

ANÁLISE ESTATÍSTICA DE DADOS EXPERIMENTAIS PROVENIENTES DE DIFERENTES TÉCNICAS ANALÍTICAS

Rodrigo Andrade Tavares¹, Jorge Eduardo de Souza Sarkis², Lisbeth Kaiserlian Cordani¹ e Maurício Hiromitu Kakazu²

¹Departamento de Estatística, IME -USP
Caixa Postal 20570
CEP 01452-990, São Paulo, S.P. - Brasil
e-mail: tavares@ime.usp.br

²Departamento de Caracterização de Materiais, IPEN-CNEN/SP
Caixa Postal 11049
CEP 05422-970, São Paulo, S.P. - Brasil
e-mail: jesarkis@net.ipen.br

RESUMO

A determinação da verdadeira composição química de um material é uma tarefa extremamente difícil. Normalmente adota-se como verdadeiro, o valor mais provável obtido por meio de uma técnica analítica já estabelecida. Entretanto, em várias situações rotineiras nos laboratórios ou mesmo durante trabalhos de certificação de materiais, freqüentemente nos deparamos com resultados provenientes de diferentes técnicas analíticas cuja comparação direta pode levar a erros. Neste trabalho, serão apresentados os resultados obtidos durante um exercício de certificação de materiais realizado na Supervisão de Caracterização de Materiais do Instituto de Pesquisas Energéticas e Nucleares (IPEN) em colaboração com o Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo (IME - USP). Amostras de U_3O_8 (Óxido de Urânio) de grau nuclear foram analisadas por meio de diferentes técnicas analíticas. Será apresentada uma análise dos dados utilizando-se de métodos tradicionais (Análise de Variância - ANOVA), Robustos e Balanceados.

INTRODUÇÃO

A análise química de um material é uma das atividades mais antigas no campo científico, entretanto, apesar do desenvolvimento dos procedimentos e técnicas analíticas sofisticadas, a determinação do verdadeiro valor de um componente químico ainda representa um desafio.

No cotidiano de um laboratório, o analista tem que controlar uma série de fatores relativos a instrumentos e procedimentos, que podem alterar de forma significativa a qualidade dos resultados ali produzidos, uma vez que toda técnica analítica tem suas limitações.

Dessa forma várias metodologias são desenvolvidas de modo a garantir a qualidade dos resultados, fornecer uma avaliação do comportamento de

um procedimento de análise e estabelecer um padrão de comparação com resultados obtidos de outros laboratórios, com o auxílio de ferramentas estatísticas. A maioria destes métodos opera de forma relativa e faz uso de padrões químicos certificados e de alta pureza.

Na seleção de um material a ser certificado, vários fatores precisam ser levados em consideração, dentre os quais destacam-se as aplicações a que se destinam e a metodologia para obtenção e análise de dados a ser adotada. A partir daí preparam-se os protocolos de formação do programa, definem-se número, quantidade e qualidade dos ensaios que irão caracterizar o material química e fisicamente, assim como os procedimentos de análise estatística dos dados.

O objetivo deste trabalho consiste em dar prosseguimento ao programa de produção de padrões químicos e isotópicos em andamento no Grupo de Caracterização Isotópica da Supervisão de Caracterização Química / ME, desenvolvendo um modelo específico para certificação do material que englobe protocolos, purificação, caracterização química e física e análise estatística dos dados de laboratório, sendo este último o enfoque deste artigo.

HISTÓRICO

Com o objetivo de desenvolver um modelo estatístico adequado foram feitos dois exercícios de intercomparação com cerca de 15 participantes simulando laboratórios. O material utilizado foi o U_3O_8 (Óxido de Urânio) de grau nuclear produzido no IPEN (Instituto de Pesquisas Energéticas e Nucleares).

Nestes exercícios, foram utilizadas duas metodologias de análise de dados. A primeira delas consiste no uso de técnicas clássicas como a ANOVA (Análise de Variância) [1], precedida de testes de detecção de "outliers" que visam assegurar que o conjunto de dados respeite as hipóteses necessárias para o uso deste método (igualdade de variâncias e dados com distribuição Normal).

Entretanto, a eliminação de dados pode introduzir tendência nos resultados, pois os valores descartados podem ser frutos da própria variabilidade do material, e não oriundos de erros de medida. Abordando este ponto, foram utilizados métodos robustos (Estimadores Dod - "Distribution of differences") que não eliminam dados na estimação da variabilidade de um conjunto de observações provenientes de uma distribuição Normal e, ao contrário dos métodos clássicos, são pouco sensíveis à presença de "outliers". Maiores detalhes sobre estes métodos podem ser encontrados em Beyrich et al. [2].

No primeiro exercício, procurou-se estudar o caso mais simples em que uma mesma técnica analítica foi adotada por todos os participantes para a análise do material. Os resultados, encontrados em Cordani e Yamamoto [3], demonstraram que dentro das incertezas de uma mesma técnica analítica, métodos gráficos de análise de dados (Dotplot, Boxplot e Youden) e convencionais com eliminação de resultados podem ser utilizados com a mesma eficácia quando comparados àquela obtida com métodos robustos.

Já no segundo exercício (Cordani e Fukunaga [4]), o tempo foi introduzido como fator de perturbação, sendo observado um aumento na variabilidade dos resultados. Neste caso, a utilização de métodos convencionais de análise pode mascarar o resultado final. Assim sendo, a utilização de métodos robustos se mostrou mais adequada.

O MÉTODO BALANCEADO

Dando prosseguimento a este programa, segue-se agora uma nova etapa que consiste na análise de dados provenientes de diferentes laboratórios e executados através de técnicas analíticas distintas, que é uma das metodologias mais utilizadas em programas de certificação de materiais.

Para tal fim foi estudado o Método Balanceado, uma técnica estatística robusta descrita em Analytical Methods Committee [5]. Esta técnica consiste em propor estimadores para a média e desvio padrão que não sejam sensíveis à presença de medidas espúrias ("outliers").

Supondo que participem L laboratórios, cada um com r medidas, num experimento pode-se definir uma particular medida de um laboratório de acordo com Eq. (1):

$$x_{ki} = \mu + \alpha_k + \varepsilon_{ki},$$

onde: α_k = erro do k -ésimo laboratório.
 ε_{ki} = erro da i -ésima medida do k -ésimo laboratório. (1)

$$\begin{aligned} \text{VAR}(\alpha_k) &= \sigma_L^2 & k &= 1, \dots, L. \\ \text{VAR}(\varepsilon_{ki}) &= \sigma_e^2 & i &= 1, \dots, r. \end{aligned}$$

Semelhantemente, a média de um laboratório pode ser dada pela Eq. (2):

$$\mu_k = \mu + \alpha_k \quad (2)$$

Primeiramente, deve-se estimar σ_e^2 , que representa a variabilidade interna dos laboratórios e suas respectivas médias μ_k , para em seguida proceder a estimação de σ_L^2 , que representa a variabilidade entre os laboratórios.

Como todos os laboratórios são tratados da mesma forma pelo método, a descrição abaixo corresponde a um processo de cálculo iterativo aplicado a cada laboratório.

Propõe-se então, que a estimação seja feita através de uma média ponderada das observações x_{ki} de cada laboratório, onde cada valor terá um peso w_{ki} obtido seguinte forma, conforme a Eq. (3):

$$w_{ki} = \begin{cases} 1, & \text{se } |x_{ki} - \mu_k| \leq c\sigma_e \\ \frac{c\sigma_e}{|x_{ki} - \mu_k|}, & \text{se } |x_{ki} - \mu_k| \geq c\sigma_e \end{cases} \quad (3)$$

onde σ_e é um desvio padrão robusto e c é uma constante com valores entre 1 e 2, sendo que o valor 1.5 é o que será utilizado neste trabalho, uma vez que é o mais recomendado. Portanto, de acordo com a Eq. (3), nota-se que, se uma determinada observação está a uma distância da média do laboratório a qual pertence superior a c vezes σ_e , seu valor será "atenuado" recebendo um peso inferior a

1, que diminui à medida que for maior a distância entre esta observação e a média do respectivo laboratório.

É importante ressaltar que a estimação de σ_e é feita simultaneamente à de μ_k , ou seja, a cada passo deste processo iterativo, ambos são estimados.

Pode-se definir então uma observação ponderada segundo a Eq. (4):

$$x_{ki}' = w_{ki} \cdot x_{ki} \quad (4)$$

O procedimento para estimar σ_e é muito semelhante ao descrito anteriormente, sendo que em cada passo:

$$\sigma_e = \sqrt{\frac{\sum (x_{ki}' - \mu_k)^2}{n\beta}} \quad (5)$$

onde n é o número de medidas do laboratório e β uma constante escolhida de tal forma a se obter a resposta compatível com dados normalmente distribuídos. Os valores de β estão tabelados em Analytical Methods Committee [5].

O valor inicial para μ_k pode ser dado pela média ou pela mediana dos dados originais, enquanto que a sugestão dada em Analytical Methods Committee [5] para o valor inicial de σ_e é:

$$\text{mediana} \frac{(|x_{ki} - \mu_k|)}{0.6745} \quad (6)$$

O processo só termina quando houver estabilização dos valores, que ocorre quando a diferença entre duas iterações for menor que um valor fixado. Neste exercício, será utilizado 10^{-2} como condição de parada.

O passo seguinte corresponde a analisar as médias estimadas de acordo com o que foi descrito para estimar o valor certificado μ e a variabilidade do material σ_L^2 , onde:

$$\sigma_L^2 = \sigma_L^2 + \frac{\sigma_e^2}{r} \quad (7)$$

Visto que σ_e^2 já foi estimado, resta σ_L^2 , de acordo com a Eq. (7). Esta estimação é feita analogamente ao processo descrito anteriormente, sendo que as L médias estimadas de cada laboratório são tratadas como observações de um único laboratório.

Com a finalidade de comparar este método com os demais já utilizados em exercícios anteriores, foram medidas porcentagens de pureza do U_3O_8 em 6 laboratórios do IPEN, através de diferentes técnicas analíticas. Sendo

que cada um dos laboratórios realizou 8 medidas, totalizando 48 observações apresentadas na Tabela 1:

TABELA 1. Porcentagem de pureza de U_3O_8 .

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6
98.60%	97.52%	99.65%	68.65%	100.20%	99.87%
99.80%	97.73%	99.76%	73.02%	99.60%	99.49%
99.80%	97.83%	99.65%	101.88%	100.20%	99.84%
100.00%	97.36%	99.76%	101.66%	99.50%	99.89%
99.80%	97.62%	99.53%	100.50%	100.40%	99.62%
99.80%	97.46%	99.76%	101.46%	99.50%	99.79%
99.80%	97.62%	99.53%	101.97%	102.30%	99.94%
100.00%	97.46%	99.65%	101.93%	99.80%	99.49%

As porcentagens superiores a 100 são frutos de contaminação no processo de medida, portanto apesar de teoricamente não serem esperadas, na prática podem ocorrer. Os dados serão analisados através da ANOVA, do Dod e pelo Método Balanceado.

ANÁLISE PELA ANOVA

Inicialmente, tem-se uma análise gráfica dos dados através de um Dotplot por laboratório, apresentado na Fig. 1:

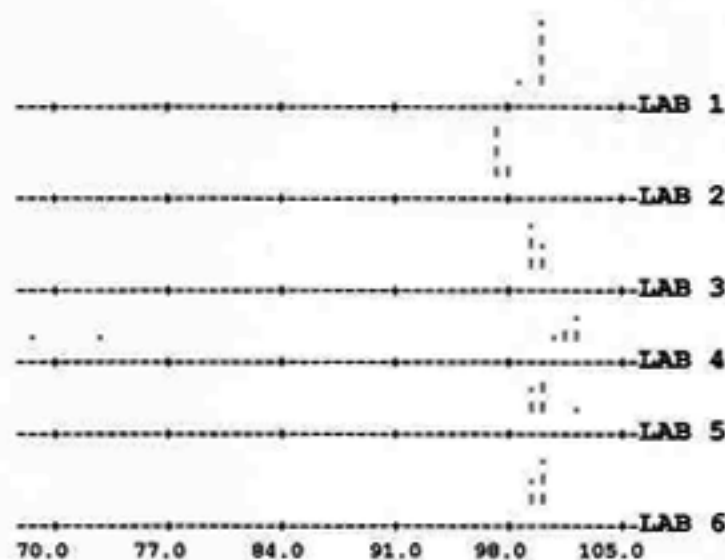


Figura 1. Dotplot das porcentagens de pureza de U_3O_8 .

É possível perceber que de um modo geral os dados se encontram pouco dispersos, entretanto o Laboratório 4 apresenta duas observações bem distantes que são candidatas a "outliers".

Na Tabela 2, são apresentadas as médias e desvio padrões amostrais por laboratório:

TABELA 2. Médias e desvios padrões amostrais.

Laboratório	Média	Desvio Padrão
1	99.70%	0.45%
2	97.58%	0.15%
3	99.66%	0.10%
4	93.88%	14.28%
5	100.19%	0.92%
6	99.74%	0.18%

Pode-se notar pela Tabela 2 que a variabilidade do Laboratório 4 é muito grande; isto se deve à presença de valores extremos que afetam medidas como média e desvio padrão.

Como se trata de uma análise clássica, foram feitos os testes de Dixon [6] para detecção de "outliers" e de Cochran [7] para igualdade de variâncias, ambos com um nível de significância de 5%. Após esta etapa, foram eliminadas as duas observações do Laboratório 4 que foram comentadas no parágrafo anterior, uma do Laboratório 1 e uma do Laboratório 5, sendo que estas últimas duas não se mostraram muito aberrantes graficamente, além do que, se tivessem ocorrido por exemplo no Laboratório 4, talvez não fossem excluídas da análise. As observações descartadas encontram-se em destaque na Tabela 1. Nenhum laboratório foi eliminado pelo teste de Cochran.

A diferença causada pela retirada destes "outliers" pode ser vista na Tabela 3, onde se encontram as médias e desvios padrões dos laboratórios após a eliminação destas medidas:

TABELA 3. Médias e desvios padrões após a eliminação dos "outliers".

Laboratório	Média	Desvio Padrão
1	99.85%	0.10%
2	97.58%	0.15%
3	99.66%	0.10%
4	101.57%	0.56%
5	99.89%	0.38%
6	99.74%	0.18%

Nota-se claramente que o descarte de alguns dados implicou em uma mudança nestas medidas descritivas, principalmente na variabilidade do Laboratório 4 que diminuiu muito.

Finalmente, após aplicar a ANOVA para os 44 valores restantes, foram obtidas as seguintes estimativas:

- Porcentagem média de pureza: 99.63%

- Desvio Padrão: 0.51%

ANÁLISE PELO DOD

Utilizando o conjunto original de dados, aplicou-se a técnica robusta Dod. Esta técnica não aplica critérios de eliminação de observações e estima apenas a variabilidade das medidas.

Com o auxílio de um programa computacional desenvolvido na linguagem Pascal, disponível no Centro de Estatística Aplicada (CEA - IME - USP), foi obtida a seguinte estimativa para a variabilidade, através do estimador DodM:

- Desvio Padrão: 0.23%

ANÁLISE PELO MÉTODO BALANCEADO

Para aplicação desta técnica, será utilizado um programa computacional desenvolvido em linguagem Fortran que se encontra em Analytical Methods Committee [5]. Primeiramente, a Tabela 4 apresenta as médias e os desvios padrões dos laboratórios, obtidos pelo processo iterativo do Método Balanceado. Abaixo da tabela estão as estimativas da média e do desvio padrão do conjunto.

TABELA 4. Médias e desvios padrões da última iteração do método balanceado.

Laboratório	Média	Desvio Padrão
1	99.77%	0.26%
2	97.58%	0.15%
3	99.61%	0.10%
4	101.41%	0.53%
5	99.97%	0.43%
6	99.74%	0.18%

- Porcentagem média de pureza: 99.69%
- Desvio Padrão: 0.18%

Os resultados desta Tabela 4 se mostram, descritivamente, muito semelhantes aos apresentados na Tabela 3, indicando que métodos que não eliminam observações podem ser usados como alternativa.

DISCUSSÃO

Com relação ao valor médio de pureza, os resultados da ANOVA e do Método Balanceado ficaram muito próximos, indicando que de fato os métodos robustos

permitem conclusões semelhantes ao clássico sem a eliminação de dados, no que diz respeito à esta estimativa.

A variabilidade da pureza apresentou maiores diferenças, principalmente quando são usados métodos robustos. É importante ressaltar que mesmo não excluindo dado algum, a variabilidade estimada por tais métodos foi menor do que a obtida pela ANOVA. Isto se deve ao fato de que os métodos robustos reduzem a influência de dados que se afastam da média, os quais provocam um efeito significativo sobre a estimativa desvio padrão.

Outro ponto importante é que o Método Balanceado atribuiu, em sua iteração final, pesos diferentes da unidade para 5 observações, sendo que 4 delas coincidem com as mesmas excluídas pelos testes de detecção de "outliers". Este fato indica que há uma certa coerência neste método, com respeito à detecção de observações suspeitas.

Deve ser ressaltado que a variabilidade estimada pela ANOVA está aceitável, mediante o erro das técnicas laboratoriais utilizadas. Cabe ainda dizer que, os resultados da variabilidade dos métodos robustos ficaram bem abaixo do clássico, devido ao fato de que os valores medidos pelos laboratórios, a menos dos "outliers", estão muito próximos, chegando ao ponto de alguns laboratórios apresentarem 2 ou 3 valores diferentes dentre 8 medidas.

Este é um fato preocupante, pois existem muitas fontes de erro num método, seja ele qual for, para existirem tantos valores repetidos dentro do mesmo laboratório. Talvez isto se deva à mentalidade de que um bom laboratório é aquele que "apresenta os resultados mais iguais possíveis". Este comportamento, aonde existir, deve ser abandonado, pois ficou claro nas análises realizadas que as variabilidades, tanto do material como da técnica, devem ser respeitadas, visto que esta é a representação realista do problema. Além disso, existem os modelos estatísticos que incorporam de maneira adequada tais variabilidades.

Com toda certeza, a qualidade do processo de medidas laboratoriais, no que diz respeito a técnicos e instrumental, implica em resultados compatíveis com a realidade das técnicas e dos materiais.

CONCLUSÃO

A partir dos exercícios realizados ao longo dos últimos 3 anos, pode-se concluir que os métodos que não eliminam dados se mostram como boas alternativas de análise e que os valores ditos espúrios devem ser encarados como parte da variabilidade do material e da técnica.

Percebeu-se também que o Método Balanceado reduz gradualmente a influência dos "outliers" na estimativa da variabilidade e que a eliminação de dados pode introduzir tendência nos resultados. Além disso, descartar observações nada mais é do que perda de

informação, o que implica em diminuição do poder das conclusões de qualquer análise.

Ficou claro que o "outlier" depende de um contexto, pois um valor pode ser considerado "outlier" num laboratório, ao mesmo tempo que se fosse considerado em outro não o seria.

Portanto, métodos alternativos ao clássico devem ser mais explorados daqui para frente com a finalidade de conhecer cada vez mais suas vantagens e limitações, uma vez que tudo o que foi feito até agora representa apenas um começo da exploração de modelos estatísticos novos. Também deve existir em maior escala o controle da qualidade dos laboratórios, com vistas a melhorar ainda mais o desempenho de todas as partes do processo de certificação de materiais.

REFERÊNCIAS

- [1] Neter, J., Wasserman, W., Kutner, M. H. (1990). **Applied linear statistical models: regression, analysis of variance and experimental design**. 3 ed. Homewood: Richard D. Irwin. 1181p.
- [2] Beyrich, W., Golly, W. (1989). **Evaluation of IDA-80 data by the Dod method**. Relatório técnico KfK 4157, EUR 10533EN-PWA 64/89, Karlsruhe, Germany. 20p.
- [3] Cordani, L. K., Yamamoto, W. H. (1992). **Programa de certificação de materiais estratégicos**. São Paulo, IME-USP, 42p. RAE-CEA-9215.
- [4] Cordani, L. K., Fukunaga, E. T. (1993). **Análise estatística de dados provenientes de programa de certificação química de compostos de urânio**. São Paulo, IME-USP, 42p. RAE-CEA-9316.
- [5] Analytical Methods Committee (1989). **Robust statistics - How not to reject outliers. Part 1 - Basic concepts. Part 2 - Intercomparative trials**. Analyst, 114, 1693-1702.
- [6] Dixon, W. J. (1953). **Processing data for outliers**. Biometrics, 9, 74-89.
- [7] Winer, B. J. (1971). **Statistical principles in experimental design**. 2ed. New York : McGraw-Hill. 907p.

ABSTRACT

The determination of the true material's chemical composition is a very difficult task. Usually, the most probable value obtained through an analytical technique

already established, is accepted as the true value. However, in many laboratory situations or even during material certification works, we often find results from different analytical techniques, whose direct comparison may cause wrong conclusions. In this work, will be presented the results of a material certification exercise performed at the Supervisão de Caracterização de Materiais do Instituto de Pesquisas Energéticas e Nucleares (IPEN) in cooperation with the Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo (IME - USP). Samples of U_3O_8 (Uranium Oxide) were analyzed by different analytical techniques. A statistical analysis using Analysis of Variance, as well as Robust and Weighted Methods, will be discussed.