

# Validação dos Métodos de Calibração nos Laboratórios Nacionais de Metrologia<sup>a</sup>

EDUARDA FILIPE

[E-mail: efilipe@mail.ipq.pt](mailto:efilipe@mail.ipq.pt)

Instituto Português da Qualidade, Rua António Gião, 2, 2829-513 Caparica, Portugal

## Sumário

Os Laboratórios Nacionais de Metrologia (LNMs), topo do sistema nacional de medição mediante calibração asseguram a rastreabilidade ao Sistema Internacional de Unidades (SI). Para cumprir os requisitos da norma NPEN ISO/IEC 17025 de 2000, nomeadamente o ponto 5.4.5, estes laboratórios devem validar os métodos não normalizados e os métodos criados/desenvolvidos pelo laboratório para a disseminação das unidades. Como consequência os LNMs devem descrever na sua documentação da Qualidade o modo como efectuam a validação dos seus métodos.

É descrito um procedimento de validação para os métodos de calibração e apresentado um exemplo de validação dos resultados de uma comparação de duas células de pontos fixos termométricos, aplicando a análise da variância aos resultados obtidos, num planeamento de experiências hierarquizado que permite inferir estatisticamente sobre os resultados que pertencem à mesma população estatística e calcular as componentes de incerteza-padrão avaliadas pelo método "Tipo A".

## Introdução

O Acordo de Reconhecimento Mútuo (MRA) [1] foi assinado em 14 de Outubro de 1999 pelos países signatários da Convenção do Metro com o objectivo de: "estabelecer o grau de equivalência dos padrões metrológicos nacionais mantidos pelos LNMs; providenciar o reconhecimento mútuo dos certificados de calibração e medição emitidos pelos LNMs e fornecer aos governos e outras entidades uma base tecnicamente sólida, para outros acordos mais abrangentes, relacionados com o comércio internacional e as actividades de regulamentação". Para atingir estes objectivos, os LNMs devem participar em Comparações internacionais de padrões designadas por Comparações-chave e Comparações suplementares<sup>b</sup> e implementar até 2004, um Sistema da Qualidade e de demonstração de competência. As capacidades de medição e de calibração dos LNMs reconhecidas constarão de uma base de dados gerida pelo BIPM que estará finalizada até ao final do corrente ano. Os países do chamado "mundo desenvolvido" possuem uma infra-estrutura metrológica suportada por padrões primários rastreados directamente às unidades SI [2] ou por padrões secundários rastreados a outros laboratórios nacionais. Um LNM tem como objectivos: desenvolver, manter e disseminar os padrões nacionais de medida apropriados às necessidades nacionais; a calibração dos padrões de referência dos laboratórios acreditados e de outras entidades e a organização da cadeia hierarquizada dos padrões de referência das entidades metrológicas do país.

O grau de equivalência obtido a partir dos resultados das comparações internacionais de padrões conferirá reconhecimento à Rastreabilidade Nacional assegurada pelo LNM. Para que ao

---

<sup>a</sup> Comunicação a apresentar no 1º Encontro da Sociedade Portuguesa de Metrologia, SPMet "Validação de Métodos e de Software" no IPQ, Caparica, 30 de Novembro de 2004.

<sup>b</sup> Classificação das comparações internacionais de padrões; esta classificação é definida pelos *Comités* Consultivos do CIPM (*Comité* Internacional dos Pesos e Medidas)

Laboratório seja conferida esta rastreabilidade é essencial também que o LNM demonstre a sua competência através da implementação de um Sistema da Qualidade.

A EUROMET<sup>c</sup> que constitui, no âmbito do MRA, a Organização Regional de Metrologia Europeia (RMO), escolheu a norma NPEN ISO/IEC 17025 de 2000 [3] – Requisitos Gerais de Competência para Laboratórios de Ensaio e Calibração, como a norma a utilizar na demonstração de competência exigida pelo MRA, que, apesar de difícil aplicabilidade nas actividades de investigação é aplicável às actividades de disseminação, por calibração ou por medição, das unidades.

Para cumprir os requisitos desta norma os LNMs devem: definir os objectivos da qualidade; garantir que os ensaios, a certificação dos materiais de referência e as calibrações são executados segundo os procedimentos técnicos, os requisitos dos clientes e com o uso das boas práticas laboratoriais; executar a validação dos métodos e dos resultados; assegurar, pela revisão pela Direcção do Laboratório, a realização dos objectivos do plano de actividades; avaliar a performance dos processos a partir nomeadamente dos relatórios de auditorias, dos inquéritos de satisfação de clientes; assegurar que a informação decorrente das reclamações, das acções correctivas e preventivas e dos trabalhos dos grupos de melhoria são implementados.

Neste artigo, será efectuada uma breve abordagem ao conceito de validação e ao cumprimento do ponto 5.4.5 da norma. É descrito um procedimento de validação para os métodos de calibração e apresentado um exemplo de validação dos resultados de uma comparação de duas células de pontos fixos termométricos, aplicando o Planeamento de Experiências, técnica estatística que permite fazer inferências sobre os resultados, nomeadamente de comparação de métodos e de medições em condições de repetibilidade/reproducibilidade.

## 1. Conceitos e Princípios Gerais

A norma NP EN ISO 9000: 2000 [4] define "validação" como "confirmação, através de evidência objectiva, de que foram satisfeitos os requisitos para uma utilização ou aplicação específicas<sup>d</sup>". A norma NP EN ISO IEC 17025, no ponto 5.4.5, refere que devem ser validados os métodos internos, os métodos não-normalizados ou normalizados usados fora do seu âmbito e suas extensões ou modificações. Na validação do método será ainda considerada a sua aptidão para satisfazer os requisitos definidos e demonstrada a sua capacidade técnica.

Este requisito não constava da norma anterior (NP EN 45001) pelo que o modo de abordar o processo de validação está a ser avaliado pelos laboratórios. No caso dos métodos normalizados os laboratórios devem considerá-los validados na sua gama de utilização e aptos para serem utilizados [5].

No processo de validação é feita a avaliação do método em termos da sua representatividade, repetibilidade<sup>e</sup> e reprodutibilidade<sup>f</sup> [6]. A validação dos métodos de ensaio e calibração abrange também de um modo geral a sua incerteza.

---

<sup>c</sup> EUROMET – Colaboração Europeia nos Padrões de Medição.

<sup>d</sup> A ISO 9000:2000 inclui na definição as seguintes notas: Nota 1: O termo "validado" é utilizado para designar o estado correspondente e Nota 2: As condições de utilização podem ser reais ou simuladas.

<sup>e</sup> [VIM 3.6] - Repetibilidade dos resultados (de uma medição): Aproximação entre os resultados de medições sucessivas da mesma mensuranda efectuadas nas mesmas condições de medição. NOTAS: 1. Estas condições são chamadas condições de repetibilidade. 2. As condições de repetibilidade incluem:- mesmo procedimento de medição;- mesmo observador;- mesmo instrumento de medição, usado nas mesmas condições;- mesmo local;- repetição num curto intervalo de tempo. 3. A repetibilidade pode exprimir-se quantitativamente em termos das características da dispersão dos resultados.

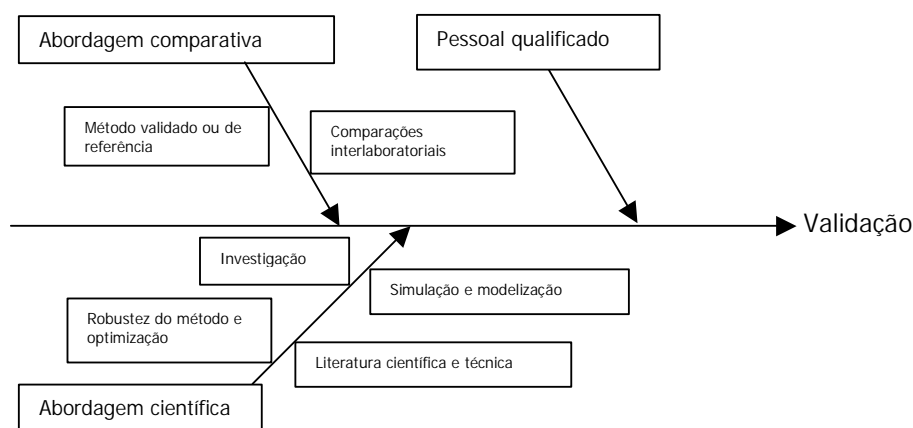
Para desenvolver um método “representativo”, é necessário conhecimento e prática de utilização dos resultados. Para todos os casos, devem ser avaliadas todas as grandezas de influência relevantes que podem afectar os resultados das calibrações e ensaios quer sejam de origem instrumental, ambiental, técnica ou humana.

## 2. Procedimentos para a Validação de Métodos

O modo como é efectuado o processo de validação depende essencialmente se o método a utilizar é interno: novo ou utilizado por outros laboratórios, se o método é de referência e em todos os casos onde foram efectuadas modificações desses métodos. Os métodos validados devem também corresponder às necessidades expressas pelo cliente (gama, incerteza etc.).

Os procedimentos técnicos de calibração e ensaio devem incluir a descrição da validação do respectivo método e também devem descrever a validação dos resultados obtidos nas calibrações e ensaios. Os estudos a realizar para a validação e a referir nos procedimentos técnicos, são de seguida descritos e classificados pela sua abordagem. Esta classificação é referida no Relatório da EUROLAB [5] por “abordagem científica ou comparativa” (ver Fig. 1).

Na primeira abordagem, abordagem científica, a avaliação da representatividade, repetibilidade e reprodutibilidade do método é efectuada através da informação científica e técnica publicada ou pela investigação efectuada no laboratório; pela base científica do método; pela modelação e simulação; pelos estudos de interferências e fontes de erro; pela optimização das condições operatórias e/ou robustez do método; pelo estudo dos parâmetros característicos do método (por exemplo: campo de aplicação, exactidão, repetibilidade, reprodutibilidade, incerteza, etc.).



Na abordagem comparativa, a reprodutibilidade do método é avaliada comparando os seus resultados aos obtidos por outros métodos já validados. O processo é validado se os resultados obtidos pertencerem ao intervalo da incerteza da comparação. As comparações podem ser efectuadas com métodos normalizados; métodos de referência; padrões internacionais ou

---

<sup>f</sup> [VIM 3.7] - Reprodutibilidade dos resultados (de uma medição) Aproximação entre os resultados das medições da mesma mensuranda efectuada com alteração das condições de medição. NOTAS: 1. Uma informação válida de reprodutibilidade exige à especificação das condições alteradas. 2. As condições alteradas podem incluir:- princípio de medição;- método de medição;- observador;- instrumento de medição;- padrão de referência;- local;- condições de utilização;- tempo. 3. A reprodutibilidade pode exprimir-se quantitativamente em termos das características da dispersão dos resultados. 4. Os resultados aqui são usualmente entendidos como resultados corrigidos.

materiais de referência certificados e/ou através de comparações interlaboratoriais com os LNM dos outros países.

A qualificação e avaliação da competência dos recursos humanos que efectuem as medições deve constar nos documentos da qualidade e referida nos procedimentos técnicos. Finalmente a validação será obtida através do uso combinado dos procedimentos acima descritos.

### 3. Validação de Resultados

#### 3.1 Conceitos Gerais

O Planeamento de experiências é uma técnica estatística utilizada na melhoria e optimização de um processo (ver Fig. 2). É um método estatístico activo que permite ao experimentador programar "tratamentos" e medições de modo a optimizar a sua eficiência [7]. Este método permite através da variação das entradas e observando a variação correspondente nas saídas fazer a inferência, através da rejeição da *hipótese nula* ( $H_0$ ) sobre quais as saídas estatisticamente diferentes, para um nível de significância  $\alpha$  também conhecido por "risco do produtor"<sup>g</sup>.

De modo inverso, pode ser utilizado para testar a homogeneidade de uma amostra<sup>h</sup>, para o mesmo nível de significância, identificar quais os resultados que podem ser considerados "outliers", o modo de os tratar em termos de *SPC*<sup>i</sup> e qual a incerteza associada à amostra. O procedimento é a bem conhecida análise da variância (*ANOVA*).

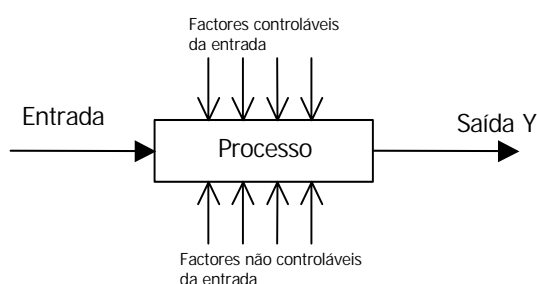


Figura 2- Modelo geral de um processo [7]

A análise da variância é definida [8] como "a técnica que consiste na separação da variação total de uma variável de resposta em componentes significativos, associados a fontes específicas de variação". Neste artigo é descrita a técnica ANOVA para um planeamento hierarquizado (ou aninhado<sup>j</sup>) em três níveis. O processo para estimar os componentes de variância associados a cada factor e um exemplo de aplicação são descritos para a avaliação de componentes de

incerteza-padrão pelo método tipo A<sup>k</sup> [9] na comparação de duas células do ponto triplo da água.

#### 3.2 Exemplo de Aplicação do Planeamento de Experiências

Para realizar a Escala Internacional de Temperatura de 1990 (EIT90), os Laboratórios Nacionais de Temperatura possuem normalmente mais do que uma célula para cada ponto fixo de definição da escala<sup>l</sup>. O laboratório pode considerar uma das células como a sua célula de referência e o seu valor de referência, recebendo a outra(s) o papel de padrão(ões) de trabalho ou considerar como valor de referência a média dos valores das células. Em qualquer dos casos,

<sup>g</sup> Probabilidade do erro *tipo I* ou probabilidade de rejeitar a hipótese nula e esta ser verdadeira.

<sup>h</sup> Neste caso, amostra é o conjunto das medições.

<sup>i</sup> SPC – *Statistical Process Control* - Controlo Estatístico de Processo.

<sup>j</sup> Nested design

<sup>k</sup> A avaliação de Tipo A da incerteza-padrão é o método de avaliação da incerteza pela análise estatística de séries de observações.

<sup>l</sup> A escala de temperatura em vigor, EIT90, estabelece valores de temperatura a estados reprodutíveis de equilíbrio térmico - pontos fixos termométricos - e define quais os termómetros a utilizar na interpolação entre esses valores.

as células devem ser regularmente comparadas e efectuado o cálculo da incerteza destas comparações. Situação similar existe quando este laboratório compara o seu valor de referência com o valor de um padrão viajante durante uma comparação interlaboratorial.

Para obter as diferenças entre as células são normalmente utilizados dois ou três termómetros. As medições de repetibilidade são efectuadas diariamente no “patamar” de equilíbrio térmico e a experiência é repetida em dias subsequentes. Todo este procedimento (*run*) poderá ser repetido algum tempo depois.

O cálculo da incerteza deverá ter em conta estas fontes de variabilidade função do tempo ( $t$ ), a repetibilidade a “curto-prazo”, a repetibilidade dia-a-dia ou a “médio-prazo” e a variação aleatória a “longo-prazo” dos resultados [9]. Estes componentes de incerteza “Tipo A” são avaliados pela análise estatística dos dados obtidos na experiência utilizando o planeamento hierarquizado de experiências, modelo “componentes de variância ou de efeitos aleatórios”.

Nesta “experiência” os factores são os termómetros utilizados, as medições feitas em dias subsequentes e as efectuadas em cada “run”. Estes factores são considerados *amostras* aleatórias da *população* que queremos estudar.

### 3.2.1 O Planeamento Hierarquizado. Modelo Geral

O planeamento hierarquizado é definido [8, 10] como “o planeamento de experiências no qual cada nível de um dado factor aparece só num nível de um outro factor”. Tem como objectivo deduzir os valores dos componentes de variância dos factores envolvidos que não podem ser medidos directamente. Os factores (ver Fig. 3) estão hierarquizados como numa “árvore” e qualquer “caminho” do “tronco” aos “ramos extremos” encontrará o mesmo número de “nós”.

Neste planeamento cada nível é analisado com o modelo de componentes de variância ou de efeitos aleatórios “encaixado” no nível subsequente.

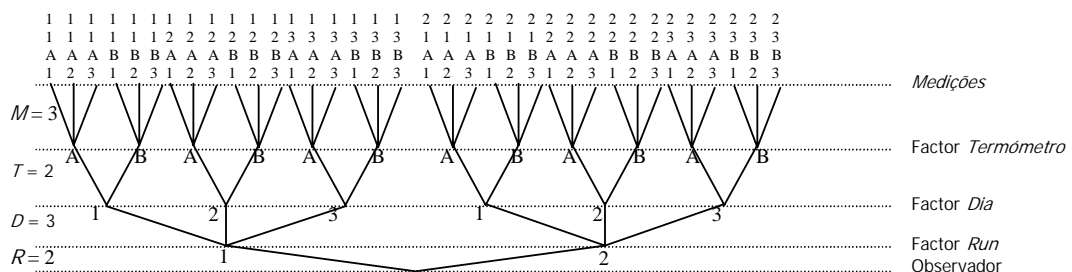


Figura 3. Planeamento de experiências hierarquizado

Considere-se, em primeiro lugar, um só factor. Neste modelo [7, 11], qualquer observação aleatória representada por  $y_{ij}$ , representa a  $j^{\text{ésima}}$  observação tomada no nível- $i$  do factor. O modelo matemático que descreve o conjunto de dados é:

$$y_{ij} = M_i + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (i = 1, 2, \dots, a \quad e \quad j = 1, 2, \dots, n) \quad (1)$$

onde  $M_i$  é o valor esperado do grupo de observações do nível- $i$ ,  $\mu$  a média global,  $\tau_i$  uma variável aleatória que corresponde ao efeito do nível- $i$  do factor e  $\varepsilon_{ij}$  o componente aleatório do erro associado à  $j^{\text{ésima}}$  observação do nível- $i$ .

Para testar as hipóteses, assume-se que os erros e os efeitos dos níveis do factor seguem uma distribuição normal e independentemente distribuída, respectivamente com média igual a zero e variância  $\sigma^2$  ou  $\varepsilon_{ij} \sim N(0, \sigma^2)$  e média igual a zero e variância  $\sigma_\tau^2$  ou  $\tau_i \sim N(0, \sigma_\tau^2)$ . A variância de cada observação é composta pela soma dos componentes da variância de acordo com:

$$\sigma_y^2 = \sigma_\tau^2 + \sigma^2 \quad (2)$$

O teste é unilateral e as hipóteses são:

$$\begin{aligned} H_0 : \sigma_\tau^2 &= 0 \\ H_1 : \sigma_\tau^2 &> 0 \end{aligned} \quad (3)$$

Isto é, se a hipótese nula for verdadeira, todos os efeitos dos níveis do factor são “iguais” a zero e cada observação é composta pela média global mais o erro aleatório  $\varepsilon_{ij} \sim N(0, \sigma^2)$ . A soma de quadrados total ( $SS_T$ ) dos desvios das observações  $y_{ij}$  em relação à média global  $\bar{y}$ , “medida” da variabilidade total dos dados, pode ser expressa por:

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y})^2 &= \sum_{i=1}^a \sum_{j=1}^n [(\bar{y}_i - \bar{y}) + (y_{ij} - \bar{y}_i)]^2 = n \sum_{i=1}^a (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 + \\ &+ 2 \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_i - \bar{y})(y_{ij} - \bar{y}_i) \end{aligned} \quad (4)$$

∴

$$SS_T = SS_{Factor} + SS_{Erro}$$

Como o produto cruzado é igual a zero [11],  $SS_T$  pode ser expressa em duas parcelas, a soma dos quadrados dos desvios das médias de cada nível do factor e a média global ( $SS_{Factor}$ ), uma “medida” das diferença entre os grupos de níveis e a soma dos quadrados dos desvios das observações dentro de cada grupo de nível do factor da média do grupo ( $SS_{Erro}$ ), devida ao erro aleatório. Dividindo cada soma dos quadrados pelos respectivos graus de liberdade, obtemos as correspondentes médias de quadrados<sup>m</sup> ( $MS$ ):

$$\begin{aligned} MS_{Factor} &= \frac{n}{a-1} \sum_{i=1}^a (\bar{y}_i - \bar{y})^2 \\ MS_{Erro} &= \frac{\sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{a(n-1)} = \sigma^2 \end{aligned} \quad (5)$$

A média de quadrados entre os grupos ( $MS_{Factor}$ ) é um estimador não enviesado da variância  $\sigma^2$ , se  $H_0$  for verdadeira, ou sobrestima  $\sigma^2$  (ver Eq. 7) se esta for falsa. A média de quadrados dentro do grupo ( $MS_{Erro}$ ) é sempre um estimador não enviesado da variância  $\sigma^2$ . Para testar a hipótese, utilizamos a estatística:

$$F_0 = \frac{MS_{Factor}}{MS_{Erro}} \sim F_{\alpha, a-1, a(n-1)} \quad (6)$$

Onde  $F^n$  é a distribuição amostral *Fisher-Snedcor* com  $a$  e  $a \times (n-1)$  graus de liberdade.

<sup>m</sup> Ou desvio quadrático médio.

<sup>n</sup> Distribuição amostral  $F$ : Se duas variáveis aleatórias independentes com distribuição do Qui-quadrado  $\chi_u^2$  e  $\chi_v^2$  com  $u$  e  $v$  graus de liberdade, o seu *ratio* distribui-se como *F-Snedcor* com os graus de liberdade  $u$  do numerador e  $v$  do denominador.

Se  $F_0 > F_{\alpha, a-1, a(n-1)}$ , rejeitamos a hipótese nula e conclui-se que a variância  $\sigma_\tau^2$  é significativamente diferente de zero, para um nível de significância  $\alpha$ . O valor esperado do  $MS_{Factor}$  é [12] para amostras de igual dimensão:

$$E(MS_{Factor}) = E\left[\frac{n}{a-1} \sum_{i=1}^a (\bar{y}_i - \bar{y})^2\right] = \sigma^2 + n\sigma_\tau^2 \quad (7)$$

pelo que o componente da variância do factor é obtido por:

$$\sigma_\tau^2 = \frac{E(MS_{Factor}) - \sigma^2}{n} \quad (8)$$

Considere-se agora o planeamento hierarquizado da Fig. 2. O modelo matemático é:

$$y_{rdtm} = \mu + P_r + \Delta_d + T_t + \varepsilon_{rdtm} \quad (9)$$

onde  $y_{rdtm}$  é a  $(rdtm)^{ésima}$  observação,  $\mu$  a média global,  $P_r$  o efeito aleatório dos níveis do factor  $R$ ,  $\Delta_d$  o efeito aleatório dos níveis  $D$ ,  $T_t$  o efeito aleatório dos níveis  $T$  e  $\varepsilon_{rdtm}$  a componente aleatória do erro.

É assumido que os erros e os efeitos dos níveis são normais e independentemente distribuídos com média igual a zero e variância  $\sigma^2$  ou  $\varepsilon_{ij} \sim N(0, \sigma^2)$  e com médias iguais a zero e variâncias  $\sigma_r^2$ ,  $\sigma_d^2$  e  $\sigma_t^2$ . A variância de qualquer observação é composta pela soma dos componentes da variância e o número total de medições,  $N$ , é obtida pelo produto das dimensões dos factores ( $N=R \times D \times T \times M$ ). Podemos expressar a variabilidade total dos dados [13, 14] por:

$$\sum_r \sum_d \sum_t \sum_m (y_{rdtm} - \bar{y})^2 = \sum_r DTM (\bar{y}_r - \bar{y})^2 + \sum_r \sum_d T.M. (\bar{y}_{rd} - \bar{y}_r)^2 + \sum_r \sum_d \sum_t M. (\bar{y}_{rdt} - \bar{y}_{rd})^2 + \sum_r \sum_d \sum_t \sum_m (y_{rdtm} - \bar{y}_{rdt})^2 \quad (10)$$

or

$$SS_T = SS_R + SS_{D|R} + SS_{T|DR} + SS_E$$

Assim, a variabilidade total dos dados é decomposta pela soma dos quadrados dos desvios correspondentes ao factor  $R$  ( $SS_R$ ), ou o efeito do factor  $R$  ( $run$ ), mais a soma dos quadrados dos desvios do factor  $D$  (dia) para o mesmo  $R$  ( $SS_{D|R}$ ), mais a soma dos quadrados do factor  $T$  (termómetro) para o mesmo  $D$  e o mesmo  $R$  ( $SS_{T|DR}$ ) e finalmente  $SS_{E_i}$  a variação residual. Dividindo pelos respectivos graus de liberdade, obtemos as médias de quadrados dos factores hierarquizados, os quais são estimadores da variância  $\sigma^2$ , se não existir variabilidade significativa devida aos factores. Os estimadores dos componentes de variância são obtidos igualando as médias de quadrados aos seus valores esperados e resolvendo as correspondentes equações:

$$\begin{aligned} E(MS_R) &= E\left[\frac{SS_R}{R-1}\right] = \sigma^2 + M\sigma_\tau^2 + TM\sigma_D^2 + DTM\sigma_R^2 \\ E(MS_{D|R}) &= E\left[\frac{SS_{D|R}}{R(D-1)}\right] = \sigma^2 + M\sigma_\tau^2 + TM\sigma_D^2 \\ E(MS_{T|DR}) &= E\left[\frac{SS_{T|DR}}{RD(T-1)}\right] = \sigma^2 + M\sigma_\tau^2 \\ E(MS_E) &= E\left[\frac{SS_E}{RDT(M-1)}\right] = \sigma^2 \end{aligned} \quad (11)$$

### 3.2.2 Comparação de Duas Células para o Ponto Triplo da Água num Planeamento Hierarquizado com Três Factores

Na comparação das duas células para o ponto triplo da água ( $t = 0,01\text{ }^{\circ}\text{C}$ ), JA e HS, utilizou-se dois termómetros padrão de resistência de platina de 25 ohm (SPRTs) A e B. Depois da preparação dos mantos de gelo, as células foram mantidas num banho termorregulado a uma temperatura de  $0,007\text{ }^{\circ}\text{C}$ . Este banho pode manter até quatro células no ponto triplo da água durante várias semanas. A preparação foi efectuada de acordo com o procedimento do laboratório, 48 horas antes do início das medições, de modo a permitir a estabilização das células. Foram efectuadas seis medições diárias das diferenças entre as células com os dois SPRTs e este conjunto de medições foi repetido durante três dias consecutivos. Duas semanas mais tarde, as células foram novamente preparadas. Um *run* completo foi repetido (*Run 2*).

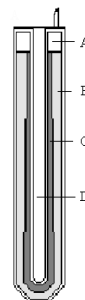


Figura 3 Célula do ponto triplo da água. A - Vapor de água; B - Água na fase líquida; C - "Manto" de gelo em torno do poço D - "Poço" para introduzir o termómetro (SPRT).

Nesta experiência hierarquizada, consideraram-se os efeitos correspondentes aos *runs* (Factor-*R*) com dois níveis ( $R = 2$ ), os efeitos correspondentes aos dias (Factor-*D*) com três níveis ( $D = 3$ ) para o mesmo *run*, os efeitos dos *termómetros* (Factor-*T*) com dois níveis ( $T = 2$ ) para o mesmo dia e para o mesmo *run* e a variação entre medições ( $M = 3$ ) para o mesmo termómetro, o mesmo dia e o mesmo *run* ou a variação residual. Considere-se os dados da Tabela 1:

Tabela 1: Resultados das medições do planeamento hierarquizado de experiências a três níveis.

SPRTs Medições		Medições- Diferenças $\Delta t$ (mK)											
		Run 1						Run 2					
		Dia 1		Dia 2		Dia 3		Dia 1		Dia 2		Dia 3	
		A11	B11	A21	B21	A31	B32	A12	B12	A22	B22	A32	B32
1		0,08	0,10	0,08	0,11	-0,07	-0,08	0,02	0,04	0,05	0,06	0,04	0,04
2		-0,04	0,03	-0,05	0,09	0,01	0,04	0,00	0,00	0,02	0,12	0,04	0,03
3		-0,04	0,05	-0,04	0,00	-0,05	0,04	0,05	0,07	0,00	0,00	0,01	0,00

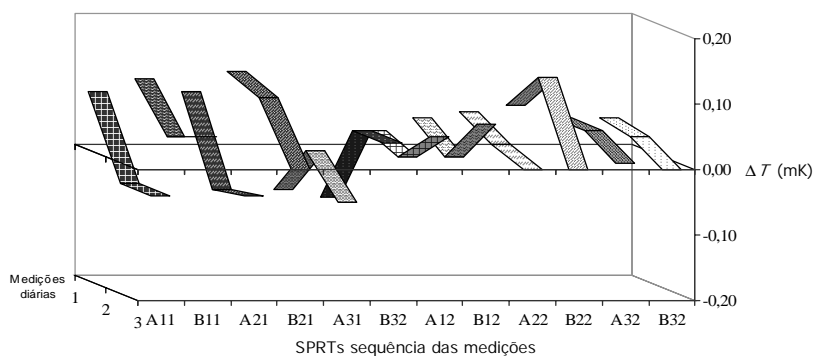


Figura 4: Representação esquemática das diferenças de temperatura observadas.

A análise de variância é normalmente representada numa tabela "Tabela ANOVA", que indica as soma dos quadrados, as médias de quadrados, os valores esperados das médias de quadrados e a estatística  $F_0$  calculada através dos *ratios* das médias de quadrados dos factores subsequentes envolvidos.



Desta tabela *ANOVA* (ver Tabela 2), obtiveram-se valores para  $F_0$  que serão comparados com a distribuição  $F$  para  $\alpha = 5\%$  e 1 e 4 graus de liberdade,  $F_{0,05, 1, 4} = 7,7086$  para o efeito do *run*, 4 e 6 graus de liberdade  $F_{0,05, 4, 6} = 4,5337$  para o efeito dos dias e 6 e 24 graus de liberdade  $F_{0,05, 6, 24} = 2,5082$  para o efeito dos termómetros.

Tabela 2: Tabela de análise de variância, para o exemplo da comparação de duas células do ponto triplo da água num planeamento hierarquizado de experiências com três factores.

Fonte de variação	Soma dos quadrados	Graus de liberdade	Medias de quadrados	$F_0$	Valor esperado das médias de quadrados
<i>Runs</i>	0,0030	1	0,0030	1,1610	$\sigma^2 + 3\sigma_T^2 + 6\sigma_D^2 + 18\sigma_R^2$
Dias	0,0104	4	0,0026	0,9133	$\sigma^2 + 3\sigma_T^2 + 6\sigma_D^2$
Termómetros	0,0171	6	0,0029	1,2168	$\sigma^2 + 3\sigma_T^2$
Medições	0,0563	24	0,0023		$\sigma^2$
Total	0,0868	35			

Verifica-se que todos os valores  $F_0$  são inferiores aos valores  $F$  e a hipótese nula nunca é rejeitada. Consequentemente, todas as médias de quadrados são estimadores da variância, os resultados pertencem à mesma população estatística e poderemos considerá-los validados.

Se a hipótese nula fosse rejeitada, deveriam ser investigados os resultados das medições com a aplicação, nomeadamente, de técnicas estatísticas de detecção de “outliers” [15]. A partir dos valores esperados, podem-se estimar os componentes da variância dos diferentes factores. Alguns autores [7, 12, 14] só determinam os componentes de variância no caso da rejeição da hipótese nula. Outros [10, 13] calculam estes componentes independentemente da validação dos resultados, considerando os componentes da variância como informação relevante da experiência.

Igualando as médias de quadrados aos valores esperados, estimam-se então os componentes de incerteza-padrão avaliados pelo método Tipo A (ver Tabela 3), que reflectem os componentes de variância aleatórios devidos aos efeitos dos factores.

Tabela 3: Balanço de incerteza para os componentes avaliados pelo método Tipo A

Componentes da incerteza-padrão avaliados pelo método tipo A	Variância (mK) <sup>2</sup>	Desvio-padrão (mK)
<i>Runs</i>	0,00002	0,005
Dias (variância negativa aproximar a zero)*	0	0
Termómetros	0,00017	0,013
Medições	0,00234	0,048
Total	0,00254	0,050

(\*) É frequente neste modelo e nos casos em que o efeito do factor é pouco significativo obter valores negativos no cálculo da variância. Nestes casos aproxima-se a zero o valor da variância respectiva. [16]

#### 4. Observações Finais - Conclusões

Foi descrita uma abordagem prática, para responder ao ponto 5.4.5 da norma NPEN ISO/CEI 17025 escolhida pela RMO - EUROMET para os Sistemas da Qualidade dos Laboratórios Nacionais de Metrologia. Foi também referido o que deveria ser incluído na documentação da Qualidade para a validação dos métodos e dos resultados.

Foi descrito o planeamento de experiências hierarquizado como uma técnica para validar os resultados de uma comparação, identificar e avaliar os componentes de incerteza-padrão "Tipo A" dos efeitos aleatórios função do tempo ( $t$ ).

Uma aplicação deste planeamento foi utilizada para descrever a análise dos componentes de variância num planeamento hierarquizado de três factores para uma comparação realizada com ensaios replicados a curto, médio e a longo termo, facilmente calculados numa "folha de cálculo" tipo "Excel".

---

#### Referências Bibliográficas

- 1 CIPM - Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates issued by National Metrology Institutes, Bureau International des Poids et Mesures, 1<sup>st</sup> ed., 1999.
- 2 CIPM - The International System of Units. Bureau International des Poids et Mesures, 1<sup>st</sup> ed., Sèvres, 1998.
- 3 ISO IEC - Requisitos Gerais de Competência para Laboratórios de Ensaio e Calibração. Versão Portuguesa (IPQ 2001) da ISO IEC 17025:2000. International Organization for Standardization, International Electrotechnical Committee, Genève, 2000.
- 4 ISO - Sistemas de gestão da qualidade: Fundamentos e vocabulário. Versão Portuguesa (IPQ 2001) da ISO 9000:2000. International Organization for Standardization, ISO 9000:2000, Genève, 2000.
- 5 EUROLAB - Validation of Test Methods. General Principles and Concepts. EUROLAB Internal Report EL1545/96, Brussels, 1996.
- 6 BIPM et al – Vocabulário Internacional de Metrologia. Versão Portuguesa (IPQ 1996) do International vocabulary of basic and general terms in metrology (VIM). International Organization for Standardization, Genève, 1993. ISBN 92-67-01075-1.
- 7 Montgomery, D. – *Introduction to Statistical Quality Control*. Singapore: John Wiley & Sons, 3<sup>rd</sup> ed., 1996. ISBN 0-471-51988-X, pp. 496-499.
- 8 ISO 3534-3, Statistics – Vocabulary and Symbols – Part 3: Design of Experiments, 2nd ed., Genève, International Organization for Standardization, 1999, pp. 31 (2.6) and 40-42 (3.4).
- 9 BIPM et al – Guide to the Expression of Uncertainty in Measurement (GUM), 2nd ed., International Organization for Standardization, Genève, 1995, pp 11, 83-87.
- 10 ISO/CD/TS 21749 Measurement uncertainty for metrological applications — Simple replication and nested experiments Genève, International Organization for Standardization, Reference number of working document: ISO/TC 69/SC 6 N 487, 2003-02-06.
- 11 Guimarães R. C., Cabral J. S., *Estatística*, 1st ed., Amadora: Mc-Graw Hill de Portugal, 1999, ISBN 972-8298-45-5, pp. 444-480.
- 12 Murteira, B., *Probabilidades e Estatística*. Volume II, 2nd ed., Amadora, Mc-Graw Hill de Portugal, 1990, ISBN 972-9241-07-4, pp. 361-364.

- 
- 13 Box, G. E. P., Hunter, W. G., Hunter J. S., *Statistics for Experimenters. An Introduction to Design, Data Analysis and Model Building*, 1st ed., New York, John Wiley & Sons, 1978, ISBN 0-471-09315-7, pp. 571-582.
- 14 Poirier J., "Analyse de la Variance et de la Régression. Plans d'Experience", *Techniques de l'Ingenieur*, **R1**, 1993, pp. R260-1 to R260-23.
- 15 Filipe E., "Validation of Calibration Methods – A Practical Approach", Advanced Mathematical and Computational Tools for Metrology VI, editors P. Ciarlini *et al* , World Scientific Publishing Co., 2004, Singapore, ISBN 981-238-904-0, pp. 285-292.
- 16 Milliken, G.A., Johnson D. E., *Analysis of Messy Data*. Vol. I: Designed Experiments. 1<sup>st</sup> ed., London, Chapman & Hall, 1997. ISBN 041-299-081-4, pp. 261.