riu i Rusell



---

**ÍNDICE**
**Capítulo 1. Introducción.**

1.1. Objetivo de la tesis doctoral .....	3
1.2. Estructura de la tesis doctoral .....	4
1.3. Notación .....	5
1.4. Validación de métodos analíticos .....	7
1.5. Trazabilidad, exactitud y veracidad .....	9
1.6. Jerarquía de referencias .....	10
1.7. Verificación de la trazabilidad en medidas químicas .....	14
1.7.1. Verificación de la trazabilidad en un intervalo reducido de concentraciones .....	15
1.7.1.1. Verificación de la trazabilidad utilizando un método de referencia ..	16
1.7.1.2. Verificación de la trazabilidad utilizando un valor de referencia ..	19
1.7.2. Verificación de la trazabilidad en un intervalo amplio de concentraciones .....	21
1.7.3. Verificación periódica de la trazabilidad a través de ensayos de aptitud .....	24
1.8. Precisión de un procedimiento analítico: reproducibilidad, precisión intermedia y repetibilidad .....	26
1.8.1. Cálculo de la precisión intermedia de series diferentes. Diseños anidados .....	27
1.9. Incertidumbre .....	32
1.9.1. Relación de la incertidumbre con otros parámetros metrológicos .....	33
1.9.1.1. Error e incertidumbre .....	33
1.9.1.2. Incertidumbre, veracidad y trazabilidad .....	35
1.9.1.3. Incertidumbre y precisión .....	35
1.10. Cálculo de la incertidumbre. Aproximaciones existentes .....	36
1.10.1. Aproximación ISO .....	38
1.10.2. Aproximación propuesta por el <i>Analytical Methods Committee</i> .....	43
1.10.3. Discusión crítica de las aproximaciones existentes para calcular la incertidumbre .....	47

---

1.11.Referencias .....	51
------------------------	----

**Capítulo 2. Cálculo de la incertidumbre en métodos analíticos de rutina validados a un nivel de concentración.**

2.1. Introducción .....	59
2.2. Fundamento teórico de la aproximación para el cálculo global de la incertidumbre de muestras de rutina aprovechando información de la validación del método .....	60
2.3. Estimating uncertainties of analytical results using information from the validation process (Analytica Chimica Acta, <b>391</b> (1999) 173-185) .....	65
2.4. Cálculo de los términos de incertidumbre utilizando información de la validación del método .....	90
2.4.1. Incertidumbre del procedimiento .....	90
2.4.1.1. Mejora de la estimación de la incertidumbre del procedimiento .....	91
2.4.2. Incertidumbre de la verificación de la trazabilidad .....	92
2.4.2.1. Incertidumbre del valor de referencia .....	93
2.4.2.2. Incertidumbre del valor medio del método analítico .....	95
2.4.3. Incertidumbre del muestreo, submuestreo y pretratamientos .....	96
2.4.4. Otros términos de incertidumbre .....	100
2.4.4.1. Utilización de estudios de robustez y técnicas de regresión para calcular la incertidumbre de factores no considerados en la precisión intermedia .....	100
2.4.4.2. Utilización de muestras adicionadas para calcular la incertidumbre de la variabilidad de la matriz de las muestras de rutina .....	105
2.5. Comparación de aproximación propuesta con la aproximación ISO y la aproximación del <i>Analytical Methods Committee</i> .....	108
2.6. Conclusiones .....	109
2.7. Referencias .....	109

---

**Capítulo 3.** *Cálculo de la incertidumbre en métodos analíticos de rutina validados en un intervalo amplio de concentraciones utilizando muestras adicionadas.*

3.1. Introducción	115
3.2. Aseguramiento de la trazabilidad de los resultados utilizando muestras adicionadas	118
3.2.1. Cálculo del sesgo proporcional. Estudios de recuperación	118
3.2.1.1. Las muestras adicionadas tienen una concentración inicial de analito	121
3.2.1.2. Las muestras adicionadas no contienen el analito	123
3.2.2. Cálculo del sesgo constante. Método de Youden	125
3.3. Cálculo de la incertidumbre en métodos de rutina validados a diferentes niveles de concentración	129
3.4. Measurement uncertainty in analytical methods in which trueness is assessed from recovery assays (Analytica Chimica Acta, <b>440</b> (2001) 171-184)	138
3.5. Estimation of measurement uncertainty using regression techniques and spiked samples (Analytica Chimica Acta, <b>446</b> (2001) 133-145)	163
3.6. Cálculo de la incertidumbre en métodos de rutina validados en un intervalo amplio de concentraciones con blancos de muestras adicionados	189
3.6.1. La precisión es aproximadamente constante con la concentración	189
3.6.2. La precisión varía linealmente con la concentración	192
3.7. Conclusiones	196
3.8. Referencias	198

---

**Capítulo 4. Incorporación del sesgo no significativo en el cálculo de la incertidumbre de los resultados.**

4.1. Introducción .....	201
4.2. Should non significant bias be included in the uncertainty budget? (Accreditation and Quality Assurance, <b>7</b> (2002) 90-94) .....	205
4.3. Effect of non significant proportional bias in the final measurement uncertainty (Analyst, submitted for publication) .....	215
4.4. Conclusiones .....	232
4.5. Referencias .....	233

**Capítulo 5. Conclusiones.**

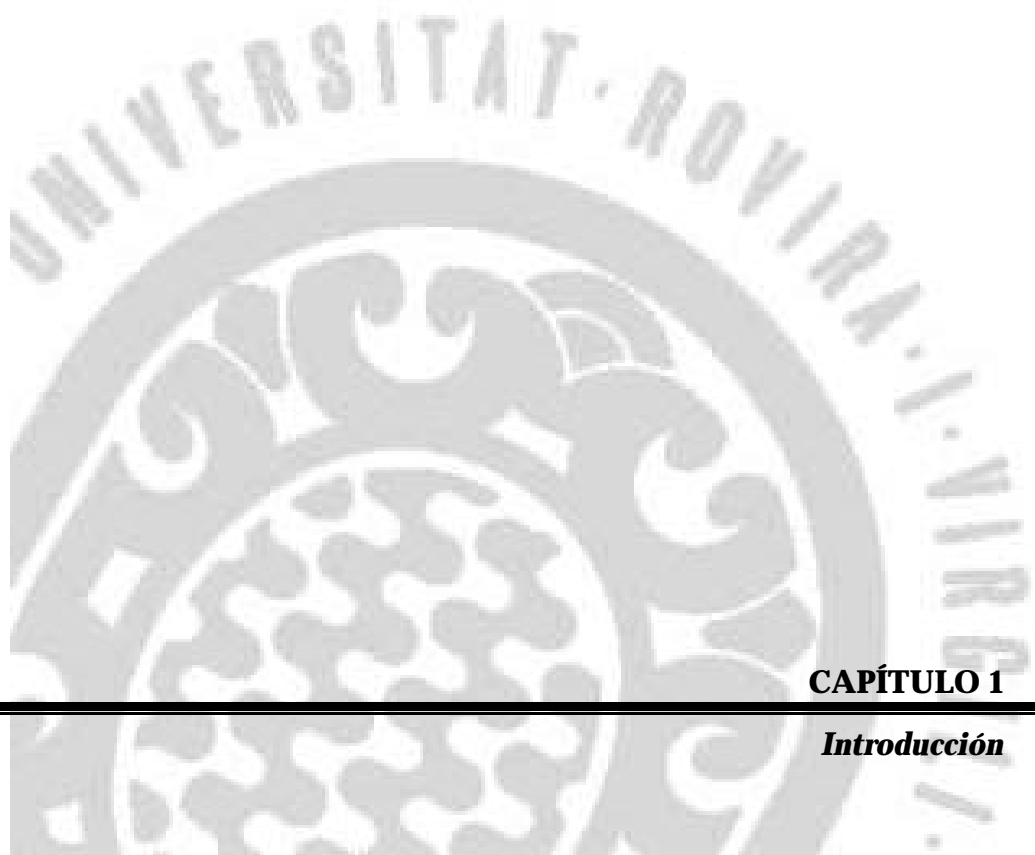
5.1. Conclusiones .....	237
5.2. Perspectivas futuras .....	241

**Apéndices**

<b>Apéndice 1.</b> Evaluating uncertainty in routine analysis (Trends in Analytical Chemistry, <b>18</b> (1999) 577-584) .....	245
<b>Apéndice 2.</b> Critical discussion on the procedures to estimate uncertainty in chemical measurements (Química Analítica, <b>19</b> (2000) 85- 94) .....	261

<b>Glosario</b> .....	285
-----------------------	-----





## **CAPÍTULO 1**

---

### ***Introducción***



## 1.1. Objetivo de la tesis doctoral

Esta tesis doctoral pretende desarrollar metodologías fácilmente aplicables para calcular la incertidumbre de resultados obtenidos con métodos analíticos que se utilizan de forma rutinaria en los laboratorios de análisis. Estas metodologías están basadas en calcular la incertidumbre globalmente aprovechando la información generada en la validación del método y, especialmente, la información generada durante la verificación de la exactitud del método. La ventaja de estas metodologías es que requieren poco trabajo adicional ya que los métodos analíticos siempre deben validarse antes de aplicarlos al análisis de muestras de rutina. Con esta finalidad, se han planteado los siguientes objetivos:

- Desarrollo de una metodología para calcular la incertidumbre de métodos de rutina que se aplican en un intervalo de concentraciones restringido y cuya exactitud se ha verificado a un nivel de concentración utilizando diversas referencias como materiales de referencia certificados, métodos de referencia o ejercicios interlaboratorio.
  
- Desarrollo de metodologías para calcular la incertidumbre de métodos de rutina que se aplican en un intervalo amplio de concentraciones y cuya exactitud se ha verificado utilizando muestras adicionadas y estudios de recuperación.
  
- Estudio de diversas estrategias para evitar la subestimación de la incertidumbre de los resultados analíticos al concluir erróneamente en la verificación de la exactitud que el sesgo del método no es significativo.

## **1.2. Estructura de la tesis doctoral**

La memoria de esta tesis se estructura en cinco capítulos. En el primer capítulo se recogen los fundamentos teóricos que se han utilizado en el desarrollo de esta tesis doctoral. Tras una breve introducción sobre la validación de métodos analíticos, se revisan las diferentes metodologías para verificar la trazabilidad en medidas químicas y se estudian los diseños experimentales para estimar las diferentes medidas de precisión asociadas a un método analítico. Posteriormente, se introduce el concepto de incertidumbre y se estudia su relación con otros parámetros metrológicos como la precisión y la veracidad. Finalmente, se revisan críticamente las diversas aproximaciones existentes en la literatura para el cálculo de la incertidumbre.

En el segundo y tercer capítulos se presentan metodologías para calcular la incertidumbre de resultados obtenidos con métodos analíticos que se usan de forma rutinaria en el laboratorio. En concreto, en el segundo capítulo se desarrolla el cálculo de la incertidumbre de resultados analíticos generados por métodos analíticos validados en un intervalo de concentraciones restringido. Para ello, se utiliza fundamentalmente la información obtenida al verificar la exactitud del método a un nivel de concentración. Además, también se estudia cómo incorporar en el cálculo de la incertidumbre la información obtenida a partir de los gráficos de control, los estudios de precisión y los estudios de robustez.

En el tercer capítulo se presentan diversas metodologías para verificar la trazabilidad y calcular la incertidumbre de resultados obtenidos con métodos analíticos que se aplican a un intervalo amplio de concentraciones. En concreto, se estudian las diversas alternativas para verificar la trazabilidad utilizando estudios de recuperación y muestras adicionadas que pueden tener o no una concentración inicial de analito. Posteriormente, se proponen diversas metodologías para calcular la incertidumbre de los resultados incorporando la información generada en los estudios de recuperación y durante la validación del método. Dichas metodologías dependen de cómo se haya verificado la trazabilidad de los resultados y de cómo varíe la precisión del método con la concentración.

En el cuarto capítulo se estudia cuánto se subestima la incertidumbre de los resultados analíticos cuando se concluye erróneamente en la verificación de la trazabilidad que el sesgo constante y/o proporcional del método no es significativo. Además, se estudian diversas estrategias para evitar la subestimación de la incertidumbre. Estos estudios se han realizado cuando la trazabilidad se verifica a un nivel de concentración o bien en un intervalo amplio de concentraciones.

En el quinto capítulo se presentan las conclusiones generales extraídas de esta tesis y las posibles líneas de investigación futuras. Finalmente, en el anexo 1 se presenta el artículo “*Evaluating uncertainty in routine analysis*” publicado en la revista *Trends in Analytical Chemistry* y en el anexo 2 se presenta el artículo “*Critical discussion on the procedures to estimate uncertainty in chemical measurements*” publicado en la revista *Química Analítica*. Estos artículos figuran en el anexo de la tesis porque parte de su contenido está incluido en el artículo presentado en el segundo capítulo, así como en la introducción del capítulo 2.

### 1.3. Notación

#### ***Símbolos que empiezan por una letra del alfabeto latino***

$a$	Ordenada en el origen de la recta de regresión.
$b$	Pendiente de la recta de regresión.
$c_{ref}$	Concentración de referencia.
$F_{\alpha, v1, v2}$	Valor de la distribución $F$ de Fischer de una cola para un nivel de significancia $\alpha$ y $v1$ y $v2$ grados de libertad.
$F_{hom}$	Factor de homogeneidad.
$k$	Factor de cobertura.
$MS$	Cuadrados medios (“ <i>Mean Squares</i> ”).
$s_a$	Estimación de la desviación estándar de la ordenada en el origen de la recta.

$S_b$	Estimación de la desviación estándar de la pendiente de la recta.
$S_e$	Estimación de la desviación estándar de los residuales de la recta.
$S_I$	Estimación de la desviación estándar intermedia del procedimiento analítico.
$S_{\text{serie}}$	Estimación de la desviación estándar entre series del procedimiento analítico.
$S_r$	Estimación de la desviación estándar de la repetibilidad del procedimiento analítico.
$S_R$	Estimación de la desviación estándar de la reproducibilidad del procedimiento analítico.
$t_{\alpha/2, \nu}$	Valor de la distribución $t$ de Student de dos colas para un nivel de significancia $\alpha$ y $\nu$ grados de libertad.
$t_{\beta, \nu}$	Valor de la distribución $t$ de Student de una cola para un nivel de significancia $\beta$ y $\nu$ grados de libertad.
$u$	Incertidumbre estándar.
$U$	Incertidumbre expandida.
$RSD$	Desviación estándar relativa.
$R$	Recuperación del método.

### ***Símbolos que empiezan por una letra del alfabeto griego***

$\alpha$	Error de primera especie o de tipo I
$\beta$	Error de segunda especie o de tipo II
$\lambda$	Sesgo mínimo que el analista quiere detectar con una probabilidad $\beta$ .
$\mu$	Concentración verdadera de la muestra.
$\sigma$	Desviación estándar verdadera.
$\sigma_I$	Desviación estándar intermedia verdadera del procedimiento analítico.
$\sigma_{\text{serie}}$	Desviación estándar entre series verdadera del procedimiento analítico.
$\sigma_r$	Desviación estándar verdadera de la repetibilidad del procedimiento analítico.

---

$\sigma_R$	Desviación estándar verdadera de la reproducibilidad del procedimiento analítico.
$\varepsilon$	Error aleatorio dentro de una misma serie.
$\delta_{serie}$	Sesgo de la serie.
$\delta_{raz}$	Sesgo del método analítico.
$\delta_{cte}$	Sesgo constante del método.
$\delta_{prop}$	Sesgo proporcional del método.
$\delta_{pret}$	Sesgo debido a los pretratamientos.
$\delta_{otros}$	Sesgo debido a otros términos.
$v_{eff}$	Grados de libertad efectivos.
$v$	Grados de libertad.

## 1.4. Validación de métodos analíticos

Hoy en día, los laboratorios deben demostrar que sus métodos analíticos proporcionan resultados fiables y adecuados para su finalidad o propósito perseguido<sup>1,2</sup> ya que muchas de las decisiones que se toman están basadas en la información que estos resultados proporcionan. La validación de las metodologías, junto a otras actividades englobadas en el control del aseguramiento de la calidad, permite demostrar a los laboratorios que sus métodos analíticos proporcionan resultados fiables.

Validar un método analítico consiste en verificar y documentar su validez, esto es, su adecuación a unos determinados requisitos, previamente establecidos por el usuario, para poder resolver un problema analítico particular. Estos requisitos son los que definen los parámetros o criterios de calidad que debe poseer el método a utilizar para resolver el problema analítico. La Figura 1.1 muestra que estos criterios de calidad pueden ser de tipo estadístico o bien de tipo operativo/económico.

<b>Tipo estadístico</b>	<b>Tipo operativo/económico</b>
✓ Veracidad, Trazabilidad	✓ Inversión
✓ Precisión, Incertidumbre	✓ Mantenimiento
✓ Representatividad	✓ Rapidez
✓ Sensibilidad	✓ Facilidad de uso
✓ Selectividad	✓ Simplicidad
✓ Límite de detección	
✓ Límite de cuantificación	
✓ Robustez	

**Figura 1.1.** Parámetros de calidad de un método analítico.

Según la norma ISO/IEC 17025,<sup>1</sup> los laboratorios deben validar todos los métodos que se utilicen en el laboratorio, tanto los desarrollados por ellos mismos como aquellos procedentes de fuentes bibliográficas o desarrollados por otros laboratorios. Además, también es necesario que el laboratorio valide los métodos de referencia aunque, en este caso, no es necesario que el laboratorio realice una validación completa. Asimismo, el laboratorio debe validar *todo* el procedimiento\* analítico teniendo en cuenta el intervalo de concentraciones y de matrices de las muestras de rutina. Los criterios de calidad que al menos deben verificarse son la trazabilidad, la precisión y la incertidumbre de los resultados obtenidos con el método ya que, de esta forma, se obtienen resultados trazables y comparables.

---

\* El procedimiento analítico es el resultado de la adaptación de un método analítico, disponible en la bibliografía para determinar un analito, al análisis de muestras caracterizadas por la naturaleza de la matriz.



## 1.5. Trazabilidad, exactitud y veracidad

El Vocabulario Internacional de Metrología, VIM<sup>3</sup>, define la trazabilidad como "la propiedad del resultado de una medida que le permite relacionarlo con referencias determinadas, generalmente nacionales o internacionales, a través de una cadena ininterrumpida de comparaciones todas ellas con incertidumbres conocidas". Esta cadena de comparaciones nos permite tener confianza en los resultados obtenidos ya que, como nuestros resultados son comparables a una referencia, podemos asegurar que no se comete ningún error sistemático significativo en ninguna de las etapas del método analítico. Además, la verificación de la trazabilidad permite asegurar que nuestros resultados son comparables a los obtenidos por otros laboratorios.

En el campo de las medidas físicas, la trazabilidad se consigue mediante el uso de patrones certificados e instrumentos calibrados que se interconectan entre sí hasta llegar al patrón físico fundamental, por ejemplo, el kilogramo, el metro o el segundo<sup>4</sup>. No obstante, en el análisis químico es muy difícil llegar a través de una cadena ininterrumpida de comparaciones a la unidad fundamental de materia en el sistema internacional (SI) de unidades, *el mol*, ya que en la mayoría de los procedimientos analíticos no es posible asegurar la trazabilidad de cada una de sus etapas<sup>5,6</sup>. Esto es debido a que el tipo de muestra puede afectar mucho a los resultados<sup>7</sup>. Además, la naturaleza de los errores es diferente en medidas químicas y físicas: en las físicas predominan los errores sistemáticos mientras que en las químicas predominan los aleatorios. La alternativa en el análisis químico consiste en trazar globalmente los resultados obtenidos con el procedimiento analítico a un valor de referencia<sup>6,8-10</sup> y evaluar si los resultados obtenidos con dicho procedimiento son comparables a los proporcionados por la referencia utilizada.

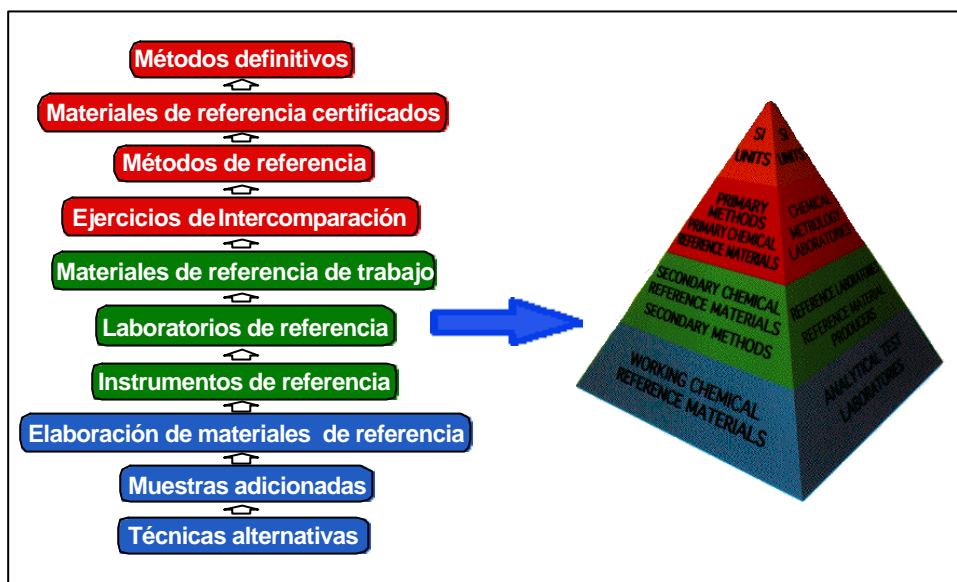
En la verificación de la trazabilidad de los resultados, implícitamente se está comprobando la ausencia de sesgo, o lo que es lo mismo, de un error sistemático en los resultados. El concepto de trazabilidad es, pues, muy similar al concepto de exactitud que, tradicionalmente, se había considerado como la ausencia de errores sistemáticos en el procedimiento. Sin embargo, según la norma ISO 3534<sup>11</sup>, la exactitud se define como el grado de concordancia entre el resultado de un ensayo y el valor de referencia aceptado. Por tanto, además de considerar los errores

sistemáticos, el término exactitud considera también los errores aleatorios ya que éstos siempre están presentes en el resultado de una medida. Es decir, exactitud es suma de dos conceptos: veracidad y precisión (*trueness+precision*)<sup>12</sup>. Veracidad, también según la norma ISO 3534, es el grado de concordancia entre el valor medio obtenido a partir de una serie de resultados de ensayo y un valor de referencia aceptado. Es decir, un resultado es veraz si está libre de error sistemático. En la práctica, la veracidad se verifica utilizando referencias. Por tanto, decir que un resultado es veraz es equivalente a afirmar que el resultado es trazable a la referencia utilizada. Por otro lado, la precisión, según la misma norma ISO 3534, es el grado de concordancia entre ensayos independientes obtenidos bajo unas condiciones estipuladas. Mediante la precisión se evalúan los errores aleatorios. Por lo tanto, un resultado es exacto si simultáneamente es veraz (se encuentra libre de errores sistemáticos) y preciso (los errores aleatorios son aceptables). Para evaluar si los errores aleatorios son aceptables se suelen aplicar tests estadísticos para comparar la varianza asociada a nuestro método con la varianza asociada a otro método o bien con la publicada en la bibliografía. Sin embargo, no tiene sentido aplicar estos tests estadísticos si previamente no se conocen las condiciones en que se han obtenido estas varianzas. Esto es debido a que la precisión obtenida depende de las condiciones en que se obtienen los resultados (por ejemplo, la precisión de un método es mucho mejor cuando los resultados se obtienen en un mismo día dentro de un mismo laboratorio que no cuando se obtienen por diferentes laboratorios).

## 1.6. Jerarquía de referencias

Como se ha dicho anteriormente, la trazabilidad de los resultados analíticos se suele verificar globalmente utilizando un valor de referencia. Este valor de referencia puede ser, por ejemplo, la concentración de un material de referencia certificado (CRM), el valor de consenso obtenido en un ejercicio interlaboratorio o la concentración encontrada con un método de referencia. Cada una de estas referencias tiene, teóricamente, un nivel de trazabilidad a la unidad fundamental, el mol. La Figura 1.2 muestra la relación entre la pirámide metrológica propuesta

por Fleming *et al*<sup>13</sup> y diversas referencias utilizadas usualmente para verificar la trazabilidad en medidas químicas, y que se describen a continuación.



**Figura 1.2.** Relación entre la pirámide metroológica y diversas referencias utilizadas en medidas químicas para verificar la trazabilidad.

#### *Métodos definitivos o primarios*

Los métodos definitivos son aquellos que están directamente ligados al sistema internacional de unidades. Son la espectrometría de masas con dilución isotópica, la volumetría, la coulombimetría, la gravimetría, un grupo de métodos coligativos y la espectroscopia de resonancia magnética nuclear<sup>14,15</sup>. La utilización de uno de estos métodos supone la mejor referencia posible siempre que éstos sean aplicados en condiciones rigurosas de garantías de calidad. Sin embargo, estas condiciones rigurosas, junto con la dificultad de aplicación de alguno de ellos o el reducido ámbito de aplicación de otros, son el motivo por el cual, a pesar de ser teóricamente la mejor referencia posible, los métodos definitivos se utilizan muy poco en la práctica para verificar la trazabilidad.

#### *Materiales de referencia certificados*

Un material de referencia certificado<sup>16</sup> es un material o sustancia que tiene certificadas una o varias de sus propiedades por procedimientos técnicamente válidos llevados a cabo por un organismo competente de manera que permite su uso para calibrar un aparato o instrumento, validar un método analítico o asignar valores a un material o sistema. Un material de referencia debe cumplir una serie de requisitos para poderlo considerar como tal: debe garantizarse la trazabilidad del valor de la concentración o el valor del parámetro que se quiera determinar, debe ser homogéneo y debe ser estable durante un tiempo razonable. Además, debería ser lo más parecido posible a las muestras reales que se analizarán con el método analítico y debe conocerse la incertidumbre asociada al valor de referencia.

Los materiales de referencia pueden ocupar diferentes niveles en la jerarquía de referencias dependiendo de cómo se haya obtenido el valor de referencia<sup>17</sup>. Los materiales de referencia primarios (PRM, *primary reference materials*) son los que tienen un nivel de trazabilidad mayor ya que sus valores de referencia se obtienen utilizando métodos definitivos. Los materiales de referencia certificados (CRM, *certified reference materials*) tienen un nivel de trazabilidad menor ya que el valor de referencia no se obtiene utilizando métodos primarios sino que se obtiene a partir de los resultados proporcionados por laboratorios de referencia que han utilizado varios métodos de referencia. Por último, los materiales de referencia de trabajo (WRM, *working reference materials*) son los que tienen un nivel de trazabilidad menor ya que la concentración de referencia la obtiene un laboratorio acreditado que ha analizado el WRM con un solo método previamente validado.

Los principales inconvenientes de los materiales de referencia son su elevado precio y que, debido a la variedad de matrices de las muestras reales, variedad de mensurandos y variedad de niveles de concentración, sólo pueden utilizarse para verificar la trazabilidad de entre un 5-10 % de determinaciones analíticas.

#### *Ejercicios interlaboratorio*

El nivel de trazabilidad obtenido al utilizar como referencia el valor de consenso generado en un ejercicio interlaboratorio depende de cómo sea el tipo de ejercicio

en que participa el laboratorio. Aunque existen numerosas variantes, se considera que hay tres tipos de ejercicios interlaboratorio:

- **Ensayos de aptitud:** el objetivo de estos ejercicios es comparar la competencia de los laboratorios cuando cada uno de ellos analiza la muestra utilizando su propio método analítico<sup>18</sup>. Estos ejercicios se suelen realizar periódicamente ya que constituyen una parte esencial dentro de los requisitos necesarios para la acreditación de los laboratorios <sup>19,20</sup>. Los ensayos de aptitud son los que tienen un nivel de trazabilidad menor ya que los laboratorios participantes no tienen que demostrar su competencia para participar en estos ejercicios.
  
- **Ejercicios colaborativos:** el objetivo de estos ejercicios es comprobar la idoneidad de una nueva metodología analítica o de un método recientemente modificado. Para ello, laboratorios que previamente han debido demostrar su competencia, analizan una muestra concreta siguiendo tan fielmente como sea posible el procedimiento analítico propuesto<sup>21,22</sup>. A partir de los resultados obtenidos en el ejercicio se determina la repetibilidad y la reproducibilidad del método<sup>23</sup> y si el método tiene un error sistemático<sup>24</sup>. El nivel de trazabilidad de estos ejercicios es mayor que el de los ensayos de aptitud.
  
- **Ejercicios de certificación:** el objetivo de estos ejercicios es establecer el valor de referencia y la incertidumbre de la propiedad certificada en un material de referencia. Estos ejercicios son los que tienen un nivel de trazabilidad mayor ya que, laboratorios que previamente han demostrado su competencia, analizan la muestra utilizando métodos analíticos cuya exactitud se haya verificado previamente<sup>25</sup>. La BCR<sup>26</sup>, la ISO<sup>27</sup> y la LGC<sup>28</sup> recomiendan que, siempre que sea posible, se utilicen varios métodos reconocidos por su exactitud y que además sean independientes entre sí. Además, cada uno de estos métodos debería aplicarse en diferentes laboratorios. De esta forma, se evita que el valor de referencia tenga un sesgo debido al método utilizado<sup>29</sup>.

#### *Métodos de referencia*

La verificación de la trazabilidad en este caso se lleva a cabo analizando muestras representativas con el método que se quiere validar y con el método de referencia

que suele ser un método normalizado o un método oficial de análisis (validado por alguna organización de reconocido prestigio utilizando ejercicios interlaboratorio de tipo colaborativo). Esta comparación será válida siempre y cuando el método de referencia se aplique en condiciones de aseguramiento de la calidad. El caso más usual es que las muestras representativas se analicen en un solo laboratorio (“*within-laboratory validation*”)<sup>30,31</sup>. Sin embargo, también pueden ser analizadas por varios laboratorios (“*inter-laboratory validation*”)<sup>32</sup> consiguiendo, de esta forma, un nivel de trazabilidad mayor.

#### *Muestras adicionadas*

Esta referencia es la que posee un nivel de trazabilidad menor ya que tiene la desventaja de que el analito adicionado puede tener un comportamiento diferente al que se encuentra en la muestra<sup>33-35</sup>. No obstante, es una de las referencias más utilizadas ya que muchas veces no se dispone de referencias con un nivel de trazabilidad mayor.

### **1.7. Verificación de la trazabilidad en medidas químicas**

Como se ha dicho anteriormente, la trazabilidad de los resultados analíticos se verifica utilizando valores de referencia. En este proceso debe comprobarse que los resultados obtenidos al analizar la muestra de referencia son comparables a su valor de referencia asignado. En realidad, siempre habrá una diferencia entre ambos valores debida a errores aleatorios y también a un posible error sistemático. Para comprobar si esta diferencia es únicamente debida a errores aleatorios y que, por tanto, los resultados obtenidos con el procedimiento analítico son trazables, es necesario aplicar tests estadísticos. Sin embargo, la aplicación de estos tests lleva consigo el riesgo de cometer dos tipos de errores. Estos errores son: 1) concluir que los resultados no son trazables cuando en realidad sí lo son y 2) concluir que los resultados son trazables cuando en realidad no lo son. El error cometido en 1) se conoce como error de primera especie o error  $\alpha$ . Al llevar a cabo los tests estadísticos, se escoge la probabilidad  $\alpha$  de cometer este error. Por otro lado, el error cometido en 2) se conoce como error de segunda especie o error  $\beta$ . La probabilidad de cometer este error depende de varios factores:

- La probabilidad  $\alpha$  escogida: cuanto menor sea  $\alpha$  mayor es la probabilidad de cometer un error  $\beta$ . Normalmente, el valor de  $\alpha$  escogido es 0.05. Este valor significa que hay un 5% de probabilidad de decir que el procedimiento analítico no es trazable cuando en realidad sí lo es. Esta probabilidad de equivocarse es debida a los errores aleatorios.
- El error sistemático,  $\lambda$ , que se quiere detectar en el test estadístico: cuanto menor sea éste, mayor es la probabilidad de cometer un error  $\beta$ .
- La precisión del procedimiento analítico y la incertidumbre del valor de referencia: cuanto peor sea la precisión del procedimiento y mayor sea la incertidumbre del valor de referencia, mayor es la probabilidad de cometer un error  $\beta$ .
- El número de veces que se ha analizado la muestra de referencia con el procedimiento: la probabilidad de cometer un error  $\beta$  disminuye a medida que aumenta el número de análisis.

En la práctica, la trazabilidad se puede verificar de diversas maneras dependiendo de: 1) que se verifique en un intervalo reducido o en un intervalo amplio de concentraciones; y 2) del tipo de referencia utilizado. Una vez que se haya verificado la trazabilidad, el laboratorio podrá trasladar dicha trazabilidad a los resultados que obtenga sobre muestras futuras. Para ello, deberá cumplirse que las muestras analizadas con el procedimiento sean similares a la muestra utilizada como referencia y que el laboratorio actúe bajo condiciones de aseguramiento de la calidad<sup>36,37</sup>. En los siguientes subapartados, se explica cómo verificar la trazabilidad en los casos más frecuentes.

### **1.7.1. Verificación de la trazabilidad en un intervalo reducido de concentraciones**

La trazabilidad de los resultados puede verificarse a un nivel de concentración cuando el método se aplica en un intervalo reducido de concentraciones. En este caso, se asume que el sesgo es el mismo para todo el intervalo de concentraciones de las muestras de rutina. Si este sesgo no es significativo, el procedimiento es

veraz y, en el caso de que se use una referencia nacional o internacional, trazable a dicha referencia. El procedimiento que debe realizarse para verificar la trazabilidad depende del tipo de referencia utilizado. Podemos distinguir dos casos:

- a) La trazabilidad de los resultados se verifica utilizando un método analítico de referencia (es decir, un método definitivo, un método estándar o bien otro tipo de método). El método de referencia debería ser independiente del método que se quiere validar.
- b) La trazabilidad de los resultados se verifica utilizando un valor de referencia (es decir, utilizando un material de referencia, el valor consenso de un ejercicio interlaboratorio o bien la cantidad de analito adicionada).

### **1.7.1.1. Verificación de la trazabilidad utilizando un método de referencia**

En la norma ISO 5725-6<sup>32</sup> se describe cómo verificar la veracidad de un método alternativo frente a un método de referencia utilizando la información generada cuando varios laboratorios analizan la muestra de validación con ambos métodos. Esta aproximación tiene el inconveniente de que no puede aplicarse cuando un laboratorio individual quiere verificar la trazabilidad de sus resultados. Esto ha hecho que Kuttatharmmakul *et al*<sup>30,31</sup> hayan adaptado la aproximación propuesta por la norma ISO 5725-6 para que pueda aplicarse a un laboratorio individual. Esta aproximación propone analizar la muestra con ambos métodos en condiciones intermedias de precisión, es decir, variando todos aquellos factores (día, analista, instrumento,...) que puedan afectar a los resultados. Además, verifica la trazabilidad del método considerando tanto la probabilidad  $\alpha$  de concluir erróneamente que el método no es trazable como la probabilidad  $\beta$  de concluir erróneamente que el método es trazable. Para verificar la trazabilidad con esta aproximación es necesario: 1) Calcular las veces que la muestra de validación debe analizarse con ambos métodos para detectar un sesgo mínimo,  $\lambda$ , con una probabilidad  $\beta$ ; y 2) Una vez que se ha analizado la muestra de validación, verificar que el método alternativo no tiene un sesgo significativo.

1) *Determinación de las veces que debe analizarse la muestra de validación con ambos métodos.* Se utiliza la siguiente expresión:



$$I \geq (t_{\alpha/2} + t_{\beta}) \sqrt{\frac{(p_A - 1) \cdot s_{I,A}^2 + (p_B - 1) \cdot s_{I,B}^2}{p_A + p_B - 2} \cdot \left( \frac{1}{p_A} + \frac{1}{p_B} \right)} \quad (1.1)$$

donde  $\lambda$  corresponde al sesgo mínimo que el analista quiere detectar con una probabilidad  $\beta$ .  $p_A$  y  $p_B$  corresponden, respectivamente, a las veces que debe analizarse la muestra de validación con el método de referencia y con el método alternativo;  $t_{\alpha/2}$  es el valor de  $t$  tabulado de dos colas para un nivel de significancia  $\alpha$  y  $p_A + p_B - 2$  grados de libertad;  $t_{\beta}$  es el valor de  $t$  tabulado de una cola para un nivel de significancia  $\beta$  y  $p_A + p_B - 2$  grados de libertad; y  $s_{I,A}^2$  y  $s_{I,B}^2$  corresponden, respectivamente, a las varianzas asociadas a la precisión intermedia del método de referencia y del método alternativo.

Esta expresión está basada en un test  $t$  de comparación de dos medias donde se ha asumido que las varianzas intermedias verdaderas de ambos métodos son iguales ya que normalmente no se dispone de una estimación de la precisión intermedia del método alternativo. En este caso, lo que se hace es sustituir  $s_{I,B}$  por  $s_{I,A}$ . Además, se recomienda que la muestra de validación se analice el mismo número de veces con ambos métodos. En este caso,  $p_A = p_B$  y la Ec. 1.1 se simplifica a:

$$I \geq (t_{\alpha/2} + t_{\beta}) \sqrt{\frac{2 \cdot s_{I,A}^2}{p_A}} \quad (1.2)$$

En esta expresión la única incógnita es  $p_A$  ya que el analista fija el sesgo mínimo que quiere detectar,  $\lambda$ , y la probabilidad de error  $\alpha$  y  $\beta$ . Por tanto, despejando en la Ec. 1.2 se obtiene el número de veces,  $p_A$ , que debe analizarse la muestra de validación con ambos métodos.

2) Verificar que el método alternativo no tiene un sesgo significativo. Una vez que se ha analizado la muestra de validación con ambos métodos, debe utilizarse un test  $t$  para verificar que el método alternativo no tiene un sesgo significativo:

$$t_{\text{cal}} = \frac{|\bar{x}_A - \bar{x}_B|}{s_d} \quad (1.3)$$

donde  $\bar{x}_A$  y  $\bar{x}_B$  corresponden, respectivamente, a la media de los resultados obtenidos al analizar la muestra de validación  $p_A$  veces con el método de referencia y  $p_B$  veces con el método alternativo (donde, normalmente  $p_A=p_B$ ).  $s_d$  es la desviación estándar asociada a la diferencia de ambas medias y su expresión depende de que las varianzas asociadas a ambos valores medios sean o no comparables. Por tanto, debe hacerse un test  $F$  para ver si existen diferencias significativas entre las varianzas asociadas a la precisión intermedia de ambos métodos. En el caso de que no haya diferencias significativas,  $s_d$  se calcula como:

$$s_d = \sqrt{\frac{(p_A - 1) \cdot s_{I,A}^2 + (p_B - 1) \cdot s_{I,B}^2}{p_A + p_B - 2}} \quad (1.4)$$

Si hay diferencias significativas entre ambas varianzas,  $s_d$  se calcula como:

$$s_d = \sqrt{\frac{s_{I,A}^2}{p_A} + \frac{s_{I,B}^2}{p_B}} \quad (1.5)$$

El valor de  $t_{\text{cal}}$  obtenido debe compararse con el valor de  $t$  tabulado,  $t_{\text{tab}}$  de dos colas para un nivel de significancia  $\alpha$  y los grados de libertad,  $\nu$ , asociados a  $s_d$ . En el caso de que las varianzas sean comparables, los grados de libertad de  $s_d$  corresponden a  $p_A+p_B-2$ . Si no son comparables, los grados de libertad deben calcularse con la aproximación de Welch-Satterthwaite<sup>38-40</sup>:

$$\nu = \frac{s_d^4}{\frac{s_{I,A}^4}{p_A^2 \cdot (p_A - 1)} + \frac{s_{I,B}^4}{p_B^2 \cdot (p_B - 1)}} \quad (1.6)$$

Si  $t_{\text{tab}} < t_{\text{cal}}$ , no hay diferencias significativas entre las medias de los resultados obtenidos con ambos métodos. Por tanto, el método alternativo no tiene un error sistemático significativo y es trazable al método de referencia. Además, en el caso

de que el método tenga un sesgo mínimo  $\lambda$ , se puede asegurar que la probabilidad de concluir erróneamente que el método es trazable es inferior o igual al error  $\beta$  fijado al calcular las veces que debe analizarse la muestra de validación con ambos métodos.

### 1.7.1.2. Verificación de la trazabilidad utilizando un valor de referencia

La aproximación propuesta por Kuttatharmmakul *et al*<sup>30,31</sup> también puede aplicarse cuando la trazabilidad se verifica utilizando un valor de referencia. Este es el caso de que se utilice un material de referencia certificado (CRM), el valor de consenso obtenido en un ejercicio interlaboratorio o bien la cantidad de analito adicionada. A continuación se explica cómo adaptar la aproximación propuesta por Kuttatharmmakul *et al* para: 1) calcular las veces que debe analizarse la muestra de referencia para poder detectar un sesgo mínimo,  $\lambda$ ; y 2) verificar que los resultados proporcionados por el laboratorio al analizar la muestra de referencia no tienen un sesgo significativo.

1) *Determinación de las veces que debe analizarse la muestra de referencia.* En este caso, se debe disponer de una estimación de la precisión intermedia del método que se valida. Para calcular el número de veces,  $p_B$ , que debe analizarse la muestra de referencia se utiliza la siguiente expresión:

$$I \geq (t_{a/2} + t_b) \sqrt{u_{\text{ref}}^2 + \frac{S_{\text{I,B}}^2}{p_B}} \quad (1.7)$$

En esta expresión la única incógnita es  $p_B$  ya que debe conocerse la precisión intermedia del método,  $s_{\text{I,B}}$ , y la incertidumbre del valor de referencia,  $u_{\text{ref}}$ . Al igual que en el apartado 1.7.1.1, el analista debe fijar el sesgo mínimo,  $\lambda$ , y la probabilidad de error  $\alpha$  y  $\beta$ .

La incertidumbre del valor de referencia,  $u_{\text{ref}}$ , depende de la referencia utilizada. Si es un CRM, se calcula como:

$$u_{\text{ref}} = \frac{U_{\text{CRM}}}{k} \quad (1.8)$$

donde  $U_{\text{CRM}}$  es la incertidumbre asociada a la concentración de referencia y  $k$  es el factor de cobertura (normalmente  $k=2$ ). Si la referencia es el valor de consenso de un ejercicio interlaboratorio,  $u_{\text{ref}}$  se calcula como:

$$u_{\text{ref}} = \frac{s_{\text{lab}}}{\sqrt{n_{\text{lab}}}} \quad (1.9)$$

donde  $n_{\text{lab}}$  es el número de laboratorios participantes del ejercicio interlaboratorio con los que se ha calculado el valor de consenso y  $s_{\text{lab}}$  es la desviación estándar de los resultados proporcionados por los laboratorios participantes ( $u_{\text{ref}}$  se ha calculado asumiendo que cada laboratorio proporciona sólo un resultado y que el valor de consenso se ha obtenido como el valor medio de los resultados de los laboratorios participantes después de eliminar los laboratorios que proporcionan resultados discrepantes).

2) Verificar que los resultados obtenidos al analizar la muestra de referencia no tienen un sesgo significativo. Una vez que se ha analizado la muestra de referencia  $p_B$  veces con el método analítico, debe utilizarse un test  $t$  para verificar que el método analítico no tiene un sesgo significativo:

$$t_{\text{cal}} = \frac{|\bar{x}_{\text{ref}} - \bar{x}_B|}{\sqrt{u_{\text{ref}}^2 + \frac{s_{\text{I,B}}^2}{p_B}}} \quad (1.10)$$

donde  $x_{\text{ref}}$  es el valor de referencia y  $\bar{x}_B$  es la media de los resultados obtenidos cuando la muestra de referencia se analiza  $p_B$  veces con el método analítico. La incertidumbre  $u_{\text{ref}}$  se calcula con la Ec. 1.8 si la referencia utilizada es un CRM y con la Ec. 1.9 si la referencia utilizada es el valor de consenso de un ejercicio interlaboratorio.

El valor de  $t_{\text{cal}}$  obtenido debe compararse con el valor de  $t$  tabulado,  $t_{\text{tab}}$ , de dos colas para un nivel de significancia  $\alpha$  y los grados de libertad,  $\nu$ , calculados con la aproximación de Welch-Satterthwaite<sup>38-40</sup>:

$$\mathbf{n} = \frac{\left( u_{\text{ref}}^2 + \frac{S_{\text{I,B}}^2}{p_{\text{B}}} \right)^2}{\frac{u_{\text{ref}}^4}{\mathbf{n}_{\text{ref}}} + \frac{S_{\text{I,B}}^4}{p_{\text{B}}^2 \cdot (p_{\text{B}} - 1)}} \quad (1.11)$$

donde  $\nu_{\text{ref}}$  corresponde a los grados de libertad asociados al valor de referencia. En el caso de que se utilice un valor de consenso corresponden a  $m_{\text{ab}} - 1$  (siendo  $m_{\text{ab}}$  el número de laboratorios con los que se ha calculado el valor de consenso del ejercicio interlaboratorio).

Si  $t_{\text{tab}} < t_{\text{cal}}$ , no hay diferencias significativas entre el valor de referencia y el valor medio obtenido al analizar la muestra de referencia con el método analítico. Por tanto, el método analítico no tiene un error sistemático significativo y es trazable a la referencia utilizada (CRM o valor de consenso). Además, en el caso de que el método tenga un sesgo mínimo  $\lambda$ , se puede asegurar que la probabilidad de concluir erróneamente que el método es trazable es inferior o igual al error  $\beta$  fijado para calcular las veces,  $p_{\text{B}}$ , que debe analizarse la muestra de referencia.

### 1.7.2. Verificación de la trazabilidad en un intervalo amplio de concentraciones

Cuando la trazabilidad de los resultados debe verificarse en un intervalo amplio de concentraciones es necesario utilizar más de una muestra de referencia, pues no se puede asumir que el error sistemático sea el mismo en todo el intervalo de aplicación del método. Esto es debido a que puede haber dos tipos de errores sistemáticos: uno constante (no depende de la concentración de la muestra) y uno proporcional (depende de la concentración de la muestra y se suele expresar como un factor de recuperación).

Las muestras de referencia deben estar distribuidas a lo largo del intervalo de concentraciones donde se quiere verificar la trazabilidad de los resultados. Normalmente, el tipo de referencias utilizadas son muestras adicionadas a diversos niveles de concentración aunque también puede utilizarse otro tipo de referencias como, por ejemplo, los métodos de referencia o los CRM. La trazabilidad puede verificarse de dos formas<sup>41</sup>: 1) utilizando el método de la recuperación media; y 2) utilizando técnicas de regresión.

#### 1) Método de la recuperación media

La trazabilidad se verifica calculando la recuperación para cada una de las  $n$  muestras de referencia<sup>34,42</sup>. Posteriormente, debe calcularse el valor medio de las  $n$  recuperaciones calculadas,  $\bar{R}$ , así como su desviación estándar,  $s_R$ . El método es trazable si la recuperación no difiere significativamente de 1. Esto se verifica con un test  $t$ :

$$t_{\text{cal}} = \frac{|R - 1|}{s_R / \sqrt{n}} \quad (1.12)$$

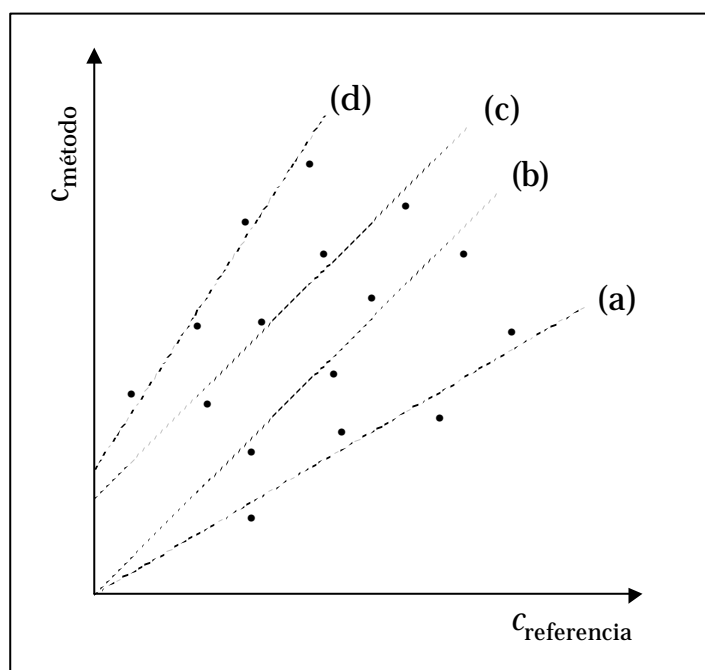
Este valor debe compararse con el valor de  $t$  tabulado,  $t_{\text{tab}}$ , de dos colas para un nivel de significancia  $\alpha$  y  $n-1$  grados de libertad. Si  $t_{\text{cal}} < t_{\text{tab}}$ , la recuperación no difiere significativamente de 1 y, por tanto, el método no tiene un error proporcional. En este caso, los resultados proporcionados por el método analítico son trazables a la referencia utilizada. No obstante, debe tenerse en cuenta que esta aproximación no considera que el método analítico pueda tener un sesgo constante<sup>33</sup>.

#### 2) Técnicas de regresión

La trazabilidad de los resultados se verifica representando la concentración encontrada con el método analítico frente a la concentración de referencia. La Fig. 1.3 muestra las diferentes situaciones que se pueden dar. En el caso de que no haya errores sistemáticos, los datos deben ajustarse a una recta con pendiente unidad y ordenada en el origen cero (caso (b) de la Fig. 1.3). La presencia de un sesgo proporcional hace que la recta tenga una pendiente distinta de la unidad (casos (d)

y (a) de la Fig. 1.3). Por último, la presencia de un sesgo constante hace que la recta tenga una ordenada diferente de cero (casos (d) y (c) de la Fig. 1.3).

A partir de una regresión por mínimos cuadrados, pueden ajustarse los datos a una recta. Como hay errores aleatorios, los datos no se ajustarán exactamente a una recta  $y=x$ . Por tanto, a partir de tests estadísticos debe verificarse si las diferencias son debidas únicamente a errores aleatorios o también a errores sistemáticos. A partir de tests individuales sobre la pendiente y la ordenada en el origen se puede comprobar la ausencia de sesgo proporcional y constante<sup>43</sup>. Sin embargo, debido a la covarianza entre la pendiente y la ordenada en el origen, los tests individuales pueden dar resultados incorrectos cuando se quiere verificar la ausencia de sesgo proporcional y constante. Esto hace que la ausencia de ambos sesgos deba verificarse conjuntamente utilizando un test conjunto de la pendiente y la ordenada en el origen<sup>44</sup>.



**Figura 1.3.** Diferentes situaciones que pueden producirse cuando la trazabilidad del método se verifica representando la concentración encontrada con el método analítico ( $C_{\text{método}}$ ) frente a la concentración de referencia ( $C_{\text{referencia}}$ ).

El inconveniente de verificar la trazabilidad utilizando la regresión por mínimos cuadrados es que se asume que la precisión del método se mantiene constante en el intervalo de concentraciones y que la concentración de referencia está libre de error. En algunos casos, no se puede asumir que la precisión se mantenga constante en el intervalo de concentraciones. Por tanto, es mejor utilizar la regresión por mínimos cuadrados ponderados (WLS)<sup>45</sup>. En el caso de que la incertidumbre de la concentración de referencia no sea despreciable frente a la incertidumbre de la concentración encontrada, deben usarse otras técnicas de regresión que consideren error en los dos ejes. Una de estas técnicas es la regresión por mínimos cuadrados bivariantes (BLS)<sup>46</sup>. Además, para este tipo de regresión se ha desarrollado un test conjunto para la ordenada en el origen y la pendiente<sup>47</sup> que verifica conjuntamente la ausencia de sesgo constante y proporcional.

### **1.7.3. Verificación periódica de la trazabilidad a través de ensayos de aptitud**

Los ensayos de aptitud constituyen una parte esencial dentro de los requisitos necesarios para la acreditación ya que permiten al laboratorio comprobar periódicamente su competencia en el análisis de un determinado analito (o analitos) en un determinado tipo de matriz<sup>48,49</sup>. La competencia de cada uno de los laboratorios se evalúa generalmente a través de su coeficiente  $z_i$  (*z-score*) que se calcula con la siguiente expresión:

$$z_i = \frac{x_i - \bar{x}}{s} \quad (1.13)$$

donde  $x_i$  es la concentración encontrada por el laboratorio participante,  $\bar{x}$  es la concentración asignada a la muestra y  $s$  es su desviación estándar. Para obtener un valor de *z-score* fiable es necesario tener una buena estimación de la concentración de la muestra,  $\bar{x}$ , y de la desviación estándar,  $s$ . La mejor forma de obtener la concentración de la muestra es a partir de los resultados obtenidos por un conjunto de laboratorios expertos<sup>45</sup>. No obstante, esto suele ser muy costoso y lo que se hace es obtener la concentración de la muestra como el valor consenso obtenido en el ejercicio de aptitud. Este valor consenso puede calcularse como el valor medio



obtenido después de haber eliminado los laboratorios discrepantes o bien a partir de una media robusta. La desviación estándar,  $s$ , puede obtenerse como la desviación estándar de los resultados obtenidos en el ejercicio de aptitud después de haber eliminado los laboratorios discrepantes. Sin embargo, es mejor obtener este valor de  $s$  como la reproducibilidad del método obtenida en un estudio colaborativo<sup>48,49</sup>, como la desviación estándar obtenida a partir de la expresión de Horwitz<sup>50</sup> o bien a partir de estimadores robustos de la dispersión de los resultados (por ejemplo, el estimador  $M$  de Huber<sup>48</sup>).

La competencia del laboratorio se evalúa a través del valor absoluto de su  $z$ -score:

- Si  $|z_i| < 2$ : la competencia del laboratorio es satisfactoria
- Si  $2 < |z_i| < 3$ : la competencia del laboratorio es cuestionable
- Si  $|z_i| > 3$ : la competencia del laboratorio es no satisfactoria.

Los  $z$ -scores pueden combinarse entre sí en el caso de que el laboratorio necesite un solo índice de calidad que comprenda diferentes análisis (diferentes analitos y/o diferentes matrices). Sin embargo, el *Analytical Methods Committee* no recomienda el uso de los  $z$ -scores combinados ya que se pueden enmascarar  $z$ -scores no satisfactorios<sup>48</sup>. Normalmente se suele utilizar el  $SSZ$  (“*sum of squared scores*”) ya que tienen la ventaja de que sigue una distribución  $\chi^2$  con  $m$  grados de libertad (donde  $m$  es el número de  $z$ -scores combinados<sup>45</sup>). El  $SSZ$  se calcula como la suma de la suma de los  $z$ -scores al cuadrado:

$$SSZ = \sum_i z_i^2 \quad (1.14)$$

La competencia del laboratorio es satisfactoria si su valor de  $SSZ$  es menor que el valor de  $\chi^2$  tabulado para  $m$  grados de libertad y  $\alpha=0.0455$ <sup>45</sup>. Se utiliza el valor de  $\alpha=0.0455$  porque es el que corresponde al valor tabulado de dos colas de  $z=2$  utilizado para evaluar si la competencia de un laboratorio es satisfactoria al analizar un solo analito.

## **1.8. Precisión de un procedimiento analítico: reproducibilidad, precisión intermedia y repetibilidad**

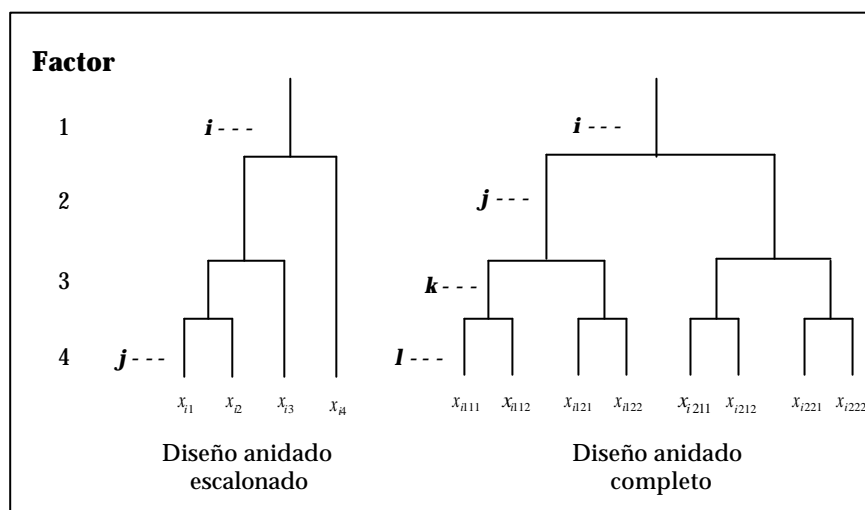
La precisión, según la norma ISO 3534<sup>11</sup>, es el grado de concordancia entre ensayos independientes obtenidos bajo unas condiciones estipuladas. Estas condiciones dependen de los factores que se varíen entre cada uno de los ensayos. Por ejemplo, algunos de los factores que se pueden variar son: el laboratorio, el analista, el equipo, la calibración del equipo, los reactivos y el día en que se hace el ensayo. Dependiendo de los factores que se varíen, pueden obtenerse tres tipos de precisión: la reproducibilidad, la precisión intermedia y la repetibilidad<sup>51</sup>.

Las dos medidas de precisión extremas son la reproducibilidad y la repetibilidad. La reproducibilidad proporciona la mayor variabilidad de los resultados ya que se obtiene cuando se varían todos los posibles factores (incluido el laboratorio) que puedan afectar a un resultado. Por otro lado, la repetibilidad proporciona la menor variabilidad que puede haber entre los resultados ya que los ensayos se obtienen en intervalos cortos de tiempo sin variar ningún factor (mismo analista, equipo, reactivos, etc.). La repetibilidad y la reproducibilidad de un método pueden calcularse a partir de ejercicios colaborativos siguiendo la metodología propuesta por la norma ISO 5725-2<sup>23</sup>. La precisión intermedia se obtiene cuando dentro de un laboratorio se varían uno o más factores entre cada uno de los ensayos<sup>51</sup>. Hay diferentes tipos de precisión intermedia dependiendo de los factores que se varíen. Por ejemplo, si los ensayos se hacen en días diferentes se obtiene la precisión intermedia de días diferentes "*time-different intermediate precision*" y si se hacen variando el analista y el día se obtiene la precisión intermedia de días y analistas diferentes "*[time+operator]-different intermediate precision*". En el caso de que se varíen todos los factores que puedan afectar a los resultados en un laboratorio (día, analista, instrumento, calibrado, etc. ) se obtiene la precisión intermedia de días, analistas, instrumentos, calibrados, etc. diferentes "*[time+operator+instrument+calibration+...]-different intermediate precision*". A este tipo de precisión intermedia también se le llama precisión intermedia de series diferentes "*run-different intermediate precision*" ya que se obtiene cuando se varían en las diferentes series todos aquellos factores que puedan afectar a la variabilidad de los resultados. Este es el tipo de precisión más importante porque proporciona una estimación de la

dispersión máxima que pueden tener los resultados obtenidos en un laboratorio para un determinado procedimiento analítico.

### 1.8.1. Cálculo de la precisión intermedia de series diferentes. Diseños anidados

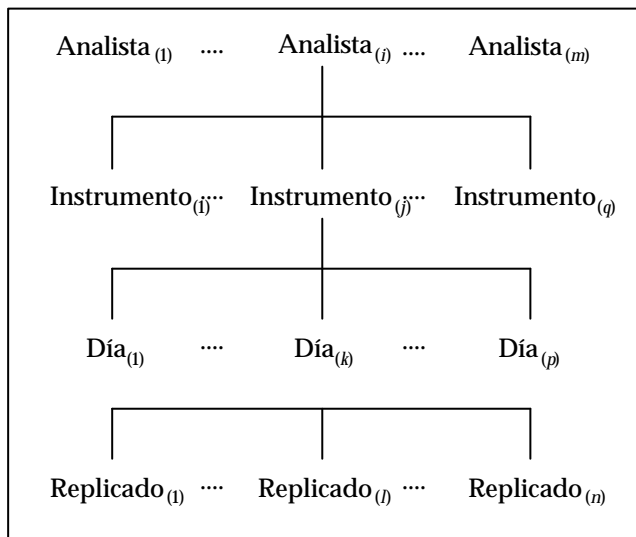
Una de las formas de calcular la precisión intermedia de un procedimiento analítico es analizando una muestra  $n$  veces variando entre cada ensayo todos aquellos factores que puedan afectar a los resultados. Normalmente, se recomienda que  $n$  sea al menos 15<sup>51</sup>. Esta aproximación es correcta cuando se quiere obtener la precisión intermedia de días diferentes. Sin embargo, si se quiere obtener la precisión intermedia debida a variar varios factores, es mejor utilizar los diseños anidados<sup>51</sup>. La ventaja de estos diseños es que proporcionan una buena estimación de la precisión intermedia ya que los factores se varían de forma ordenada. Además, otra de las ventajas es que con estos diseños se puede saber cuáles son los factores que influyen más en la variabilidad de los resultados ya que proporcionan la precisión intermedia asociada a cada uno de los factores variados en el diseño anidado.



**Figura 1.4.** Comparación del número de experimentos necesario en los diseños anidados escalonados y completos para estudiar cuatro factores de variabilidad.

Hay dos tipos de diseños anidados: los diseños anidados completos “*fully-nested designs*” y los diseños anidados escalonados “*staggered-nested designs*”<sup>51</sup>. Como se puede ver en la Fig. 1.4 la ventaja de los diseños anidados escalonados es que se necesita hacer menos ensayos que en los anidados completos. Ahora bien, la desventaja de estos diseños es que son más complejos y que las estimaciones de precisión intermedia obtenidas tienen más incertidumbre debido al menor número de ensayos. Nos centraremos en los diseños anidados completos ya que son los que se han aplicado en esta tesis.

Los factores de variabilidad estudiados en un diseño anidado deben colocarse de forma que los factores que estén más afectados por errores sistemáticos estén más arriba en el diseño y los menos afectados por errores sistemáticos más abajo. Normalmente, los factores que suelen variarse para obtener la precisión intermedia de series diferentes son el analista, el instrumento, el día y el replicado. La Fig. 1.5 muestra cómo se ordenarían estos factores en el diseño anidado.



**Figura 1.5.** Disposición de los factores analista, instrumento, día y replicado en un diseño anidado.

Siguiendo el diseño de la Fig. 1.5, una muestra homogénea y estable debería ser analizada por  $m$  analistas. A su vez, cada analista debería analizar dicha muestra en  $q$  instrumentos diferentes. En cada uno de esos instrumentos, se debería

analizar la muestra en  $p$  días diferentes y, en cada uno de esos días, deberían hacerse  $n$  replicados de la muestra. Para evitar infraestimar el efecto del día sobre la variabilidad de los resultados, las medidas obtenidas con los diferentes instrumentos deberían realizarse en diferentes días (es decir, no puede analizarse la muestra en un mismo día con dos instrumentos)<sup>30</sup>.

**Tabla 1.1.** Tabla ANOVA de un diseño completo anidado de 4 factores (g.l. son los grados de libertad).  $x_{ijkl}$  es el resultado obtenido por el analista  $i$  con el instrumento  $j$  en el día  $k$  y en el replicado  $l$ .  $\bar{x}_{ijk}$  es el valor medio de todos los replicados  $x_{ijkl}$  obtenidos por el analista  $i$  con el instrumento  $j$  y en el día  $k$ .  $\bar{x}_{ij}$  es el valor medio de los resultados  $\bar{x}_{ijk}$  obtenidos por el analista  $i$  con el instrumento  $j$ .  $\bar{x}_i$  es el valor medio de los resultados  $\bar{x}_{ij}$  obtenidos por el analista  $i$  en los distintos instrumentos. Finalmente,  $\bar{x}$  corresponde al valor medio de los resultados  $\bar{x}_i$  obtenidos por los distintos analistas.

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Analista	$MS_A = \frac{q \cdot p \cdot n \cdot \sum_{i=1}^m (\bar{x}_i - \bar{x})^2}{m - 1}$	$m - 1$	$q \cdot p \cdot n \cdot \sigma_A^2 + p \cdot n \cdot \sigma_{In}^2 + n \cdot \sigma_D^2 + \sigma_r^2$
Instrumento	$MS_{In} = \frac{p \cdot n \cdot \sum_{j=1}^q \sum_{i=1}^m (\bar{x}_{ij} - \bar{x}_i)^2}{m \cdot (q - 1)}$	$m \cdot (q - 1)$	$p \cdot n \cdot \sigma_{In}^2 + n \cdot \sigma_D^2 + \sigma_r^2$
Día	$MS_D = \frac{n \cdot \sum_{k=1}^p \sum_{j=1}^q \sum_{i=1}^m (\bar{x}_{ijk} - \bar{x}_{ij})^2}{m \cdot q \cdot (p - 1)}$	$m \cdot q \cdot (p - 1)$	$n \cdot \sigma_D^2 + \sigma_r^2$
Replicado	$MS_E = \frac{\sum_{i=1}^m \sum_{j=1}^q \sum_{k=1}^p \sum_{l=1}^n (x_{ijkl} - \bar{x}_{ijk})^2}{m \cdot q \cdot p \cdot (n - 1)}$	$m \cdot q \cdot p \cdot (n - 1)$	$\sigma_r^2$

Cada uno de los resultados obtenidos siguiendo este diseño anidado puede verse como la suma de 6 componentes:

$$y = \mu + \delta + A + In + D + e \tag{1.15}$$

donde  $\mu$  es el valor verdadero,  $\delta$  es el sesgo del método,  $A$  es el sesgo debido al analista,  $In$  es el sesgo debido al instrumento,  $D$  el sesgo debido al día y  $\epsilon$  es el error aleatorio.

La varianza asociada a la precisión intermedia de series diferentes,  $\sigma_{I(\text{serie})}^2$ , que, en este caso, corresponde a la varianza asociada a la precisión intermedia de analistas, instrumentos y días diferentes,  $\sigma_{I(A+In+D)}^2$ , se obtiene como la suma de:

$$\hat{\sigma}_{I(\text{serie})}^2 = \hat{\sigma}_{I(A+In+D)}^2 = \hat{\sigma}_A^2 + \hat{\sigma}_{In}^2 + \hat{\sigma}_D^2 + \hat{\sigma}_r^2 \quad (1.16)$$

donde  $\sigma_A^2$  es la varianza entre analistas,  $\sigma_{In}^2$  es la varianza entre instrumentos,  $\sigma_D^2$  es la varianza entre días y  $\sigma_r^2$  es la varianza asociada a la repetibilidad del método. Todas estas varianzas pueden obtenerse haciendo un análisis de la varianza (ANOVA)<sup>52</sup> a los resultados obtenidos a partir del diseño anidado propuesto en la Fig. 1.5. La Tabla 1.1 muestra el ANOVA asociado a un diseño anidado completo de cuatro factores. La Tabla 1.2 muestra los distintos tipos de precisión (expresados como varianzas) que pueden calcularse a partir del ANOVA de la Tabla 1.1.

**Tabla 1.2.** Cálculo de las varianzas obtenidas a partir del ANOVA de los resultados obtenidos en un diseño anidado de cuatro factores.

Varianza	Expresión
Repetibilidad, $s_r^2$	$MS_E$
Entre días, $s_D^2$	$\frac{MS_D - MS_E}{n}$
Entre instrumentos, $s_{In}^2$	$\frac{MS_{In} - MS_D}{p \cdot n}$
Entre analistas, $s_A^2$	$\frac{MS_A - MS_{In}}{q \cdot p \cdot n}$
Intermedia de días diferentes, $s_{I(D)}^2$	$s_r^2 + s_D^2$
Intermedia de días e instrumentos diferentes, $s_{I(D+In)}^2$	$s_r^2 + s_D^2 + s_{In}^2$
Intermedia de días, instrumentos y analistas diferentes, $s_{I(D+In+A)}^2$	$s_r^2 + s_D^2 + s_{In}^2 + s_A^2$

Sin embargo, el número de analistas y de instrumentos que hay en un laboratorio no suele ser mayor de dos o tres. Esto hace que el número de grados de libertad

asociados a las varianzas entre analistas,  $\sigma_A^2$ , y entre instrumentos,  $\sigma_{In}^2$ , sea muy pequeño y que, por tanto, no se obtenga una buena estimación de la precisión intermedia,  $\sigma_{(serie)}^2$ , a partir de la Ec. 1.16. Para solucionar este problema, Kuttatharmakul *et al*<sup>30</sup> han propuesto que, en vez de calcular estas varianzas individualmente, se calcule de forma conjunta la varianza entre días, instrumentos y analistas,  $\sigma_{AlnD}^2$ . El inconveniente de esta aproximación es que se puede infraestimar la variabilidad de los resultados debida a los instrumentos y a los analistas ya que estos factores no se cambian en cada una de las medidas utilizadas para calcular  $\sigma_{AlnD}^2$ . No obstante, es la mejor forma de calcular la precisión intermedia en el caso de que el número de analistas e instrumentos sea pequeño. Las Tabla 1.3 y la Tabla 1.4 muestran las expresiones propuestas por Kuttatharmakul *et al*.

**Tabla 1.3.** Tabla ANOVA propuesta por Kuttatharmakul *et al*<sup>30</sup> para un diseño completo anidado de 4 factores (g.l. son los grados de libertad).

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Analista+ Instrumento+Día	$MS_{AlnD} = \frac{n \cdot \sum_{i=1}^m \sum_{j=1}^q \sum_{k=1}^p (\bar{x}_{ijk} - \bar{\bar{x}})^2}{m \cdot q \cdot p - 1}$	$m \cdot q \cdot p - 1$	$n \cdot \sigma_{AlnD}^2 + \sigma_r^2$
Día	$MS_D = \frac{n \cdot \sum_{i=1}^m \sum_{j=1}^q \sum_{k=1}^p (\bar{x}_{ijk} - \bar{x}_{ij})^2}{m \cdot q \cdot (p - 1)}$	$m \cdot q \cdot (p - 1)$	$n \cdot \sigma_D^2 + \sigma_r^2$
Replicado	$MS_E = \frac{\sum_{i=1}^m \sum_{j=1}^q \sum_{k=1}^p \sum_{l=1}^n (x_{ijkl} - \bar{x}_{ijk})^2}{m \cdot q \cdot p \cdot (n - 1)}$	$m \cdot q \cdot p \cdot (n - 1)$	$\sigma_r^2$

Además de los estudios de precisión, también hay otros procedimientos alternativos para determinar la precisión de un método. Por ejemplo, a partir de los gráficos de control<sup>36</sup> también puede obtenerse la precisión intermedia del método ya que la muestra de control se va analizando a lo largo del tiempo. Esto hace que los resultados del gráfico de control contemplen toda la variabilidad que puede afectar al método analítico a pesar de que los factores no se hayan variado de forma ordenada siguiendo un diseño anidado. Asimismo, en la verificación de la trazabilidad también se puede obtener la precisión intermedia del método ya que

la muestra de referencia debería analizarse variando todos los factores que afecten a los resultados.

**Tabla 1.4.** Cálculo de las varianzas obtenidas a partir del ANOVA de la Tabla 1.3.

Varianza	Expresión
Repetibilidad, $s_r^2$	$MS_E$
Entre días, $s_D^2$	$\frac{MS_D - MS_E}{n}$
Entre analistas, días e instrumentos, $s_{AInD}^2$	$\frac{MS_{AInD} - MS_D}{n}$
Intermedia de días diferentes, $s_{I(D)}^2$	$s_r^2 + s_D^2$
Intermedia de días, instrumentos y analistas diferentes, $s_{I(D+In+A)}^2$	$s_r^2 + s_{AInD}^2$

## 1.9. Incertidumbre

Hoy en día es cada vez más importante que los resultados analíticos vayan acompañados de su incertidumbre. Así lo establece la nueva norma EN ISO/IEC 17025<sup>1</sup> que sustituye a las normas EN 45001<sup>53</sup> y a la ISO/IEC Guide 25:1990<sup>54</sup>. En esta norma no sólo se subraya la necesidad de estimar la incertidumbre de ensayo sino que también deben calcularse las incertidumbres asociadas a las calibraciones internas (es decir, las calibraciones que realiza el propio laboratorio).

La guía ISO 3534-1<sup>11</sup>, define incertidumbre como “una estimación unida al resultado de un ensayo que caracteriza el intervalo de valores dentro de los cuales se afirma que está el valor verdadero”. Esta definición tiene poca aplicación práctica ya que el “valor verdadero” no puede conocerse. Esto ha hecho que el VIM<sup>3</sup> evite el término “valor verdadero” en su nueva definición y defina la incertidumbre como “un parámetro, asociado al resultado de una medida, que



caracteriza el intervalo de valores que puede ser razonablemente atribuido al mensurando”. Ahora bien, esta nueva definición evita el término “valor verdadero” pero traslada la dificultad a la interpretación del término “mensurando”. Según el VIM el mensurando es “la magnitud sujeta a medida”<sup>3</sup>. Por tanto, en análisis químico un mensurando se refiere al analito o a la propiedad que estamos determinando. Sin embargo, el mensurando debe interpretarse correctamente para considerar todas las fuentes de incertidumbre. Por ejemplo, si queremos determinar la concentración de zinc en un mineral, habrá que incluir más o menos términos de incertidumbre en función de cómo se defina el mensurando. Es decir, ¿se quiere determinar el contenido de zinc que hay en todo el mineral o bien en la submuestra que ha llegado al laboratorio? En el primer caso habría que incluir la incertidumbre del muestreo mientras que en el segundo no se tendría que incluir<sup>55</sup>.

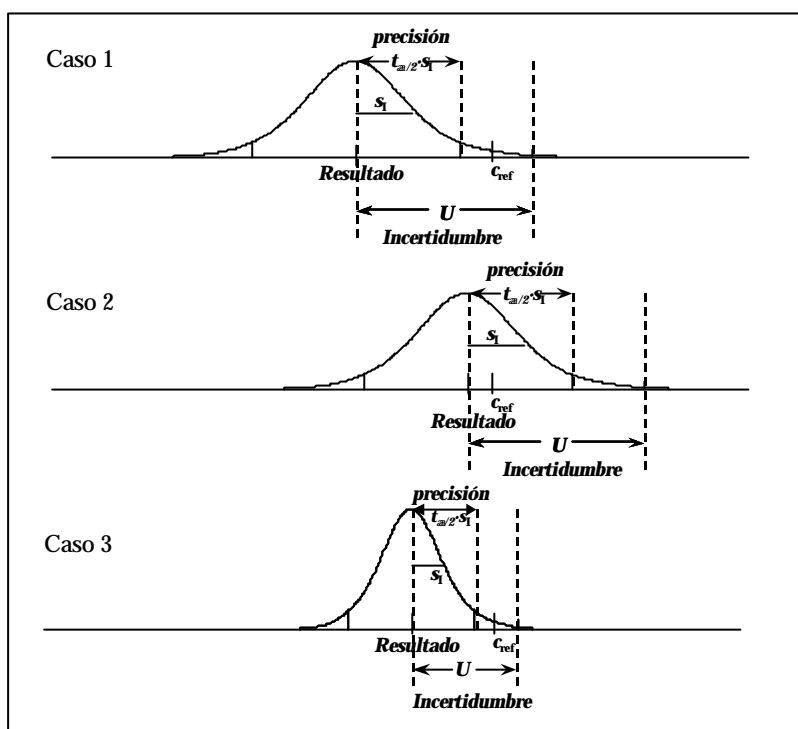
El concepto de incertidumbre refleja, pues, duda acerca de la exactitud del resultado obtenido una vez que se han evaluado todas las posibles fuentes de error y que se han aplicado las correcciones oportunas. Es decir, la incertidumbre proporciona una idea de la calidad del resultado ya que indica cuánto puede alejarse un resultado del valor considerado verdadero. Por tanto, los resultados siempre deben ir acompañados de su incertidumbre para que se puedan tomar decisiones basadas en dichos resultados<sup>56,57</sup>.

## **1.9.1. Relación de la incertidumbre con otros parámetros metrológicos**

### **1.9.1.1. Error e incertidumbre**

El VIM<sup>3</sup> define el error como “la diferencia entre el resultado obtenido y el valor verdadero del mensurando”. La incertidumbre y el error están relacionados entre sí ya que la incertidumbre debe considerar todas las posibles fuentes de error del proceso de medida. De todas formas, hay importantes diferencias entre ambos conceptos<sup>58</sup>. Por ejemplo, puede darse el caso de que un resultado tenga un error despreciable ya que, por casualidad, este resultado puede estar muy próximo al

valor verdadero. Ahora bien, la incertidumbre de este resultado puede ser muy elevada simplemente porque el analista está inseguro del resultado que ha obtenido debido al gran número de fuentes de error que puede tener el método analítico.



**Figura 1.6.** Resultados obtenidos al analizar un CRM. En los casos 1 y 2 se utiliza el mismo método analítico mientras que en el caso 3 se utiliza un método más preciso.  $c_{ref}$  es la concentración certificada del CRM y  $s_1$  la desviación estándar asociada a la precisión intermedia de series diferentes.

Por otro lado, el error cometido al analizar varias veces una muestra con un método analítico no es siempre el mismo ya que los errores aleatorios hacen que el error cometido en cada uno de los análisis sea diferente. Sin embargo, la incertidumbre de todos los resultados obtenidos al analizar esa muestra es siempre la misma ya que se utiliza el mismo método analítico. Por tanto, si la incertidumbre se ha calculado para un método analítico y un tipo de muestra determinado, todos los resultados obtenidos para todas las muestras de ese tipo que se analicen con ese método tendrán la misma incertidumbre pero no tienen por qué tener el mismo error asociado.

La diferencia entre error e incertidumbre se muestra en la Fig. 1.6. En esta figura se muestran los resultados de analizar un material de referencia certificado (CRM) con un valor de referencia,  $c_{ref}$ . En los casos 1 y 2, el CRM se ha analizado con el mismo método analítico mientras que en el caso 3 se ha utilizado otro método más preciso para analizar el CRM. Se observa que el error cometido en el caso 1 es mucho mayor que el cometido en el caso 2 pero que la incertidumbre asociada a analizar el CRM en ambos casos es la misma porque se ha utilizado el mismo método analítico.

### **1.9.1.2. Incertidumbre, veracidad y trazabilidad**

La veracidad de un resultado se define como el grado de concordancia entre el valor medio obtenido a partir de una serie de resultados de ensayo y un valor de referencia aceptado<sup>11</sup>. La incertidumbre y la veracidad están muy relacionadas entre sí ya que, si no se ha verificado la veracidad de un resultado, no se puede garantizar que se hayan corregido todos los posibles errores sistemáticos del resultado y, por tanto, es imposible asegurar que el intervalo *Valor estimado*  $\pm$  *Incertidumbre* contenga al valor considerado verdadero con una determinada probabilidad.

Por tanto, al expresar un resultado analítico como *Valor estimado*  $\pm$  *Incertidumbre*, el analista debería verificar que el *Valor estimado* no tiene un error sistemático. Además, si la veracidad del resultado se ha verificado utilizando un estándar nacional o internacional (CRM, método de referencia, etc.), también se verifica la trazabilidad del resultado frente al estándar utilizado. En este caso, la incertidumbre y la trazabilidad también están relacionadas entre sí.

### **1.9.1.3. Incertidumbre y precisión**

La precisión intermedia de series diferentes,  $s_{(serie)}$ , tiene en cuenta la variabilidad de los resultados debida a las condiciones en que se hace el análisis (día, replicado,

analista, calibrado,...). Sin embargo, la incertidumbre no solamente incluye esta variabilidad sino que también debe incluir el error asociado a la estimación de errores sistemáticos<sup>59-61</sup>. Por tanto, la incertidumbre debe incluir un término asociado a la precisión intermedia del procedimiento y otro término asociado a verificar si el método tiene o no un error sistemático. Esto hace que la incertidumbre siempre sea mayor que la variabilidad de los resultados debida a la precisión intermedia.

Además, como la incertidumbre incluye un término asociado a verificar la veracidad del método, es decir, a comprobar la ausencia de sesgo, el intervalo *Valor estimado ± Incertidumbre* incluye al valor aceptado de referencia con una cierta probabilidad. Ahora bien, el intervalo de confianza asociado a la precisión del procedimiento,  $\text{Valor estimado} \pm t_{\alpha/2} \cdot s_1$ , no tiene por qué incluir al valor de referencia con una probabilidad  $1-\alpha$ .

La Fig. 1.6 muestra las diferencias entre precisión e incertidumbre. Se observa que en todos los casos el intervalo de confianza asociado a la precisión intermedia es menor que el asociado a la incertidumbre de los resultados. Esto es debido a que la incertidumbre también incluye el término debido a verificar la ausencia de sesgo y, en el caso en que sea necesario, otros términos asociados a tratamientos previos o al muestreo. Además, también se observa que el intervalo asociado a la precisión sólo incluye al valor de referencia en uno de los casos mientras que el intervalo asociado a la incertidumbre lo incluye en los tres casos. Finalmente, el caso 3 muestra que, normalmente, cuanto más preciso es el método, menor es la incertidumbre de los resultados.

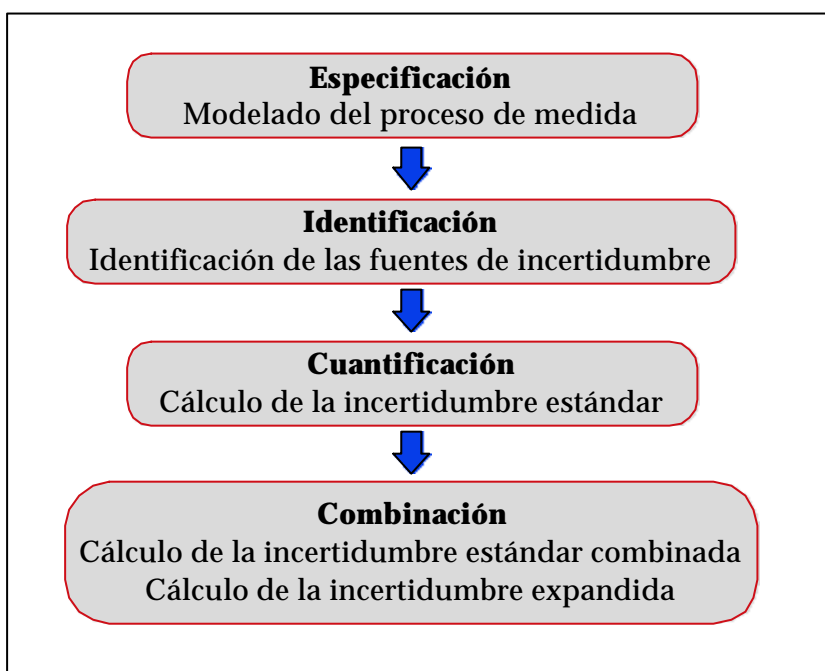
## 1.10. Cálculo de la incertidumbre. Aproximaciones existentes

Ya se ha visto la importancia de calcular la incertidumbre de los resultados. No obstante, el cálculo de la incertidumbre no es sencillo debido al elevado número de fuentes de error presentes en un procedimiento analítico. Esto ha hecho que se hayan propuesto varias aproximaciones para calcular la incertidumbre. En 1993 la ISO desarrolló una aproximación para calcular la incertidumbre en medidas

físicas<sup>59</sup>. Esta aproximación está basada en identificar y cuantificar cada uno de los componentes de incertidumbre del proceso de medida. Dos años después, Eurachem adaptó esta aproximación a medidas químicas<sup>60</sup>. Sin embargo, la adaptación que hizo Eurachem fue muy criticada y debatida<sup>62,63</sup> ya que la adaptación se hizo sin considerar las diferencias que hay entre los procesos de medida químicos y los procesos de medida físicos. Eurachem proponía calcular la incertidumbre tal y como se hace en las medidas físicas, es decir, identificando y cuantificando cada uno de los componentes de incertidumbre presentes en el proceso de medida químico. Sin embargo, esta metodología es muy costosa sino inviable en muchos de los procedimientos analíticos. Esto hizo que se fueran imponiendo otras metodologías basadas en calcular la incertidumbre más globalmente, es decir, agrupando términos de incertidumbre siempre que fuera posible. La primera de estas “aproximaciones globales” la propuso el *Analytical Methods Committee* (AMC) en 1995<sup>55</sup>. Esta aproximación, también conocida como “top-down”, está basada en utilizar la información de ejercicios colaborativos. Otras aproximaciones globales se han ido proponiendo desde entonces. Por ejemplo, en 1997 el “*Nordic Committee on Food Analysis*” (NMKL) adaptó la aproximación “top-down” a un laboratorio individual<sup>64</sup>. En 1998 y 1999 se fueron proponiendo otras aproximaciones globales basadas en calcular la incertidumbre utilizando la información obtenida en la validación de los métodos analíticos<sup>65-71</sup>. Finalmente, el debate y las críticas generadas hicieron que, en 1999, Eurachem contemplase en un documento preliminar<sup>72</sup> la posibilidad de calcular la incertidumbre utilizando información de la validación y que, un año después, Eurachem incluyera esa posibilidad en la segunda edición de su guía para el cálculo de la incertidumbre<sup>73</sup>. Esta nueva filosofía del cálculo global de la incertidumbre se está imponiendo cada vez más<sup>74-81</sup>. Asimismo, el objetivo de esta tesis ha sido calcular la incertidumbre globalmente pero aprovechando la información generada durante la verificación de la trazabilidad y en otras etapas de la validación del método.

### 1.10.1. Aproximación ISO

Comúnmente, este método recibe el nombre de “*bottom-up*” debido a que divide el proceso de medida químico en sus partes fundamentales, cada una de las partes fundamentales se subdivide a la vez en contribuciones más pequeñas, etc. Posteriormente, se calcula la incertidumbre de cada una de las partes y se combinan para obtener la incertidumbre global del proceso de medida químico. El procedimiento para cuantificar la incertidumbre total está dividido en cuatro etapas y se muestra en la Fig. 1.7:



**Figura 1.7.** Etapas seguidas por la ISO para calcular la incertidumbre.

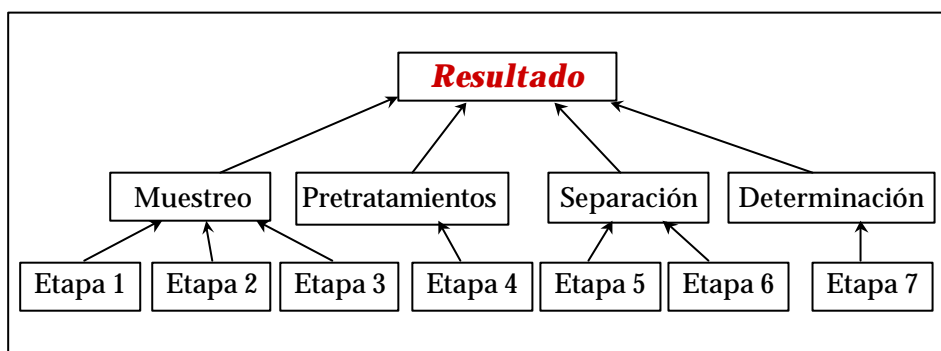
#### **Especificación**

En esta etapa debe modelarse el proceso de medida. Es decir, debe establecerse cuál es la relación que hay entre el resultado analítico y los parámetros de los que depende. Por ejemplo, en la valoración de ácido clorhídrico con hidróxido sódico esta relación vendría dada por:

$$N_{\text{HCl}} = \frac{N_{\text{NaOH}} \cdot V_{\text{NaOH}}}{V_{\text{HCl}}} = \frac{m_{\text{KHP}} \cdot P_{\text{KHP}} \cdot V_{\text{NaOH}}}{PM_{\text{KHP}} \cdot V_{\text{NaOH(KHP)}} \cdot V_{\text{HCl}}} \quad (1.17)$$

donde  $N_{\text{NaOH}}$  representa la normalidad de la solución valorante de hidróxido sódico,  $V_{\text{NaOH}}$  el volumen gastado de solución valorante y  $V_{\text{HCl}}$  el volumen de muestra valorada. En la Ec. 1.17 se ha supuesto que el hidróxido sódico se ha estandarizado frente a una solución de ftalato ácido de potasio patrón primario. Por tanto,  $N_{\text{NaOH}}$  depende a su vez del peso de ftalato ácido de potasio,  $m_{\text{KHP}}$ , de su pureza,  $P_{\text{KHP}}$ , de su peso molecular,  $PM_{\text{KHP}}$ , y del volumen de la solución de NaOH gastado durante la valoración del ftalato ácido de potasio,  $V_{\text{NaOH(KHP)}}$ .

En el caso de procedimientos analíticos más complejos, no es posible establecer una relación donde estén presentes todos los parámetros de los que depende el resultado analítico. Por tanto, lo que debe hacerse es dividir el procedimiento analítico en sus diferentes etapas tal y como se muestra en la Fig. 1.8:



**Figura 1.8.** División del procedimiento analítico en cada una de sus etapas.

### **Identificación**

Una vez modelado el proceso de medida, deben identificarse todas las fuentes de incertidumbre independientemente de la importancia que pueda tener cada una de ellas en la incertidumbre final de los resultados. Sin pretender ser exhaustivos, algunas de las fuentes de incertidumbre están asociadas a la heterogeneidad de la muestra, a la calibración de los instrumentos, a la pureza de los reactivos, a las condiciones ambientales y a los errores de los analistas. En el ejemplo anterior de la

valoración, la incertidumbre del volumen de NaOH gastado está asociada a la calibración de la bureta, a la temperatura y al error del analista.

### **Cuantificación**

En esta etapa deben cuantificarse todas las fuentes de incertidumbre identificadas en la etapa anterior. Estas fuentes de incertidumbre pueden cuantificarse de tres formas diferentes<sup>60</sup>:

- a) Experimentalmente en el laboratorio. Por ejemplo haciendo replicados en el laboratorio.
- b) Usando información obtenida en trabajos experimentales previos.
- c) Usando otro tipo de información disponible (tolerancia del material volumétrico, certificados de calibración, etc.) o bien la experiencia del analista.

La cuantificación de la incertidumbre a partir del análisis estadístico de los resultados da lugar a incertidumbres de “tipo A”<sup>59,60</sup>. Estos tipos de incertidumbre suelen estar asociados a la variabilidad de los resultados debida a los errores aleatorios. Por ejemplo, la repetibilidad, la precisión intermedia y la reproducibilidad son incertidumbres de “tipo A”.

Sin embargo, hay veces que no es posible calcular un componente de incertidumbre de forma experimental. En este caso, la incertidumbre debe evaluarse a partir de trabajos anteriores o bien a partir del criterio del analista. Las incertidumbres obtenidas de esta forma son de tipo B. Estas incertidumbres se obtienen también a partir de la información proporcionada por los proveedores. Algunos ejemplos son la tolerancia del material volumétrico y los certificados de calibración.

En otros casos la única forma de estimar la incertidumbre está basada en la opinión del analista. Por ejemplo, la incertidumbre debida a las diferencias de composición de la matriz del material de referencia y de la muestra analizada puede estimarse a partir de los conocimientos que tiene el analista sobre la muestra y el procedimiento analítico. No obstante, estas estimaciones tienen el inconveniente de que no son completamente objetivas.



*Cálculo de la incertidumbre estándar.* Antes de combinar las incertidumbres individuales para obtener la incertidumbre total, es necesario expresar todas las incertidumbres como “incertidumbres estándar”. En el caso de las incertidumbres de “tipo A”, la desviación estándar de los resultados equivale a la “incertidumbre estándar”.

Generalmente las incertidumbres de “tipo B” suelen estar expresadas como un intervalo de confianza. En este caso, es necesario convertir estas incertidumbres a “incertidumbres estándar” y, para ello, es necesario conocer la distribución en que se basan estas incertidumbres así como el nivel de significación del intervalo de confianza.

Por ejemplo, en los certificados de calibración o en los materiales de referencia, el proveedor suele proporcionar el valor de  $k$  con el que ha calculado la incertidumbre (normalmente  $k=2$ ). La incertidumbre estándar se calcula en este caso dividiendo el valor de incertidumbre proporcionada por el proveedor por el valor de  $k$ .

Ahora bien, muchas veces no se dispone de la información sobre la distribución como es el caso de las tolerancias del material volumétrico. En estos casos se asume que es igual de probable que el valor esté en cualquier lugar dentro de los límites del intervalo fijado por la tolerancia. Esto corresponde a una distribución rectangular y, por tanto, la incertidumbre estándar se obtiene dividiendo el valor de la tolerancia por  $\sqrt{3}$ <sup>59,60,82</sup>. Si es más probable que el valor esté en el centro del intervalo, se suele asumir una distribución triangular. En este caso, la incertidumbre estándar se obtiene dividiendo el valor de la tolerancia por  $\sqrt{6}$ .

### **Combinación**

Los componentes de incertidumbre individuales deben combinarse siguiendo la ley de propagación de errores<sup>59</sup>. De esta forma, se obtiene la incertidumbre total estándar,  $u(c)$ :

$$u(c) = \sqrt{\sum_{i=1}^n \left( \frac{\partial c}{\partial x_i} \cdot u(x_i) \right)^2 + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial c}{\partial x_i} \cdot \frac{\partial c}{\partial x_j} \cdot \text{cov}(x_{ij})} \quad (1.18)$$

donde  $c$  es el resultado obtenido con el procedimiento analítico,  $x_i$  es el parámetro  $i$  del que depende el resultado,  $u(x_i)$  es la incertidumbre estándar de  $x_i$  y  $\text{cov}(x_{ij})$  es la covarianza que hay entre los parámetros  $x_i$  y  $x_j$ .

Finalmente, el cálculo de la incertidumbre expandida,  $U(c)$ , proporciona un intervalo de confianza donde se encuentra el valor verdadero con una determinada probabilidad. Esta incertidumbre se obtiene multiplicando la incertidumbre estándar por un factor de cobertura  $k$ :

$$U(c) = k \cdot u(c) \quad (1.19)$$

El factor de cobertura  $k$  depende de la probabilidad con la que queremos que se encuentre el valor verdadero dentro del intervalo  $c \pm U(c)$  así como de la distribución y de los grados de libertad asociados a  $u(c)$ . Normalmente, se utiliza el valor de  $k=2$ . Este valor asume una distribución normal y una probabilidad aproximada del 95% de contener el valor verdadero. Sin embargo, en el caso de que la incertidumbre del resultado esté basada en observaciones estadísticas con pocos grados de libertad, es más correcto utilizar el valor de  $t$  tabulado de dos colas para el nivel de significación escogido y los grados de libertad efectivos,  $v_{\text{eff}}$ , calculados con la aproximación de Welch-Satterthwaite<sup>38-40</sup>:

$$v_{\text{eff}} = \frac{u(c)^4}{\sum_{i=1}^n \frac{u(x_i)^4}{v_i}} \quad (1.20)$$

donde  $v_i$  es el número de grados de libertad asociados a  $u(x_i)$ .

La ISO recomienda calcular la incertidumbre expandida utilizando la aproximación de Satterthwaite aún cuando se combinen incertidumbres de “Tipo B”. Sin embargo, la aproximación de Welch-Satterthwaite se propuso inicialmente

para combinar únicamente varianzas de distribuciones normales<sup>38-40</sup>. Recientemente, Hall *et al* han comprobado que esta aproximación también proporciona buenos resultados cuando se combinan incertidumbres de “Tipo B”<sup>83</sup>.

### 1.10.2. Aproximación propuesta por el *Analytical Methods Committee*

Esta aproximación está basada en calcular la incertidumbre utilizando la información generada en los ejercicios colaborativos<sup>55</sup>. A esta aproximación también se la conoce como “top-down” debido a que cuantifica la incertidumbre desde un “nivel superior” ya que considera al laboratorio como un miembro más de un conjunto de laboratorios. Desde este nivel, los errores cometidos por un laboratorio individual se convierten en errores aleatorios y la incertidumbre puede calcularse de forma global sin tener que identificar y cuantificar cada una de las fuentes de error del procedimiento analítico.

La Figura 1.9 muestra que la aproximación “top-down” agrupa todas las posibles fuentes de error de un resultado en sólo 4 términos: el sesgo del método,  $\delta_{\text{met}}$ , el sesgo del laboratorio,  $\delta_{\text{lab}}$ , el sesgo de la serie,  $\delta_{\text{serie}}$ , es decir, el sesgo debido a las condiciones en que se analiza la muestra (analista, día, reactivos, etc. ) y, por último, el error aleatorio,  $\varepsilon$ . Por tanto, un resultado analítico,  $c$ , obtenido al analizar una muestra homogénea puede expresarse como:

$$c = \mu + \delta_{\text{met}} + \delta_{\text{lab}} + \delta_{\text{serie}} + \varepsilon \quad (1.21)$$

donde  $\mu$  es el valor verdadero. Como puede verse en la Fig. 1.9, el sesgo del laboratorio y el de la serie se ven como errores aleatorios con una varianza asociada ya que los resultados del ejercicio colaborativo se obtienen en distintos laboratorios y en distintas series. Sin embargo, el sesgo del método se ve como un error sistemático ya que todos los laboratorios analizan la muestra con el mismo método analítico.

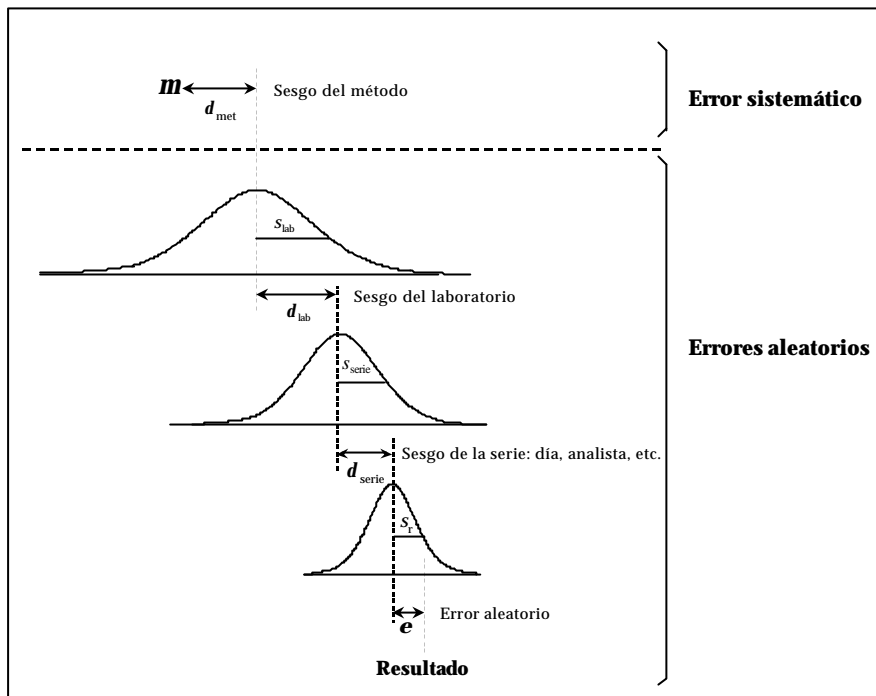


Figura 1.9. Tipos de errores asociados a un resultado según la aproximación “top-down”.

La incertidumbre de un resultado futuro se calcula aplicando la ley de propagación de errores a la Ec. 1.21:

$$U(c) = k \cdot \sqrt{u(\delta_{met})^2 + u_{lab+serie+rep}^2} \quad (1.22)$$

donde  $u(\delta_{met})$  es la incertidumbre del sesgo del método,  $u_{lab+serie+rep}$  considera la incertidumbre del laboratorio, la serie y los errores aleatorios y  $k$  es el factor de cobertura. El AMC recomienda utilizar un valor de  $k=2$  de forma que haya aproximadamente una probabilidad del 95% de que el valor verdadero esté dentro del intervalo  $c \pm U(c)$ . A continuación se explica cómo calcular los dos términos de incertidumbre de la Ec. 1.22.

**Incertidumbre del sesgo del método**

La incertidumbre del sesgo del método sólo puede calcularse en el caso de que se disponga del valor de referencia,  $x_{\text{ref}}$ , de la muestra analizada en el ejercicio colaborativo. En este caso, el sesgo del método se calcula como:

$$\ddot{a}_{\text{met}} = |x_{\text{cons}} - x_{\text{ref}}| \quad (1.23)$$

donde  $x_{\text{cons}}$  es el valor de consenso obtenido en el ejercicio colaborativo. La incertidumbre de dicho sesgo,  $u(\ddot{a}_{\text{met}})$ , se calcula como:

$$u(\ddot{a}_{\text{met}}) = \sqrt{u(x_{\text{cons}})^2 + u(x_{\text{ref}})^2} \quad (1.24)$$

donde  $u(x_{\text{cons}})$  y  $u(x_{\text{ref}})$  son, respectivamente, la incertidumbre estándar del valor de consenso y del valor de referencia. Una vez calculado el sesgo del método y su incertidumbre, debería comprobarse que el método no tiene un sesgo significativo:

$$\ddot{a}_{\text{met}} \leq 2 \cdot u(\ddot{a}_{\text{met}}) \quad (1.25)$$

En el caso de que el sesgo fuera significativo, el método analítico debería revisarse antes de ser aplicado por los laboratorios para analizar muestras de rutina. En algunos métodos analíticos como, por ejemplo, la determinación del índice de octano en gasolina, el sesgo del método es cero ya que el analito viene definido por el propio método.

**Incertidumbre del laboratorio, la serie y los errores aleatorios**

La incertidumbre del laboratorio, la serie y los errores aleatorios,  $u_{\text{lab+serie+rep}}$ , se obtiene a partir de los resultados obtenidos en el propio ejercicio colaborativo. El caso más usual<sup>21-23,25</sup> es que cada laboratorio analice dos veces la muestra en condiciones de repetibilidad. A partir del ANOVA de los resultados de los laboratorios puede estimarse la varianza asociada a la repetibilidad del método,  $s_r^2$ , y la varianza entre laboratorios y entre series,  $s_{\text{lab+serie}}^2$ . La incertidumbre  $u_{\text{lab+serie+rep}}$  se calcularía como:

$$u_{\text{lab+serie+rep}} = \sqrt{S_{\text{lab+serie}}^2 + S_r^2} \quad (1.26)$$

donde el primer término considera de forma conjunta la incertidumbre del laboratorio y de la serie y el segundo término considera la incertidumbre de los errores aleatorios.

Otra posibilidad, es que los laboratorios analicen la muestra una sola vez. En este caso, la varianza de los resultados del ejercicio colaborativo proporciona la varianza asociada a la reproducibilidad del método,  $s_R^2$ , y la incertidumbre  $u_{\text{lab+serie+rep}}$  se calcularía como:

$$u_{\text{lab+serie+rep}} = S_R \quad (1.27)$$

### **Cálculo de la reproducibilidad con la ecuación de Horwitz**

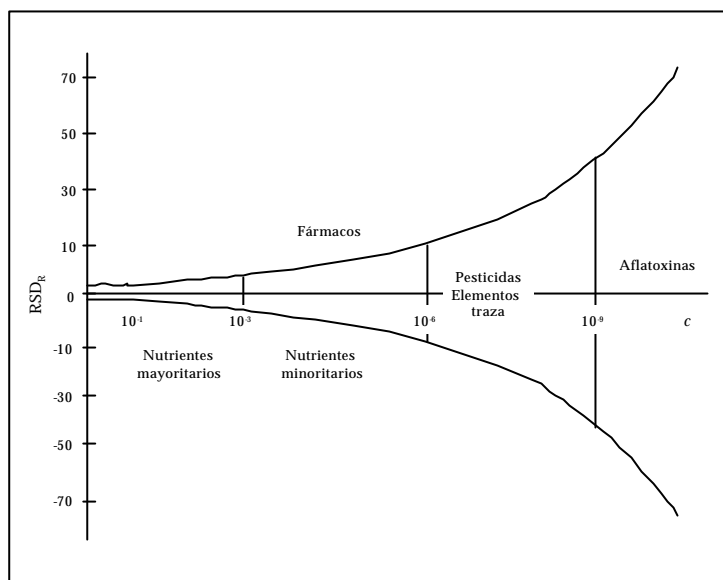
Muchas veces no se dispone de información de ejercicios colaborativos para un tipo de muestra y método analítico. En este caso se puede calcular la reproducibilidad de la Ec. 1.27 utilizando la ecuación de Horwitz<sup>84-89</sup>. Esta ecuación modela la reproducibilidad de un método en función de la concentración:

$$\%RSD_R = 2^{1-0.5 \cdot \log c} \quad (1.28)$$

donde  $\%RSD_R$  corresponde a la reproducibilidad del método expresada como un tanto por ciento de desviación estándar relativa y  $c$  corresponde a la concentración de la muestra en tanto por uno.

Esta expresión la dedujo Horwitz empíricamente examinando los resultados generados en un gran número de ejercicios colaborativos organizados por la AOAC<sup>84</sup>. Posteriormente Horwitz *et al* examinaron la aplicabilidad de esta expresión en el caso de ejercicios colaborativos donde se analizaron drogas en un intervalo de concentraciones del 0.1 al 100%<sup>85,86</sup> y en el caso de ejercicios colaborativos de metales traza con un intervalo de porcentajes de entre  $10^{-1}$  y  $10^{-8}$ %<sup>87,88</sup>. La Fig. 1.10 muestra que, según la ecuación de Horwitz, la

reproducibilidad estándar relativa va aumentando a medida que disminuye la concentración.



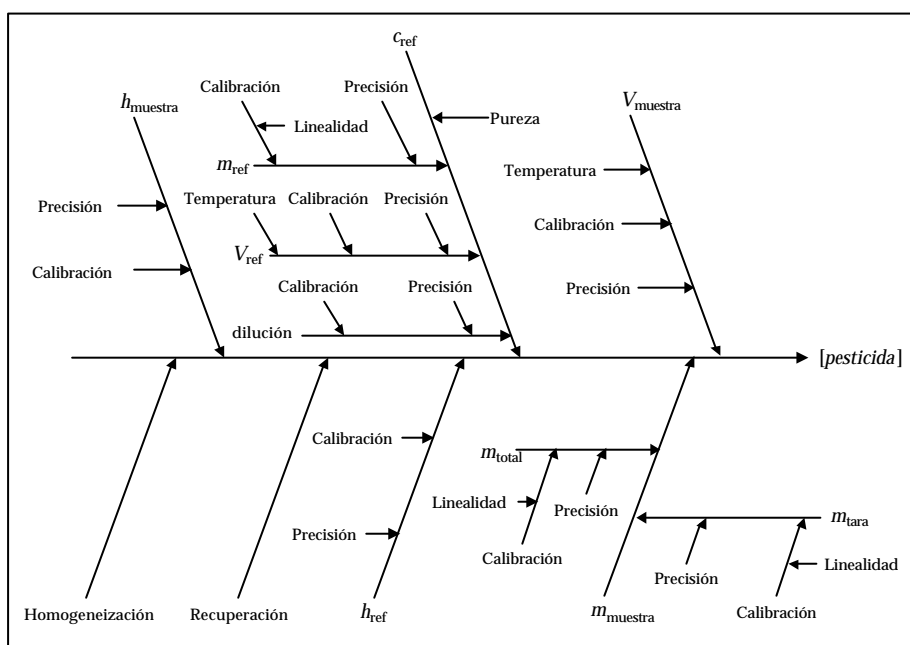
**Figura 1.10.** Variación de la reproducibilidad estándar relativa,  $RSD_R$ , en función de la concentración,  $c$ , expresada como tanto por uno.

### 1.10.3. Discusión crítica de las aproximaciones existentes para calcular la incertidumbre

La aproximación ISO tiene la ventaja de que, como se han tenido que identificar y cuantificar todas las fuentes de incertidumbre del procedimiento analítico, se puede disminuir la incertidumbre de los resultados mejorando aquellas partes del procedimiento que contribuyan más a la incertidumbre final del resultado<sup>90</sup>. Además, la incertidumbre de algunas etapas del procedimiento analítico como, por ejemplo, pesar o diluir la muestra, puede aprovecharse para calcular la incertidumbre de otros procedimientos que también tengan dichas etapas.

Sin embargo, el haber adaptado la metodología propuesta por la ISO a las medidas químicas hace que esta aproximación presente algunas desventajas,<sup>62,63</sup> debido a

que los procesos de medida químicos tienen más etapas que los procesos de medida físicos y a que es más difícil identificar y cuantificar todas las posibles fuentes de error presentes en un proceso de medida químico que en uno físico. Esto hace que sea muy costoso calcular la incertidumbre de medidas químicas siguiendo la metodología propuesta por la ISO. Así lo muestra el ejemplo de la valoración ácido-base propuesto en la guía de Eurachem<sup>60</sup> y modificado por Pueyo *et al* en 1996<sup>92</sup>. Del mismo modo, la aplicación de esta aproximación a la técnica de espectroscopia de masas por dilución isotópica también puede llegar a ser desalentadora<sup>93</sup>. Asimismo, la aplicación de esta metodología en métodos no primarios también es muy difícil<sup>94-98</sup> y no siempre posible ya que muchas veces no puede asegurarse que alguna fuente de error importante se haya pasado por alto (por ejemplo, una interferencia de la matriz imprevista)<sup>91</sup>.

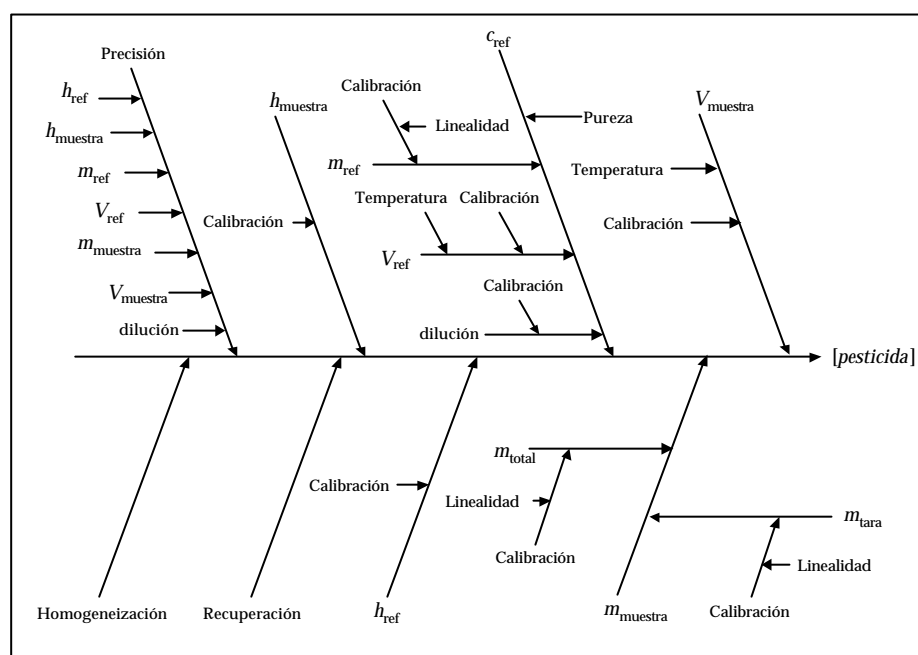


**Figura 1.11.** Diagrama causa y efecto de la determinación de pesticidas en pan.  $h_{muestra}$  y  $h_{ref}$  son, respectivamente, las alturas de los picos cromatográficos obtenidos al analizar la muestra y el patrón.

Las desventajas de la aproximación ISO han hecho que Eurachem propusiera en un documento preliminar<sup>72</sup> la posibilidad de calcular la incertidumbre agrupando



factores y aprovechando información generada en el proceso de aseguramiento de la calidad de los resultados analíticos. Finalmente, esta metodología la incluyó Eurachem en su segunda edición de la guía Eurachem para el cálculo de la incertidumbre<sup>73</sup>. En concreto, se propone utilizar la metodología propuesta por Ellison *et al*<sup>42,71,74,75</sup> que está basada en utilizar los diagramas de causa y efecto<sup>99</sup> para identificar de forma estructurada todas las fuentes de incertidumbre. La Fig. 1.11 muestra el diagrama de causa y efecto que ha propuesto Eurachem<sup>73</sup> para la determinación de pesticidas organofosforados en pan por cromatografía de gases.



**Figura 1.12.** Diagrama de causa y efecto de la determinación de pesticidas en pan después de la “etapa de reconciliación”.

Como se puede ver en la Fig. 1.11 los diagramas de causa y efecto se estructuran en ramas principales que provienen de los términos que aparecen en la etapa de especificación. A su vez, cada rama principal tiene otras ramas secundarias que consideran todos aquellos factores que puedan afectar al término considerado en la rama principal. Posteriormente, en la “etapa de reconciliación” debe identificarse si hay alguna fuente de incertidumbre que se haya considerado más de una vez. Además, en el caso de que se use información existente del método (ejercicios

interlaboratorio, estudios de recuperación, de precisión, etc. ) deben suprimirse del diagrama aquellos factores que se hayan tenido en cuenta en la información utilizada. La Fig. 1.12 muestra el diagrama de causa y efecto después de la “etapa de reconciliación” en el caso de la determinación de pesticidas en pan. En este ejemplo, Eurachem utilizó información de la precisión del método analítico. Como se puede ver, las fuentes de incertidumbre asociadas a la precisión (de la balanza, del cromatógrafo, del volumen de muestra y de la dilución) quedan agrupadas en una nueva rama del diagrama que corresponde a la precisión del método analítico.

A diferencia de la aproximación ISO, la ventaja de la aproximación propuesta por el AMC es que calcula la incertidumbre globalmente, y, por tanto, no tiene que identificar y cuantificar cada una de las fuentes de error de un procedimiento analítico. No obstante, esta aproximación es difícilmente aplicable ya que muchas veces no se dispone de información sobre ejercicios colaborativos. En este caso, se debe recurrir a calcular la incertidumbre utilizando la expresión de Horwitz<sup>84</sup>. La desventaja de esta aproximación es que la incertidumbre calculada puede tener poco que ver con la de un laboratorio individual<sup>58</sup>. En primer lugar, los laboratorios que participan en ejercicios colaborativos no se suelen escoger al azar sino que se escogen aquellos laboratorios que han demostrado previamente su competencia. Además la precisión del método se calcula después de haber eliminado los laboratorios discrepantes con el test de Cochran y el test de Grubbs<sup>21-23</sup>. Por tanto, la precisión obtenida en estos estudios suele ser mejor que la que puede tener un laboratorio individual. Por otro lado, hay dos causas que pueden tener el efecto contrario, es decir, que proporcionen una precisión peor que la de un laboratorio individual. La primera causa es que la muestra analizada en el ejercicio no sea totalmente homogénea. La segunda causa es que los ejercicios colaborativos suelen hacerse con métodos nuevos o recientemente modificados. Por tanto, esto puede hacer que laboratorios que utilicen rutinariamente dicho método tengan mejores valores de precisión que la obtenida a partir del ejercicio colaborativo.

Las desventajas de la aproximación “top-down” han hecho que se hayan ido proponiendo otras aproximaciones globales basadas en utilizar la información de un laboratorio individual. En 1997 el NMKL<sup>64</sup> adaptó la aproximación “top-down”

a un solo laboratorio y calculó la incertidumbre como la precisión intermedia del procedimiento. Sin embargo, esta aproximación tiene el inconveniente que no considera la incertidumbre asociada a posibles errores sistemáticos. Actualmente se han ido proponiendo otras aproximaciones globales para calcular la incertidumbre. Alder *et al* proponen calcular la incertidumbre globalmente utilizando la información obtenida en los ejercicios de aptitud<sup>79</sup>. Para ello es necesario que el laboratorio haya demostrado previamente su competencia (a partir de su *z-score*, de estudios de recuperación y del análisis de materiales de referencia) y que la precisión del laboratorio sea menor que la reproducibilidad obtenida en el ensayo de aptitud<sup>79</sup>. Ahora bien, la mayoría de las aproximaciones globales que se están proponiendo actualmente son las que están basadas en utilizar la información obtenida en la validación del método<sup>42,65,71-75,78,80,81</sup>. En el siguiente capítulo se profundizará en una aproximación propuesta para calcular la incertidumbre globalmente aprovechando, sobre todo, la información generada durante la verificación de la trazabilidad y también la obtenida durante otras etapas de la validación del método (estudios de precisión y robustez).

## 1.11. Referencias

1. UNE-EN ISO/IEC 17025, *Requisitos generales relativos a la competencia de los laboratorios de ensayo y calibración*, AENOR, Madrid (2000)
2. EURACHEM, *The fitness for Purpose of Analytical Methods. A Laboratory Guide to method Validation and Related Topics*, EURACHEM Secretariat, Teddington, Middlesex (1998)
3. BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *International Vocabulary of basic and general terms in Metrology, VIM*, ISO, Geneva (1993)
4. The Internacional System of Units (SI), 7th Ed. Bureau Internacional des Poids et Mesures, Sèvres (1998)  
<http://www.bipm.fr/pdf/si-brochure.pdf>
5. M. Valcárcel, A. Ríos, *Analyst*, **120** (1995) 2291-2297
6. M. Valcárcel, A. Ríos, *Trends in Analytical Chemistry*, **18** (1999) 570-576
7. A. Ríos, M. Valcárcel, *Accreditation and Quality Assurance*, **3** (1998) 14-19
8. M. Thompson, *Analyst*, **121** (1996) 285-288
9. M. Thompson, *Analyst*, **122** (1997) 1201-1205
10. B. King, *Analyst*, **122** (1997) 197-204
11. International Organization for Standardization, *Statistics, Vocabulary and symbols*, ISO 3534-1, ISO, Geneva (1993)

12. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-1, ISO, Geneva (1994)
13. J. Fleming, B. King, E. Prichard, M. Sargent, B. Scott, R. Walker, *Tracing the chemical links*, VAM (1994)
14. T.J. Quinn, *Metrologia*, **34** (1997) 61-65
15. B. King, *Química Analítica*, **19** (2000) 67-75
16. International Organization for Standardization, *Terms and definitions used in connection with reference materials*, ISO Guide 30, ISO, Geneva (1992)
17. X.R. Pan, *Metrologia*, **34** (1997) 61-65
18. ISO, IEC, *Proficiency testing by interlaboratory comparisons: Part 1: Development and operation of proficiency testing schemes*, Guide 43-1, ISO, Geneva (1997)
19. M. Thompson, R. Wood, *Journal of AOAC International*, **76** (1993) 926-940
20. Analytical Methods Committee, *Analyst*, **117** (1992) 97-104
21. Collaborative Study Guidelines, *Journal of AOAC International*, **78** (1995) 143A-160A
22. International Union of Pure and Applied Chemistry (IUPAC), *Pure and Applied Chemistry*, **67** (1995) 649-666
23. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-2, ISO, Geneva (1994)
24. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-4, ISO, Geneva (1994)
25. E. Hund, D.L. Massart, J. Smeyers-Verbeke, *Analytica Chimica Acta*, **423** (2000) 145-165
26. Commission of the European Communities, Community Bureau of Reference (BCR), *New reference materials, improvements of methods and measurements*, EUR 9921 EN, Brussels (1985)
27. International Organization for Standardization, *Certification of reference materials-General and statistical principles*, ISO Guide 35, ISO, Geneva (1989)
28. Laboratory of the Government Chemist, *Certified Reference Materials Catalogue*, Issue 4, LGC, Teddington (1997)
29. H. Marchandise, E. Colinet, *Fresenius Journal of Analytical Chemistry*, **316** (1983) 669-672
30. S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Analytica Chimica Acta*, **391** (1999) 203-225
31. S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Chemometrics and Intelligent Laboratory Systems*, **52** (2000) 61-73
32. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-6, ISO, Geneva (1994)
33. M. Parkany, *The use of recovery factors in trace analysis*, Royal Society of Chemistry, Cambridge (1996)
34. V.J. Barwick, S.L.R. Ellison, *Analyst*, **124** (1999) 981-990
35. *Harmonised Guidelines for the use of recovery information in analytical measurement*, Technical Report Resulting from the Symposium on Harmonisation of Internal Quality Assurance Systems for Analytical Laboratories, Orlando, USA (1996)

36. M. Thompson, R. Wood, *Pure and Applied Chemistry*, **67** (1995) 649-666
37. Analytical Methods Committee, *Analyst*, **120** (1995) 29-34
38. B. L. Welch, *Biometrika*, **2** (1937) 110-114
39. F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309-316
40. F.E. Satterthwaite, *Biometrics*, **2** (1946) 110-114
41. A.G. González, M.A. Herrador, A.G. Asuero, *Talanta*, **48** (1999) 729-736
42. V.J. Barwick, S.L.R. Ellison, *Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data*, VAM Project 3.2.1 (2000)
43. N. Draper, H. Smith, *Applied Regression Analysis*, 2<sup>nd</sup> Ed. Wiley, Nueva York (1981)
44. J. Mandel, F.J. Linnig, *Analytical Chemistry*, **29** (1957) 743-749
45. D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier (1997)
46. J. Riu, F.X. Rius, *Journal of Chemometrics*, **9** (1995) 343-362
47. J. Riu, F.X. Rius, *Analytical Chemistry*, **68** (1996) 1851-1857
48. M. Thompson, P.J. Lowthian, *Analyst*, **118** (1993) 1495-1500
49. W. Horwitz, L.R. Kamps, K.W. Boyer, *Journal of AOAC International*, **63** (1980) 1344-1354
50. Analytical Methods Committee, *Analyst*, **122** (1997) 495-497
51. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-3, ISO, Geneva (1994)
52. H. Scheffé, *Analysis of variance*, Wiley, Nueva York (1953)
53. UNE 66-501-91 (Norma Europea EN 45001), *Criterios generales para el funcionamiento de los laboratorios de ensayo*, AENOR, Madrid (1991)
54. ISO, IEC, *General requirements for the competence of calibration and testing laboratories*, ISO/IEC Guide 25, Geneva (1990)
55. Analytical Methods Committee, *Analyst*, **120** (1995) 2303-2308
56. P. de Bièvre, *Accreditation and Quality Assurance*, **2** (1997) 269
57. A. Williams, *Accreditation and Quality Assurance*, **1** (1996) 14-17
58. S.L.R. Ellison, W. Wegscheider, A. Williams, *Analytical Chemistry*, **1** (1997) 607A-613 A
59. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva (1993)
60. EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK (1995)
61. J.S. Kane, *Analyst*, **122** (1997) 1283-1288
62. W. Horwitz, R. Albert, *Analyst*, **122** (1997) 615-617
63. W. Horwitz, *Journal of AOAC International*, **81** (1998) 785-794
64. Nordic Committee on Food Analysis, *Estimation and Expression of measurement uncertainty in chemical analysis*, procedure No.5, NMKL (1997)
65. W. Hässelbarth, *Accreditation and Quality Assurance*, **3** (1998) 418-422
66. S.L.R. Ellison, A. Williams, *Accreditation and Quality Assurance*, **3** (1998) 6-10

67. S.L.R. Ellison, V.J. Barwick, *Analyst*, **123** (1998) 1387-1392
68. V.J. Barwick, S.L.R. Ellison, *Analytical Communications*, **35** (1998) 377-383
69. R.J.N.B. Silva, M.F.G.F.C. Camoes, J.S. Barros, *Analytica Chimica Acta*, **393** (1999) 167-175
70. B. Jülischer, P. Gowik, S. Uhlig, *Analyst*, **124** (1999) 537-545
71. V. Barwick, S.L.R. Ellison, B. Fairman, *Analytica Chimica Acta*, **394** (1999) 281-291
72. EURACHEM, *Quantifying uncertainty in analytical measurement*, Draft: EURACHEM Workshop, 2<sup>nd</sup> Edition, Helsinki (1999)
73. EURACHEM/CITAC, *Quantifying uncertainty in analytical measurement*, EURACHEM/CITAC Guide, 2<sup>nd</sup> Edition (2000)
74. V.J. Barwick, S.L.R. Ellison, *Accreditation and Quality Assurance*, **5** (2000) 47-53
75. V.J. Barwick, S.L.R. Ellison, M.J.K. Rafferty, S.G. Rattanjit, *Accreditation and Quality Assurance*, **5** (2000) 104-113
76. J. Moser, W. Wegscheider, C. Sperka-Gottlieb, *Fresenius Journal of Analytical Chemistry*, **370** (2001) 679-689
77. R.G. Visser, *Accreditation and Quality Assurance*, **7** (2002) 124-125
78. T.P.J. Linsinger, M. Führer, W. Kandler, R. Schuhmacher, *Analyst*, **126** (2001) 211-216
79. L. Alder, W. Korth, A.L. Patey, H.A. Schee, S. Schoenweiss, *Journal of AOAC International*, **84** (2001) 1569-1578
80. J. Kristiansen, *Clinical Chemistry and Laboratory Medicine*, **39** (2001) 920-931
81. L. Cuadros, M.E. Hernández, E. Almansa, F.J. Egea, F.J. Arrebola, J.L. Martínez, *Analytica Chimica Acta*, **454** (2002) 297-314
82. C.F. Dietrich, *Uncertainty, Calibration and Probability. The statistic of Scientific and Industrial Measurement*, 2<sup>nd</sup> Ed. Adam Hilger, Bristol, (1991)
83. B.D. Hall, R. Willink, *Metrologia*, **38** (2001) 9-15
84. W. Horwitz, *Analytical Chemistry*, **54** (1982) 67A-75A
85. W. Horwitz, R.J. Albert, *Journal of AOAC International*, **67** (1984) 81-90
86. W. Horwitz, R.J. Albert, *Journal of AOAC International*, **67** (1984) 648-652
87. K.W. Boyer, W. Horwitz, R. Albert, *Analytical Chemistry*, **57** (1985) 454-459
88. W. Horwitz, R. Albert, *Fresenius Journal of Analytical Chemistry*, **351** (1995) 507-513
89. R. Albert, W. Horwitz, *Analytical Chemistry*, **69** (1997) 789-790
90. R. Brix, S.H. Hansen, V. Barwick, J. Tjornelund, *Analyst*, **127** (2002) 140-143
91. E. Hund, D.L. Massart, J. Smeyers-Verbeke, *Trends in Analytical Chemistry*, **20** (2001) 394-405
92. M. Pueyo, J. Obiols, E. Vilalta, *Analytical Communications*, **33** (1996) 205-208
93. A. Dobney, H. Klinkenberg, F. Souren, W. van Born, *Analytica Chimica Acta*, **420** (2000) 89-94
94. V.J. Barwick, *Journal of Chromatography A*, **849** (1999) 13-33
95. R. Ramachandran, *Analyst*, **124** (1999) 1099-1103
96. C.S.J. Wolff Briche, C. Harrington, T. Catterick, B. Fairman, *Analytica Chimica Acta*, **437** (2000) 1-10

97. D.B. Hibbert, J. Jiang, M.I. Mulholland, *Analytica Chimica Acta*, **443** (2001) 205-214
98. F. Cordeiro, G. Bordin, A.R. Rodríguez, J.P. Hart, *Analyst*, **126** (2001) 2178-2185
99. International Standard for Standardization, *Total Quality Management Part 2. Guidelines for Quality Improvement*, ISO 9004-4, ISO, Geneva (1993)







## **CAPÍTULO 2**

---

### ***Cálculo de la incertidumbre en métodos analíticos de rutina validados a un nivel de concentración***



## 2.1. Introducción

Después de revisar en el capítulo anterior las diversas aproximaciones para el cálculo de la incertidumbre asociada a los resultados químicos, se ha visto que la dificultad de calcular la incertidumbre siguiendo la aproximación propuesta por la ISO ha motivado que actualmente se estén imponiendo otras aproximaciones basadas en calcular la incertidumbre de una forma global, es decir, agrupando términos de incertidumbre siempre que sea posible. Estas aproximaciones están basadas en utilizar la información generada en el proceso de validación del método y durante el control interno del mismo.

En este capítulo se presenta una aproximación para calcular globalmente la incertidumbre de resultados obtenidos con métodos analíticos que se usan de forma rutinaria en el laboratorio y que se aplican en un intervalo restringido de concentraciones. Esta aproximación está basada en utilizar la información generada durante la validación del método y, especialmente, la obtenida durante la verificación de la trazabilidad de los resultados obtenidos con dicho método. La ventaja de esta aproximación es que supone poco trabajo adicional ya que, tal y como se ha visto en el capítulo anterior, no tiene sentido calcular la incertidumbre de los resultados si previamente no se ha verificado su trazabilidad.

En el apartado 2.2. de este capítulo se explica el fundamento teórico de la aproximación propuesta para el cálculo global de la incertidumbre en muestras de rutina que tienen un intervalo restringido de concentraciones. En el apartado 2.3. se presenta el artículo "*Estimating uncertainties using information from the validation process*" publicado en la revista *Analytica Chimica Acta*. En este artículo se detalla el cálculo de la incertidumbre utilizando fundamentalmente la información obtenida durante la verificación de la trazabilidad de los resultados. En el apartado 2.4. se profundiza sobre cómo calcular los diversos términos de incertidumbre parciales descritos en este artículo utilizando tanto la información de la verificación de la trazabilidad como la obtenida en los gráficos de control, en los estudios de precisión y en los de robustez. En el apartado 2.5. se compara la aproximación propuesta con la aproximación ISO y la aproximación del AMC. Finalmente, en el

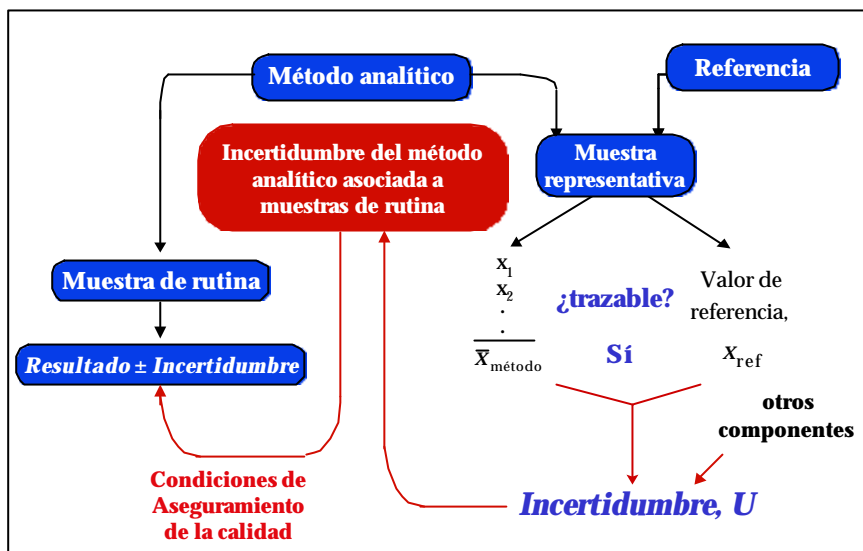
apartado 2.6. se exponen las conclusiones de este capítulo y en el apartado 2.7. la bibliografía.

La aproximación propuesta para el cálculo global de la incertidumbre también se ha presentado en el artículo “*Evaluating uncertainty in routine analysis*” publicado en la revista *Trends in Analytical Chemistry* y en el artículo “*Critical discussion on the procedures to estimate uncertainty in chemical measurements*” publicado en la revista *Química Analítica* que figuran, respectivamente, en los anexos 1 y 2 de esta Tesis. En estos artículos también se describe críticamente las aproximaciones propuestas por la ISO y por el *Analytical Methods Committee* para el cálculo de la incertidumbre. Además, en el artículo “*Critical discussion on the procedures to estimate uncertainty in chemical measurements*” también se explica el concepto de incertidumbre y su relación con otros parámetros metrológicos como son la trazabilidad, la exactitud, la veracidad y la precisión. Por último, también se sugiere la posibilidad de incluir información del aseguramiento de la calidad para calcular la incertidumbre.

## **2.2. Fundamento teórico de la aproximación para el cálculo global de la incertidumbre de muestras de rutina aprovechando información de la validación del método**

En el capítulo anterior se ha visto que no tiene sentido calcular la incertidumbre de los resultados si previamente no se ha verificado la trazabilidad de los mismos. La Fig. 2.1 muestra que durante la verificación de la trazabilidad se genera información que puede utilizarse para calcular la incertidumbre de las muestras de rutina. Estrictamente, esta información está ligada a la muestra de referencia analizada durante la verificación de la trazabilidad. No obstante, esta información puede utilizarse para calcular la incertidumbre de muestras de rutina ya que la trazabilidad debe verificarse utilizando muestras de referencia que sean lo más parecidas posibles a las muestras de rutina que se analizarán con el método analítico. Además, para que la incertidumbre obtenida sea representativa de las muestras de rutina es necesario que el laboratorio esté bajo condiciones de aseguramiento de la calidad<sup>1-3</sup>. De esta forma, el laboratorio puede asegurar que: 1)

los resultados obtenidos al analizar las muestras de rutina siguen siendo trazables; y 2) que la incertidumbre calculada sigue siendo válida.

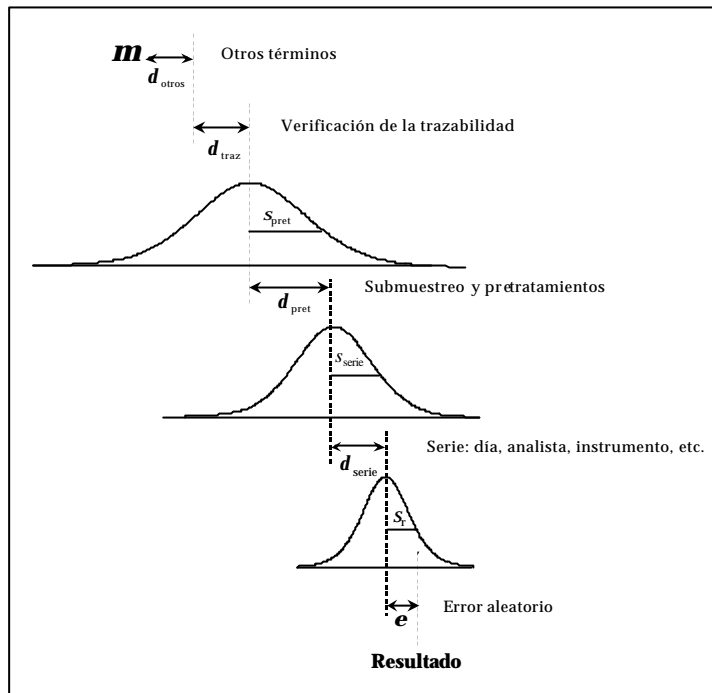


**Figura 2.1.** Cálculo de la incertidumbre de muestras de rutina utilizando la información generada al verificar la trazabilidad de los resultados.

Para calcular globalmente la incertidumbre de muestras de rutina debe expresarse el resultado analítico como la suma de<sup>4,5</sup>:

$$\text{Resultado} = \bar{x} + \ddot{a}_{\text{otros}} + \ddot{a}_{\text{traz}} + \ddot{a}_{\text{pret}} + \ddot{a}_{\text{serie}} + \ddot{a} \quad (2.1)$$

donde  $\varepsilon$  es el error aleatorio dentro de una misma serie,  $\delta_{\text{serie}}$  es el sesgo de la serie (es decir, el sesgo debido a las condiciones en que se analiza la muestra: día, analista, instrumento, etc.),  $\delta_{\text{pret}}$  es el sesgo debido a la heterogeneidad de la muestra y a los tratamientos previos,  $\delta_{\text{traz}}$  es el sesgo del método y se calcula durante la verificación de la trazabilidad de los resultados generados por el método,  $\delta_{\text{otros}}$  es el sesgo asociado a otros términos que no se han considerado previamente y  $\mu$  es el valor verdadero. Esta forma de expresar el resultado se la conoce como la “escalera de los errores”<sup>6</sup> ya que estos errores tienen una estructura jerarquizada. La Fig. 2.2 muestra cómo afectan estos errores a un resultado analítico.



**Figura 2.2.** Errores asociados a un resultado según la aproximación basada en calcular la incertidumbre utilizando la información obtenida al verificar la trazabilidad en un método analítico a un nivel de concentración.

Para poder obtener la concentración verdadera de la muestra, se debería poder corregir el resultado por cada uno de los errores cometidos. Sin embargo, el sesgo de la serie y el error aleatorio no pueden conocerse. Únicamente se puede conocer una estimación de la varianza entre series,  $s_{serie}^2$ , y de la varianza asociada a la repetibilidad del método,  $s_r^2$ . Estas varianzas están asociadas, respectivamente, a la distribución de valores del error aleatorio y del sesgo de la serie. Del mismo modo, el sesgo debido a la heterogeneidad y a los pretratamientos tampoco puede conocerse. Únicamente, se puede conocer una estimación de la varianza,  $s_{pret}^2$ , asociada a la distribución de valores de los sesgos debidos a los pretratamientos y a la heterogeneidad de las muestras. Asimismo, no se puede conocer el sesgo verdadero del método. Sólo se puede conocer una estimación del sesgo del método

y su incertidumbre asociada,  $u_{\text{traz}}$ . Por último, el sesgo de otros términos tampoco puede conocerse. En este caso, se suele tener una estimación asociada a este sesgo.

Como sólo se puede obtener una estimación del valor verdadero del resultado, es necesario incluir como términos de incertidumbre los errores asociados al error aleatorio y a los sesgos de la serie, del método, de los pretratamientos y de otros términos. De esta forma, se puede asegurar que el intervalo  $\text{Valor estimado} \pm \text{Incertidumbre}$  contiene al valor verdadero con una cierta probabilidad ya que la incertidumbre incluye todas las posibles fuentes de error. Para calcular la incertidumbre debe aplicarse la ley de propagación de errores a la Ec. 2.1. La Fig. 2.3 muestra la expresión de la incertidumbre estándar que se obtiene aplicando la ley de propagación de errores. Los términos de la ecuación de la Fig. 2.3 se han reordenado de forma diferente que en la Ec. 2.1 por conveniencia ya que en primer lugar se explican aquellos términos que pueden calcularse utilizando la información de la verificación de la trazabilidad, dejando para el final los que se evalúan por otros medios.

$$u = \sqrt{u(\mathbf{e})^2 + u(\ddot{a}_{\text{serie}})^2 + u(\ddot{a}_{\text{traz}})^2 + u(\ddot{a}_{\text{pret}})^2 + u(\ddot{a}_{\text{otros}})^2}$$

**Figura 2.3.** Expresión para el cálculo de la incertidumbre de las muestras de rutina.

La incertidumbre del procedimiento,  $u_{\text{proc}}$ , considera la variabilidad experimental debida a la variación dentro de una serie de análisis,  $\varepsilon$ , y a la variación entre series,  $\delta_{\text{serie}}$ , y se calcula a partir de las varianzas de la repetibilidad,  $s_r^2$ , y de la variabilidad entre series,  $s_{\text{serie}}^2$ . Estas varianzas pueden obtenerse durante la verificación de la trazabilidad de los resultados, en estudios de precisión o bien utilizando la información de los gráficos de control. La incertidumbre de la

verificación de la trazabilidad,  $u_{\text{traz}}$ , considera la incertidumbre del sesgo del método calculado durante la verificación de la trazabilidad. Estrictamente, un método analítico puede tener un sesgo constante y un sesgo proporcional. En este caso se asume que el sesgo es el mismo en todo el intervalo de aplicación del método ya que las muestras de rutina tienen un intervalo de concentraciones restringido. La incertidumbre del submuestreo y pretratamientos,  $u_{\text{pret}}$ , considera la incertidumbre de la heterogeneidad de la muestra y de pretratamientos que se realizan sobre las muestras de rutina pero no en la muestra utilizada para verificar la trazabilidad de los resultados. Esta incertidumbre se puede obtener analizando una muestra de rutina variando todos los factores que afecten al submuestreo y los pretratamientos. Por último,  $u_{\text{otros}}$  corresponde a la incertidumbre de otros términos asociados a diferencias entre la muestra de rutina y la muestra utilizada para verificar la trazabilidad, a niveles más altos de trazabilidad y a factores de variabilidad que no se han tenido en cuenta en el cálculo de la precisión intermedia. Estos términos de incertidumbre pueden calcularse utilizando la información generada en los estudios de robustez o bien utilizando muestras adicionales.

La incertidumbre expandida,  $U$ , se obtiene multiplicando la incertidumbre estándar,  $u$ , por el factor de cobertura  $k=2$ . De esta forma, se puede asegurar que el intervalo *Valor estimado*  $\pm U$  contiene la concentración verdadera,  $\mu$ , con aproximadamente un 95% de probabilidad. Sin embargo, si los distintos términos de incertidumbre tienen pocos grados de libertad asociados, es más correcto utilizar el valor de  $t$  tabulado de dos colas para el nivel de significación escogido y los grados de libertad efectivos calculados con la aproximación de Welch-Satterthwaite<sup>7,8</sup>.

En los siguientes apartados se explica cómo utilizar la información generada en la validación del método para calcular globalmente la incertidumbre con la expresión propuesta en la Fig. 2.3.



### **2.3. Estimating uncertainties of analytical results using information from the validation process (*Analytica Chimica Acta*, **391** (1999) 173-185)**

Alicia Maroto, Jordi Riu, Ricard Boqué and F. Xavier Rius

Department of Analytical and Organic Chemistry. Rovira i Virgili University.  
Pl. Imperial Tàrraco, 1. 43005-Tarragona. Spain.

#### **ABSTRACT**

A new approach for calculating uncertainties of analytical results based on the information from the validation process is proposed. This approach complements the existing approaches proposed to date and can be applied to any validated analytical method. The precision estimates generated during the process of assessment of the accuracy take into account the uncertainties of preprocessing steps and analytical measurement steps as long as the different factors that influence these steps are representatively varied in the whole validation process. Since the accuracy of an analytical method should be always assessed before applying it to future working samples, little extra work needs to be done to estimate the final uncertainty. Other sources of uncertainty not previously considered (e.g. uncertainty associated to sampling, to differences between the test-sample and the working sample, etc.) are subsequently included and mathematically combined with the uncertainty arising from the assessment of the accuracy to provide the overall uncertainty. These ideas are illustrated with a case study of the determination of copper in contaminated land.

## 1. INTRODUCTION

These days analysts are increasingly being urged to ensure that their results are both reliable and comparable. The estimation of uncertainty is therefore one of the main focuses of interest in the field of measurement chemistry [1]. The main reason for this is that the analytical result is a piece of information which is used for various purposes. Important decisions are often based on these results, so both the traceability and the degree of confidence (i.e. the uncertainty), given as parameters that define the quality of these results, must be demonstrated.

The International Vocabulary of basic and general terms in Metrology, VIM, states in its latest definition that uncertainty is “a parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand” [2]. This definition must be interpreted if a link between traceability and uncertainty is to be established. If traceability has not been assessed against a high quality reference in metrological terms, the correctness of all the possible systematic effects of the measurements cannot be guaranteed, and it is therefore impossible to ensure that the confidence interval comprises the true value. So, if the quantitative analytical result is expressed as Value 1  $\pm$  Uncertainty, every analyst should verify that Value 1 is an unbiased value which is traceable, whenever possible, to an internationally recognised reference.

The main approaches for calculating uncertainty which have been proposed to date are the ISO approach (commonly known as “*bottom-up*”) and the Analytical Methods Committee approach (commonly known as “*top-down*”). The ISO approach was originally proposed for quantifying uncertainty in physical measurements [3] and was subsequently adapted by EURACHEM [4] for chemical measurements. It is based on identifying, quantifying and combining all the sources of uncertainty of the measurement. Although the ISO approach improves knowledge of the measurement procedure, the difficulty of applying it in chemical measurements hampers its widespread use. On the other hand, the approach proposed by the Analytical Methods Committee [5] is based on information gathered from interlaboratory exercises. In this approach the laboratory is seen

from a higher level (i.e. as a member of the population of laboratories). Consequently, systematic and random errors within individual laboratories become random errors when they are considered from this “higher level”. The “*top-down*” approach takes into account factors of the measurement that are intrinsically impossible to know; however, the lack of interlaboratory information for many measurands in a diversity of matrices and the fact that a collaborative trial carried out in the past may not be related to the present conditions in the laboratory often means that it is difficult to use.

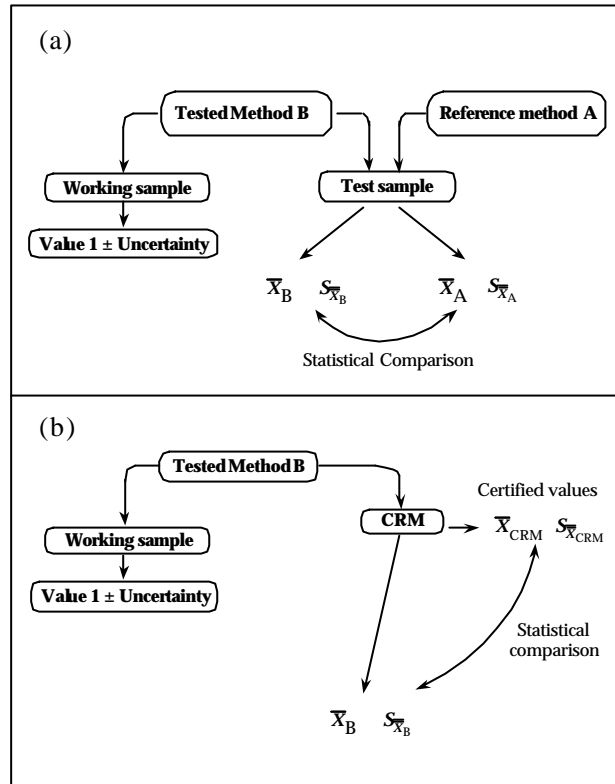
To complement the existing approaches for calculating uncertainties, we propose a new procedure that takes advantage of the precision estimates generated in assessing the accuracy (trueness and precision) of the analytical methods. Since the accuracy of all the methods should always be assessed before applying them to future working samples, this approach means that little extra work needs to be done when calculating the uncertainty of the final result. Following the error propagation procedure, we combine the uncertainty arising from the assessment of accuracy with other components not previously considered, such as the uncertainty arising from sampling or lack of homogeneity, to obtain the overall uncertainty.

## **2. ASSESSMENT OF ACCURACY**

The accuracy of the measurement methods should always be assessed before applying them to future working samples. Verifying accuracy is one of the basic steps in the process of achieving traceability. The international concept of traceability [2] and its interpretation in chemical analysis [6-10] can be applied to chemical measurement processes by comparing the measurement method statistically to an appropriate reference. The measurement method, which will be applied to future samples to be analysed in the same conditions, is therefore compared for precision and trueness to the reference results.

Fig. 1 compares the results obtained with the measurement method to two different types of references: a reference measurement method and a certified

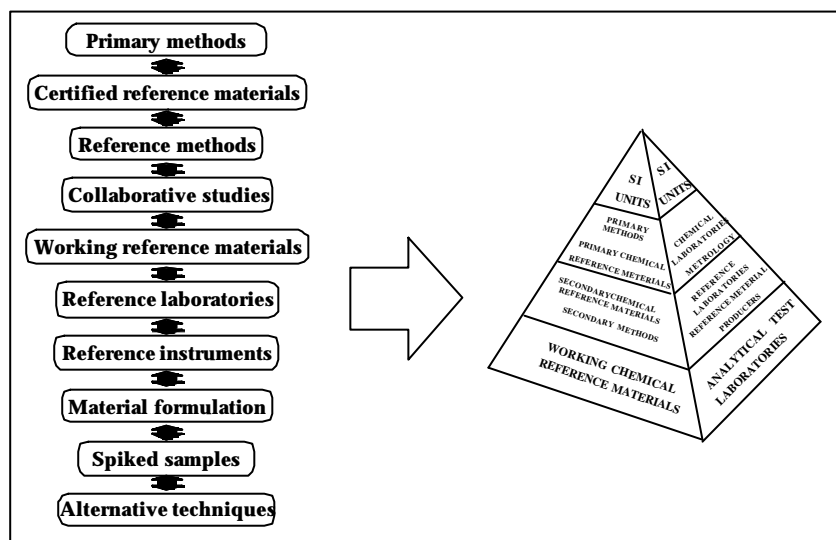
reference material (CRM). However, other references, whose metrological importance is shown in Fig. 2, could also be used.



**Figure 1.** Plan showing the assessment of the accuracy of a tested method B. a) Comparison against a reference method A. b) Comparison with a certified reference material, CRM.

Fig. 1a compares the accuracy of the tested method B against a reference method A. This is a common situation in many laboratories, where an alternative method has been developed to substitute a reference method in routine analysis. One of the most advanced guidelines for comparing two methods can be found in ISO 5725-6 [11]. However, it is not applicable in a single laboratory because it is based on interlaboratory information. Therefore, a recently developed procedure will be used [12]. This is based on validating a method by generating different precision estimates in a laboratory. If the accuracy of method B has to be assessed against reference method A in a single laboratory, the experimental design should take

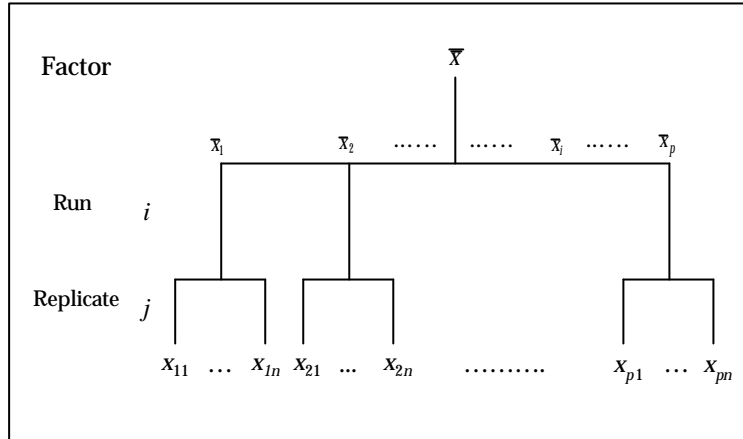
into account all the possible factors of variation. The experimenter must decide how many factors are to be varied representatively and how many analyses are to be carried out. The number of factors depends on the resources (number of analysts, instruments, time, etc.) required by the measurement process and the experimenter's knowledge of the method. Moreover, the number of analyses to be made should be chosen by considering the minimum bias and the minimum ratio of precision measures that the analyst wishes to detect and the  $\alpha$  probability of falsely concluding the presence of bias and the  $\beta$  probability of falsely concluding the absence of bias. The sample to be analysed (i.e. the test sample), should be homogeneous, stable and as representative as possible of the samples to be subsequently analysed.



**Figure 2.** Hierarchy of some relevant references in chemical analysis and their relationship to the classical metrological pyramid.

The experimental design proposed to vary the chosen factors representatively is a fully-nested design [13]. As stated in [13], the factors affected mainly by systematic effects must be in the highest ranks, while the factors affected mainly by random effects must be in the lowest ranks. For instance, if the factors studied are the repeatability and the run on which the measurement is performed, the runs will be in a higher rank than the repeatability because the repeatability is less affected by systematic errors. Moreover, in this case nesting the run-effect in the replicates is

not possible. A fully-nested design which takes into account the variation of the runs and the replicates is shown in Fig. 3.



**Figure 3.** Two-factor fully-nested design. The factors studied are the runs and the replicates.  $p$  is the number of runs on which the measurement is carried out and  $n$  is the number of replicates performed every run.  $\bar{x}_i$  is the mean of the  $j$  replicate measurements performed on run  $i$ . The grand mean,  $\bar{x}$ , is calculated as the mean of the mean values obtained on the different runs.

The proposed experimental design provides precision estimates of the different factors studied. Following the proposed two-factor fully-nested design, the precision estimates generated are the repeatability standard deviation,  $s_r$ , the between-run standard deviation,  $s_{run}$ , the run-different intermediate standard deviation,  $s_{i(run)}$ , and the standard deviation of the means,  $s_{\bar{x}}$ . The typical analysis of variance (ANOVA) table and the calculation of the different variance terms is shown in Tables 1 and 2. All the important sources of variation of the measurement result (e.g. the analyst, the calibration, the day, etc.) have been varied between each run. The run-different intermediate standard deviation can therefore be considered as an estimate of the [time + analyst + calibration + ...]-different intermediate precision [13]. As a result, the run-different intermediate standard deviation includes all the important factors of variation of the measurement result. However, it must be pointed out that the estimation of the between-run standard deviation may not be as good as the one that would have been obtained if a fully-nested design had been carried out by including all the important sources of variation like

time, analyst, calibration, etc. as factors in the experimental design. However, in this latter case, the cost in terms of work increases considerably.

**Table 1.** ANOVA table for a two-factor fully-nested design.

Source	Mean Squares	Degrees of freedom	Expected mean square
run	$MS_{\text{RUN}}$	$p-1$	$n\sigma_{\text{run}}^2 + \sigma_r^2$
residual	$MS_E$	$p \cdot (n-1)$	$\sigma_r^2$
total		$p \cdot n-1$	

**Table 2.** Calculation of variances for a two-factor fully-nested design.

Variance	Expression	Degrees of freedom
repeatability variance $s_r^2$	$MS_E$	$p \cdot (n-1)$
between-run variance $s_{\text{run}}^2$	$\frac{MS_{\text{RUN}} - MS_E}{n}$	
run-different intermediate variance $s_1^2$	$s_r^2 + s_{\text{run}}^2$	
variance of the means $s_x^2$	$\frac{MS_{\text{RUN}}}{n}$	$p-1$

The precision estimates and the grand means obtained for both methods are then used to assess the accuracy of method B. It is important to note that the different runs within which the results of both methods have been obtained need not (necessarily) be the same (i.e. laboratory A obtained the results in 10 consecutive days while laboratory B did it in 20 days, because they analysed a sample every two days).

Before comparing the two methods, the consistency of the individual results should be checked by graphical analysis and/or tests for outliers [14]. Moreover, the normality of the measurement results should be always checked.

To assess the precision of method B, all the intermediate precision estimates must be compared statistically with those of the reference method by performing an  $F$ -test. For instance, when the runs and the repeatability are the factors varied, the repeatability variance and the run-different intermediate variance of both methods must be compared. Both methods should be as precise as possible because the probability of falsely concluding the trueness of method B against method A increases when the precision of both methods diminishes.

Trueness can be assessed by checking whether the difference between the grand mean for reference method A,  $\bar{x}_A$ , and the grand mean of method B,  $\bar{x}_B$ , is not statistically significant:

$$|\bar{x}_A - \bar{x}_B| \leq t_{\alpha/2} \cdot s_d \quad (1)$$

where  $s_d$  depends on the variance of the means of both methods. Therefore, an  $F$ -test is performed to compare the variances of the means of both methods,  $s_{\bar{x}_A}^2$  and  $s_{\bar{x}_B}^2$ . If the difference between these variances is not significant,  $s_d$  is given by Eq. (2a):

$$s_d = \sqrt{\left( \frac{(p_A - 1) \cdot s_{x_A}^2 + (p_B - 1) \cdot s_{x_B}^2}{p_A + p_B - 2} \right) \left( \frac{1}{p_A} + \frac{1}{p_B} \right)} \quad (2a)$$

If the difference between the variances of the means is statistically significant,  $s_d$  is given by Eq (2b):

$$s_d = \sqrt{\frac{s_{x_A}^2}{p_A} + \frac{s_{x_B}^2}{p_B}} \quad (2b)$$

If the fully-nested design of Fig. 3 is followed,  $p$  corresponds to the number of runs on which the analysis is carried out. In this case, the variance of the means,  $s_{\bar{x}}^2$  is calculated as:



$$s_{\bar{x}}^2 = \frac{\sum_{i=1}^p (\bar{x}_i - \bar{\bar{x}})^2}{p-1} \quad (3)$$

where  $\bar{x}_i$  is the mean of the replicate measurements carried out in the run  $i$ .

If the difference between the grand mean of reference method A and the grand mean of method B is not significant, the trueness of method B against method A is demonstrated, but only when the factors considered in the experimental design are varied. Therefore, method B is traceable to method A in the conditions of variation studied.

The accuracy of a method can be checked by means other than just a reference method, as Fig. 2 shows. For instance, in Fig. 1b the process of comparing method B against a CRM is shown. In this case, the CRM is analysed with method B by varying all the factors that can affect the measurement method representatively. Finally, the results obtained,  $\bar{x}_B$  and  $s_{x_B}$  are compared statistically to the certified reference values,  $\bar{x}_A$  and  $s_{x_A}$ . If the reference used in assessing accuracy has a low level of traceability, higher levels can be achieved by linking this reference to a recognised interlaboratory exercise, to a primary method, to a CRM or, in selected cases, to the SI units. This latter path is still not established in many practical instances, however. Moreover, as the number of steps in the traceability chain increases, the measurement result becomes less precise due to the addition of uncertainty terms and the propagation of errors. It is always advisable, therefore, to trace the procedure to the highest available reference. All future analyses using tested method B in conditions of routine work will be traceable to the reference used in assessing accuracy as long as the working samples are similar to the one used in the internal validation, and the quality assurance conditions are implemented in the laboratory effectively.

### 3. CALCULATION OF UNCERTAINTY

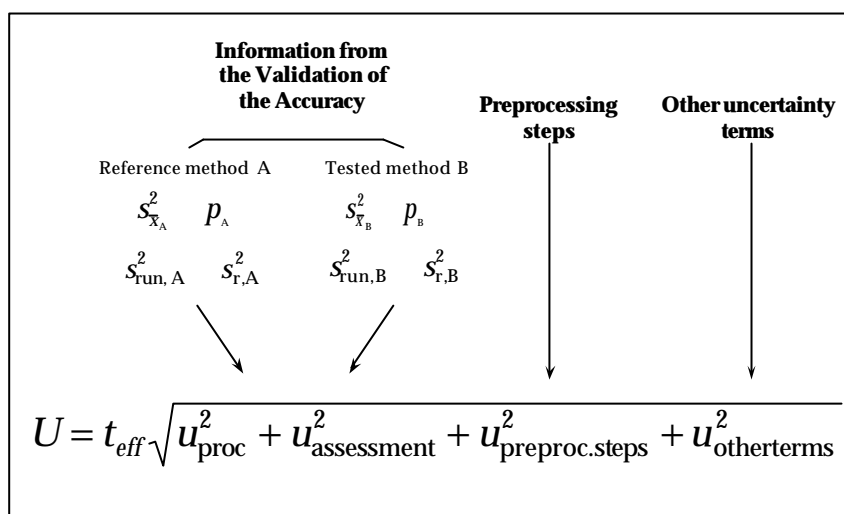
While verifying the accuracy of the measurement, the analyst generates information about different intermediate precision estimates. Since these estimates have been obtained by varying the factors that affect the measurement result representatively, they take into account the uncertainties of different steps in the analytical process. For instance, uncertainties due to the separation and preconcentration of the analyte of interest, matrix effects, environmental conditions, instruments or operators are considered as long as the factors that influence them are varied representatively. However, there may be other sources of variation which have not been considered in the assessment of accuracy. These sources may be related to global factors or local factors.

The *global factors* are inherent to the measurement process. Therefore, these sources of uncertainty must always be considered in the overall uncertainty. They may be related to uncertainties not taken into account while assessing the accuracy or to the bias of the reference (estimated by linking this reference to another one with higher levels of traceability). For instance, if the reagent's suppliers had not been representatively varied while assessing accuracy, possible future changes in the supplier would introduce a new source of uncertainty that should subsequently be considered in the overall uncertainty of all the samples to be analysed in the future. The uncertainty arising from laboratory bias or from method bias should also be taken into account when the higher level of traceability achieved corresponds to a reference method carried out in the same laboratory as the method to be tested. When the higher level is related to a CRM, the source of uncertainty that should be considered is the uncertainty arising from the bias of this reference material. Finally, all the sources of uncertainty arising from the different links of the established chain of traceability must also be considered a component of uncertainty in all the samples to be analysed in the future.

The *local factors* are characteristic of the sample to be analysed. These factors therefore vary according to the type of sample analysed. One of the main sources of uncertainty related to local factors are the differences in sampling and pretreatment procedures between the test sample and the working sample. Other

possible sources of uncertainty are the differences in the measurand value or in the matrix of the test sample and the working sample.

All the different uncertainty terms previously mentioned must be calculated or estimated and mathematically combined to obtain the overall uncertainty value of the analytical result. This process is illustrated in Fig. 4. In this scheme, the uncertainty components have been grouped into four terms:



**Figure 4.** Scheme showing the process for calculating the overall uncertainty of chemical measurement results using the information from the validation of the accuracy.

### 3.1. Uncertainty arising from the analytical procedure

The uncertainty which arises from the experimental variation when applying the analytical procedure to obtain future measurement results is associated to the different values that the run bias and the random error of method B can take on within the laboratory. If every sample is analysed in  $p_s$  different runs and, in each run  $n_s$  replicates are carried out, an estimation of the uncertainty of the grand mean,  $\bar{x}_s$ , of these measurements is given by the between-run variance, the repeatability variance and the number of runs and replicates carried out within each run.

$$u_{\text{proc}} = \sqrt{\frac{s_{\text{run}}^2}{p_s \cdot n_s} + \frac{s_r^2}{n_s}} \quad (4)$$

However, the result of a future working sample is seldom obtained as the grand mean of measurement results performed in  $p_s$  different runs and with  $n_s$  replicates in each run. Normally, the result is obtained with only one analysis or at most as the mean of three replicates. In this case, information about the between-run variance and repeatability variance cannot be obtained because of the small number of replicates and runs carried out. However, the between-run variance and repeatability variance obtained in assessing the accuracy of method B can be used as estimates of the  $s_{\text{run}}^2$  and  $s_r^2$  related to the working sample. This can be done whenever the working sample is similar to the test sample in the assessment of accuracy and the quality assurance conditions are effectively implemented in the laboratory. Therefore, the uncertainty of the procedure can be calculated with Eq. 4.

In the most common case i.e.  $n_s=1$  and  $p_s=1$ , equation 4 becomes:

$$u_{\text{proc}} = \sqrt{s_{\text{run}}^2 + s_r^2} = \sqrt{\frac{MS_{\text{RUN}} - MS_E}{n} + MS_E} = \sqrt{\frac{MS_{\text{RUN}}}{n} + \left(1 - \frac{1}{n}\right) \cdot MS_E} \quad (5)$$

and its associated number of effective degrees of freedom,  $\nu_{\text{proc}}$ , are estimated with the Welch-Satterthwaite approach [16]

$$\nu_{\text{proc}} = \frac{u_{\text{proc}}^4}{\frac{\left(\frac{MS_{\text{RUN}}}{n}\right)^2}{p-1} + \frac{\left(\left(1 - \frac{1}{n}\right)MS_E\right)^2}{p \cdot (n-1)}} \quad (6)$$

### 3.2. Uncertainty arising from the assessment of accuracy

In assessing accuracy, what is checked is whether the laboratory and/or the method bias is statistically significant. The type of bias checked depends on the metrological quality of the reference used. If the bias is not statistically significant, the method is traceable to the reference but there still remains a source of uncertainty in the estimation of this bias. If the reference used is a reference method A, the estimated bias of method B is calculated from the difference between the grand means obtained from both methods when analysing the same test sample:

$$\hat{d}_{\text{method}} = \bar{x}_B - \bar{x}_A \quad (7)$$

The standard uncertainty associated to the estimated method bias is given by  $s_d$ .

$$u_{\text{assessment}} = s_d \quad (8)$$

This standard deviation is estimated from the variance of the method to be tested and from the variance of the reference (if available) used in the assessment of accuracy. If the difference between variances is statistically not significant,  $s_d$  is calculated using Eq. 2a with  $p_A+p_B-2$  degrees of freedom. If the difference between variances is statistically significant,  $s_d$  is estimated by Eq. 2b. Since in this case  $s_d$  is a compound variance, the degrees of freedom are obtained by applying the Welch-Satterthwaite approach:

$$\mathbf{n}_{\text{assessment}} = \frac{s_d^4}{\frac{\left(\frac{s_{\bar{x}_A}^2}{p_A}\right)^2}{p_A - 1} + \frac{\left(\frac{s_{\bar{x}_B}^2}{p_B}\right)^2}{p_B - 1}} \quad (9)$$

### 3.3. Uncertainty arising from preprocessing steps

This component is necessary when the sources of uncertainty arise from subsampling and from preprocessing steps applied to future working samples that have not been considered in the process of the assessment of accuracy. Field sampling uncertainty is not considered here because the uncertainty estimated normally corresponds to the sample that has reached the laboratory once the field sampling has been performed. However, the component of uncertainty due to field sampling is usually very high and exceeds that from other steps, so it should be taken into account whenever the laboratory has the responsibility for sampling. To estimate this uncertainty is rather complex but it can be done by collaborative trials in which all the factors that can affect sampling (protocol, sampler, replicate, etc.) are varied. This subject has been extensively treated by Ramsey *et al* [15].

The uncertainty arising from subsampling i.e. from the heterogeneity of the material, has not been considered when assessing the accuracy. In this process, the test sample or reference material must be homogeneous by definition. However, future working samples are not expected to be homogeneous, and so the uncertainty arising from the heterogeneity of the sample once in the laboratory must be included in this term. The uncertainty arising from other preprocessing steps (filtration, weighing, drying, etc.) not carried out on the test sample used in assessing the accuracy must also be included in this term. The uncertainty arising from heterogeneity and/or preprocessing steps can be estimated from a sample with the same characteristics as the working samples to be analysed subsequently. The subsampling and/or the preprocessing steps must be carried out by representatively changing all the factors that can affect them (e.g. operator, material, time, etc.). Then the analytical results of each portion of the sample subsampled and/or pretreated are obtained under repeatability conditions. If only one individual measurement is taken, the contribution of the analytical measurement repeatability can be subtracted by using the information of the repeatability variance from the assessment of the accuracy,

$$s_{\text{preproc.steps}}^2 = s_{\text{I(preproc.steps)}}^2 - s_{\text{r}}^2 \quad (10)$$

where  $s_{I(\text{preproc.steps})}^2$  is the preprocessing step-different intermediate variance and  $s_{\text{preproc.steps}}^2$  is the between-preprocessing steps variance.

However, if each portion of the sample subsampled and/or pretreated is analysed at least twice under repeatability conditions, the analysis of variance (ANOVA) can be applied to separate the repeatability variance of the analytical measurement from the variance arising from the subsampling and/or the preprocessing steps,  $s_{\text{preproc.steps}}^2$ . Like in the assessment of accuracy, the experimental design adopted in this case is a fully-nested design.

Once the variance of the between-preprocessing steps has been obtained, Eq. 11 estimates the standard uncertainty due to the preprocessing steps of future working samples with the same characteristics as the sample analysed:

$$u_{\text{preproc.steps}} = s_{\text{preproc.steps}} \quad (11)$$

where the degrees of freedom depend on the number of portions of the sample subsampled and/or pretreated.

### **3.4. Uncertainty arising from other sources**

The last term in the equation in Fig. 4 contains all the other components of uncertainty that have not yet been considered. Most of these components can be considered as “type B” uncertainties [3,4] e.g. associated to differences between the working sample and the test sample, or to the bias of the reference (estimated by linking the reference to other references which have higher levels of traceability). Computation of some of these terms is not straightforward and they should therefore be evaluated with the experience and knowledge of the analyst, bibliography or any other source of reliable information in mind.

The combined standard uncertainty of this term is estimated as the square root of the sum of the variances of the different components:

$$u_{\text{other terms}} = \sqrt{\sum_i s_i^2} \quad (12)$$

Since  $u_{\text{other terms}}$  is an estimate of a compound variance, the degrees of freedom,  $\nu_{\text{other terms}}$ , associated to this standard combined uncertainty are estimated using the Welch-Satterthwaite approach:

$$\nu_{\text{other terms}} = \frac{(u_{\text{other terms}}^2)^2}{\sum_i \frac{(s_i^2)^2}{\nu_i}} \quad (13)$$

where  $\nu_i$  are the degrees of freedom associated to each individual variance term,  $s_i^2$ . In many cases, however, there is no information about the degrees of freedom of the components related to this last term. In such a case, a coverage factor,  $k$ , can be used instead of the two-sided  $t$ -tabulated value [3].

### 3.5. Estimating overall uncertainty

Once the variances of each term and their associated degrees of freedom have been estimated, the overall standard uncertainty,  $u$ , is estimated by combining these variances.

$$u = \sqrt{u_{\text{proc}}^2 + u_{\text{assessment}}^2 + u_{\text{preproc.steps}}^2 + u_{\text{other terms}}^2} \quad (14)$$

The overall expanded uncertainty is estimated by multiplying the overall standard uncertainty by the two-sided  $t_{\text{eff}}$ -tabulated value, at a given  $\alpha$  probability and  $\nu_{\text{eff}}$  degrees of freedom.



---

$$U = t_{a/2, \text{eff}} \cdot \sqrt{u_{\text{proc}}^2 + u_{\text{assessment}}^2 + u_{\text{preproc.steps}}^2 + u_{\text{other terms}}^2} \quad (15)$$

where the effective degrees of freedom,  $v_{\text{eff}}$ , are estimated using the Welch-Satterthwaite approach [16,17]. The overall expanded uncertainty is a confidence interval within which the conventional true value is considered to lie with selected probabilities of error.

Although the overall uncertainty has been estimated from precision information from test samples ( $u_{\text{proc}}$  and  $u_{\text{assessment}}$  in assessing accuracy and  $u_{\text{preproc.steps}}$  from a sample similar to the working samples to be analysed), the uncertainty calculated with Eq. 15 is a reliable estimate of the uncertainty related to future working samples. The only requirement is that the samples used to estimate precision have similar matrix and level of analyte concentration as the future working samples. The uncertainty arising from possible differences in the analyte concentration and/or in the matrix type between the test sample and the working sample should be considered and included, whenever possible, as type B uncertainties. Further analyses of working samples using the same procedure, whenever this is under statistical control, would provide similar uncertainty values. Therefore, just as traceability establishes a link between the measurement method and the results corresponding to the analysed samples, the estimated uncertainty can be associated to the measured quantities of similar samples in the future.

#### **4. A CASE STUDY: DETERMINING COPPER IN CONTAMINATED LANDS**

A tested (candidate) method B, whose accuracy against a reference method A had previously been demonstrated, is used to determine copper content in contaminated lands. Method B treats the sample with a mixture of nitric and perchloric acids, dissolves the analyte in diluted hydrochloric acid and then analyses it by ICP-AES. The reference method A is based on treating the sample with the same acids but the copper content of the final acidic solution is determined by AAS.

The main sources of uncertainty arise from the experimental variation, the estimation of method bias, the estimation of laboratory bias (estimated by linking the reference method A to a CRM) and the lack of homogeneity of the sample to be analysed.

**Table 3.** Analytical results obtained for method A and method B in the assessment of accuracy by carrying out a two-factor fully-nested design. The copper content is expressed as  $\mu\text{g}\cdot\text{g}^{-1}$ .

Method A			Method B		
Run	Mean	Replicate	Run	Mean	Replicate
1	26.15	27.2 25.1	1	29.55	30.1 29.0
2	29.15	28.3 30.0	2	31.40	32.6 30.2
3	24.75	26.5 23.0	3	29.45	28.9 30.0
4	24.65	24.1 25.2	4	27.75	29.1 26.4
5	31.20	32.1 30.3	5	26.60	27.2 26.0
6	30.25	31.5 29.0	6	30.60	31.2 30.0
7	31.80	32.6 31.0	7	31.65	32.5 30.8
8	28.90	29.5 28.3	8	31.05	32.1 30.0
9	29.75	30.2 29.3	9	27.35	28.2 26.5
10	26.15	27.3 25.0	10	32.10	33.0 31.2
$\bar{x}_A = 28.27$ $s_{\bar{x}_A}^2 = 6.98$ $s_{r,A}^2 = 2.02$ $s_{\text{run}A}^2 = 5.97$			$\bar{x}_B = 29.75$ $s_{\bar{x}_B}^2 = 3.79$ $s_{r,B}^2 = 1.59$ $s_{\text{run}B}^2 = 2.99$		

The uncertainty arising from the experimental variation caused by applying the analytical procedure to obtain future measurement results and the uncertainty from estimating the method bias are calculated from the information generated in validating (assessing the accuracy of) method B against method A. Method B was validated with a test sample which is very similar to the contaminated land

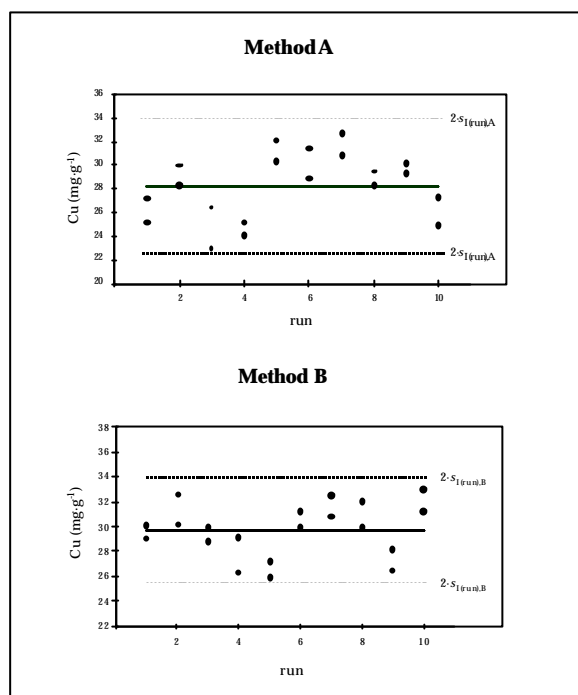
samples (i.e. the level of the measurand and the matrix are similar in both samples), so most of the sources of uncertainty arising from analytical measurement and pretreatment steps can be estimated with the information from the assessment of the accuracy. A two-factor fully-nested design was followed for both methods. The factors representatively varied were the replicate and the run. Between each run, factors such as the analyst, the calibration and the day were varied. Therefore, information about the repeatability standard deviation and the between-run standard deviation was obtained.

#### **4.1. Assessment of accuracy**

The accuracy of method B was assessed against method A with a homogenous and stable test sample. The measurements were carried out following a two-factor fully-nested design. From the requirements, two measurements were made in repeatability conditions on ten different runs. The results, together with the grand means,  $\bar{x}_A$  and  $\bar{x}_B$ , the variance of the means,  $s_{\bar{x}_A}^2$  and  $s_{\bar{x}_B}^2$ , the between-run variance,  $s_{\text{run},A}^2$  and  $s_{\text{run},B}^2$ , and the repeatability variance,  $s_{r,A}^2$  and  $s_{r,B}^2$ , for both methods, are shown in Table 3. A graphical presentation of the data from Table 3 for methods A and B are shown in Fig. 5. It can be seen that for both methods all the individual results are within the confidence interval defined by twice the run-different intermediate standard deviation of the analytical method.

##### **4.1.1. Comparison of precision**

After checking the absence of outlying results and the normality of the data, the repeatability variance and the run-different intermediate variance of each method were compared by *F*-tests [11,12]. There was no evidence that both methods have different repeatabilities and different run-different intermediate precision at 95% significance.



**Figure 5.** Graphical presentation of the analytical results obtained for method A and method B in the assessment of accuracy.

#### 4.1.2. Comparison of trueness

The variance of the means of both methods must also be compared by means of an *F*-test to find out which expression must be used for calculating *s<sub>d</sub>*.

$$F_{\text{cal}} = \frac{S_{\bar{x}_A}^2}{S_{\bar{x}_B}^2} = 1.84 \quad (16)$$

This value was compared to the two-sided tabulated value at probability  $\alpha=0.05$  and  $\nu_A=\nu_B=9$  degrees of freedom. The tabulated *F*-value is 4.03. As  $F_{\text{cal}} \leq F(0.975, 9,9)$ , the difference between the variance of the means of both methods is not significant. Therefore, Eq. 2a was used to calculate *s<sub>d</sub>* and a value of 1.04 was obtained.

The trueness of method B was checked by calculating the following statistic which is based on Eq. 1:

$$t_{\text{cal}} = \frac{|\bar{x}_A - \bar{x}_B|}{s_d} = 1.42 \quad (17)$$

Finally, the  $t_{\text{cal}}$  was compared to the two-sided  $t$ -value, at a probability  $\alpha=0.05$  and  $\nu=18$  degrees of freedom. The  $t$ -tabulated value is 2.10. Since  $t_{\text{cal}} < t(0.975, 18)$ , the difference between the two grand means was found to be not significant. Therefore, since the accuracy of method B has been assessed against method A, method B is traceable to reference method A.

## **4.2. Estimating uncertainty**

### **4.2.1. Uncertainty arising from the analytical procedure**

The standard uncertainty arising from the experimental variation of future measurements is given by Eq. 4. Since each portion of the working sample (i.e. the contaminated land), was analysed only in one run and with one analysis (i.e.  $p_s=1$  and  $n_s=1$ ), the standard uncertainty is obtained by combining the repeatability variance and the between-run variance obtained in assessing the accuracy for method B. The standard uncertainty was  $2.14 \mu\text{g}\cdot\text{g}^{-1}$  and its associated degrees of freedom (estimated from equation 6) were 13.

### **4.2.2. Uncertainty arising from the assessment of accuracy**

Since the traceability of method B was assessed against method A, the method bias was not found to be statistically significant. The standard uncertainty arising from the assessment of accuracy was estimated with  $s_d$ . From Eq. 2a, the value for  $u_{\text{assessment}}$  was  $1.04 \mu\text{g}\cdot\text{g}^{-1}$  with 18 degrees of freedom.

#### **4.2.3. Uncertainty arising from preprocessing steps**

Since the samples of soil are not homogeneous and the lack of homogeneity was not considered in the assessment of accuracy, an uncertainty term related to the heterogeneity of the sample was included in the overall uncertainty. This term was estimated from the information generated previously with a sample similar to the working sample. An estimate of the between-preprocessing steps standard deviation,  $s_{\text{preproc.steps}}$ , was obtained by taking  $n_{\text{ps}}=10$  different portions of the sample. The different portions subsampled were analysed under repeatability conditions (i.e. the sample pretreatments and measurement were performed in the same run). The standard deviation of the results was  $s_{\text{I(preproc.steps)}}=4.7 \mu\text{g}\cdot\text{g}^{-1}$  with 9 degrees of freedom. Since the standard deviation obtained is much larger than the repeatability of the analytical method, the repeatability variance can be neglected when it is compared to the preprocessing step-different intermediate variance.

#### **4.2.4. Uncertainty arising from other sources**

To obtain a higher level of traceability, method A was traced to a CRM which had similar properties to the test sample and the contaminated land. This provided an estimate of the laboratory bias of method A.

The copper content of the CRM was analysed with method A. As in assessing the accuracy of method B, the experimental design was a two-factor fully-nested design. The factors varied in the fully-nested design were the replicate and the run. Between each run, factors such as the analyst, the day, the calibration, etc. were also varied. The grand mean,  $\bar{x}_A=30.46 \mu\text{g}\cdot\text{g}^{-1}$ , and the variance of the means,  $s_{\bar{x}_A}^2=2.74$ , were used to assess the traceability of method A to the CRM. The certified values from the results of ten laboratories were  $\bar{x}_{\text{CRM}}=31.4 \mu\text{g}\cdot\text{g}^{-1}$  and  $s_{\bar{x}_{\text{CRM}}}^2=5.2$ . The uncertainty,  $u_{\text{other items}}$ , related to this higher link in traceability was estimated as  $0.85 \mu\text{g}\cdot\text{g}^{-1}$ , with 20 degrees of freedom.

#### 4.2.5. *Estimating the overall uncertainty*

The overall uncertainty was estimated by combining the uncertainties arising from the analytical procedure, the assessment of accuracy, the preprocessing steps of the sample and higher levels of traceability and then applying Eq. 15:

$$U = t_{\text{eff},\alpha/2} \sqrt{u_{\text{proc}}^2 + u_{\text{assessment}}^2 + u_{\text{preproc.steps}}^2 + u_{\text{other terms}}^2} = 11.44 \mu\text{g}\cdot\text{g}^{-1}$$

where  $t_{\text{eff}}$  is the two-sided tabulated  $t$ -value for  $\alpha=0.05$  and  $\nu_{\text{eff}}= 14$ .

Nevertheless, this estimated uncertainty does not take into account every component. For instance, because of a lack of references for a higher metrological level than the CRM, the uncertainty arising from the bias of the CRM has not been considered in the term which corresponds to other sources. Moreover, sources of uncertainty arising from some preprocessing steps or from some analytical measurement steps, influenced by factors of variation not considered in the assessment of accuracy, may have been overlooked.

## 5. DISCUSSION

We have presented a new way of calculating the uncertainty in analytical results as a complement to the two main approaches proposed to date: the ISO approach and the Analytical Methods Committee approach. The new proposal takes advantage of the information generated in assessing the accuracy of a given analytical procedure. The advantage of evaluating the uncertainty with this new approach is that the effort made to check the accuracy can also be used to calculate the uncertainty. Therefore, not a great deal of extra work needs to be done. The main benefits of this approach are its conceptual simplicity, its low cost and its universal application.

Compared to the ISO approach, the new proposal has a more holistic character. It means that it is no longer necessary to know the various individual steps into

which the analytical method can be broken down because the analytical procedure is taken as a whole.

The main disadvantages of the approach proposed by the Analytical Methods Committee (i.e. the lack of interlaboratory information for the specific type of matrix required and the fact that the interlaboratory study may not be related to the conditions of the laboratory where the result was obtained) are also overcome by this new approach. However, if the interlaboratory information is available and is related to the present conditions of the laboratory, the “top-down” approach considers components of uncertainty, such as laboratory bias or changes in supplier, that must sometimes be taken into account in the new approach as “Type B” uncertainties.

This new approach also has some drawbacks. The main restrictions are related to constraints in validating the accuracy of measurement methods. Since in practice the accuracy of the analytical methods can only be assessed by a limited number of references placed at high levels in the traceability hierarchy, the analyst has often to resort to references placed at low levels in the traceability hierarchy. The analyst should be aware that the lower the level the reference is placed at, the more terms need to be added to the uncertainty budget in order to include the uncertainties arising from systematic effects. For instance, if the accuracy of a method is assessed against a reference method carried out in the same laboratory and not linked to references which have higher levels of traceability (i.e. not linked to a CRM or to a collaborative trial) uncertainty terms such as laboratory bias and reference method bias should be included in the overall uncertainty. Moreover, the analyst must be aware that the factors chosen to be varied in the assessment of accuracy may not include some of the sources of variation of the measurement method. Depending on the number of factors chosen, therefore, other terms of uncertainty which were not previously taken into account in assessing the accuracy must be included. However, the extra work needed to evaluate them is not comparable to the work needed to estimate the uncertainty with the ISO approach. This makes this approach useful in routine analysis or if no information about the different sources of uncertainty is available.



More work still needs to be done to evaluate the component of uncertainty arising from differences in the level of the measurand and the type of matrix between the test sample and the work sample. Future studies must focus on developing detailed procedures and guidelines which describe the experimental steps and statistics required to test the accuracy of measurement methods and their traceability to different references of high metrological quality.

## 6. REFERENCES

- [1] CEN/CENELEC, European Standard EN 45001: *General Criteria for the operation of testing laboratories*, Brussels, Belgium, 1989
- [2] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *International Vocabulary of basic and general terms in Metrology, VIM*, ISO, Geneva, 1993
- [3] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993
- [4] EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK, 1995
- [5] Analytical Methods Committee, *Analyst* **120** (1995) 2303-2308
- [6] P. de Bièvre, R. Kaarls, H.S. Peiser, S.D. Raspberry, W.P. Reed, *Accred. Qual. Assur.* **2** (1997) 168-179
- [7] P. de Bièvre, R. Kaarls, H.S. Peiser, S.D. Raspberry, W.P. Reed, *Accred. Qual. Assur.* **2** (1997) 270-274
- [8] M. Valcárcel, A. Ríos, *Analyst* **120** (1995) 2291-2297
- [9] M. Thompson, *Analyst* **126** (1996) 285-288
- [10] B. King, *Analyst* **122** (1997) 197-204
- [11] International Organization for Standardization, *ISO 5725-6. Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva, 1994
- [12] D.L. Massart, J. Smeyers-Verbeke, *Robust process analytical methods for industrial practice and their validation*, Report Project SMT4-CT95-2031, 1996
- [13] International Organization for Standardization, *ISO 5725-3. Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva, 1994
- [14] V. Barnett, T. Lewis, *Outliers in Statistical Data*, 3<sup>rd</sup> Ed. Wiley, Chichester, 1994
- [15] M. Ramsey, A. Argyraki, M. Thompson, *Analyst* **120** (1995) 2309-2312
- [16] F.E. Satterthwaite, *Psychometrika* **6** (1941) 309-316
- [17] C.F. Dietrich, *Uncertainty, Calibration and Probability. The statistic of Scientific and Industrial Measurement*, 2<sup>nd</sup> Ed. Adam Hilger, Bristol, 1991

## 2.4. Cálculo de los términos de incertidumbre utilizando información de la validación del método

En este apartado se explica cómo calcular globalmente la incertidumbre utilizando los diferentes tipos de información obtenidos durante la validación del método. En el apartado anterior se ha calculado la incertidumbre utilizando fundamentalmente la información obtenida en la verificación de la trazabilidad. Sin embargo, los gráficos de control, los estudios de precisión y los de robustez también proporcionan información útil para calcular la incertidumbre. A continuación se detalla cómo incorporar esta información en el cálculo de los diferentes términos parciales de la incertidumbre.

### 2.4.1. Incertidumbre del procedimiento ( $u_{\text{proc}}$ )

La incertidumbre del procedimiento considera la variabilidad experimental debida a los errores aleatorios y a las condiciones en las que se realiza el análisis (día, analista, calibrado, etc.). Este término se calcula a partir de la precisión del procedimiento analítico:

$$u_{\text{proc}} = \sqrt{\frac{s_{\text{serie}}^2}{p_s} + \frac{s_r^2}{p_s \cdot n_s}} \quad (2.2)$$

donde  $s_{\text{serie}}$  es la desviación estándar entre series del procedimiento,  $s_r$  la desviación estándar de la repetibilidad del procedimiento,  $p_s$  es el número de veces que se analiza la muestra de rutina en diferentes series y  $n_s$  el número de replicados que se hace en cada una de estas series (normalmente  $p_s=1$  y  $n_s=1$  ó 2).

En el apartado anterior la incertidumbre del procedimiento se calcula asumiendo que se conocen la desviación estándar entre series,  $s_{\text{serie}}$ , y la repetibilidad,  $s_r$ , del procedimiento. Sin embargo, en ocasiones sólo se conoce la desviación estándar intermedia del procedimiento,  $s_i$ . En este caso, la incertidumbre del procedimiento se calcula como:

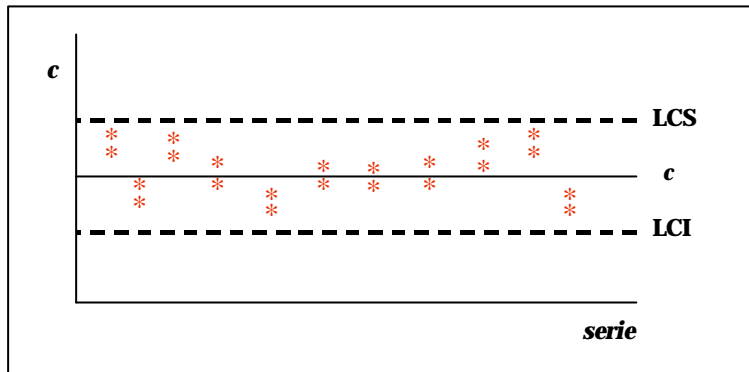
$$u_{\text{proc}} = \frac{S_1}{\sqrt{p_s}} \quad (2.3)$$

La desventaja de esta expresión es que no se puede disminuir la incertidumbre del resultado aunque la muestra de rutina se analice varias veces en condiciones de repetibilidad (es decir, si  $n_s > 1$ ).

#### 2.4.1.1. Mejora de la estimación de la incertidumbre del procedimiento

En el apartado anterior, la precisión del procedimiento se calcula utilizando únicamente la precisión obtenida durante la verificación de la trazabilidad. No obstante, si la muestra de referencia no se analiza un número suficiente de veces en condiciones intermedias puede obtenerse una mala estimación de la incertidumbre del procedimiento (según la IUPAC<sup>9</sup> la referencia debe analizarse al menos 7 veces y, preferiblemente, más de 10 veces). En estos casos, se puede utilizar la información obtenida en los estudios de precisión del método (explicados en el Capítulo 1) o la suministrada por los gráficos de control. Los gráficos de control proporcionan información sobre la precisión intermedia del procedimiento ya que, como la muestra de control se analiza a lo largo de tiempo, se obtiene una precisión que considera todos aquellos factores que afectan al procedimiento analítico (día, analista, calibrado, etc.). Si además se quiere obtener la repetibilidad y la variabilidad entre series del procedimiento, es necesario hacer un duplicado cada vez que se analice la muestra de control (ver Fig. 2.4). A partir del ANOVA de los resultados, se puede obtener una estimación de la repetibilidad, de la precisión entre series y de la precisión intermedia (ver apartado 1.8 del capítulo anterior).

Las varianzas calculadas a partir de diversos estudios (es decir, estudios de precisión, gráficos de control y/o verificación de la trazabilidad), pueden combinarse para dar lugar a una varianza conjunta (o “*pooled variance*”). De esta forma, se obtiene una estimación mejor de la incertidumbre del procedimiento ya que las varianzas asociadas a los distintos tipos de precisión del procedimiento (repetibilidad, precisión entre series y precisión intermedia) poseen más grados de libertad asociados.



**Figura 2.4.** Gráfico de control con dos replicados en cada serie. LCS y LCI son, respectivamente, los límites de control superior e inferior y  $c$  es la concentración asignada a la muestra de control.

### 2.4.2. Incertidumbre de la verificación de la trazabilidad ( $u_{\text{traz}}$ )

Al verificar la trazabilidad de los resultados, siempre existe una incertidumbre al afirmar que los resultados son trazables. Es decir, aunque los resultados resulten ser trazables, nunca se tendrá una seguridad del 100% en esta afirmación. Verificar la trazabilidad implica calcular el sesgo del método. Tal y como se ha visto en el apartado 1.7. del capítulo anterior un método puede tener dos tipos de sesgos: un sesgo constante y un sesgo proporcional. No obstante, si el método analítico tiene un intervalo de concentraciones restringido, se puede asumir que el sesgo del método es el mismo en todo el intervalo de aplicación del método. Para calcular este sesgo debe analizarse varias veces en condiciones intermedias una muestra de referencia que sea similar a las muestras de rutina en términos de concentración y de matriz. De esta forma, puede calcularse el sesgo del método como la diferencia entre el valor de referencia,  $\bar{x}_{\text{ref}}$ , y el valor obtenido,  $\bar{x}_{\text{método}}$ , al analizar varias veces la muestra de referencia:

$$\text{sesgo} = \left| \bar{x}_{\text{ref}} - \bar{x}_{\text{método}} \right| \quad (2.4)$$

La incertidumbre de la verificación de la trazabilidad corresponde a la incertidumbre del sesgo calculado. Esta incertidumbre,  $u_{\text{traz}}$ , depende del tipo de

referencia que se ha utilizado para verificar la trazabilidad de los resultados. En el apartado anterior se ha explicado cómo calcular la incertidumbre cuando la trazabilidad se verifica con un método de referencia. Esta metodología también es aplicable cuando se utilizan como referencia los métodos definitivos o bien otro tipo de métodos. No obstante, no es aplicable cuando la trazabilidad se verifica utilizando un valor de referencia como, por ejemplo, el valor asignado a un material de referencia, el valor de consenso de un ejercicio interlaboratorio o bien la cantidad de analito adicionada. Esto es debido a que para verificar la trazabilidad con un método de referencia debe hacerse un test  $F$  para evaluar si las varianzas de la precisión intermedia de ambos métodos son o no comparables. Sin embargo, este test no debe realizarse cuando la trazabilidad se verifica utilizando un valor de referencia. A continuación, se explica cómo calcular la incertidumbre de la verificación de la trazabilidad,  $u_{\text{traz}}$ , cuando la trazabilidad se verifica utilizando un valor de referencia. Para calcular esta incertidumbre, debe aplicarse la ley de propagación de errores a la Ec. 2.4:

$$u_{\text{traz}} = \sqrt{u_{\text{ref}}^2 + u_{\bar{x}_{\text{método}}}^2} \quad (2.5)$$

donde  $u_{\text{ref}}$  corresponde a la incertidumbre del valor de referencia y  $u_{\bar{x}_{\text{método}}}$  es la incertidumbre del valor medio obtenido con el método analítico al analizar la muestra de referencia.

#### **2.4.2.1. Incertidumbre del valor de referencia**

La incertidumbre del valor de referencia,  $u_{\text{ref}}$ , depende de que la referencia sea un material de referencia, el valor de consenso de un ejercicio interlaboratorio o bien una cantidad de analito adicionada:

##### Material de referencia

La incertidumbre del valor de referencia asignado a un material de referencia se calcula como:

$$u_{\text{ref}} = \frac{U_{\text{RM}}}{k} \quad (2.6)$$

donde  $U_{\text{RM}}$  es la incertidumbre del valor de referencia proporcionado por el fabricante y  $k$  es el factor de cobertura. Normalmente, debe utilizarse un valor de  $k=2$  ya que la incertidumbre del valor de referencia suele estar calculada para un 95% de confianza.

#### Valor de consenso de un ejercicio interlaboratorio

La incertidumbre asociada al valor de consenso depende de cómo se haya obtenido el valor de consenso. En el caso de que se utilicen tests de *outliers* para eliminar los laboratorios discrepantes, el valor de consenso se obtiene a partir del valor medio de los laboratorios no eliminados. Si cada laboratorio ha proporcionado un solo resultado, la incertidumbre del valor de consenso se calcula como:

$$u_{\text{ref}} = \frac{s_{\text{lab}}}{\sqrt{n_{\text{lab}}}} \quad (2.7)$$

donde  $s_{\text{lab}}$  es la desviación estándar de los resultados proporcionados por los laboratorios no eliminados y  $n_{\text{lab}}$  es el número de laboratorios no eliminados. Otra posibilidad es que cada laboratorio proporcione  $n$  replicados. En este caso, la incertidumbre se obtiene aplicando el análisis de la varianza a los resultados proporcionados por los laboratorios no identificados como discrepantes<sup>10</sup>.

En vez de utilizar los tests de *outliers*, hay otros protocolos<sup>11</sup> que recomiendan calcular el valor de consenso y la dispersión de los resultados utilizando la estadística robusta. En este caso, la incertidumbre del valor de consenso depende del estimador robusto utilizado.

#### Cantidad de analito adicionada

La incertidumbre del valor de referencia debe calcularse aplicando la ley de propagación de errores a la expresión utilizada para calcular la cantidad de analito adicionada. El caso más usual es que se adicione un volumen,  $V_0$ , de una

disolución con una concentración de analito,  $c_0$ , a un volumen de muestra,  $V$ . En este caso la concentración de analito adicionado se calcula como:

$$c_a = \frac{V_0 \cdot c_0}{V} \quad (2.8)$$

La incertidumbre se calcula siguiendo la ley de propagación de errores tal y como se ha descrito en el apartado 1.10.1. del Capítulo 1:

$$u_{\text{ref}} = c_a \cdot \sqrt{\frac{u_{V_0}^2}{V_0^2} + \frac{u_{c_0}^2}{c_0^2} + \frac{u_V^2}{V^2}} \quad (2.9)$$

donde  $u_{V_0}$  es la incertidumbre estándar del volumen adicionado,  $u_{c_0}$  es la incertidumbre estándar de la concentración de analito,  $c_0$ , y  $u_V$  es la incertidumbre estándar del volumen de muestra,  $V$ .

#### 2.4.2.2. Incertidumbre del valor medio del método analítico

Esta incertidumbre depende de cómo se ha analizado la muestra de referencia con el método analítico. Si se ha seguido el diseño anidado de dos factores propuesto en el apartado 2.3. (es decir, que se analice la muestra de referencia en  $p$  series y que se hagan  $n$  réplicas en cada una de estas series), la incertidumbre se calcula como:

$$u_{\bar{x}_{\text{método}}} = \sqrt{\frac{s_{\text{serie}}^2}{p} + \frac{s_r^2}{p \cdot n}} \quad (2.10)$$

donde  $s_{\text{serie}}$  es la desviación estándar entre series del procedimiento y  $s_r$  es la desviación estándar de la repetibilidad del procedimiento. Estas desviaciones estándar pueden calcularse a partir de los resultados obtenidos al analizar la muestra de referencia siguiendo la metodología propuesta en el apartado anterior. Sin embargo, la precisión obtenida a partir de estudios de precisión o de gráficos

de control también puede utilizarse para calcular estas desviaciones estándar. Es aconsejable utilizar esta información cuando la muestra de referencia se ha analizado pocas veces (la IUPAC<sup>9</sup> recomienda analizar la muestra en al menos 7 series diferentes y, preferiblemente, en más de 10 series).

Si no se desea seguir diseños anidados, otra forma de verificar la trazabilidad de los resultados consiste en analizar  $p$  veces en condiciones intermedias la muestra de referencia. En este caso, la incertidumbre se calcula como:

$$u_{x_{\text{método}}} = \frac{s_I}{\sqrt{p}} \quad (2.11)$$

donde  $s_I$  es la desviación estándar intermedia del procedimiento y puede calcularse como la desviación estándar de los  $p$  resultados obtenidos. Esta desviación estándar también puede obtenerse utilizando gráficos de control y estudios de precisión. Al igual que en el caso anterior, es aconsejable utilizar esta información si la muestra de referencia se ha analizado pocas veces.

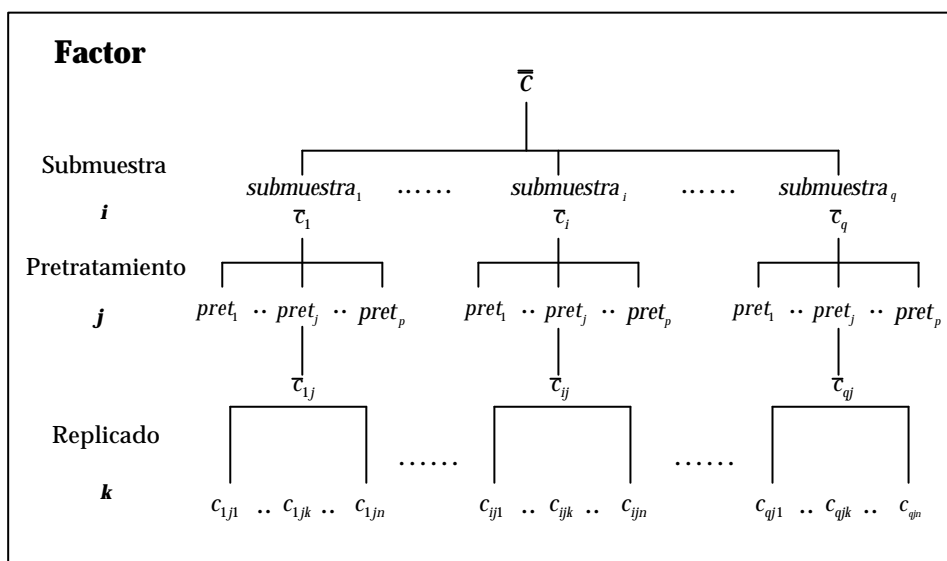
### **2.4.3. Incertidumbre del muestreo, submuestreo y pretratamientos**

La incertidumbre del muestreo depende de la heterogeneidad de la muestra y de los errores cometidos en alguna etapa del proceso de muestreo que den lugar, por ejemplo, a una contaminación de la muestra o a la volatilización del analito que se determina. Esta incertidumbre suele ser una de las fuentes mayores de incertidumbre del resultado final<sup>12-14</sup>. Asimismo, la incertidumbre del submuestreo y de los pretratamientos también puede contribuir de forma significativa a la incertidumbre de los resultados<sup>15-18</sup>. Esta incertidumbre debe calcularse cuando la muestra no es homogénea o bien cuando a la muestra de rutina se la somete a pretratamientos que no se han realizado sobre la muestra de referencia utilizada para verificar la trazabilidad de los resultados generados por el método analítico.



La incertidumbre del muestreo puede calcularse siguiendo la metodología propuesta por Heikka *et al*<sup>19,20</sup>. Esta metodología está basada en modelar la heterogeneidad de la muestra utilizando las series temporales propuestas en la teoría de muestreo desarrollada por Gy<sup>21</sup>. Asimismo, Ramsey ha propuesto calcular la incertidumbre del muestreo utilizando la información generada en ejercicios interlaboratorio donde varios laboratorios llevan a cabo el muestreo. Después, el organismo que ha organizado el ejercicio interlaboratorio debe analizar por replicado cada una de las muestras tomadas por los laboratorios participantes. A partir del análisis de la varianza, se obtiene la incertidumbre asociada al muestreo<sup>22-24</sup>.

En el apartado 2.3. se propone calcular la incertidumbre del submuestreo y los pretratamientos utilizando una muestra representativa de las muestras de rutina. Para ello, deben analizarse  $q$  porciones de la muestra variando todos aquellos factores que afecten al submuestreo y a los pretratamientos. Estos factores pueden variarse de forma ordenada siguiendo diseños anidados. La Fig. 2.5 muestra un diseño anidado de 3 factores para analizar diferentes porciones de muestra.



**Figura 2.5.** Diseño anidado de tres factores para calcular la incertidumbre del submuestreo y de los pretratamientos en el análisis de una muestra heterogénea.

Este diseño ha sido aplicado por Pau *et al*<sup>15</sup> para calcular la incertidumbre en la determinación de calcio y zinc en piensos por absorción atómica. La ventaja de este diseño es que a partir del ANOVA<sup>25</sup> de los resultados se puede calcular la incertidumbre del submuestreo y la incertidumbre de los pretratamientos por separado. La Tabla 2.1 muestra el ANOVA asociado a los resultados de la Fig. 2.5. La Tabla 2.2 muestra las varianzas que pueden calcularse a partir del ANOVA de la Tabla 2.1.

**Tabla 2.1.** Tabla ANOVA del diseño anidado de 3 factores de la Fig. 2.5 (g.l. son los grados de libertad).  $c_{ijk}$  es el replicado  $k$  obtenido al analizar la submuestra  $i$  con el pretratamiento  $j$ .  $\bar{c}_j$  es el valor medio de todos los replicados obtenidos al analizar la submuestra  $i$  con el pretratamiento  $j$ .  $\tau_i$  es el valor medio de los resultados obtenidos con la submuestra  $i$  y  $\bar{\tau}$  la gran media.

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Submuestra	$MS_{\text{sub}} = \frac{p \cdot n \cdot \sum_{i=1}^q (\tau_i - \bar{\tau})^2}{q-1}$	$q-1$	$p \cdot n \cdot \sigma_{\text{sub}}^2 + n \cdot \sigma_{\text{pret}}^2 + \sigma_{\text{prec}}^2$
Pretratamiento	$MS_{\text{pret}} = \frac{n \cdot \sum_{i=1}^q \sum_{j=1}^p (\tau_{ij} - \tau_i)^2}{q \cdot (p-1)}$	$q \cdot (p-1)$	$n \cdot \sigma_{\text{pret}}^2 + \sigma_{\text{prec}}^2$
Precisión	$MS_{\text{prec}} = \frac{\sum_{i=1}^q \sum_{j=1}^p \sum_{k=1}^n (\bar{c}_{ijk} - \bar{c}_{ij})^2}{q \cdot p \cdot (n-1)}$	$q \cdot p \cdot (n-1)$	$\sigma_{\text{prec}}^2$

**Tabla 2.2.** Cálculo de las varianzas obtenidas a partir del ANOVA de la Tabla 2.1.

Varianza	Expresión
Precisión, $s_{\text{prec}}^2$	$MS_{\text{prec}}$
Entre pretratamientos, $s_{\text{pret}}^2$	$\frac{MS_{\text{pret}} - MS_{\text{prec}}}{n}$
Entre submuestras, $s_{\text{sub}}^2$	$\frac{MS_{\text{sub}} - MS_{\text{pret}}}{p \cdot n}$

Las varianzas de la Tabla 2.2 proporcionan la incertidumbre del submuestreo,  $u_{\text{sub}}$ , y la incertidumbre de los pretratamientos,  $u_{\text{pret}}$  (es decir,  $u_{\text{sub}}=s_{\text{sub}}$  y  $u_{\text{pret}}=s_{\text{pret}}$ ). La desventaja de esta metodología es que estos diseños requieren un mayor número de análisis. Una alternativa a este diseño es calcular globalmente la incertidumbre del submuestreo y los pretratamientos. Para ello, puede seguirse un diseño anidado de dos factores tal y como se propone en el apartado 2.3. Este diseño consiste en hacer dos replicados al analizar cada porción de muestra a la que se le han hecho los pretratamientos. Estos dos replicados deben hacerse en condiciones de repetibilidad únicamente si todas las porciones de muestra se analizan con el procedimiento analítico en las mismas condiciones (mismo día, analista, etc.). Si cada una de las porciones se analiza en diferentes condiciones, los replicados deben hacerse en condiciones de precisión intermedia (es decir, diferente día, analista, etc.). A partir del ANOVA de los resultados de este diseño, se obtiene una incertidumbre global del submuestreo y los pretratamientos.

Una metodología más simple es que cada porción se analice una sola vez con el procedimiento analítico. En este caso, la variabilidad de los resultados de cada porción es debida al submuestreo y a los pretratamientos pero también al propio procedimiento analítico. Por tanto, para no sobreestimar la incertidumbre es necesario sustraer la variabilidad debida al procedimiento analítico:

$$u_{\text{sub-pret}} = \sqrt{s_{\text{sub-pret}}^2 - s_{\text{condiciones}}^2} \quad (2.12)$$

donde  $s_{\text{sub-pret}}^2$  es la varianza de los resultados de las diferentes porciones y  $s_{\text{condiciones}}^2$  considera la variabilidad de los resultados debida al procedimiento analítico. Esta varianza depende de cómo se han analizado cada una de las porciones con el procedimiento. Si se han analizado en condiciones de repetibilidad, corresponde a la repetibilidad del procedimiento ( $s_{\text{condiciones}}^2 = s_r^2$ ) y si se han analizado en condiciones intermedias corresponde a la precisión intermedia del procedimiento ( $s_{\text{condiciones}}^2 = s_i^2$ ).

#### **2.4.4. Otros términos de incertidumbre ( $u_{\text{otros}}$ )**

Estos términos están asociados a la variabilidad de la matriz de las muestras de rutina, a factores de variabilidad no considerados en el cálculo de la precisión intermedia y a niveles más altos de trazabilidad. Este último componente debe incluirse si la referencia utilizada en la verificación de la trazabilidad se ha trazado a su vez a una referencia situada en un nivel superior de la pirámide metrológica (ver Fig. 1.2). A continuación, se explica cómo aplicar los estudios de robustez y las técnicas de regresión para calcular la incertidumbre de términos no considerados en la precisión intermedia. Por último, se explica cómo utilizar las muestras adicionales para calcular la incertidumbre debida a la variabilidad de la matriz de las muestras de rutina.

##### **2.4.4.1. Utilización de estudios de robustez y técnicas de regresión para calcular la incertidumbre de factores no considerados en la precisión intermedia**

La precisión intermedia del método analítico debería obtenerse variando representativamente todos aquellos factores que pueden afectar a los resultados analíticos. De esta forma, la incertidumbre del procedimiento incluye todas aquellas fuentes de incertidumbre que pueden afectar a los resultados. Sin embargo, hay veces que algún factor no se varía suficientemente como, por ejemplo, el tiempo de extracción o la composición de la fase móvil de un método cromatográfico. En estos casos, la incertidumbre del procedimiento no considera la incertidumbre asociada a los factores que no se han variado representativamente y, por tanto, es necesario incluir la incertidumbre de estos factores en la incertidumbre final del resultado. Para calcular la incertidumbre de estos factores, Ellison *et al*<sup>26-28</sup> han propuesto utilizar la información obtenida en los estudios de robustez. La robustez es un parámetro de calidad importante que se evalúa en las primeras etapas del desarrollo de un método analítico. Este parámetro mide la capacidad de un método analítico de no ser afectado por pequeñas pero deliberadas variaciones de algunos parámetros del método. Por tanto, proporciona una indicación de la fiabilidad del método durante el análisis de muestras de rutina<sup>29</sup>.

Los estudios de robustez deben realizarse en una muestra que sea representativa de las muestras de rutina. En primer lugar, deben identificarse todos aquellos factores que pueden afectar al método analítico. En métodos cromatográficos, algunos de estos factores son la temperatura, la velocidad de flujo, la composición de la fase móvil, etc. Después, debe establecerse cuál es el valor nominal de estos factores así como sus valores extremos. La Tabla 2.3 muestra un ejemplo de los valores extremos de los factores estudiados para evaluar la robustez de un método cromatográfico.

**Tabla 2.3.** Valores máximos y mínimos propuestos para estudiar la robustez de un método cromatográfico.  $\delta_{\text{test}}$  corresponde a la diferencia entre el valor máximo y el valor mínimo.

Factor	Mínimo (-)	Máximo (+)	$\delta_{\text{test}}$
Temperatura (°C)	30	70	40
Flujo (ml/min)	0.9	1.1	0.2
%THF	49	80	31

Posteriormente, se debe escoger cuál es el diseño experimental que debe utilizarse para evaluar la influencia de estos factores. Normalmente, se utilizan los diseños de experimentos de Plakett-Burman<sup>26,30</sup>. Con estos diseños, es necesario realizar al menos  $n+1$  experimentos para determinar el efecto de  $n$  factores. Un ejemplo de estos diseños se muestra en la Tabla 2.4:

**Tabla 2.4.** Diseño de Plakett-Burman para evaluar el efecto de 3 factores. Los signos (+) y (-) indican respectivamente el valor máximo y mínimo que puede tomar cada uno de los factores.

Experimento	Factor			Resultado
	A (Temperatura)	B (Flujo)	C (%THF)	
1	+	+	+	$y_1$
2	-	+	-	$y_2$
3	+	-	-	$y_3$
4	-	-	+	$y_4$

El efecto de cada uno de los factores se calcula como la diferencia entre el valor medio de los resultados obtenidos al fijar el factor en su valor máximo (+) y el

valor medio de los resultados obtenidos al fijar el factor en su valor mínimo (-). Por ejemplo, el efecto del factor A de la Tabla 2.4 se calcula como:

$$D_A = \left[ \frac{(y_1 + y_3)}{2} - \frac{(y_2 + y_4)}{2} \right] \quad (2.13)$$

Una vez calculado el efecto del factor A, debe evaluarse mediante un test *t* si el factor A produce variaciones significativas en los resultados:

$$t = \frac{D_A \cdot \sqrt{n}}{s_1 \cdot \sqrt{2}} \quad (2.14)$$

donde *n* es el número de experimentos llevados a cabo en cada nivel para determinar  $D_A$  (en este caso,  $n=2$ ) y  $s_1$  es la desviación estándar intermedia del método. El valor de *t* obtenido debe compararse con el valor de *t* tabulado,  $t_{tab}$ , de dos colas para *p*-1 grados de libertad (donde *p* es el número de análisis que se han realizado para obtener la precisión intermedia,  $s_1$ ).

La variación que proporciona el factor A en los resultados obtenidos al analizar muestras de rutina puede calcularse como:

$$\Delta_A = \frac{D_A}{\ddot{a}_{test}} \cdot \ddot{a}_{real} \quad (2.15)$$

donde  $\delta_{est}$  es lo que se varía el factor en el estudio de robustez (se calcula como la diferencia del factor en su valor máximo y mínimo, ver Tabla 2.3) y  $\delta_{real}$  es lo que varía el factor cuando el método analítico está bajo condiciones de aseguramiento de la calidad. Por ejemplo, en el caso de que se quiera calcular la incertidumbre de la temperatura (ver Tabla 2.3),  $\delta_{est}$  corresponde a 40°C (es decir, a 70 menos 30).  $\delta_{real}$  corresponde a la variabilidad que puede tener la temperatura durante el análisis de las muestras de rutina. Este valor puede obtenerse a partir de las especificaciones del fabricante. Si, por ejemplo, el fabricante garantiza la temperatura del horno con un intervalo de  $\pm 0.5^\circ\text{C}$ ,  $\delta_{real}$  correspondería a 1°C. La

incertidumbre del factor A se obtiene aplicando la ley de propagación de errores a la Ec. 2.15:

$$u(\Delta_A) = \sqrt{\left(\frac{\ddot{a}_{\text{real}}}{\ddot{a}_{\text{test}}}\right)^2 \cdot u_{D_A}^2 + \left(\frac{D_A}{\ddot{a}_{\text{test}}}\right)^2 \cdot u_{\ddot{a}_{\text{real}}}^2} = \sqrt{\left(\frac{\ddot{a}_{\text{real}}}{\ddot{a}_{\text{test}}}\right)^2 \cdot \left(\frac{s_I \cdot \sqrt{2}}{\sqrt{n}}\right)^2 + \left(\frac{D_A}{\ddot{a}_{\text{test}}}\right)^2 \cdot \left(\frac{\ddot{a}_{\text{real}}}{\sqrt{3}}\right)^2} \quad (2.16)$$

Esta expresión debe aplicarse cuando el factor A es significativo. Sin embargo, la expresión propuesta por Ellison *et al*<sup>26</sup> cuando el factor es significativo únicamente considera el segundo término de la Ec. 2.16.

Si el factor no es significativo, la incertidumbre se calcula considerando únicamente el primer término de la Ec. 2.16:

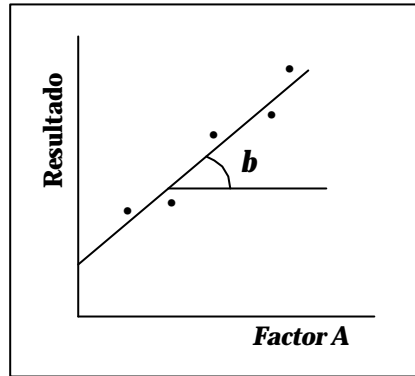
$$u(\Delta_A) = \sqrt{\left(\frac{\ddot{a}_{\text{real}}}{\ddot{a}_{\text{test}}}\right)^2 \cdot u_{D_A}^2} = \frac{\ddot{a}_{\text{real}} \cdot s_I \cdot \sqrt{2}}{\ddot{a}_{\text{test}} \cdot \sqrt{n}} \quad (2.17)$$

Ellison *et al* proporcionan una expresión similar para calcular la incertidumbre cuando el factor A no es significativo:

$$u(\Delta_A) = \frac{t_{\text{tab}} \cdot \ddot{a}_{\text{real}} \cdot s_I \cdot \sqrt{2}}{1.96 \cdot \ddot{a}_{\text{test}} \cdot \sqrt{n}} \quad (2.18)$$

La única diferencia es el factor  $t_{\text{tab}}/1.96$ . Este factor se incluye para considerar los grados de libertad asociados a la incertidumbre de A. Sin embargo, para considerar estos grados de libertad, es mejor utilizar la aproximación de Satterthwaite para calcular los grados de libertad efectivos de la incertidumbre final de los resultados<sup>7,8,31,32</sup> (ver Capítulo 1).

Las técnicas de regresión también pueden utilizarse para calcular la incertidumbre de factores no considerados en la precisión intermedia. Para ello, si se quiere calcular la incertidumbre de un factor A, es necesario analizar una muestra similar a las muestras de rutina a varios niveles de dicho factor. De esta forma, se puede relacionar el resultado obtenido con el valor del factor A (ver Fig. 2.6).



**Figura 2.6.** Recta de regresión a la que se ajustan los resultados obtenidos al variar el factor A.

La incertidumbre del factor A se puede calcular a partir de la pendiente de la recta,  $b$ , ajustada por mínimos cuadrados. En este caso, la variabilidad que proporciona ese factor en los resultados se calcula como:

$$\Delta_A = b \cdot \ddot{a}_{\text{real}} \quad (2.19)$$

donde  $\ddot{a}_{\text{real}}$  es lo que varía el factor cuando el método analítico está bajo condiciones de aseguramiento de la calidad.

La incertidumbre del factor se calcula aplicando la ley de propagación de errores a la Ec. 2.19:

$$u_{\Delta_A} = \sqrt{\left(\ddot{a}_{\text{real}} \cdot s_b\right)^2 + \left(b \cdot \frac{\ddot{a}_{\text{real}}}{\sqrt{3}}\right)^2} \quad (2.20)$$

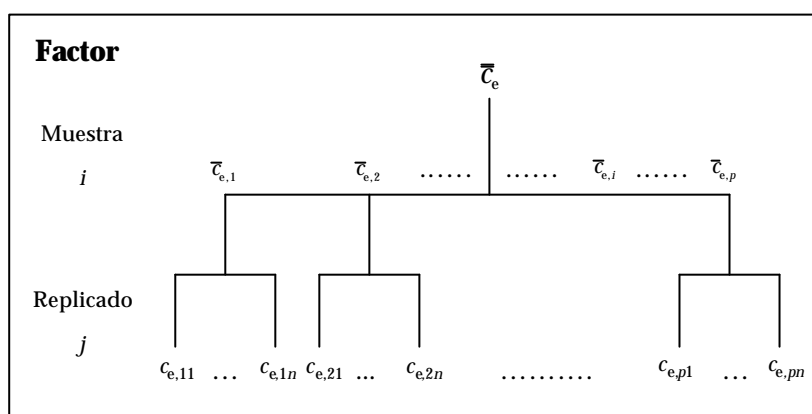
donde  $s_b$  es la desviación estándar de la pendiente de la recta. Cowles *et al*<sup>33</sup> proponen esta metodología como una alternativa a los estudios de robustez ya que utilizar técnicas de regresión permite variar el factor estudiado en un intervalo de variabilidad mayor. De esta forma, se obtienen incertidumbres más realistas. Esto es debido a que en los estudios de robustez los factores se varían en un intervalo más restringido. Por tanto, es más probable concluir erróneamente que un factor no es significativo. No obstante, la desventaja de utilizar técnicas de regresión es



que supone más trabajo adicional que los estudios de robustez cuando se quiere determinar la incertidumbre de más de un factor. La expresión propuesta por Cowles *et al*<sup>33</sup> únicamente considera el segundo término de la Ec. 2.20. Sin embargo, el término asociado a la incertidumbre de la pendiente de la recta también debe considerarse.

#### 2.4.4.2. Utilización de muestras adicionales para calcular la incertidumbre de la variabilidad de la matriz de las muestras de rutina

La incertidumbre asociada a los diferentes tipos de matriz de las muestras de rutina puede calcularse adicionando una misma cantidad de analito a diferentes tipos de muestras que sean representativas de la variabilidad de las muestras de rutina. La Fig. 2.7 muestra el diseño anidado de dos factores que puede seguirse para calcular la incertidumbre debida a la matriz.



**Figura 2.7.** Diseño anidado de 2 factores para calcular la incertidumbre de la variabilidad de la matriz de las muestras de rutina.

En este diseño se adiciona una cantidad de analito  $c_a$  a  $p$  muestras representativas de las matrices de las muestras de rutina. Después debe analizarse cada una de estas muestras antes y después de adicionar el analito. De esta forma, se puede calcular la concentración de analito encontrada,  $c_e$ , en cada una de las muestras:

$$c_e = c_{0+a} - c_0 \quad (2.21)$$

donde  $c_{0+a}$  es la concentración de la muestra después de adicionar el analito y  $c_0$  es la concentración inicial de la muestra. Si la matriz no afecta a los resultados del método, debería encontrarse la misma cantidad de analito en cada una de las  $p$  muestras. Por tanto, a partir de la variabilidad de las concentraciones encontradas para cada una de las muestras se puede obtener la incertidumbre de la variabilidad de la matriz de las muestras de rutina. Sin embargo, hay que tener en cuenta que esta variabilidad no es debida únicamente a la matriz de las muestras de rutina. Además, existe una variabilidad de los resultados debida a los errores aleatorios y a las condiciones en que se hacen los análisis (es decir, al día, al analista, etc.). Por tanto, para no sobreestimar la incertidumbre de la matriz, es necesario sustraer la variabilidad debida a los errores aleatorios y a las condiciones en que se hace la medida. Para ello, puede aplicarse el ANOVA a los resultados obtenidos en la Fig. 2.7. La Tabla 2.5 muestra el ANOVA asociado a los resultados de la Fig. 2.7 y la Tabla 2.6 las diferentes varianzas que pueden calcularse a partir del ANOVA de la Tabla 2.5.

**Tabla 2.5.** Tabla ANOVA del diseño anidado de 2 factores de la Fig. 2.7 (g.l. son los grados de libertad).  $c_{e,ij}$  es la concentración encontrada con el replicado  $j$  al analizar la muestra adicionada  $i$ .  $\bar{c}_{e,i}$  es el valor medio de todas las concentraciones encontradas al analizar la muestra adicionada  $i$ .  $\bar{c}_e$  es la media de todos los  $\bar{c}_{e,i}$ .

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Muestra	$MS_{\text{muestra}} = \frac{n \cdot \sum_{i=1}^p (\bar{c}_{e,i} - \bar{c}_e)^2}{p-1}$	$p-1$	$n\sigma_{\text{muestra}}^2 + \sigma_{\text{prec}}^2$
Replicado	$MS_{\text{prec}} = \frac{\sum_{i=1}^p \sum_{j=1}^n (c_{e,ij} - \bar{c}_{e,i})^2}{p(n-1)}$	$p(n-1)$	$\sigma_{\text{prec}}^2$

La incertidumbre de la variabilidad de la matriz corresponde a la varianza entre muestras:

$$u_{\text{matriz}} = s_{\text{muestra}} \tag{2.22}$$

**Tabla 2.6.** Cálculo de las varianzas obtenidas a partir del ANOVA de la Tabla 2.5.

Varianza	Expresión
Precisión, $s_{\text{prec}}^2$	$MS_{\text{prec}}$
Entre muestras, $s_{\text{muestra}}^2$	$\frac{MS_{\text{muestra}} - MS_{\text{prec}}}{n}$

Una alternativa más sencilla al ANOVA consiste en analizar una sola vez cada una de las  $p$  muestras adicionadas. En este caso, la incertidumbre de la matriz se calcula como:

$$u_{\text{matriz}} = \sqrt{s_{\text{total}}^2 - s_{\text{condiciones}}^2} \quad (2.23)$$

donde  $s_{\text{total}}^2$  corresponde a la varianza de la concentración encontrada,  $c_e$ , para cada una de las  $p$  muestras,  $s_{\text{condiciones}}^2$  corresponde a la varianza asociada a las condiciones en que se analizan las muestras y depende de que las muestras adicionadas tengan o no una concentración inicial de analito. Si tienen una concentración inicial de analito,  $s_{\text{condiciones}}^2$  depende de en qué condiciones se analiza la muestra adicionada y la muestra sin adicionar. Normalmente, se analizan en condiciones de repetibilidad. En este caso, el sesgo de la serie se compensa ya que es el mismo al analizar la muestra adicionada y la muestra sin adicionar. Por tanto, la variabilidad de los resultados sólo es debida a la repetibilidad del método y  $s_{\text{condiciones}}^2 = 2 \cdot s_r^2$ . Si la muestra adicionada y la muestra sin adicionar se analizan en diferentes series, el sesgo debido a la serie no se compensa y la variabilidad de los resultados está asociada a la precisión intermedia del método (es decir,  $s_{\text{condiciones}}^2 = 2 \cdot s_I^2$ ). Si las adiciones se hacen a muestras que no contienen una concentración de analito,  $s_{\text{condiciones}}^2$  depende de las condiciones en que se analicen las muestras adicionadas. Si todas ellas se analizan en la misma serie, la variabilidad de los resultados sólo es debida a la repetibilidad del método y  $s_{\text{condiciones}}^2 = s_r^2$ . Si se analizan en diferentes series, la variabilidad de los resultados es debida a la precisión intermedia del método y  $s_{\text{condiciones}}^2 = s_I^2$ .

## **2.5. Comparación de la aproximación propuesta con la aproximación ISO y la aproximación del *Analytical Methods Committee***

Comparada con la aproximación ISO, la aproximación propuesta tiene la ventaja de que requiere poco trabajo adicional para calcular la incertidumbre ya que la exactitud del método siempre debe verificarse antes de aplicarlo al análisis de muestras de rutina. Este trabajo adicional puede ser debido al cálculo de la incertidumbre del submuestreo y pretratamientos y también a factores que no se hayan variado representativamente en el cálculo de la precisión intermedia. De todas formas, este trabajo adicional es mucho menor que el necesario para calcular la incertidumbre con la aproximación ISO<sup>8,34</sup>. Otra de las ventajas es que, a diferencia de la aproximación ISO, esta estrategia es sencilla y fácilmente aplicable. No obstante, el hecho de que la nueva aproximación calcule la incertidumbre por bloques hace que no sea tan inmediato el identificar las etapas críticas del método analítico y que, por tanto, sea más complicado disminuir la incertidumbre de los resultados en el caso de que fuera necesario. La aproximación ISO permite identificar las etapas críticas del procedimiento ya que todas las fuentes de incertidumbre del procedimiento analítico deben cuantificarse individualmente.

El fundamento teórico de la nueva aproximación es consistente con la aproximación ISO ya que se cuantifican todas las fuentes de incertidumbre y se combinan siguiendo la ley de propagación de errores. La diferencia entre ambas aproximaciones es que la ISO cuantifica individualmente cada uno de los componentes de incertidumbre del procedimiento analítico mientras que la nueva aproximación calcula la incertidumbre globalmente. Recientemente, Eurachem<sup>35</sup> ha incluido esta estrategia de cálculo global de la incertidumbre en su adaptación de la aproximación ISO para medidas químicas. Por tanto, esta nueva estrategia se acerca más a la nueva aproximación propuesta en este capítulo.

La nueva aproximación es conceptualmente muy similar a la aproximación propuesta por el *Analytical Methods Committee* (AMC)<sup>36</sup> ya que ambas están basadas en el cálculo global de la incertidumbre. Sin embargo, a diferencia de la estrategia

del AMC, la incertidumbre calculada tiene qué ver con la del laboratorio individual ya que se ha obtenido utilizando la información generada por dicho laboratorio y no por otros laboratorios que han participado en un ejercicio colaborativo.

## 2.6. Conclusiones

En este capítulo se ha presentado una aproximación para calcular globalmente la incertidumbre que está basada en utilizar la información generada al verificar la trazabilidad de los resultados. Además, la información generada en los estudios de precisión y robustez y en los gráficos de control también proporciona información útil para calcular la incertidumbre. Esta aproximación es sencilla, fácilmente aplicable y consistente con la aproximación ISO. Además, requiere poco esfuerzo adicional ya que los métodos siempre deben validarse antes de analizar las muestras de rutina.

La aproximación propuesta es válida para resultados obtenidos con métodos de rutina con un intervalo restringido de concentraciones que sea similar a la concentración de la muestra de referencia utilizada en la verificación de la exactitud del método. Por tanto, no es aplicable en el caso de métodos donde el intervalo de aplicación es de más de un orden de magnitud. En el próximo capítulo se desarrollará una estrategia basada en utilizar la información generada en estudios de recuperación para calcular la incertidumbre de resultados obtenidos con métodos cuyo intervalo de concentraciones es más amplio.

## 2.7. Referencias

1. UNE-EN ISO 9000, *Sistemas de gestión de la calidad. Fundamentos y vocabulario*, ISO (2000)
2. UNE-EN ISO 9001, *Sistemas de gestión de la calidad. Requisitos*, ISO (2000)
3. UNE-EN ISO 9004, *Sistemas de gestión de la calidad. Directrices para la mejora del desempeño*, ISO (2000)
4. A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Trends in Analytical Chemistry*, **18** (1999) 577-584

5. A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Química Analítica*, **19** (2000) 85-94
6. M. Thompson, *Analyst*, **125** (2000) 2020-2025
7. F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309-316
8. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva (1993)
9. International Union of Pure and Applied Chemistry, *Harmonised Guidelines for In-House Validation of Methods of Analysis (Technical report)*, IUPAC (1999)  
[http://www.iupac.org/projects/1997/501\\_8\\_97.html](http://www.iupac.org/projects/1997/501_8_97.html)
10. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-2, ISO, Geneva (1994)
11. Analytical Methods Committee, *Analyst*, **114** (1989) 1693-1702
12. M. Thompson, M.H. Ramsey, *Analyst*, **120** (1995) 261-270
13. M.H. Ramsey, A. Argyraki, M. Thompson, *Analyst*, **120** (1995) 2309-2312
14. M.H. Ramsey, *Analyst*, **122** (1997) 1255-1260
15. J.L. Pau, R. Boqué, F.X. Rius, *Journal of AOAC International*, enviado
16. G. Stringari, I. Pancheri, F. Möller, O. Failla, *Accreditation and Quality Assurance*, **3** (1998) 122-126
17. A. Henrion, *Accreditation and Quality Assurance*, **3** (1998) 122-126
18. A.S. Johanson, T.B. Whitaker, W.M. Hagler, F.G. Giesbretch, J.H. Young, D.T. Bowman, *Journal of AOAC International*, **83** (2000) 1264-1269
19. R. Heikka, P. Minkkinen, *Analytica Chimica Acta*, **346** (1997) 277-283
20. R. Heikka, P. Minkkinen, *Chemometrics and Intelligent Laboratory Systems*, **44** (1998) 205-211
21. P.M. Gy, *Trends in Analytical Chemistry*, **14** (1995) 67-76
22. M. Ramsey, A. Argyraki, *The Science of the Total Environment*, **198** (1997) 243-257
23. M. Ramsey, A. Argyraki, *Analyst*, **120** (1995) 1353-1356
24. M. Ramsey, A. Argyraki, *Analyst*, **120** (1995) 2309-2312
25. H. Scheffé, *Analysis of variance*, Wiley, Nueva York (1953)
26. V.J. Barwick, S.L.R. Ellison, *Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data*, VAM Project 3.2.1 (2000)
27. V.J. Barwick, S.L.R. Ellison, *Accreditation and Quality Assurance*, **5** (2000) 47-53
28. V.J. Barwick, S.L.R. Ellison, M.J.Q. Rafferty, R.S. Gill, *Accreditation and Quality Assurance*, **5** (2000) 104-113
29. ICH, *Harmonised Tripartite Guideline prepared within the Third International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use*, Text on Validation of Analytical Procedures (1994)  
<http://www.ifpma.org/ich1.html>
30. W.J. Youden, E.H. Steiner, *Statistical Manual of the Association of Official Analytical Chemists*, AOAC, Arlington, VA (1975)

31. C.F. Dietrich, *Uncertainty, Calibration and Probability. The Statistic of Scientific and Industrial Measurement*, 2<sup>nd</sup> Ed. Adam Hilger, Bristol, (1991)
32. B.D. Hall, R. Willink, *Metrologia*, **38** (2001) 9-15
33. J.R. Cowles, S. Daily, S.L.R. Ellison, W.A. Hardcastle, C. Williams, *Accreditation and Quality Assurance*, **6** (2001) 368-371
34. EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK (1995)
35. EURACHEM/CITAC, *Quantifying uncertainty in analytical measurement*, EURACHEM/CITAC Guide, 2<sup>nd</sup> Ed. (2000)
36. Analytical Methods Committee, *Analyst*, **120** (1995) 2303-2308







## **CAPÍTULO 3**

---

***Cálculo de la incertidumbre en métodos analíticos de rutina validados en un intervalo amplio de concentraciones utilizando muestras adicionadas***



### 3.1. Introducción

En este capítulo se presenta una metodología para calcular la incertidumbre de resultados obtenidos con métodos de rutina que se aplican en un intervalo amplio de concentraciones. Esta metodología es conceptualmente similar a la presentada en el capítulo anterior ya que está basada en calcular la incertidumbre globalmente utilizando la información generada en la validación del método y, especialmente, la información generada en el proceso de verificación de la trazabilidad.

Verificar la exactitud de un método implica calcular su sesgo. Un método analítico puede tener dos tipos de sesgo: un sesgo proporcional (que varía con la concentración y se expresa en términos de recuperación) y un sesgo constante. En el capítulo anterior se ha descrito la verificación de la trazabilidad asumiendo que el sesgo era constante ya que el método de rutina se aplicaba en un intervalo restringido de concentraciones. Sin embargo, esto no es generalizable cuando el método se aplica en un intervalo amplio de concentraciones. En este caso, la trazabilidad de los resultados debe verificarse calculando tanto el sesgo proporcional como el constante. Para ello, es necesario utilizar muestras de referencia a varios niveles de concentración que sean representativas de las muestras de rutina. Salvo en algunas excepciones, como la fluorescencia de rayos X (XRF) donde las referencias utilizadas suelen ser materiales de referencia, normalmente se verifica la trazabilidad utilizando muestras adicionadas ya que no se suele disponer de referencias con un nivel de trazabilidad mayor. El inconveniente de las muestras adicionadas es que el comportamiento del analito adicionado puede ser diferente del del analito nativo<sup>1,2</sup> más aún en el caso de que las muestras no sean líquidas o bien en métodos analíticos donde no haya una total digestión de la muestra.

La verificación de la trazabilidad utilizando muestras adicionadas puede variar dependiendo de que las muestras adicionadas contengan una concentración inicial de analito o bien de que se disponga de blancos de muestras. Si se dispone de blancos de muestras puede calcularse tanto el sesgo proporcional como el constante utilizando muestras adicionadas. Sin embargo, si las muestras tienen una concentración inicial de analito únicamente puede calcularse el sesgo proporcional

del método. Por tanto, en este caso es necesario utilizar el método de Youden<sup>3</sup> para verificar la trazabilidad de los resultados ya que el método puede tener una recuperación del 100% y, aún así, puede tener un sesgo constante<sup>4</sup>. El método de Youden consiste en analizar diferentes cantidades de una muestra. Si el método no tiene un sesgo constante se deberían obtener los mismos resultados independientemente de la cantidad de muestra analizada.

Por otro lado, el cálculo de la recuperación del método analítico también depende de que las adiciones se hagan a muestras con concentración inicial de analito o bien a muestras que no contienen el analito. Si las muestras adicionadas tienen concentración inicial de analito, la recuperación puede calcularse utilizando el método de adiciones estándar o bien el método de la recuperación media. Sin embargo, si las muestras adicionadas no contienen el analito, la recuperación sólo puede calcularse con el método de adiciones estándar.

Una vez verificada la trazabilidad, puede calcularse la incertidumbre de los resultados obtenidos con dicho método. En el capítulo anterior se ha explicado que la incertidumbre debe incluir al menos los componentes asociados a verificar la trazabilidad de los resultados y a la precisión del método. Además, también puede ser necesario incluir otros términos como, por ejemplo, los debidos al submuestreo y a los pretratamientos. Para calcular la incertidumbre debida a la precisión del método es necesario estudiar cómo varía la precisión con la concentración. En el capítulo anterior se ha asumido que la precisión era constante ya que el intervalo de aplicación del método era restringido. Sin embargo, cuando la incertidumbre se calcula en un intervalo amplio de concentraciones debe estudiarse cómo varía la precisión del método en función de la concentración. En este capítulo se ha calculado la incertidumbre asumiendo que la precisión era aproximadamente constante en el intervalo de aplicación del método o bien que variaba linealmente con la concentración (es decir, asumiendo una desviación estándar relativa constante). Cabe decir que la relación entre la precisión (expresada como desviación estándar) y la concentración puede seguir otro tipo de modelos<sup>5-8</sup> los cuales no se han considerado en este capítulo aunque es generalizable.

En el apartado 3.2. de este capítulo se explica cómo verificar la trazabilidad utilizando muestras adicionadas con o sin concentración inicial de analito. En el apartado 3.3. se explica cómo calcular la incertidumbre en un intervalo amplio de concentraciones utilizando la información obtenida al verificar la trazabilidad de los resultados con muestras adicionadas. El cálculo de la incertidumbre depende de cómo se verifica la trazabilidad y de cómo varía la precisión con la concentración. En el apartado 3.4 se presenta el artículo “*Measurement uncertainty in analytical methods in which trueness is assessed from recovery assays*” publicado en la revista *Analytica Chimica Acta*. En este artículo se detalla cómo calcular la incertidumbre cuando la recuperación se calcula con el método de la recuperación media y la precisión varía linealmente con la concentración. En el apartado 3.5 se presenta el artículo “*Estimation of measurement uncertainty by using regression techniques and spiked samples*” publicado en la revista *Analytica Chimica Acta*. En este artículo se explica cómo calcular la incertidumbre cuando la trazabilidad de los resultados se verifica utilizando técnicas de regresión y la precisión es aproximadamente constante en el intervalo de aplicación del método. En concreto, se proponen dos alternativas para verificar la trazabilidad de los resultados. La primera de ellas consiste en expresar los resultados de la recta de adiciones estándar y de la recta de Youden en términos de respuesta instrumental mientras que la segunda consiste en expresarlos en términos de concentración. Posteriormente, se detallan expresiones para calcular la incertidumbre para cada una de las dos alternativas propuestas para verificar la trazabilidad de los resultados.

En el apartado 3.6 se explica cómo calcular la incertidumbre cuando la trazabilidad se verifica utilizando blancos de muestras adicionadas. Esta metodología también es válida cuando la trazabilidad se verifica utilizando otras referencias (materiales de referencia, métodos de referencia, etc.) siempre que la incertidumbre de los valores de referencia sea despreciable. En este apartado se explica cómo calcular la incertidumbre utilizando la información de la recta de adiciones estándar cuando la precisión se mantiene aproximadamente constante con la concentración y cuando varía linealmente con la concentración. Finalmente, en el apartado 3.7 se exponen las conclusiones de este capítulo y en el apartado 3.8 la bibliografía.

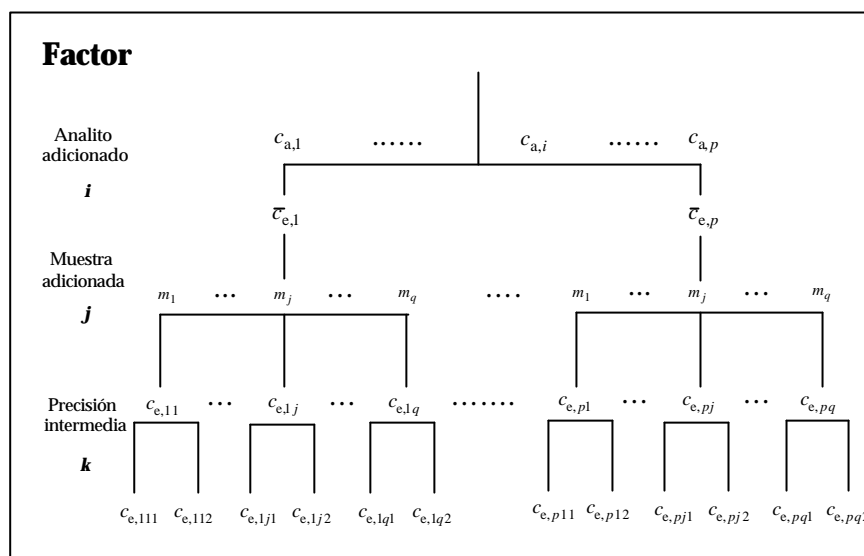
## **3.2. Aseguramiento de la trazabilidad de los resultados utilizando muestras adicionadas**

Para verificar la trazabilidad en un intervalo amplio de concentración es necesario evaluar dos tipos de sesgo: el sesgo proporcional (en términos de recuperación) y el sesgo constante. El cálculo de ambos sesgos depende de que las adiciones se hagan a blancos de muestras o bien a muestras que contienen el analito. Si las adiciones se hacen a blancos de muestras, se puede calcular tanto el sesgo proporcional como el sesgo constante a partir de la recta de adiciones estándar. Si las adiciones se hacen a muestras que contienen el analito, únicamente se puede calcular el sesgo proporcional a partir de las muestras adicionadas. El sesgo constante debe calcularse mediante el método de Youden. A continuación se explica cómo calcular ambos sesgos.

### **3.2.1. Cálculo del sesgo proporcional. Estudios de recuperación**

El sesgo proporcional se evalúa en términos de recuperación. La IUPAC define la recuperación como “la proporción de la cantidad de analito, presente o añadida a una muestra, que se extrae y se presenta para su medida”<sup>1</sup>. Actualmente, el término de recuperación se utiliza con dos usos diferenciados: a) Para expresar la proporción de analito recuperada en una etapa de preconcentración o de extracción del método analítico; o bien b) para expresar la relación entre analito encontrado con el método analítico frente al valor de referencia. El primer uso de la recuperación debería distinguirse claramente del segundo ya que tener una recuperación del 100% no implica recuperar el 100% de analito en cada una de las etapas de preconcentración o extracción del método. Para distinguir entre ambos conceptos la IUPAC ha recomendado en un documento preliminar<sup>9</sup> distinguir entre ambos tipos de recuperación. Para el uso (a) recomienda utilizar el término *recuperación* mientras que para el uso (b) recomienda utilizar el término *recuperación aparente*. Por tanto, según la nomenclatura propuesta por la IUPAC, el sesgo proporcional no se evalúa en términos de recuperación sino que se evalúa en términos de recuperación aparente. Actualmente, no existen protocolos bien

definidos para calcular y aplicar la recuperación aparente a los resultados. Corregir los resultados depende sobre todo del área de aplicación<sup>1</sup>. Sin embargo, es importante subrayar que los resultados deben ser exactos y comparables. Por tanto, los resultados deberían corregirse por la recuperación aparente en el caso de que ésta sea significativamente diferente del 100%<sup>10</sup>. Además, la incertidumbre de la recuperación aparente debe incluirse en la incertidumbre final de los resultados para asegurar que los resultados analíticos sean comparables.



**Figura 3.1.** Diseño anidado 3 factores para calcular la recuperación utilizando  $n$  muestras adicionadas. El analito se adiciona a  $p$  niveles de concentración. En cada nivel de concentración se adiciona a  $q$  tipos de muestras diferentes. Cada una de las muestras adicionadas debe analizarse dos veces en condiciones intermedias.  $c_{e,1j1}$  es la concentración encontrada en el replicado 1 al adicionar una cantidad de analito  $c_{a,1}$  a un tipo de muestra  $m_j$ .  $c_{e,1j}$  es el valor medio de los replicados  $c_{e,1j1}$  y  $c_{e,1j2}$ .

Para evaluar el sesgo proporcional del método es necesario calcular la recuperación (o recuperación aparente según la IUPAC) variando representativamente la concentración adicionada y el intervalo de matrices de las muestras de rutina. Para ello, pueden utilizarse diseños anidados de tres factores (ver Fig. 3.1). Según este diseño, el analito se adiciona a diferentes tipos de muestras que son representativas de las matrices de las muestras de rutina. Esta adición debe hacerse preferiblemente al menos a tres niveles de concentración (en

los extremos y en el valor medio del intervalo de concentraciones de las muestras de rutina). Posteriormente, se debería analizar cada muestra adicionada dos veces en condiciones intermedias. Este diseño anidado permite obtener información útil para el cálculo de la incertidumbre ya que a partir del ANOVA puede obtenerse la incertidumbre debida a la variabilidad de las muestras de rutina e información sobre cómo varía la precisión con la concentración. En este apartado se estudia cómo calcular la recuperación aparente cuando las referencias utilizadas son muestras adicionadas con concentración inicial de analito o bien cuando se dispone de blancos de muestras.

Una de las limitaciones de la técnica de adiciones estándar es que se puede obtener una estimación incorrecta del sesgo proporcional del método si se produce un efecto interactivo entre la matriz y el analito<sup>11-14</sup>. Este efecto provoca que la señal analítica por parte del analito presente en la muestra no sólo dependa de la cantidad del mismo, sino que también dependa de la cantidad de matriz presente en la muestra. Para poder calcular correctamente el sesgo proporcional debe llegarse a la “saturación del efecto interactivo”<sup>11-15</sup>. Esta situación se alcanza a partir de una cantidad determinada de matriz y consiste en que el efecto interactivo es constante y, por consiguiente, ya no depende de la relación entre la cantidad de analito y matriz. Por tanto, para poder obtener una buena estimación del sesgo proporcional debería comprobarse previamente que se ha alcanzado la situación de saturación del efecto interactivo. Para ello, debe comprobarse la linealidad de la recta de adiciones estándar<sup>12,13</sup>. Esta alternativa es válida siempre y cuando el rango de analito adicionado sea suficientemente amplio. Otra alternativa es evaluar con un test *t* si hay diferencias significativas entre las pendientes de dos rectas de adiciones estándar construidas adicionando analito a dos cantidades diferentes de muestra<sup>12,13</sup>. Si ambas pendientes no son significativamente diferentes es que se ha alcanzado la situación de saturación del efecto interactivo. En el caso de que se quieran comparar las pendientes de tres o más rectas de adiciones estándar debe utilizarse el análisis de la covarianza (ANCOVA)<sup>15,16</sup>.



**3.2.1.1. Las muestras adicionadas tienen una concentración inicial de analito**

La concentración encontrada,  $c_e$ , de cada una de las muestras adicionadas se calcula como:

$$c_e = c_{0+a} - c_0 \quad (3.1)$$

donde  $c_0$  es la concentración inicial de analito en la muestra y  $c_{0+a}$  es la concentración de la muestra adicionada.

Una vez que se ha obtenido la concentración encontrada de cada una de las muestras adicionadas, puede calcularse la recuperación del método. Cuando las muestras adicionadas contienen una concentración inicial de analito, la recuperación del método puede calcularse utilizando el método de la recuperación media o bien el método de las adiciones estándar.

Método de la recuperación media

Este método consiste en calcular la recuperación del método como el valor medio de las recuperaciones obtenidas para cada uno de los tipos de muestras adicionadas. Siguiendo el diseño de la Fig. 3.1, la recuperación media se calcula como:

$$\bar{R} = \frac{\sum_{i=1}^p \sum_{j=1}^q R_{ij}}{n} \quad (3.2)$$

donde  $R_{ij}$  es la recuperación obtenida para la muestra  $j$  adicionada con una cantidad  $i$  de analito y se calcula como el cociente entre la concentración encontrada y la concentración adicionada (es decir,  $R_{ij} = c_{e,ij} / c_{a,ij}$ ).  $n$  es el número total de muestras adicionadas analizadas (es decir,  $n = p \cdot q$ ). La incertidumbre de la recuperación media puede calcularse a partir de la desviación estándar,  $s(R)$ , de todas las recuperaciones,  $R_{ij}$ , obtenidas para cada una de las muestras adicionadas:

$$u(R) = \frac{s(R)}{\sqrt{n}} \quad (3.3)$$

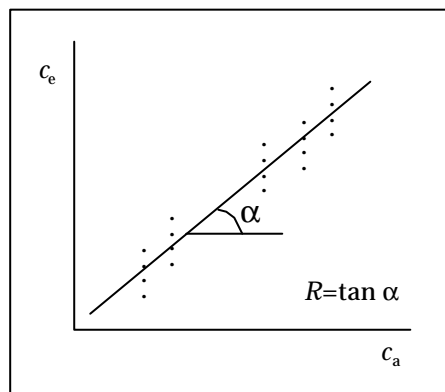
Posteriormente, debe verificarse si el sesgo proporcional del método es significativo o no. La recuperación media no difiere significativamente de 1 si se cumple la siguiente inecuación:

$$|\bar{R} - 1| \leq t_{\alpha/2, n-1} \cdot u(R) \quad (3.4)$$

donde  $t_{\alpha/2, n-1}$  es el valor de  $t$  tabulado de dos colas para un nivel de significancia  $\alpha$  y  $n-1$  grados de libertad.

Método de las adiciones estándar

Una vez que se ha obtenido la concentración encontrada para cada una de las muestras adicionadas (ver Fig. 3.1), puede obtenerse la recta de adiciones estándar representando la concentración encontrada frente a la concentración adicionada (ver Fig. 3.2). La pendiente de la recta de adiciones estándar proporciona una estimación de la recuperación del método.



**Figura 3.2.** Recta de adiciones estándar.  $c_e$  es la concentración encontrada y  $c_a$  la concentración adicionada. La pendiente de la recta corresponde a la recuperación,  $R$ , del método.

Para obtener la pendiente de la recta deben utilizarse técnicas de regresión. La técnica de regresión utilizada depende de que la precisión se mantenga

aproximadamente constante o bien varíe en el intervalo de concentraciones del método. Si se mantiene aproximadamente constante se debe utilizar el método de regresión por mínimos cuadrados clásicos (“*Ordinary Least Squares*”, OLS)<sup>17,18</sup> y si varía con la concentración debe utilizarse el método de regresión por mínimos cuadrados ponderados (“*Weighted Least Squares*”, WLS)<sup>17,18</sup>. En ambos casos se obtiene una pendiente,  $b$ , con una desviación estándar asociada,  $s_b$ .

Al igual que en el método de la recuperación media, debe realizarse un test  $t$  para evaluar si el método tiene un sesgo proporcional significativo. La recuperación no difiere significativamente de 1 si se cumple la siguiente inecuación:

$$|\bar{R} - 1| \leq t_{\alpha/2, n-2} \cdot s_b \quad (3.5)$$

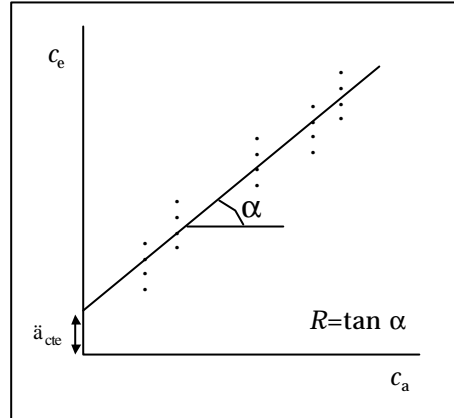
donde  $t_{\alpha/2, n-2}$  es el valor de  $t$  tabulado de dos colas para un nivel de significancia  $\alpha$  y  $n-2$  grados de libertad.

### 3.2.1.2. Las muestras adicionadas no contienen el analito

Cuando el analito se adiciona a blancos de muestras, puede calcularse tanto el sesgo proporcional como el constante a partir de las muestras adicionadas. En este caso, la recuperación del método se calcula utilizando el método de adiciones estándar. El método de la recuperación media no puede aplicarse ya que la recuperación se vería afectada por un posible sesgo constante del método<sup>3</sup>. Por tanto, podría obtenerse una mala estimación de la recuperación. Además, el método de la recuperación media no permite obtener el sesgo constante del método a partir de las muestras adicionadas.

Como las muestras adicionadas no contienen una concentración inicial de analito, la concentración encontrada corresponde a la concentración obtenida al analizar la muestra adicionada. La Fig. 3.3 muestra la recta de adiciones obtenida al representar la concentración encontrada frente a la concentración adicionada. Tal y como muestra la figura, la pendiente de la recta proporciona una estimación de la

recuperación del método y la ordenada en el origen proporciona una estimación del sesgo constante del método.



**Figura 3.3.** Recta de adiciones estándar.  $c_e$  es la concentración encontrada y  $c_a$  la concentración adicionada. La pendiente de la recta corresponde a la recuperación del método,  $R$ , y la ordenada corresponde al sesgo constante del método,  $\delta_{cte}$ .

Para poder obtener la recta de adiciones estándar deben utilizarse técnicas de regresión. Al igual que cuando las muestras tienen una concentración inicial de analito, debe aplicarse el método de regresión por mínimos cuadrados clásicos (OLS) cuando la precisión se mantiene aproximadamente constante y el método de regresión por mínimos cuadrados ponderados (WLS) si la precisión varía con la concentración del método.

Para verificar la trazabilidad de los resultados, debe verificarse que la pendiente de la recta no difiere significativamente de uno y que la ordenada de la recta no difiere significativamente de cero. Esto puede hacerse mediante tests individuales de cada parámetro, aunque es recomendable realizar un test conjunto<sup>17-19</sup> de la pendiente y de la ordenada de la recta ya que, de esta forma, se tiene en cuenta la correlación entre la pendiente y la ordenada de la recta. El valor de  $F$  calculado se obtiene a partir de la Ec. 3.6:

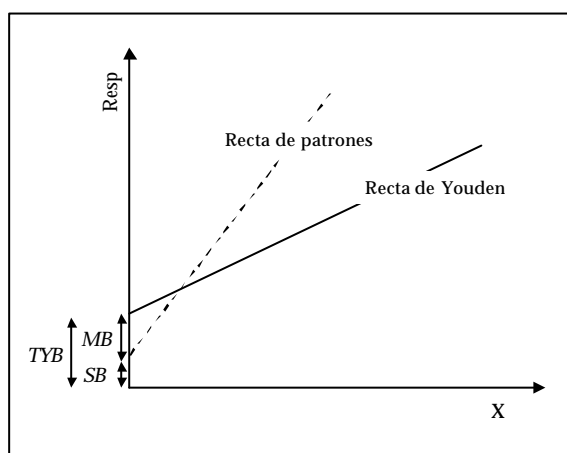
$$F = \frac{\mathbf{d}_{cte}^2 + 2 \cdot \tau_a \cdot \mathbf{d}_{cte} \cdot (R - 1) + \frac{\sum_{i=1}^p \sum_{j=1}^q c_{a,ij}^2}{n} \cdot (1 - R)^2}{2 \cdot s_e^2 / n} \quad (3.6)$$

donde  $c_{a,ij}$  es la concentración adicionada  $i$  al tipo de muestra  $j$ ;  $\tau_a$  es el valor medio de todos los  $c_{a,ij}$ ,  $p$  son los niveles de concentración en que se adiciona el analito;  $q$  el número de matrices a las que se adiciona el analito a cada nivel de concentración;  $n$  es el número de puntos de la recta de adiciones ( $n=p \cdot q$ );  $s_e$  es la desviación estándar de los residuales de la recta;  $R$  es la recuperación y corresponde a la pendiente de la recta; y  $\delta_{cte}$  es el sesgo constante y corresponde a la ordenada en el origen de la recta. El valor de  $F$  calculado debe compararse con el de  $F$  tabulado de una distribución de 2 y  $n-2$  grados de libertad. Si el valor de  $F$  calculado es menor que el  $F$  tabulado, el sesgo constante y proporcional del método no son significativos. Por tanto, los resultados obtenidos con el método son trazables. Si el valor de  $F$  calculado es mayor que el de  $F$  tabulado, los resultados obtenidos con el método no son trazables. Por tanto, debe revisarse el método o bien corregir los resultados por la recuperación y el sesgo constante obtenidos con la recta de adiciones estándar.

### 3.2.2. Cálculo del sesgo constante. Método de Youden

Cuando las muestras adicionadas contienen una concentración inicial de analito, no se puede verificar la ausencia de sesgo constante a partir de los estudios de recuperación. Por tanto, es necesario utilizar el método de Youden<sup>3,12,13,15,20-25</sup> para verificar la ausencia de sesgo constante. Este método consiste en analizar diferentes cantidades (masas o volúmenes) de una muestra. Por ejemplo, si se quiere evaluar la presencia de un sesgo constante en la determinación de hierro (III) en aguas podrían añadirse volúmenes de 0, 5, 10, 25 y 50 ml de una muestra de agua en un matraz de 100 ml y enrasar el matraz con agua desionizada<sup>20</sup>. Estas cantidades de muestra deberían analizarse en condiciones intermedias de precisión. Posteriormente, la recta de Youden puede obtenerse representando la respuesta instrumental obtenida frente al volumen analizado o bien representando la

concentración obtenida frente al volumen analizado. La Fig. 3.4 muestra la relación entre la recta de calibrado y la recta de Youden obtenida al representar la respuesta instrumental frente al volumen de muestra<sup>20</sup>. La ordenada de la recta de Youden corresponde al blanco total de Youden (“Total Youden Blank”, TYB). En esta ordenada quedan incluidas todas las contribuciones constantes de la señal analítica, es decir, todas aquellas contribuciones que no varían con la cantidad de muestra analizada y que, por el hecho de haber estado medidas en presencia del analito y de la matriz, proporciona una estimación del verdadero blanco total del procedimiento<sup>23</sup>.



**Figura 3.4.** Representación conjunta de la recta de patrones y de la recta de Youden obtenida al representar la cantidad de muestra analizada frente a la respuesta instrumental. En el eje X se representa bien concentración de analito para la recta de patrones o cantidad de muestra para la recta de Youden.

La Fig. 3.4 muestra que el blanco total de Youden, *TYB*, se puede ver como la suma de dos contribuciones: el blanco instrumental (“System Blank”, *SB*) que corresponde a la ordenada de la recta de patrones y el blanco del método (“Method Blank”, *MB*)<sup>23</sup> que está relacionado con el sesgo constante del método. Dicho sesgo puede obtenerse con la siguiente expresión:

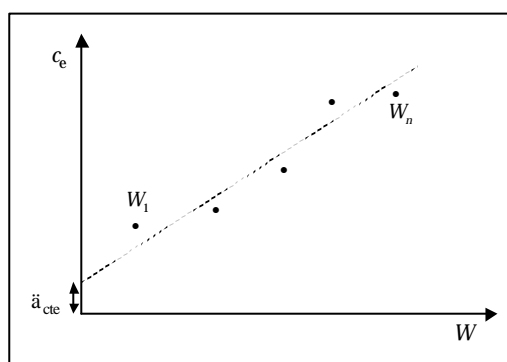
$$d_{cte} = \frac{TYB - SB}{b} \quad (3.7)$$

donde  $b$  corresponde a la pendiente de la recta de patrones. La incertidumbre del sesgo constante se obtiene aplicando la ley de propagación de errores a la Ec. 3.7:

$$u_{d_{cte}} = \frac{1}{b} \sqrt{s_{TYB}^2 + s_{SB}^2 + d_{cte} \cdot s_b^2 - 2 \cdot d_{cte} \cdot \text{cov}(SB, b)} \quad (3.8)$$

donde  $s_{TYB}$  es la desviación estándar de la ordenada de la recta de Youden,  $s_{SB}$  es la desviación estándar de la ordenada de la recta de patrones,  $s_b$  es la desviación estándar de la pendiente de la recta de patrones y  $\text{cov}(SB, b)$  es la covarianza entre la pendiente y la ordenada de la recta de patrones.

El sesgo constante del método también puede estimarse representando la recta de Youden como la concentración obtenida frente a la cantidad de muestra analizada. En este caso, la ordenada de la recta de Youden proporciona una estimación del sesgo constante (ver Fig. 3.5).



**Figura 3.5.** Recta de Youden donde se representa la cantidad de muestra analizada,  $W$ , frente a la concentración obtenida. La ordenada de la recta corresponde a una estimación del sesgo constante,  $\delta_{cte}$ .

El método de Youden proporciona una buena estimación del sesgo constante si se ha alcanzado la saturación del efecto interactivo<sup>12,13,15,25</sup> que se produce cuando el efecto matriz es el mismo independientemente de la cantidad de muestra analizada. Por tanto, debe comprobarse si se ha alcanzado la saturación del efecto interactivo comprobando con un test  $t$  que no hay diferencias significativas entre la

pendiente de la recta de adiciones estándar obtenida al adicionar analito a la cantidad mínima analizada ( $W_1$  en la Fig. 3.5) y la obtenida al adicionar analito a la cantidad máxima analizada ( $W_n$  en la Fig. 3.5).

Para obtener la ordenada de la recta de Youden deben utilizarse técnicas de regresión. La técnica de regresión utilizada depende de que la precisión se mantenga aproximadamente constante o bien varíe en el intervalo de concentraciones del método. Si se mantiene aproximadamente constante se debe utilizar el método de regresión por mínimos cuadrados clásicos (“*Ordinary Least Squares*”, OLS) y si varía con la concentración debe utilizarse el método de regresión por mínimos cuadrados ponderados (“*Weighted Least Squares*”, WLS). En ambos casos se obtiene una ordenada,  $a$ , con una desviación estándar asociada,  $s_a$ .

Una forma más sencilla de aplicar el método de Youden es analizando únicamente dos cantidades de muestra<sup>26</sup>. Si no hay un sesgo constante, debería obtenerse la misma proporción de analito al analizar ambas cantidades de muestra. Por tanto, el sesgo constante,  $\delta_{cte}$ , puede calcularse a partir de la siguiente expresión:

$$\frac{\bar{x}_n - \mathbf{d}_{cte}}{W_n} = \frac{\bar{x}_m - \mathbf{d}_{cte}}{W_m} \quad (3.9)$$

donde  $\bar{x}_n$  y  $\bar{x}_m$  son, respectivamente, el valor medio de los resultados obtenidos al analizar las cantidades de muestra,  $W_n$  y  $W_m$ . Esta aplicación del método de Youden se ha utilizado en el apartado 3.4. En este caso la incertidumbre debe obtenerse aplicando la ley de propagación de errores (ver Ec. 16 del apartado 3.4).

Una vez que se ha calculado el sesgo constante del método, debe realizarse un test  $t$  para evaluar si el método tiene un sesgo constante significativo. Este sesgo no es significativo si se cumple la siguiente inecuación:

$$|\mathbf{d}_{cte}| \leq t_{\alpha/2, \nu} \cdot u_{\mathbf{d}_{cte}} \quad (3.10)$$

donde  $t_{\alpha/2, \nu}$  es el valor de  $t$  tabulado de dos colas para un nivel de significancia  $\alpha$  y  $\nu$  grados de libertad. La incertidumbre del sesgo constante y los grados de libertad



dependen de cómo se haya obtenido la recta de Youden. Si se ha obtenido como respuesta instrumental frente a cantidad de muestra, la incertidumbre se obtiene con la Ec. 3.8 y los grados de libertad se obtienen con la aproximación de Welch-Satterthwaite<sup>27</sup>. Si se ha obtenido como concentración encontrada frente a cantidad de muestra, la incertidumbre corresponde a la desviación estándar de la ordenada de la recta de Youden y los grados de libertad corresponden a  $n-2$  (siendo  $n$  el número de cantidades de muestra,  $W$ , analizadas en la recta de Youden). Si el sesgo constante es significativo, el método debe revisarse o bien deben corregirse los resultados por el sesgo constante.

### 3.3. Cálculo de la incertidumbre en métodos de rutina validados a diferentes niveles de concentración

La incertidumbre de muestras de rutina se puede calcular utilizando la información generada al verificar la trazabilidad a varios niveles de concentración. La forma más usual de verificar la trazabilidad a varios niveles es utilizando como referencias muestras adicionadas. No obstante, también pueden utilizarse otras referencias con un nivel de trazabilidad mayor como los CRM o los métodos de referencia. En el capítulo anterior se ha detallado cómo calcular la incertidumbre en un intervalo restringido de concentraciones, es decir, cuando se asume que la incertidumbre no varía en el intervalo de aplicación del método. Para ello, debía expresarse el resultado,  $c$ , como la suma de:

$$c = \mu + \delta_{\text{otros}} + \delta_{\text{traz}} + \delta_{\text{pret}} + \delta_{\text{serie}} + \epsilon \quad (3.11)$$

donde  $\epsilon$  es el error aleatorio dentro de una misma serie,  $\delta_{\text{serie}}$  es el sesgo de la serie,  $\delta_{\text{pret}}$  es el sesgo debido a la heterogeneidad de la muestra y a tratamientos previos,  $\delta_{\text{traz}}$  es el sesgo del método,  $\delta_{\text{otros}}$  es el sesgo asociado a otros términos que no se han considerado previamente y  $\mu$  es el valor verdadero. Aplicando la ley de propagación de errores a la Ec. 3.11 se obtenía en la Fig. 2.3 la expresión para calcular globalmente la incertidumbre:

$$u = \sqrt{u_{\text{proc}}^2 + u_{\text{traz}}^2 + u_{\text{pret}}^2 + u_{\text{otros}}^2} \quad (3.12)$$

donde  $u_{\text{proc}}$  es la incertidumbre del procedimiento y está asociada a la precisión intermedia del método;  $u_{\text{traz}}$  es la incertidumbre debida a verificar la trazabilidad de los resultados;  $u_{\text{pret}}$  es la incertidumbre debida a la heterogeneidad y a pretratamientos; y  $u_{\text{otros}}$  es la incertidumbre de otros términos.

Para calcular la incertidumbre en un intervalo amplio de concentraciones es necesario expresar el resultado,  $c$ , de una muestra de rutina asumiendo que el sesgo del método,  $\delta_{\text{raz}}$ , corresponde a la suma de dos sesgos: un sesgo constante,  $\delta_{\text{cte}}$ , y un sesgo proporcional,  $\delta_{\text{prop}}$ .

$$c = \hat{\mu} + \hat{a}_{\text{otros}} + \hat{a}_{\text{prop}} + \hat{a}_{\text{cte}} + \hat{a}_{\text{pret}} + \hat{a}_{\text{serie}} + \hat{a} \quad (3.13)$$

El sesgo proporcional,  $\delta_{\text{prop}}$ , puede expresarse en términos de recuperación,  $R$ :

$$R = \frac{c_e}{c_a} \approx \frac{\hat{\mu} + \hat{a}_{\text{prop}}}{\hat{\mu}} \quad (3.14)$$

La concentración adicionada,  $c_a$ , es una estimación de la concentración verdadera añadida a la muestra,  $\mu$ , y la concentración encontrada,  $c_e$ , corresponde a una estimación de la suma de  $\mu$  y del sesgo proporcional. Por tanto, la recuperación y el sesgo proporcional están relacionados por la siguiente expresión:

$$R \hat{\mu} = \hat{\mu} + \hat{a}_{\text{prop}} \quad (3.15)$$

Del mismo modo, el sesgo de los pretratamientos,  $\delta_{\text{pret}}$ , puede expresarse en términos de un factor de homogeneidad,  $F_{\text{hom}}^{28}$  que considera el error asociado al submuestreo y a los pretratamientos. Este factor se expresa como:

$$F_{\text{hom}} = \frac{\hat{\mu}}{c_{\text{pret}}} \quad (3.16)$$

donde  $\mu$  es la concentración verdadera de toda la muestra y  $c_{\text{pret}}$  es la concentración verdadera de la porción de muestra submuestreada para el análisis. Esta concentración puede expresarse como la suma de la concentración verdadera y del sesgo de los pretratamientos:  $c_{\text{pret}} = \mu + \delta_{\text{pret}}$ . Por tanto, el sesgo de los pretratamientos y la concentración verdadera están relacionados por la siguiente expresión:

$$\frac{\hat{c}}{F_{\text{hom}}} = \hat{c} + \hat{a}_{\text{pret}} \quad (3.17)$$

Por tanto, la Ec. 3.13 puede expresarse como:

$$c = \frac{R}{F_{\text{hom}}} \hat{c} + \hat{a}_{\text{otros}} + \hat{a}_{\text{cte}} + \hat{a}_{\text{serie}} + \hat{a} \quad (3.18)$$

Para poder obtener la concentración verdadera,  $\mu$ , se debería poder corregir el resultado por el factor de homogeneidad, por la recuperación, por los diferentes tipos de sesgos y por el error aleatorio:

$$\mu = F_{\text{hom}} \cdot \frac{c - \hat{a}_{\text{otros}} - \hat{a}_{\text{cte}} - \hat{a}_{\text{serie}} - \hat{a}}{R} \quad (3.19)$$

En la práctica sólo se dispone de una estimación de la recuperación y del sesgo constante. Además, el factor de homogeneidad de cada muestra tampoco puede conocerse y se asume que es igual a uno. Sólo puede conocerse la incertidumbre de dicho factor. Del mismo modo, sólo puede conocerse la incertidumbre del sesgo de otros términos, del sesgo de la serie y del error aleatorio. Por tanto, en lugar de obtener la concentración verdadera de la muestra, sólo puede obtenerse una estimación de la concentración verdadera:

$$c_{\text{corr}} = F_{\text{hom}} \cdot \frac{c - \hat{a}_{\text{cte}}}{R} \quad (3.20)$$

Como no se puede obtener la concentración verdadera (lo cual supondría corregir por todos los errores sistemáticos y aleatorios), es necesario calcular la

incertidumbre de la concentración,  $c_{\text{corr}}$ , incluyendo todas las fuentes de error que afectan al resultado (es decir, la incertidumbre asociada a estimar el sesgo constante y la recuperación, la incertidumbre del factor de homogeneidad y la incertidumbre del sesgo de otros términos, de la serie y del error aleatorio). De este modo, se puede asegurar que el intervalo  $c_{\text{corr}} \pm \text{Incertidumbre}$  contiene la concentración verdadera,  $\mu$ , con una cierta probabilidad. Para calcular la incertidumbre, es necesario aplicar la ley de propagación de errores a la Ec. 3.19. La Fig. 3.6 muestra la expresión obtenida para el cálculo de la incertidumbre estándar en un intervalo amplio de concentraciones:

The diagram shows the equation for the standard uncertainty  $u(c_{\text{corr}})$  with arrows indicating the contribution of various uncertainty sources:

$$u(c_{\text{corr}}) = \frac{1}{R} \sqrt{u(\ddot{a}_{\text{serie}})^2 + u(\mathbf{e})^2 + u(\ddot{a}_{\text{ct}})^2 + (c_{\text{corr}} \cdot u(R))^2 + (R \cdot c_{\text{corr}} \cdot u(F_{\text{hom}}))^2 + u(\ddot{a}_{\text{otros}})^2}$$

The sources and their corresponding terms are:

- Procedimiento** ( $u_{\text{proc}}$ ) points to  $u(\mathbf{e})^2$ .
- Verificación de la trazabilidad** ( $u_{\text{traz}}$ ) points to  $u(\ddot{a}_{\text{ct}})^2$ .
- Submuestreo y pretratamientos** ( $u_{\text{pret}}$ ) points to  $(R \cdot c_{\text{corr}} \cdot u(F_{\text{hom}}))^2$ .
- Otros términos** ( $u_{\text{otros}}$ ) points to  $u(\ddot{a}_{\text{otros}})^2$ .

**Figura 3.6.** Cálculo de la incertidumbre en un intervalo amplio de concentraciones y relación con los términos de la Ec. 3.12.

En la Fig. 3.6 se observa que la incertidumbre depende inversamente de la recuperación del método. Por tanto, cuanto menor sea la recuperación del método mayor es la incertidumbre de los resultados. Asimismo, se observa una dependencia lineal de la incertidumbre con la concentración a través de los términos de incertidumbre de la verificación de la trazabilidad,  $u_{\text{traz}}$ , del submuestreo y pretratamientos,  $u_{\text{pret}}$ , y, en muchos casos, de la incertidumbre del procedimiento,  $u_{\text{proc}}$ .

La incertidumbre expandida,  $U(c_{\text{corr}})$ , se obtiene multiplicando la incertidumbre estándar por el factor de cobertura  $k=2$ . De esta forma, se puede asegurar que el intervalo  $c_{\text{corr}} \pm U(c_{\text{corr}})$  contiene la concentración verdadera,  $\mu$ , con aproximadamente un 95% de probabilidad. Sin embargo, si los distintos términos de incertidumbre tienen pocos grados de libertad asociados, es más correcto utilizar el valor de  $t$  tabulado de dos colas para el nivel de significación escogido y

los grados de libertad efectivos calculados con la aproximación de Welch-Satterthwaite<sup>27</sup>.

A continuación se va a detallar el cálculo de los diversos componentes de la Fig. 3.6:

Incertidumbre del procedimiento,  $u_{proc}$

La incertidumbre del procedimiento considera la variabilidad experimental debida al sesgo de la serie y al error aleatorio:

$$u_{proc} = \frac{1}{R} \sqrt{u(\hat{\alpha}_{serie})^2 + u(\hat{\alpha})^2} \quad (3.21)$$

Esta incertidumbre corresponde a la precisión del método en condiciones intermedias y se calcula utilizando la Ec. 2.2:

$$u_{proc} = \frac{1}{R} \sqrt{\frac{S_{serie}^2}{p_s} + \frac{S_r^2}{p_s \cdot n_s}} \quad (3.22)$$

Si únicamente se conoce la precisión intermedia del método se calcula utilizando la Ec. 2.3:

$$u_{proc} = \frac{1}{R} \sqrt{\frac{S_r^2}{p_s}} \quad (3.23)$$

La única diferencia respecto a las Ec. 2.2 y 2.3 es que en las Ec. 3.22 y 3.23 debe utilizarse la precisión del método obtenida para el nivel de concentración de la muestra de rutina. Por tanto, el cálculo de la incertidumbre del procedimiento depende de cómo varíe la precisión con la concentración. Para calcular la incertidumbre del procedimiento puede utilizarse la información obtenida en estudios de precisión o bien la información obtenida al verificar la trazabilidad con estudios de recuperación y muestras adicionadas. En los siguientes apartados de este capítulo se explica cómo calcular esta incertidumbre utilizando la información obtenida a partir de muestras adicionadas. Para ello, se estudia el caso en que la

precisión es aproximadamente constante y el caso en que la desviación estándar relativa del método es aproximadamente constante.

Incertidumbre de la verificación de la trazabilidad,  $u_{\text{traz}}$

La incertidumbre de la verificación de la trazabilidad considera la incertidumbre del sesgo constante y de la recuperación:

$$u_{\text{traz}} = \frac{1}{R} \sqrt{u(\delta_{\text{cte}})^2 + (c_{\text{corr}} \cdot u(R))^2} \quad (3.24)$$

donde  $u(\delta_{\text{cte}})$  corresponde a la incertidumbre del sesgo constante y  $u(R)$  corresponde a la incertidumbre de la recuperación del método. Estos términos se calculan utilizando la información de la verificación de la trazabilidad con muestras adicionales. El cálculo de esta incertidumbre depende de que el analito se adicione a blancos de muestras o bien a muestras que contienen una concentración inicial de analito (ver apartado 3.2). Los apartados 3.4 y 3.5 detallan cómo calcular esta incertidumbre cuando el analito se adiciona a muestras con concentración inicial de analito y el apartado 3.6 detalla cómo calcularla cuando el analito se adiciona a blancos de muestras.

Incertidumbre del submuestreo y pretratamientos,  $u_{\text{pret}}$

Esta incertidumbre debe calcularse en el caso de que las muestras de rutina no sean homogéneas o bien cuando a las muestras de rutina se las somete a pretratamientos que no se han realizado sobre las muestras adicionales.

$$u_{\text{pret}} = c_{\text{corr}} \cdot u(F_{\text{hom}}) \quad (3.25)$$

Para calcular la incertidumbre del factor de homogeneidad,  $u(F_{\text{hom}})$ , debe aplicarse la ley de propagación de errores a la expresión del factor de homogeneidad (Ec. 3.16):

$$u(F_{\text{hom}}) = \frac{F_{\text{hom}} \cdot u(c_{\text{pret}})}{c_{\text{pret}}} \approx \frac{u(c_{\text{pret}})}{c_{\text{pret}}} \quad (3.26)$$

La incertidumbre del factor de homogeneidad se calcula asumiendo que  $F_{\text{hom}}=1$ . Para calcular la incertidumbre del factor de homogeneidad debe seguirse la metodología descrita en el capítulo 2. Para ello, deben analizarse  $q$  porciones de una muestra que sea representativa de las muestras de rutina. Cada una de estas porciones debe analizarse variando todos los factores que afecten al submuestreo y a los pretratamientos. A partir de estos resultados se obtiene la desviación estándar del submuestreo y de los pretratamientos,  $s_{\text{pret}}$ , y el valor medio de la concentración de la muestra,  $\bar{c}$ , que corresponde a una estimación de  $c_{\text{pret}}$ . La incertidumbre  $u(c_{\text{pret}})$  se calcula a partir de la desviación estándar  $s_{\text{pret}}$ :

$$u(c_{\text{pret}}) = \frac{s_{\text{pret}}}{\sqrt{N_p}} \quad (3.27)$$

donde  $N_p$  es el número de porciones analizadas de la muestra de rutina (normalmente,  $N_p=1$ ). Por tanto, la incertidumbre del factor de homogeneidad se calcula como:

$$u(F_{\text{hom}}) = \frac{s_{\text{pret}}}{\bar{c} \cdot \sqrt{N_p}} \quad (3.28)$$

La incertidumbre del factor de homogeneidad también puede calcularse si se analizan dos porciones de varias muestras. De esta forma, si el laboratorio analiza muestras por duplicado puede aprovechar dicha información. Por ejemplo, éste es el caso de los laboratorios de Sanidad Pública que suelen repetir el análisis de aquellas muestras de rutina que proporcionan resultados cercanos al límite establecido. Para calcular la incertidumbre, debe calcularse para cada una de las muestras la desviación estándar,  $s_{\text{pret},i}$ , y el valor medio,  $\bar{c}_i$ , de los resultados obtenidos al analizar las diferentes porciones de la muestra. Posteriormente, debe calcularse la desviación estándar relativa de los pretratamientos,  $RSD_{\text{pret},i}$ , para cada una de las muestras:

$$RSD_{\text{pret},i} = \frac{s_{\text{pret},i}}{\bar{c}_i} \quad (3.29)$$

Si la incertidumbre de los pretratamientos varía linealmente con la concentración, se puede calcular una  $RSD$  ponderada de los pretratamientos,  $RSD_{\text{pret}}$ :

$$RSD_{\text{pret}} = \sqrt{\frac{\sum_{i=1}^p (n_i - 1) \cdot RSD_{\text{pret},i}^2}{\sum_{i=1}^p (n_i - 1)}} \quad (3.30)$$

donde  $RSD_{\text{pret},i}$  es la  $RSD$  obtenida para la muestra  $i$  analizada,  $n_i$  es el número de porciones analizadas de la muestra  $i$  y  $p$  es el número de muestras analizadas por duplicado. Los grados de libertad de  $RSD_{\text{pret}}$  corresponden a  $\sum_{i=1}^p (n_i - 1)$ . La incertidumbre del factor de homogeneidad se calcula entonces como:

$$u(F_{\text{hom}}) = \frac{RSD_{\text{pret}}}{\sqrt{N_p}} \quad (3.31)$$

donde  $N_p$  es el número de porciones analizadas de la muestra de rutina (normalmente,  $N_p=1$ ).

La incertidumbre del factor de homogeneidad calculada en las Ec. 3.28 y 3.31 es debida a la heterogeneidad y a los pretratamientos de la muestra pero también al propio procedimiento analítico. Por tanto, para no sobreestimar la incertidumbre es necesario substrair de la varianza de los pretratamientos,  $s_{\text{pret}}^2$ , la varianza debida a la variabilidad del procedimiento analítico siguiendo la metodología descrita en el apartado 2.4.3.

*Incetidumbre de otros términos.  $u_{\text{otros}}$*

Estos términos están asociados a niveles más altos de trazabilidad, a la variabilidad de las matrices de las muestras de rutina y a factores de variabilidad no contemplados al obtener la precisión intermedia. La incertidumbre de factores de variabilidad no variados representativamente al obtener la precisión intermedia puede obtenerse utilizando los estudios de robustez tal y como se ha descrito en el



capítulo 2. La incertidumbre de la variabilidad de las matrices de las muestras de rutina se puede calcular utilizando la información obtenida al verificar la trazabilidad con muestras adicionadas. Para ello, debe aplicarse el análisis de la varianza a los resultados obtenidos al analizar las muestras adicionadas. El cálculo de esta incertidumbre se detalla en los apartados 3.4, 3.5 y 3.6.

### **3.4. Measurement uncertainty in analytical methods in which trueness is assessed from recovery assays (Analytica Chimica Acta, 440 (2001) 171-184)**

Alicia Maroto, Ricard Boqué, Jordi Riu, F. Xavier Rius

Department of Analytical and Organic Chemistry. Institute of Advanced Studies.  
Rovira i Virgili University of Tarragona.  
Pl. Imperial Tàrraco, 1. 43005 Tarragona, Catalonia. SPAIN.

#### **SUMMARY**

We propose a new procedure for estimating the uncertainty in quantitative routine analysis. This procedure uses the information generated when the trueness of the analytical method is assessed from recovery assays. In this paper, we assess trueness by estimating proportional bias (in terms of recovery) and constant bias separately. The advantage of the procedure is that little extra work needs to be done to estimate the measurement uncertainty associated to routine samples. This uncertainty is considered to be correct whenever the samples used in the recovery assays are representative of the future routine samples (in terms of matrix and analyte concentration). Moreover, these samples should be analysed by varying all the factors that can affect the analytical method. If they are analysed in this fashion, the precision estimates generated in the recovery assays take into account the variability of the routine samples and also all the sources of variability of the analytical method. Other terms related to the sample heterogeneity, sample pretreatments or factors not representatively varied in the recovery assays should only be subsequently included when necessary. The ideas presented are applied to calculate the uncertainty of results obtained when analysing sulphides in wine by HS-SPME-GC.

**Keywords:** Recovery, Uncertainty, Assessment of trueness, Constant bias

## 1. INTRODUCTION

Analytical results must be both reliable and comparable because they are often used as a piece of valuable information for a certain aim. Therefore, analysts are increasingly impelled to validate analytical procedures and to estimate the uncertainty associated to the results these procedures provide. One of the constraints in method validation is the lack of references which have a high level of traceability such as certified reference materials or reference methods of analysis. As a result, the analyst often has to resort to references with a lower level of traceability, such as spiked samples.

One of the most important parameters in the validation of an analytical method is the assessment of trueness. Checking the trueness of an analytical procedure involves estimating its bias. Two types of bias may be present: a proportional bias (which varies with concentration and is expressed as recovery) and a constant bias. When an analytical procedure is to be validated in a range of concentrations, both types of bias must be estimated separately so that the trueness of the results obtained throughout the concentration range can be guaranteed.

Various approaches for estimating recovery and its uncertainty have been proposed. For instance, EURACHEM [1] proposes that recovery should be estimated by using spiked samples of various types. However, the expression EURACHEM proposes for calculating the uncertainty of recovery does not take into account that the variability of recovery can depend on the type of sample analysed. Other approaches, which estimate the recovery in spiked samples at different concentration levels, have also been proposed [2]. These approaches use either the method of the "averaged recovery" or the regression analysis of added analyte versus found analyte to estimate the recovery. However, neither do the approaches proposed in [2] consider the uncertainty which arises from the variability of recovery and which depends on the type of sample analysed. Barwick and Ellison [3] have recently presented an approach for estimating the recovery which includes this variability in the uncertainty budget. However, none of these approaches consider the fact that there may be a constant bias in the analytical method. As a result, the trueness of the results cannot be guaranteed

because the recovery could be 100 % and, even so, a systematic error could be present [4]. Moreover, these approaches do not calculate the uncertainty of the results of future routine samples once the method has been validated.

This paper presents a new procedure so that the information generated in the estimation of recovery (proportional bias) and constant bias of a method of analysis can be used to calculate the uncertainty associated to future routine samples. Maroto *et al.* have already proposed a conceptually similar approach in which the trueness of the analytical method is assessed against a reference material or a reference method at one level of concentration [5,6]. This approach can only be applied to routine samples with characteristics (i.e. concentration level and matrix type) similar to the sample used in the assessment of trueness. If the characteristics are not similar, other terms of uncertainty should be included to take these differences into account. However, this is not necessary in the methodology proposed in this paper because the uncertainty arising from the variability of both the concentration level and the matrix is considered in the uncertainty budget.

Finally, the ideas proposed in this paper are used to calculate the uncertainty of the results of a method for determining sulphides in wine by HS-SPME-GC [7].

## **2. ASSESSMENT OF TRUENESS**

The trueness of an analytical method should always be assessed before it is applied to future routine samples. Assessing trueness implies estimating separately the proportional bias (in terms of recovery) and the constant bias of the analytical method. Moreover, the uncertainty of both biases should also be calculated to see whether the biases calculated are statistically significant or not.

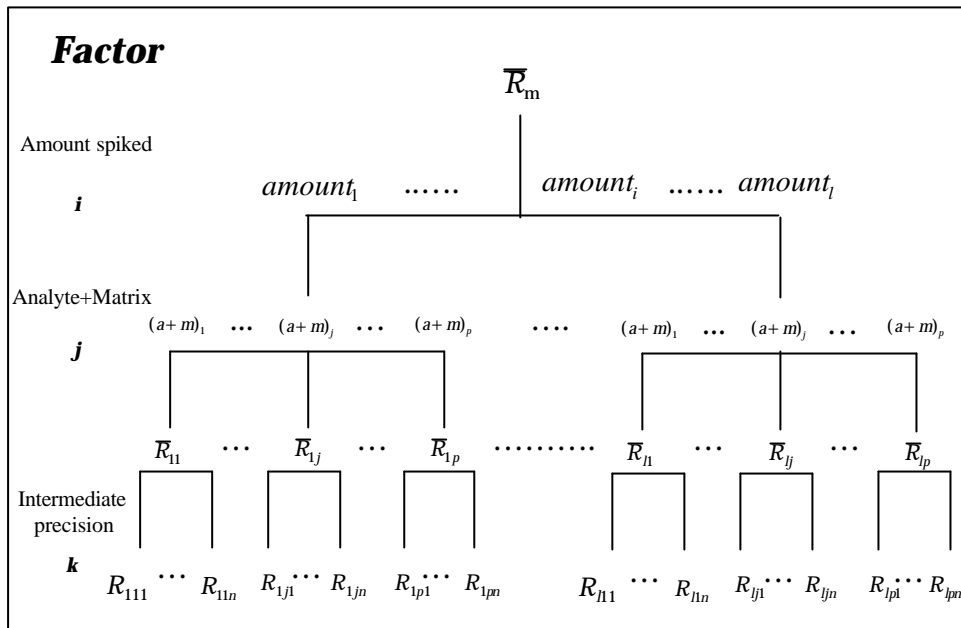
### **2.1. Estimation of recovery**

Proportional bias should be estimated from recovery assays using reference samples which are representative of the routine samples. However, these reference samples are seldom available and often the analyst has to resort to spiked samples.

This has the disadvantage that an incorrect estimation of recovery may be obtained if the spiked analyte has a different behaviour than the native analyte. However, a spike can be a good representation of the native analyte in case of liquid samples and in analytical methods that involve the total dissolution or destruction of the sample. Otherwise, some approaches such as comparing the spiked recovery with the recovery from a less representative reference material, should be used to investigate whether the behaviour of the spiked analyte is different from that of the native analyte [3].

Recovery can be calculated either with the method of the averaged recovery or with regression techniques [2]. In this paper, recovery will be estimated using the averaged recovery method. This method consists of estimating a mean recovery value,  $\bar{R}_m$ , which is the result of estimating the recovery for each amount of analyte (or analytes if a set of similar compounds are analysed with the procedure) added to a set of different matrices which are representative of the future routine samples. The analyte (or analytes) should be spiked in the different representative matrices at two levels of concentration at least, ideally at both extremes of the concentration range. For each amount of analyte spiked, recovery is calculated in each matrix type by analysing the sample before and after the addition of the analyte. The analysis is carried out under intermediate precision conditions, i.e. by varying within the laboratory all the factors that affect the analytical procedure (day, reagents, calibration, analyst, changes in some steps of the analytical procedure, etc.). This recovery should be calculated at least twice so that the intermediate precision of the method and the variability of recovery with the matrix, type of analyte and level of concentration can be estimated. From all these measurements, an overall recovery,  $\bar{R}_m$ , and its standard deviation,  $s(\bar{R}_m)$ , can be calculated.

Moreover, the experiments can be designed so that the recovery assays can provide additional information. One of the possible experimental designs for estimating recovery for a variable amount  $i$  of spiked analyte is shown in Fig. 1. This design corresponds to a three-factor fully-nested design [8]. The factors varied are the amount of analyte added, the factor analyte+matrix and the intermediate precision.



**Figure 1.** Experimental design for estimating recovery when  $l$  different amounts of analyte are added to representative matrices. Three factors of variation are studied, i.e. the amount added, the analyte+matrix factor and the intermediate precision. For each amount added, recovery is estimated in  $p$  samples in which the analyte and/or the matrix have been representatively varied. In each of these  $p$  samples, recovery is estimated  $n$  times under intermediate precision conditions.  $R_{ijk}$  is the  $k^{\text{th}}$  recovery obtained under intermediate precision conditions for an amount  $i$  added to the  $j^{\text{th}}$  sample.  $\bar{R}_{ij}$  is the mean value of  $R_{ijk}$  recoveries.  $\bar{R}_i$  is the mean of all the  $\bar{R}_{ij}$  and, finally,  $\bar{R}_m$  is the grand mean of all the recoveries obtained in the experimental design.

The design shown in Fig. 1 can be simplified if the amount spiked and the matrix+analyte factor are studied as a whole. Moreover, this design can also be made more complex if there are sources of variation, such as changes in some steps of the analytical procedure, which are studied separately as other factors of the fully-nested design.

Once the absence of outlier mean recoveries has been checked for each of the factors studied in the design, the overall recovery,  $\bar{R}_m$ , can be estimated as the mean of all the  $\bar{R}_i$  obtained for the  $l$  different amounts of the analytes spiked:

$$\bar{\mathcal{R}}_m = \frac{\sum_{i=1}^I \bar{\mathcal{R}}_i}{I} \quad (1)$$

The overall recovery,  $\bar{\mathcal{R}}_m$ , is an estimation of the “method recovery” [3] because it has been obtained by representatively varying the factors (i.e. matrix, concentration and analyte) that can affect recovery. Therefore, proportional bias can be estimated in terms of the overall recovery.

Once the overall recovery has been estimated, the proportional bias is checked to see whether it is statistically significant or not. This is done by seeing whether the overall recovery,  $\bar{\mathcal{R}}_m$ , is statistically different from one, within the limits of its uncertainty. If  $\bar{\mathcal{R}}_m$  is not statistically different from unity, the following inequality should hold:

$$|\bar{\mathcal{R}}_m - 1| \leq t_{\alpha/2, \text{eff}} \cdot u(\bar{\mathcal{R}}_m) \quad (2)$$

where  $u(\bar{\mathcal{R}}_m)$  is the uncertainty of the overall recovery (calculated in section 4.1) and  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value for the effective degrees of freedom [9] associated to  $u(\bar{\mathcal{R}}_m)$ .

If the recovery is not statistically different from one, it can be concluded that the procedure does not have any significant proportional bias. Therefore, it is not necessary to apply the recovery factor to future results obtained with the procedure. If the recovery is statistically different from one, the recovery factor should be applied to future results. In any case, an uncertainty coming from the recovery estimation is always present and has to be included in the final uncertainty budget.

## 2.2. Estimation of constant bias

The constant bias of the method cannot be estimated from the spiked samples because the samples already contain the analytes. Therefore, a method known as the Youden Method [10] will be used to estimate the constant bias. The Youden method consists of analysing two or more different amounts (weights or volumes) of a sample which is representative of the routine samples. If there is no bias, both analyses should give the same result. For greater simplicity, the Youden Method will be applied to the analysis of only two different amounts of a routine sample [11]. In this case, the constant bias can be estimated from:

$$\frac{\bar{x}_n - \mathbf{d}_{ct}}{W_n} = \frac{\bar{x}_m - \mathbf{d}_{ct}}{W_m} \quad (3)$$

where  $\bar{x}_n$  and  $\bar{x}_m$  are the mean of the analytical results obtained when the amounts of sample,  $W_n$  and  $W_m$ , respectively, are analysed and  $\delta_{ct}$  is the constant bias. From Eq. (3) the constant bias is calculated as:

$$\mathbf{d}_{ct} = \frac{W_m \cdot \bar{x}_n - W_n \cdot \bar{x}_m}{W_m - W_n} \quad (4)$$

The constant bias is commonly known as the Youden Blank and it is considered to be an estimation of the true sample blank [12] because it is determined with routine samples when both the native analyte and the matrix are present.

Once the constant bias has been estimated, the lack of constant bias of the analytical procedure must be verified by checking if the constant bias estimated is statistically significant:

$$|\mathbf{d}_{ct}| \leq t_{\alpha/2, eff} \cdot u(\mathbf{d}_{ct}) \quad (5)$$

If inequality 5 is satisfied, the analytical procedure does not have any significant constant bias. However, the uncertainty due to the estimation of this bias,  $u(\delta_{ct})$ , must be included in the overall uncertainty of future results obtained with the



analytical procedure. This uncertainty is calculated in section 4.2. If the constant bias is significant the procedure should be revised.

### 3. UNCERTAINTY OF ROUTINE SAMPLES ANALYSED AFTER THE TRUENESS OF THE ANALYTICAL METHOD IS VERIFIED

The uncertainty of future routine samples can be estimated from the information generated when the trueness of the analytical method is assessed by recovery assays. This uncertainty is considered to be correct whenever the samples used in the recovery assays representatively span the different types of matrices and analyte concentrations of the routine samples. The laboratory is also assumed to have implemented and carried out an effective policy of quality assurance procedures.

A result,  $x_{\text{fut}}$ , obtained for a future routine sample can be expressed as:

$$x_{\text{fut}} = R \cdot \mu_{\text{fut}} + d_{\text{ct}} + d_{\text{run}} + e \quad (6)$$

where  $\mu_{\text{fut}}$  is the true concentration of the routine sample,  $R$  is the recovery factor,  $d_{\text{ct}}$  is the constant bias,  $d_{\text{run}}$  is the run bias and  $e$  is the random error. Since we are interested in estimating the true concentration,  $\mu_{\text{fut}}$ , the result obtained should be corrected by the constant and run biases, recovery and random error. However, the run bias and the random error are unknowable quantities. Only the repeatability variance,  $s_r^2$ , and the between-run variance,  $s_{\text{run}}^2$ , can be known. These variances are, respectively, associated to the distribution of the values that the random error and the run bias can take. The sum of these variances corresponds to the run-different intermediate variance,  $s_i^2$ . The true values of the constant bias and recovery are also unknown. Only their estimations, and their related uncertainties, are available. As a result, the true concentration of a routine sample cannot be obtained, only a corrected concentration,  $x_{\text{corr}}$ :

$$x_{\text{corr}} = F_{\text{hom}} \cdot \frac{x_{\text{fut}} - \hat{a}_{\text{ct}} - \hat{a}_{\text{run}} - \hat{a}}{R} \quad (7)$$

Here the factor  $F_{\text{hom}}$  has been included. It considers the sources of variability that are present in the analysis of routine samples but that are not taken into account in the assessment of trueness. These sources of variability are often due to the lack of homogeneity of the routine samples or to sample pretreatments not carried out during the assessment of trueness. This term makes Eq. (7) more general. For practical purposes,  $F_{\text{hom}}$  will, in most cases, be 1. However, its uncertainty may not be negligible and should be taken into account, as will be shown below.

The true concentration,  $\mu_{\text{fit}}$ , will lie in the interval  $x_{\text{corr}} \pm \text{Uncertainty}$  with a chosen probability whenever all the sources of uncertainty are considered. The sources of uncertainty are due to various factors: the estimation of constant bias,  $u(\mathbf{d}_{\text{ct}})$ , and recovery,  $u(R)$ , the uncertainty of the run bias, which is given by the between-run standard deviation,  $s_{\text{run}}$ , the uncertainty of the random error, which is given by the repeatability standard deviation,  $s_r$ , and, sometimes, the lack of homogeneity of the sample,  $u(F_{\text{hom}})$ . The standard uncertainty associated to the corrected result,  $x_{\text{corr}}$ , is then estimated by:

$$u(x_{\text{corr}}) = \sqrt{\frac{(x_{\text{corr}} \cdot u(F_{\text{hom}}))^2}{F_{\text{hom}}^2} + \frac{F_{\text{hom}}^2}{R^2} \cdot \left( \frac{s_{\text{run}}^2}{p_s} + \frac{s_r^2}{n_s \cdot p_s} \right) + \frac{F_{\text{hom}}^2 \cdot u(\mathbf{a}_{\text{ct}})^2}{R^2} + \frac{(F_{\text{hom}} \cdot x_{\text{corr}} \cdot u(R))^2}{R^2}} \quad (8a)$$

where  $p_s$  is the number of runs in which the routine sample has been analysed and  $n_s$  the number of times which the routine sample has been analysed within each run (normally  $p_s=1$ ). Here we suppose that the between-run standard deviation considers all the factors of variation of the analytical method in the normal analysis of the routine samples. However, the between-run standard deviation may be obtained even though all the factors that affect analytical results are not representatively varied. If this is the case, the uncertainty of the factors not representatively varied should also be included in Eq. 8a. These other sources of uncertainty can be obtained either from manufacturer's specifications or from ruggedness tests performed during method validation [13].

The between-run and the repeatability variances can be expressed as a function of their relative standard deviation if the constant bias of the method is negligible. Taking  $F_{\text{hom}} = 1$  and expressing, respectively, the between-run variance and the repeatability variance as a function of the between-run standard deviation,  $RSD_{\text{run}}$ , and of the repeatability standard deviation,  $RSD_r$ , Eq. 8a becomes:

$$u(x_{\text{corr}}) = \frac{1}{R} \cdot \sqrt{(x_{\text{corr}} \cdot R)^2 \cdot (u(F_{\text{hom}}))^2 + x_{\text{fut}}^2 \cdot \left( \frac{RSD_{\text{run}}^2}{p_s} + \frac{RSD_r^2}{n_s \cdot p_s} \right) + u(\hat{a}_{\text{ct}})^2 + x_{\text{corr}}^2 \cdot u(R)^2} \quad (8b)$$

The first term corresponds to the uncertainty arising from pretreatments and/or the lack of homogeneity of the sample. The second one considers the uncertainty arising from the experimental variability of the method, and the third one takes into account the uncertainty associated to the estimation of the constant bias. The fourth one is due to the uncertainty in the estimation of recovery. Note that all the terms except  $u(\hat{a}_{\text{ct}})$  are concentration dependent.

If the repeatability and the between-run precision are not known, Eq. 8b can be expressed in terms of the run-different intermediate precision:

$$u(x_{\text{corr}}) = \frac{1}{R} \cdot \sqrt{(x_{\text{corr}} \cdot R)^2 \cdot (u(F_{\text{hom}}))^2 + \frac{x_{\text{fut}}^2 \cdot RSD_{\text{I}}^2}{p_s} + u(\hat{a}_{\text{ct}})^2 + x_{\text{corr}}^2 \cdot u(R)^2} \quad (8c)$$

where  $RSD_{\text{I}}$  is the relative run-different intermediate standard deviation. Compared with Eq. 8b, Eq. 8c has the disadvantage that uncertainty is not diminished when the sample is analysed several times under repeatability conditions. Eq. 8b and Eq. 8c show that the lower the recovery, the higher the overall uncertainty is.

The overall expanded uncertainty,  $U(x_{\text{corr}})$ , is obtained by multiplying the standard uncertainty by the coverage factor  $k$ :

$$U(x_{\text{corr}}) = k \cdot u(x_{\text{corr}}) \quad (9)$$

However, if the degrees of freedom associated to the various terms of estimated uncertainty are low, the two-sided  $t$  effective tabulated value should be used [1,9,14].

#### 4. PRACTICAL ESTIMATION OF THE COMPONENTS OF UNCERTAINTY

The uncertainty arising from the lack of homogeneity of the sample and/or pretreatment steps can be calculated by means of the analysis of variance (ANOVA), as recommended by the IUPAC for testing a material for sufficient homogeneity [15]. The recovery and its uncertainty, as well as the constant bias and its uncertainty, are calculated using the information generated in the assessment of trueness. Finally, the precision of the method is obtained either from the information generated in the assessment of trueness or from historical data. The calculation of the last three terms of Eq. (8) is explained in the sections below.

##### 4.1. Uncertainty of recovery

The true recovery,  $R$ , of a given routine sample can be described as the sum of three components:

$$R = R_m + \Delta R_{\text{analyte+matrix}} + \Delta R_{\text{conc}} \quad (10)$$

The first component is known as the “method recovery” [3] and the last two components are associated to the variation of recovery depending on the type of sample analysed: the second one considers the variation of recovery with analyte and matrix whereas the third one considers the variation of recovery with concentration.

The overall recovery,  $\bar{R}_m$ , is calculated when trueness is assessed. This is an estimation of the “method recovery” and can be applied to correct future results if necessary. The last two components of Eq. (10) are unknowable quantities for future routine samples. Only the variance associated to the variability of recovery with matrix and analyte,  $u(\Delta R_{\text{analyte+matrix}})^2$ , and the variance associated to the variability of recovery with concentration,  $u(\Delta R_{\text{conc}})^2$ , can be known. These variances are, respectively, associated to the distribution of values that  $\Delta R_{\text{analyte+matrix}}$  and  $\Delta R_{\text{conc}}$  can take depending on the type of sample analysed. Therefore, they should be included as components of uncertainty to account for differences in recovery for a particular sample. The uncertainty of recovery is then calculated as:

$$u(R) = \sqrt{u(\bar{R}_m)^2 + u(\Delta R_{\text{analyte+matrix}})^2 + u(\Delta R_{\text{conc}})^2} \quad (11)$$

where  $u(\bar{R}_m)$  is the uncertainty of the overall recovery and is estimated in section 4.1.1. The last two uncertainties of Eq. (11) can be obtained from the analysis of variance (ANOVA) of the recoveries obtained in the experimental design in Fig 1. The expressions for calculating these uncertainties are shown in Tables 1 and 2.

**Table 1.** ANOVA table and calculation of variances for the design proposed in Fig. 1.

Source	Mean squares	Degrees of freedom	Expected mean square
Concentration	$MS_{\text{conc}} = \frac{p \cdot n \cdot \sum_{i=1}^l (\bar{R}_i - \bar{R})^2}{l-1}$	$l-1$	$p \cdot n \cdot \sigma_{\text{conc}}^2 + n \cdot \sigma_{\text{a+m}}^2 + \sigma_1^2$
Analyte+matrix	$MS_{\text{a+m}} = \frac{n \cdot \sum_{i=1}^l \sum_{j=1}^p (\bar{R}_{ij} - \bar{R}_i)^2}{l \cdot (p-1)}$	$l \cdot (p-1)$	$n \cdot \sigma_{\text{a+m}}^2 + \sigma_1^2$
Intermediate precision	$MS_1 = \frac{\sum_{i=1}^l \sum_{j=1}^p \sum_{k=1}^n (R_{ijk} - \bar{R}_{ij})^2}{l \cdot p \cdot (n-1)}$	$l \cdot p \cdot (n-1)$	$\sigma_1^2$
Total		$l \cdot p \cdot n - 1$	

**Table 2.** Calculation of variances for the experimental design proposed

Variance	Expression	Degrees of freedom
$u(\Delta R_{\text{conc}})^2$	$\frac{MS_{\text{conc}} - MS_{\text{a+m}}}{p \cdot n}$	
$u(\Delta R_{\text{analyte +matrix}})^2$	$\frac{MS_{\text{a+m}} - MS_{\text{I}}}{n}$	
$u(R_i)^2$	$MS_{\text{I}}$	$l \cdot p \cdot (n-1)$

**4.1.1. Uncertainty of the overall recovery,  $u(\bar{R}_m)$**

This uncertainty can be estimated as:

$$u(\bar{R}_m) = \sqrt{\frac{\sum_{i=1}^l u(\bar{R}_i)^2}{l^2}} \quad (12)$$

To calculate this uncertainty, the uncertainty of each mean recovery,  $u(\bar{R}_i)$ , must first be calculated. This uncertainty can be estimated in two different ways:

a) As the standard deviation of  $\bar{R}_i$ :

$$u(\bar{R}_i)^2 = \frac{\sum_{j=1}^p (\bar{R}_{ij} - \bar{R}_i)^2}{p \cdot (p-1)} \quad (13)$$

b) Using information about the intermediate precision of the method

The mean recovery,  $\bar{R}_i$ , for each amount  $i$  of analyte added, is obtained as:

$$\bar{R}_i = \frac{\sum_{j=1}^p \sum_{k=1}^n R_{ijk}}{p \cdot n} \quad (14)$$

The uncertainty of the mean recovery,  $u(\bar{R}_i)$ , can be calculated by applying the error propagation law to Eq. (14):

$$u(\bar{R}_i)^2 = \frac{(RSD_1 \cdot \bar{x}_i)^2}{x_{a,i}^2 \cdot n \cdot p} \quad (15)$$

where  $\bar{x}_i$  is the mean concentration obtained when analysing the samples spiked with an amount  $x_{a,i}$ . This is a good approach for estimating  $u(\bar{R}_i)$  and should be used when the number of mean recoveries,  $\bar{R}_{ij}$ , is not high and when intermediate precision information is available from historical data.

It is important to note that the uncertainty of the overall recovery is obtained by assuming that the uncertainty associated to the amount of analyte spiked is negligible. If it is not negligible, this uncertainty (which is mainly due to material purity, mass and volumetric glassware) should also be included in Eq. 12.

#### 4.2. Uncertainty of constant bias

The uncertainty associated to the constant bias estimated using the Youden method is obtained by applying the error propagation law to Eq. (4). This uncertainty depends on the uncertainty of the analytical results,  $\bar{x}_n$  and  $\bar{x}_m$ , and on the uncertainty associated to the amounts of the sample analysed,  $W_n$  and  $W_m$ . If the uncertainty associated to the amounts of sample analysed is negligible in comparison to the uncertainty of the analytical results,  $\bar{x}_n$  and  $\bar{x}_m$ , the uncertainty of the constant bias  $u(\delta_{ct})$  is calculated as:

$$u(\mathbf{d}_{ct}) = \frac{1}{W_m - W_n} \cdot \sqrt{(W_m \cdot u(\bar{x}_n))^2 + (W_n \cdot u(\bar{x}_m))^2} \quad (16)$$

where  $u(\bar{x}_n)$  and  $u(\bar{x}_m)$  are the uncertainties of the analytical results  $\bar{x}_n$  and  $\bar{x}_m$ , respectively. These uncertainties depend on the intermediate precision of the method at the level of concentration of the sample and on the number of analysis

carried out to obtain the analytical result. For instance, if  $\bar{x}_n$  is obtained as the mean of  $n$  analysis, the uncertainty  $u(\bar{x}_n)$  is calculated as  $RSD_1 \cdot \bar{x}_n / \sqrt{n}$ .

### 4.3. Uncertainty arising from the precision of the method

This uncertainty considers the variation of the analytical method in the normal analysis of the routine samples. Therefore, this uncertainty depends on the intermediate precision of the analytical method at the level of concentration of the routine sample analysed. The intermediate precision should be factored into the repeatability and the between-run precision when routine samples are analysed more than once under repeatability conditions. Both types of precision can be estimated in precision studies using two-factor fully-nested designs [5,8].

Intermediate precision should consider all the factors of variability which may affect the analytical results. However, if this is not the case, the uncertainty of the factors not representatively varied should also be included. These other sources of uncertainty can be obtained either from manufacturer's specifications or from ruggedness tests performed in method validation [13].

The intermediate precision can be obtained either from historical data or from the information generated during the recovery assays. In this latter case, a good estimation for the intermediate precision can be obtained for each concentration corresponding to the amount of analyte spiked whenever the concentration of the native analyte is negligible. At each level, the intermediate standard deviation can be calculated as:

$$S_{i,i}^2 = \frac{\sum_{j=1}^p \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2}{p \cdot (n-1)} \quad (17)$$

where  $x_{ijk}$  is the  $k$ th uncorrected analytical result (i.e. the result obtained without being corrected by recovery and constant bias) obtained for a given analyte contained in a sample  $j$  when an amount  $x_{a,i}$  is added to the sample. The



intermediate precision, expressed as relative standard deviation, is obtained for each amount  $i$  of analyte spiked as:

$$RSD_{1,i} = \frac{S_{1,i}}{\bar{x}_i} \quad (18)$$

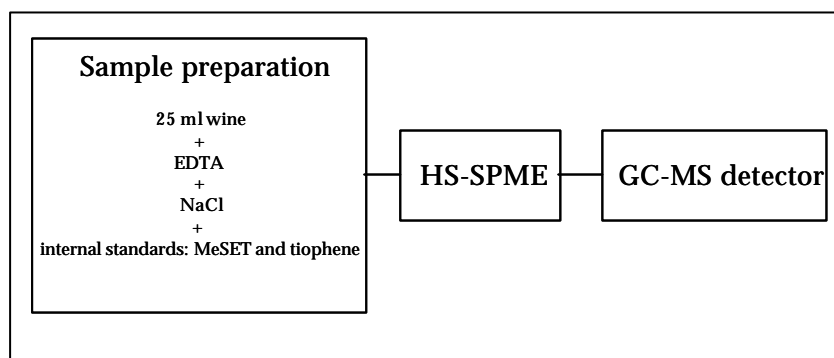
where  $\bar{x}_i$  is the mean of the uncorrected analytical results when an amount  $x_{a,i}$  is added to the sample. If the relative standard deviation remains constant in the concentration range of the recovery assays, the relative intermediate standard deviation can also be obtained as the variability of recovery arising from the precision of the method,  $u(R_i)/\bar{R}_m$ , (the expression of  $u(R_i)$  is shown in Tables 1 and 2). In this case,  $(s_{1(\text{sample})}/x_{a,i})^2$  should also be negligible in comparison to  $u(R_i)^2$  (where  $s_{1(\text{sample})}$  is the intermediate precision associated to the sample analysed before the analyte is added).

## 5. PRACTICAL EXAMPLE: ESTIMATION OF UNCERTAINTY IN THE DETERMINATION OF SULPHIDES IN WINE BY HS-SPME-GC

To illustrate the advice proposed in this paper, the results obtained in the validation of a GC method for the determination of sulphides in wine [7] were used to calculate the uncertainty of future samples analysed with this method. The analytical method consisted several steps (Fig. 2). In the second step, after an initial sample preparation, the sulphides were extracted by headspace-solid phase microextraction (HS-SPME). Two different types of Carboxen-polydimethylsiloxane (CAR-PDMS) fibres were used to extract the sulphides. Afterwards, the extracted sulphides were analysed with a Hewlett-Packard 5890 (series II) gas chromatograph equipped with an HP-5972 mass-selective detector.

Table 3 shows the results of the recovery assays for the following monosulphides: methyl thioacetate (MeSAc), diethyl sulphide (EtSEt), methyl-*n*-propyl sulphide (MeSPr) and ethyl thioacetate (EtSAc). As can be seen, the recovery was estimated in red and white wines for three different amounts of each of the spiked analytes,

i.e. 0.5, 2.5 and 25  $\mu\text{g/l}$ . The wines used in the recovery assays already contained these sulphides at trace levels because of the degradation of the sulphur constituents of some amino acids. As a result, the Youden Method was used to check the lack of constant bias.



**Figure 2.** Scheme of the analytical procedure for estimating sulphides in wine [7]

**Table 3.** Recovery estimated for spiked amounts of 0.5, 2.5 and 25  $\mu\text{g/l}$ . For each amount added, the recovery,  $R$ , of each analyte is estimated in white wine and red wine using two different fibers in the SPME. Recovery is estimated three times with each fiber under intermediate precision conditions. The standard deviation of these three recoveries,  $s(R)$ , is also shown.

	White wine				Red wine			
	Fiber 1		Fiber 2		Fiber 1		Fiber 2	
	$R$	$s(R)$	$R$	$s(R)$	$R$	$s(R)$	$R$	$s(R)$
$x_a (\mu\text{g/l}) = 0.5$								
MeSAc	0.90	0.05	0.98	0.11	1.12	0.28	1.14	0.11
EtSEt	1.01	0.24	1.09	0.17	0.97	0.19	1.19	0.17
MeSPr	0.95	0.15	0.97	0.02	1.02	0.02	1.07	0.08
EtSAc	0.96	0.11	1.10	0.20	1.16	0.16	1.15	0.25
$x_a (\mu\text{g/l}) = 2.5$								
MeSAc	1.03	0.08	1.09	0.08	1.07	0.18	0.98	0.08
EtSEt	0.99	0.03	0.96	0.17	1.05	0.05	1.07	0.07
MeSPr	1.00	0.07	1.15	0.07	1.04	0.05	1.02	0.15
EtSAc	1.10	0.02	1.00	0.22	1.15	0.12	0.99	0.08
$x_a (\mu\text{g/l}) = 25$								
MeSAc	1.00	0.26	1.18	0.09	1.12	0.09	1.19	0.18
EtSEt	0.95	0.21	0.82	0.04	0.95	0.14	1.18	0.19
MeSPr	0.98	0.21	0.82	0.07	0.95	0.12	1.21	0.23
EtSAc	1.04	0.23	1.02	0.25	1.02	0.11	1.14	0.19

## 5.1. Assessment of trueness

Assessing trueness involves estimating proportional bias (in terms of the overall recovery) and constant bias (with the Youden Method). Then both biases are checked to see whether they are statistically significant or not.

### 5.1.1. Estimating the overall recovery

Since the analyte is simply dissolved in the wine sample, we assumed that the analyte spiked behaves as the native analyte and, as a result, that the recovery obtained with the spiked samples was a good estimation of the real recovery. Eq. (1) was used to estimate the mean of all the recoveries in Table 3 for all the sulphide concentrations added. This value,  $\bar{R}_m$ , is 1.0425. Before calculating the uncertainty of the overall recovery, the square uncertainty of the recovery obtained at each level of concentration,  $u(\bar{R}_i)^2$ , was obtained following Eq. (13). In this case,  $\bar{R}_{ij}$  corresponded to the mean recovery obtained with the two fibers for a given analyte spiked with an amount  $i$ .  $p$  was equal to 8, i.e. there are four analytes spiked in two different matrices. Table 4 shows the values of  $u(\bar{R}_i)^2$  obtained for the three amounts spiked. The uncertainty of the overall recovery,  $u(\bar{R}_m)$ , was then calculated with Eq. (12) (here  $l=3$ ). This uncertainty, expressed as a variance, is shown in Table 4. The uncertainty associated to the amount spiked was not included because it was considered to be negligible.

**Table 4** Components of uncertainty of Eq. (11) using the information generated in the recovery assays. The overall uncertainty associated to recovery,  $u(R)$ , is also shown.

Source	$x_a$	Uncertainty		
		0.5	2.5	25
Estimation of $\bar{R}_m$	$u(\bar{R}_i)^2$	$6.95 \cdot 10^{-4}$	$1.34 \cdot 10^{-4}$	$1.13 \cdot 10^{-4}$
	$u(\bar{R}_m)^2$		$2.17 \cdot 10^{-4}$	
$u(\Delta R_{\text{analyte+matrix}})^2$			$1.35 \cdot 10^{-3}$	
$u(\Delta R_{\text{conc}})^2$			0	
Overall uncertainty			$3.96 \cdot 10^{-2}$	
$u(R)$				

A significance test was used to determine whether the overall recovery estimated was significantly different from 1. The test statistic  $t$  was calculated with Eq. (2):

$$t = \frac{|1.0425 - 1|}{1.47 \cdot 10^{-2}} = 2.88$$

This value was higher than the two-sided  $z$ -value for a probability  $\alpha=0.05$ , i.e. 1.96 (the  $z$  tabulated value was used instead of the  $t$  tabulated value because of the considerable number of degrees of freedom associated to the uncertainty of recovery). Therefore, the recovery estimated,  $\bar{R}_m$ , was found to be statistically significant and, as a result, it was applied to correct the analytical results of future routine samples.

### **5.1.2. Estimating constant bias**

Constant bias was estimated with the Youden method. Two volumes of red wine (25 and 10 ml) were analysed. The 25 ml of red wine had already been analysed six times in the recovery assays (three times with each fibre). The mean value of the results was 0.48  $\mu\text{g/l}$ . The 10 ml of red wine was analysed twice. The mean value was 0.18  $\mu\text{g/l}$ . The constant bias estimated with Eq.(4) was 0.02. The uncertainty estimated with Eq. (16) was  $9.22 \cdot 10^{-2}$ .

Once the uncertainty of the constant bias had been estimated, a significance test was used to determine whether the constant bias was statistically significant. The test statistic  $t$  was calculated using the following equation:

$$t = \frac{|0.02|}{9.22 \cdot 10^{-2}} = 0.22$$

This value was lower than the two-sided  $z$ -value for a probability  $\alpha=0.05$ , i.e. 1.96. (The  $z$  tabulated value was also used because of the high number of degrees of freedom associated to the uncertainty of the constant bias, i.e. 96). Therefore, the constant bias estimated was not statistically significant.

**Table 5.** Mean squares, degrees of freedom (d.f) and variance (Var) associated to the four factors varied in the recovery assays.

Source	Mean squares	d.f	Expected mean square	Var
Concentration	$MS_{\text{conc}} = \frac{8 \cdot 2 \cdot 3 \cdot \sum_{i=1}^3 (\bar{R}_i - \bar{R}_m)^2}{3 - 1} = 2.08 \cdot 10^{-3}$	2	$8 \cdot 2 \cdot 3 \cdot \sigma_{\text{conc}}^2 + 2 \cdot 3 \cdot \sigma_{\text{a+m}}^2 + 3 \cdot \sigma_f^2 + \sigma_1^2$	0
Analyte + matrix	$MS_{\text{a+m}} = \frac{2 \cdot 3 \cdot \sum_{i=1}^3 \sum_{j=1}^8 (\bar{R}_{ij} - \bar{R}_i)^2}{3 \cdot (8 - 1)} = 3.13 \cdot 10^{-2}$	21	$2 \cdot 3 \cdot \sigma_{\text{a+m}}^2 + 3 \cdot \sigma_f^2 + \sigma_1^2$	$1.35 \cdot 10^3$
Fiber	$MS_f = \frac{\sum_{i=1}^3 \sum_{j=1}^8 \sum_{k=1}^2 (R_{ijk} - \bar{R}_{ij})^2}{3 \cdot 8 \cdot (2 - 1)} = 2.32 \cdot 10^{-2}$	24	$3 \cdot \sigma_f^2 + \sigma_1^2$	0
Intermediate precision	$MS_1 = \frac{\sum_{i=1}^3 \sum_{j=1}^8 \sum_{k=1}^2 \sum_{r=1}^3 (R_{ijk} - R_{ijkr})^2}{3 \cdot 8 \cdot 2 \cdot (3 - 1)} = 2.32 \cdot 10^{-2}$	96	$\sigma_1^2$	$2.33 \cdot 10^2$

## 5.2. Uncertainty of future routine samples

The information generated while the trueness of the analytical method is verified can be used to estimate the uncertainty of results obtained when the analytical method is applied to future routine samples. Uncertainty was calculated with equation Eq. (8c) because an estimation of the repeatability and the run-different variances were not available. The uncertainty associated to the lack of homogeneity of the sample,  $u(F_{\text{hom}})$ , was not calculated because the samples of wine were very homogeneous. Moreover, the pretreatments carried out on the routine samples are the same as the ones carried out during the validation of the method. Therefore, the contribution of  $u(F_{\text{hom}})$  is negligible in comparison to the other terms of the uncertainty budget. The estimation of the uncertainties arising from recovery, constant bias and precision are described below.

### 5.2.1. Uncertainty of recovery, $u(R)$

Since recovery should be applied to routine samples with different matrices and concentration levels, the uncertainty of the variability of recovery which depends on the sample analysed also has to be calculated and included in the uncertainty

budget of recovery. To estimate this uncertainty, four factors of variation were considered in the experimental design: the concentration, the factor analyte+matrix, the fiber used to extract the sulphides and the intermediate precision. The fiber was also studied as a factor of variation to determine whether the uncertainty of results could be diminished when a sample was analysed in the same run with two different fibers. Table 5 shows the variances obtained from this design together with all the components of uncertainty of recovery. As can be seen, the estimated variance associated to the variability of the fiber is zero. Therefore, a change in the fiber causes negligible variation in the results. As a result, the uncertainty of future results is not diminished if a future sample is analysed twice in the same run with two different fibers. Finally, the uncertainty of the recovery applied to future samples,  $u(R)$ , was estimated with Eq. (11). This value is shown in Table 4.

### **5.2.2. Uncertainty of constant bias**

This uncertainty has been calculated in section 5.1.2.

### **5.2.3. Uncertainty arising from the precision of the method**

This term is expressed as a relative intermediate standard deviation. It is estimated from the recovery assays as  $u(R_1)/\bar{R}_m$ . The last row in Table 5 shows the value of  $u(R_1)^2$ .

### **5.2.4. Estimating the overall uncertainty**

The uncertainty was calculated for concentrations of 0.5, 2.5 and 25  $\mu\text{g}/\text{l}$ . It was assumed that the routine sample was analysed once and twice in two different runs (i.e. for  $p_s=1$  and  $p_s=2$ ). The uncertainty values calculated, together with the values of the last three terms of Eq. (8c), are shown in Table 6. The overall uncertainty,  $U$ , was calculated using the coverage factor  $k$  instead of the  $t$  effective tabulated value. We can see that the uncertainty values obtained are very high. This is because the analytical method is currently under development and, as a result, the precision of the method needs to be improved to diminish uncertainty.

Fig. 3 shows the contribution to the measurement uncertainty of precision, recovery and constant bias when the routine sample is analysed once and twice. This is shown for the three concentrations studied.

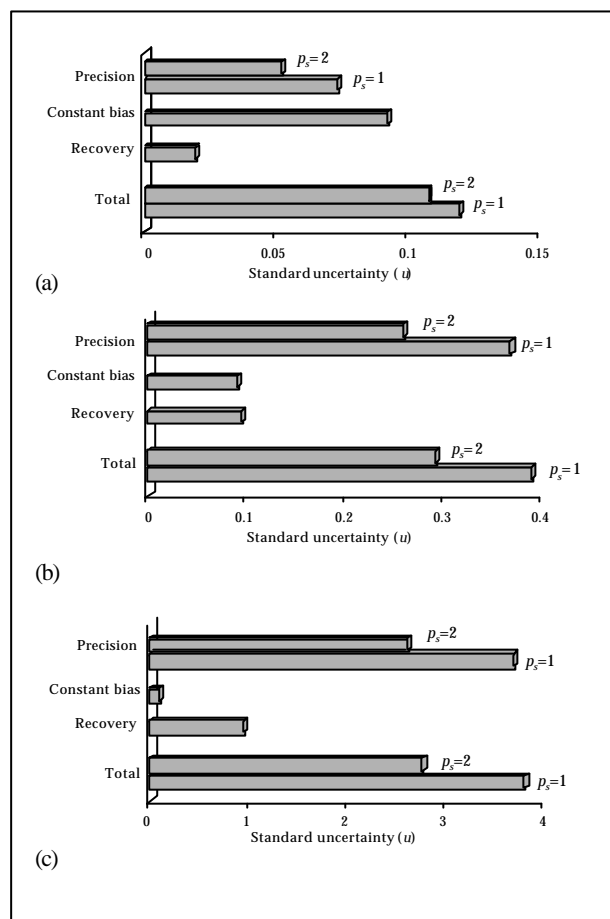
**Table 6.** Uncertainty values estimated for concentrations of sulphides of 0.5, 2.5 and 25 µg/l when the routine sample is analysed once ( $p_s=1$ ) or twice ( $p_s=2$ ). The relative uncertainty values together with the contribution of each term of Eq. (8c) to the overall uncertainty are also shown.

Sulphide µg/l	$\frac{(u(R_t) \cdot x_{\text{fut}})^2}{R_m^4}$		$\frac{u(d_t)^2}{R_m^2}$	$\frac{x_{\text{corr}}^2 \cdot u(R)^2}{R_m^2}$	<b>U</b> (k=2) ( $p_s=1$ )	<b>U / <math>x_{\text{corr}}</math></b> (%) ( $p_s=1$ )	<b>U</b> (k=2) ( $p_s=2$ )	<b>U / <math>x_{\text{corr}}</math></b> (%) ( $p_s=2$ )
	( $p_s=1$ )	( $p_s=2$ )						
0.5	4.92·10 <sup>-3</sup>	2.47·10 <sup>-3</sup>	8.51·10 <sup>3</sup>	3.60·10 <sup>-4</sup>	2.35·10 <sup>-1</sup>	46.97	2.13·10 <sup>-1</sup>	42.59
2.5	1.23·10 <sup>-1</sup>	6.16·10 <sup>-2</sup>	8.51·10 <sup>3</sup>	9.01·10 <sup>-3</sup>	7.50·10 <sup>-1</sup>	29.99	5.63·10 <sup>-1</sup>	22.50
25	1.23·10 <sup>1</sup>	6.16	8.51·10 <sup>3</sup>	9.01·10 <sup>-1</sup>	7.27	29.07	5.32	21.27

Fig. 3 shows that, for concentrations of 2.5 and 25 µg/l, the term with the highest contribution is the one associated to the precision of the method. The constant bias, however, is the most important contribution when the concentration of sulphide is 0.5 µg/l but it is not important for concentrations of 2.5 and 25 µg/l. This is because the constant bias was estimated with a sample whose concentration level was similar to 0.5 µg/l. Therefore, the uncertainty of the constant bias is related to an intermediate precision of a level of concentration of 0.5 µg/l. As a result, the contribution of the constant bias is important for concentrations similar to 0.5 µg/l but not for higher concentrations. The contribution of the proportional bias to the overall uncertainty was not important at the three levels of concentration studied.

The relative uncertainties,  $U/x_{\text{corr}}$ , obtained at the three concentration levels are similar to the relative within-laboratory variability values predicted by Horwitz. These values range between 1/2 and 2/3 of the between-laboratory variability given by the Horwitz expression [16]. If the within-laboratory standard deviation is calculated assuming 2/3 of the relative standard deviation predicted by

Horwitz, the concentration studied should have relative standard deviations of 33.3%, 26% and 18.6% for concentrations of 0.5, 2.5 and 25  $\mu\text{g/l}$ , respectively.



**Figure 3.** Contribution of precision, recovery and constant bias to the measurement uncertainty in the determination of sulphides in wine [7]. These uncertainties, as well as the total uncertainty of a routine sample, are expressed as standard uncertainties and have been calculated when the routine sample is analysed once ( $p_s=1$ ) or twice ( $p_s=2$ ). This contribution is shown for the three levels of concentration studied: 0.5  $\mu\text{g/l}$  (Fig. 3a); 2.5  $\mu\text{g/l}$  (Fig. 3b) and 25  $\mu\text{g/l}$  (Fig. 3c).

Fig. 3 also shows that uncertainty decreases considerably when the routine sample is analysed twice ( $p_s=2$ ) for concentrations of 2.5 and 25  $\mu\text{g/l}$ . This decrease is not so important at a concentration of 0.5  $\mu\text{g/l}$  because the uncertainty arising from the



experimental variation is not the most important contribution. Therefore, if it is in your interests to decrease the uncertainty of a result, the sample should be analysed more than once. However, if the routine samples have a level of concentration of 0.5 µg/l or lower, it is better to decrease the uncertainty arising from the estimation of the constant bias than to increase the number of analyses of the routine sample. This uncertainty can be decreased if the two weights of the sample are analysed more times.

## **6. CONCLUSIONS**

We have proposed a new procedure for calculating the uncertainty of analytical results using information generated during the process of estimating proportional and constant bias. Compared to other approaches proposed so far [1], this approach has the advantage that no extra work of any importance needs to be done because most of the uncertainty estimations come from the validation process. The uncertainty estimated with this approach can be applied to all future routine samples analysed with the analytical method because the variability of these samples is included as a component of uncertainty. Only other terms related to the lack of homogeneity, to pretreatments or to factors not representatively varied when estimating intermediate precision need be subsequently included. Moreover, the approach verifies the lack of constant bias and includes this term in the uncertainty budget. We evaluate the importance of each of the terms of uncertainty using data from the validation of a GC method and show that the uncertainty associated to the constant bias can be very important when the routine sample has a concentration similar to or lower than the sample used to estimate this constant bias. Therefore, it is better to check the absence of constant bias using a sample with a concentration which falls in the lower end of the concentration range. Finally, we show how the uncertainty can be significantly diminished by just analysing the sample twice under run-different intermediate precision conditions. Future studies will focus on how to apply regression techniques for calculating the measurement uncertainty of methods in which trueness is assessed with spiked samples at different levels of concentration.

## 7. ACKNOWLEDGEMENTS

The authors would like to thank the DGICYT (project num. BP96-1008) for financial support and the CIRIT from the Catalan Government for providing Alicia Maroto's doctoral fellowship. Montse Mestres, from the Group of Wine and Food Analytical Chemistry of the Rovira i Virgili University is gratefully acknowledged for providing the experimental data.

## 8. REFERENCES

- [1] EURACHEM, *Quantifying uncertainty in analytical measurement*, Draft: EURACHEM Workshop, 2<sup>nd</sup> Edition, Helsinki, 1999
- [2] A.G. González, M.A. Herrador, A.G. Asuero, *Talanta* **48** (1999) 729-736
- [3] V.J. Barwick, S.L.R. Ellison, *Analyst* **124** (1999) 981-990
- [4] M. Parkany, *The use of recovery factors in trace analysis*, Royal Society of Chemistry, Cambridge, 1996
- [5] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Anal. Chim. Acta* **391** (1999) 173-185
- [6] A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Trends Anal. Chem.* **18/9-10** (1999) 577-584
- [7] M. Mestres, C. Sala, M.P. Martí, O. Busto, J. Guasch, *J. Chromatogr. A* **835** (1999) 137-144
- [8] International Organization for Standardization, ISO 5725, *Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva, 1994
- [9] F.E. Satterthwaite, *Psychometrika* **6** (1941) 309-316
- [10] W.J. Youden, *Anal. Chem.* **19** (1947) 946-950
- [11] W.R. Kelly, J.D. Fasset, *Anal. Chem.* **55** (1983) 1040-1044.
- [12] M. Cardone, *Anal. Chem.* **58** (1986) 438-445
- [13] V.J. Barwick, S.L.R. Ellison, *Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data*, VAM Project 3.2.1, 2000
- [14] BIPM, IEC, IFCC, IUPAC, IUPAP, OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993
- [15] IUPAC, ISO, AOAC, *The International Harmonized Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories*, *Pure & Appl. Chem.* **65** (1993) 2123-2144
- [16] K.W. Boyer, W. Horwitz, R. Albert, *Anal. Chem.* **57** (1985) 454-459

### **3.5. Estimation of measurement uncertainty by using regression techniques and spiked samples (*Analytica Chimica Acta* 446 (2001) 133-145)**

Alicia Maroto, Ricard Boqué, Jordi Riu, F. Xavier Rius

Department of Analytical and Organic Chemistry. Institute of Advanced Studies.

Rovira i Virgili University of Tarragona.

Pl. Imperial Tàrraco, 1. 43005 Tarragona, Catalonia. SPAIN.

#### **SUMMARY**

In this paper, we describe how to assess the trueness of analytical procedures using spiked samples and regression techniques. We show then how to calculate the uncertainty of future samples using the information generated in this process. Two types of bias are calculated in the assessment of trueness: a proportional bias (normally expressed in terms of recovery) and a constant bias. Since in many cases blank samples are not available, we propose a method for assessing trueness when samples that already contain the native analyte are spiked at several levels of concentration. Only proportional bias can be estimated from spiked samples, while constant bias must be estimated by the Youden Method, i.e. by taking different amounts of sample. We propose two ways of assessing trueness. The first expresses results as instrumental responses and the second expresses results as concentration. We then present expressions for calculating uncertainty that cover these two situations. Finally, we use the expressions to calculate uncertainty from validation data of the analysis of esters in wine by SPME-GC.

**Keywords:** Assessment of trueness, Uncertainty, Recovery, Constant bias, Spiked samples, Regression

## 1. INTRODUCTION

Analytical results are often used as pieces of information for relevant purposes. Analysts are therefore increasingly required to provide not only the numerical values of the analytical results, but also their uncertainty. Uncertainty can be obtained either by calculating all the sources of uncertainty individually or by grouping different sources of uncertainty whenever possible. The first way is known as the “bottom-up” approach and was proposed by ISO [1,2]. However, identifying and quantifying all the sources of uncertainty individually is not straightforward, so other, more holistic, approaches based on calculating uncertainty using information from the validation process [3-7] have been proposed. Recently, Eurachem [8] has incorporated a holistic strategy as one way to calculate uncertainty.

The approach proposed by Maroto *et al.* [3,4] uses the information obtained when the trueness of the analytical procedure is verified within a laboratory following the procedure developed by Kuttatharmmakul *et al* [9,10]. This is based on evaluating different intermediate precision estimates which consider all the factors, such as changes in operator, equipment and time, that contribute to the variability of the analytical results. The approach proposed in [3,4] can be used when routine samples have similar levels of concentration because the bias of the analytical procedure is assumed to be constant throughout the concentration range. However, this cannot be assumed when the routine samples vary within a range of concentration. If this is the case, trueness should be verified using samples that cover the whole concentration range of the routine samples. Ellison *et al.* [5-7,11] recently proposed a new protocol for assessing trueness and calculating uncertainty in wide ranges of concentration. This is to verify trueness in terms of a “method recovery”. Therefore, the bias of the analytical procedure is only assumed to be proportional. However, there may be two types of bias: proportional bias and constant bias.

To overcome this problem, another approach has recently been proposed [12] that calculates uncertainty in wide ranges of concentration and assumes that both types of bias may be present. In this latter approach, recovery is estimated with the

method of averaged recovery, and constant bias is estimated with the Youden Method [13].

In this paper, we propose assessing trueness by also employing spiked samples but using regression techniques to estimate recovery and constant bias. Different situations are studied: a calibration curve can be either used or not used to express the results of the analyte to be determined (i.e. results can be expressed as concentration or as instrumental responses). Moreover, in both situations the samples can be analysed under repeatability or under intermediate precision conditions within the laboratory. We have used this approach to calculate the uncertainty of the results of an analytical procedure for determining esters in wine by Solid Phase Microextraction and Gas Chromatography (SPME-GC) [14].

## **2. UNCERTAINTY AND VALIDATION OF ANALYTICAL PROCEDURES**

Analytical procedures should be validated before they are used to analyse routine samples. In this process, the systematic errors are estimated in the assessment of trueness. Usually, biased analytical procedures should be modified to eliminate these systematic errors. However, when dealing with spiked samples and recovery estimation, analytical results may be corrected for these errors so that the final results are traceable. Moreover, the uncertainty of these results should also be calculated as a measure of their reliability. Some components of this uncertainty can be obtained using information generated when the analytical procedure is validated within the laboratory. Uncertainty should then consider all the sources of error of the analytical results and can be calculated in a general way by grouping all these sources in four terms:

$$U = t_{\alpha/2, \text{eff}} \cdot \sqrt{u_{\text{precision}}^2 + u_{\text{trueness}}^2 + u_{\text{pretreatments}}^2 + u_{\text{other terms}}^2} \quad (1)$$

where  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value for the effective degrees of freedom [15], and can be replaced by the coverage factor  $k$ . The first term of uncertainty,

$u_{\text{precision}}$ , considers the variation of the analytical procedure in the analysis of the routine samples. Therefore, this uncertainty depends on the intermediate precision of the procedure. Moreover, this term also takes into account the fact that results depend on the matrix of the routine samples. The second term of uncertainty,  $u_{\text{trueness}}$ , considers the uncertainty caused by systematic errors, i.e. constant and proportional bias, estimated in the assessment of trueness. The third term,  $u_{\text{pretreatments}}$ , considers the uncertainty caused by the lack of homogeneity of the sample and sample pretreatments not carried out when the samples are analysed in the assessment of trueness. Finally, the fourth term,  $u_{\text{other terms}}$ , considers other terms of uncertainty associated with higher levels of traceability or factors not representatively varied in the precision studies. The uncertainty associated with these factors can be estimated with ruggedness studies [6-7, 11]. In this paper we show how to calculate the first two terms of uncertainty when analytical procedures are validated using spiked samples in a wide range of concentration and regression techniques. These terms are calculated from information obtained in the precision studies and in the assessment of trueness.

### 3. PRECISION STUDIES

In this paper, precision is assumed to be approximately the same across the concentration range in which the analytical procedure is validated. If this is so, precision can be estimated simply by analysing a test sample that lies within the concentration range studied. Otherwise, other test samples covering the concentration range should be analysed to estimate precision at different levels of concentration.

The experimental design we have proposed to analyse the test sample is a two-factor fully-nested design [9,16]. With this design, information about the intermediate and the repeatability precisions can be obtained. It consists of analysing a test sample in  $p$  different runs. Within each run, the sample is analysed  $n$  times under repeatability conditions. The test sample must be stable, homogeneous, and as similar as possible to the future routine samples. All the

important sources of variation (for instance, the analyst) should be varied between each run.

The typical ANOVA table and the expressions for calculating the between-run variance, repeatability variance and intermediate standard deviation are shown in Tables 1 and 2. As we will see later, precision information can also be obtained from the information generated while assessing trueness.

**Table 1.** ANOVA table and calculation of variances for a two-factor fully-nested design.  $p$  is the number of runs on which the sample is analysed and  $n$  is the number of replicates performed at every run.  $r_{ij}$  is the instrumental response of the sample analysed in the  $j$ th replicate and the  $i$ th run,  $\bar{r}_i$  is the mean of the  $j$  replicate responses performed on run  $i$ . The grand mean,  $\bar{r}$ , is calculated as the mean of the mean values obtained in the  $p$  different runs.

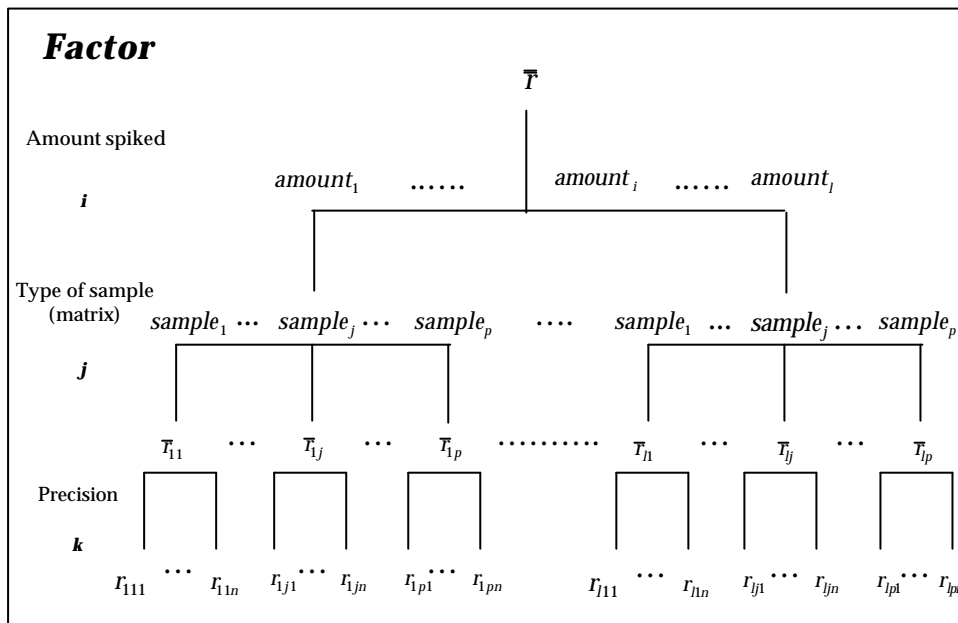
Source	Mean squares	Degrees of freedom	Expected mean square
Run	$MS_{\text{run}} = \frac{n \cdot \sum_{i=1}^p (\bar{r}_i - \bar{r})^2}{p-1}$	$p-1$	$n \cdot s_{\text{run}}^2 + s_r^2$
Repeatability	$MS_r = \frac{\sum_{i=1}^p \sum_{j=1}^n (r_{ij} - \bar{r}_i)^2}{p \cdot (n-1)}$	$p \cdot (n-1)$	$s_r^2$
Total		$p \cdot n-1$	

**Table 2.** Calculation of variances for the two-factor fully-nested design

Variance	Expression	Degrees of freedom
Repeatability variance, $s_r^2$	$MS_r$	$p \cdot (n-1)$
Between-run variance, $s_{\text{run}}^2$	$\frac{MS_{\text{run}} - MS_r}{n}$	
Intermediate variance, $s_l^2$	$s_r^2 + s_{\text{run}}^2$	

#### 4. ASSESSMENT OF TRUENESS

In the assessment of trueness proportional and constant bias are calculated from spiked samples. The calculation of these biases depends on whether blank samples (i.e. samples free from the analyte of interest) are available or not. If they are available, both types of biases can be estimated from the spiked samples. If the analyte is already present in the samples, only the proportional bias, expressed as recovery, can be calculated from the spiked samples. Constant bias must be calculated using the Youden Method [13]. This paper studies the latter case, i.e. when no blank samples are available. We will show that results can be expressed either as instrumental responses or, if a standard curve is used, as concentrations. This is the most usual way to express results in a laboratory of analysis.



**Figure 1.** Experimental design for obtaining information about matrix variability and precision from the results obtained with spiked samples.  $l$  are the amounts of analyte added,  $p$  is the number of different sample types of samples spiked and  $n$  is the number of analyses performed on each sample.  $r_{ijk}$  is the result for the  $k$ th analysis of the  $j$ th sample spiked with an amount of analyte  $i$ . These results are calculated as the difference between the instrumental responses after and before spiking each type of sample.  $\bar{r}_{ij}$  is the mean of the  $k$  results obtained for the  $j$ th sample spiked with an amount of sample  $i$ .  $\bar{r}_i$  is the mean of all the results obtained for an amount  $i$  of analyte.



#### 4.1. Standard Additon Method (SAM). Calculation of proportional bias

Different amounts of analyte are spiked to a set of different matrices with similar levels of concentration of the analyte to be studied and which are representative of the future routine samples. Each spiked sample should be analysed at least twice so that the precision of the analytical procedure and the variability of results with the matrix can be obtained. Fig. 1 shows the proposed experimental design for obtaining information of the between-matrix variance,  $s^2_{\text{matrix}}$ , and the variance associated to precision,  $s^2$ . Each result is obtained as the difference between the instrumental response after analysing the spiked sample once and the average instrumental response after analysing the sample containing only the native analyte several times. If this is so, the error of analysing the samples containing the native analyte is negligible. Consequently, a good estimate can be obtained for matrix variability.

**Table 3.** ANOVA table for the experimental design proposed in Fig. 1.

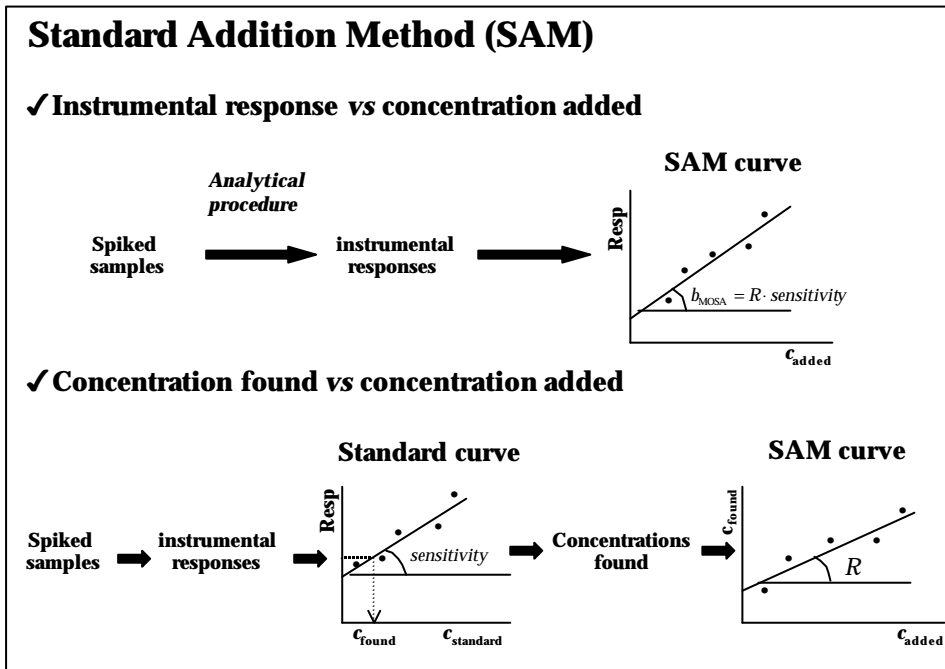
Source	Mean squares	Degrees of freedom	Expected mean square
Matrix	$MS_{\text{matrix}} = \frac{n \cdot \sum_{i=1}^l \sum_{j=1}^p (r_{ij} - \bar{r}_i)^2}{l \cdot (p-1)}$	$l(p-1)$	$n \cdot s^2_{\text{matrix}} + s^2$
Precision	$MS = \frac{\sum_{i=1}^l \sum_{j=1}^p \sum_{k=1}^n (r_{ijk} - r_{ij})^2}{l \cdot p \cdot (n-1)}$	$l \cdot p \cdot (n-1)$	$\sigma^2$

**Table 4.** Calculation of variances for the experimental design proposed in Fig. 1.

Variance	Expression	Degrees of freedom
Variance of precision, $s^2$	$MS$	$l \cdot p \cdot (n-1)$
Matrix variance, $s_{\text{matrix}}^2$	$\frac{MS_{\text{matrix}} - MS}{n}$	

Tables 3 and 4 show the expressions for calculating these variances. The variance,  $s^2$ , depends on how the spiked samples are analysed: a repeatability variance,  $s_r^2$ , is obtained if they are analysed under repeatability conditions and an intermediate variance,  $s_i^2$ , is obtained if they are analysed under intermediate precision conditions.

Fig. 2 shows that the SAM results can be expressed either as instrumental responses or, if a standard curve is used, as concentrations. Therefore, the SAM can be performed by plotting either the instrumental response versus the concentration added or the concentration found versus the concentration added.



**Figure 2.** Two ways of performing the Standard Addition Method: 1) plotting the instrumental response against the concentration added; and 2) plotting the concentration found against the concentration added.

**4.1.1. The instrumental response is plotted against the concentration added**

The slope of the SAM curve,  $b_{\text{SAM}}$ , is an estimate of the product of the sensitivity of the analytical procedure (which corresponds to the slope of a standard curve,  $b_{\text{SC}}$ ) and the “method recovery”,  $R$ , i.e.  $b_{\text{SAM}} = R \cdot b_{\text{SC}}$ . This does not verify whether the proportional bias is significant or not.

**4.1.2. The concentration found is plotted against the concentration added**

In this case, the slope of the SAM curve is an estimate of the “method recovery” [11] because the spiked samples cover the set of matrices and concentrations of the routine samples representatively. Once we have obtained the “method recovery”,  $R$ , and its uncertainty,  $u(R)$ , we can evaluate whether the proportional bias is significant or not. Proportional bias is not significant if:

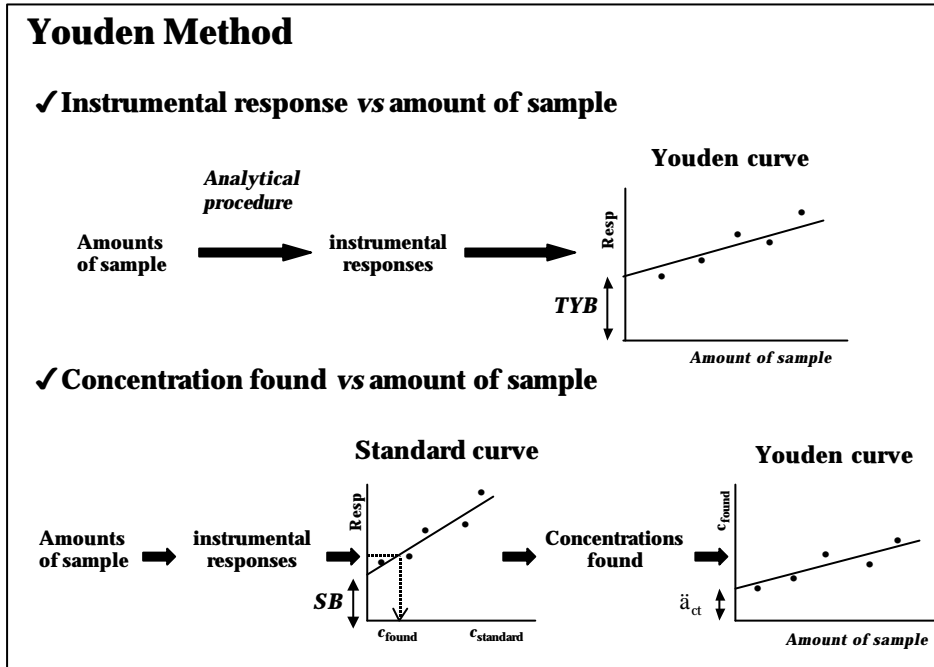
$$|R - 1| \leq t_{\alpha/2, \text{eff}} \cdot u(R) \quad (2)$$

where  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value for the effective degrees of freedom [15] associated with  $u(R)$ , and can be replaced by the coverage factor  $k$ . The uncertainty  $u(R)$  is calculated in Section 5.2.1.2.

**4.2. Youden Method. Calculation of constant bias**

The Youden Method [13] analyses various amounts ( $w_i$ ) of a test sample under conditions of repeatability or intermediate precision. An instrumental response or, if a standard curve is used, a concentration is found for each amount of sample. Fig. 3 shows that the Youden plot can then be defined as a sample response curve, i.e. instrumental responses plotted against sample amounts, or as a sample concentration curve, i.e. concentrations found plotted against sample amounts. The Youden Method provides a good estimate of constant bias whenever the matrix effect is the same for all amounts of sample. This can be assumed if the variance of the residuals of the Youden plot does not differ significantly from the repeatability variance (if the amounts are analysed under repeatability conditions) or from the

intermediate variance (if they are analysed under intermediate precision conditions).



**Figure 3.** Two ways of performing the Youden Method: 1) plotting the instrumental response against the amount of sample; and 2) plotting the concentration found against the amount of sample.

**4.2.1. Sample response curve**

The intercept of the Youden plot is an estimate of the Total Youden Blank (*TYB*) [17]. This is known as the true sample blank because it is determined when both the native analyte and the matrix are present. The *TYB* is the sum of the System Blank (*SB*) (which corresponds to the intercept of a standard curve) and the Youden Blank (*YB*) (associated with constant bias, i.e.  $\ddot{a}_{ct} = YB/b_{SC}$ ), i.e.  $TYB=SB+YB$ . As in Section 4.1.1, this does not verify whether the constant bias of the analytical procedure is significant or not.

#### **4.2.2. Sample concentration curve**

Here, the intercept of the Youden plot is an estimate of the constant bias. Once we have obtained the constant bias,  $\delta_{ct}$ , and its uncertainty,  $u(\delta_{ct})$ , we can evaluate whether the constant bias is significant or not. Constant bias is not significant if:

$$|\mathbf{d}_{ct}| \leq t_{\alpha/2, eff} \cdot u(\mathbf{d}_{ct}) \quad (3)$$

where  $t_{\alpha/2, eff}$  is the two-sided  $t$  tabulated value for the effective degrees of freedom [15] associated with  $u(\delta_{ct})$ , and can be replaced by the coverage factor  $k$ . The uncertainty of constant bias is calculated in Section 5.2.1.3.

### **5. UNCERTAINTY OF FUTURE ROUTINE SAMPLES**

Once constant and proportional biases of the analytical procedure have been evaluated, this information can be used to calculate the uncertainty of all the future routine samples determined with this procedure. This is possible when the samples used to assess trueness are representative of the future routine samples and when the analytical procedure is under statistical control. The expression for calculating uncertainty will depend on whether the results of routine samples are expressed in terms of instrumental response or in terms of concentration (if a standard curve is used). Moreover, in both situations, the expression also depends on the conditions in which the samples have been analysed in the assessment of trueness, i.e. repeatability or intermediate conditions.

#### **5.1. The results of routine samples are expressed as instrumental responses**

Here the estimated concentration of a routine sample, *conc*, is obtained as:

$$conc = \frac{Resp - TYB}{b_{SAM}} \quad (4)$$

where  $Resp$  is the instrumental response of the routine sample,  $b_{SAM}$  is the slope of the SAM curve obtained by plotting the instrumental response against the concentration added (Section 4.1.1), and  $TYB$  is the intercept of the Youden curve obtained by plotting the instrumental response against the amount of sample (Section 4.2.1).

The concentration,  $conc$ , is corrected by constant and proportional bias because, as we explained in Section 4, the  $TYB$  and the slope  $b_{SAM}$  are associated with constant and proportional bias, respectively.

The uncertainty of this concentration is obtained by applying the error propagation law to Eq. 4:

$$u = \frac{1}{b_{MOSA}} \cdot \sqrt{(conc \cdot u(b_{SAM}))^2 + u(TYB)^2 + u(Resp)^2} \quad (5)$$

where  $u$  is the standard uncertainty. The first two terms of this equation consider the uncertainty associated with the assessment of trueness,  $u_{trueness}$ .  $u(b_{SAM})$ ,  $u(TYB)$  and  $u(Resp)$  are the uncertainties associated with the slope of the standard additions curve, the Total Youden Blank and the precision,  $u_{precision}$ , respectively.

### 5.1.1. Practical estimation of the components of uncertainty

5.1.1.1. *Uncertainty of the instrumental response,  $u(Resp)$ .* This uncertainty is estimated as the sum of two components:

$$u(Resp) = \sqrt{s_{matrix}^2 + s_I^2} \quad (6)$$

The first component,  $s_{matrix}^2$ , is the between-matrix variance, which takes into account the fact that results vary according to the matrix of the routine sample; the second component,  $s_I^2$ , is the intermediate variance, which considers the variation

of the analytical results in the normal analysis of the routine samples. The first component can be estimated from the results of the spiked samples. The second component can also be estimated from these results if the spiked samples are analysed under intermediate precision conditions. Otherwise, it can be estimated from precision studies. Tables 3 and 4 show the expressions for calculating these variances from the results of the assessment of trueness with spiked samples.

5.1.1.2. *Uncertainty of the slope of the SAM curve,  $u(b_{SAM})$ .* This is estimated as the standard deviation of the slope of the standard additions curve (see Appendix A).

$$u(b_{SAM}) = s(b_{SAM(resp)}) \quad (7)$$

The expression for calculating this uncertainty does not depend on the conditions in which the spiked samples are analysed, i.e. under repeatability or intermediate conditions.

5.1.1.3. *Uncertainty of the Total Youden Blank,  $u(TYB)$ .* This uncertainty is calculated as:

$$u(TYB) = \sqrt{s(a_{YOU(resp)})^2 + u_{conditions}^2} \quad (8)$$

where  $s(a_{YOU(resp)})$  is the standard deviation of the intercept of the Youden curve (see Appendix A).  $u_{conditions}$  is the uncertainty associated with how the sample amounts are analysed. If they are analysed under intermediate conditions,  $u_{conditions}=0$ . If they are analysed under repeatability conditions,  $u_{conditions}=s_{run}$  (where  $s_{run}$  is the between-run standard deviation, which can be obtained from precision studies).

## 5.2. The results of routine samples are expressed as concentration

The concentration of the routine samples, *conc*, is obtained by using a standard calibration curve. This concentration is calculated as:

$$conc = \frac{c_{\text{found}} - \mathbf{d}_{\text{ct}}}{R} \quad (9)$$

where  $c_{\text{found}}$  is the concentration found with the analytical procedure after having converted the instrumental response into concentration with a standard curve,  $R$  is the “method recovery” and  $\mathbf{d}_{\text{ct}}$  is the constant bias. The concentration of future samples is obtained by correcting results by constant and proportional bias according to Eq. 9. However, some authors [18] suggest including the uncertainty of these two types of bias, not correcting analytical results and letting  $conc = c_{\text{found}}$ .

The standard uncertainty of the concentration,  $u$ , is obtained by applying the error propagation law to Eq. 9:

$$u = \frac{1}{R} \cdot \sqrt{(conc \cdot u(R))^2 + u(\mathbf{d}_{\text{ct}})^2 + u(c_{\text{found}})^2} \quad (10)$$

The first two terms of this equation consider the uncertainty associated with the assessment of trueness,  $u_{\text{trueness}}$ :  $u(R)$  is the uncertainty of the “method recovery” and  $u(\mathbf{d}_{\text{ct}})$  is the uncertainty of constant bias. The third term,  $u(c_{\text{found}})$ , is the uncertainty of the concentration found for the routine sample with the standard curve and considers the uncertainty associated with precision,  $u_{\text{precision}}$ .

### 5.2.1. Practical estimation of the components of uncertainty

5.2.1.1. *Uncertainty of the concentration found,  $u(c_{\text{found}})$ .* This uncertainty is the sum of two components:

$$u(c_{\text{found}}) = \sqrt{\left(\frac{S_{\text{matrix}}}{b_{\text{SC}}}\right)^2 + u(\text{pred})^2} \quad (11)$$

The first component takes into account the fact that results vary according to the matrix of the routine sample. The second component,  $u(\text{pred})$ , is the uncertainty of predicting a concentration with a standard curve. Tables 3 and 4 show the



expression for calculating the between-matrix variance,  $s^2_{\text{matrix}}$ . The uncertainty arising from prediction,  $u(\text{pred})$ , is calculated as:

$$u(\text{pred}) = \frac{1}{b_{\text{SC}}} \cdot \sqrt{s_1^2 + s_{e,\text{SC}}^2 \cdot \left( \frac{1}{n_{\text{SC}}} + \frac{(c_{\text{found}} - \bar{c})^2}{\sum (c_i - \bar{c})^2} \right) + u_{\text{conditions}}^2} \quad (12)$$

where  $s_{e,\text{SC}}$  is the residual standard deviation of the standard curve,  $c_i$  is the concentration of the standard  $i$ ,  $n_{\text{SC}}$  is the number of standards and  $\bar{c}$  is the mean of all the standards analysed.  $s_1^2$  is the intermediate variance, which can be estimated from the results of the spiked samples if they are analysed under intermediate conditions. Otherwise, it can be estimated from precision studies. Finally,  $u_{\text{conditions}}$  is the uncertainty associated with how the standards of the standard curve are analysed. If they are analysed under intermediate conditions,  $u_{\text{conditions}}=0$ . If they are analysed under repeatability conditions,  $u_{\text{conditions}}=s_{\text{run}}$ .

5.2.1.2. *Uncertainty of the “method recovery”,  $u(R)$ .* This is obtained as the sum of two components:

$$u(R) = \sqrt{s(b_{\text{SAM}(\text{conc})})^2 + \left( \frac{R}{b_{\text{SC}}} \right)^2 \cdot s(b_{\text{SC}})^2} \quad (13)$$

The first component considers the uncertainty of the slope of the standard additions curve and is calculated as the standard deviation of the slope of the standard additions curve (Section 4.1.2). The expression for calculating this standard deviation is shown in Appendix A. The second term considers the uncertainty of the error committed when the instrumental responses of the spiked samples are converted into concentrations with a standard curve. The expression for calculating the standard deviation of the slope of the standard curve,  $s(b_{\text{SC}})$ , is shown in Appendix A. The expression for calculating the uncertainty of recovery does not depend on the conditions in which the spiked samples are analysed, i.e. under repeatability or intermediate conditions.

5.2.1.3. *Uncertainty of the constant bias.* This uncertainty is calculated as:

$$u(\ddot{a}_{ct}) = \sqrt{s(a_{YOU(conc)})^2 + u_{conditions}^2 + u(SC)^2} \quad (14)$$

where  $s(a_{YOU(conc)})$  is the standard deviation of the intercept of the Youden curve obtained when concentration is plotted against the amount of sample (section 4.2.2, see Appendix A).  $u_{conditions}$  is the uncertainty associated with how the amounts of sample and the standards of the standard curve are analysed. If they are analysed under intermediate conditions,  $u_{conditions}=0$ . If they are analysed under repeatability conditions,  $u_{conditions}=s_{run} \cdot \sqrt{2}$ . Finally,  $u(SC)$  is the uncertainty associated with converting the instrumental responses of the amounts analysed into the concentration found, using the standard curve. It is calculated as:

$$u(SC) = \sqrt{\frac{s(a_{SC})^2}{b_{SC}^2} + \frac{s(b_{SC})^2 \cdot a_{YOU}^2}{b_{SC}^2} - 2 \cdot \frac{a_{YOU}}{b_{SC}^2} \cdot \text{cov}(a_{SC}, b_{SC})} \quad (15)$$

where  $s(a_{SC})$  is the standard deviation of the intercept of the standard curve,  $s(b_{SC})$  is the uncertainty of the slope of the standard curve,  $\text{cov}(a_{SC}, b_{SC})$  is the covariance of the intercept and the slope of the standard curve and  $a_{YOU}$  is the intercept of the Youden concentration curve. These expressions are shown in Appendix A.

## 6. PRACTICAL EXAMPLE: ESTIMATING UNCERTAINTY WHEN DETERMINING ESTERS IN WINE BY SPME-GC

To illustrate the procedures described in this paper, validation data of an analytical procedure for determining esters in wine [14] were used to calculate the uncertainty of future routine samples. The procedure consisted of extracting the esters by solid phase microextraction (SPME) and analysing them with a Hewlett-Packard 5890 (series II) gas chromatograph equipped with an HP-5972 mass-selective detector.

Red, white and rosé wine were spiked with the following esters: ethyl acetate, isobutyl acetate, ethyl butyrate, isoamyl acetate, ethyl hexanoate, hexyl acetate, ethyl octanoate, ethyl decanoate and 2-phenylethyl acetate. The samples were analysed three times before the esters were added and twice after. All samples were analysed under intermediate precision conditions. Methyl octanoate was used as internal standard. The Youden Method was used to know whether the procedure had a constant bias.

## 6.1. Assessment of trueness

### 6.1.1. Standard Addition Method (SAM). Estimating proportional bias

The analysis of the spiked samples provides information about proportional bias and precision. Table 5 shows the values of the between-matrix and the intermediate variances calculated with the expressions given in Tables 3 and 4.

**Table 5.** Between-matrix and intermediate variances calculated for each of the esters added in red, white and rosé wine. The expressions used for calculating these variances are given in Tables 3 and 4.

<b>Ester</b>	<b>Intermediate variance, <math>s_1^2</math></b>	<b>Between-matrix variance, <math>s_{\text{matrix}}^2</math></b>
Ethyl acetate	$1.32 \cdot 10^{-2}$	$5.28 \cdot 10^{-2}$
Isobutyl acetate	$2.40 \cdot 10^{-5}$	$1.98 \cdot 10^{-4}$
Ethyl butyrate	$3.77 \cdot 10^{-4}$	$8.26 \cdot 10^{-4}$
Isoamyl acetate	$2.91 \cdot 10^{-2}$	$3.24 \cdot 10^{-2}$
Ethyl hexanoate	$1.00 \cdot 10^{-1}$	$2.57 \cdot 10^{-2}$
Hexyl acetate	$7.01 \cdot 10^{-3}$	$2.81 \cdot 10^{-3}$
Ethyl octanoate	1.55	1.51
Ethyl decanoate	1.86	5.34
2-phenylethyl acetate	$2.09 \cdot 10^{-4}$	$1.79 \cdot 10^{-4}$

Each result was calculated as the difference between the result after analysing the spiked sample once and the average result after analysing the sample containing the native esters three times.

The results were expressed as instrumental responses and, after using a calibration curve, as concentration found. Table 6 shows the slope,  $b_{\text{SAM}}$ , and its uncertainty,

$u(b_{SAM})$ , obtained by plotting the instrumental response against the concentration added. Table 7 shows the recovery,  $R$ , and its uncertainty,  $u(R)$ , obtained when analytical results are expressed as concentration found.

**Table 6.** Slope,  $b_{SAM}$ , and its uncertainty,  $u(b_{SAM})$ , of the standard additions curve (Section 4.1.1) for each of the esters spiked. The Total Youden Blank and its uncertainty,  $u(TYB)$ , obtained with the Youden curve (Section 4.2.1) are also shown.

Ester	$b_{SAM}$	Uncertainty $u(b_{SAM})$	Total Youden Blank, TYB	Uncertainty, $u(TYB)$
Ethyl acetate	$1.21 \cdot 10^{-3}$	$7.18 \cdot 10^{-5}$	$2.09 \cdot 10^{-2}$	$3.06 \cdot 10^{-2}$
Isobuthyl acetate	$1.98 \cdot 10^{-2}$	$6.67 \cdot 10^{-4}$	$1.37 \cdot 10^{-2}$	$1.07 \cdot 10^{-2}$
Ethyl butirate	$2.39 \cdot 10^{-2}$	$1.39 \cdot 10^{-3}$	$1.15 \cdot 10^{-1}$	$7.67 \cdot 10^{-3}$
Isoamyl acetate	$6.58 \cdot 10^{-2}$	$3.79 \cdot 10^{-3}$	$1.08 \cdot 10^{-2}$	$5.72 \cdot 10^{-2}$
Ethyl hexanoate	$2.86 \cdot 10^{-1}$	$1.72 \cdot 10^{-2}$	$1.92 \cdot 10^{-2}$	$6.58 \cdot 10^{-2}$
Hexyl acetate	$2.68 \cdot 10^{-2}$	$8.64 \cdot 10^{-3}$	$-1.43 \cdot 10^{-1}$	$1.31 \cdot 10^{-1}$
Ethyl octanoate	1.58	$6.05 \cdot 10^{-2}$	$-7.70 \cdot 10^{-1}$	1.38
Ethyl decanoate	1.69	$1.31 \cdot 10^{-1}$	$4.88 \cdot 10^{-1}$	$5.56 \cdot 10^{-1}$
2-phenylethyl acetate	$2.94 \cdot 10^{-2}$	$8.29 \cdot 10^{-4}$	$-1.69 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$

**Table 7.** Recovery,  $R$ , and its uncertainty,  $u(R)$ , obtained with the standard additions curve (Section 4.1.2) and when recovery is calculated with the method of averaged recovery. The constant bias,  $d_{ct}$ , and its uncertainty,  $u(d_{ct})$ , obtained with the Youden curve (Section 4.2.2) are also shown.

	Standard additions		Average recovery		Youden curve	
	$R$	$u(R)$	$R$	$u(R)$	$d_{ct}$	$u(d_{ct})$
Ethyl acetate	1.05	$8.21 \cdot 10^{-2}$	0.91	$1.14 \cdot 10^{-1}$	40.45	102.18
Isobuthyl acetate	0.93	$5.45 \cdot 10^{-2}$	0.90	$7.59 \cdot 10^{-2}$	$8.04 \cdot 10^{-1}$	$6.75 \cdot 10^{-1}$
Ethyl butirate	0.97	$8.91 \cdot 10^{-2}$	0.87	$1.24 \cdot 10^{-1}$	$8.47 \cdot 10^{-1}$	$8.19 \cdot 10^{-1}$
Isoamyl acetate	0.92	$7.22 \cdot 10^{-2}$	0.82	$1.06 \cdot 10^{-1}$	$1.11 \cdot 10^{-1}$	1.25
Ethyl hexanoate	0.97	$7.10 \cdot 10^{-2}$	0.86	$1.03 \cdot 10^{-1}$	$2.88 \cdot 10^{-2}$	$2.86 \cdot 10^{-1}$
Hexyl acetate	0.91	$4.41 \cdot 10^{-2}$	0.87	$6.71 \cdot 10^{-2}$	$-8.02 \cdot 10^{-1}$	$5.29 \cdot 10^{-1}$
Ethyl octanoate	1.06	$3.02 \cdot 10^{-2}$	0.99	$3.95 \cdot 10^{-2}$	$-9.48 \cdot 10^{-2}$	1.32
Ethyl decanoate	1.01	$8.49 \cdot 10^{-2}$	1.01	$1.17 \cdot 10^{-1}$	$4.86 \cdot 10^{-2}$	$3.25 \cdot 10^{-1}$
2-phenylethyl acetate	1.00	$4.05 \cdot 10^{-2}$	0.97	$5.60 \cdot 10^{-2}$	$-5.04 \cdot 10^{-1}$	$7.09 \cdot 10^{-1}$

To compare the results of this method (with those of others), recovery was also calculated using the method of averaged recovery [11,12]. Recovery was calculated for each spiked sample and the overall recovery was then estimated as the mean of the  $n$  recoveries calculated. The uncertainty of this average recovery was calculated using the precision information from the results of the spiked samples:

$$u(R) = \frac{1}{n} \cdot \sqrt{\frac{S_1^2 + S_{\text{matrix}}^2}{b_{\text{SC}}^2 \cdot \sum_{i=1}^n c_{\text{ad},i}^2}} \quad (16)$$

Table 7 shows the average recovery and its uncertainty for each ester.

### 6.1.2. Youden Method. Estimating constant bias

The esters of five different amounts,  $w_i$ , of red wine were analysed under intermediate precision conditions. The analytical results were expressed as instrumental responses and, after using a calibration curve, as concentration found. Table 6 shows the Total Youden Blank,  $TYB$ , and its uncertainty,  $u(TYB)$ , when results are expressed as instrumental responses. Table 7 shows the constant bias,  $\delta_t$ , and its uncertainty,  $u(\delta_t)$ , when results are expressed as concentration found.

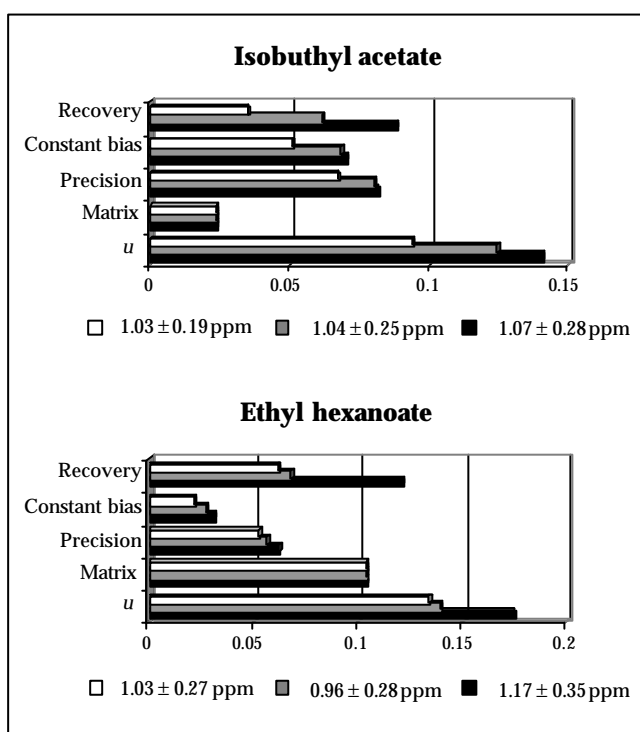
The variance of the residuals of the Youden plot was compared with the variance associated with the intermediate precision of the method. Since the difference between the variances was not statistically significant for the esters determined, we assumed that the matrix effect was the same for all the amounts of sample and, therefore, that a correct estimation of the constant bias was obtained from the Youden plot.

**Table 8.** Concentration, together with its uncertainty, obtained with the three procedures for all the esters analysed. Results are expressed as ppm.

	Regression techniques		Method of averaged recovery
	Instrumental responses	Concentration found	
Ethyl acetate	169±45	155±52	179±63
Isobutyl acetate	1.03±0.19	1.04±0.25	1.07±0.28
Ethyl butyrate	0.96±0.30	0.92±0.38	0.92±0.44
Isoamyl acetate	3.31±0.82	3.27±0.89	3.67±1.10
Ethyl hexanoate	1.03±0.27	0.96±0.28	1.17±0.35
Hexyl acetate	0.86±0.13	0.96±0.16	1.01±0.19
Ethyl octanoate	3.17±0.28	3.35±0.37	3.58±0.43
Ethyl decanoate	1.16±0.35	1.13±0.37	1.16±0.41
2-phenylethyl acetate	1.41±0.19	1.42±0.23	1.47±0.26

### 6.1.3. Uncertainty of future routine samples

This uncertainty was calculated in three different ways: a) when results are expressed as instrumental responses; b) when results are expressed as concentration; and c) when recovery was estimated with the method of averaged recovery. Fig. 4 shows the contribution made by recovery, constant bias, precision and matrix variability to the measurement uncertainty for the isobutyl acetate and the ethyl hexanoate. Fig. 4 also shows total uncertainty, expressed as standard uncertainty, and the concentration, together with its uncertainty obtained with the three procedures. Table 8 shows the concentration, together with its uncertainty, for all the esters and for the three procedures.



**Figure 4.** Contribution of the different components (recovery, constant bias, precision and matrix variability) to the measurement uncertainty of isobutyl acetate and ethyl hexanoate. *u* is the standard uncertainty, calculated in three different ways: □ results are expressed as instrumental responses; ■ results are expressed as concentration; ■ recovery is estimated with the method of averaged recovery. The concentrations, together with their associated uncertainties, are also shown for the three procedures .

## 7. DISCUSSION

Fig. 4 and Table 8 show that estimating recovery with the method of averaged recovery provides higher uncertainties than estimating it with either of the two regression procedures. Uncertainties are lowest when results are expressed as instrumental responses (situation a). The uncertainty obtained when the results are expressed as concentration (situation b) is higher because there is an additional uncertainty from converting the instrumental response into concentration with a standard curve. The difference between the two procedures depends on the uncertainty of the standard curves used to convert instrumental responses into concentration. This uncertainty depends on the precision of the analytical procedure: the higher the uncertainty due to precision, the higher the uncertainty of the standard curve. It also depends on the contribution to the measurement uncertainty of other terms such as matrix variability. If this contribution is high, the uncertainties of the two procedures are likely to be similar.

Fig. 4 shows that, for isobutyl acetate, situation a) produces slight lower uncertainty values than situation b). This is because the uncertainty of the standard curves used is not negligible since the uncertainty introduced by the component of precision is an important source of uncertainty. On the other hand, situations a) and b) provide almost the same results for ethyl hexanoate. This is because, for this ester, the matrix variability is the most important source of uncertainty. Fig. 4 shows that constant bias, which is not usually considered in the uncertainty budget [11], is not negligible for these esters and should therefore be taken into account.

When low uncertainties are important, it is better not to use a standard curve to convert responses into concentration, and to use Eq. 4 for quantification. However, the disadvantage of this procedure is that trueness must be reassessed when we suspect that the behaviour of the instrumental response has changed. This procedure is therefore suitable for analytical procedures in which the instrumental response does not change with time, as in the example in this paper, in which the use of internal standards makes the instrumental response highly stable. This is also true for HPLC methods without derivatization.

On the other hand, results should be converted to concentration found in analytical procedures in which the instrumental response changes with time. This, however, produces higher uncertainty values, as in HPLC methods, for example, that have reactions of derivatization.

## **8. CONCLUSIONS**

We have proposed a method for calculating the uncertainty of the results of routine samples using the information generated when the trueness of the analytical procedure is assessed using spiked samples and regression techniques. We recommend using it when the analytical procedure must be validated in a wide range of concentration. This method, which involves estimating the constant and proportional biases of the analytical procedure, produces lower uncertainties than other methods. Moreover, we have shown that the contribution of constant bias (usually neglected) to the final uncertainty budget is significant for the esters we studied. Future studies in this area will focus on calculating uncertainty when precision varies within the concentration range, when the uncertainty of the reference values is not negligible, and when blank samples are available.

## **ACKNOWLEDGEMENTS**

The authors would like to thank the DGICYT (project num. BP96-1008) for financial support and the CIRIT from the Catalan Government for providing Alicia Maroto's doctoral fellowship. Pilar Martí, from the Group of Wine and Food Analytical Chemistry of the Rovira i Virgili University is gratefully acknowledged for providing the experimental data.



---

**NOMENCLATURE**

$a_{SC}$	intercept of the standard curve
$a_{YOU(resp)}$	intercept of the Youden curve of Section 4.2.1
$b_{SAM(resp)}$	slope of the standard additions curve of Section 4.1.1
$b_{SC}$	slope of the standard curve
$c_i$	concentration of the standard $i$ used in the standard curve
$c_{ad,i}$	concentration added to the $i$ spiked sample
$\bar{c}_{ad}$	mean of all the concentrations added
$\bar{c}$	mean of all the standards used in the standard curve
$cov(a_{SC}, b_{SC})$	covariance of the intercept and the slope of the standard curve
$\delta_t$	constant bias
$n_{SC}$	number of standards used in the standard curve
$n_{YOU}$	number of amounts of sample analysed
$R$	recovery of the analytical procedure
$s(a_{SC})$	standard deviation of the intercept of the standard curve
$s(a_{YOU(resp)})$	standard deviation of the intercept of the Youden curve of Section 4.2.1
$s(b_{SAM(resp)})$	standard deviation of the slope of the standard additions curve of Section 4.1.1
$s(b_{SC})$	standard deviation of the slope of the standard curve
$s_{e,SAM(resp)}$	residual standard deviation of the standard additions curve of Section 4.1.1
$s_{e,SC}$	residual standard deviation of the standard curve
$s_{e,YOU(resp)}$	residual standard deviation of the Youden curve of Section 4.2.1
$s_I$	intermediate standard deviation
$s_{matrix}$	between-matrix standard deviation
$s_r$	repeatability standard deviation
$s_{run}$	between-run standard deviation
$SB$	System Blank
$TYB$	Total Youden Blank
$u(\delta_t)$	uncertainty of constant bias
$u(R)$	uncertainty of recovery
$w_i$	amount $i$ of sample analysed

$\bar{w}$	mean of all the amounts analysed
YB	Youden Blank

## APPENDIX A

Standard deviation of the slope of the standard additions curve of Section 4.1.1:

$$s(b_{\text{SAM}(\text{resp})}) = \frac{S_{e,\text{SAM}(\text{resp})}}{\sqrt{\sum (c_{\text{ad},i} - \bar{c}_{\text{ad}})^2}}$$

Standard deviation of the intercept of the Youden curve of Section 4.2.1:

$$s(a_{\text{YOU}(\text{resp})}) = S_{e,\text{You}(\text{resp})} \cdot \sqrt{\frac{\sum w_i^2}{n_{\text{You}} \cdot \sum (w_i - \bar{w})^2}}$$

Standard deviation of the slope of the standard additions curve of Section 4.1.2:

$$s(b_{\text{SAM}(\text{conc})}) = \sqrt{\frac{S_{e,\text{SAM}(\text{conc})}^2}{\sum (c_{\text{ad},i} - \bar{c}_{\text{ad}})^2}}$$

Standard deviation of the slope of the standard curve:

$$s(b_{\text{SC}}) = \frac{S_{e,\text{SC}}}{\sqrt{\sum (c_i - \bar{c})^2}}$$

Standard deviation of the intercept of the Youden curve of Section 4.2.2:

$$s(a_{\text{YOU}(\text{conc})}) = S_{e,\text{You}(\text{conc})} \cdot \sqrt{\frac{\sum w_i^2}{n_{\text{You}} \cdot \sum (w_i - \bar{w})^2}}$$

Standard deviation of the intercept of the standard curve:

$$s(a_{SC}) = s_{e,SC} \cdot \sqrt{\frac{\sum c_i^2}{n_{SC} \cdot \sum (c_i - \bar{c})^2}}$$

Covariance between the slope and the intercept of the standard curve:

$$\text{cov}(a_{SC}, b_{SC}) = \frac{\bar{c} \cdot s_{e,SC}^2}{\sum (c_i - \bar{c})^2}$$

## REFERENCES

- [1] BIPM, IEC, IFCC, IUPAC, IUPAP, OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993
- [2] EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK, 1995
- [3] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Anal. Chim. Acta* **391** (1999) 173-185
- [4] A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Trends Anal. Chem.* **18/9-10** (1999) 577-584
- [5] S.L.R. Ellison, A. Williams, *Accreditation Quality Assurance* **3** (1998) 6-10
- [6] V.J. Barwick, S.L.R. Ellison, *Accreditation Quality Assurance* **5** (2000) 47-53
- [7] V.J. Barwick, S.L.R. Ellison, M.J.K. Rafferty, R.S. Gill, *Accreditation Quality Assurance* **5** (2000) 104-113
- [8] EURACHEM, *Quantifying uncertainty in analytical measurement*, Draft: EURACHEM Workshop, 2<sup>nd</sup> Edition, Helsinki, 1999
- [9] S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Anal. Chim. Acta* **391** (1999) 203-225
- [10] S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Chemometr. Intell. Lab. Syst.* **52** (2000) 61-73
- [11] V.J. Barwick, S.L.R. Ellison, *Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data*, VAM Project 3.2.1, 2000
- [12] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Anal. Chim. Acta* **440** (2001) 171-184
- [13] W.J. Youden, *Anal. Chem.* **19** (1947) 946-950
- [14] P. Martí, O. Busto, J. Guasch, unpublished work
- [15] F.E. Satterthwaite, *Psychometrika* **6** (1941) 309-316
- [16] International Organization for Standardization, *ISO 5725. Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva, 1994
- [17] M. Cardone, *Anal. Chem.* **58** (1986) 438-445

- [18] International Union of Pure and Applied Chemistry, *Harmonised Guidelines for the use of recovery information in analytical measurement*, Resulting from the Symposium on Harmonisation of Internal Quality Assurance Systems for Analytical Laboratories, Orlando, USA, 1996

### **3.6. Cálculo de la incertidumbre en métodos de rutina validados en un intervalo amplio de concentraciones con blancos de muestras adicionados**

En este apartado se explica cómo calcular la incertidumbre utilizando la información generada al verificar la trazabilidad utilizando blancos de muestras adicionados. Al igual que cuando se adicionan muestras con concentración inicial de analito, el cálculo de la incertidumbre depende de cómo varía la precisión del método con la concentración. A continuación se detalla cómo calcular la incertidumbre cuando: 1) la precisión se mantiene aproximadamente constante en el intervalo de aplicación del método; y 2) cuando la precisión varía linealmente en el intervalo de aplicación del método.

#### **3.6.1. La precisión se mantiene aproximadamente constante en el intervalo de aplicación del método**

Para calcular la incertidumbre utilizando blancos de muestras adicionados, es recomendable seguir el diseño experimental propuesto en la Fig. 3.1 para realizar los estudios de recuperación. A partir del ANOVA de los resultados obtenidos en los estudios de recuperación puede obtenerse la incertidumbre debida al procedimiento, a la verificación de la trazabilidad y a la variabilidad de las matrices de las muestras de rutina. A continuación se detalla la metodología que debe seguirse para calcular estos términos de incertidumbre:

##### *ANOVA de las muestras adicionadas*

La incertidumbre del procedimiento y de la variabilidad de las matrices de las muestras de rutina se puede calcular a partir del ANOVA de las concentraciones encontradas obtenidas al analizar las muestras adicionadas. Para ello, es necesario que las adiciones se realicen siguiendo el diseño propuesto en la Fig. 3.1. La Tabla 3.1 muestra la Tabla ANOVA asociada al diseño experimental de la Fig. 3.1 y la Tabla 3.2 las varianzas que pueden calcularse a partir del ANOVA.

**Tabla 3.1.** Tabla ANOVA para calcular la incertidumbre de la matriz y del procedimiento a partir del diseño experimental de la Fig. 3.1.

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Matriz	$MS_{\text{matriz}} = \frac{2 \cdot \sum_{i=1}^p \sum_{j=1}^q (c_{e,ij} - \bar{c}_{e,i})^2}{p \cdot (q-1)}$	$p \cdot (q-1)$	$2 \cdot s_{\text{matriz}}^2 + s_1^2$
Precisión	$MS_1 = \frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^2 (c_{e,ijk} - c_{e,ij})^2}{p \cdot q \cdot (2-1)}$	$p \cdot q \cdot (2-1)$	$s_1^2$

**Tabla 3.2.** Cálculo de las varianzas de la matriz y de la precisión intermedia a partir del ANOVA de la Tabla 3.1.

Varianza	Expresión
Precisión intermedia, $s_1^2$	$MS_1$
Matriz, $s_{\text{matriz}}^2$	$\frac{MS_{\text{matriz}} - MS_1}{2}$

Incertidumbre del procedimiento

La incertidumbre del procedimiento se calcula a partir de la varianza de la precisión,  $s_1^2$ , obtenida en la Tabla 3.2. Esta varianza corresponde a la precisión intermedia del método ya que las muestras adicionadas se han analizado en condiciones intermedias. Por tanto, la incertidumbre del procedimiento puede calcularse utilizando la Ec. 3.23. Como la precisión del método se mantiene aproximadamente constante con la concentración, la incertidumbre del procedimiento es la misma para todo el intervalo de aplicación del método.

Incertidumbre de la variabilidad de las matrices de las muestras de rutina

La incertidumbre de la variabilidad de las muestras de rutina puede calcularse utilizando la información obtenida al analizar las muestras adicionadas ya que el analito se adiciona a diferentes tipos de muestras que son representativas de la

variabilidad de las muestras de rutina. Esta incertidumbre se calcula a partir de la varianza de la matriz,  $s_{\text{matriz}}^2$ , obtenida en la Tabla 3.2:

$$u_{\text{matriz}} = \frac{s_{\text{matriz}}}{R} \quad (3.32)$$

#### Incertidumbre de la verificación de la trazabilidad

Como la precisión es aproximadamente constante, la recta de adiciones estándar se obtiene aplicando la regresión por mínimos cuadrados clásicos (OLS). Tal y como se menciona en el apartado 3.2.1.2, la pendiente de la recta proporciona una estimación de la recuperación del método,  $R$ , y la ordenada de la recta proporciona una estimación del sesgo constante,  $\delta_{\text{cte}}$ . Para calcular la incertidumbre de la verificación de la trazabilidad, debe tenerse en cuenta la correlación que hay entre la pendiente y la ordenada de una recta ( $\text{cov}(a,b) \approx \text{cov}(\delta_{\text{cte}}, R)$ ). Por tanto, debe añadirse a la Ec. 3.24 un término que tenga en cuenta esta correlación:

$$u_{\text{traz}} = \frac{1}{R} \sqrt{u(\hat{\delta}_{\text{cte}})^2 + (c_{\text{corr}} \cdot u(R))^2 - 2 \cdot c_{\text{corr}} \cdot \text{cov}(\hat{\delta}_{\text{cte}}, R)} \quad (3.33)$$

donde  $u(\hat{\delta}_{\text{cte}})$  es la desviación estándar de la ordenada en el origen,  $s_a$ ;  $u(R)$  es la desviación estándar de la pendiente,  $s_b$ ; y  $\text{cov}(\hat{\delta}_{\text{cte}}, R)$  es la covarianza entre la pendiente y la ordenada en el origen. Esta expresión puede simplificarse a<sup>17</sup>:

$$u_{\text{traz}} = \frac{s_e}{R} \sqrt{\frac{1}{n} + \frac{(c_{\text{corr}} - \bar{c}_a)^2}{\sum \sum (c_{a,ij} - \bar{c}_a)^2}} \quad (3.34)$$

donde  $s_e$  es la desviación estándar de los residuales de la recta,  $c_{a,ij}$  es la concentración adicionada  $i$  al tipo de muestra  $j$ ,  $\bar{c}_a$  es el valor medio de todos los  $c_{a,ij}$  y  $n$  es el número de puntos de la recta de adiciones.

Cálculo de la incertidumbre total

La incertidumbre de los resultados se obtiene sustituyendo las ecuaciones 3.23, 3.32 y 3.34 en la expresión propuesta para el cálculo global de la incertidumbre (Ec. 3.12):

$$u(c_{\text{corr}}) = \frac{1}{R} \sqrt{\frac{s_{\text{I}}^2}{p_s} + s_e^2 \cdot \left( \frac{1}{n} + \frac{(c_{\text{corr}} - \bar{c}_a)^2}{\sum \sum (c_{a,ij} - \bar{c}_a)^2} \right) + s_{\text{matriz}}^2} \quad (3.35)$$

El primer término corresponde a la incertidumbre del procedimiento, el segundo a la incertidumbre de la verificación de la trazabilidad y el tercero a la incertidumbre de otros términos debidos a la variabilidad de las matrices de las muestras de rutina. En esta expresión se ha supuesto que no hay incertidumbre debida al submuestreo y/o pretratamientos. En el caso de que fuera necesario incluir este término de incertidumbre, debería utilizarse la metodología propuesta en el apartado 3.3. Asimismo, se ha supuesto que no hay otros términos de incertidumbre debidos a niveles más altos de trazabilidad o a factores que no se han variado representativamente en el cálculo de la precisión intermedia. Tal y como se menciona en el apartado 3.3, la incertidumbre expandida se obtiene multiplicando la incertidumbre estándar,  $u$ , por el factor de cobertura,  $k=2$ . No obstante, si los distintos términos de incertidumbre tienen pocos grados de libertad asociados es mejor utilizar el valor de  $t$  tabulado para  $\alpha=0.05$  y los grados de libertad efectivos calculados con la aproximación de Welch-Satterthwaite<sup>27</sup>.

**3.6.2. La precisión varía linealmente en el intervalo de aplicación del método**

En este apartado se calcula la incertidumbre asumiendo que la  $RSD$  del método se mantiene aproximadamente constante en el intervalo de aplicación del método. Al igual que en el apartado 3.6.1, la incertidumbre se calcula utilizando la información obtenida cuando se sigue el diseño propuesto en la Fig. 3.1 para realizar los estudios de recuperación. A partir del ANOVA de los resultados obtenidos en los estudios de recuperación puede obtenerse la incertidumbre debida al



procedimiento, a la verificación de la trazabilidad y a la variabilidad de las matrices de las muestras de rutina. A continuación se detalla la metodología que debe seguirse para calcular estos términos de incertidumbre:

#### ANOVA de las muestras adicionadas

Para realizar el ANOVA es necesario que los resultados sean homoscedásticos, es decir, que siempre tengan asociada la misma varianza. En este caso, los resultados son heteroscedásticos ya que la precisión varía linealmente con la concentración. Como la desviación estándar relativa, *RSD*, sí que se mantiene constante, se puede realizar el ANOVA de las recuperaciones obtenidas para cada muestra adicionada. De esta forma, los resultados sí que son homoscedásticos. La Tabla 3.3 muestra la tabla ANOVA y la Tabla 3.4 muestra las *RSD* que pueden calcularse a partir del ANOVA.

**Tabla 3.3.** Tabla ANOVA para calcular la incertidumbre de la matriz y del procedimiento a partir del diseño experimental propuesto en la Fig. 3.1.

Fuente	Cuadrados medios	g.l.	Cuadrados medios esperados
Matriz	$MS_{\text{matriz}} = \frac{2 \cdot \sum_{i=1}^p \sum_{j=1}^q (R_{ij} - \bar{R}_i)^2}{p \cdot (q-1)}$	$p \cdot (q-1)$	$2 \cdot RSD_{\text{matriz}}^2 + RSD_1^2$
Precisión	$MS_1 = \frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^2 (R_{ijk} - R_{ij})^2}{p \cdot q \cdot (2-1)}$	$p \cdot q \cdot (2-1)$	$RSD_1^2$

**Tabla 3.4.** Cálculo de la *RSD* de la matriz y de la precisión intermedia a partir del ANOVA de la Tabla 3.3.

RSD	Expresión
Precisión intermedia, $RSD_1$	$\sqrt{MS_1}$
Matriz, $RSD_{\text{matriz}}$	$\sqrt{\frac{MS_{\text{matriz}} - MS_1}{2}}$

Incertidumbre del procedimiento

Como las muestras adicionadas se han analizado en condiciones intermedias, la incertidumbre del procedimiento puede calcularse utilizando la desviación estándar relativa,  $RSD_I$ , obtenida en la Tabla 3.4:

$$u_{\text{proc}} = \frac{1}{R} \sqrt{\frac{(c_{\text{corr}} \cdot RSD_I)^2}{p_s}} \quad (3.36)$$

donde  $p_s$  es el número de veces que se analiza la muestra de rutina en condiciones intermedias (normalmente,  $p_s=1$ ).

Incertidumbre de la variabilidad de las matrices de las muestras de rutina

La incertidumbre de la variabilidad de las muestras de rutina puede calcularse utilizando la información obtenida al analizar las muestras adicionadas ya que el analito se adiciona a diferentes tipos de muestras que son representativas de la variabilidad de las muestras de rutina. Se calcula a partir de la  $RSD$  de la matriz,  $RSD_{\text{matriz}}$ , obtenida en la Tabla 3.4:

$$u_{\text{matriz}} = \frac{c_{\text{corr}} \cdot RSD_{\text{matriz}}}{R} \quad (3.37)$$

Incertidumbre de la verificación de la trazabilidad

Como la precisión varía con la concentración, la recta de adiciones estándar debe obtenerse aplicando la regresión por mínimos cuadrados ponderados (WLS). Para ello, es necesario calcular los pesos de cada una de las muestras adicionadas. Estos pesos pueden calcularse utilizando la información generada en el ANOVA. Si cada una de las muestras adicionadas se ha analizado  $n$  veces (donde  $n$  generalmente es 2) en condiciones intermedias, el peso de cada muestra puede calcularse como:

$$w_{ij} = \frac{1}{\left( c_{a,ij} \cdot \left( RSD_{\text{matriz}} + \frac{RSD_I}{n} \right) \right)^2} \quad (3.38)$$

Tal y como se menciona en el apartado 3.2.1.2, la pendiente de la recta proporciona una estimación de la recuperación del método y la ordenada de la recta proporciona una estimación del sesgo constante. Al igual que en el apartado 3.6.1, la incertidumbre de la verificación de la trazabilidad debe calcularse considerando la correlación que hay entre la pendiente y la ordenada de una recta. Por tanto, debe utilizarse la Ec. 3.33. En el caso de una recta WLS, esta ecuación se expresa como<sup>17</sup>:

$$u_{\text{traz}} = \frac{s_e}{R} \sqrt{\frac{1}{\sum \sum w_{ij}} + \frac{\left( c_{\text{corr}} - \frac{\sum \sum w_{ij} \cdot c_{a,ij}}{\sum \sum w_{ij}} \right)^2}{\sum \sum w_{ij} \cdot \sum \sum w_{ij} \cdot c_{a,ij}^2 - \left( \sum \sum w_{ij} \cdot c_{a,ij} \right)^2}} \quad (3.39)$$

donde  $c_{a,ij}$  es la concentración adicionada  $i$  al tipo de muestra  $j$  y  $s_e$  es la desviación estándar de los residuales de la recta y se calcula como:

$$s_e = \sqrt{\frac{\sum \sum w_{ij} \cdot (c_{e,ij} - \hat{c}_{e,ij})^2}{n - 2}} \quad (3.40)$$

donde  $c_{e,ij}$  es la concentración obtenida con el método al analizar la muestra adicionada  $c_{a,ij}$ ,  $\hat{c}_{e,ij}$  es la concentración predicha con la recta para la cantidad adicionada  $c_{a,ij}$  y  $n$  es el número de puntos de la recta de adiciones.

#### Cálculo de la incertidumbre total

La incertidumbre de los resultados se obtiene sustituyendo las ecuaciones 3.36, 3.37 y 3.39 en la expresión propuesta para el cálculo global de la incertidumbre (Ec. 3.12):

$$u(c_{\text{corr}}) = \frac{c_{\text{corr}}}{R} \sqrt{\frac{RSD_I^2}{p_s} + \frac{s_e^2}{c_{\text{corr}}^2} \left( \frac{1}{\sum \sum w_{ij}} + \frac{\left( c_{\text{corr}} - \frac{\sum \sum w_{ij} \cdot c_{a,ij}}{\sum \sum w_{ij}} \right)^2}{\sum \sum w_{ij} \cdot \sum \sum w_{ij} \cdot c_{a,ij}^2 - \left( \sum \sum w_{ij} \cdot c_{a,ij} \right)^2} \right)} + RSD_{\text{matriz}}^2} \quad (3.41)$$

El primer término corresponde a la incertidumbre del procedimiento, el segundo a la incertidumbre de la verificación de la trazabilidad y el tercero a la incertidumbre de otros términos debidos a la variabilidad de las matrices de las muestras de rutina. Al igual que en la Ec. 3.35, se ha supuesto que no hay incertidumbre debida al submuestreo y/o pretratamientos. En el caso de que fuera necesario incluir este término de incertidumbre, debería utilizarse la metodología propuesta en el apartado 3.3. Asimismo, se ha supuesto que no hay otros términos de incertidumbre debidos a niveles más altos de trazabilidad o a factores que no se han variado representativamente en el cálculo de la precisión intermedia. Tal y como se menciona en el apartado 3.3, la incertidumbre expandida se obtiene multiplicando la incertidumbre estándar,  $u$ , por el factor de cobertura,  $k=2$ . No obstante, si los distintos términos de incertidumbre tienen pocos grados de libertad asociados es mejor utilizar el valor de  $t$  tabulado para  $\alpha=0.05$  y los grados de libertad efectivos calculados con la aproximación de Welch-Satterthwaite<sup>27</sup>.

### 3.7. Conclusiones

En este capítulo se ha presentado una aproximación para calcular globalmente la incertidumbre de métodos analíticos cuya exactitud se ha verificado en un intervalo amplio de concentraciones utilizando muestras adicionadas. Esta aproximación es conceptualmente similar a la presentada en el capítulo anterior ya que está basada en calcular la incertidumbre utilizando sobre todo la información obtenida al verificar la trazabilidad de los resultados.

La trazabilidad de los resultados puede verificarse utilizando únicamente muestras adicionadas si el analito se adiciona a blancos de muestras. En caso contrario, es necesario utilizar el método de Youden para verificar la ausencia de un sesgo constante. Por otro lado, el cálculo de la recuperación del método también depende de que el analito se adicione a blancos de muestras o bien a muestras con concentración inicial. Si el analito se adiciona a blancos de muestras, la recuperación debe calcularse con el método de adiciones estándar. En caso contrario, la recuperación puede calcularse con el método de adiciones estándar o bien con el método de la recuperación media. Es aconsejable utilizar el método de la recuperación media cuando la precisión varía con la concentración ya que este método proporciona incertidumbres menores que el de adiciones estándar.

La aproximación propuesta para calcular la incertidumbre en un intervalo amplio de concentraciones depende de cómo se verifica la trazabilidad de los resultados y de cómo varía la precisión con la concentración. En este capítulo se ha estudiado cómo calcular la incertidumbre cuando el analito se adiciona a blancos de muestras o bien a muestras que ya contienen el analito. En ambos casos, se han propuesto expresiones cuando la precisión varía linealmente con la concentración o bien cuando se mantiene constante. Estas expresiones se han obtenido suponiendo que la incertidumbre de la cantidad adicionada es despreciable.

Las expresiones propuestas cuando la trazabilidad se verifica utilizando blancos de muestras adicionados pueden aplicarse para calcular la incertidumbre cuando la trazabilidad se verifica con otras referencias (como, por ejemplo, materiales de referencia, métodos de referencia o bien cuando se adiciona un isótopo del analito). No obstante, para que puedan aplicarse es necesario que la incertidumbre del valor de referencia sea despreciable. Es preciso estudiar cómo calcular la incertidumbre cuando la precisión varía exponencialmente con la concentración y cuando la incertidumbre de la concentración de referencia no es despreciable.

### 3.8. Referencias

1. International Union of Pure and Applied Chemistry, *Pure and Applied Chemistry*, **71** (1999) 337-348
2. J.K. Taylor, *Analytical Chemistry*, **55** (1983) 600A-608A
3. W.J. Youden, *Analytical Chemistry*, **19** (1947) 946-950
4. M. Parkany, *The use of recovery factors in trace analysis*, Royal Society of Chemistry, Cambridge (1996)
5. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-2, ISO, Geneva (1994)
6. W. Horwitz, *Analytical Chemistry*, **54** (1982) 67A-75A
7. M. Thompson, *Analyst*, **113** (1988) 1579-1587
8. D.M. Rocke, S. Lorenzato, *Technometrics*, **37** (1995) 176-184
9. International Union of Pure and Applied Chemistry, *Recommendations for the use of the term "recovery" in analytical procedures*, Draft, IUPAC (2002)  
<http://www.iupac.org/>
10. Analytical Methods Committee, *AMC Technical Brief*, **1** (2000)
11. J.F. Tyson, *Analyst*, **109** (1984) 313-317
12. R. Ferrús, F. Torrades, *Analytical Chemistry*, **60** (1988) 1281-1285
13. R. Ferrús, F. Torrades, *Química Analítica*, **11** (1992) 79-99
14. R.A. Mostyn, A.F. Cunningham, *Analytical Chemistry*, **38** (1966) 121-123
15. A.M. García Campaña, L. Cuadros Rodríguez, J. Aybar Muñoz, F. Alés Barrero, *Journal of AOAC International*, **80** (1997) 657-664
16. G.W. Snedecor, W.G. Cochran, *Statistical Methods*, 6th Ed., Iowa State University Press (1967)
17. N. Draper, H. Smith, *Applied Regression Analysis*, 2<sup>nd</sup> Ed. Wiley, Nueva York (1981)
18. D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier (1997)
19. J. Mandel, F.J. Linnig, *Analytical Chemistry*, **29** (1957) 743-749
20. R. Ferrús, *Anàlisi química general. Pràctiques*, Edicions UPC (1995)
21. M. J. Cardone, *Journal of AOAC International*, **66** (1983) 1257-1282
22. M. J. Cardone, *Journal of AOAC International*, **66** (1983) 1283-1294
23. M. Cardone, *Analytical Chemistry*, **58** (1989) 433-438
24. L. Cuadros, A.M. García Campaña, F. Alés Barrero, C. Jiménez Linares, M. Román Ceba, *Journal of AOAC International*, **78** (1995) 471-475
25. R. C. Castells, M. A. Castillo, *Analytica Chimica Acta*, **423** (2000) 179-185
26. W.R. Kelly, J.D. Fasset, *Analytical Chemistry*, **55** (1983) 1040-1044.
27. F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309-316
28. EURACHEM/CITAC, *Quantifying uncertainty in analytical measurement*, EURACHEM/CITAC Guide, 2<sup>nd</sup> Ed. (2000)



## **CAPÍTULO 4**

---

***Incorporación del sesgo no significativo en el cálculo de la incertidumbre de los resultados***



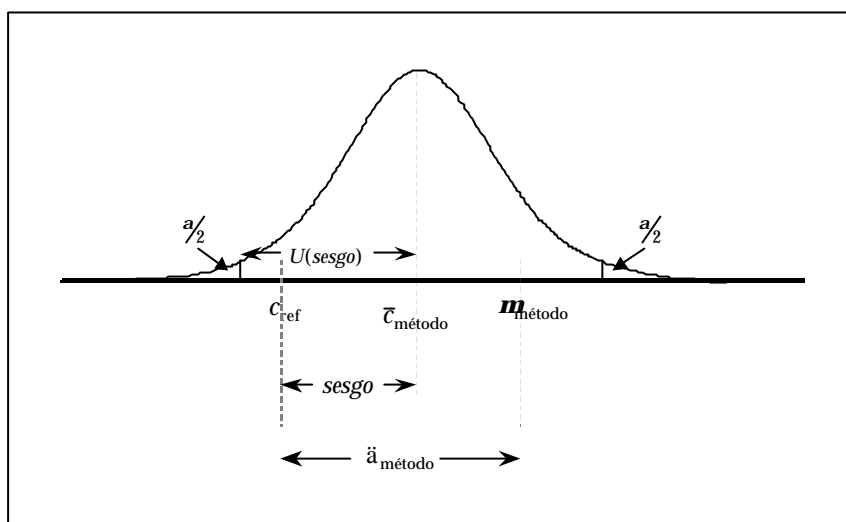


## 4.1. Introducción

En los capítulos anteriores se ha visto que no se debería calcular la incertidumbre de los resultados si previamente no se ha verificado la trazabilidad de los mismos. Verificar la trazabilidad de los resultados implica calcular el sesgo del método y su incertidumbre asociada. Para que los resultados sean trazables, el sesgo del método debería ser cero. No obstante, durante la verificación de la trazabilidad nunca se obtiene un sesgo que sea exactamente cero debido a la presencia de errores aleatorios. Por tanto, debe comprobarse con un test estadístico si el sesgo calculado no difiere estadísticamente de cero. Si el sesgo no difiere estadísticamente de cero, se concluye que el sesgo no es significativo y que los resultados son trazables a la referencia utilizada. Sin embargo, como el sesgo calculado tiene una incertidumbre asociada, siempre existe una incertidumbre al afirmar que los resultados son trazables. Es decir, aunque los resultados resulten ser trazables, nunca se tendrá una seguridad del 100% en esta afirmación. Esto hace que en los capítulos anteriores se haya incluido en la incertidumbre de los resultados un término asociado a afirmar que los resultados son trazables.

Estrictamente, la incertidumbre calculada en los capítulos anteriores corresponde a la incertidumbre de los resultados corregidos por el sesgo. Sin embargo, los resultados no se corrigen cuando el sesgo del método no es significativo. No obstante, debido a la presencia de errores aleatorios, siempre hay una probabilidad de concluir erróneamente que el sesgo del método no es significativo. Por ejemplo, la Fig. 4.1 muestra cómo se puede concluir erróneamente que los resultados son trazables a un valor de referencia,  $c_{\text{ref}}$ . Esta figura muestra el proceso de la verificación de la trazabilidad a un nivel de concentración utilizando un material de referencia. El sesgo del método se calcula como la diferencia entre el valor de referencia y el valor medio obtenido al analizar varias veces la muestra de referencia,  $\bar{c}_{\text{método}}$  (es decir,  $\text{sesgo} = c_{\text{ref}} - \bar{c}_{\text{método}}$ ). La Fig. 4.1 muestra que el sesgo calculado es menor que su incertidumbre expandida,  $U(\text{sesgo})$ . Por tanto, se concluye que el sesgo no es significativo y que los resultados son trazables. Sin embargo, el método tiene en realidad un sesgo,  $\delta_{\text{método}}$ , que, tal y como muestra la Fig. 4.1, corresponde a la diferencia entre el valor de referencia,  $c_{\text{ref}}$ , y el valor que se obtendría,  $\mu_{\text{método}}$ , si la muestra de referencia se analizara infinitas veces con el

método analítico. Este tipo de error corresponde a un error de tipo  $\beta$  o falso negativo ya que se afirma que el método no tiene un sesgo cuando en realidad sí que está sesgado.



**Figura 4.1.** Verificación de la trazabilidad en un método sesgado.  $\delta_{método}$  es el sesgo verdadero del método,  $sesgo$  es el sesgo calculado durante la verificación de la trazabilidad y  $U(sesgo)$  es su incertidumbre expandida.

De hecho, desde un punto de vista químico, siempre es de esperar que el método analítico tenga algún tipo de sesgo<sup>1</sup>. Esto ha hecho que, tal y como se ha visto en el apartado 1.7, Kuttatharmmakul *et al*<sup>1,2</sup> hayan propuesto una aproximación para fijar la probabilidad de error  $\beta$  de no detectar un valor mínimo,  $\lambda$ , de sesgo del método analítico. Esta aproximación está basada en determinar el número de veces que el analista debe analizar la muestra de referencia para que se pueda detectar el sesgo  $\lambda$  con una probabilidad  $1-\beta$ . Sin embargo, incluso en este caso, un sesgo no significativo no implica que el método no esté sesgado, simplemente el sesgo cae dentro de los márgenes de incertidumbre.

Si se concluye erróneamente que los resultados son trazables, se puede subestimar la incertidumbre de los resultados ya que, como el método está sesgado, los

resultados deberían corregirse por el sesgo calculado a pesar de haber concluido con el test estadístico que el sesgo no era significativo. En este capítulo, se estudia si se subestima la incertidumbre de los resultados cuando se concluye erróneamente en la verificación de la trazabilidad que el sesgo constante y/o proporcional del método analítico no es significativo. Además, también se estudian diversas estrategias para evitar la subestimación de la incertidumbre. Para ello, se han estudiado dos casos diferenciados: a) cuando el método se aplica en un intervalo restringido de concentraciones y, por tanto, la trazabilidad se verifica a un nivel de concentración asumiendo que el sesgo es constante en todo el intervalo de aplicación (apartado 4.2); y b) cuando la trazabilidad se verifica en un intervalo amplio de concentraciones utilizando estudios de recuperación (apartado 4.3). En ambos casos, se ha aplicado el método de Monte-Carlo<sup>3</sup> para simular el proceso de la verificación de la trazabilidad y, posteriormente, el análisis de muestras de rutina. Estas simulaciones se realizaron cubriendo todas las posibles situaciones que pueden producirse en la práctica, es decir, la presencia o ausencia de otros términos de incertidumbre y diferentes valores del sesgo y de su incertidumbre.

En el apartado 4.2. de este capítulo se presenta el artículo “*Should non-significant bias be included in the uncertainty budget?*” publicado en la revista *Accreditation and Quality Assurance*. En este artículo se estudia la subestimación de la incertidumbre cuando al verificar la trazabilidad a un nivel de concentración con un material de referencia se concluye erróneamente que los resultados son trazables. Además, se estudia si esta subestimación puede evitarse aplicando diversas estrategias que se han propuesto en el campo de medidas físicas<sup>4-6</sup> para incluir el sesgo como componente de incertidumbre cuando los resultados no se corrigen por errores sistemáticos.

En el apartado 4.3. de este capítulo se presenta el artículo “*Effect of non-significant proportional bias in the final measurement uncertainty*” enviado para su publicación a la revista *Analyst*. En este artículo se estudia la subestimación de la incertidumbre cuando la trazabilidad se verifica en un intervalo amplio de concentraciones en términos de recuperación y se concluye erróneamente que los resultados son trazables. Además, también se estudia cómo incluir el sesgo proporcional en la incertidumbre final de los resultados para evitar que la incertidumbre esté

subestimada. Por último, en el apartado 4.4. se exponen las conclusiones de este capítulo y en el apartado 4.5. la bibliografía.

## **4.2. Should non-significant bias be included in the uncertainty budget? (Accreditation and Quality Assurance, 7 (2002) 90-94)**

Alicia Maroto, Ricard Boqué, Jordi Riu, F. Xavier Rius

Department of Analytical and Organic Chemistry. Institute of Advanced Studies.

Rovira i Virgili University of Tarragona.

Pl. Imperial Tàrraco, 1. 43005 Tarragona, Catalonia. SPAIN.

### **ABSTRACT**

The bias of an analytical procedure is calculated in the assessment of trueness. If this experimental bias is not significant, we assume that the procedure is unbiased and, consequently, the results obtained with this procedure are not corrected for this bias. However, when assessing trueness there is always a probability of incorrectly concluding that the experimental bias is not significant. Therefore, non significant experimental bias should be included as a component of uncertainty. In this paper, we study if it is necessary to include always this term and which is the best approach to include this bias in the uncertainty budget. To answer these questions, we have used the Monte-Carlo method to simulate the assessment of trueness of biased procedures and the future results these procedures provide. The results show that non significant experimental bias should be included as a component of uncertainty when the uncertainty of this bias represents at least a 30% of the overall uncertainty.

**Keywords:** Bias, Uncertainty, Assessment of Trueness

## 1. INTRODUCTION

One of the most important steps in the validation of an analytical procedure is the assessment of trueness. In this process, the experimental bias of the analytical procedure is estimated. If this bias is statistically not significant, we assume that the procedure is unbiased and, consequently, results are not corrected for the experimental bias. However, are we sure that the procedure does not have any bias?. In fact, when assessing trueness there is always a probability of incorrectly concluding that the experimental bias is statistically not significant. As a result, this probability should be included (expressed as a quantity related to the experimental bias) as a component of the uncertainty of the results obtained with the analytical procedure. However, several questions arise, i.e. is it always necessary to include this component of the uncertainty? Moreover, if it is necessary, how should this non significant experimental bias be included?

Non significant experimental bias has not been included so far as a component of uncertainty in chemical measurements. However, different approaches have been proposed to include bias as a component of uncertainty when physical measurements are not corrected for systematic errors [1]. In this paper we study whether these approaches can be applied to include non significant experimental bias in the uncertainty budget of chemical measurements and whether it is necessary to include always this term. To answer these questions, we have simulated the process of assessment of trueness of biased analytical procedures and, subsequently, the future results these procedures provide. We simulated these results covering most of the possible situations that may happen in practice.

## 2. ASSESSMENT OF TRUENESS

Checking the trueness of an analytical procedure involves estimating its experimental bias. If the routine samples have similar levels of concentration, we can assume that we have the same bias in the whole concentration range and, consequently, the experimental bias can be estimated using one reference sample with a concentration similar to the routine samples. If this is the case, the experimental bias is calculated as

the difference between the reference value,  $c_{\text{ref}}$ , and the mean value,  $\bar{c}_{\text{found}}$ , obtained when the reference sample is analysed  $p$  times under intermediate precision conditions (i.e. varying the day, operator, etc.), i.e.  $\text{bias} = c_{\text{ref}} - \bar{c}_{\text{found}}$ . The experimental bias is not significant if:

$$\text{bias} \leq t_{\alpha/2, \text{eff}} \cdot u(\text{bias}) \quad (1)$$

where  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value for the effective degrees of freedom,  $\nu_{\text{eff}}$ , [2] associated with  $u(\text{bias})$ , and can be replaced by the coverage factor  $k$  if the effective degrees of freedom are large enough [3,4]. The uncertainty of the experimental bias,  $u(\text{bias})$ , depends on the reference used to assess trueness. If a certified reference material (CRM) is used, this uncertainty is calculated as:

$$u(\text{bias}) = \sqrt{\frac{s_1^2}{p} + u(c_{\text{ref}})^2} \quad (2)$$

where  $s_1$  is the standard deviation of the  $p$  results obtained when analysing the CRM and  $u(c_{\text{ref}})$  is the standard uncertainty of the CRM (i.e.  $U(c_{\text{ref}})/k$ , where  $k$  is normally equal to 2 and  $U(c_{\text{ref}})$  is the uncertainty of the CRM provided by the manufacturer).

If the experimental bias is significant, the procedure should subsequently be revised in order to identify and eliminate the systematic errors which produced the bias. Otherwise, we assume that the procedure is unbiased and, consequently, we do not correct results for the experimental bias. However, several questions arise in this latter case because, from a chemical point of view, some bias is always to be expected in an analytical procedure.

### 3. CALCULATION OF UNCERTAINTY

Uncertainty can be obtained either by calculating all the sources of uncertainty individually [3,4] or by grouping different sources of uncertainty whenever possible

[5-9]. In this paper, the latter strategy is followed to calculate uncertainty using information obtained in the process of assessment of trueness [5,6]:

$$u = \sqrt{u(\text{proc})^2 + u(\text{trueness})^2 + u(\text{pret})^2 + u(\text{other terms})^2} \quad (3)$$

where  $u$  is the standard uncertainty [3],  $u(\text{proc})$  is the uncertainty of the procedure and corresponds to the intermediate standard deviation of the procedure, i.e.  $u(\text{proc}) = s_i$ .  $u(\text{trueness})$  is the uncertainty of the experimental bias and corresponds to  $u(\text{bias})$ .  $u(\text{pret})$  is the uncertainty associated to subsampling and to sample pretreatments not considered in the assessment of trueness. Finally,  $u(\text{other terms})$  considers other terms of uncertainty due to factors not representatively varied when estimating precision. In this paper, this latter term will be considered to be negligible. The overall expanded uncertainty,  $U$ , is then calculated by multiplying the standard uncertainty,  $u$ , by the two-sided  $t$  tabulated value,  $t_{\alpha/2, \text{eff}}$ , for the effective degrees of freedom,  $\nu_{\text{eff}}$  [2], i.e.  $U = t_{\alpha/2, \text{eff}} \cdot u$ . A coverage factor of  $k=2$  is recommended for most purposes when the effective degrees of freedom,  $\nu_{\text{eff}}$ , are large enough. This value represents a level of confidence of approximately 95%. Strictly, the uncertainty calculated in Eq. 3 corresponds to results of future samples obtained after correcting the concentration found for the experimental bias. However, analytical results are never corrected for non significant experimental bias. As a result, this bias should be included as a component of uncertainty because the procedure may have a true bias.

#### **4. APPROACHES FOR INCLUDING NON SIGNIFICANT EXPERIMENTAL BIAS IN THE UNCERTAINTY BUDGET**

Different approaches have been proposed in the field of physical measurements to include bias as a component of uncertainty when results are not corrected for systematic errors [1]. In this paper, we will study whether these approaches can be applied to include non significant experimental bias in chemical measurements. The first approach consists of including this bias as another component of uncertainty and simply to add it in the usual root-sum-of-squares (RSS) manner, i.e.  $U(\text{RSS}u) = t_{\alpha/2, \text{eff}} \sqrt{u^2 + \text{bias}^2}$ . The second approach sums this bias in a RSS manner



with the expanded uncertainty,  $U$ , i.e.  $U(RSSU) = \sqrt{U^2 + bias^2}$ . The third procedure consists of adding this bias to the expanded uncertainty. This approach is denoted as SUMU [1] and is equivalent to correcting the results:

$$U_+ = U + bias; U_- = U - bias \quad (4)$$

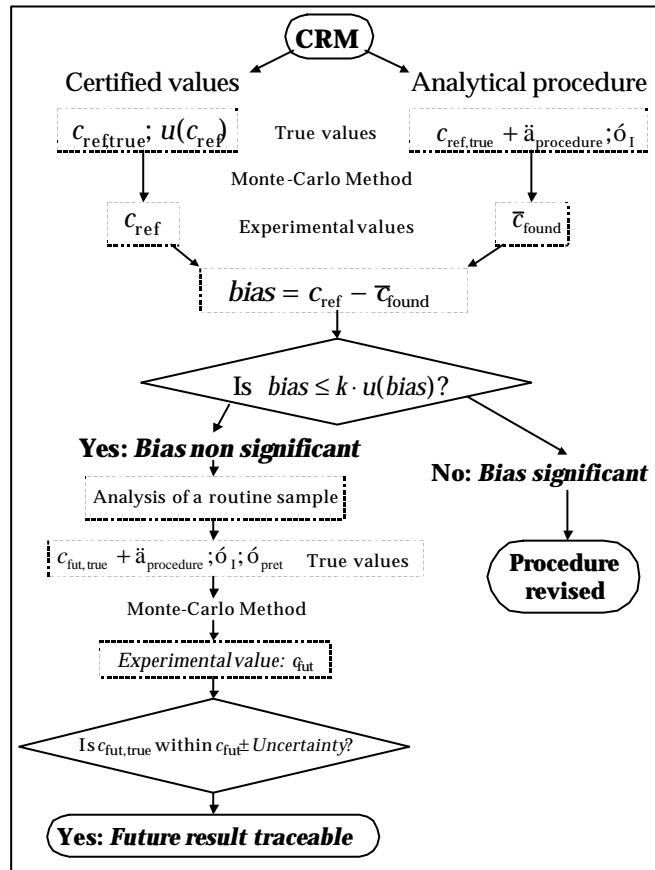
Finally, the last procedure to be studied consists of adding the absolute value of the experimental bias to the expanded uncertainty, i.e.  $U(bias) = U + |bias|$ .

## 5. NUMERICAL EXAMPLE: BIAS IN THE ASSESSMENT OF TRUENESS

To investigate the effects of including the non significant experimental bias as a component of uncertainty, the process of assessment of trueness was simulated with the Monte-Carlo method (see Fig. 1). A true reference value,  $c_{ref,true}$ , together with its standard uncertainty,  $u(c_{ref})$ , was assigned to a "hypothetical CRM". A true bias,  $\delta_{procedure}$ , together with a true intermediate standard deviation,  $\sigma_i$ , was associated to the analytical procedure. Moreover, the possibility of having other steps in the analytical procedure (i.e. pretreatments and/or subsampling) not carried out when analysing the CRM was also studied. In this case, an additional true standard deviation,  $\sigma_{pret}$ , was also added to the future results obtained with the analytical procedure.

In the assessment of trueness, we simulated the certified reference value,  $c_{ref}$ , and the  $p$  results obtained when analysing the CRM. The true value and the standard deviation used to simulate the  $p$  results were, respectively,  $c_{ref,true} + \delta_{procedure}$  and  $\sigma_i$ . Results were simulated assuming that they followed a normal distribution. After this, we calculated the mean of these results,  $\bar{c}_{found}$ , and the experimental bias. The uncertainty of this bias,  $u(bias)$ , was calculated with Eq. 2. Afterwards, it was checked with Eq. 1 whether this bias was statistically significant or not. If it was not significant, a future concentration obtained for a routine sample,  $c_{fut}$ , was simulated with the Monte-Carlo method. The true value and the true variance used to simulate this result were, respectively,  $c_{fut,true}$  and  $S_{fut}$ . The variance,  $S_{fut}$ , corresponded to  $S_I + S_{pret}$ . After this

we calculated the uncertainty of this future result following section 3. The experimental bias was included as a component of uncertainty using the four approaches explained in section 4. Then we checked whether these approaches included the true concentration of the routine sample,  $c_{\text{fut,true}}$ , within the interval  $c_{\text{fut}} \pm \text{Uncertainty}$ .



**Figure 1.** Scheme of the process of assessment of trueness simulated with the Monte-Carlo method. Future results of routine sample were simulated if the experimental bias was identified as non significant.

This process was simulated 300,000 times for 25 different values of  $\delta_{\text{procedure}}$ . After this, we calculated the percentage of times that the experimental bias was found to be non significant. This corresponded to the probability of  $\beta$  error (or probability of false negative) because in the assessment of trueness we state that the method is unbiased

when in fact is biased. We also calculated the percentage of times that, once the experimental bias was identified as non significant, the different approaches studied included the true concentration of the routine sample,  $c_{\text{rut,true}}$ , within the interval  $c_{\text{rut}} \pm \text{Uncertainty}$ . We simulated this process for three different cases which cover different situations that may happen in practice (i.e. presence/absence of pretreatment steps and different number of replicated analysis of the CRM). Table 1 shows, for the three cases, the values of  $\sigma$ ,  $\sigma_{\text{pret}}$ ,  $u(c_{\text{ref}})$  and  $p$  used for simulating the results.

**Table 1.** Three different situations simulated in the assessment of trueness.

	$p$	$u(c_{\text{ref}})$	$\sigma$	$\sigma_{\text{pret}}$	$u$	$u(\text{bias})$	$\frac{u(\text{bias})}{u} \cdot 100$
Case 1	5	0.2	1.5	0	1.64	0.70	43
Case 2	10	0.2	1.5	0	1.59	0.51	32
Case 3	10	0.2	1.5	3	3.39	0.51	15

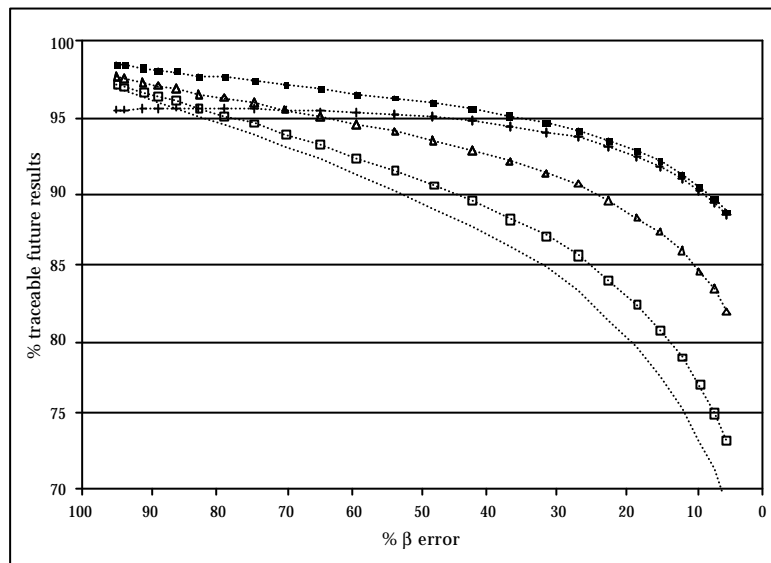
## 6. RESULTS AND DISCUSSION

Table 2 shows the probability of  $\beta$  error committed for different values of the true bias of the analytical procedure,  $\delta_{\text{procedure}}$ , for the three cases studied. We can see that the higher is  $\delta_{\text{procedure}}$ , the lower is the probability of  $\beta$  error. This is because the higher is the true bias, the more likely is to detect that the procedure is biased. The probability of  $\beta$  error depends also on  $\sigma$ ,  $u(c_{\text{ref}})$  and  $p$ . We can see that for the same values of  $\sigma$ ,  $u(c_{\text{ref}})$  and  $\delta_{\text{procedure}}$ , the lower is  $p$ , the higher is the probability of  $\beta$  error.

**Table 2.** True bias of the analytical procedure,  $\delta_{\text{procedure}}$ , and percentage of times that the experimental bias is identified as non significant, i.e. %  $\beta$  error, for the three cases described in Table 1.

Case 1		Cases 2 and 3	
$\delta_{\text{procedure}}$	% $\beta$ error	$\delta_{\text{procedure}}$	% $\beta$ error
0	95	0	95
0.4	91	0.4	88
0.8	79	0.8	66
1.2	60	1.2	36
1.6	37	1.4	22
2	18	1.6	13
2.4	7	1.8	6

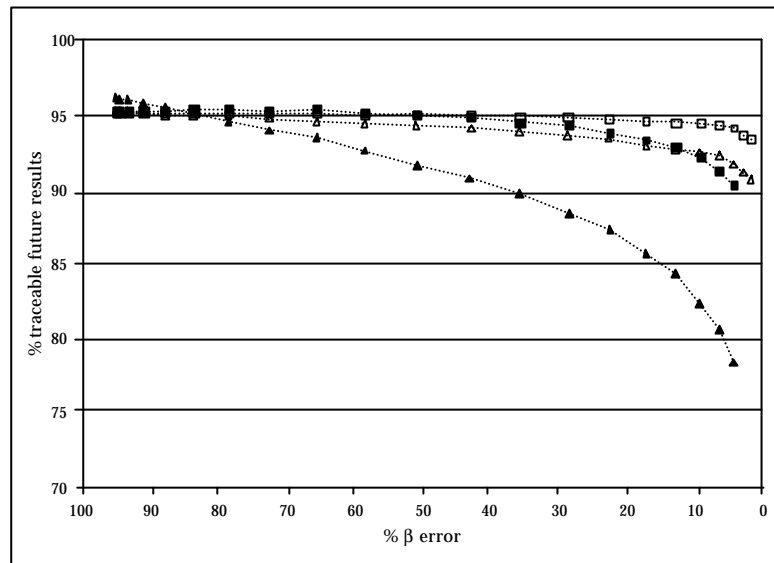
We also studied the percentage of times that, once the experimental bias was identified as non significant, the different ways of calculating uncertainty included the true concentration  $d$  of the routine sample,  $c_{fut,true}$ , within the interval  $c_{fut} \pm \text{Uncertainty}$ . Uncertainty was calculated using the  $z$ -value for a level of significance  $\alpha=5\%$ . Therefore, if the uncertainty is correctly calculated, this percentage, i.e. % traceable future results, should be 95% (i.e.  $100-\alpha\%$ ). If uncertainty is underestimated, this percentage is lower than 95% and, if it is overestimated, it is higher than 95%. This percentage was calculated and plotted as a function of the  $\beta$  error committed in the assessment of trueness.



**Figure 2.** Percentage of traceable future results for case 1 versus the probability of  $\beta$  error. Uncertainty is calculated without including non significant experimental bias: .....  $U$  and including this bias with the approaches of section 4: ..... $\Delta$ .....  $U(RSSu)$ ; ..... $\square$ .....  $U(RSSU)$ ; .....+.....  $SUMU$  and ..... $\blacksquare$ .....  $U(bias)$ .

Fig. 2 shows these results for case 1. In this case the contribution of  $u(bias)$  to the overall uncertainty,  $u(bias)/u$ , is 43%. We see that uncertainty can be very underestimated when the experimental bias is not included as a component of uncertainty. This underestimation depends on the ratio  $u(bias)/u$ , i.e.: the higher is this ratio, the higher is the underestimation of uncertainty. The best approach to include non significant experimental bias is the  $SUMU$  approach because it gives the

percentage of traceable future results closest to 95%. The uncertainty,  $U(bias)$ , is also a good approach for including this bias. However, this approach gives higher uncertainty values than the *SUMU* approach. The  $U(RSSU)$  and the  $U(RSSu)$  uncertainties are clearly inferior for including this bias because they overestimate uncertainty for higher probabilities of  $\beta$  errors and underestimate uncertainty for lower probabilities of  $\beta$  errors. Moreover, these approaches give higher uncertainty values than the *SUMU* approach.



**Figure 3.** Percentage of traceable future results for cases 2 and 3 versus the probability of  $\beta$  error. Uncertainty is calculated without including the experimental bias,  $U$ , (case 2:  $\blacktriangle$  and case 3  $\blacktriangle$ ) and with the *SUMU* approach (case 2  $\blacksquare$  and case 3  $\blacksquare$ ).

Fig. 3 shows the percentage of traceable results versus the percentage of  $\beta$  error for case 2 (i.e. 32% of contribution of  $u(bias)$  to the overall uncertainty) and for case 3 (i.e. 15% of contribution). In this Figure, uncertainty is calculated without including the experimental bias and with the *SUMU* approach. We see that the experimental bias should be included in case 2 but this is not necessary for case 3. This is because in case 3 the uncertainty of the experimental bias is negligible when compared to the overall uncertainty.

## 7. CONCLUSIONS

Non significant experimental bias should be included in the uncertainty budget when the uncertainty of this bias represents about 30% of the overall uncertainty. The higher is this contribution, the more important is to include the non significant experimental bias. On the contrary, it is not necessary to include this bias when its uncertainty has a low contribution to the overall uncertainty, i.e. 15% or lower. The best approach for including this bias is the *SUMU* approach. The uncertainty,  $U(bias)$ , also gives good results. Otherwise, we can use the uncertainty  $U(bias)$  because, opposite to the *SUMU* approach, it has the advantage that it gives a symmetric confidence interval around the estimated result. However, it gives higher uncertainty values than the *SUMU* approach.

## 9. REFERENCES

- [1] Phillips SD, Eberhardt KR, Parry B (1997) *J Res Natl Inst Stand Technol* 102: 577-585
- [2] Satterthwaite FE (1941) *Psychometrika* 6: 309-316
- [3] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML (1993) *Guide to the expression of uncertainty in measurement*, ISO, Geneva
- [4] EURACHEM (1995) *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK
- [5] Maroto A, Riu J, Boqué R, Rius FX (1999) *Anal Chim Acta* 391: 173-185
- [6] Maroto A, Boqué R, Riu J, Rius FX (1999) *Trends Anal Chem* 18/9-10: 577-584
- [7] Ellison SLR, Williams A (1998) *Accred Qual Assur* 3: 6-10
- [8] Barwick VJ, Ellison SLR (2000) *Accred Qual Assur* 5: 47-53
- [9] EURACHEM/CITAC *Guide* (2000) *Quantifying uncertainty in analytical measurement*, EURACHEM, 2<sup>nd</sup> Edition, Helsinki

### **4.3. Effect of non-significant proportional bias in the final measurement uncertainty (Analyst, submitted for publication)**

Alicia Maroto, Ricard Boqué, Jordi Riu, F. Xavier Rius

Department of Analytical and Organic Chemistry. Institute of Advanced Studies.  
Rovira i Virgili University of Tarragona.  
Pl. Imperial Tàrraco, 1. 43005 Tarragona, Catalonia. SPAIN.

#### **ABSTRACT**

The trueness of an analytical method can be assessed by calculating the proportional bias of the method in terms of apparent recovery. If the apparent recovery calculated does not differ significantly from one, we assume that the analytical method does not have any proportional bias and the trueness of the method is assessed. However, when assessing trueness there is always a probability of incorrectly concluding that the proportional bias is not significant. If this is the case, the uncertainty of the analytical results can be underestimated. This underestimation can be avoided by including the non-significant bias as a component of uncertainty. In this paper, we study if it is necessary to include always this term and which is the best approach to include it. To answer these questions we have simulated with the Monte-Carlo method the assessment of trueness of biased methods and the results these methods provide. The results show that the non-significant proportional bias should be included as a component of uncertainty when the contribution of the uncertainty of the apparent recovery represents at least a 30% of the overall uncertainty.

**Keywords:** trueness, proportional bias, uncertainty, apparent recovery

## 1. INTRODUCTION

The trueness of an analytical method can be assessed by estimating the proportional bias with recovery studies. The term recovery is defined as the proportion of the amount of analyte, present in or added to the analytical portion of the test material, which is extracted and presented for measurement [1]. At present the term recovery is used in two different contexts: 1) to express the yield of an analyte in a preconcentration or separation step of an analytical method; and 2) to denote the ratio of the concentration obtained from the analytical process via a calibration graph compared to the concentration of analyte present in or added to the test material [2]. The first use should be clearly distinguished from the second one because having a 100% of recovery does not necessary require a 100% yield for any preconcentration or separation step. For this reason, the IUPAC [2] recommends the use of two different terms to distinguish between the two different uses of recovery. The term *recovery* is recommended for the first use whereas the term *apparent recovery* is recommended for the second one. This latter use is related with the proportional bias of the method.

The trueness of an analytical method is assessed if the *apparent recovery* calculated does not differ significantly from one. If this is the case, we assume that the analytical method does not have any proportional bias and, consequently, we do not correct results for recovery. However, can we be sure that the method does not have any proportional bias? In fact, there is always a probability of incorrectly concluding that the proportional bias is not significant. Therefore, this probability should be included (expressed as a component related to the *apparent recovery* estimated) in the uncertainty budget, otherwise uncertainty may be underestimated.

Maroto *et al* have already studied different approaches to include non-significant constant bias when the trueness of the method is assessed in a restricted concentration range [3]. In the present paper, we study how to apply these approaches when the trueness of the method is assessed in large concentration ranges in terms of apparent recovery. Moreover, we also study if it is always necessary to include non-significant proportional bias in the uncertainty budget.



To answer these questions, we have used the Monte-Carlo method to simulate the process of estimating the apparent recovery of a biased method and, subsequently, the future results these methods provide. For simulating these results we have considered the presence/absence of other sources of uncertainty and different number of replicated analysis of the spiked samples.

## 2. ESTIMATION OF RECOVERY

Apparent recovery is usually estimated using samples spiked at different levels of concentration. The apparent recovery can then be estimated using regression techniques or the method of the averaged recovery [4]. In this paper, we have simulated the estimation of the apparent recovery using the averaged recovery method. This method consists of estimating the experimental apparent recovery,  $R$ , as the mean value of all the  $n$  apparent recoveries,  $R_i$ , obtained using the spiked samples. These apparent recoveries are calculated as:

$$R_i = \frac{c_{f+0} - c_0}{c_s} \quad (1)$$

where  $c_s$  is the concentration of analyte spiked to the sample,  $c_{f+0}$  is the concentration of analyte found in the spiked sample and  $c_0$  is the concentration found of the native analyte. These spiked samples should cover the range of sample matrices and analyte concentrations of the routine samples [5]. The uncertainty of the experimental apparent recovery can then be estimated as:

$$u(R) = \frac{s(R)}{\sqrt{n}} \quad (2)$$

where  $s(R)$  is the standard deviation of the  $n$  recoveries,  $R_i$ . If precision information is available and it can be assumed that the relative intermediate standard deviation,  $RSD_i$ , remains constant throughout the concentration range, the uncertainty of recovery can be calculated as [5]:

$$u(R) = \frac{R \cdot RSD_1}{\sqrt{n}} \quad (3)$$

This expression can be used if the uncertainty of  $c_0$  is negligible when compared to the uncertainty of  $c_{t+0}$ . The experimental apparent recovery is not statistically different from the theoretical value of one if the following inequality holds:

$$|1 - R| \leq t_{\alpha/2, n} \cdot u(R) \quad (4)$$

where  $t_{\alpha/2, n}$  is the two-sided  $t$  tabulated value for  $v$  degrees of freedom. Here  $v$  depends on whether the uncertainty of recovery is calculated using Eq. 2 or Eq. 3. If it is calculated with Eq. 2,  $v=n-1$ . If it is calculated with Eq. 3,  $v=p-1$  (where  $p$  is the number of analysis carried out to obtain  $RSD_1$ ).

### 3. CALCULATION OF UNCERTAINTY

A result obtained for a future routine sample can be expressed as:

$$c_{\text{corr}} = F_{\text{hom}} \cdot \frac{c}{R} \quad (5)$$

where  $c$  is the concentration obtained when analysing the sample and  $F_{\text{hom}}$  is a factor that considers the heterogeneity of the sample and the variability of the apparent recovery with concentration or matrix type. For practical purposes,  $F_{\text{hom}}$  will be one. However, its uncertainty is not negligible and should be considered. Similarly, results are not corrected for apparent recovery if this recovery does not differ significantly from one. Nonetheless, there is a current debate focused on whether results should be corrected when the apparent recovery differs significantly from one [6]. Regardless of this, the uncertainty of the apparent recovery calculated should be included in the final measurement uncertainty.

The standard uncertainty of the results,  $u$ , obtained when the method of analysis validated in-house is applied to future routine samples can be calculated by applying the error propagation law to Eq. 5:

$$u = \frac{1}{R} \cdot \sqrt{\frac{s_I^2}{N} + (c_{\text{corr}} \cdot u(R))^2 + (c_{\text{corr}} \cdot R \cdot u(F_{\text{hom}}))^2} \quad (6)$$

The first term of Eq. 6 considers the uncertainty of the concentration,  $c$ , obtained when analysing the routine sample.  $s_I$  is the intermediate precision of the method at the level of concentration of the routine sample analysed. This precision is obtained from the results obtained when varying all the factors that can affect the analytical method (i.e. operator, day, reagents, etc.).  $N$  is the number of times that the routine sample is analysed under intermediate precision conditions (usually,  $N=1$ ). The second term of Eq. 6 is the contribution due to the uncertainty of the apparent recovery. The third term of Eq. 6 is the contribution due to the uncertainty of the factor  $F_{\text{hom}}$ . This uncertainty can be calculated by analysing different portions of a sample similar to the routine samples [5]. The uncertainty of the variability of recovery with concentration or matrix type can be calculated from the analysis of variance (ANOVA) of the results obtained when the analyte is spiked at different levels of concentration to different samples which are representative of the matrix variability of the routine samples [5]. If it can be assumed that the relative intermediate standard deviation,  $RSD_I$ , remains constant throughout the concentration range and  $u(R)$  is calculated with Eq. 3, then Eq. 6 can be expressed as:

$$u = c_{\text{corr}} \sqrt{RSD_I^2 \left( \frac{1}{N} + \frac{1}{n} \right) + u(F_{\text{hom}})^2} \quad (7)$$

The overall expanded uncertainty,  $U$ , is then calculated by multiplying the combined standard uncertainty,  $u$ , by a coverage factor,  $k$ , for the effective degrees of freedom calculated with the Welch-Satterthwaite approach [7] and the significance level chosen, i.e.  $U=k \cdot u$ . A coverage factor of  $k=2$  is recommended for most purposes when the effective degrees of freedom are large enough. This value represents a level of confidence of approximately 95%. Strictly, this uncertainty

corresponds to results of future samples obtained after correcting the concentration found by the apparent recovery calculated. However, analytical results are never corrected if the apparent recovery calculated does not differ significantly from one. In this case, it is assumed that the method does not have any proportional bias and, consequently, results are not corrected for the apparent recovery. Nonetheless, there is always a probability of incorrectly concluding that the method does not have any proportional bias. If this is the case, the uncertainty of future results may be underestimated. This underestimation can be avoided by including the non-significant proportional bias (expressed as a component related to the *apparent recovery* estimated) in the uncertainty budget.

#### 4. APPROACHES FOR INCLUDING NON-SIGNIFICANT PROPORTIONAL BIAS IN THE UNCERTAINTY BUDGET

In this paper, we study three different procedures to include non-significant proportional bias in the uncertainty budget. The first one, proposed by Barwick *et al* [8], was initially developed to correct uncertainty when the apparent recovery differs significantly from one and results remain uncorrected. The overall expanded uncertainty,  $U_B$ , is calculated as:

$$U_B = k \cdot \sqrt{u^2 + c_{\text{corr}}^2 \cdot \left(\frac{1-R}{k}\right)^2} \quad (8)$$

The second approach consists of adding the non-significant proportional bias to the expanded uncertainty. This approach is denoted as the *SUMU* approach [9] and provided the best results when non-significant constant bias was included as a component of uncertainty [3]. The *SUMU* uncertainty is calculated as:

$$U_+ = U + c \cdot \left(\frac{1-R}{R}\right); U_- = U - c \cdot \left(\frac{1-R}{R}\right) \quad (9)$$

This approach amounts to correcting the results and has the disadvantage that it gives an asymmetric confidence interval around the estimated result. This

disadvantage can be overcome by adding the absolute value of the non significant proportional bias to the uncertainty:

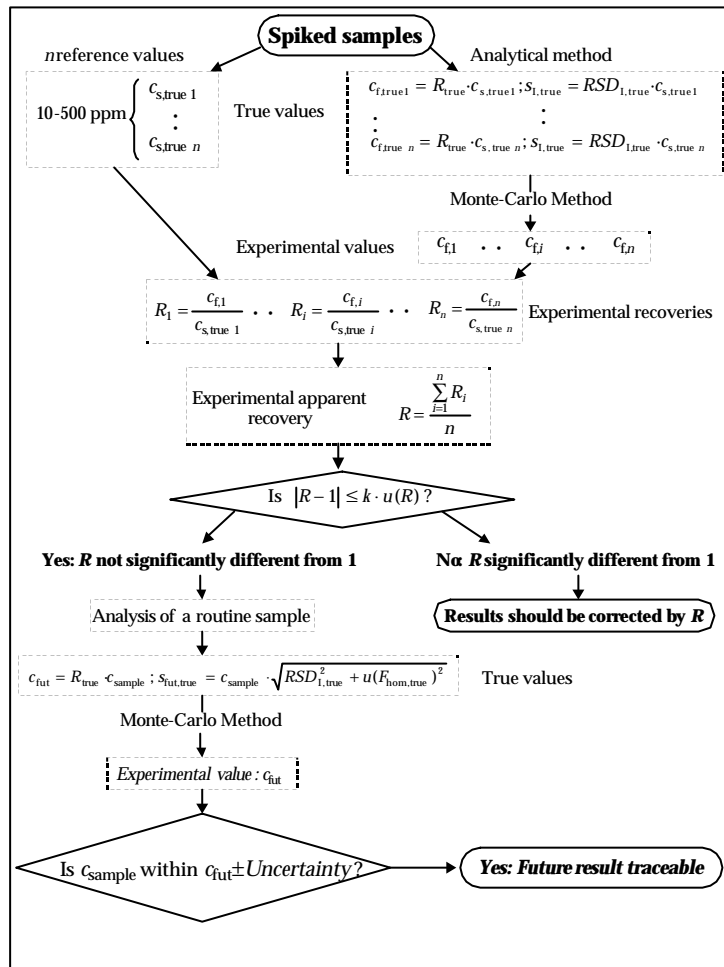
$$U(bias) = U + \left| c \left( \frac{1-R}{R} \right) \right| \quad (10)$$

This approach gives similar results than the *SUMU* approach [3] but has the advantage that it gives a symmetric confidence interval around the estimated results. However, it gives higher uncertainty values than the *SUMU* approach.

## 5. NUMERICAL EXAMPLE: ESTIMATING RECOVERY WITH SPIKED SAMPLES

To determine how much the uncertainty is affected by wrongly concluding that the proportional bias is not significant, we have simulated the process of estimating the apparent recovery in several conditions using the Monte-Carlo method [10] (see Fig. 1). In all the simulations we assumed that the uncertainty of the concentration of analyte spiked was negligible, that the spiked analyte was a good representation of the native analyte and that sample blanks were available. We simulated the recovery assays using samples spiked with concentrations ranging from 10 to 500 ppm. We assigned to the analytical method different values for the true apparent recovery,  $R_{true}$ , and a true relative intermediate standard deviation,  $RSD_{I,true}$ , equal to 0.08 (i.e. we assumed that  $RSD_{I,true}$  was approximately constant throughout the concentration range).

The concentration found,  $\alpha_{i,i}$ , obtained when analysing each spiked sample was simulated using the Monte-Carlo method. These concentrations were simulated assuming that they follow a normal distribution. The true value used to simulate  $\alpha_{i,i}$  was the product of the true apparent recovery and the spiked concentration, i.e.  $R_{true} \cdot C_{s,true} \cdot i$ . The true standard deviation used to simulate  $\alpha_{i,i}$  was the true intermediate precision of the method. This standard deviation was calculated as  $RSD_{I,true} \cdot C_{s,true} \cdot i$ .



**Figure 1.** Scheme of the process for calculating recovery using spiked samples. Results are simulated using the Monte-Carlo method. Future results of routine samples were simulated if recovery did not differ significantly from 1.

After computing the concentration found,  $c_{f,i}$ , for each spiked sample, we calculated the experimental apparent recovery for each of the spiked samples analysed (i.e.  $R_i = c_{f,i} / c_{s,true i}$ ). The experimental apparent recovery of the method,  $R$ , was then calculated as the mean value of all the  $n$  apparent recoveries  $R_i$ . Afterwards, we checked with Eq. 4 whether  $R$  significantly differed from one. The uncertainty of  $R$ ,  $u(R)$ , was obtained with Eq. 3 and using the true value of the relative intermediate standard deviation,  $RSD_{i,true}$ . Since true standard deviations

were used, a z-tabulated value of 1.96 (for  $\alpha=0.05$ ) instead of the t-tabulated value was used in Eq. 4.

If the recovery estimated did not significantly differ from one, we simulated a future result,  $c_{\text{fut}}$ , obtained when analysing a routine sample. We did not simulate any future result if  $R$  significantly differed from one because we were interested in studying what happens when it is wrongly concluded that recovery does not significantly differ from one. The future result was simulated in the Monte-Carlo process by assuming that results followed a normal distribution. These future results were simulated using a true value,  $c_{\text{sample}}$ , that could take values comprised between 10 and 500 ppm and a true standard deviation calculated as  $s_{\text{fut,true}} = c_{\text{sample}} \cdot \sqrt{RSD_{1,\text{true}}^2 + u(F_{\text{hom,true}})^2}$  (where  $u(F_{\text{hom,true}})$  is the true value of the uncertainty due to sample heterogeneity or matrix variability). The standard uncertainty,  $u$ , of the future result,  $c_{\text{fut}}$ , was calculated following Eq. 7. The non-significant proportional bias was included as a component of uncertainty using the three approaches described in the previous section. Since we used true standard deviations, the final expanded uncertainty was calculated by using a coverage factor  $k=1.96$  (for  $\alpha=0.05$ ), based on the normal distribution.

The whole process was simulated 300.000 times for different values of proportional bias expressed in terms of true apparent recovery,  $R_{\text{true}}$ . After that, we calculated for each value of  $R_{\text{true}}$  the percentage of times that the experimental recovery did not differ significantly from one. This corresponded to the probability of  $\beta$  error (or probability of false negative) because we conclude that the analytical method does not have any bias when in fact is biased. We also calculated the percentage of times that, once proportional bias was identified as non-significant, the different expressions for calculating uncertainty contained the true concentration of the sample,  $c_{\text{sample}}$ , within the interval  $c_{\text{fut}} \pm \text{Uncertainty}$ . We did simulations for five and ten spiked samples analysed (i.e.  $n=5$  and  $n=10$ ) and also for different values of the uncertainty of other terms,  $u(F_{\text{hom}})$ .

## 6. RESULTS AND DISCUSSION

### 6.1. True apparent recovery and $\beta$ error

Table 1 shows the probability of  $\beta$  error committed for different values of proportional bias expressed in terms of true apparent recovery,  $R_{\text{true}}$  (strictly, for  $R_{\text{true}}=1$  a  $\beta$  error is not committed because the method does not have any proportional bias. In this case, the probability of 95% is due to the  $\alpha$  error of 5%). Table 1 shows that the probability of committing a  $\beta$  error is very high for recoveries close to one. This is because the proportional bias is so small that it is very unlikely to detect with a significance test that the method has in fact a proportional bias. For instance, for  $n=5$ , there is a probability of 91.40% of not detecting that a true apparent recovery of 0.98 is in fact different from 1. However, this high probability of  $\beta$  error does not affect the accuracy of the analytical results because this bias is so small that hardly affects the results. We can see that the lower is the apparent recovery, the lower is the probability of  $\beta$  error. This is because the higher is the true bias, the more likely is to detect that the analytical method is biased. The probability of  $\beta$  error depends also on the intermediate precision and on the number of times that the spiked samples are analysed,  $n$ . For the same intermediate precision and the same true apparent recovery, the higher is  $n$  the lower is the probability of  $\beta$  error.

**Table 1.** True recovery of the analytical method,  $R_{\text{true}}$ , and percentage of times that the experimental recovery is identified as being not significantly different from one, i.e. % $\beta$  error, obtained when  $RSD_{\text{true}}=0.08$ . The number of spiked samples analysed to calculate recovery is  $n=5$  and  $n=10$ .

$R_{\text{true}}$	$n=5$		$n=10$	
	$R_{\text{true}}$	% $\beta$ error	$R_{\text{true}}$	% $\beta$ error
1	1	95	1	95
0.98	0.98	91.40	0.98	87.58
0.96	0.96	79.88	0.96	64.64
0.94	0.94	61.18	0.94	34.13
0.91	0.91	28.86	0.93	20.98
0.89	0.89	13.30	0.92	11.47
0.87	0.87	4.85	0.91	5.52



## 6.2. Percentage of traceable future results

We have studied the percentage of times that, once the proportional bias calculated was identified as non-significant, the uncertainty (calculated in different ways) included the true concentration of the sample,  $c_{\text{sample}}$ . Uncertainty was calculated using the  $z$ -value for a level of significance of  $\alpha=5\%$ . Therefore, if uncertainty is correctly calculated, this percentage, i.e. % traceable future results, should be 95% (i.e.  $100-\alpha$  %). If uncertainty is underestimated, this percentage is lower than 95% and, if it is overestimated, is higher than 95%.

**Table 2.** Value of the ratios  $r$  and  $f$  for the different situations simulated in Figures 2, 3 and 4.  $f$  represents the contribution of the uncertainty of other terms and is calculated with Eq. 11.  $r$  represents the contribution of the uncertainty of recovery and is calculated with Eq. 12.

$f$	$r$ (%)	
	$n=5$	$n=10$
0	41	30
0.5	37	27
1	30	22
2	20	14

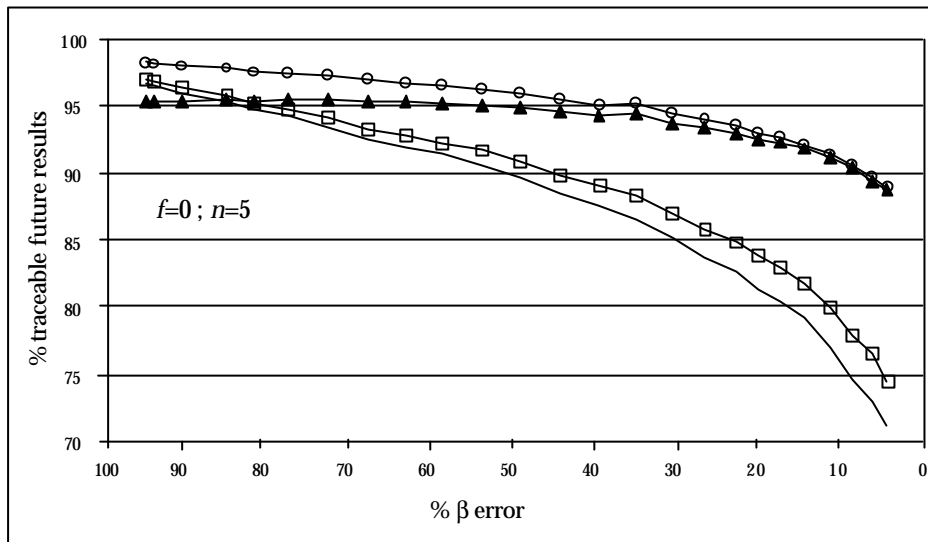
We calculated the percentage of traceable future results for the situations described in Table 2. In this table,  $f$  represents the contribution of the uncertainty of other terms,  $u(F_{\text{hom}})$ , to the intermediate precision and is calculated as:

$$f = \frac{u(F_{\text{hom}})}{RSD_1} \quad (11)$$

The percentage  $r$  represents the contribution of the uncertainty of recovery to the overall uncertainty,  $u$ , calculated with Eq. 6. This percentage is calculated as:

$$r = \frac{\frac{c_{\text{corr}} \cdot u(R)}{R}}{u} \cdot 100 \quad (12)$$

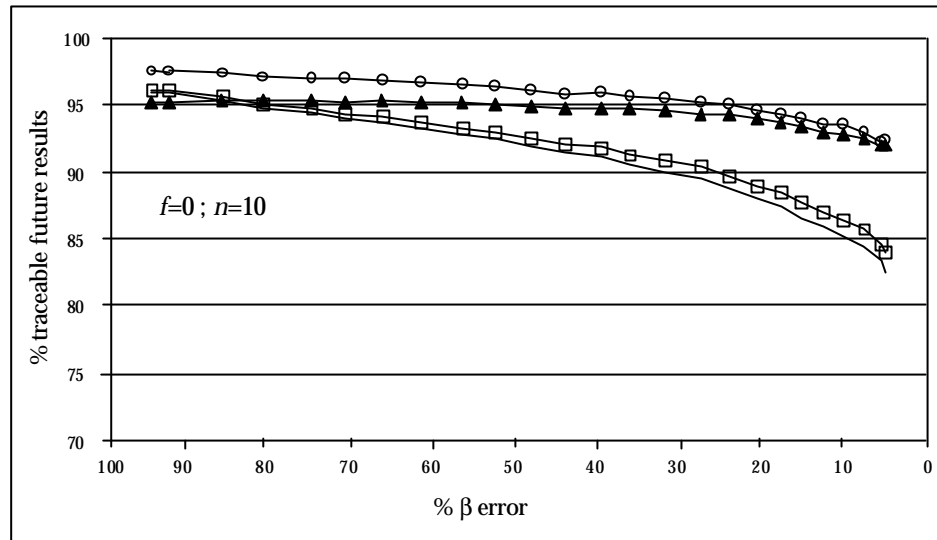
The results obtained are shown in figures 2-4. We plotted the percentage of traceable results as a function of the  $\beta$  error to make the results more general. This is because if we plot the percentage of traceable results as a function of the true apparent recovery,  $R_{true}$ , we would see that this percentage depends also on  $RSD_I$ ,  $u(F_{hom})$  and  $n$  (not shown). Moreover, for the same values of  $r$  and  $\beta$  error we have always the same percentage of traceable future results independently of the values of  $RSD_I$  and  $u(F_{hom})$ .



**Figure 2.** Percentage of traceable future results versus the probability of  $\beta$  error obtained for  $f=0$  and when the apparent recovery is calculated using 5 spiked samples. This gives a contribution of the uncertainty of the apparent recovery,  $r$ , equal to 41%. The percentage of traceable future results is calculated without including non-significant bias (—  $U$ ) and by including non-significant bias with the three approaches described: —□—  $U_B$ ; —▲—  $SUMU$  and —○—  $U(bias)$ .

Fig. 2 shows that the uncertainty calculated without including non-significant bias can be greatly underestimated for lower probabilities of  $\beta$  errors. This underestimation is partially avoided with the  $SUMU$  and the  $U(bias)$  uncertainties. On the other hand, the approach proposed by Barwick *et al* is not a good alternative to avoid this underestimation.

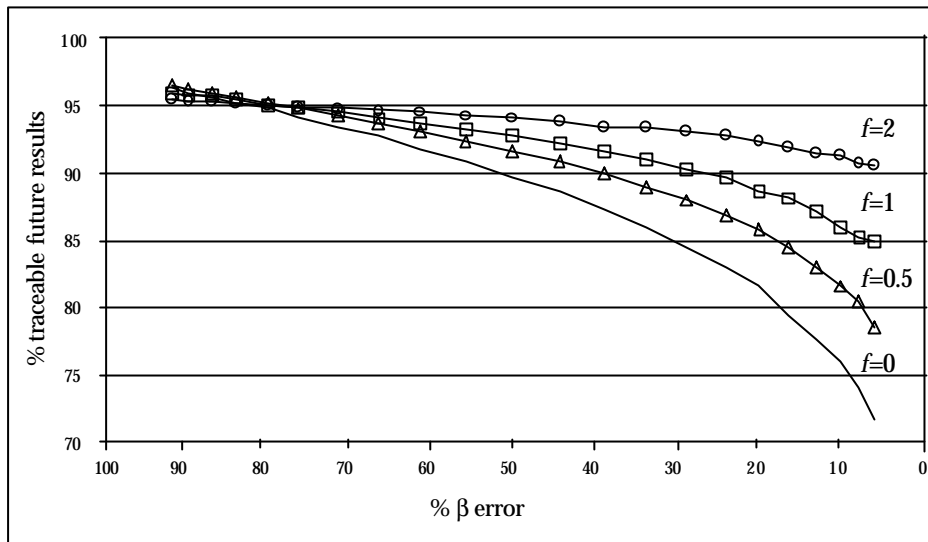
A careful look at figures 2-4 shows that the percentage of traceable results depends on: 1) probability of  $\beta$  error; and 2) contribution of the uncertainty of recovery to the overall uncertainty:



**Figure 3.** Percentage of traceable future results versus the probability of  $\beta$  error obtained for  $f=0$  and when the apparent recovery is calculated using 10 spiked samples. This gives a contribution of the uncertainty of the apparent recovery,  $r$ , equal to 30%. The percentage of traceable future results is calculated without including non-significant bias: (—)  $U$  and by including non-significant bias with the three approaches described: —□—  $U_B$ ; —▲— SUMU and —○—  $U(bias)$ .

### 6.2.1. Probability of $\beta$ error

It is observed that the lower is the probability of  $\beta$  error, the lower is the percentage of traceable future results. This is because the proportional bias increases when the probability of  $\beta$  error decreases. As a result, for lower  $\beta$  errors the percentage of traceable results can be much lower than 95%. On the other hand, for higher  $\beta$  errors the proportional bias is so small that the uncertainty of recovery is higher than the error of not correcting the result. This makes that the percentage of traceable results is higher than 95% for higher  $\beta$  errors.



**Figure 4.** Percentage of traceable results versus the probability of  $\beta$  error for 5 spiked samples analysed and for different values of the ratio  $f$ :  $f=0$  (—) ( $r=41\%$ );  $f=0.5$  ( $\blacktriangle$ ) ( $r=37\%$ );  $f=1$  ( $\blacksquare$ ) ( $r=30\%$ );  $f=2$  ( $\circ$ ) ( $r=20\%$ ). Uncertainty is calculated without including non-significant bias.

### 6.2.2. Contribution of the uncertainty of recovery to the overall uncertainty

The percentage of traceable results depends also on the contribution of the uncertainty of recovery to the overall uncertainty. This contribution is measured by the percentage  $r$  calculated with Eq. 12 (i.e. the higher is  $r$ , the higher is this contribution). Figure 4 shows that, for the same probability of  $\beta$  error, the lower is  $r$  the more close to 95% is the percentage of traceable results obtained when non-significant bias is not included in the final uncertainty. Therefore, if we do not want to include non-significant bias as a component of uncertainty, we should have lower contributions of the uncertainty of recovery. Figures 3 and 4 show that it is not necessary to include non-significant bias for  $r \leq 30\%$ .

The contribution of the uncertainty of recovery to the overall uncertainty,  $r$ , depends on: a) the number of spiked samples analysed,  $n$ ; and b) the contribution of the uncertainty of other terms,  $f$ . To see how  $r$  depends on  $n$  and  $f$ , we can express Eq. 12 as:

$$r = \frac{1}{\sqrt{n} \cdot \sqrt{1 + \frac{1}{n} + f^2}} \cdot 100 \quad (13)$$

This expression is obtained by substituting  $u(R)$  by Eq. 3,  $u$  by Eq. 7 and  $u(F_{\text{hom}})$  by  $f \cdot RSD_{\text{I}}$ . Note that this ratio does not depend on the concentration of the routine samples. Therefore, the same percentage of traceable results was obtained independently of the concentration of the routine samples (not shown).

*a) Number of spiked samples analysed,  $n$*

Eq. 13 shows that the higher is  $n$ , the lower is  $r$ . Therefore, if we do not want to include non-significant bias in the uncertainty budget, we should obtain recoveries by analysing enough number of spiked samples. Figures 2 and 3 show the results obtained when there are no other terms of uncertainty (i.e. for  $f=0$ ). Fig. 2 shows that non-significant bias should be included for  $n=5$ . On the other hand, Fig. 3 shows that this is not necessary for  $n=10$ . Therefore, if we do not want to include non-significant bias as a component of uncertainty, we should estimate recovery by analysing at least 10 spiked samples.

*b) Contribution of other terms of uncertainty,  $f$*

Fig. 4 shows the effect of other terms of uncertainty in the percentage of traceable results. We see that the higher is the contribution of other terms of uncertainty (i.e. the higher is  $f$ ), the closer to 95% is the percentage of traceable future results obtained when uncertainty is calculated without including non-significant bias. This is because the percentage  $r$  diminishes when the contribution of other terms of uncertainty increases (see Eq. 13). Therefore, if the uncertainty of other terms has an important contribution, it is not necessary to include the non-significant bias in the final uncertainty. This can be seen in Fig. 4. Even for  $n=5$ , we obtain good estimates of uncertainty if  $f \geq 1$ .

## 7. CONCLUSIONS

Non-significant bias should be included in the uncertainty budget if the contribution of the uncertainty of recovery to the final uncertainty is higher than 30%. The higher is this contribution, the more important is to include non-significant bias. The only way to avoid this underestimation is with the *SUMU* and the *U(bias)* approaches. If an increase of the uncertainty by including the non-significant bias is not desired, we should analyse the spiked samples enough times to have contributions of the uncertainty of recovery lower than 30%. This number of times depends on whether there are other terms of uncertainty (i.e. such as sample heterogeneity, matrix variability, etc.). If there are no other terms of uncertainty, recovery should be obtained by analysing at least 10 spiked samples. This agrees with the IUPAC guidelines [11] which recommends to verify traceability by preferably analysing the reference sample at least 10 times. If there are other terms of uncertainty it is not necessary to analyse so many times the spiked samples. In this case, the number of times depends on the contribution of the uncertainty of other terms to the overall uncertainty: the more important is this contribution, the less times the spiked samples need to be analysed.

## 8. REFERENCES

- [1] IUPAC, International Union of Pure and Applied Chemistry, *Pure and Applied Chemistry*, **71** (1999) 337-348
- [2] IUPAC, International Union of Pure and Applied Chemistry, *Recommendations for the use of the term "recovery" in analytical procedures*, Draft (2002)  
<<http://www.iupac.org/>>
- [3] A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Accreditation and Quality Assurance*, **7** (2002) 90-94
- [4] A.G. González, M.A. Herrador, A.G. Asuero, *Talanta*, **48** (1999) 729-736
- [5] A. Maroto, R. Boqué, J. Riu, F.X. Rius, *Analytica Chimica Acta*, **440** (2001) 171-184
- [6] Analytical Methods Committee, *AMC Technical Brief*, **1** (2000)
- [7] F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309-316
- [8] S.L.R. Ellison, V.J. Barwick, *Analyst*, **123** (1998) 1387-1392
- [9] S.D. Phillips, K.R. Eberhardt, B. Parry, *Journal of Research of the National Institute of Standards and Technology*, **102** (1997) 577-585
- [10] O. Güell, J.A. Holcombe, *Analytical Chemistry*, **62** (1990) 529A-542A

- [11] International Union of Pure and Applied Chemistry, *Harmonised Guidelines for In-House Validation of Methods of Analysis* (Technical report) (1999)  
[http://www.iupac.org/projects/1997/501\\_8\\_97.html](http://www.iupac.org/projects/1997/501_8_97.html)

## 4.4. Conclusiones

La incertidumbre de los resultados puede estar muy subestimada cuando se concluye erróneamente que el sesgo del método no es significativo. Esta subestimación depende de la contribución de la incertidumbre del sesgo del método a la incertidumbre final: cuanto mayor sea esta contribución, más subestimada está la incertidumbre. En concreto, se ha observado que la incertidumbre de los resultados puede estar muy subestimada cuando la contribución del sesgo del método es mayor del 30%. Este resultado se ha obtenido cuando la trazabilidad se verifica tanto en un intervalo restringido de concentraciones como en un intervalo amplio de concentraciones. En ambos casos, se ha observado que esta subestimación se puede evitar incluyendo el sesgo del método en la incertidumbre final de los resultados. En el caso de que no se desee incluir este término, es necesario verificar la trazabilidad de forma que la incertidumbre del sesgo del método tenga una contribución menor del 30%. Para ello, es necesario utilizar referencias que tengan poca incertidumbre y analizar las referencias un número suficiente de veces. Desde un punto de vista práctico, utilizar referencias que tengan poca incertidumbre implica un mayor coste de las referencias utilizadas. Por otro lado, analizar las referencias un número suficiente de veces implica más trabajo adicional en la verificación de la trazabilidad. Si, por ejemplo, la incertidumbre de la referencia es despreciable, la referencia debería analizarse al menos 10 veces para obtener contribuciones menores al 30%. Esto está de acuerdo con las recomendaciones de la IUPAC<sup>7</sup>, que aconsejan verificar la trazabilidad de los resultados analizando la muestra de referencia al menos 10 veces. La presencia de otros términos de incertidumbre provoca una disminución de la contribución de la incertidumbre del sesgo en la incertidumbre final. Por tanto, cuanto mayor sea la incertidumbre de otros términos menos se subestima la incertidumbre de los resultados.

En el apartado 4.3. sólo se ha estudiado el efecto de concluir erróneamente que el método no tiene un sesgo proporcional cuando la precisión del método varía linealmente con la concentración. En este caso se ha observado que la subestimación de la incertidumbre es siempre la misma independientemente de la concentración de la muestra de rutina. Sin embargo, la precisión también puede

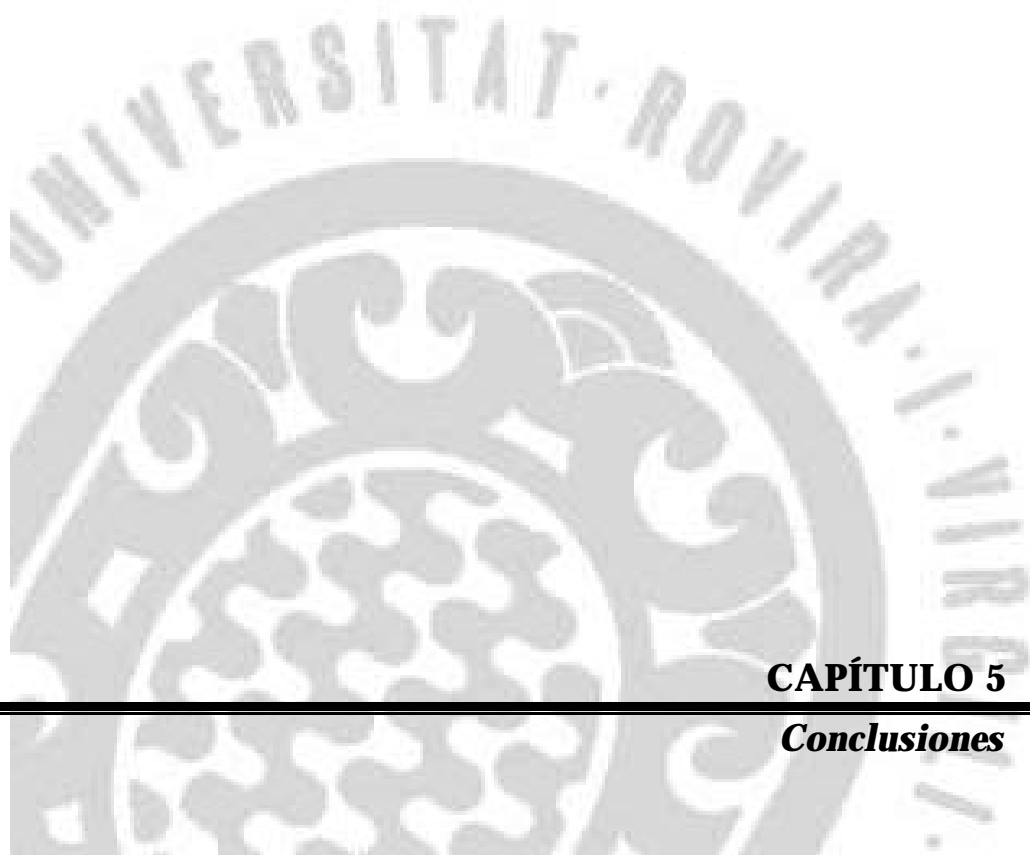


variar exponencialmente con la concentración<sup>8</sup>. En este caso, la subestimación de la incertidumbre también depende de la concentración de la muestra de rutina ya que la contribución del sesgo va variando en función de la concentración: cuanto mayor es la concentración, mayor es la contribución del sesgo. Por tanto, si la precisión varía exponencialmente con la concentración, la incertidumbre está cada vez más subestimada a medida que aumenta la concentración de las muestras de rutina. Por otro lado, la trazabilidad de los resultados se ha verificado en un intervalo amplio de concentraciones asumiendo que el sesgo del método es únicamente proporcional. No obstante, tal y como se ha visto en el capítulo anterior, la trazabilidad debe verificarse calculando tanto el sesgo proporcional como el sesgo constante. En el apartado 4.3. no se ha estudiado el efecto de un sesgo constante porque su contribución es despreciable comparada con la del sesgo proporcional (sobre todo a medida que aumenta la concentración de las muestras de rutina). Sin embargo, en el caso de que la contribución sea importante, se debería verificar que la contribución de la incertidumbre de ambos sesgos no supera el 30% de la incertidumbre total para evitar la subestimación de la incertidumbre.

## 4.5. Referencias

1. S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Chemometrics and Intelligent Laboratory Systems*, **52** (2000) 61-73
2. S. Kuttatharmmakul, D.L. Massart, J. Smeyers-Verbeke, *Analytica Chimica Acta*, **391** (1999) 203-225
3. O. Güell, J.A. Holcombe, *Analytical Chemistry*, **62** (1990) 529A-542A
4. I.H. Lira, W. Wöger, *Measurement Science and Technology*, **9** (1998) 1010-1011
5. I.H. Lira, W. Wöger, *Measurement Science and Technology*, **9** (1998) 1167-1173
6. S.D. Phillips, K.R. Eberhardt, B. Parry, *Journal of Research of the National Institute of Standards and Technology*, **102** (1997) 577-585
7. International Union of Pure and Applied Chemistry, *Harmonised Guidelines for In-House Validation of Methods of Analysis* (Technical report) (1999)  
[http://www.iupac.org/projects/1997/501\\_8\\_97.html](http://www.iupac.org/projects/1997/501_8_97.html)
8. International Organization for Standardization, *Accuracy (Trueness and Precision) of Measurement Methods and Results*, ISO Guide 5725-2, ISO, Geneva (1994)





## **CAPÍTULO 5**

### ***Conclusiones***

---



## 5.1. Conclusiones

En esta tesis doctoral se han desarrollado una serie de metodologías para calcular la incertidumbre de resultados analíticos obtenidos con métodos utilizados de forma rutinaria en el laboratorio que aprovechan la información generada durante la validación del método. Las expresiones propuestas son consistentes con las recomendaciones de la ISO, pueden aplicarse a todos los métodos de rutina y tienen la ventaja de que requieren poco trabajo adicional ya que los métodos analíticos siempre deben validarse antes de analizar las muestras de rutina.

A continuación se detallan las conclusiones que se obtienen a partir de los objetivos enunciados en la sección 1.1:

*1. Desarrollo de una metodología para calcular la incertidumbre de métodos de rutina que se aplican en un intervalo de concentraciones restringido y cuya exactitud se ha verificado a un nivel de concentración utilizando diversas referencias como materiales de referencia certificados, métodos de referencia o ejercicios interlaboratorio.*

En el capítulo 2 se ha desarrollado una aproximación donde se calcula la incertidumbre de métodos de rutina utilizando fundamentalmente la información generada al verificar la exactitud del método a un nivel de concentración. Esta aproximación sólo es válida en un intervalo restringido de concentraciones ya que asume que el sesgo y la precisión del método no varían en el intervalo de aplicación del método.

Uno de los componentes más importantes de incertidumbre es la incertidumbre del procedimiento. Este componente considera la variabilidad experimental de los resultados debida a las condiciones en que se efectúa el análisis (analista, día, calibrado, instrumento, etc.) y se calcula a partir de la precisión intermedia del método. Una forma de disminuir la incertidumbre de los resultados es disminuyendo la incertidumbre del procedimiento. Para ello, debe analizarse la muestra de rutina más de una vez en condiciones intermedias. Por ejemplo, esto puede ser necesario cuando el resultado obtenido al analizar la muestra de rutina está muy próximo al límite permitido. No obstante, las muestras de rutina se

suelen analizar una sola vez o bien se suelen efectuar dos replicados en condiciones de repetibilidad. Si se efectúan varios replicados en condiciones de repetibilidad, no puede disminuirse la incertidumbre del procedimiento a no ser que se divida la precisión intermedia en dos componentes de precisión: la varianza de la repetibilidad y la varianza entre series. De esta forma, se puede disminuir la incertidumbre dividiendo la varianza de la repetibilidad por el número de replicados realizados en condiciones de repetibilidad.

Otro de los componentes de incertidumbre que debe incluirse en la incertidumbre final está asociado a afirmar que los resultados son trazables. Es decir, aunque los resultados resulten ser trazables nunca se tendrá una seguridad del 100% en esta afirmación ya que el sesgo calculado en la verificación de la trazabilidad tiene una incertidumbre asociada. A partir de la información obtenida al verificar la trazabilidad a un nivel de concentración puede calcularse tanto la incertidumbre de la verificación de la trazabilidad como la incertidumbre del procedimiento. Además, también puede utilizarse la información de los estudios de precisión y de los gráficos de control para mejorar la estimación de la incertidumbre de ambos términos. Es aconsejable utilizar también esta información cuando la muestra de referencia se analiza pocas veces en condiciones intermedias.

La incertidumbre final de los resultados siempre contiene el término asociado al procedimiento y a la verificación de la trazabilidad. Además, puede ser necesario incluir dos componentes más en la incertidumbre final de los resultados. El primero de ellos es debido al submuestreo y a los pretratamientos de la muestra. Debe incluirse en la incertidumbre final si las muestras de rutina son heterogéneas o bien cuando se realizan pretratamientos en las muestras de rutina que no se han realizado en la muestra de referencia analizada al verificar la trazabilidad de los resultados. El cálculo de este componente de incertidumbre requiere un trabajo adicional ya que no puede obtenerse a partir de la información de la validación del método. El segundo componente de incertidumbre que puede ser necesario incluir corresponde a la incertidumbre de otros términos y es debido fundamentalmente a factores de variabilidad que no se han considerado en el cálculo de la precisión intermedia y a la variabilidad de las matrices de las muestras de rutina. El cálculo de la incertidumbre de factores de variabilidad no variados representativamente

no siempre requiere un trabajo adicional ya que puede obtenerse utilizando la información generada en los estudios de robustez. Sin embargo, sí se requiere un trabajo adicional para calcular la incertidumbre de la variabilidad de las muestras de rutina. Una forma sencilla de calcular esta incertidumbre es adicionando la misma concentración de analito a matrices que sean representativas de las muestras de rutina.

*2. Desarrollo de metodologías para calcular la incertidumbre de métodos de rutina que se aplican en un intervalo amplio de concentraciones y cuya exactitud se ha verificado utilizando muestras adicionadas y estudios de recuperación.*

En el capítulo 3 se han desarrollado una serie de metodologías que utilizan la información generada al verificar la trazabilidad con muestras adicionadas y estudios de recuperación para calcular la incertidumbre de métodos de rutina con un intervalo amplio de concentraciones.

La verificación de la trazabilidad con muestras adicionadas depende de que el analito se adicione a blancos de muestras o bien a muestras con concentración inicial de analito. Si se dispone de blancos de muestras, la trazabilidad debe verificarse con el método de adiciones estándar. De esta forma, puede calcularse el sesgo proporcional (que depende de la concentración y suele expresarse en términos de recuperación) y el sesgo constante del método. Si no se dispone de blancos de muestras, la trazabilidad no puede verificarse utilizando sólo muestras adicionadas ya que no puede calcularse el sesgo constante. Por tanto, debe utilizarse el método de Youden para calcular este sesgo. El sesgo proporcional se calcula a partir de las muestras adicionadas utilizando el método de adiciones estándar o bien el método de la recuperación media. Es aconsejable utilizar el método de la recuperación media cuando la precisión varía con la concentración ya que se obtiene una incertidumbre de la recuperación menor que con el método de adiciones estándar.

El cálculo de la incertidumbre depende de cómo se haya verificado la trazabilidad de los resultados y de cómo varíe la precisión con la concentración. En el capítulo 3 se ha estudiado cómo calcular la incertidumbre cuando el analito se adiciona a

blancos de muestras o bien a muestras con concentración inicial. Las expresiones propuestas suponen que la precisión varía linealmente con la concentración o bien que se mantiene aproximadamente constante. Además, se supone que la incertidumbre de la cantidad adicionada es despreciable. En el caso de que la trazabilidad se verifique con otras referencias (como, por ejemplo, materiales de referencia o bien adicionando un isótopo del analito), podrían aplicarse las expresiones propuestas cuando la trazabilidad se verifica utilizando blancos de muestras siempre y cuando la incertidumbre de los valores de referencia sea despreciable.

Las metodologías propuestas tienen la ventaja de que requieren poco trabajo adicional para calcular la incertidumbre ya que a partir de los estudios de recuperación con muestras adicionadas puede calcularse: la incertidumbre del procedimiento, la incertidumbre de la verificación de la trazabilidad y la incertidumbre de la variabilidad de las muestras de rutina (en términos de concentración y de matriz). Además, al igual que en el capítulo 2, también se puede utilizar la información de los estudios de precisión y de los gráficos de control para mejorar la estimación de la incertidumbre del procedimiento y de la verificación de la trazabilidad. Asimismo, los estudios de robustez pueden utilizarse para calcular la incertidumbre de factores que no se hayan variado representativamente en el cálculo de la precisión intermedia. Por tanto, el único trabajo adicional para calcular la incertidumbre es debido al cálculo de la incertidumbre del submuestreo y/o pretratamientos.

*3. Estudio de diversas estrategias para evitar la subestimación de la incertidumbre de los resultados analíticos al concluir erróneamente en la verificación de la exactitud que el sesgo del método no es significativo.*

En el capítulo 4 se ha comprobado que la incertidumbre de los resultados puede estar muy subestimada si se concluye erróneamente que los resultados son trazables. Esta subestimación depende de la contribución de la incertidumbre del sesgo en la incertidumbre final de los resultados: cuanto mayor sea esta contribución, más subestimada está la incertidumbre. En concreto, la



incertidumbre de los resultados puede estar muy subestimada si esta contribución es mayor del 30%.

Las diversas estrategias para evitar la subestimación de la incertidumbre están basadas en incluir el sesgo calculado durante la verificación de la trazabilidad. No obstante, sino se desea incluir el sesgo como un componente de incertidumbre, es necesario verificar la trazabilidad de los resultados de forma que la contribución de la incertidumbre del sesgo sea menor del 30%. Para ello, es necesario utilizar referencias con poca incertidumbre asociada y analizar las muestras de referencia un número suficiente de veces. La presencia de otros términos de incertidumbre provoca una disminución de la contribución de la incertidumbre del sesgo. Por tanto, cuanto mayor sea la incertidumbre de otros términos menos se subestima la incertidumbre de los resultados.

## **5.2. Perspectivas futuras**

Esta tesis se ha centrado únicamente en el cálculo de la incertidumbre en métodos analíticos cuantitativos de rutina. No obstante, hay otros campos dentro de la química analítica donde el cálculo de la incertidumbre está menos estudiado y donde queda bastante trabajo por realizar. Algunos ejemplos son el análisis bioquímico y el análisis cualitativo. A continuación se exponen algunos trabajos pendientes en el cálculo de la incertidumbre en métodos analíticos cuantitativos de rutina:

*1. Calcular la incertidumbre en un intervalo amplio de concentraciones utilizando materiales de referencia y técnicas de regresión*

En el capítulo 3 se han desarrollado diversas expresiones para calcular la incertidumbre en un intervalo amplio de concentraciones utilizando técnicas de regresión y muestras adicionadas. Sin embargo, estas expresiones no pueden aplicarse cuando la incertidumbre de los valores de referencia no es despreciable. Por tanto, deben desarrollarse expresiones para calcular la incertidumbre en este caso. Para ello, deben buscarse técnicas de regresión como, por ejemplo, los

mínimos cuadrados bivariantes (BLS) que, a diferencia de la regresión por mínimos cuadrados (OLS) y de la regresión por mínimos cuadrados ponderados (WLS), consideren la incertidumbre de los valores de referencia. Otra posibilidad es modificar las expresiones del intervalo de predicción en OLS y WLS para mejorar la estimación de las incertidumbres obtenidas con estas técnicas de regresión.

*2. Calcular la incertidumbre en un intervalo amplio de concentraciones cuando la precisión del método varía exponencialmente con la concentración*

En el capítulo 3 se ha calculado la incertidumbre asumiendo que la precisión se podía mantener aproximadamente constante con la concentración o bien variar linealmente con la concentración. No obstante, la precisión del método también puede variar exponencialmente siguiendo una expresión similar a la propuesta por Horwitz. Por tanto, deben establecerse metodologías para modelar la precisión con la concentración e incluir la incertidumbre asociada a predecir la precisión con este modelo en la incertidumbre final de los resultados.

*3. Recalcular la incertidumbre utilizando la información obtenida en el proceso de aseguramiento de la calidad*

La incertidumbre de los resultados se calcula utilizando la información obtenida durante la verificación de la trazabilidad de los mismos. No obstante, la incertidumbre de los resultados puede cambiar con el tiempo ya que puede variar tanto la precisión del método como su sesgo. Por tanto, deberían establecerse aproximaciones para recalcular la incertidumbre en función de los resultados que se vayan obteniendo en los gráficos de control.



**APÉNDICES**

---



**Evaluating uncertainty in routine analysis (Trends in Analytical Chemistry, 18 (1999) 577-584)**

Alicia Maroto, Ricard Boqué, Jordi Riu and F. Xavier Rius

Department of Analytical and Organic Chemistry. Rovira i Virgili University.  
Pl. Imperial Tàrraco, 1. 43005-Tarragona. Spain.

**ABSTRACT**

In this paper, we critically describe the different approaches proposed so far for calculating uncertainty in chemical measurements: 1) The ISO approach, adapted to the analytical field by EURACHEM (commonly known as “bottom-up”), and 2) the Analytical Methods Committee approach (commonly known as “top-down”), based on interlaboratory information. We also propose a new procedure, which is totally consistent with the ISO approach in the sense that all the sources of error are identified, quantified and combined, but which is conceptually more similar to the “top-down” approach due to its holistic character. This new procedure estimates uncertainty from the information generated during the process of assessing accuracy.

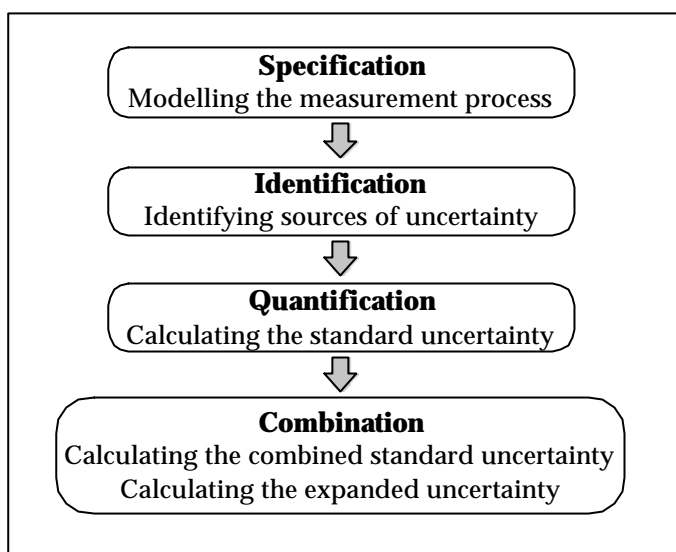
**Keywords:** Uncertainty in chemical measurements, Assessing accuracy; Validation

## 1. THE CONCEPT OF UNCERTAINTY

The latest edition of the International Vocabulary of Basic and General Terms in Metrology, VIM, states that uncertainty is “a parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand” [1]. Uncertainty is a parameter that defines the quality of the measurement results. Consequently, results without uncertainty should not be taken seriously [2]. Traceability and uncertainty are totally related concepts and, to understand the link between them, the concept of uncertainty must be interpreted. If an analytical result is expressed as *Value 1* ± *Uncertainty*, where *Uncertainty* is expressed as a confidence interval, all analysts should verify that *Value 1* is an unbiased value which is traceable, whenever possible, to an internationally recognised reference which guarantees that the confidence interval comprises the true value. Once the traceability of an analytical procedure has been demonstrated, uncertainty can then be calculated for the results of applying the chemical procedure to routine samples similar to the test sample used in the assessment of traceability. This paper reviews two approaches for calculating uncertainty, the ISO approach and the Analytical Methods Committee approach, and proposes an approach based on calculating the uncertainty from the information of the validation process.

## 2. THE ISO APPROACH FOR CALCULATING UNCERTAINTY

This approach estimates the overall uncertainty by identifying, estimating and combining all the sources of uncertainty associated with the measurement. Because of the way the overall uncertainty is estimated, it has been defined as a “bottom-up” approach [3-6]. The procedure for quantifying the overall uncertainty is summarised in the four main steps shown in Fig. 1.



**Figure 1.** Basic steps for calculating the measurement uncertainty according to the ISO approach.

The *specification step* obtains a relationship between the analytical result and the parameters on which the analytical result depends, whenever possible, in the form of an algebraic expression such as  $Value = f(x_1, \dots, x_n)$ . Normally, the difficulties of formalising the algebraic relationship between analytical results and the parameters that intervene in the analytical procedure become apparent. In this case, the specification step is done for the different steps in which the measurement procedure can be broken down, e.g. sampling, pre-processing steps, instrumental measurement, etc.

The *identification step* identifies all the sources of uncertainty present in the different steps of the analytical procedure. After this, all the components of uncertainty that are expected to have a significant contribution to the overall uncertainty are quantified (*Quantification step*). These components can be quantified from experimental work ("Type A" uncertainty) or by using previous information or the analyst's experience ("Type B" uncertainty).

The *combined standard uncertainty*,  $u$ , is obtained by using the principle of error propagation to combine all the sources of uncertainty expressed as *standard uncertainties*:

$$u(\text{Value 1}) = \sqrt{\sum_{i=1}^n \left( \frac{\partial \text{Value 1}}{\partial x_i} \cdot u(x_i) \right)^2 + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left( \frac{\partial \text{Value 1}}{\partial x_i} \cdot \frac{\partial \text{Value 1}}{\partial x_j} \cdot \text{cov}(x_{ij}) \right)} \quad (1)$$

where:

$x_i$  ( $i=1, \dots, n$ ) are the individual parameters that influence the analytical result.

$u(x_i)$  is the uncertainty of the  $i$ th parameter expressed as standard deviation

$\text{cov}(x_{ij})$ : is the covariance between  $x_i$  and  $x_j$

If all parameters contributing to the measurement result are independent, the covariance term,  $\text{cov}(x_{ij})$ , is zero.

If “Type B” uncertainties are used, it may be necessary to convert a confidence interval into a standard uncertainty. In this case, information is required about the distribution of the value and degrees of freedom. If this information is not available, the distribution can be assumed to be rectangular.

Finally, the *calculation of the expanded uncertainty* provides a confidence interval within which the value of the measurand is expected to lie. The expanded uncertainty is obtained by multiplying the combined standard uncertainty by a coverage factor,  $k$  [3,4] which assumes that the distribution of the measurement result, *Value 1* is normal. However, since the combined standard uncertainty is obtained by combining different types of distribution, it is more correct to use the  $t$  tabulated value for the level of confidence chosen and the effective degrees of freedom,  $\nu_{\text{eff}}$ , calculated using the Welch-Satterthwaite approach. [7,8]

$$\mathbf{n}_{\text{eff}} = \frac{u(\text{Value 1})^4}{\sum_{i=1}^n \frac{\left[ \left( \frac{\partial \text{Value 1}}{\partial x_i} \cdot u(x_i) \right)^2 \right]^2}{\mathbf{n}_i}} \quad (2)$$

where  $\nu_i$  corresponds to the degrees of freedom of the  $i$ th component of uncertainty.



If the effective degrees of freedom, or the degrees of freedom corresponding to the largest and dominant contribution, are large enough, the *t*-distribution can be approximated to a normal distribution. In these circumstances, the coverage factors  $k=2$  and  $k=3$  can be adopted for a level of confidence of approximately 95% and 99% respectively. In practice,  $k=2$  is recommended for most purposes.

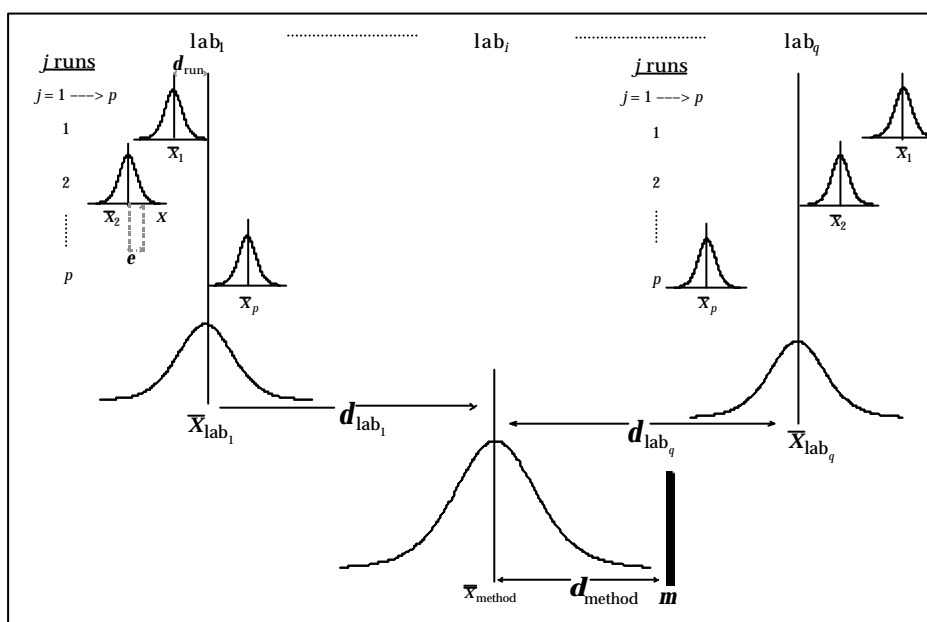
## **2.1. Advantages and drawbacks**

The ISO approach is a well structured methodology which was originally developed to estimate the uncertainty in physical measurements [3] and subsequently adapted to chemical measurement processes by EURACHEM [4]. This approach helps to improve knowledge of analytical techniques and principles since all the components that contribute to uncertainty and their magnitude must be estimated. Furthermore, the uncertainty of chemical measurements can be reduced by optimising those steps that significantly contribute to the overall uncertainty.

However, the adaptation of this methodology to chemical measurements, which are completely different to the physical ones, causes important problems [9,10]. On the one hand, the number of potential sources of error in chemical measurements and the structure of the errors, mainly of the random type, make it difficult to provide valid estimates of chemical uncertainty by budget analysis. Moreover, devising experiments that can accurately characterise uncertainties in method bias and other systematic effects is extremely complex. Therefore, the analyst is impelled to make estimations which are based on prior data or experience ("Type B" uncertainties) and which, despite being based on expert knowledge and common sense, may be considerably different from the estimations of other analysts due to the difficulties of fully understanding chemical measurement methods. What's more, the application of this approach in routine chemical measurements involves significant costs in time and effort. This often prevents its widespread application.

### 3. THE ANALYTICAL METHODS COMMITTEE APPROACH FOR CALCULATING UNCERTAINTY

The Analytical Methods Committee approach (commonly known as “top-down”) quantifies uncertainty using information generated when different laboratories are involved in method performance exercises [11]. In this approach the laboratory is seen from a “higher level”, i.e. as a member of the population of laboratories. As a result, systematic and random errors that occur within individual laboratories become random errors when they are considered from this “higher level” (see Fig. 2).



**Figure 2.** A “top-down” view of the experimental design of measurements for calculating uncertainty. Different runs ( $j=1, \dots, p$ ) of an analytical method are performed in several laboratories ( $i=1, \dots, q$ ) to provide a mean value of the method,  $\bar{x}_{\text{method}}$ . The random error of individual results,  $\varepsilon$ , the run bias,  $\delta_{\text{run}}$ , the laboratory bias,  $\delta_{\text{lab}}$ , and the method bias,  $\delta_{\text{method}}$ , are shown.

When an analytical procedure is applied to a homogeneous sample, the results obtained can be expressed as the sum of [11]:

$$\text{Value 1} = \mu + \delta_{\text{method}} + \delta_{\text{lab}} + \delta_{\text{run}} + \epsilon \quad (3)$$

where *Value 1* is the result obtained,  $\mu$  is the true value,  $\delta_{\text{method}}$  is the method bias,  $\delta_{\text{lab}}$  is the laboratory bias,  $\delta_{\text{run}}$  is the bias due to the conditions where the measurement is made (day, analyst, etc.) and  $\epsilon$  is the random error.

In the “top-down” approach, the consensus value is obtained from the analytical results of a population of laboratories. From this higher metrological level, the laboratory and the run bias are random variables with an associated variance. The method bias is seen as a systematic error from this level.

Since uncertainty must include all factors that can affect the analytical results (i.e. sources of variation arising from systematic errors and random errors) the sources of uncertainty depend on the different values that  $\delta_{\text{lab}}$ ,  $\delta_{\text{run}}$  and  $\epsilon$  can have and on the uncertainty arising from the estimation of  $\delta_{\text{method}}$ . Therefore, the uncertainty of a future measurement is expressed as:

$$U(\text{Value 1}) = t_{\alpha/2, \text{eff}} \sqrt{s_{\epsilon}^2 + s_{\text{run}}^2 + s_{\text{lab}}^2 + u_{\text{method}}^2} \quad (4)$$

where  $\sigma_{\epsilon}^2$ ,  $\sigma_{\text{run}}^2$  and  $\sigma_{\text{lab}}^2$  are the true variances associated to the different values that the random error,  $\epsilon$ , the run bias,  $\delta_{\text{run}}$ , and the laboratory bias,  $\delta_{\text{lab}}$ , can take respectively. All these variances can be estimated from the results of a collaborative trial, whereas the uncertainty related to the estimation of the method bias,  $u_{\text{method}}$ , can only be estimated if there is reference value available.  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value, at a given  $\alpha$  probability and  $\nu_{\text{eff}}$  degrees of freedom. The effective degrees of freedom,  $\nu_{\text{eff}}$ , are estimated using the Welch-Satterthwaite approach [7]. The original article in which this approach was proposed [11] uses the coverage factor,  $k$ , instead of using the effective  $t$  tabulated value. However, it is more correct to use  $t_{\alpha/2, \text{eff}}$  whenever different  $t$ -distributions are combined [8].

Since the bias of the method is a systematic error when the laboratory is seen from this metrological level, an estimation of the method bias is given by the difference

between a reference value (e.g. a certified reference material) and the consensus value,  $\bar{x}_{\text{method}}$ , obtained by the laboratories in the collaborative trial:

$$\hat{d}_{\text{method}} = \bar{x}_{\text{method}} - \bar{x}_{\text{ref}} \quad (5)$$

The reference value,  $\bar{x}_{\text{ref}}$ , is not always available. The uncertainty associated to this estimation,  $u_{\text{method}}$ , is obtained by combining the variance of the means of the interlaboratory results,  $s_{\bar{x}_{\text{method}}}^2$ , and the variance related to the reference value,  $s_{\bar{x}_{\text{ref}}}^2$ :

$$u_{\text{method}} = \sqrt{\frac{s_{\bar{x}_{\text{method}}}^2}{q} + s_{\bar{x}_{\text{ref}}}^2} \quad (6)$$

where  $q$  is the number of laboratories in the collaborative trial. Since the variance related to the reference value is normally expected to be lower than the variance of the grand mean of a collaborative trial, in equation 6 it has been assumed that the difference between these variances is statistically significant. Otherwise, a pooled variance should be calculated.

This uncertainty must be included as a component of the overall uncertainty of the measurement result. Moreover, if the method bias estimated is statistically significant, the analytical procedure used by the laboratories should be revised.

This expression does not include other effects, such as the uncertainty arising from sampling, pre-processing steps to homogenise the sample or differences between the matrix composition of the sample analysed in the interlaboratory exercise and the samples analysed in routine measurements.

### 3.1. Advantages and drawbacks

The “top-down” approach can only be applied if interlaboratory information is available. Therefore, on many occasions it cannot be applied because a considerable number of measurands do not have interlaboratory exercises for

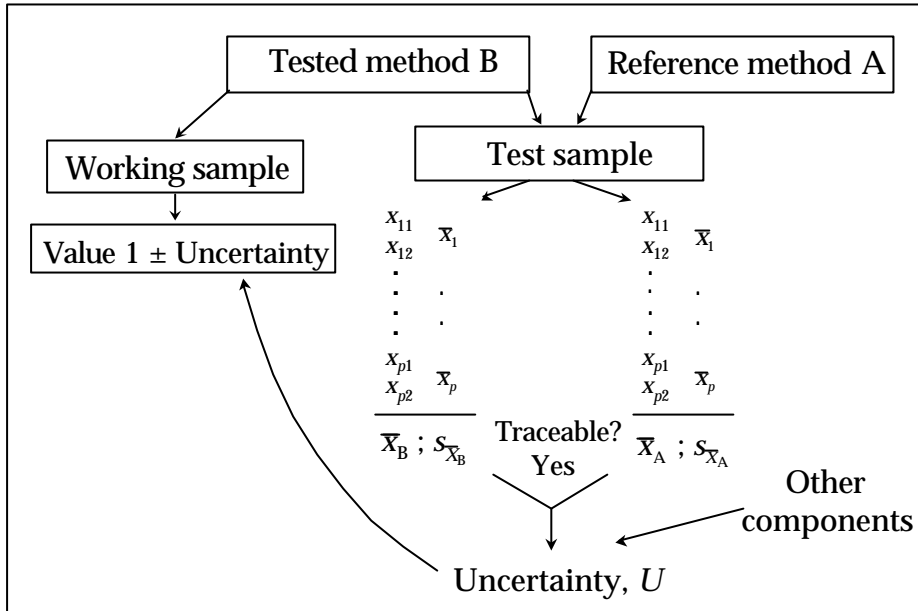
many matrices. However, when uncertainty is estimated with this approach, individual factors which contribute to the variability of chemical measurements but which are intrinsically unknowable (despite the knowledge of the analyst about the measurement method and procedure) are included. Moreover, uncertainty can be estimated for different concentrations because the reproducibility of the method can be modelled against concentration with the Horwitz function [12].

In this approach, the laboratory that has obtained the measurement result is considered to be a random variable. Therefore, the uncertainty estimated with this approach should only be applied if the laboratory that obtains the measurement result varies randomly and not when the laboratory is a fixed effect. If we are to calculate the uncertainty of the results obtained in routine analysis by a given laboratory (i.e. the laboratory is a fixed effect), the uncertainty estimated with this approach may have little to do with the uncertainty of the measurement results. For instance, if the laboratory bias is very small, the uncertainty may be overestimated. However, if the laboratory has one of the largest biases of the population, the uncertainty may be considerably underestimated. Moreover, the uncertainty of an individual laboratory may be estimated incorrectly because of differences between the within-laboratory precision of the individual laboratory and the average value of the within-laboratory precision estimated in the collaborative trial.

#### **4. QUANTIFYING UNCERTAINTIES USING THE INFORMATION FROM THE VALIDATION PROCESS**

This approach estimates the uncertainty of future measurement results using the information from the process of assessing the accuracy of an analytical method in a laboratory. The accuracy of an analytical method can be tested with results obtained by different laboratories [13] or from within a laboratory [14]. In Figure 3, the accuracy of a tested method B is assessed in a laboratory against a reference method A. In this process, information is obtained about the different precision estimates of the analytical method that will subsequently be applied to routine

samples. As can be seen in the figure, components of uncertainty that are not considered in this validation process are subsequently included to obtain the overall uncertainty.



**Figure 3.** Assessment of accuracy of a tested method B against a reference method A. The precision information obtained in this process is used to estimate the uncertainty of future routine samples.

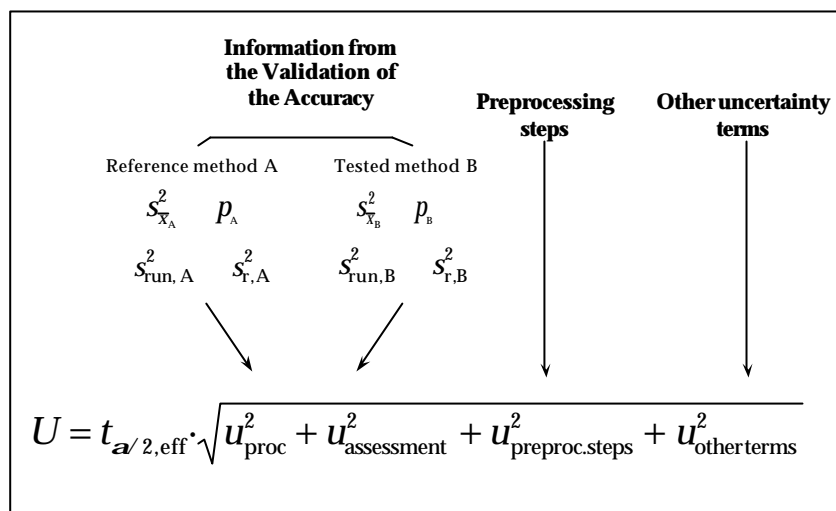
This approach does not strictly estimate the uncertainty of a future routine sample because the information generated is related to a test sample analysed in the past. However, if the test sample is homogeneous, stable and representative (i.e. similar to the routine samples), the estimation can be considered to be good whenever the quality assurance procedures are implemented in the laboratory effectively.

With this approach, a future result can be expressed as:

$$Value1 = m + d_{\text{other terms}} + d_{\text{preproc.steps}} + d_{\text{assessment}} + d_{\text{run}} + e \quad (7)$$

where *Value 1* is the result obtained,  $\mu$  is the true value of the result,  $\varepsilon$  is the random error,  $\delta_{\text{run}}$  is the bias due to the conditions in which the measurement was

carried out (day, analyst, instrument, etc.),  $\delta_{\text{assessment}}$  is the bias estimated in the assessment of accuracy,  $\delta_{\text{other terms}}$  is the bias associated to higher levels of traceability or to other biases not previously considered.



**Figure 4.** Expression of the overall uncertainty using information from the validation of the accuracy

The sources of uncertainty of the measurement result depend on  $\delta_{\text{other terms}}$ ,  $\delta_{\text{preproc.steps}}$ ,  $\delta_{\text{assessment}}$ ,  $\delta_{\text{run}}$  and  $\varepsilon$ . The expression of the overall uncertainty is shown in Figure 4.  $u_{\text{proc}}$  is the uncertainty associated to the analytical procedure, i.e. to  $\delta_{\text{run}}$  and  $\varepsilon$ ,  $u_{\text{assessment}}$  is the uncertainty associated to the estimation of  $\delta_{\text{assessment}}$ , and  $u_{\text{other terms}}$  and  $u_{\text{preproc.steps}}$  are associated to the estimations of  $\delta_{\text{other terms}}$  and  $\delta_{\text{preproc.steps}}$  respectively.  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value, at a given  $\alpha$  probability and  $\nu_{\text{eff}}$  degrees of freedom.  $s_{\bar{x}}^2$  is the variance of the means,  $p$  is the number of runs carried out to estimate the grand mean ( $\bar{x}$ ),  $s_{\text{run}}^2$  is the between-run variance and  $s_r^2$  is the repeatability variance. These last two variances can be estimated by using an appropriate experimental design [13].

Since accuracy is assessed within the laboratory, the bias related to other terms, such as higher levels of traceability, and the bias related to the assessment of accuracy,  $\delta_{\text{assessment}}$ , are systematic errors whereas the run bias and the random error are random variables.

From the information generated from assessing accuracy, the uncertainty related to the analytical procedure (i.e. due to the application of the analytical method to future routine samples) and the uncertainty due to the bias estimated in the assessment of accuracy,  $\delta_{\text{assessment}}$ , can be estimated.

#### 4.1. Uncertainty arising from the analytical procedure

The uncertainty which arises from the experimental variation when applying the analytical procedure to obtain future measurement results is associated to the different values that the run bias and the random error can take within the laboratory. If every routine sample is analysed in  $p_s$  different runs and in each run  $n_s$  replicates are carried out, the uncertainty of the grand mean of these measurements,  $\bar{x}_s$ , is given by the between-run variance, the repeatability variance and the number of runs and replicates carried out within each run.

$$u_{proc} = \sqrt{\frac{s_{run}^2}{p_s \cdot n_s} + \frac{s_r^2}{n_s}} \quad (8)$$

However, the result of a future routine sample is seldom obtained as the grand mean of measurement results in  $p_s$  different runs and with  $n_s$  replicates within each run. Normally, the result is obtained with only one analysis (i.e.  $p_s=n_s=1$ ) or, at most, as the mean of three replicates (i.e.  $p_s=1; n_s=3$ ). In both cases, information about the between-run variance and about the repeatability variance cannot be obtained because of the small number of replicates and runs carried out. However, the between-run variance and the repeatability variance obtained by assessing the accuracy of the analytical method can be used as estimates of the  $s_{run}^2$  and  $s_r^2$  related to the routine sample analysed. Therefore, the uncertainty of the procedure can be calculated with equation 8.



## 4.2. Uncertainty arising from the assessment of accuracy

The bias estimated in the assessment of accuracy is calculated from the difference between the mean of the analytical results of a test sample (obtained by varying representatively the run and the replicate) and the reference value assigned to this test sample. In Figure 3 the reference value is estimated using a reference method. The bias calculated depends on the metrological level of the reference used. For instance, if the reference used is a reference method carried out within the laboratory, only the method bias is estimated. If the reference used is a CRM, the bias estimated is the laboratory and the method bias.

In assessing the accuracy, the run and the replicate can be representatively varied by following a two-factor fully-nested design [15]. The grand mean calculated from the analytical results can be seen as:

$$\bar{x} = \bar{x}_{\text{ref}} + \mathbf{d}_{\text{assessment}} \quad (9)$$

where  $\bar{x}$  is the grand mean of the set of results obtained in the fully-nested design ( $\bar{x}_B$  in Fig. 3),  $\bar{x}_{\text{ref}}$  is the reference value ( $\bar{x}_A$  in Fig. 3) and  $\mathbf{d}_{\text{assessment}}$  is the bias estimated in the assessment of accuracy.

The uncertainty related to the bias estimated in the assessment of accuracy,  $u_{\text{assessment}}$ , is obtained by combining the variance of the reference value and the variance of the grand mean obtained for the test sample with the tested method.

$$u_{\text{assessment}} = \sqrt{s_{\bar{x}_{\text{ref}}}^2 + \frac{s_{\bar{x}}^2}{p}} \quad (10)$$

In Figure 3,  $s_{\bar{x}}^2$  and  $p$  correspond to  $s_{\bar{x}_B}^2$  and  $p_B$  respectively.  $s_{\bar{x}_{\text{ref}}}^2$  is the variance associated to the reference value. In Figure 3 this variance is calculated as:

$$S_{\bar{x}_{ref}}^2 = \frac{S_{x_A}^2}{P_A} \quad (11)$$

In Eq. 10, it has been supposed that the difference between the variance of the reference value and the variance of the grand mean is statistically significant. If it is not, a pooled variance should be calculated.

If the bias estimated is not statistically different from zero, the method is traceable to the reference value. However, if it is significant, the method is not traceable and the analytical procedure must be revised. It should be stressed that, even if method is traceable, there is still a component of uncertainty due to the assessment of traceability itself.

### **4.3. Uncertainty arising from preprocessing steps**

This term includes the uncertainty arising from the subsampling and/or from the preprocessing steps not carried out in the assessment of accuracy. It can be estimated with a sample which has the same characteristics as the routine samples which are to be subsequently analysed. The subsampling and/or the preprocessing steps must be carried out by representatively changing all the factors that can affect them, i.e. operator, material, time, etc. After this, the analytical results of each portion of the sample subsampled and/or pretreated are obtained under repeatability conditions. Finally, the variances associated to these steps can be calculated from an analysis of variance.

### **4.4. Uncertainty associated to other terms**

This includes all the sources of uncertainty that have not been taken into account in the other terms of uncertainty (e.g. the uncertainty associated to the estimation of bias related to higher levels of traceability). Like the bias estimated in the assessment of accuracy, the uncertainty of higher traceability levels is calculated by combining the variance of both references. If the bias estimated is statistically not

different from zero, the reference used in the assessment of accuracy is traceable to the higher reference.

#### **4.5. Advantages and drawbacks**

The new proposal uses the information generated in the process of assessing the accuracy of a given analytical procedure. Evaluating the uncertainty with this new approach has the advantage that the effort made to check the accuracy can also be used to calculate the uncertainty of future routine measurements. Therefore, no extra work of importance needs to be done. The main benefits of this approach are, then, its conceptual and practical simplicity, its low cost and its universal application.

Compared to the ISO approach, the new proposal is more holistic. This means that it is no longer necessary to know the various individual steps into which the analytical method can be broken down because the analytical procedure is taken as a whole.

The new approach, however, also has some drawbacks, most of which are due to constraints in the validation of the accuracy of measurement methods. Since in practice the accuracy of the analytical methods can only be assessed by a limited number of references, which are high in the traceability hierarchy, the analyst often has to resort to references which are at lower levels. The analyst should be aware that the lower the reference is in the hierarchy, the more terms need to be added to the uncertainty budget in order to include the uncertainties arising from systematic effects. Moreover, the analyst must be aware that the factors chosen to be varied in the assessment of accuracy cannot include some of the sources of variation of the measurement method. Therefore, depending on the number of factors chosen, other terms of uncertainty not previously taken into account in the assessment of accuracy must be included.

## 5. ACKNOWLEDGMENTS

The authors would like to thank the DGICYT (project num. BP96-1008) for financial support and the CIRIT from the Catalan Government for providing Alicia Maroto's doctoral fellowship.

## 6. REFERENCES

- [1] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *International Vocabulary of Basic and General Terms in Metrology, VIM*, ISO, Geneva, 1993
- [2] P. de Bièvre, *Accred. Qual. Assur.*, **2** (1997) 269
- [3] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993
- [4] EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY, UK, 1995
- [5] A. Williams, *Accred. Qual. Assur.*, **1** (1996) 14
- [6] S. Ellison, W. Wegscheider, A. Williams, *Anal. Chem.*, **1** (1997) 607 A
- [7] F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309
- [8] C.F. Dietrich, *Uncertainty, Calibration and Probability. The statistic of Scientific and Industrial Measurement*, 2<sup>nd</sup> Ed. Adam Hilger, Bristol, 1991
- [9] W. Horwitz, *Journal of AOAC International*, **81** (1998) 785
- [10] W. Horwitz, R. Albert, *Analyst*, **122** (1997) 615
- [11] Analytical Methods Committee, *Analyst*, **120** (1995) 2303
- [12] W. Horwitz, *Analytical Chemistry*, **54** (1982) 67 A
- [13] International Organization for Standardization, *ISO 5725. Accuracy (trueness and precision) of measurement methods and results*. ISO. Geneva, 1994
- [14] D.L. Massart, J. Smeyers-Verbeke, *Robust process analytical methods for industrial practice and their validation*. Report Project SMT4-CT95-2031, 1996
- [15] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Anal. Chim. Acta*, **391** (1999) 173

**Critical discussion on the procedures to estimate uncertainties in chemical measurements (*Química Analítica*, **19** (2000) 85-94)**

Alicia Maroto, Ricard Boqué, Jordi Riu and F. Xavier Rius

Department of Analytical and Organic Chemistry. Rovira i Virgili University.  
Pl. Imperial Tàrraco, 1. 43005-Tarragona. Spain.

**ABSTRACT**

In this paper, we discuss the concept of uncertainty in analytical measurements and its relationship with other related metrological parameters, such as accuracy or precision. Subsequently, we briefly explain and critically review the main approaches proposed so far for calculating uncertainty in analytical measurements: 1) The ISO approach, adapted to the analytical field by EURACHEM (commonly known as “bottom-up”), 2) the Analytical Methods Committee approach (commonly known as “top-down”), based on interlaboratory information, and 3) a recently proposed approach based on using the information generated during the process of assessing the accuracy within a laboratory.

**Keywords:** uncertainty, analytical measurement, assessment of accuracy

## INTRODUCTION

Analytical results are used as pieces of information for various purposes. Since important decisions are often based on these results, analysts are impelled to provide not only the numerical values of the analytical results but also proof of their accuracy and the degree of confidence (i.e. the uncertainty). Uncertainty is, then, a parameter that defines the quality of a given result. Without uncertainty, results obtained by different laboratories or by different methods cannot be compared and decisions derived from these analytical results cannot be taken. For instance, decisions about compliance with limit values specified in a regulation or a commercial specification cannot be taken without the uncertainty of the result. Therefore, measurement results must always be expressed as *Result  $\pm$  Uncertainty* so that the measurement can be used for a certain aim. Results without uncertainty should not be trusted.

Potential users take important decisions based on the uncertainty of measurements so they will be more confident if it is estimated correctly. Some users, however, are less confident of results reported with uncertainty because they do not fully understand the concept. Even so, uncertainty should always be calculated and never underestimated to prevent incorrect decisions or unnecessary discrepancies between results.

Uncertainty is often confused with other related metrological parameters such as precision and accuracy. Therefore, the first objective of this paper is to clarify the relationships and the differences between these concepts.

Estimating uncertainty in analytical measurements is not straightforward due to the large number of errors and sources of variability to report in an analytical procedure. Because of this, and because of the increasing necessity to report analytical results with their associated uncertainty, the estimation of uncertainty has become one of the focuses of interest in the field of analytical measurement [1].

The approach used to estimate uncertainty depends on the way in which all the sources of uncertainty of an analytical measurement are evaluated. The most

popular approach is the ISO approach (also known as the “bottom-up” approach [2-5]). In this approach, all the sources of uncertainty are identified, estimated and combined to estimate the overall uncertainty. Two other holistic approaches have also been proposed: the approach of the Analytical Methods Committee [6] (known as “top-down”), which uses the information generated in collaborative trials, and another approach [7,8] which uses information from the assessment of accuracy. The second objective of this paper is to briefly explain and critically review these approaches.

## **UNCERTAINTY AND ITS RELATIONSHIP WITH OTHER METROLOGICAL PARAMETERS**

### **The uncertainty concept**

The ISO standard 3534-1 defines uncertainty as “an estimate attached to a test result which characterises the range of values within which the true value is asserted to lie” [9]. This definition refers to “the true value”, an unknowable quantity. Therefore, it is placed in the theoretical domain and has limited practical significance.

The latest edition of the international vocabulary of basic and general terms in metrology, VIM, avoids the term “true value” and states that uncertainty is “a parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand” [10].

Uncertainty reflects the lack of exact knowledge of the result of the measurement. This lack of exact knowledge is due to uncertainty about the correctness of the stated result that still remains, that is, doubts about how well the result of the measurement represents the value of the quantity being measured, when all the known and suspected components of error have been evaluated and the appropriate corrections have been applied. Therefore, uncertainty could be

understood as a parameter that characterises the range of values within which the value of the quantity being measured is expected to lie.

### **Accuracy, error and uncertainty**

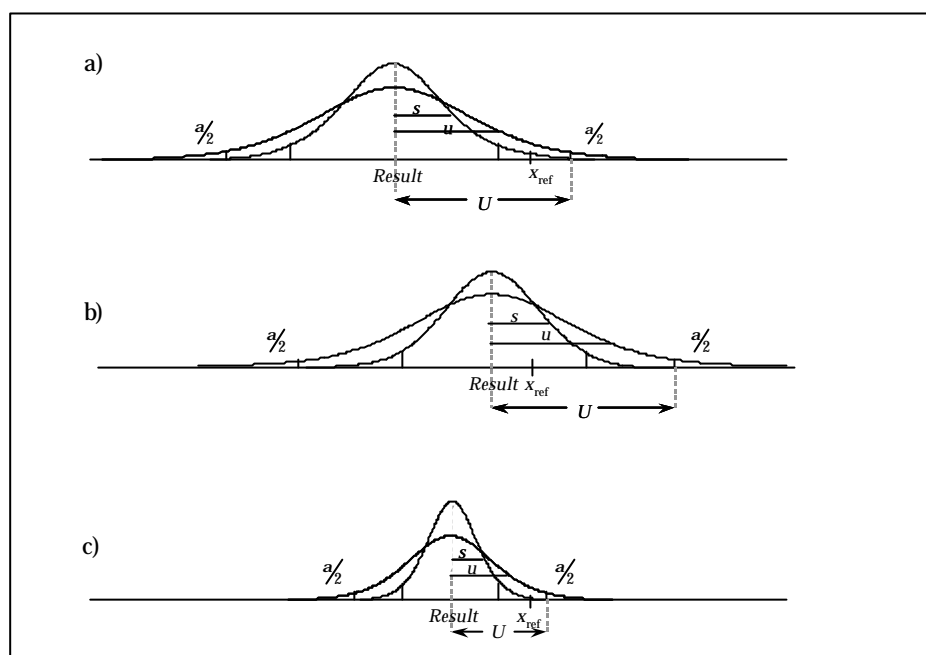
Accuracy is defined as the closeness of the agreement between the result of a measurement and the true value of the measurand [10] and it is assessed in practice by verifying trueness and quantifying precision. The difference between the result of a measurement minus the true value of a measurand is termed as “error” [10]. Error simply refers to a position and indicates the distance between these two values. To illustrate the difference between the concepts of error and uncertainty, the result of an analysis may be very close to the true value of the measurand and have a negligible error. Nevertheless, the uncertainty may be very large, simply because the analyst is unsure of the size of the error. Likewise, if the uncertainty is estimated for an analytical procedure and a well defined sample type, this uncertainty can be used for the same procedures and sample types in the future. These differences are illustrated in Figure 1. The uncertainty in Figure 1a and Figure 1b is the same because the same analytical procedure and sample type is used. However, the error committed in Figure 1a is much higher than the one in Figure 1b. In this Figure, the accepted reference value,  $x_{ref}$ , is used as an estimation of the true value of the measurand.

### **Uncertainty, trueness and traceability**

Trueness is the closeness of agreement between the average value of a large set of test results and an accepted reference value (i.e. a value that serves as an agreed-upon reference for comparison) [9]. The uncertainty concept must be interpreted if a link between uncertainty and trueness is to be established. If the trueness has not been assessed against an accepted reference value, the correctness of all the possible systematic effects of the measurements cannot be guaranteed, and it is therefore impossible to ensure that the confidence interval comprises the accepted reference value. So, if the quantitative analytical result is expressed as *Result* ±



Uncertainty, every analyst should verify that *Result* is an unbiased value and, whenever possible, a value which is traceable to stated references, usually national or international standards. In this case, a link between uncertainty and traceability is also established.



**Figure 1.** Intermediate precision (represented as a standard deviation,  $s$ ) and uncertainty (expressed as standard uncertainty,  $u$ ) associated to the *Result* of an analytical measurement.  $t_{\alpha/2} \cdot s$  and  $U$  (expanded uncertainty) are the confidence intervals associated to  $s$  and  $u$ , respectively.  $x_{\text{ref}}$  is the accepted reference value. See text for explanatory comments.

## Uncertainty and precision

Precision refers to the closeness of agreement among the results of a series of replicate measurements [10]. Depending on the conditions of the measurement (i.e. the factors varied in the measurement), it is evaluated in terms of repeatability, intermediate precision or reproducibility. Precision includes the variability of the factors which are randomly varied in the measurement. On the other hand, uncertainty does not only include all the random contributions that affect the

measurement but also the contributions derived from the estimation of systematic errors. Therefore, uncertainty must include a term of precision estimated by representatively varying all the factors that affect the measurement and another term arising from the assessment of accuracy itself because in this process a systematic error which is not statistically significant is estimated. As a result, uncertainty is larger than precision.

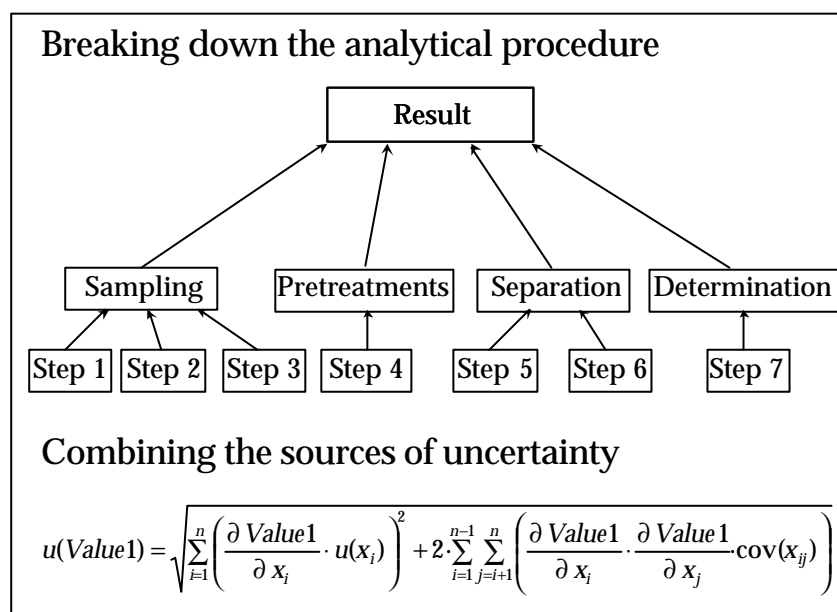
Moreover, since uncertainty includes components arising from the assessment of accuracy, the confidence interval  $Result \pm Uncertainty$  includes the accepted reference value whereas the confidence interval associated to precision in a laboratory, i.e.  $Result \pm t_{\alpha/2} \cdot s$  (where  $s$  is an intermediate standard deviation which considers all the random sources of variation within the laboratory), may not comprise the accepted reference value. Figure 1 shows these concepts. In this figure, a working sample has been analysed with two different analytical procedures under intermediate precision conditions so that all the factors that affect the measurement result (replicate, day, operator, etc.) are included. Since both procedures should be previously traced to the accepted reference value, the analytical results, expressed as  $Result$ , are traceable to the considered true value,  $x_{ref}$ . In figure 1a and 1b, the sample is analysed with the same analytical procedure, and in Figure 1c a more precise analytical procedure is used. In all three cases, the uncertainty associated to the measurement result,  $U$ , is larger than the confidence interval associated to the precision and comprises the accepted reference value,  $x_{ref}$ , because it includes all the sources of variation arising from random components (estimated with intermediate precision) and the component of uncertainty arising from the assessment of accuracy (i.e. from the estimation of a non significant systematic error). However, the confidence interval defined by intermediate precision comprises the accepted reference value only in Figure 1b. This is because precision does not include the term which arises from assessing accuracy. Figure 1b shows that the measurement result is closer to the accepted reference value than in 1a, but that both results have the same precision and uncertainty because they have been obtained with the same analytical method. Finally, Figure 1c shows that higher precision than in previous cases normally means lower uncertainty values.

## THE ISO APPROACH FOR CALCULATING UNCERTAINTY

This approach estimates the overall uncertainty by identifying, estimating and combining all the sources of uncertainty associated with the measurement. Because of the way the overall uncertainty is estimated, it has been defined as a “bottom-up” approach [2-5]. This approach involves four main steps:

### Specification step

In this step a relationship is established between the analytical result and the parameters on which the analytical result depends, such as  $Result = f(x_1, \dots, x_n)$ . Normally, the difficulties of formalising the algebraic relationship between analytical results and the parameters that intervene in the analytical procedure become apparent. In this case, the analytical procedure is divided into the different steps in which the measurement can be broken down (see Figure 2). After this, the *specification* is done for each of the steps.



**Figure 2.** Scheme of the ISO approach for calculating uncertainty.

### **Identification step**

This step identifies all the sources of uncertainty in the different steps shown in Figure 2.

### **Quantification step**

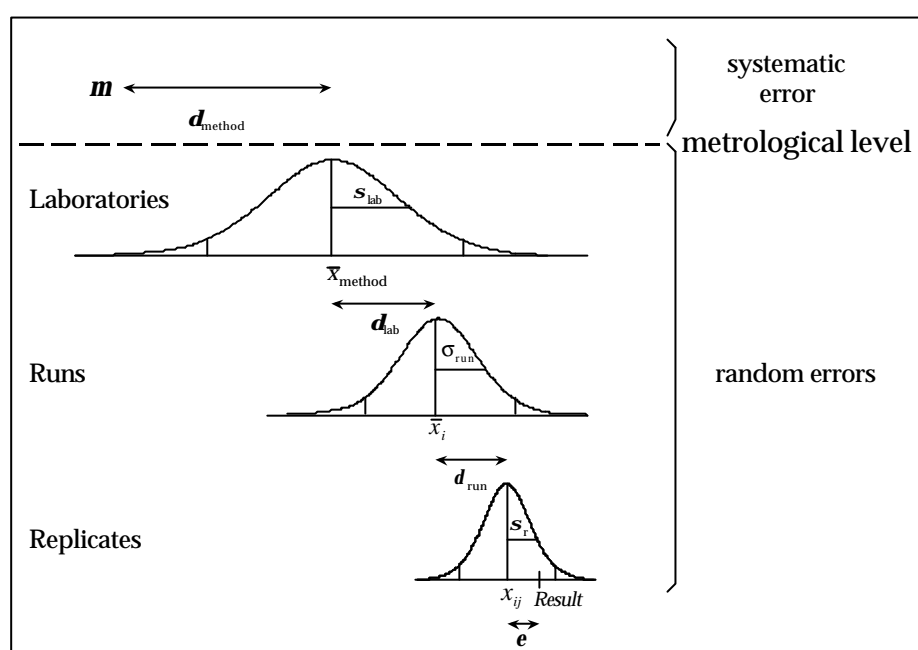
In this step, all the components of uncertainty that are expected to make a significant contribution to the overall uncertainty are quantified as *standard uncertainties* (i.e. the uncertainties are expressed as standard deviations). These components can be estimated from experimental work (“Type A” uncertainty) or by using previous information or the analyst’s experience (“Type B” uncertainty).

### **Combination step**

In this step, the *combined standard uncertainty*,  $u$ , is obtained by using the principle of error propagation to combine all the sources of uncertainty expressed as *standard uncertainties*. Calculating the expanded uncertainty then provides a confidence interval within which the accepted reference value is expected to lie. This confidence interval can be obtained by multiplying the combined standard uncertainty by a coverage factor,  $k$  [2,3] which assumes that the distribution of the measurement result, *Result*, is normal. However, when different types of distribution are combined, it is more correct to use the effective  $t$  tabulated value for the level of confidence chosen and the effective degrees of freedom,  $\nu_{\text{eff}}$ , calculated using the Welch-Satterthwaite approach [11,12]. If the effective degrees of freedom are large enough, the  $t$ -distribution can be approximated to a normal distribution. In these circumstances, the coverage factors can be adopted. In practice,  $k=2$  is recommended for most purposes.

## THE ANALYTICAL METHODS COMMITTEE APPROACH FOR CALCULATING UNCERTAINTY

The Analytical Methods Committee approach (commonly known as “top-down”) uses information generated by different laboratories participating in method performance exercises to quantify uncertainty [6]. The laboratory is seen from a “higher level”, i.e. as a member of a population of laboratories. As a result, systematic and random errors that occur within individual laboratories become random errors when they are considered from this “higher level” (see Fig. 3).



**Figure 3.** Random and systematic errors of an individual result, *Result*, when the laboratory is seen as a member of a population of laboratories. The metrological level is the level of traceability assessed when the analytical result is obtained as the consensus value of the participating laboratories in the collaborative trial.  $\mu$  represents the accepted reference value.

The analytical result of a homogeneous sample is expressed as the sum of random and systematic errors:

$$Result = m + d_{method} + d_{lab} + d_{run} + e \quad (1)$$

where *Result* is the analytical result obtained,  $\mu$  is the true value,  $\delta_{\text{method}}$  is the method bias,  $\delta_{\text{lab}}$  is the laboratory bias,  $\delta_{\text{run}}$  is the bias due to the conditions in which the measurement is made (day, analyst, etc.) and  $\varepsilon$  is the random error.

As can be seen in Figure 3, the laboratory and the run bias are random variables with an associated variance whereas the method bias is seen as a systematic error. Since uncertainty must include all the factors that can affect the analytical results (i.e. sources of variation arising from systematic and random errors), the uncertainty of a future measurement is expressed as:

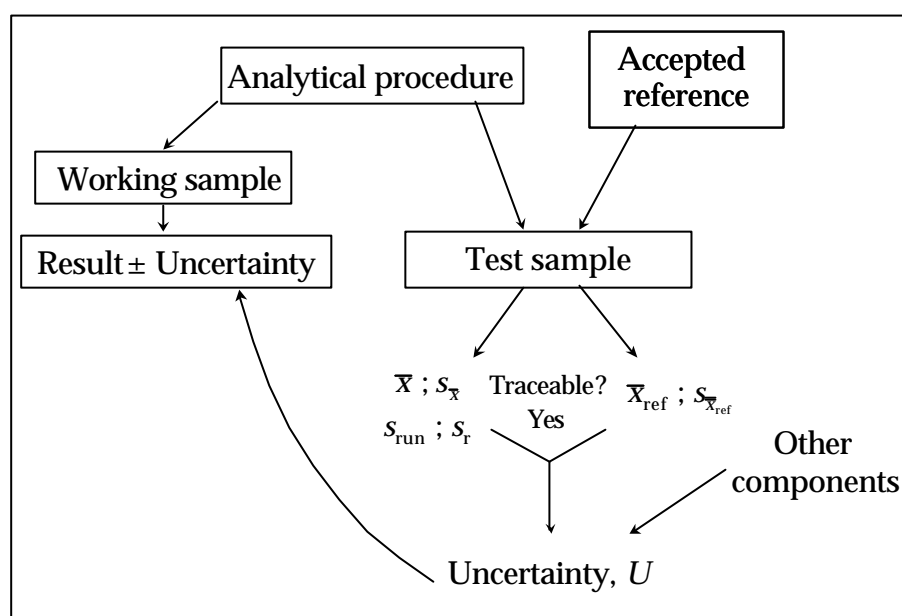
$$U(\text{Result}) = t_{\alpha/2, \text{eff}} \sqrt{\mathbf{s}_e^2 + \mathbf{s}_{\text{run}}^2 + \mathbf{s}_{\text{lab}}^2 + u_{\text{method}}^2} \quad (2)$$

where  $\sigma_e^2$ ,  $\sigma_{\text{run}}^2$  and  $\sigma_{\text{lab}}^2$  are the true variances associated to the different values that the random error, the run bias and the laboratory bias can take respectively. All these variances can be estimated from the results of a collaborative trial, whereas the uncertainty related to the estimation of the method bias,  $u_{\text{method}}$ , can only be estimated if there is an accepted reference value available.  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value, at a given  $\alpha$  probability and  $\nu_{\text{eff}}$  degrees of freedom. The effective degrees of freedom,  $\nu_{\text{eff}}$ , are estimated using the Welch-Satterthwaite approach [11]. The original article in which this approach was proposed [6] uses the coverage factor,  $k$ , instead of the effective  $t$  tabulated value. However, it is more correct to use  $t_{\alpha/2, \text{eff}}$  whenever different  $t$ -distributions are combined [12].

This expression does not include other effects, such as the uncertainty arising from sampling, pre-processing steps to homogenise the sample or differences between the matrix composition of the sample analysed in the interlaboratory exercise and the samples analysed in routine measurements.

## ESTIMATING UNCERTAINTY USING INFORMATION FROM THE ASSESSMENT OF ACCURACY

This approach estimates the uncertainty of future measurement results using the information generated when assessing the accuracy of an analytical procedure in a laboratory [7,8]. A similar procedure for estimating uncertainty has also been proposed in [13]. In this process, information is obtained about the different within-laboratory precision estimates of the analytical procedure that will subsequently be applied to routine samples. As can be seen in figure 4, components of uncertainty that are not considered in this validation process are subsequently included to obtain the overall uncertainty.



**Figure 4.** Assessment of accuracy of an analytical procedure within a laboratory. The information obtained is used as a component of uncertainty of the future working samples.

This approach does not strictly estimate the uncertainty of a future routine sample because the information generated is related to a test sample analysed in the past. However, if the test sample is stable and representative (i.e. similar to the routine

samples), the estimation can be considered to be good whenever the suitable quality assurance procedures are effectively implemented in the laboratory.

### Assessment of accuracy

The accuracy of an analytical procedure must be assessed before it is applied to future working samples. The assessment of accuracy involves checking the trueness (i.e. the lack of systematic errors) of the analytical procedure. In this process, references must be used to make a statistical comparison of the analytical procedure and the reference result. Trueness is checked by estimating the difference between the average result obtained with a test sample (analysed by varying representatively all the sources of variation of the analytical procedure) and the accepted reference value assigned to this test sample. Fully-nested designs can be used to vary representatively all the sources of variation (day, analyst, replicate, etc.) [7,14]. The grand mean calculated from the analytical results can be seen as:

$$\bar{x} = \bar{x}_{\text{ref}} + \hat{\boldsymbol{d}}_{\text{assessment}} \quad (3)$$

where  $\bar{x}$  is the grand mean of the set of results obtained with the analytical procedure,  $\bar{x}_{\text{ref}}$  is the accepted reference value and  $\hat{\boldsymbol{d}}_{\text{assessment}}$  is the bias estimated in the assessment of accuracy. If the bias estimated is not statistically significant, the trueness of the analytical procedure is assessed.

Different types of references can be used in the assessment of accuracy. However, not all the references are of the same level in the traceability hierarchy [7]. The higher the reference is in the traceability hierarchy, the better it is, because the absence of more possible systematic errors can be checked as a whole. Therefore, of the references available, the one used should be that which is on the highest level in the traceability hierarchy. The following references can be used to assess accuracy:



**a) Primary methods.** Primary methods are on the highest level of the traceability hierarchy [15]. The results they provide are directly connected to the SI units. These methods can be completely described and understood. Their uncertainty can be written down in terms of the SI units whenever the laboratory is under quality assurance conditions. As a result, the measurement results provided by a primary method are accepted without having to be referenced to any standard. However, these methods can seldom be applied because the *Comité Consultatif pour la Quantité de Matière* (CCQM) identifies very few analytical methods as having the potential to be primary methods. These methods are: isotope dilution mass spectrometry, coulometry, gravimetry, titrimetry and a group of colligative methods (i.e. freezing point depression).

The analytical result of a test sample obtained with a primary method is compared to the mean value obtained by analysing the same sample with the analytical procedure when all the sources of variation are varied representatively. If the difference is not statistically significant, the analytical procedure is traceable to the SI units through the primary method.

**b) Reference materials.** Reference materials can be on different levels in the traceability hierarchy depending on how the reference has been obtained. A hierarchy for reference materials has been proposed in [16]. Primary reference materials (PRM) are placed on the highest level of the reference materials traceability hierarchy. They are certified by primary methods carried out by a national metrology laboratory. Therefore, they are traced back to the SI units. Certified reference materials (CRM) and working reference materials (WRM) are placed on a lower level because the accepted reference values are obtained as the mean of measurement results not obtained with primary methods. Therefore, they cannot be traced to the SI units.

Reference materials are in short supply because a considerable number of analytes need to be analysed in a diversity of matrices. Therefore, only 5-10% of analytical procedures can be traced to a reference material.

**c) Reference methods.** A reference method can be used to obtain the accepted reference value of a test sample. This method can be performed within a laboratory (“in-laboratory” validation) [17] or by different laboratories (“interlaboratory” validation) [14]. In both cases the test sample is analysed with the reference method and the analytical procedure to be traced by representatively varying all the sources of variation.

“In-laboratory” validation provides information about the intermediate precision of both methods. The assessment of trueness only checks the absence of a method bias within the laboratory. It does not, however, check the absence of a possible laboratory bias because the measurement results are only obtained in one laboratory. “Inter-laboratory” validation provides information about the reproducibility of the methods. Moreover, because the results are provided by several laboratories, the bias of both the laboratory and the method is checked. Therefore, the traceability level assessed in “inter-laboratory” validation is higher than the level assessed in the “in-laboratory” validation.

**d) Interlaboratory trials.** A laboratory can trace an analytical procedure by participating in an interlaboratory trial. A consensus value with a given uncertainty is obtained from the trial data after the outlying laboratories have been rejected.

If the laboratory is not detected as an outlier, the mean result of the individual laboratory should be statistically compared to the consensus value. This process checks the absence of a laboratory bias. Moreover, if the participating laboratories have used different analytical procedures, the absence of the method bias is also checked. The level of traceability assessed is lower than the one assessed using a reference method because the consensus value is obtained with analytical procedures which may not have been previously validated. In this case, the traceability is only related to the consensus value obtained by the participating laboratories.

**e) Spiked samples.** Accuracy is frequently assessed by adding to the test sample a known amount of a compound, the chemical behaviour of which is supposed to be

the same as the analyte of the working samples to be analysed. The compounds added can be (a) the analyte itself, (b) an isotopically-modified version of the analyte, or (c) an internal standard, i.e. a different analyte whose chemical properties are closely related to the chemical properties of the analyte to be analysed.

Spiked samples have a much lower level of traceability than the references discussed above because the analyte added may not be totally representative of the behaviour of the analyte present in the sample (i.e. the native analyte). For instance, the analyte added may not be so strongly bound to the matrix as the native analyte. If this is the case, the bias estimated when assessing accuracy with the added analyte will probably be lower than the true bias of the native analyte. Therefore, the accuracy can only be guaranteed for the added analyte but not for the native analyte. As a result, some systematic errors may have been overlooked because the analyte added may poorly represent the native analyte.

The best compound to be added is an isotope of the native analyte because its chemical properties are the same and the initial amount of the native analyte need not be known to assess accuracy. Constant and proportional systematic errors can be detected by statistically comparing the amount of the isotope of the analyte found to the amount added.

If the same analyte is added, the concentration of the native analyte in the sample needs to be known whenever a blank of the sample is not available. In this case the accuracy is assessed by statistically comparing the analyte added with the difference between the analyte found in the spiked sample and the analyte found in the unspiked sample. If the analyte is spiked at different concentrations, the accuracy can be assessed using regression techniques [18]. The level of traceability when the same analyte is added is not as high as when an isotope is added because only proportional systematic errors can be detected. Constant systematic errors are not detected whenever a blank of the sample is not available.

Internal standards have the lowest level of traceability because their chemical properties are not the same as the ones of the analyte to be analysed. Therefore, they should only be used when the pure native analyte is not available.

### Calculation of uncertainty

As can be seen in Figure 5, a future result can be expressed as the true value plus the sum of random errors and systematic errors:

$$Result = \mu + d_{\text{other terms}} + d_{\text{assessment}} + d_{\text{preproc.steps}} + d_{\text{run}} + \varepsilon \quad (4)$$

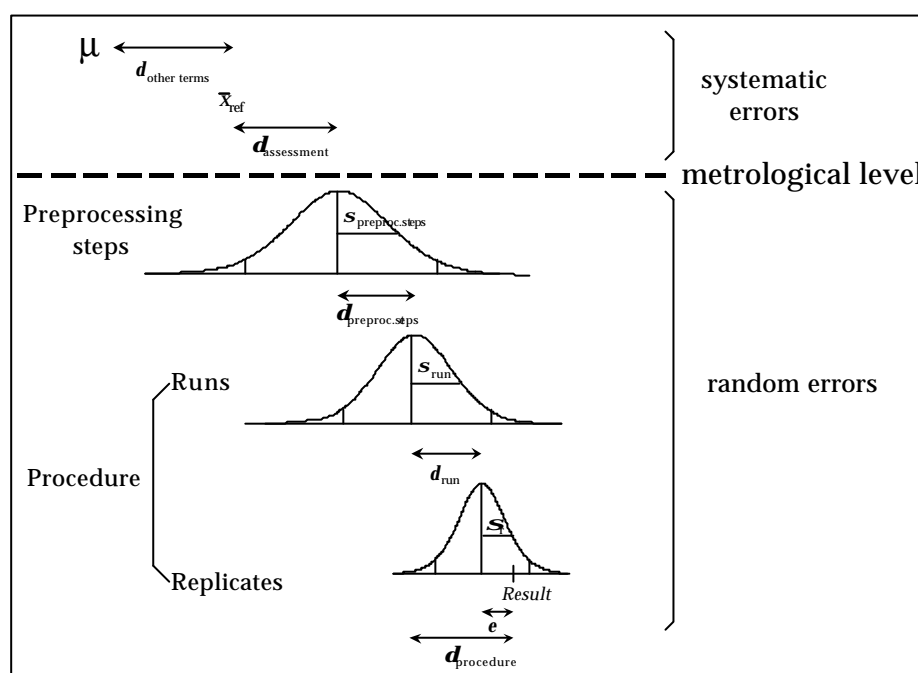
where *Result* is the analytical result obtained,  $\mu$  is the true value of the result,  $\varepsilon$  is the random error,  $\delta_{\text{run}}$  is the bias due to the conditions in which the measurement was carried out (day, analyst, instrument, etc.),  $\delta_{\text{preproc.steps}}$  is the bias due to preprocessing steps carried out in the working sample but not in the test sample used in the assessment of accuracy,  $\delta_{\text{assessment}}$  is the bias estimated in the assessment of accuracy and  $\delta_{\text{other terms}}$  is the bias associated to higher levels of traceability or to other biases not previously considered.

The sources of uncertainty of the measurement result depend on  $\delta_{\text{other terms}}$ ,  $\delta_{\text{preproc.steps}}$ ,  $\delta_{\text{assessment}}$ ,  $\delta_{\text{run}}$  and  $\varepsilon$ . Therefore, the overall uncertainty is expressed as:

$$U(Result) = t_{\alpha/2, \text{eff}} \cdot \sqrt{u_{\text{proc}}^2 + u_{\text{assessment}}^2 + u_{\text{preproc.steps}}^2 + u_{\text{other terms}}^2} \quad (5)$$

where  $u_{\text{proc}}$  is the uncertainty associated to the analytical procedure. This uncertainty arises from the experimental variation when the analytical procedure is applied to obtain future measurement results and is associated to the different values that the run bias and the random error can take within the laboratory.  $u_{\text{assessment}}$  is the uncertainty associated to the estimation of  $\delta_{\text{assessment}}$ . The bias estimated depends on the level of traceability of the reference used. The higher the reference is in the traceability hierarchy, the more systematic errors are considered in assessing accuracy. Therefore, fewer terms then need to be added to the

uncertainty budget for all the possible systematic errors of the analytical procedure to be considered.  $u_{\text{other terms}}$  is associated to the estimation of  $\delta_{\text{other terms}}$ . This includes all the sources of uncertainty that have not been taken into account in the other terms of uncertainty (e.g. the uncertainty associated to the estimation of bias related to higher levels of traceability).  $u_{\text{preproc.steps}}$  is associated to the estimation of  $\delta_{\text{preproc.steps}}$ . This term includes the uncertainty arising from the subsampling and/or from the preprocessing steps not carried out in the assessment of accuracy. It can be estimated with a sample which has the same characteristics as the routine samples which are to be subsequently analysed.  $t_{\alpha/2, \text{eff}}$  is the two-sided  $t$  tabulated value, at a given  $\alpha$  probability and  $\nu_{\text{eff}}$  degrees of freedom.



**Figure 5.** Random and systematic errors of an individual result, *Result*, obtained by an individual laboratory. The metrological level (i.e. the level of traceability) assessed when the analytical result is obtained as the mean of results obtained within a laboratory is lower than the one obtained in the collaborative trial.

As can be seen in Figure 5, the bias related to other terms, such as higher levels of traceability, and the bias related to the assessment of accuracy,  $\delta_{\text{assessment}}$ , are

considered to be systematic errors whereas the run bias and the random error are considered to be random variables.

From the information generated when assessing accuracy, we can estimate the uncertainty related to the analytical procedure (i.e. due to the application of the analytical method to future routine samples) and the uncertainty due to the bias estimated in the assessment of accuracy,  $\delta_{\text{assessment}}$ .

### **Using information from the validation process to implement the quality procedures**

The quality control procedures should be effectively implemented in the laboratory in order to ensure that the interval expressed as *Result*  $\pm$  *Uncertainty* comprises the accepted reference value with a given probability. This can usually be ensured by checking that the analytical procedure is under statistical control. Therefore, trueness and the stability of precision of the analytical procedure are commonly checked using control charts [14, 18-20]. The information used to build the initial control charts can be obtained from the grand mean and precision estimates of the analytical procedure estimated in assessing the accuracy.

When carrying out the quality control procedures systematically in the laboratory, more information is obtained about the analytical procedure. This information can be used to recalculate the terms of uncertainty which arise from the assessment of accuracy and the procedure. Since the grand mean and the precision are estimated better, the uncertainty arising from the assessment of accuracy is expected to diminish. Moreover, the *t* effective tabulated value diminishes and tends to the *z*-value because the degrees of freedom increase.

## **CRITICAL DISCUSSION OF THE DIFFERENT APPROACHES FOR CALCULATING UNCERTAINTY**

The ISO approach, which was originally developed to estimate the uncertainty in physical measurements [2], was adapted to chemical measurement processes by EURACHEM [3]. The application of this approach to chemical measurements increases our knowledge of the measurement process because all the possible sources of error and their magnitude must be known to estimate the uncertainty. This greater knowledge of the measurement process means that the measurement uncertainty can be reduced by optimising those steps that significantly contribute to the overall uncertainty. The whole ISO procedure may initially involve a lot of work, because all the sources of uncertainty arising from the different steps of the measurement process must be identified and quantified. However, some of the uncertainties of the general steps in the measurement process, such as weighting, dilution steps, etc., can eventually be used for different analytical procedures.

Nevertheless, there are considerable problems involved in adapting this methodology to chemical measurements [21,22]. On the one hand, the considerable number of steps and sources of error in chemical measurements make it difficult to correctly estimate the overall uncertainty by budget analysis. The uncertainty may be underestimated because of the difficulties of knowing all the sources of error that affect the measurement process. Moreover, estimating all the sources of uncertainty in a chemical measurement can be extremely complex and time consuming. Furthermore, on most occasions it is not possible to experimentally estimate all the sources of uncertainty and the analyst is impelled to use "Type B" uncertainties. Although this type of uncertainty is based on prior data or on the analyst's knowledge of the measurement process, "Type B" uncertainty estimations may not be correct because understanding the whole chemical measurement process is not straightforward. Therefore, the disadvantages of this approach often prevent it from being applied to chemical measurements.

To overcome these limitations, EURACHEM attempts to simplify the originally ISO approach proposed for physical measurements by (1) not estimating negligible effects and (2) where possible grouping effects. For instance, Ellison [5] recently

proposed a methodology, which is based on grouping some of the effects, to estimate the overall uncertainty. This methodology is totally consistent with the ISO approach in the sense that all the factors that affect the measurement result are identified and structured by using *cause and effect analysis*. The effects that have been taken into account by existing data (e.g. interlaboratory data, recovery data, etc.) or by experiments are then identified in the *reconciliation step*. In this way, overestimating the overall uncertainty by considering effects more than once is avoided.

The “top-down” approach and the one that uses information from validation data are totally consistent with the ISO recommendations for calculating uncertainty. Unlike the “bottom-up” approach, these approaches are more holistic because they estimate uncertainty by taking the measurement procedure as a whole. Therefore, it is not necessary to identify and estimate all the individual sources of error in an analytical procedure to estimate the overall uncertainty. Moreover, they take into account sources of uncertainty that may have been overlooked when estimating the uncertainty with the “bottom-up” approach due to difficulties of fully understanding the analytical measurement process.

However, these two approaches also have some drawbacks. The “top-down” approach can only be applied if interlaboratory information is available. Therefore, it often cannot be applied because a considerable number of measurands do not have interlaboratory exercises for many matrices. Moreover, the collaborative trials are usually concentration/matrix specific. Therefore, the uncertainty estimated is not useful when the laboratory has to estimate the uncertainty attached to samples which can have different matrices (e.g. oils, wines, fats, etc.) but it is useful for estimating uncertainty for different concentrations because the reproducibility of the method can be modelled against concentration with the Horwitz function [23]. If this approach is to be used, other components of uncertainty arising from subsampling, preprocessing steps, differences in the matrices, etc. must be included in the overall uncertainty. However, the additional work is much less than that required to estimate uncertainty with the “bottom-up” approach.



The uncertainty estimated with the “top-down” approach may have little to do with the uncertainty of the measurement results of a given laboratory. First of all, the uncertainty of an individual laboratory may be estimated incorrectly because of differences between the within-laboratory precision of the individual laboratory and the average value of the within-laboratory precision estimated in the collaborative trial. Secondly, the uncertainty estimated with the “top-down” approach may be highly overestimated whenever some of the participating laboratories in the collaborative trial have a significant laboratory bias without having been detected as outliers. In this case, the uncertainty includes a significant laboratory bias which means that the uncertainty attached to the results of an individual traceable laboratory will be overestimated. Therefore, to prevent the possibility of the uncertainty being overestimated, laboratories which are not traceable to the consensus value should not take part in the final calculations derived from the collaborative trial. By not taking part, the final uncertainty with this approach is that of an individual laboratory whenever the within-laboratory precision estimated in the collaborative trial is the same as that of the individual laboratory.

The uncertainties estimated with the “top-down” approach and the ISO approach were compared in [24] for different analyses of food. The results show that both approaches give similar precision estimates. However, the two approaches could not usually be compared because the “bottom-up” approach considered the uncertainty arising from the variability of the matrix whereas the interlaboratory studies did not. In these cases, the “top-down” approach underestimated the uncertainty.

Some of the limitations of the “top-down” approach can be overcome by the approach that uses information from the validation process. The uncertainty estimated with this approach is totally related with the individual laboratory whenever all the sources of variability are varied representatively when assessing accuracy. Since the accuracy (trueness+precision) of an analytical procedure must always be assessed before analysing future routine samples, the information for estimating uncertainty is also always available. Likewise, the effort required to estimate sources of uncertainty not considered in the validation process

(preprocessing steps, subsampling, etc) is considerably less than the effort needed to estimate uncertainty with the ISO approach. Therefore, the main benefits of this approach are its conceptual and practical simplicity, its lower cost and its universal application.

This approach, however, also has some drawbacks, most of which are due to the limited number of references available with a high level in the traceability hierarchy. As a consequence, analysts are often forced to use references at a low level in the traceability hierarchy. When they do, more terms must be added to the uncertainty budget to include uncertainties arising from systematic effects which are not taken into account when the reference used is low in the traceability hierarchy. Moreover, the factors varied in the assessment of accuracy may not include some of the sources of variation in the measurement method. Therefore, depending on the number of factors chosen, other terms of uncertainty not previously considered in the assessment of accuracy must be subsequently included in the uncertainty budget.

The results obtained with the “top-down” approach and the approach based on using information from the validation process have also been compared [25] using the interlaboratory data of vanadium content in steel [26]. The uncertainties calculated with both approaches show that they give similar results whenever almost all the laboratories participating in the interlaboratory trial are traceable to the consensus value. The “top-down” uncertainty was overestimated when the proportion of traceable participating laboratories decreased.

## **ACKNOWLEDGMENTS**

The authors would like to thank the DGICYT (project num. BP96-1008) for financial support and the CIRIT from the Catalan Government for providing Alicia Maroto’s doctoral fellowship.

---

**REFERENCES**

- [1] CEN/CENELEC, European Standard EN 45001, *General Criteria for the operation of testing laboratories*, Brussels, Belgium (1989)
- [2] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *Guide to the expression of uncertainty in measurement*, ISO, Geneva (1993)
- [3] EURACHEM, *Quantifying uncertainty in analytical measurements*, EURACHEM Secretariat, P.O. Box 46, Teddington, Middlesex, TW11 0LY. UK (1995)
- [4] S. Ellison, W. Wegscheider, A. Williams, *Anal. Chem.*, **1** (1997) 607 A - 613 A
- [5] S. Ellison, V.J. Barwick, *Analyst*, **123** (1998) 1387-1392
- [6] Analytical Methods Committee, *Analyst*, **120** (1995) 2303-2308
- [7] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Analytica Chimica Acta*, **391** (1999) 173-185
- [8] A. Maroto, J. Riu, R. Boqué, F.X. Rius, *Trends in Analytical Chemistry*, **18/9-10** (1999) 577-584
- [9] International Organization for Standardization, ISO 3534-1, *Statistics - Vocabulary and symbols*, ISO, Geneva, (1993)
- [10] BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *International Vocabulary of Basic and General Terms in Metrology*, VIM, ISO, Geneva (1993)
- [11] F.E. Satterthwaite, *Psychometrika*, **6** (1941) 309-316
- [12] C.F. Dietrich, *Uncertainty, Calibration and Probability. The statistic of Scientific and Industrial Measurement*, 2<sup>nd</sup> Ed. Adam Hilger, Bristol (1991)
- [13] W. Hässelbarth, *Accreditation and Quality Assurance*, **3** (1998) 418-422
- [14] International Organization for Standardization, ISO 5725, *Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva (1994)
- [15] T.J. Quinn, *Metrologia*, **34** (1997) 61-65
- [16] X.R. Pan, *Metrologia*, **34** (1997) 35-39
- [17] D.L. Massart, J. Smeyers-Verbeke, *Robust process analytical methods for industrial practice and their validation*, Report Project SMT4-CT95-2031 (1996)
- [18] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, (1997)
- [19] International Union of Pure and Applied Chemistry, *Pure & Appl. Chem.*, **67** (1995) 649-666
- [20] Analytical Methods Committee, *Analyst*, **120** (1995) 29-34
- [21] W. Horwitz, *Journal of AOAC International*, **81**, 4 (1998) 785-794
- [22] W. Horwitz, R. Albert, *Analyst*, **122**, (1997) 615-617
- [23] W. Horwitz, *Anal. Chem.*, **54** (1982) 67 A - 76 A
- [24] P. Brereton, S. Anderson, P. Willets, S. Ellison, V. Barwick, M. Thompson, *An investigation into the application of measurement uncertainty in food analysis*, Report FD 96/103, (1997)
- [25] A. Maroto, R. Boqué, J. Riu, F.X. Rius, Internal Report (1999)

[26] ISO/TC 17, Steel/SC1, *Methods of determination of chemical composition*, (1985)

---

**GLOSARIO**

AMC	<i>Analytical Methods Committee</i>
ANOVA	Análisis de la varianza <i>Analysis of variance</i>
ANCOVA	Análisis de la covarianza <i>Analysis of covariance</i>
AOAC	<i>Association of Analytical Communities</i>
BCR	<i>Community Bureau of Reference</i>
BLS	Mínimos cuadrados bivariantes <i>Bivariate least squares</i>
CRM	Material de referencia certificado <i>Certified reference material</i>
EN	<i>European Norm</i>
GUM	<i>Guide to the expression of uncertainty in measurement</i>
IEC	<i>International Electrotechnical Commission</i>
ISO	<i>International Organisation for Standardization</i>
IUPAC	<i>International Union of Pure and Applied Chemistry</i>
LGC	<i>Laboratory of the Government Chemist</i>
MB	Blanco del método <i>Method Blank</i>
NMKVL	<i>Nordic Committee on Food Analysis</i>
OLS	Mínimos cuadrados <i>Ordinary least squares</i>
PRM	Material de referencia primario <i>Primary reference material</i>
SPME-GC	Microextracción en fase sólida y cromatografía de gases <i>Solid Phase Microextraction and Gas Chromatography</i>
SAM	Método de adiciones estándar <i>Standard Addition Method</i>
SB	Blanco instrumental <i>System Blank</i>

TYB	Blanco Total de Youden <i>Total Youden Blank</i>
UNE	Una norma española
VIM	<i>International Vocabulary of basic and general terms in Metrology</i>
WRM	Material de referencia de trabajo <i>Working reference material</i>
WLS	Mínimos cuadrados ponderados <i>Weighted least squares</i>



