Available online at www.sciencedirect.com

ScienceDirect

Annals of Nuclear Energy 35 (2008) 534–538

annals of NUCLEAR ENERGY

www.elsevier.com/locate/anucene

Technical Note

Multivariate analysis with covariance matrix applied to separative power modeling of a gas centrifuge

Elder Migliavacca *, Delvonei A. Andrade

Instituto de Pesquisas Energéticas e Nucleares, IPEN/CNEN-SP, Centro de Engenharia Nuclear (CEN), Av. Prof. Lineu Prestes, 2242, 05508-000 Cidade Universitária, São Paulo, SP, Brazil

Received 3 October 2006; accepted 19 July 2007
Available online 6 September 2007

Abstract

In this work, the least-squares methodology with covariance matrix is applied to determine a data curve fitting to obtain a performance function for the separative power δU of an ultracentrifuge as a function of variables that are experimentally controlled. The experimental data refer to 460 experiments on the ultracentrifugation process for uranium isotope separation. The experimental uncertainties related to these independent variables are considered in the calculation of the experimental separative power values, determining an experimental data input covariance matrix. The process variables, which significantly influence the δU values, are chosen to give information on the ultracentrifuge behaviour when submitted to several levels of feed flow rate F, cut θ and pressure in the product line, P_{p}. After the model goodness-of-fit validation, a residual analysis is carried out to verify the assumed basis concerning its randomness and independence and mainly the existence of residual heteroscedasticity with any explained regression model variable.

1. Introduction

The nuclear fuel cycle has several stages, from the mining of the uranium to the final assembly of the fuel elements used in a PWR reactor. The uranium is a mineral found in the nature in the form of a mixture of isotopes $^{235}UF_6$ and $^{238}UF_6$, in the proportion of 0.71% and 99.28% in mass, but only the isotope $^{235}UF_6$ is fissile, in the conditions used for energy generation, demanding the increase of the concentration of the fissile isotope $^{235}UF_6$ to the desirable levels. The uranium enrichment is the most critical stage in the nuclear fuel cycle. The objective of this work is to obtain models which relate the separative performance of an ultracentrifuge and the controlled variables in the separation process, through the application of the least square method with covariance matrix, in a data set representing 460 experiments, with varied conditions of feeding flow F, cut θ and pressure in the product line P_{p}. These variables are taken into account with the propagation of their associated experimental uncertainties.

2. Description of the process

A gas ultracentrifuge, shown in Fig. 1 is composed of a long, thin vertical cylinder (rotor), rotating around its axis at a high velocity, inside a case under vacuum. The process gas, assumed to be a binary isotopic mixture with $^{235}UF_6$ and $^{238}UF_6$, inside the cylinder, is subjected to a centrifuge force that establishes a pressure gradient in the radial direction, increasing from the center to the rotor wall (Jordan, 1980). That pressure distribution is slightly dissimilar to the different isotopes because it is proportional to the mass. This results in a partial separation of the feed F, into two fractions: an enriched one (product) and another depleted (waste) in the desired isotope ($^{235}UF_6$). The ultracentrifuge performance and its production capacity evaluation are usually done by means of the required work to isotope separation, which is proportional to the amount of processed material and to the obtained separation degree. Denoting by F, P and W, the streams of feed, product and waste

* Corresponding author. Tel.: +55 11 3817 7597; fax: +55 11 3816 9423.
E-mail address: emgvacca@usp.br (E. Migliavacca).

0306-4549/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.anucene.2007.07.008

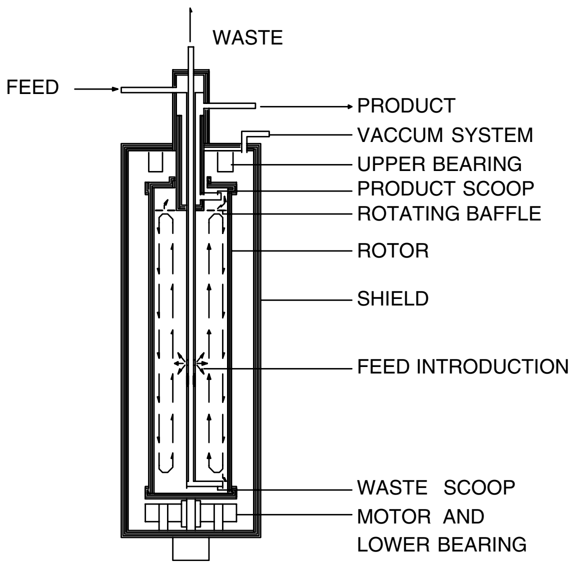

Fig. 1. Countercurrent ultracentrifuge design.

and by z, y and x, the respective isotope desired compositions, the dependent variable that best defines the separative efficiency of any isotope separation unit, is the separative power or capacity δU, given by the following expression:

$$\delta U = P(2y-1)\ln\frac{y}{1-y} + W(2x-1)\ln\frac{x}{1-x} - F(2z-1)\ln\frac{z}{1-z}, \quad (1)$$

where F, P and W are the operational variables and the response variables are the abundance rations of product $R_p = y/(1-y)$ and waste $R_w = x/(1-x)$.

3. Experimental methodology

An isotopic separation test is carried out by the operation of an ultracentrifuge in a bench plant shown in Fig. 2. The ultracentrifuge receives an input of a binary isotopic mixture with $^{235}UF_6$ and $^{238}UF_6$ as feed flow F, and allows the extractions of the product flow P and waste flow W. Samples are collected for verification of the separation obtained by the measures of the abundances rations of the enriched and depleted streams, R_p and R_w, respectively, allowing to calculate the separative power δU, given by Eq. (1). Defining the cut θ as the relation between the product and feed flow and fixing the product pressure line P_p, several groups of data are generated with the variation of the cut θ and the feed flow F. Each of them is denominated a separation experiment, resulting in an ultracentrifuge performance function like $\delta U(F, \theta, P_p)$.

4. Least squares method with covariance matrix

The measurements of R_p, R_w, P and W, involved in the determination of separative power δU, provide correlated uncertainties and define a covariance among them. These statistical uncertainties are propagated in Eq. (1) in order

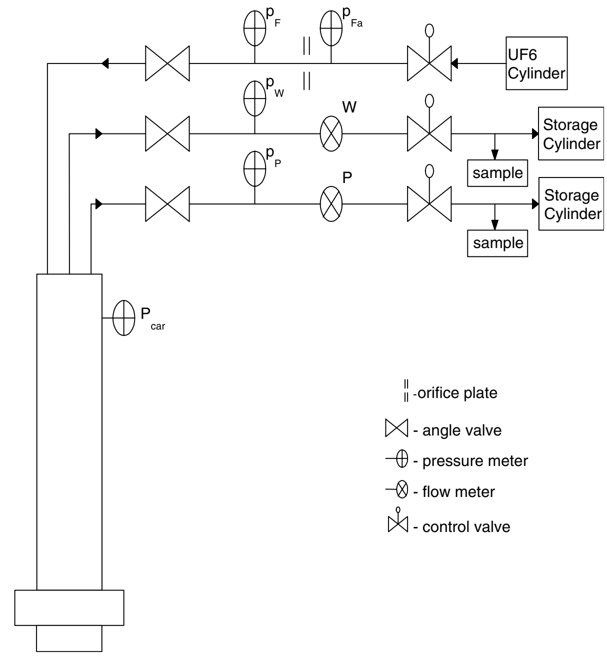

Fig. 2. Experimental bench plant design.

to obtain the δU final uncertainty using the expression (Cowan, 1998):

$$(\sigma_{\delta U})^2 \approx \sum_{i=1}^{n}\left(\frac{\partial \delta U}{\partial x_i}\right)^2 \sigma_i^2, \quad (2)$$

where x_i are the independent variables R_p, R_w, P and W; σ_i express their respective variances. R_p and R_w variances are directly given by mass spectrometry analysis while the P and W variances are calculated from mass flow meters calibration curves. Each δU experimental data covariance matrix element is calculated by the expression:

$$(V_{\delta U})_{ij} = \sum_{l=1}^{L} \rho_{ijl} e_{il} e_{jl} \quad (i, j = 1, n), \quad (3)$$

where e_{il}, e_{jl} are the partial uncertainties magnitudes of any independent variable R_p, R_w, P and W; ρ_{ijl} represents the micro correlations between these variable measurements due to each attribute l (Smith, 1991). These micro correlations values are safely determined by the experienced process analyst.

5. Data curve fitting

The δU experimental data fitting through a performance function of the kind $\delta U(F, \theta, P_p)$ is obtained due to δU and (F, θ, P_p) relation, that may be written as a second order polynomial given by

$$Y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \beta_{ij} x_i x_j + \sum \beta_{iij} x_i^2 x_j + \sum \beta_{iijj} x_i^2 x_j^2 \quad i \neq j, \quad (4)$$

where Y is the response δU, β_i are the equation coefficients, x_i and x_j are the controlled variables (F, θ, P_p). This equation is used to evaluate the linear, quadratic and interaction effects of these variables providing the project matrix A, which contains all the fitted model explained variables. Eq. (4) is a linear function in the β_i parameters and, although we can perform the least-squares method to any function, in this case the chi-square (χ^2) and estimators resulting values have desired properties: the estimators and their variances can be analytically obtained; they will be unbiased with minimum variance no matter the number of experiments and the experimental data distribution function. According to the least-squares method with covariance matrix, the best possible solution is the one which minimizes the χ^2. The χ^2 value for this particular problem is given by (Smith, 1981, 1993)

$$\chi^2 = (\delta U_{\text{exp}} - \delta U_{\text{calc}})^t V_{\delta U}^{-1} (\delta U_{\text{exp}} - \delta U_{\text{calc}}), \tag{5}$$

where $\delta U_{\text{calc}} = A\beta$, and β are the coefficients estimates vector of the fitted equation. Under the following conditions: (i) the δU experimental data is distributed according to a normal with a known covariance matrix, which allows to use the chi-square statistic, (ii) the fitted function, Eq. (4) is linear in the coefficients β_i, allowing to obtain an analytical solution for Eq. (5) and (iii) the functional form of the fitted function, Eq. (4), is corrected; it is possible to obtain the minimum deviation between the experimental and predicted values, so the quadratic form χ^2 should be distributed in conformity with the chi-square tables, allowing to evaluate the model goodness-of-fit (Cowan, 1998). The desired least-squares solution is given by

$$\beta = V_\beta A^t V_{\delta U}^{-1} \delta U_{\text{exp}}, \tag{6}$$

where the covariance matrix for the solution β is given by

$$V_\beta = (A^t V_{\delta U}^{-1} A)^{-1} \tag{7}$$

that gives the coefficients estimates variances and covariances of the experimental data fitted curves. In this case, a FORTRAN program (Migliavacca, 2004) was used.

6. Results and discussion

The experimental data performed with only one ultracentrifuge covered the whole domain of interest, consisting of eight values of feed flow F, seven values of cut θ and five levels of product pressure line P_p. Due to the confidentiality, inherent to the process development, the sensitive data were normalized, with all variables related to arbitrary units. The abundance rations R_p and R_w, obtained by gaseous spectrometry analyses, allowed the calculation of the separation power δU. After a detailed verification of the data, the same ones were checked, initially eliminating the experiments with inadequacy of samples and the incongruous ones in the values of the uncertainties in the abundance rations and, later on, eliminating the experiments that did not match the material balance.

6.1. Evaluation of the regression model

The model should be tested after its coefficients have been estimated and several specification tests can be used. The evaluation of the regression model is performed by two fundamental stages, the tests in the estimated parameters and the analysis of the residues. The test in the estimated coefficients is based on the null hypothesis that the theoretical value of a certain coefficient is zero. The t-Value is obtained by

$$t\text{-Value} = \text{estimate}/\text{standard deviation} \tag{8}$$

and through statistical tables, the level of significance of each coefficient of the equation is calculated. Those presenting p-Value $\leqslant 0.05$, are accepted as statistically significant. Those are typical values in this kind of analysis.

6.2. Regression models

Applying the method described, the first model obtained, Model 1, presented the term β_{112} with p-Value > 0.05, causing its elimination and the proposal of the Model 2, reduced model that describes the response variable δU and the controlled variables (F, θ, P_p) with all the coefficients statistically significant (Table 1).

In the next step, a residual analysis of the model is performed where it is initially verified whether there is a random scattering of the residues around zero. Through a detailed analysis of the data, three physically inconsistent values (negative values) of the separation power predicted by the model are identified. Those values were referring to extreme values of the independent variables, that is, high experimental values of F, P_p and θ, whose elimination can be justified. With that, a new regression curve was generated through the same adjustment procedure and finally yielding to the ultimate model, Model 3, represented by Table 2.

6.3. Residual analysis

The normality of the residues of the Model 3 was accepted because the percentage of values, in the range of their average taking into account two pattern deviation was 95.15%. The verification of the heteroscedasticity existence was made through the Spearman and Park tests. The correlation of Spearman test assumes that the variables can be ranked in two orderly series. In this case, it is calculated the Spearman correlation between the absolute value of the residues and each explanatory variable of the model, being tested the hypothesis that these correlations are null, through the value t, from a Student distribution. The Park test presented a regression with $R^2 = 0.0454$ and not any significant parameter, accepting the null hypothesis of homoscedasticity of the residues. On the other hand, the Spearman correlation test drove to the rejection of the same hypothesis, due to the slightly correlation existence among the residues of the regression with two explanatory variables: θ^2 and $\theta^2 F^2$.

6.4. Correction of the residues heteroscedasticity

Once it is not totally satisfied the premise of the non existence of the heteroscedasticity of the residues for the

Table 1
Estimates of coefficients of the variables in second-order polynomials and the associated statistical tests for the obtained models

Coefficient	Variable	Model 1			Model 2			Model 3		
		Estimate	*t*-Value	*p*-Value	Estimate	*t*-Value	*p*-Value	Estimate	*t*-Value	*p*-Value
β_0	Constant	0.2472	4.8528	0.000	0.2452	4.8162	0.000	0.1711	3.7598	0.000
β_1	θ	2.3944	14.0410	0.000	2.6032	22.0950	0.000	2.8038	26.6738	0.000
β_2	P_p	0.2844	3.4541	0.000	0.2285	3.0291	0.003	0.3830	8.3328	0.000
β_{11}	θ^2	−2.8879	−11.4831	0.000	−3.2914	−40.7340	0.000	−3.3685	−42.3127	0.000
β_{22}	P_p^2	−0.4287	−11.0172	0.000	−0.3916	−12.1740	0.000	−0.4533	−24.5722	0.000
β_{23}	P_pF	7.33E−04	2.2116	0.027	7.14E−04	2.1545	0.032			
β_{122}	θP_p^2	0.1420	2.4301	0.015	4.57E−02	3.4083	0.000			
β_{223}	P_p^2F	3.03E−03	16.4288	0.000	3.05E−03	16.5582	0.000	3.52E−03	29.9228	0.000
β_{112}	θ^2P_p	−0.2891	−1.6940	*0.091*						
β_{1133}	θ^2F^2	3.08E−06	2.8299	0.004	3.05E−06	2.8089	0.005	3.99E−06	3.6234	0.000
β_{2233}	$P_p^2F^2$	−9.21E−06	−23.2400	0.000	−9.24E−06	−23.3200	0.000	−9.55E−06	−24.2044	0.000

Table 2
Estimated coefficients for the Model 3

Coefficient	Variable	Value	B_0	β_1	β_2	β_{11}	β_{22}	β_{223}	β_{1133}	β_{2233}
B_0	Constant	0.1711(455)	*0.002072*	−0.002942	0.000530	−0.001475	0.001383	8.680E−09	−1.162E−09	1.781E−08
B_1	θ	2.8038(1051)	−0.6148	*0.011049*	0.000081	0.000183	−0.004584	−5.373E−08	9.455E−09	−2.532E−06
B_2	P_p	0.3830(184)	0.6309	0.0420	*0.000340*	−0.000761	0.000080	−1.678E−09	2.519E−09	−8.462E−07
β_{11}	θ^2	−3.3685(460)	−0.7048	0.0378	−0.8971	*0.002113*	−0.000068	−3.181E−10	4.054E−10	−9.014E−08
β_{22}	P_p^2	−0.4533(796)	0.3817	−0.5478	0.0544	−0.0187	*0.006338*	−2.972E−08	3.788E−09	−1.244E−06
β_{223}	P_p^2F	3.52E−03(1.1E−6)	0.1733	−0.4647	−0.0827	−0.0063	−0.3393	*1.210E−12*	−1.676E−13	4.353E−11
β_{1133}	θ^2F^2	3.99E−06(3.9E−7)	−0.0647	0.2281	0.3462	0.0224	0.1206	−0.3864	*1.556E−13*	−4.499E−11
β_{2233}	$P_p^2F^2$	−9.55E−06(1.2E−4)	0.0033	−0.2049	−0.3901	−0.0167	−0.1329	0.3366	−0.9703	*1.382E−08*
χ^2/v	0.94									

The values in the parentheses are the standard deviation in the last digits. Covariances between coefficients are shown in the upper triangle (including the main diagonal). Correlations are shown in the lower triangle. The last row gives the reduced χ^2.

proposed model, some type of correction should be used so that the residues become homoscedastic, or to have constant variances, giving credibility to the used statistical tests. Gujarati (2000) presented several types of transformations that can be applied to the experimental data to turn the residues homoscedastic. Box-Hunter (1978) shows that the tendency to a linear relationship of the residues with some explanatory variable of the model or even with the values predicted by the model, makes it possible to seek a convenient transformation of the experimental data, so that the variances of the residues become constant. Eq. (9) shows the transformation of the data

$$\varphi = (\delta U_{\text{exp}})^{\lambda} \tag{9}$$

and it was applied to the proposed statistical model. Using the experimental values of the separation power δU as dependent variable and the same explanatory variables of the Model 3, the program Boxcox.stb (STATSOFT, 1998) allowed to obtain the value of $\lambda = 0.3158$. Calculating the values of the δU experimental data according to the expression:

$$\delta U_{\text{transf.}} = (\delta U_{\text{exp}})^{\lambda=0.3158} \tag{10}$$

and the new experimental data covariance matrix according to the expression

$$(\sigma_{\delta U})_{\text{transf.}} \approx \lambda * (\delta U_{\text{exp}})^{\lambda-1} * (\sigma_{\delta U})_{\text{exp}} \tag{11}$$

a new regression was accomplished, yielding to new estimates of the coefficients of the representative model. However, applying the Spearman correlation test to the residues of the transformed model, it did not allow the acceptance of the null hypothesis of homoscedasticity between the residues and the explanatory variables. Such result can be interpreted as consequence that the tests for verification of the homoscedasticity were built for small samples. When applied to larger volume of data, they become very rigorous and any small variation has stronger probability of being considered. Parallelly, studies on the same experimental database indicated a degradation of part of the data, resultant of slow progressive alterations in the conditions of accomplishment of the experiments (Andrade, 2004). Due to these considerations the transformed Model 3 was discarded.

6.5. Model 3 goodness-of-fit

The quality of the adjustment of a regression model can be evaluated through the analysis of several statistical parameters, such as the variance–covariance coefficients matrix, the correlation coefficient, and the determination coefficient. The Model 3 determined a matrix with small variances of the coefficients, implying in small variability of the representative model. The correlation coefficient ($R = 0.9627$) represents a suitable linear correspondence

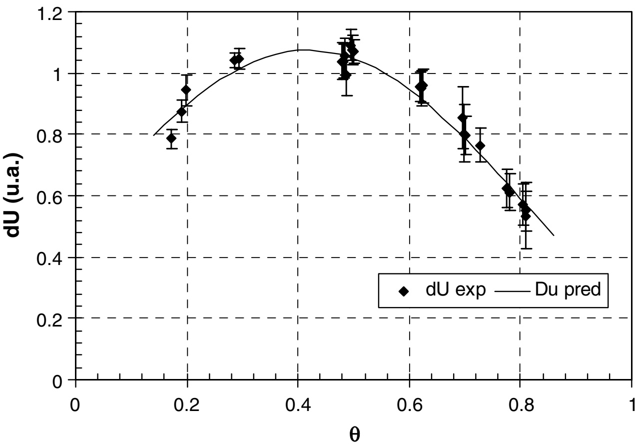

Fig. 3. δU response curve against the cut θ (P_p and F are fixed).

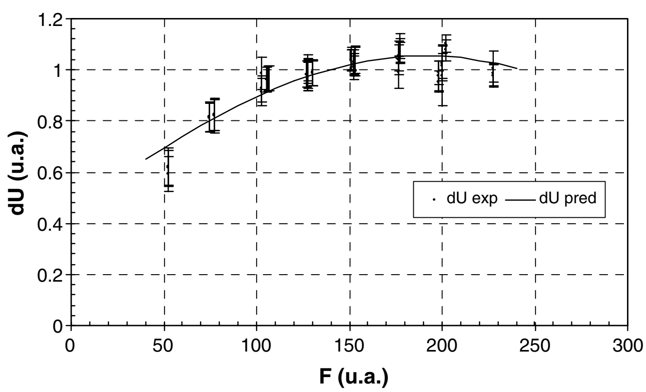

Fig. 4. δU response curve against the feed flow F. (θ and F are fixed).

among the experimental values and the values predicted by the Model 3 and finally, the determination coefficient ($R^2 = 0.9268$) means the amount of 92.68% of the total variance of the experimental data be explained by the model, with the remaining variance attributed to the variability of the data.

6.6. Graphic verification of the curve adjusted to the experimental points

If the regression model represents an appropriate mathematical relationship between the separation power δU and the process controlled variables, then necessarily, the theoretical curve should fit within the experimental points. Throughout the final theoretical model, represented by the Table 2, we can obtain the δU variation curves against the control variables θ and F presented in Figs. 3 and 4.

7. Conclusion

As a conclusion, the least squares method with covariance matrix demonstrated to be an efficient tool in the fitting curve determination of the ultracentrifuge separative power δU as a function of the process controlled variables. The value of the reduced χ^2, $\chi^2/V = 0.94$, indicated a good agreement between the dispersion of the experimental data of the δU and the estimates of the uncertainties contained in its covariance matrix. It is shown in the graphs of the response curves, Figs. 3 and 4, that the theoretical model is in a good agreement with the experimental data.

References

Andrade, M.C.V., 2004. Aplicação de redes neurais para a detecção de erros grosseiros em dados de processo em separação de isótopos de urânio por ultracentrifugação. São Paulo, SP, 2004. Dissertação (Mestrado) – Instituto de Pesquisas Energéticas e Nucleares.

Box, G.E.P., Hunter, W.G., Hunter, J.S., 1978. Statistics for Experimenters. [S.l.]. John Wiley & Sons.

Cowan, G., 1998. Statistical Data Analysis. Oxford University Press, New York.

Gujarati, D.N., 2000. Econometria Básica São Paulo. Pearson Education do Brazil.

Jordan, I., 1980. Separação dos isótopos de urânio pelo processo da centrifugação em fase gasosa. Informação IPEN 3 – IPEN-Inf-3.

Migliavacca, E., 2004. Modelagem do Comportamento Separativo de Ultracentrífugas por Regressão Multivariada com Matriz de Covariância, São Paulo, SP, 2004. Dissertação (Mestrado) – Instituto de Pesquisas Energéticas e Nucleares.

Smith, D.L., 1981. Covariance Matrices and Applications to the Field of Nuclear Data. Data and Measurements Serials. Argonne National Laboratory, Nov. (ANL/NDM-62).

Smith, D.L., 1991. Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology. American Nuclear Society, Inc., Illinois.

Smith, D.L., 1993. A Least-Squares Computational "Tool Kit". Nuclear Data and Measurements Series. Argonne National Laboratory, Apr. (ANL/NDM-128).

STATSOFT, 1998. Statistica for WINDOWS [Computer program manual]. StatSoft, Inc., Tulsa, OK.